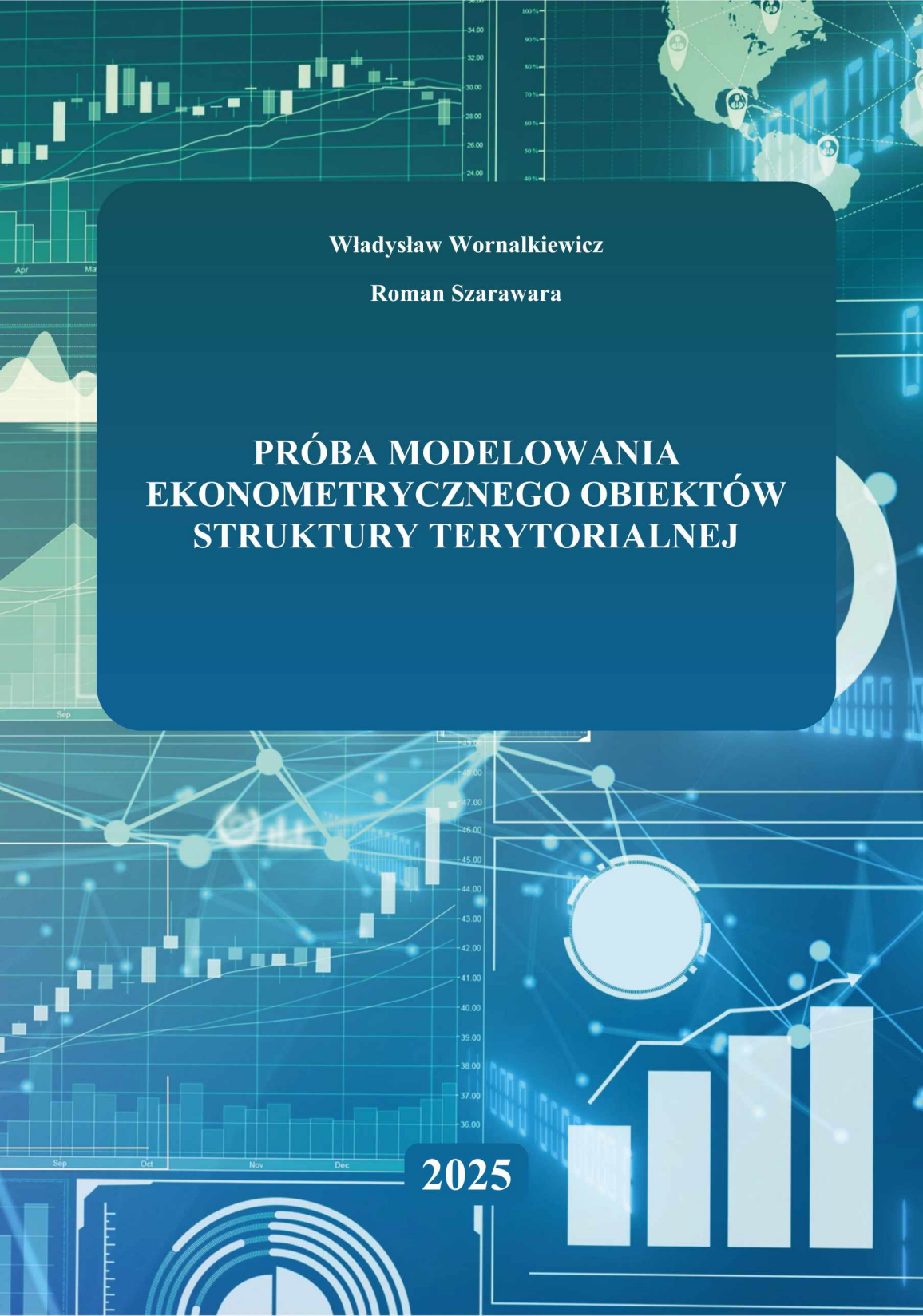


The background is a complex collage of financial and data-related graphics. It includes candlestick charts, line graphs, bar charts, and a world map with location pins. The color palette is primarily blue and green, with white text. The overall theme is economic analysis and data visualization.

Władysław Wornalkiewicz

Roman Szarawara

PRÓBA MODELOWANIA EKONOMETRYCZNEGO OBIEKTÓW STRUKTURY TERYTORIALNEJ

2025

Władysław Wornalkiewicz

Roman Szarawara

**PRÓBA MODELOWANIA
EKONOMETRYCZNEGO OBIEKTÓW
STRUKTURY TERYTORIALNEJ**

2025

UDK 332.12:330.43](075.8)

**Zalecane do publikacji przez radę akademicką
Poltawski Instytut Ekonomii i Prawa
(protokół nr 2 z dnia 28 października 2025 r.)**

Recenzenci:

*Prof. Akademii Śląskiej w Katowicach dr **Tetyana Nestorenko** (Poland);*

*Doktor nauk ekonomicznych, profesor, kierownik katedry finanse, bankowość i ubezpieczenia
Poltawski Państwowy Uniwersytet Rolniczy **Oleksija Zorya** (Ukraina)*

Wornalkiewicz W., Szarawara R. Próba modelowania ekonometrycznego obiektów struktury terytorialnej. Połtawa: Połtawski Instytut Ekonomii i Prawa Uniwersytetu „Ukraina”, 2025. 215 s.

ISBN 978-966-388-729-6

W Polsce obiektami struktury terytorialnej są regiony, województwa, podregiony, powiaty i gminy. Na określonym ich poziomie występuje rywalizacja, zwłaszcza w zakresie produktu brutto wytworzonych wyrobów i świadczonych usług. W opracowaniu wskazano na możliwość sklasyfikowania obiektów i ułożenia ich według pożądanego rankingu, biorąc pod uwagę cechy wpływające na ich efektywność, czy też nieefektywność działania. Najczęściej efektywnością ekonomiczną nazywamy rezultat działalności gospodarczej, określony przez stosunek uzyskanego efektu do nakładów. Niska efektywność ekonomiczna powoduje zwykle wzrost cen, zwiększenie zapotrzebowania na energię, materiały i powierzchnie produkcyjne - bez wzrostu wyników produkcyjnych. W niniejszym podręczniku podjęto się też budowy modelu produktu brutto podregionów, przez analogię do sformułowanego wcześniej modelu dla województw korzystając też z danych czasowo-przekrojowych. Natomiast inspiracją do budowy modelu dochodu budżetu powiatu lub miasta na prawach powiatu, na jednego mieszkańca w zł na przykładzie województwa dolnośląskiego, był adekwatny model liniowy wielowymiarowy z efektami indywidualnymi opracowany dla województw Polski. W tym celu dokonano konwersji utworzonej bazy danych w Excelu w format wymagany przez Gretl w celu wyłonienia tylko zmiennych istotnych. Procedura w tym względzie jest iteracyjna, podobna do zaprezentowanej przy modelowaniu PKB województw. Elementem wieńczącym postępowanie jest przeprowadzenie weryfikacji modelu poprzez sprawdzenie istotności parametrów.

UDK 332.12:330.43](075.8)

ISBN 978-966-388-729-6

© Połtawski Instytut Ekonomii i Prawa, 2025
© Wornalkiewicz W., Szarawara R, 2025
@Uniwersytet „Ukraina”, 2025

Wstęp	5
1. Miary kształtujące efektywność produktu krajowego brutto	6
1.1. Pojęcie efektywności.....	6
1.2. Energochłonność PKB.....	8
1.3. Efektywność inwestycji w PKB.....	10
1.4. Efektywność gospodarowania województw.....	12
1.5. Efektywność sektora publicznego.....	14
1.6. Efektywność przedsiębiorstw przemysłowych i rolnych.....	15
1.7. Efektywność kapitału trwałego.....	15
2. Propozycja określenia efektywności gospodarowania województw	17
2.1. Wprowadzenie.....	17
2.2. Baza danych statystycznych.....	18
2.3. Sformatowanie plików wejściowych.....	20
2.4. Charakterystyka zmiennych.....	22
2.5. Macierz korelacji.....	23
2.6. Określenie modelu ekonometrycznego.....	26
2.7. Zastosowanie programu DEAP.....	28
3. Etapy i procedury budowy modelu produktu brutto podregionów	34
3.1. Wstęp.....	34
3.2. Podejście do modelowania.....	36
3.3. Modelowanie z zastosowaniem Gretl-a.....	38
3.4. Porównanie modelu z danymi empirycznymi.....	42
4. Zastosowanie miary Penerose’a do oceny różnic w rozwoju podregionów	46
4.1. Wstęp.....	46
4.2. Macierze „częstkowe” jako próby.....	47
5. Modelowanie dochodu budżetu powiatów województwa dolnośląskiego	60
5.1. Wstęp.....	60
5.2. Baza danych statystycznych.....	61
5.3. Import danych.....	68
5.4. Generowanie zmiennych.....	69
6. Wyznaczanie efektywności technicznej programem DEAP	76
6.1. Wprowadzenie.....	76
6.2. Opis programu DEAP.....	76
6.3. Dane empiryczne.....	78
6.4. Przykład zastosowania.....	79
6.5. Uruchomienie programu.....	81
7. Wyznaczanie efektywności modelem CCR z karami	85
7.1. Wstęp.....	85
7.2. Rozwiązanie optymalne zadania CCR ukierunkowanego na nakłady.....	89
7.3. Luzy nakładów i rezultatów.....	93
7.4. Dwuetapowa korekta efektywności początkowej.....	97

8. Ocena nieefektywności gospodarowania zasobami ludności	102
8.1. Wstęp	102
8.2. Dostęp do danych statystycznych.....	104
8.3. Zamierzenia początkowe i przyjęte założenia	107
8.4. Modelowanie ekonometryczne cechy złożonej Mio	108
8.5. Zastosowanie opcji <i>Malmquist</i>	110
8.6. Sugestia wyznaczenia wag znaczeń społecznych.....	115
9. Budowa i weryfikacja modelu ekonometrycznego w wybranych programach.....	122
9.1. Wstęp	122
9.2. Budowa modelu ekonometrycznego w Gretl	123
9.3. Przygotowanie do realizacji w programie R	126
9.4. Weryfikacja podstawowa modelu końcowego	131
10. Dobór pakietu integracji modułów wsparcia rolnictwa	137
10.1. Wstęp	137
10.2. Zintegrowany system informatyczny	138
10.3. System zarządzania dokumentacją	140
10.4. Organizacja KOWR.....	142
10.5. Istniejące i projektowane zastosowanie informatyki.....	144
10.6. Zarządzanie wiedzą i organizacja baz danych.....	148
10.7. Potrzeba dalszego przystosowania stosowanych modułów do potrzeb użytkowników	149
10.8. Przeprowadzenie wywiadu i zakres potrzeb integracji eksploatowanych modułów	152
11. Wielokryterialny ranking obiektów	156
11.1. Zastosowane wzory i oznaczenia	156
11.2. Baza danych statystycznych	157
11.3. Unormowanie zmiennych.....	158
11.4. Obliczenie wartości zmiennych syntetycznych.....	158
12. Zastosowanie taksonomii wrocławskiej	161
12.1. Wstęp	161
12.2. Przygotowanie danych testowych	162
12.3. Odległości euklidesowe	167
13. Badanie występowania wielowymiarowego rozkładu normalnego	175
13.1. Wstęp	175
13.2. Zastosowanie zmodyfikowanej odległości Mahalanobisa	175
13.3. Test Shapiro-Wilka.....	179
13.4. Wykrywanie obserwacji nietypowych w sytuacji nieznanego rozkładu cech.....	182
14. Dostęp do publikacji w latach 2008-2025 (Władysław Wornalkiewicz)	186
Bibliografia	214

W polskiej strukturze terytorialnej wyodrębnia się regiony, województwa, podregiony, powiaty i gminy. Między tymi obiektami, na określonym poziomie, występuje rywalizacja, zwłaszcza w zakresie produktu brutto wytworzonych wyrobów i świadczonych usług. W niniejszym opracowaniu wskazano na możliwość sklasyfikowania obiektów i ułożenia ich według pożądanego rankingu, przy czym warto w tym względzie zwrócić uwagę na rozdziały:

- *Wielokryterialny ranking obiektów,*
- *Zastosowanie taksonomii wrocławskiej,*
- *Zastosowanie miary Penderosa'a do budowy różnic w rozwoju podregionów.*

Podjęto również próbę oceny nieefektywności gospodarowania zasobami ludności.

W wielowymiarowym modelowaniu produktu brutto określonych obiektów struktury terytorialnej zastosowano program DEAP. Posłużono się nim w celu określenia i umiejscowienia obiektów w szeregu pod kątem efektywności działania. Bliższe informacje dotyczące procedur budowy modeli gospodarowania znajdujemy w następujących rozdziałach tej publikacji:

- *Propozycja określenia efektywności gospodarowania województw,*
- *Etapy i procedury budowy modelu produktu brutto podregionów,*
- *Modelowanie dochodu budżetu powiatów województwa dolnośląskiego.*

Dla bliżej zainteresowanych, poruszaną w tej pracy tematyką, proponuje się skorzystanie i zaimplementowanie do innych zastosowań sposobów podanych w rozdziałach:

- *Wyznaczenie efektywności technicznej programem DEAP,*
- *Wyznaczenie efektywności modelem CCR z karami.*

Ponadto, naszym zdaniem w doskonaleniu się w procesie modelowania ekonometrycznego, w tym obiektów struktury terytorialnej, pomocne mogą być teksty zawarte w rozdziałach:

- *Budowa i weryfikacja modelu ekonometrycznego w wybranych programach,*
- *Badanie występowania wielowymiarowego rozkładu normalnego,*
- *Dobór pakietu integracji modułów wsparcia rolnictwa.*

Trzeba dodać jeszcze, że opracowanie stanowi rodzaj monografii skupiającej zaimplementowane różne artykuły, powstałe w czasie, stąd też mogą tu wystąpić powtórzenia, zwłaszcza w zakresie danych statystycznych.

Projektującym systemy informatyczne, zwłaszcza zintegrowane wiadomo, że rozległa jest problematyka implementacji pakietów dla potrzeb zarządzania. Pewne elementy z zestawionego dorobku bibliograficznego, w formie linków do wydanych publikacji zwartych i pozycji w monografiach, stanowią uwieńczenie tej pracy, w rozdziale „Dostęp do publikacji w latach 2008-2025” do lektury których zachęcamy Czytelników.

1. Miary kształtujące efektywność produktu krajowego brutto



1.1. Pojęcie efektywności

Zdefiniowanie pojęcia efektywności spotykamy w różnych źródłach, począwszy od słowników, encyklopedii jak też w materiałach publikowanych na łamach Internetu. Najczęściej efektywnością ekonomiczną nazywamy rezultat działalności gospodarczej, określony przez stosunek uzyskanego efektu do nakładów^{1 2}. Na uwagę zasługuje tzw. efektywność Kaldora-Hicksa, która jest jedną z metod oceny ekonomicznej efektywności alokacji zasobów. W myśl jej kryteriów rozwiązanie jest efektywne, gdy w wyniku jego zastosowania jeden podmiot zyskuje więcej, niż traci inny, a jednocześnie istnieje sposób kompensacji strat przez podmiot zyskujący na rzecz podmiotu tracącego³. Efektywność Kaldora-Hicksa jest rozwinięciem i uogólnieniem efektywności Pareta. Efektywność Pareta nie dopuszcza sytuacji w której którykolwiek z podmiotów straci na wprowadzeniu rozwiązania. Zatem **efektywność ekonomiczna to** sposób pomiaru skuteczności i celowości danej działalności gospodarczej, wyrażający się porównaniem wartości uzyskanych efektów do nakładu czynników użytych do ich uzyskania.

Niska efektywność ekonomiczna powoduje zwykle wzrost cen, zwiększenie zapotrzebowania na energię, materiały i powierzchnie produkcyjne - bez wzrostu wyników produkcyjnych. Wyższa efektywność ekonomiczna stwarza możliwości obniżenia kosztów wytwarzania, zwiększa wyniki produkcyjne i zysk, co z kolei umożliwia inwestowanie albo wzrost indywidualnych wynagrodzeń. Efektywność ekonomiczna stanowi jeden z podstawowych sposobów oceny podejmowanych działań gospodarczych⁴. Pojęcie efektywności odnoszone jest często do określonego przedsięwzięcia, przykładowo wprowadzenia systemu informatycznego. W tym przypadku miarą efektywności jest wskaźnik charakteryzujący stosunek korzyści osiągniętych dzięki uruchomieniu i eksploatacji systemu informatycznego do poniesionych na ten cel nakładów⁵.

Korzyści z uzyskania określonego produktu, czy też przykładowo z wprowadzenia systemu informatycznego dzielą się na wymierne i niewymierne oraz bezpośrednie i pośrednie. Do korzyści wymiernych zaliczamy:

- zmniejszenie wartości zapasów materiałowych,
- skrócenie cyklu realizacji typowych zadań administracyjnych,
- wzrost przeciętnej wydajności pracy pracowników zaplecza inżynieryjno-technicznego.

Natomiast jako korzyści niewymierne uznaje się:

- większą wiarygodność informacji sprawozdawczych,
- zmniejszenie uciążliwości niektórych czynności administracyjnych,
- wyższą kulturę obsługi klientów.

Korzyści bezpośrednie odnoszą się do stanowisk pracy lub obsługiwanych procesów. Przykładowo w skali mikroekonomicznej wprowadzenie systemu komputerowego kierowania przewozami może:

- zmniejszyć procent pustych przebiegów samochodów,
- skrócić łączną długość przejazdów przy utrzymaniu tych samych zadań transportowych.

¹ Słownik języka polskiego, PWN, Warszawa 1988.

² Leksykon Naukowo-techniczny, Wydawnictwo Naukowo-Techniczne, Warszawa 1972.

³ Encyklopedia Wikipedia, Internet

⁴ Encyklopedia WIEM, Internet

⁵ Encyklopedia organizacji i zarządzania, PWE, Warszawa 1982.

Trzymając się powyższego przykładu, to korzyścią pośrednią jest wyższa suma premii uzyskanych przez załogę służby transportowej przedsiębiorstwa. Istotne przy ocenie efektywności uzyskanego produktu, czy też wprowadzenia określonego systemu jest ustalenie hierarchii ważności korzyści różnych typów. Również ocena wydatków na określony cel wymaga wyodrębnienia nakładów bezpośrednich, przykładowo w odniesieniu do systemu informatycznego koszty zakupu sprzętu, za instalowania sieci i pośrednich kosztów tj. na reorganizację określonego obiektu, instytucji, gałęzi gospodarki⁶.

Wskaźniki wzrostu gospodarczego mają charakter ilościowy i jakościowy. Jest to zespół odpowiednio przygotowanych i przetworzonych danych statystycznych przedstawiających pewne wycinki życia gospodarczego kraju. Część z tych wskaźników stanowi elementy budowanych miar efektywności PKB, bowiem syntetyczny pomiar tej efektywności jest zadaniem złożonym. Z tego względu w niniejszym opracowaniu dokonano przeglądu opublikowanych materiałów w Internecie, w których autorzy podejmują próbę znalezienia tych miar. Jednak najpierw zachodzi potrzeba przytoczenia podstawowych pojęć z nimi związanych (tab. 1.1).

Tab. 1.1. Podstawowe pojęcia występujące w miarach efektywności PKB

Lp.	Symbol	Nazwa	Określenie
1	PPI	Indeks cen producentów	Miara zmiany cen produkcji przemysłowej.
2	CPI	Indeks wzrostu cen	Miara całorocznego wzrostu cen towarów i usług konsumpcyjnych.
3		Inflacja (patrz tab. 1.2)	Wzrost średniej ceny dóbr, odnoszący się do pewnego okresu. Przeciwnieństwem inflacji jest deflacja.
4		Koszty pracy	Wielkość dochodu z wypracowanego przez pracownika.
5		Koszty zatrudnienia	Koszty płac plus narzuty na płace, ubezpieczenia społeczne i zdrowotne, podatki dla państwa.
6	PPP	Parytet siły nabywczej	Opiera się na teorii jednej ceny, która mówi, że za określoną liczbę jednostek danej waluty można kupić w każdym kraju w tym samym czasie dokładnie tyle samo dóbr. Polega na dokonaniu przeliczenia kursu waluty według siły nabywczej. Zasadniczą sprawą jest zebranie danych o cenach z zagregowanej listy towarów, która zawiera towary porównywalne i reprezentatywne dla analizowanych krajów.
7		PKB nominalny	Oblicza się według bieżącej wartości pieniądza.
8		PKB realny	Oblicza się według realnej wartości pieniądza, a bez wpływu inflacji. Podzieleniu podlega PKB nominalne przez indeks cen. PKB realny najczęściej przedstawiany jest w cenach stałych z wybranego roku bazowego.
9	PKBER	PKB wg kursów wymiany	PKB przelicza się według bieżącego kursu wymiany (zazwyczaj na dolary amerykańskie).
10	PKBPPP	PKB wg siły nabywczej	PKB przelicza się według parytetu siły nabywczej.
11		PKB per capita	PKB w przeliczeniu na osobę (mieszkańca) kraju. Jest syntetyczną oceną poziomu rozwoju gospodarczego kraju.
12	PKB	Produkt krajowy brutto	Zagregowana wartość dóbr i usług finalnych wytworzonych na terenie danego kraju w określonej jednostce czasu (np. roku).
13	PNB	Produkt narodowy brutto	Wartość wszystkich dóbr i usług wytworzonych przez obywateli kraju oraz przez osoby prawne z siedzibą na jego terenie niezależnie od tego, czy podmioty te działają w kraju, czy za granicą.
14		Szara strefa	Wielkość produkcji niezarejestrowanej oraz produkcji gospodarstw domowych przeznaczanej na własne potrzeby.
15		Wartość dodana	Wartość dodana = Produkcja całkowita - wartość dóbr i usług zużytych do tej produkcji

Źródło: Opracowanie własne na podstawie encyklopedii internetowej *Wikipedia*.

⁶ Ibidem

Produkt krajowy brutto (PKB) jest to miernik produkcji wytworzonej na obszarze danego kraju, można powiedzieć, że jest:

- sumą wydatków gospodarstw domowych na zakup dóbr i usług konsumpcyjnych,
- wydatków sektora prywatnego na zakup dóbr i usług inwestycyjnych,
- wydatków państwa na zakup dóbr i usług,
- salda bilansu handlu zagranicznego.

PKB może być wyrażony w cenach rynkowych, zawierających podatki pośrednie lub - po ich potrąceniu - w cenach czynników wytwórczych (cenach netto). PKB różni się od produktu narodowego brutto (PNB) tym, że nie uwzględnia dochodu z lokat zagranicznych. Produkt krajowy brutto możemy przedstawić w postaci równania:

$$PKB = \text{produkcja globalna kraju} - \text{zużycie pośrednie} = \text{suma wartości dodanej ze wszystkich gałęzi gospodarki narodowej}$$

Natomiast od strony popytowej:

$$PKB = \text{konsumpcja} + \text{inwestycje} + \text{wydatki rządowe} + \text{eksport} - \text{import} + \text{zmiana stanu zapasów}$$

Zmiana stanu zapasów jest pomijana ze względu na małą wartość.

Od strony dochodowej produkt krajowy brutto zapiszemy równaniem:

$$PKB = \text{dochody z pracy} + \text{dochody z kapitału} + \text{dochody państwa} + \text{amortyzacja}$$

Tab. 1.2. Inflacja w Polsce w okresie 1950-2008 (nawiązanie do pkt. 3 tabeli 1.1)

Rok	Inflacja	Rok	Inflacja	Rok	Inflacja	Rok	Inflacja	Rok	Inflacja	Rok	Inflacja
1950	6,5%	1960	1,8%	1970	1,1%	1980	1%	1990	585,8%	2000	10,1%
1951	9,6%	1961	0,7%	1971	-0,1%	1981	21,2%	1991	70,3%	2001	5,5%
1952	14,4%	1962	2,5%	1972	0,0%	1982	100,8%	1992	43,0%	2002	1,9%
1953	41,9%	1963	0,8%	1973	2,8%	1983	22,1%	1993	35,3%	2003	0,8%
1954	-6,3%	1964	1,2%	1974	7,1%	1984	15,0%	1994	32,2%	2004	3,5%
1955	-2,4%	1965	0,9%	1975	3,0%	1985	15,1%	1995	27,8%	2005	2,1%
1956	-1,0%	1966	1,2%	1976	4,4%	1986	17,7%	1996	19,9%	2006	1,0%
1957	5,4%	1967	1,5%	1977	4,9%	1987	25,2%	1997	14,9%	2007	2,5%
1958	2,7%	1968	1,6%	1978	8,1%	1988	60,2%	1998	11,8%	2008	4,2%
1959	1,1%	1969	1,4%	1979	7,0%	1989	251,1%	1999	7,3%		

Źródło: Internet, encyklopedia Wikipedia.

1.2. Energochłonność PKB

Jak wiemy produkt krajowy brutto jest jedną z podstawowych miar tempa wzrostu gospodarczego. Dyskusję na temat kosztów jego uzyskania strat oraz efektów spotykamy w różnych materiałach publikowanych na łamach Internetu. Przegląd tych głosów może być pomocny w modelowaniu ekonometrycznym wpływu czynników kształtujących efektywność produktu krajowego brutto. Tempo wzrostu gospodarczego w regionie Europy Wschodniej, należy obecnie do najwyższych na świecie⁷. Wzrost ten przekracza tempo osiągane przez gospodarki „starych” państw Unii Europejskiej. Uznaje się jednak, że jest on znacznie bardziej uciążliwy dla środowiska naturalnego i może prowadzić do niekorzystnych zmian klimatycznych. Dla zmniejszenia emisji gazów cieplarnianych zachodzi potrzeba zwiększenia efektywności energetycznej w nowych krajach UE⁸. Zdaniem ekspertów zużycie energii w tym regionie cechuje się silnym zanieczyszczaniem środowiska, a przy tym jest wielce nieefektywne ekonomicznie. Polska emituje znacznie więcej dwutlenku węgla na jednostkę zużytej energii niż kraje „starej” UE. Związane jest to ze infrastrukturą, przestarzałymi metody produkcji i znacznym uzależnieniem energetyki od węgla.

⁷ Internet, hasło: efektywność PKB, *Zbyt drogi PKB*.

⁸ Internet, hasło j. w., *Zwiększyć efektywność energetyczną*.

Koszty produkcji przekraczają niekiedy ceny dostawy, co nie zachęca do unowocześniania tego sektora gospodarki. Europejski Bank Odbudowy i Rozwoju (EBOR) przeznacza obecnie 10 % swoich środków finansowych na inwestycje zwiększające efektywność energetyczną kredytobiorców z obszaru „chłodnych” państw UE. Stopniowo nastąpi przechodzenie na czystsza i efektywniejszą energię. Finansowane będą projekty dużych elektrowni zasilanych wiatrem i inną energią odnawialną. Przyłączenie się Polski do przedsięwzięcia ograniczenia gazów cieplarnianych oznacza spełnianie w przyszłości norm w tym zakresie. W badaniu efektywności energochłonności na produkt krajowy brutto posłużono się porównaniem przykładowo zużycia energii elektrycznej z wartościami średnimi krajów tzw. starej piętnastki Unii Europejskiej (UE-15)⁹.

Tab. 1.3. Dane i wskaźniki dotyczące zużycia energii elektrycznej

Lp.	Wielkość/wskaźnik	Polska	UE-15
1	Ludność [mln]	38,157	386,064
2	PKB liczony wg:		
	a) kursów wymiany (<i>Exchange Rate – ER</i>) [mld \$ER]	303,959	12 770,319
	b) parytetu siły nabywczej (<i>Purchasing Power Parity – PPP</i>) [mld \$PPP]	530,153	11 749,948
3	PKB na mieszkańca wg:		
	a) kursów wymiany walut [\$ER/os.]	7 966	33 078
	b) parytetu siły nabywczej [\$PPP/os.]	13 894	30 435
4	Zużycie energii elektrycznej:		
	a) brutto [TWh]	144,173	2 863,542
	b) finalnej [TWh]	98,835	2 443,911
	c) przez sektor energii [TWh / % zużycia brutto]	30,775 / 21,35%	239,421 / 8,36%
	d) straty sieciowe [TWh / % zużycia brutto]	14,563 / 10,10%	180,210 / 6,29%
5	Zużycie energii elektrycznej na mieszkańca:		
	a) brutto [kWh/os.]	3 778	7 417
	b) finalnej [kWh/os.]	2 590	6 330
6	Zużycie energii elektrycznej na jednostkę PKB wg kursów wymiany:		
	a) brutto [kWh/\$ER]	0,474	0,224
	b) finalnej [kWh/\$ER]	0,325	0,191
7	Zużycie energii elektrycznej na jednostkę PKB wg parytetu siły nabywczej:		
	a) brutto [kWh/\$PPP]	0,272	0,244
	b) finalnej [kWh/\$PPP]	0,186	0,208
8	Emisja CO ₂ przez sektor elektroenergetyki [tCO ₂ /MWh]	1,151	0,416

Źródło: Dane z Internetu dotyczące porównanie wskaźników zużycia energii elektrycznej w Polsce z wartościami średnimi dla krajów „starej piętnastki” UE (EU-15), dane z 2005 r., wielkości PKB wyrażone w \$ z 2000 r. Wielkości i wskaźniki zaczerpnięto z:

- Energy: Yearly statistics 2005. 2007 Edition. Eurostat Statistical books. European Commission. Eurostat. <http://epp.eurostat.ec.europa.eu>.

- Mały rocznik statystyczny Polski. 2007. GUS, Warszawa 2007.

Porównując kolumny 3 i 4 tabeli 1.3 zwróćmy uwagę na wskaźniki zaznaczone kursywą. Widzimy znaczne różnice w zużyciu energii elektrycznej przez sektor energii, straty związane z przesyłem energii, a także na zużycie na jednostkę PKB. Tak więc bardzo niekorzystnie wygląda porównanie wskaźników emisji CO₂ Polski i krajów UE-15. Ponadto spoglądając na tabelę 1.3 rodzi się

⁹ Internet, hasło: efektywność, PKB, artykuł: *Gdzie leży klucz do poprawy efektywności zużycia energii elektrycznej w Polsce?*

pytanie: *Jaka wartość PKB jest reprezentatywna dla potrzeb oceny energochłonności - liczona według kursów wymiany walut [SER] czy też parytetu siły nabywczej [SPPP]?¹⁰*

Wielkości produktu krajowego brutto liczonego według kursów wymiany (PKBER) różnią się znacznie od wielkości obliczonych według parytetu siły nabywczej (PKBPPP) – szczególnie wówczas gdy porównuje się ze sobą kraje o różnym poziomie rozwoju gospodarczego. Dla podanego okresu dla Polski PKB liczony według parytetu siły nabywczej jest 1,74- krotnie większy od liczonego według kursów wymiany. Zdaniem autora omawianego artykułu PKB liczone według kursów wymiany walut nie odzwierciedla prawidłowo poziomu zamożności społeczeństw. Kurs waluty danego kraju w stosunku do innych walut jest bowiem wypadkową aktualnej „siły” jego gospodarki (w tym: stanu równowagi budżetowej, bilansu handlowego i płatniczego) oraz polityki pieniężnej banku centralnego (stóp procentowych). Przykładem jest fakt, że w ostatnich latach, na skutek radykalnego osłabienia waluty amerykańskiej w stosunku do innych głównych walut, znacznie wzrosły nominalne wartości PKB różnych krajów wyrażone w USD. Tak więc właściwą dla oceny energochłonności gospodarki miarą dochodu narodowego jest PKB obliczony według parytetu siły nabywczej (PKBPPP).

1.3. Efektywność inwestycji w PKB

Wysoki udział inwestycji w produkcie krajowym brutto jest dowodem modernizacji kraju i zwiększa jego konkurencyjność na rynku międzynarodowym, a mierzony jest stopą inwestycji¹¹. Ta syntetyczna miara (wskaźnik) stanowiąca procent PKB służy do oceny aktywności inwestycyjnej gospodarki. Wysokość stopy inwestycji, konieczna do określonego wzrostu gospodarczego, może być różna w czasie. Tendencja wzrostowa tego wskaźnika przyczynia się do wzrostu gospodarczego. Anna Szymczak dokonała analizy tego procesu na tle wybranych krajów UE w latach 2000-2006¹². Zauważyła, że spośród 25. krajów UE Polska zajmuje ostatnie miejsce pod względem wielkości PKB *per capita*, a przyczyną jest relatywnie zbyt niski poziom inwestycji. Czynnikiem determinującym wzrost produkcji jest kapitał rzeczowy, który zależy od poziomu inwestycji. Majątek zużywa się, a wielkość inwestycji powinna przyczynić się do utrzymania zasobów. W poniższej tabeli 1.4 pokazano miary charakteryzujące udział i efektywność inwestycji w odniesieniu do produktu krajowego brutto i w układzie grup państw Unii Europejskiej według jej historycznego rozwoju, przy czym: UE 15 – „stara” piętnastka krajów, UE 25 – rozszerzenie o nowe kraje, w tym z Europy Wschodniej, UE 27 – dalsze rozszerzenie.

Tab. 1.4. Proponowane miary efektywności inwestycji

Kraj	Średnia (1998-2004)	2005	2006	2007
Stopa inwestycji [%]				
UE 15	19,9	19,9	20,5	21,1
UE 25	20,04	20,0	20,7	21,3
Polska	21,13	18,2	19,7	22,3
Dynamika realnego PKB				
UE 15	2,3	1,7	2,8	2,6
UE 25	2,4	1,8	3,0	2,8
Polska	3,6	3,6	6,2	6,5

¹⁰ Kielbasa W., Internet, HYDROENERGO.

¹¹ Szymczak A., *Próba oceny makroekonomicznej efektywności inwestycji w wybranych gospodarkach UE*, Rozdział 8. Opracowanie opublikowane w Internecie, hasło: efektywność, PKB.

¹² Ibidem.

Kapitałochłonność oraz efektywność inwestycji (1998-2007)				
	Kapitałochłonność	Efektywność		
UE 15	8,40	0,119		
UE 25	8,24	0,121		
Polska	7,07	0,246		
Stopa inwestycji publicznych				
	Średnia (1998-2004)	2005	2006	
UE 15	2,3	2,2	2,4	
UE 25	2,3	2,2	2,5	
Polska	3,3	3,4	3,9	
Średnioroczna stopa inwestycji i średnioroczna stopa zatrudnienia (1998-2006) w [%]				
	Stopa inwestycji		Stopa zatrudnienia	
UE 25	20,04		Okolo 1,00	
Polska	18,00		-0,40	
Produktywność pracy = PKB w PPS na 1. zatrudnionego [%] , dane przybliżone				
	1998			2007
UE 15	118			116
UE 25	102			101
UE 27	100			100
Polska	45			60

Źródło: Wybrane fragmenty tabel i danych z wykresów publikacji internetowej¹³ PKB na mieszkańca wg PPS: Użycie Standardów Zdolności Nabywczej (*Purchasing Power. Standards*) eliminuje różnice w cenach w poszczególnych krajach, ... (ec.europa.eu/consumers/cons_info/cons_diary2007.../cp_pl.pdf).

Z punktu widzenia modelowania ekonometrycznego interesujące jest zbadanie zależności pomiędzy stopą inwestycji a dynamiką realnego PKB. Według tabeli 1.4 średnioroczna stopa inwestycji zbliżona jest do średniej w UE, a średnioroczny wzrost gospodarczy mierzony dynamiką realnego PKB jest około dwa razy szybszy niż w UE. Warto tu zaznaczyć, że dla inwestycji rzeczowych efekt wpływu na wzrost gospodarczy występuje nie tylko w okresie inwestowania, ale także podczas eksploatacji określonych obiektów. Celowe jest zatem przyjęcie dłuższego przedziału uśrednienia stopy inwestycji, przykładowo dziesięcioletniego. Jeśli bowiem wysoka stopa inwestycji wpływa na wysoki poziom PKB, świadczy to o lepszym wykorzystaniu zdobytego w wyniku inwestycji potencjału gospodarczego.

Kryterium efektywności przedsięwzięć inwestycyjnych firm prywatnych jest maksymalizacja dochodu zainwestowanego kapitału. Należy przeprowadzać analizę kosztów i korzyści. Pomocny w ocenie wpływu inwestycji na gospodarkę jest współczynnik kapitałochłonności wyrażający stosunek inwestycji do przyrostu dochodu narodowego:

$$K = \frac{I}{\Delta Y}$$

Odwrótność współczynnika K jest ogólną miarą efektywności inwestycji. Zatem wzrost PKB uzależniony jest od oddanych do użytku inwestycji oraz od ich efektywności.

Polska wykorzystując środki z UE przyczynia się do wzrostu infrastruktury inwestycji publicznych, stanowiących udział środków publicznych na inwestycje w stosunku do PKB, a efekty tych działań będą zauważone w najbliższej przyszłości. Korzyścią społeczną wdrożonej inwestycji w gospodarce może być wzrost zatrudnienia, występuje bowiem istotna korelacja między powstaniem nowych miejsc pracy a wzrostem gospodarczym mierzonym PKB. Należy jednak pamiętać, że przy pewnym poziomie inwestycji może wystąpić bez powiększenia zatrudnienia wzrost gospodarczy, ale dotyczy to inwestycji modernizacyjnych. Inwestycje tego typu przyczyniają się przede wszystkim do

¹³ Szymczak A., *Próba oceny makroekonomicznej efektywności inwestycji w wybranych gospodarkach UE*, op. cit.

wzrostu produktywności produkcji. Korzyść jaką jest przyrost produktywności pracy może prowadzić do: wzrostu zysku, obniżenia cen produktów, wzrostu płacy, skrócenia czasu pracy.

W Polsce w porównaniu lat 1998 oraz 2007 nastąpił 15% wzrost produktywności pracy, lecz nie towarzyszył temu wzrost zatrudnienia. Doprowadzić do może do wzrostu konkurencyjności naszej gospodarki na rynku międzynarodowym, lecz utrzyma się bezrobocie.

Przy uwolnionym kursie walutowym wzrost produktu krajowego brutto powoduje wzrost kursu waluty krajowej w stosunku do walut obcych. Jeśli kurs jest sztywny to istotny wzrost PKB danego kraju w stosunku do innych gospodarek stanowi presję na rewaluację kursu. Wzrost PKB oznacza zazwyczaj dobry stan gospodarki, następuje przyływ inwestycji zagranicznych i wzrost eksportu co powoduje zwiększenie popytu na walutę narodową ze strony zagranicy, czego efektem jest wzrost jej kursu.

1.4. Efektywność gospodarowania województw¹⁴

Poziom gospodarowania województwa można mierzyć poprzez wielkość produktu krajowego brutto, przy czym poziom rozwoju gospodarczego przez *PKB per capita*. Ta ostatnia wielkość nie uwzględnia w pełni poziomu jakości życia społeczeństwa. PKB umożliwia bowiem wstępną analizę gospodarności danego województwa. Ale jak określić aktywność oraz przedsiębiorczość województw i od jakich czynników one zależą. Aneta Becker proponuje w tym względzie nieparametryczną metodę granicznej analizy danych DEA do oceny efektywności gospodarowania województw. Metoda ta należy do programowania liniowego i pozwala utworzyć krzywą efektywności danej zbiorowości na podstawie jej najbardziej efektywnych jednostek. Stosowane metody oceny efektywności opierają się na podejściu:

- wskaźnikowym (budowanie relacji pomiędzy różnymi wielkościami) i porównanie z przyjętymi bazami odniesienia,
- parametrycznym, sprowadzającym się do oceny efektywności funkcji produkcji,
- nieparametrycznym, nie uwzględniającym wpływu czynnika losowego, potencjalnych błędów pomiaru, analizy zależności pomiędzy nakładami i wynikami (np. metoda DEA).

W DEA zastosowano następujący model efektywności nazwany efektywnością techniczną: przy czym μ_r, ν_i są to wagi określające ważność odpowiednio efektów/nakładów.

$$\text{Efektywność} = \frac{\sum_{r=1}^s \mu_r \text{Efekt}}{\sum_{i=1}^m \nu_i \text{Nakład}}$$

W trakcie obliczeń są wyszukiwane wagi maksymalizujące efektywność każdego obiektu. Rozwiązanie problemu optymalizacyjnego polega na wyznaczeniu odległości rzeczywistego poziomu produkcji obiektu od teoretycznie stwierdzonej granicy efektywności. Obiekty uważa się za efektywne technicznie, jeżeli znajdują się na krzywej efektywności. Przedmiotem analizy jest produktywność. Klasyfikując modele DEA stosuje się jednocześnie dwa kryteria: rodzaj efektów skali oraz orientację modelu. Pierwsze kryterium wskazuje, czy minimalizowane są nakłady, czy maksymalizowane efekty. Drugie kryterium określa, jakie założenia dotyczące efektów skali zostały przyjęte w modelu (zmienne, stałe czy nierosnące).

Materiał empiryczny przyjęty do badań przeprowadzonych przez autorkę pochodzi z roczników statystycznych GUS-u i dotyczy 2005 roku¹⁵. Zmiennymi obejmującymi podstawowe cechy makroekonomiczne są:

¹⁴ Internet, hasło: efektywność PKB, Becker Aneta: *Analiza efektywności gospodarowania województw Polski*.

¹⁵ Ibidem.

- X_1 - produkt krajowy brutto [zł/1 mieszkańca],
 X_2 - przeciętne zatrudnienie [osoby],
 X_3 - stopa bezrobocia [%],
 X_4 - nakłady inwestycyjne [mln zł],
 X_5 - nakłady na działalność innowacyjną w przemyśle [mln zł],
 X_6 - globalna produkcja rolnicza [%],
 X_7 - produkcja sprzedana przemysłu [mln zł],
 X_8 - przeciętne wynagrodzenie miesięczne [zł],
 X_9 - inflacja (CPI - wskaźnik cen towarów i usług konsumpcyjnych).

A teraz zrobimy porównanie współczynników zmienności V_i otrzymane dla poszczególnych cech z arbitralnie przyjętą wartością krytyczną $V^* = 0,2$. Zmiennymi spełniającymi nierówność $V_i \leq V^*$ są: przeciętne wynagrodzenie miesięczne (X_8), inflacja (X_9). Zmienne te zostały wyeliminowane z dalszego badania przeprowadzonego dla próby 16. województw.

Dokonano klasyfikacji zmiennych. Rolę efektu pełni produkt krajowy brutto (X_1). Niepożądanym efektem była stopa bezrobocia (X_3), którą do modelu wprowadzono w formie $1/y$, gdyż odwrotność tego typu efektu jest traktowana jako nakład. Zmienne X_2 , X_4 - X_7 potraktowano jako nakłady. Zastosowano model CCR (stałych efektów skali) zakładając maksymalizację wyniku przy zachowaniu niezmiennych nakładów. Badania przeprowadzono w programie *DEAP Version 2.1.18* (<http://www.une.edu.au/economics/emet/cepa.htm>).

Tab. 1.5. Efektywność województw według DEA

Lp.	Nazwa	Efektywność
1	Dolnośląskie	0,890
2	Kujawsko-pomorskie	0,939
3	Lubelskie	0,625
4	Lubuskie	1,000
5	Łódzkie	0,670
6	Małopolskie	0,535
7	Mazowieckie	0,975
8	Opolskie	1,000
9	Podkarpackie	0,606
10	Podlaskie	1,000
11	Pomorskie	0,860
12	Śląskie	0,849
13	Świętokrzyskie	0,874
14	Warmińsko-mazurskie	0,844
15	Wielkopolskie	0,658
16	Zachodniopomorskie	1,000
	Średnia:	0,833

Źródło: http://www.ae.katowice.pl/images/user/File/katedra_ekonomii/A.BeckerANALIZA_EFEKTYWNOSCI_GO-SPODAROWANIA_WOJEWODZTW_POLSKI.pdf

Bazując na powyższej tabeli przeprowadzono analizę wyników, przy czym współczynnik = 1 odpowiada efektywności, wynik < 1 oznacza nieefektywność techniczną. Województwa, które uzyskały współczynnik efektywności równy 1, leżą na tzw. krzywej efektywności. Natomiast przykładowo w województwie dolnośląskim PKB powinno być wyższe od rzeczywistego (w 2005 r.) o 11%. Następnie dokonano obliczeń modelu ukierunkowanego na nakłady (przy niezminionej wartości PKB) i założeniu minimalizacji nakładów oraz zachowaniu niezmiennego wyniku.

Tab. 1.6. Wielkości przekroczonych nakładów w województwach w procentach według DEA

Lp.	Nazwa	Inwestycje	Innowacje	Rolnictwo	Przemysł	Zatrudnienie	Bezrobocie
1	Dolnośląskie	53,48	64,02	0,75	56,98	44,30	1,62
2	Kujawsko-pomorskie	16,87	58,03	3,64	36,94	24,96	4,02
3	Lubelskie	38,91	71,36	5,57	37,48	54,75	9,32
4	Łódzkie	55,90	32,96	4,91	52,67	52,08	7,10
5	Małopolskie	64,50	77,24	2,28	66,62	66,75	13,23
6	Mazowieckie	71,80	82,71	7,08	74,84	67,94	0,91
7	Podkarpackie	42,68	75,21	1,42	46,02	50,87	14,83
8	Pomorskie	42,73	68,26	0,52	54,89	42,52	4,06
9	Śląskie	74,37	93,87	0,47	86,25	78,93	5,33
10	Świętokrzyskie	12,56	60,18	1,22	12,56	23,66	5,72
11	Warmińsko-mazurskie	23,44	45,22	2,71	15,58	28,91	2,21
12	Wielkopolskie	56,75	70,97	12,99	68,91	54,74	7,14

Źródło: Jak dla tabeli 1.5.

Gdzie:

inwestycje – nakłady inwestycyjne,
 innowacje – nakłady na działalność innowacyjną w przemyśle,
 rolnictwo – globalna produkcja rolnictwa,
 przemysł – produkcja sprzedana przemysłu,
 zatrudnienie – przeciętne zatrudnienie,
 bezrobocie – stopa bezrobocia.

Estymacja miar w metodzie DEA nie ma charakteru statystycznego. DEA koncentruje się na zależnościach technicznych i z tego względu stosowana jest raczej do mierzenia produktywności. Pozwala jednak wyłonić wzorcowe obiekty w określonym roku z punktu widzenia efektywności nakładów.

1.5. Efektywność sektora publicznego¹⁶

Jeśli występuje kryzys to jest on impulsem do zrewidowania stanu finansów państwa oraz wydajności sektora publicznego. Z tego też względu istotna ocena efektywności gospodarowania środkami publicznymi, a w tym struktury i kierunków wydatków. Zarówno dochody jak i wydatki publiczne, w relacji do PKB są w naszym kraju większe niż w innych krajach nowej UE. Istotne są wydatki w edukację, które w Polsce wynoszą 5,7% rocznego PKB. Natomiast w UE-27 średnio 5,29%. Zdaniem autorów publikacji występuje mimo wysokich wydatków niski poziom kształcenia, czego powodem może być nieefektywny system finansowania. Niska jest aktywność zawodowa Polaków. W 2007 roku wiek produkcyjny osiągnęło 64,4% ogółu społeczeństwa, przy czym zaledwie 55,2% spośród nich świadczyło pracę. Występuje niski wskaźnik zatrudnienia osób powyżej 55-go roku życia (w Polsce 28%, UE-15 w przedziale 45-85%).

Wydatki publiczne na tzw. wczesną dezaktywizację zawodową w 2006 roku wyniosły w 4,7% PKB, a na ochronę zdrowia 4,6% PKB. Wysokie są dotacje na świadczenia socjalne serwowane przez KRUS. W rolnictwie pracuje 15,4% pracujących, a efekt ich pracy jest 5,23% PKB.

Sytuację w sferze wynagrodzeń wpływa negatywnie na sektor publiczny. Pracownicy sektora prywatnego, finansujący wydatki publiczne zarabiają mniej, przykładowo w roku 2008 w sektorze publicznym zarabiano o 16% więcej niż średnio w kraju. Przedsiębiorstwa przemysłowe stanowiące własność sektora publicznego są dwukrotnie mniej wydajne niż przedsiębiorstwa sektora prywatnego. Natomiast w 2008 roku ich pracownicy otrzymywali o 37% wyższe wynagrodzenia niż w sektorze prywatnym. Trzeba stwierdzić, że opóźniona prywatyzacja jest kosztowna.

¹⁶ Internet hasło: efektywność PKB, Pomianek T., Skica T., *Efektywność sektora publicznego w Polsce. Próba oceny.*

1.6. Efektywność przedsiębiorstw przemysłowych i rolnych

Miarą struktury polskiej gospodarki w okresie transformacji jest wskaźnik struktury udziału w tworzeniu wartości dodanej PKB i poziomu zatrudnienia. W strukturze tej wyodrębniono usługi, przemysł i budownictwo oraz rolnictwo i leśnictwo¹⁷. Autor publikacji analizuje okres 1994-2000.

Tab. 1.7. Struktura tworzenia wartości dodanej PKB

Wyszczególnienie	1994	1995	1996	1997	1998	1999	2000	2001	2002
Przemysł i budownictwo	33,1	31,6	30,1	29,3	27,6	27,1	26,7	26,4	26,1
Handel i usługi	59,4	61,5	63,5	65,2	67,7	69,0	69,6	69,9	70,7
Rolnictwo, leśnictwo, łowiectwo i rybołówstwo	7,5	6,9	6,4	5,4	4,7	3,9	3,7	3,7	3,2

Źródło: Internet http://www.jard.edu.pl/pub/13_2_2003.pdf.

Na podstawie tabeli 1.7 widzimy wyraźne zwiększenie się udziału handlu i usług kosztem innych sektorów gospodarki. Ze szczegółowej analizy rynku przeprowadzonej przez Witolda Wielickiego wynika że jest to następstwem rozwoju rynku w zakresie pośrednictwa finansowego oraz w obsłudze nieruchomości i firm. Następuje też zmiana struktury zatrudnienia na rzecz handlu i usług. Udział ten przykładowo w Wielkiej Brytanii w 1999 roku wynosił 71%. Udział rolnictwa, leśnictwa, łowiectwa i rybołówstwa w strukturze zatrudnienia utrzymuje się na niezmiennym poziomie. Następuje zmniejszanie się udziału zatrudnionych w przemyśle i budownictwie, co świadczy o przemieszczaniu się pracowników. Występuje istotna różnica w zakresie udziału rolnictwa w tworzeniu PKB i zatrudnieniu np. w Holandii jest 3% przez 2% zatrudnionych¹⁸.

1.7. Efektywność kapitału trwałego

Podstawowym czynnikiem pobudzającym wzrost gospodarczy w krótkim okresie jest popyt konsumpcyjny i inwestycyjny, krajowy i zagraniczny, natomiast w długim okresie dostateczna podaż i efektywność czynników wytwórczych. Czynnikiem wzrostu jest też handel zagraniczny. Tempo wzrostu (lub spadku) produktu krajowego brutto jest (w przybliżeniu) sumą wskaźników obrazujących zmiany stanu zatrudnienia oraz wydajności pracy. Jednym z najważniejszych źródeł wzrostu wydajności jest postęp techniczny oraz dynamika inwestycji¹⁹. Można wymienić cztery podstawowe czynniki wzrostu gospodarczego:

- praca (podaż pracy, dyscyplina pracy, wykształcenie, kwalifikacje motywacja);
- zasoby naturalne (ziemia, zasoby mineralne, paliwa, jakość środowiska);
- kapitał (środki wykorzystywane w procesie produkcji), który dzielimy na: rzeczowy (maszyny, fabryki, drogi), finansowy, ludzki, technologia (nauka, technika, zarządzanie, przedsiębiorczość).

Wzrost gospodarczy jest rezultatem wzrostu efektywności kapitału. Efektywność kapitału trwałego jest to ilość rzeczywistego (realnego) PKB przypadającego na jednostkę zaangażowanego kapitału trwałego. Wzrost ilości kapitału umożliwia wytworzenie większej produkcji²⁰. Elementem istotnym jest postęp technologii wytwarzania, która mówi nam ile możemy wytworzyć przy użyciu określonej ilości pracy i kapitału. Technologia obejmuje wszystko to co wpływa na wydajność pracy lub produktywność kapitału.

¹⁷ Internet, Wielicki W., *Efektywność przedsiębiorstw przemysłowych i rolnych*.

¹⁸ Ibidem.

¹⁹ Internet: pl.Wikipedia.org/wiki/.

²⁰ Internet: <http://wgrit.ae.jgora.pl/sm/forum/>.

Zaskakiwani jesteśmy nieraz informacjami o wzroście produktu krajowego brutto. Miara ta podawana przez GUS różni się zazwyczaj od wartości sygnalizowanej w statystyce europejskiej (*Eurostat*). Nasuwa się pytanie, czy sygnalizowany wzrost PKB jest objawem zaradności, przykładowo określonego województwa, czy też wynika z zabiegu obliczeniowego związanego z kursem walut lub innych czynników. Propozycje wielkości i wskaźników mogących posłużyć do oceny składników, a później całości efektywności PKB podano w tabeli 1.8.

Tab. 1.8: Wielkości i wskaźniki oceny efektywności PKB

Lp.	Cecha/Czynnik	Wielkość/Wskaźnik
1	Energochłonność gospodarki (mierzona zużyciem energii elektrycznej)	PKB liczony według parytetu siły nabywczej (PKBPPP) Zużycie energii elektrycznej na jednostkę PKB wg parytetu siły nabywczej: brutto [kWh/\$PPP] Emisja CO ₂ przez sektor elektroenergetyki [tCO ₂ /MWh]
2	Inwestycje	Stopa inwestycji (% udziału w PKB) Średnioroczna stopa inwestycji (horyzont np. 10 lat) Stopa inwestycji publicznych Średnioroczna stopa zatrudnienia Wzrost zatrudnienia Produktywność pracy (produkcji) = PKB w PPS na 1 zatrudnionego [%] Kapitałochłonność Współczynnik kapitałochłonności wyrażający stosunek inwestycji do przyrostu dochodu narodowego Odwrotność współczynnika kapitałochłonności jest ogólną miarą efektywności Dynamika realnego PKB
3	Kurs walutowy	Wzrost kursu waluty krajowej w stosunku do walut obcych
4	Gospodarność (zaradność województw)	PKB <i>per capita</i> <i>Efektywność techniczna</i> = suma ważona ważności efektów / suma ważona ważności nakładów Efekty: PKB, stopa bezrobocia w formie 1/y Nakłady: przeciętne zatrudnienie [osoby], nakłady inwestycyjne [mln zł], nakłady na działalność innowacyjną w przemyśle [mln zł], globalna produkcja rolnicza [%], produkcja sprzedana przemysłu [mln zł] Produktywność (j. w.)
5	Gospodarność sektora publicznego	Dochody jak i wydatki publiczne w relacji do PKB Porównanie sfery wynagrodzeń z sektorem prywatnym Opóźnienia w prywatyzacji
6	Gospodarność przedsiębiorstw przemysłowych i rolnych	Wskaźnik struktury udziału w tworzeniu wartości dodanej PKB i poziomu zatrudnienia
7	Gospodarowanie kapitałem trwałym	PKB realny przypadającego na jednostkę zaangażowanego kapitału trwałego Wydajność pracy Produktywność kapitału

Źródło: Opracowanie własne.

2. Propozycja określenia efektywności gospodarowania województw



2.1. Wprowadzenie

Z próbą technicznego sklasyfikowania 16. województw pod względem ich gospodarności przy zastosowaniu programu *DEAP* spotykamy się w publikacji A. Becker²¹. Użyto następujące zmienne makroekonomiczne:

- X_1 – produkt krajowy brutto [zł/1. mieszkańca],
- X_2 – przeciętne zatrudnienie [osoby],
- X_3 – stopa bezrobocia [%],
- X_4 – nakłady inwestycyjne [mln zł],
- X_5 – nakłady na działalność innowacyjną w przemyśle [mln zł],
- X_6 – globalna produkcja rolnicza [%],
- X_7 – produkcja sprzedana przemysłu [mln zł].

W programie *DEAP* stosowany jest model efektywności nazwany efektywnością techniczną: przy czym μ_r, v_i są to wagi określające ważność odpowiednio efektów/nakładów.

$$\text{Efektywność} = \frac{\sum_{r=1}^s \mu_r \text{Efekt}}{\sum_{i=1}^m v_i \text{Nakład}}$$

W trakcie realizacji powyższego programu wyszukiwane są wagi μ_r, v_i maksymalizujące efektywność każdego obiektu, przykładowo województwa. Obiekty uważa się za efektywne technicznie, jeżeli znajdują się na tzw. krzywej efektywności. We wspomnianej publikacji rolę efektu pełni produkt krajowy brutto (X_1). Niepożądanym efektem jest stopa bezrobocia (X_3), którą do modelu wprowadzono w formie $1/y$. Odwrotność tego typu efektu potraktowano jako nakład. Zmienne X_2, X_4 – X_7 uznano jako nakłady.

Program obejmuje dwie opcje odnoszące się do szeregów obserwacji zmiennych wejścia/wyjścia obiektów, z rozróżnieniem bazowania na nakładach czy efektach, a trzecia opcja wykorzystuje dane panelowe, czyli szeregi zmiennych dla obiektów w jednakowych okresach czasowych. Szersze wprowadzenie do programu *DEAP* oraz przykład zastosowania programu *DEAP* do sklasyfikowania pod względem efektywności technicznej występuje we wcześniejszych artykułach autora^{22 23}. Ze względu na możliwości szerszego zastosowania metod wykorzystywanych przez *DEAP* do innych celów niż klasyfikowanie obiektów według czynników przyjętych jako cechy gospodarności, przedstawiam teraz sposób zainstalowania tego programu na własnym komputerze.

Jak już wspomniano program *DEAP* pozyskany został z Internetu ze strony: www.une.edu.au/economics/emet/cepa.htm. Najpierw tworzony jest katalog *WinRun* na dysku twardym np. D:. Pliki wzorcowe, podręcznik oraz program ściągany jest w postaci skompresowanej z rozszerzeniem *zip* i wymaga następnie „rozpakowania”. Program *DEAP Version 2.1* bazuje na pliku wsadowym danych wejściowych oraz pliku instrukcji w formacie tekstowym. Uruchomienie programu *DEAP.EXE* następuje poprzez podwójne kliknięcie na jego ikonę, a następnie wpisanie z

²¹ Internet, hasło: efektywność + PKB, Becker A.: *Analiza efektywności gospodarowania województw Polski*

²² Wornalkiewicz W., *Miary kształtujące efektywność produktu krajowego brutto*, WSZiA Opole.

²³ Wornalkiewicz W., *Wyznaczanie efektywności technicznej obiektów programem DEAP*, WSZiA Opole.

klawiatury zbioru instrukcji. Zastosowanie *DEAP* wymaga rozpakowania - zainstalowania na dysku następujących zbiorów:

- program *DEAP.EXE*,
- plik startujący *DEAP.000*,
- pliki tekstowe danych,
- pliki tekstowe instrukcji,
- pliki wyjściowe.

Plik *DEAP.000* zapamiętuje parametry realizacji programu w celu ich użycia w późniejszym okresie. Zbiór danych jak i instrukcji zakładany jest przed uruchomieniem programu. W pakiecie programu *DEAP* występują cztery komplety różnych przykładów zdefiniowania zbiorów tekstowych bazujących na określonych zmiennych wejścia/wyjścia, a mianowicie:

eg1-dta.txt (5 obserwacji - *firm*, 1 wyjście, 2 wejścia); *eg1-ins.txt*; *eg1-out.txt*; opcja *CRS*;

eg2-dta.txt (5 obserwacji, 1 wyjście, 1 wejście); *eg2-ins.txt*, *eg2-out.txt*; opcja *VRS*;

eg3-dta.txt (jak w *eg1-dta.txt*), *eg3-ins.txt*, *eg3-out.txt*; opcja z cenami dla wartości zmiennych wejścia, opcja *CRS*;

eg4-dta.txt (panel 5 obserwacji po 3 okresy - lata), *eg4-ins.txt*, *eg4-out.txt*; opcja *CRS/VRS* z zaznaczeniem panelu – opcja *malmquist*.

Opcja *CRS* dotyczy maksymalizacji funkcji celu licznika, a *VRS* mianownika - we wzorze na efektywność (gospodarność) techniczną obiektu. Nasze zainteresowanie skierowane jest na opcję *malmquist* dotyczącą danych panelowych.

2.2. Baza danych statystycznych

Celem niniejszego opracowania jest przede wszystkim określenie efektywności technicznej województw na szerszej platformie zmiennych, jak też przestrzeni czasowej (lata 1999-2007), a ponadto przy zastosowaniu programu *Gretl* wykonanie:

- statystyki opisowej zmiennych;
- macierzy korelacji;
- weryfikacji zmiennych objaśniających przy założeniu, że zmienną objaśnianą jest produkt krajowy brutto;
- zbudowanie równania regresji liniowej wielowymiarowej.

Dane panelowe zebrano na podstawie roczników statystycznych województw z lat 2000-2008. Następujące zmienne dobrano intuicyjnie biorąc pod uwagę ich rolę jako efektu czy też nakładu²⁴:

Efekty:

X_1 – produkt krajowy brutto (ceny bieżące) na 1 mieszkańca w zł.

Dalsze zmienne wchodzące do wzoru na efektywność techniczną w postaci $1/X_i$:

X_2 – emisja pyłowych przemysłowych zanieczyszczeń powietrza z zakładów szczególnie uciążliwych dla czystości powietrza w tys. t.,

X_3 – emisja gazowych (bez dwutlenku węgla) przemysłowych zanieczyszczeń powietrza z zakładów szczególnie uciążliwych dla czystości powietrza w tys. t.,

X_4 – grunty zdewastowane i zdegradowane wymagające rekultywacji i zagospodarowania (stan w dniu 31 XII) w ha,

X_5 – odpady wytworzone (w ciągu roku, z wyłączeniem odpadów komunalnych) w tys. t.,

X_6 – ścieki przemysłowe i komunalne odprowadzone do wód lub do ziemi w hm³,

X_7 – przestępstwa stwierdzone w zakończonych postępowaniach przygotowawczych w tys.,

X_8 – stopa bezrobocia rejestrowanego (stan w dniu 31 XII) w %.

²⁴ Internet: <http://www.une.edu.au/economics/emet//cepa.htm>.

Nakłady:

X_9 – pracujący w przemyśle (stan w dniu 31 XII) w tys.

X_{10} – pracujący w budownictwie (stan w dniu 31 XII) w tys.

X_{11} – pracujący w transporcie, gospodarce magazynowej i łączności (stan w dniu 31 XII) w tys.

X_{12} – przeciętne miesięczne wynagrodzenie w przemyśle brutto w zł,

X_{13} – przeciętne miesięczne wynagrodzenie w budownictwie brutto w zł,

X_{14} – przeciętne miesięczne wynagrodzenie w transporcie, gospodarce magazynowej i łączności brutto w zł,

X_{15} – nakłady na działalność badawczą i rozwojową (ceny bieżące) w mln zł,

X_{16} – użytki rolne (stan w czerwcu) w tys. ha,

X_{17} – pozyskanie drewna (grubizny) w dekametrach sześciennych,

X_{18} – produkcja sprzedana przemysłu (ceny bieżące) w mln zł,

X_{19} – produkcja budowlano-montażowa (ceny bieżące) w mln zł.

X_{20} – nakłady inwestycyjne w przemyśle (ceny bieżące) w mln zł,

X_{21} – nakłady inwestycyjne w budownictwie (ceny bieżące) w mln zł,

X_{22} – nakłady inwestycyjne w transporcie, gospodarce magazynowej i łączności (ceny bieżące) w mln zł.

X_{23} – nakłady inwestycyjne na ochronę środowiska w mln zł.

Odmienne są przygotowania co do formatu danych wejściowych. Program *Gretl* może korzystać z zaimportowanych danych z *Excels* (z przecinkiem) ułożonych według 16. województw oraz 9. okresów czasowych (lata 1999-2007). Natomiast *DEAP* wymaga plików tekstowych (z kropką dziesiętną) wprowadzanych najlepiej z programu *NOTATNIK* posortowanych według okresów czasowych, a następnie według województw. Ze względu na obszerność danych panelowych (obiektoowo-czasowych) wszystkie szeregi danych wczytano do kolumn arkusza kalkulacyjnego *Excel*, a dopiero później sporządzono odpowiednie pliki wejściowe do *Gretl*-a czy też *DEAP*. Arkusz kalkulacyjny danych panelowych obejmuje macierz o 144 wierszach oraz 23 kolumnach (zmiennych). Fragment zestawienia danych empirycznych obejmujący 2 województwa oraz 10 zmiennych podano w tab. 2.1.

Tab. 2.1. Dane empiryczne zmiennych dla województwa łódzkiego i mazowieckiego

Lp.	Województwo	Rok	Okres	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	łódzkie	1999	1	12682	14,6	380,8	4384	5171,6	188,3	78,4	14,3	264,4	47,6
2		2000	2	14497	12,1	327,9	4544	5807,8	158,6	78,4	16,3	240,9	43,7
3		2001	3	15723	10,3	298,0	4516	5263,7	153,0	85,0	18,1	237,5	37,7
4		2002	4	17535	7,1	264,3	4440	5041,8	130,8	88,5	18,4	227,3	34,7
5		2003	5	18492	7,7	232,2	4488	4881,6	134,5	93,3	20,7	223,5	32,8
6		2004	6	19498	7,4	230,2	4484	4884,3	124,0	93,0	19,5	226,3	35,5
7		2005	7	22274	7,9	220,8	4558	4449,5	127,7	86,7	17,9	223,5	37,3
8		2006	8	23666	7,2	201,1	4614	4120,7	127,4	79,6	14,7	230,0	41,0
9	18876	2007	9	25521	7,8	180,8	4635	3895,2	129,4	73,2	11,2	241,2	46,7
10	mazowieckie	1999	1	20920	14,2	206,5	5100	3692,3	2130,2	182,4	9,5	407,4	151,6
11		2000	2	23760	13,9	188,3	5061	4383,5	2165,2	182,4	10,8	389,1	140,8
12		2001	3	26871	13,7	201,8	5244	5059,9	2128,9	195,6	13,0	367,3	136,3
13		2002	4	30283	12,0	189,0	5238	4916,2	2185,1	194,1	13,8	354,0	119,5
14		2003	5	31115	13,4	210,4	5296	5917,7	2363,7	205,6	15,4	345,5	106,8
15		2004	6	32722	12,5	208,8	4344	5842,3	2458,4	208,6	14,7	347,8	101,2
16		2005	7	36636	11,3	217,1	4273	5855,9	2524,2	206,0	13,8	336,4	105,4
17		2006	8	40817	11,3	232,2	4248	5773,3	2745,6	189,1	11,8	348,8	115,9
18	31945	2007	9	44381	9,9	196,2	4218	6788,4	2700,6	159,9	9,0	362,3	129,3

Źródło: Opracowanie własne na podstawie roczników statystycznych województw.

Dla każdego z 16. województw obliczono wartość średnią zmiennej $X1$:

- 1 - łódzkie,
- 2 - mazowieckie,
- 3 - małopolskie,
- 4 - śląskie,
- 5 - lubelskie,
- 6 - podkarpackie,
- 7 - podlaskie,
- 8 - świętokrzyskie,
- 9 - lubuskie,
- 10 - wielkopolskie,
- 11 - zachodniopomorskie,
- 12 - dolnośląskie,
- 13 - opolskie,
- 14 - kujawsko-pomorskie,
- 15 - pomorskie,
- 16 - warmińsko-mazurskie.

2.3. Sformatowanie plików wejściowych

Procedura postępowania w zakresie przeniesienia danych statystycznych z arkusza kalkulacyjnego *Excel* jako bazy danych użytkownika do programu *Gretl* (bezpłatnie dostępnego w Internecie) jest następująca:

1. Skopiowanie wcześniej wprowadzonych manualnie danych na wyodrębniony skoroszyt o jednym arkuszu nadając nazwę np. *Gretl.xls*. Rozpoczynamy od komórki A1 wpisując w pierwszym wierszu nazwy pól kolejno: W (symbol liczbowy województwa), T (symbol cyfrowy kolejnych okresów – lat 1999–2007), zmienne $X1$ – $X23$, n (numer obserwacji od 1–144). Cyfra 144 stanowi iloczyn ilości obiektów – 16 (województw) i 9 okresów.

2. Wywołanie programu *Gretl*.

3. Pobranie pliku *Gretl.xls* poprzez polecenie: *Plik/Otwórz dane/Import/Excel*. Wskazanie pliku w podkatalogu np. *D:/MODELE-P-S*, a następnie polecenie: *Otwórz*.

4. Podanie w oknie dialogowym początku zakresu importu: *Kolumna 1 wiersz 1 Arkusz 1 OK*. Pojawia się informacja: *Importowane dane są traktowane jako niedatowane (dane przekrojowe). Czy chcesz interpretować dane jako szereg czasowy lub dane panelowe? Tak*

Okno *Kreator struktury danych* Wybieramy *Dane panelowe Do przodu*.

Okno *Kreator struktury danych panelowych* Wybieramy *Stos szeregów czasowych Do przodu*

Okno kreator struktury danych: liczba przekrojów (obiektów) w jednostce czasu – 16, liczba jednostek czasu w przekroju – 9 Do przodu.

Okno *Potwierdź strukturę danych*: *Dane panelowe (stos szeregów czasowych) 16 jednostek przekroju w 9 okresach. Zastosuj*.

Powstał zbiór *Gretl.xls** o strukturze:

$$0 \mid \text{const} \mid 1 \mid W \mid 2 \mid T \mid 3 \mid X1 \mid \dots \mid 25 \mid X23 \mid 26 \mid n$$

Program przeniósł wyróżnienia województw W , okresy T , wartości zmiennych $X1$ – $X23$ oraz n dodając numery obserwacji jako 1.1 - 16.9. Fragment zaimportowanych danych przez *Gretl* dla województwa łódzkiego i trzech zmiennych jest następujący:

Obs	W	T	X1	X2	X3
1:1	1	1	12682	14,5752	380,7771
1:2	1	2	14497	12,1000	327,9000
1:3	1	3	15723	10,3000	298,0000
1:4	1	4	17535	7,1000	264,3000
1:5	1	5	18492	7,7000	232,2000
1:6	1	6	19498	7,4000	230,2000
1:7	1	7	22274	7,9000	220,8000
1:8	1	8	23666	7,2000	201,1000
1:9	1	9	25521	7,8000	180,8000

Następnie skopiowano zaimportowany plik danych *Gretl.xls** jako plik *Gretl-a* pod nazwą *Gretl.gdt*, który posłuży nam między innymi do estymacji parametrów modelu ekonometrycznego.

Kolejnym krokiem było doprowadzenie danych pliku *Gretl.xls* do postaci tekstowej wymaganej przez *DEAP*, tak jakbyśmy wprowadzali dane bezpośrednio do programu *NOTATNIK*. Jak zaznaczono we wprowadzeniu efektem jest zmienna X_1 oraz zmienne wchodzące do wzoru na efektywność techniczną w postaci $1/X_i$: $X_2, X_3, X_4, X_5, X_6, X_7, X_8$. Nakłady stanowią zmienne: $X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{14}, X_{15}, X_{16}, X_{17}, X_{18}, X_{19}, X_{20}, X_{21}, X_{22}, X_{23}$. Procedura postępowania związana z przygotowaniem i importem pliku z *Excelsa* do podkatalogu programu *DEAP* jest następująca:

- skopiowanie skoroszytu *Gretl.xls* na *DEAP.xls*,
- pozostawienie tylko jednego arkusza *Gretl-a* z danymi,
- wprowadzenie dodatkowych kolumn dla obliczenia zmiennych o efekcie $1/x_i$, a mianowicie: $X2E, X3E, X4E, X5E, X6E, X7E, X8E$,
- obliczenie odwrotności zmiennych $X2-X8$ z dokładnością do 5. miejsc po przecinku,
- skopiowanie tylko wartości powyższego szeregu,
- usunięcie zbędnych kolumn oraz wiersza nagłówkowego,
- zamiana wartości numerycznych na tekstowe poprzez zaznaczenie danych i wywołanie menu *Format/komórka/tekstowe*
- automatyczna zamiana przecinka w wartościach danych zmiennych na kropkę dziesiętną korzystając z menu *Edycja/zamień*,
- sprawdzenie czy kolumny zmiennych są w kolejności najpierw dotyczące efektów później nakładów,
- posortowanie narastające zbioru w układzie najpierw okres T , a później numer województwa W ,
- usunięcie kolumny obiektów – województw oraz kolumny okresów,
- zapisanie tak przygotowanego arkusza jako *DEAP2.xls*,
- zapisanie zbioru *Excelsa* jako tekstowego z tabulacją kolumn w podkatalogu *D:\WinRun\DEAP-xp1\EAP5-dta.txt* nadając mu kolejną nazwę typu *EAPn* w stosunku do istniejących już plików.

Dalej widzimy fragment zapisanego zbioru *EAP5-dta.txt* dla okresu 1. i 16. województw w zakresie w kolejności od lewej zmiennych $X1, X2E, X3E, X4E$ wywołanego programem *NOTATNIK*.

12682	0.0686096931774521	0.00262620835129003	0.000228102189781022
20920	0.0702286645317153	0.00484335617460105	0.000196078431372549
13029	0.0600297747682851	0.00507944248039335	0.000275862068965517
16014	0.0232401403704478	0.00169459356867849	0.000159872102318145
10383	0.0995460699211595	0.0221213488713688	0.000261028452101279
10866	0.185949644836178	0.0309916074726964	0.000240789790512882
10928	0.247770069375619	0.0619425173439049	0.000361663652802893
11056	0.122393028493097	0.0127873313350997	0.000337837837837838
13062	0.143020594965675	0.0376369986751776	0.000714796283059328
15141	0.0670555890833501	0.00523871789713673	0.000095047999239616

13980	0.145547695979973	0.0103962639985695	0.00029171528588098
14290	0.0501303388810908	0.00911460706928924	0.000119731800766
12635	0.0965884943785496	0.018639884880071	0.000251130085384229
13193	0.0695603784084586	0.0126473415288106	0.00022386389075
14129	0.136664297818838	0.01822190637584	0.000363240101
10986	0.137723973612087	0.0590245601194657	0.000207856994387861

2.4. Charakterystyka zmiennych

Program *Gretl* umożliwia szybkie otrzymywanie statystyki opisowej w odniesieniu do zmiennych X_1 - X_{23} w ujęciu panelowym, a w ramach niej określenie:

- średniej;
- mediany;
- wartości minimalnej oraz maksymalnej (*Min*, *Max*);
- odchylenia standardowego (*s*);
- współczynnika zmienności (*V*);
- współczynnik skośności (*S*) i współczynnik kurtozy (*K*).

Informacje te podaje poniższy raport z *Gretl*-a dla 144 (1.1-16.9) obserwacji próby danych panelowych.

Statystyki opisowe, wykorzystane obserwacje 1:1 - 16:9

X_i	<i>Średnia</i>	<i>Mediana</i>	<i>Min</i>	<i>Max</i>
X1	19144	18360	10383	44381
X2	8,6977	6,9960	1,4000	43,029
X3	126,90	64,800	9,5000	754,10
X4	4280,4	3641,0	1234,0	10852
X5	7712,1	3379,3	399,20	49038
X6	577,44	197,60	36,000	2745,6
X7	83,875	71,250	26,700	208,60
X8	17,578	16,750	7,8000	30,300
X9	188,85	159,95	58,600	600,04
X10	46,337	33,700	11,000	367,26
X11	46,366	36,912	16,200	152,00
X12	2079,5	2010,4	1490,0	3230,7
X13	1731,0	1663,6	1305,0	3042,3
X14	2182,6	2131,4	1570,2	3424,9
X15	323,48	145,05	12,700	2742,3
X16	1061,2	1012,6	457,90	2419,3
X17	1775,8	1706,7	688,30	3899,7
X18	38943	24888	7463,3	1,8165E+05
X19	4927,1	3197,1	1159,0	23230
X20	2716,3	1905,0	558,40	11999
X21	687,52	486,03	76,568	3270,5
X22	872,13	432,16	134,00	7798,8
X23	399,44	318,64	86,545	1993,7

X_i	<i>s</i>	<i>V</i>	<i>S</i>	<i>K</i>
X1	5766,8	0,30124	1,3310	2,9203
X2	7,0371	0,80908	2,2147	6,2462
X3	159,48	1,2568	2,4567	5,9287

X4	2177,9	0,50882	1,5628	2,2389
X5	12180	1,5793	2,2257	3,4764
X6	713,65	1,2359	1,3893	0,71205
X7	48,629	0,57978	1,0088	0,13337
X8	4,8636	0,27669	0,49439	-0,19435
X9	121,56	0,64368	1,2253	0,92177
X10	41,091	0,88678	3,8767	24,570
X11	31,360	0,67637	1,6441	2,2576
X12	358,65	0,17247	0,89701	0,62085
X13	261,10	0,15084	1,8483	5,0775
X14	309,06	0,14160	1,0880	2,1168
X15	526,09	1,6263	3,1145	9,1345
X16	474,79	0,44739	1,0554	0,58792
X17	751,39	0,42314	0,61059	-0,40331
X18	36224	0,93019	1,9042	3,1951
X19	4333,9	0,87961	1,9968	3,9843
X20	2095,9	0,77162	1,4636	2,1660
X21	642,22	0,93412	1,9280	3,7895
X22	1425,7	1,6347	3,4583	11,378
X23	307,01	0,76860	2,1417	6,4796

Źródło: Opracowanie własne programem *Gretl*.

Wyznaczenie procentowe współczynnika zmienności V_j na podstawie powyższego raportu stanowi wstępny krok zbadania wystarczającego zróżnicowania danej j -tej zmiennej objaśniającej. Określany jest w procentach i stanowi moduł z ilorazu odchylenia standardowego i średniego poziomu wartości rozpatrywanej zmiennej:

$$V_j = \frac{\left| \frac{S_j}{\bar{x}_j} \right|}{1} * 100$$

Z raportu *Gretl*-a dotyczącego statystyki opisowej wynika, że dla 23. zmiennych $X1$ - $X23$ współczynnik zmienności przekracza założoną wartość krytyczną $V^* = 10\%$, czyli zmienne te wstępnie wchodzi do modelu ekonometrycznego. Zdaniem A. Becker bazując na tym kryterium wszystkie zmienne mogą wejść również do badania efektywności technicznej gospodarowania województw²⁵.

2.5. Macierz korelacji

Kryterium wstępne według współczynnika zmienności nie pozwoliło nam na selekcję zmiennych do budowy modelu ekonometrycznego. Posłużmy się więc analizą współczynników korelacji. W tym celu przyjmijmy zmienną $X1$ jako zmienną objaśnianą, a pozostałe zmienne $X2$ - $X23$ jako objaśniające. Program *Gretl* umożliwia uzyskanie macierzy korelacji na podstawie zaimportowanych danych tj. pliku *Gretl.gdt*. Procedura w tym zakresie jest następująca: *Plik/Otwórz dane/Pliki użytkownika* lub bezpośrednio wskazanie pliku założonej bazy danych *Gretl.gdt* wywołując menu:

Dane/Zaznacz wszystko; Widok/Macierz korelacji.

Współczynniki korelacji liniowej uzyskujemy dla próby 1-114 obserwacji szeregu danych panelowych. *Gretl* podaje nam wartość krytyczną $r^* = 0,1637$ przy dwustronnym 5% obszarze krytycznym dla $n = 144$. Wartość tą możemy obliczyć także ze wzoru:

²⁵ Internet, hasło: efektywność + PKB, Becker A.: *Analiza efektywności gospodarowania województw Polski*.

$$r^* = \sqrt{\frac{t_\alpha^2}{t_\alpha^2 + n - 2}}$$

$$r^* = \sqrt{\frac{(1,9759)^2}{(1,9759)^2 + 144 - 2}} \approx 0,1637$$

przy czym:

$t_\alpha = 1,9759$ - statystyka wyznaczona na podstawie rozkładu t -Studenta dla poziomu istotności $\alpha = 0,05$ oraz ilości stopni swobody $k = iss = n - 2 = 144 - 2 = 142$. Dla wyznaczenia wartości krytycznej przyjęto wartość istniejącą w tablicy $k = 150$.

Odpowiedzią programu *Grell* jest raport:

X1	X2	X3	X4	X5
1,0000	0,1298	0,3336	0,2072	0,2435 X1
1,0000	0,8315	0,4083	0,7959 X2	
1,0000	0,3127	0,7383 X3		
1,0000	0,3761 X4			
1,0000 X5				
X6	X7	X8	X9	X10
0,5130	0,5927	-0,2680	0,4790	0,4593 X1
0,2018	0,7155	-0,3236	0,8390	0,6100 X2
0,1903	0,7189	-0,3672	0,8664	0,5542 X3
0,3478	0,4427	-0,1644	0,5655	0,3709 X4
-0,0348	0,6298	-0,1681	0,6984	0,4112 X5
1,0000	0,5033	-0,2623	0,4081	0,5455 X6
1,0000	-0,3008	0,8894	0,7878 X7	
1,0000	-0,4592	-0,4572 X8		
1,0000	0,7666 X9			
1,0000 X10				
X11	X12	X13	X14	X15
0,6408	0,8701	0,8989	0,8007	0,6779 X1
0,6484	0,2751	0,0566	0,0010	0,2954 X2
0,6409	0,4830	0,1986	0,1328	0,3160 X3
0,3861	0,1482	0,1597	0,1248	0,1800 X4
0,4828	0,5133	0,1515	0,0838	0,1188 X5
0,6403	0,3050	0,4966	0,5062	0,6837 X6
0,9333	0,5981	0,5250	0,5693	0,7319 X7
-0,4626	-0,3157	-0,4152	-0,0595	-0,4223 X8
0,8621	0,5226	0,4004	0,3194	0,5658 X9
0,8497	0,4431	0,4912	0,4604	0,7477 X10
1,0000	0,5835	0,6372	0,5905	0,8751 X11
1,0000	0,8326	0,6852	0,5203 X12	
1,0000	0,8062	0,7435 X13		
1,0000	0,6984 X14			
1,0000 X15				
X16	X17	X18	X19	X20
0,2852	0,2382	0,7944	0,7785	0,6866 X1
0,0417	-0,2618	0,5628	0,4992	0,6631 X2
-0,0400	-0,2290	0,7100	0,5623	0,7693 X3
0,4949	0,2257	0,3944	0,3644	0,5177 X4

-0,1791	-0,0212	0,5168	0,4009	0,6840 X5
0,5974	0,1522	0,5588	0,6647	0,4311 X6
0,3872	0,0048	0,8777	0,8551	0,8433 X7
-0,2771	0,3938	-0,4504	-0,5000	-0,4838 X8
0,2917	-0,0607	0,8496	0,7843	0,9060 X9
0,4728	-0,0732	0,7406	0,7885	0,7198 X10
0,5301	-0,0176	0,9161	0,9486	0,8458 X11
0,0409	0,1121	0,7700	0,7003	0,7223 X12
0,4066	0,1623	0,7473	0,7888	0,6200 X13
0,4166	0,3287	0,6448	0,6933	0,4693 X14
0,6364	-0,0636	0,7968	0,9017	0,6316 X15
1,0000	0,1290	0,3880	0,5213	0,2905 X16
1,0000	0,0332	0,0576	0,0555 X17	
1,0000	0,9438	0,9139 X18		
1,0000	0,8649 X19			
1,0000 X20				
X21	X22	X23		
0,7361	0,6230	0,4635 X1		
0,3074	0,2986	0,7179 X2		
0,3896	0,2753	0,7008 X3		
0,3802	0,1796	0,4367 X4		
0,3009	0,1087	0,5981 X5		
0,5551	0,6707	0,4448 X6		
0,6916	0,6931	0,7746 X7		
-0,5092	-0,3860	-0,4696 X8		
0,6468	0,5411	0,8574 X9		
0,7027	0,7885	0,7585 X10		
0,7786	0,8566	0,8114 X11		
0,6686	0,4660	0,4843 X12		
0,7690	0,7003	0,4146 X13		
0,6575	0,6635	0,3311 X14		
0,7732	0,9601	0,5936 X15		
0,4927	0,6485	0,2999 X16		
0,1319	-0,0212	-0,0222 X17		
0,7937	0,7438	0,7547 X18		
0,8792	0,8662	0,7715 X19		
0,7809	0,5899	0,8544 X20		
1,0000	0,7575	0,6989 X21		
1,0000	0,5919 X22			
1,0000 X23				

Postępowanie w zakresie wyboru zmiennych objaśniających jest następujące²⁶:

a) Sprawdzamy, czy $|r_j| \leq r^*$ i eliminujemy zmienne X_j nieistotnie skorelowane z $Y = X1$: Kierując się tym kryterium i przeglądając w macierzy korelacji pary relacji zmiennej objaśnianej $X1$

$$|r_j| \leq r^* = 0,1637$$

²⁶ Dziechciarz J., *Ekometria Metody, przykłady, zadania*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2003.

względem odpowiadających zmiennych objaśniających $X2$ - $X23$ możemy wyeliminować tylko zmienną $X2$.

b) Z pozostałych 21 potencjalnych zmiennych objaśniających $X3$ - $X23$ wybieramy zmienną X_h wg kryterium:

$$r_h = \max_j (r_j)$$

Zmienną najsilniej skorelowaną ze zmienną $X1$ jest zmienna: $X13$ – przeciętne miesięczne wynagrodzenie w budownictwie brutto w zł dla której $r_{13} = 0,8989$.

c) Badamy relacje powyższej zmiennej z pozostałymi zmiennymi i eliminujemy te dla których $|r_{hi}| > r^*$ gdyż są one zbyt silnie skorelowane ze zmienną X_h . Pozostałe zmienne $X4$ oraz $X5$.

d) Wracamy do punktu b) i wybieramy zmienną $X5$ o współczynniku korelacji względem $X1$ równym 0,2435.

e) Realizując ponownie etap wg c) badamy relację zmiennej $X5$ ze zmienną $X4$. Ponieważ współczynnik korelacji równy 0,3761 $>$ 0,1637 pomijamy zmienną $X4$ w modelu.

Zatem wstępnie wg analizy współczynników korelacji liniowy model ekonometryczny ma postać:

$$\hat{X}_1 = a_0 + a_1 X_{13} + a_2 X_5$$

A teraz zobaczymy jaki model zaproponuje *Gretl* stosując metodę *a posteriori*, czyli stopniowego eliminowania określonej zmiennej kierując się testem *t*-Studenta statystycznej istotności parametrów.

2.6. Określenie modelu ekonometrycznego

Parametry liniowego modelu ekonometrycznego możemy oszacować stosując klasyczną metodę najmniejszych kwadratów (KMNK) w *Gretl*. Skorzystajmy z pliku *Gretl.gdt* wywołując menu:

Dane/Zaznacz wszystko; Model/Klasyczna metoda najmniejszych kwadratów

W oknie ze specyfikacją modelu definiujemy jako zmienną zależną (objaśnianą) $X1$ oraz określamy wstępnie zmienne niezależne (objaśniające) $X2$ - $X23$. W rezultacie postępowania *a posteriori* (16 kroków) otrzymujemy z *Gretl* raport z parametrami modelu regresji liniowej zwane współczynnikami oraz statystykę opisową modelu, którego fragment zamieszczono poniżej:

Model 16: Estymacja Panelowe KMNK z wykorzystaniem 144 obserwacji. Włączono 16 jednostek danych przekrojowych. Szereg czasowy długości = 9. Zmienna zależna: $X1$.

	Współczynnik	Błąd standardowy	t-Studenta	Wartość p	
const	-14175,2	1722,40	-8,230	1,41E-013	***
X5	-0,124107	0,0246434	-5,036	1,49E-06	***
X6	0,685282	0,288771	2,373	0,0190	**
X8	188,507	36,3130	5,191	7,51E-07	***
X10	-14,1471	5,89481	-2,400	0,0178	**
X12	7,70218	1,11035	6,937	1,52E-010	***
X13	6,94873	1,51385	4,590	1,00E-05	***
X18	0,0238954	0,0132364	1,805	0,0733	*
X20	0,827128	0,226411	3,653	0,0004	***

Średnia arytmetyczna zmiennej zależnej = 19143,6

Odchylenie standardowe zmiennej zależnej = 5766,83

Suma kwadratów reszt = 3,99299e+008

Błąd standardowy reszt = 1719,82

Współczynnik determinacji $R^2 = 0,91604$

Skorygowany współczynnik determinacji $R^2 = 0,91106$

Statystyka $F(8, 135) = 184,106$ (wartość $p < 0,00001$)

Logarytm wiarygodności = -1272,48

Kryterium informacyjne Akaike'a (AIC) = 2562,95.

Źródło: Opracowanie własne programem *Gretl*.

Widzimy zachowanie w modelu ekonometrycznym zmiennych $X5$ oraz $X13$ określonych wstępnie w wyniku analizy współczynników korelacji. W raporcie występuje też określenie błędu standardowego $D(a_i)$ oszacowania danego parametru a_i oraz obliczenie statystyki t -Studenta wg wzoru:

$$t_{obl} = \frac{a_i}{D(a_i)}$$

Program sprawdza poniższy warunek:

$$|t_{obl}| > t_\alpha$$

przy czym:

t_α - wartość krytyczna wg rozkładu t -Studenta dla poziomu istotności α . Jak już wiemy dla poziomu istotności $\alpha = 0,05$ oraz ilości stopni swobody $iss = 142$ $t_\alpha = 1,9759$.

Warto zauważyć, że *Gretl* wprowadza dogodne oznaczenia informujące o poziomie istotności, a mianowicie:

*** - $\alpha = 0,01$; ** - $\alpha = 0,05$; * - $\alpha = 0,1$; puste – brak statystycznej istotności ($\alpha > 0,1$).

Program *Gretl* wskazuje na wysoki współczynnik determinacji $R^2 = 0,91604$. Wyliminujemy jeszcze ze względu na poziom istotności oznaczony jedną gwiazdką odpowiadający kryterium t -Studenta zmienną $X18$. Po tej czynności uzyskujemy z *Gretl* następujący model:

Model 17: Estymacja Panelowe KMNK z wykorzystaniem 144 obserwacji. Włączono 16 jednostek danych przekrojowych. Szereg czasowy długości = 9. Zmienna zależna: $X1$.

Współczynnik Błąd standardowy t -Studenta Wartość p

const	-15763,3	1493,05	-10,56	2,11E-019	***
$X5$	-0,129212	0,0246830	-5,235	6,12E-07	***
$X6$	0,810494	0,282636	2,868	0,0048	***
$X8$	196,194	36,3607	5,396	2,94E-07	***
$X10$	-11,3945	5,74129	-1,985	0,0492	**
$X12$	8,05910	1,10165	7,316	2,00E-011	***
$X13$	7,37759	1,50746	4,894	2,75E-06	***
$X20$	1,09901	0,170458	6,447	1,83E-09	***

Średnia arytmetyczna zmiennej zależnej = 19143,6

OchYLENIE standardowe zmiennej zależnej = 5766,83

Suma kwadratów reszt = 4,08938e+008

Błąd standardowy reszt = 1734,04

Współczynnik determinacji $R^2 = 0,91401$

Skorygowany współczynnik determinacji $R^2 = 0,90958$

Statystyka $F(7, 136) = 206,511$ (wartość $p < 0,00001$)

Logarytm wiarygodności = -1274,19

Kryterium informacyjne Akaike'a (AIC) = 2564,39.

Możemy zatem określić model jako:

$$\hat{X}_1 = -15763,3 - 0,1292X_5 + 0,8105X_6 + 196,194X_8 - 11,3945X_{10} + 8,0591X_{12} + 7,3776X_{13} + 1,099X_{20}$$

gdzie:

- X_1 – produkt krajowy brutto (ceny bieżące) na 1 mieszkańca w zł,
- X_5 – odpady wytworzone (w ciągu roku, z wyłączeniem odpadów komunalnych) w tys. t.,
- X_6 – ścieki przemysłowe i komunalne odprowadzone do wód lub do ziemi w hm³,
- X_8 – stopa bezrobocia rejestrowanego (stan w dniu 31 XII) w %,
- X_{10} – pracujący w budownictwie (stan w dniu 31 XII) w tys.,
- X_{12} – przeciętne miesięczne wynagrodzenie w przemyśle brutto w zł,
- X_{13} – przeciętne miesięczne wynagrodzenie w budownictwie brutto w zł,
- X_{20} – nakłady inwestycyjne w przemyśle (ceny bieżące) w mln zł.

2.7. Zastosowanie programu DEAP

Plik tekstowy instrukcji tworzymy poprzez skopiowanie odpowiedniego pliku wzorcowego w naszym przypadku *eg4-ins.txt* i nadanie nowemu plikowi nazwy własnej *EAP5-ins.txt*, a następnie dokonujemy zmiany w parametrach, a mianowicie:

- kopiujemy przykładowy zbiór wyjścia *eg4-out.txt*, zapisujemy pod własną nazwą *EAP5-out.txt*,
- czyszcimy zbiór *EAP5-out.txt* tak aby był pusty,
- zmieniamy nazwę zbioru danych na *EAP5-dta.txt*,
- zmieniamy ilość (firm) województw na 16, a ilość okresów na 9,
- podajemy ilość zmiennych licznika (wyjściowych) 8, oraz zmiennych mianownika (wejściowych) 15,
- podajemy symbol 1 oznaczający zorientowanie się modelu na zmienne wyjściowe,
- określamy metodę obliczeń jako CRS, czyli bazującą na maksymalizacji wartości licznika w funkcji efektywności technicznej,
- zaznaczamy cyfrą 2, że mamy do czynienia z bazą danych panelowych *MALMQUIST-DEA*.

Lista instrukcji zapisana w pliku *EAP5-ins.txt*

<i>EAP5-dta.txt</i>	DATA FILE NAME
<i>EAP5-out.txt</i>	OUTPUT FILE NAME
16	NUMBER OF FIRMS
9	NUMBER OF TIME PERIODS
8	NUMBER OF OUTPUTS
15	NUMBER OF INPUTS
1	0=INPUT AND 1=OUTPUT ORIENTATED
0	0=CRS AND 1=VRS
2	0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA, 3=DEA(1-STAGE), 4=DEA(2-STAGE)

Źródło: Opracowanie własne z zastosowaniem pliku przykładowego w programie *DEAP wersja 2.1*.

Wywołujemy program *DEAP.EXE* zapisany podczas instalacji z Internetu w podkatalogu *D:/WinRun/DEAPA-xp1*. Po podwójnym kliknięciu na ikonie tego programu pojawia się tekst w który wpisujemy nazwę zbioru instrukcji *EAP5-ins.txt*. System zapisuje wyniki w pustym zbiorze rezultatów: *EAP5-out.txt* i naciskamy klawisz ENTER. Zbiór *EAP5-out.txt* po zapisaniu wyników jest bardzo obszerny. Zawiera raporty wskaźników efektywności odnoszące się do kolejnych z 9. lat, a w ramach nich do poszczególnych 16. województw. Ponadto dla poszczególnych lat i województw podane są wskaźniki dla okresu bieżącego *t* oraz wcześniejszego *t-1* oraz następnego *t+1* z uwzględnieniem opcji CRS. W kolumnie czwartej podano dodatkowo wartości wskaźników dla opcji VRS.

Przykład fragmentu odnoszącego się tylko do pierwszego roku (okresu) oraz 16. województw listy rezultatów podano w poniższym fragmencie raportu programu *DEAP*. Dla pierwszego zapisu

nie występuje okres wcześniejszy stąd wartości zerowe w kolumnie $t-1$. Interesuje nas przede wszystkim kolumna t . Na zakończenie podlisty dla 16. województw danego okresu przykładowo pierwszego program oblicza efektywność średnią roku, która dla okresu pierwszego wynosi: 0,999. Łącznie zatem województwa nie osiągnęły pełnej efektywności gospodarowania w roku 1999. Wpływ na to miały województwa lubelskie (5) oraz dolnośląskie (12). W tabeli współczynnik „1” odpowiada efektywności technicznej według przyjętych miar (zmiennych) w zakresie efektów i nakładów. Wynik mniejszy od jeden stosując opcję CRS oznacza bowiem nieefektywność techniczną²⁷. Fragment listy rezultatów dla roku 1999 i 16. województw (firm) jest następujący:

Results from DEAP Version 2.1 Instruction file = *EAP5-ins.txt* Data file = *EAP5-dta.txt* Output oriented *Malmquist DEA* (dane panelowe) DISTANCES SUMMARY year = 1 (1999)

firm	CRS		VRS	

	<i>t</i> -1	<i>t</i>	<i>t</i> +1	
1	0.000	1.000	1.170	1.000
2	0.000	1.000	1.175	1.000
3	0.000	1.000	1.360	1.000
4	0.000	1.000	1.351	1.000
5	0.000	0.985	1.221	1.000
6	0.000	1.000	1.197	1.000
7	0.000	1.000	1.342	1.000
8	0.000	1.000	1.466	1.000
9	0.000	1.000	2.101	1.000
10	0.000	1.000	1.321	1.000
11	0.000	1.000	1.085	1.000
12	0.000	0.992	1.020	0.998
13	0.000	1.000	1.313	1.000
14	0.000	1.000	1.049	1.000
15	0.000	1.000	1.090	1.000
16	0.000	1.000	1.586	1.000
mean	0.000	0.999	1.303	1.000

Źródło: Opracowanie własne programem *DEAP Version 2.1*. (fragment zbioru *EAP5-out.txt*).

Zainteresuje nas z pewnością okres 9, czyli rok 2007 (t) dla którego powyższy fragment listy rezultatów jest następujący:

firm	CRS		VRS	

	<i>t</i> -1	<i>t</i>	<i>t</i> +1	
1	1.193	1.000	0.000	1.000
2	1.124	1.000	0.000	1.000
3	1.389	1.000	0.000	1.000
4	1.421	1.000	0.000	1.000
5	1.062	0.928	0.000	1.000
6	1.076	1.000	0.000	1.000
7	2.268	1.000	0.000	1.000
8	1.275	1.000	0.000	1.000

²⁷ Internet, hasło: efektywność + PKB, Becker A.: *Analiza efektywności gospodarowania województw Polski*.

9	1.578	1.000	0.000	1.000
10	1.422	1.000	0.000	1.000
11	0.970	0.957	0.000	1.000
12	1.056	0.980	0.000	1.000
13	1.379	1.000	0.000	1.000
14	1.038	1.000	0.000	1.000
15	1.108	0.970	0.000	1.000
16	1.187	1.000	0.000	1.000
mean	1.284	0.990	0.000	1.000

W stosunku do roku 1999 nastąpiło łączne obniżenie efektywności technicznej 16. województw z 0,999 na 0,990. Wpływ na to miała niewystarczająca efektywność gospodarowania województw: lubelskiego (5), zachodniopomorskiego (11), dolnośląskiego (12), pomorskiego (15).

W liście rezultatów znajdują się także raporty porównawcze *MALMQUIST INDEX SUMMARY* współczynników efektywności technicznej *effch* bieżącego okresu do współczynników okresu wcześniejszego, które rozpoczynają się od okresu drugiego dla każdego województwa. Ten fragment tabeli ze względu na obszerność nie zamieszczono, natomiast w poniższej tabeli podano jedynie końcowe zestawienie sumaryczne średnich rocznych z porównań w układzie okresów 2-9 dla 16. województw. Obliczone zostały one jako średnie geometryczne ze względu na stosunkowy charakter obliczonych miar łańcuchowych. W raportach DEAP zamiast przecinka występuje kropka dziesiętna. Średnie geometryczne roczne dla okresów województw (*MALMQUIST INDEX SUMMARY OF ANNUAL MEANS*):

year	2	3	4	5	6	7	8	9	means
effch	1.000	0.997	1.001	1.001	1.002	0.992	0.993	1.005	0.999

Źródło: Opracowanie własne programem *DEAP Version 2.1*.

Oprócz zestawienia sumarycznego porównawczego według okresów w liście rezultatów znajduje się również raport w układzie poszczególnych 16. województw stanowiący średnie geometryczne dla województw (*MALMQUIST INDEX SUMMARY OF FIRM MEANS*):

firm	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	mean
effch	1.000	1.000	1.000	1.000	0.993	1.000	1.000	1.000	1.000	1.000	0.995	0.998	1.000	1.000	0.996	1.000	0.999

Źródło: Opracowanie własne programem *DEAP Version 2.1*.

Zatem zarówno dla wszystkich okresów, jak też województw występuje nieznaczna efektywność gospodarowania (< 1) biorą pod uwagę założone kryteria 9. zmiennych stanowiących efekty jak też 15 zmiennych przyjętych jako nakłady. Wróćmy jednak do zmiennych opracowanego modelu ekonometrycznego i zastosujmy dla nich model efektywności technicznej w programie *DEAP*, przy czym jako efekty występują zmienne:

X_1 – produkt krajowy brutto (ceny bieżące) na 1 mieszkańca w zł,

X_5 – odpady wytworzone (w ciągu roku, z wyłączeniem odpadów komunalnych) w tys. t.,

X_6 – ścieki przemysłowe i komunalne odprowadzone do wód lub do ziemi w hm³,

X_8 – stopa bezrobocia rejestrowanego (stan w dniu 31 XII) w %,

przy czym zmienne X_5 , X_6 , X_8 wejdą do wzoru efektywności jako $1/x$.

Nakładami są:

X_{10} – pracujący w budownictwie (stan w dniu 31 XII) w tys.,

X_{12} – przeciętne miesięczne wynagrodzenie w przemyśle brutto w zł,

X_{13} – przeciętne miesięczne wynagrodzenie w budownictwie brutto w zł,

X_{20} – nakłady inwestycyjne w przemyśle (ceny bieżące) w mln zł.

Kopiujemy teraz plik *DEAP2.xls* jako *DEAP3.xls*. Pozostawiamy jako efekty zmienną X_1 oraz X_5, X_6, X_8 w postaci odwróconej ($1/x$), a jako nakłady zmienne $X_{10}, X_{12}, X_{13}, X_{20}$ (symbole zmiennych wg zastosowania w *Excelu*). Zapisujemy plik jako *EAP7-dta.txt* w postaci tekstowej (tekst rozdzielony znakami tabulacji). Sprawdzamy zawartość pliku programem NOTATNIK. Następnie kopiujemy plik *EAP5-out.txt* jak *EAP6-out.txt* i czyścimy zawartość utworzonego pliku. Dokonujemy korekty pliku sterującego po wcześniejszym skopiowaniu *EAP5-ins.txt* na *EAP6-ins.txt*, którego zawartość jest następująca:

```
EAP6-dta.txt      DATA FILE NAME
EAP6-out.txt      OUTPUT FILE NAME
16               NUMBER OF FIRMS
9                NUMBER OF TIME PERIODS
4                NUMBER OF OUTPUTS
4                NUMBER OF INPUTS
1                0=INPUT AND 1=OUTPUT ORIENTATED
0                0=CRS AND 1=VRS
2                0=DEA(MULTI-STAGE), 1=COST-DEA,
                  2=MALMQUIST-DEA, 3=DEA(1-STAGE), 4=DEA(2-STAGE)
```

Analogicznie jak dla wcześniejszego przykładu (8 zmiennych wyjściowych oraz 15 zmiennych wejściowych) otrzymujemy:

a) Fragmenty listy rezultatów dla roku 1999 (1) oraz dla roku 2007 (9).

Results from *DEAP* Version 2.1 Instruction file = *EAP6-ins.txt* Data file = *EAP7-dta.txt* Output oriented *Malmquist DEA* DISTANCES SUMMARY year = 1

firm	CRS		VRS	

	<i>t</i> -1	<i>t</i>	<i>t</i> +1	
1	0.000	0.954	1.043	1.000
2	0.000	1.000	1.148	1.000
3	0.000	1.000	1.305	1.000
4	0.000	1.000	1.215	1.000
5	0.000	0.936	1.142	1.000
6	0.000	0.887	1.011	0.932
7	0.000	1.000	1.206	1.000
8	0.000	0.844	0.909	0.873
9	0.000	1.000	1.118	1.000
10	0.000	1.000	1.308	1.000
11	0.000	1.000	1.039	1.000
12	0.000	0.870	0.894	0.879
13	0.000	1.000	1.110	1.000
14	0.000	0.952	0.996	0.952
15	0.000	0.981	1.044	0.989
16	0.000	1.000	1.311	1.000
mean	0.000	0.964	1.113	0.977

year = 9

```
firm      CRS      VRS
*****
```

	<i>t-1</i>	<i>t</i>	<i>t+1</i>	
1	1.089	0.977	0.000	1.000
2	1.075	1.000	0.000	1.000
3	1.190	0.982	0.000	0.999
4	1.075	0.960	0.000	0.965
5	1.012	0.913	0.000	1.000
6	0.919	0.794	0.000	0.827
7	1.170	1.000	0.000	1.000
8	1.006	0.950	0.000	1.000
9	1.146	1.000	0.000	1.000
10	1.365	1.000	0.000	1.000
11	0.942	0.948	0.000	0.950
12	0.941	0.917	0.000	0.952
13	1.099	1.000	0.000	1.000
14	0.930	0.897	0.000	0.954
15	1.031	0.947	0.000	1.000
16	0.877	0.879	0.000	1.000
mean	1.054	0.948	0.000	0.978

b) Zestawienie bazujące na danych panelowych (*MALMQUIST*) w układzie rocznym (*ANUAL*) i województw (*FIRM*) z podaniem wartości średnich (*means*)

MALMQUIST INDEX SUMMARY OF ANNUAL MEANS

year	2	3	4	5	6	7	8	9	mean
effch	1.009	1.002	0.993	1.010	1.002	0.987	0.993	0.987	0.998

MALMQUIST INDEX SUMMARY OF FIRM MEANS

Firm	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	mean
effch	1.003	1.000	0.998	0.995	0.997	0.986	1.000	1.015	1.000	1.000	0.993	1.007	1.000	0.993	0.996	0.984	0.998

Według kryterium bazującego na zmiennych modelu ekonometrycznego osiem województw nie osiągnęło założonej jako 1 efektywności, należą do nich województwa: 3 - małopolskie, 4 - śląskie, 5 - lubelskie, 6 - podkarpackie, 11 - zachodniopomorskie, 14 - kujawsko-pomorskie, 15 - pomorskie, 16 - warmińsko-mazurskie. Porównując z wcześniejszym kryterium nadal pozostają województwa o symbolach: 5, 11, 15. Gdy przyjmimy miarę efektywności technicznej (*effch*) jako rangę to po posortowaniu malejącym ranking województw jest następujący:

firm	8	12	1	2	7	9	10	13	3	5	15	4	11	14	6	16
effch	1,015	1,007	1,003	1,000	1,000	1,000	1,000	1,000	0,998	0,997	0,996	0,995	0,993	0,993	0,986	0,984

Najwyżej usytuowane w kolejności są województwa: 8 - świętokrzyskie, 12 - dolnośląskie, 1 - łódzkie. Osiem województw osiągnęło co najmniej efektywność 1 a 8 powinno podnieść swoją efektywność.

* * *

Zaprezentowano przykład określenia efektywności technicznej gospodarności województw. Jest to podłoże do badania efektywności ekonomicznej obiektów struktury terytorialnej. Wprowadzenie cen do algorytmu programu *DEAP* umożliwi taką realizację. Dobór zmiennych przyjęty przez autora intuicyjnie może budzić zastrzeżenia, lecz celem podstawowym niniejszego opracowania było

przedstawienie procedur postępowania w przypadku zastosowania dostępnych w Internecie programów. Program *Gretl* służy do modelowania ekonometrycznego, natomiast *DEAP* do określania rangi obiektów, przy czym efektywność mierzona jest stosunkiem sumy ważonej efektów do sumy ważonej nakładów. Analogicznie można rozwiązać dowolne zagadnienie, gdyż program *DEAP* poprzez definiowanie zbiorów oraz parametrów wejścia/wyjścia jest otwarty na użytkownika. Zaletą programu *DEAP* jest możliwość wprowadzania na wejściu danych panelowych w postaci szeregów obiektowo-czasowo-przekrojowych.

3. Etapy i procedury budowy modelu produktu brutto podregionów



3.1. Wstęp

Produkt krajowy brutto jest końcowym rezultatem wszystkich podmiotów gospodarki narodowej. Produkt ten równa się sumie wartości dodanej brutto wszystkich sektorów własności oraz sektorów instytucjonalnych powiększonej o podatki od przedmiotów i pomniejszonej o dotacje do produktów²⁸. Produkt krajowy zarówno brutto jak i netto liczony jest w cenach rynkowych (bieżących).

Z dniem 1.01.1999 r. wprowadzono trójstopniowy podział terytorialny państwa (gminy, powiaty, województwa). W dniu 31.12.2006 r. było 16 województw, 314 powiatów, 65 miast na prawach powiatu oraz 2478 gmin, a ramach gmin jednostki pomocnicze - sołectwa, których liczba wynosiła 40348. Obecnie dla celów sprawozdawczych występuje 6 regionów grupujących określone województwa oraz 45 podregionów obejmujących jednostki szczebla powiatowego. Nomenklatura jednostek terytorialnych dla celów statystycznych odpowiada nomenklaturze w Unii Europejskiej.

W przykładowym zestawieniu statystycznym *Podstawowe dane o regionach w roku ...* występują 22 kolumny z informacjami charakteryzującymi nie tylko poszczególne regiony, lecz również w ramach nich podregiony. Spośród tych informacji wybrano takie, które tworzą szeregi o długości siedmiu okresów (lata 2000-2006) zmiennych egzogenicznych (objaśniających).

Dane czasowo-przekrojowe dotyczące podregionów dostępne są w rocznikach statystycznych lat 2000-2007. Dla realizacji procedury budowy modelu produktu brutto podregionów zmiennym nadano następujące symbole:

- Y - produkt brutto (ceny bieżące) na 1 mieszkańca w roku poprzednim w zł;
- X_1 - ludność na 1 km² powierzchni ogólnej;
- X_2 - przyrost naturalny na 1000 ludności;
- X_3 - saldo migracji wewnętrznych i zagranicznych na pobyt stały (na 1000 ludności);
- X_4 - przestępstwa stwierdzone na 1000 ludności;
- X_5 - pracujący w tys. (rolnictwo, łowiectwo i leśnictwo, rybactwo);
- X_6 - pracujący w tys. (przemysł i budownictwo);
- X_7 - pracujący w tys. (usługi rynkowe i nierynkowe);
- X_8 - bezrobotni zarejestrowani ogółem w tys.;
- X_9 - mieszkania oddane do użytkowania na 1000 ludności;
- X_{10} - miejsca noclegowe w turystycznych obiektach zbiorowego zakwaterowania w tys.;
- X_{11} - užitki rolne w tys. ha w roku poprzednim.

Jako zmienną objaśnianą przyjęto produkt brutto, natomiast zmiennymi objaśniającymi liniowego modelu ekonometrycznego są X_1 - X_{11} . Dane zaczerpnięto z tabel: wybrane dane o podregionach (lata 2000÷2003), podstawowe dane o regionach (lata 2004÷2007). Ponieważ niektóre informacje za ostatni rok podawane są na podstawie danych nie ostatecznych (a więc mogą ulec zmianie w roku następnym), dla celów badawczych zestawiono dane za okres 2000-2006 (dla poszczególnych 45 podregionów) rozpoczynając od roku 2006 wstecz.

W okresie sprawozdawczym 2000-2007 nastąpiły zmiany w strukturze administracyjnej kraju, stąd też niektóre brakujące dane dla podregionów województwa śląskiego przyjęto takie same jak dla roku sąsiedniego. Wystąpiły też zmiany w jednostkach miary. Przykładem są užitki rolne określone obecnie w tys. ha. Przy przeliczaniu na 1000 ludności w danych statystycznych przyjmowany jest

²⁸ Rocznik Statystyczny Rzeczypospolitej Polskiej 2007. Dział: Rachunki narodowe.

stan na 31 grudnia. Warto przytoczyć za rocznikiem statystycznym z roku 2007 niektóre pojęcia występujące w powyższych zmiennych:

- *przestępstwo stwierdzone* jest to zdarzenie, co do którego w zakończonym postępowaniu przygotowawczym potwierdzono, że jest przestępstwem;
- *przyrost naturalny ludności* stanowi różnicę między liczbą urodzeń żywych i zgonów w danym okresie;
- *przyrost rzeczywisty ludności* jest sumą przyrostu naturalnego i salda migracji;
- *migracje wewnętrzne ludności* są to zmiany miejsca zamieszkania w kraju;
- *migracje zagraniczne ludności* są to wyjazdy poza granicę kraju zamieszkania, a także przyjazdy do kraju.

Dane o migracjach zagranicznych na popyt stały bazują na informacjach ewidencyjnych gmin (imigracja) oraz o wymeldowaniach osób wyjeżdżających z Polski (emigracja), a dane o pracujących dotyczą osób wykonujących pracę przynoszącą zarobek lub dochód. Natomiast dane o bezrobotnych zarejestrowanych obejmują osoby, które zgodnie z ustawą z dnia 21.04.2004 o promocji zatrudnienia i instytucjach rynku pracy, określane są jako bezrobotne. Przy czym pod pojęciem bezrobotnego rozumie się osobę niezatrudnioną i niewykonującą innej pracy zarobkowej, zdolną i gotową do podjęcia zatrudnienia, nieuczącą się (z wyjątkiem szkół dla dorosłych)²⁹.

Sformułowanie modelu ekonometrycznego dla podregionów wymaga wielu lat obserwacji, co przy dynamicznie zmieniającej się sytuacji gospodarczej wydaje się niecelowe. Z tego względu posłużono się poszczególnymi 45 jednostkami administracyjnymi (podregionami) obserwowanymi przez okresy 7-mio letnie. Powstał w ten sposób panel danych o $45 \times 7 = 315$ obserwacjach. Fragment tego panelu obejmującego *województwo łódzkie* zaprezentowano w tabeli 3.1. Jeżeli w rocznikach statystycznych użytki rolne występowały w % to dokonano przeliczenia na tys. ha w sposób następujący: powierzchnia ogółem w km^2 , użytki rolne w km^2 (powierzchnia ogółem $\times$ przelicznik w %), przeliczenie na hektary ($1 \text{ km}^2 = 100 \text{ ha}$), pomnożenie przez 1000.

Tab. 3.1. Dane czasowo-przekrojowe dotyczące podregionów województwa łódzkiego

Lp.	Region/Województwo/Podregion	Rok	Okres	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	%
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1	centralny/łódzkie/łódzki	2000	1	11442	110,0	-2,3	0,4	223,9	164,2	75,3	83,7	77,6	1,1	9,4	629,9	72,2
		2001	2	12790	109,0	-2,6	0,1	269,6	164,1	75,3	83,7	86,1	1,5	5,0	629,9	72,3
		2002	3	14350	108,0	-2,8	0,4	278,8	93,6	66,4	81,0	87,4	1,6	5,2	575,5	
		2003	4	15259	108,0	-2,9	0,8	287,4	93,3	67,7	80,5	88,3	3,2	4,9	564,1	66,9
		2004	5	15955	108,0	-2,4	0,9	287,0	93,5	68,5	81,1	83,2	2,1	4,9	564,1	66,9
		2005	6	18713	107,0	-3,0	0,6	275,0	93,5	72,5	81,0	77,6	2,0	4,8	552,7	
		2006	7	19837	107,0	-2,5	0,5	256,0	93,5	76,5	82,5	65,3	1,8	4,7	552,7	
2	piotrkowsko-skierniewicki	2000	1	12254	97,0	-0,9	-1,0	216,3	177,3	76,7	79,3	75,0	1,7	8,6	611,8	66,5
		2001	2	13478	97,0	-1,3	-1,0	265,5	176,8	74,9	74,9	82,6	2,2	9,2	610,9	76,0
		2002	3	15289	86,0	-1,1	-1,1	268,5	96,5	70,2	77,8	82,3	2,3	8,2	548,1	
		2003	4	16060	95,0	-1,7	-1,0	275,0	96,1	69,6	78,1	79,6	4,9	7,0	553,8	60,2
		2004	5	17034	95,0	-1,4	-1,4	266,0	96,2	71,8	79,6	73,4	2,5	7,3	553,8	60,2
		2005	6	20040	95,0	-1,3	-1,1	246,0	96,2	72,6	81,6	66,2	2,3	7,0	514,2	
		2006	7	21657	94,7	-0,8	-1,8	257,0	96,2	73,7	82,9	56,7	2,3	7,5	514,2	
3	miasto Łódź	2000	1	20645	2694,0	-7,0	-0,9	472,4	3,9	73,5	140,7	55,5	1,5	5,7	11,0	37,3
		2001	2	21767	2672,0	-6,7	-0,9	450,0	3,9	72,6	141,5	61,9	2,1	4,8	10,6	36,0
		2002	3	23924	2671,0	-6,3	-1,2	489,3	2,2	64,1	135,7	65,1	1,6	4,6	23,1	
		2003	4	25072	2647,0	-6,2	-1,2	538,7	2,1	60,0	137,7	65,2	1,6	3,3	21,2	27,7
		2004	5	26531	2629,0	-5,8	-1,5	550,0	2,1	59,1	134,5	61,7	1,8	3,5	21,2	27,7
		2005	6	29104	2608,0	-6,1	-1,8	512,0	2,1	59,3	138,9	54,7	0,9	3,5	19,2	
		2006	7	30602	2593,0	-6,5	-2,4	434,0	2,1	60,9	147,9	38,7	1,6	3,6	19,2	

Źródło: Opracowanie własne na podstawie roczników statystycznych z lat 2000-2006.

W poszczególnych podregionach występuje znaczne zróżnicowanie wysokości produktu brutto w przeliczeniu na 1. mieszkańca. Możemy to szczególnie zauważyć po wprowadzeniu wartości średniej podregionów Polski i wykonaniu wykresu.

²⁹ Rocznik Statystyczny Rzeczypospolitej Polskiej 2007. Dział: Rynek pracy.

3.2. Podejście do modelowania

W literaturze występują różne sposoby sformułowania i weryfikacji liniowego modelu ekonometrycznego opartego o dane panelowe. Najczęściej sugeruje się rozwiązania ze wspomaganiami komputerowym. Jednym z przykładów jest określenie modelu po wcześniejszym wyznaczeniu wartości średnich z obserwacji zmiennych (patrz tabela 3.2). Zabieg ten ma na celu wyeliminowanie problemu z odwracaniem wysokiego rzędu macierzy.

Tab. 3.2. Wartości średnie zmiennych dla podregionów

Lp.	Podregion/Wartość średnia zmiennej	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁
1	Łódź	15478	108,1	-2,6	0,5	268,2	113,7	71,7	81,9	80,8	1,9	5,6	581,3
2	piotrkowsko-skierniewicki	16544,6	94,2	-1,2	-1,2	256,3	119,3	72,8	79,2	73,7	2,6	7,8	558,1
3	miasto Łódź	25377,9	2644,9	-6,4	-1,4	492,3	2,6	64,2	139,6	57,5	1,6	4,1	17,9
4	ciechanowsko-płocki	20354,1	81,4	0,1	-1,7	282,8	80,7	44,0	61,8	64,8	2,6	2,9	535,5
5	ostrolecko-siedlecki	14160,1	62,6	0,7	-2,3	278,8	126,0	35,2	64,9	64,1	2,1	5,3	750,9
6	warszawski	20122,1	163,4	-0,1	8,1	347,0	88,2	89,9	132,9	66,7	6,0	4,7	460,6
7	radomski	14743,6	104,6	0,5	-1,8	311,9	98,9	143,3	64,9	81,8	2,1	3,2	405,6
8	miasto Warszawa	60932,7	3271,3	-2,5	3,8	536,1	139,2	157,6	602,4	55,4	8,1	19,8	88,4
9	krakowsko-tarnowski	14488,3	189,6	0,6	1,1	262,4	145,2	91,2	116,0	83,8	3,1	6,0	427,4
10	nowosądecki	12391,9	147,7	3,9	-0,1	230,5	119,5	55,6	98,1	79,7	3,0	43,7	331,5
11	miasto Kraków	32526,4	2302,3	-1,1	1,8	608,2	2,9	71,7	178,8	32,6	6,1	14,2	20,9
12	częstochowski (północnośląski)	18488,7	177,3	-2,4	-0,5	298,3	44,6	62,4	73,8	47,5	2,3	4,4	198,1
13	bielsko-biały (południowośląski)	21405,3	288,9	0,5	0,7	300,8	38,4	100,5	105,6	47,8	3,8	21,4	133,9
14	centralny śląski	25433,8	860,1	-1,8	-3,1	456,8	22,3	302,4	378,8	172,8	1,3	8,3	160,7
15	hybnycko-jastrzębski	21225,8	474,2	1,3	-3,8	353,1	8,3	66,3	66,2	34,7	1,8	1,9	66,1
16	białkopodlaski	12878,1	70,5	-0,6	-2,9	306,4	55,6	26,1	41,7	30,4	1,9	6,2	326,1
17	chełmsko-zamojski	12266,1	71,6	-1,5	-2,8	248,4	128,5	27,6	58,6	56,3	1,5	6,1	582,3
18	lubelski	16052,6	123,7	-0,2	-1,3	292,4	162,5	71,3	142,6	81,5	2,4	9,0	663,3
19	rzeczowsko-tarnobrzelski	16164	154,6	1,9	-0,9	232,9	122,6	99,5	126,8	88,7	2,6	7,0	399,3
20	króśnieńsko-przemyski	12890,4	91,3	1,2	-1,7	187,5	117,2	57,8	93,8	86,5	1,9	13,2	424,2
21	białostocko-suwalski	16364,4	60,2	-0,9	-0,9	290,2	108,5	47,8	103,9	54,2	3,2	11,2	781,0
22	łomżyński	12990,3	58,9	0,9	-3,5	239,8	55,6	12,3	24,9	23,6	1,4	2,0	360,6
23	świętokrzyski	15974,6	110,9	-1,4	-1,9	263,5	185,3	81,0	131,9	121,6	1,8	8,8	627,4
24	gorzowski	18755	62,4	1,3	-0,4	418,0	14,3	30,3	45,8	34,3	3,2	8,3	211,4
25	zielenogórski	18169,3	80,2	0,7	-1,1	409,1	20,9	45,4	71,8	61,5	2,5	10,7	289,8
26	pilski	16908,6	63,3	2,2	-1,3	330,4	27,8	35,4	37,8	35,3	1,9	4,3	309,0
27	poznański	18518	99,6	2,1	4,2	262,1	79,6	114,0	110,6	65,1	3,4	17,9	694,5
28	kaliski	16073,4	112,3	0,7	-0,6	243,9	86,2	75,0	74,3	58,9	2,2	4,0	486,0
29	koniński	16534,4	98,3	1,1	-1,2	286,8	52,6	36,5	37,8	42,4	2,2	6,0	296,1
30	miasto Poznań	42297,9	2187,4	-1,3	-2,4	667,3	4,6	64,8	157,0	18,4	5,0	7,2	29,1
31	szczeciński	21570,7	88,6	0,3	-0,2	424,6	31,9	73,1	138,0	98,6	3,2	61,3	623,1
32	koszaliński	16983,9	57,6	1,2	-1,8	372,5	22,4	35,4	60,6	72,3	2,6	54,0	410,5
33	jeleniogórsko-walbrzyski	16728,4	129,0	-1,9	-2,2	385,1	46,8	101,4	130,8	138,4	1,5	37,4	497,2
34	legnicki	26157,6	121,3	1,1	-2,6	461,5	19,7	52,4	61,2	44,0	2,0	2,4	208,4
35	wrocławski	15930	85,1	0,9	3,6	236,4	30,2	33,6	33,8	34,8	3,4	3,2	285,9
36	miasto Wrocław	30598,4	2171,4	-2,1	1,0	586,9	2,0	50,4	140,0	30,4	5,4	7,1	56,2
37	opolski	17184,6	112,7	-0,7	-3,7	331,6	65,6	76,3	108,7	73,9	1,4	7,1	563,1
38	bydgoski	19971,7	114,6	0,5	-0,6	378,8	53,7	86,8	125,0	88,9	2,2	14,6	496,3
39	toruńsko-włocławski	17075,6	116,7	0,7	-1,1	355,1	87,1	76,7	111,0	105,0	2,5	11,9	599,8
40	ślupski	15745,3	58,7	2,8	-2,1	422,8	21,4	36,2	50,3	58,0	2,6	34,1	301,8
41	gdański	15276,6	99,1	4,3	2,5	369,7	50,3	63,9	77,2	79,1	4,3	38,6	496,5
42	Gdańsk-Gdynia-Sopot	30127,1	1812,9	-0,8	-0,8	672,3	3,1	64,0	151,2	30,2	4,9	16,3	66,1
43	elbląski	15207,6	71,5	2,1	-2,7	336,9	31,8	39,0	51,2	61,7	2,1	8,6	378,6
44	olsztyński	17814,9	59,7	1,7	-1,3	396,8	30,1	42,4	78,6	62,1	3,5	22,0	486,6
45	elcki	14091,6	45,2	2,2	-2,7	395,5	19,4	14,7	24,5	36,9	2,7	9,2	257,8

Źródło: Opracowanie własne w Excel-u, gdzie np. $y = \bar{y}$.

W tabeli 3.2 uśredniono na podstawie okresów zmienną objaśnianą Y oraz zmienne objaśniające X_1 - X_{11} . Następnie możemy zastosować funkcje Excel-a dla wyznaczenia wektora parametrów modelu liniowego, który w postaci ogólnej określany jest następującym równaniem macierzowym:

$$a = (X^T X)^{-1} X^T y$$

Gdzie dla pełnych szeregów czasowo-przekrojowych:

X - macierz zmiennych o wymiarach $N \times K = 315 \times 11$,

X^T - macierz transponowana o wymiarach 11×315 ,

$(X^T X)^{-1}$ - macierz odwrotna o wymiarach 11×11 ,

y - wektor zmiennej objaśnianej o wymiarach 315 x 1.

W określeniu modelu na bazie danych panelowych podregionów korzysta się z estymatora *fixed effect* (FE), który zakłada, że efekty indywidualne β_i dotyczące poszczególnych 45 podregionów można oszacować. Wzór w postaci macierzowej na ocenę parametrów wektora β jest następujący³⁰:

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^N \tilde{x}_i^T \tilde{x}_i \right)^{-1} \sum_{i=1}^N (\tilde{x}_i^T \tilde{y}_i) \quad (3.1)$$

przy czym zastosowano tu podstawienia:

$$\tilde{y}_i = y_i - c\bar{y}_i \quad \tilde{x}_i = x_i - c\bar{x}_i^T$$

Wektor oszacowanych efektów β_{FE} dotyczy poszczególnych jedenastu zmiennych objaśniających. Obliczenia poprzedzone muszą być określeniem różnic wartości obserwacji od średnich. Podane w powyższym zapisie macierzowym określenie c to zmienna zerojedynkowa odnosząca się tylko do określonego podregionu. Przekształcenie jako wektor $\hat{\beta}_{FE}$ umożliwia zastosowanie klasycznej metody najmniejszych kwadratów *KMNK*, a w ramach niej następujących funkcji programu *Excel* na obszarach adresowych macierzy lub wektorów:

TRANSPONUJ(...) transpozycja macierzy X ,

MACIERZ.ILOCZYN(...) – iloczyn macierzy transponowanej X^T oraz macierzy zmiennych X ,

MACIERZ.ODW(...) – macierz odwrotna macierzy powstałej z iloczynu $(X^T X)$.

Po określeniu wektora parametrów stojących przy zmiennych objaśniających estymator 45 efektów indywidualnych dla poszczególnych podregionów otrzymujemy ze z poniższego wzoru³¹:

$$\hat{\beta}_i = \bar{y}_i - \bar{x}_i^T \hat{\beta}_{FE} \quad (3.2)$$

W obliczeniach oprócz funkcji rachunku macierzowego możemy skorzystać alternatywnie z funkcji *REGLINP* w odniesieniu do całego panelu 315 obserwacji. Korzystając z tej funkcji nie przyjmujemy występowanie parametru stałego, gdyż będzie on rozczłonkowany jako efekty indywidualne wpływu danych podregionów.

Sformułowanie modelu po powyższej procedurze umożliwia porównanie wartości empirycznych i modelowych zmiennej objaśnianej. Ponadto możemy obliczyć procentowy błąd względny *MAPE* dla 315 obserwacji zmiennej Y dla całego panelu oraz poszczególnych obiektów (podregionów) korzystając ze wzoru:

$$MAPE = \sum_i^n |(y - \hat{y}) / y| * 100 \quad (3.3)$$

Kontynuowanie manualne ze wspomaganiem komputerowym funkcjami *Excel*-a jest zbyt mozolne i z tego powodu w dalszej procedurze budowy modelu ekonometrycznego produktu brutto podregionów posłużono się programem *Gretl*. Na uwagę zasługuje możliwość korzystania w *Gretl* z danych panelowych. Jak już wspomniano z danymi panelowymi mamy do czynienia, gdy obserwujemy dane na kilku obiektach, obszarach i okresach czasowych. Tadeusz Kufel jako przykład podaje „...badanie ekonomicznej efektywności ludności ..., stopy bezrobocia w powiatach w Polsce w latach 1998-2002” oraz rozwiązuje ten przykład w *Gretl*³².

³⁰ Osińska M. (red. naukowa): *Ekonometria współczesna*. Towarzystwo Naukowe Organizacji i Kierownictwa. Toruń 2007.

³¹ Kufel T., *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*, Wydanie 3, WN PWN, Warszawa 2011.

³² Ibidem, rozdz. 12.

3.3. Modelowanie z zastosowaniem *Gretl*-a

Jak wiemy program *Gretl* umożliwia szybkie otrzymywanie statystyki opisowej w odniesieniu do zmiennej Y jak też do zmiennych objaśniających X_1 - X_{11} , a w ramach niej określenie średniej, mediany, wartości minimalnej oraz maksymalnej, odchylenia standardowego, współczynnika zmienności, skośności i kurtozy. Na uwagę zasługuje współczynnik zmienności w %, który tu przypomnimy sobie:

$$V_j = \left| \frac{S_j}{\bar{x}_j} \right| * 100$$

Wyznaczenie V_j stanowi wstępny krok zbadania wystarczającego zróżnicowania danej j -tej zmiennej objaśniającej. Określany jest jako moduł z ilorazu odchylenia standardowego i średniego poziomu wartości rozpatrywanej zmiennej.

Tab. 3.3. Współczynniki zmienności zmiennych objaśniających

Zmienna	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
V_j	226	65	153	106	102	79	88	50	59	104	56

Źródło: Opracowanie własne z zastosowaniem programu *Gretl*.

Z tabeli wynika, że dla danych panelowych współczynnik zmienności znacznie przekracza wartość krytyczną $V^* = 10\%$, czyli wszystkie zmienne pełnego panelu o 315 obserwacjach wchodzą wstępnie do budowy modelu ekonometrycznego. W selekcji zmiennych objaśniających możemy następnie posłużyć się jedną z metod doboru zmiennych objaśniających, przykładowo analizą korelacji. Niezbędna w tym zakresie jest tablica współczynników korelacji między zmienną objaśnianą Y a zmiennymi objaśniającymi X_j , jak również współczynników korelacji wzajemnych między zmiennymi objaśniającymi (patrz tabela 3.4). Skorzystano w tym względzie z *Gretl*-a, który numeruje obserwacje wg kolejności obiektów - podregionów, a w ramach tego kolejne okresy czasowe (1:1 - 45:7).

Tab. 3.4. Współczynniki korelacji

Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	Y
1	0,7994	-0,3757	0,2317	0,6465	-0,2653	0,2443	0,7321	-0,2088	0,6093	-0,0020	-0,5392	Y
	1	-0,5590	0,1954	0,7204	-0,1893	0,2158	0,6571	-0,2385	0,5230	-0,0627	-0,6477	X_1
		1	0,0206	-0,3258	0,0075	-0,2053	-0,3964	-0,0760	-0,0874	0,2933	0,2210	X_2
			1	0,0651	0,0700	0,1299	0,2747	-0,0440	0,5859	0,1063	0,0042	X_3
				1	-0,4647	0,1039	0,3960	-0,1952	0,4664	0,1166	-0,6063	X_4
					1	0,1777	0,1373	0,2919	-0,1017	-0,0747	0,5152	X_5
						1	0,6309	0,5659	0,0389	0,0220	-0,0550	X_6
							1	0,3115	0,4366	0,1074	-0,2461	X_7
								1	-0,2591	0,2885	0,4271	X_8
									1	0,1298	-0,3273	X_9
										1	0,1441	X_{10}
											1	X_{11}

Źródło: Opracowanie własne z zastosowaniem programu *Gretl*.

Program *Gretl* oblicza także wartość krytyczną (przy dwustronnym 5% obszarze krytycznym) która wynosi $r^* = 0,1105$ dla $n = 315$ obserwacji. Wartość tą możemy także obliczyć ze wzoru:

$$r^* = \sqrt{\frac{t_{\alpha}^2}{t_{\alpha}^2 + n - 2}} = \sqrt{\frac{(1,979)^2}{(1,979)^2 + 315 - 2}} = 0,1105$$

Przy czym:

$t_{\alpha} = 1,979$ - statystyka wyznaczona na podstawie rozkładu t -Studenta dla poziomu istotności $\alpha = 0,05$ oraz ilości stopni swobody $iss = n - 2 = 315 - 2 = 313$

Przeglądamy powyższą tabelę i sprawdzamy warunek:

$$|r_j| \leq r^* = 0,1105$$

Według tego kryterium możemy wyeliminować zmienną X_{10} o $r_{10} = -0,002$ ze zbioru zmiennych objaśniających. Zmienna najsilniej skorelowana ze zmienną Y jest zmienna X_1 o $r_1 = 0,7994$.

Możemy posłużyć się jeszcze innymi manualnymi metodami, lecz rozpatrzenie wszystkich kombinacji zmiennych jest pracochłonne. Warto jednak choć wymienić te metody, a więc:

- metoda wskaźników zmiennych objaśniających,
- bazowanie na mierze błędów *ex ante*,
- metoda pojemności nośników informacji,
- metoda grafowa i inne.

Wróćmy jednak do naszego programu *Gretl*. Pierwszym jednak krokiem jest przygotowanie danych panelowych. W tym względzie dogodnie jest uczynić to w *Excel*-u korzystając z jednego arkusza skoroszytu i oznaczając kolejne zmienne jako $X1$ - $X11$. W arkuszu kalkulacyjnym rozpoczynamy zawsze od komórki $A1$ w pierwszym wierszu, a więc wpisujemy Y i w kolejności $X1$ - $X11$. Warto zaznaczyć, że znak precyzji przecinek jest akceptowalny przez *Gretl* (w wersji programu z grudnia 2008 roku). Program ten jest prawie co kwartałnie modyfikowany oraz rozszerzany przez ekonometryków w miarę rozwoju metod ekonometrycznych. Import pliku *Excel*-a do *Gretl*-a następuje po kolejnym wywołaniu opcji menu:

Plik/ otwórz dane/import/Excel.

wybranie pliku (w tym przykładzie *Dane-gretl.xls* z podkatalogu na dysku D).

Zastosowanie klasycznej metody najmniejszych kwadratów *KMNK* przy formułowaniu modelu na danych panelowych jest ograniczone ze względów merytorycznych. Zdaniem T. Kufla³³ możemy zastosować *KMNK*, gdy wprowadzimy zmienne zerojedynkowe w naszym przykładzie do poszczególnych 45. obiektów (podregionów) panelu o 315 obserwacjach, przy czym 1 oznacza uwzględnienie danego okresu czasowego jednostki panelu (podregionu), a cyfra 0 pominie.

Zabieg ten stosowany jest także przez innych autorów między innymi w *Metodzie Kleina* w której zastosowano sezonową zmienną zerojedynkową. W naszym przypadku umożliwia on uzyskanie efektów indywidualnych β_{FE} w odniesieniu do poszczególnych podregionów. Wcześniej jednak musimy dokonać dołączenia „podmacierzy” zerojedynkowej do macierzy zmiennych objaśniających. Dokonujemy to poprzez polecenie:

Dodawanie zmiennych/zmienne 0-1 jednostek panelu.

Kolejny krok to polecenie *Model/klasyczna metoda najmniejszych kwadratów*. W specyfikacji zmiennych określamy zmienną objaśnianą (zależną) oraz zmienne objaśniające (niezależne) $X1$ - $X11$ bez stałej *const*, gdyż jej rolę przejmą efekty indywidualne odniesione do podregionów.

Moglibyśmy już teraz zrezygnować z zmiennej X_{10} , której nieistotność wykazała wstępna analiza korelacji.

Efektom pracy programu *Gretl* jest poniższy *Model 1* stanowiący estymację danych panelowych z wykorzystaniem 315 obserwacji. Włączono do niego 45 jednostek danych przekrojowych (podregionów) o szeregach czasowych długości 7 (lat).

Tab. 3.5. Parametry modelu panelowego oraz statystyka *t*-Studenta

Zmienna	Współczynnik (parametr)	Błąd standardowy	t-Studenta	Wartość p	
X_1	-9,43147	1,93873	-4,8648	<0,00001	***
X_2	2489,94	398,217	6,2527	<0,00001	***
X_3	-930,436	235,347	-3,9535	0,00010	***
X_4	-4,38485	3,88392	-1,1290	0,25995	
X_5	-33,9051	3,35581	-10,1034	<0,00001	***

³³ Kufl T., *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*, op. cit. s. 167.

X6	-8,46421	5,28443	-1,6017	0,11043	
X7	77,376	16,2766	4,7538	<0,00001	***
X8	-71,8521	16,6996	-4,3026	0,00002	***
X9	403,339	157,835	2,5554	0,01118	**
X10	-141,224	85,0459	-1,6606	0,09801	*
X11	-36,3181	6,24222	-5,8181	<0,00001	***
du 1	49802,3	3862,25	12,8946	<0,00001	***
du 2	44621,4	3512,42	12,7039	<0,00001	***
du 3	61595,6	7245,92	8,5007	<0,00001	***
du 4	42305,9	3409,37	12,4087	<0,00001	***
du 5	43383,3	4412,54	9,8318	<0,00001	***
du 6	44168,9	4056,78	10,8877	<0,00001	***
du 7	33819,9	2800,85	12,0749	<0,00001	***
du 8	70167,8	14011,9	5,0077	<0,00001	***
du 9	34797,3	3020,25	11,5214	<0,00001	***
du 10	24652,8	3859,02	6,3884	<0,00001	***
du 11	50856,7	6827,43	7,4489	<0,00001	***
du 12	33244	2430,98	13,6752	<0,00001	***
du 13	28033,7	2851,16	9,8324	<0,00001	***
du 14	29606	7143,92	4,1442	0,00005	***
du 15	20550,4	2646,26	7,7659	<0,00001	***
du 16	24720,8	2673,92	9,2451	<0,00001	***
du 17	40802,2	3751,68	10,8757	<0,00001	***
du 18	43016,9	4161,7	10,3364	<0,00001	***
du 19	29115,1	2902,21	10,0321	<0,00001	***
du 20	30041,1	2618,63	11,4720	<0,00001	***
du 21	48196,6	4829,41	9,9798	<0,00001	***
du 22	23776,5	2692,25	8,8315	<0,00001	***
du 23	48567,1	3937,48	12,3346	<0,00001	***
du 24	24855,6	2414,78	10,2931	<0,00001	***
du 25	28965,3	2614	11,0808	<0,00001	***
du 26	24208,9	2502,71	9,6731	<0,00001	***
du 27	45489,9	3911,89	11,6286	<0,00001	***
du 28	35148,1	3013,94	11,6618	<0,00001	***
du 29	27726	2261,49	12,2601	<0,00001	***
du 30	56659,7	6380,39	8,8803	<0,00001	***
du 31	51356,1	5468,97	9,3905	<0,00001	***
du 32	37517,2	4641,81	8,0824	<0,00001	***
du 33	47212,5	4257,06	11,0904	<0,00001	***
du 34	30772,7	2721,12	11,3089	<0,00001	***
du 35	29420,5	2170,14	13,5570	<0,00001	***
du 36	52595,5	6204,47	8,4770	<0,00001	***
du 37	38651,7	3810,65	10,1431	<0,00001	***
du 38	39295	3380,15	11,6252	<0,00001	***
du 39	41995,9	3660,14	11,4739	<0,00001	***
du 40	25112	3581,53	7,0115	<0,00001	***
du 41	33129,6	3830,45	8,6490	<0,00001	***
du 42	45174,2	5783,76	7,8105	<0,00001	***
du 43	25566,6	2814,73	9,0831	<0,00001	***
du 44	33930,2	3375,82	10,0510	<0,00001	***
du 45	19498,7	2710,71	7,1932	<0,00001	***

Źródło: Opracowanie własne w Grettl.

Przeglądając model widzimy poszczególne parametry zwane w *Gretl* współczynnikami odnoszące się do zmiennych objaśniających, a także stanowiące efekty indywidualne stojące przy zmiennych zerojedynkowych (*du_1* do *du_45*). W raporcie występuje też określenie błędu standardowego oszacowania danego parametru, obliczenie statystyki *t*-Studenta. Jak już nadmieniono, program wprowadza dogodne oznaczenia informujące o statystycznej istotności danego parametru a mianowicie:

*** - $\alpha = 0,01$; ** - $\alpha = 0,05$; * - $\alpha = 0,1$; puste – brak statystycznej istotności ($\alpha > 0,1$).

Dużym udogodnieniem jest również informacja na zakończenie *Modelu 1* o treści:

Wyluczając stałą, największą wartość p jest dla zmiennej X4.

Uwaga ta wskazuje zmienną nieistotną statystycznie dla której prawdopodobieństwo błędu jest największe i wynosi $p = 0,25995$, a statystyka $t_{obl} = 1,129$, czyli jest najmniejsza.

Teraz możemy przystąpić w *Gretl* do usuwania zmiennych nieistotnych statystycznie posługując się w kolejnych iteracjach *eliminacją a posteriori*. W tej eliminacji kierujemy się właśnie wskazaniami zmiennej na zakończeniu raportu z poprzedniego modelu.

Tak postępując otrzymano *Model 2*, gdzie *Gretl* zaproponował usunięcie X_6 , *Model 3* z sugestią eliminacji zmiennej X_{10} , a więc takiej jaką wskazała wcześniej wstępna analiza korelacji. Ostatnim jest zamieszczony niżej *Model 4*, który nie zawiera już propozycji eliminacji zmiennej objaśniającej. Tylko przy zmiennej X_9 występują dwie gwiazdki co świadczy o oszacowaniu na poziomie istotności $\alpha = 0,05$.

Tab. 3.6. Propozycja końcowa modelu ekonometrycznego podregionów

Zmienna	Współczynnik	Błąd standardowy	t-Studenta	Wartość p	
X1	-9,96098	1,93834	-5,1389	<0,00001	***
X2	2583,1	380,316	6,7920	<0,00001	***
X3	-1031,14	226,898	-4,5445	<0,00001	***
X5	-34,3597	3,33367	-10,3069	<0,00001	***
X7	69,1179	15,9021	4,3465	0,00002	***
X8	-72,249	16,6168	-4,3480	0,00002	***
	402,411	158,493	2,5390	0,01170	**
X11	-40,4479	5,85304	-6,9106	<0,00001	***
du 1	50753,3	3714,09	13,6651	<0,00001	***
du 2	44862,5	3323,07	13,5003	<0,00001	***
du 3	61411,8	7225,55	8,4993	<0,00001	***
du 4	42936,8	2996,06	14,3311	<0,00001	***
du 5	44576,3	3949,65	11,2861	<0,00001	***
du 6	45208,5	3866,8	11,6914	<0,00001	***
du 7	32896,1	2342,31	14,0443	<0,00001	***
du 8	71474,5	13955,7	5,1215	<0,00001	***
du 9	35011	2782,9	12,5807	<0,00001	***
du 10	18970	2369,25	8,0068	<0,00001	***
du 11	48660	6585,54	7,3889	<0,00001	***
du 12	32524,9	2276,87	14,2849	<0,00001	***
du 13	24479,9	2284,64	10,7150	<0,00001	***
du 14	28044,6	7040,7	3,9832	0,00009	***
du 15	18769,1	2046,03	9,1734	<0,00001	***
du 16	23874,4	2237,98	10,6678	<0,00001	***
du 17	41493,7	3506,6	11,8330	<0,00001	***
du 18	43829,1	3835,53	11,4271	<0,00001	***
du 19	28862,3	2620,12	11,0156	<0,00001	***
du 20	29258,1	2424,43	12,0680	<0,00001	***
du 21	49117	4554,93	10,7833	<0,00001	***
du 22	23666,6	2230,06	10,6125	<0,00001	***

du 23	49287,8	3742	13,1715	<0,00001	***
du 24	22743,1	1583,36	14,3638	<0,00001	***
du 25	26977,1	1955,16	13,7978	<0,00001	***
du 26	23167,7	1807,71	12,8160	<0,00001	***
du 27	44968,6	3744,55	12,0091	<0,00001	***
du 28	35487,8	2671,93	13,2817	<0,00001	***
du 29	26708,8	1799,86	14,8393	<0,00001	***
du 30	54632,3	5943,02	9,1927	<0,00001	***
du 31	43981,4	3638,11	12,0891	<0,00001	***
du 32	29930,8	2341,18	12,7845	<0,00001	***
du 33	42608,8	3484,75	12,2272	<0,00001	***
du 34	29055,6	1715,72	16,9349	<0,00001	***
du 35	29463,8	1949,56	15,1130	<0,00001	***
du 36	51455,4	5980,38	8,6040	<0,00001	***
du 37	38584,2	3429,38	11,2511	<0,00001	***
du 38	37931,2	2986,35	12,7015	<0,00001	***
du 39	41479,5	3285,02	12,6269	<0,00001	***
du 40	19387,4	1980,46	9,7894	<0,00001	***
du 41	28157,9	2750,7	10,2366	<0,00001	***
du 42	41870,1	5251,86	7,9724	<0,00001	***
du 43	24148,1	2191,24	11,0203	<0,00001	***
du 44	31177,6	2707,83	11,5139	<0,00001	***
du 45	17183,1	1825,61	9,4122	<0,00001	***

Źródło: Opracowanie własne w *Gretl*.

3.4. Porównanie modelu z danymi empirycznymi

Warto jeszcze zwrócić uwagę na miary pomocnicze dobroci dopasowania modelu do specyficznych danych empirycznych jakimi są dane panelowe (przestrzenno-czasowe), a należą do nich:

- współczynnik determinacji (R^2),
- skorygowany współczynnik determinacji ($\bar{R}^2$),
- kryterium informacyjne Akaike'a (AIC),
- kryterium Schwarza (BIC).

Tab. 3.7. Porównanie modeli według czterech kryteriów

Miara	Model 1	Model 2	Model 3	Model 4	MIN
R^2	0,94794	0,94769	0,94720	0,94669	Model 4
$\bar{R}^2$	0,93689	0,93682	0,93647	0,93611	Model 4
AIC	5825,00	5824,55	5825,49	5826,49	Model 2
BIC	6035,15	6030,94	6028,13	6025,38	Model 4

Źródło: Opracowanie własne na podstawie raportów *Gretl*-a czterech kolejnych modeli.

Najlepszym według czterech kryteriów jest *Model 4*, czyli taki jaki uzyskaliśmy w *Gretl* poprzez eliminację *a posteriori*. Na wydruku *Modelu 4* znajduje się także statystyka F (*Fishera-Snedecora*), która umożliwia całościową ocenę przydatności zaprojektowanego modelu ekonometrycznego dla $(8 + 45 = 53)$ zmiennych objaśniających:

$$F = 89,4765 \text{ (wartość } p < 0,00001).$$

Mała wartość prawdopodobieństwa błędu przemawia za słusnością zróżnicowania doboru parametrów dla podregionów. Pozostaje nam jeszcze ocena wygasłych już prognoz stosując miarę *MAPE* (zob. tabela 3.8 oraz 3.9) dla przykładowego województwa łódzkiego oraz mazowieckiego.

Tab. 3.8. Porównanie danych empirycznych i modelowych oraz określenie błędu *MAPE* dla województwa łódzkiego w ramach regionu centralnego

Lp.	Region/Województwo/ Podregion	Rok	Okres	y	$\hat{y}$	$y - \hat{y}$	$ (y - \hat{y}) / y * 100$	β_i	Śred- nia z kol. 8
1	2	3	4	5	6	7	8	9	10
1	centralny/łódzkie/ łódzki	2000	1	11442	12804,6	-1362,6	11,9	50753,3	7,6
		2001	2	12790	11895,7	894,3	7,0		
		2002	3	14350	15466,4	-1116,4	7,8		
		2003	4	15259	15811,3	-552,3	3,6		
		2004	5	15955	16960,2	-1005,2	6,3		
		2005	6	18713	16548,2	2164,8	11,6		
		2006	7	19837	18854,7	982,3	5,0		
2	piotrkowsko-skiemiewicki	2000	1	12254	12512,3	-258,3	2,1	44862,5	6,2
		2001	2	13478	10878,7	2599,3	19,3		
		2002	3	15289	17170,3	-1881,3	12,3		
		2003	4	16060	16471,3	-411,3	2,6		
		2004	5	17034	17241,1	-207,1	1,2		
		2005	6	20040	19371,4	668,6	3,3		
		2006	7	21657	22164,0	-507,0	2,3		
3	miasto Łódź	2000	1	20645	23161,7	-2516,7	12,2	61411,8	5,7
		2001	2	21767	24006,2	-2239,2	10,3		
		2002	3	23924	24079,7	-155,7	0,7		
		2003	4	25072	24790,4	281,6	1,1		
		2004	5	26531	26424,4	106,6	0,4		
		2005	6	29104	26694,6	2409,4	8,3		
		2006	7	30602	28489,2	2112,8	6,9		

Źródło: Opracowanie własne w Excel-u.

Tab. 3.9. Porównanie danych empirycznych i modelowych oraz określenie błędu *MAPE* dla województwa mazowieckiego

Lp.	Region/Województwo/ Podregion	Rok	Okres	y	$\hat{y}$	$y - \hat{y}$	$ (y - \hat{y}) / y * 100$	β_i	Śred- nia z kol. 8
1	2	3	4	5	6	7	8	9	10
4	.../mazowieckie/ ciechanowsko-płocki	2000	1	14676	18309,9	-3633,9	24,8	42936,8	13,9
		2001	2	17371	17949,9	-578,9	3,3		
		2002	3	18032	21054,6	-3022,6	16,8		
		2003	4	18036	20813,4	-2777,4	15,4		
		2004	5	19800	19775,5	24,5	0,1		
		2005	6	26813	21188,3	5624,7	21,0		
		2006	7	27751	23393,1	4357,9	15,7		
5	ostrołęcko-siedlecki	2000	1	9842	11375,7	-1533,7	15,6	44576,3	6,9

		2001	2	11449	10302,7	1146,3	10,0		
		2002	3	14008	14739,0	-731,0	5,2		
		2003	4	14009	14357,1	-348,1	2,5		
		2004	5	14200	14763,5	-563,5	4,0		
		2005	6	16930	16338,5	591,5	3,5		
		2006	7	18683	17237,7	1445,3	7,7		
6	warszawski	2000	1	14089	15899,6	-1810,6	12,9	45208,5	6,5
		2001	2	17366	15487,3	1878,7	10,8		
		2002	3	20040	18393,1	1646,9	8,2		
		2003	4	20346	20190,8	155,2	0,8		
		2004	5	20520	21423,5	-903,5	4,4		
		2005	6	22949	24427,3	-1478,3	6,4		
		2006	7	25545	25039,9	505,1	2,0		
7	radomski	2000	1	9789	12635,8	-2846,8	29,1	32896,1	9,5
		2001	2	11827	10626,9	1200,1	10,1		
		2002	3	14879	15652,4	-773,4	5,2		
		2003	4	14680	14607,2	72,8	0,5		
		2004	5	15464	16204,5	-740,5	4,8		
		2005	6	17546	16340,2	1205,8	6,9		
		2006	7	19020	17140,9	1879,1	9,9		
8	miasto Warszawa	2000	1	48217	53589,8	-5372,8	11,1	71474,5	4,1
		2001	2	52630	51547,6	1082,4	2,1		
		2002	3	58449	61157,8	-2708,8	4,6		
		2003	4	59196	59245,8	-49,8	0,1		
		2004	5	62896	63332,3	-436,3	0,7		
		2005	6	68140	67131,0	1009,0	1,5		
		2006	7	77001	70531,4	6469,6	8,4		
Średnia dla regionu centralnego w %:							7,5		

Gdzie: $\hat{y}$ - wartość teoretyczna obliczona z uwzględnieniem parametrów i wartości zmiennych objaśniających, β_i - efekt indywidualny podregionu.

Źródło: Opracowanie własne w Excel-u.

W dotychczasowych rozważaniach przyjęliśmy *a priori*, że modelem adekwatnym jest model oparty o efekty indywidualne podregionów. Może jednak wystąpić odniesienie do efektów losowych lub wystarczającym jest model bazujący na *KMNC* i stałej *const*. Aby to rozstrzygnąć program *Gretl* proponuje diagnostykę realizowaną poleceniem:

Testy/test diagnostyczny panelu.

Test ten przeprowadzamy po realizacji modelu wykonanego *KMNC* na pełnych danych panelowych, tj. 45. jednostek w przekroju dla siedmiu okresów i stałej *const*, a ponadto przy występowaniu wszystkich potencjalnych zmiennych objaśniających $X1$ - $X11$.

W sytuacji wspólnej stałej dla wszystkich podregionów program *Gretl* proponuje następujący model zamieszczony w tabeli 3.10.

Tab. 3.10. Oszacowanie parametrów równania zmiennej Y przy wspólnej stałej $const$

Zmienna	Współczynnik	Błąd standardowy	Wartość p
const:	37647	(3134,2)	[0,00000]
$X1$	-9,4315	(1,9387)	[0,00000]
$X2$	2489,9	(398,22)	[0,00000]
$X3$	-930,44	(235,35)	[0,00010]
$X4$	-4,3849	(3,8839)	[0,25995]
$X5$	-33,905	(3,3558)	[0,00000]
$X6$	-8,4642	(5,2844)	[0,11043]
$X7$	77,376	(16,277)	[0,00000]
$X8$	-71,852	(16,7)	[0,00002]
$X9$	403,34	(157,84)	[0,01118]
$X10$	-141,22	(85,046)	[0,09801]
$X11$	-36,318	(6,2422)	[0,00000]

Źródło: Opracowanie własne w *Gretl*.

Odpowiada także na pytania:

Czy model ustalony ze stałą $const$ jest właściwszy niż z efektami indywidualnymi (niełosowymi) jednostek przekrojowych, czyli uwzględniający zróżnicowanie wyrazu wolnego?

Następuje badanie przez program *Gretl* łącznej istotności zróżnicowania wyrazu wolnego bazując na teście F otrzymując statystykę:

$$F = 11,5759 \text{ z wartością } p = 3,27261e-040.$$

Niska wartość p oznacza odrzucenie hipotezy H_0 , że model panelowy ze stałą $const$ zbudowany według *KMNC* jest poprawny, wobec hipotezy H_1 , że model o ustalonych efektach indywidualnych odniesionych do podregionów jest właściwszy.

Czy model ustalony ze stałą $const$ jest właściwszy niż z efektami losowymi odniesionymi do siedmiu okresów czasowych grup panelu?

W tym względzie program *Gretl* bazuje na teście Breuscha-Pagana o statystyce:

$$LM = 150,016 \text{ z wartością } p = \text{prob}(\chi^2(1) > 150,016) = 1,7198e-034.$$

Biorąc pod uwagę niską wartość p należy odrzucić hipotezę H_0 , że model ze stałą $const$ jest poprawny, wobec hipotezy H_1 , że model o losowych efektach jest właściwszy.

Mamy więc dylemat, który model efektów obrać niełosowe, czy losowe. W podjęciu decyzji pomaga nam test *Hausmana* proponowany w tym celu przez program *Gretl*, który podaje statystykę:

$$H = 75,631 \text{ z wartością } p = \text{prob}(\chi^2(11) > 75,631) = 1,02556e-011.$$

Występuje tu niska wartość p co pozwala nam na odrzucenie hipotezy zerowej o poprawności modelu z losowymi efektami, wobec hipotezy alternatywnej o poprawności modelu z ustalonymi efektami. Tak więc możemy przypuszczać, że zbudowany model produktu brutto wg cen bieżących w przeliczeniu na 1. mieszkańca z efektami indywidualnymi dla podregionów w okresach rocznych jest poprawny.

* * *

Niniejsze opracowanie należy traktować jako propozycję możliwości szybkiego sformułowania innych modeli miar ekonomicznych bazując na danych przestrzenno-czasowych. Może to być załącznikiem do sformułowania dalszych podstawowych charakterystyk rynków, a w przyszłości zbudowania pełnego modelu predykcji zjawisk ekonomicznych podregionów.

Procedurę postępowania możemy także wykorzystać do formułowania modeli ekonometrycznych zjawisk ekonomicznych występujących w innych obiektach struktury terytorialnej przykładowo do województw.

4. Zastosowanie miary Penrose'a do oceny różnic w rozwoju podregionów



4.1. Wstęp

Odległość Penrose'a stanowi jedną z miar odległości między populacjami a określona jest wzorem³⁴:

$$d_{kl} = \frac{1}{p} \sum_{j=1}^p \frac{(\mu_{jk} - \mu_{jl})^2}{\sigma_j^2}$$

Gdzie:

σ_j^2 - wariancje $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$

k oraz l - populacje ($k, l = 1, \dots, m$) cech, czyli zmiennych $X_1, X_2, \dots, X_p$.

Formuła miary Penrose'a nie uwzględnia wzajemnego skorelowania cech. Rozkłady tych zmiennych p -wymiarowych mogą się różnić w poszczególnych populacjach o numerze k ($k = 1, \dots, m$) i odpowiadają im wektory wartości oczekiwanych:

$$\mu_k = \begin{bmatrix} \mu_{1k} \\ \mu_{2k} \\ \dots \\ \mu_{pk} \end{bmatrix}$$

oraz macierze kowariancji Σ_k zmiennych $X_1, X_2, \dots, X_p$.

Dla przedstawienia procedury stosowania miary Penrose'a bazujemy na danych statystycznych dotyczących poszczególnych 45. podregionów Polski (zob. tabela 2.6 zamieszczona - ze względu na obszerność - na zakończenie tego opisu). Dla wygody posługiwania się programem *Excel* zmiennym nadano symbole:

- Y – produkt brutto (ceny bieżące) na 1. mieszkańca w roku poprzednim w zł;
- X_1 – ludność na 1 km² powierzchni ogólnej;
- X_2 – przyrost naturalny na 1000 ludności;
- X_3 – saldo migracji wewnętrznych i zagranicznych na pobyt stały (na 1000 ludności);
- X_4 – przestępstwa stwierdzone na 1000 ludności;
- X_5 – pracujący w tys. (rolnictwo, łowiectwo i leśnictwo, rybactwo);
- X_6 – pracujący w tys. (przemysł i budownictwo);
- X_7 – pracujący w tys. (usługi rynkowe i nierynkowe);
- X_8 – bezrobotni zarejestrowani ogółem w tys.;
- X_9 – mieszkania oddane do użytkowania na 1000 ludności;
- X_{10} – miejsca noclegowe w turystycznych obiektach zbiorowego zakwaterowania w tys.;
- X_{11} – užitki rolne w tys. ha w roku poprzednim.

W tym przykładzie pominięto cechę Y - produkt brutto (ceny bieżące) na 1. mieszkańca w roku poprzednim w zł, gdyż jest to zmienna zależna modelu ekonometrycznego produktu krajowego brutto podregionów³⁵. Ponadto przyjęto tylko cechy mające te same odniesienie na 1000 ludności, a należą do nich:

³⁴ Ostasiewicz W. (red), *Statystyczne metody analizy danych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 1999, rozdz. 2.4.2.

³⁵ Wornalkiewicz W., *Modele ekonometryczne PKB obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013, rozdz. 7. *Etapy i procedury budowy modelu produktu brutto podregionów*.

X_2 - przyrost naturalny,

X_3 - saldo migracji wewnętrznych i zagranicznych na pobyt stały,

X_4 - przestępstwa stwierdzone,

X_9 - mieszkania oddane do użytkowania.

Sprawdźmy wzajemne skorelowanie przyjętych czterech cech³⁶:

	X_2	X_3	X_4	X_9
X_2	1	0,02006	-0,3258	-0,0874
X_3		1	0,0651	0,5859
X_4			1	0,4664
X_9				1

Nie występują zatem cechy istotnie - silnie wzajemnie skorelowane, co pozwala nam na mierzenie odległości między dwoma populacjami miarą Penderosa'a.

4.2. Macierze „cząstkowe” jako próby

Dostępne wcześniej zebrane dane statystyczne GUS-u dotyczące podregionów obejmują siedem lat (2000-2006). Przyjmijmy, że poszczególne lata oraz cztery cechy X_2 , X_3 , X_4 , X_9 tworzą macierze o wymiarach (45 x 4) nazwane jako X_1 - X_7 , które potraktujemy jako próby dla 45. podregionów. W tym celu tworzymy w *Excelu* arkusz danych (*Dane-gretl-Penerosa.xls*) obejmujący wymienione zmienne oraz poszczególne macierze. Następnie sortujemy tabelę rosnąco według lat (2000-2006). Zawartość pierwszej macierzy X_1 dla roku 2000 podano w tabeli 4.1.

Tab. 4.1. Macierz X_1

-2,3	0,4	223,9	1,1
-0,9	-1,0	216,3	1,7
-7,0	-0,9	472,4	1,5
0,5	-1,4	222,5	2,0
1,5	-1,6	229,9	1,8
-0,4	7,0	294,7	5,3
1,4	-1,3	261,2	1,8
-3,7	2,1	575,0	8,9
1,4	0,7	223,6	1,7
5,2	0,1	182,3	2,0
-1,5	1,3	603,8	4,6
-2,0	-0,6	276,2	1,2
1,2	-1,4	257,9	1,9
-1,7	-4,4	440,5	0,5
1,4	-3,5	348,4	1,5
0,5	-2,3	246,1	1,1
-0,8	-2,1	220,0	0,9
0,2	-0,7	258,2	1,6
2,9	-0,6	186,1	1,5
2,2	-1,4	135,0	1,2
-0,4	-0,6	290,4	3,8
2,0	-3,4	230,1	1,2

³⁶ Ibidem, tabela 7.4.

-0,6	-1,6	250,5	1,4
1,1	-0,7	442,3	2,7
1,1	-0,5	388,3	1,7
2,3	-1,4	291,3	1,6
1,9	2,8	226,3	2,4
0,9	-0,4	183,4	1,4
1,8	-0,9	237,6	1,3
-2,3	-0,6	689,8	3,9
0,7	0,1	446,7	2,8
1,8	-1,2	364,1	1,7
-1,4	-1,6	365,5	1,3
1,7	-1,4	440,3	1,2
1,2	1,9	187,4	1,7
-2,6	0,3	599,2	4,3
0,0	-3,4	281,1	1,0
0,5	-0,3	352,5	1,7
1,1	-0,7	316,2	1,9
3,4	-1,3	330,1	2,2
4,7	1,3	288,1	2,8
-1,1	-0,3	776,2	2,5
2,8	-1,4	296,2	1,9
2,3	-1,4	378,0	3,1
2,8	-2,7	325,1	2,5

Źródło: Opracowanie własne w *Excelu* na podstawie roczników statystycznych GUS-u, lata 2001-2007.

Jak już wspomniałem takich macierzy jest siedem odrębnie dla roku 2000, 2001, 2002, 2003, 2004, 2005, 2006. Poszczególne w kolejności wiersze odpowiadają następującym 45. podregionom (czytany najpierw pionowo od „łódzki” do „rybnicko-jastrzębski” i dalej analogicznie druga kolumna oraz trzecia pokazana poniżej.

łódzki	białskopodlaski	szczeciński
piotrkowsko-skierniewicki	chełmsko-zamojski	koszaliński
miasto Łódź	lubelski	jeleniogorski-wałbrzyski
ciechanowsko-płocki	rzeszowsko-tarnobrzeski	legnicki
ostrołęcko-siedlecki	krośniesko-przemyski	wrocławski
warszawski	białostocko-suwański	miasto Wrocław
radomski	łomżyński	opolski
miasto Warszawa	świętokrzyski	bydgoski
krakowsko-tarnowski	gorzowski	toruńsko-włocławski
nowosądecki	zielonogórski	słupski
miasto Kraków	piłski	gdański
częstochoński	poznański	Gdańsk-Gdynia-Sopot
bielsko-bialski	kaliski	elbląski
centralny śląski	koniński	olsztyński
rybnicko-jastrzębski	miasto Poznań	elcki

Dane dotyczące podregionów potraktujemy przykładowo jako próby o stałej liczności ($n = n_1, n_2, n_3, n_4, n_5, n_6, n_7 = 45$) opisujące za pomocą czterech cech X_2, X_3, X_4, X_9 rozwój podregionów w poszczególnych latach 2000-2006. Dla obliczenia wartości średnich oraz wariancji danej cechy x_{ij} określonej kolumny macierzy k próby skorzystajmy z funkcji *Excela* ŚREDNIA(...) oraz WARIANCJA.POPULACJI(...). Obliczenia przeprowadzane są według wzorów³⁷:

- średnia arytmetyczna cechy X_j ($j = 1, \dots, p$):

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$$

- wariancja cechy w ramach danej k -tej próby:

$$s_j^2 = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$$

Obliczone w *Excelu* średnie arytmetyczne siedmiu prób dla lat 2000-2006 ($k = 1-7$) i 45. obserwacji (podregionów) podano w tabeli 4.2.

Tab. 4.2. Średnie arytmetyczne cech X_2, X_3, X_4, X_9

X_2	X_3	X_4	X_9	k
0,53	-0,69	330,02	2,17	1
0,40	-0,67	359,62	2,66	2
0,12	-0,69	361,75	2,49	3
-0,14	-0,61	381,59	4,12	4
0,01	-0,52	378,89	2,81	5
0,11	-0,68	356,51	2,93	6
0,31	-1,36	333,49	3,01	7

Źródło: Opracowanie własne w *Excelu*.

Wiersze poziome stanowią wektory średnich $\bar{x}_1$ do $\bar{x}_7$ poszczególnych prób. Przykładowo wektor $\bar{x}_1$ jest następujący:

$$\bar{x}_1 = \begin{bmatrix} 0,53 \\ -0,69 \\ 330,02 \\ 2,17 \end{bmatrix}$$

Na podstawie macierzy obserwacji cech podregionów dla prób $k = 1 - 7$ analogicznie jak dla średnich arytmetycznych obliczono w *Excelu* wariancje cech (zob. tabela 4.3).

Tab. 4.3. Wariancje cech X_2, X_3, X_4, X_9

X_2	X_3	X_4	X_9	k
4,80	3,37	19507,77	2,04	1
4,23	4,13	18690,81	3,78	2
3,89	5,19	16785,12	1,95	3
3,33	5,58	13941,79	3,12	4
3,04	6,01	14158,81	1,37	5
3,27	6,44	10630,38	2,45	6
3,34	8,49	7639,01	3,07	7

Źródło: Opracowanie własne w *Excelu*.

³⁷ Ostasiewicz W. (red), *Statystyczne metody analizy danych*, op. cit., rozdz.2.2.2. Wzory średnich oraz macierze kowariancji i korelacji.

Z tabeli 2.3. wynika, że dla próby pierwszej (rok 2000) wariancje czterech cech są następujące:

s_{21}^2	s_{31}^2	s_{41}^2	s_{91}^2
4,80	3,37	19505,77	2,04

Oznaczenie symbolami wariancji dla kolejnych prób jest analogiczne. W kolejnym kroku obliczamy wariancje „sumaryczne” cech. Stanowią one główną przekątną macierzy kowariancji na podstawie wzoru³⁸:

$$C = \sum_{k=1}^m n_k C_k \left(\sum_{k=1}^m n_k - m \right)$$

Gdzie:

n_k - liczebność próby k ($k = 1, 2, \dots, m$),

C_k - macierz wariancji i kowariancji w k -tej próbie.

Kowariancję cech X_j oraz X_k w próbie obliczamy według wzoru:

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{22} \\ \dots & \dots & \dots & \dots \\ c_{p1} & c_{p2} & \dots & c_{p2} \end{bmatrix}$$

Powyższą macierz C danych czasowo-przekrojowych zwaną macierzą wariancji i kowariancji z próby dla czterech cech o elementach $c_{jk} = s_j^2$ ($j, k = 1, \dots, p$), czyli wariancji na głównej przekątnej można przedstawić jako:

$$c_{jk} = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

Przykładowo pierwszy element głównej przekątnej c_{jk} macierzy C , czyli wariancję cechy $X_{j=2}$ dla siedmiu prób k o stałej liczności $n = 45$ obliczamy następująco³⁹:

$$s_2^2 = c_{11} = (n_1 s_{21}^2 + n_2 s_{22}^2 + n_3 s_{23}^2 + n_4 s_{24}^2 + n_5 s_{25}^2 + n_6 s_{26}^2 + n_7 s_{27}^2) / (n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7 - 7) \\ = 45 * (s_{21}^2 + s_{22}^2 + s_{23}^2 + s_{24}^2 + s_{25}^2 + s_{26}^2 + s_{27}^2) / 308$$

W tym względzie przeprowadzmy obliczenia w *Excelu* sumując dane w kolumnach tabeli 2.3 dla cech X_2, X_3, X_4 i X_9 i przemnażając przez współczynnik $(45/308)$ otrzymamy:

s_2^2	s_3^2	s_4^2	s_9^2
3,78	5,73	14808,17	2,60

Teraz przystąpimy do obliczenia miar Penrose’a, przy czym zastąpimy wektory wartości oczekiwanych oraz wariancji w populacji estymatorami (oszacowaniami) z prób k i l ($k, l = 1, 2, \dots, m = 7$). Są nimi wektory średnich arytmetycznych $\bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_9$ oraz obliczone wcześniej wariancje $s_2^2, s_3^2, s_4^2, s_9^2$ naszych cech X_2, X_3, X_4 oraz X_9 . Zatem wzór na ocenę różnic w rozwoju podregionów przy przyjętych przykładowo $p = 4$ cechach jest następujący:

$$d_{kl} = \frac{1}{p} \sum_{j=1}^p \frac{(\bar{x}_{jk} - \bar{x}_{jl})^2}{s_j^2} = \frac{1}{4} \sum_{j=1}^4 \frac{(\bar{x}_{jk} - \bar{x}_{jl})^2}{s_j^2}$$

³⁸ Ostasiewicz W. (red), *Statystyczne metody analizy danych*, op. cit., rozdz.2.4. *Odległości między obiektami, populacjami i próbami*.

³⁹ Ibidem, s. 61.

Przykładowo przy porównaniu roku 2000 oraz 2001 ($k = 1$ i $l = 2$) odległość (miara) Penrose'a dla naszego przykładu wyrażona jest wzorem⁴⁰:

$$d_{12} = \frac{1}{4} \left[(\bar{x}_{21} - \bar{x}_{22})^2 / s_2^2 + (\bar{x}_{31} - \bar{x}_{32})^2 / s_3^2 + (\bar{x}_{41} - \bar{x}_{42})^2 / s_4^2 + (\bar{x}_{91} - \bar{x}_{92})^2 / s_9^2 \right]$$

Dla lat (2000 -2006) powstanie macierz kwadratowa odległości o wymiarach (7x7). Fragment arkusza pomocniczego w formie założonego w *Excelu* pliku *Dane-grejl-Penrosa.xls* zawiera tabela 4.4. Obejmuje ona rok 2000 oraz 2001. W pierwszym wierszu wpisano wariancje s^2 dotyczące cech X_2, X_3, X_4 oraz X_9 , które wcześniej uśrednione są stałe dla poszczególnych siedmiu prób lat 2000-2006. Macierz obliczeń jest symetryczna, a na przekątnej tej macierzy w odniesieniu do układu dwóch zmiennych znajdują się obliczone wartości według przykładowej podanej formuły (zob. powiększony zapis 0,004):

$$=((\$C\$3-B9)^2)/\$C\$1.$$

Odległość Penrose'a jest średnią arytmetyczną z obliczeń cząstkowych dotyczących par zmiennych (zob. komórke B3):

$$=(C4+D5+E6+F7)/4.$$

Skrótem „Śr.” oznaczono średnią arytmetyczną danej zmiennej w określonej próbie.

Tab. 4.4. Fragment arkusza pomocniczego do obliczania odległości Penrose'a

⇒

W/K	A	B	C	D	E	F	G	H	J	K	L	M
1		s^2	3,78	5,73	14808,17	2,60		s^2	3,78	5,73	14808,17	2,60
2	Rok	2000	Śr. X_2	Śr. X_3	Śr. X_4	Śr. X_9	Rok	2001	Śr. X_2	Śr. X_3	Śr. X_4	Śr. X_9
3	2000 - d	0	0,53	-0,69	330,02	2,17	2000 - d	0,0383	0,40	-0,67	359,62	2,66
4	Śr. X_2	0,53	0				Śr. X_2	0,53				
5	Śr. X_3	-0,69		0			Śr. X_3	-0,69				
6	Śr. X_4	330,02			0		Śr. X_4	330,02				
7	Śr. X_9	2,17				0	Śr. X_9	2,17				
8	2001 - d	0,0383					2001 - d	0				
9	Śr. X_2	0,40	0,004				Śr. X_2	0,40	0			
10	Śr. X_3	-0,67		0,000			Śr. X_3	-0,67		0		
11	Śr. X_4	359,62			0,059		Śr. X_4	359,62			0	
12	Śr. X_9	2,66				0,089	Śr. X_9	2,66				0
13	2002 - d	0,0379					2002 - d	0,0078				
14	Śr. X_2	0,12	0,044				Śr. X_2	0,12	0,021			
15	Śr. X_3	-0,69		0,000			Śr. X_3	-0,69		0,000		
16	Śr. X_4	361,75			0,068		Śr. X_4	361,75			0,000	
17	Śr. X_9	2,49				0,039	Śr. X_9	2,49				0,010
18	2003	0,4396					2003	0,2341				
19	Śr. X_2	-0,14	0,119				Śr. X_2	-0,14	0,078			
20	Śr. X_3	-0,61		0,001			Śr. X_3	-0,61		0,001		
21	Śr. X_4	381,59			0,180		Śr. X_4	381,59			0,033	
22	Śr. X_9	4,12				1,458	Śr. X_9	4,12				0,825

⁴⁰ Ibidem, s. 61.

23	2004	0,0990					2004	0,0199				
24	Śr. X_2	0,01	0,072				Śr. X_2	0,01	0,041			
25	Śr. X_3	-0,52		0,005			Śr. X_3	-0,52		0,004		
26	Śr. X_4	378,89			0,161		Śr. X_4	378,89			0,025	
27	Śr. X_9	2,81				0,158	Śr. X_9	2,81				0,010
28	2005	0,0781					2005	0,0128				
29	Śr. X_2	0,11	0,047				Śr. X_2	0,11	0,022			
30	Śr. X_3	-0,68		0,000			Śr. X_3	-0,68		0,000		
31	Śr. X_4	356,51			0,047		Śr. X_4	356,51			0,001	
32	Śr. X_9	2,93				0,218	Śr. X_9	2,93				0,028
33	2006	0,0898					2006	0,0445				
34	Śr. X_2	0,31	0,013				Śr. X_2	0,31	0,002			
35	Śr. X_3	-1,36		0,078			Śr. X_3	-1,36		0,082		
36	Śr. X_4	333,49			0,001		Śr. X_4	333,49			0,046	
37	Śr. X_9	3,01				0,267	Śr. X_9	3,01				0,047

Źródło: Opracowanie własne w Excelu.

Pełną macierz odległości Penrose'a podano w tabeli 4.5.

Tab. 4.5. Macierz odległości Penrose'a wyznaczonych dla cech (zmiennych) X_2 , X_3 , X_4 i X_9 .

Rok	2000	2001	2002	2003	2004	2005	2006
2000	0	0,0383	0,0379	0,4396	0,0990	0,0781	0,0898
2001	0,0383	0	0,0078	0,2341	0,0199	0,0128	0,0445
2002	0,0379	0,0078	0	0,2661	0,0170	0,0186	0,0604
2003	0,4396	0,2341	0,2661	0	0,1662	0,1520	0,1960
2004	0,0990	0,0199	0,0170	0,1662	0	0,0115	0,0751
2005	0,0781	0,0128	0,0186	0,1520	0,0115	0	0,0323
2006	0,0898	0,0445	0,0604	0,1960	0,0751	0,0323	0

Źródło: Opracowanie własne w Excelu.

Przeglądając kolumny poszczególnych lat widzimy, że wyraźna różnica w rozwoju wszystkich podregionów Polski - ze względu na niżej podane cechy na 1000 ludności - nastąpiła w stosunku do roku 2003:

X_2 - przyrost naturalny,

X_3 - saldo migracji wewnętrznych i zagranicznych na pobyt stały,

X_4 - przestępstwa stwierdzone,

X_9 - mieszkania oddane do użytkowania.

Jesli jednak rozpatrzmy rok 2006 (z pominięciem roku 2003) to następuje nierównomierny spadek różnicy w rozwoju wszystkich podregionów (zob. odniesienie roku 2006 do 2000 oraz 2006 do 2005). Oczywiście dyskusyjny może być dobór cech, ale w tym materiale chciano przede wszystkim przedstawić procedurę postępowania określania miary Penrose'a z zastosowaniem możliwości arkusza kalkulacyjnego Excel.

Dane empiryczne 11. cech (zmiennych: X_1 - X_{11}) objaśniających oraz zmiennej objaśnianej Y dla 7 lat oraz 45. podregionów zaprezentowano w tabeli 4.6.

Tab. 4.6. Dane empiryczne dotyczące podregionów⁴¹

T	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	Rok	Podregion
1	12790	110,0	-2,3	0,4	223,9	164,2	75,3	83,7	77,6	1,1	9,4	629,9	2000	łódzki
2	14350	109,0	-2,6	0,1	269,6	164,1	75,3	83,7	86,1	1,5	5,0	629,9	2001	
3	15259	108,0	-2,8	0,4	278,8	93,6	66,4	81,0	87,4	1,6	5,2	575,5	2002	
4	15955	108,0	-2,9	0,8	287,4	93,3	67,7	80,5	88,3	3,2	4,9	564,1	2003	
5	18713	108,0	-2,4	0,9	287,0	93,5	68,5	81,1	83,2	2,1	4,9	564,1	2004	
6	19837	107,0	-3,0	0,6	275,0	93,5	72,5	81,0	77,6	2,0	4,8	552,7	2005	
7	21681	107,0	-2,5	0,5	256,0	93,5	76,5	82,5	65,3	1,8	4,7	552,7	2006	
1	13478	97,0	-0,9	-1,0	216,3	177,3	76,7	79,3	75,0	1,7	8,6	611,8	2000	piotrkowsko-skierniewicki
2	15289	97,0	-1,3	-1,0	265,5	176,8	74,9	74,9	82,6	2,2	9,2	610,9	2001	
3	16060	86,0	-1,1	-1,1	268,5	96,5	70,2	77,8	82,3	2,3	8,2	548,1	2002	
4	17034	95,0	-1,7	-1,0	275,0	96,1	69,6	78,1	79,6	4,9	7,0	553,8	2003	
5	20040	95,0	-1,4	-1,4	266,0	96,2	71,8	79,6	73,4	2,5	7,3	553,8	2004	
6	21657	95,0	-1,3	-1,1	246,0	96,2	72,6	81,6	66,2	2,3	7,0	514,2	2005	
7	22849	94,7	-0,8	-1,8	257,0	96,2	73,7	82,9	56,7	2,3	7,5	514,2	2006	
1	21767	2694,0	-7,0	-0,9	472,4	3,9	73,5	140,7	55,5	1,5	5,7	11,0	2000	miasto Łódź
2	23924	2672,0	-6,7	-0,9	450,0	3,9	72,6	141,5	61,9	2,1	4,8	10,6	2001	
3	25072	2671,0	-6,3	-1,2	489,3	2,2	64,1	135,7	65,1	1,6	4,6	23,1	2002	
4	26531	2647,0	-6,2	-1,2	538,7	2,1	60,0	137,7	65,2	1,6	3,3	21,2	2003	
5	29104	2629,0	-5,8	-1,5	550,0	2,1	59,1	134,5	61,7	1,8	3,5	21,2	2004	
6	30602	2608,0	-6,1	-1,8	512,0	2,1	59,3	138,9	54,7	0,9	3,5	19,2	2005	
7	33269	2593,0	-6,5	-2,4	434,0	2,1	60,9	147,9	38,7	1,6	3,6	19,2	2006	
1	17371	83,0	0,5	-1,4	222,5	117,8	47,0	62,2	61,4	2,0	2,9	572,5	2000	ciechanowsko-płocki
2	18032	83,0	0,5	-1,6	268,0	117,8	44,1	60,8	67,9	2,0	3,2	572,5	2001	
3	18036	81,0	0,6	-1,6	292,5	66,3	43,3	60,2	71,6	2,2	2,9	540,7	2002	
4	19800	81,0	-0,1	-1,8	318,5	65,5	41,2	60,7	71,3	3,7	2,7	524,1	2003	
5	26813	81,0	-0,4	-1,5	316,0	65,8	42,6	61,1	67,0	3,0	2,8	524,1	2004	
6	27751	80,5	-0,4	-2,0	289,0	65,9	43,6	63,0	62,4	2,4	2,6	507,4	2005	
7	29265	80,5	0,1	-1,9	273,0	65,9	46,5	64,4	51,9	2,8	3,0	507,4	2006	
1	11449	64,0	1,5	-1,6	229,9	181,8	37,6	65,6	61,0	1,8	6,0	808,1	2000	ostrolecko-siedlecki
2	14008	64,0	1,2	-1,8	261,3	181,5	35,1	64,0	68,4	2,0	5,5	806,9	2001	
3	14009	62,0	0,9	-2,2	257,1	103,9	34,2	64,3	69,7	1,6	5,0	749,0	2002	
4	14200	62,0	0,1	-2,7	304,5	103,4	34,8	63,9	68,1	2,5	5,0	731,7	2003	
5	16930	62,0	0,4	-2,1	303,0	103,7	33,1	64,1	65,8	2,7	5,3	731,7	2004	
6	18683	62,1	0,5	-2,6	296,0	103,7	34,2	65,7	62,6	2,1	5,3	714,3	2005	
7	20452	62,0	0,5	-2,8	300,0	103,6	37,2	67,0	53,2	1,9	4,9	714,3	2006	
1	17366	159,0	-0,4	7,0	294,7	139,6	97,0	127,9	57,1	5,3	3,9	532,2	2000	warszawski
2	20040	160,0	-0,3	7,0	345,6	141,1	94,2	131,2	70,8	5,6	4,1	531,4	2001	
3	20346	162,0	-0,7	7,8	341,2	69,5	87,3	128,2	75,7	4,5	4,0	449,1	2002	
4	20520	164,0	-0,5	8,4	352,7	66,4	86,2	129,0	76,0	7,5	5,6	435,0	2003	
5	22949	165,0	0,1	8,5	366,0	67,1	85,8	131,2	72,8	6,1	5,2	435,0	2004	
6	25545	166,0	0,3	8,6	377,0	67,0	88,6	140,9	57,4	6,7	4,8	420,8	2005	
7	27965	168,0	0,8	9,2	352,0	66,9	90,5	141,7	57,1	6,4	5,0	420,8	2006	
1	11827	106,0	1,4	-1,3	261,2	147,3	413,2	64,9	76,4	1,8	3,8	464,6	2000	radomski
2	14879	106,0	0,9	-1,4	293,2	148,8	398,0	64,9	85,2	1,4	3,3	463,9	2001	
3	14680	104,0	0,7	-1,9	301,7	79,4	39,3	64,5	86,7	1,0	3,4	391,7	2002	
4	15464	104,0	-0,3	-1,8	342,2	78,9	37,1	64,8	85,7	4,0	3,2	383,7	2003	
5	17546	104,0	0,5	-1,9	346,0	79,2	37,6	64,7	82,8	2,1	3,0	383,7	2004	

⁴¹ W opracowaniu danych bazowano na aneksie A2. *Dane czasowo-przekrojowe do modelu PKB podregionów* zawartym w książce: Wornakiewicz W., *Modele ekonometryczne PKB obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013.

6	19020	104,0	0,3	-2,0	330,0	79,3	38,2	65,7	81,7	2,3	2,9	375,7	2005	
7	20343	104,0	0,2	-2,6	309,0	79,3	39,5	64,5	73,9	2,2	2,9	375,7	2006	
1	52630	3258,0	-3,7	2,1	575,0	485,5	190,1	614,1	34,0	8,9	16,4	14,8	2000	miasto Warszawa
2	58449	3257,0	-3,6	2,9	567,5	467,1	176,7	597,1	53,8	10,1	18,2	14,7	2001	
3	59196	3265,0	-3,2	3,8	528,5	5,1	161,5	592,6	65,1	7,7	20,0	118,4	2002	
4	62896	3269,0	-3,0	4,3	537,5	3,8	149,3	564,2	62,5	7,3	19,8	118,0	2003	
5	68140	3275,0	-2,1	4,1	543,0	4,1	144,0	596,9	64,6	6,1	21,9	118,0	2004	
6	77001	3284,0	-1,5	4,9	535,0	4,2	138,0	622,0	58,4	8,5	21,3	117,5	2005	
7	83933	3291,0	-0,6	4,8	466,0	4,4	143,6	630,1	49,4	8,1	21,0	117,5	2006	
1	12723	190,0	1,4	0,7	223,6	245,5	99,4	117,4	83,0	1,7	6,8	492,4	2000	krakowsko-tarnowski
2	13603	190,0	0,9	0,4	54,0	244,8	93,9	114,7	94,2	1,9	6,1	491,6	2001	
3	14376	189,0	0,8	1,0	266,0	105,5	88,3	113,2	91,4	2,4	5,7	428,2	2002	
4	15275	189,0	0,1	1,6	396,2	105,0	87,6	114,4	91,6	8,0	6,1	406,0	2003	
5	16749	189,0	0,5	1,5	327,0	105,2	88,9	115,3	84,9	2,5	5,7	406,0	2004	
6	17657	190,0	0,1	1,4	293,0	105,1	89,1	117,5	78,3	2,5	5,4	383,7	2005	
7	19368	190,0	0,4	0,9	277,0	105,0	91,3	119,7	63,3	2,4	5,9	383,7	2006	
1	10313	147,0	5,2	0,1	182,3	228,0	56,8	99,0	77,7	2,0	49,4	375,2	2000	nowosądecki
2	11926	148,0	4,5	0,0	222,1	227,8	53,1	98,2	89,1	2,1	46,0	374,5	2001	
3	12214	147,0	3,9	-0,1	212,1	76,2	53,8	96,5	86,7	2,7	41,4	329,5	2002	
4	12958	147,0	3,6	0,1	229,0	76,0	54,4	96,8	86,5	7,3	44,5	316,7	2003	
5	14086	148,0	3,3	0,1	266,0	76,1	54,6	96,9	81,8	2,6	44,2	316,7	2004	
6	15155	148,0	3,5	-0,1	261,0	76,1	57,4	99,0	74,5	2,4	40,8	303,9	2005	
7	16180	149,0	3,3	-0,7	241,0	76,1	59,0	100,1	61,7	2,2	39,8	303,9	2006	
1	29871	2269,0	-1,5	1,3	603,8	4,5	85,9	176,4	66,1	4,6	10,3	14,6	2000	miasto Kraków
2	29998	2268,0	-1,5	2,9	582,3	4,4	77,3	174,3	29,7	7,4	11,7	14,6	2001	
3	31490	2317,0	-1,3	2,5	610,1	2,5	70,1	169,6	30,1	4,6	15,6	26,0	2002	
4	32559	2318,0	-1,3	1,6	647,1	2,2	68,6	174,0	29,9	5,4	14,5	23,8	2003	
5	37473	2317,0	-0,8	1,5	687,0	2,3	66,6	179,6	26,9	5,9	15,3	23,8	2004	
6	39964	2315,0	-0,8	2,0	597,0	2,3	65,5	185,1	25,3	6,0	15,3	21,6	2005	
7	44529	2314,0	-0,4	1,1	530,0	2,4	67,7	192,4	20,3	8,7	17,0	21,6	2006	
1	15351	179,0	-2,0	-0,6	276,2	98,9	93,2	111,4	65,5	1,2	6,6	323,9	2000	częstochowski
2	15351	179,0	-2,4	-0,8	320,0	98,8	89,5	112,0	76,0	1,4	5,9	322,1	2001	
3	18259	178,0	-2,5	-0,1	270,6	23,1	49,5	59,4	40,4	2,1	3,7	158,4	2002	
4	18884	177,0	-2,4	-0,3	303,1	22,8	48,8	58,2	40,6	4,4	3,6	149,9	2003	
5	21906	177,0	-2,5	-0,2	312,0	22,8	50,9	57,7	40,0	2,7	3,5	149,9	2004	
6	22844	176,0	-2,8	-0,6	307,0	22,8	51,2	58,7	37,0	2,1	3,9	141,4	2005	
7	23272	175,0	-2,5	-1,0	299,0	22,8	53,8	59,0	32,7	2,4	3,6	141,4	2006	
1	18364	324,0	1,2	-1,4	257,9	103,9	203,5	173,7	78,7	1,9	29,6	255,3	2000	bielsko-bialski
2	18364	324,0	0,9	-1,2	273,0	103,2	192,7	179,9	94,1	2,0	23,2	254,8	2001	
3	21314	274,0	0,2	1,1	291,5	12,4	59,5	77,0	36,5	4,1	17,7	101,3	2002	
4	22353	274,0	-0,3	1,4	308,1	12,3	61,0	76,4	35,0	7,9	20,5	88,1	2003	
5	24807	275,0	0,5	1,9	336,0	12,3	61,3	76,3	34,1	3,9	20,0	88,1	2004	
6	25513	275,0	0,3	1,8	329,0	12,1	61,4	78,0	31,3	3,5	19,5	74,8	2005	
7	27301	276,0	1,0	1,0	310,0	12,2	64,0	78,2	25,0	3,3	19,1	74,8	2006	
1	22498	1729,0	-1,7	-4,4	440,5	10,3	273,3	322,4	115,6	0,5	6,3	39,5	2000	centralny śląski
2	22498	1719,0	-2,0	-3,5	469,1	10,0	256,5	323,2	143,2	0,8	5,7	38,5	2001	
3	24501	520,0	-1,8	-2,8	435,4	28,1	333,5	401,5	214,3	1,0	6,9	216,1	2002	
4	25819	518,0	-1,9	-2,7	463,4	27,2	314,5	393,4	210,4	2,0	9,0	210,5	2003	
5	29213	512,0	-1,8	-2,5	488,0	27,2	306,1	397,7	199,4	1,7	9,8	210,5	2004	
6	29906	513,0	-1,8	-2,6	461,0	26,9	314,9	401,6	180,4	1,4	9,9	204,8	2005	
7	32205	510,0	-1,5	-3,3	440,0	26,4	318,2	411,6	146,0	1,4	10,8	204,8	2006	

1	19162	477,0	1,4	-3,5	348,4	8,5	69,6	65,8	39,0	1,5	1,9	67,9	2000	rybnicko-jastrzębski
2	19162	477,0	1,4	-3,5	348,4	8,5	69,6	65,8	39,0	1,5	1,9	67,9	2001	
3	19162	477,0	1,4	-3,5	348,4	8,5	69,6	65,8	39,0	1,5	1,9	67,9	2002	
4	19700	475,0	0,9	-3,7	347,3	8,3	66,7	64,0	39,5	2,4	2,1	66,4	2003	
5	24852	474,0	1,1	-3,8	361,0	8,2	64,2	65,2	36,2	1,8	2,0	66,4	2004	
6	24789	473,0	1,4	-3,4	365,0	8,3	63,6	67,2	32,5	1,6	2,0	64,9	2005	
7	24859	472,0	1,7	-4,6	344,0	8,2	67,3	68,7	26,1	1,6	1,7	64,9	2006	
1	10172	54,0	0,5	-2,3	246,1	76,1	14,2	93,5	23,7	1,1	7,0	381,9	2000	bialskopodlaski
2	11680	54,0	0,3	-2,5	326,1	76,1	13,0	30,0	26,4	1,4	6,6	376,0	2001	
3	12386	53,0	-0,6	-2,6	347,3	42,9	12,7	30,1	26,4	1,9	6,6	327,9	2002	
4	12885	52,0	-0,8	-3,4	334,4	42,8	12,3	29,7	25,3	2,9	6,2	318,5	2003	
5	14373	52,0	-0,6	-3,0	307,0	42,9	12,1	29,9	24,3	2,2	6,8	318,5	2004	
6	14963	52,0	-0,4	-3,8	288,0	42,9	12,4	30,1	23,5	2,4	5,5	309,1	2005	
7	16143	51,8	-0,2	-4,3	266,0	42,9	12,7	30,5	21,5	1,6	5,4	309,1	2006	
1	10164	73,0	-0,8	-2,1	220,0	190,9	30,6	61,7	54,0	0,9	6,3	644,7	2000	chełmsko-zamojski
2	11689	73,0	-0,7	-2,1	250,6	190,4	28,0	58,8	60,6	1,6	5,6	643,8	2001	
3	12249	72,0	-2,0	-2,7	276,0	103,7	26,4	58,0	60,3	1,4	6,4	579,9	2002	
4	12827	71,0	-1,7	-2,6	270,9	103,6	26,5	57,1	59,9	2,5	6,7	561,3	2003	
5	14469	71,0	-1,8	-2,9	255,0	103,7	26,3	57,7	56,7	1,1	6,6	561,3	2004	
6	15213	70,8	-2,0	-3,0	236,0	103,6	27,5	58,4	53,1	1,7	5,9	542,7	2005	
7	16408	70,4	-1,8	-3,9	230,0	103,8	27,7	58,8	49,7	1,5	5,2	542,7	2006	
1	13759	125,0	0,2	-0,7	258,2	240,7	79,6	145,2	82,1	1,6	11,1	693,2	2000	lubelski
2	15183	125,0	0,0	-1,2	297,0	240,8	74,3	140,9	92,0	2,8	9,6	692,2	2001	
3	15920	124,0	-0,4	-1,0	299,2	131,1	70,0	137,5	92,2	1,6	9,1	665,1	2002	
4	16765	123,0	-0,2	-1,4	297,6	131,1	69,1	137,7	89,3	2,8	8,3	653,8	2003	
5	18653	123,0	-0,5	-1,3	310,0	131,2	68,7	142,1	63,8	2,2	8,5	653,8	2004	
6	19557	123,0	-0,3	-1,4	304,0	131,3	68,1	145,7	80,2	2,8	8,1	642,5	2005	
7	20739	123,0	0,1	-2,2	281,0	131,2	69,3	149,0	70,6	2,7	8,5	642,5	2006	
1	13684	156,0	2,9	-0,6	186,1	226,7	106,7	124,5	93,2	1,5	8,0	437,2	2000	rzeszowsko-tarnobrzeski
2	15174	156,0	2,2	-0,8	236,2	225,8	101,6	124,2	99,1	1,4	7,8	444,0	2001	
3	16099	154,0	2,0	-1,2	262,3	81,3	97,1	123,7	95,8	2,1	8,4	406,9	2002	
4	17132	154,0	1,4	-0,7	271,0	81,1	96,0	125,1	92,3	5,6	6,0	386,9	2003	
5	18811	154,0	1,7	-0,6	246,0	81,2	97,1	127,0	85,9	2,6	6,3	386,9	2004	
6	19796	154,0	1,5	-0,9	222,0	81,1	96,8	130,2	82,7	2,2	6,1	366,8	2005	
7	21107	154,0	1,6	-1,4	207,0	81,1	101,1	132,8	72,2	2,6	6,7	366,8	2006	
1	11299	92,0	2,2	-1,4	135,0	219,4	61,9	96,3	89,0	1,2	14,2	505,1	2000	króśniesko-przemyski
2	12280	92,0	1,5	-1,4	178,9	218,8	59,3	92,7	96,0	1,2	14,3	504,0	2001	
3	12707	91,0	1,2	-1,6	206,3	76,5	54,9	92,7	91,7	1,8	14,6	401,0	2002	
4	13353	91,0	1,0	-1,7	202,1	76,3	55,9	91,9	90,2	4,4	12,2	393,6	2003	
5	14520	91,0	0,6	-1,5	212,0	76,5	56,6	92,4	84,4	1,7	11,0	393,6	2004	
6	15321	91,0	0,6	-1,6	193,0	76,5	57,2	94,4	81,2	1,6	12,4	386,1	2005	
7	16456	90,9	1,0	-2,5	185,0	76,4	58,5	96,1	73,0	1,7	13,5	386,1	2006	
1	14082	61,0	-0,4	-0,6	290,4	154,4	53,0	106,4	54,3	3,8	13,6	831,3	2000	bialostocko-suwański
2	15501	61,0	-0,5	-0,5	284,5	154,0	48,5	104,3	60,0	3,1	11,5	831,3	2001	
3	16548	60,0	-0,8	-0,8	294,1	90,2	45,7	101,9	60,5	2,6	12,1	784,1	2002	
4	16999	60,0	-1,2	-1,0	324,4	90,0	46,6	102,0	57,2	4,1	11,0	764,8	2003	
5	18915	60,0	-1,2	-0,8	317,0	90,4	46,5	103,2	52,7	3,2	10,5	764,8	2004	
6	19972	59,9	-1,1	-0,9	279,0	90,4	45,8	104,6	51,3	3,1	9,9	745,5	2005	
7	21297	59,7	-1,0	-1,4	242,0	90,2	48,7	104,9	43,3	2,8	10,1	745,5	2006	
1	10585	60,0	2,0	-3,4	230,1	74,6	13,4	25,7	24,9	1,2	2,5	368,4	2000	łomżyński
2	12524	60,0	1,2	-3,1	262,4	74,5	12,6	25,1	26,5	0,9	2,1	368,4	2001	
3	13339	59,0	1,7	-3,3	245,4	47,9	12,1	24,6	25,6	1,0	1,9	365,7	2002	

4	13536	59,0	0,4	-3,2	251,0	47,9	12,3	24,0	24,1	1,8	1,9	358,9	2003	
5	15590	58,0	0,0	-2,9	263,0	48,1	11,9	24,5	23,4	1,6	2,1	358,9	2004	
6	16498	58,2	0,4	-3,7	226,0	48,1	11,7	24,8	21,9	1,4	1,9	352,0	2005	
7	17803	58,0	0,3	-5,0	201,0	48,1	11,8	25,4	18,5	1,7	1,6	352,0	2006	
1	13868	113,0	-0,6	-1,6	250,5	290,4	92,3	135,8	118,5	1,4	9,8	731,9	2000	świętokrzyski
2	14843	113,0	-0,8	-1,8	275,1	289,8	85,6	131,2	131,1	1,4	9,3	731,9	2001	
3	15777	111,0	-1,3	-1,9	30,1	143,6	78,2	128,8	131,3	2,1	9,6	629,3	2002	
4	16911	110,0	-1,6	-1,9	329,1	143,1	76,3	128,8	126,5	3,5	8,5	592,9	2003	
5	18714	110,0	-1,6	-1,8	329,0	143,6	75,4	130,9	126,3	1,5	8,3	592,9	2004	
6	19274	110,0	-1,7	-1,7	311,0	143,5	77,8	133,8	117,8	1,6	8,3	556,5	2005	
7	21130	109,0	-1,9	-2,7	320,0	143,5	81,4	133,9	99,4	1,3	7,9	556,5	2006	
1	16536	63,0	1,1	-0,7	442,3	22,9	32,0	46,7	31,3	2,7	7,9	234,5	2000	gorzowski
2	17875	63,0	1,7	-0,5	423,4	22,6	30,7	46,0	36,6	3,1	8,6	234,5	2001	
3	18633	62,0	1,1	-0,6	419,1	11,0	27,5	45,0	39,5	3,0	8,3	210,2	2002	
4	18748	62,0	0,9	0,2	437,4	10,8	28,2	45,0	39,6	4,0	7,9	203,5	2003	
5	21548	62,0	1,3	0,0	422,0	10,9	28,7	44,9	36,7	3,1	9,4	203,5	2004	
6	22813	62,4	1,4	-0,1	399,0	10,9	30,5	46,1	31,9	3,6	8,0	196,8	2005	
7	24620	62,4	1,3	-1,3	383,0	10,9	34,2	46,6	24,3	3,1	7,9	196,8	2006	
1	15517	81,0	1,1	-0,5	388,3	34,9	49,2	73,1	57,9	1,7	12,0	319,0	2000	zielonogórski
2	16880	81,0	1,2	-1,4	427,8	34,4	46,9	72,4	65,8	2,2	10,4	319,8	2001	
3	17377	80,0	0,6	-1,2	433,7	15,7	42,0	68,9	70,0	2,0	9,7	270,1	2002	
4	18183	80,0	-0,1	-1,1	425,2	15,2	42,0	69,2	68,5	3,7	10,2	276,7	2003	
5	21697	80,0	0,7	-0,6	417,0	15,6	43,8	70,7	62,7	2,2	10,9	276,7	2004	
6	23501	79,7	0,4	-0,9	387,0	15,4	45,9	73,6	57,2	3,1	9,9	283,3	2005	
7	29279	79,6	1,0	-1,7	385,0	15,4	47,9	74,7	48,5	2,8	11,6	283,3	2006	
1	14284	64,0	2,3	-1,4	291,3	38,3	36,6	38,7	33,2	1,6	4,4	335,2	2000	piłski
2	15899	64,0	2,2	-1,1	335,3	38,5	34,9	38,4	40,3	1,7	5,1	333,9	2001	
3	16379	63,0	2,5	-1,4	356,1	23,5	34,1	36,5	40,3	1,7	4,2	313,2	2002	
4	16931	63,0	1,9	-0,9	349,3	23,5	33,7	36,8	39,4	2,3	4,4	301,2	2003	
5	20361	63,0	1,8	-1,0	343,0	23,8	34,4	37,1	36,1	2,2	4,2	301,2	2004	
6	21504	62,9	2,3	-1,1	325,0	23,6	36,4	38,1	32,3	1,5	4,0	289,1	2005	
7	21931	62,9	2,3	-2,4	313,0	23,7	37,8	39,0	25,5	2,0	3,7	289,1	2006	
1	16933	98,0	1,9	2,8	226,3	110,7	112,0	106,8	50,5	2,4	20,1	732,7	2000	poznański
2	18374	98,0	2,1	3,2	254,1	108,5	107,5	105,0	67,1	2,4	19,1	730,4	2001	
3	11521	99,0	1,7	4,0	269,0	69,1	104,5	106,0	74,9	2,9	18,5	687,4	2002	
4	20096	99,0	1,7	4,3	279,4	67,4	112,6	110,3	76,0	4,3	18,1	680,9	2003	
5	23441	100,0	2,0	4,9	270,0	67,4	116,9	110,2	72,6	3,8	18,0	680,9	2004	
6	24350	101,0	2,4	5,0	283,0	67,0	121,3	114,3	64,6	4,3	15,9	674,4	2005	
7	26754	102,0	2,7	4,9	253,0	67,0	123,1	121,6	50,3	3,6	15,8	674,4	2006	
1	13468	113,0	0,9	-0,4	183,4	119,0	74,9	74,7	57,3	1,4	4,0	496,7	2000	kaliski
2	15127	113,0	0,5	-0,4	226,2	118,6	73,3	72,8	66,0	1,8	4,4	496,7	2001	
3	15451	112,0	0,6	-0,7	237,9	73,4	71,1	72,1	66,9	2,0	4,2	484,6	2002	
4	16195	112,0	0,6	-0,5	255,0	72,8	71,6	72,3	65,4	2,9	3,9	482,2	2003	
5	19489	112,0	0,2	-0,4	270,0	73,0	76,2	73,7	58,7	2,7	3,8	482,2	2004	
6	20639	112,0	1,2	-0,6	272,0	73,0	78,0	78,2	54,0	2,6	4,3	479,7	2005	
7	22109	112,0	1,2	-1,3	263,0	73,4	80,0	76,0	43,9	2,1	3,6	479,7	2006	
1	13819	99,0	1,8	-0,9	237,6	78,1	38,6	38,3	41,4	1,3	7,7	320,1	2000	koniński
2	15917	99,0	1,4	-1,1	286,9	77,9	37,4	37,2	46,8	1,4	7,2	319,6	2001	
3	16206	98,0	1,0	-1,4	294,3	42,6	37,0	36,9	45,3	2,1	5,3	291,2	2002	
4	17228	98,0	0,6	-1,2	321,0	42,4	34,2	37,5	45,2	3,4	6,0	287,4	2003	
5	19548	98,0	0,9	-1,0	304,0	42,4	36,1	37,4	43,3	2,9	4,8	287,4	2004	
6	20744	98,1	1,0	-1,2	301,0	42,5	35,8	38,2	40,9	2,2	5,5	283,6	2005	

7	21214	98,1	1,1	-1,8	263,0	42,6	36,4	39,4	33,6	2,2	5,8	283,6	2006	
1	37517	2200,0	-2,3	-0,6	689,8	20,3	69,0	154,4	11,0	3,9	6,6	8,2	2000	miasto Poznań
2	39847	2189,0	-1,9	-0,9	755,0	2,2	67,9	154,0	17,1	5,8	6,9	8,1	2001	
3	41013	2211,0	-2,1	-1,2	711,4	2,1	62,7	156,8	21,7	5,0	7,2	42,9	2002	
4	42508	2197,0	-1,4	-2,7	725,1	1,8	63,8	150,4	21,9	5,3	7,1	38,4	2003	
5	49125	2184,0	-1,2	-3,6	657,0	2,1	66,1	159,7	21,5	4,4	7,3	38,4	2004	
6	53279	2173,0	0,0	-3,7	611,0	1,9	62,3	160,3	19,6	5,9	7,5	33,9	2005	
7	56081	2158,0	0,0	-4,4	522,0	1,8	61,6	163,6	15,8	5,0	7,7	33,9	2006	
1	19338	90,0	0,7	0,1	446,7	50,0	84,2	146,6	80,9	2,8	71,9	644,9	2000	szczeciński
2	21226	90,0	0,6	-0,1	483,7	48,1	79,8	142,6	97,2	3,5	64,2	642,4	2001	
3	21947	88,0	0,0	-0,3	450,5	26,3	71,4	134,6	110,7	3,0	64,3	616,0	2002	
4	21753	88,0	0,1	-0,4	414,4	24,6	67,2	133,2	112,9	3,8	64,0	615,1	2003	
5	23662	88,0	0,1	-0,3	409,0	24,6	68,2	134,1	107,0	3,3	56,9	615,1	2004	
6	25370	88,1	0,4	-0,4	399,0	24,7	69,5	135,7	100,2	3,3	51,6	614,1	2005	
7	26502	88,1	0,3	-0,3	369,0	24,9	71,5	139,6	81,1	2,7	56,2	614,1	2006	
1	14130	59,0	1,8	-1,2	364,1	34,9	40,2	63,9	69,2	1,7	58,5	469,2	2000	koszaliński
2	15710	59,0	1,9	-1,5	412,0	34,8	36,5	59,6	78,1	2,6	52,0	471,3	2001	
3	16946	57,0	1,4	-1,7	384,7	17,9	33,0	59,2	78,9	2,2	52,6	397,6	2002	
4	17762	57,0	1,1	-1,9	384,0	17,0	33,2	58,9	77,9	2,5	57,4	388,5	2003	
5	20398	57,0	0,9	-1,7	372,0	17,4	33,9	59,3	75,7	2,8	51,4	388,5	2004	
6	21239	57,0	0,6	-2,1	352,0	17,4	34,7	61,2	68,6	3,0	53,7	379,3	2005	
7	23136	57,0	0,8	-2,4	339,0	17,5	36,2	62,2	57,7	3,6	52,5	379,3	2006	
1	14231	134,0	-1,4	-1,6	365,5	79,4	112,1	139,3	133,0	1,3	41,2	571,4	2000	jeleniogorski-walbrzyski
2	16125	133,0	-1,2	-1,8	362,8	78,4	103,4	134,4	151,5	1,3	43,4	571,4	2001	
3	16545	128,0	-1,9	-2,2	385,4	34,6	100,1	129,8	153,3	1,3	40,2	482,8	2002	
4	17576	128,0	-2,1	-2,1	403,8	33,9	98,0	127,9	151,6	2,3	36,6	470,0	2003	
5	19526	127,0	-2,2	-2,1	400,0	33,8	98,4	126,4	141,0	1,4	35,1	470,0	2004	
6	20331	127,0	-2,0	-2,4	388,0	33,8	98,4	127,8	130,4	1,5	33,2	457,2	2005	
7	21728	126,0	-2,2	-3,2	390,0	33,9	99,3	130,1	108,1	1,5	32,4	457,2	2006	
1	22828	124,0	1,7	-1,4	440,3	32,1	53,4	64,5	44,4	1,2	2,8	247,3	2000	legnicki
2	22656	124,0	1,2	-2,0	437,1	31,8	52,8	62,0	48,5	2,2	2,3	247,3	2001	
3	25099	121,0	1,3	-2,7	476,2	14,9	50,4	60,7	50,6	1,7	2,3	203,8	2002	
4	25934	120,0	0,4	-2,7	484,2	14,8	51,2	59,4	50,2	2,9	2,4	194,7	2003	
5	30681	120,0	0,8	-2,6	499,0	14,8	52,9	58,7	44,3	2,0	2,4	194,7	2004	
6	37026	120,0	1,1	-3,1	463,0	14,8	52,5	60,3	39,3	2,0	2,4	185,6	2005	
7	44366	120,0	1,4	-3,4	431,0	14,7	53,9	62,4	30,5	2,2	2,5	185,6	2006	
1	13280	85,0	1,2	1,9	187,4	47,6	30,2	32,4	32,5	1,7	3,8	332,2	2000	wrocławski
2	14342	85,0	1,3	2,1	204,3	47,3	30,5	34,5	38,9	2,3	3,4	330,2	2001	
3	16281	84,0	1,0	3,6	231,6	23,4	31,3	32,4	40,0	2,3	3,2	274,0	2002	
4	17281	85,0	0,1	4,0	243,2	23,2	31,9	33,5	39,6	5,2	3,2	268,9	2003	
5	19032	85,0	0,9	4,4	263,0	23,2	33,2	33,6	36,5	3,7	2,9	268,9	2004	
6	20098	85,4	0,8	3,4	256,0	23,2	35,8	35,0	32,2	4,5	2,8	263,7	2005	
7	21634	86,1	1,3	5,8	269,0	23,2	42,6	35,1	23,6	4,1	2,8	263,7	2006	
1	27064	2165,0	-2,6	0,3	599,2	3,3	60,8	140,7	21,7	4,3	6,5	12,7	2000	miasto Wrocław
2	29072	2165,0	-2,0	2,6	556,8	3,3	54,3	139,2	28,5	8,8	6,4	12,5	2001	
3	31143	2181,0	-2,4	1,0	544,8	1,6	48,4	133,8	35,4	5,5	7,3	78,8	2002	
4	31274	2177,0	-2,4	0,3	618,3	1,5	45,5	135,7	37,0	6,1	7,4	74,5	2003	
5	34351	2173,0	-2,2	1,2	638,0	1,4	45,5	137,7	35,4	3,2	7,3	74,5	2004	
6	36000	2172,0	-1,8	2,2	598,0	1,4	48,6	142,3	31,4	3,1	7,2	70,1	2005	
7	40457	2167,0	-1,5	-0,5	553,0	1,4	49,3	150,3	23,3	6,9	7,3	70,1	2006	
1	15149	115,0	0,0	-3,4	281,1	105,5	88,6	113,5	69,3	1,0	6,4	579,8	2000	opolski
2	15780	115,0	-0,2	-3,5	317,7	104,5	79,3	108,9	78,6	1,2	7,4	579,8	2001	

3	16738	113,0	-0,8	-4,3	367,2	50,2	73,8	106,0	83,4	1,2	6,0	560,1	2002	
4	17173	112,0	-1,2	-4,0	359,1	49,9	72,3	106,7	81,6	2,6	7,0	557,0	2003	
5	20785	112,0	-0,9	-3,2	354,0	49,8	72,0	106,6	74,6	1,2	7,9	557,0	2004	
6	21347	111,0	-0,9	-3,0	330,0	49,9	73,3	108,3	69,4	1,1	7,5	553,9	2005	
7	22347	111,0	-0,9	-4,6	312,0	49,7	74,5	111,2	60,1	1,3	7,5	553,9	2006	
1	17289	116,0	0,5	-0,3	352,5	76,5	98,5	126,2	81,8	1,7	16,5	533,0	2000	bydgoski
2	18962	116,0	0,9	-0,5	390,3	75,2	91,6	126,4	94,2	2,4	15,3	532,1	2001	
3	20212	114,0	0,5	-0,4	399,0	45,7	85,2	122,5	97,9	2,0	13,3	503,2	2002	
4	20814	114,0	0,3	-0,4	414,1	44,8	80,1	123,3	97,8	3,5	15,0	485,4	2003	
5	23034	114,0	0,3	-0,5	399,0	44,7	82,4	122,4	91,9	1,7	14,8	485,4	2004	
6	23962	114,0	0,6	-0,9	369,0	44,5	82,9	125,4	86,0	1,8	14,1	467,6	2005	
7	25949	114,0	0,6	-1,4	328,0	44,3	87,1	129,1	72,7	2,0	13,1	467,6	2006	
1	14537	118,0	1,1	-0,7	316,2	123,3	85,0	113,6	99,4	1,9	13,5	629,5	2000	toruńsko-włocławski
2	16205	118,0	0,9	-0,9	363,9	122,9	79,1	110,7	111,0	2,6	12,2	629,5	2001	
3	17118	116,0	0,5	-1,1	362,8	73,1	73,6	107,5	113,7	2,1	11,4	602,4	2002	
4	17587	116,0	0,6	-0,9	388,7	72,6	72,7	107,7	114,2	3,1	10,5	590,4	2003	
5	20285	116,0	0,6	-0,8	385,0	72,6	74,0	109,4	107,7	2,7	10,9	590,4	2004	
6	21043	117,0	0,4	-1,1	343,0	72,7	74,6	112,8	102,0	2,5	12,2	578,4	2005	
7	22719	116,0	0,8	-2,0	326,0	72,7	78,0	115,5	87,3	2,8	12,4	578,4	2006	
1	13327	60,0	3,4	-1,3	330,1	32,6	38,9	50,3	55,4	2,2	37,2	347,9	2000	ślupski
2	15340	60,0	3,3	-2,2	456,0	32,2	36,3	49,8	63,4	3,3	34,5	346,2	2001	
3	15800	58,0	2,9	-2,2	442,6	17,1	34,2	49,1	65,4	1,6	33,6	285,0	2002	
4	16556	58,0	2,7	-1,7	456,6	16,7	34,3	49,1	63,2	3,2	34,4	284,0	2003	
5	18082	58,0	2,5	-2,1	421,0	17,0	34,1	50,2	59,0	2,8	34,9	284,0	2004	
6	19713	58,4	2,3	-2,1	429,0	17,2	36,8	50,9	54,4	2,7	33,2	282,9	2005	
7	21705	58,4	2,8	-3,4	424,0	17,1	38,9	52,5	45,4	2,5	30,7	282,9	2006	
1	12985	98,0	4,7	1,3	288,1	74,2	64,7	74,2	69,7	2,8	43,8	548,6	2000	gdański
2	14404	99,0	4,6	2,4	380,7	73,8	63,1	76,5	81,1	3,3	43,3	550,6	2001	
3	15408	98,0	4,1	2,1	395,2	41,6	60,5	76,4	87,5	3,4	37,2	493,8	2002	
4	16140	99,0	3,8	2,3	410,1	40,6	60,8	76,3	90,1	7,8	36,3	478,3	2003	
5	17626	99,0	4,1	3,6	375,0	40,8	61,8	77,3	86,7	4,1	36,1	478,3	2004	
6	18408	100,0	4,3	2,8	380,0	40,6	65,0	78,4	77,2	4,5	36,5	462,8	2005	
7	20049	101,0	4,5	2,7	359,0	40,7	71,2	81,3	61,7	4,1	37,2	462,8	2006	
1	26834	1817,0	-1,1	-0,3	776,2	5,1	74,4	151,6	22,2	2,5	16,5	90,8	2000	Gdańsk-Gdynia-Sopot
2	28105	1814,0	-0,6	-0,5	767,2	4,4	72,8	151,8	31,3	3,2	16,2	90,8	2001	
3	29392	1823,0	-1,2	-0,3	728,6	2,8	64,9	146,4	38,8	5,9	17,3	90,8	2002	
4	29719	1821,0	-0,8	-0,5	683,4	2,4	60,0	148,3	38,2	6,4	16,3	62,0	2003	
5	34733	1815,0	-0,9	-0,7	722,0	2,3	58,2	150,6	33,9	4,3	16,2	62,0	2004	
6	37731	1810,0	-0,5	-0,9	563,0	2,3	58,6	153,4	28,3	5,5	15,6	33,1	2005	
7	40491	1790,0	-0,3	-2,6	466,0	2,2	59,0	156,1	18,8	6,4	16,0	33,1	2006	
1	15090	73,0	2,8	-1,4	296,2	46,1	38,7	52,0	61,4	1,9	10,1	444,0	2000	elbląski
2	13289	73,0	2,8	-2,4	353,2	46,1	38,5	51,9	69,2	1,5	14,4	444,7	2001	
3	14489	71,0	2,1	-2,8	343,9	26,6	36,4	49,8	69,1	1,6	7,9	358,8	2002	
4	15453	71,0	2,1	-2,4	359,8	25,9	37,4	50,2	65,8	3,0	6,4	353,4	2003	
5	17649	71,0	1,7	-2,5	357,0	26,1	38,7	50,6	61,7	2,5	5,7	353,4	2004	
6	18427	70,8	1,6	-3,3	332,0	25,9	40,9	52,0	57,0	2,2	7,2	348,0	2005	
7	19580	70,8	1,8	-3,9	316,0	25,8	42,7	51,7	47,6	2,0	8,2	348,0	2006	
1	15090	61,0	2,3	-1,4	378,0	44,7	45,2	78,8	59,2	3,1	25,2	549,7	2000	olsztyński
2	15858	61,0	1,8	-0,9	450,1	44,4	43,3	78,1	68,4	3,4	27,8	547,6	2001	
3	17459	59,0	1,7	-0,8	455,7	24,7	40,5	77,2	68,7	2,6	18,8	515,3	2002	
4	18993	59,0	1,3	-0,6	405,5	24,1	39,8	77,3	66,6	4,5	19,9	470,7	2003	
5	21231	59,0	1,4	-0,9	376,0	24,2	41,5	78,2	63,3	3,4	20,4	470,7	2004	

6	22346	59,3	1,4	-1,2	359,0	24,3	42,5	79,4	58,7	3,6	21,0	426,1	2005	
7	23577	59,3	1,7	-3,1	353,0	24,2	44,2	81,0	49,5	3,9	20,7	426,1	2006	
1	10827	46,0	2,8	-2,7	325,1	28,3	15,2	24,2	37,5	2,5	13,6	311,6	2000	elcki
2	11666	46,0	2,4	-3,0	437,8	28,1	15,2	23,7	40,8	1,7	6,7	311,0	2001	
3	21072	45,0	2,2	-3,6	436,8	15,9	13,7	23,2	38,9	1,7	7,3	254,3	2002	
4	13122	45,0	2,0	-1,7	422,7	15,8	14,3	23,7	38,0	2,9	9,6	239,5	2003	
5	15610	45,0	1,7	-2,2	404,0	15,8	13,5	24,7	37,3	2,6	8,7	239,5	2004	
6	16429	44,8	2,2	-1,9	372,0	15,9	15,4	25,8	35,3	3,7	9,2	224,6	2005	
7	18128	44,7	1,8	-3,8	370,0	15,8	15,8	26,4	30,5	3,5	9,4	224,6	2006	

Źródło: Roczniki statystyczne Rzeczypospolitej Polskiej, GUS Warszawa, lata 2001-2007.

5. Model dochodu budżetu powiatów województwa dolnośląskiego



5.1. Wstęp

Wejściem do budowy modelu dochodu budżetu powiatu lub miasta na prawach powiatu, na jednego mieszkańca w zł na przykładzie województwa dolnośląskiego, był adekwatny model liniowy wielowymiarowy z efektami indywidualnymi opracowany przez autora dla województw Polski⁴². Zmienną objaśnianą w tym modelu był:

Y - produkt krajowy brutto (ceny bieżące) na 1 mieszkańca w roku poprzednim w zł).
a parametry tego modelu są następujące:

Tab. 5.1. Parametry modelu liniowego wielowymiarowego dla województw Polski

CX1	CX2	CX3	CX4	CX5	CX6	CX7	CX8
92272,71	85367,19	102278,32	155780,77	78881,90	73088,73	62862,98	72209,13
CX9	CX10	CX11	CX12	CX13	CX14	CX15	CX16
58460,39	92101,27	75187,89	90705,06	73565,03	83168,33	81386,92	66332,57
X1	X3	X4	X5A	X6	X7	X8	X9
-372,93	401,18	-12,61	-4,57	-22,81	81,46	-22,17	174,81
X10	X11						
-65,29	-23,47						

gdzie:

CX1÷CX16 – efekty indywidualne odniesione do 16. województw Polski: łódzkie, mazowieckie, małopolskie, śląskie, lubelskie, podkarpackie, podlaskie, świętokrzyskie, lubuskie, wielkopolskie, zachodniopomorskie, dolnośląskie, opolskie, kujawsko-pomorskie, pomorskie, marmińsko-mazurskie.

X₁ - ludność na 1 km² powierzchni ogółem,

X₃ - saldo migracji wewnętrznych i zagranicznych na pobyt stały (na 1000 ludności),

X₄ - przestępstwa stwierdzone na 10 tys. ludności,

X_{5A} - pracujący w tys. (rolnictwo, łowiectwo i leśnictwo, rybactwo),

X₆ - pracujący w tys. (przemysł i budownictwo),

X₇ - pracujący w tys. (usługi rynkowe i nierynkowe),

X₈ - bezrobotni zarejestrowani ogółem w tys.,

X₉ - mieszkania oddane do użytkowania na 1000 ludności,

X₁₀ - miejsca noclegowe w turystycznych obiektach zbiorowego zakwaterowania w tys.,

X₁₁ – użytki rolne w tys. ha w roku poprzednim.

Zmienna objaśniana X_2 została wyeliminowana na etapie formułowania w *Excelu* modelu PKB dla województw.

⁴² Wornalkiewicz W., *Zastosowanie programu Excel do sformułowania modelu produktu brutto województw*, WSZiA Opole, 2009.

5.2. Baza danych statystycznych

Jak już wspomniano w prezentowanym temacie ograniczono się do powiatów województwa dolnośląskiego o strukturze terytorialnej obowiązującej od 1.01.2008 roku i dostępnych danych. Przykładowe wybrane dane powiatów za rok 2000 oraz 2007, wprowadzono bezpośrednio do arkusza kalkulacyjnego *Excel*, celem ich późniejszego importu do programu *Gretl* zamieszczono w tabeli 5.2⁴³.

Tab. 5.2. Dane empiryczne powiatów województwa dolnośląskiego

n	Y1	X1	X3	X5	X6	X7	X8	X9	X10	T	Rok	P.	Nr	Powiat
1	431,70	69,1	0,7	0,50	7,40	7,60	8,90	4,5	6,91	1	2000	J-W	101	bolesławiecki
2	493,72	69,2	1,0	0,46	6,88	7,54	9,22	2,1	5,38	2	2001			
3	461,28	67,6	-1,7	0,43	6,54	7,47	8,57	2,8	8,70	3	2002			
4	394,96	68,0	0,2	0,35	6,23	7,82	8,80	4,9	8,24	4	2003			
5	440,75	68,0	-0,7	0,33	6,92	7,25	8,00	1,9	6,84	5	2004			
6	467,79	68,0	-1,5	0,32	7,40	7,68	7,10	2,6	6,95	6	2005			
7	541,00	68,0	-0,5	0,30	7,90	7,50	5,30	1,9	6,73	7	2006			
8	550,31	68,0	-1,0	0,32	8,64	7,84	4,00	2,0	6,18	8	2007			
9	464,90	233,3	-2,1	0,34	9,61	7,95	10,60	0,9	2,49	1	2000	J-W	102	dzierżoniowski
10	525,96	232,6	-2,2	0,30	8,53	7,94	12,02	1,5	2,67	2	2001			
11	509,57	222,4	-2,9	0,26	9,02	7,93	12,23	1,7	2,89	3	2002			
12	447,81	221,0	-1,7	0,21	8,36	7,53	12,40	2,7	2,72	4	2003			
13	500,60	221,0	-1,9	0,25	7,98	7,47	12,10	1,1	2,97	5	2004			
14	525,40	220,0	-2,0	0,24	7,40	7,36	12,30	1,6	2,93	6	2005			
15	543,70	219,0	-2,2	0,22	7,20	7,48	10,50	1,3	3,03	7	2006			
16	629,25	218,0	-2,0	0,23	7,37	7,60	8,10	1,3	5,89	8	2007			
17	628,40	94,1	-1,8	0,39	3,88	4,23	6,30	0,8	7,24	1	2000	J-W	103	jaworski
18	659,19	93,7	-3,2	0,39	3,59	4,04	9,94	1,2	19,61	2	2001			
19	643,48	90,6	-2,2	0,39	3,75	3,86	7,17	1,1	20,63	3	2002			
20	554,62	90,0	-1,8	0,38	3,46	3,76	7,00	1,5	14,40	4	2003			
21	616,09	90,0	-2,2	0,39	3,28	3,63	6,20	0,9	13,69	5	2004			
22	624,46	90,0	-2,5	0,38	3,28	3,94	5,80	0,7	14,39	6	2005			
23	665,79	89,0	-4,6	0,37	3,55	3,98	4,70	0,9	14,54	7	2006			
24	736,82	89,0	-2,7	0,36	3,45	4,09	3,50	2,1	13,61	8	2007			
25	498,60	106,7	2,7	0,27	5,52	5,61	5,50	0,8	256,60	1	2000	J-W	104	jeleniogórski
26	560,41	106,7	0,5	0,26	4,58	5,53	6,30	1,8	268,73	2	2001			
27	543,98	102,0	1,5	0,25	4,71	5,46	6,50	1,6	257,12	3	2002			
28	573,10	102,0	3,1	0,20	4,20	5,10	6,40	2,3	253,21	4	2003			

⁴³ Kufel T., *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETl*. Wydawnictwo Naukowe PWN Warszawa 2007.

29	663,79	102,0	2,1	0,21	4,33	4,96	5,60	2,8	240,60	5	2004			
30	649,14	102,0	1,7	0,22	3,97	5,21	5,10	2,3	237,50	6	2005			
31	666,23	102,0	0,2	0,22	4,48	5,00	4,40	1,6	242,80	7	2006			
32	780,68	102,0	2,6	0,22	4,59	5,29	2,90	3,3	233,20	8	2007			
33	492,70	123,3	-3,7	0,14	5,37	4,09	5,30	1,8	17,30	1	2000	J-W	105	kamiennogórski
34	571,84	123,1	-2,3	0,13	4,30	4,15	5,97	0,4	19,59	2	2001			
35	539,43	119,2	-4,0	0,12	4,01	4,21	5,56	0,5	19,47	3	2002			
36	508,20	119,0	-2,7	0,11	3,86	3,53	5,40	2,5	15,49	4	2003			
37	537,90	118,0	-2,9	0,11	4,16	3,63	4,80	2,6	15,88	5	2004			
38	577,65	118,0	-4,2	0,11	4,18	3,71	4,3	1,0	9,48	6	2005			
39	573,55	117,0	-5,3	0,10	4,09	3,61	3,50	0,9	9,65	7	2006			
40	595,01	116,0	-3,5	0,10	4,24	3,68	2,20	1,2	7,98	8	2007			
41	565,50	109,6	-1,7	0,68	11,38	19,04	19,10	1,4	57,70	1	2000	J-W	106	kłódzki
42	692,69	109,2	-1,7	0,63	9,67	18,55	21,99	1,5	70,85	2	2001			
43	628,29	103,6	-2,6	0,59	9,32	18,07	22,24	1,2	72,43	3	2002			
44	568,33	103,0	-3,4	0,54	8,74	17,52	22,80	1,9	53,63	4	2003			
45	617,45	102,0	-2,2	0,54	8,13	17,23	21,30	1,8	39,86	5	2004			
46	649,06	102,0	-3,0	0,51	7,71	17,38	20,50	1,8	47,25	6	2005			
47	649,50	101,0	-3,4	0,53	7,74	18,13	16,80	1,5	39,39	7	2006			
48	714,15	100,0	-2,6	0,53	8,05	17,72	13,60	1,7	47,99	8	2007			
49	554,70	139,7	-2,2	0,14	3,91	5,55	6,50	0,9	50,66	1	2000	J-W	107	łubański
50	606,71	139,1	-2,7	0,16	3,42	5,31	7,14	0,6	49,38	2	2001			
51	564,24	135,0	-3,5	0,18	3,41	5,08	6,60	1,0	34,08	3	2002			
52	497,13	135,0	-2,5	0,14	3,47	5,19	6,40	2,6	39,86	4	2003			
53	550,04	134,0	-3,0	0,13	3,40	5,07	6,10	0,8	41,90	5	2004			
54	581,85	134,0	-2,2	0,14	3,34	5,33	5,60	1,2	42,06	6	2005			
55	602,77	133,0	-2,8	0,13	3,35	5,42	4,80	0,9	41,88	7	2006			
56	673,18	132,0	-3,0	0,13	3,11	5,26	4,00	1,5	47,74	8	2007			
57	690,70	72,3	-0,8	0,12	2,46	3,92	5,90	0,7	8,75	1	2000	J-W	108	lwówecki
58	719,55	72,1	-1,5	0,12	2,32	3,80	6,20	1,7	12,94	2	2001			
59	693,65	68,6	-1,1	0,12	2,36	3,69	6,01	0,8	8,99	3	2002			
60	594,25	68,0	-3,0	0,10	2,36	3,54	5,70	3,5	7,98	4	2003			
61	670,27	68,0	1,0	0,12	2,46	3,62	5,20	1,3	7,98	5	2004			
62	691,72	68,0	-2,2	0,15	2,62	3,63	5,00	0,7	7,51	6	2005			
63	749,43	68,0	-3,8	0,13	2,77	3,80	4,50	1,0	8,96	7	2006			
64	855,98	67,0	-1,5	0,14	2,82	3,84	3,50	1,2	13,06	8	2007			
65	582,60	74,1	0,7	0,75	2,64	3,11	4,10	0,6	4,21	1	2000	J-W	109	strzeliński

66	614,54	73,9	-2,4	0,63	2,33	3,16	4,50	0,4	3,72	2	2001			
67	651,64	71,8	-0,1	0,51	1,92	3,22	5,08	1,3	3,83	3	2002			
68	493,76	72,0	-2,8	0,36	2,14	3,10	5,10	2,5	3,97	4	2003			
69	571,37	71,0	-1,9	0,36	1,97	3,27	5,00	0,5	4,11	5	2004			
70	630,98	71,0	-2,8	0,39	2,23	3,28	4,70	0,9	4,25	6	2005			
71	665,19	71,0	-2,8	0,54	2,37	3,59	4,00	1,3	4,29	7	2006			
72	802,65	71,0	-1,6	0,48	2,36	3,86	2,90	1,0	4,29	8	2007			
73	518,80	226,1	-1,4	1,01	17,06	14,43	16,10	1,8	3,00	1	2000	J-W	110	świdnicki
74	631,55	225,7	-1,6	0,90	15,55	14,34	18,47	1,0	5,39	2	2001			
75	581,49	217,8	-1,2	0,89	14,66	14,25	17,76	1,6	4,20	3	2002			
76	550,75	217,0	-1,4	0,80	13,71	13,99	17,90	2,0	4,23	4	2003			
77	541,26	217,0	-2,4	0,71	13,90	14,39	16,20	0,8	4,42	5	2004			
78	615,14	216,0	-1,5	0,71	14,26	14,43	14,30	1,6	4,40	6	2005			
79	638,45	216,0	-2,5	0,69	15,11	14,30	11,80	2,0	4,22	7	2006			
80	704,92	216,0	-2,2	0,70	16,60	14,50	6,90	1,9	4,49	8	2007			
81	278,60	144,5	-0,2	0,63	14,24	22,03	17,70	1,0	7,75	1	2000	J-W	111	wałbrzyski
82	433,36	144,1	-1,4	0,60	14,46	21,57	19,17	0,4	7,69	2	2001			
83	334,86	139,1	-2,1	0,57	14,68	21,11	20,64	0,9	7,63	3	2002			
84	431,84	366,0	-3,5	0,54	14,91	20,65	22,10	0,4	7,57	4	2003			
85	503,92	363,0	-3,6	0,47	15,59	20,44	21,00	0,9	7,63	5	2004			
86	529,71	361,0	-3,5	0,52	15,73	20,75	19,10	0,6	5,06	6	2005			
87	547,42	357,0	-4,0	0,48	15,69	20,93	16,30	1,0	5,59	7	2006			
88	609,77	354,0	-4,5	0,47	16,94	21,89	10,70	0,9	6,20	8	2007			
89	667,00	90,8	-2,6	0,57	4,40	6,83	7,60	0,7	5,66	1	2000	J-W	112	ząbkowski
90	726,88	90,6	0,1	0,53	4,13	6,48	8,42	0,6	5,66	2	2001			
91	687,87	87,8	-3,1	0,50	3,66	6,14	8,90	1,4	5,78	3	2002			
92	637,01	87,0	-1,8	0,47	3,80	6,03	8,60	2,8	5,25	4	2003			
93	672,67	87,0	-1,7	0,46	3,85	5,59	8,20	1,1	4,88	5	2004			
94	739,90	87,0	-3,1	0,44	4,01	5,25	8,00	1,3	4,03	6	2005			
95	777,35	86,0	-3,3	0,42	3,89	5,59	6,80	1,4	2,84	7	2006			
96	813,45	86,0	-1,2	0,33	4,04	5,73	4,30	1,6	5,55	8	2007			
97	486,50	116,6	-2,3	0,39	13,75	11,76	8,10	1,2	10,79	1	2000	J-W	113	zgorzelecki
98	546,42	119,4	-2,4	0,37	14,25	11,56	9,18	2,1	8,69	2	2001			
99	520,26	114,5	-1,8	0,36	14,28	11,36	9,58	1,6	8,79	3	2002			
100	441,27	114,0	-1,8	0,28	13,07	10,65	9,80	2,9	8,32	4	2003			
101	519,99	114,0	-2,2	0,25	12,18	10,47	9,60	1,0	8,12	5	2004			
102	538,92	113,0	-2,3	0,25	11,86	10,49	8,30	1,3	8,86	6	2005			

103	624,55	112,0	-6,0	0,23	11,60	10,97	6,20	0,8	9,51	7	2006			
104	642,95	112,0	-3,6	0,25	10,71	11,54	4,10	0,7	9,27	8	2007			
105	519,00	82,5	-3,0	0,36	3,29	3,85	5,60	1,0	4,09	1	2000	J-W	114	złotoryjski
106	571,95	82,2	-2,7	0,36	2,81	3,87	6,37	0,6	2,33	2	2001			
107	605,41	80,3	-2,7	0,35	2,68	3,89	6,72	2,0	4,09	3	2002			
108	537,87	80,0	-2,4	0,28	2,26	3,66	6,40	1,8	3,84	4	2003			
109	569,63	80,0	-1,4	0,31	2,30	3,59	6,10	1,2	3,59	5	2004			
110	637,74	80,0	-2,4	0,30	2,11	3,69	5,7	1,1	3,33	6	2005			
111	603,22	79,0	-4,6	0,32	1,94	3,74	4,70	1,1	3,35	7	2006			
112	934,39	79,0	-2,4	0,30	2,09	3,81	3,90	1,5	3,35	8	2007			
113	2173,60	857,9	-2,4	0,12	7,17	17,91	5,60	0,9	32,16	1	2000	J-W	115	Jelenia Góra
114	2233,22	852,7	-0,9	0,10	7,91	17,28	6,64	2,1	26,13	2	2001			
115	2185,97	820,1	-2,9	0,09	8,06	16,65	6,85	0,8	24,32	3	2002			
116	2034,57	808,0	-3,6	0,07	7,44	15,09	6,70	2,2	29,28	4	2003			
117	2326,24	804,0	-4,6	0,09	7,99	15,02	5,70	2,9	29,04	5	2004			
118	2581,61	803,0	-3,1	0,07	8,35	14,98	4,60	2,4	25,08	6	2005			
119	2889,53	792,0	-2,9	0,07	7,75	15,38	4,00	3,1	27,28	7	2006			
120	3561,37	785,0	-4,1	0,07	7,58	15,75	2,70	4,3	25,52	8	2007			
121	602,00	207,2	-1,7	0,25	8,23	11,12	7,80	1,9	3,70	1	2000	L	201	głogowski
122	675,45	207,4	-3,3	0,24	8,00	10,55	8,66	1,5	4,30	2	2001			
123	637,95	198,3	-5,0	0,23	7,61	9,99	9,20	2,3	4,15	3	2002			
124	572,80	198,0	-4,2	0,26	7,17	9,57	9,00	3,3	3,25	4	2003			
125	623,80	198,0	-4,0	0,31	7,23	9,76	8,00	2,6	2,49	5	2004			
126	704,98	198,0	-3,7	0,30	7,08	10,02	7,30	2,3	1,73	6	2005			
127	750,29	198,0	-4,5	0,32	7,26	10,12	5,80	3,6	1,73	7	2006			
128	810,01	198,0	-5,2	0,32	7,52	11,06	4,10	2,5	1,73	8	2007			
129	576,30	51,5	-1,6	0,37	1,91	2,62	3,80	1,0	0,50	1	2000	L	202	górowski
130	614,90	51,4	-4,6	0,35	1,87	2,57	4,20	1,4	0,50	2	2001			
131	555,89	49,9	-3,5	0,33	1,74	2,52	4,26	0,1	0,52	3	2002			
132	550,99	50,0	-2,4	0,31	1,54	2,45	4,60	0,8	0,52	4	2003			
133	584,46	50,0	-2,4	0,24	1,55	2,61	4,30	0,8	0,52	5	2004			
134	670,16	50,0	-5,0	0,23	1,44	2,53	4,00	0,6	0,52	6	2005			
135	715,64	49,0	-4,6	0,22	1,34	2,64	3,40	0,4	0,49	7	2006			
136	738,94	49,0	-3,2	0,22	1,25	2,73	2,40	0,5	0,52	8	2007			
137	572,10	72,8	-1,8	0,45	3,04	3,71	6,00	0,5	16,75	1	2000	L	203	legnicki
138	614,91	72,8	0,5	0,41	2,89	3,60	6,43	0,5	9,63	2	2001			
139	569,66	71,2	0,9	0,38	2,52	3,49	6,49	1,1	9,28	3	2002			

140	568,70	71,0	1,0	0,32	2,53	3,35	6,20	4,2	8,43	4	2003			
141	699,88	71,0	2,4	0,35	2,54	3,21	5,70	1,8	7,98	5	2004			
142	634,60	71,0	0,5	0,36	2,71	3,03	5,20	2,1	8,83	6	2005			
143	673,32	71,0	0,5	0,32	2,57	3,21	3,90	2,1	8,25	7	2006			
144	781,75	72,0	3,1	0,29	2,63	3,58	2,70	1,9	8,65	8	2007			
145	431,40	154,6	-4,1	0,30	10,52	14,58	7,20	1,2	2,87	1	2000	L	204	łubiński
146	485,68	154,5	-4,0	0,31	10,25	14,39	7,92	2,0	3,36	2	2001			
147	449,08	149,5	-3,0	0,33	9,41	14,21	9,14	1,8	3,04	3	2002			
148	357,15	149,0	-4,8	0,27	8,71	13,42	9,10	3,3	3,49	4	2003			
149	486,62	149,0	-4,4	0,26	9,67	13,37	7,10	1,6	3,94	5	2004			
150	559,64	149,0	-4,8	0,24	9,49	14,17	5,60	2,2	3,69	6	2005			
151	559,48	148,0	-4,8	0,21	9,68	14,51	4,10	1,6	3,74	7	2006			
152	719,07	148,0	-4,6	0,21	10,12	15,27	2,80	2,7	3,53	8	2007			
153	367,60	80,2	-0,4	0,31	16,41	6,58	6,00	1,3	3,01	1	2000	L	205	połkowski
154	426,73	80,3	-0,8	0,26	16,51	6,18	6,59	2,3	3,26	2	2001			
155	378,37	78,6	-2,7	0,22	16,48	5,78	6,54	1,1	3,48	3	2002			
156	302,71	78,0	-2,3	0,30	17,20	6,10	6,40	2,1	6,38	4	2003			
157	370,16	78,0	-2,7	0,24	17,18	6,98	5,60	1,3	6,22	5	2004			
158	437,22	78,0	-3,7	0,22	17,38	7,20	5,10	2,1	6,22	6	2005			
159	607,28	78,0	-2,4	0,23	18,18	7,29	4,10	1,2	5,92	7	2006			
160	649,22	79,0	-1,1	0,22	19,17	8,51	3,10	2,7	6,13	8	2007			
161	579,10	73,7	1,2	0,26	3,87	5,07	4,20	1,7	10,05	1	2000	L	206	wołowski
162	611,87	73,7	0,1	0,23	3,72	5,39	4,61	2,5	10,01	2	2001			
163	568,92	70,8	-2,5	0,20	3,59	5,71	4,45	0,9	7,83	3	2002			
164	510,85	71,0	-1,2	0,22	4,24	4,94	4,50	2,7	7,00	4	2003			
165	587,76	71,0	-1,2	0,20	4,59	4,71	4,10	1,6	8,15	5	2004			
166	617,20	70,0	-1,1	0,20	4,49	4,61	6,40	2,9	8,97	6	2005			
167	757,59	70,0	-1,5	0,20	4,64	4,76	3,10	1,5	7,99	7	2006			
168	676,56	70,0	0,3	0,19	4,17	4,84	2,60	2,5	7,99	8	2007			
169	1974,60	1938,6	0,1	0,15	9,39	20,16	9,50	0,7	4,85	1	2000	L	207	Legnica
170	2170,80	1934,8	-1,0	0,15	9,48	19,32	10,08	3,9	2,98	2	2001			
171	2230,50	1899,0	-2,0	0,14	9,05	18,49	10,55	2,4	4,58	3	2002			
172	2069,24	1902,0	-2,3	0,13	8,98	17,69	10,40	2,6	5,75	4	2003			
173	2279,94	1895,0	-3,1	0,11	10,17	17,52	9,50	3,0	7,93	5	2004			
174	2547,04	1879,0	-2,9	0,14	9,93	18,23	8,40	1,6	5,08	6	2005			
175	2877,12	1869,0	-3,9	0,15	10,18	19,27	6,10	3,1	7,36	7	2006			
176	3061,93	1861,0	-3,8	0,15	10,64	19,61	3,50	2,6	7,71	8	2007			

177	1117,10	52,4	-1,2	0,48	2,79	3,03	3,00	0,6	30,25	1	2000	W	301	milicki
178	1169,64	52,5	-0,8	0,46	2,70	3,08	3,41	1,9	24,71	2	2001			
179	897,52	51,7	-1,5	0,44	2,87	3,13	3,40	0,9	23,69	3	2002			
180	915,98	52,0	-2,4	0,40	2,96	2,84	3,70	3,3	20,57	4	2003			
181	818,92	52,0	-2,2	0,36	3,00	2,94	3,40	1,4	24,23	5	2004			
182	839,56	52,0	-2,9	0,37	3,04	2,79	3,10	1,2	26,29	6	2005			
183	932,59	52,0	-5,8	0,35	3,14	2,81	2,40	2,4	26,19	7	2006			
184	1074,02	52,0	-3,0	0,36	2,92	2,92	2,00	1,4	20,54	8	2007			
185	515,70	100,1	-0,7	0,49	8,03	8,38	9,00	0,7	6,08	1	2000	W	302	oleśnicki
186	582,97	100,3	-0,9	0,46	8,24	8,2	10,64	1,7	4,25	2	2001			
187	544,35	98,3	-0,6	0,43	7,95	8,02	10,71	2,2	4,27	3	2002			
188	480,54	98,0	-0,2	0,40	8,43	7,97	10,10	3,6	4,18	4	2003			
189	549,10	98,0	-1,1	0,5	8,74	8,10	9,00	2,3	3,50	5	2004			
190	613,98	98,0	-1,0	0,45	9,34	8,31	7,80	2,2	3,51	6	2005			
191	654,42	98,0	-0,5	0,45	10,51	8,54	5,90	1,4	3,82	7	2006			
192	731,47	99,0	0,1	0,43	11,31	8,66	4,10	2,1	3,92	8	2007			
193	459,30	138,0	0,3	0,39	6,01	6,20	6,10	3,5	3,55	1	2000	W	303	oławski
194	519,98	138,5	0,7	0,36	5,42	6,13	7,04	3,0	2,61	2	2001			
195	432,95	135,2	0,8	0,34	5,64	6,07	7,49	1,7	2,91	3	2002			
196	383,93	135,0	0,7	0,21	5,65	5,84	6,10	3,6	2,58	4	2003			
197	430,09	136,0	0,6	0,20	6,59	5,61	6,10	1,9	1,37	5	2004			
198	477,39	136,0	-0,5	0,19	8,45	5,76	5,10	4,3	0,17	6	2005			
199	523,37	136,0	-0,6	0,17	9,90	6,03	3,40	1,9	0,83	7	2006			
200	547,48	136,0	2,5	0,19	11,04	7,37	2,30	4,1	1,29	8	2007			
201	388,10	70,2	1,5	0,38	3,17	3,35	4,10	0,8	1,72	1	2000	W	304	średzki
202	451,64	70,3	0,5	0,34	3,76	3,25	4,78	1,0	3,09	2	2001			
203	414,97	69,2	2,9	0,31	4,06	3,15	4,37	0,9	3,39	3	2002			
204	381,41	70,0	4,9	0,28	4,42	3,40	4,20	3,3	3,26	4	2003			
205	431,76	70,0	4,4	0,30	4,50	3,20	4,10	2,7	3,46	5	2004			
206	484,73	70,0	1,4	0,30	4,42	3,18	3,60	3,0	3,60	6	2005			
207	559,69	70,0	6,5	0,29	5,29	3,22	2,80	2,8	3,54	7	2006			
208	552,64	70,0	5,5	0,32	5,66	3,42	1,80	3,2	3,57	8	2007			
209	476,10	74,7	2,9	0,68	4,34	5,28	5,80	1,3	11,48	1	2000	W	305	trzebnicki
210	495,32	75,1	5,0	0,62	4,23	5,22	7,20	2,4	9,64	2	2001			
211	441,89	74,5	4,1	0,56	4,18	5,16	7,20	2,3	9,83	3	2002			
212	408,67	75,0	3,1	0,50	3,78	5,12	7,30	6,4	10,03	4	2003			
213	454,80	75,0	4,7	0,50	3,60	5,20	6,90	4,7	7,99	5	2004			

214	479,32	75,0	1,8	0,52	3,69	5,39	3,60	3,0	7,71	6	2005			
215	518,11	76,0	3,9	0,54	3,96	5,50	5,00	3,3	7,09	7	2006			
216	573,84	76,0	6,6	0,54	4,47	5,89	3,40	3,9	7,56	8	2007			
217	327,00	84,9	6,4	1,14	5,91	6,05	4,40	2,6	8,90	1	2000	W	306	wrocławski
218	458,60	85,4	6,4	1,10	6,19	6,38	5,88	3,1	9,60	2	2001			
219	395,46	86,7	12,0	1,06	6,54	6,71	6,83	3,9	8,01	3	2002			
220	369,79	88,0	13,7	1,00	6,70	8,20	7,00	8,9	9,06	4	2003			
221	498,02	89,0	15,3	1,00	6,79	8,41	7,00	2,6	8,86	5	2004			
222	459,84	90,0	15,2	0,98	6,83	9,39	6,30	10,4	6,70	6	2005			
223	546,54	93,0	22,2	1,02	9,76	8,82	4,00	10,2	6,11	7	2006			
224	504,00	95,0	23,0	0,94	16,85	10,01	2,00	12,5	5,82	8	2007			
225	2340,60	2164,5	0,3	0,75	60,75	138,80	21,70	4,3	10,30	1	2000	MW	401	m. Wrocław
226	2340,33	2165,2	2,6	0,73	54,30	135,39	28,52	8,8	10,11	2	2001			
227	2285,40	2182,6	1,0	0,72	48,33	131,98	35,41	5,5	11,46	3	2002			
228	2321,79	2176,0	0,3	0,54	45,51	133,85	37,00	6,1	11,61	4	2003			
229	2730,90	2172,0	1,2	0,55	45,50	135,95	35,40	3,2	11,41	5	2004			
230	3367,37	2172,0	2,2	0,57	48,56	140,57	31,40	3,1	11,35	6	2005			
231	3838,85	2167,0	-0,5	0,40	49,38	148,52	23,30	6,9	11,51	7	2006			
232	4207,27	2161,0	-0,4	0,63	51,32	159,25	13,50	5,2	12,31	8	2007			

Źródło: Rocznik statystyczny Województwa Dolnośląskiego. Tablice przeglądowe. Tab. II. Wybrane dane o podregionach (NTS 3) i powiatach (NTS 4) w r. Lata 2000÷2008. Rocznik statystyczny Województwa Dolnośląskiego. Aneks. Tab. I. Wybrane dane w układzie podregionów (NTS 3) i powiatów (NTS 4) w 2007 r. Tablice przeglądowe. III. Powiaty na tle województwa w R.; Dział IV. Powierzchnia i ludność.

W powyższej tabeli, stanowiącej bazę danych dla opracowania programu modelu ekonometrycznego, dla przejrzystości nie powtarzano informacji:

- nadanego przez autora symbolu literowego i cyfrowego podregionu oraz nazwy powiatu, przy czym: J-W – Podregion jeleniogórsko-walbrzyski (symbol cyfrowy 1), L – Podregion legnicki (2), W – Podregion wrocławski (3), MW – Podregion m. Wrocław (4);
- symboli cyfrowych kolejnych powiatów w ramach podregionów w porządku narastającym np. 101,;
- nazw powiatów oraz miast na prawach powiatu, do których należą: Jelenia Góra, Legnica, Wrocław.

Baza danych panelowych obejmuje 29 obiektów (26 powiatów oraz 3 miasta). W ramach każdego obiektu występuje szereg czasowo-przekrojowy (8 okresów oraz 8 zmiennych objaśniających X1, X3, X5÷X10).

Zmienna objaśniana *Y* - *produkt krajowy brutto* nie występuje w roczniku statystycznych województwa dolnośląskiego w części dotyczącej powiatów. Z tego względu przyjęto jako zmienną objaśnianą:

Y1 - dochód budżetu powiatu lub miasta na prawach powiatu na 1 mieszkańca w zł.

W *Rocznikach Statystycznych Województwa Dolnośląskiego* lat 2000÷2008 występują:

- zmiany w strukturze sprawozdawczości podyktowane między innymi korektą struktury terytorialnej;

- niejednorodność agregacji wartości zmiennej X_4 – przestępstwa stwierdzone liczone dla niektórych lat łącznie z miastami na prawach powiatu i z tego względu pominięto tą zmienną w formułowaniu modelu dla powiatów;

- brak zmiennej X_{11} - użytki rolne w tys. ha w roku poprzednim i zmienna ta też nie jest uwzględniona w budowie modelu dochodu budżetu powiatu na 1 mieszkańca;

- braki wartości danych zmiennych dla niektórych lat; luki te wypełniono obliczając średnie arytmetyczne z lat sąsiednich, czy też przyjmując równomierny przyrost między wartościami określonych lat; uzupełnienia te zaznaczono zacienianiem;

- X_1 obliczono jako iloraz ludności w tys. wg powyższego źródła i powierzchni w km^2 powiatu;

- X_{10} na tysiąc mieszkańców powiatu obliczono jako iloraz ilości miejsc noclegowych w turystycznych obiektach zbiorowego zakwaterowania i ilości ludności ogółem w tys.;

- X_5 , X_6 , X_7 dla jednolitości z opracowanym modelem dla 16. województw Polski dokonano zamiany pracujących w % w stosunku do ogółem pracujących w tys. na pracujących w tys. poniższych grup branż: (rolnictwo, łowiectwo i leśnictwo oraz rybactwo), przemysł i budownictwo, usługi rynkowe + usługi nierynkowe.

5.3. Import danych

Procedura postępowania w zakresie przeniesienia danych statystycznych z arkusza kalkulacyjnego *Excel* jako bazy danych użytkownika w programie Gretl dostępnego w Internecie jest następująca:

1. Skopiowanie danych statystycznych w *Excelu* na wyodrębniony skoroszyt o jednym arkuszu nadając nazwę przykładowo BAZAPOWIAT.xls. Rozpoczynamy od komórki A1 wpisując w pierwszym wierszu nazwy pól kolejno: Y_1 , X_1 , X_3 - X_{10} , n (numer obserwacji).

2. Wywołanie programu *Gretl*.

3. Pobranie pliku BAZAPOWIAT.xls poprzez polecenie: *Plik/Otwórz dane/Import/Excel...* Wskazanie pliku w podkatalogu np. D:/POWIAT a następnie *Otwórz*

4. Podanie w oknie dialogowym początku zakresu importu *Kolumna 1 wiersz 1 Arkusz 1 OK*. Pojawia się informacja: *Importowane dane są traktowane jako niedatowane (dane przekrojowe). Czy chcesz interpretować dane jako szereg czasowy lub dane panelowe? Tak*

5. Okno „Kreator struktury danych” Wybieramy *Dane panelowe Do przodu*.

6. Okno „Kreator struktury danych panelowych” Wybieramy *Stos szeregów czasowych Do przodu*.

7. Okno kreator struktury danych: *Liczba przekrojów (obiektów) w jednostce czasu – 29. Liczba jednostek czasu w przekroju – 8 Do przodu*.

8. Okno „Potwierdź strukturę danych”: *Dane panelowe (stos szeregów czasowych) 29 jednostek przekroju w 8 okresach. Zastosuj*.

Powstał zbiór BAZAPOWIAT.xls* o strukturze:

0	const	1	Y_1	2	X_1	3	X_3	4	X_5	5	X_6	6	X_7	7	X_8	8	X_9	9	X_{10}	10	n
---	-------	---	-------	---	-------	---	-------	---	-------	---	-------	---	-------	---	-------	---	-------	---	----------	----	-----

Program przeniósł wartości zmiennych oraz n dodając numery obserwacji jako 1.1-29.8. Fragment zaimportowanych danych jest następujący:

Obs	Y1	X1	X3	X5	X6
1:1	431,70	69,1	0,7	0,50	7,40
1:2	493,72	69,2	1,0	0,46	6,88
1:3	461,28	67,6	-1,7	0,43	6,54
1:4	394,96	68,0	0,2	0,35	6,23
1:5	440,75	68,0	-0,7	0,33	6,92
1:6	467,79	68,0	-1,5	0,32	7,40
1:7	541,00	68,0	-0,5	0,30	7,90
1:8	550,31	68,0	-1,0	0,32	8,64

Kolejnym krokiem jest skopiowanie zaimportowanej bazy danych jako plik *GRET*L-a z rozszerzeniem *gdt*.

5.4. Generowanie zmiennych

W celu uzyskania modelu z efektami indywidualnymi odnoszącymi się do 29. powiatów zamiast jednej stałej *const* zachodzi potrzeba wygenerowania przez Greta dodatkowych zmiennych zero-jedynkowych o symbolach (*du_1*) do (*du_29*). Wartość jeden występuje w odniesieniu do danego powiatu panelu a pozostałe wartości w kolumnie wygenerowanej zmiennej mają wartość zero. Realizacja powyższego przebiega poleceniem: *Dodawanie zmiennych/zmienna 0-1 panelu*.

Zastosowanie KMNK

Wygenerowanie zmiennych umożliwia zastosowanie klasycznej metody najmniejszych kwadratów dla otrzymania modelu liniowego wielowymiarowego z efektami indywidualnymi. W tym względzie stosujemy następujące postępowanie:

- polecenie: *Model/klasyczna metoda najmniejszych kwadratów*;

- wypełnienie okna dialogowego: *Zmienna zależna*: Y1; *Zmienne niezależne*: X1, X3, X5-X10; *Zmienne 0-1*: (*du_1*) do (*du_29*);

- usuwamy stałą *const*, gdyż chcemy wprowadzić efekty indywidualne. OK

Rezultatem pracy Greta jest poniższy Model 1 stanowiący estymację KMNK z wykorzystaniem 232 obserwacji. Włączono 29 jednostek danych przekrojowych o szeregach czasowych długości = 8, przy czym zmienną zależną jest Y1:

	Współczynnik	Błąd stand.	t-Studenta	Wartość p
X1	-0,383157	0,449798	-0,8518	0,3953
X3	-4,60022	7,08700	-0,6491	0,5170
X5	-227,888	178,796	-1,275	0,2040
X6	-35,1590	9,23800	-3,806	0,0002 ***
X7	50,1244	7,16091	7,000	4,03E-011 ***
X8	-34,1350	6,37158	-5,357	2,37E-07 ***
X9	13,2335	9,85926	1,342	0,1811
X10	1,83079	2,69744	0,6787	0,4981
du_1	661,947	142,399	4,649	6,15E-06 ***
du_2	915,740	183,393	4,993	1,31E-06 ***
du_3	850,961	116,231	7,321	6,29E-012 ***
du_4	312,781	657,830	0,4755	0,6350
du_5	677,220	107,569	6,296	1,98E-09 ***
du_6	760,271	273,665	2,778	0,0060 ***
du_7	612,688	147,054	4,166	4,65E-05 ***

du_8	801,170	88,1330	9,090	1,13E-016	***
du_9	802,561	116,064	6,915	6,53E-011	***
du_10	1149,74	281,232	4,088	6,35E-05	***
du_11	753,560	328,387	2,295	0,0228	**
du_12	918,848	134,016	6,856	9,09E-011	***
du_13	770,428	196,153	3,928	0,0001	***
du_14	781,951	105,382	7,420	3,52E-012	***
du_15	2388,74	391,189	6,106	5,41E-09	***
du_16	760,500	179,933	4,227	3,64E-05	***
du_17	742,017	89,6161	8,280	1,93E-014	***
du_18	817,366	99,6552	8,202	3,13E-014	***
du_19	425,287	198,597	2,141	0,0335	**
du_20	939,504	197,358	4,760	3,76E-06	***
du_21	684,038	97,9563	6,983	4,43E-011	***
du_22	2796,53	854,357	3,273	0,0013	***
du_23	1063,92	114,076	9,326	2,45E-017	***
du_24	877,721	163,870	5,356	2,38E-07	***
du_25	678,898	131,320	5,170	5,78E-07	***
du_26	655,457	93,4101	7,017	3,65E-011	***
du_27	665,631	132,209	5,035	1,09E-06	***
du_28	747,902	219,365	3,409	0,0008	***
du_29	-493,357	1534,96	-0,3214	0,7482	

Średnia arytmetyczna zmiennej zależnej = 793,831

Odczylenie standardowe zmiennej zależnej = 659,125

Suma kwadratów reszt = 3,7467e+006

Błąd standardowy reszt = 138,614

Współczynnik determinacji $R^2 = 0,96267$

Skorygowany współczynnik determinacji $R^2 = 0,95577$

Statystyka $F(36, 195) = 139,671$ (wartość $p < 0,00001$)

Logarytm wiarygodności = -1453,19

Kryterium informacyjne Akaike'a (AIC) = 2980,39

Wylączając stałą, największa wartość p jest dla zmiennej 39 (du_29)

Program *Gretl* proponuje nam usunięcie stałej dotyczącej miasta Wrocławia, jako efektu du_29. Ponieważ efekty chcemy zachować dla eliminacji zmiennych objaśniających posłużymy się statystyką t -Studenta oraz wartością prawdopodobieństwa.

Eliminacja a posteriori

Metoda ta polega na budowaniu kolejnych modeli poprzez eliminowanie zmiennych o największym prawdopodobieństwie lub najmniejszej statystyce t -Studenta. Największa wartość $p = 0,5170$ występuje przy zmiennej $X3$. Po wyłączeniu tej zmiennej otrzymujemy Model 2, zaprezentowany poniżej bez zmiennych (du_1) do (du_29).

	Współczynnik	Błąd stand.	t -Studenta	Wartość p

X1	-0,357144	0,447347	-0,7984	0,4256
X5	-229,467	178,515	-1,285	0,2002
X6	-37,0184	8,76965	-4,221	3,72E-05 ***
X7	49,2093	7,01039	7,019	3,56E-011 ***

X8	-35,3546	6,07924	-5,816	2,42E-08	***
X9	11,1255	9,29531	1,197	0,2328	
X10	1,99804	2,68113	0,7452	0,4570	

Średnia arytmetyczna zmiennej zależnej = 793,831
Odchylenie standardowe zmiennej zależnej = 659,125
Suma kwadratów reszt = 3,7548e+006
Błąd standardowy reszt = 138,409
Współczynnik determinacji $R^2 = 0,96259$
Skorygowany współczynnik determinacji $R^2 = 0,95590$
Statystyka $F(35, 196) = 144,075$ (wartość $p < 0,00001$)
Logarytm wiarygodności = -1453,44
Kryterium informacyjne Akaike'a (AIC) = 2978,89

Interesuje nas na pewno jaka jest wartość krytyczna t_α przy założonym poziomie istotności $\alpha = 0,1$ oraz ilości stopni swobody określonych dla naszego panelu wg zależności: $iss = n - m - k = 232 - 7 - 29 = 196$. Przy czym: $n = 232$ obserwacje panelu, $m = 7$ zmiennych objaśniających modelu, $k = 29$ – ilość obiektów panelu.

Na podstawie tablicy wartości krytycznych t -Studenta dla $\alpha = 0,1$ i $iss = 196$, $t_\alpha = 1,6525$. Biorąc pod uwagę kryterium testu t -Studenta $|t_{\text{obs}}| > t_\alpha$ traktujące o istotności statystycznej danego parametru (współczynnika) a_i , który wyraża wpływ danej zmiennej objaśniającej na zmienną zależną (objaśnianą) kryterium spełniają tylko zmienne X7, X8. Największą wartość p ma zmienna X10 i ją eliminujemy otrzymując Model 3.

	Współczynnik	Błąd stand.	t-Studenta	Wartość p
--	--------------	-------------	------------	-----------

X1	-0,369660	0,446527	-0,8279	0,4088
X5	-224,518	178,190	-1,260	0,2092
X6	-35,9630	8,64477	-4,160	4,75E-05 ***
X7	50,3106	6,84510	7,350	5,17E-012 ***
X8	-34,1542	5,85534	-5,833	2,20E-08 ***
X9	10,0028	9,16206	1,092	0,2763

Średnia arytmetyczna zmiennej zależnej = 793,831
Odchylenie standardowe zmiennej zależnej = 659,125
Suma kwadratów reszt = 3,76544e+006
Błąd standardowy reszt = 138,253
Współczynnik determinacji $R^2 = 0,96248$
Skorygowany współczynnik $R^2 = 0,95600$
Statystyka $F(34, 197) = 148,631$ (wartość $p < 0,00001$)
Logarytm wiarygodności = -1453,77
Kryterium informacyjne Akaike'a (AIC) = 2977,54.

Zwróćmy uwagę na stopniowe zmniejszanie się kryterium informacyjnego Akaike'a bazującego na logarytmie wiarygodności w dotychczasowych modelach 1, 2, 3.

W kolejnym kroku wyeliminujemy X1 wykazujące największe prawdopodobieństwo błędu i uzyskamy w ten sposób Model 4.

	Współczynnik	Błąd stand.	t-Studenta	Wartość p
--	--------------	-------------	------------	-----------

X5	-198,390	175,233	-1,132	0,2589
X6	-37,2243	8,50269	-4,378	1,94E-05 ***

X7	50,1547	6,83707	7,336	5,54E-012 ***
X8	-34,8459	5,79082	-6,017	8,44E-09 ***
X9	10,8572	9,09650	1,194	0,2341

Średnia arytmetyczna zmiennej zależnej = 793,831

Odchylenie standardowe zmiennej zależnej = 659,125

Suma kwadratów reszt = 3,77853e+006

Błąd standardowy reszt = 138,143

Współczynnik determinacji $R^2 = 0,96235$

Skorygowany współczynnik $R^2 = 0,95607$

Statystyka $F(33, 198) = 153,358$ (wartość $p < 0,00001$)

Logarytm wiarygodności = -1454,17

Kryterium informacyjne Akaike'a (AIC) = 2976,35.

Mimo usuwania kolejnych zmiennych objaśniających współczynnik determinacji który jest miarą pomocniczą dobroci dopasowania modelu do danych empirycznych jest na wysokim poziomie $R^2 = 0,96235$. Nasuwa się nam następna zmienna do eliminacji metodą *a posteriori*, jest nią $X5$, zatem kolejny model jest następujący:

	Współczynnik	Błąd stand.	t-Studenta	Wartość p
X6	-40,2737	8,07051	-4,990	1,31E-06 ***
X7	49,6058	6,82469	7,269	8,10E-012 ***
X8	-36,7772	5,53776	-6,641	2,90E-010 ***
X9	13,8990	8,69683	1,598	0,1116

Średnia arytmetyczna zmiennej zależnej = 793,831

Odchylenie standardowe zmiennej zależnej = 659,125

Suma kwadratów reszt = 3,803e+006

Błąd standardowy reszt = 138,241

Współczynnik determinacji R -kwadrat = 0,96211

Skorygowany współczynnik R -kwadrat = 0,95601

Statystyka $F(32, 199) = 157,887$ (wartość $p < 0,00001$)

Logarytm wiarygodności = -1454,92

Kryterium informacyjne Akaike'a (AIC) = 2975,85.

Program Gretl podaje nam jeszcze dodatkowe oznaczenie jako podpowiedź w eliminacji. Jest nią system gwiazdek podawany na końcu każdej linii zmiennej określające poziom istotności: *** - $\alpha = 0,01$, ** - $\alpha = 0,05$, * - $\alpha = 0,1$ puste – brak statystycznej istotności ($\alpha > 0,1$). Tak więc należy usunąć zmienną $X9$, co daje Model 5 jako końcowy naszej procedury eliminacji *a posteriori* zmiennych objaśniających.

	Współczynnik	Błąd stand.	t-Studenta	Wartość p
X6	-36,8497	7,81112	-4,718	4,47E-06 ***
X7	49,8971	6,84871	7,286	7,24E-012 ***
X8	-36,3471	5,55267	-6,546	4,87E-010 ***
du_1	632,944	122,752	5,156	6,04E-07 ***
du_2	847,922	144,761	5,857	1,91E-08 ***
du_3	804,470	84,7242	9,495	6,91E-018 ***
du_4	715,611	90,6469	7,894	1,89E-013 ***
du_5	684,437	80,8454	8,466	5,40E-015 ***

du_6	784,918	240,835	3,259	0,0013	***
du_7	655,990	88,2441	7,434	3,02E-012	***
du_8	805,853	75,7373	10,64	3,07E-021	***
du_9	704,219	70,1387	10,04	1,81E-019	***
du_10	982,111	229,493	4,279	2,90E-05	***
du_11	631,927	280,480	2,253	0,0253	**
du_12	840,844	100,727	8,348	1,14E-014	***
du_13	749,362	171,034	4,381	1,90E-05	***
du_14	731,072	77,2452	9,464	8,48E-018	***
du_15	2180,69	160,022	13,63	2,43E-030	***
du_16	708,331	136,964	5,172	5,61E-07	***
du_17	695,874	63,5443	10,95	3,62E-022	***
du_18	762,190	75,2866	10,12	1,03E-019	***
du_19	394,693	164,321	2,402	0,0172	**
du_20	937,066	173,404	5,404	1,84E-07	***
du_21	671,785	83,5968	8,036	7,91E-014	***
du_22	2131,56	197,313	10,80	1,00E-021	***
du_23	1042,63	66,5715	15,66	1,31E-036	***
du_24	811,017	139,345	5,820	2,31E-08	***
du_25	634,742	108,443	5,853	1,95E-08	***
du_26	592,565	76,1661	7,780	3,79E-013	***
du_27	573,670	90,6155	6,331	1,57E-09	***
du_28	545,175	120,931	4,508	1,11E-05	***
du_29	-1196,26	1188,03	-1,007	0,3152	

Średnia arytmetyczna zmiennej zależnej = 793,831

Odchylenie standardowe zmiennej zależnej = 659,125

Suma kwadratów reszt = 3,85181e+006

Błąd standardowy reszt = 138,777

Współczynnik determinacji $R^2 = 0,96162$

Skorygowany współczynnik $R^2 = 0,95567$

Statystyka $F(31, 200) = 161,642$ (wartość $p < 0,00001$)

Logarytm wiarygodności = -1456,4

Kryterium informacyjne Akaike'a (AIC) = 2976,8.

Tak więc sprawdziliśmy „Model 1” o 8. zmiennych objaśniających do postaci „Modelu 5” o trzech zmiennych, które są następujące:

X6 - pracujący w tys. (przemysł, budownictwo),

X7 - pracujący w tys. (usługi rynkowe i nierynkowe),

X8 - bezrobotni zarejestrowani ogółem w tys.

Przypomnijmy sobie jeszcze zmienną objaśnianą: Y1 - dochód budżetu powiatu lub miasta na pracach powiatu na 1 mieszkańca w zł.

Test normalności rozkładu reszt

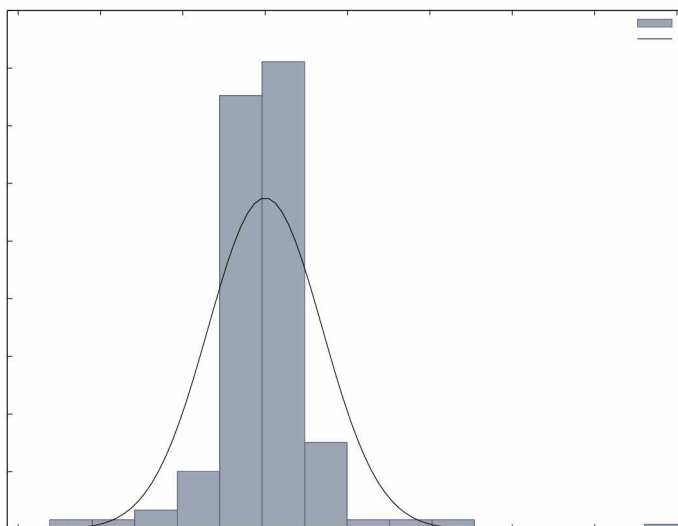
Program *Gretl* dysponuje szeregiem funkcji do weryfikacji modelu opracowanego na bazie danych panelowych klasyczną metodą najmniejszych kwadratów. Szersze ich omówienie spotykamy w pracy T. Kufla podanej we wcześniejszym przypisie. Dla przykładu sprawdzimy normalność rozkładu reszt. Wymaga to wywołania polecenia przy aktywnym modelu końcowym w naszym przykładzie piętego. Korzystamy z polecenia: *Testy/Test normalności rozkładu reszt*. W efekcie otrzymujemy

rozkład częstości dla testowej zmiennej oznaczonej jako uhat1 i 1-232 obserwacji naszego panelu. Program tworzy 15 przedziałów i jest to liczba, która jest określona jako pierwiastek z 232. Ponadto oblicza średnią = -1,78861e-014, która jest bliska zeru oraz odchylenie standardowe = 138,777.

Przedziały	Średnia	Liczba	Częstość	Skumulowana
< -420,00	-471,57	2	0,86%	0,86%
-420,00 - -316,87	-368,44	2	0,86%	1,72%
-316,87 - -213,74	-265,31	4	1,72%	3,45%
-213,74 - -110,61	-162,18	12	5,17%	8,62%
-110,61 - -7,4793	-59,045	90	38,79%	47,41%
-7,4793 - 95,652	44,086	97	41,81%	89,22%
95,652 - 198,78	147,22	18	7,76%	96,98%
198,78 - 301,91	250,35	2	0,86%	97,84%
301,91 - 405,04	353,48	2	0,86%	98,71%
405,04 - 508,18	456,61	2	0,86%	99,57%
508,18 - 611,31	559,74	0	0,00%	99,57%
611,31 - 714,44	662,87	0	0,00%	99,57%
714,44 - 817,57	766,00	0	0,00%	99,57%
817,57 - 920,70	869,13	0	0,00%	99,57%
>= 920,70	972,26	1	0,43%	100,00%

Hipoteza zerowa: dystrybucja empiryczna posiada rozkład normalny. Test Doornika-Hansena (1994) - transformowana skośność i kurtoza: Chi-kwadrat(2) = 157,524 z wartością $p = 0,00000$.

Prawie zerowa wartość prawdopodobieństwa popełnienia błędu wskazuje, że rozkład reszt modelu jako różnicy między wartościami empirycznymi a teoretycznymi jest rozkładem normalnym. Obrazem graficznym rozkładu częstości jest rysunek 2.1.



Rys. 2.1. Wykres rozkładu reszt

Zebranie danych statystycznych dotyczących powiatów przykładowego województwa dolnośląskiego dla potrzeb modelu ekonometrycznego w układzie jednolitości informacji panelu było zadaniem pracochłonnym. Jak wspomniano okres 2000-2008 to istotne zmiany w strukturze terytorialnej jak i sposobie agregacji danych. Nadzieją na praktyczne zastosowanie modelu to jego ograniczenie tylko do trzech istotnych zmiennych objaśniających oraz wyprofilowanie się struktury sprawozdawczości GUS-u po wejściu do Unii Europejskiej. Na uwagę zasługuje program *Gretl* - dostępny w Internecie, który jest kompatybilny z powszechnie stosowanym arkuszem kalkulacyjnym *Excel*, służącym między innymi do zestawienia i obróbki danych. Polecenia programu *Gretl* umożliwiają szybkie i sprawne budowanie modelu ekonometrycznego. Autor ma nadzieję, że niniejszy materiał będzie stanowił wygodną instrukcję w formułowaniu modeli zjawisk ekonomicznych dotyczących nie tylko powiatów.

6. Wyznaczenie efektywności technicznej programem DEAP



6.1. Wprowadzenie

W dobie kryzysu gospodarczego pracodawcy coraz częściej poszukują obiektywnej metody klasyfikowania zaradności swoich obiektów gospodarczych. Słyszycie głosy nawet z małych firm o potrzebie obiektywnego sposobu wynagradzania swoich grup a nawet pojedynczych pracowników z sfery marketingu oraz produkcji. Przykładem jest firma zajmująca się pozycjonowaniem stron WWW. Zadała pytanie, czy istnieje oprogramowanie metody optymalizującej wyselekcjonowanie pracowników marketingu, których cechuje zaradność w firmie, czyli tych co pracują efektywnie z punktu widzenia określonych miar wejścia/wyjścia, jak również wskazanie tych co powinni podciągnąć się w swej pracy. Dotychczasowe bowiem miary dobierane subiektywnie rodzą okresowe konflikty między grupami branżowymi firmy internetowej.

Było to spojrzenie w skali mikroekonomicznej. Analogiczne rozumowanie możemy odnieść do obiektów struktury terytorialnej w skali makroekonomicznej tj. podregionów, województw, powiatów. W tym względzie z próbą analizy efektywności tzw. technicznej gospodarowania województw Polski spotykamy się w publikacji A. Becker⁴⁴. Zastosowano tam analizę DEA obsługiwana programem *DEAP Version 2.1*⁴⁵. W wymienionej publikacji skupiono się na elementach merytorycznych analizy mniej natomiast uwagi poświęcono samemu programowi *DEAP*. Ponieważ zastosowanie tego programu może być szersze w niniejszym opracowaniu podano przykład jego zastosowania, lecz w ujęciu instrukcji postępowania.

6.2. Opis programu DEAP

Program komputerowy *DEAP* został napisany w języku *Fortran* i może być eksploatowany na mikrokomputerach IBM PC. Bazuje na definicjach zbiorów *DOS*-a lecz może być również stosowany w środowisku *Windows*. Opis programu *DEAP* występuje w poradniku wydanym przez Uniwersytet Nowej Anglii w Australii⁴⁶. Wymieniona publikacja opisuje program komputerowy, który został napisany do prowadzenia tzw. analizy zawijaniem *DEA* dla celów obliczania wydajności produkcji, która może mieć jednak szersze zastosowanie. Coell zastosował w programie metody opracowane między innymi przez Rolf Fare oraz Shawna Grosskopfa. Program obejmuje dwie opcje odnoszące się do szeregów obserwacji zmiennych wejścia/wyjścia obiektów, z rozróżnieniem bazowania na nakładach czy efektach, a trzecia opcja wykorzystuje dane panelowe, czyli szeregi zmiennych dla obiektów w jednakowych okresach czasowych.

Analiza *DEA* wiąże się z wykorzystaniem metody programowania liniowego dla określenia efektywności danego obiektu w stosunku do teoretycznego określonego przez program. Wymieniony autor uważa, że *DEA* stanowi nowoczesny sposób pomiaru efektywności przykładowo zastosowanej określonej technologii w danym przedsiębiorstwie, gdzie możemy wyróżnić wiele oczekiwanych efektów oraz występują różne rodzajowo nakłady. Coell powołując się na innych autorów zaproponował, aby efektywność firmy składa się z dwóch części:

- *sprawności technicznej (skuteczności)*, która jest obrazem zdolności firmy do uzyskania maksymalnego wyjścia z realizacji wymagań podanych w zbiorze danych wejściowych;

⁴⁴ Internet, hasło: efektywność + PKB, Becker A.: *Analiza efektywności gospodarowania województw Polski*.

⁴⁵ Internet: <http://www.une.edu.au/economics/emet/cepa.htm>.

⁴⁶ Coell T.J.: *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program*, CEPA Working Papers, Department of Econometrics University of New England, Australia.

- skuteczności przydziału, która odzwierciedla zdolność firmy do korzystania z zasobów w optymalnej proporcjach.

Zdaniem Coella po uwzględnieniu cen dla zmiennych te dwa środki można następnie łączyć w celu uzyskania pomiaru tzw. wydajności ekonomicznej.

Po tym wprowadzeniu wróćmy do naszego programu. Zainstalowałem go z Internetu po dojściu do strony: www.une.edu.au/economics/emet/cepa.htm. Najpierw tworzony jest katalog *WinRun* na dysku twardym np. *D:* Pliki wzorcowe, podręcznik oraz program ściągany jest w postaci skompresowanej z rozszerzeniem *zip* i wymaga następnie „rozpakowania”. Program *DEAP Version 2.1* bazuje on na pliku wsadowym danych wejściowych oraz pliku instrukcji w formacie tekstowym. Uruchomienie programu *DEAP.EXE* następuje poprzez podwójne kliknięcie na jego ikonie, a następnie wpisanie z klawiatury zbioru instrukcji. Uruchomienie programu *DEAP* wymaga jak już wspomniano wcześniejszego rozpakowania (zainstalowania) na dysku następujących zbiorów:

- program *DEAP.EXE*,
- plik startujący *DEAP.000*,
- pliki tekstowe danych,
- pliki tekstowe instrukcji,
- pliki wyjściowe.

Plik *DEAP.000* zapamiętuje parametry realizacji programu w celu ich użycia w późniejszym okresie. Zbiór danych jak i instrukcji zakładany jest przed uruchomieniem programu. W pakiecie programu *DEAP* występują cztery komplety różnych przykładów zdefiniowania zbiorów tekstowych bazujących na określonych zmiennych wejścia/wyjścia:

- *eg1-dta.txt* (5 obserwacji - firm, 1 wyjście, 2 wejścia); *eg1-ins.txt*; *eg1-out.txt*; opcja *CRS*;
- *eg2-dta.txt* (5 obserwacji, 1 wyjście, 1 wejście); *eg2-ins.txt*; *eg2-out.txt*; opcja *VRS*;
- *eg3-dta.txt* (jak w *eg1-dta.txt*), *eg3-ins.txt*; *eg3-out.txt*; opcja z cenami dla wartości zmiennych wejścia, opcja *CRS*;
- *eg4-dta.txt* (panel 5 obserwacji po 3 okresy - lata), *eg4-ins.txt*; *eg4-out.txt*; opcja *CRS/VRS* z zaznaczeniem panelu - opcja *malmquist*.

Opcja *CRS* dotyczy maksymalizacji funkcji celu licznika, a *VRS* mianownika - we wzorze na efektywność (gospodarność) techniczną obiektu⁴⁷:

$$\text{Efektywność} = \frac{\sum_{r=1}^s \mu_r \text{Efekt}}{\sum_{i=1}^m v_i \text{Nakład}}$$

przy czym μ_r, v_i są to wagi określające ważność odpowiednio efektów/nakładów.

W niektórych publikacjach zamieszczonych w Internecie program *DEAP* stosowany jest do sklasyfikowania określonych firm, oddziałów banków pod względem ich efektywności technicznej. Program bazuje na analizie komputerowej odległości miar obserwacji od najlepszych dla teoretycznie ustalonej firmy. Program sam ustala powyższe wagi po kątem maksymalizacji efektywności każdego obiektu. W standardowym rozwiązaniu problemu optymalizacyjnego następuje wyznaczenie odległości rzeczywistego poziomu produkcji od teoretycznie stwierdzonej granicy efektywności. *DEA* pozwala na ustalenie krzywej efektywności na której znajdują się wszystkie najbardziej efektywne jednostki badanej zbiorowości. Jeśli obiekty traktowane jako jednostki decyzyjne (DMU) znajdują się poza krzywą efektywności, są uważane jako nieefektywne technicznie.

⁴⁷ Becker A.: Internet *Analiza efektywności* ..., op. cit.

6.3. Dane empiryczne

Program *DEAP* wymaga plików tekstowych. Można je założyć bezpośrednio w edytorze tekstów przykładowo w *NOTATNIK*-u. Często jednak zachodzi potrzeba importu obszernych plików z innych programów, przykładowo z arkusza kalkulacyjnego *Excel*. Dla zaprezentowania procedury posłużono się danymi panelowymi z wcześniej założonej bazy dla potrzeb modelu ekonometrycznego produktu brutto *PKB* dla podregionów. Dane zaczerpnięto z tabel w rocznikach statystycznych 2000-2007: wybrane dane o podregionach (lata 2000÷2003), podstawowe dane o regionach (lata 2004÷2007). Dane obejmują one następujące zmienne, którym nadano symbole⁴⁸:

- Y - produkt brutto (ceny bieżące) na 1. mieszkańca w roku poprzednim w zł;
- X_1 - ludność na 1 km² powierzchni ogólnej;
- X_2 - przyrost naturalny na 1000 ludności;
- X_3 - saldo migracji wewnętrznych i zagranicznych na pobyt stały (na 1000 ludności);
- X_4 - przestępstwa stwierdzone na 1000 ludności;
- X_5 - pracujący w tys. (rolnictwo, łowiectwo i leśnictwo, rybactwo);
- X_6 - pracujący w tys. (przemysł i budownictwo);
- X_7 - pracujący w tys. (usługi rynkowe i nierynkowe);
- X_8 - bezrobotni zarejestrowani ogółem w tys.;
- X_9 - mieszkania oddane do użytkowania na 1000 ludności;
- X_{10} - miejsca noclegowe w turystycznych obiektach zbiorowego zakwaterowania w tys.;
- X_{11} - užitki rolne w tys. ha w roku poprzednim.

Dla powyższych zmiennych współczynnik zmienności przekracza wartość krytyczną 10%. Tak więc wstępnie wszystkie zmienne mogą wejść do badania efektywności podregionów.

Ze względu na zastosowanie programowania liniowego w *DEAP* pominięto zmienne przyjmujące wartości ujemne, a są nimi X_2 oraz X_3 . Jako efekt uznano zmienną Y , natomiast jako nakłady wartości zmiennych X_5 , X_6 , X_7 , X_9 , X_{10} , X_{11} . Pominięto zmienne X_4 i X_8 jako efekty negatywne. Rozwiązanie wersję z ich udziałem pozostawia się Czytelnikowi, z tym że należy przyjąć je w postaci $1/y$, gdyż zdaniem A. Becker odwrótność tego typu efektu można traktować jako nakład⁴⁹.

Procedura postępowania związana z importem pliku z *Excelsa* do podkatalogu programu *DEAP* jest następująca:

- skopiowanie skoroszytu,
- pozostawienie tylko jednego arkusza z danymi,
- usunięcie wiersza nagłówka,
- zamiana wartości numerycznych na tekstowe poprzez zaznaczenie danych i wywołanie menu *Format/komórka/tekstowe*,
- automatyczna zamiana przecinka w wartościach danych zmiennych na kropkę dziesiętną korzystając z menu *Edycja/zamień*,
- przemieszczenie kolumn zmiennych w kolejności najpierw dotyczące efektów później nakładów,
- wprowadzenie kolumny numeru obiektu (podregionu) oraz kolumny kolejnego okresu (roku),
- posortowanie narastające zbioru w układzie najpierw okres a później numer podregionu,
- usunięcie kolumny obiektów oraz kolumny okresów,
- zapisanie tak przygotowanego zbioru z rozszerzeniem *xls*,
- zapisanie zbioru jako tekstowego *DOS*-a z rozszerzeniem *txt* w podkatalogu programu *DEAP* nadając mu nazwę typu *egn-dta.txt*.

⁴⁸ Wornalkiewicz W, *Etapy i procedury budowy modelu produktu brutto podregionów*, Artykuł, wydawnictwo WSZiA Opole, 2009.

⁴⁹ Internet hasło: efektywność+PKB, Becker A.: *Analiza efektywności gospodarowania województw Polski*.

Zastosowany materiał empiryczny obejmuje 45 podregionów, które wymieniono w tabeli 17.1.

Tab. 6.1. Podregiony

Lp.	Nazwa	Lp.	Nazwa
1	łódzki	24	gorzowski
2	piotrkowsko-skierniewicki	25	zielonogórski
3	miasto Łódź	26	pilski
4	ciechanowsko-płocki	27	poznański
5	ostrolęcko-siedlecki	28	kaliski
6	warszawski	29	koniński
7	radomski	30	miasto Poznań
8	miasto Warszawa	31	szczeciński
9	krakowsko-tarnowski	32	koszaliński
10	nowosądecki	33	jeleniogórsko-wałbrzyski
11	miasto Kraków	34	legnicki
12	częstochowski (północnośląski)	35	wrocławski
13	bielsko-bialski (południowośląski)	36	miasto Wrocław
14	centralny śląski	37	opolski
15	rybnicko-jastrzębski ⁴⁰	38	bydgoski
16	Białkopodlaski ⁴¹	39	toruńsko-włocławski
17	chełmsko-zamojski ⁴²	40	słupski
18	lubelski	41	gdański
19	rzeszowsko-tamobrzeski	42	Gdańsk-Gdynia-Sopot
20	krośnieńsko-przemyski	43	elbląski
21	białostocko-suwański	44	olsztyński
22	łomżyński	45	ełcki
23	świętokrzyski		

6.4. Przykład zastosowania

W naszym przykładzie analizy efektywności technicznej podregionów przyjęliśmy jako efekt: Y - produkt krajowy brutto PKB (ceny bieżące) na 1. mieszkańca w roku poprzednim w zł, natomiast jako nakłady:

- X_1 - ludność na 1 km² powierzchni ogólnej;
- X_5 – pracujący w tys. (rolnictwo, łowiectwo i leśnictwo, rybactwo);
- X_6 - pracujący w tys. (przemysł i budownictwo);
- X_7 - pracujący w tys. (usługi rynkowe i nierynkowe);
- X_9 - mieszkania oddane do użytkowania na 1000 ludności;
- X_{10} - miejsca noclegowe w turystycznych obiektach zbiorowego zakwaterowania w tys.;
- X_{11} - użytki rolne w tys. ha w roku poprzednim.

W poniższej tabeli efekt (PKB) zapisany jest w pierwszej kolumnie, a w siedmiu kolejnych nakłady. Zwróćmy uwagę na wymagane wypoźycjonowanie kolumn do lewej strony i zastosowanie kropki dziesiętnej. Dzięki posortowaniu danych według kolejności: numer okresu (roku) rosnąco, numer podregionu rosnąco, dane w tabeli wejściowej - stanowiącej plik nazwany *EAP4-dta.txt* - ułożone są kolejnymi latami 1-7 a w ramach nich dla 1-45 podregionów.

Tab. 6.2. Lista danych wejściowych dla okresu 1. i 45. podregionów

11442	110	164.2	75.3	83.7	1.1	9.4	629.9
12254	97	177.3	76.7	79.3	1.7	8.6	611.8
20645	2694	3.9	73.5	140.7	1.5	5.7	11
14676	83	117.8	47	62.2	2	2.9	572.5

9842	64	181.8	37.6	65.6	1.8	6	808.1
14089	159	139.6	97	127.9	5.3	3.9	532.2
9789	106	147.3	413.2	64.9	1.8	3.8	464.6
48217	3258	485.5	190.1	614.1	8.9	16.4	14.8
11035	190	245.5	99.4	117.4	1.7	6.8	492.4
10091	147	228	56.8	99	2	49.4	375.2
26330	2269	4.5	85.9	176.4	4.6	10.3	14.6
13688	179	98.9	93.2	111.4	1.2	6.6	323.9
16081	324	103.9	203.5	173.7	1.9	29.6	255.3
20666	1729	10.3	273.306	322.392	0.5	6.3	39.5

19162	477	8.5	69.6	65.8	1.5	1.9	67.9
9631	54	76.1	14.2	93.5	1.1	7	381.9
9252	73	190.9	30.6	61.7	0.9	6.3	644.7
12531	125	240.7	79.6	145.2	1.6	11.1	693.2
12452	156	226.7	106.7	124.5	1.5	8	437.2
10753	92	219.4	61.9	96.3	1.2	14.2	505.1
12534	61	154.4	53	106.4	3.8	13.6	831.3
8860	60	74.6	13.4	25.7	1.2	2.5	368.4
12435	113	290.4	92.3	135.8	1.4	9.8	731.9
15132	63	22.9	32	46.7	2.7	7.9	234.5
14030	81	34.9	49.2	73.1	1.7	12	319
13002	64	38.3	36.6	38.7	1.6	4.4	335.2
14911	98	110.7	112	106.8	2.4	20.1	732.7
12145	113	119	74.9	74.7	1.4	4	496.7
12279	99	78.1	38.6	38.3	1.3	7.7	320.1
32796	2200	20.3	69	154.4	3.9	6.6	8.2
17699	90	50	84.2	146.6	2.8	71.9	644.9
12702	59	34.9	40.2	63.9	1.7	58.5	469.2
12765	134	79.4	112.1	139.3	1.3	41.2	571.4
18879	124	32.1	53.4	64.5	1.2	2.8	247.3
11196	85	47.6	30.2	32.4	1.7	3.8	332.2
25285	2165	3.3	60.8	140.7	4.3	6.5	12.7
13320	115	105.5	88.6	113.5	1	6.4	579.8
15529	116	76.5	98.5	126.2	1.7	16.5	533
12754	118	123.3	85	113.6	1.9	13.5	629.5
11399	60	32.6	38.9	50.3	2.2	37.2	347.9
11965	98	74.2	64.7	74.2	2.8	43.8	548.6
24376	1817	5.1	74.4	151.6	2.5	16.5	90.8
12056	73	46.1	38.7	52	1.9	10.1	444
13727	61	44.7	45.2	78.8	3.1	25.2	549.7
9915	46	28.3	15.2	24.2	2.5	13.6	311.6

Źródło: Opracowanie własne z zastosowaniem programu *Excel*; (fragment jednego członu zbioru tekstowego *EAP4-dta.txt*), dziesiątki wierszy zaznaczono czcionką pogrubioną.

Plik tekstowy instrukcji tworzymy poprzez skopiowanie odpowiedniego pliku wzorcowego w naszym przypadku *eg4-ins.txt* i nadanie nowemu plikowi nazwy własnej np. *eg1H-ins.txt*, a następnie dokonujemy zmiany w parametrach, a mianowicie:

- wprowadzamy własną nazwę zbioru danych na *EAP4-dta.txt*;
- wcześniej kopiujemy przykładowy zbiór wyjścia *eg4-out.txt*, zapisujemy pod własną nazwą *eg1H-out.txt*;
- zmieniamy ilość podregionów (firm) na 45, a ilość okresów na 7;
- podajemy ilość zmiennych wyjściowych - 1, oraz zmiennych wejściowych - 7;
- podajemy symbol **1** oznaczający zorientowanie się modelu na zmienną wyjściową;
- określamy metodę obliczeń jako *CRS*, czyli bazującą na maksymalizacji wartości licznika w funkcji efektywności technicznej;
- zaznaczamy cyfra **2**, że mamy do czynienia z bazą danych panelowych (*MALMQUIST-DEA*).

Tab. 6.3. Lista instrukcji zapisana w pliku *eg1H-ins.txt*

EAP4-dta.txt	DATA FILE NAME
eg1H-out.txt	OUTPUT FILE NAME
45	NUMBER OF FIRMS
7	NUMBER OF TIME PERIODS
1	NUMBER OF OUTPUTS
7	NUMBER OF INPUTS
1	0=INPUT AND 1=OUTPUT ORIENTATED
0	0=CRS AND 1=VRS
2	0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA, 3=DEA(1-STAGE), 4=DEA(2-STAGE)

Źródło: Opracowanie własne z zastosowaniem pliku przykładowego w programie *DEAP* wersja 2.1.

6.5. Uruchomienie programu

Korzystamy z programu *DEAP.EXE* zapisanego podczas instalacji z Internetu w podkatalogu *D:\WinRun\DEAPA-xp1*. Jak już wspomniano po podwójnym kliknięciu na ikonie tego programu wpisujemy nazwę zbioru instrukcji: *eg1H-ins.txt*. System zapisuje wyniki w pustym zbiorze rezultatów: *eg1H-out.txt*. Zbiór ten po zapisaniu wyników jest bardzo obszerny, gdyż zawiera raporty wskaźników efektywności odnoszące się do kolejnych z siedmiu lat, a w ramach nich do poszczególnych 45. podregionów. Ponadto dla poszczególnych lat i podregionów podane są wskaźniki dla okresu bieżącego *t* oraz wcześniejszego *t-1* oraz następnego *t+1* z uwzględnieniem opcji *CRS* a w kolumnie czwartej podano dodatkowo wartości wskaźników dla opcji *VRS*. Przykład fragmentu odnoszącego się tylko do pierwszego roku (okresu) oraz 45. podregionów listy rezultatów podano w tabeli 17.4. Dla pierwszego okresu nie występuje okres wcześniejszy stąd wartości zerowe w kolumnie *t-1*. Interesuje nas przede wszystkim kolumna *t*. Na zakończenie podlisty dla 45. podregionów danego okresu przykładowo pierwszy program oblicza efektywność średnią roku, która dla okresu pierwszego wynosi: 0,819. W tabeli współczynnik 1 odpowiada efektywności, a wynik mniejszy od jeden oznacza nieefektywność techniczną⁵⁰.

Tab. 6.4. Lista rezultatów z realizacji programu

Results from DEAP Version 2.1
Instruction file = eg1H-ins.txt
Data file = EAP4-dta.txt
Output orientated Malmquist DEA

⁵⁰ Internet hasło: efektywność+PKB, Becker A.: *Analiza efektywności gospodarowania województw Polski*.

DISTANCES SUMMARY				
year = 1				
firm crs te rel to tech in yr vrs				
no.	*****			
	t-1	t	t+1	te
1	0.000	0.678	0.833	0.725
2	0.000	0.708	0.668	0.725
3	0.000	1.000	1.222	1.000
4	0.000	1.000	0.864	1.000
5	0.000	0.733	0.681	0.734
6	0.000	0.568	0.463	0.734
7	0.000	0.543	0.504	0.559
8	0.000	0.983	0.857	1.000
9	0.000	0.411	0.549	0.566
10	0.000	0.450	0.463	0.528
11	0.000	0.969	0.787	1.000
12	0.000	0.709	0.941	0.720
13	0.000	0.626	0.725	0.775
14	0.000	1.000	1.470	1.000
15	0.000	1.000	1.000	1.000
16	0.000	1.000	0.928	1.000
17	0.000	0.793	0.849	1.000
18	0.000	0.616	0.648	0.663
19	0.000	0.527	0.706	0.649
20	0.000	0.715	0.747	0.724
21	0.000	0.855	0.783	0.863
22	0.000	1.000	0.818	1.000
23	0.000	0.682	0.729	0.688
24	0.000	1.000	0.968	1.000
25	0.000	0.925	0.865	0.929
26	0.000	1.000	0.939	1.000
27	0.000	0.755	0.700	0.853
28	0.000	0.666	0.740	0.672
29	0.000	0.996	0.847	1.000
30	0.000	1.000	1.164	1.000
31	0.000	0.897	0.851	1.000
32	0.000	1.000	0.948	1.000
33	0.000	0.625	0.816	0.673
34	0.000	1.000	1.410	1.000
35	0.000	1.000	0.845	1.000
36	0.000	1.000	0.852	1.000
37	0.000	0.842	1.043	1.000
38	0.000	0.793	0.816	0.836
39	0.000	0.623	0.597	0.684
40	0.000	0.831	0.789	0.843
41	0.000	0.577	0.539	0.686
42	0.000	1.000	1.028	1.000
43	0.000	0.814	0.747	0.825
44	0.000	0.937	0.857	0.946
45	0.000	1.000	0.916	1.000
mean	0.000	0.819	0.834	0.858
[Note that t-1 in year 1 and t+1 in the final year are not defined]				

Źródło: Opracowanie własne w programie *DEAP Version 2.1* (fragment zbioru *eg1H-out.txt*).

Ponadto w liście rezultatów znajdują się raporty porównawcze *MALMQUIST INDEX SUMMARY* współczynników efektywności technicznej *effch* bieżącego okresu do współczynników okresu wcześniejszego, które rozpoczynają się od okresu drugiego dla każdego podregionu. Ten fragment tabeli ze względu na obszerność nie zamieszczono w niniejszym materiale, natomiast w poniższej tabeli podano zestawienie sumaryczne średnich rocznych z porównań w układzie okresów 2-7 dla 45. podregionów.

Tab. 6.5. Średnie geometryczne roczne dla okresów podregionów

MALMQUIST INDEX SUMMARY OF ANNUAL MEANS	
year	effch
2	1.038
3	0.995
4	0.869
5	1.163
6	0.949
7	0.995
mean	0.998

Źródło: Opracowanie własne programem *DEAP Version 2.1*.

Oprócz zestawienia sumarycznego porównawczego według okresów w liście rezultatów znajduje się również raport w układzie poszczególnych 45. podregionów.

Tab. 6.6. Średnie geometryczne dla podregionów *MALMQUIST INDEX SUMMARY OF FIRM MEANS*

firm	effch	firm	effch
1	0.993	24	1.000
2	1.000	25	0.989
3	1.000	26	1.000
4	1.000	27	0.985
5	1.033	28	0.981
6	0.972	29	0.975
7	1.019	30	1.000
8	0.986	31	0.992
9	1.007	32	1.000
10	0.984	33	1.039
11	1.005	34	1.000
12	1.007	35	0.990
13	1.081	36	0.977
14	1.000	37	1.020
15	1.000	38	0.982
16	0.998	39	0.983
17	0.982	40	1.022
18	0.966	41	0.981
19	0.974	42	0.970
20	0.956	43	0.993
21	1.013	44	1.011
22	1.000	45	1.000
23	1.039	mean	0.998

[Note that all *Malmquist* index averages are geometric means - Należy pamiętać, że wszystkie średnie indeksu *Malmquist* są średnimi geometrycznymi]

Źródło: Opracowanie własne w Excelu po zastosowaniu programu *DEAP Version 2.1*.

* * *

Zaprezentowany dobór danych empirycznych dla oceny efektywności technicznej podregionów może budzić zastrzeżenia, lecz jest to tylko przykład dla zademonstrowania procedury uzyskania rezultatów na podstawie danych wejściowych i sterujących. Analogicznie można rozwiązać dowolne zagadnienie, gdyż program *DEAP* poprzez definiowanie zbiorów oraz parametrów wejścia/wyjścia jest otwarty na użytkownika. Wprowadzenie cen w odniesieniu do składników efektów czy też nakładów umożliwia określenie efektywności ekonomicznej. Zaletą programu *DEAP* jest możliwość wprowadzania na wejściu danych panelowych w postaci szeregów obiektowo-czasowo-przekrojowych.

7. Określenie efektywności modelem CCR z karami

7.1. Wstęp

Nawiązując do postulatów jakie powinny spełniać dane empiryczne na podstawie których wyznaczana jest efektywność technologiczna modelami z grupy DEA bazowano na danych statystycznych dotyczących województw Polski^{51 52}. Jako rezultaty przyjęto:

$C_o - X_5$ – odpady wytworzone (w ciągu roku, z wyłączeniem odpadów komunalnych) w tys. t.; dane te stanowiące między innymi efekt negatywny realizacji produkcji przemysłu jak i produkcji budowlano-montażowej w budownictwie zapisano w formie różnicy od stałej C_o ; stałą tą wyznaczono na podstawie szeregu obserwacji 9. okresów czasowych (lata 1999-2007) osobno dla każdego z 16. województw zaokrąglając subiektywnie do liczb całkowitych;

X_{18} – (produkcja p) - produkcja sprzedana przemysłu (ceny bieżące) w mln zł,

X_{19} – (produkcja b) - produkcja budowlano-montażowa (ceny bieżące) w mln zł.

Dla celów testowania modeli DEA, w tym modelu CCR z karami za luzy w nakładach i rezultatach przyjęto dane:

X_9 – (pracujący p) - pracujący w przemyśle (stan w dniu 31 XII) w tys.,

X_{10} – (pracujący b) - pracujący w budownictwie (stan w dniu 31 XII) w tys.,

X_{20} – (nakłady inw. p) - nakłady inwestycyjne w przemyśle (ceny bieżące) w mln zł,

X_{21} – (nakłady inw. b) - nakłady inwestycyjne w budownictwie (ceny bieżące) w mln zł.

Tab. 7.1. Dane obiektowo-czasowe w zakresie potrzeb modelu CCR

W	T	C-X5	X18	X19	X9	X10	X20	X21	X5	MAX(X5 _o)	C	C-X5
1	1	728,4	26540,7	4337,5	264,4	47,6	2434,6	273,5	5171,6			728,4
1	2	92,2	30492,2	3739,4	240,9	43,7	2670,4	270,3	5807,8			92,2
1	3	636,3	32060,4	2988,4	237,5	37,7	2687,7	277,6	5263,7			636,3
1	4	858,2	32845,0	2875,6	227,3	34,7	2290,1	321,2	5041,8			858,2
1	5	1018,4	33607,4	2936,4	223,5	32,8	2482,6	357,3	4881,6			1018,4
1	6	1015,7	35742,0	3337,6	226,3	35,5	3189,8	581,3	4884,3			1015,7
1	7	1450,5	37002,0	4301,2	223,5	37,3	3901,8	1072,3	4449,5			1450,5
1	8	1779,3	39976,0	4669,6	230,0	41,0	3850,0	1070,0	4120,7			1779,3
1	9	2004,8	44450,0	6650,0	241,2	46,7	5479,1	1148,4	3895,2	5807,8	5900	2004,8
2	1	3107,7	85240,9	15421,2	407,4	151,6	6779,5	3216,8	3692,3			3107,7
2	2	2416,5	102269,6	19407,8	389,1	140,8	5870,1	3270,5	4383,5			2416,5
2	3	1740,1	107640,2	14575,8	367,3	367,3	5638,3	2405,7	5059,9			1740,1
2	4	1883,8	112511,5	16464,4	354,0	119,5	5086,4	2034,6	4916,2			1883,8
2	5	882,3	119369,0	14798,0	345,5	106,8	5884,7	1310,6	5917,7			882,3
2	6	957,7	134602,0	18876,8	347,8	101,2	6357,6	1500,7	5842,3			957,7
2	7	944,1	144362,0	16134,4	336,4	105,4	6326,1	1596,0	5855,9			944,1
2	8	1026,7	166349,0	20601,1	348,8	115,9	6856,5	2230,0	5773,3			1026,7
2	9	11,6	181651,0	23230,0	362,3	129,3	7602,7	2949,7	6788,4	6788,4	6800	11,6
3	1	613,5	30215,1	5021,6	264,9	78,5	2997,4	510,9	10086,5			613,5
3	2	695,5	35787,6	5454,3	230,8	72,0	3076,1	555,0	10004,5			695,5
3	3	819,0	34488,4	5862,5	216,8	64,9	2297,5	586,3	9881,0			819,0
3	4	839,4	36892,1	5414,4	213,6	59,9	2231,3	798,7	9860,6			839,4

⁵¹ Wornalkiewicz W., Rozwiązywanie modelu CCR programem WinQSB, artykuł, WSZiA Opole.

⁵² Wornalkiewicz W., Modele i efektywność PKB, artykuł, WSZiA Opole.

3	5	95,5	40117,4	5639,7	213,1	58,0	2345,9	812,7	10604,5			95,5
3	6	620,1	43428,0	5613,7	215,2	56,0	3049,8	679,8	10079,9			620,1
3	7	1215,6	47063,0	6736,4	209,1	59,4	2765,4	874,9	9484,4			1215,6
3	8	2023,3	51109,0	8115,3	214,8	64,4	3810,8	1399,9	8676,7			2023,3
3	9	1673,1	58306,0	10613,0	225,1	71,9	4919,6	1822,1	9026,9	10604,5	10700	1673,1
4	1	61,7	75024,3	9604,5	600,0	141,3	7128,0	761,4	49038,3			61,7
4	2	2211,3	83990,1	9286,4	537,5	122,6	5907,2	966,6	46888,7			2211,3
4	3	3225,8	85813,6	9008,7	502,5	114,7	5393,9	817,3	45874,2			3225,8
4	4	6384,4	89082,5	8080,7	491,9	103,6	6123,1	629,4	42715,6			6384,4
4	5	6660,1	100527,8	8683,0	481,9	90,4	6482,3	630,8	42439,9			6660,1
4	6	5640,8	129781,0	9417,0	479,2	85,8	6130,4	847,1	43459,2			5640,8
4	7	6827,0	126142,0	10823,4	477,8	69,7	7025,7	1377,3	42273,0			6827,0
4	8	7421,9	140493,0	12251,8	489,0	97,6	9516,4	1566,5	41678,1			7421,9
4	9	7691,8	158013,0	16014,0	503,6	106,7	11999,2	2168,2	41408,2	49038,3	49100	7691,8
5	1	1069,6	12250,0	1990,2	127,2	36,1	1194,9	211,1	4130,4			1069,6
5	2	957,4	14468,7	2215,6	118,5	30,3	1282,5	278,3	4242,6			957,4
5	3	1529,6	13925,2	2020,2	110,3	28,8	1249,2	257,3	3670,4			1529,6
5	4	1355,7	14093,6	2190,5	108,2	26,5	1055,5	270,8	3844,3			1355,7
5	5	1067,6	15331,3	2070,0	108,6	24,6	1146,4	396,5	4132,4			1067,6
5	6	1127,8	17221,0	2595,3	108,1	23,5	1176,5	529,6	4072,2			1127,8
5	7	44,6	17492,0	2383,9	106,2	24,5	1299,8	386,9	5155,4			44,6
5	8	315,7	18043,0	3535,9	107,3	27,5	1347,5	708,2	4884,3			315,7
5	9	77,5	24811,0	3916,0	109,7	31,2	1517,8	819,6	5122,5	5155,4	5200	77,5
6	1	76,3	16088,4	1986,3	177,2	40,1	1794,5	173,9	1323,7			76,3
6	2	50,8	19274,3	2209,9	159,8	35,0	1567,0	184,6	1349,2			50,8
6	3	329,1	20043,3	2293,2	152,9	32,4	1382,1	184,3	1070,9			329,1
6	4	331,7	19733,3	1920,3	150,4	28,5	1537,2	168,1	1068,3			331,7
6	5	280,1	21411,4	2398,9	152,6	25,5	1700,3	318,2	1119,9			280,1
6	6	230,7	24175,0	2285,1	154,1	25,3	1716,4	418,0	1169,3			230,7
6	7	163,7	23938,0	2360,5	151,5	25,9	1986,2	419,8	1236,3			163,7
6	8	63,1	27443,0	2822,9	155,7	29,9	2026,1	714,7	1336,9			63,1
6	9	123,9	30441,0	3023,0	164,8	32,9	2607,8	597,5	1276,1	1349,2	1400	123,9
7	1	400,9	7463,3	1283,0	73,4	18,8	591,5	174,7	699,1			400,9
7	2	382,2	8661,1	1559,5	65,9	16,8	581,9	200,8	717,8			382,2
7	3	355,4	9242,8	1274,0	62,1	15,5	558,4	181,5	744,6			355,4
7	4	292,9	9527,7	1283,9	60,5	14,4	587,1	153,3	807,1			292,9
7	5	172,5	10275,6	2352,8	62,0	13,8	700,8	173,3	927,5			172,5
7	6	212,6	11754,0	1485,7	61,1	12,9	796,0	209,5	887,4			212,6
7	7	172,1	11739,0	1656,9	58,6	13,2	980,1	245,0	927,9			172,1
7	8	142,6	12593,0	1831,0	61,1	14,8	1222,1	278,9	957,4			142,6
7	9	42,8	15264,0	2470,0	66,4	16,8	1493,5	234,7	1057,2	1057,2	1100	42,8
8	1	295,2	9651,8	1822,1	96,8	28,6	1217,0	137,2	1704,8			295,2
8	2	86,1	11347,2	1955,7	87,3	23,8	1163,6	195,3	1913,9			86,1
8	3	289,4	11295,4	1520,3	81,3	22,6	956,5	169,3	1710,6			289,4
8	4	614,0	11183,5	1410,6	78,4	20,6	932,3	156,8	1386,0			614,0
8	5	596,2	13215,8	1580,0	77,4	19,6	801,0	184,8	1403,8			596,2
8	6	239,0	15008,0	1582,8	77,9	17,6	1068,0	204,0	1761,0			239,0
8	7	401,4	15407,0	2134,6	78,6	18,1	1222,1	316,7	1598,6			401,4
8	8	420,4	17589,0	2081,8	83,1	19,3	1190,0	375,0	1579,6			420,4
8	9	362,3	21211,0	2900,0	89,0	21,8	1652,2	439,8	1637,7	1913,9	2000	362,3
9	1	7,6	9691,9	1445,0	87,0	21,0	1391,4	178,8	902,4			7,6
9	2	45,5	11092,2	1584,9	79,6	17,2	893,5	207,6	864,5			45,5

9	3	107,6	10887,1	1724,8	75,8	15,5	1064,0	209,2	802,4			107,6
9	4	125,0	11124,1	1291,9	71,6	13,4	909,5	235,4	785,0			125,0
9	5	360,9	12147,8	1478,1	72,8	11,6	1205,5	214,7	549,1			360,9
9	6	286,9	17221,0	2595,3	76,7	11,0	1065,4	225,7	623,1			286,9
9	7	394,1	14507,0	1843,9	78,2	12,9	1154,3	597,1	515,9			394,1
9	8	201,4	15897,0	1952,0	84,1	14,3	1216,3	581,4	708,6			201,4
9	9	183,0	18044,0	2481,0	87,1	16,7	1863,3	625,7	727,0	902,4	910	183,0
10	1	1592,0	45964,2	6035,3	351,5	85,8	3946,2	629,6	3608,0			1592,0
10	2	1549,2	50742,7	6767,6	328,3	77,0	4036,7	1032,9	3650,8			1549,2
10	3	1494,8	52747,7	6746,1	315,3	69,3	4081,2	876,2	3705,2			1494,8
10	4	1176,1	54771,2	6998,1	311,5	66,5	3539,1	1310,4	4023,9			1176,1
10	5	1468,2	63293,7	8218,2	320,1	63,3	4388,8	1973,7	3731,8			1468,2
10	6	895,9	75782,0	6886,7	338,6	59,9	5212,2	1405,3	4304,1			895,9
10	7	3,7	78868,0	7819,5	334,6	64,0	4943,5	1239,1	5196,3			3,7
10	8	836,1	89218,0	9921,3	339,9	70,7	5158,0	1465,7	4363,9			836,1
10	9	348,4	95799,0	11654,0	357,1	81,4	6271,9	1336,9	4851,6	5196,3	5200	348,4
11	1	1631,7	15872,1	3117,5	134,2	38,3	1428,7	265,1	5368,3			1631,7
11	2	1701,4	17956,1	3735,9	121,6	36,0	1578,3	413,0	5298,6			1701,4
11	3	2008,5	17345,3	3306,9	112,5	32,7	1029,6	478,5	4991,5			2008,5
11	4	2680,7	16033,8	2589,9	105,8	29,5	1049,0	354,9	4319,3			2680,7
11	5	1723,1	18471,1	3128,2	103,9	25,5	1363,3	306,5	5276,9			1723,1
11	6	1818,2	20114,0	2955,6	107,4	24,3	1900,3	275,7	5181,8			1818,2
11	7	1574,0	21293,0	3377,3	106,6	26,7	1635,8	350,2	5426,0			1574,0
11	8	25,3	22202,0	4619,5	110,1	30,5	2481,1	565,1	6974,7			25,3
11	9	159,0	25620,0	5404,0	114,3	34,6	2533,5	721,9	6841,0	6974,7	7000	159,0
12	1	2669,9	28904,0	6341,6	270,3	74,0	4426,1	655,3	32530,1			2669,9
12	2	2836,5	36304,1	6082,6	246,1	66,7	3528,2	874,4	32363,5			2836,5
12	3	1569,2	36428,5	6068,9	231,0	57,5	4393,8	616,7	33630,8			1569,2
12	4	3225,6	37410,5	5199,0	230,1	51,0	4084,8	631,4	31974,4			3225,6
12	5	2378,1	43439,3	4588,3	229,3	47,4	3944,8	897,8	32821,9			2378,1
12	6	493,5	50751,0	5690,4	235,8	45,0	4582,6	1031,1	34706,5			493,5
12	7	317,9	55070,0	6664,8	234,8	47,9	4922,8	1239,3	34882,1			317,9
12	8	19,3	66854,0	6697,9	244,5	52,3	6353,3	1844,5	35180,7			19,3
12	9	900,3	78596,0	10271,0	261,2	59,8	7063,2	2139,8	34299,7	35180,7	35200	900,3

13	1	461,0	11043,9	1616,7	92,3	25,9	1050,9	549,9	2139,0			461,0
13	2	89,9	12892,5	1845,1	83,0	21,9	719,2	517,1	2510,1			89,9
13	3	271,7	12947,4	1249,5	74,7	19,4	682,9	328,1	2328,3			271,7
13	4	667,7	12718,9	1984,9	72,1	18,2	749,3	93,7	1932,3			667,7
13	5	775,1	13956,2	1216,0	72,3	15,3	780,6	76,6	1824,9			775,1
13	6	783,4	18087,0	1159,0	72,2	15,6	950,6	114,2	1816,6			783,4
13	7	957,0	17147,0	1744,3	72,4	16,0	1197,5	174,1	1643,0			957,0
13	8	753,2	18335,0	1791,5	73,7	17,8	1117,5	237,4	1846,8			753,2
13	9	1063,5	20275,0	2571,0	77,7	20,2	1638,1	282,9	1536,5	2510,1	2600	1063,5
14	1	145,4	22720,0	2924,9	187,8	42,7	1839,0	315,1	3054,6			145,4
14	2	103,1	24828,7	2887,6	177,7	40,4	1853,1	316,8	3096,9			103,1
14	3	450,5	24947,8	3266,0	164,8	35,2	2112,3	208,9	2749,5			450,5
14	4	413,9	25124,1	2920,7	160,1	32,3	2030,7	258,6	2786,1			413,9
14	5	904,9	28812,7	2539,7	157,0	29,3	1635,3	310,2	2295,1			904,9
14	6	649,7	32059,0	3037,9	161,3	28,2	1938,6	256,3	2550,3			649,7
14	7	49,4	31717,0	3411,0	158,9	29,6	2287,7	477,0	3150,6			49,4
14	8	1111,4	33013,0	3454,2	165,3	34,1	2420,9	770,0	2088,6			1111,4

14	9	615,6	39116,0	5493,0	173,6	39,0	3048,2	916,1	2584,4	3150,6	3200	615,6
15	1	193,6	24179,4	3938,7	187,0	56,3	2336,8	394,1	2406,4			193,6
15	2	623,9	29953,6	4696,9	174,0	48,9	2113,9	615,3	1976,1			623,9
15	3	723,2	30298,1	5222,3	168,2	44,6	1643,1	658,4	1876,8			723,2
15	4	610,0	31282,2	4660,1	163,0	39,6	1515,9	556,2	1990,0			610,0
15	5	493,7	33779,0	4139,9	160,7	36,2	1527,6	399,3	2106,3			493,7
15	6	502,9	39676,0	4162,8	162,9	33,3	1800,9	493,6	2097,1			502,9
15	7	402,1	41424,0	4385,4	164,3	36,2	1991,7	636,7	2197,9			402,1
15	8	60,1	48903,0	5369,7	172,0	41,7	2212,6	897,2	2539,9			60,1
15	9	133,1	53048,0	7576,0	179,3	48,4	2719,1	2570,1	2466,9	2539,9	2600	133,1
16	1	49600,8	10901,9	1822,1	107,5	28,2	880,7	137,8	399,2			500,8
16	2	49540,4	12561,6	1516,3	95,3	23,2	848,5	174,5	459,6			440,4
16	3	49549,8	13106,8	1691,4	91,5	21,1	712,0	149,9	450,2			449,8
16	4	49556,7	13573,9	1612,9	89,4	18,4	874,8	162,5	443,3			456,7
16	5	49481,9	14906,7	1775,4	91,4	16,7	980,2	220,0	518,1			381,9
16	6	49405,6	15809,0	1932,7	94,4	15,9	1001,8	299,9	594,4			305,6
16	7	49390,7	15540,0	2786,1	97,1	17,8	1554,6	324,2	609,3			290,7
16	8	49246,8	16701,0	2432,5	100,9	19,9	1909,7	510,5	753,2			146,8
16	9	49105,2	20330,0	4003,0	105,4	24,1	1935,7	833,4	894,8	894,8	900	5,2

Źródło: Opracowanie własne w *Excelu* na podstawie roczników statystycznych województw (2000-2008).

Czcionką pogrubioną zaznaczono nazwy szeregów danych biorące udział w obliczaniu efektywności technologicznej modelem CCR. Szeregi te kopiujemy, a następnie wklejamy z transpozycją na nowy arkusz kalkulacyjny *Excela*. Nastąpiła w ten sposób zamiana i tak w kolumnach występują obiekty (województwa), a w ramach nich okresy czasowe. Natomiast wiersze stanowią dane najpierw rezultaty (*C-X5*, *X18*, *X19*) a w kolejności nakłady (*X9*, *X10*, *X20*, *X21*).

Program *WinQSB* w wersji posiadanej przez autora wymagał zamiany przecinka na kropkę dziesiętną. Realizujemy to w *Excelu* wywołując menu: *Edycja/zamień ...* W wyniku zmiany formatowania (z liczbowego na tekstowe) *Excel* interpretuje jednak liczby całkowite jako przykładowo 55070,0. Wartości danych liczbowych wcześniej sprowadzonych do jednego znaku po przecinku zapisuje w postaci przed sformatowaniem np. 1031.081. Aby korzystać z modułu programowania liniowego i całkowitoliczbowego to niedoskonałości musimy zatem wcześniej poprawić manualnie. Efekt pracy w odniesieniu do przykładowego województwa łódzkiego podaje tabela 7.2.

Tab. 7.2. Fragment danych dla województwa łódzkiego (1) oraz 9 okresów czasowych (1999-2007)

W	1	1	1	1	1	1	1	1	1
T	1	2	3	4	5	6	7	8	9
C-X5	728.4	92.2	636.3	858.2	1018.4	1015.7	1450.5	1779.3	2004.8
X18	26540.7	30492.2	32060.4	32845.0	33607.4	35742.0	37002.0	39976.0	44450.0
X19	4337.5	3739.4	2988.4	2875.6	2936.4	3337.6	4301.2	4669.6	6650.0
X9	264.4	240.9	237.5	227.3	223.5	226.3	223.5	230.0	241.2
X10	47.6	43.7	37.7	34.7	32.8	35.5	37.3	41.0	46.7
X20	2434.6	2670.4	2687.7	2290.1	2482.6	3189.8	3901.763	3850.0	5479.1
X21	273.5	270.3	277.6	321.3	357.3	581.28	1072.3	1070.0	1148.4

Źródło: Opracowanie własne w *Excelu* na podstawie danych tabeli 7.1.

Zastosowanie modelu CCR wymaga jeszcze przesortowania danych w kolumnach najpierw rosnąco według okresów czasowych *T*, a następnie rosnąco obiektami *W*. W tym względzie dokonujemy ponownej transpozycji tabeli danych po czym w *Excelu* korzystamy z menu: *Dane/sortuj ...*

Tab. 7.3. Fragment posortowanego zbioru danych dla okresu $T = 1$ i województw $W = 1-16$

T	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
W	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
C-	728.4	3107.	613.5	61.7	1069.	76.3	400.9	295.2	7.6	1592.	1631.	2669.	461.0	145.4	193.6	49600
X5		7			6					0	7	9				.8
X1	26540	85240	30215	75024	12250	16088	7463.	9651.	9691.	45964	15872	28904	11043	22720	24179	10901
8	.7	.9	.1	.3	.0	.4	3	8	9	.2	.1	.0	.9	.0	.4	.9
X1	4337.	15421	5021.	9604.	1990.	1986.	1283.	1822.	1445.	6035.	3117.	6341.	1616.	2924.	3938.	1822.
9	5	.2	6	5	2	3	0	1	0	3	5	6	7	9	7	1
X9	264.4	407.4	264.9	600.0	127.2	177.2	73.4	96.8	87.0	351.5	134.2	270.3	92.3	187.8	187.0	107.5
X1	47.6	151.6	78.5	141.3	36.1	40.2	18.8	28.6	21.0	85.8	38.3	74.0	25.9	42.7	56.3	28.2
0																
X2	2434.	6779.	2997.	7128.	1194.	1794.	591.5	1217.	1391.	3946.	1428.	4426.	1050.	1839.	2336.	880.7
0	6	5	4	0	9	5		0	4	2	7	1	9	0	8	
X2	273.5	3216.	510.9	761.4	211.1	173.9	174.7	137.2	178.8	629.6	265.1	655.3	549.9	315.1	394.1	137.8
1		8														

Uwaga: Ze względu na ograniczoność formatu A4 dłuższe wartości danych zamieszczono w dwóch wierszach. Linia przerywana rozgranicza dane dotyczące rezultatów od nakładów.

Źródło: Opracowanie własne w *Excelu* według pełnej tabeli 2.2 i po ponownej jej transpozycji.

7.2. Rozwiązywanie optymalne zadania CCR ukierunkowanego na nakłady

Rozwiązywanie zadania CCR ukierunkowanego na nakłady przeprowadzono korzystając z modułu LPILP programu *WinQSB*. Procedura postępowania w zakresie formułowania poszczególnych cząstkowych zadań decyzyjnych oraz realizacji w programie *WinQSB* została wcześniej przedstawiona w opracowaniu *Rozwiązywanie modelu CCR programem WinQSB*⁵³. Warto tu jednak przypomnieć sobie model zadania decyzyjnego i wykonać odniesienie do rozszerzonych danych.

Sformułowanie zadania decyzyjnego w *WinQSB* oraz wartości danych dla roku 2006 i poszczególnych obiektów (województw) pokazano na rysunku 2.1. Rok 2006 przyjęto na początek celowo, aby nawiązać do rozwiązania cząstkowego we wcześniejszej pracy⁵⁴. Forma zapisu algebraicznego ukierunkowanego na nakłady standardowego zadania modelu CCR, którą zaczerpnięto z monografii Guzika⁵⁵ jest następująca:

- dane źródłowe - wielkości nakładów oraz rezultatów w poszczególnych obiektach:

$$x_{rj} \text{ oraz } y_{rj} \quad (j = 1, \dots, J; r = 1, \dots, R; n = 1, \dots, N)$$

w naszym przykładzie mamy:

$J = 16$ obiektów decyzyjnych (województw),

$R = 3$ rezultaty ($C-X5$, $X18$, $X19$); dane $X5$ w formie " $C-X5$ " pełnią rolę destymulanty,

$N = 4$ nakłady ($X9$, $X10$, $X20$, $X21$).

- zmienne decyzyjne; w modelu CCR w odróżnieniu od modelu ekonometrycznego zmiennymi są parametry oszacowywane przez algorytm simpleks modelu CCR, a należą do nich:

$\lambda_{o1}, \lambda_{o2}, \dots, \lambda_{oJ}$	Wagi intensywności w technologii wspólnej zorientowanej na obiekt o -ty
θ_o	Mnożnik nakładów obiektu o -tego

Ze względu na możliwości zapisu znaków w module LPILP programu *WinQSB* zmienne decyzyjne oznaczono jako LO-1 do LO16, przy czym litera L odpowiada greckiej λ (lambda), a w miejsce

⁵³ Wornalkiewicz W., *Rozwiązywanie modelu CCR programem WinQSB*, WSZiA Opole.

⁵⁴ Ibidem, s. 5.

⁵⁵ Guzik B., *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, 2009, rozdz. 3.3.

literę "O" wpisujemy numer obiektu, przykładowo 12 dla którego obecnie szacowane są parametry modelu CCR. Mnożnik nakładów obiektu o -tego θ_o oznaczono jako T-O (patrz rysunek 2.1).

Funkcja celu - minimalizacja mnożnika poziomu nakładów:

$$\theta_o = \sum_{j=1}^J w_{oj} \lambda_{oj} + w_0 \theta_o = \theta_o \rightarrow \min$$

gdzie:

$w_{oj} = 0$, natomiast $w_0 = 1$.

Warunki ograniczające: nakłady technologii wspólne są nie większe od możliwie najmniejszej części nakładów poniesionych przez obiekt o -ty:

$$\sum_{j=1}^J x_{nj} \lambda_{oj} \leq \theta_o x_{no} \quad (dla \ n = 1, \dots, N)$$

Jest to warunek ograniczający w tzw. zapisie normalnym. Ze względu na kopiowanie macierzy danych z *Excela* zastosowano format matrix. Wymagało to przeniesienia $\theta_o x_{no}$ ze znakiem ujemnym na lewą stronę równania ograniczeń. W przypadku ukierunkowania CCR na nakłady dotyczy to ograniczeń związanych z nakładami ($Cn - Xn$), czyli *C4-X9*, *C5-X10*, *C6-20*, *C7-21*), przy czym Cn to skrót od słowa "constraint" z dopisanym numerem kolejnym ograniczenia.

- rezultaty technologii wspólnej są nie mniejsze od rezultatów osiągniętych przez obiekt o -ty:

$$\sum_{j=1}^J y_{rj} \lambda_{oj} \geq y_{ro} \quad (dla \ n = 1, \dots, R)$$

Na rysunku 2.1 występują wartości dla ograniczeń: C1-C-X5, C2-X18 oraz C3-X19.

- nakłady technologii wspólnej nie przekraczają nakładów obiektu o -tego: $\theta_o \leq 1$.

Warunek ten w odniesieniu do obiektu 12 (województwo dolnośląskie) wprowadzono dając mu symbol C8-T-12 (patrz rysunek 2.1).

- warunki znakowe, tzw. ograniczenia brzegowe: $\theta_o; \lambda_{o1}, \lambda_{o1}, \dots, \lambda_{oJ} \geq 0$

Typ danych w przykładzie zilustrowanym na rysunku 7.1 jest "continuous", czyli ciągły z przecinkiem zapisywanym jak już wiemy w *WinQSB* w postaci kropki dziesiętnej. Rozwiązanie zadania podano w tabeli 7.4.

Tab. 7.4. Rozwiązanie sumaryczne zadania cząstkowego modelu CCR ukierunkowanego na nakłady dla obiektu 12.

2010-08-25 09:42:42	Decision Variable	Solution Value	Unit Cost or Profit C(j)	Total Contribution	Reduced Cost	Basis Status
1	L12-1	0	0	0	0,2517	at bound
2	L12-2	0,1424	0	0	0	basic
3	L12-3	0	0	0	0,5486	at bound
4	L12-4	0,3072	0	0	0	basic
5	L12-5	0	0	0	0,2841	at bound
6	L12-6	0	0	0	0,2067	at bound
7	L12-7	0	0	0	0,1149	at bound
8	L12-8	0	0	0	0,1348	at bound
9	L12-9	0	0	0	0,0617	at bound
10	L12-10	0	0	0	0,1645	at bound
11	L12-11	0	0	0	0,2854	at bound
12	L12-12	0	0	0	0,1111	at bound
13	L12-13	0	0	0	0,0966	at bound
14	L12-14	0	0	0	0,2129	at bound
15	L12-15	0	0	0	0,1466	at bound
16	L12-16	0	0	0	0,1576	at bound

17	T-12	0,8889	1, 000	0,8889	0	basic
	Objective Function	(Min.) =		0,8889		

Źródło: Opracowanie własne modulem LPILP programu WinQSB po edycji w programie Excel.

W rozwiązaniu zadania skorzystano z menu modułu "Programowanie liniowe i całkowitoliczbowe LPILP": *Solve and Analyze/solve the Problem*. Uzyskano raport kombinowany, który można zastąpić rozwiązaniem sumarycznym wywołując menu: *File/Save as*. Rozwiązanie zapisano z rozszerzeniem "dat" co umożliwia jego edycję pod *Excelem*.

Rys. 7.1. Ekranu modułu LPILP sformułowania zadania decyzyjnego rozwiązywanego modelem CCR ukierunkowanym na nakłady dla obiektu 12.

Dla naszego obiektu optymalny mnożnik $\hat{\theta}_{12} = 0,8889$ stanowiący jednocześnie wskaźnik efektywności jest mniejszy od 1. Zwróćmy uwagę, że wartości optymalne zaznaczamy ozdobnikiem w postaci "daszku". Według modelu CCR ukierunkowanego na nakłady obiekt 12. nie jest zatem w pełni efektywny, a jego stopień nieefektywności wynosi:

$$1 - \hat{\theta}_o = 1 - 0,8889 = 0,1111 \text{ czyli } \approx 11\%$$

Zaprezentowano tu szczegółowo przykład zadania cząstkowego odniesionego do roku 2006 oraz obiektu 12. Ponieważ interesują nas również wskaźniki efektywności technologicznej w ramach tego samego roku dla pozostałych 15. obiektów (województw) to musimy ponownie 15-krotnie wywołać menu rozwiązania. W tym celu należy uprzednio każdorazowo zastąpić dane w kolumnach T-12 oraz R.H.S. na wielkości odpowiadające danemu obiektowi.

Rozwiązanie modelu CCR ukierunkowanego na nakłady dla roku 2006 i wszystkich 16 obiektów zamieszczono w tabeli 7.5. Wyszczególniono efektywności $\hat{\theta}_o = TO$ oraz optymalne współczynniki benchmarkingowe $\hat{\lambda}_{of} = LO - j$ stanowiące odniesienia do konkurentów (pozostałych obiektów).

Tab. 7.5. Rozwiązanie optymalne zadania decyzyjnego według modelu CCR ukierunkowanego na nakłady dla roku 2006

Obiekt <i>o</i>	$\hat{\theta}_o$	$\hat{\lambda}_{oj}$															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0,7995		0,0916		0,2271												
2	1		1														
3	1			1													
4	1				1												
5	0,9443		0,1377	0,0862													
6	0,6391		0,1279		0,0439												
7	0,7171		0,0844		0,0075												
8	0,7854		0,0527	0,0451	0,0046									0,3199			
9	0,7986		0,0857		0,0153												
10	0,8783		0,3513		0,2191												
11	0,8849		0,2242														
12	0,8889		0,1424		0,3072												
13	1													1			
14	0,7809		0,0628	0,0206	0,0737									0,6082			
15	0,9110		0,2940														
16	0,6931		0,1158		0,0038												

Gdzie: Współczynniki benchmarkingowe $\hat{\lambda}_{oj}$ dotyczą *o*-tego obiektu i odniesienia do *j*-tej technologii konkurentów (innych obiektów)

Źródło: Opracowanie własne modulem LPILP programu *WinQSB*; pola puste dla efektów i współczynników mają wartości zero

Zgodnie z modelem CCR ukierunkowanego na nakłady efektywne w sensie Farrela oraz z punktu widzenia przyjętych danych są województwa⁵⁶: 2 - mazowieckie, 3 - małopolskie, 4 - śląskie, 13 - opolskie. Najmniej efektywne biorąc pod uwagę wartość optymalną $\hat{\theta}_o$ którą traktujemy jako wskaźnik efektywności technologicznej *o*-tego obiektu jest województwo o numerze sześć - podkarpackie.

Aby przyspieszyć prace, rozwiązaniem alternatywnym będzie sformułowanie zadania integralnego i uzyskanie w jednym przebiegu przetwarzania wszystkich wyników jednocześnie⁵⁷.

W sposób analogiczny jak dla roku 2006 możemy postąpić z kolejnymi latami okresu czasowego 1999-2005 oraz 2007. Ze względu na ramy niniejszego materiału zaprezentowano jedynie makietę do zapisu efektu końcowego prac w postaci zestawienia procentowych wskaźników efektywności technologicznej (patrz tabela 7.6).

Tab. 7.6. Procentowe wskaźniki efektywności technologicznej województw w latach 1999-2007

Nr W	T/W	1999	2000	2001	2002	2003	2004	2005	2006	2007
1	łódzkie									
2	mazowiecki									
3	małopolskie									
4	śląskie									
5	lubelskie									
6	podkarpackie									
7	podlaskie									
8	świętokrzyskie									
9	lubuskie									
10	wielkopolskie									

⁵⁶ Guzik B., *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*, op. cit.

⁵⁷ Wornalkiewicz W., *Rozwiązywanie modelu CCR ...*, op. cit.

11	zachodniopomorskie														
12	dołnośląskie												88,89		
13	opolskie														
14	kujawsko-pomorskie														
15	pomorskie														
16	warmińsko-mazurskie														

Gdzie: W - obiekt (województwo), T - okres czasowy (rok)

Źródło: Opracowanie własne.

7.3. Luzy nakładów i rezultatów

Aby skorygować efektywność technologiczną początkową obliczoną modelem CCR ukierunkowanym na nakłady poprzez wprowadzenie kar zachodzi potrzeba określenia tzw. luzów nakładów i rezultatów. Brana po uwagę jest efektywność ustalona według koncepcji Farrela, którą nazywa się θ -efektywność⁵⁸. Jak wiemy już w formułowaniu zadania decyzyjnego ograniczenie dla rezultatów y oraz nakładów x zapisujemy jako:

$$\sum_{j=1}^J y_{rj} \lambda_{oj} \geq y_{ro} \quad (dla n = 1, \dots, R)$$

$$\sum_{j=1}^J x_{nj} \lambda_{oj} \leq \theta_o x_{no} \quad (dla n = 1, \dots, N)$$

Po rozwiązaniu optymalnym zadania według modelu CCR ukierunkowanym na nakłady, oznaczonym " $\wedge$ ", a dotyczącym technologii wspólnej powyższe warunki przyjmą postać:

$$\hat{x}_{no} = \sum_{j=1}^J x_{nj} \hat{\lambda}_{oj} \leq \hat{\theta}_o x_{no} \quad \hat{y}_{ro} = \sum_{j=1}^J y_{rj} \hat{\lambda}_{oj} \geq y_{ro}$$

$$S_{no} = x_{no}^* - \hat{x}_{no} = \hat{\theta}_o x_{no} - \hat{x}_{no} \quad (1 \leq n \leq N; 1 \leq o \leq J)$$

Przyjmijmy określenie deficytu (niedoboru) rezultatu D_{ro} oraz nadwyżki nakładów S_{no} obiektu o -tego:

Alternatywnym rozwiązaniem jest przyjęcie jako nadwyżki nakładów:

$$D_{ro} = \hat{y}_{ro} - y_{ro} \quad (1 \leq r \leq R; 1 \leq o \leq J)$$

$$S_{no}^{emp} = x_{no} - \hat{x}_{no}$$

Kontynuując nasz temat, obliczenie luzów przedstawimy na przykładzie roku 2006 obiektu 12. (województwo dołnośląskie). Najpierw jednak musimy przygotować tabele pomocnicze. Fragment danych empirycznych zaczerpnięty z tabeli 7.1 po sortowaniu rosnąco według okresów, a następnie województw podano w tabeli 7.7.

Tab. 7.7. Dane okresu $t = 8$. (rok 2006) w zakresie potrzeb modelu CCR ukierunkowanego na nakłady

t	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
w	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C-1779.	1026.7	2023.	7421.9	315.7	63.1	142.6	420.4	201.4	836.1	25.3	19.3	753.2	1111.	60.1	146.8
X53		3											4		
X139976	16634	51109	14049	18043	27443	12593	17589	15897	89218	22202	66854	18335	33013	48903	16701
8.0	9.0	.0	3.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
X14669.	20601.	8115.	12251.	3535.	2822.	1831.	2081.	1952.	9921.	4619.	6697.	1791.	3454.	5369.	2432.
9.6	1	3	8	9	9	0	8	0	3	5	9	5	2	7	5
X9230.0	348.8	214.8	489.0	107.3	155.7	61.1	83.1	84.1	339.9	110.1	244.5	73.7	165.3	172.0	100.9
X141.0	115.9	64.4	97.6	27.5	29.9	14.8	19.3	14.3	70.7	30.5	52.3	17.8	34.1	41.7	19.9
0															

⁵⁸ Guzik B., *Podstawowe modele DEA ...*, op. cit., rozdz. 14.

X2	3850.	6856.5	3810.	9516.4	1347.	2026.	1222.	1190.	1216.	5158.	2481.	6353.	1117.	2420.	2212.	1909.
0	0		8		5	1	1	0	3	0	1	3	5	9	6	7
X2	1070.	2230.0	1399.	1566.5	708.2	714.7	278.9	375.0	581.4	1465.	565.1	1844.	237.4	770.0	897.2	510.5
1	0		9							7		5				

Źródło: Opracowanie własne w Excelu.

Tab. 7.8. Fragment rozwiązanie optymalnego zadania decyzyjnego według modelu CCR ukierunkowanego na nakłady dla roku 2006

Obiekt o	$\hat{\theta}_o$	$\hat{\lambda}_{oj}$			
		2	3	4	13
1	0,7995	0,0916		0,2271	
2	1	1			
3	1		1		
4	1			1	
5	0,9443	0,1377	0,0862		
6	0,6391	0,1279		0,0439	
7	0,7171	0,0844		0,0075	
8	0,7854	0,0527	0,0451	0,0046	0,3199
9	0,7986	0,0857		0,0153	
10	0,8783	0,3513		0,2191	
11	0,8849	0,2242			
12	0,8889	0,1424		0,3072	
13	1				1
14	0,7809	0,0628	0,0206	0,0737	0,6082
15	0,9110	0,2940			
16	0,6931	0,1158		0,0038	

Źródło: Opracowanie własne w Excelu; pominięto wartości zerowe występujące w tabeli 7.5.

Widzimy, że w rozwiązaniu optymalnym technologię wspólną stanowią odniesienia współczynników benchmarkingowych $\hat{\lambda}_{oj}$ do obiektów 2, 3, 4 oraz 13, czyli do obiektów uznanych jako efektywne. Obliczenie luzów rezultatów oraz nakładów może odbywać się:

- manualnie poprzez zastosowanie wzorów na: D_{ro} , S_{no} ,
- poprzez zastosowanie arkusza kalkulacyjnego Excel,
- przez skorzystanie z danych w raporcie kombinowanym generowanym przez moduł LPILP programu WinQSB.

Manualne obliczenia w odniesieniu do 16. obiektów są bardzo pracochłonne. Obliczenia w Excelu przy korzystaniu z tabeli 7.7 oraz 7.8 są mimo wszystko czasochłonne. Podejście w tym zakresie zostanie omówione na podstawie rysunku 7.2.

Rys. 7.2. Rozplanowanie arkusza kalkulacyjnego do obliczeń luzów na podstawie danych rezultatów i nakładów oraz wyników θ oraz λ rozwiązania optymalnego modelu CCR ukierunkowanego na nakłady

Dokonano obliczeń nakładów dla obiektów 1, 2, 3, 4. Dalsze obliczenia jako ćwiczenia pozostawia się Czytelnikowi. Już z wstępnej obserwacji możemy zauważyć:

- występowanie wartości zerowych dla obiektów efektywnych,
- występowanie małych wartości oraz niedopuszczalnych wartości ujemnych ze względu na ograniczenia znakowe zoptymalizowanego zadania decyzyjnego, co wiąże się ze zbyt niską precyzją obliczeń.

W obliczeniu przykładowo wartości 7167,529 dla rezultatu "Produkcja p" zastosowano wyrażenie *Excela*:

$$=(D3*\$D\$12+E3*\$E\$12+F3*\$F\$12+G3*\$G\$12+H3*\$H\$12+I3*\$I\$12+J3*\$J\$12+K3*\$K\$12+L3*\$L\$12+M3*\$M\$12+N3*\$N\$12+O3*\$O\$12+P3*\$P\$12+Q3*\$Q\$12+R3*\$R\$12+S3*\$S\$12)-D3$$

Zapis w nawiasach bazujący na wzorze $\hat{y}_{ro} = \sum_{j=1}^J y_{rj} \hat{\lambda}_{oj} \geq y_{ro}$ to odniesienie do rozwiązania opty-

malnego z użyciem odpowiednich wartości danych oraz współczynników (patrz rysunek 2.2).

Przytoczmy jeszcze przykład wyrażenia na obliczenie wartości 40,883 dla nakładu "Pracujący

$$p"$$
 opartego o wzór $\hat{x}_{no} = \sum_{j=1}^J x_{nj} \hat{\lambda}_{oj} \leq \hat{\theta}_o x_{no}$:

$$=(\$C\$12*D5)-(D5*\$D\$12+E5*\$E\$12+F5*\$F\$12+G5*\$G\$12+H5*\$H\$12+I5*\$I\$12+J5*\$J\$12+K5*\$K\$12+L5*\$L\$12+M5*\$M\$12+N5*\$N\$12+O5*\$O\$12+P5*\$P\$12+Q5*\$Q\$12+R5*\$R\$12+S5*\$S\$12)$$

Wystąpienie wartości bardzo małych oraz ujemnych skłoniło autora do realizacji od początku obliczania luzów bazując na raporcie kombinowanym programu *WinQSB*. Ilustracją tego raportu jest rysunek 7.3 wykonany dla danych roku 2006 oraz obiektu 1 (województwo łódzkie).

Raport obejmuje między innymi:

- rozwiązanie sumaryczne z podaniem θ -efektywności Farrela (tu oznaczonym T-12, aby nie zmieniać nazwy z poprzedniego ćwiczenia odniesionego do obiektu 12-tego) oraz współczynników benchmarkingowych λ zapisanych jako L12-1 - L12-16,

- określenie luzów w odniesieniu do rezultatów oraz nakładów (*Slack or Surplus*).

Porównanie wyników uzyskanych z *Excelsa* oraz z programu *WinQSB* podaje tabela 7.9. Możemy zaobserwować większą dokładność obliczeń i prawidłowe reagowanie modułu LPILP programu *WinQSB*, który przed generowaniem raportu sprawdza warunki ograniczające zadania decyzyjnego. Ponadto *WinQSB* oblicza luz na ograniczeniu $\hat{\theta}$, który wynosi $1 - 0,7995 = 0,2005$.

Decision Variable	Solution Value	Unit Cost or Profit c[i]	Total Contribution	Reduced Cost	Basis Status	Allowable Min. c[i]	Allowable Max. c[i]
1 L12-1	0	0	0	0,2005	at bound	-0,2005	M
2 L12-2	0,0916	0	0	0	basic	-2,4975	0,5413
3 L12-3	0	0	0	0,2910	at bound	-0,2910	M
4 L12-4	0,2271	0	0	0	basic	-0,5993	0,8628
5 L12-5	0	0	0	0,1712	at bound	-0,1712	M
6 L12-6	0	0	0	0,3499	at bound	-0,3499	M
7 L12-7	0	0	0	0,1045	at bound	-0,1045	M
8 L12-8	0	0	0	0,1526	at bound	-0,1526	M
9 L12-9	0	0	0	0,0702	at bound	-0,0702	M
10 L12-10	0	0	0	0,3279	at bound	-0,3279	M
11 L12-11	0	0	0	0,1311	at bound	-0,1311	M
12 L12-12	0	0	0	0,3888	at bound	-0,3888	M
13 L12-13	0	0	0	0,1282	at bound	-0,1282	M
14 L12-14	0	0	0	0,2613	at bound	-0,2613	M
15 L12-15	0	0	0	0,3016	at bound	-0,3016	M
16 L12-16	0	0	0	0,1490	at bound	-0,1490	M
17 T-12	0,7995	1,0000	0,7995	0	basic	0	M
Objective	Function	(Min.) =	0,7995				

Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Shadow Price	Allowable Min. RHS	Allowable Max. RHS
1 C1-C-X5	1 779,3000	>=	1 779,3000	0	0,0001	604,7423	2 496,4660
2 C2-X18	47 143,1600	>=	39 976,8000	7 167,1570	0	M	47 143,1600
3 C3-X19	4 669,6000	>=	4 669,6000	0	0,0001	3 747,2640	6 187,0790
4 C4-X9	-40,9007	<=	0	40,9007	0	-40,9007	M
5 C5-X10	0,0000	<=	0	0	-0,0244	-0,2188	3,0794
6 C6-X20	-289,1615	<=	0	289,1615	0	-289,1615	M
7 C7-X21	-295,4805	<=	0	295,4805	0	-295,4805	M
8 C8-T-12	0,7995	<=	1,0000	0,2005	0	0,7995	M

Rys. 7.3. Raport kombinowany stanowiący rozwiązanie optymalne modelu CCR ukierunkowanego na nakłady odniesionego do obiektu 1.

Tab. 7.9. Luzy rezultatów oraz nakładów dla przykładowego 1-go obiektu uzyskane w dwóch programach

Dana	Nazwa	J.m.	Wg Excelsa	Wg WinQSB
C-X5	Odpady	tys. t	0,259	0
X18	Produkcja p	mln zł	7167,529	7167,1570
X19	Produkcja b	mln zł	-0,155	0
X9	Pracujący p	tys.	40,883	40,9007
X10	Pracujący b	tys.	-0,002	0
X20	Nakłady inv. p	mln zł	288,845	289,1615
X21	Nakłady inv. b	mln zł	295,445	295,4805
C8-T-12				0,2005

Źródło: Opracowanie własne.

W tej sytuacji dalsze prace przeprowadzono korzystając z luzów zapisanych w kolejnych raportach kombinowanych zadań cząstkowych dla obiektów (województw) 1 -16, a wynik końcowy prezentuje tabela 7.10.

Tab. 7.10. Luzy rezultatów R i nakładów N w rozwiązaniu optymalnym modelu CCR ukierunkowanym na nakłady w obiekcie o-tym w roku 2006

y	R	Obiekt o																
x	N	J. m.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
C-X5	Od-pady	tys. t	0	0	0	0	0	394,21 28	0	0	0	1150,6 22	204,92 27	2407,2 66	0	0	241,72 75	0
X18	Produkcyjap	mln zł	7167,1 57	0	0	0	9266,1 15	0	2505,2 18	0	500,26 56	0	15099, 37	0	0	0	0	3096,8 97
X19	Produkcyjab	mln zł	0	0	0	0	0	349,60 87	0	0	0	0	0	0	0	0	685,57 65	0
X9	Pracującyp	tys.	40,900 7	0	0	0	34,788 4	33,418 4	10,693 9	11,36 5	29,809 2	68,879 6	19,211 9	17,432 1	0	21,871	54,151 1	27,691 7
X10	Pracującyb	tys.	0	0	0	0	4,4607	0	0,0965	0	0	0	0,9999	0	0	0	3,9163	0
X20	Nakłady inv. p	mln zł	289,16 15	0	0	0	0	0	226,02 19	0	238,53 96	36,856 3	658,00 32	1747,4 28	0	0	0	493,60 4
X21	Nakłady inv. b	mln zł	295,48 05	0	0	0	241,07 92	102,76 47	0	30,66 9	249,33 93	160,78 22	0	840,79 08	0	172,50 48	161,77 09	89,615 2

Źródło: Opracowanie własne na podstawie raportów kombinowanych programu WinQSB.

W obiekcie ósmym oraz czternastym nie występowały niedobory rezultatów, a występowały jedynie nadwyżki nakładów pracujących w przemyśle oraz nakładów inwestycyjnych w budownictwie. W obiektach 1, 5-7, 9-12, 15,16 występowała rozrzutność niektórych nakładów jak też i rezultatów.

7.4. Dwuetapowa korekta efektywności początkowej

Przystąpmy teraz do skorygowania podanej w tabeli 2.8 θ -efektywności.

Obiekt o	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\hat{\theta}_o$	0,7995	1	1	1	0,9443	0,6391	0,7171	0,7854	0,7986	0,8783	0,8849	0,8889	1	0,7809	0,9110	0,6931

W pierwszym etapie obliczamy kary jednostkowe (zwane również wagami) typu A, według wzorów⁵⁹:

- dla nakładu n -tego:

$$s_{no} = \frac{1}{(N+R)} \frac{1}{x_{no}}$$

- dla rezultatu r -tego:

$$d_{ro} = \frac{1}{(N+R)} \frac{1}{y_{ro}}$$

⁵⁹ Guzik B., *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej* ..., op. cit.

Gdzie:

N - ilość nakładów,

R - ilość rezultatów,

x_{no} - wartość n -tego nakładu w o -tym obiekcie,

y_{ro} - wartość r -tego rezultatu w o -tym obiekcie,

1 - największa wartość współczynnika θ -efektywności według Farrela.

W naszym przykładzie $N = 3$, a $R = 4$. Obliczone na podstawie wartości w tabeli 2.7 kary jednostkowe za luzy nakładów i rezultatów dla województw o numerach 1-16 w roku 2006 podano w tabeli 7.11.

Tab. 7.11. Wartości danych oraz kary (wagi) jednostkowe typu A

Dane	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
C-X5	1779,3	1026,7	2023,3	7421,9	315,7	63,1	142,6	420,4	201,4	836,1	25,3	19,3	753,2	1111,4	60,1	146,8
d_{ro}	0,0000	0,00013	0,00007	0,00001	0,0004	0,0022	0,0010	0,0003	0,0007	0,0001	0,0056	0,0074	0,0001	0,0001	0,0023	0,0009
	80	9	1	9	53	64	02	40	09	71	47	02	90	29	77	73
X18	39976	166349	51109	140493	18043	27443	12593	17589	15897	89218	22202	66854	18335	33013	48903	16701
d_{ro}	0,0000	0,00000	0,00000	0,00000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	04	1	3	1	08	05	11	08	09	02	06	02	08	04	03	09
X19	4669,6	20601,1	8115,3	12251,8	3535,9	2822,9	1831	2081,8	1952	9921,3	4619,5	6697,9	1791,5	3454,2	5369,7	2432,5
d_{ro}	0,0000	0,00000	0,00001	0,00001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	31	7	8	2	40	51	78	69	73	14	31	21	80	41	27	59
X9	230	348,8	214,8	489	107,3	155,7	61,1	83,1	84,1	339,9	110,1	244,5	73,7	165,3	172	100,9
s_{no}	0,0006	0,00041	0,00066	0,00029	0,0013	0,0009	0,0023	0,0017	0,0016	0,0004	0,0012	0,0005	0,0019	0,0008	0,0008	0,0014
	21	0	5	2	31	18	38	19	99	20	98	84	38	64	31	16
X10	41	115,9	64,4	97,6	27,5	29,9	14,8	19,3	14,3	70,7	30,5	52,3	17,8	34,1	41,7	19,9
s_{no}	0,0034	0,00123	0,00221	0,00146	0,0051	0,0047	0,0096	0,0074	0,0099	0,0020	0,0046	0,0027	0,0080	0,0041	0,0034	0,0071
	84	3	8	4	95	78	53	02	90	21	84	31	26	89	26	79
X20	3850	6856,5	3810,8	9516,4	1347,5	2026,1	1222,1	1190	1216,3	5158	2481,1	6353,3	1117,5	2420,9	2212,6	1909,7
s_{no}	0,0000	0,00002	0,00003	0,00001	0,0001	0,0000	0,0001	0,0001	0,0001	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
	37	1	7	5	06	71	17	20	17	28	58	22	28	59	65	75
X21	1070	2230	1399,9	1566,5	708,2	714,7	278,9	375	581,4	1465,7	565,1	1844,5	237,4	770	897,2	510,5
s_{no}	0,0001	0,00006	0,00010	0,00009	0,0002	0,0002	0,0005	0,0003	0,0002	0,0000	0,0002	0,0000	0,0006	0,0001	0,0001	0,0002
	34	4	2	1	02	00	12	81	46	97	53	77	02	86	59	80

Źródło: Opracowanie własne w Excelu; cyfry zapisane w drugim wierszu stanowią kontynuację wartości danej.

Teraz pomnożymy odpowiednie komórki luzów z tabeli 7.10 oraz kar z tabeli 7.11 dla obliczenia dla poszczególnych danych cząstkowych korekt. Następnie obliczymy sumy - korekty efektywności danego obiektu ze względu na nadwyżki nakładów oraz deficyty rezultatów na podstawie wzoru⁶⁰:

$$K_o = K_o^S + K_o^D = \sum_{n=1}^N S_{no} s_{no} + \sum_{r=1}^R D_{ro} d_{ro}$$

Korekta θ -efektywności dla o -tego obiektu stanowi zatem sumę sum iloczynów luzów i kar w odniesieniu do nakładów oraz rezultatów.

Tab. 7.12. Obliczenia pomocnicze korekty efektywności danego obiektu W

W	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
D_{ro}	0	0	0	0	0	394,21	0	0	0	1150,6	204,92	2407,26	0	0	241,72	0
x_{so}						28				22	27	6			75	
d_{ro}	0,000	0,0001	0,0000	0,0000	0,0004	0,0022	0,0010	0,0003	0,0007	0,0001	0,0056	0,00740	0,0001	0,0001	0,0023	0,0009
	080	39	71	19	53	64	02	40	09	71	47	2	90	29	77	73
d'	0,000	0,0000	0,0000	0,0000	0,0000	0,8924	0,0000	0,0000	0,1965	1,1571	17,8184	0,0000	0,0000	0,0000	0,5745	0,0000
	000	00	00	00	00	90	00	00	00	97	02	01	00	00	84	00
D_{X180}	1767,157	0	0	0	9266,15	0	2505,2	0	500,26	0	15099,37	0	0	0	0	3096,8
	157				15		18		56							97
d_{ro}	0,000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,00000	0,0000	0,0000	0,0000	0,0000
	004	01	03	01	08	05	11	08	09	02	06	2	08	04	03	09
d'	0,025	0,0000	0,0000	0,0000	0,0733	0,0000	0,0284	0,0000	0,0044	0,0000	0,0971	0,00000	0,0000	0,0000	0,0000	0,0264
	612	00	00	00	65	00	20	00	96	00	56	0	00	00	00	90

⁶⁰ Ibidem, rozdz. 14.4 "Korygowanie efektywności początkowej".

D_{X190}	0	0	0	0	0	349,60	0	0	0	0	0	0	0	0	0	685,57	0
						87										65	
d_{70}	0,000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,00002	0,0000	0,0000	0,0000	0,0000	0,0000
	031	07	18	12	40	51	78	69	73	14	31	1	80	41	27	59	
d'	0,000	0,0000	0,0000	0,0000	0,0000	0,0176	0,0000	0,0000	0,0000	0,0000	0,0000	0,00000	0,0000	0,0000	0,0182	0,0000	
	000	00	00	00	00	92	00	00	00	00	00	0	00	00	39	00	
S_{X190}	40,90	0	0	0	34,788	33,418	10,693	11,365	29,809	68,879	19,211	17,4321	0	21,871	54,151	27,691	
	07				4	4	9	2	6	6	9				1	7	
s_{no}	0,000	0,0004	0,0006	0,0002	0,0013	0,0009	0,0023	0,0017	0,0016	0,0004	0,0012	0,00058	0,0019	0,0008	0,0008	0,0014	
	621	10	65	92	31	18	38	19	99	20	98	4	38	64	31	16	
S'	0,025	0,0000	0,0000	0,0000	0,0463	0,0306	0,0250	0,0195	0,0506	0,0289	0,0249	0,01018	0,0000	0,0189	0,0449	0,0392	
	404	00	00	00	17	62	03	38	36	50	28	5	00	02	76	07	
S_{X190}	0	0	0	0	4,4607	0	0,0965	0	0	0	0,9999	0	0	0	3,9163	0	
s_{no}	0,003	0,0012	0,0022	0,0014	0,0051	0,0047	0,0096	0,0074	0,0099	0,0020	0,0046	0,00273	0,0080	0,0041	0,0034	0,0071	
	484	33	18	64	95	78	53	02	90	21	84	1	26	89	26	79	
S'	0,000	0,0000	0,0000	0,0000	0,0231	0,0000	0,0009	0,0000	0,0000	0,0000	0,0046	0,00000	0,0000	0,0000	0,0134	0,0000	
	000	00	00	00	72	00	31	00	00	00	83	0	00	00	17	00	
S_{X190}	289,1	0	0	0	0	0	226,02	0	238,53	36,856	658,00	1747,42	0	0	0	493,60	
	615						19		96	3	32	8				4	
s_{no}	0,000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0001	0,0001	0,0001	0,0000	0,0000	0,00002	0,0001	0,0000	0,0000	0,0000	
	037	21	37	15	06	71	17	20	17	28	58	2	28	59	65	75	
S'	0,010	0,0000	0,0000	0,0000	0,0000	0,0264	0,0000	0,0000	0,0280	0,0010	0,0378	0,03929	0,0000	0,0000	0,0000	0,0369	
	730	00	00	00	00	00	21	00	17	21	87	2	00	00	00	25	
S_{X190}	295,4	0	0	0	241,07	102,76	0	30,669	249,33	160,78	0	840,790	0	172,50	161,77	89,615	
	805				92	47			93	22	8			48	09	2	
s_{no}	0,000	0,0000	0,0001	0,0000	0,0002	0,0002	0,0005	0,0003	0,0002	0,0000	0,0002	0,00007	0,0006	0,0001	0,0001	0,0002	
	134	64	02	91	02	00	12	81	46	97	53	7	02	86	59	80	
S'	0,039	0,0000	0,0000	0,0000	0,0486	0,0205	0,0000	0,0116	0,0612	0,0156	0,0000	0,06512	0,0000	0,0320	0,0257	0,0250	
	450	00	00	00	30	41	00	83	66	71	00	0	00	05	58	78	
K_o	0,101	0,0000	0,0000	0,0000	0,1914	0,9613	0,0807	0,0312	0,1444	0,2422	1,3217	17,9329	0,0000	0,0509	0,6769	0,1276	
	196	00	00	00	85	85	75	21	14	38	55	98	00	06	74	99	
w	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

Źródło: Opracowanie własne w Excelu.

W kolejności obliczymy skorygowana efektywność obiektu o -tego z zależności:

$$\hat{\kappa}_o = \hat{\theta}_o - K_o$$

Tab. 7.13. Skorygowane efektywności obiektów według korekt typu A

W	$\hat{\theta}_o$	Miejsce wg $\hat{\theta}_o$	K_o (typ A)	$\hat{\kappa}_o(A)$	Efektywność wstępna	Miejsce wg $\hat{\kappa}_o(A)$
1	0,7995	10	0,101196	0,698304	0,698304	8
2	1	1	0	1	1	1
3	1	1	0	1	1	1
4	1	1	0	1	1	1
5	0,9443	5	0,191485	0,752815	0,752815	6
6	0,6391	16	0,961385	-0,322285	0	14
7	0,7171	14	0,080775	0,636325	0,636325	10
8	0,7854	12	0,031221	0,754179	0,754179	5
9	0,7986	11	0,144414	0,654186	0,654186	9
10	0,8783	9	0,242238	0,636062	0,636062	10
11	0,8849	8	1,321755	-0,436855	0	14
12	0,8889	7	17,932998	-17,044098	0	14
13	1	1	0	1	1	1
14	0,7809	13	0,050906	0,729994	0,729994	7
15	0,911	6	0,676974	0,234026	0,234026	13
16	0,6931	15	0,127699	0,565401	0,565401	12

Źródło: Opracowanie własne w Excelu.

Dla naszego przykładu empirycznego w obiektach 6, 11,12 wystąpiły stosunkowo duże kary jednostkowe (patrz tabela 7.12), przy odpowiednio małym poziomie efektywności początkowej $\hat{\theta}_o$, co spowodowało, że efektywność skorygowana jest niedopuszczalnie ujemna. W takiej sytuacji przyjmowana jest jako zero⁶¹. Spowodowało to zmianę miejsca w rankingu (patrz tabela 7.13). Przyśkipamy zatem do etapu drugiego - kar typu B - potraceria w relacji do efektywności początkowej i zmodyfikujemy zastosowane już wcześniej wzory⁶²:

- dla nakładu n -tego:

$$s_{no} = \hat{\theta}_o \frac{1}{(N+R)} \frac{1}{x_{no}}$$

- dla rezultatu r -tego:

$$d_{ro} = \hat{\theta}_o \frac{1}{(N+R)} \frac{1}{y_{ro}}$$

Korektę całkowitą z uwzględnieniem kar typu B określamy ze wzoru:

$$K_o^B = \hat{\theta}_o K_o^A$$

Zatem skorygowaną efektywność obliczymy według:

$$\hat{\kappa}_o = \hat{\theta}_o - K_o^B$$

Tab. 7.14. Skorygowane efektywności obiektów według korekt typu B

W	$\hat{\theta}_0$	K_o^A	$\hat{\kappa}_o^A$	K_o^B	$\hat{\kappa}_o^B$	Miejsce wg $\hat{\kappa}(typuB)$	Miejsce wg $\hat{\kappa}_o typu(A)$
1	0,7995	0,101196	0,698304	0,080906	0,718594	8	8
2	1	0	1	0	1	1	1
3	1	0	1	0	1	1	1
4	1	0	1	0	1	1	1
5	0,9443	0,191485	0,752815	0,180819	0,763481	5	6
6	0,6391	0,961385	-0,322285	0,614421	0,024679	14	14
7	0,7171	0,080775	0,636325	0,057924	0,659176	11	10
8	0,7854	0,031221	0,754179	0,024521	0,760879	6	5
9	0,7986	0,144414	0,654186	0,115329	0,683271	9	9
10	0,8783	0,242238	0,636062	0,212758	0,665542	10	10
11	0,8849	1,321755	-0,436855	1,169621	0	15	14
12	0,8889	17,932998	-17,044098	15,940642	0	15	14
13	1	0	1	0,000000	1,000000	1	1
14	0,7809	0,050906	0,729994	0,039753	0,741147	7	7
15	0,911	0,676974	0,234026	0,616723	0,294277	13	13
16	0,6931	0,127699	0,565401	0,088508	0,604592	12	12

Źródło: Opracowanie własne w Excelu.

W wyniku modyfikacji kar nastąpiła nieznaczna korekta w rankingu obiektów (województw), chociaż nadal trzeba było zakwalifikować obiekty 11. oraz 12. jako miejsce zerowe. Interesujące pozostaje przeanalizowanie korygowania efektywności początkowej bez danej postaci ($C-X5$), lecz to pozostawia się jako ćwiczenie Czytelnikowi.

* * *

Ilustracja modelu CCR z karami za luzy nakładów i rezultatów na danych statystycznych wymaga pracochłonnego testowania różnych sposobów ustalenia rankingu obiektów. W doborze szeregów danych określonego roku zachodzi pytanie co jest stymulantą, a co destymulantą. Wydaje się,

⁶¹ Guzik B., *Podstawowe modele DEA ...*, op. cit., s. 275.

⁶² Ibidem, s. 276.

że aparat teoretyczny zaprezentowany w monografii Guzika wymaga jeszcze dalszego uzupełnienia, zwłaszcza, gdy dane są w postaci $(C-X)$ lub (I/X) ⁶³.

Dużym udogodnieniem jest jednak moduł LPILP "Programowanie liniowe i całkowitoliczbowe" programu *WinQSB*, który nie tylko szybko podaje rozwiązanie optymalne, lecz również w raporcie kombinowanym daje materiał do analizy jego wrażliwości.

⁶³ Ibidem

8. Ocena nieefektywności gospodarowania zasobami ludności



8.1. Wstęp

Efektywność mierzona jest stosunkiem uzyskanych efektów do poniesionych nakładów. Różne są współczynniki mierzenia efektywności czy też nieefektywności, której zmienne mogą być w postaci ilościowej, wartościowej z uwzględnieniem cen jednostkowych poniesionych nakładów, czy też różnego rodzaju efektów. Narzędziem, które umożliwia klasyfikowanie obiektów pod względem efektywności jest program *DEAP*. W wersji 2.1 jest on dostępny w Internecie i pozwala na korzystanie z szeregów czasowych zmiennych, jak też z szeregów czasowo-przestrzennych zwanych panelowymi. Daje też możliwość wprowadzania cen informacyjnych i wyliczania tzw. efektywności ekonomicznej. Gdy szeregi zmiennych są w postaci ilościowej to program *DEAP* oblicza tzw. efektywność techniczną. W ramach programu *DEAP* wyodrębnia się również efektywność alokacji związana ze strukturą nakładów.

Jeśli mamy kilka szeregów czasowych reprezentujących wyjście (*output*) oraz kilka szeregów czasowych dotyczących wejścia (*input*), czyli nakładów i chcemy dokonać optymalizacji tak skonstruowanego ilorazu, to przed każdą ze zmiennych stoją parametry. Optymalizację takiego zadania możemy zrealizować z punktu widzenia celu jakim jest zorientowanie się na wyjście lub wejście. Program *DEAP* sam ustala odpowiednie wielkości parametrów przy optymalizacji licznika we wzorze na efektywność, który w postaci ogólnej możemy zapisać jako:

$$Efektywność = \frac{\sum_{r=1}^s \mu_r Efekt}{\sum_{i=1}^m v_i Nakład}$$

W ramach programu *DEAP* mamy do wyboru opcję *CRS* dotyczącą maksymalizacji funkcji celu licznika, a *PRS* mianownika we wzorze na obliczaną efektywność (gospodarność) techniczną obiektu⁶⁴. Wymienione we wzorze parametry μ_r, v_i są to wagi określające ważność odpowiednich efektów/nakładów.

Omawiany program dysponuje jeszcze opcją *COST* wyznaczania efektywności ekonomicznej oraz opcją *MALMQUIST* do skorzystania z danych panelowych i wyznaczania dystansu danego okresu lub obiektu wobec rozwiązania optymalnego wyobrażonego w postaci krzywej - izokwanty. Program bazuje bowiem na analizie komputerowej odległości miar obserwacji od najlepszych dla teoretycznie ustalonej dla obiektu np. przedsiębiorstwa. Wagi ustalone są po kątem maksymalizacji efektywności każdego obiektu. W standardowym rozwiązaniu problemu optymalizacyjnego następuje wyznaczenie odległości rzeczywistego poziomu efektywności produkcji obiektu od teoretycznie stwierdzonej granicy efektywności.

Program *DEAP* bazuje na metodzie zwanej *DEA*, która pozwala na ustalenie krzywej efektywności o wartości 1, na której znajdują się wszystkie najbardziej efektywne lub nieefektywne jednostki badanej zbiorowości. Jeśli obiekty traktowane są jako jednostki decyzyjne (DMU) to znajdują się poza krzywą efektywności, mają wartości różne od 1 i uważane są jako nieefektywne technicznie. Tak więc program *DEAP* służy do ustalenia efektywności technicznej produkcji mając wykonaną produkcję wyrobów finalnych lub części zamiennych i poniesione nakłady w postaci

⁶⁴ Becker A., *Analiza efektywności działania województw Polski*, Internet, Katedra Zastosowań Matematyki, Akademia Rolnicza w Szczecinie.

materiałochłonności, robocizny, energochłonności określonych procesów wytwarzania. Szersze odniesienie to tematyki efektywności ze zorientowaniem na obiekty struktury terytorialnej spotykamy w książce Mariana Duczmala⁶⁵. Pomocnym w niniejszym opracowaniu był też *Przewodnik po DEAP*⁶⁶. Podstawy teoretyczne modeli *DEA* znajdujemy w książce Bogusława Guzika pt. *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*⁶⁷. Zachęcony możliwościami programu *DEAP* podjąłem próbę innego jego zastosowania niż w sferze produkcji, a mianowicie skoncentrowałem swoją uwagę na zasobach ludności, a zwłaszcza na nieefektywności gospodarowania nimi. Do rozważań przyjęto ogólny model ekonometryczny liniowy wielowymiarowy, jaki wygenerowałem przy zastosowaniu funkcji *REGLINP Excela*⁶⁸ w postaci:

$$Mio = f(Em, M, R, U, Z),$$

Gdzie w odniesieniu do ludności, 12. powiatów województwa opolskiego, szeregu 15. obserwacji rocznych (2002-2016) poszczególne zmienne oznaczają:

Mio - (*Migracja* - *odpływ*) → *żonaci/zamężne*;

Em - *Emigracja* (*żonaci/zamężne*);

M - *Małżeństwa* → *grupa wieku mężczyzny*, przedziały: 20-24, 25-29, 30-34;

R - *Rozwody* → według *grupa wieku mężczyzny w dniu wniesienia powództwa*, przedziały: 20-24, 25-29, 30-34;

U - *Urodzenia* → według *pleć dziecka, chłopiec*;

Z - *Zgony* → według *grupa wieku*, przedziały: 0 dni; 1-27 dni; 28, 29 dni, 11 miesięcy; 1-4 lata; 5-9 lat; 10-14 lat).

Zamierzeniem początkowym było uwzględnienie również w badaniu efektywności gospodarowania ludnością szerszej platformy zmiennych w określeniu efektywności zasobów ludności, a mianowicie liczby zatrudnionych w podstawowych sektorach gospodarki oraz liczby absolwentów różnego poziomu szkół. Po stronie efektów, czyli licznika wzoru na efektywność znalazły się zmienne:

- zatrudnienie-przemysł,
- zatrudnienie-budownictwo,
- bezrobotni,
- migracja-dopływ,
- migracja-odpływ,
- emigracje,
- imigracje.

Natomiast w mianowniku zaproponowano zmienne:

- urodzenia,
- absolwenci (szkoła podstawowa),
- absolwenci (szkoła zawodowa),
- absolwenci (technikum),
- absolwenci (szkoła ogólnokształcąca),
- absolwenci (szkoła policealna),
- dyplomanci (szkoła wyższa),

⁶⁵ Duczmal M., *Efektywność działania obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013, ISBN 978-83-62683-37-6.

⁶⁶ Coelli T., *Przewodnik po DEAP wersji 2.1, Analiza zwijania [enevelopment] danych*, Centrum Analizy Efektywności i Produktywności, [www.une.edu.au/econometrics/cepa.htm].

⁶⁷ Guzik B., *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2009.

⁶⁸ Wornalkiewicz W., *Prognozowanie migracji ludności z uwzględnieniem wag harmonicznych*, s. 212-233, monografia zbiorowa „Uwarunkowania demograficzne rozwoju społecznego i ekonomicznego miasta Nysa i powiatu nyskiego”, Opole-Nysa 2018, Wyższa Szkoła Zarządzania i Administracji w Opolu, ISBN 978-83-946765-0-6.

- małżeństwa,
- separacje orzeczone,
- rozwody,
- zgony.

Ramy niniejszej publikacji nie pozwalają zaprezentować szerszemu ogółowi postępowania związanego z przygotowaniem tak obszernych danych statystycznych wejścia/wyjścia oraz procedury korzystania z programu *DEAP*. Program ten generuje bowiem wielostronicowe dokładne raporty. Są one pokazane w układzie różnych technik modelowania wyznaczania efektywności i z tego względu jako uproszczone założenia wejściowe przyjęto:

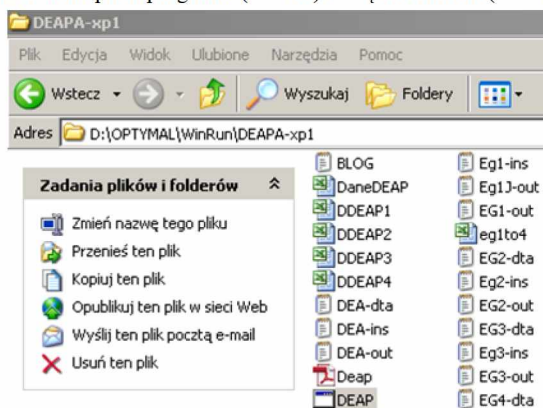
- skorzystanie z *Bazy Demografia* prowadzonej w Internecie przez GUS,
- jedną zmienną wyjścia tj. migracja (odpływ),
- pięć zmiennych liczbowych wejścia: emigracja, małżeństwa, rozwody, urodzenia, zgony.

Program *DEAP wersja 2.1* został napisany w *Fortranie*, jeszcze pod starym systemem operacyjnym *DOS*, ale pracuje również pod *Windows*em. Wymaga trzech plików tekstowych z kropką dziesiętną dla cyfr wypozycjonowanych w kolumnach do lewej, a mianowicie pliku danych, pliku sterującego oraz pustego pliku do zapisu raportu, czyli efektu realizacji procedury optymalizacyjnej.

Zaprezentowane w niniejszym materiale rozwiązania jak i sam tytuł mogą prowokować Czytelnika, ale zamierzeniem moim było wskazanie na możliwości implementacji *DEAP* także do sfery społecznej, a w tym do oceny gospodarowania ludnością określonego powiatu województwa opolskiego w sytuacji wzmożonego procesu migracji i emigracji ludności.

8.2. Dostęp do danych statystycznych

Jak już wspomniałem program *DEAP* o rozszerzeniu *exe* dostępny jest w Internecie i zawiera pliki przykładowe, które można skopiować i przystosować do własnych uwarunkowań zadania decyzyjnego. Zainstalowałem go na ścieżce *D:\OPTYMAL\WinRun\DEAPA-xp1* - zobacz rysunek 8.1. Widzimy tu omawiany program oraz pliki tekstowe danych np. *EG3-dta*, instrukcji *EG3-ins*, plik rezultatu *EG3-out* o rozszerzeniu *txt*. Zwróćmy jeszcze uwagę na własności aplikacji programowej *DEAP*, a mianowicie rozmiar pliku programu (548 kB) i datę utworzenia (3.02.2014).



Źródło: Opracowanie własne.

Rys. 8.1. Program *DEAP* i pliki pomocnicze

Jak już nadmieniałem *DEAP* wymaga trzech plików tekstowych, które można utworzyć w także w arkuszu kalkulacyjnym *Excel*, pamiętając o jego sformatowaniu wyjściowym jako zapisanych wielkości w kolumnach do lewej, przy użyciu kropki dziesiętnej. Opis programu *DEAP* występuje w

poradniku wydanym przez Uniwersytet w Brisbane w Australii⁶⁹. Opisuje program komputerowy, który został napisany do prowadzenia tzw. *analizy zawijaniem DEA* dla celów obliczania wydajności produkcji. Coelli zastosował w programie metody opracowane między innymi przez Rolf Fare oraz Shawna Grosskopfa. Jak już wspomniałem program obejmuje dwie opcje odnoszące się do szeregow obserwacji zmiennych wejścia/wyjścia obiektów, z rozróżnieniem bazowania na nakładach czy efektach, a trzecia opcja wykorzystuje dane panelowe, czyli szeregi zmiennych dla obiektów w jednakowych okresach czasowych.

Analiza DEA wiąże się z wykorzystaniem metody programowania liniowego dla określenia efektywności danego obiektu w stosunku do teoretycznego określonego przez program. Uważa się, że metody *DEA* stanowią nowoczesny sposób pomiaru efektywności, przykładowo zastosowanej określonej technologii w danym przedsiębiorstwie, gdzie możemy wyróżnić wiele oczekiwanych efektów oraz występują różne rodzajowo nakłady. Coelli, powołując się na innych autorów zaproponował, aby efektywność firmy składa się z dwóch części:

- *sprawności technicznej (skuteczności)*, która jest obrazem zdolności firmy do uzyskania maksymalnego wyjścia z realizacji wymagań podanych w zbiorze danych wejściowych;
- *skuteczności przydziału*, która odzwierciedla zdolność firmy do korzystania z zasobów w optymalnej proporcjach.

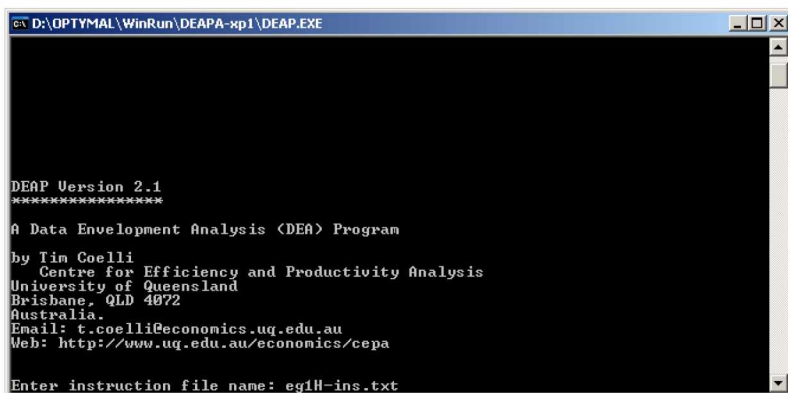
Zdaniem autora po uwzględnieniu cen dla zmiennych te dwa sposoby można następnie łączyć w celu uzyskania pomiaru tzw. *wydajności ekonomicznej*.

Po tych rozważaniach, wróćmy do naszego programu. Jak już nadmieniałem, zainstalowałem go z Internetu ze strony: www.une.edu.au/economics/emet//cepa.htm. Najpierw tworzony jest katalog *WinRun* na dysku twardym np. *D:*. Pliki wzorcowe, podręcznik oraz program ściągany jest w postaci skompresowanej z rozszerzeniem *zip* i wymaga następnie „rozpakowania”. Program *DEAP Version 2.1* bazuje on na pliku wsadowym danych wejściowych oraz pliku instrukcji w formacie tekstowym. Uruchomienie programu *DEAP* wymaga jak już wspomniano wcześniejszego rozpakowania (zainstalowania) na dysku następujących zbiorów:

- program *DEAP.EXE*,
- plik startujący *DEAP.000*, który zapamiętuje parametry realizacji programu w celu ich użycia w późniejszym okresie.
- pliki tekstowe danych,
- pliki tekstowe instrukcji,
- pliki wyjściowe.

Uruchomienie programu *DEAP.EXE* następuje poprzez podwójne kliknięcie na jego ikonie, a następnie wpisanie z klawiatury zbioru instrukcji (sterowania) o nazwie użytkownika np. *Eg1H-ins.txt* i naciśnięciu klawisza *ENTER* (zob. rysunek 8.2).

⁶⁹ Coelli T.J., *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program*, CEPA Working Papers, Department of Econometrics University of New England, Australia.



Źródło: Opracowanie własne.

Rys. 8.2. Okno wejściowe program *DEAP*

Jeśli plik danych jest poprawny, a plik wyjścia pusty to generowany jest obszerny raport rezultatu procesu optymalizacji uwzględniający zadany w pliku sterującym model efektywności i wskazane w pliku danych szeregi czasowe lub panelowe zmiennych efektu i nakładów. Zwróćmy uwagę na informacje o autorze programu dotyczącego tzw. analizy zawijania (*Data Envelopment Analysis - DEA*), czyli wspomnianym już Timie Coelli z Uniwersytetu Queensland w Brisbane w Australii. Dla mimo wszystko pracochłonnego zebrania danych z okresu (2002-2016) skorzystałem z funkcji „Generowanie zestawień” *Bazy Demografia* dostępnej na stronie: demografia.stat.gov.pl (zob. rys. 8.3).



Źródło: Opracowanie własne.

Rys. 8.3. Okno *Bazy Demografia*

Postępowanie związane z uzyskaniem szeregów panelowych (czasowo-przestrzennych) dwunastu powiatów województwa opolskiego i piętnastu lat danych statystycznych GUS-u było następujące:

- wybranie zakresu tematycznego (II. *RUCH NATURALNY LUDNOŚCI* lub III. *MIGRACJE LUDNOŚCI*);

- wskazanie kolejnego roku np. 2002;
- wybór jednostki terytorialnej, najpierw POLSKA i klikamy na (+) rozwijając listę województw,
- wybieramy opolskie,
- klikamy na jedno z dwunastu powiatów, np. Brzeski.

Wymienione postępowanie obrazują rysunki 8.4 oraz 8.5.

1 Wybierz zakres tematyczny	2 Wybierz rok
I. STAN I STRUKTURA LUDNOŚCI - Stan ludności II. RUCH NATURALNY LUDNOŚCI - Małżeństwa - Separacje orzeczone - Separacje zniesione - Rozwody - Urodzenia - Zgony III. MIGRACJE LUDNOŚCI - Emigracje - Imigracje - Migracje - napływ - Migracje - odpływ	2016 2015 2014 2013 2012 2011 2010 2009 2008 2007 2006 2005 2004 2003 2002

Źródło: Opracowanie własne.

Rys. 8.4. Okno wyboru zakresu tematycznego i roku w *Bazie Demografia*

- Opolskie
 - Brzeski - powiat
 - Głubczycki - powiat
 - Kędzierzyńsko-kozielski - powiat
 - Kluczborski - powiat
 - Krapkowicki - powiat
 - Namysławski - powiat
 - Nyski - powiat
 - Oleski - powiat
 - Opolski - powiat
 - Prudnicki - powiat
 - Strzelecki - powiat
 - Opole - miasto na prawach powiatu

Źródło: Opracowanie własne.

Rys. 8.5. Okno wyboru powiatu w ramach wyboru jednostki terytorialnej *Bazy Demografia*

8.3. Zamierzenia początkowe i przyjęte założenia

Jak już wspominałem we *Wstępie* zamierzałem pokazać ustalenie nieefektywności gospodarowania zasobami ludności na szerszej płaszczyźnie zmiennych stanowiących licznik „Efekty” współczynnika efektywności jak i mianownik „Nakłady” w tym samym współczynniku. Zakres danych tego typu wykracza poza *Bazę Demografia* i część danych musi być pozyskana i zestawiona z roczników statystycznych województwa opolskiego. Można je później zestawić w szeregi panelowe poszczególnych 12. powiatów i okresu czasowego (2002-2016). Było by to aż 15 tablic i taka praca zasługuje na odrębną szerszą publikację. Cechy (zmienne) przyczyn skutku, czyli (*Migracji - odpływ*) przyjęto intuicyjnie i należą do nich: *Emigracja, Małżeństwa, Rozwody, Urodzenia, Zgony*.

Bazując na wcześniejszym moim artykule, dla przejrzystości tego materiału powtórzę, że migracja (odpływ) ludności z określonego powiatu jest funkcją liniową wielowymiarową czynników⁷⁰:

⁷⁰ Wornalkiewicz W., *Prognozowanie migracji ludności z uwzględnieniem wag harmonicznnych*, op. cit.

Migracja - odpływ (żonaci/zamężne) = f[*Emigracja* (żonaci/zamężne); *Małżeństwa* (grupa wieku mężczyzny, przedziały: 20-24, 25-29, 30-34); *Rozwody* (wg grupa wieku mężczyzny w dniu wniesienia powództwa, przedziały: 20-24, 25-29, 30-34); *Urodzenia* (wg płeć dziecka, chłopiec); *Zgony* (wg grupa wieku, przedziały: 0 dni; 1-27 dni; 28, 29 dni, 11 miesięcy; 1-4 lata; 5-9 lat; 10-14 lat)]. Po zastosowaniu funkcji *REGLINP Excela* dla danych powiatu nyskiego uzyskano model:

$$Mio = 269,4 + 2,079 EM - 0,22 M - 0,37 R + 0,943 U + 0,035 Z$$

przy współczynniku $R^2 \approx 0,7$ wskazującym na w miarę dobre dopasowanie modelu regresji do danych empirycznych.

Zebrane dane panelowe w narastającym układzie lat (2002-2016), a w ramach nich powiatów i występujących dla nich zmiennych (*Mio*, *Em*, *M*, *R*, *U*, 4.) ze względu na obszerność zamieszczono po zakończeniu niniejszego opracowania (zob. tabela 1).

Dane panelowe obejmujące (12 x 15 = 180) wierszy zapisano w pliku *Dane-DEAP.xls* w podkatalogu *D:\Artykul-deap*. Procedura postępowania związana z importem pliku z *Excela* do podkatalogu *D:\OPTIMAL\WinRun\DEAPA-xp1* środowiska instalacyjnego programu *DEAP.EXE* jest następująca⁷¹:

- skopiowanie skoroszytu;
- pozostawienie w ramach skoroszytu tylko jednego arkusza z danymi;
- usunięcie wiersza nagłówka i formatowania liniami kolumn oraz wierszy;
- zamiana wartości numerycznych na tekstowe poprzez zaznaczenie danych i wywołanie menu

Format/komórka/tekstowe;

- automatyczna zamiana przecinka w wartościach danych zmiennych na kropkę dziesiętną korzystając z menu *Edycja/zamień*;
- usunięcie kolumn: okresy P, nazwa powiatu, rok;
- przemieszczenie kolumn zmiennych w kolejności najpierw dotyczące efektów później nakładów;
- zapisanie tak przygotowanego zbioru jako kopii z rozszerzeniem *xls*;
- zapisanie zbioru jako tekstowego *DOS-a* z rozszerzeniem *txt* w wymienionym wcześniej podkatalogu programu *DEAP* nadając mu nazwę typu eg7-dta.txt.

8.4. Modelowanie ekonometryczne cechy złożonej *Mio*

Nadmieniłem już o opracowaniu w moim wcześniejszym artykule modelu liniowego wielowymiarowego dla powiatu nyskiego, podając tylko jego postać końcową po estymacji funkcją *REGLINP Excela*. Jednak mając zgromadzone dane proponuję przedstawić procedurę w tym zakresie na danych statystycznych demograficznych powiatu „Opole - miasto na prawach powiatu” (zob. tabela 8.2).

Tab. 8.2. Dane demograficzne powiatu „Opole - miasto na prawach powiatu”

Powiat	Lp.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Max
Opole - miasto na prawach powiatu	Rok	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	
Migracja- odpływ	<i>Mio</i>	617	683	638	645	753	855	701	750	697	686	702	708	549	497	492	855
Emigracja	<i>Em</i>	44	23	26	74	125	81	97	55	37	49	68	80	69	0	9	125
Małżeństwa	<i>M</i>	458	409	439	414	523	537	562	571	500	433	440	428	372	394	362	571

⁷¹ W opracowaniu bazowano na niektórych procedurach zamieszczonych w książce: Duczmal M., *Efektywność działania obiektów struktury terytorialnej*, op. cit.

Rozwody	R	104	117	130	135	135	117	103	102	75	74	73	83	89	79	63	135
Urodzenia	U	447	473	496	495	533	492	542	586	612	524	549	515	542	574	607	612
Zgony	Z	12	8	8	11	9	6	10	7	8	4	13	4	6	9	10	13

Źródło: Opracowanie własne na podstawie danych statystycznych w Baza Demografia.

W tabeli tej skorzystano też z funkcji [=MAX(...)] *Excela* dla wskazania wartości maksymalnej w danym wierszu (zob. czcionka pogrubiona). Jednak do potrzeb *REGLINP* wygodniej jest dokonać transpozycji macierzy danych, tak aby szeregi czasowe zmiennych (*Mio*, *Em*, *M*, *R*, *U*, *Z*) były ułożone kolumnami. Zaznaczono również pogrubieniem cyfr kolumny okresów (5-8) z dominującą przeważnie wartością danej cechy. Po tej operacji i wywołaniu funkcji *REGLINP* w oknie dialogowym (rysunek 8.6) wpisujemy przedziały adresów dla:

- zmiennej objaśnianej *Mio*, traktowanej jako *y*;
- macierzy zmiennych objaśniających [*Em M R U Z*];
- deklaracji występowania stałej w modelu ekonometrycznym, poprzez wpisanie „PRAWDA”;
- żądania wygenerowania podstawowej statystyki opisowej, wpisując „PRAWDA”.

Argumenty funkcji

REGLINP

Znane_y: N38:N52 = {-7,93659626976197;-0,601053150035}

Znane_x: O38:S52 = {44\458\104\447\12;23\409\117\473\8}

Stała: PRAWDA = PRAWDA

Statystyka: PRAWDA = PRAWDA

Zwraca statystykę opisującą trend liniowy, dopasowany do znanych punktów danych, dopasowując linię prostą przy użyciu metody najmniejszych kwadratów.

Statystyka - wartość logiczna: zwraca dodatkowe statystyki regresji = PRAWDA; zwraca współczynniki m i stałą b = FAŁSZ lub pominięta.

Wynik formuły = -7,93659627

[Pomoc dotycząca tej funkcji](#)

OK Anuluj

Źródło: Opracowanie własne.

Rys. 8.6. Okno wskazania argumentów funkcji *REGLINP*

Omawiana funkcja generuje na wyjściu tabelę 8.3 zawierającą:

- parametry ($a_0 - a_5$), stojące przy poszczególnych zmiennych (*Em - Z*);
- błędy standardowe oszacowania parametrów;
- współczynnik determinacji $R^2 = 0,7728$, jako ocena dobroci dopasowania modelu do danych statystycznych;
- odchylenie standardowe reszt S_e , między wartościami modelowymi a rzeczywistymi;
- statystykę Fishera (*F*), oceniająca na podstawie rozkładu *F* całość modelu;
- ilość stopni swobody ($iss = n - m - 1 = 15 - 5 - 1 = 9$), gdzie *m* – liczba zmiennych objaśniających;
- *ESS* - estymowana sumę kwadratów,
- *RSS* - resztową sumę kwadratów;
- *#N/D* - nie występowanie danych dla danego modelu w tabeli standardowej.

Tab. 8.3. Tabela wygenerowana przez funkcję *REGLINP*

a_5	a_4	a_3	a_2	a_1	a_0	Opis
-7,937	-0,601	-0,302	1,0003	0,6216	589,9	parametry
5,8886	0,434	0,9327	0,2966	0,5732	274,04	błąd standardowy
0,7728	57,818	#N/D!	#N/D!	#N/D!	#N/D!	R^2 , S_e
6,1219	9	#N/D!	#N/D!	#N/D!	#N/D!	F , iss
102327	30087	#N/D!	#N/D!	#N/D!	#N/D!	ESS , RSS

Źródło: Opracowanie własne w Excelu.

Zatem model w postaci ogólnej jest następujący: $y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5$, czemu odpowiada:

$$Mio = 589,9 + 0,6216 Em + 1,0003 M - 0,302 R - 0,601 U - 7,9337 Z.$$

Tak więc w powiecie „Opole - miasto na prawach powiatu” na odpływ ludności poza teren mają dodatni wpływ sprzyjające warunki dla udających się na emigrację i liczba zawieranych małżeństw, a ujemny zmienne R , U oraz Z .

Oczywiście cecha złożona Mio w sensie nieefektywności technicznej to mamotrawstwo najcenniejszego zasobu aktualnej i potencjalnej siły wytwórczej rozpatrywanego powiatu. W analogiczny sposób możemy określić modele ekonometryczne dla pozostałych powiatów województwa opolskiego.

8.5. Zastosowanie opcji *Malmquist*

Program *DEAP* umożliwia przetwarzanie danych statystycznych ułożonych dla zestawu cech (zmiennych) wejścia/wyjścia narastająco według okresów (lat), a następnie narastająco według obiektów zwanych DMU, a w naszym przykładzie reprezentowanych przez powiaty. Plik danych wejściowych panelowych nazwano (*eg7-dta.txt*), danych sterujących (*eg7-ins.txt*), a plik pusty oczekujący na rezultaty z optymalizacji ma nazwę użytkownika (*eg7-out.txt*). Wszystkie te pliki zapisane zostały w środowisku programu *DEAP.EXE*, a więc na ścieżce: *D:\OPTIMAL\WinRun\DEAP-xp1*. Fragment początkowy pliku tekstowego danych wejściowych panelowych pokazano na rys. 8. 7.

DMU	Em	M	R	U	Z
564	27	363	65	426	13
334	43	208	13	230	6
398	285	373	46	391	12
414	70	260	20	305	7
321	272	247	28	291	4
256	14	204	19	222	4
862	36	461	65	640	12
359	111	281	15	298	4
628	178	467	44	525	10
274	121	221	23	254	6
362	309	307	26	339	4
617	44	458	104	447	12

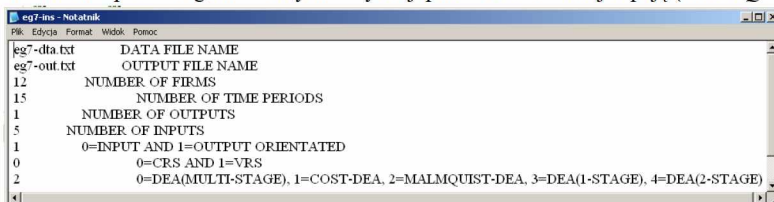
Źródło: Opracowanie własne.

Rys. 8.7. Dane okresu 1. w ramach pliku (*eg7-dta.txt*)

W pierwszej kolumnie cyfr jest zmienna Mio , a w kolejnych pięć zmiennych Em , M , R , U , Z . Plik danych sterujących (*eg7-ins.txt*) zawiera następujące parametry dla programu *DEAP.EXE* (zob. rys. 8.8):

- nazwę pliku wejściowego (*DATA FILE NAME*);

- nazwę pliku wyjściowego (*OUTPUT FILE NAME*);
- liczbę powiatów (*NUMBER OF FIRM*);
- liczbę okresów czasowych - lat (*NUMBER OF OUTPUTS*);
- liczbę zmiennych wejściowych w mianowniku wzoru na nieefektywność *Mio* (*NUMBER OF INPUTS*);
- orientację na wyjście, czyli zmienną *Mio*;
- wskazanie na uproszczoną metodę przeprowadzania obliczeń;
- wskazanie sposobu zgromadzonych danych tj. panelowe i realizacja opcją (*MALMQUIST-DEA*).



Źródło: Opracowanie własne.

Rys. 8.8. Plik sterujący (*eg7-ins.txt*)

Po naciśnięciu klawisza „ENTER” zaglądamy do pustego wcześniej pliku rezultatu (*eg7-out.txt*) i znajdujemy rozwiązanie, którego początek dla roku 1., czyli 2002 pokazano na rysunku 4.9. Widzimy tu określenie sumarycznych odległości (*DISTANCE SUMMARY*) poszczególnych efektywności technicznych (*TE*) powiatów obliczone według metody *CRS TE* i *VRS TE*, względem teoretycznie określonej izokwanty jako rozwiązania optymalnego. Dla *CRS TE* (komputerowo zaznaczonego małymi literami)) występuje oprócz *t* odniesienie do roku poprzedniego (*t-1*) oraz następnego (*t+1*). Na zakończenie fragmentu listy (zob. rysunek 8.9) program wylicza średnie geometryczne dla okresu pierwszego i dwunastu powiatów.

DISTANCES SUMMARY				
year = 1				
firm no.	crs	te rel to tech in yr	vrs	te
	t-1	t	t+1	
1	0.000	0.979	1.000	1.000
2	0.000	1.000	1.678	1.000
3	0.000	0.701	0.714	0.734
4	0.000	0.979	1.345	1.000
5	0.000	0.944	0.925	1.000
6	0.000	0.956	1.000	1.000
7	0.000	1.000	1.245	1.000
8	0.000	1.000	1.601	1.000
9	0.000	0.894	0.978	0.962
10	0.000	0.761	0.841	0.780
11	0.000	1.000	1.065	1.000
12	0.000	1.000	0.892	1.000
mean	0.000	0.935	1.107	0.956

Źródło: Opracowanie własne.

Rys. 8.9. Fragment pliku wyjściowego (*eg7-out.txt*) dla okresu - roku 1 (2002)

Długa jest lista z wyszczególnieniem (*DISTANCES SUMMARY*) bo obejmuje specyfikację dla 15. okresów - lat (2002-2016) i z tego powodu pomińmy okresy 2-14 i zobaczmy jeszcze okres 15, czyli rok 2016 i to tylko metodę *CRS TE* oraz okres bieżący *t*.

Powiat	1	2	3	4	5	6	7	8	9	10	11	12	Średnia
CRS	0,875	1,000	0,786	1,000	0,907	1,000	0,919	1,000	0,932	0,864	1,000	1,000	0,940
TE													

W ramach raportu rezultatu z wyszczególnieniem *DISTANCE SUMMARY* program *DEAP.EXE* generuje jeszcze:

- zestawienie porównawcze kolejnych lat (2002-2016) w stosunku do roku wcześniejszego nazwane (*MALMQUIST INDEX SUMMARY*) dla 12. powiatów z rozróżnieniem metod (*EFFCH*, *TECHCH*, *PECH*, *SECH*, *TFPCH*), przy czym przedstawienie każdej z tych metod obliczania nieefektywności może stanowić przedmiot odrębnej publikacji;

- zestawienie porównawcze kolejnych lat (2-15) o nazwie (*MALMQUIST INDEX SUMMARY OF ANNUAL MEANS*) z określeniem na zakończenie listy średniej geometrycznej dla lat i całego okresu, również z rozróżnieniem metod obliczeniowych (*EFFCH*, *TECHCH*, *PECH*, *SECH*, *TFPCH*);

- zestawienie sumaryczne oceniające poszczególne powiaty (1-12) nazwane (*MALMQUIST INDEX SUMMARY OF FIRM*), dla którego dane odniesione do obliczeń metodą *EFFCH* są następujące:

Powiat	1	2	3	4	5	6	7	8	9	10	11	12	Średnia geom.
EFFCH	0,992	1,000	1,008	1,002	0,997	1,003	0,994	1,000	1,003	1,009	1,000	1,000	1,001

Na zakończenie tego zestawienia program oblicza również średnie geometryczne wyników dla poszczególnych metod, w tym dla *EFFCH*. Warto w tym miejscu nadmienić, że dla szeregu szczegółowego średnia geometryczna jest pierwiastkiem *n*-tego stopnia z iloczynu *n* wyrazów obserwacji zmiennej *X*⁷²:

$$G_x = \sqrt[n]{\prod_{i=1}^n x_i}$$

prawdźmy to na przykładzie średniej geometrycznej obliczonej dla indeksu *Malmquist* dla 12. powiatów województwa polskiego stosując funkcję potęgi ($1/n = 1/12$):

$$= \text{POTĘGA}(1,007893167; 1/12) \approx 1,001$$

przy czym na podstawie (*MALMQUIST INDEX SUMMARY OF FIRM MEANS*) - zob. wcześniejsze wyszczególnienie:

$$1,007893167 = 0,992 \cdot 1 \cdot 1,008 \cdot 1,002 \cdot 0,997 \cdot 1,003 \cdot 0,994 \cdot 1 \cdot 1,003 \cdot 1 \cdot 1 \cdot 1$$

Korzystając z wygenerowanego pliku rezultatu (*ef7-out.txt*) w tabeli 8.4. wyszczególniono posortowane malejąco współczynniki efektywności technicznej według metody *EFFCH* bieżącego okresu do współczynników okresu wcześniejszego (rozpoczynające od okresu drugiego). Jak już wspominałem pełny raport rezultatu oprócz *EFFCH* obejmuje także inne współczynniki, a mianowicie: *TECHCH*, *PECH*, *SECH*, *TFPCH*, jak już nadmieniałem zwane też modelami lub sposobami wyznaczania efektywności.

⁷² Wornalkiewicz W., *Formułowanie modeli ekonometrycznych na potrzeby zarządzania*, Część I, Środowiska programowe statystyki opisowej, Wydawnictwo Instytut Śląski, Opole 2014.

Tab. 8.4. Średnie geometryczne roczne dla powiatów
(*MALMQUIST INDEX SUMMARY OF ANNUAL MEANS*)

Rok	11	15	7	14	5	3	9	6	12	4	2	8	13	10
Oznaczenie	2012	2016	2008	2015	2006	2004	2010	2007	2013	2005	2003	2009	2014	2011
EF-FCH	1,13 0	1,08 1	1,06 0	1,05 7	1,05 0	1,03 8	0,99 6	0,98 7	0,98 2	0,95 6	0,95 4	0,95 3	0,91 3	0,88 0

Źródło: Opracowanie własne w programie DEAP.

Raport w układzie poszczególnych powiatów obejmuje współczynniki nieefektywności (*EF-FCH*, *TECHCH*, *PECH*, *SECH*, *TFPCH*). W załączonej tabeli 8.5 pokazano tylko *EFFCH*, będące współczynnikami nieefektywności technicznej, posortowane malejąco jako średnie geometryczne dla poszczególnych 12. powiatów.

Tab. 8.5. Średnie geometryczne dla powiatów
(*MALMQUIST INDEX SUMMARY OF FIRM MEANS*)

Powiat	Nazwa	EFFCH
10	Prudnicki	1,009
3	Kędzierzyńsko-kozielski	1,008
6	Namysłowski	1,003
9	Opolski	1,003
4	Kluczborski	1,002
2	Głubczycki	1,000
8	Oleski	1,000
11	Strzelecki	1,000
12	Opole - miasto na prawach powiatu	1,000
5	Krapkowicki	0,997
7	Nyski	0,994
1	Brzeski	0,992

Źródło: Opracowanie własne w programie DEAP.

W raporcie rezultatu znajdujemy też w ramach roku 15. (2016) współczynniki nieefektywności odniesione do bieżących okresów t według modeli *CRS TE* oraz *VRS TE*. Pełny raport obejmuje ponadto dla modelu *CRS TE* również okresy $(t - 1)$ oraz $(t + 1)$.

Powiat	1	2	3	4	5	6	7	8	9	10	11	12	Średnia
<i>CRS TE</i>	0,875	1,000	0,786	1,000	0,907	1,000	0,919	1,000	0,932	0,864	1,000	1,000	0,940
<i>VRS TE</i>	1,000	1,000	0,825	1,000	1,000	1,000	1,000	1,000	1,000	0,867	1,000	1,000	0,974

Wygenerowany raport rezultatu w obszarze dystansu odległości (*DISTANCE SUMMARY*) umożliwił zestawienie - zwane indeksem dla okresu t w układzie rosnącym współczynników *CRS TE* oraz *VRS TE* dla roku 2016:

Powiat	3	10	1	5	7	9	2	4	6	8	11	12
<i>CRS TE</i>	0,786	0,864	0,875	0,907	0,919	0,932	1,000	1,000	1,000	1,000	1,000	1,000

Powiat	3	10	1	2	4	5	6	7	8	9	11	12
<i>VRS TE</i>	0,825	0,867	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000

Warto w tym miejscu skupić uwagę na interpretacji graficznej współczynników ogólnie nazywanych *CRS DEA* i *VRS DEA*. W tym celu dla okresu 15. (2016) najpierw obliczymy współczynniki korelacji (*R*) między zmienną objaśnianą *Mio*, a zmiennymi objaśniającymi (*Em*, *M*, *R*, *U*, *Z*), korzystając z funkcji *Excela*:

=WSP.KORELACJI(\$C\$2:\$C\$13;D2:D13).

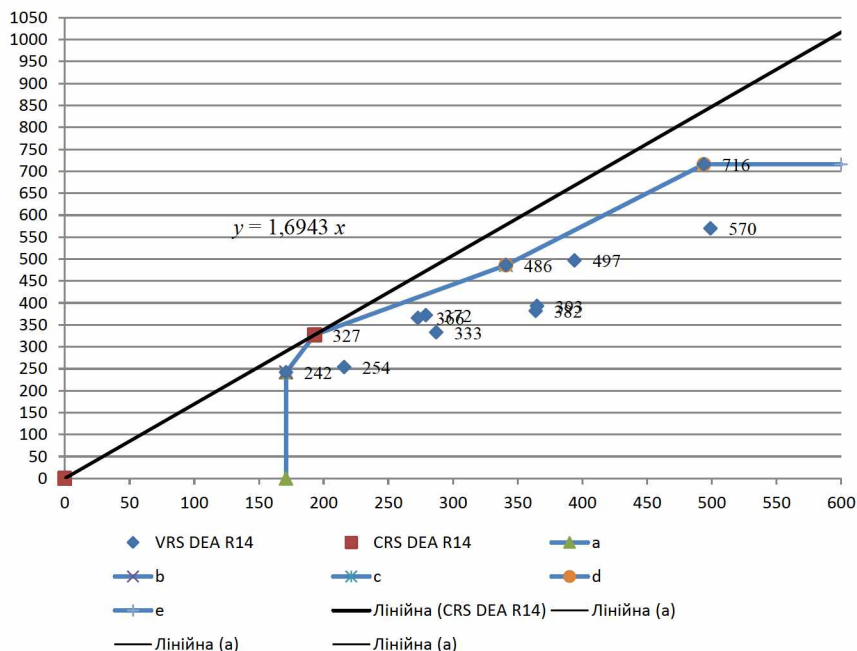
K	L	M	N	O	P
	Em	M	R	U	Z
Mio	0,29	0,88	0,82	0,87	0,77

Widzimy, że najwyższy $R = 0,88$ ma relacja ($Mio \rightarrow M$). W kolejności naszego działania, w wyniku testowania przykładowych okresów 13, 14, 15 okazało się, że najlepsze zobrazowanie graficzne uproszczonych izokwant zwanych według metod *CRS DEA* i *VRS DEA* występuje dla okresu 14 (zob. rysunek 8.10).

Na osi odciętej jako x występuje zmienna M , a na osi rzędnej jako y wartości Mio . Linia łamana odcinkowa obrazująca kształtowanie się współczynnika według *VSE DEA* łączy skrajne punkty wyznaczone przez współrzędne (x , y). Linie pomocnicze (a-e) wyznaczono korzystając z wartości:

x	171	171	193	341	494	600
y	0	242	327	486	716	716

Linie regresji $y = 1,6943 x$ wyznaczono w *Excelu* jako łącząca punkty: (0, 0); (327, 193).



Źródło: Opracowanie własne w *Excelu*.

Rys. 8.10. Interpretacja graficzna budowy uproszczonych izokwant metod *CRS DEA* i *VRS DEA* dla okresu 14

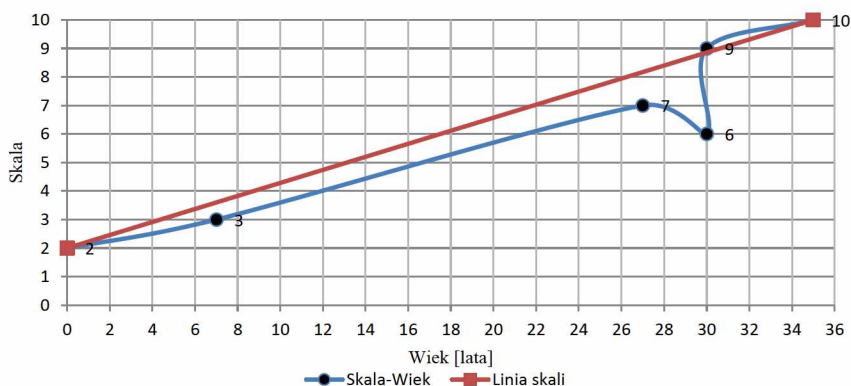
8.6. Sugestia wyznaczenia wag znaczeń społecznych

Przyjmijmy „znaczenie społeczne” cech w skali (1:10). Przyjęcie dla wszystkich obiektów – powiatów jednakowych wartości znaczeń (wag, które mogą być cenami jednostkowymi) nie wpływa na zmianę ustawienia indeksu np. *EFFCH (MALMQUIST INDEX SUMMARY OF FIRM MEANS)*. Zachodzi zatem potrzeba ich rozróżnienia dla poszczególnych powiatów zapisując je po zmiennych oraz stosując model (*COST - CRS DEA*), czyli standardowo zwany efektywność kosztów:

Cecha	U	Z	M	R	Em	Ems
y	2	3	7	6	9	10
x	0	7	27	30	30	35

Gdzie: y - skala [1-10], x - wiek [lata], Ems - emigracja na stałe (założenie teoretyczne).

Podane założenie możemy przedstawić ogólnie na wykresie (zob. rysunek 8.11) w zależności od przyjętego intuicyjnie wieku osoby w ramach danej cechy pobranej ze zbioru danych statystycznych *Baza Demografia*. Dla urodzeń chłopców wiek jest zero, a waga 2, zgony dotyczą chłopców w wieku (0-14), przyjęto więc 7 z wagą 3. Wiek zawieranych małżeństw (przedział: 20-34) dla celów testowania opcji (*COST - CRS DEA*) programu *DEAP* przyjęto 27. Natomiast wiek rozwiedzionych (zameężnych/zameężne) oraz wyjeżdżających na emigrację około 30. Skalę 10 zarezerwowano dla tych, którzy wyjeżdżają na stałe na emigrację przy około 30 latach.



Źródło: Opracowanie własne w *Excelu*.

Rys. 8.11. Zobrazowanie graficzne skali od wieku dla zmiennych: U, Z, M, R, Em i Ems

Dla rozróżnienia poszczególnych powiatów rozpatrzmy przykład dla roku 2016 przy dwunastu powiatach, zmiennych (*Mio*, *Em*, *M*, *R*, *U*, *Z*) i rozwiązując zagadnienie klasyfikacji opcją (*COST - CRS DEA*). Dla celu niniejszego opracowania wprowadźmy pojęcie „znaczenie informacyjne”, dające analogie do standardowej w *DEAP* „ceny informacyjnej” dla cech (zmiennych), dążąc do wyznaczenia wag:

$$w = \text{skala} \times (\text{wartość cechy} / \text{suma cechy dla okresu}).$$

Przykładowo dla powiatu brzeskiego, zmiennej *Em* i skali 9, obliczone znaczenie społeczne będzie następujące (zob. komórka H3 na rysunku 8.12):

$$w_b = \text{skala} \times (Em / \text{Suma } Em) = 9 \times (18 / 415) = 0,39.$$

Pełny efekt naszej pracy w *Excelu* dla poszczególnych powiatów województwa opolskiego i roku 2016 przedstawiono na rysunku 8.12.

Po naciśnięciu klawisza „ENTER” w pliku wyjściowym pojawia się bardzo obszerne wyszczególnienie obliczenia współczynników nieefektywności, z którego zademonstrowano tylko fragment w układzie sumarycznym, dotyczącym poszczególnych powiatów województwa opolskiego. Program *DEAP* zestawia tu trzy typy nieefektywności (zob. rysunek 8.15):

CRS TE - nieefektywność techniczna według modelu *CRS*,

CRS AE - nieefektywność alokacji (*CRS CE/CRS TE*),

CRS CE - nieefektywność według skali – ceny informacyjnej (kosztowej).

Przykładowo dla powiatu 1. $CRS AE = 0,855/0,875 = 0,977$.

Tab. 8.6. Fragment pliku (*eg8-out.txt*) nieefektywność gospodarowania zasobami ludności w powiatach województwa opolskiego

Powiat	<i>CRS TE</i>	<i>CRS AE</i>	<i>CRS CE</i>
1	0.875	0.977	0.855
2	1.000	1.000	1.000
3	0.786	0.669	0.526
4	1.000	0.815	0.815
5	0.907	0.741	0.672
6	1.000	0.732	0.732
7	0.919	0.893	0.820
8	1.000	0.000	0.000
9	0.932	0.684	0.637
10	0.864	0.885	0.765
11	1.000	0.434	0.434
12	1.000	0.732	0.732
Średnia:	0.940	0.714	0.666

Źródło: Opracowanie własne w programie *DEAP*.

* * *

Posługiwanie się programem *DEAP* w wersji 2.1 wymaga znajomości metod modelowania obliczeń różnych współczynników oraz przyjętych uproszczeń w wyznaczaniu rozwiązań zbliżonych do teoretycznie optymalnych. Zamierzeniem moim było zachęcenie Czytelnika do głębszego spenetrowania opcji tego programu. Ponadto zachęcam do coraz szerszego stosowania *DEAP* do nowych obszarów branżowych klasyfikowania obiektów, przy dążeniu do określenia efektywności, czy też nieefektywności ich działania w przedziale czasowym. W przedstawionym opracowaniu celowo nie podjęto się interpretacji zamieszczonych współczynników, aby zachęcić do tego innych badaczy.

Tab. 8.1. Baza danych panelowych do oceny nieefektywności gospodarowania zasobami ludności powiatów województwa opolskiego

Okres	P	<i>Mio</i>	<i>Em</i>	<i>M</i>	<i>R</i>	<i>U</i>	<i>Z</i>	Nazwa powiatu	Rok
1	1	564	27	363	65	426	13	Brzeski	2002
1	2	334	43	208	13	230	6	Głubczycki	2002
1	3	398	285	373	46	391	12	Kędzierzyńsko-kozielski	2002
1	4	414	70	260	20	305	7	Kluczborski	2002
1	5	321	272	247	28	291	4	Krapkowicki	2002
1	6	256	14	204	19	222	4	Namysławski	2002
1	7	862	36	461	65	640	12	Nyski	2002
1	8	359	111	281	15	298	4	Oleski	2002
1	9	628	178	467	44	525	10	Opolski	2002
1	10	274	121	221	23	254	6	Prudnicki	2002

1	11	362	309	307	26	339	4	Strzelecki	2002
1	12	617	44	458	104	447	12	Opole - miasto na prawach powiatu	2002
2	1	500	17	379	66	448	3	Brzeski	2003
2	2	255	75	220	19	201	4	Głubczycki	2003
2	3	427	230	378	72	409	11	Kędzierzyńsko-kozielski	2003
2	4	321	54	251	31	302	5	Kluczborski	2003
2	5	311	207	263	29	296	2	Krapkowicki	2003
2	6	287	17	195	20	182	4	Namysławski	2003
2	7	786	24	601	77	608	6	Nyski	2003
2	8	323	109	273	20	269	8	Oleski	2003
2	9	552	291	507	49	482	9	Opolski	2003
2	10	233	155	243	38	237	5	Prudnicki	2003
2	11	349	244	332	30	315	7	Strzelecki	2003
2	12	683	23	409	117	473	8	Opole - miasto na prawach powiatu	2003
3	1	475	9	371	65	457	6	Brzeski	2004
3	2	326	34	197	25	239	4	Głubczycki	2004
3	3	418	201	373	80	405	5	Kędzierzyńsko-kozielski	2004
3	4	354	45	306	42	299	4	Kluczborski	2004
3	5	322	150	245	43	275	5	Krapkowicki	2004
3	6	260	13	218	25	201	1	Namysławski	2004
3	7	768	13	564	95	628	11	Nyski	2004
3	8	323	91	256	25	244	4	Oleski	2004
3	9	636	308	451	48	484	3	Opolski	2004
3	10	319	102	220	30	270	5	Prudnicki	2004
3	11	326	257	312	46	325	5	Strzelecki	2004
3	12	638	26	439	130	496	8	Opole - miasto na prawach powiatu	2004
4	1	519	14	382	74	454	7	Brzeski	2005
4	2	324	29	233	22	217	6	Głubczycki	2005
4	3	367	262	421	88	430	10	Kędzierzyńsko-kozielski	2005
4	4	372	67	289	39	283	2	Kluczborski	2005
4	5	316	211	302	50	296	5	Krapkowicki	2005
4	6	244	11	226	37	215	1	Namysławski	2005
4	7	702	30	590	122	605	8	Nyski	2005
4	8	320	156	312	34	284	5	Oleski	2005
4	9	581	264	537	66	489	16	Opolski	2005
4	10	250	110	245	31	248	4	Prudnicki	2005
4	11	365	215	325	40	312	1	Strzelecki	2005
4	12	645	74	414	135	495	11	Opole - miasto na prawach powiatu	2005
5	1	654	79	451	78	476	13	Brzeski	2006
5	2	333	34	268	28	225	3	Głubczycki	2006
5	3	399	213	469	93	419	7	Kędzierzyńsko-kozielski	2006
5	4	431	91	363	57	302	5	Kluczborski	2006
5	5	355	128	264	65	284	7	Krapkowicki	2006
5	6	296	30	239	26	221	1	Namysławski	2006
5	7	963	131	695	144	613	12	Nyski	2006
5	8	403	136	317	35	277	3	Oleski	2006
5	9	648	299	540	86	515	7	Opolski	2006
5	10	322	147	271	43	256	1	Prudnicki	2006
5	11	369	301	355	39	325	4	Strzelecki	2006
5	12	753	125	523	135	533	9	Opole - miasto na prawach powiatu	2006
6	1	683	44	516	80	479	11	Brzeski	2007
6	2	377	29	280	26	200	6	Głubczycki	2007

6	3	477	231	451	80	411	6	Kędzierzyńsko-kozielski	2007
6	4	496	78	289	41	320	4	Kluczborski	2007
6	5	408	155	308	42	262	6	Krapkowicki	2007
6	6	323	26	244	22	204	5	Namysłowski	2007
6	7	882	104	755	101	635	10	Nyski	2007
6	8	441	441	339	35	298	3	Oleski	2007
6	9	702	295	521	57	503	3	Opolski	2007
6	10	329	169	253	41	264	3	Prudnicki	2007
6	11	435	287	388	41	332	6	Strzelecki	2007
6	12	855	81	537	117	492	6	Opole - miasto na prawach powiatu	2007
7	1	587	33	578	54	546	7	Brzeski	2008
7	2	315	43	278	26	263	4	Głubczycki	2008
7	3	406	178	469	60	457	3	Kędzierzyńsko-kozielski	2008
7	4	377	135	356	43	275	6	Kluczborski	2008
7	5	308	123	324	46	273	9	Krapkowicki	2008
7	6	269	14	317	28	249	4	Namysłowski	2008
7	7	867	100	801	81	666	19	Nyski	2008
7	8	373	101	324	27	283	4	Oleski	2008
7	9	632	238	606	67	582	12	Opolski	2008
7	10	280	111	287	32	261	3	Prudnicki	2008
7	11	377	253	365	39	293	4	Strzelecki	2008
7	12	701	97	562	103	542	10	Opole - miasto na prawach powiatu	2008
8	1	575	28	442	65	528	5	Brzeski	2009
8	2	350	25	293	21	235	6	Głubczycki	2009
8	3	428	111	462	65	440	6	Kędzierzyńsko-kozielski	2009
8	4	423	60	347	41	332	1	Kluczborski	2009
8	5	357	66	333	45	291	9	Krapkowicki	2009
8	6	300	20	247	33	241	3	Namysłowski	2009
8	7	812	41	758	68	689	17	Nyski	2009
8	8	370	61	332	26	399	6	Oleski	2009
8	9	640	141	606	68	558	3	Opolski	2009
8	10	308	79	277	29	284	5	Prudnicki	2009
8	11	370	147	365	36	340	9	Strzelecki	2009
8	12	750	55	571	102	586	7	Opole - miasto na prawach powiatu	2009
9	1	596	30	440	66	479	6	Brzeski	2010
9	2	349	22	264	29	240	2	Głubczycki	2010
9	3	401	104	403	62	422	8	Kędzierzyńsko-kozielski	2010
9	4	391	33	310	40	359	2	Kluczborski	2010
9	5	388	69	308	40	277	4	Krapkowicki	2010
9	6	296	16	215	17	226	1	Namysłowski	2010
9	7	881	48	679	67	668	10	Nyski	2010
9	8	422	48	348	26	513	6	Oleski	2010
9	9	691	98	621	56	597	2	Opolski	2010
9	10	283	68	284	28	270	8	Prudnicki	2010
9	11	409	144	377	36	356	5	Strzelecki	2010
9	12	697	37	500	75	612	8	Opole - miasto na prawach powiatu	2010
10	1	606	25	388	58	465	7	Brzeski	2011
10	2	383	28	222	21	200	1	Głubczycki	2011
10	3	427	134	386	62	376	4	Kędzierzyńsko-kozielski	2011
10	4	394	48	287	32	277	2	Kluczborski	2011
10	5	319	108	290	31	287	3	Krapkowicki	2011
10	6	291	12	201	32	216	2	Namysłowski	2011

10	7	764	65	610	101	652	9	Nyski	2011
10	8	390	60	341	29	283	1	Oleski	2011
10	9	638	163	575	85	568	5	Opolski	2011
10	10	283	89	267	21	272	3	Prudnicki	2011
10	11	416	155	347	45	296	3	Strzelecki	2011
10	12	686	49	433	74	524	4	Opole - miasto na prawach powiatu	2011
11	1	514	35	399	68	472	6	Brzeski	2012
11	2	327	22	239	22	211	3	Głubczycki	2012
11	3	442	114	416	55	431	5	Kędzierzyńsko-kozielski	2012
11	4	383	74	324	43	307	7	Kluczborski	2012
11	5	355	102	293	37	274	2	Krapkowicki	2012
11	6	263	14	186	23	221	3	Namysłowski	2012
11	7	770	49	622	85	617	6	Nyski	2012
11	8	380	70	319	16	316	3	Oleski	2012
11	9	611	170	533	60	599	8	Opolski	2012
11	10	305	43	232	25	266	3	Prudnicki	2012
11	11	385	130	370	44	366	8	Strzelecki	2012
11	12	702	68	440	73	549	13	Opole - miasto na prawach powiatu	2012
12	1	589	57	316	61	416	6	Brzeski	2013
12	2	366	41	191	33	219	2	Głubczycki	2013
12	3	447	132	367	65	381	1	Kędzierzyńsko-kozielski	2013
12	4	357	75	264	43	287	6	Kluczborski	2013
12	5	319	95	258	36	294	8	Krapkowicki	2013
12	6	322	19	177	35	188	3	Namysłowski	2013
12	7	855	93	541	69	527	5	Nyski	2013
12	8	417	81	280	22	294	3	Oleski	2013
12	9	652	240	481	57	581	6	Opolski	2013
12	10	309	68	203	32	234	2	Prudnicki	2013
12	11	411	202	307	38	366	8	Strzelecki	2013
12	12	708	80	428	83	515	4	Opole - miasto na prawach powiatu	2013
13	1	500	54	338	68	410	8	Brzeski	2014
13	2	354	28	220	27	189	2	Głubczycki	2014
13	3	378	123	356	44	430	5	Kędzierzyńsko-kozielski	2014
13	4	361	43	269	43	306	5	Kluczborski	2014
13	5	332	74	255	30	285	2	Krapkowicki	2014
13	6	295	33	156	23	184	3	Namysłowski	2014
13	7	797	69	515	83	585	9	Nyski	2014
13	8	412	62	284	22	318	2	Oleski	2014
13	9	596	183	515	52	558	7	Opolski	2014
13	10	268	42	225	29	272	5	Prudnicki	2014
13	11	358	140	343	37	355	3	Strzelecki	2014
13	12	549	69	372	89	542	6	Opole - miasto na prawach powiatu	2014
14	1	486	0	341	69	429	3	Brzeski	2015
14	2	327	0	193	27	189	2	Głubczycki	2015
14	3	382	0	364	52	425	6	Kędzierzyńsko-kozielski	2015
14	4	372	0	279	37	296	5	Kluczborski	2015
14	5	333	0	287	30	273	2	Krapkowicki	2015
14	6	242	0	171	32	192	3	Namysłowski	2015
14	7	716	54	494	69	493	5	Nyski	2015
14	8	366	0	273	31	288	4	Oleski	2015
14	9	570	0	499	61	600	8	Opolski	2015
14	10	254	0	216	19	249	3	Prudnicki	2015

14	11	393	0	365	43	366	2	Strzelecki	2015
14	12	497	0	394	79	574	9	Opole - miasto na prawach powiatu	2015
15	1	459	18	325	50	403	5	Brzeski	2016
15	2	276	9	171	26	198	0	Głubczycki	2016
15	3	352	74	322	35	395	6	Kędzierzyńsko-kozielski	2016
15	4	353	7	279	25	306	5	Kluczborski	2016
15	5	292	38	248	24	273	3	Krapkowicki	2016
15	6	208	3	159	33	231	1	Namysławski	2016
15	7	670	39	512	62	563	6	Nyski	2016
15	8	370	16	293	30	289	0	Oleski	2016
15	9	565	70	526	41	558	8	Opolski	2016
15	10	259	23	200	26	240	2	Prudnicki	2016
15	11	373	109	306	23	366	4	Strzelecki	2016
15	12	492	9	362	63	607	10	Opole - miasto na prawach powiatu	2016

Źródło: Opracowanie własne w *Excelu*.

9. Budowa i weryfikacja modelu ekonometrycznego w wybranych programach



9.1. Wstęp

Coraz powszechniejsze staje się stosowanie platformy programowej *R* do zagadnień statystyki i ekonometrii. Z tego względu pokazano możliwość sformułowania poleceniami języka *R* modelu ekonometrycznego bazującego na danych statystycznych zawartych w roczniku statystycznym GUS-u⁷³. Ponadto przeprowadzono podstawowe przykładowe testy weryfikacji zbudowanego modelu. Oparto się na fragmencie obserwacji zmiennych z okresu 1992-2007 dotyczących liczby bezrobotnych (zmienna objaśniana) oraz zmiennych objaśniających przyjętych w modelowaniu tej zmiennej⁷⁴. Szereg zaktualizowano o dane z roczników ostatnich oraz poszerzono o obserwacje dotyczące lat 2008-2011 oraz oznaczono w notacji wygodnej do przetwarzania komputerowego (zob. tabela 9.1).

Tab. 9.1. Wartości obserwacji zmiennych

Rok	<i>t</i>	<i>X3</i>	<i>X9</i>	<i>X10</i>	<i>X11</i>	<i>X12</i>	<i>X13</i>	<i>X14</i>	<i>X25</i>	<i>X38</i>	<i>X48</i>	<i>X49</i>	<i>X50</i>	<i>X51</i>	<i>X59</i>	<i>X78</i>
1992	1	2394	3467,6	823,0	1042,1	794,7	476,7	287,7	295,16	138,9	404,70	1921,9	18430	4626,8	41,1	5,7
1993	2	2595	3391,8	721,9	999,4	746,5	381,1	239,3	384,25	134,8	550,29	1856,5	18206	4504,2	40,1	5,4
1994	3	2375	3361,4	678,8	996,1	728,1	366,4	235,4	515,49	134,5	708,30	1750,0	18206	4472,6	45,8	5,4
1995	4	2233	3461,1	689,2	1078,6	723,4	414,4	243,5	704,99	131,6	878,29	1648,3	18075	4775,1	59,2	5,4
1996	5	1961	3436,0	684,2	1119,7	721,9	425,3	250,5	900,58	126,7	1112,95	1565,0	19224	4258,9	70,5	5,3
1997	6	1737	3433,4	723,5	1244,8	730,7	487,4	273,3	1088,03	121,9	1300,00	1514,7	19771	4437,4	77,4	5,3
1998	7	1827	3378,7	736,9	1321,1	725,2	552,5	253,0	1277,06	115,7	1527,52	1452,6	20288	4243,5	83,3	5,4
1999	8	2641	3138,4	710,4	1318,4	686,3	583,0	286,9	1428,01	112,5	1700,54	1406,5	20800	4138,6	86,4	5,7
2000	9	2760	2955,0	661,9	1325,0	654,9	614,2	286,7	2622,76	111,1	1904,57	1360,3	21037	3750,8	78,4	5,5
2001	10	3186	2820,6	627,8	1295,6	530,4	637,3	252,1	2821,30	108,0	1897,60	1310,4	21376	3905,8	75,1	5,0
2002	11	3431	2670,5	545,5	1291,1	601,9	671,3	260,0	2930,70	102,6	1886,10	1284,3	21659	3708,1	85,7	5,0
2003	12	3329	2639,1	496,4	1308,0	586,3	699,5	261,8	3017,64	104,2	1918,71	1268,6	22052	3877,2	78,4	5,1
2004	13	3230	2663,1	453,1	1295,4	583,0	700,6	264,6	3111,04	104,0	1966,72	1229,1	22804	3779,1	98,0	5,0
2005	14	3045	2665,4	483,6	1360,7	577,7	695,8	265,2	3252,02	103,8	1954,20	1219,4	26565	3855,3	104,3	5,4
2006	15	2344	2714,3	511,5	1394,2	596,7	723,3	267,6	3421,05	104,9	2106,28	1221,5	27547	3967,0	88,8	5,9
2007	16	1619	2842,4	576,8	1482,8	621,0	780,3	277,0	3720,10	107,9	2290,78	1280,1	27713	3807,0	84,7	6,5
2008	17	1211	2890,7	671,2	1633,5	792,0	743,1	406,7	3986,50	110,1	2558,46	1212,2	28425	3792,2	80,6	6,8
2009	18	1411	2725,4	706,6	1611,5	777,5	740,1	407,4	4232,68	105,4	2706,47	1194,7	28684	3873,6	80,6	6,6
2010	19	1699	2696,1	674,4	1598,7	758,2	767,0	412,0	4388,31	103,9	2805,48	1197,9	29774	4237,3	81,5	5,9
2011	20	2146	2714,2	705,2	1580,3	787,5	775,1	412,2	4633,95	105,6	2872,24	1202,0	29383	3857,1	82,0	5,4

Źródło: Opracowanie własne na podstawie roczników statystycznych GUS-u (lata 1992-2011).

Ze względu na zmianę struktury (sekcje i działy) w roczniku 2011 wartości dla lat 2009, 2010, 2011 i zmiennych *X12*-*X14* pogrupowano następująco:

X12 - transport, gospodarka magazynowa i łączność:

- transport i gospodarka magazynowa,
- informacja i komunikacja.

⁷³ Opracowanie bazuje na rozdziale 2.15 w książce autora pt. *Formułowanie modeli ekonometrycznych na potrzeby zarządzania*, Wydawnictwo Instytut Śląski, Opole 2015.

⁷⁴ *Przejawy wielowymiarowości współczesnego zarządzania - formy i instrumenty ekonomiczno-społeczne*, red. naukowa T. Pokusa, artykuł W. Wornalkiewicz *Model ekonometryczny zmiennej „liczba bezrobotnych”*, Wydawnictwo Instytut Śląski, Opole 2011.

X13 - obsługa nieruchomości i firm:

- obsługa rynku nieruchomości,
- zakwaterowanie i gastronomia,
- administrowanie i działalność wspierająca,
- pozostała działalność usługowa.

X14 - działalność usługowa komunalna, społeczna i indywidualna:

- działalność finansowa i ubezpieczeniowa,
- działalność związana z kulturą, rozrywką i rekreacją.

Niniejsze opracowanie nawiązuje również do pracy autora pt. „*Modele ekonometryczne PKB obiektów struktury terytorialnej*”. Zestawiono w nim szeregi czasowe zmiennych objaśniających (X1-X79) przyjętych z poszczególnych rynków do badania ich wpływu na zmienną objaśnianą Y jaką jest produkt krajowy brutto, według cen bieżących w przeliczeniu na jednego mieszkańca⁷⁵. Dane statystyczne zapisano w formie pliku arkusza kalkulacyjnego *Excel*, skąd zostały importowane do programu *Gretl* jako plik *Korelacja.gdt*. Plik ten po rozszerzeniu obserwacji o lata 2008-2011 w zakresie zmiennych X3, X9-X14, X25, X38, X48-X51, X59, X78 o nazwie *Korelacja-roz-X3.gdt* stanowił bazę do wstępnego sformułowania zmodyfikowanego modelu zmiennej X3 - „liczba bezrobotnych” z zastosowaniem programu *Gretl*⁷⁶. Dla ciągłości rozważań w tabeli 2.1 zachowano symbole wybranych zmiennych, a ponadto dla ewentualnego sprawdzenia wartości zmiennych wymieniono także „ścieżki” dojścia do danych w Rocznikach Statystycznych Rzeczypospolitej Polskiej⁷⁷:

- rynek pracy/aktywność ekonomiczna ludności w wieku 15 i więcej - przeciętne w roku tys./bezrobotni (X3);
- przeciętne zatrudnienie w tys.: przemysł (X9), budownictwo (X10), handel i naprawy (X11), transport, gospodarka magazynowa i łączność (X12), obsługa nieruchomości i firm (X13), działalność usługowa komunalna, społeczna i indywidualna, pozostałe (X14);
- koszty prac w zł/ogółem na jednego zatrudnionego przeciętnie miesięcznie (X25);
- rynek płacy/wynagrodzenia w mln/wskaźnik - przeciętne miesięczne wynagrodzenie nominalne brutto - rok poprzedni = 100 (X38);
- budżety gospodarstw domowych/ przeciętne miesięczne wydatki na 1 gospodarstwo domowe w zł (X48),
- rynek usług komunalnych (w odniesieniu do gospodarstw domowych): zużycie wody z wodociągów w ciągu roku w hm³ (X49), zużycie energii elektrycznej w ciągu roku w GWh (X50), zużycie gazu w sieci w ciągu roku w hm³ (X51),
- rynek edukacji i wychowania (szkoły, uczniowie i studenci w dziesiątkach)/absolwenci - policealne w tys. (X59),
- wskaźnik zawartych małżeństw na 1000 ludności (X78).

9.2. Budowa modelu ekonometrycznego w *Gretl*

Dla sformułowania modelu ekonometrycznego „*liczby bezrobotnych*” - po rozszerzeniu bazy danych o wartości z lat 2008-2011 do 20. obserwacji, zapisanych jako plik *Gretl*-a *Korelacja-roz-X3.gdt* wywołujemy menu: *Model/klasyczna metoda najmniejszych kwadratów*. Następnie wskazujemy zmienną zależną (objaśnianą) X3 oraz zmienne objaśniające X9-X14, X25, X38, X48-X51, X59,

⁷⁵ Wornalkiewicz W., *Modele ekonometryczne PKB obiektów struktury terytorialnej*, tab. 3.3, Wydawnictwo Instytut Śląski, Opole 2013.

⁷⁶ Początkowy model zaproponował autor w artykule *Model ekonometryczny zmiennej „liczba bezrobotnych”* w ramach monografii *Przejawy wielowymiarowości współczesnego zarządzania - formy i instrumenty ekonomiczno-społeczne*, op. cit. s. 172.

⁷⁷ Ibidem.

X78. Rezultatem pracy programu jest raport *Modelu1* zamieszczony w tabeli 9.2 obejmujący estymację parametrów modelu zwanych współczynnikami.

Tab. 9.2. Oszacowanie parametrów modelu X3 dla 20 obserwacji (Model 1)

Zmienna	Współczynnik	Błąd stand.	Statystyka t	Wartość p
const	11037,4	6207,87	1,7780	0,13555
X9	-5,18169	2,50564	-2,0680	0,09348
X10	2,30576	2,81819	0,8182	0,45048
X11	2,05037	4,56698	0,4490	0,67225
X12	-0,983977	3,34782	-0,2939	0,78063
X13	0,599386	2,6107	0,2296	0,82751
X14	-2,87111	8,04704	-0,3568	0,73581
X25	-0,663503	0,506457	-1,3101	0,24712
X38	82,6385	79,3045	1,0420	0,34513
X48	0,250775	0,882549	0,2841	0,78769
X49	0,143198	3,20708	0,0447	0,96611
X50	-0,0689788	0,0748771	-0,9212	0,39920
X51	-0,191685	0,423229	-0,4529	0,66959
X59	14,6661	15,441	0,9498	0,38582
X78	-604,484	302,113	-2,0009	0,10183

Źródło: Opracowanie własne w programie *Gretl*.

Ponadto otrzymujemy również statystykę opisową obejmującą cechy:

- średnia arytmetyczna zmiennej zależnej = 2358,7;
- odchylenie standardowe zmiennej zależnej = 664,528;
- suma kwadratów reszt = 205105;
- błąd standardowy reszt = 202,536;
- współczynnik determinacji $R^2 = 0,975555$;
- skorygowany współczynnik determinacji = 0,907108;
- statystyka $F(14, 5) = 14,2527$ (wartość $p = 0,00419$);
- logarytm wiarygodności = -120,734;
- kryterium informacyjne Akaike'a = 271,468;
- kryterium bayesowskie Schwarz'a = 286,404.

Na zakończenie raportu występuje podpowiedz, którą ze zmiennych usunąć dla poprawy modelu: „Wyłączając stałą największa wartość p jest dla zmiennej X49”. Dalej w kolejnych iteracjach kierując się podpowiedziami *Gretl*-a usuwamy kolejno X49, X48, X13, X12, X51, X11, X14, X59, X38, X10, X50, czyli zmienne dla których prawdopodobieństwo popełnienia błędu jest duże. W efekcie końcowym procesu iteracyjnego uzyskano następujący model.

Tab. 9.3. Oszacowanie końcowe parametrów modelu

Zmienna	Współczynnik	Błąd stand.	Statystyka t	Wartość p
const	16749	1061,19	15,7832	<0,00001 ***
X9	-3,317	0,369783	-8,9701	<0,00001 ***
X25	-0,668989	0,0910932	-7,3440	<0,00001 ***
X78	-501,059	107,708	-4,6520	0,00027 ***

Źródło: Opracowanie własne w programie *Gretl*.

Zwróćmy uwagę na wybrane elementy statystyki opisowej wskazujące na dobre dopasowanie modelu do danych empirycznych, a mianowicie:

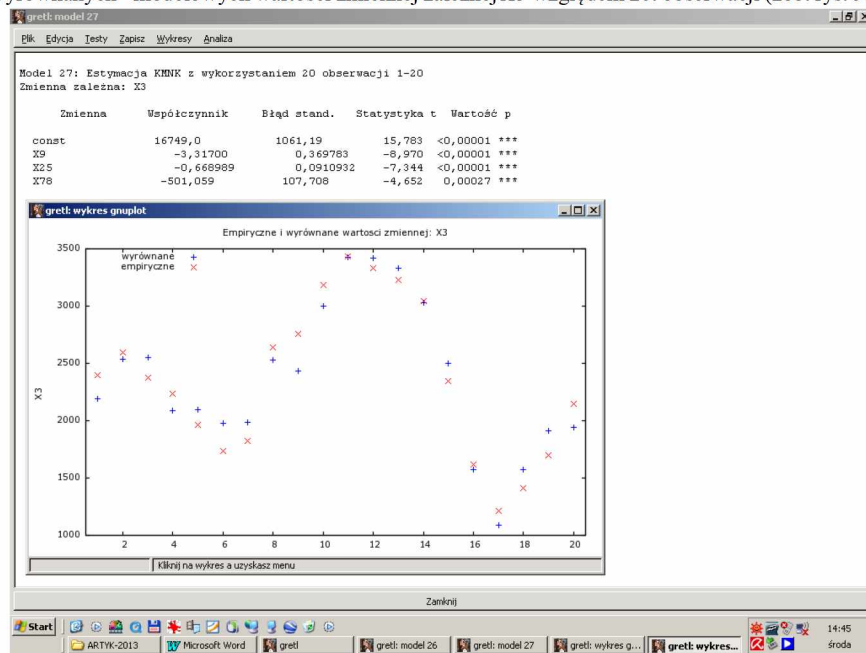
- suma kwadratów reszt = 519489;
- błąd standardowy reszt = 180,189;
- współczynnik determinacji $R^2 = 0,938085$
- skorygowany współczynnik determinacji = 0,926476;
- statystyka $F(3, 16) = 80,8061$ (wartość $p < 0,00001$);
- logarytm wiarygodności = -130,027;
- kryterium informacyjne Akaike'a = 268,055;
- kryterium bayesowskie Schwarz'a = 272,038.

W tabeli 9.2 oraz 9.3 warto zauważyć graficzne informowanie przez *Gretl* - w postaci gwiazdek - o poziomie α istotności danego parametru: *** $\rightarrow \alpha = 0,01$; ** $\rightarrow \alpha = 0,05$; * $\rightarrow \alpha = 0,1$; puste $\rightarrow \alpha > 0,1$. Tak więc występuje niski poziom istotności - popełnienia błędu w oszacowaniu parametrów oraz wysoki skorygowany współczynnik determinacji $\bar{R}^2 = 0,926476$.

Wprowadzenie dodatkowych czterech dalszych obserwacji wpłynęło na istotną zmianę struktury modelu liniowego wielowymiarowego w stosunku do wcześniejszego przy 16. obserwacjach, który w tabeli cytowanego artykułu był następujący⁷⁸:

$$X_3^* = 17753,3 - 4,33449X_9 + 3,19675X_{10} - 2,29688X_{13} - 0,88772X_{48} - 0,116366X_{50} + 16,1492X_{59}$$

Warto tu przypomnieć zmienne występujące w naszym rozszerzonym modelu końcowym (zob. tabela 9.3): bezrobotni (X_3), przeciętne zatrudnienie w tysiącach: przemysł (X_9), koszty pracy ogółem na 1. zatrudnionego - przeciętne miesięczne (X_{25}), ilość zawartych małżeństw na 1000 ludności (X_{78}). Przy aktywnym modelu program *Gretl* umożliwia wykonanie wykresu empirycznych oraz wyrównanych - modelowych wartości zmiennej zależnej X_3 względem 20. obserwacji (zob. rys. 9.1).



Źródło: Opracowanie własne.

Rys. 9.1. Widok ekranu z raportem po estymacji parametrów oraz wykresem graficznym wartości empirycznych i modelowych

⁷⁸ Artykuł autora: *Model ekonometryczny zmiennej „liczba bezrobotnych”*, op. cit. s. 178.

9.3. Przygotowanie do realizacji w programie R

Procedura postępowania mająca na celu sformułowanie pliku danych w postaci tzw. tabeli danych do potrzeb budowy modelu ekonometrycznego przykładowej zmiennej X_3 jest następująca:

1. Skopiowanie tabeli 9.1 zamieszczonej w niniejszej pracy, a zapisanej w formacie *Worda* do skoroszytu *Excels* i nadanie mu nazwy *Dane.xls* umieszczając go na przykładowej ścieżce *D:\AR-TYK-2013*.

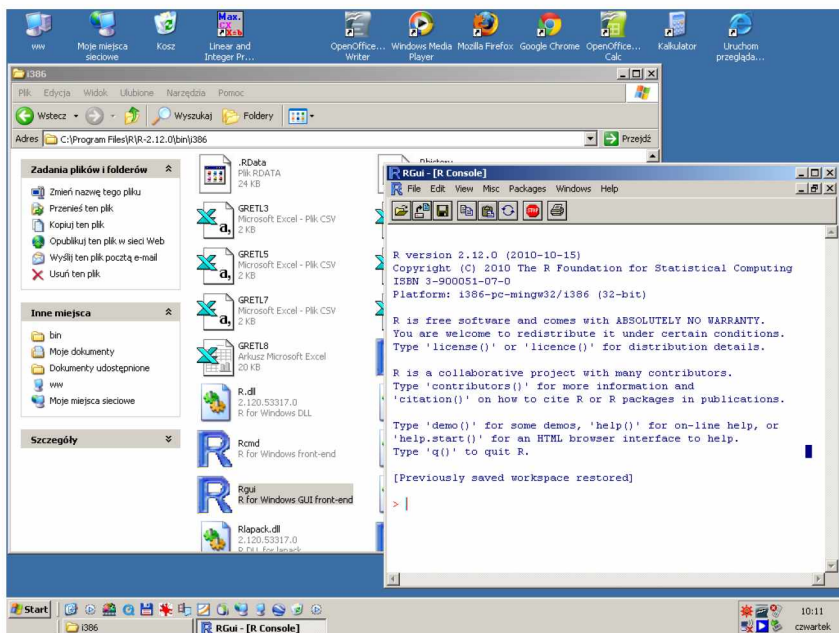
2. Pozostawienie tylko jednego arkusza w ramach skoroszytu i nadanie mu nazwy „Dane-R”.

3. Zmniejszenie szerokości kolumn szeregów danych ze standardowej 8.43 na 7, celem zmniejszenia danych na ekranie.

4. Sformatowanie zmiennych X_9 - X_{11} , X_{38} , X_{48} , X_{49} , X_{51} , X_{59} , X_{78} z dokładnością do jednego miejsca po przecinku a zmienną X_{25} dwa miejsca po przecinku. Zabieg ten wynika z faktu występowania supresji zer po skopiowaniu tabeli danych z *Worda* do *Excels*.

5. Usunięcie kolumn „Rok” i „l”.

Mając już przygotowane dane przystępujemy do zainicjowania środowiska platformy programowej R zawierającej obszerną strukturę programów⁷⁹.



Źródło: Opracowanie własne w programie R.

Rys. 9.2. Obraz ekranów wywołania programu *Rgui* trybu okienkowego

Postępowanie w tym zakresie sprowadza się do pozyskania aktualnej wersji szeroko rozumianego programu R z Internetu i zapisanie go na ścieżce *D:\Program Files\R\R-2.12.0*⁸⁰. W wyniku instalacji utworzone zostały podkatalogi: *bin* - z plikami wykonywalnymi; *doc* - zawierający doku-

⁷⁹ Charakterystykę tego środowiska opisano w książce: *Wstęp do ekonometrii i badań operacyjnych ...*, wydanie II rozszerzone, Womalkiewicz W., "Wybrane modele ekonometryczne", Wydawnictwo Instytut Śląski, Opole 2012, rozdz. 1.16.

⁸⁰ Wersja R-2-12.0 była aktualna w 2012 roku w trakcie pisania wydania II rozszerzonego książki „Wstęp do ekonometrii i badań operacyjnych ...”, op. cit.

mentację pakietu *R*; *library* - biblioteka z pakietami programowymi, które można później wywołać funkcją: *install.packages*.

Następnie przystępujemy do otwarcia podkatalogów *bin* i w kolejności *i386*. Tryb okienkowy wywołujemy otwierając program *Rgui.exe*. Zwróćmy uwagę na zamieszczoną rysunku 2.2, z widocznymi fragmentami ekranów pulpitu, wywołanego trybu okienkowego *Rgui* (*R Console*) ze znakiem zachęty „>” do wprowadzania komend w języku *R*.

Jednak wygodniejszy w początkującym kontakcie ze środowiskiem programowym *R* jest tryb nadzorowany zwany także windowsowym. Instalujemy go poleceniem - komendą: *library(Rcmdr)*. Powoduje to przywołanie pakietu *Rcmdr* o ile jest on już wcześniej w bibliotece. Jeśli nie występuje to doinstalowujemy go z Internetu poleceniem: *install.packages („Rcmdr”,dependencies=TRUE)*. Opcja „*dependencies=TRUE*” oznacza pobranie również podpakietów niezbędnych do działania *Rcmdr*⁸¹. Alternatywnym rozwiązaniem jest skorzystanie z trybu okienkowego *Rgui* i przywołanie programu *Rcmdr* poprzez menu: *Packages/install package(s)*. Obraz ekranu trybu windowsowego z programem *Rcmdr* określony jak *R Commander* pokazano na rysunku 9.3. Zwróćmy uwagę na menu, które obejmuje między innymi grupy funkcji:

File - otwarcie i zapisanie zbioru poleceń tzw. zbioru skryptowego;

Edit - kopiowanie/wklejanie;

Data - założenie zbioru, załadowanie istniejącego, import z....;

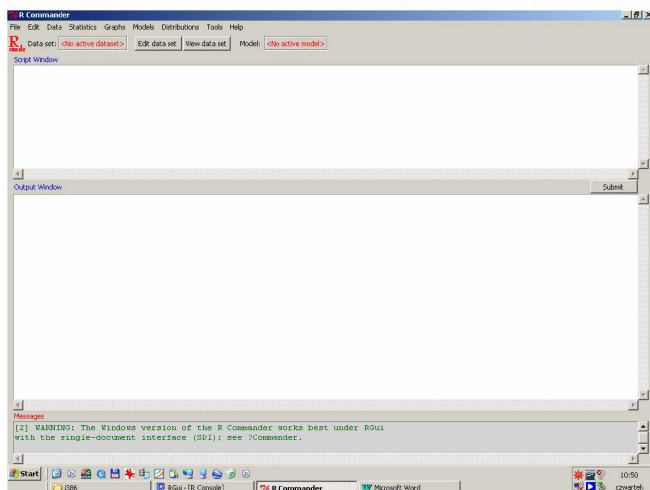
Statistics - podstawowa statystyka opisowa, testy sumaryczne.

Graphs - możliwość wykonania wykresów różnego typu dla pojedynczej zmiennej jak i grupowych;

Models - analiza regresji oraz *Distributions* - tablice statystyczne, *Tools* - ładowanie pakietu i różne opcje;

Help - opis do poleceń programowych.

Na rysunku 9.3 widzimy trzy podokna opisowe: *Script Window* - skryptowe (tekstów poleceń), *Output Window* - realizacji - raportów programu na daną sekwencję poleceń, *Messages* - komunikatów - informacji o błędach w trakcie rozpoznania składni poleceń a następnie ich realizacji.



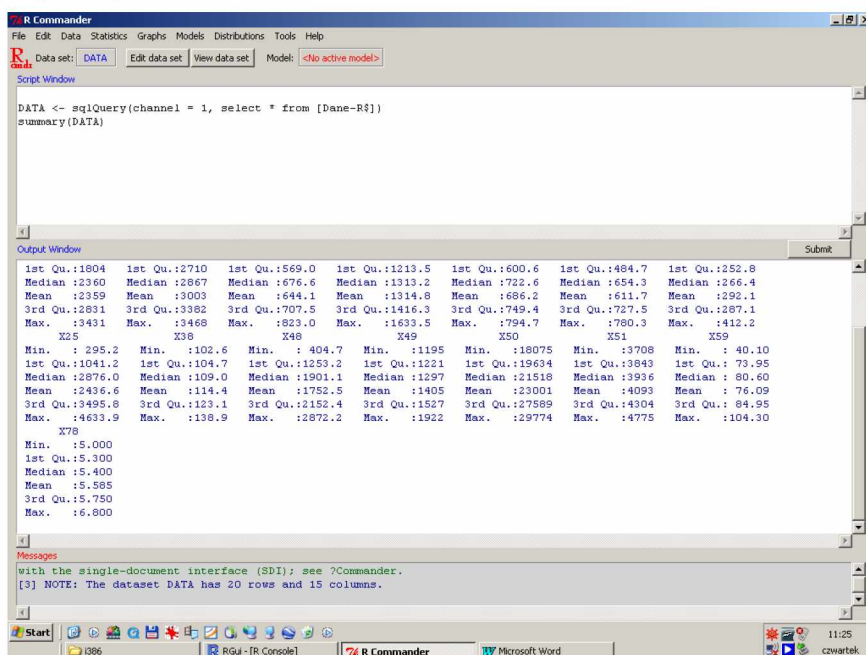
Źródło: Opracowanie własne korzystając ze środowiska programowego *R*.

Rys. 9.3. Widok ekranu trybu windowsowego

⁸¹ Ibidem, s. 377.

Jesteśmy teraz w trybie windowsowym o dogodnym dla użytkownika interfejsie. Umożliwia to nam dokonanie importu naszego pliku danych *Dane.xls* z Excela umieszczonego w podkatalogu *D:\ARTYK-2013* korzystając z menu: *Data/Import data/from Excel, Access or dBase data set ...* Określiśmy nazwę obiektu jako *DATA* i OK. Pojawia się okno „Otwieranie” i na ścieżce *D:\ARTYK-2013* wskazujemy do pobrania plik *Dane.xls*. W pierwszym oknie tj. skryptowym pojawia się polecenie w języku *R*, gdyż program dokonuje konwersji naszych „ruchów” w menu na sekwencję języka *R*: *DATA <- sqlQuery(channel = 1, select * from [Dane-R\$])*.

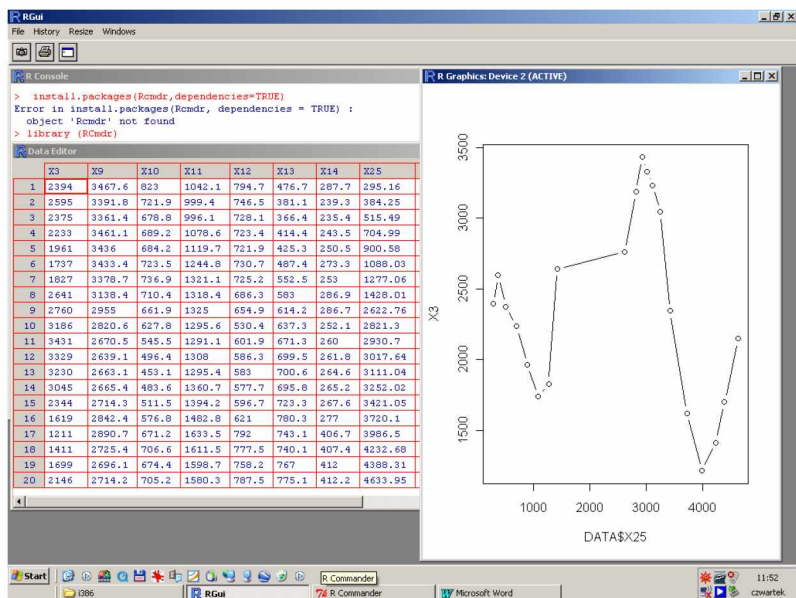
Odpowiedzią programu w oknie wyjścia *Output Window* jest potwierdzenie. Natomiast w oknie komunikatów jest informacja, że traktowany jako tabela danych, importowany plik *Dane.xls* zawiera 20 wierszy (oprócz nagłówka) i 15 kolumn. Po zakończeniu importu mamy aktywną tabelę danych i przystępujemy przykładowo do sporządzenia raportu sumarycznego podstawowych statystyk opisowych naszych 15. zmiennych. korzystając z menu: *Statistics/Summaries/Active dataset*. Odpowiada to poleceniu: *summary(DATA)*. Raport z realizacji tego polecenia ukazał się w oknie *Output Window* (zob. rysunek 9.4). Widzimy tu obliczone dla każdej zmiennej wartości z 20. obserwacji: *Min.* - minimum, *1 st Qu.* - kwartyl pierwszy, *Median* - mediana, *Mean* - średnia arytmetyczna, *3 rd Qu.* - kwartyl trzeci, *Max.* - maksimum.



Źródło: Opracowanie własne korzystając ze środowiska programowego *R*.

Rys. 9.4. Podstawowa statystyka opisowa

Zwróćmy uwagę, że wartości poszczególnych zmiennych z przecinkiem zamienia program na kropkę dziesiętną. Będąc w trybie windowsowym sporządzmy przykładowo wykres liniowy zależności zmiennych *X3 ~ X25*. Ukazano go na tle fragmentu tabeli danych oraz widoku ekranu okna trybu okienkowego *Rgui* (zob. rysunek 9.5), a realizację tę umożliwia menu: *Graphs/Line graph...*



Źródło: Opracowanie własne w programie R.

Rys. 9.5. Podstawa statystyka opisowa i wykres zależności zmiennych $X3 \rightarrow X25$.

Przystąpmy teraz do oszacowania parametrów modelu regresji liniowej wielowymiarowej, poprzez wcześniejsze uaktywnienie tabeli danych przyciskiem DATA (zob. rysunek 9.4), a następnie wywołanie menu: *Statistics/Fit models/Linear model* ...Pojawia się okno dialogowe w którym określamy ogólnie składnię argumentu „formula funkcji lm dla modelu”⁸²:

W odniesieniu do trzech zmiennych naszego końcowego modelu uzyskanego programem *Gretl* model przyjmuje postać:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + \dots$$

$$\hat{X}_3 = \hat{Y} = b_0 + b_1X_9 + b_2X_{25} + b_3X_{78}$$

Program *R* w oknie *Output Windows* dla podprogramu lm pokazuje sposób określenia opcji formula, określa statystykę opisową reszt modelu zmiennej zależnej $X3$, a ponadto oszacowanie parametrów modelu liniowego ze zmiennymi $X9$, $X25$, $X78$. Sposób zapisu wartości jest w notacji inżynierskiej, przykładowo: $1.675e+04 = 1,675 \cdot 10^4 = 16750$ (zob. tabela 9.4).

Tab. 9.4. Oszacowanie parametrów dla trzech zmiennych modelu końcowego w programie *R* i dodatkowe informacje wejściowe oraz podstawowa statystyka opisowa

lm(formula = X3 ~ X9 + X25 + X78 , data = DATA)					
Residuals:					
Min	1Q	Median	3Q	Max	
-239.89	-155.26	12.21	128.91	323.18	
Coefficients:					
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.675e+04	1.061e+03	15.783	3.55e-11	***

⁸² Walesiak M., Gatnar E. (red. nauk.), *Statystyczna analiza danych z wykorzystaniem programu R*, Wydawnictwo Naukowe PWN, Warszawa 2009, s. 143.

X9	-3.317e+00	3.698e-01	-8.970	1.22e-07	***
X25	-6.690e-01	9.109e-02	-7.344	1.65e-06	***
X78	-5.011e+02	1.077e+02	-4.652	0.000266	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Residual standard error: 180.2 on 16 degrees of freedom					
Multiple R-squared: 0.9381; Adjusted R-squared: 0.9265					
F-statistic: 80.81 on 3 and 16 DF; p-value: 7.009e-10					

Gdzie:

Intercept - stała *const*,

Estimate - współczynnik (oszacowany parametr),

Std. Error - standardowy błąd oszacowania

t value - wartość statystyki *t*-Studenta,

Pr(>|t|) - poziom istotności w odniesieniu do parametru,

Signif. Codes - oznaczenie graficzne poziomu istotności,

Residual standard error - błąd standardowy reszt,

Multiple R-squared - współczynnik determinacji R^2 ,

Adjusted R-squared - skorygowany współczynnik determinacji,

F-statistic - statystyka *F*,

p-value - poziom istotności modelu.

Źródło: Opracowanie własne programem *lm* pakietu *R*.

Porównanie wyników z rezultatem uzyskanym z Gretl-a (zob. tabela 2.3) wskazuje na identyczność estymacji i uzyskanej statystyki opisowej. Zauważmy, że w *R* występuje też sygnalizacja graficzna błędów istotności poszczególnych parametrów modelu w postaci gwiazdek.

Zobaczmy jeszcze jak radzi sobie środowisko *R* z oszacowaniem regresji zmiennej *X3* z pozostałymi 14. zmiennymi objaśniającymi, które stanowiły wejście do modelu 1. zrealizowanego wcześniej programem *Gretl* (zob. tabela 2.2). Możemy to zrealizować korzystając z menu programu *R* w trybie windowsowym lub wprowadzając do okna skryptowego następujące polecenie:

```
LinearModel.3 <- lm(X3 ~ X9 + X10 + X11 + X12 + X13 + X14 + X25 + X38 + X48 + X49 +
X50 + X51 + X59 + X78, data=DATA)
```

Odpowiedzią komputera jest raport z zestawieniem sumarycznym modelu regresji liniowej wielowymiarowej nazwany *LinearModel.3* - zob. tabela 9.5.

Tab. 9.5. Oszacowanie parametrów dla czternastu zmiennych modelu wstępnego w programie *R*, dodatkowe informacje wejściowe oraz podstawowa statystyka opisowa

> summary(LinearModel.3)										

Call:										
lm(formula = X3 ~ X9 + X10 + X11 + X12 + X13 + X14 + X25 + X38 +										
X48 + X49 + X50 + X51 + X59 + X78, data = DATA)										

Residuals:										
1	2	3	4	5	6	7	8	9	10	11
-46.41	118.79	-175.16	178.07	-24.31	-120.37	-60.73	104.19	108.12	-43.20	48.21
12	13	14	15	16	17	18	19	20		
-13.98	-121.87	41.02	77.33	-55.06	106.79	-170.12	-34.03	82.72		

Coefficients:					
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	9833.04163	6662.18868	1.476	0.2000	
X9	-5.31161	2.37295	-2.238	0.0754	.
X10	2.34751	2.65491	0.884	0.4171	
X11	2.20456	4.46969	0.493	0.6428	
X12	-0.85295	3.11914	-0.273	0.7954	
X13	0.61113	2.51962	0.243	0.8180	
X14	-3.76527	7.68213	-0.490	0.6448	
X25	-0.63737	0.50188	-1.270	0.2600	
X38	88.29286	76.31622	1.157	0.2996	
X48	0.36000	0.88948	0.405	0.7024	
X49	0.46933	3.17212	0.148	0.8882	
X50	-0.07832	0.07615	-1.029	0.3508	
X51	-0.16798	0.42598	-0.394	0.7096	
X59	17.09851	15.97375	1.070	0.3334	
X78	-585.81937	297.79532	-1.967	0.1063	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Residual standard error: 199.1 on 5 degrees of freedom					
Multiple R-squared: 0.9764; Adjusted R-squared: 0.9102					
F-statistic: 14.76 on 14 and 5 DF; p-value: 0.003864					

Źródło: Opracowanie własne programem **lm** pakietu *R*.

Gdy porównamy z wynikami *Modelu 1*, uzyskanymi z *Gretl*-a to występują jednak nieznaczące różnice liczbowe w oszacowaniu parametrów modelu, być może jest to kwestia dokładności przeprowadzenia obliczeń. Ze względu na ramy tego opracowania sprawdzimy tylko trzy wiersze z *Modelem 1*. *Gretl*-a.

Zmienna	Współczynnik	Współczynnik wg <i>Gretl</i> -a	Błąd stand.	Statystyka t	Wartość p
const	11037,4	9833,0	6207,87	1,7780	0,13555
X9	-5,18169	-5,31161	2,50564	-2,0680	0,09348
X10	2,30576	2,34751	2,81819	0,8182	0,45048
X11	2,05037	2,20456	4,56698	0,4490	0,67225

Niedogodnością w raporcie z programu *R* jest brak podpowiedzi, którą ze zmiennych (oprócz stałej) usunąć w następnej iteracji. Tu musimy się sami kierować maksymalną wartością popełnienia błędu zapisaną w kolumnie *Wartość p*. Według tabeli 9.5 jest nią zmienna *X49* o $p = 0,8882$, czyli taka sama jak w *Modelu 1*, uzyskanym z *Gretl*-a. Pozostaje nam teraz zorientowanie się jakie możliwości daje menu programu *R* w zakresie weryfikacji uzyskanego modelu końcowego.

9.4. Weryfikacja podstawowa modelu końcowego

Przyjmijmy sytuację, że etap weryfikacji modelu podanego w tabeli 2.4 realizujemy po zamknięciu systemu komputerowego. W takiej sytuacji musimy ponownie zainicjować tryb okienkowy wywołując program *Rgui* ze ścieżki: *D:\Program Files\R\Bin\i386*. Następnie instalujemy program trybu windowsowego poleceniem: *library(Rcmdr)*. Pojawia się okno *R Commander* na którym uaktywniamy odpowiednią tabelę danych przyciskiem *Select Data Set*. Zaznaczamy plik *DATA i*

naciskamy OK. Pojawia się przycisk DATA z czcionką w kolorze niebieskim. W kolejności korzystamy z menu: *Statistics/Fit models/Linear Model ...*

Wpisujemy formułę naszego modelu: $X3 \sim X9 + X25 + X78$ i pojawia się przycisk również ze czcionką w kolorze niebieskim o treści: *LinearModel.1*. Oznacza to uaktywnienie naszego modelu regresji liniowej o trzech zmiennych objaśniających. Możemy teraz przystąpić do weryfikacji modelu korzystając w tym względzie z pięciu testów podstawowych proponowanych przez program R wywołując menu *Models/Numerical diagnostics*, a mianowicie:

1. *Variance-inflation factor* - badanie, czy w modelu nie występuje problem współliniowości.
 2. *Breuch-Pagan test for heteroscedasticity* - badanie, czy w próbie na podstawie której estymowano model zmiennej objaśnianej $X3$ nie występuje niejednorodność wariancji składników losowych.
 3. *Durbin-Watson test for autocorrelation* - badanie czy w modelu nie występuje autokorelacja reszt pierwszego stopnia.
 4. *RESET test for nonlinearity* - badanie nieliniowości modelu.
 5. *Bonferroni outlier test* - sprawdzenie czy w zbiorze danych występują obserwacje nietypowe.
- Ponadto wywołując menu: *Statistics/summaries/Shapiro-Wilk test normality* możemy sprawdzić, czy składnik losowy modelu zmiennej objaśnianej $X3$, w przybliżeniu reprezentowany przez reszty ma rozkład normalny.

Wywołanie poszczególnych poleceń menu generuje odpowiednie komendy, które są następnie zapisywane w oknie *Script Window*. Natomiast odpowiedzi programu (raporty) z kropką dziesiętną występują w oknie *Output Window* z powtórzeniem polecenia (zob. rysunku 9.6). Odnosnie *badania współliniowości* to polecenie i informacja wyniku testu są następujące:

```
Vif(LinearModel.1)
      X9      X25      X78
8.936837  10.425633  1.859964
```

W modelu występuje problem współliniowości - nieznaczny, gdyż kryterium jest $VIF_2 > 10$ ⁸³. Miara ta zwaną czynnikiem inflacji wariancji dla j -tego estymatora parametru obliczana jest według wzoru:

$$VIF_j = \frac{1}{(1 - R_j^2)}$$

Gdzie: R_j^2 - współczynnik determinacji modelu, w którym zmienna X_j , jest zmienną objaśnianą, a pozostałe z naszych trzech zmiennymi objaśniającymi.

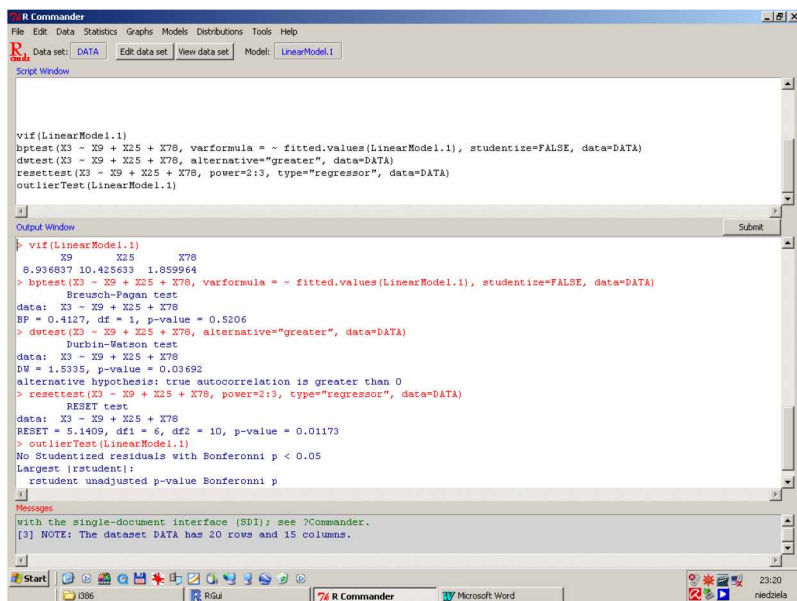
Niejednorodność wariancji składników losowych w próbie, czyli *heteroskedastyczność* występuje przeważnie w modelach bazujących na danych przekrojowych, a takim jest nasz model zmiennej $X3$. Na rysunku 9.6 występuje test Breuscha-Pagana, a polecenie i odpowiedź programu są następujące:

```
bptest(X3 ~ X9 + X25 + X78)
Breusch-Pagan test
data: X3 ~ X9 + X25 + X78
BP = 0,4127, df = 1, p-value= 0,5206.
```

Wartość $p\text{-value} = 0,5206 \geq \alpha = 0,05$, co wskazuje na *homoskedastyczność* składnika losowego⁸⁴, czyli nie występuje niejednorodność wariancji składników losowych.

⁸³ Walesiak M., Gatnar E. (red. nauk.), *Statystyczna analiza danych z wykorzystaniem programu R*, op. cit., s. 148.

⁸⁴ Ibidem, s. 149.



Źródło: Opracowanie własne w programie R.

Rys. 9.6. Widok poleceń oraz reakcji programu R na żądania wywołane poprzez menu

W badaniu autokorelacji pierwszego stopnia lub inaczej rzędu $AR(1)$ zastosowano test Durbina-Watsona przy założonym poziomie istotności $\alpha = 0,05$. W raporcie występuje polecenie i następująca odpowiedź (zob. rysunek 9.6):

```
dwtest(X3 ~ X9 + X25 + X78, alternative="greater", data=DATA)
```

```
data: X3 ~ X9 + X25 + X78
```

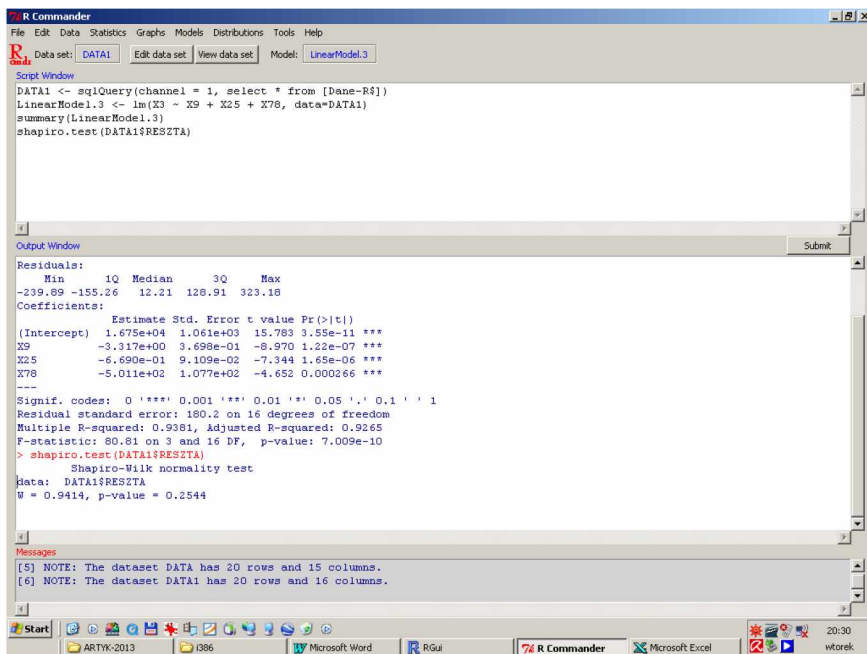
```
DW = 1.5335, p-value = 0.03692
```

```
alternative hypothesis: true autocorrelation is greater than 0.
```

Test ten wskazuje jednak na występowanie autokorelacji reszt pierwszego stopnia, ponieważ $\alpha = 0,05 > p\text{-value} = 0,03692$. Przyczyny tego zjawiska są różne, a jedną z nich jest dobór zmiennych objaśniających w okresach czasowych.

Sprawdzenie czy składnik losowy ma rozkład normalny. W tym względzie zastosujemy test Shapiro-Wilka w odniesieniu do obliczonych reszt ($X_3 - \hat{X}_3$) nazwanych w zbiorze *DATA1* jako *RESZTA*. Musi to być poprzedzone obliczeniem w *Excelu* wartości teoretycznych modelu końcowego z programu R oraz reszt, więc przypomnijmy sobie ten model:

$$\hat{X}_3 = 16750 - 3,317 X_9 - 0,669 X_{25} - 501,1 X_{78}$$



Źródło: Opracowanie własne w programie R.

Rys. 9.7. Widok poleceń oraz reakcji programu R na żądania estymacji modelu i przeprowadzenia testu na normalność składnika losowego

Na rysunku 9.7 wymieniono polecenia wygenerowania tabeli danych *DANE1* oraz odpowiedź programu na żądanie wykonania testu: *Shapiro-Wilk normality test* na podstawie zmiennej *RESZTA* przy domyślnej wartości poziomu istotności $\alpha = 0,05$. Widzimy, że $p\text{-value} = \mathbf{0,2544} > \alpha = 0,05$, zatem nie ma podstaw do odrzucenia hipotezy o normalności składnika losowego.

W badaniu występowania w zbiorze danych obserwacji nietypowych, które cechują duże reszty, a po ich wyeliminowaniu następuje poprawa współczynnika determinacji (R^2) wykorzystuje się funkcję **rstudent**:

```
outlierTest(LinearModel.1)
No Studentized residuals with Bonferonni p < 0.05
Largest |rstudent|:
rstudent unadjusted    p-value    Bonferonni p
9                2.012186    0.062517      NA
```

Program wskazał obserwację o numerze **9** jako nietypową ($p\text{-value} > p = \mathbf{0,05}$). Zachodzi więc potrzeba ponownego oszacowania parametrów modelu końcowego bez tej obserwacji. Z ciekawości porównajmy wartości tej obserwacji ze średnią (*mean*) poszczególnych zmiennych modelu.

Rok	t	X3	X9	X10	X11	X12	X13	X14	X25	X38	X48	X49	X50	X51	X59	X78
Średnia	-	2358,7	3003,3	644,1	1314,9	682,2	611,7	292,1	2436,6	114,4	1752,5	1404,8	23001	4093,4	76,1	5,6
2000	9	2760	2955,0	661,9	1325,0	654,9	614,2	286,7	2622,76	111,1	1904,57	1360,3	21037	3750,8	78,4	5,5

Postępując analogicznie w wyniku kolejnych iteracji wyeliminowano jeszcze obserwacje **19**, **18** oraz **6** uzyskując w efekcie poprawę współczynnika determinacji na $R^2 = 0,9596$ i oszacowanie parametrów modelu:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.583e+04	9.232e+02	17.146	8.35e-10	***
X9	-3.054e+00	3.292e-01	-9.277	8.01e-07	***
X25	-6.009e-01	8.041e-02	-7.473	7.50e-06	***
X78	-5.029e+02	9.754e+01	-5.156	0.000238	***

oraz raport z realizacji funkcji wyłonienia obserwacji wpływowych:

No Studentized residuals with Bonferonni p < 0.05

Largest |rstudent|:

rstudent unadjusted	p-value	Bonferonni p	
6	-1.671100	0.12288	NA

Funkcja *rstudent* nie proponuje kolejnej obserwacji jako nietypowej do usunięcia z tabeli danych.

Na zakończenie tego opracowania zobaczmy jak można uzyskać w trybie windowsowym, programem *R*, statystykę opisową obejmującą: średnią - *mean*, odchylenie standardowe - *sd*, kwantyle, skorzystać z menu: *Statistics/Summaries/Numerical summaries ...*

Tab. 9.6. Polecenie oraz wartości podstawowej statystyki poszczególnych zmiennych modelu wejściowego

<pre>> numSummary(DATA[,c("X3", "X9", "X10", "X11", "X12", "X13", "X14", "X25", "X38", "X48", "X49", "X50", "X51", "X59", + "X78")], statistics=c("mean", "sd", "quantiles"), quantiles=c(0,.25,.5,.75,1))</pre>								
	mean	sd	0%	25%	50%	75%	100%	n
X3	2358.700	664.527541	1211.00	1804.500	2359.500	2831.250	3431.00	20
X9	3003.260	334.192300	2639.10	2709.675	2866.550	3381.975	3467.60	20
X10	644.095	99.467289	453.10	568.975	676.600	707.550	823.00	20
X11	1314.850	199.076998	996.10	1213.525	1313.200	1416.350	1633.50	20
X12	686.195	84.393243	530.40	600.600	722.650	749.425	794.70	20
X13	611.720	140.931830	366.40	484.725	654.300	727.500	780.30	20
X14	292.145	62.042816	235.40	252.775	266.400	287.100	412.20	20
X25	2436.581	1465.266731	295.16	1041.168	2876.000	3495.813	4633.95	20
X38	114.405	12.278414	102.60	104.725	109.050	123.100	138.90	20
X48	1752.510	736.167686	404.70	1253.237	1901.085	2152.405	2872.24	20
X49	1404.800	230.951980	1194.70	1220.975	1297.350	1527.275	1921.90	20
X50	23000.950	4237.527425	18075.00	19634.250	21517.500	27588.500	29774.00	20
X51	4093.380	328.385402	3708.10	3843.225	3936.400	4303.525	4775.10	20
X59	76.095	17.219923	40.10	73.950	80.600	84.950	104.30	20
X78	5.585	0.523425	5.00	5.300	5.400	5.750	6.80	20

Źródło: Opracowanie własne w programie *R*.

Ponadto przy aktywnej tabeli danych przykładowo *DATA*, modelu regresji liniowej możemy skorzystać z polecenia **fix(DATA)** do pokazania zawartości tablicy danych wejściowych (zob. rysunek 9.8).

	X3	X9	X10	X11	X12	X13	X14	X25	X38	X48	X49	X50	X51	X59	X78
1	2394	3467.6	823	1042.1	794.7	476.7	287.7	295.16	138.9	404.7	1921.9	10430	4626.8	41.1	5.7
2	2595	3391.8	721.9	999.4	746.5	381.1	239.3	384.25	134.8	550.29	1856.5	18206	4504.2	40.1	5.4
3	2375	3361.4	678.8	996.1	728.1	366.4	235.4	515.49	134.5	708.3	1750	18206	4472.6	45.8	5.4
4	2233	3461.1	689.2	1078.6	723.4	414.4	243.5	704.99	131.6	878.29	1648.3	18075	4775.1	59.2	5.4
5	1961	3436	684.2	1119.7	721.9	425.3	250.5	900.58	126.7	1112.95	1565	19224	4258.9	70.5	5.3
6	1737	3433.4	723.5	1244.8	730.7	487.4	273.3	1088.03	121.9	1300	1514.7	19771	4437.4	77.4	5.3
7	1827	3378.7	736.9	1321.1	725.2	552.5	253	1277.06	115.7	1527.52	1452.6	20288	4243.5	83.3	5.4
8	2641	3138.4	710.4	1318.4	686.3	583	286.9	1428.01	112.5	1700.54	1406.5	20800	4138.6	86.4	5.7
9	2760	2955	661.9	1325	654.9	614.2	286.7	2622.76	111.1	1904.57	1360.3	21037	3750.8	78.4	5.5
10	3186	2820.6	627.8	1295.6	530.4	637.3	252.1	2821.3	108	1897.6	1310.4	21376	3905.8	75.1	5
11	3431	2670.5	545.5	1291.1	601.9	671.3	260	2930.7	102.6	1886.1	1284.3	21659	3708.1	85.7	5
12	3329	2639.1	496.4	1308	586.3	699.5	261.8	3017.64	104.2	1918.71	1268.6	22052	3877.2	78.4	5.1
13	3230	2663.1	453.1	1295.4	583	700.6	264.6	3111.04	104	1966.72	1229.1	22804	3779.1	98	5
14	3045	2665.4	483.6	1360.7	577.7	695.8	265.2	3252.02	103.8	1954.2	1219.4	26565	3855.3	104.3	5.4
15	2344	2714.3	511.5	1394.2	596.7	723.3	267.6	3421.05	104.9	2106.28	1221.5	27547	3967	88.8	5.9
16	1619	2842.4	576.8	1482.8	621	780.3	277	3720.1	107.9	2290.78	1280.1	27713	3807	84.7	6.5
17	1211	2890.7	671.2	1633.5	792	743.1	406.7	3986.5	110.1	2558.46	1212.2	28425	3792.2	80.6	6.8
18	1411	2725.4	706.6	1611.5	777.5	740.1	407.4	4232.68	105.4	2706.47	1194.7	28684	3873.6	80.6	6.6
19	1699	2696.1	674.4	1598.7	758.2	767	412	4388.31	103.9	2805.48	1197.9	29774	4237.3	81.5	5.9
20	2146	2714.2	705.2	1580.3	787.5	775.1	412.2	4633.95	105.6	2872.24	1202	29383	3857.1	82	5.4
21															
22															
23															
24															
25															
26															
27															
28															
29															
30															

Źródło: Opracowanie własne w programie R.

Rys. 9.8. Zawartość tabeli danych wejściowych

* * *

Celem niniejszego opracowania było pokazanie możliwości korzystania z trybu windowsowego programów wywołanych w środowisku pakietu programowego R. Odnośnie budowy oraz weryfikacji modelu ekonometrycznego w „Statystycznej analizie danych ...” prezentowane są przykłady realizacji w trybie *scriptowym* (tekstowym). Wymaga to jednak szczegółowej znajomości struktury i komend języka nazwanego również R.

Ponieważ opracowanie adresowane jest dla studentów, zwłaszcza studiów zaocznych kierunku zarządzanie, którzy łączą pracę zawodową z nauką, szybkie i sprawne posługiwanie się rozbudowanym menu umożliwi dogodną estymację parametrów jak i sprawdzenie poprawności zbudowanego modelu ekonometrycznego.

Ponadto może stanowić alternatywne, obok innych programów komputerowych, obszerne narzędzie badań statystycznych - w różnych branżach obejmujące nowe elementy z tej dziedziny.

10. Dobór pakietu integracji modułów wsparcia rolnictwa



10.1. Wstęp

Współcześnie systemy wspomagające zarządzanie firmą są niezbędnym narzędziem do funkcjonowania każdego zakładu. Służą do usprawnienia pracy pracowników i ułatwienia komunikacji z kontrahentami, z innymi oddziałami, centralą oraz instytucjami. Ponadto znacznie ułatwiają realizację wykonywanych zadań, dostarczają niezbędnych informacji i wywierają pozytywny wpływ na spójność przedsiębiorstwa. Tak więc wymagania jakie stawia klient, zapisy w przepisach prawnych czy wręcz sam pracownik zmuszają pracodawców do wdrażania techniki IT. Zatem nowoczesne rozwiązania systemowe wykorzystujące sieć globalną Internet stały się koniecznością i stanowią niezbędne minimum dla dobrze prosperującego przedsiębiorstwa.

Instytucje państwowe posiadając ograniczenia budżetowe i specyficzną strukturę zależności występujących pomiędzy oddziałem, a centralą oraz wyznaczniki zawarte w aktach normatywnych (jak zamówienia publiczne) nie mają zbyt wiele możliwości rozwoju w zakresie informatyki. Jednakże pracownicy, jak i kontrahenci chcą sprawnej i zrozumiałej współpracy. Ponadto rzadkością w jednostkach budżetowych jest zastosowanie strategii zarządzania wiedzą.

Opracowanie niniejsze bazuje na wykonanej pod kierunkiem autora pracy magisterskiej⁸⁵. W badaniu wykorzystano opinie pracowników wyrażone na podstawie ankiet oraz dopowiedzenie wątpliwych kwestii określone w przeprowadzonym wywiadzie. Materiał empiryczny pochodzi z badania opinii, przeprowadzonego wśród pracowników Krajowego Ośrodka Wsparcia Rolnictwa - Oddział w Opolu (KOWR OT Opole).

Respondentami byli najbardziej aktywni pracownicy, którzy chętnie dzielą się wiedzą. W pierwszej kolejności przeprowadzona została ankieta wśród 40. pracowników, których staż pracy jest powyżej 5 lat, a kolejno dwie osoby na szczeblu kierowniczym udzieliły wypowiedzi i podzieliły się ich wizją sprawniejszego funkcjonowania systemów i wpływu możliwości zmian na lepsze wyniki pracy podległych im komórek organizacyjnych. Badanie ankietowe wykazało konieczność zmian systemowych, możliwości i zasadności ich przeprowadzenia zwłaszcza na poziomie integracji stosowanych aplikacji programowych, w zakresie obsługi nieruchomości z zasobu Skarbu Państwa. W niniejszym opracowaniu pracy magisterskiej występuje szereg skrótów, które już teraz na wstępie warto objaśnić:

- API - *Application programming interface* (interfejs programowania aplikacji),
- ARIMR - Agencja Restrukturyzacji i Modernizacji Rolnictwa,
- BI - *Business Intelligence* - narzędzie analityczne do prowadzenia analiz controllingowych,
- CRM - *Customer Relationship Management* (zarządzanie relacjami z klientami),
- DMS - System zarządzania dokumentami,
- e-Doręczenia - Elektroniczny odpowiednik listu poleconego za potwierdzeniem odbioru,
- E-obieg - Elektroniczny obieg dokumentów,
- E-Puap - Elektroniczna Platforma Usług Administracji Publicznej,
- EPR - *Enterprise Resource Planning* (planowanie zasobów),
- EZD - Elektroniczne zarządzanie dokumentacją,
- IC - *Inventory Control* (Sterowanie zapasami),

⁸⁵ Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, Akademia Nauk Stosowanych Wyższej Szkoły Zarządzania i Administracji w Opolu, 2024.

ICT - Technologia informacyjno – komunikacyjna (*Information and communication technologies*),
 IT - *Information Technology*,
 JRWA - Jednolity rzeczowy wykaz akt,
 MRP - *Material Requirements Planning* (Planowanie zapotrzebowania materiałowego),
 OCR - Automatyzacja procesów wprowadzania dokumentów,
 SCM - *Supply Chain Management* (System zarządzanie przepływem towarów),
 SRM - *Supplier Relationship Management* (Zarządzanie relacjami z dostawcami),
 TQM - *Total Quality Management* (Kompleksowe zarządzanie jakością),
 VPN - *Virtual Private Network* (Wirtualna sieć prywatna),
 WMS - *Warehouse Management System* (System do zarządzania magazynem),
 WRSP - Własność Rolna Skarbu Państwa

10.2. Zintegrowany system informatyczny

Każde przedsiębiorstwo, do sprawnego funkcjonowania potrzebuje integracji różnych funkcji, takich jak finanse, zasoby ludzkie, produkcja, sprzedaż czy marketing. Integracja dokonuje się poprzez udostępnianie informacji na temat działalności organizacji na różnych poziomach zarządzania. Niezawodnym sposobem unifikacji procesów biznesowych oraz systemów i technologii informatycznych w przedsiębiorstwie jest stosowanie systemów informatycznych.

Systemy informatyczne wspierają zadania podstawowe, kontrolne procesów biznesowych czy zarządzanie dokumentami elektronicznymi. Te systemy to również narzędzie przewagi konkurencyjnej przedsiębiorstwa, bowiem stanowią wydzieloną, skomputeryzowaną część systemu informacyjnego⁸⁶. Encyklopedia *Britannica* definiuje „*information system*” jako „*zintegrowany zestaw składników do zbierania, przechowywania, przetwarzania i komunikowania informacji*”, na który składa się *hardware*, programy, bazy danych, telekomunikacja, zasoby ludzkie i procedury⁸⁷. Natomiast z innego punktu rozpatrywania, system informacyjny to nauki o przetwarzaniu informacji.

W naukach o zarządzaniu informacja jest bowiem pojmowana jako element przekazywania wiedzy od odbiorcy informacji, umożliwiający zwiększenie pewności działania odbiorcy tejże informacji. Informacja jest zatem wartością dodaną w systemie przekazywania komunikatu od odbiorcy. Danymi są fakty poddane interpretacji, kwantyfikowane lub jakościowe przedstawione w określonej formie. Informacje zatem stanowią przetworzone dane, które zostały wstępnie zinterpretowane.

Systemy informatyczne jako sztuczne służą do przetwarzania danych w sposób zautomatyzowany i schematyczny. Są to zbiory powiązanych ze sobą elementów, którego funkcją jest przetwarzanie danych przy użyciu techniki komputerowej⁸⁸. Z realnego punktu widzenia system informatyczny organizacji można zobrazować w postaci uporządkowanych elementów takich jak⁸⁹:

- kadra pracownicza obsługująca system,
- dane i informacje,
- rozwiązania organizacyjne,
- narzędzia i urządzenia technologii informatycznej,
- zbiór metainformacji (informacji o informacji),
- relacje pomiędzy elementami systemu informatycznego,
- otoczenie i budowa systemu informatycznego.

System zarządzania to system złożony z zasad, procedur, metod, opisów stanowisk, narzędzi pracy, ludzi oraz relacji pomiędzy nimi. Zadaniem tych zasad jest osiągnięcie wyznaczonych celów

⁸⁶ Kisielnicki J., Sroka H., *Systemy informacyjne biznesu*, Placet, Warszawa 2005, s. 18.

⁸⁷ <https://prawo.it.pl/2010/12/19/system-informatyczny-system-informacyjny-uslug-a-informatyczna/>.

⁸⁸ Ibidem, s. 9.

⁸⁹ Wornalkiewicz W., *Procesy informacyjne w zarządzaniu*, Wydawnictwo WSZiA, Opole 2020 s. 10.

poprzez ciągłe doskonalenie wszystkich jego elementów. Systemu zarządzania powinien zmierzać do poprawy jakości własnych wyrobów, usług, także obsługi klienta i wszystkich innych, zdefiniowanych elementów. Celem wdrożenia systemu zarządzania jest zmniejszenie ilości nieprawidłowości, niezgodności i braków wewnętrznych. System zarządzania usprawnia obieg informacji i dokumentów w instytucji, buduje świadomość oraz kulturę organizacyjną. Ponadto wprowadza przejrzysty podział zadań, uprawnień i odpowiedzialności wszystkich pracowników.

Rozpowszechnienie Internetu przyczyniło się w znacznym stopniu do zmian w sposobie funkcjonowania przedsiębiorstw, dostarczając nowych narzędzi pozwalających sprostać oczekiwaniom biznesowym. Takie zmiany wymuszają stosowanie w firmach systemów przetwarzania informacji, lecz skuteczne zarządzanie informacją wymaga zaprojektowania i wdrożenia odpowiedniej infrastruktury teleinformatycznej.

Podstawowym narzędziem organizacyjnym, umożliwiającym każdemu przedsiębiorstwu podnoszenie poziomu jakości wytwarzanych wyrobów czy też świadczonych usług, poprzez ciągłe doskonalenie metod i sposobów działania jest zgodność z normami ISO (*International Organization for Standardization*). Główną normą jest ISO 9001 jako międzynarodowy standard poświęcony systemowi zarządzania jakością. Stosowanie wymagań normy ISO 9001 stanowi podstawę wdrożenia zarządzania przez jakość TQM (*Total Quality Management*). Natomiast system CRM (*Customer Relationship Management*) to strategia biznesowa oparta o systemy informatyczne gromadzące dane klienta ze wszystkich źródeł, zarówno wewnętrznych jak i zewnętrznych w celu ich analizy.

Zintegrowane systemy informatyczne (ZSI) są zaawansowanymi rozwiązaniami wspomagającymi procesy zarządzania w przedsiębiorstwach i instytucjach. ZSI to system komputerowy (zbiór programów) obejmujący obszary działania: finanse i księgowość, sprzedaż i zaopatrzenie, konstrukcja i technologia, produkcja, utrzymanie ruchu i remonty, inne⁹⁰. ZSIZ (zintegrowane informatyczne systemy zarządzania) są najczęściej wdrażane jako środowisko aplikacyjne integrujące różne obszary funkcjonowania przedsiębiorstw na poziomie operacyjnym. ZSI najczęściej wspiera działania przedsiębiorstwa opartego na metodologiach zarządzania produkcją takimi jak:

EPR (*Extended Producer Responsibility*),

SCM (*Supply Chain Management*),

WSM (*Warehouse management system*).

Metodologia EPR (rozszerzona odpowiedzialność producenta) to systemy, które gromadzą wiele modułów dzięki czemu pozwalają na objęcie nimi wszystkich aspektów firmy poprzez logistykę, magazyny, a także zarządzanie i raportowanie⁹¹.

Natomiast zintegrowane aplikacje pakietu SAP ERP łączą wszystkie obszary przedsiębiorstwa w ramach inteligentnego pakietu rozwiązań dostępnych na całkowicie cyfrowej platformie, wypierając wcześniejsze wersje. Ideę budowy tego rozwiązania informatycznego przedstawiono na rysunku 10.1⁹².



Źródło: <https://hicon.com/pl/blog/sap-erp-biznesowe-kompedium-wiedzy/>.

Rys. 10.1. Budowa SAP ERP

⁹⁰ Ibidem.

⁹¹ Wornalkiewicz W., *Wdrożenie zintegrowanego systemu informatycznego wspomagającego zarządzanie*, WSZiA Opole, 2015. s. 16-17.

⁹² <https://www.sap.com/poland/about/what-is-sap.html>.

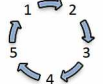
Struktura systemu ERP zaoferowana przez SAP przechodzi przez kolejne etapy. W bazie danych (dla wszystkich modułów) przechowywane są informacje wykorzystywane przez oprogramowanie. Informacje można przeglądać, aktualizować, archiwizować i usuwać. Potrzebne jest tu jednak wprowadzenie zabezpieczenia, które nie pozwoli osobom nieautoryzowanym na dostęp do tej najgłębszej warstwy oprogramowania.

W warstwie aplikacji przetwarzane są żądania użytkownika (wprowadzanie zmian, przywoływanie danych, generowaniem raportów). Aby żądanie zostało zidentyfikowane jak najszybciej, jest ono przesyłane zawsze do tego serwera, który w danej chwili ma najmniejsze obciążenie. Warstwa prezentacji to interfejs do pracy użytkownika. Jego żądania wysyłane są z warstwy prezentacji, a następnie są przetwarzane przez warstwę aplikacji. Kolejno wymagane dane w następnym kroku są pobierane z bazy danych, przetwarzane i wyświetlone z powrotem w warstwie prezentacji. W taki sposób każdy użytkownik może pracować z systemem, bez zagłębiania w bardziej skomplikowane warstwy oprogramowania.

10.3. System zarządzania dokumentacją

Każde przedsiębiorstwo wymaga efektywnego zarządzania dokumentami. *System Zarządzania Dokumentami (DMS)* jest rozwiązaniem dla firm i instytucji, w których istnieje do przetworzenia duża ilość dokumentów w wersji papierowej. Dzięki tym systemom możliwe jest zbudowanie profesjonalnego cyfrowego archiwum i sprawne zarządzanie dokumentami. DMS zapewnią wsparcie w digitalizacji firmy i odejście od realizowania obiegu w formie papierowej, czy załatwiania spraw e-mailem, czy poprzez program *Excel*. Usprawnienie działania związane jest z wprowadzaniem, opisywaniem oraz akceptacją dokumentów różnego typu takich jak faktury, umowy, korespondencja. DMS usprawnia bowiem procesy i komunikację wewnętrzną i zewnętrzną w określonej organizacji. Możliwości systemów klasy DMS wyszczególniono w tabeli 3.1.

Tab. 10.1. Możliwości systemów DMS

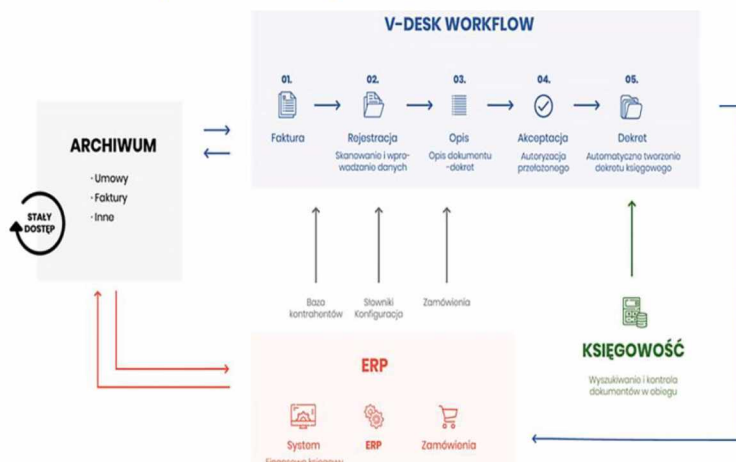
Opis zdolności	Nazwa zdolności	Symbol
Swobodny dostęp do dokumentów z dowolnego miejsca dzięki możliwości pracy na aplikacjach mobilnych, przeglądarkach WWW i aplikacji <i>desktop</i> .	Mobilny dostęp	
Bieżący dostęp do zadań dzięki systemowi powiadomień mobilnych i e-mailowych.	Powiadomienia	
Sprawniejsza realizacja procesów poprzez definiowanie czasu, automatyczna eskalacja i raportowanie informacji.	Kontrola czasu	
Tworzenie i zmiana firmowych procesów dzięki wbudowanemu graficznemu konfiguratorowi procesów.	Platforma <i>low-code</i>	
Możliwość rozszerzenia standardowych obszarów EPR o nowe procesy zachowując integrację danych.	Współpraca z EPR	
Automatyzacja procesów wprowadzania dokumentów księgowych.	OCR	
Elektroniczne archiwum z pełnym dostępem do skanów załączników.	Elektroniczne archiwum	
Możliwość zastosowania własnych schematów obiegu dokumentów wraz z opeją wprowadzania reguł biznesowych i podpowiadaniem kolejnych kroków.	Ścieżka adaptacji	

Dostęp do pełnej informacji o procesie: osoby odpowiedzialne, czas realizacji komentarze i komplet danych.	Historia procesu	
--	------------------	--

Źródło: Opracowanie na podstawie - <https://primaco.pl/comarch-dms-elektroniczny-obieg-dokumentow/>.

W obecnych czasach najlepszym rozwiązaniem do zarządzania dokumentacją w firmie wydaje się być elektroniczny obieg dokumentów (*E-obieg*), czyli proces przechodzenia dokumentów przez różne etapy pracy w formie cyfrowej. Etapami takich procesów jest ich przetwarzanie, przesyłanie, przechowywanie, weryfikowanie i podpisywanie elektronicznie, a także przekazywanie do kolejnych etapów lub użytkowników w sposób zautomatyzowany. *E-obieg* dokumentów ma na celu zastąpienie tradycyjnych procesów opartych na „papierze” i fizycznych dokumentach oraz minimalizację kosztów i czasu związanych z obsługą. Dzięki elektronicznemu przetwarzaniu dokumentów można zwiększyć efektywność pracy, przyspieszyć procesy biznesowe jak i zmniejszyć ryzyko utraty nieautoryzowanego dostępu do poufnych informacji⁹³.

E-obieg umożliwia przetwarzanie przez systemy zarządzania dokumentami (DMS) lub systemy *workflow*, które pozwalają na śledzenie postępu pracy, przypisywanie zadań do konkretnych osób i automatyzację przepływu dokumentów między różnymi etapami pracy. Dzięki temu możliwe jest również wykorzystanie narzędzi takich jak szablony dokumentów, przypomnienia o terminach czy automatyczne generowanie raportów⁹⁴. Wymieniony tu *Workflow* (*work flow*, „przepływ pracy”), w szerszym znaczeniu oznacza sposób przepływu informacji pomiędzy różnymi obiektami biorącymi udział w jej przetwarzaniu. W węższym sensie jest to określenie sposobu przepływu dokumentów pomiędzy pracownikami wykonującymi pewien zalgorytmizowany zespół czynności, odbywający się według danej liczby scenariuszy⁹⁵. Graficzne przedstawienie systemu obiegu dokumentów w procesie zarządzania zaprezentowano na rysunku 10.2.



Źródło: <https://primesoft.pl/elektroniczny-obieg-dokumentow/?gclid>.

Rys. 10.2. Graficzne przedstawienie obiegu dokumentów przetwarzanych przez system zarządzania

W celu usprawnienia funkcjonowania administracji rządowej RP i świadczenie usług, poprzez ich elektroniczną, Rozporządzeniem Prezesa Rady Ministrów z dnia 18 stycznia 2011 r. w sprawie

⁹³ <https://primesoft.pl/elektroniczny-obieg-dokumentow/?gclid>

⁹⁴ Ibidem.

⁹⁵ <https://pl.wikipedia.org/wiki/Workflow>.

instrukcji kancelaryjnej, jednolitych rzeczowych wykazów akt oraz instrukcji w sprawie organizacji i zakresu działania archiwów zakładowych Dz. U. 2011 nr 14 poz. 67, przedstawiono wytyczne w zakresie systemu zarządzania dokumentacją, umożliwiające kompleksowe załatwianie i rozstrzyganie spraw, w tym z zakresu postępowania administracyjnego. Cel główny i cele szczegółowe zostaną osiągnięte poprzez użytkowanie w jednostkach administracji rządowej RP jednolitego systemu EZD PUW zintegrowanego z platformą ePUAP⁹⁶.

Pod względem technicznym system EZD można wdrożyć w infrastrukturze lokalnej podmiotu lub wykorzystując usługi chmurowe. Decyzja o trybie instalacji powinna być podyktowana wcześniejszą analizą wdrożeniową, możliwościami technicznymi placówki oraz kompetencjami pracowników. Należy przy tym pamiętać, że niezależnie od wybranej metody może być też konieczne wyposażenie podmiotu w sprzęt biurowy (skanery, drukarki i czytniki kodów kreskowych) oraz skonfigurowanie stacji roboczych⁹⁷. Wymienione rozwiązanie chmurowe, potocznie zwane *chmurą*, to rozległa sieć serwerów zdalnych znajdujących się w różnych miejscach na świecie. Serwery są połączone i działają jak jeden ekosystem. Pełnią one różne funkcje: przechowują dane i umożliwiają zarządzanie nimi, obsługują aplikacje oraz dostarczają zawartość lub usługi, takie jak strumieniowe przesyłanie materiałów wideo, poczta internetowa, oprogramowanie biurowe i sieci społecznościowe⁹⁸.

EZD jest także skutecznym narzędziem kierownictwa do zarządzania jednostką. Urzędy i instytucje, w których EZD obowiązuje jako podstawowy system kancelaryjny, funkcjonują sprawniej, podnosząc kulturę organizacyjną pracy. System EZD RP jako jednolite narzędzie dla administracji publicznej jest wspólną wartością korzystających z niego urzędów. Dotyczy to ponownego wykorzystania produktów wytwarzanych przez polską administrację w ramach realizowanych zadań publicznych. Oprócz wspólnego systemu są to procedury, dokumentacja, aplikacje oraz współdzielenie bazy wiedzy tworzonej dzięki współpracy pomiędzy wieloma jednostkami.

Integracja systemu EZD RP z systemami zewnętrznymi jest realizowana poprzez implementację dedykowanych interfejsów API (interfejs programowania aplikacji) dla głównych systemów używanych w administracji. Dodatkowo API umożliwiać będzie integrację systemu EZD RP z systemami dziedzicznymi wykorzystywanymi przez instytucje⁹⁹.

Technologia informacyjno - komunikacyjna ICT (*information and communication technologies*) wywiera obecnie istotny wpływ na administrację publiczną i jest jednym z kluczowych elementów zwiększenia efektywności jej funkcjonowania. Rolą jej jest rozwój administracyjnych usług publicznych oraz poziom rozwoju *e-usług* administracji publicznej w Polsce. Szczególną uwagę zwrócić trzeba na funkcjonowanie *Elektronicznej Platformy Usług Publicznych*, która jest podstawowym elementem koncepcji zarządzania *one-stop-government*.

10.4. Organizacja KOWR

Krajowy Ośrodek Wsparcia Rolnictwa (KOWR) został utworzony 1.09.2017 roku na mocy ustawy z dnia 10 lutego 2017 r. (Dz. U. 2017 poz. 623). W tym też terminie przestały istnieć dwie agencje rolne Agencja Nieruchomości Rolnych i Agencja Rynku Rolnego, których zadania przejął w znacznej części KOWR. Krajowy Ośrodek Wsparcia Rolnictwa jest instytucją powiemniczą, której Skarb Państwa powierzył wykonywanie prawa własności i innych praw rzeczowych na jego rzecz w stosunku do mienia państwowego w rolnictwie.

⁹⁶ <https://ezd.gov.pl/www/poig7/opis>.

⁹⁷ <https://podrecznik.ezdrp.gov.pl/modele-wdrazania-ezd-rp/>.

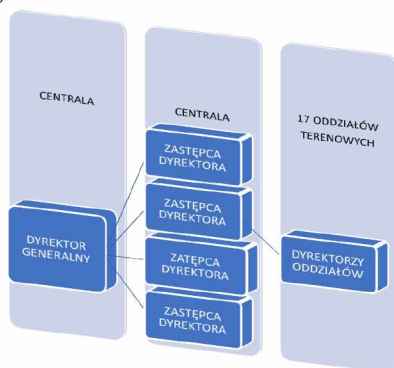
⁹⁸ <https://azure.microsoft.com/pl-pl/resources/cloud-computing-dictionary/what-is-the-cloud/>.

⁹⁹ <https://www.gov.pl/web/ezd-rp/o-ezd-rp>.

W oparciu o treść art. 27 *Ustawy o gospodarowaniu nieruchomościami rolnymi Skarbu Państwa* organizacja KOWR jest jedynym podmiotem prowadzącym sprzedaż i nabywanie nieruchomości z Zasobu Własności Rolnej Skarbu Państwa¹⁰⁰. Jak wynika natomiast z treści art. 2. *Ustawy o Krajowym Ośrodku Wsparcia Rolnictwa* (Dz.U.2023 poz. 1933) ośrodek ten jest państwową osobą prawną będącą agencją wykonawczą w rozumieniu Ustawy z dnia 27.08.2009 r. o finansach publicznych (Dz. U. z 2023 r. poz. 1270)¹⁰¹. Do końca 1995 r., kiedy zasadnicza część tego zadania została zrealizowana, Agencja przejęła do Zasobu Własności Rolnej Skarbu Państwa około 4 414 tys. ha, z których ponad 83 % pochodziło z 1 666 zlikwidowanych państwowych przedsiębiorstw gospodarki rolnej PGR o przeciętnej powierzchni 2 250 ha.

Przejęte państwowe gospodarstwa rolne podlegały restrukturyzacji, w ramach której wydzielano do zaferowania w drodze sprzedaży lub dzierżawy zarówno większe obszary nieruchomości posiadające bazę budynkową, jak też grunty niezabudowane, służące głównie powiększaniu istniejących gospodarstw indywidualnych. Restrukturyzacja prowadziła również do wydzielenia gruntów na inne cele, np. gruntów niezbędnych gminom na realizację zadań własnych, czy też gruntów nierolnych. Opracowywane programy restrukturyzacji przedstawiające sposób zagospodarowania mienia pochodzącego z byłych PGR (państwowych gospodarstw rolnych), przedkładane były do zaopiniowania gminom oraz konsultowane z organizacjami społeczno-zawodowymi rolników¹⁰².

Strukturę KOWR stanowi Centrala mieszcząca się w Warszawie oraz 17 oddziałów i filii w całym kraju (zob. rysunek 10.3).



Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 6.

Rys. 10.3. Struktura organizacyjna KOWR

Natomiast struktura organizacyjna oddziału badanego została przedstawiona na rysunku 10.4. Głównym zadaniem oddziałów jest obrót nieruchomościami rolnymi pochodzącymi z Zasobu Skarbu Państwa, przy czym badany oddział w Opolu zatrudnia około 80 osób.

¹⁰⁰ Ustawa z dnia 19.10.1991 r. o gospodarowaniu nieruchomościami rolnymi Skarbu Państwa, Dz.U. 2023, poz. 1933.

¹⁰¹ Ustawa z dnia 10.02.2017 r. O Krajowym Ośrodku Wsparcia Rolnictwa, Dz. U. 2023 poz. 1308.

¹⁰² <https://www.kowr.gov.pl/>.



Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 7.

Rys. 10.4. Struktura organizacyjna KOWR OT Opole

10.5. Istniejące i projektowane zastosowanie informatyki

Projekt *Pasportyzacja polskiej żywności* przewiduje budowę cyfrowego systemu gwarantującego efektywne monitorowanie i identyfikowanie informacji o produktach rolno-spożywczych w łańcuchu dostaw, tj. „od pola do stołu” (produkcja - konsumpcja). W przypadku wdrożenia system IT (*information technology*) udostępni to szerokie spektrum informacji o produkcji żywnościowym, przede wszystkim dzięki innowacyjnym funkcjonalnościom, a także integracji z aktualnie działającymi bazami referencyjnymi. Tym samym polski sektor rolno-spożywczy może zyskać narzędzie, w którym znajdują się wiarygodne, rzetelne i niepodrabialne dane o produkcji rolno-spożywczej¹⁰³.



Źródło: <https://www.nask.pl/pl/aktualnosci/4378.Pasportyzacja-polskiej-zywnosci-nowatorski-projekt-w-skali-swiatowej.html>.

Rys. 10.5. Graficzna prezentacja reklamy „*Pasportyzacji polskiej żywności*”

Wymieniony tu NASK to państwowy instytut badawczy, którego misją jest poszukiwanie i wdrażanie rozwiązań, służących rozwojowi sieci teleinformatycznych w Polsce oraz poprawie ich efektywności i bezpieczeństwa. Instytut prowadzi badania naukowe, prace rozwojowe, a także działalność operacyjną na rzecz bezpieczeństwa polskiej cywilnej cyberprzestrzeni. Ważnym elementem działalności NASK jest też edukacja użytkowników oraz promowanie koncepcji społeczeństwa informacyjnego, głównie w celu ochrony dzieci i młodzieży przed zagrożeniami, związanymi z użytkowaniem nowych technologii¹⁰⁴.

Przewidywany do wdrożenia *System Satelitarne Monitorowania Upraw Rolnych* (Projekt S2MUR) obejmuje zaprojektowanie i budowę systemu IT wykorzystującego dane satelitarne, meteorologiczne, glebowe i statystyczne do całorocznego monitorowania upraw rolnych na terenie całego kraju¹⁰⁵.

¹⁰³ <https://www.gov.pl/web/kowr/pasportyzacja-polskiej-zywnosci>.

¹⁰⁴ <https://www.nask.pl/>.

¹⁰⁵ <https://www.gov.pl/web/kowr/projekt-s2mur-stan-realizacji>.

Do realizacji celów i wykonywania powierzonych zadań związanych z zarządzaniem Zasobem w każdym z 17. Oddziałów Terenowych oraz Centrali działa lokalna instancja systemu informatycznego, zwanego ZSI. System powstał w 2000 roku jako rozwiązanie dedykowane, wspierające zadania statutowe oraz działalność bieżącą i administracyjną jednostki terenowej. Każda jednostka pracuje na swojej instalacji, bazie danych i niezależnej konfiguracji. Dane, które są przetwarzane w systemie pochodzą głównie z konwersji z systemów użytkowanych przez Agencję Własności Rolnej Skarbu Państwa przed powstaniem systemu.

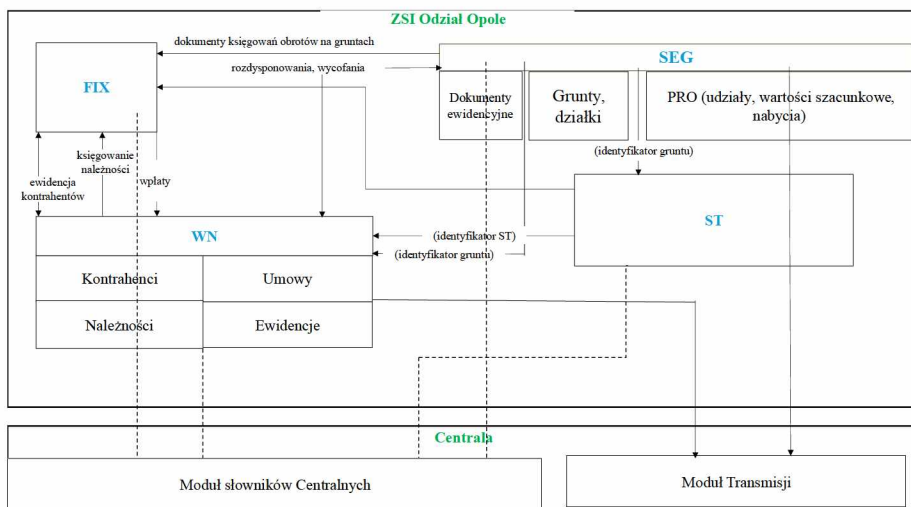
KOWR jest właścicielem autorskich praw majątkowych do podsystemów SEG i WN oraz modułów KE, PU, SC i TR. Pozostałe systemy FIX i ST użytkowane są na mocy licencji. System ZSI składa się z następujących powiązanych i współpracujących ze sobą podsystemów lokalnych:

- *System Ewidencji Gruntów* (SEG),
- *Windykacja Należności* (WN),
- *Finanse i Księgowość* (FIX),
- *Środki Trwałe* (ST).

Dodatkowo udostępnione są wspomagające moduły międzysystemowe, a mianowicie: *Słowniki Centralne* (SC), *Transmisja* (TR), *Konwersja umów i innych ewidencji* (KE), *Porządkowanie Umów* (PU), *Integracja SEG z portalem* (ISP). W oddziałach funkcjonują także aplikacje lokalne powiązane z ZSI służące do sprawnej realizacji zadań KOWR w szczególności rozliczania zobowiązań podatkowych. Bazy danych w kontekście ZSI są logicznie rozłączone dla każdego oddziału, choć schemat bazy jest taki sam dla wszystkich.

Infrastruktura serwerowa została scentralizowana, z jednostek terenowych (oddziałów) została przeniesiona do serwerowni centralnej. Obecnie systemy jednostek terenowych wykorzystują zasoby centralnej zwirtualizowanej infrastruktury udostępnionej i zarządzanej w Centrali KOWR. Aktualny model architektury systemów wspierających gospodarowanie Zasobem (gruntów) oparty jest na architekturze rozproszonej, lokalnie autonomicznej w zakresie aplikacyjnym i bazodanowym. Oznacza to, że każda jednostka terenowa posiada swoją instalację zestawu aplikacji i baz danych składających się na całość ZSI obsługującego Zasób.

W każdej jednostce terenowej powtórzony jest w całości pełen schemat logicznych powiązań pomiędzy komponentami systemu. Każda instalacja ZSI jest zintegrowana z *Centralnymi Słownikami*. W przypadku wystąpienia zdarzeń zmiany właściwości oddziałowej dokonuje się transmisji do centralnego modułu *Transmisji*, skąd oddział docelowy pobiera dane do swojej ewidencji zasobu. Schemat architektury logicznej *Zintegrowanego Systemu Informatycznego* funkcjonującego w Oddziale KOWR w Opolu pokazano na rysunku 10.6.



Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 11.

Rys. 10.6. Schemat architektury logicznej ZSI w KOWR

Podstawowym podsystemem ZSI Oddziału KOWR w Opolu jest WN, który współpracuje z lokalną instalacją bazy danych umów, dokumentów, należności, rozłączną z pozostałymi bazami lokalnymi w zakresie przetwarzania danych. Podstawowym zadaniem WN jest obsługa należności wynikających z umów sprzedaży, najmu, dzierżawy, użytkowania, administrowania i użyczenia majątku z Zasobu Własności Rolnej Skarbu Państwa, będącego w dyspozycji danego oddziału terenowego, a głównymi funkcjami WN są¹⁰⁶:

- ewidencja umów z kontrahentami,
- rozliczanie umów,
- inne ewidencje,
- obsługa dokumentów i księgowości,
- wystawianie faktur VAT,
- dokonywanie memoriałowych naliczeń zadany okres oraz generowanie dokumentów,
- obsługa pomocy publicznej,
- raportowanie,
- kontrola zgodności sald na kontach 215 pomiędzy systemami WN i FIX,
- słowniki,
- planowanie.

Kolejnym podsystemem głównym lokalnym współpracującym z lokalną instalacją bazy danych finansowo - księgowych rozłączną z pozostałymi bazami lokalnymi w zakresie przetwarzania danych jest podsystem FIX, przy czym główne funkcje systemu FIX są następujące:

- definiowanie planu kont;
- rejestracja (w tym import dokumentów z innych systemów ZSI, księgowanie korekty dokumentów);
- ewidencja danych dotyczących kontrahentów (osób fizycznych, prawnych, banków, sądów);
- prowadzenie rejestru VAT;

¹⁰⁶ Instrukcja użytkownika WN.

- prowadzenie analiz i rozliczeń transakcji rozrachunkowych;
- rozliczenia kosztów działalności;
- automatyczne generowanie przelewów bankowych, w tym integracja z modulem bankowości elektronicznej innego producenta;
- obsługa wyciągów bankowych i kas;
- obsługa scenariuszy księgowania dokumentów zewnętrznych dla wszystkich modułów ZSI
- raportowanie (dostępny generator własnych raportów i sprawozdań finansowych) i prowadzenie analiz z możliwością eksportu do arkuszy kalkulacyjnych;
- współpraca z innymi systemami ZSI w oparciu o definiowane na poziomie FIX i na poziomie systemów dziedzicznych scenariusze księgowania - scenariusze te mogą w szczególnych przypadkach (WN) być zdefiniowane dla całej organizacji (scenariusze centralne).

Integralną częścią podsystemu FIX jest dedykowany moduł kontrahentów pozwalający na zarządzanie informacjami o kontrahentach ZSI. Zbiór kontrahentów jest używany przez wszystkie moduły ZSI.

Następnym podsystemem jest *Ewidencja Gruntów* (SEG). Podstawowym zadaniem podsystemu SEG jest ewidencjonowanie rzeczowe i księgowe gruntów Zasobu Własności Rolnej Skarbu Państwa oraz wspomaganie procesów zarządzania tymi gruntami. Księgowaniem w systemie SEG objęte są dokumenty oraz dane dotyczące działek i użytków zaewidencjonowane w systemie po 1 stycznia 2002 roku. Baza systemu SEG składa się z części ewidencyjnej, nazywanej również SEG, księgowej (PRO), przejściowej z SEG do FIX, której nadano nazwę SPRZEG.

Część SEG zawiera dane dotyczące dokumentów, działek i użytków na działkach oraz ich powiązań z dokumentami wprowadzonymi przez użytkownika systemu w ramach rejestracji określonych operacji. W tej części ewidencjonowana jest pełna informacja rzeczowa o gruntach i o zmianach na tych gruntach poczynając od rejestracji ich przejęcia lub nabycia do Zasobu. Struktura części SEG jest jedyną w systemie SEG, do której użytkownik ma bezpośredni dostęp. W części PRO zapisywane są dane rzeczowe (udziały i powierzchnie) oraz księgowe (wartości szacunkowe i/lub wartości nabycia) dotyczące działek i użytków. Stany działek i użytków w strukturze PRO generowane są wyłącznie programowo na określony przez użytkownika miesiąc rozrachunkowy. Jak już nadmieniono, część SPRZEG stanowi strukturę przejściową wiążącą podsystem SEG z podsystemem FIX i zawiera analitykę obrotów dotyczących gruntów dla systemu FIX. Struktura SPRZEG wypełniana jest wyłącznie programowo poleceniem użytkownika aplikacji SEG zaksięgowania określonego dokumentu lub grupy dokumentów, przy czym generowane są odpowiednie dokumenty księgowe powiązane z właściwymi scenariuszami księgowania na określone konta w systemie FIX.

Ostatnim podsystemem jest *Środki Trwale* (ST), a jego podstawowe zadanie to obsługa wielozakładowego ewidencjonowania majątku rzeczowego oraz wspomaganie procesów zarządzania tym majątkiem.

Zintegrowany System Informatyczny (ZSI) obejmuje *Słowniki Centralne* (SC) do systemów SEG, WN, FIX zlokalizowane w Centrali KOWR i *Moduł Transmisji* (TR). Systemy SEG, WN, FIX posiadają również własne słowniki lokalne. *Moduł aktualizacji słowników centralnych* umożliwia korektę wybranego zakresu danych. TR przeznaczony jest do przenoszenia pomiędzy Oddziałami Terenowymi danych działek ewidencyjnych, umów i rozrachunków z kontrahentami, w sytuacji gdy następuje zmiana właściwości terytorialnej dla nieruchomości¹⁰⁷.

W 2021 roku w KOWR wdrożona została elektroniczna obsługa procesu obsługi dokumentacji. Minister Cyfryzacji i Wojewoda Podlaski zawarli porozumienie którego przedmiotem jest wdrażanie systemu EZD PUW. Minister cyfryzacji powierzył realizację zadań podległej jednostce

¹⁰⁷ Instrukcja użytkownika SEG.

Naukowej i Akademickiej Sieci Komputerowej Państwowy Instytut Badawczy - NASK. Przy współpracy NASK wdrożono jednolity system rozwijany na zasadach niekomercyjnych, będący narzędziem wymiany informacji oraz usprawniania funkcjonowania Urzędów. NASK zapewnia bieżące wsparcie i utrzymanie pod względem merytorycznym i technicznym (*Help Desk*).

Trzeba jeszcze dodać, że *Elektroniczne zarządzanie dokumentacją* (EZD) to system teleinformatyczny umożliwiający wykonanie czynności kancelaryjnych, dokumentowanie przebiegu załatwiania spraw oraz gromadzenie i tworzenie dokumentów elektronicznych. Zatem EZD jest narzędziem wspierającym obsługę kancelarii w systemie tradycyjnym.

10.6. Zarządzanie wiedzą i organizacja baz danych

Komunikacja w określonej organizacji jest istotną strategią zarządzania. Natomiast zarządzanie wiedzą to sposób na rozwój interpersonalny, a zatem rozwój całej organizacji. W ujęciu praktycznym wiedzę można określić jako „użytkową informację”, która to pozwala podejmować trafniejsze decyzje oraz wносить istotny wkład w procesy twórczego myślenia oraz dialogu, zachodzące w organizacji¹⁰⁸. Wiedza to zbiór reguł i informacji pozwalający na wnioskowanie na podstawie danych bądź sytuacji¹⁰⁹. W latach 90-tych powstaje koncepcja *Knowledge - Based Economy*, czyli gospodarka oparta na wiedzy. Wiedza jest ściśle powiązana z takimi elementami zarządzania, jak: kultura organizacyjna, style kierowania, procedury działania, warunki pracy, kwalifikacje pracowników.

Platforma intranetowa stanowi swojego rodzaju bazę informacyjną KOWAR-u, pozwalającą pracownikom na wymianę informacji, korzystanie z poczty elektronicznej. Sieć lokalna, czyli *intranet* jest rozwiązaniem pozwalającym w jednym miejscu skoncentrować wiele systemów oraz informacji wykorzystywanych w codziennym funkcjonowaniu tej organizacji¹¹⁰. Trzeba tu nadmienić, że *intranet* w KOWR jest platformą programową umożliwiającą pracownikom czerpać informacje w dowolnym czasie.

Można stwierdzić że przyczyny ludzkiej nieracjonalności tkwią w wadliwości ludzkiego myślenia oraz braku odpowiedniej wiedzy. Należy podkreślić, że odpowiedzialność i kompetencje muszą iść w parze. Konieczne jest przekształcenie wiedzy pracownika w jego odpowiedzialność we wnętrzu przedsiębiorstwa i na jego stanowisku pracy¹¹¹. Wiedza staje się głównym źródłem kompetencji, siły i bogactwa i musi być skierowana na doskonalenie jakości, które jest procesem odnoszącym się do wszystkich obszarów działania organizacji. Wiedza jest to płynna mieszanka wyrażonego doświadczenia, wartości, informacji wypływających z kontekstu i eksperckiej wnikliwości, które dostarczają podstaw do oceny i przyswajania nowych doświadczeń i informacji¹¹².

Znaczącą rolę w KOWR w tworzeniu wiedzy odgrywa menedżer. Menedżerowie to głównie kierownicy działów bezpośrednio podlegający dyrektorowi Oddziału. Menedżer sygnalizuje pracownikom problemy, oni zaś sami poszukują rozwiązania i w ten sposób generują wiedzę bez zewnętrznej ingerencji.

Wszędzie gromadzone są ogromne ilości danych, które z kolei muszą gdzieś być przechowywane i przetwarzane. Najwięksi giganci na rynku, to firmy, które zarabiają dzięki danym. Wymienić tu możemy *Google*, *Amazon* czy *Facebook*. Konieczne jest więc usystematyzowanie tego w zorganizowane bazy danych. Przy dużej ilości danych niezbędny już jest program do obsługi tych danych, czyli *System Zarządzania Bazą danych* (SZBD), a przykładami tego typu systemów są:

¹⁰⁸ Jashapara A., tłum. Sawicki J., *Zarządzanie wiedzą*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2006, s. 34.

¹⁰⁹ Strojny M., *Teoria i praktyka zarządzania wiedzą, Ekonomika i Organizacja Przedsiębiorstwa*, 2000, nr. 10.

¹¹⁰ Borowiecki R., Kwieciński M., *Informacja i Wiedza w Zintegrowanym Systemie Zarządzania*, Redakcja Naukowa, Zakamycze 2004, s. 28.

¹¹¹ Ibidem, s. 35.

¹¹² https://www.dbc.wroc.pl/Content/36488/Morawski_Zarządzanie_wiedza_organizacja.pdf.

Microsoft SQL Server - przeznaczony głównie dla dużych firm korporacyjnych i obsługi dużych baz danych;

Microsoft Access - kompletne narzędzie do zarządzania danymi przeznaczone dla małych i średnich firm; istnieje możliwość migracji do systemu *SQL Server*;

Oracle - SZBD przeznaczony dla małych, średnich, dużych i korporacyjnych firm;

IBM Informix - SZBD przeznaczone dla małych, średnich, dużych i korporacyjnych firm;

Sybase;

MySQL- darmowy system, stosowany szczególnie do obsługi stron internetowych;

PostgreSQL - darmowy, stosowany szczególnie do obsługi stron internetowych.

Baza danych musi zawierać taki zestaw tabel, aby udało się w nich przechowywać niezbędne informacje. Dodatkowo baza danych powinna dbać o bezpieczeństwo danych, by zapobiegać nieautoryzowanemu sięganiu do nich¹¹³. Głównymi komponentami typowego serwera bazy danych są co najmniej jeden procesor, przestrzeń dyskowa oraz pamięć. Systemem zastosowanym w KOWR jest *Oracle* z językiem PL/SQL. *Oracle PL/SQL* jest systemem dedykowanym dla jednostek państwowych RP o dużych możliwościach, służący do zarządzania relacyjnymi bazami danych¹¹⁴. Trzeba dodać jeszcze, że PL/SQL jest to rozszerzenie programistyczne języka SQL (*structured query language*) udostępniane przez Oracle i przeznaczone wyłącznie do obsługi produktów tej firmy¹¹⁵.

Projekt bazy danych, zawiera diagramy związków encji. Wyróżniamy np. encję EA_KONTRAHENCI, która będzie zawierała informacje o klientach, encję EA_ADRES, która zawierała informacje o adresach przypisanych do klienta oraz EA_TYP jako tabelę satelitarną zawierającą typ klienta (np. rolnik, spółka). Identyfikacja rekordów w tabelach następuje przy pomocy unikalnych kluczy podstawowych. W tabeli KONTARHENT może to być np. PESEL.

SZBD różnie radzą sobie z obsługą transakcji i niekiedy obsługa ta musi być wspomagana programistycznie przez projektanta bazy danych. Tworzenie bazy danych w obecnym czasie jest bowiem obowiązkiem przedsiębiorstwa, gwarantującym istnienie informacji, która to jest warunkiem jego egzystencji i rozwoju.

10.7. Potrzeba dalszego przystosowania stosowanych modułów do potrzeb użytkowników

Przesłanką do zastosowania badań w organizacji KOWR było rosnące niezadowolenie pracowników i niechęć do pracy z systemami komputerowymi, spowodowane narzekaniami na opóźnienia w realizacji zadań, które przedkładało się na relacje z przełożonymi i współpracownikami. Pierwszym etapem w przeprowadzonym badaniu ankietowym był wybór grupy badawczej. W badaniu uczestniczyło 37. pracowników etatowych, których staż pracy obejmował minimum trzy lata. Poprzez ankietę zgromadzono dane, których analiza pozwoliła wysnuć wnioski i określić zakres implementacji systemów informatycznych ze wskazaniem ograniczeń.

Wyniki badania ankietowego wykorzystujące wiedzę pracowników pozwoliły stworzyć model usprawniający moduły informatyczne. Ponadto dodatkowa analiza przeprowadzonych wywiadów z menadżerami potwierdziła, że spostrzeżenia i oczekiwania pracowników (respondentów) w punktu widzenia kierownictwa są słuszne i jak najbardziej zasadne.

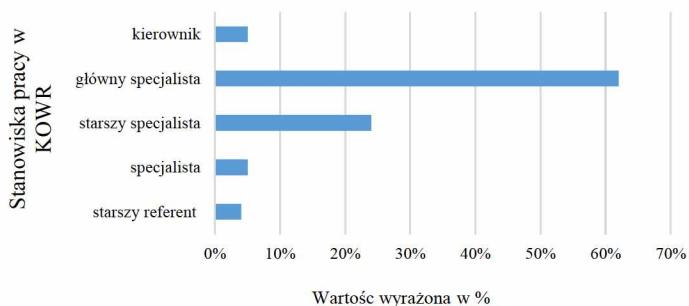
Jak już nadmieniono, w badaniu ankietowym uczestniczyło 37. respondentów, co stanowi 46% liczby pracowników KOWR OT w Opolu. Najliczniejszą grupę stanowią pracownicy wydziału WKUR i GZ (Wydział Kształtowania Ustroju Rolnego i Gospodarowania Zasobem). Drugą grupą, najmniej liczną, lecz najbardziej istotną stanowią pracownicy wydziału WFKiW (Wydział

¹¹³ Bryla B., Loney K., *Oracle Database 11g, Podręcznik administratora baz danych*, Helion 2010, s. 22.

¹¹⁴ Urman S., Harmand R., McLaughlin M., *Oracle database 10g, Programowanie w języku PL/SQL*, Helion 2008, s. 17.

¹¹⁵ Ibidem, s. 28.

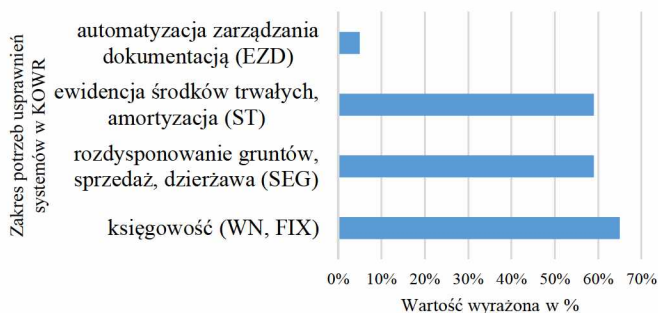
Finansowo-Księgowy i Windykacji). Ten wydział bazuje na wszystkich systemach ZSI w KOWR, gdyż ewidencjonuje każde zdarzenie mające finansowe skutki. Prowadzi ewidencję, rejestry, bazy danych, raporty, analizy i sprawozdania. Ostatnia sekcja biorąca udział w badaniu to WOP (Wydział Organizacyjno-Prawny). Liczebność grup biorących udział w badaniu ankietytowym według stanowisk w KOWR OT Opole pokazano na rysunku 10.7. Przeprowadzona ankieta, oprócz metryczki dotyczącej płci, stażu pracy, wykształcenia, komórki organizacyjnej oraz zajmowanego stanowiska przez respondentów (pytania 1-5), obejmowała dalsze pytania 6-12) dotyczące potrzeby udoskonalenia użytkowanych przez respondentów modułów programowych.



Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 14.

Rys. 10.7. Liczebność grup respondentów biorących udział w badaniu w odniesieniu do zajmowanego stanowiska pracy

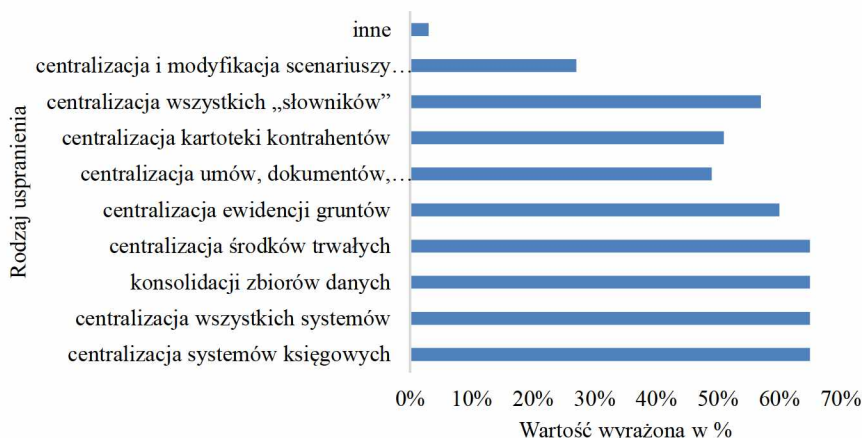
Na pytanie 6. *Które systemy wymagają usprawnień?* uczestnicy mieli do wyboru odpowiedzi: księgowość (WN, FIX), rozdysponowanie gruntów, sprzedaż, dzierżawa (SEG), ewidencja środków trwałych, amortyzacja (ST), automatyzacja zarządzania dokumentacją (EZD) - zob. rysunek 10.8.



Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 15.

Rys. 10.8. Systemy wymagające usprawnień

W pytaniu 7. dotyczącym *rodzaju usprawnień* w obrębie systemów ankietowani do wyboru mieli odpowiedzi: centralizacja systemów księgowych, centralizacja wszystkich systemów; konsolidacji zbiorów danych; centralizacja środków trwałych; centralizacja ewidencji gruntów; centralizacja umów, dokumentów, rozrachunków z kontrahentami; centralizacja kartoteki kontrahentów; centralizacja wszystkich „słowników”; centralizacja i modyfikacja scenariuszy księgowych; inne (zob. rysunek 10.9).



Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 16.

Rys. 10.9. Rodzaj usprawnień proponowanych dla KOWR

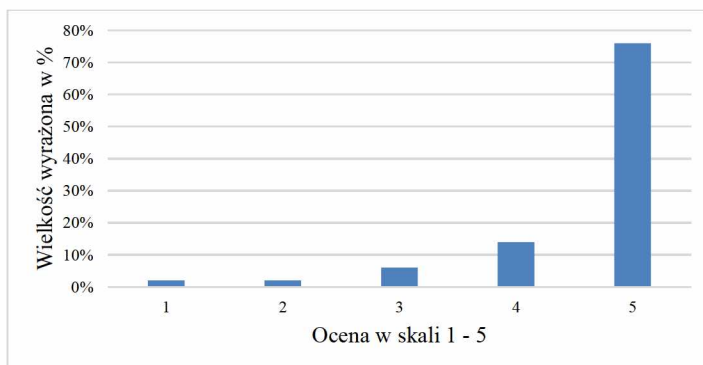
Następnym pytaniem 8. było *Czy systemy informatyczne wymagają utrzymania wielu zbiorów i różnorodnych aplikacji?* Możliwość wyboru odpowiedzi to: tak, nie, nie wiem. Wśród badanych 81% wskazało odpowiedź nie, 5% wskazało tak, 14% wskazało nie wiem. Potwierdza się zatem, iż występowanie dużej liczby zbiorów oraz aplikacji w praktycznym obiegu informacji (także w odniesieniu do propozycji wprowadzenia powiadomień sms) nie jest słuszne, a jest kłopotliwe i niepotrzebnie wprowadza chaos i dezorganizację.

Na pytanie 9. *Czy system sieciowy (Intranet) jest wystarczający do komunikacji w pracy?* Możliwe odpowiedzi to tak lub nie. Wyniki uzyskane od ankietowanych stanowią przybliżone wartości, 54% osób odpowiedziało tak, a 46% nie.

W jakim zakresie system sieciowy nie spełnia funkcji komunikacji? To kolejne pytanie 10. podane w przeprowadzonej ankietyzacji. Odnosi się do pytania 9. i uzupełnia wypowiedzi 46% (17 osób) badanych, dla których *Intranet* jest niewystarczającym systemem komunikacji. Najczęściej wskazana odpowiedź to informacje zamieszczane są z opóźnieniem.

W zapytaniu 11. *Wskaz inne źródła do zdobywania i upowszechniania wiedzy (możliwość kilkakrotnego wyboru)* badani mieli do wyboru odpowiedzi: *Internet*; pozyskanie wiedzy od menedżera (burza mózgów); inne (opisz jakie). Dla 46% opiniodawców to *Internet* jest świetnym źródłem uzupełniania wiedzy, 41% pozyskuje wiedzę od menedżera, a 13% z innych źródeł. Wśród najwyższej punktowanej i prawie jednogłośnie odpowiedzi czerpania wiedzy z innych źródeł było czerpanie informacji od pracowników innych oddziałów terenowych.

Ostatnie pytanie było prośbą o ocenę wcześniejszych propozycji i brzmiało: *Czy podane sugestie wpłyną na usprawnienie funkcjonowania systemów?* (5 to ocena najwyższa), skala: 1; 2; 3; 4; 5. Pracownicy w 76% wskazali ocenę 5, w 14% ocenę 4, w 6% ocenę 3, w 2% ocenę 1 i 2.



Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 17.

Rys. 10.10. Ocena wpływu sugestii na usprawnienia funkcjonowania systemów

10.8. Przeprowadzenie wywiadu i zakres potrzeb integracji eksploatowanych modułów

Dnia 5.02.2024 r. w budynku KOWR OT w Opolu przeprowadzono wywiad na temat: *Jak funkcjonują systemy informatyczne w KOWR?* Pytania kierowane były przede wszystkim do kierowniczki Wydziału Finansowo-Księgowego i Windykacji. Wybór tego menedżera był istotny dla sprawy z racji posiadanego doświadczenia i długoletniej pracy w KOWR. Zdaniem osoby z którą przeprowadzono wywiad przede wszystkim trzeba poprawić systemy SEG, WN, FIX i ST. Usprawnienia powinny dotyczyć kartotek kontrahentów, scenariuszy księgowych oraz środków trwałych. Jedynie system zarządzania dokumentacją (EZD) wydaje się bez zastrzeżeń - stwierdziła.

Systemy wspierające działalność statutową KOWR są rozproszone terytorialnie, autonomiczne, a dane w nich przetwarzane są niezależne, mogą nie być jednoznaczne i w pełni porównywalne pomiędzy jednostkami. W sytuacji gdy analityka danych musi objąć więcej niż jedną jednostkę terytorialną zebranie, skonsolidowanie i porównanie danych staje się wyzwaniem. Jednocześnie utrzymywanie rozproszonego środowiska informatycznego jest nieefektywne procesowo i kosztowo w kontekście jednostek realizujących ten sam zakres zadań i nie konkurujących ze sobą nawzajem. Wymagania stawiane w KOWR jako całości organizacji wobec centralizacji dotyczą:

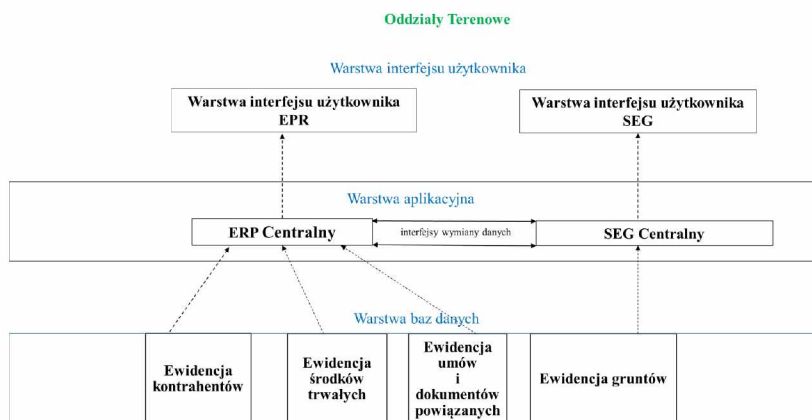
- ograniczenia liczności utrzymywanych aplikacji i bazy danych;
- centralizacji zarządzania aplikacjami i bazami danych;
- zwiększenia poziomu spójności danych, tj. zapewnienia, że stan bazy danych jest zgodny ze stanem reprezentowanego przez nią fragmentu rzeczywistości;
- konsolidacji zbiorów danych i zagwarantowania interoperacyjności systemu na poziomie całej organizacji;
- umożliwienia sprawnej analizy i raportowania w ujęciu krajowym, nie w perspektywie jednostki terenowej;
- zagwarantowania niezakłóconej ciągłości działań operacyjnych w trakcie centralizacji.

Dokonanie udoskonalenia stosowanych systemów, a przede wszystkim ich integracja w formie pakietu ZSI klasy ERP zapewni pozyskanie odpowiedniej, wiarygodnej, i szybkiej informacji zarządczej oraz umożliwi przeprowadzenie centralizacji zasobów gruntów rolnych prowadzonych w imieniu Skarbu Państwa i sprowadzonych do pełnej centralizacji:

- kartoteki kontrahentów,
- wszystkich słowników,

- modyfikacji scenariuszy księgowych,
- środków trwałych,
- ewidencji gruntów.

Centralizacji i migracji podlegać będą dane 17. oddziałów terenowych KOWAR-u. Migracja danych przeprowadzona zostanie ze struktur rozproszonych lokalnie do modułów centralnych. W ramach całej organizacji KOWR, model architektury systemów wspomagających gospodarowanie zasobem gruntów, zwanych krótko Zasobem, zakłada centralizację aplikacyjną i bazodanową, poprzez konsolidację danych z instalacji lokalnych do wspólnej instalacji centralnej. W docelowym modelu każda jednostka terenowa będzie miała odpowiedni poziom dostępu do zakresu danych, którymi zarządza oraz do danych, które może analizować i przeglądać. Schemat takie rozwiązania zaprezentowano na rysunku 10.11.



Źródło: Źródło: Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, op. cit., rys. 18.

Rys. 10.11. Schemat scentralizowanej architektury systemów wspierających gospodarowanie Zasobem

W systemie zintegrowanym ze wspólną bazą danych, będą zaprowadzone wspólne kartoteki kontrahentów, umów, dokumentów, środków trwałych, rozrachunków, umożliwiające realizację zadań KOWR zarówno w perspektywie oddziałowej jak i krajowej. Natomiast zakres funkcjonalny podsystemu zastępującego SEG nie ulegnie zmianie. Model architektury sprzętowej zakłada że, pracownik musi posiadać dostęp do Portalu Pracownika poprzez przeglądarkę internetową przy czym:

- musi mieć możliwość pracy zdalnej w projektowanym systemie z wykorzystaniem protokołu VPN (*Virtual Private Network*),
- wszystkie dane w systemie muszą być obsługiwane w bazie danych umożliwiające dostęp do danych za pomocą języka SQL,
- musi zostać zapewniona współpraca ze stacjami roboczymi z systemami operacyjnymi klasy Windows 7/8/10/11 (wersje 32, 64 i bitowe). Ponadto wymagana jest współpraca z oprogramowaniem biurowym MS Office 2007/210/2013/2016/365.

Konieczne jest posiadanie oprogramowania wirtualizacyjnego firmy *Vmware* oraz systemu pocztowego *MS Exchange*. W przypadku interfejsu *web* musi on być wyświetlany przez przeglądarki Internetowe: *Explorer*, *FireFox*, *Google Chrom*, *Safari*, *Microsoft Edge* w najnowszych wersjach. Dane muszą być eksportowane i importowane w formatach: txt, csv, xls, doc, xml, pdf, xlsx, docx.

System musi zapewniać obsługę kodów kreskowych formacie „CODE 128”, QR przy współpracy z posiadanymi drukarkami: TSC TTP-245 plus, TSC-TTP-247 oraz czytnikami Honewell HHP 6100. Celowe jest wykorzystanie obszarów i licencji modułów istniejącego oprogramowania posiadanego w organizacji KOWR.

W wyniku przeprowadzonego rozpoznania rynku w zakresie software klasy ERP jako godne uwagi uznano pakiety dużych firm informatycznych, mających doświadczenie w projektowaniu i wdrażaniu, a w szczególności przedsiębiorstw:

Asseco Business Solution (ABS),
Comarch,
Oracle Polska,
Simple,
SAP Polska.

Przed wyborem założono kryteria jakie musi spełnić pożądany system standardowy, a ponadto trzeba uwzględnić jeszcze rozwiązania wdrażane i eksploatowane w jednostkach administracji publicznej (Ministerstwo Finansów, Agencja Rezerw Materiałowych, Służba Celna, ARIMR). Jako założenia projektowe wybrano zintegrowane systemy informatyczne wspierające zarządzanie klasy ERP, a mianowicie:

SAP S/4HANA,
Microsoft Dynamics 365,
Fusion ERP Cloud.

Należy jednak podkreślić, że spełnienie wszystkich kryteriów i wymagań funkcjonalnych w przypadku systemów ERP dla każdej z jednostek administracyjnych organizacji KOWR nie jest możliwe. Oferowane bowiem na rynku komputerowym standardowe pakiety klasy ERP posiadają zróżnicowane struktury i standardy rozwiązań programowych w zależności od przyjętego przez siebie rozłożenia akcentów merytorycznych oraz polityki rozwoju systemu producenta.

Proces wdrożenia systemu zintegrowanego należy podzielić na etapy i rozpatrzyć decyzyjnie alternatywne rozwiązania. W wyborze należy kierować się rozwiązaniami wykorzystywanymi obecnie w jednostkach administracji publicznej. Oparte one były przeważnie o analizę największych dostawców systemów w Polsce. Istotne jest również zapoznanie się z postępowaniami przetargowymi na zakup systemów przeprowadzonymi w dużych podmiotach publicznych. Biorąc to pod uwagę wyłoniono trzy rozwiązania do szczegółowej analizy:

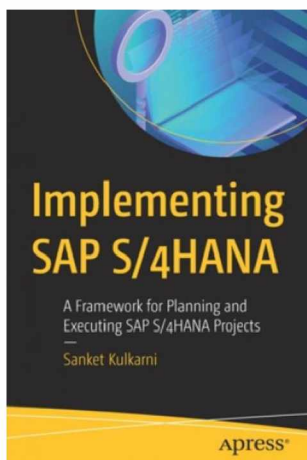
Fusion EPR Cloud - producent Oracle. Rozwiązanie chmurowe z możliwością instalacji na infrastrukturze klienta (*Cloud Customer*). System wdrażany przez Asseco Poland, *Sygnity Business Solution*, ATOS.

Microsoft Dynamics 365 - producent Microsoft. Rozbudowany system klasy EPR wspierający większość obszarów funkcjonalnych. Rozwiązanie korzysta z interfejsu użytkownika uruchamianego w przeglądarce internetowej. W Polsce wdrażany przez: Arcus, Asseco Poland, *Sygnity Business Solution*.

SAP S/4HANA - producent SAP. Rozbudowany system klasy EPR, który stanowi nową wersję systemu SAP, opartą na zaawansowanej platformie *inmemory* SAP HANA oraz interfejsie SAP Fiori. Wdrażany jest przez producentów: Asseco Poland, *Sygnity*. Bliższe informacje dotyczące przystosowania tego pakietu do warunków obiektu inwestora można znaleźć w publikacji „*Implementing SAP S/4HANA*”, której okładkę pokazano na rysunku 12¹¹⁶. Implementacja *SAP S/4 HANA* jest niezbędną wielofunkcyjną aplikacją bazodanową dla branży biznesowej. Kompletny przewodnik dotyczący implementacji tej aplikacji znajduje się na stronie:

<https://www.impactqa.com/blog/sap-s-4-hana-implementation-a-complete-guide/>.

¹¹⁶ https://www.libristo.pl/pl/ksiazka/implementing-sap-s-4hana_21353804.



* * *

W wyniku przeprowadzonej analizy dostawców systemów zintegrowanych opracowano szacunkową wycenę systemów najlepiej spełniających wymagania KOWR. Opracowano koszt zakupu trzech proponowanych systemów w podziale na koszt licencji oprogramowania oraz poszczególnych części wdrożenia (zob. tabela 10.2).

Tab. 10.2. Szacunkowy koszt zakupu rekomendowanych systemów (netto) w PLN

Zakres	<i>SAP S/4HANA</i>	<i>Microsoft Dynamics 365</i>	<i>Fusion EPR Cloud</i>
Licencje oprogramowania	2 200 000	2 000 000	2 400 000
Usługi wdrożenia (100%)	10 000 000	9 000 000	11 000 000
Analiza przedwdrożeniowa (10%)	1 000 000	900 000	1 100 000
Konfiguracja systemu (45%)	4 500 000	4 050 000	4 950 000
Migracja danych (15%)	1 500 000	1 350 000	1 650 000
Dokumentacja (8%)	800 000	720 000	880 000
Szkolenia (15%)	1 500 000	1 350 000	1 650 000
Prowadzenie projektu (5%)	500 000	450 000	550 000
Stabilizacja systemu (2%)	200 000	180 000	220 000
Suma kosztów	12 200 000	11 000 000	13 400 000

Źródło: Opracowanie własne na podstawie oferowanych szacunkowych cen rynkowych w roku 2024.

Na zakończenie dodam, że wskazując postępowanie w zakresie wyboru i wdrożenia systemu integrującego funkcjonalności eksploatowanych modułów w organizacji KOWR, przy uwzględnieniu kosztów oraz konieczności przejścia poprzez procedurę Zamówień Publicznych, wydaje się za najbardziej celowe zakupienie i wdrożenie systemu *SAP S/4HANA* firmy SAP.

11. Wielokryterialny ranking obiektów



11.1. Zastosowane wzory i oznaczenia

Wytworzenie określonej wartości *produktu krajowego brutto* możemy uznać za *zjawisko złożone* zależne między innymi od innych cech ekonomicznych, czy też dotyczących ochrony środowiska. Jako przykład dla porównania wielokryterialnego obiektów przyjęto 16. województw Polski. Najpierw dla danego obiektu musimy sporządzić oceny z punktu widzenia wielu cech - zmiennych, obliczyć zmienną syntetyczną, zwaną też agregatową, a w kolejności dla wszystkich obiektów zbudować ranking. Ze względu na to, że zmienne mają odmienne miana i wykazują różny rząd wielkości zachodzi potrzeba ich unormowania. W budowie rankingu obiektów z punktu widzenia zjawiska złożonego wyróżniamy następujące zbiory zmiennych¹¹⁷:

W – wyjściowych,

X – diagnostycznych (po procesie weryfikacji zmiennych),

Y – opisujących,

Z – unormowanych,

Q – syntetycznych.

Do weryfikacji zmiennych objaśniających zjawiska złożonego formułowanego przez zmienną objaśnianą możemy zastosować jedną z metod opisanych w części pierwszej „Zagadnienia wstępne i i wybrane modele ekonometryczne” (rozdział 1.6. „Wybór kombinacji zmiennych objaśniających”)¹¹⁸. Dla potrzeb rankingu obiektów zbiór zmiennych diagnostycznych dzielimy na podzbiory: stymulanty, destymulanty, nominanty. Wzrost wartości stymulanty wpływa na wzrost zjawiska złożonego, a wzrost destymulanty na jego spadek. W praktyce nominanty tworzą określony przedział liczbowy. Istnieją różne grupy metod unormowania zmiennych, między innymi grupa oparta o formuły bazujące na przekształceniu ilorazowym, które dotyczą stymulant i mają następujące odniesienia dotyczące zmiennych:

a) odchylenie standardowe (stosowane w standaryzacji) dla $S(X_j) > 0$:

$$z_{ij} = \frac{x_{ij} - X_j}{S(X_j)} \quad \text{oraz} \quad z_{ij} = \frac{x_{ij}}{S(X_j)}$$

b) rozstęp (wykorzystywany w unitaryzacji) przy warunku: $\max_i x_{ij} > \min_i x_{ij}$

$$z_{ij} = \frac{x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \quad \text{oraz} \quad z_{ij} = \frac{x_{ij} - \bar{X}_j}{\max_i x_{ij} - \min_i x_{ij}} \quad \text{oraz} \quad z_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}$$

c) maksymalna lub minimalna wartość zmiennej przy warunku: $\max_i x_{ij} \neq 0$

$$z_{ij} = \frac{x_{ij}}{\max_i x_{ij}} \quad \text{oraz} \quad z_{ij} = \frac{x_{ij}}{\min_i x_{ij}}$$

¹¹⁷ W opracowaniu tego rozdziału bazowano na pojęciach i wzorach zamieszczonych w rozdziale 10. pracy pod red. naukową Karola Kukuly *Badania operacyjne w przykładach i zadaniach*, Wydawnictwo Naukowe PWN, Warszawa 2004.

¹¹⁸ Duczmał M., Wornalkiewicz W., *Wstęp do ekonometrii i badań operacyjnych*, Wydawnictwo Instytut Śląski w Opolu, Opole 2010.

d) średnia dla warunku: $\bar{X}_j \neq 0$

$$z_{ij} = \frac{x_{ij}}{\bar{X}_j}$$

e) suma realizacji zmiennej dla warunku: $\sum_{i=1}^r x_{ij} \neq 0$

$$z_{ij} = \frac{x_{ij}}{\sum_{i=1}^r x_{ij}}$$

f) długość wektora realizacji zmiennej dla warunku: $\sum_{i=1}^r x_{ij}^2 > 0$

$$z_{ij} = \frac{x_{ij}}{\left[\sum_{i=1}^r x_{ij}^2 \right]^{0,5}}$$

W badaniu wykorzystuje się jeden z ostatnich okresów zebranego materiału statystycznego. Z tego względu nie ma potrzeby oznaczania obserwacji x_{ij} zmiennej X_j w obiekcie O_i symbolem t , a

$$[x_{ij}] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1s} \\ x_{21} & x_{22} & \dots & x_{2s} \\ \dots & \dots & \dots & \dots \\ x_{r1} & x_{r2} & \dots & x_{rs} \end{bmatrix}$$

macierz dwuwymiarową przyjmuje postać:

11.2. Baza danych statystycznych

W przykładowej prezentacji budowy rankingu województw posłużono się danymi z roku 2007 zebranymi na podstawie roczników statystycznych poszczególnych województw¹¹⁹. Zarówno zmienna objaśniana jak i zmienne objaśniające wchodzą w skład opracowanego przez autora modelu poniższego ekonometrycznego PKB:

$\hat{X}_1 = -15763,3 - 0,1292X_5 + 0,8105X_6 + 196,194X_8 - 11,3945X_{10} + 8,0591X_{12} + 7,3776X_{13} + 1,099X_{20}$
gdzie:

X_1 – produkt krajowy brutto (ceny bieżące) na 1 mieszkańca w zł;

X_5 – odpady wytworzone (w ciągu roku, z wyłączeniem odpadów komunalnych) w tys. t.;

X_6 – ścieki przemysłowe i komunalne odprowadzone do wód lub do ziemi w hm³;

X_8 – stopa bezrobocia rejestrowanego (stan w dniu 31 XII) w %;

X_{10} – pracujący w budownictwie (stan w dniu 31 XII) w tys.;

X_{12} – przeciętne miesięczne wynagrodzenie w przemyśle brutto w zł;

X_{13} – przeciętne miesięczne wynagrodzenie w budownictwie brutto w zł;

X_{20} – nakłady inwestycyjne w przemyśle (ceny bieżące) w mln zł.

Wartości zmiennych w odniesieniu do roku 2007 i 16. województw podaje tabela 11.1.

Tab. 11.1. Wartości zmiennych modelu PKB

W	Województwo	X ₁	X ₅	X ₆	X ₈	X ₁₀	X ₁₂	X ₁₃	X ₂₀
1	łódzkie	25521	3895,2	129,4	11,2	46,7	2262,39	1928,20	5479,1
2	mazowieckie	44381	6788,4	2700,6	9,0	129,3	3128,25	3042,25	7602,7
3	małopolskie	24111	9026,9	764,8	8,7	71,9	2581,55	2069,37	4919,6
4	śląskie	29497	41408,2	372,3	9,2	106,7	3230,68	2183,35	11999,2

¹¹⁹ Duczmal M., *Efektywność działania obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski w Opolu, Opole 2013, rozdział 5.6. *Określenie modelu ekonometrycznego*.

5	lubelskie	18779	5122,5	157,6	13,0	31,2	2343,07	1927,93	1517,8
6	podkarpackie	19024	1276,1	201,8	14,2	32,9	2300,37	1919,83	2607,8
7	podlaskie	20396	1057,2	41,2	10,4	16,8	2339,81	2136,29	1493,5
8	świętokrzyskie	21130	1637,7	1042,4	14,9	21,8	2415,57	1797,03	1652,2
9	lubuskie	24733	727,0	36,0	14,0	16,7	2257,42	1784,03	1863,3
10	wielkopolskie	29279	4851,6	1694,4	7,8	81,4	2481,16	2116,09	6271,9
11	zachodniopomorskie	25324	6841,0	1645,9	16,4	34,6	2407,72	1907,41	2533,5
12	dolnośląskie	29739	34299,7	231,7	11,4	59,8	3024,88	2293,91	7063,2
13	opolskie	22347	1536,5	91,7	11,9	20,2	2549,77	1973,60	1638,1
14	kujawsko-pomorskie	24301	2584,4	117,9	14,9	39,0	2281,94	1984,52	3048,2
15	pomorskie	27373	2466,9	220,9	10,7	48,4	2732,64	2271,08	2719,1
16	marmińsko-mazurskie	21005	894,8	66,2	18,7	24,1	2158,18	1946,59	1935,7

Źródło: Opracowanie własne.

Ponieważ model ekonometryczny jest efektem wcześniejszej weryfikacji zmiennych objaśniających, przyjmijmy, że na produkt krajowy brutto (X_1) jako stymulanty wpływają zmienne X_{10} , X_{12} , X_{13} , X_{20} , destymulantami są X_5 , X_6 , a nominantą jest zmienna X_8 przyjmująca wartości przykładowo z przedziału $<5\%, 12\%>$.

11.3. Unormowanie zmiennych

W procesie transformacji (unormowania) zmiennych zastosowano spośród zmiennych diagnostycznych dla stymulant (S) metodę standaryzacji zerowej (czyli trzecią spośród wymienionych w punkcie b). Natomiast destymulanty (D) unormowano według wzoru:

$$z_{ij} = \frac{\max_j x_{ij} - x_{ij}}{\max_j x_{ij} - \min_j x_{ij}} \quad X_j \in D$$

Ponieważ istnieje przedział wartości nominalnych $<c_1, c_2> = <5, 12>$ unormowanie przepro-

$$z_{ij} = \begin{cases} \frac{x_{ij} - \min_j x_{ij}}{c_{1j} - \min_j x_{ij}} & \text{gdy } x_{ij} < c_{1j} \\ 1 & \text{gdy } c_{1j} \leq x_{ij} \leq c_{2j} \\ \frac{x_{ij} - \max_j x_{ij}}{c_{2j} - \max_j x_{ij}} & \text{gdy } x_{ij} > c_{2j} \end{cases} \quad X_j \in N$$

wadzono według formuły:

Za pomocą metody unitaryzacji zerowej otrzymujemy przedział unormowania $<0, 1>$.

11.4. Obliczenie wartości zmiennych syntetycznych

Mając unormowane cechy (zmienne) diagnostyczne możemy przystąpić do agregacji ocen wielokryterialnych każdego z szesnastu województw. W tym względzie możemy posłużyć się dwoma sposobami:

- sumowanie ocen dla $r = 16$ obiektów oraz $i = 8$ cech unormowanych:

$$Q_i = \sum_{j=1}^s z_{ij}$$

- obliczenie średniej z unormowanych ocen:

$$Q_i = \frac{1}{s} \sum_{j=1}^s z_{ij}$$

Przyjmujemy dla naszego przykładu wzór pierwszy, a następnie po uporządkowaniu nierosnącym wartości Q_i otrzymujemy ranking obiektów. Pomocna w podziale obiektów na trzy podgrupy jest stała U :

$$U = \frac{\max_i Q_i - \min_i Q_i}{3}$$

$Q_i \in (\max_i Q_i - U, \max_i Q_i - U > \text{najlepsze}$

$Q_i \in (\max_i Q_i - 2U, \max_i Q_i - U > \text{przecietne}$

$Q_i \in (\min_i Q_i - U, \max_i Q_i - 2U > \text{najgorsze}$

Jeśli Q_i jest w poniższym przedziale to:

Tab. 11.2. Tabela pomocnicza do obliczenia wartości zmiennej syntetycznej

Obiekt <i>i</i> -ty	Zmienne diagnostyczne i unormowane														Q _i
	X _{i5}	Z _{i5}	X _{i6}	Z _{i6}	X _{i8}	Z _{i8}	X _{i10}	Z _{i10}	X _{i12}	Z _{i12}	X _{i13}	Z _{i13}	X _{i20}	Z _{i20}	
	D				N		S								
W ₁	3895,2	0,92	129,4	0,96	11,2	1,00	46,7	0,27	2262,3 9	0,10	1928,2 0	0,11	5479,1	0,38	3,74
W ₂	6788,4	0,85	2700,6	0,00	9,0	1,00	129,3	1,00	3128,2 5	0,90	3042,2 5	1,00	7602,7	0,58	5,34
W ₃	9026,9	0,80	764,8	0,73	8,7	1,00	71,9	0,49	2581,5 5	0,39	2069,3 7	0,23	4919,6	0,33	3,96
W ₄	41408,2	0,00	372,3	0,87	9,2	1,00	106,7	0,80	3230,6 8	1,00	2183,3 5	0,32	11999,2	1,00	4,99
W ₅	5122,5	0,89	157,6	0,95	13,0	0,85	31,2	0,13	2343,0 7	0,17	1927,9 3	0,11	1517,8	0,00	3,11
W ₆	1276,1	0,99	201,8	0,94	14,2	0,67	32,9	0,14	2300,3 7	0,13	1919,8 3	0,11	2607,8	0,11	3,09
W ₇	1057,2	0,99	41,2	1,00	10,4	1,00	16,8	0,00	2339,8 1	0,17	2136,2 9	0,28	1493,5	0,00	3,44
W ₈	1637,7	0,98	1042,4	0,62	14,9	0,57	21,8	0,05	2415,5 7	0,24	1797,0 3	0,01	1652,2	0,02	2,48
W ₉	727,0	1,00	36,0	1,00	14,0	0,70	16,7	0,00	2257,4 2	0,09	1784,0 3	0,00	1863,3	0,04	2,83
W ₁₀	4851,6	0,90	1694,4	0,38	7,8	1,00	81,4	0,57	2481,1 6	0,30	2116,0 9	0,26	6271,9	0,45	3,87
W ₁₁	6841,0	0,85	1645,9	0,40	16,4	0,34	34,6	0,16	2407,7 2	0,23	1907,4 1	0,10	2533,5	0,10	2,18
W ₁₂	34299,7	0,17	231,7	0,93	11,4	1,00	59,8	0,38	3024,8 8	0,81	2293,9 1	0,41	7063,2	0,53	4,23
W ₁₃	1536,5	0,98	91,7	0,98	11,9	1,00	20,2	0,03	2549,7 7	0,37	1973,6 0	0,15	1638,1	0,01	3,52
W ₁₄	2584,4	0,95	117,9	0,97	14,9	0,57	39,0	0,20	2281,9 4	0,12	1984,5 2	0,16	3048,2	0,15	3,11
W ₁₅	2466,9	0,96	220,9	0,93	10,7	1,00	48,4	0,28	2732,6 4	0,54	2271,0 8	0,39	2719,1	0,12	4,21
W ₁₆	894,8	1,00	66,2	0,99	18,7	0,00	24,1	0,07	2158,1 8	0,00	1946,5 9	0,13	1935,7	0,04	2,22
Min	727,0		36,0		7,8		16,7		2158,2		1784,0		1493,5		
Max	41408,2		2700,6		18,7		129,3		3230,7		3042,3		11999,2		
Rozstęp	40681,2		2664,6		10,9		112,6		1072,5		1258,2		10505,7		

Gdzie: D - destymulanta, S - stymulanta, N - nominanta;

Q_i - zmienna agregatowa (syntetyczna) województwa *i*-tego;

c_{18}, c_{28} - dolna oraz górna wartość przedziału zmiennej nominalnej;

s - liczba zmiennych unormowanych.

Źródło: Opracowanie własne w Excelu.

Tab. 11.3. Ranking województw w świetle zmiennych objaśniających modelu ekonometrycznego PKB

Nazwa	Symbol	Q_i	Podgrupa
mazowieckie	W ₂	5,34	N
śląskie	W ₄	4,99	N
dolnośląskie	W ₁₂	4,23	N
pomorskie	W ₁₅	4,21	N
małopolskie	W ₃	3,96	P
wielkopolskie	W ₁₀	3,87	P
łódzkie	W ₁	3,74	P
opolskie	W ₁₃	3,52	P
podlaskie	W ₇	3,44	P
lubelskie	W ₅	3,11	G
kujawsko-pomorskie	W ₁₄	3,11	G
podkarpackie	W ₆	3,09	G
lubuskie	W ₉	2,83	G
świętokrzyskie	W ₈	2,48	G
marmińsko-mazurskie	W ₁₆	2,22	G
zachodniopomorskie	W ₁₁	2,18	G
min		2,18	
max		5,34	
U		1,05	
$\max Q_i - U$		4,28	
$\max Q_i - 2U$		3,23	

Gdzie: N – najlepsze, P – przeciętne, G – najgorsze.

Źródło: Opracowanie własne w Excelu.

12. Zastosowanie taksonomii wrocławskiej



12.1. Wstęp

Zagadnienie podziału określonej zbiorowości na grupy według podobieństwa kilku cech (zmiennych) jest tematem wielu publikacji. Jednakże dla celów dydaktycznych stanowią one skróty pewnych badań i z tego względu zawierają dość dużo niedomówień. Z tego powodu w niniejszym materiale podjęto się szczegółowego przedstawienia zastosowania taksonomii wrocławskiej z użyciem tego funkcjonalności arkusza kalkulacyjnego *Excel*.

Bazowano na danych statystycznych zebranych na etapie modelowania ekonometrycznego liniowego produktu krajowego brutto 45. podregionów Polski¹²⁰. Są to dane panelowe, czyli czasowo-przekrojowe, gdyż obejmują lata 2000-2006, a w ramach tych lat rozpatrywanych jest 11 następujących cech statystycznych:

- X_1 - ludność na 1 km² powierzchni ogólnej;
- X_2 - przyrost naturalny na 1000 ludności;
- X_3 - saldo migracji wewnętrznych i zagranicznych na pobyt stały (na 1000 ludności);
- X_4 - przestępstwa stwierdzone na 1000 ludności;
- X_5 - pracujący w tys. (rolnictwo, łowiectwo i leśnictwo, rybactwo);
- X_6 - pracujący w tys. (przemysł i budownictwo);
- X_7 - pracujący w tys. (usługi rynkowe i nierynkowe);
- X_8 - bezrobotni zarejestrowani ogółem w tys.;
- X_9 - mieszkania oddane do użytkowania na 1000 ludności;
- X_{10} - miejsca noclegowe w turystycznych obiektach zbiorowego zakwaterowania w tys.;
- X_{11} - użytki rolne w tys. ha w roku poprzednim.

Stanowią one zmienne objaśniające zmiennej objaśnianej Y (produktu krajowego brutto na jednego mieszkańca w cenach bieżących w roku poprzednim w zł.). Produkt krajowy brutto jest końcowym rezultatem wszystkich podmiotów gospodarki narodowej i równa się sumie wartości dodanej brutto wszystkich sektorów własności oraz sektorów instytucjonalnych powiększonej o podatki od przedmiotów i pomniejszonej o dotacje do produktów¹²¹. Produkt krajowy zarówno brutto jak i netto liczony jest w cenach rynkowych (bieżących). Obecnie dla celów sprawozdawczych występuje 6 regionów grupujących określone województwa oraz 45 podregionów obejmujących jednostki szczebla powiatowego.

Nomenklatura jednostek terytorialnych dla celów statystycznych odpowiada nomenklaturze w Unii Europejskiej. W przykładowym zestawieniu statystycznym *Podstawowe dane o regionach w roku* występują 22 kolumny z informacjami charakteryzującymi nie tylko poszczególne regiony, lecz również w ramach nich podregiony. Spośród tych informacji wybrano takie które tworzą szeregi o długości siedmiu okresów (lata 2000-2006) zmiennych egzogenicznych (objaśniających). Dane czasowo-przekrojowe dotyczące podregionów dostępne są w rocznikach statystycznych lat 2000-2007.

¹²⁰ Wornalkiewicz W., *Modele ekonometryczne PKB obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013, rozdział 7. *Etap i procedury budowy modelu produktu brutto podregionów*.

¹²¹ Rocznik Statystyczny Rzeczypospolitej Polskiej 2007. Dział: Rachunki narodowe.

12.2. Przygotowanie danych testowych

Szeregi panelowe zmiennych X_1 - X_{11} w latach 2000-2006 (zapisane w pliku *Dane.xls* na dysku D:), były podstawą opracowania liniowego produktu krajowego brutto. Pobrano je z następujących tabel roczników statystycznych: wybrane dane o podregionach (lata 2000-2003), podstawowe dane o regionach (lata 2004-2007). Dla celu przedstawienia procedury taksonomii wrocławskiej wystarczyło wybrać tylko jeden rok i dlatego przykładowo przyjęto rok 2006, zapisując go w skoroszycie wieloarkuszowym - pliku *Tabele pomocnicze-1.xls*. Poszczególne podregiony i wartości zmiennych pokazano w tabeli 12.1.

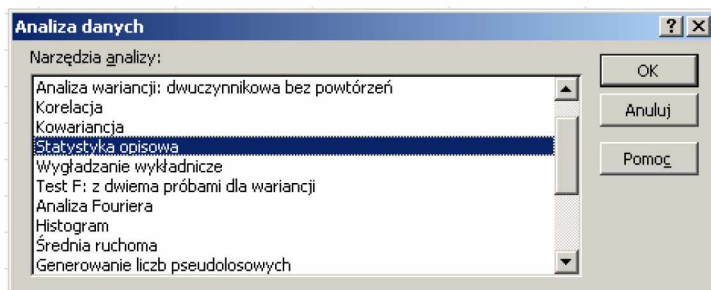
Tab. 12.1. Zmienne objaśniające (rok 2006)

Lp.	Podregion	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
1	łódzki	107,0	-2,5	0,5	256,0	93,5	76,5	82,5	65,3	1,8	4,7	552,7
2	piotrkowsko-skierniewicki	94,7	-0,8	-1,8	257,0	96,2	73,7	82,9	56,7	2,3	7,5	514,2
3	miasto Łódź	2593,0	-6,5	-2,4	434,0	2,1	60,9	147,9	38,7	1,6	3,6	19,2
4	ciechanowsko-płocki	80,5	0,1	-1,9	273,0	65,9	46,5	64,4	51,9	2,8	3,0	507,4
5	ostrołęcko-siedlecki	62,0	0,5	-2,8	300,0	103,6	37,2	67,0	53,2	1,9	4,9	714,3
6	warszawski	168,0	0,8	9,2	352,0	66,9	90,5	141,7	57,1	6,4	5,0	420,8
7	radomski	104,0	0,2	-2,6	309,0	79,3	39,5	64,5	73,9	2,2	2,9	375,7
8	miasto Warszawa	3291,0	-0,6	4,8	466,0	4,4	143,6	630,1	49,4	8,1	21,0	117,5
9	krakowsko-tarnowski	190,0	0,4	0,9	277,0	105,0	91,3	119,7	63,3	2,4	5,9	383,7
10	nowosądecki	149,0	3,3	-0,7	241,0	76,1	59,0	100,1	61,7	2,2	39,8	303,9
11	miasto Kraków	2314,0	-0,4	1,1	530,0	2,4	67,7	192,4	20,3	8,7	17,0	21,6
12	częstochowski (północnośląski)	175,0	-2,5	-1,0	299,0	22,8	53,8	59,0	32,7	2,4	3,6	141,4
13	bielsko-bialski (południowośląski)	276,0	1,0	1,0	310,0	12,2	64,0	78,2	25,0	3,3	19,1	74,8
14	centralny śląski	510,0	-1,5	-3,3	440,0	26,4	318,2	411,6	146,0	1,4	10,8	204,8
15	rybnicko-jastrzębski	472,0	1,7	-4,6	344,0	8,2	67,3	68,7	26,1	1,6	1,7	64,9
16	bialskopodlaski	51,8	-0,2	-4,3	266,0	42,9	12,7	30,5	21,5	1,6	5,4	309,1
17	chełmsko-zamojski	70,4	-1,8	-3,9	230,0	103,8	27,7	58,8	49,7	1,5	5,2	542,7
18	lubelski	123,0	0,1	-2,2	281,0	131,2	69,3	149,0	70,6	2,7	8,5	642,5
19	rzyszowsko-tarnobrzeczski	154,0	1,6	-1,4	207,0	81,1	101,1	132,8	72,2	2,6	6,7	366,8
20	krośnieńsko-przemyski	90,9	1,0	-2,5	185,0	76,4	58,5	96,1	73,0	1,7	13,5	386,1
21	białostocko-suwański	59,7	-1,0	-1,4	242,0	90,2	48,7	104,9	43,3	2,8	10,1	745,5
22	łomżyński	58,0	0,3	-5,0	201,0	48,1	11,8	25,4	18,5	1,7	1,6	352,0
23	świętokrzyski	109,0	-1,9	-2,7	320,0	143,5	81,4	133,9	99,4	1,3	7,9	556,5
24	gorzowski	62,4	1,3	-1,3	383,0	10,9	34,2	46,6	24,3	3,1	7,9	196,8
25	zielonogórski	79,6	1,0	-1,7	385,0	15,4	47,9	74,7	48,5	2,8	11,6	283,3
26	pilski	62,9	2,3	-2,4	313,0	23,7	37,8	39,0	25,5	2,0	3,7	289,1
27	poznański	102,0	2,7	4,9	253,0	67,0	123,1	121,6	50,3	3,6	15,8	674,4
28	kaliski	112,0	1,2	-1,3	263,0	73,4	80,0	76,0	43,9	2,1	3,6	479,7
29	koniński	98,1	1,1	-1,8	263,0	42,6	36,4	39,4	33,6	2,2	5,8	283,6
30	miasto Poznań	2158,0	0,0	-4,4	522,0	1,8	61,6	163,6	15,8	5,0	7,7	33,9
31	szczęciński	88,1	0,3	-0,3	369,0	24,9	71,5	139,6	81,1	2,7	56,2	614,1
32	koszaliński	57,0	0,8	-2,4	339,0	17,5	36,2	62,2	57,7	3,6	52,5	379,3
33	jeleniogórsko-walbrzyski	126,0	-2,2	-3,2	390,0	33,9	99,3	130,1	108,1	1,5	32,4	457,2
34	legnicki	120,0	1,4	-3,4	431,0	14,7	53,9	62,4	30,5	2,2	2,5	185,6
35	wrocławski	86,1	1,3	5,8	269,0	23,2	42,6	35,1	23,6	4,1	2,8	263,7
36	miasto Wrocław	2167,0	-1,5	-0,5	553,0	1,4	49,3	150,3	23,3	6,9	7,3	70,1
37	opolski	111,0	-0,9	-4,6	312,0	49,7	74,5	111,2	60,1	1,3	7,5	553,9
38	bydgoski	114,0	0,6	-1,4	328,0	44,3	87,1	129,1	72,7	2,0	13,1	467,6
39	toruńsko-włocławski	116,0	0,8	-2,0	326,0	72,7	78,0	115,5	87,3	2,8	12,4	578,4
40	ślupski	58,4	2,8	-3,4	424,0	17,1	38,9	52,5	45,4	2,5	30,7	282,9
41	gdański	101,0	4,5	2,7	359,0	40,7	71,2	81,3	61,7	4,1	37,2	462,8
42	Gdańsk-Gdynia-Sopot	1790,0	-0,3	-2,6	466,0	2,2	59,0	156,1	18,8	6,4	16,0	33,1

43	elbląski	70,8	1,8	-3,9	316,0	25,8	42,7	51,7	47,6	2,0	8,2	348,0
44	olsztyński	59,3	1,7	-3,1	353,0	24,2	44,2	81,0	49,5	3,9	20,7	426,1
45	elcki	44,7	1,8	-3,8	370,0	15,8	15,8	26,4	30,5	3,5	9,4	224,6

Źródło: Opracowanie własne na podstawie tabeli A.2.1. „Dane czasowo przekrojowe panelu podregionów” w zakresie tylko roku 2006¹²².

Dalsze prace rozpoczniemy od statystyki opisowej dla naszego zestawu zmiennych, zapisanych dla wygody w *Excelu* jako X_1 - X_{11} . Skorzystajmy więc z zakładki „Analiza danych” w menu głównym *Excela* (zob. rysunek 12.1).



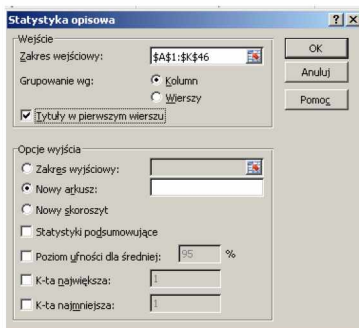
Źródło: Opracowanie własne w *Excelu*.

Rys. 12.1. Zakładka „Analiza danych”

Po „OK” pojawia się nam okno dialogowe w którym określamy:

- zakres wejściowy zmiennych (X_1 - X_{11}) zapisany w postaci adresowania bezwzględnego tj. $\$A\$1:\$K\46 ,
- zaznaczenie „*Tytuły w pierwszym wierszu*”,
- zaznaczenie opcji wyjścia „*Nowy arkusz*”.

Ponadto program Excel wymaga wskazania „*Statystyki podsumowujące*” (zob. rysunek 12.2).



Źródło: Opracowanie własne w *Excelu*.

Rys. 12.2. Okno dialogowe „Statystyka opisowa”

Otrzymany raport statystyki opisowej w *Excelu* na nowym arkuszu, w ramach tego samego skoroszytu „*Tabele pomocnicze-1.xls*”, wymagał ujednoliconego sformatowania liczb oraz szerokości kolumn (zob. rysunek 12.3).

¹²² Wornalkiewicz W., *Modele ekonometryczne PKB obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013, op. cit.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Cecha	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
2	Średnia	424,2	0,3	-1,4	333,5	47,2	66,4	110,8	51,3	3,0	12,8	353,5
3	Błąd standardowy	119,6	0,3	0,4	13,2	5,7	7,0	15,1	4,0	0,3	1,9	29,7
4	Mediana	107,0	0,5	-2,0	316,0	40,7	59,0	82,5	49,5	2,4	7,9	366,8
5	Tryb	#N/D!	0,8	-2,4	466,0	#N/D!	59,0	#N/D!	61,7	2,8	3,6	#N/D!
6	Odczylenie standardowe	802,3	1,8	2,9	88,4	38,0	46,9	101,5	26,6	1,8	12,9	199,5
7	Wariancja próbki	643622,2	3,4	8,7	7812,6	1446,8	2198,9	10295,4	706,0	3,1	166,8	39808,3
8	Kurtosis	4,6	3,2	3,4	0,0	-0,4	18,9	16,8	2,5	2,9	3,6	-0,8
9	Skośność	2,4	-1,0	1,7	0,7	0,7	3,7	3,7	1,2	1,8	2,0	0,0
10	Zakres	3246,3	11,0	14,2	368,0	142,1	306,4	604,7	130,2	7,4	54,6	726,3
11	Minimum	44,7	-6,5	-5,0	185,0	1,4	11,8	25,4	15,8	1,3	1,6	19,2
12	Maksimum	3291,0	4,5	9,2	553,0	143,5	318,2	630,1	146,0	8,7	56,2	745,5
13	Suma	19087,4	13,8	-61,1	15007,0	2125,1	2986,1	4987,5	2309,3	135,3	574,4	15906,3
14	Licznik	45,0	45,0	45,0	45,0	45,0	45,0	45,0	45,0	45,0	45,0	45,0

Źródło: Opracowanie własne w Excelu.

Rys. 12.3. Statystyka opisowa zmiennych X_1 - X_{11}

W celu wyłonienia niezbędnych zmiennych objaśniających sprawdzimy współczynnik zmienności będący modulem stosunku odchylenia standardowego do średniej danej zmiennej.

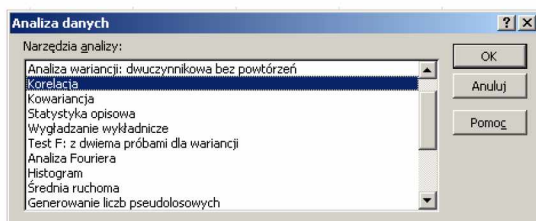
$$V_j = \left| \frac{S_j}{\bar{x}_j} \right|$$

Wyznaczenie na podstawie rysunku 12.3. wartości współczynnika zmienności V_j stanowi wstępny krok zbadania wystarczającego zróżnicowania danej j -tej zmiennej objaśniającej, który odpowiednio dla zmiennych X_1 - X_{11} jest w przybliżeniu następujący:

$$V_j | 0,5 | 0,2 | 0,5 | 3,8 | 1,2 | 1,4 | 1,1 | 1,9 | 1,7 | 1,0 | 1,8 |$$

Zatem wszystkie zmienne spełniają warunek: $V^* > 0,1$, tak więc obiekty (podregiony) uczestniczące w badaniu wykazują istotne zróżnicowanie pod względem 11 cech, zwanych tu zmiennymi.

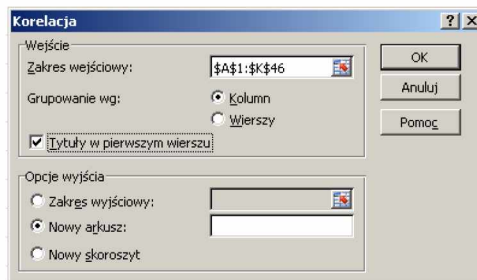
Zmienne do taksonomii wrocławskiej powinny wykazywać słabe skorelowanie względem siebie, sprawdzimy więc skorelowanie korzystając z funkcji „Korelacja” w ramach zakładki „Analiza danych” (zob. rysunek 12.4).



Źródło: Opracowanie własne w Excelu.

Rys. 12.4. Wskazanie funkcji „Korelacja”

W kolejnym kroku pojawia się nam okno dialogowe „Korelacja” w którym podajemy: zakres wejściowy danych, zaznaczenie tytułów w pierwszym wierszu oraz opcję wyjścia „Nowy arkusz” (zob. rysunek 12.5).



Źródło: Opracowanie własne w Excelu.

Rys. 12.5. Okno dialogowe „Korelacja”

W odpowiedzi uzyskujemy zestawienie symetryczne współczynników korelacji pomiędzy zmiennymi X_1 - X_{11} (zob. rysunek 12.6). Czcionką pogrubioną zaznaczono relacje zmiennych, dla których współczynniki korelacji są zbyt duże i dlatego te zmienne stopniowo wyeliminujemy z dalszych badań.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
X_1	1,000										
X_2	-0,418	1,000									
X_3	0,153	0,160	1,000								
X_4	0,709	-0,187	0,021	1,000							
X_5	-0,468	-0,056	0,019	-0,669	1,000						
X_6	0,185	-0,163	0,220	0,208	0,024	1,000					
X_7	0,650	-0,263	0,296	0,441	-0,144	0,724	1,000				
X_8	-0,273	-0,149	-0,021	-0,138	0,448	0,671	0,348	1,000			
X_9	0,676	0,080	0,553	0,609	-0,414	0,041	0,457	-0,358	1,000		
X_{10}	-0,019	0,272	0,114	0,173	-0,206	0,069	0,139	0,274	0,159	1,000	
X_{11}	-0,616	0,104	0,042	-0,607	0,766	-0,014	-0,206	0,475	-0,413	0,107	1,000

Źródło: Opracowanie własne w Excelu.

Rys. 12.6. Współczynniki korelacji między zmiennymi X_1 - X_{11}

Widzimy, że dużą korelację względem siebie wykazują: X_1 - X_4 (0,709); X_5 - X_{11} (0,766); X_6 - X_7 (0,724). W tej sytuacji spróbujmy teraz pominąć zmienną X_1 (zob. rysunek 12.7).

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
X_2	1,000									
X_3	0,160	1,000								
X_4	-0,187	0,021	1,000							
X_5	-0,056	0,019	-0,669	1,000						
X_6	-0,163	0,220	0,208	0,024	1,000					
X_7	-0,263	0,296	0,441	-0,144	0,724	1,000				
X_8	-0,149	-0,021	-0,138	0,448	0,671	0,348	1,000			
X_9	0,080	0,553	0,609	-0,414	0,041	0,457	-0,358	1,000		
X_{10}	0,272	0,114	0,173	-0,206	0,069	0,139	0,274	0,159	1,000	
X_{11}	0,104	0,042	-0,607	0,766	-0,014	-0,206	0,475	-0,413	0,107	1,000

Źródło: Opracowanie własne w Excelu.

Rys. 12.7. Współczynniki korelacji po pominięciu X_1

Teraz zbyt dużą korelację wykazują nadal: X_5 - X_{11} (0,766), X_6 - X_7 (0,724), tak więc pozostaje nam pominięcie tych relacji i obliczamy ponownie korelację wzajemną zmiennych. Oczywiście mogliśmy to m uczynić w jednym etapie.

	X_2	X_3	X_4	X_7	X_8	X_9	X_{10}	X_{11}
X_2	1,000							
X_3	0,160	1,000						
X_4	-0,187	0,021	1,000					
X_7	-0,263	0,296	0,441	1,000				
X_8	-0,149	-0,021	-0,138	0,348	1,000			
X_9	0,080	0,553	0,609	0,457	-0,358	1,000		
X_{10}	0,272	0,114	0,173	0,139	0,274	0,159	1,000	
X_{11}	0,104	0,042	-0,607	-0,206	0,475	-0,413	0,107	1,000

Źródło: Opracowanie własne w Excelu.

Rys. 12.8. Współczynniki korelacji po pominięciu X_1 , X_5 , X_6

Po tych zmianach dobrane zmienne wykazują raczej słabą korelację, bo poniżej 0,7.

Przyjęte zmienne X_2 - X_4 , X_7 - X_{11} wykazują znaczne zróżnicowanie pod względem wartości i dlatego potrzebne jest przeprowadzenie ich standaryzacji poprzez odjęcie od wartości danej zmiennej jej średniej i podzielenie przez odchylenie standardowe. Wymaga to jednak, dla nowego zestawu zmiennych zapisanej w dogodnej notacji *Excela*, ponownego określenia statystyki opisowej. Na rysunku 12.9, dla zaoszczędzenia miejsca zamieszczono tylko średnią i odchylenie standardowe.

Cecha	X_2	X_3	X_4	X_7	X_8	X_9	X_{10}	X_{11}
Średnia	0,3	-1,4	333,5	110,8	51,3	3,0	12,8	353,5
Odchylenie standardowe	1,8	2,9	88,4	101,5	26,6	1,8	12,9	199,5

Źródło: Opracowanie własne w Excelu.

Rys. 12.9. Średnie i odchylenia standardowe zmiennych X_2 - X_4 , X_7 - X_{11}

Macierz wartości zmiennych zestandaryzowanych zapisanych jako Z_2 - Z_4 , Z_7 - Z_{11} pokazana została na rysunku 12.10. Dla korzystania z *Excela* zmienne zestandaryzowane zapisano bez indeksowania cyfr. Widzimy tu formułę obliczeniową zapisaną w komórce „I4”, przy czym: A4 - wartość zmiennej X_2 dla podregionu 1, \$I\$1 - średnia zmiennej X_2 (pobrana z wiersza 1), \$I\$2 - odchylenie standardowe zmiennej X_2 .

I4		▼		fx		=(A4-\$I\$1)/\$I\$2															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P					
2									1,8	2,9	88,4	101,5	26,6	1,8	12,9	199,5					
3	X2	X3	X4	X7	X8	X9	X10	X11	Z2	Z3	Z4	Z7	Z8	Z9	Z10	Z11					
4	-2,5	0,5	256,0	82,5	65,3	1,8	4,7	592,7	-1,5	0,6	-0,9	-0,3	0,5	-0,7	-0,6	1,0					
5	-0,8	-1,8	297,0	82,9	56,7	2,3	7,5	514,2	-0,8	-0,2	-0,9	-0,3	0,2	-0,4	-0,4	0,8					
6	-6,5	-2,4	434,0	147,9	38,7	1,6	3,6	19,2	-3,7	-0,4	1,1	0,4	-0,5	-0,8	-0,7	-1,7					
7	0,1	-1,9	273,0	64,4	51,9	2,8	3,0	507,4	-0,1	-0,2	-0,7	-0,5	0,0	-0,1	-0,8	0,8					
8	0,5	-3,8	300,0	67,0	53,2	1,9	4,9	714,3	0,1	-0,5	-0,4	-0,4	0,1	-0,6	-0,6	1,8					
9	0,8	9,2	352,0	141,7	57,1	6,4	3,0	420,8	0,3	3,6	0,2	0,3	0,2	1,9	-0,6	0,3					
10	0,2	-2,6	309,0	64,5	73,9	2,2	2,9	375,7	-0,1	-0,4	-0,5	-0,5	0,8	-0,5	-0,8	0,1					
11	-0,6	4,8	466,0	630,1	49,4	8,1	21,0	117,5	-0,5	2,1	1,5	5,1	-0,1	2,9	0,6	-1,2					
12	0,4	0,9	277,0	119,7	69,3	2,4	5,9	383,7	0,1	0,8	-0,6	0,1	0,5	-0,3	-0,5	0,2					
13	3,3	-0,7	241,0	109,1	61,7	2,2	39,8	303,9	1,8	0,2	-1,0	-0,1	0,4	-0,5	2,1	-0,2					
14	-0,4	1,1	530,0	192,4	20,3	8,7	17,0	21,6	-0,4	0,8	2,2	0,8	-1,2	3,2	0,3	-1,7					
15	-2,5	-1,0	299,0	59,0	32,7	2,4	3,6	141,4	-1,5	0,1	-0,4	-0,5	-0,7	-0,3	-0,7	-1,1					
16	1,0	1,0	310,0	78,2	25,0	3,3	19,1	74,8	0,4	0,8	-0,3	-0,3	-1,0	0,2	0,5	-1,4					
17	-1,5	-3,3	440,0	411,6	146,0	1,4	10,8	204,8	-1,0	-0,7	1,2	3,0	3,6	-0,9	-0,2	-0,7					
18	-1,7	-4,6	344,0	68,7	26,1	1,6	1,7	64,9	0,8	-1,1	0,1	-0,4	-0,9	-0,8	-0,9	-1,4					
19	-0,2	-4,3	246,0	30,5	21,3	1,6	5,4	309,1	-0,3	-1,0	-0,8	-0,8	-1,1	-0,8	-0,6	-0,2					
20	-1,8	-3,9	230,0	58,8	49,7	1,5	5,2	542,7	-1,1	-0,9	-1,2	-0,5	-0,1	-0,9	-0,6	0,9					
21	0,1	-2,2	281,0	149,0	70,6	2,7	8,5	642,5	-0,1	-0,3	-0,6	0,4	0,7	-0,2	-0,3	1,4					
22	1,6	-1,4	207,0	132,8	72,2	2,6	6,7	366,8	0,7	0,0	-1,4	0,2	0,8	-0,2	-0,5	0,1					
23	1,0	-2,3	185,0	96,1	79,0	1,7	13,5	386,1	0,4	-0,4	-1,7	-0,1	0,8	-0,7	0,1	0,2					
24	-1,0	-1,4	242,0	104,9	49,3	2,8	10,1	745,5	-0,7	0,0	-1,0	-0,1	-0,3	-0,1	-0,2	2,0					
25	0,3	-3,0	201,0	25,4	18,5	1,7	1,6	392,0	0,0	-1,2	-1,5	-0,8	-1,2	-0,7	-0,9	0,0					
26	-1,9	-2,7	320,0	135,9	99,4	1,3	7,9	556,5	-1,2	-0,5	-0,2	0,2	1,8	-1,0	-0,4	1,0					
27	1,3	-1,3	383,0	46,6	24,3	3,1	7,9	196,8	0,5	0,0	0,6	-0,6	-1,0	0,1	-0,4	-0,8					
28	1,0	-1,7	385,0	74,7	48,5	2,8	11,6	283,3	0,4	-0,1	0,6	-0,4	-0,1	-0,1	-0,1	-0,4					
29	2,3	-2,4	313,0	39,0	25,5	2,0	3,7	289,1	1,1	-0,4	-0,2	-0,7	-1,0	-0,6	-0,7	-0,3					
30	2,7	4,9	253,0	121,6	50,3	3,6	15,8	674,4	1,3	2,1	-0,9	0,1	0,0	0,3	0,2	1,6					
31	1,2	-1,3	263,0	78,0	43,9	2,1	3,6	479,7	0,5	0,0	-0,8	-0,3	-0,3	-0,5	-0,7	0,6					
32	1,1	-1,8	249,0	39,4	33,8	2,2	5,8	283,6	0,4	-0,2	-0,8	-0,7	-0,7	-0,3	-0,5	-0,4					
33	0,0	-4,4	522,0	163,6	15,8	5,0	7,7	33,9	-0,2	-1,0	2,1	0,5	-1,3	1,1	-0,4	-1,6					
34	0,3	-0,3	369,0	139,6	81,1	2,7	56,2	614,1	0,0	0,4	0,4	0,3	1,1	-0,2	3,4	1,3					
35	0,8	-2,4	339,0	62,2	57,7	3,6	52,5	379,3	0,3	-0,4	0,1	-0,5	0,2	0,3	3,1	0,1					
36	-2,2	-3,2	390,0	130,1	108,1	1,5	32,4	467,2	-1,4	-0,6	0,6	0,2	2,1	-0,9	1,5	0,5					
37	1,4	-3,4	431,0	62,4	30,5	2,2	2,5	185,6	0,6	-0,7	1,1	-0,5	-0,8	-0,5	-0,8	-0,8					
38	1,3	3,8	269,0	35,1	23,6	4,1	2,8	263,7	0,5	2,4	-0,7	-0,7	-1,0	0,6	-0,8	-0,4					
39	-1,5	-0,5	553,0	150,3	23,3	6,9	7,3	70,1	-1,0	0,3	2,5	0,4	-1,1	2,2	-0,4	-1,4					
40	-0,9	-4,6	312,0	111,2	60,1	1,3	7,5	563,9	-0,7	-1,1	-0,2	0,0	0,3	-1,0	-0,4	1,0					
41	0,6	-1,4	328,0	129,1	72,7	2,0	13,1	467,6	0,2	0,0	-0,1	0,2	0,8	-0,6	0,0	0,6					
42	0,8	-2,0	326,0	115,5	87,3	2,8	12,4	578,4	0,3	-0,2	-0,1	0,0	1,4	-0,1	0,0	1,1					
43	2,8	-3,4	424,0	52,5	45,4	2,5	30,7	282,9	1,3	-0,7	1,0	-0,6	-0,2	-0,3	1,4	-0,4					
44	4,5	2,7	359,0	81,3	61,7	4,1	37,2	462,8	2,3	1,4	0,3	-0,3	0,4	0,6	1,9	0,5					
45	-0,3	-2,6	466,0	136,1	18,8	6,4	16,0	33,1	-0,3	-0,4	1,5	0,4	-1,2	1,9	0,3	-1,6					
46	1,8	-3,9	316,0	51,7	47,6	2,0	8,2	348,0	0,8	-0,9	-0,2	-0,6	-0,1	-0,6	-0,4	0,0					
47	1,7	-3,1	353,0	81,0	49,5	3,9	20,7	426,1	0,8	-0,6	0,2	-0,3	-0,1	0,5	0,6	0,4					
48	1,8	-3,8	370,0	26,4	30,5	3,5	9,4	224,6	0,8	-0,8	0,4	-0,8	-0,8	0,3	-0,3	-0,6					

Źródło: Opracowanie własne w Excelu.

Rys. 12.10. Macierz danych wejściowych i zestandaryzowanych

W Excelu rozróżniamy wartość komórki i jej zawartość. Na rysunku 18.10 są wartości pokazane z dokładnością do jednego miejsca po przecinku, natomiast zawartość jest z standardowo z dokładnością do 14 miejsc po przecinku np. dla komórki I4 wynosi (-1,51955091270367). Podano to dlatego, gdyż niekiedy Czytelnicy sprawdzają manualnie na kalkulatorze wyniki i mają zastrzeżenia.

12.3. Odległości euklidesowe

Różne są sposoby określenia tzw. odległości między badanymi obiektami ze względu na cechy. Dysponując narzędziem jakim jest Excel zdecydowano się na obliczanie odległości euklidesowych (d_{ij}) dla wielu cech następującym wzorem, będącym analogią do twierdzenia Pitagorasa:

$$d_{ij} = \sqrt{\sum_{j=1}^m (z_{ij} - z_{lj})^2}$$

Gdzie: m - liczba badanych cech (zmiennych); z_{ij} - wartość zestandaryzowana zmiennej j dla obiektu i ; z_{lj} - wartość zestandaryzowana zmiennej l dla obiektu i , względem której określamy odległość euklidesową.

	L3										f ₃ =PIERWIASTEK(((SBS2-B3)*2)+((SCS2-C3)*2)+((SDS2-D3)*2)+((SE2-E3)*2)+((SFS2-F3)*2)+((SGS2-G3)*2)+((SHS2-H3)*2)+((SIS2-I3)*2))																											
	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ			
1	P	23	24	25	26	27	28	29	30	31	P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
2	1	-1.5	-0.6	-0.8	-0.3	0.5	-0.7	-0.6	1.0	1	0.000																											
3	2	-0.6	-0.2	-0.6	-0.3	0.2	-0.4	-0.4	0.8	2	1.314	0.000																										
4	3	-0.7	-0.4	1.1	0.4	-0.5	-0.8	-0.7	-1.7	3	4.278	4.566	0.000																									
5	4	-0.1	-0.2	-0.7	-0.5	0.0	-0.1	-0.8	0.8	4	1.832	0.734	4.846	0.000																								
6	5	0.1	-0.5	-0.4	-0.4	0.1	-0.6	-0.6	1.8	5	2.245	1.459	5.458	1.242	0.000																							
7	6	0.3	0.4	0.2	0.3	0.2	1.8	-0.4	0.3	6	4.550	4.668	4.603	4.483	5.112	0.000																						
8	7	-0.1	-0.4	-0.3	-0.5	0.8	-0.5	-0.8	0.1	7	2.144	1.333	4.472	1.209	1.892	4.802	0.000																					
9	8	-0.5	2.1	1.8	5.1	-0.1	2.8	0.6	-1.2	8	7.570	7.452	7.387	7.475	8.006	5.700	7.503	0.000																				
10	9	0.1	0.8	-0.6	0.1	0.5	-0.3	-0.5	0.2	9	1.876	1.397	4.783	1.378	2.212	3.733	1.441	6.759	0.000																			
11	10	1.6	0.2	-1.0	-0.1	0.4	-0.5	2.1	-0.2	10	4.370	3.540	6.650	3.582	3.876	5.303	3.528	7.483	3.167	0.000																		
12	11	-0.4	0.8	2.2	0.8	-1.2	3.2	0.3	-1.7	11	6.175	5.768	5.594	5.546	6.283	4.556	5.602	4.729	5.309	6.075	0.000																	
13	12	-1.5	0.1	-0.4	-0.5	-0.7	-0.3	-0.7	-1.1	12	2.535	2.367	2.840	2.471	3.457	4.910	2.499	2.263	2.477	4.493	5.448	0.000																
14	13	0.4	0.8	-0.3	-0.3	0.2	0.2	0.2	-1.4	13	5.741	3.101	4.792	2.960	3.870	4.134	3.053	6.583	2.488	2.953	4.202	2.449	0.000															
15	14	-1.0	-0.7	1.2	3.0	0.4	-0.9	-0.2	-0.7	14	5.417	5.404	5.820	5.664	6.852	6.599	4.845	6.361	6.072	6.130	7.054	5.866	6.297	0.000														
16	15	0.8	-1.1	0.1	-0.4	-0.9	-0.8	-0.9	-1.4	15	4.171	3.337	4.713	3.827	3.573	5.886	4.658	7.796	3.103	4.000	4.976	2.714	2.579	6.096	0.000													
17	16	-0.3	-1.0	-0.8	-0.8	-1.1	-0.8	-0.6	-0.2	16	2.950	2.024	4.422	1.897	2.502	5.794	2.209	1.159	2.613	3.890	5.625	2.002	2.751	6.404	1.897	0.000												
18	17	-1.1	-0.9	-1.2	-0.5	-0.1	-0.9	-0.6	0.9	17	1.701	1.130	4.442	1.538	1.770	5.694	1.972	8.239	2.424	4.237	5.566	2.540	3.804	5.831	3.453	1.874	0.000											
19	18	-0.1	-0.3	-0.6	0.4	0.7	-0.2	-0.3	1.4	18	1.978	1.224	2.244	1.361	1.276	4.657	1.699	1.102	1.745	3.556	5.832	3.390	3.720	4.925	3.794	2.929	1.877	0.000										
20	19	0.7	0.0	-1.4	-0.2	0.8	-0.2	-0.5	0.1	19	2.657	1.785	5.555	1.695	2.414	4.565	1.634	1.184	1.346	2.834	5.846	3.180	2.965	5.164	3.123	2.726	2.659	1.843	0.000									
21	20	0.4	-0.4	-1.1	-0.1	0.8	-0.1	0.1	0.2	20	2.557	1.675	5.507	1.848	2.361	5.235	1.735	1.789	1.792	2.627	6.314	3.193	1.815	5.324	3.241	2.521	2.175	1.936	0.991	0.000								
22	21	-0.7	-0.0	-1.0	-0.3	-0.1	-0.2	0.2	20	1.806	1.351	2.580	1.580	1.428	4.772	2.531	2.570	2.318	4.032	6.255	3.536	3.852	6.128	4.209	2.748	1.705	1.487	2.681	2.583	0.000								
23	22	0.0	-0.2	-1.5	-0.8	-1.1	-0.7	-0.2	0.0	22	2.245	2.324	5.111	2.125	2.459	6.073	3.400	5.983	6.001	4.128	6.087	2.671	3.273	6.869	3.355	3.807	2.008	3.128	2.760	2.543	2.894	0.000						
24	23	-1.2	-0.2	0.2	0.2	1.6	-1.0	-0.4	1.6	23	1.945	2.038	4.530	2.473	2.448	5.470	1.982	2.599	2.511	4.390	6.688	3.488	4.478	3.257	4.318	3.571	2.305	1.845	2.814	2.022	2.719	3.949	0.000					
25	24	0.5	0.0	0.4	-0.6	-1.0	0.1	-0.4	0.8	24	3.600	2.775	4.606	3.312	4.481	4.445	7.024	2.442	3.564	4.128	3.247	1.698	6.170	1.712	2.134	3.374	3.308	2.963	2.243	3.544	2.769	4.692	0.000					
26	25	0.4	-0.1	0.6	-0.4	-0.1	-0.1	-0.1	-0.4	25	3.018	2.161	4.484	1.892	2.523	4.387	1.651	6.825	1.839	3.084	4.350	2.431	1.943	5.271	2.070	2.223	2.804	2.486	2.372	2.605	3.051	2.889	3.131	1.118	0.000			

Źródło: Opracowanie własne w Excelu.

Rys. 12.11. Fragment macierzy odległości euklidesowych między zmiennymi Z_i i Z_l

Na rysunku 18.11 pokazano macierz zmiennych zestandaryzowanych Z_2 - Z_4 , Z_7 - Z_{11} oraz macierz symetryczną odległości euklidesowych, przy czym na głównej przekątnej występują zera. Zwróćmy uwagę na przykładową komórkę „L3” dla której wyświetlona jest formuła obliczeniowa wynikająca z wcześniej podanego wzoru na odległość d_{ij} .

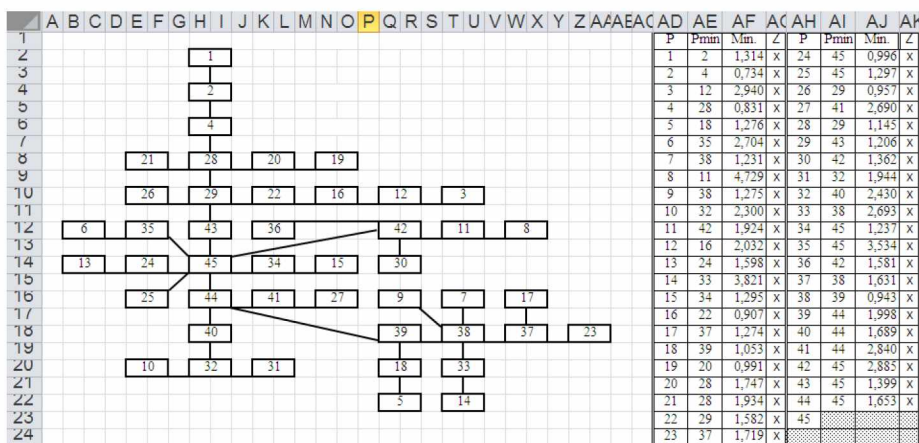
	L47										f ₄										=MIN(L3:L46)											
	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE		
27	26	1.1	-0.4	-0.2	-0.7	-1.0	-0.6	-0.7	-0.3	26	3.511	2.493	5.278	2.026	2.590	5.040	2.210	7.782	2.332	3.387	5.263	2.778	2.256	4.604	1.487	1.624	2.940	3.007	2.441	2.636		
28	27	1.3	2.1	-0.9	0.1	0.0	0.3	0.2	1.6	27	3.573	3.250	6.955	3.093	3.241	3.059	3.699	6.965	2.622	3.390	5.797	4.638	3.637	6.765	4.965	4.415	4.222	3.025	2.014	3.437		
29	28	0.5	0.0	-0.6	-0.3	-0.3	-0.5	-0.7	0.6	28	2.284	1.256	5.216	0.831	1.454	4.518	1.524	7.546	1.334	3.246	5.885	2.702	2.725	5.889	2.654	1.832	1.965	1.715	1.534	1.747		
30	29	0.4	-0.2	-0.8	-0.7	-0.7	-0.5	-0.5	-0.4	29	2.810	1.833	4.878	1.504	2.389	4.782	1.791	7.652	1.774	3.168	5.329	2.151	1.995	6.173	1.852	1.260	2.315	2.611	1.918	2.012		
31	30	-0.2	-1.0	2.1	0.5	-1.3	-1.1	-0.4	-1.6	30	5.276	4.591	4.288	4.332	4.929	5.664	4.183	6.107	4.427	5.362	2.459	3.890	3.460	6.047	3.139	3.973	5.003	4.823	4.869	5.206		
32	31	0.0	0.4	0.4	0.3	1.1	-0.2	3.4	1.3	31	4.575	4.230	6.559	4.527	4.380	5.686	4.498	7.221	4.274	3.090	6.258	5.400	4.613	5.735	5.773	5.259	4.844	3.904	4.510	4.227		
33	32	0.3	-0.4	-0.1	0.6	0.2	0.3	3.1	0.1	32	4.444	3.651	6.072	4.010	4.194	5.675	3.996	7.359	3.954	2.800	5.300	4.552	3.469	6.211	4.631	4.247	4.396	3.862	4.026	3.709		
34	33	-1.4	-0.6	0.6	0.2	2.1	-0.9	1.5	0.5	33	3.402	3.320	4.732	3.765	3.754	6.034	3.191	7.455	3.607	4.079	6.500	4.233	4.654	3.821	4.936	4.449	3.650	3.156	3.907	3.546		
35	34	0.6	-0.7	1.1	-0.5	-0.8	-0.5	-0.8	-0.8	34	3.918	3.077	4.476	2.705	3.208	5.277	2.445	7.358	2.893	4.090	4.409	2.726	2.562	5.817	1.296	2.256	3.483	3.477	3.289	3.517		
36	35	0.5	1.4	-0.7	-0.7	-1.0	0.6	-0.8	-0.4	35	3.738	3.533	5.841	3.241	4.104	2.704	3.718	7.006	2.698	4.245	4.636	3.339	2.402	7.305	4.042	3.808	4.360	4.078	3.441	3.647		
37	36	-1.0	0.3	2.5	0.4	-1.1	-1.2	-0.4	-1.4	36	5.998	5.026	4.357	4.812	5.508	4.732	4.776	5.423	4.724	6.024	2.355	4.021	3.888	6.376	4.508	4.950	5.607	5.251	4.554	4.896		
38	37	-0.7	-1.1	-0.2	0.0	3.3	-1.0	-0.4	1.0	37	2.066	1.921	4.443	1.590	1.423	5.644	1.575	7.756	2.299	3.958	6.029	2.889	3.791	4.974	3.188	2.174	1.274	1.454	2.469	2.205		
39	38	0.2	0.0	-0.1	0.2	0.8	-0.6	0.0	0.6	38	2.191	1.442	4.863	1.537	1.783	4.472	1.231	6.909	1.275	2.894	4.445	2.992	2.974	4.534	3.147	2.710	2.337	1.235	1.669	1.762		
40	39	0.3	-0.2	-0.1	0.0	1.4	-0.1	0.0	1.1	39	2.454	1.767	5.418	1.786	1.767	4.580	1.542	1.134	1.943	3.227	5.657	3.890	3.665	4.577	3.787	3.282	2.602	1.069	1.920	2.057		
41	40	1.3	-0.7	1.0	-0.6	-0.2	-0.3	1.4	-0.4	40	5.504	5.538	5.730	3.359	3.528	5.511	3.129	7.418	3.380	2.516	4.752	4.005	2.756	5.935	2.896	3.01	4.122	3.97	3.491	3.444		
42	41	2.3	1.4	0.3	-0.3	0.4	0.6	1.9	0.5	41	4.951	4.280	7.553	4.101	4.257	4.149	4.166	6.981	3.641	3.232	5.526	5.284	3.477	6.661	4.873	4.941	5.245	4.019	3.781	4.016		
43	42	-0.3	-0.4	1.5	0.4	-1.2	1.8	0.3	-1.6	42	5.108	4.460	4.498	4.246	5.001	4.956	4.235	5.444	4.222	4.944	1.924	3.875	6.828	5.620	3.598	4.113	5.012	4.720	4.671	5.031		
44	43	0.8	-0.9	-0.2	-0.3	-0.1	-0.6	-0.4	0.0	43	3.128	1.968	5.113	1.610	2.043	5.244	1.645	7.743	2.082	3.019	2.625	2.828	2.622	5.717	1.830	2.417	2.311	1.971	1.987			
45	44	0.8	-0.6	-0.2	-0.3	-0.1	0.2	0.6	0.4	44	3.402	2.100	5.367	2.049	2.385	4.645	2.165	5.904	2.380	2.598	3.871	3.311	2.529	5.020	1.861	2.642	2.044	2.261	2.420	2.438		
46	45	0.8	-0.6	0.4	-0.8	-0.6	0.3	-0.3	-0.6	45	3.872	2.819	4.968	2.381	3.015	5.085	2.599	7.425	2.773	3.470	4.184	2.802	2.179	6.211	1.584	2.031	3.295	3.259	2.907	3.095		
47	Min 1.314 0.734 2.940 0.831 1.276 2.704 1.231 4.719 1.275 2.300 1.924 0.832 1.698 1.821 1.296 0.907 1.274 1.063 0.991 1.747																															
48	P 2 4 12 28 18 35 38 38 32 42 16 24 33 34 22 37 39 20																															
49	P 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20																															
50																																

#	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE
27	3,241	1,949	3,982	1,331	1,633	0,000																				26
28	3,037	4,542	4,390	3,854	3,531	3,688	0,000																			27
29	1,924	1,988	2,934	2,217	1,864	1,530	2,815	0,000																		28
30	2,739	1,882	3,507	1,576	1,631	0,987	3,471	1,145	0,000																	29
31	5,203	4,621	5,274	2,706	3,002	3,720	5,877	4,444	3,969	0,000																30
32	4,244	5,735	4,190	4,925	4,122	5,174	4,247	4,615	4,910	5,941	0,000															31
33	4,134	4,722	4,422	3,861	3,307	4,191	4,339	4,064	3,951	4,906	1,944	0,000														32
34	3,901	5,019	2,156	4,556	3,533	4,792	5,097	4,134	4,393	5,296	2,893	3,327	0,000													33
35	3,933	2,922	3,992	1,149	1,405	1,586	4,590	2,567	2,075	2,507	5,304	4,348	4,512	0,000												34
36	3,915	4,051	5,055	2,841	3,205	3,118	2,787	3,002	2,833	4,908	5,692	5,039	5,842	3,815	0,000											35
37	5,493	5,612	5,631	3,481	3,697	4,707	5,851	5,074	4,741	1,953	6,117	5,327	5,607	3,697	4,695	0,000										36
38	1,873	2,543	1,719	3,127	2,384	2,787	4,153	1,944	2,447	4,482	4,313	4,041	2,960	2,969	4,616	5,235	0,000									37
39	2,275	3,154	1,888	2,627	1,631	2,503	3,039	1,626	1,822	4,293	3,510	3,348	2,693	2,765	3,731	4,815	1,631	0,000								38
40	2,320	3,602	1,813	3,231	2,226	3,072	3,110	2,073	2,783	4,728	3,505	3,521	2,696	3,341	4,178	5,131	1,891	0,943	0,000							39
41	4,142	3,871	4,260	2,331	1,638	2,607	4,213	3,172	2,902	3,578	3,540	2,430	3,838	2,437	4,436	4,552	3,458	2,735	3,119	0,000						40
42	4,476	5,280	5,070	3,716	3,381	3,950	2,690	3,836	4,007	5,374	3,232	2,965	4,782	4,413	3,937	5,552	4,801	3,435	3,507	2,831	0,000					41
43	4,992	4,717	5,318	2,751	3,023	3,885	5,411	4,412	3,921	1,902	5,474	4,349	5,194	3,069	4,401	1,581	4,663	4,241	4,624	3,587	4,860	0,000				42
44	2,865	2,012	3,144	1,777	1,336	1,118	5,740	1,373	1,206	3,795	4,538	3,658	3,888	1,750	3,714	4,804	1,996	1,771	2,234	2,246	3,812	3,832	0,000			43
45	2,789	3,078	3,364	2,014	1,378	2,170	3,291	2,094	2,150	3,497	3,487	2,568	3,523	2,391	3,718	4,179	2,512	1,815	1,995	1,689	2,840	3,274	1,606	0,000		44
46	3,987	2,547	4,079	0,996	1,297	1,335	4,180	2,290	1,674	2,780	4,908	3,681	4,490	1,237	3,634	3,822	2,978	2,672	3,100	2,047	3,855	2,886	1,399	1,653	0,000	45
47	1,934	1,882	1,719	0,996	1,297	0,987	2,690	1,145	1,206	1,362	1,944	2,430	2,693	1,237	3,634	1,581	1,631	0,943	1,995	1,689	2,840	2,886	1,399	1,653		
48	28	29	37	45	45	29	41	29	43	42	32	40	38	45	45	42	38	39	44	44	44	45	45	45		
49																										
50	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	

Źródło: Opracowanie własne w Excelu.

Rys. 12.13. Fragment macierzy odległości euklidesowych między zmiennymi Z_i i Z_j (cd. rys.12)

Na podstawie macierzy odległości utwórzmy, również w Excelu, dendryt wrocławski. Podane w prostokątach na rysunku 12.14. numery umożliwiają identyfikację podregionu. Łączono dany podregion z innym wykazującym najmniejszą odległość euklidesową wynikającą z rysunków 12.11 - 12.13.



Źródło: Opracowanie własne w Excelu.

Rys. 12.14. Dendryt wrocławski i tabela odległości minimalnych między obiektami (podregionami)

W tabeli 12.14 poszczególne skróty oznaczają: P - numer podregionu, P_{min} - numer podregionu do którego jest najkrótsza odległość euklidesowa uwzględniająca zmienne zestandaryzowane Z₂-Z₄, Z₇-Z₁₁; Min. - wartość minimalnej odległości euklidesowej do innego podregionu, x - znacznik pomocniczy przy konstruowaniu dendrytu wrocławskiego.

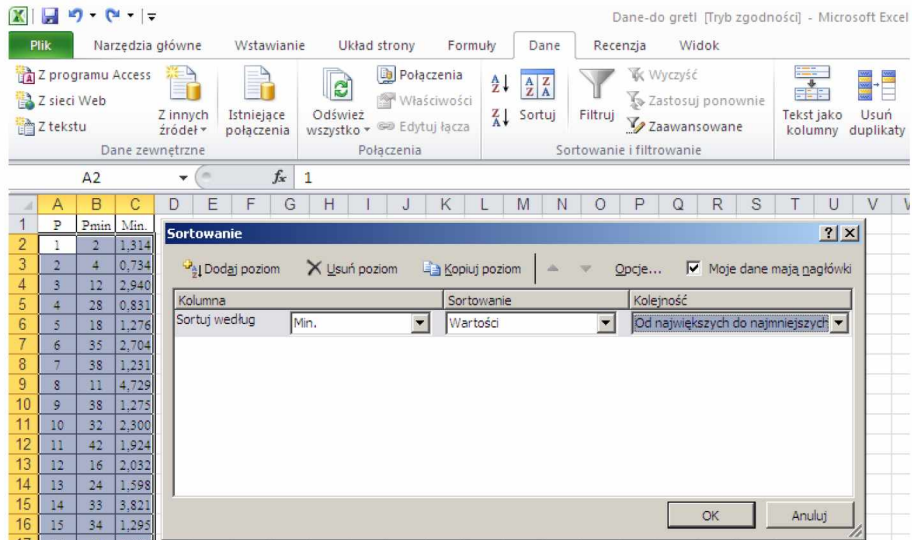
Powstały po wprowadzeniu wiązań graf niezorientowany jest spójny, ponieważ każde dwa różne wierzchołki są połączone nieprzerwanym ciągiem wiązań. Uzyskanie dendrytu wrocławskiego pozwala nam w kolejnym etapie jego podział na „K” skupisk, zwanych też grupami, ze względu na podobieństwo badanych cech (zmiennych zestandaryzowanych). Realizacja tego zadania wymaga jednak wcześniejszego posortowania malejąco tabeli odległości, przy czym zestawiamy

zamieszczona na rysunku 12.14. tabele pionowo w jednym szeregu na arkuszu *Excela*. W tej tabeli mamy kolumny:

P - numer podregionu,

$P_{\min}$ - numer podregionu minimalnej odległości euklidesowej,

$M_{\min}$ - odległość euklidesowa (zob. rysunek 12.15).



Źródło: Opracowanie własne w *Excelu*.

Rys. 12.15. Fragment szeregu odległości i okno dialogowe ustawienia sortowania

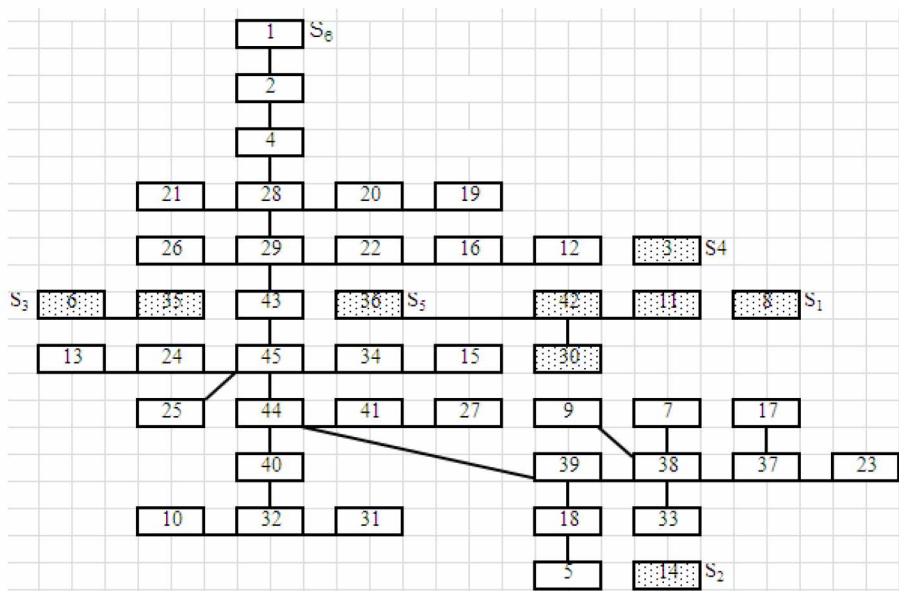
W oknie dialogowym zaznaczono: „*Moje dane mają nagłówki*”, sortowanie według „*Wartość*” w kolumnie „ $M_{\min}$ ”, kolejność „*Od największych do najmniejszych*”. W efekcie sortowania, po odpowiedniej edycji, uzyskujemy zestawienie odległości euklidesowych pomiędzy podanymi obiektami - podregionami (zob. tabela 12.2).

Tab. 12.2. Odległości minimalne (malejąco)

P	8	14	35	3	42	41	6	33	27	32	10	12	39	31	21	11	20	23	40	44	37	1	22
$P_{\min}$	11	33	45	12	45	44	35	38	41	40	32	16	44	32	28	42	28	37	44	45	38	2	29
M_i	4, 73	3, 82	3, 53	2, 94	2, 89	2, 84	2, 7	2, 69	2, 69	2, 43	2, 3	2, 03	2	1, 94	1, 93	1, 92	1, 75	1, 72	1, 69	1, 65	1, 63	1, 6	1, 58
Lp	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	2	23
P	36	43	30	1	25	15	5	9	17	34	7	29	28	18	24	19	26	38	16	4	2	4	
$P_{\min}$	42	45	42	2	45	34	18	38	37	45	38	43	29	39	45	20	29	39	22	28	4	4	
M_i	1, 58	1, 4	1, 36	1, 31	1, 3	1, 29	1, 28	1, 28	1, 27	1, 24	1, 23	1, 21	1, 14	1, 05	1	0, 99	0, 96	0, 94	0, 91	0, 83	0, 73	0	
Lp	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	4	

Źródło: Opracowanie własne w *Excelu*.

Dla lepszego zaprezentowania wyników na formacie A4, zaokrąglono $Min.$ do dwóch miejsc po przecinku. Usuńmy jako eksperci $k - 1 = 5$ najdłuższych wiązań. Uzyskano $k = 6$ skupisk (grup podregionów) podobnych pod względem wartości zmiennych zestandaryzowanych Z_2-Z_4, Z_7-Z_{11} , lecz licznie zróżnicowanych (zob. rysunek 12.16)¹²³:



Źródło: Opracowanie własne w Excelu.

Rys. 12.16. Dendryt wrocławski po usunięciu pięciu najdłuższych wiązań

Poprzez usunięcie pięciu najdłuższych wiązań nastąpiło wyodrębnienie sześciu skupisk (grup podregionów), a są nimi:

- S₁: (8) miasto Warszawa;
- S₂: (14) centralny śląski;
- S₃: (35) wrocławski, (6) warszawski;
- S₄: (3) miasto Łódź;
- S₅: (42) Gdynia-Gdańsk-Sopot, (36) miasto Wrocław, (11) miasto Kraków, (30) miasto Poznań;
- S₆: pozostałe 36 podregionów nie wymienione w S₁-S₅.

Temat ustalenia liczby grup zwanych też klasami K jest przedmiotem różnych rozważań naukowych. Hellwig proponuje wyjście od obliczenia odległości krytycznej według wzoru¹²⁴:

$$d^* = \bar{d} + k * s_d$$

Gdzie:

$\bar{d}$ - średnia arytmetyczna wszystkich odległości zestandaryzowanych z grafu spójnego (dendrytu wrocławskiego) równa dla naszego przykładu 1,803;

S_d - odchylenie standardowe powyższych odległości równe 0,855;

k - stała przyjmowana z przedziału (0-2).

¹²³ W opracowaniu niniejszej części materiału uwzględniono sugestie zawarte w publikacji internetowej „Zastosowanie taksonomii wrocławskiej do oceny poziomu rozwoju społeczno-gospodarczego karpaccich gmin górskich ONW, Metody ilościowe w badaniach ekonomicznych, tom XVII, 2016, strony: 7-15.

¹²⁴ Wornalkiewicz W., *Formułowanie modeli ekonometrycznych na potrzeby zarządzania*, Wydawnictwo Instytut Śląski w Opolu, s. 269.

Przyjmijmy dla testowania, że stała k zmienia się co 0,5, tak więc odległości krytyczne obliczone według wzoru na d^* są następujące:

$\bar{d}$	1,803	S_d	0,855		
k	0	0,5	1	1,5	2
d^*	1,803	2,231	2,658	3,086	3,513

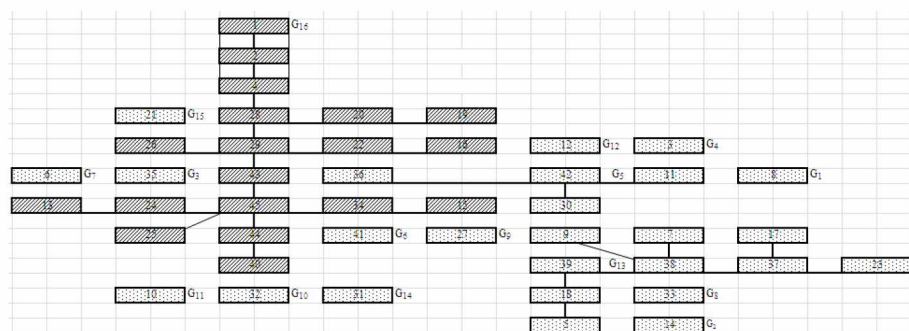
Przyjmując maksymalne kryterium $k = 2$ to należy z grafu usunąć wiązadła o długości $>$ niż 3,513, czyli niżej wymienione:

P	8	14	35
$P_{\min}$	11	33	45
Min.	4,73	3,82	3,53

Ze względów praktycznych procesu zarządzania podregionami, podzbiór - skupisko 6 jest zbyt obszerny do zarządzania, spróbujemy zatem usunąć np. 15. najdłuższych wiązań, czyli 10 dalszych z rozwiązania pokazanego wcześniej na rysunku 184.16.

P	8	14	35	3	42	41	6	33	27	32	10	12	39	31	21
$P_{\min}$	11	33	45	12	45	44	35	38	41	40	32	16	44	32	28
Min.	4,73	3,82	3,53	2,94	2,89	2,84	2,7	2,69	2,69	2,43	2,3	2,03	2	1,94	1,93
Lp.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Efekt naszego działania w postaci skorygowanego dendrytu wrocławskiego możemy zobaczyć na rysunku 12.17.



Źródło: Opracowanie własne w Excelu.

Rys. 12.17. Grupy podregionów po usunięciu 15. najdłuższych wiązań

Wydzielono 16 grup ze względu na badane cechy X_2 - X_4 , X_7 - X_{11} , przy czym poszczególne grupy obejmują podregiony:

- G₁: miasto Warszawa (8); G₂: centralny śląski (14); G₃: wrocławski (35); G₄: miasto Łódź (3);
- G₅: Gdańsk-Gdynia-Sopot (42), miasto Kraków (11), miasto Wrocław (36), miasto Poznań (30);
- G₆: gdański (41); G₇: warszawski (6); G₈: jeleniogórsko-wałbrzyski (33); G₉: poznański (27);
- G₁₀: koszaliński (32); G₁₁: nowosądecki (10); G₁₂: częstochowski (12);
- G₁₃: toruńsko-włocławski (39), bydgoski (38), opolski (37), świętokrzyski (23), krakowsko-tarnowski (9), radomski (7), chełmsko-zamojski (17), lubelskie (18), ostrołęcko-siedlecki (5);
- G₁₄: szczeciński (31); G₁₅: białostocko-suwański (21);
- G₁₆: łódzki (1), piotrkowsko-skierniewicki (2), ciechanowsko-płocki (4), kaliski (28), rzeszowsko-tarnobrzyski (19), krośnieńsko-przemyski (20), pilski (26); koniński (29), łomżyński (22),

bielsko-podlaski (16), elbląski (43), bielsko-bialski (13), gorzowski (24), elcki (45), legnicki (34), rybnicko-jastrzębski (15), zielonogórski (25), olsztyński (44), słupski (40).

Dla poszczególnych 16 grup obliczono wartości średnie ($\bar{X}_r$) zmiennych X_2 - X_4 , X_7 - X_{11} (zob. tabela 12.3).

Tab. 12.3. Średnie zmiennych dla grup G₁-G₁₆

Gr.	X_2	X_3	X_4	X_7	X_8	X_9	X_{10}	X_{11}	P
G ₁	-0,6	4,8	466,0	630,1	49,4	8,1	21,0	117,5	8
G ₂	-1,5	-3,3	440,0	411,6	146,0	1,4	10,8	204,8	14
G ₃	1,3	5,8	269,0	35,1	23,6	4,1	2,8	263,7	35
G ₄	-6,5	-2,4	434,0	147,9	38,7	1,6	3,6	19,2	3
G ₅	-0,3	-2,6	466,0	156,1	18,8	6,4	16,0	33,1	42
	-0,4	1,1	530,0	192,4	20,3	8,7	17,0	21,6	11
	-1,5	-0,5	553,0	150,3	23,3	6,9	7,3	70,1	36
	0,0	-4,4	522,0	163,6	15,8	5,0	7,7	33,9	30
	-0,6	-1,6	517,8	165,6	19,6	6,8	12,0	39,7	$\bar{X}_r$
G ₆	4,5	2,7	359,0	81,3	61,7	4,1	37,2	462,8	41
G ₇	0,8	9,2	352,0	141,7	57,1	6,4	5,0	420,8	6
G ₈	-2,2	-3,2	390,0	130,1	108,1	1,5	32,4	457,2	33
G ₉	2,7	4,9	253,0	121,6	50,3	3,6	15,8	674,4	27
G ₁₀	0,8	-2,4	339,0	62,2	57,7	3,6	52,5	379,3	32
G ₁₁	3,3	-0,7	241,0	100,1	61,7	2,2	39,8	303,9	10
G ₁₂	-2,5	-1,0	299,0	59,0	32,7	2,4	3,6	141,4	12
G ₁₃	0,8	-2,0	326,0	115,5	87,3	2,8	12,4	578,4	39
	0,6	-1,4	328,0	129,1	72,7	2,0	13,1	467,6	38
	-0,9	-4,6	312,0	111,2	60,1	1,3	7,5	553,9	37
	-1,9	-2,7	320,0	133,9	99,4	1,3	7,9	556,5	23
	0,4	0,9	277,0	119,7	63,3	2,4	5,9	383,7	9
	0,2	-2,6	309,0	64,5	73,9	2,2	2,9	375,7	7
	-1,8	-3,9	230,0	58,8	49,7	1,5	5,2	542,7	17
	0,1	-2,2	281,0	149,0	70,6	2,7	8,5	642,5	18
	0,5	-2,8	300,0	67,0	53,2	1,9	4,9	714,3	5
	-0,2	-2,4	298,1	105,4	70,0	2,0	7,6	535,0	$\bar{X}_r$
G ₁₄	0,3	-0,3	369,0	139,6	81,1	2,7	56,2	614,1	31
G ₁₅	1,7	-4,6	344,0	68,7	26,1	1,6	1,7	64,9	21
G ₁₆	-2,5	0,5	256,0	82,5	65,3	1,8	4,7	552,7	1
	-0,8	-1,8	257,0	82,9	56,7	2,3	7,5	514,2	2
	0,1	-1,9	273,0	64,4	51,9	2,8	3,0	507,4	4
	1,2	-1,3	263,0	76,0	43,9	2,1	3,6	479,7	28
	1,6	-1,4	207,0	132,8	72,2	2,6	6,7	366,8	19
	1,0	-2,5	185,0	96,1	73,0	1,7	13,5	386,1	20
	2,3	-2,4	313,0	39,0	25,5	2,0	3,7	289,1	26
	1,1	-1,8	263,0	39,4	33,6	2,2	5,8	283,6	29

0,3	-5,0	201,0	25,4	18,5	1,7	1,6	352,0	22
-0,2	-4,3	266,0	30,5	21,5	1,6	5,4	309,1	16
1,8	-3,9	316,0	51,7	47,6	2,0	8,2	348,0	43
1,0	1,0	310,0	78,2	25,0	3,3	19,1	74,8	13
1,3	-1,3	383,0	46,6	24,3	3,1	7,9	196,8	24
1,8	-3,8	370,0	26,4	30,5	3,5	9,4	224,6	45
1,4	-3,4	431,0	62,4	30,5	2,2	2,5	185,6	34
1,7	-4,6	344,0	68,7	26,1	1,6	1,7	64,9	15
1,0	-1,7	385,0	74,7	48,5	2,8	11,6	283,3	25
1,7	-3,1	353,0	81,0	49,5	3,9	20,7	426,1	44
2,8	-3,4	424,0	52,5	45,4	2,5	30,7	282,9	40
1,0	-2,4	305,3	63,7	41,6	2,4	8,8	322,5	Śr.

Źródło: Opracowanie własne w Excelu.

* * *

Zaprezentowany sposób klasyfikacji obiektów (w naszym przykładzie podregionów) według metody „*Taksonomia wrocławska*” może być kanwą do wykorzystania procedury do innych zastosowań, zwłaszcza w medycynie, biologii, demografii, edukacji, usługach. Zwróćmy uwagę na występowanie dużych aglomeracji miejskich jako samodzielnych grup. Grupę G_{16} stanowią podregiony dla których średnia cech (zmiennych) odbiega znacznie np. od grupy G_1 , czyli „miasto Warszawa”.

Średnia/Cecha	X_2	X_3	X_4	X_7	X_8	X_9	X_{10}	X_{11}
G_{16}	1,0	-2,4	305,3	63,7	41,6	2,4	8,8	322,5
G_1	-0,6	4,8	466,0	639,1	49,4	8,1	21,0	117,5

Objaśnienie poszczególnych zmiennych podano na początku tego materiału. Porównując średnie w obserwowanym roku 2006 (ze względu na wcześniej zgromadzone dane) nastąpił spadek przyrostu naturalnego na 1000 ludności w G_1 . W podregionach grupy G_{16} wystąpiło znaczne ujemne saldo migracji wewnętrznych i zagranicznych na popyt stały (na 1000 ludności).

Mniejsze podregiony (Grupa G_{16}) charakteryzują się mniejszą liczbą stwierdzonych przestępstw na 1000 ludności. Występuje duża dysproporcja pracujących w tys. w obszarze usług rynkowych i pozarynkowych (zob. X_7) oraz prawie czterokrotnie większa liczba mieszkań oddanych do użytkowania na 1000 ludności, gdy porównamy podregion „miasto Warszawa”, a średnia podregionów grupy G_{16} .

Zachęcam Czytelnika do przeprowadzenia klasyfikacji podregionów np. dla roku 2018 lub podjęcie trudu opracowania klasyfikacji obiektów z obszaru własnej pracy badawczej.

13. Badanie występowania wielowymiarowego rozkładu normalnego



13.1. Wstęp

W wyeliminowaniu obserwacji nietypowych, lecz o rozkładzie wielowymiarowym stosowane są następujące statystyczne metody analizy danych:

- wykrywanie losowych obserwacji poprzez zastosowanie zmodyfikowanej odległości Mahalanobisa,

- metoda Gnanadesikana-Kettenringa (GK),

- metoda Campbella.

W celu poznania tych trzech wymienionych metod można skorzystać z opracowanych makroinstrukcji zastosowanych w arkuszu kalkulacyjnym *Excel*¹²⁵. Dla przybliżenia pierwszej z metod wyłonienia obserwacji nietypowych przeprowadzono testowanie komputerowe. Zabazowano na szeregach obserwacji czterech cech (zmiennych) odniesionych na 1000 ludności dotyczących podregionów a mianowicie¹²⁶:

X_2 – przyrost naturalny,

X_3 – saldo migracji wewnętrznych i zagranicznych na pobyt stały,

X_4 – przestępstwa stwierdzone,

X_9 – mieszkania oddane do użytkowania.

Przy prezentacji metody wykrywania losowych obserwacji nietypowych podano w dalszej części tekstu także formuły *Excels* służące obliczeniom cząstkowym w arkuszu kalkulacyjnym. Zastosowanie tej metody wymaga spełnienia założenia o wielowymiarowym rozkładzie normalnym obserwowanych cech statystycznych, a mianowicie: X_2 , X_3 , X_4 , X_9 . Na zadanym poziomie istotności α wykluczamy przypadki braku zgodności rozkładu empirycznego z wielowymiarowym rozkładem normalnym. W tym względzie stosujemy następujące testy: Shapiro-Wilka, skośności oraz spłaszczenia, macierzy kowariancji, wektora średnich i macierzy kowariancji. Zanim przejdziemy do zaprezentowania jednego z tych testów a mianowicie testu Shapiro-Wilka, zwanego też *metodą badania występowania wielowymiarowego rozkładu normalnego* prześledźmy wyłonienie obserwacji nietypowych przy założeniu *a priori* istnienia takiego rozkładu.

13.2. Zastosowanie zmodyfikowanej odległości Mahalanobisa¹²⁷

Przyjmijmy, że dane dotyczące 45. podregionów stanowią populację o wielowymiarowym rozkładzie normalnym dotyczących zmiennych: $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, gdzie¹²⁸:

$\boldsymbol{\mu}$ - wektor wartości oczekiwanych,

$\boldsymbol{\Sigma}$ - wektor odchyleń standardowych.

¹²⁵ Miszczak W., *Statystyczne metody analizy danych Materiały do ćwiczeń*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 1999.

¹²⁶ Womalkiewicz W., *Modele ekonometryczne PKB obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013, tab. 7.1.

¹²⁷ W niniejszym opracowaniu bazowano na fragmencie książki: Duczmał M., *Metody optymalizacji w zarządzaniu*, rozdział 24. *Wyłonienie operacji nietypowych*, Wydawnictwo Instytut Śląski, Opole 2014.

¹²⁸ Miszczak W., *Statystyczne metody analizy danych Materiały do ćwiczeń*, op. cit., s. 11.

Bazując na tym rozkładzie statystyka empiryczna F_i jest przybliżeniem statystyki teoretycznej F -Snedecora (F_α) o m i $(n - m)$ stopniach swobody. Statystykę empiryczną obliczamy według wzoru:

$$F_i = \frac{(n - m)n}{(n^2 - 1)m} D_i^2$$

gdzie w naszym przykładzie: $n = 45$ – ilość obserwacji 4-ro wymiarowych, $m = 4$ – ilość zmiennych.

Kwadrat odległości Mahalanobisa między wektorem danej obserwacji $\mathbf{x}_i$ a wektorem średnich $\bar{\mathbf{x}}$ poszczególnych cech wyznaczamy według wzoru:

$$D_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$$

gdzie:

$\mathbf{x}_i$ - wektor wierszowy czterech zmiennych obserwacji i -tej,

$\bar{\mathbf{x}}$ - wektor wierszowy czterech średnich zmiennych,

$\mathbf{S}^{-1}$ - macierz odwrotna macierzy kowariancji,

T - transpozycja różnicy wektorów.

Według Miszczaka „Za nietypową uważa się taką obserwację, w odniesieniu do której zmodyfikowana odległość Mahalanobisa od środka ciężkości obserwacji jest tak duża, że praktycznie mało prawdopodobne staje się uzyskanie jej w analogicznej próbie”. Kryterium do tej oceny i przyjętego poziomu istotności $\alpha = 0,05$ jest następujące:

$$P(F_i > F_\alpha) = \alpha.$$

Tak więc, jeżeli wartość krytyczna $F_{\alpha=0,05}$ pobrana z tablicy rozkładu F -Snedecora jest mniejsza od wartości obliczeniowej F_i to obserwację możemy uznać za nietypową.

Skorzystajmy zatem z propozycji wstępnie przygotowanego pliku arkusza kalkulacyjnego w *Excelu* o nazwie *SadMakro.xls*¹²⁹. Uruchomienie opcji „Dane losowe” wymaga jednak wcześniejszego wprowadzenia 45. obserwacji szeregu czterech zmiennych począwszy od komórki A4, przy czym w następujących kolumnach zapisano tylko wartości zmiennych: kol. A – X_2 , kol. B – X_3 , kol. C – X_4 , kol. D – X_5 .

Arkusz kalkulacyjny programu *Excel* stanowiący skoroszyt kilku arkuszy cząstkowych nazwano *SadMakro2.xls*. Ponieważ dane używać można wielokrotnie dla różnych opcji, a w tym do testowania istnienia wielowymiarowego rozkładu normalnego to z tego powodu zapisane są dodatkowo w osobnym arkuszu cząstkowym „Arkusz na dane”. Można je zatem skopiować do arkuszy uruchamianych opcji. Widok ekranu zawierającego przyciski makropolecen z 24. pierwszymi obserwacjami pokazano na rysunku 13.1.

¹²⁹ Ibidem, s. 9.

Źródło: Opracowanie w Excelu bazujące na pliku *SadMakro.xls*.

Rys. 13.1. Fragment szeregu wartości obserwacji zmiennych w kolejności X_2, X_3, X_4, X_9 wprowadzonych do arkusza „Obserwacje nietypowe”.

Wprowadzone dane możemy wyczyścić przyciskiem „Wyczyść arkusz”. Po naciśnięciu przycisku „Dane losowe” otrzymaliśmy obliczone kwadraty odległości Mahalanobisa D_i^2 oraz statystykę F_i dla poszczególnych 45. obserwacji dotyczących podregionów od łódzkiego po ełcki (zob. rys. 13.2).

Źródło: Opracowanie bazujące na pliku *SadMakro.xls*.

Rys. 13.2. Wyłonienie obserwacji nietypowych makropoleceniem „Dane losowe” cz. 1.

Uruchomione makropolecenie podaje nam również wartość krytyczną $F_\alpha = 2,5999682$ rozkładu F -Snedecora przy warunkach:

$$\alpha = 0,05, m = 4, n - m = 45 - 4 = 41.$$

Kontynuacją widoku rozwiązania jest rysunek 1.3. Wyszukajmy teraz obserwacje nietypowe, czyli takie dla których

$$F_\alpha < F_i.$$

Warunek ten spełniają podregiony dla których wartość empiryczna statystyki F -Snedecora wynosi:

- miasto Łódź – 3,221234218,
- Gdańsk-Gdynia-Sopot – 3,56406862,
- Warszawski – 4,112186012,
- miasto Warszawa – 6,41488523.

Na rysunku 31.3 występują też obliczenia elementów składowych kwadratu odległości Mahalanobisa a mianowicie:

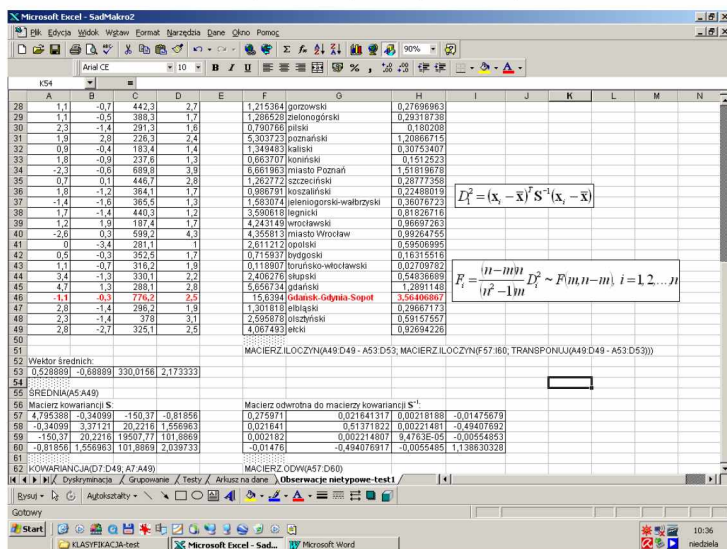
- wektor średnich $\bar{\mathbf{x}}$ dla zmiennych zapisanych w kolumnach A, B, C oraz D obliczonych funkcją *Excelsa*: ŚREDNIA(A5:A49);
- macierz kowariancji $\mathbf{S}$ ustalona przy wykorzystaniu funkcji statystycznej *Excelsa*: KOWARIANCJA(D5:D49;A5:A49);
- macierz odwrotna macierzy kowariancji: MACIERZ.ODW(A57:D60).

Pokazana na rysunku 13.3 formuła złożonej funkcji macierzowej *Excelsa* na obliczenie kwadratu odległości Mahalanobisa w odniesieniu do przykładowej obserwacji dotyczącej podregionu łódzkiego jest następująca:

$$MACIERZ.ILOCZYN(A49 : D49 - A53 : D53; MACIERZ.ILOCZYN(F57 : I60; TRANSPONUJ(A49 : D49 - A53 : D53)))$$

Odpowiada to operacji na wektorach różnic i macierzy odwrotnej w układzie:

$$D_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})\mathbf{S}^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})^T$$



Źródło: Opracowanie w *Excelsa* bazujące na pliku *SadMakro.xls*.

Rys. 13.3. Wyłonienie obserwacji nietypowych makropoleceniem „Dane losowe” cz. 2.

Wartości w kolumnie H stanowią zmodyfikowanie kwadratów odległości Mahalanobisa poprzez pomnożenie D_i^2 przez współczynnik:

$$\frac{(n-m)n}{(n^2-1)m}$$

Przykładowo dla operacji 45. statystyka empiryczna wynosi:

$$F_i = \frac{(45-4)45}{(45^2-1)4} 4,067493 = 0,92694226.$$

Trzeba tu nadmienić, że w sytuacji wpisywania formuł macierzowych w starszych wersjach *Excels* (bez udziału przycisku makropolecenia w skoroszybie *SadMakro2.xls*) musimy pamiętać o użyciu jednocześnie trzech klawiszy:

<CTRL>+<Shift>+<Enter>.

13.3. Test Shapiro-Wilka

Dla zbadania z ryzykiem $\alpha = 0,05$ istnienia wielowymiarowego rozkładu normalnego N naszych czterech cech przyjętych w metodzie zastosowania zmodyfikowanej odległości Mahalanobisa wychodzimy z hipotezy zerowej:

$$H_0 : \mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

gdzie:

$\boldsymbol{\mu}$ - wektor wartości oczekiwanych poszczególnych czterech cech o szeregach $n = 45$ obserwacji: $\mathbf{x}_1, \dots, \mathbf{x}_n$;

$\boldsymbol{\Sigma}$ - wektor odchyleń standardowych czterech cech.

Dla realizacji procedury testu Shapiro-Wilka stosujemy makropolecenie zawarte w pliku *SadMakro.xls* oraz uwzględnimy kroki postępowania zawarte w rozdziale 8.1¹³⁰. Wywołujemy arkusz „Testy” i kopiujemy szeregi zmiennych X_2, X_3, X_4, X_9 bez nagłówka z arkusza „Arkusz na dane” począwszy od komórki A4:D48. Zaznaczamy myszą zakres komórek i naciskamy OK.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Test Shapiro-Wilka	Test skosność	Test spłaszczenia	Test dla wektora średnich	Test macierzy kowariancji	Test średnich i kowariancji				Wyczytać arkusz					
2															
3															
4	-2,3	0,4	223,9	1,1											
5	-0,9	-1,0	216,3	1,7											
6	-7,0	-0,9	472,4	1,6											
7	0,6	-1,4	222,5	2,0											
8	1,5	-1,6	229,9	1,8											
9	-0,4	7,0	294,7	5,3											
10	1,4	-1,3	261,2	1,8											
11	-3,7	2,1	675,0	8,9											
12	1,4	0,7	223,6	1,7											
13	5,2	0,1	182,3	2,0											
14	-1,5	1,3	603,8	4,6											
15	-2,0	-0,8	276,2	1,2											
16	1,2	-1,4	257,9	1,5											
17	-1,7	-4,4	440,5	0,5											
18	1,4	-3,5	348,4	1,5											
19	0,5	-2,3	246,1	1,1											
20	-0,8	-2,1	220,0	0,9											
21	0,2	-0,7	268,2	1,6											
22	2,9	-0,6	186,1	1,5											
23	2,2	-1,4	135,0	1,2											
24	-0,4	-0,6	230,4	3,6											
25	2,0	-3,4	230,1	1,2											
26	-0,6	-1,6	250,5	1,4											
27	1,1	-0,7	442,3	2,7											
28	1,1	-0,5	389,3	1,7											
29	2,3	-1,4	291,3	1,6											

Źródło: Opracowanie w Excelu bazujące na pliku *SadMakro.xls*.

Rys. 13.4. Ekran inicjujący do testu Shapiro-Wilka.

¹³⁰ Miszczak W., *Statystyczne metody analizy danych Materiały do ćwiczeń*, op. cit., s. 46.

Uruchamiamy makropolecenie „Test Shapiro-Wilka” i podobnie jak dla zastosowanego obliczania zmodyfikowanej odległości Mahalanobisa następuje określenie:

- wektora średnich $\bar{\mathbf{x}}$,
- macierzy kowariancji $\mathbf{S}$,
- odwrotnej macierzy do macierzy kowariancji $\mathbf{S}^{-1}$.

Ponadto dla wszystkich $n = 45$ obserwacji obliczane są kwadraty odległości Mahalanobisa od wektora wierszowego:

$$\bar{\mathbf{x}} : d_M^2(\bar{\mathbf{x}}_i, \bar{\mathbf{x}})$$

oraz statystyki empiryczne U_i dla $i = 1, 2, \dots, n$ obserwacji według wzoru:

$$U_i = (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$$

Przy czym $\mathbf{x}_k$ to obserwacja dla której kwadrat odległości Mahalanobisa $d_M^2(\bar{\mathbf{x}}_k, \bar{\mathbf{x}})$ jest największy.

Jest to obserwacja ósma (podregion Miasto Warszawa) o wartości zmiennych:

- przyrost naturalny $x_2 = -3,7$,
- saldo migracji wewnętrznych i zagranicznych na pobyt stały $x_3 = 2,1$,
- przestępstwa stwierdzone $x_4 = 575$,
- mieszkania oddane do użytkowania $x_9 = 8,9$.

oraz maksymalnej wartości odległości Mahalanobisa równej 28,149018 (zob. rysunek 13.5).

Makro testu Shapiro-Wilka układu narastająco statystyki empiryczne (rys. 13.5) a następnie oblicza kolejną statystykę wynikową W :

$$W = \frac{\left(\sum_{i=1}^h a_{(i,n)} (U_{(n-1+i)} - U_{(i)}) \right)^2}{n d_M^2(\mathbf{x}_k, \bar{\mathbf{x}})}$$

We wzorze tym dla nieparzystej liczby obserwacji (45): $h = \frac{n-1}{2} = \frac{45-1}{2} = 22$.

Wartości współczynników $a_{(i,n)}$ wprowadzamy manualnie do testu na podstawie poniższego zestawienia współczynników dla testu Shapiro-Wilka¹³¹ dla $n = 45$ i kolejnych i 1-22:

i	1	2	3	4	5	6	7	8	9	10	11
a	0,3850	0,2651	0,2313	0,2065	0,1865	0,1695	0,1545	0,1410	0,1286	0,1170	0,1062
i	12	13	14	15	16	17	18	19	20	21	22
a	0,0959	0,0860	0,0765	0,0673	0,0584	0,0497	0,0412	0,0328	0,0245	0,0163	0,0081

Makropolecenie testu Shapiro-Wilka wyznacza wartość statystyki empirycznej W , która wynosi 0,683908 (zob. rysunek 13.5). Wartość teoretyczną $W_\alpha = 0,9174$ dla $m = 4$ zmienne, $\alpha = 0,05$ i $n = 45$ obserwacji odczytujemy z tablicy 2 „Wartości krytyczne dla testu wielowymiarowej normalności Shapiro-Wilka”¹³². Teraz sprawdzamy Warunek prawdziwości hipotezy zerowej:

$$P(W < W_\alpha) = \alpha$$

Hipotezę H_0 odrzucamy, gdyż $W = 0,683908 < W_\alpha = 0,9174$. Tak więc zaobserwowana próba nie pochodzi z czterowymiarowego rozkładu normalnego i wstępne wyłonienie obserwacji nietypowych poprzez zastosowanie zmodyfikowanej odległości Mahalanobisa nie było słuszne. Pozostaje więc nam skorzystanie z metody wykrywania obserwacji nietypowych w sytuacji nieznanego rozkładu cech, zwanej metodą Gnanadesikana-Ketteringa (GK).

¹³¹ Ibidem, s. 52.

¹³² Ibidem, s. 56.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
34	0,7	0,1	446,7	2,8	1,262772			-0,12916		-0,22167					
35	1,8	-1,2	364,1	1,7	0,986791			-0,06169		-0,20927					
36	-1,4	-1,6	365,5	1,3	1,583074			-2,36498		-0,12916					
37	1,7	-1,4	440,3	1,2	3,590618			-6,50619		0,331099					
38	1,2	1,9	187,4	1,7	4,243149			-4,09019		0,448127					
39	-2,6	0,3	699,2	4,3	4,355813			6,555842		0,516817					
40	0	-3,4	281,1	1	2,611212			-0,75071		0,999136					
41	0,5	-0,3	362,5	1,7	0,715937			-3,28554		2,023946					
42	1,1	-0,7	316,2	1,9	0,118907			-1,49262		3,063067					
43	3,4	-1,3	330,1	2,2	2,486276			-0,91647		3,630157					
44	4,7	1,3	208,1	2,6	5,656734			-1,82135		4,20473					
45	-1,1	-0,3	776,2	2,5	15,6394			-5,49343		5,744263					
46	2,8	-1,4	296,2	1,9	1,301818			-1,2833		5,896736					
47	2,3	-1,4	378	3,1	2,593878			3,630157		6,659842					
48	2,8	-2,7	325,1	2,6	4,067493			3,083087		9,283317					
49										26,14902					
50	Wektor średnich:														
51															
52	0,528889	-0,68889	330,0156	2,173333											
53															
54	Macierz kowariancji:					Macierz odwrotna do macierzy kowariancji:									
55															
56	4,795388	-0,34099	-150,37	-0,81856		0,275971	0,021641	0,002182		-0,01476					
57	-0,34099	3,37121	20,2216	1,556963		0,021641	0,513718	0,002215		-0,49408					
58	-150,37	20,2216	19607,77	101,8869		0,002182	0,002215	9,48E-05		-0,00555					
59	-0,81856	1,556963	101,8869	2,039733		-0,01476	-0,49408	-0,00555		1,13063					
60															
61															
62															
63															
64															
65															

Źródło: Opracowanie w Excelu bazujące na pliku *SadMakro.xls*.

Rys. 13.6. Raport z testu Shapiro-Wilka c.d.

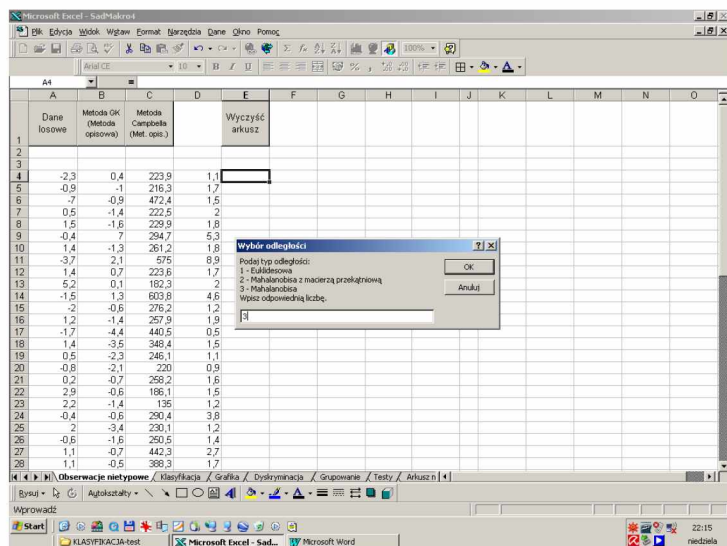
13.4. Wykrywanie obserwacji nietypowych w sytuacji nieznanego rozkładu cech

Przyjmijmy teraz, że rozkład wektora czterech cech podregionów jest nieznan. W wyniku przeprowadzenia prób nad użyciem kolejnej metody wykrywania obserwacji nietypowych posłużono się plikiem *SadMakro4.xls* stanowiącym kopię *SadMakro2.xls* bez danych.

Z arkusza „Arkusz na dane” kopiujemy szeregi 45. obserwacji a następnie naciskamy przycisk „Metoda GK”. Makropolecenie żąda od nas zaznaczenia myszą zakresu komórek zawierających dane liczbowe w postaci adresowania bezwzględnego (\$A\$4:\$D\$48). W kolejności wybieramy typ obliczanej odległości danej obserwacji od środka ciężkości zbioru obserwacji spośród:

- 1 - Euklidesowa,
- 2 - Mahalonobisa z macierzą przekątniową,
- 3 - Mahalonobisa.

Dla porównania z wynikami już uzyskanymi metodą wykrywania losowych obserwacji nietypowych przyjmijmy opcję 3 (zob. rysunek 13.7).



Źródło: Opracowanie w Excelu bazujące na pliku *SadMakro.xls*.

Rys. 13.7. Wybór typu odległości w metodzie GK.

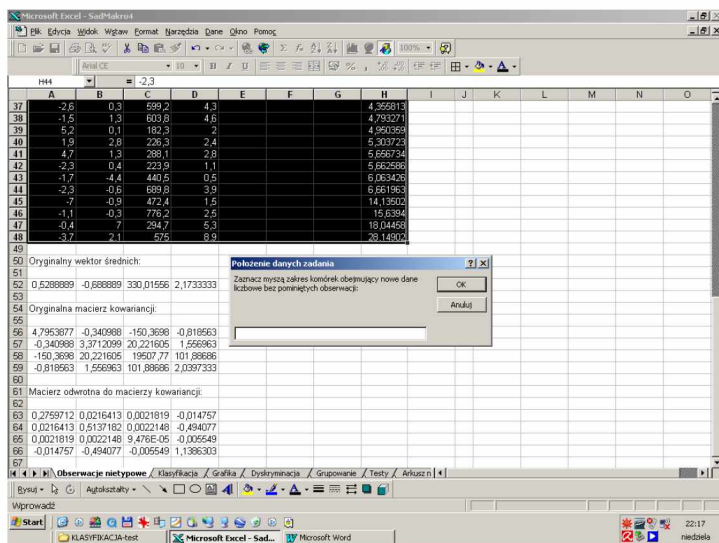
Po naciśnięciu OK uzyskujemy tabelę obserwacji posortowaną narastająco według kwadratu odległości Mahalanobisa D_i^2 oraz dodatkowo w odniesieniu do oryginalnego – wstępnego zestawu 45. obserwacji następujące zestawienia:

- wektor średnich,
- macierz kowariancji,
- macierz odwrotna do macierzy kowariancji.

Prezentuje to widok ekranu (zob. rys. 13.8). Z ciekawości zobaczmy, czy wyliczenie D_i^2 jest identyczne jak dla metody wykrywania losowych obserwacji na przykładzie obserwacji ósmej (zob. rys. 13.3, wiersz 12 danych oraz rysunek 13.8, wiersz 48 danych) dla których zmienne mają wartości:

$$\begin{array}{c|c|c|c} x_2 & x_3 & x_4 & x_9 \\ \hline -3,7 & 2,1 & 575 & 8,9 \end{array}$$

Wartość D_i^2 w kolumnie F, w wierszu 12. wynosi 28,149018 a w wierszu 48 kolumny H – 28,14902 (zob. rysunki 13.2 oraz 13.8). Zatem wyniki możemy uznać za identyczne, gdyż jest to kwestia precyzji obliczenia.



Źródło: Opracowanie w Excelu bazujące na pliku *SadMakro.xls*.

Rys. 13.8. Raport makropolecenia metody GK.

Widzimy, biorąc pod uwagę D_i^2 , że wyraźnie różniące są cztery ostatnie obserwacje (zob. rysunek 13.8) o następujących wartościach:

Podregion	Obserwacja wg wykresu 2	D_i^2
3	Miasto Łódź	14,13502
42	Gdańsk-Gdynia-Sopot	15,63940
6	Warszawski	18,04458
8	Miasto Warszawa	28,14902

Tak więc rozwiązanie potwierdza wyniki otrzymane według metody wykrywania obserwacji nietypowych z rozkładem wielowymiarowym normalnym. Z rysunku 13.8 wynika, że makropolecenie zachęca nas do kolejnej iteracji bez pominiętych obserwacji – w naszym przypadku czterech ostatnich. Próba korzystania z tego makropolecenia wykazała celowość realizacji nieco innej procedury a mianowicie:

- skopiowanie 41. obserwacji tj. bez pominiętych,
- wyczyszczenie arkusza (przycisk „Wyczyść dane”),
- wklejenie danych,
- ponowne uruchomienie przyciskiem „Metoda GK”.

W rezultacie dla 44. obserwacji uzyskujemy na nowo obliczone wartości D_i^2 oraz:

- wektor średnich traktujących 44 obserwacje jako oryginalne (wstępne), chociaż stanowią zmodyfikowany zestaw danych;
- macierz kowariancji;
- macierz odwrotną do macierzy kowariancji.

Zwróćmy uwagę na obserwacje 21, (podregion: białostocko-suwański) o cechach (zob. rys. 13.9 – wiersz 44):

x_2	x_3	x_4	x_9	D_i^2
-0,4	-0,6	290,4	3,8	13,39715

O ile cztery podregiony pierwszej iteracji wyraźnie różniły się w sensie dodatnim cech to podregion 21. wykazuje „odstawianie” odwrotne i z tego względu możemy daną obserwację potraktować jako nietypową (zob. rysunek 13.9, wiersz 44 na arkuszu kalkulacyjnym).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
26	-0,8	-2,1	220	0,9				2,99187							
27	-0,9	-1	216,3	1,7				3,314378							
28	0	-3,4	281,1	1				3,325723							
29	1,4	-3,5	348,4	1,5				3,76224							
30	2,3	-1,4	376	3,1				3,904644							
31	2	-3,4	230,1	1,2				3,941055							
32	-2	-0,6	276,2	1,2				4,40642							
33	2,8	-2,7	325,1	2,5				5,497993							
34	1,2	1,9	187,4	1,7				6,001014							
35	5,2	0,1	182,3	2				6,159944							
36	4,7	1,3	288,1	2,8				7,149147							
37	1,9	2,8	226,3	2,4				7,824752							
38	-2,3	0,4	223,9	1,1				7,974108							
39	1,7	-1,4	440,3	1,2				8,402961							
40	-2,6	0,3	599,2	4,3				9,615208							
41	-1,5	1,3	603,8	4,6				10,05986							
42	-2,3	-0,6	609,6	3,9				10,45414							
43	-1,7	-4,4	440,5	0,5				10,88434							
44	-0,4	-0,6	290,4	3,8				13,39715							
45															
46	Originalny wektor średnich:														
47															
48	0,8780488	-0,94878	310,54634	1,9414634											
49															
50	Originalna macierz kowariancji:														
51								Macierz odwrotna do macierzy kowariancji:							
52	3,2368352	0,1550268	-89,90093	-0,16348				0,3992392	0,042029	0,003808	-0,26663				
53	0,1550268	1,906167	3,5942118	0,5981202				0,042029	0,7634562	0,004492	-0,913924				
54	-89,90093	3,5942118	14187,564	88,595683				0,003808	0,0044917	0,000176	-0,017034				
55	-0,16348	0,5981202	88,595683	0,8292564				-0,26663	-0,913924	-0,017034	3,2215632				
56															

Źródło: Opracowanie w Excelu bazujące na pliku *SadMakro.xls*.

Rys. 13.9. Raport makropolecenia metody GK dla zestawu danych bez obserwacji pominiętych

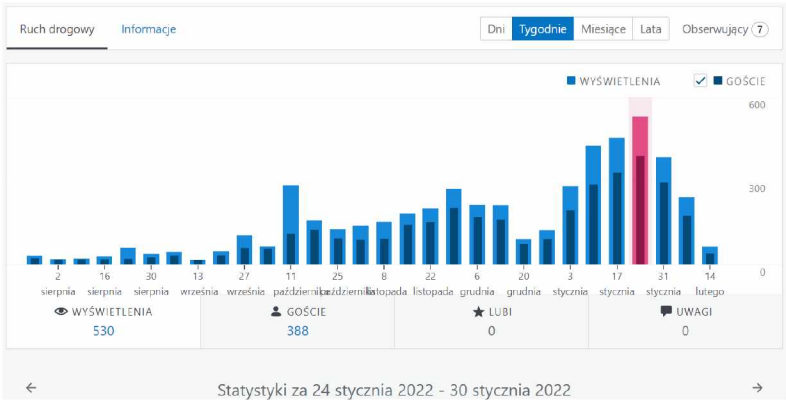
* * *

Kwadrat odległości Mahalanobisa nie jest jednak rażąco różniący się od wcześniejszego o wartości 10,88434 (podregion 14: centralny Śląski). Trzeba tu zauważyć, że w opisowej metodzie GK następuje tylko wskazanie nietypowości. Użytkownik musi zatem sam rozstrzygnąć, które obserwacje uzna za nietypowe.

14. Dostęp do publikacji w latach 2008-2025 (Władysław Wornalkiewicz)



Coraz częściej spotykam się z dużym zainteresowaniem internautów moimi publikacjami i wpisami, a świadczy o tym tygodniowa statystyka w *WordPress* na blogu: <https://wornalkiewicz.wordpress.com/> (zob. rysunek 14.1).



Źródło: Opracowanie własne na podstawie blogu „Procesy informacyjne w teorii i praktyce”.
Rys. 14.1. Statystyka w układzie tygodni na blogu moim blogu

Oprócz tego forum, współpracy z sympatykami metod ilościowych w domenie „Google” w Internecie, znajdują się pozycje ze zmiankami o moim dorobku naukowym na stronie WWW Akademii Nauk Stosowanych WSZiA w Opolu. Dalej zostaną zaprezentowane wybrane z tej strony moje pozycje publikacji zarówno jako książki, monografie indywidualne oraz artykuły w monografiach zbiorowych. Zaprezentowane zostaną publikacje sygnalizowane w domenie Google po wywołaniu autora „Władysław Wornalkiewicz”. Całość opracowano w układzie kolejności narastającej lat (2008-2024).

2008	
1. „Ekonomiczno-społeczne problemy współczesnego zarządzania i komunikacji”, monografia zbiorowa recenzowana, Opole: Wydawnictwo Instytut Śląski, 2008, 255 s., artykuł „Dobór metod i parametrów w budowaniu modeli ekonometrycznych”, strony: 190-204, ISBN 978-83-88980-69-5, 978-83-7511-098-2. Link: https://wordpress.com/media/wornalkiewicz.wordpress.com .	

2010

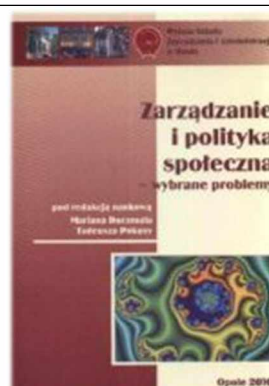
1. „*Wstęp do ekonometrii i badań operacyjnych Zbiór przykładów z zastosowaniem mikrokomputera*”, monografia zbiorowa recenzowana (podręcznik akademicki), współautorstwo - Marian Duczmal, Opole: Wydawnictwo Instytut Śląski, 2010, 597 s., ISBN 978-8388980-94-7, 978-83-62105-41-0.

Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Wstep_do_ekonometrii_i_badan_operacyjnych.pdf.



2. „*Zarządzanie i polityka społeczna - wybrane problemy*”, monografia zbiorowa recenzowana, Opole: Wydawnictwo Instytut Śląski, 2010, 443 s., artykuł „*Estymacja modeli wielorównaniowych w GRETL*” strony: 255-276, ISBN 978-83-88980-90-9, 978-83-62105-23-6.

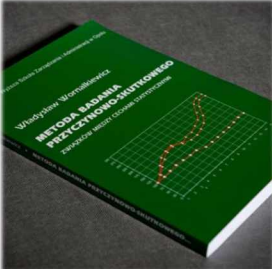
Link: <https://wordpress.com/media/wornalkiewicz.wordpress.com>.

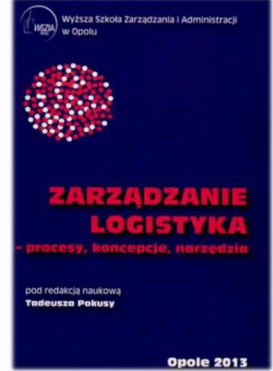
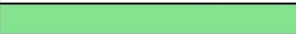


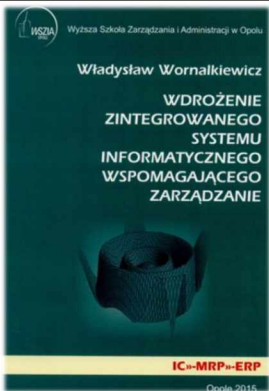
3. „*Nowoczesne zarządzanie - wybrane problemy ekonomiczno-społeczne*”, monografia zbiorowa recenzowana, Opole: Wydawnictwo Instytut Śląski, 2010, 242 s., artykuł: „*Propozycja określenia efektywności województw*”, strony: 189-208, ISBN 978-83-88980-98-5, 978-83-62105-53-3.

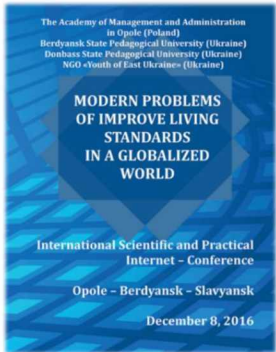
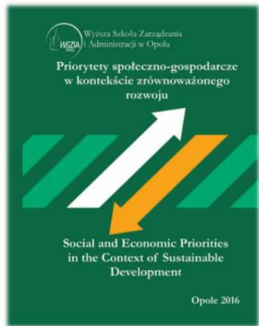
Link: <https://integro.bg.polsl.pl/172200374055/ksiazka/nowoczesne-zarzadzanie-wybrane-problemy-ekonomiczno-spoeczne?bibFilter=17>.

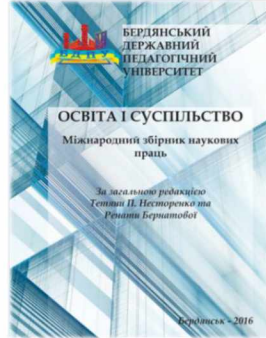
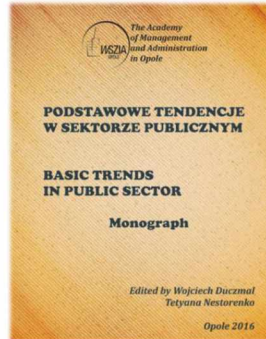


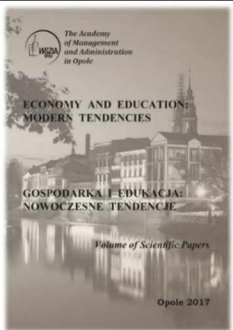
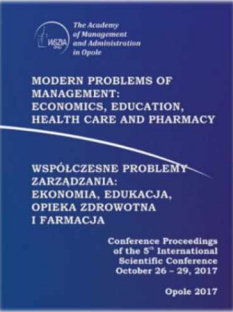
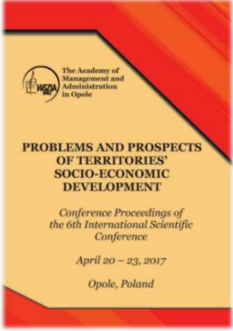
2011 1. „<i>Przejawy wielowymiarowości współczesnego zarządzania - formy i instrumenty ekonomiczno-społeczne</i>”, monografia zbiorowa recenzowana, Opole: Wydawnictwo Instytut Śląski, 2011, 394 s., artykuł: „<i>Model ekonometryczny zmiennej „liczba bezrobotnych”</i>”, strony: 172-186, ISBN 978-83-62683-20-8, 978 - 83-7511-135-4. Link: https://wordpress.com/media/wornalkiewicz.wordpress.com.	
2012 1. „<i>Wstęp do ekonometrii i badań operacyjnych Zbiór przykładów z zastosowaniem mikrokomputera</i>”, wydanie II rozszerzone, część I, monografia zbiorowa recenzowana (podręcznik akademicki), współautorstwo - Marian Duczmał, Opole: Wydawnictwo Instytut Śląski, 2012, 411 s., ISBN 978-83-62683-30-7, 978-83-7511-259-0. 2. „<i>Wstęp do ekonometrii i badań operacyjnych Zbiór przykładów z zastosowaniem mikrokomputera</i>”, wydanie II rozszerzone, część II, monografia zbiorowa recenzowana (podręcznik akademicki), współautorstwo - Marian Duczmał, Opole: Wydawnictwo Instytut Śląski, 2012, 223 s., ISBN 978-83-62683-30-7, 978-83-7511-259-0. 3. „<i>Zarządzanie Logistyka Finanse - Problemy innowacyjności i instrumenty analizy</i>”, monografia zbiorowa recenzowana, Opole: Wyższa Szkoła Zarządzania i Administracji w Opolu, 2012, 287 s., artykuły: „<i>Model wielorównaniowy PKB</i>”, strony: 223-249, „<i>Etapy i procedury budowy modelu produktu brutto podregionów</i>”, strony: 261-279, ISBN Link: https://w.bibliotece.pl/1798765/Zarz%C4%85dzanie+logistyka+finanse.	
2013 1. „<i>Metoda badania przyczynowo-skutkowego związków między cechami statystycznymi</i>”, skrypt nr 1/2013 - pomocniczy do przedmiotu <i>Ekonometria</i>, Opole: Wyższa Szkoła Zarządzania i Administracji w Opolu, 2013, 252 s., ISBN 978-83-62683-44-4, 978-83-7511-187-3. Link: https://wordpress.com/media/wornalkiewicz.wordpress.com; https://books.google.pl/books/about/Metoda_badania_przyczynowo_skutkowego_zw.html?id=0d0DoQEA-CAAJ&redir_esc=y.	

2. „<i>Modele ekonometryczne PKB obiektów struktury terytorialnej</i>”, monografia indywidualna recenzowana, Opole: Wydawnictwo Instytut Śląski, 2013, 343 s., ISBN 978-83-62683-36-9, 978-83-7511-170-5. Link: https://integro.bs.katowice.pl/32403087968/wornalkiewicz-wladyslaw/modele-ekonometryczne-pkb-obiektow-struktury-terytorialnej.	
3. „<i>Zarządzanie Logistyka - procesy, koncepcje, narzędzia</i>”, monografia zbiorowa recenzowana, Opole: Wydawnictwo Instytut Śląski, 2013, 304 s., artykuł „<i>Programowanie wielokryterialne z zastosowaniem WinQSB</i>”, strony: 168-184, ISBN 978-83-62683-49-9, 978-83-7511-192-7. Link: https://integro.ciniba.edu.pl/integro/192504222555/ksika/zarzdanielogistykakoncepcjeprocenyarnarzdzia.	
4. „<i>Společno-ekonomiczne uwarunkowania zarządzania i administracji - innowacyjność, komunikacja</i>”, monografia zbiorowa recenzowana, Opole: Wydawnictwo Instytut Śląski, 2013, 251 s., artykuł „<i>Wybór lokalizacji obiektu z zastosowaniem Expert Choice</i>”, strony: 124-140, ISBN 978-83-62683-53-6, 978-83-7511-103-4. Link: https://wordpress.com/media/wornalkiewicz.wordpress.com.	
2014	
1. „<i>Formułowanie modeli ekonometrycznych na potrzeby zarządzania</i>” cz. I, monografia indywidualna recenzowana (podręcznik akademicki), Opole: Wydawnictwo Instytut Śląski, 2014, 665 s., ISBN 978-83-62683-64-2, 978-83-7511-210-8. 2. „<i>Formułowanie modeli ekonometrycznych na potrzeby zarządzania</i>” cz. II, monografia indywidualna recenzowana (podręcznik akademicki), Opole: Wydawnictwo Instytut Śląski, 2015, 665 s., ISBN 978-83-62683-64-2, 978-83-7511-210-8. Link: https://integro.bg.polsl.pl/172600881896/wornalkiewicz-wladyslaw/formulowanie-modeli-ekonometrycznych-na-potrzeby-zarzadzania.	

2015 1. „<i>Wdrożenie zintegrowanego systemu informatycznego wspomagającego zarządzanie</i>”, monografia indywidualna recenzowana (podręcznik akademicki), Opole: Wydawnictwo Instytut Śląski, 2015, 370 s., ISBN 978-83-62683-67-3, 978-83-7511-226-9. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Womalkiewicz_Wdrozenie_zintegrowanego_sys-temu_informatycznego.pdf.	
2. „<i>Ekonomiczno-społeczne uwarunkowania rozwoju gospodarczego - zarządzanie informacją i nowymi technologiami</i>”, monografia zbiorowa recenzowana, Opole: Wydawnictwo „Instytut Śląski”, 2015, 507 s., artykuły i strony: „<i>Przejawy wdrożenia systemów informatycznych</i>” (85-108), „<i>Modelowanie procesów zarządzania</i>” (124-146), współautorstwo: Ryszard Broszkiewicz, „<i>EDI w procesie logistycznym</i>” (278-301), „<i>Analityka biznesowa</i>” (356-375). Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/2_2021.pdf.	
3. „<i>Koncepcje i narzędzia współczesnej logistyki - monografia upowszechniająca projekt „WSZiA w Opolu w perspektywie realizacji Strategii Europa 2020 Wzmocnienie potencjału dydaktycznego Uczelni w obszarze kształcenia kadr dla sektora logistyki poprzez wdrożenie programu rozwojowego”</i>”, praca zbiorowa, Opole: Wydawnictwo Instytut Śląski, 102 s., artykuł „<i>Implementacja systemów klasy ERP w logistyce</i>”, ISBN 978-83-62683-75-8, 978-83-7511-235-1. Link: https://wordpress.com/media/wornalkiewicz.wordpress.com.	
2016 1. „<i>Wprowadzenie do projektowania systemów informatycznych zarządzania</i>” Część 1, monografia indywidualna recenzowana (podręcznik), Opole: Wydawnictwo Wyższej Szkoły Zarządzania i Administracji w Opolu, 2016, 328 s., ISBN 978-83-62683-97-0, 978-83-7511-243-6. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Projektowanie_systemow_informatycznych_zarzadzania.pdf. 2. „<i>Wprowadzenie do projektowania systemów informatycznych zarządzania</i>” część 2, monografia indywidualna recenzowana	

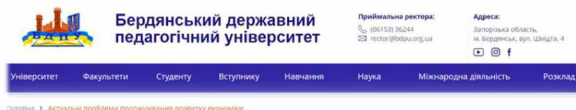
(podręcznik), Opole: Wydawnictwo Wyższej Szkoły Zarządzania i Administracji w Opolu, 2016, 567 s., ISBN 978-83-62683-97-0, 978-83-7511-243-6. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Wprowadzenie_do_projektowania_Czesc2.pdf.	
3. „<i>Conference Proceedings of the International Scientific Internet-Conference Modern Problems of Improve Living Standards in a Globalized World</i>”, materiały pokonferencyjne - monografia zbiorowa recenzowana, Opole - Berdyansk - Slavyansk), 2016, (electronic edition), 534 s., artykuł „<i>Rozwiązanie problemu transportowego metodą VAM</i>” (22-28), ISBN 978-83-62683-871. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2016_modern_problems_of_imprve_living_standards_in_a_globalized_world_slavyansk.pdf.	
4. „<i>Social and Economic Priorities in the Context of Sustainable Development</i>”, monografia zbiorowa recenzowana, Opole: The Academy of Management and Administration in Opole, 2016, 444 s., artykuł „<i>Product promotion and company image in Internet</i>” (<i>Promocja produktu i wizerunku firmy w Internecie</i>), strony: 138-148, , ISBN 978 - 83 - 62683 -78 - 9. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2016_priorytet_spoleczno_gospodarcze_w_kontekscie_zrownowazonego_rozwoju.pdf.	
5. „<i>20 lat Wyższa Szkoła Zarządzania i Administracji w Opolu - interdyscyplinarność badań</i>”, monografia zbiorowa recenzowana jubileuszowa, Opole: Wydawnictwo Instytut Śląski, 2016, 375 s., artykuł: „<i>Wyszukiwarki i media społecznościowe</i>”, strony: 320-333, ISBN 978-83-62683-86-4, 978-83-7511-247-4. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2016_wyzsza_szkola_zarzadzania_i_administracji_w_opolu_interdyscyplinarnosc_badan.pdf.	

6. „СОЦІАЛЬНО-ЕКОНОМІЧНИЙ РОЗВИТОК: РОЛЬ ІНФОРМАЦІЇ ТА НОВИХ ТЕХНОЛОГІЙ”, monografia zbiorowa recenzowana, Бердянський державний педагогічний університет, 2016, 295 s., artykuł „<i>Symulacja biznesowa</i>”, strony: 197-208, ISBN 978-617-7291-44-1. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2016_%D0%9E%D0%A6%D0%86%D0%90%D0%9B%D0%AC%D0%9D%D0%9E_%D0%95%D0%9A%D0%9E%D0%9D%D0%9E%D0%9C%D0%86%D0%A7%D0%9D%D0%98%D0%99.pdf.	
7. „Освіта і суспільство. Міжнародний збірник наукових праць”, Бердянський державний педагогічний університет, 2016, 418 s., artykuł „<i>Ranking metod ilościowych w Internecie</i>”, strony: 275-285, ISBN 978-617-7291-80-9 (електронне видання). Link: https://bdpu.org.ua/wp-content/uploads/2019/10/Papers_Berdiansk_2016.pdf.	
8. „<i>Basic Trends in Public Sector</i>” („<i>Podstawowe tendencje w sektorze publicznym</i>”), monografia zbiorowa recenzowana, artykuł: <i>Designing of managerial consoles (Projektowanie pulpitów menadżerskich)</i>, strony: 58-74. Opole: The Academy of Management and Administration in Opole, 2016, 200 s., ISBN 978-83-62683-79-6. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2016_podstawowe_tendencje_w_sektorze_publicznym_duczmal_nestoreka-1.pdf.	

2017 1. „<i>ECONOMY AND EDUCATION: MODERN TENDENCIES</i>” - <i>GOSPODARKA I EDUKACJA: NOWOCZESNE TENDENCJE</i>, Volume of Scientific Papers, monografia zbiorowa recenzowana, Wyższa Szkoła Zarządzania i Administracji w Opolu, 2017, 362 s., artykuł „<i>Prognostowanie z wykorzystaniem zasady postarzania informacji</i>”, strony: 39-43, ISBN 978 - 83 - 62683 - 24 - 6. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2017_economy_andeducation_modern_tendencies.pdf.	
2. „<i>Modern Problems of Management: Economics, Education, Health Care and Pharmacy</i>”, Conference Proceedings of the 5 th International Scientific Conference, 232 s., Opole, The Academy of Management and Administration in Opole, 2017, artykuł „<i>Zastosowanie konwertera plików</i>”, strony: 25-27, 978-83-62683-23-9. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2017_modern_problems_of_management_economics_education_health_care_and_pharmacy.pdf.	
3. „<i>Popularyzacja wybranych metod ilościowych w Internecie</i>”, monografia indywidualna, Wyższa Szkoła Zarządzania i administracji w Opolu, Opole, 2017, 351 s. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2017_popularyzacja_wybranych_metod_ilosciowych_w_intemeice_wornalkiewicz.pdf.	
4. „<i>Problems and Prospects of Territories' Socio-Economic Development</i>”, Conference Proceedings of the 6 th International Scientific Conference, materiały pokonferencyjne, monografia zbiorowa recenzowana, The Academy of Management and Administration in Opole, 2017, 272 s., artykuł: „<i>Optymalizacja marszrutyacji przewozów z zastosowaniem funkcji Excelsa</i>”, strony: 41-43, ISBN 978-83-62683-10-9. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2017_problems_and_prodprrcts_pf_territories_socio_economic_development.pdf.	

5. „Актуальні проблеми прогнозування розвитку економіки...” (*„Bieżące problemy prognozowania ekonomiki Ukrainy”*), zbiór prac naukowych - monografia zbiorowa recenzowana, 370 s., КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА (Kijowski Narodowy Uniwersytet im. Tarasa Szewczenki) i inne, 2017, „Optymalizacja przewozów z zastosowaniem funkcji Excela” strony: 154-169, ISBN 978-617-7291-98-4.

Link: <https://bdpu.org.ua/actual-problems-of-forecasting-economic-development/>.

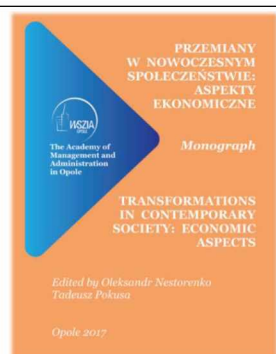


Актуальні проблеми прогнозування розвитку економіки



6. „Transformations in contemporary society: economic aspects” (*Przemiany w nowoczesnym społeczeństwie: aspekty ekonomiczne*), monografia zbiorowa recenzowana, Wyższa Szkoła Zarządzania i Administracji w Opolu, 2017, 348 s., artykuł: „Forecasting using the multiplicative model” (*„Prognozowanie z zastosowaniem modelu multiplikatywnego”*), strony: 205-212, ISBN 978-83-62683-96-3.

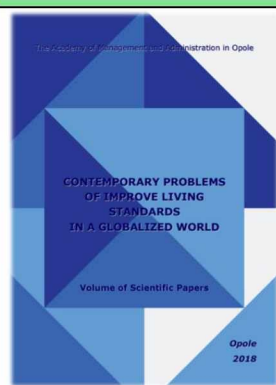
Link: https://www.wsia.opole.pl/wp-content/uploads/2020/09/2017_przemiany_w_nowoczesnym_spoleczenstwie_aspekty_ekonomiczne_nosferanko_pokusa.pdf.

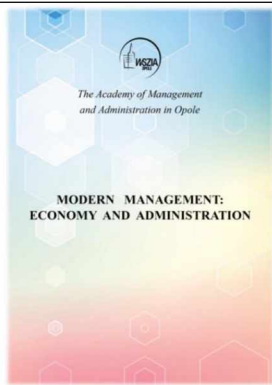
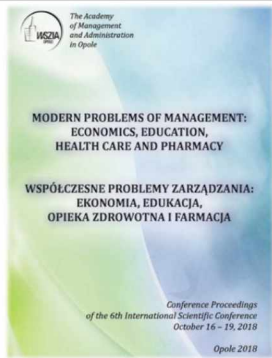
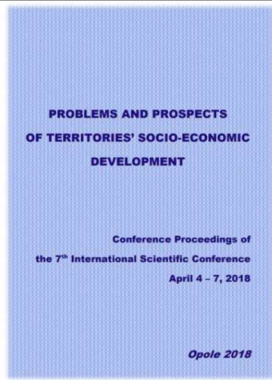


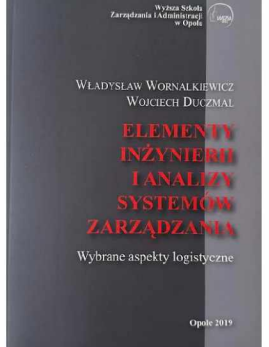
2018

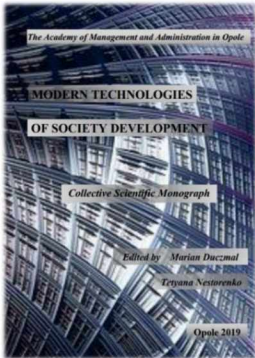
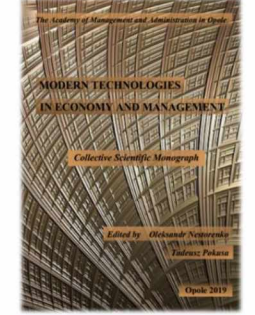
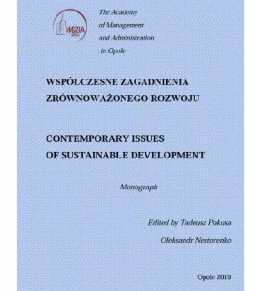
1. „Contemporary Problems of Improve Living Standards in a Globalized World”, Volume of Scientific Papers, monografia zbiorowa recenzowana, The Academy of Management and Administration in Opole, Opole, 2018, electronic edition, 770 s., artykuł „Echa migracji w wybranej literaturze (Stosowane metody i modele)”, strony: 516-527, ISBN 978-83-946765-1-3.

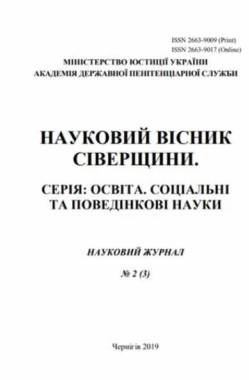
Link: https://www.wsia.opole.pl/wp-content/uploads/2020/09/2018_contemporary_problems_of_improve_living_in_a_globalized_world.pdf.

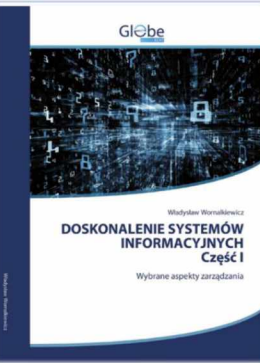
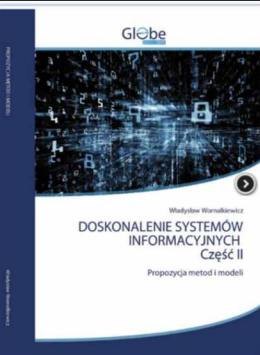


2. „<i>MODERN MANAGEMENT: ECONOMY AND ADMINISTRATION</i>” (<i>NOWOCZESNE ZARZĄDZANIE: EKONOMIA I ADMINISTRACJA</i>), monografia zbiorowa recenzowana, The Academy of Management and Administration in Opole, Opole, 2018, 218 s., artykuł „<i>Applications used in designing websites</i>” (<i>Aplikacje stosowane w projektowaniu stron www</i>) , strony: 154-160, ISBN 978-83-62-683-27-7. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2018_modern_management_economy_and_administration.pdf.	
3. „<i>Modern Problems of Management: Economics, Education, Health Care and Pharmacy</i>” (<i>Współczesne problemy zarządzania: Ekonomia, edukacja, opieka zdrowotna i farmacja</i>), Conference Proceedings of the 6th International Scientific Conference, Opole, The Academy of Management and Administration in Opole, 2018, 116 s., artykuł „<i>Potrzeba utworzenia piramidy wskaźników procesów ludnościowych</i>”, strony: 22-25, ISBN 978-83-946765-1-3. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2018_modern_problems_of_management_economics_education_health_nad_pharmacy.pdf.	
4. „<i>Problems and Prospects of Territories' Socio-Economic Development</i>” Conference Proceedings of the 7 th International Scientific Conference, Opole, materiały pokonferencyjne, monografia zbiorowa recenzowana, The Academy of Management and Administration in Opole, 2018, 345 s., artykuł „<i>Ocena nieefektywności gospodarowania zasobami ludności</i>”, strony: 70-72, ISBN 978 - 83 - 62683 - 25 - 3. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2018_problems_andprospects_of_territories_socio_economic_development.pdf.	

5. „Uwarunkowania demograficzne rozwoju społecznego i ekonomicznego miasta Nysa i powiatu nyskiego”, Opole-Nysa, Wyższa Szkoła Zarządzania i Administracji w Opolu, 2018, 260 s., artykuł: „Prognostowanie migracji ludności z uwzględnieniem wag harmonicznych”, strony: 212-233, ISBN 978-83-946765-0-6.	 WYŻSZA SZKOŁA ZARZĄDZANIA I ADMINISTRACJI W OPOLE PANSTWOWA WYŻSZA SZKOŁA ZARZĄDZANIA W NYSIE UWARUNKOWANIA DEMOGRAFICZNE ROZWOJU SPOŁECZNEGO I EKONOMICZNEGO MIASTA NYSY I POWIATU NYSKIEGO POD REDAKCJĄ NAJWIĘKSZĄ ROBERTA RAJZONOWICZA TAMARA POLESKI TOMASZA DZIUBANOWSKA Opole-Nysa 2018
2019 1. „Elementy inżynierii i analizy danych systemów zarządzania Wybrane aspekty logistyczne”, monografia zbiorowa recenzowana (podręcznik akademicki), współautorstwo - Wojciech Duczmał, Opole, Wydawnictwo Wyższej Szkoły Zarządzania i Administracji w Opolu, 2019, 341 s., ISBN 978-83-946765-8-2. Link: https://www.wsza.opole.pl/wp-content/uploads/2020/05/Elementy_inzynierii_i_analazy_systemow_zarzadzania.pdf.	 WYŻSZA SZKOŁA ZARZĄDZANIA I ADMINISTRACJI W OPOLE WŁADYSŁAW WORNALKIEWICZ WOJCIECH DUCZMAŁ ELEMENTY INŻYNIERII I ANALIZY SYSTEMÓW ZARZĄDZANIA Wybrane aspekty logistyczne Opole 2019
2. „Improving the Life Quality: View of Scientists”, Volume of Scientific Papers, monografia zbiorowa recenzowana, The Academy of Management and Administration in Opole, Opole, 2019, electronic edition, 660 s., artykuły: „Pozyskiwanie danych o odległościach dla potrzeb zagadnienia transportowego” strony: 118-136, „Prognostowanie zewnętrznych usług transportowych” strony: 137-156, „Zastosowanie taksonomii wrocławskiej” (157-188), ISBN 978-83- 946765-3-7. Link: https://www.wsza.opole.pl/wp-content/uploads/2020/09/2019_improving_the_life_quality_view_of_scientists.pdf.	 The Academy of Management and Administration in Opole Improving the Life Quality: View of Scientists Volume of Scientific Papers Opole 2019

3 „<i>Modern Technologies of Society Development</i>”, Collective Scientific Monograph, monografia zbiorowa recenzowana, Opole: The Academy of Management and Administration in Opole, 2019, 410 s., artykuł: „<i>Verification of the research when comparing in pairs</i>” (<i>Weryfikacja badania przy porównywaniu parami</i>), strony: 119-129, ISBN 978- 83-946765-6-8. Link: https://www.wsia.opole.pl/wp-content/uploads/2020/09/2019_modern_technologies_in_education.pdf.	
4. „<i>Modern Technologies in Education</i>”, Collective Scientific Monograph, monografia zbiorowa recenzowana, Opole: The Academy of Management and Administration in Opole, 2019, 495 s., artykuł: „<i>Converting PDF to DOC</i>” (<i>Konwersja pliku formatu PDF na DOC</i>), strony: 184-194, ISBN 978-83- 946765-5-1. Link: https://www.wsia.opole.pl/wp-content/uploads/2020/09/2019_modern_technologies_in_education_dauczan.pdf.	
5. „<i>Modern Technologies in Economy and Management</i>”, Collective Scientific Monograph, monografia zbiorowa recenzowana, Opole: The Academy of Management and Administration in Opole, 2019, 493 s., artykuł: „<i>Mortality modeling</i>” (<i>Modelowanie umieralności</i>), strony: 148-160, ISBN 978-83-946765-4-4. Link: https://www.wsia.opole.pl/wp-content/uploads/2020/09/2019_moderntechnologies_in_economy_and_management.pdf.	
6. „<i>Contemporary issues of sustainable development</i>”, monografia zbiorowa recenzowana, Opole: The Academy of Management and Administration in Opole, 2019, 452 s., artykuł: „<i>Technology of optimization solutions in decision-making task</i>” (<i>Techniki rozwiązań optymalizacyjnych zadania decyzyjnego</i>), strony: 63-82, ISBN 978 - 83 - 946765 - 7 - 5. Link: https://www.wsia.opole.pl/wp-content/uploads/2020/09/2019_wspolczesne_zagadnienia_zrownowazonego_rozwoju.pdf.	

7. „Наука III тисячоліття: пошуки, проблеми”, перспективи розвитку, 2019, Збірник тез, БЕРДЯНСЬКИЙ ДЕРЖАВНИЙ ПЕДАГОГІЧНИЙ УНІВЕРСИТЕТ РАДА МОЛОДИХ УЧЕНИХ, 291 s., artykuł „Conversion of Audio File to Text File” (<i>Konwersja pliku audio na tekstowy</i>), strony: 199-200, УДК 378:001. (063) Н 34.	
Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2019_%D0%9D%D0%90%D0%A3%D0%9A%D0%9E%D0%92%D0%98%D0%99_%D0%92%D0%86%D0%A1%D0%9D%D0%98%D0%A.pdf.	
8. „Науковий вісник Сіверщини. Серія: Освіта”, Соціальні та поведінкові науки: науковий журнал / Академія Державної пенітенціарної служби”. Чернігів: Академія ДПТС, 2019. № 2 (3). 208 с. artykuł „Przyszłość → magistrala drogowa Hamburg-Szanghaj”, strony: 178-207, ISSN 2663-9009 (Print), ISSN 2663-9017 (Online).	
Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2019_%D0%9D%D0%90%D0%A3%D0%9A%D0%9E%D0%92%D0%98%D0%99_%D0%92%D0%86%D0%A1%D0%9D%D0%98%D0%A.pdf.	
9. „Współpraca specjalizowanych systemów informatycznych Implementacja i integracja wybranych modułów”, monografia indywidualna, wydawca: GlobeEdit (International Book Market Service Ltd., Member of OmniScriptum Publishing Group), Republic of Moldova, druk: Printforce - United Kingdom, 2019, 52 s., ISBN 978-613-42041-4. Link: https://www.morebooks.de/store/gb/book/wsp%C3%B3w-praca-specjalizowanych-system%C3%B3w-informatycznych/isbn/978-613-9-42041-4.	

2020 1. „<i>Narzędzia wdrażania systemów informatycznych</i>”, monografia indywidualna, wydawca: GlobeEdit (International Book Market Service Ltd. , Member of OmniScriptum Publishing Group), Republic of Moldova, druk: Printforce - United Kingdom, 2020, 140 s., ISBN 978-620-0-61025-6. Link: https://www.morebooks.de/gb/p_978-620-0-61025-6; https://www.globeedit.com; https://www.morebooks.de/store/gb/book/narz%C4%99dzia-wdra%C5%BCania-system%C3%B3w-informatycznych/isbn/978-620-0-61025-6.	
2. „<i>Doskonalenie systemów informacyjnych</i>” Część I „<i>Wybrane aspekty zarządzania</i>”, monografia indywidualna, wydawca: GlobeEdit (International Book Market Service Ltd., Member of OmniScriptum Publishing Group), Republic of Moldova, druk: Printforce - United Kingdom, 2020, 388 s., ISBN 978-620-0-59233-0. Link: https://www.morebooks.de/store/gb/book/doskonalenie-system%C3%B3w-informacyjnych-cz%C4%99%C5%9B%C4%87-i/isbn/978-620-0-59233-0.	
3. „<i>Doskonalenie systemów informacyjnych</i>” Część II „<i>Propozycja metod i modeli</i>”, monografia indywidualna, wydawca: GlobeEdit (International Book Market Service Ltd., Member of OmniScriptum Publishing Group), Republic of Moldova, druk: Printforce - United Kingdom, 2020, 452 s., ISBN 978-620-0-59542-3. Link: https://www.morebooks.de/store/gb/book/doskonalenie-system%C3%B3w-informacyjnych-cz%C4%99%C5%9B%C4%87-ii/isbn/978-620-0-59542-3.	

4. „*Osvita i suspilstvo V*: Міжнародний збірник наукових праць”, Бердянський державний педагогічний університет, Ополе: видавництво Вищої школи управління і адміністрації в Ополе, 422 с., 2020, artykuły: „*MOBILE APPLICATIONS IN LOGISTICS*” („*Aplikacje mobilne w logistyce*”) (strony: 277-287), „*METADATA EDITING PROGRAMS*”, „*Programy edycji metadanych*” (strony: 288-299), ISBN 978-83-66567-00-9.



Link: https://www.wszia.opole.pl/wp-content/uploads/2020/09/2020_%D0%9E%D0%A1%D0%92%D0%86%D0%A2%D0%90_%D0%86_%D0%A1%D0%A3%D0%A1%D0%9F%D0%86%D0%9B%D0%AC%D0%A1%D0%A2%D0%92%D0%9E_V.pdf

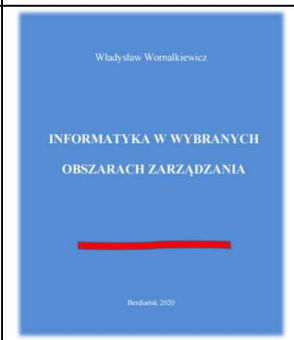
5. „*Implementacja informatyczna wybranych metod ilościowych*”, monografia indywidualna recenzowana, Bratysława: Wydawnictwo DENAKYR, s. r. o., 2020, 500 s., ISBN 978-80-973568-0-4.

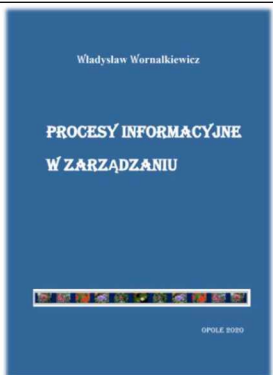
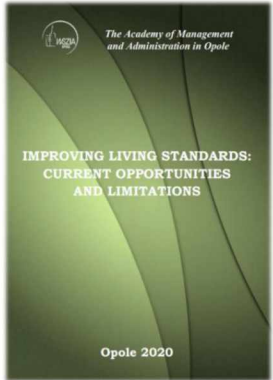
Link: <https://womalkiewicz.files.wordpress.com/2020/02/implementacja-informatyczna-wybranych-metod-ilosc59bcionych.pdf>.

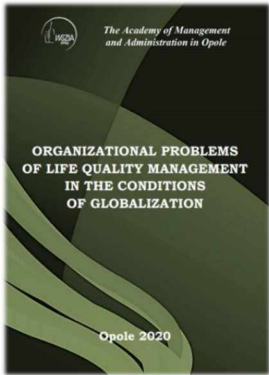
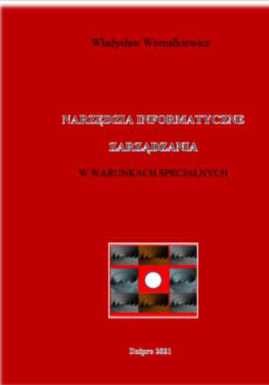
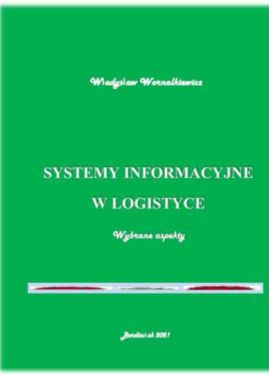


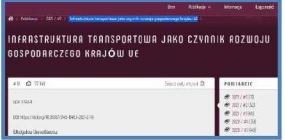
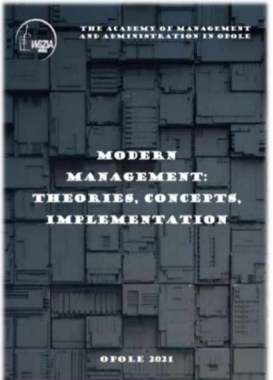
6. „*Informatyka w wybranych obszarach zarządzania*”, monografia indywidualna recenzowana, Berdiańsk: Wydawca „Svidler A.L.”, 2020, 450 s., ISBN 978-617-627-145-1.

Link: <https://womalkiewicz.files.wordpress.com/2020/10/informatyka-w-wybranych-obszarach-zarzadzania.pdf>.

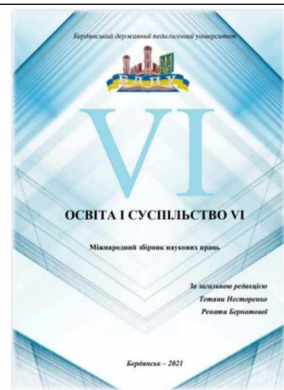


7. „<i>Procesy informacyjne w zarządzaniu</i>”, monografia indywidualna recenzowana, Wydawnictwo - Wyższa Szkoła Zarządzania i Administracji w Opolu 2020, 373 s., ISBN 978-83-665-22-1. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Procesy_informacyjne_w_zarzadzaniu.pdf.	
8. „<i>Journal of Modern Economic Research</i>”, Bratysława, monografia zbiorowa recenzowana, artykuł „<i>Innovative logistic solutions</i>” (<i>Innowacyjne rozwiązania logistyczne</i>), strony: 53-63, współautorstwo: Maksym Kutsenko, 2020, ISSN 2644-4380 nadruk; 2644-6332 online. Link: https://denakypublishing.science/index.php/jmer/article/view/40.	
9. „<i>Improving living standards: current opportunities and limitations</i>”, monografia zbiorowa recenzowana, 594 s., Opole: The Academy of Management and Administration in Opole, 2020, konferencja, artykuł „<i>POS systems</i>” (<i>Systemy klasy POS</i>), s. 139-167, ISBN 978 - 83 - 66567 - 21 - 4. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Zbirnyk_Osvita-i-suspilstvo-VI_new.pdf.	

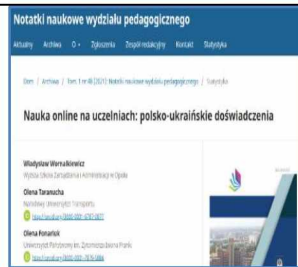
10. „Organizational problems of life quality management in the conditions of globalization”, Opole: The Academy of Management and Administration in Opole, 428 s., 2020, artykuł „Existing and intended logistic projects” (<i>Istniejące i zamierzone przedsięwzięcia logistyczne</i>), strony: 309-334, ISBN 978 - 83 - 66567 - 20 - 7. Link: https://www.wsia.opole.pl/wp-content/uploads/2020/05/2_2021.pdf.	
2021 1. „Narzędzia informatyczne zarządzania w warunkach specjalnych”, monografia indywidualna recenzowana, Dnipro: Wydawca „Svidler A.L.”, 2021, 288 s., ISBN 978-617-627-168-0. Link: https://dspace.bdpu.org/handle/123456789/4082	
2. „Systemy informacyjne w logistyce Wybrane aspekty”, monografia indywidualna recenzowana, Dnipro: Wydawca „Svidler A/L.”, 2021, 376 s., ISBN 978-617-627-157-4. Link: https://www.wsia.opole.pl/wp-content/uploads/2020/05/Systemy_informacyjne_w_logistyce_Wornalkiewicz.pdf.	

3. „Економічні науки”, monografia zbiorowa recenzowana, Вісник Хмельницького національного університету, 2021, № 5, Том I, artykuł: <i>THE OVERSIZED CARGO FORWARDING: ECONOMIC AND LEGAL ASPECTS</i>, UDC 330, strony: 175-180, ISSN 2307-5740, współautorstwo: YURII KRAVCHYK, ALLA KAPLUNOVSKA. Link: vknk-es-2021-n-5-298-175-180[1].pdf (khnu.km.ua). 3. „Економічні науки”, monografia zbiorowa recenzowana, Вісник Хмельницького національного університету, 2021, № 5, Том I, artykuł: <i>THE OVERSIZED CARGO FORWARDING: ECONOMIC AND LEGAL ASPECTS</i>, UDC 330, strony: 175-180, ISSN 2307-5740, współautorstwo: YURII KRAVCHYK, ALLA KAPLUNOVSKA. Link: vknk-es-2021-n-5-298-175-180[1].pdf (khnu.km.ua).	
4. „Ukrainian Journal of Applied Economics”, 2021, Volume 6, Nr 2, artykuł „Transport infrastructure as a factor of the EU countries' economic development” („Infrastruktura transportowa jako czynnik rozwoju gospodarczego krajów UE”), strony: 137-146, współautorstwo: Alla Kaplunovska, Olena Padchenko, ISSN 2415-8453. Link: https://doi.org/10.36887/2415-8453-2021-2-18.	
5. Huzhou University's Multicultural Center (Wielokulturowe Centrum Huzhou - Chiny), UDC 656.021.2, artykuł „The perspective of increasing of road capacity”, („Perspektywa rozwoju dróg szosowych”), współautorstwo: Ievgen Medvediev, Seriy Soroka, 2021, strony: 12-24. Link: https://wornalkiewicz.files.wordpress.com/2021/08/universityet-chiny.pdf; http://kwh.zjhu.edu.cn. 2021 年 第 3 期	
6. „SCIENTIFIC NOTES OF THE PEDAGOGICAL DEPARTMENT”, artykuł „ONLINE LEARNING AT UNIVERSITIES: POLISH-UKRAINIAN EXPERIENCE” „Nauczenie online na uniwersytetach: polsko-ukraińskie doświadczenia”, strony: 123-132, УДК 378.147.31, 2021, współautorstwo: Olena Taranukha, Olena Fonariuk. Link: 17560-Текст статті-34489-1-10-20210721 (2).pdf; https://periodicals.karazin.ua.	
7. „Modern management: theories, concepts, implementation”, monografia zbiorowa recenzowana, Opole: The Academy of Management and Administration in Opole, 2021, 430 s., artykuł „UNBLOCKING THE "ODRA - DANUBE" WATERWAY” („Udrożnienie drogi wodnej Odra - Dunaj), strony: 121-141, ISBN 978-83-66567-24-2. Link: https://www.wsia.opole.pl/wp-content/uploads/2020/05/7_2021.pdf.	

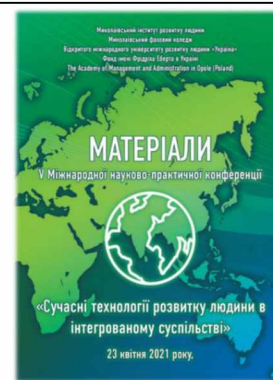
8. „ОСВІТА І СУСПІЛЬСТВО VI” Міжнародний збірник наукових праць, monografia zbiorowa recenzowana, Biediańsk: Wydawnictwo - Wyższa Szkoła Zarządzania i Administracji w Opolu, 2021, artykuł „Opportunities to make milk reception logistics more modern” („Możliwości unowocześnienie logistyki odbioru mleka”), strony: 328-343, ISBN 978-83-66567-26-9. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Zbirnyk_Osvita-i-suspilstvo-VI_new.pdf.

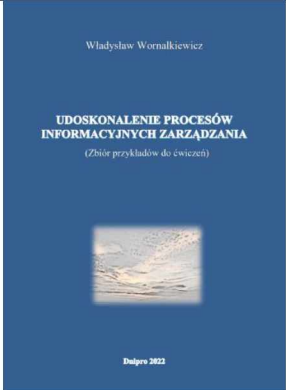
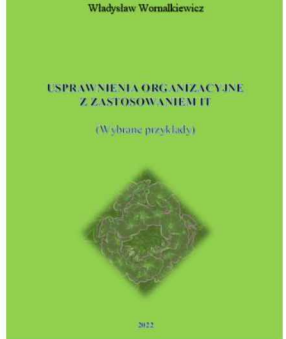


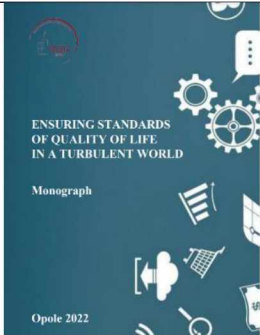
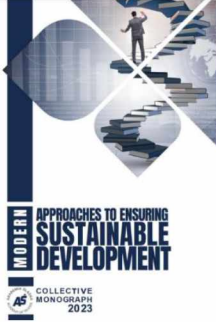
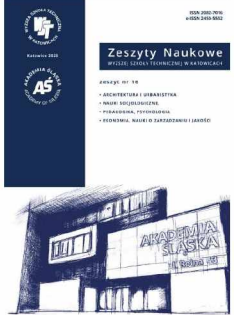
9. „SCIENTIFIC NOTES OF THE PEDAGOGICAL DEPARTMENT”, artykuł „ONLINE LEARNING AT UNIVERSITIES: POLISH-UKRAINIAN EXPERIENCE” „Nauczanie online na uniwersytetach: polsko-ukraińskie doświadczenia”, strony: 123-132, j. ang., УДК 378.147.31, 2021, współautorstwo: Olena Taranukha, Olena Fonariuk. Link: 17560-Текст статті-34489-1-10-20210721 (2).pdf, <https://periodicals.karazin.ua>; <https://periodicals.karazin.ua/pedagogy/article/view/17560>.

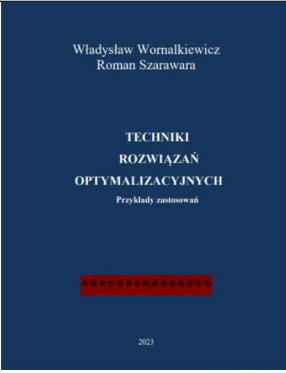


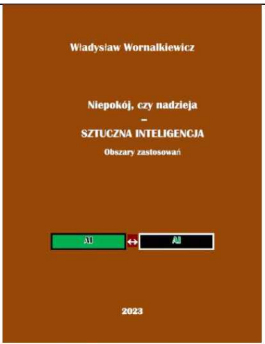
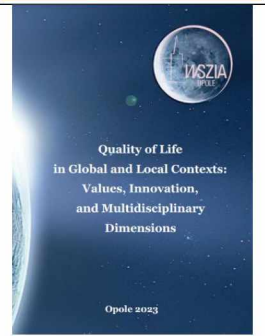
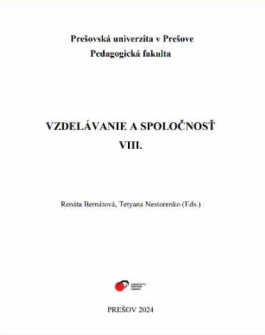
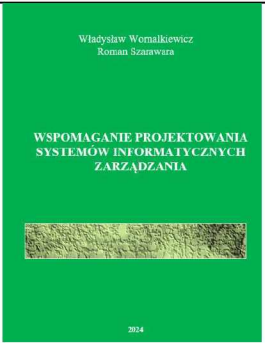
10. „СУЧАСНІ ТЕХНОЛОГІЇ РОЗВИТКУ ЛЮДИНИ В ІНТЕГРОВАНОМУ СУСПІЛЬСТВІ” Матеріали V Міжнародної науково-практичної конференції (Materiały V Międzynarodowej naukowo-praktycznej konferencji), 2021, Миколаївський інститут розвитку людини (Instytut w Mikolaiv - Ukraina), monografia zbiorowa recenzowana pokonferencyjna, 308 s., artykuł „Socio-logistical aspects of the Vistula spit dug-through”, („Społeczno-logistyczne aspekty przekopu Mierzei Wiślanej”), strony: 302-304, удк 371: 378. Link: https://www.wszia.opole.pl/wp-content/uploads/2020/05/Spoleczno_logistyczne_aspekty_przekopu.pdf.

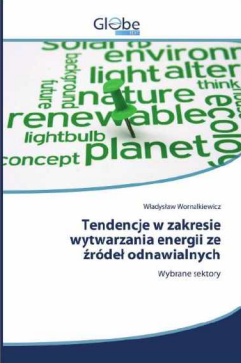


2022 1. „<i>Udoskonalenie procesów informacyjnych zarządzania (Zbiór przykładów do ćwiczeń)</i>”, monografia indywidualna recenzowana, Dnipro: Wydawca „Svidler A.L.”, 241 s., ISBN 978-617-627-170-3.	
2. „<i>Usprawnienia organizacyjne z zastosowaniem IT (Wybrane przykłady)</i>”, monografia indywidualna recenzowana, Kijów: Wydawca „Majster Knyg”, s. 340, ISBN 978-617-7652-59-4. Link do strony internetowej Narodowego Uniwersytetu Gospodarki Miejskiej im. O.M. Beketova w Charkowie.: http://eprints.kname.edu.ua/61802/1/Monografia_7.pdf	
3. „<i>Zasilanie alternatywne pojazdów samochodowych</i>”, Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach, 2022, nr 15, s. 30, ISSN 2082-7016, eISSN 2450-5552.	
4. Artykuły w ramach monografii EDUKACJA I SPOŁECZEŃSTWO VII, zbiór artykułów naukowych, Opole 2022, strony: 278-320, ISBN 978-83-66567-41-2: - o „<i>Modelowanie biznesowe z zastosowaniem UML</i>”, stron: 16, - o „<i>Potrzeba scalania systemów klasy ERP</i>”, stron: 12, - o „<i>Wspomaganie dystrybucji systemem Dynamics NAV</i>”, s.: 15. Wydanie przez: ANS Opole, Wydział Pedagogiczny Uniwersytetu w Preszowie (Słowacja), Państwowy Uniwersytet Pedagogiczny w Bierańsku (Ukraina).	

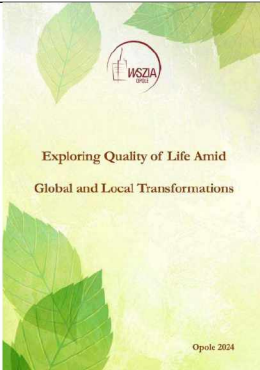
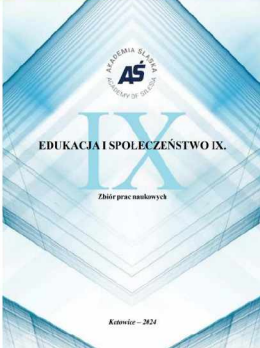
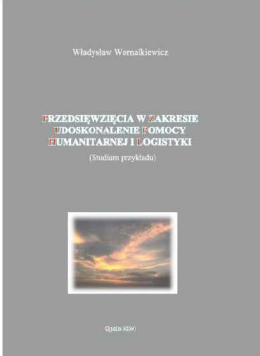
5. „<i>Projects in the field of 5G network construction</i>” (<i>Przedsięwzięcia w zakresie budowy sieci 5G</i>), s. 277-293, <i>Ensuring Standards of Quality of Life in a Turbulent World</i>. Monograph. The Academy of Applied Sciences – Academy of Management and Administration in Opole.	
2023	
1. Artykuły w ramach monografii zbiorowej „<i>Modern Approaches to sustainable development</i> „AŚ Katowice 2023: ○ <i>Flood Prevention (Przeciwdziałanie powodziom wiosennym)</i>, 1.12. s. 95-105, referat na konferencji. ○ <i>Organization of a rescue action in a situation earthquake (Organizacja akcji ratowniczej w sytuacji trzęsienia ziemi)</i>, 2.3. s. 285-294. Monografia pokonferencyjna 4th International Scientific Conference „<i>Role of Science and Education in Sustainable Development</i>” (<i>Znaczenie nauki i edukacji w zrównoważonym rozwoju</i>), konferencja międzynarodowa internetowa.	
2. Artykuł „<i>Procedure implementation of logistics services</i>” (<i>Implementacja procedury optymalizacji usług logistycznych</i>), <i>Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach</i>, 2023, nr 16, s. 157-170. ISSN: 2082-7016; e-ISSN: 2450-5512; DOI: 10.54264/0070.	
3. Artykuł „<i>Cyber security</i>” (<i>Cybersecurity</i>), abstract, s. 5, VII International Scientific and Practice Conference „<i>Innovative potential and legal support of social and economic development of Ukraine: the challenge of the global world</i>”, konferencja internetowa, referat na platformie ZOOM, Połtawski Instytut Ekonomii i Prawa Uniwersytetu „Ukraina”, s. 16-21, Poltava, may, 17-18, 2023, УДК 330.111.66(477):34.	

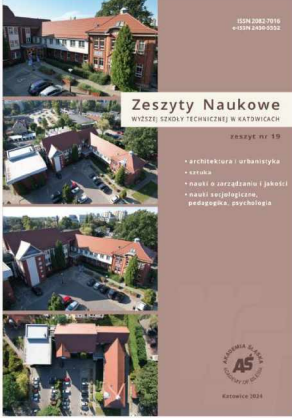
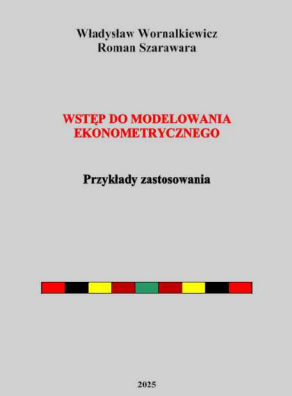
4. Artykuł „<i>Iteracyjne korzystanie ze sztucznej inteligencji</i>” - <i>Iterative use of artificial intelligence</i>, s. 12-15. Збірник матеріалів V Всеукраїнської науково-практичної конференції, м. Полтава, Полтавський інститут економіки і права, 5-6 грудня 2023 року 12 Полтавський інститут економіки і права, РОЗДІЛ 1 ПРАВОВЕ РЕГУЛЮВАННЯ СУСПІЛЬНИХ ВІДНОСИН В УКРАЇНІ В УМОВАХ ВОЄННОГО СТАНУ; prezentacja w Power Point ISBN 978-617-8368-11-1 УДК 330.111.66(477):34	
5. Презентація нового видання – підручника, виданого в межах українопольської співпраці: Wornalkiewicz W., Szarawara R. „<i>Techniki rozwiązań optymalizacyjnych. Przykłady zastosowań</i>”, Połtawa: Połtawski Instytut Ekonomii i Prawa Uniwersytetu „Ukraina”, 2023, s. 243.	
6. Wornalkiewicz W., Szarawara R., podręcznik „<i>Techniki rozwiązań optymalizacyjnych Przykłady zastosowań</i>”, Połtawski Instytut Ekonomii i Prawa Uniwersytetu „Ukraina”, s. 244, UDK 338.24:001.82(075), ISBN 978-966-388-674-9. DOI: https://doi.org/10.36994/978-966-388-674-9-2023-243	
7. Wornalkiewicz W., Szarawara R., podręcznik „<i>Techniki rozwiązań optymalizacyjnych Przykłady zastosowań</i>”, Wydawnictwo Naukowe PWN S.A., Warszawa 2024, s. 246, ISBN 978-83-01-23503-1, DOI: https://doi.org/10.53271/2024.020.	

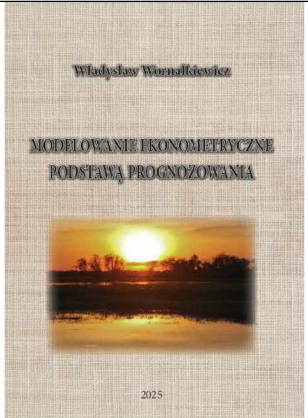
8. Wornalkiewicz W., monografia indywidualna „<i>Niepokój, czy nadzieja – SZTUCZNA INTELIGENCJA. Obszary zastosowań</i>”, s. 165., ISBN 978-617-627-171-0, Państwowy Uniwersyte Pedagogiczny w Berdiańsku, Wydawca „Svidler A.L.”, Dnipro 2023.	
9. Artykuł „<i>Artificial intelligence applications</i>” (<i>Aplikacje sztucznej inteligencji</i>) - w monografii pokonferencyjnej, s. 12, <i>Quality of Life in Global and Local Contexts: Values, Innovation, and Multidisciplinary Dimensions</i> (<i>Jakość życia w kontekście globalnym i lokalnym: wartości, innowacje i wymiary multidyscyplinarne</i>), The Academy of Applied Sciences – Academy of Management and Administration in Opole, 2023, ISBN 978-83-66567-53-5.	
2024	
1. Artykuł "<i>Metody ochrony przed cyberbezpieczeństwem</i>" s. 126-134, w Zeszytach Naukowych <i>Vzdelávane a spoločnosť</i> VIII, Prešovská univerzita v Prešove Pedagogická fakulta, PREŠOV 2024, ISBN 978-80-555-3260-8, EAN 9788055532608.	
2. Wornalkiewicz W., Szarawara R., „<i>Wspomaganie projektowania systemów informatycznych zarządzania</i>”, podręcznik, s. 408, Połtawski Instytut Ekonomii i Prawa. Uniwersytet „Ukraina”, ISBN 978-966-388-689-3, UDK 005.8:004.415(075.8),s.	

3. Artykuł „<i>Niewykorzystane możliwości energii wodnej</i>”, s. 213-217, referat, konferencja 19.04.2023, Kijów, National University of Live and Environmental Sciences of Ukraine, Katedra Technologii i Środków Transportu Kompleksu Rolno-Przemysłowego, monografia tylko w wersji elektronicznej. ISBN978-617-8368-11-1, УДК 631.17+62-52-631.3.	
4. Wornalkiewicz W., „<i>Tendencje w zakresie wytwarzania energii ze źródeł odnawialnych. Wybrane sektory</i>”, Globe Edit, ISBN 978-620-6-79987-0, s. 132.	
5. Artykuły: Konferencja „Digital Economy and Digital Society 6th International Scientific Conference, April 9-10, 2024, Akademia Śląska w Katowicach, Monografia „Exploring the Digital Landscape: Interdisciplinary Perspectives”, Edited by Olha Blaha and Iryna Ostopolets, The University of Technology in Katowice Press 2024, ISBN 978-83-969890-9-3, DOI: 10.54264/M036: - „Breaking social anxiety – green light for nuclear power plants” - „Przelamanie niepokoju społecznego – zielone światło dla elektrowni atomowych”, s. 457-477. - „Carbon-Nuclear Transformation” - „Transformacja węglowo-jądrowa”, s. 478-891, referat. - „Trend of application of AI in search engines” - „Tendencje zastosowania AI w wyszukiwarkach”, s. 892-906.	

[illegible]

9. Artykuł „<i>Współpraca mieszkańców wsi w ustalaniu priorytetu remontu dróg lokalnych</i>”, s.: 447-459, monografia po konferencji międzynarodowej z udziałem przedstawicieli Ukrainy pt. „Exploring Quality of Life Amid Global and Local Transformations”, 9th International Scientific and Practical – Conference „<i>Modern problems of Improve living standards in a globalized world</i>”, november 13-14, 2024. Prezentacja artykułu w Power Point na Konferencji.	
10. Artykuł „<i>IT technology in a small tailoring shop</i>” - „<i>Technika IT w małym zakładzie krawieckim</i>”, w ramach monografii EDUKACJA I SPOŁECZEŃSTWO IX, Zbiór prac naukowych, Katowice 2024, strony: 221-236, ISBN 978-83-972085-8-2, DOI 10.54264/M041.	
11. Podręcznik "<i>Przedsięwzięcia w zakresie udoskonalenia pomocy humanitarnej i logistyki (Studium przykładu)</i>", Uniwersytet Pedagogiczny w Berdiańsku (siedziba w Zaporozie), ISBN 978-617-627-189-5, Wydawca „Svidler A.A.”, Dnipro 2024, 478 s.	

12. Artykuł "Dobór pakietu integracji modułów wsparcia rolnictwa", Połtawski Instytut Ekonomii i Prawa, Uniwersytet „Ukraina”, wygłoszony referat w dniu 12.12.2024, monografia pokonferencyjna wydana w wersji elektronicznej. УДК 330.111.66(477):34	 ПРАВОВІ, ЕКОНОМІЧНІ ТА СОЦІОКУЛЬТУРНІ ЗАСАДИ РЕГУЛЮВАННЯ СУСПІЛЬНИХ ВІДНОСИН: СЛУЖБИ РОЛІВ ТА ВИСІВНИХ ЧАСІВ призначено дозакладні теми для обговорення економічних і соціальних аспектів розвитку аграрного сектору ЗБІРНИК МАТЕРІАЛІВ VI Всеукраїнської науково-практичної конференції з міжнародною участю 12-14 грудня 2024 р. м. Полтава Полтава – 2025
13. Artykuł „Usprawnienie obiegu dokumentów w administracji publicznej”, Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach”, półrocznik 2024, ISSN 2082-7016; eISSN 2450-5552 2024, nr 19, DOI: 10.54264/0102, 20 pkt., wersja elektroniczna. http://www.wydawnictwo.wst.pl/oferta_wydawnicza_oraz_zakup_publicacji/zeszyty_naukowe/zeszyty_naukowe_wyszej_szkoy_techicznej_w_katowicach_nr_19/141	 Zeszyty Naukowe WYŻSZEJ SZKOŁY TECHNICZNEJ W KATOWICACH zeszyt nr 19 • architektura i urbanistyka • ekonomia • nauki o zarządzaniu i jakości • nauki społeczne pedagogika, psychologia Kwartalnik 2024
2025	
1. Podręcznik: Wornalkiewicz W., Szarawara R., <i>Wstęp do modelowania ekonometrycznego. Przykłady zastosowania</i>, Instytut Ekonomii i prawa w Połtawie, s. 265, ISBN 978-966-388-702-9.	 Władysław Wornalkiewicz Roman Szarawara WSTĘP DO MODELOWANIA EKONOMETRYCZNEGO Przykłady zastosowania 2025

2. Artykuł <i>Możliwości usprawnienia procedury orzecznictwa z zastosowaniem IT</i>, materiały pokonferencyjne, Międzynarodowa naukowo-praktyczna konferencja, <i>Potencjał innowacyjny i wsparcie prawne dla rozwoju społeczno-gospodarczego Ukrainy: wyzwania globalnego świata</i>, Połtawa 20-21 maja 2025	 Національний університет з права та кримінології Полтавський інститут економіки і права  ІННОВАЦІЙНИЙ ПОТЕНЦІАЛ ТА ПРАВОВЕ ЗАБЕЗПЕЧЕННЯ СОЦІАЛЬНО-ЕКОНОМІЧНОГО РОЗВИТКУ УКРАЇНИ: ВИКЛИК ГЛОБАЛЬНОГО СВІТУ МАТЕРІАЛИ IX Міжнародної науково-практичної конференції 20-21 травня 2025 р.
3. Podręcznik „Modelowanie ekonometryczne podstawą prognozowania”, Wydział Edukacji Fizyczno-Matematycznej, Informatycznej i Technologicznej Państwowego Uniwersytetu Pedagogicznego w Berdiańsku (Zaporoże), Ukraina 2025, ISBN 978-617-555-283-4.	
4. Artykuł <i>ITS w ruchu drogowym miejskim</i>, Zeszyty Naukowe Akademii Śląskiej w Katowicach, wysłany do druku.	
5. Artykuł <i>Systemy ERP z elementami nowoczesności</i>, Połtawski Instytut Ekonomii i Prawa, wysłany do druku w monografii pokonferencyjnej.	
6. Artykuł <i>Monitorowanie spraw obywatelskich</i>, ANS-WSZiA Opole, wysłany do druku w monografii pokonferencyjnej.	
7. Monografia <i>Inspiracje do kreatywności prac dyplomowych</i>, Część A. Artykuły w latach 2008-2016, GlobeEdit, w druku.	

- B**ecker A., *Analiza efektywności działania województw Polski*, Internet, Katedra Zastosowań Matematyki, Akademia Rolnicza w Szczecinie.
- Borowiecki R., Kwieciński M., *Informacja i Wiedza w Zintegrowanym Systemie Zarządzania*, Redakcja Naukowa, Zakamycze 2004.
- Bryla B., Loney K., *Oracle Database 11g, Podręcznik administratora baz danych*, Helion 2010.
- C**oelli T., *Przewodnik po DEAP wersji 2.1, Analiza zwijania [enevelopment] danych*, Centrum Analizy Efektywności i Produktyności, [www.une.edu.au/econometrics/cepa.htm].
- Coelli T.J., *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program*, CEPA Working Papers, Department of Econometrics University of New England, Australia.
- D**uczmal M., *Efektywność działania obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013.
- Duczmal M., *Metody optymalizacji w zarządzaniu*, Wydawnictwo Instytut Śląski, Opole 2014.
- Duczmal M., Wornalkiewicz W., *Wstęp do ekonometrii i badań operacyjnych Zbiór przykładów z zastosowaniem mikrokomputera*, Wydawnictwo Instytut Śląski w Opolu, Opole 2010.
- Dziechciarz J., *Ekonometria Metody, przykłady, zadania*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2003.
- E*ncyklopedia organizacji i zarządzania, PWE, Warszawa 1982.
- Encyklopedia WIEM, Internet
- G**uzik B., *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, 2009.
- I**nstrukcja użytkownika SEG.
- Instrukcja użytkownika WN.
- J**ashapara A., tłum. Sawicki J., *Zarządzanie wiedzą*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2006.
- K**isielnicki J., Sroka H., *Systemy informacyjne biznesu*, Placet, Warszawa 2005.
- Kufel T., *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*, Wydanie 3, WN, PWN, Warszawa 2011.
- Kukuła K., *Badania operacyjne w przykładach i zadaniach*, Wydawnictwo Naukowe PWN, Warszawa 2004.
- L*eksykon Naukowo-techniczny, Wydawnictwo Naukowo-Techniczne, Warszawa 1972.
- M**iszczak W., *Statystyczne metody analizy danych Materiały do ćwiczeń*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, 1999.
- O**sińska M. (red. nauk.), *Ekonometria współczesna*. Towarzystwo Naukowe Organizacji i Kierownictwa. Toruń 2007.
- Ostasiewicz W. (red), *Statystyczne metody analizy danych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, 1999.
- P**omianek T., Skica T., *Efektywność sektora publicznego w Polsce. Próba oceny*, Internet.

Przejawy wielowymiarowości współczesnego zarządzania - formy i instrumenty ekonomiczno-społeczne, red. nauk. T. Pokusa, artykuł: W. Wornalkiewicz *Model ekonometryczny zmiennej „liczba bezrobotnych”*, Wydawnictwo Instytut Śląski, Opole 2011.

Rocznik Statystyczny Rzeczypospolitej Polskiej 2007.

Słownik języka polskiego, PWN, Warszawa 1988.

Strojny M., *Teoria i praktyka zarządzania wiedzą, Ekonomika i Organizacja Przedsiębiorstwa*, 2000, nr 10.

Szymczak A., *Próba oceny makroekonomicznej efektywności inwestycji w wybranych gospodarkach UE*, rozdz. 8, Internet.

Urman S., Harmand R., McLaughlin M., *Oracle database 10g, Programowanie w języku PL/SQL*, Helion 2008.

Ustawa z dnia 10.02.2017 r. *O Krajowym Ośrodku Wsparcia Rolnictwa*, Dz. U. 2023 poz. 1308.

Ustawa z dnia 19.10.1991 r. *o gospodarowaniu nieruchomościami rolnymi Skarbu Państwa*, Dz. U. 2023, poz. 1933.

Walesiak M., Gatnar E. (red. nauk.), *Statystyczna analiza danych z wykorzystaniem programu R*, Wydawnictwo Naukowe PWN, Warszawa 2012.

Wielicki W., *Efektywność przedsiębiorstw przemysłowych i rolnych*, Internet.

Wilczek T., *Analiza możliwości integracji systemów w zakresie nieruchomości rolnych*, praca magisterska, Akademia Nauk Stosowanych Wyższej Szkoły Zarządzania i Administracji w Opolu, 2024.

Wornalkiewicz W., *Etapy i procedury budowy modelu produktu brutto podregionów*, artykuł, Wydawnictwo WSZiA Opole, 2009.

Wornalkiewicz W., *Zastosowanie programu Excel do sformułowania modelu produktu brutto województw*, artykuł, WSZiA Opole, 2009.

Wornalkiewicz W., *Formułowanie modeli ekonometrycznych na potrzeby zarządzania*, Wydawnictwo Instytut Śląski, Opole 2015.

Wornalkiewicz W., *Miary kształtujące efektywność produktu krajowego brutto*, artykuł, WSZiA Opole.

Wornalkiewicz W., *Modele ekonometryczne PKB obiektów struktury terytorialnej*, Wydawnictwo Instytut Śląski, Opole 2013.

Wornalkiewicz W., *Modele i efektywność PKB*, artykuł, WSZiA Opole.

Wornalkiewicz W., *Procesy informacyjne w zarządzaniu*, Wydawnictwo WSZiA, Opole 2020.

Wornalkiewicz W., *Prognozowanie migracji ludności z uwzględnieniem wag harmonicznych*, monografia zbiorowa „*Uwarunkowania demograficzne rozwoju społecznego i ekonomicznego miasta Nysa i powiatu nyskiego*”, Opole-Nysa 2018, Wyższa Szkoła Zarządzania i Administracji w Opolu.

Wornalkiewicz W., *Rozwiązywanie modelu CCR programem WinQSB*, artykuł, WSZiA Opole.

Wornalkiewicz W., *Wdrożenie zintegrowanego systemu informatycznego wspomagającego zarządzanie*, WSZiA Opole, 2015.

Wornalkiewicz W., *Wstęp do ekonometrii i badań operacyjnych Zbiór przykładów z zastosowaniem mikrokomputera*, wydanie II rozszerzone, *Wybrane modele ekonometryczne*, Wydawnictwo Instytut Śląski, Opole 2012.

Wornalkiewicz W., *Wyznaczanie efektywności technicznej obiektów programem DEAP*, artykuł, WSZiA Opole.

Zastosowanie taksonomii wrocławskiej do oceny poziomu rozwoju społeczno-gospodarczego karpaczkich gmin górskich ONW, Internet, *Metody ilościowe w badaniach ekonomicznych*, tom XVII, 2016.

Wydanie Edukacyjne

**Wornalkiewicz Władysław
Szarawara Roman**

**PRÓBA MODELOWANIA EKONOMETRYCZNEGO
OBIEKTÓW STRUKTURY TERYTORIALNEJ**

Format 60x84/16
Konwencjonalnie zadrukowane arkusze 10,1
Wydanie 100 kopie
Zamówienie 99

ОП "ШвидкоДрук"
36003, м. Полтава, вул. Чорновола 2^б
Свідоцтво ФОП Іпполітова Т.С.
В02 №414529 від 27.11.2007 р.



Dr inż. prof. ANS-WSZiA w Opolu (Polska) Władysław Wornalkiewicz jest autorem książek z zakresu statystyki i ekonometrii z zastosowaniem programów komputerowych. Jego praca naukowa skupia się na testowaniu różnych metod modelowania ekonometrycznego z użyciem danych statystycznych oraz takich narzędzi programistycznych jak pakiety *Excel*, *Gretl*, *WinQSB*, *R*, *DEAP*, *Expert Choice* i innych. Jest absolwentem kilku kierunków na Politechnice Wrocławskiej, gdzie uzyskał tytuły inżyniera mechanika, magistra inżyniera organizatora produkcji, doktora nauk ekonomicznych, pedagoga Ministerstwa Edukacji Narodowej. Ukończył również program edukacyjny „Polska w procesie integracji europejskiej”.

Zatrudniony jest w Akademii Nauk Stosowanych (ANS-WSZiA) w Opolu na stanowisku Profesora Uczelni. Obecnie ma tam wykłady z przedmiotów: *Procesy informacyjne w zarządzaniu*, *Systemy informacyjne w logistyce*, *Technologie informacyjne*, *Informatyka w zarządzaniu*, *Badania operacyjne*, *Optymalizacja decyzji gospodarczych*. Ponadto prowadzi seminaria dyplomowe licencjackie i magisterskie. W dorobku naukowym ostatnich lat są następujące książki: *Wstęp do ekonometrii i badań operacyjnych*, tom I. *Wybrane modele ekonometryczne*, *Formułowanie modeli ekonometrycznych do potrzeb zarządzania* - dwa tomy (*Środowiska programowe statystyki opisowej*, *Zagadnienia ekonometrii*), *Wdrożenie zintegrowanego systemu informatycznego wspomagającego zarządzanie*, *Wprowadzenie do projektowania systemów informatycznych zarządzania* - dwie części (*Wybrane systemy zarządzania i sposoby modelowania*, *Narzędzia wspomagające projektowanie systemów*), *Elementy inżynierii i analizy systemów zarządzania Wybrane aspekty logistyczne* - rozdziały: 1-9, 16-22, *Współpraca specjalizowanych systemów informatycznych*.

W latach 2020-2021 ukazały się monografie indywidualne autora, a mianowicie:

- *Implementacja informatyczna wybranych metod ilościowych*, opublikowana przez wydawnictwo DENAKYR w Bratysławie;
- *Informatyka w wybranych obszarach zarządzania*, Państwowy Uniwersytet Pedagogiczny w Berdyansku (Ukraina);
- *Procesy informacyjne w zarządzaniu*, Wyższa Szkoła Zarządzania i Administracji w Opolu;
- *Systemy informacyjne w logistyce Wybrane aspekty*, Państwowy Uniwersytet Pedagogiczny w Berdyansku (Ukraina);
- *Narzędzia informatyczne zarządzania w warunkach specjalnych*, Państwowy Uniwersytet Pedagogiczny w Berdyansku (Ukraina);
- *Udoskonalenie procesów informacyjnych zarządzania (zbiór przykładów do ćwiczeń)*, Państwowy Uniwersytet Pedagogiczny w Berdyansku (Ukraina);
- *Usprawnienia organizacyjne z zastosowaniem IT (Wybrane przykłady)*, Państwowy Uniwersytet Pedagogiczny w Berdyansku (Ukraina);
- *Doskonalenie systemów informatycznych: część I. Wybrane aspekty zarządzania, część II. Propozycja metod i modeli*, GlobeEdit (Niemcy);
- *Narzędzia wdrażania systemów informatycznych*, GlobeEdit (Niemcy);
- *Współpraca specjalizowanych systemów informatycznych*, w GlobeEdit (Niemcy).

Efektom prac badawczych są trzy publikacje: skrypt - *Metoda badania przyczynowo-skutkowego związków między cechami statystycznymi*, książka - *Modele ekonometryczne PKB obiektów struktury terytorialnej*, książka *Popularyzacja wybranych metod ilościowych w Internecie*. Jest autorem wielu artykułów wydrukowanych w monografiach WSZiA w Opolu, opublikowanych przez Uniwersytet Pedagogiczny w Berdyansku (Ukraina), Uniwersytet Pedagogiczny w Presovie (Słowacja), Uniwersytet Technologiczny w Czerniowie (Ukraina) i w innych wydawnictwach.



Roman Sharavara – kandydat nauk ekonomicznych, profesor nadzwyczajny, profesor Katedry Prawa i Finansów. Od 2009 roku pracuje w Połtawskim Instytucie Ekonomii i Prawa Uniwersytetu „Ukraina”.

Zainteresowania naukowca obejmują problematykę funkcjonowania giełdy jako odrębnego segmentu systemu finansowego kraju. A także problematykę budowy modeli ekonomicznych i matematycznych służących optymalizacji decyzji zarządczych. Szczególną uwagę w badaniach poświęca się zagadnieniu kształtowania się nowoczesnego modelu interakcji inwestora prywatnego z emitentem papierów wartościowych.

Autor ma w swoim dorobku około 150 prac naukowo-dydaktyczno-metodologicznych dla szkolnictwa wyższego poświęconych problematyce kształtowania konkurencyjnego sektora finansowego gospodarki.

Obecnie wykłada następujące dyscypliny: *Metodologia i organizacja badań naukowych*, *Gospodarka globalna*, *Zarządzanie finansami*, *Współczesne teorie ekonomiczne*.

Od 2012 roku pełni funkcję pierwszego zastępcy dyrektora Połtawskiego Instytutu Ekonomii i Prawa Uniwersytetu „Ukraina”.