

Cena 15,00 zł
(VAT 8%)

Indeks 381306
e-ISSN 2543-8476
PL ISSN 0043-518X

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

KWIECIEŃ / APRIL
ROCZNIK / VOLUME 69

2024 | 4

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION



WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

KWIECIEŃ / APRIL
ROCZNIK / VOLUME 69

2024 | 4 (755)

RADA NAUKOWA / SCIENTIFIC COUNCIL

dr Dominik Rozkrut – przewodniczący/chairman (Uniwersytet Szczeciński, Polska), Prof. Anthony Arundel (Maastricht University, Holandia), Eric Bartelsman, PhD, Assoc. Prof. (Vrije Universiteit Amsterdam, Holandia), prof. dr hab. Czesław Domański (Uniwersytet Łódzki, Polska), prof. dr hab. Elżbieta Gołata (Uniwersytet Ekonomiczny w Poznaniu, Polska), Semen Matkovskyy, PhD, Assoc. Prof. (Ivan Franko National University of Lviv, Ukraina), prof. dr hab. Włodzimierz Okrasa (Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska), prof. dr hab. Józef Oleński (Polskie Towarzystwo Statystyczne, Polska), prof. dr hab. Tomasz Panek (Szkola Główna Handlowa w Warszawie, Polska), Juan Manuel Rodríguez Poo, PhD, Assoc. Prof. (University of Cantabria, Hiszpania), Iveta Stankovičová, BEng, PhD, Assoc. Prof. (Comenius University in Bratislava, Słowacja), prof. dr hab. Marek Walesiak (Uniwersytet Ekonomiczny we Wrocławiu, Polska), prof. dr hab. Józef Zegar (Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy, Polska)

sekretarz/secretary: Paulina Kucharska-Singh, Główny Urząd Statystyczny, Polska

KOLEGIUM REDAKCYJNE / EDITORIAL BOARD

Tudorel Andrei, PhD, Assoc. Prof. (Bucharest Academy of Economic Studies, Rumunia), mgr Renata Bielak (Główny Urząd Statystyczny, Polska), dr hab. Marek Cierpień-Wolan, prof. UR (Uniwersytet Rzeszowski, Polska), dr hab. Grażyna Dehnel, prof. UEP (Uniwersytet Ekonomiczny w Poznaniu, Polska), dr Jacek Kowalewski (Uniwersytet Ekonomiczny w Poznaniu, Polska), dr Jan Kubacki (Polskie Towarzystwo Statystyczne, Polska), dr Grażyna Marciniak (Główny Urząd Statystyczny, Polska), dr hab. Andrzej Młodak, prof. Akademii Kaliskiej (Akademia Kaliska im. Prezydenta Stanisława Wojciechowskiego, Polska), prof. dr hab. Mateusz Pipień (Uniwersytet Ekonomiczny w Krakowie, Polska), Marek Rojíček, BEng, PhD (University of Economics, Prague, Czechy), Anna Shostya, PhD, Assoc. Prof. (Pace University in New York, Stany Zjednoczone), dr hab. Małgorzata Tarczyńska-Łuniewska, prof. US (Uniwersytet Szczeciński, Polska), dr Wioletta Wrzascz (Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy, Polska), dr inż. Agnieszka Zgierska (Główny Urząd Statystyczny, Polska)

ZESPÓŁ REDAKCYJNY / EDITORIAL STAFF

redaktor naczelny / editor-in-chief: Marek Cierpień-Wolan

zastępca redaktora naczelnego / deputy editor-in-chief: Andrzej Młodak

redaktorzy tematyczni / thematic editors: Andrzej Młodak, Małgorzata Tarczyńska-Łuniewska, Wioletta Wrzascz, Agnieszka Zgierska

sekretarz/secretary: Małgorzata Zygmunt, Główny Urząd Statystyczny, Polska

ADRES REDAKCJI / EDITORIAL OFFICE ADDRESS

Główny Urząd Statystyczny / Statistics Poland, al. Niepodległości 208, 00-925 Warszawa
tel./phone +48 22 608 32 25, e-mail: redakcja.ws@stat.gov.pl

Redakcja językowa: Wydział Czasopism Naukowych, Główny Urząd Statystyczny
Language editing: Scientific Journals Division, Statistics Poland

Redakcja techniczna, skład i łamanie, opracowanie materiałów graficznych, korekta, druk i oprawa:
Zakład Wydawnictw Statystycznych – zespół pod kierunkiem Macieja Adamowicza
Technical editing, typesetting, preparation of graphic materials, proofreading, printing and binding:
Statistical Publishing Establishment – team supervised by Maciej Adamowicz



Zakład Wydawnictw
Statystycznych

Druk i oprawa / Printed and bound by:

Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment
al. Niepodległości 208, 00-925 Warszawa, zws.stat.gov.pl

Wersja elektroniczna, stanowiąca wersję pierwotną czasopisma, jest dostępna na ws.stat.gov.pl
The primary version of the journal, issued in electronic form, is available at ws.stat.gov.pl

© Copyright by Główny Urząd Statystyczny and the authors, some rights reserved. CC BY-SA 4.0 licence



Informacje w sprawie sprzedaży i prenumeraty czasopisma / Sales and subscription of the journal:
Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment
zws.stat.gov.pl
tel./phone +48 22 608 32 10, +48 22 608 38 10

SPIS TREŚCI
CONTENTS

Od redakcji	IV
From the editorial team	
Statystyka w praktyce	
Statistics in practice	
Elżbieta Gołata	
Różnice w pomiarze płodności kobiet zależnie od szacunku liczby ludności	1
Varying measurement of women's fertility depending on different estimates of a population	
Tomasz Banasik	
Ocena zależności między rozwojem branży leasingowej a wzrostem gospodarczym w Polsce	26
Assessment of the relationship between the development of the leasing industry and economic growth in Poland	
Studia interdyscyplinarne. Wyzwania badawcze	
Interdisciplinary studies. Research challenges	
Marcin Pełka, Antonio Irpino	
The application of multidimensional scaling of data to determining changes in retailer customers' preferences	41
Zastosowanie wielowymiarowego skalowania danych do określenia zmian preferencji klientów sklepów	
Dyskusje. Recenzje. Informacje	
Discussions. Reviews. Information	
Dariusz Smoliński	
Elektroniczne zbiory danych jako źródło informacji o stanie bezpieczeństwa i higieny w miejscu pracy	59
Electronic datasets as a source of information on the state of health and safety in the workplace	
Joanna Sadowy	
Wydawnictwa GUS. Marzec 2024	69
Publications of Statistics Poland. March 2024	
Dla autorów	71
For the authors	
Działy „WS” – tematyka artykułów	82
WS sections – topics of the articles	

OD REDAKCJI

W kwietniowym numerze „Wiadomości Statystycznych. The Polish Statistician” proponujemy Państwu lekturę trzech artykułów recenzowanych oraz innych opracowań.

Prof. dr hab. Elżbieta Gołata w pracy *Różnice w pomiarze płodności kobiet zależnie od szacunku liczby ludności* omawia wyniki pomiaru płodności kobiet w Polsce przeprowadzonego dla populacji kraju definiowanej na dwa sposoby: według krajowej definicji *ludności* i według definicji *ludności rezydującej*. Bada także zróżnicowanie zjawiska pod względem wieku kobiet i w ujęciu terytorialnym. W obliczeniach wykorzystuje dane za 2021 r. (z uwzględnieniem zmian płodności kobiet od generacji 1945 r.) udostępniane przez GUS i Eurostat, w tym wyniki Narodowego Spisu Powszechnego Ludności i Mieszkań 2021, oraz dane pochodzące z Human Fertility Database. Stosuje metody analizy demograficznej, przekrojowej i kohortowej, w ujęciu poprzecznym i wzdłużnym. Z badania wynika, że płodność kobiet znacznie się obniżyła w młodszych generacjach – w największym stopniu w grupie wieku 35–39 lat. Wartość współczynnika dzietności kobiet w Polsce obliczona dla ludności według definicji krajowej okazała się też niższa od wartości obliczonej dla ludności rezydującej.

Tematem artykułu dr. Tomasza Banasika jest *Ocena zależności między rozwojem branży leasingowej a wzrostem gospodarczym w Polsce*. W badaniu opartym na danych GUS i Związku Polskiego Leasingu za lata 2010–2021 autor wykorzystuje szeregi czasowe wartości PKB per capita i wartości umów leasingowych netto w przeliczeniu na mieszkańca jako miary – odpowiednio – wzrostu gospodarczego i aktywności branży leasingowej. Stwierdza, że wyniki analizy kointegracji i modelu korekty błędem wskazują na występowanie istotnej długookresowej zależności między rozwojem branży leasingowej a wzrostem gospodarczym w Polsce, niemniej potrzebne są dalsze badania w tym zakresie, z uwzględnieniem bieżących zmian społeczno-gospodarczych.

Praca dr. hab. Marcina Pełki, prof. UEW, i prof. Antonia Irpina pt. *The application of multidimensional scaling of data to determining changes in retailer customers' preferences* ma na celu wskazanie, które podejście w ramach metody skalowania wielowymiarowego: danych klasycznych, danych symbolicznych interwałowych czy danych symbolicznych histogramowych jest najodpowiedniejsze do rozpoznania zmian preferencji klientów sklepów. Autorzy analizują zmiany, jakie nastąpiły na rynku kosmetyków w trakcie pandemii COVID-19. Wykorzystują dane o zamówieniach na produkty do makijażu jednej z polskich firm złożonych w okresie 2019–2021 przez 18 niewielkich drogerii z obszaru woj. dolnośląskiego. Z badania wynika, że w latach 2020 i 2021 w porównaniu z 2019 r. zmniejszyła się popularność kosmetyków do upiększania ust i policzków, a zwiększyła – do makijażu oczu. Najbardziej efektywną metodą wychycenia i analizy zmian zachodzących w czasie na rynku okazało się skalowanie wielowymiarowe danych symbolicznych histogramowych.

Ponadto publikujemy opracowanie dr. Dariusza Smolińskiego pt. *Elektroniczne zbiory danych jako źródło informacji o stanie bezpieczeństwa i higieny w miejscu pracy*. Autor postuluje stworzenie centralnej platformy danych z zakresu bhp, która odciążałaby pracodawców, a także ułatwiłaby wykonywanie zadań instytucjom uprawnionym do analizowania warunków pracy.

W numerze jak zawsze znajdują Państwo również omówienie najnowszych publikacji GUS przygotowane przez Joannę Sadowy.

Życzymy przyjemnej lektury.

FROM THE EDITORIAL TEAM

The April issue of *Wiadomości Statystyczne. The Polish Statistician* consists of three peer-reviewed articles and other studies.

In the work entitled *Varying measurement of women's fertility depending on different estimates of a population*, Elżbieta Gołata, PhD, DSc, ProfTit, discusses the results of female fertility measurement in Poland performed for the country's population defined in two ways: according to the national definition of *population* and according to the *resident population* definition. The author also examines the variability of the phenomenon depending on women's age and region. The calculations are based on data referring to the year 2021 (considering the changes in fertility starting from the 1945 generation) provided by Statistics Poland and Eurostat, including the results of the National Population and Housing Census 2021, and data from the Human Fertility Database. The research uses cross-sectional and longitudinal demographic, period and cohort methods of analysis. The study shows that female fertility has significantly declined among the younger generations – most considerably in the 35–39 age group. Moreover, the value of the female fertility rate in Poland calculated for the population according to the national definition proved lower than the value calculated for the resident population.

Tomasz Banasik, PhD, presents an *Assessment of the relationship between the development of the leasing industry and economic growth in Poland*. The study is based on data from Statistics Poland and the Polish Leasing Association for the years 2010–2021. The author uses GDP *per capita* time series and the actual net values *per capita* of leasing contracts as measures of economic growth and the activity of the leasing industry, respectively. He concludes that the results of the cointegration analysis and the error correction mechanism indicate the existence of a significant long-lasting relationship between the development of the leasing industry and economic growth in Poland. However, the author recommends further research in this area, especially one considering the ongoing socio-economic changes.

The paper by Marcin Pełka, PhD, DSc, Professor at the Wrocław University of Economics and Business, and Professor Antonio Iripino, entitled *The application of multidimensional scaling of data to determining changes in retailer customers' preferences* aims to indicate which approach in the framework of multidimensional scaling (the scaling of classical data, symbolic interval data or symbolic histogram data) is most relevant when identifying changes in the preferences of shoppers. The authors analyse the changes that took place in the beauty market during the COVID-19 pandemic. They use data on orders for makeup products of one of the Polish cosmetics producers placed by 18 small health and beauty retailers in Lower Silesia in 2019–2021. The study shows that in 2020 and 2021, compared to 2019, the popularity of lip and cheek products decreased, while that for eye makeup increased. Multidimensional scaling of histogram symbolic data proved the most effective method for capturing and analysing changes in the market over time.

The issue also features a study by Dariusz Smoliński, PhD, entitled *Electronic datasets as a source of information on the state of health and safety in the workplace*. The author advocates the creation of a health and safety central data platform, which would ease employers' strain and support institutions authorised to analyse working conditions.

The issue concludes with Joanna Sadowy's compilation of the latest publications of Statistics Poland.

We wish you pleasant reading.

Różnice w pomiarze płodności kobiet zależnie od szacunku liczby ludności

Elżbieta Gołata^a

Streszczenie. Głównym celem artykułu jest przedstawienie różnic w pomiarze płodności kobiet w Polsce w zależności od metody szacunku liczby ludności. Dodatkowym zamierzeniem jest określenie zróżnicowania płodności według wieku oraz w ujęciu terytorialnym jako konsekwencji przyjętych definicji ludności. W analizach wykorzystano dane GUS i Eurostatu oraz dane z Human Fertility Database za 2021 r., z uwzględnieniem zmian płodności kobiet od generacji 1945 r. Wykorzystano metody analizy demograficznej, przekrojowej i kohortowej. Zastosowanie ujęcia wzdłużnego pozwoliło wskazać także zmiany wzorców płodności w kolejnych generacjach kobiet. Zgodnie z zasadą konstrukcji współczynników demograficznych odniesiono liczbę zdarzeń do ludności mogącej ich rzeczywiście doświadczyć, tj. ludności rezydującej. Przeliczono wszystkie współczynniki płodności, opierając się na szacunku liczby ludności rezydującej na podstawie wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2021.

Z analizy wynika, że współczynnik dzietności kobiet w Polsce w 2021 r. obliczony dla ludności według definicji krajowej był niższy o 4,3% od współczynnika obliczonego dla ludności rezydującej. Największe zaniżenie w skali kraju dotyczyło kobiet z grupy wieku 35–39 lat (6,3%), a w przekroju terytorialnym – zamieszkujących woj. podkarpackie (11,0%). Rozbieżności są większe dla grup wieku i regionów o większej mobilności, co wskazuje na potrzebę uwzględnienia ludności rezydującej w pomiarze natężenia procesów demograficznych.

Słowa kluczowe: płodność, ludność, ludność rezydująca, spis ludności

JEL: A10, C8, J10, J11

Varying measurement of women's fertility depending on different estimates of a population

Abstract. The primary objective of the article was to present the differences in women's fertility measurements in Poland depending on the method of estimating the population number. An additional aim was to determine the variability of fertility rates by age and geographical region as a consequence of the adopted definitions of population. The study used data from Statistics Poland, Eurostat and the Human Fertility Database for 2021, taking into account changes in women's fertility starting from the 1945 generation. Demographic, period and cohort analysis methods were employed in the study. The longitudinal approach was used to additionally identify changes in fertility patterns across generations of women. In accordance with the principle of constructing demographic coefficients, the number of events was related to the population that could actually experience them, i.e. the resident population. All partial fertility rates were

^a Uniwersytet Ekonomiczny w Poznaniu, Instytut Informatyki i Ekonomii Ilościowej, Katedra Statystyki, Polska / Poznań University of Economics and Business, Institute of Informatics and Quantitative Economics, Department of Statistics, Poland. ORCID: <https://orcid.org/0000-0001-5570-9576>.

E-mail: elzbieta.golata@ue.poznan.pl.

recalculated based on the estimated resident population according to the Population and Housing Census 2021.

The results of the analysis indicate that total fertility rate (TFR) of women in Poland in 2021 calculated for the population according to the national definition was lower by 4.3% than the rate calculated for the resident population. The most significant underestimation was observed for the 35–39 age group (by 6.3%), and in terms of geographical regions, for Podkarpacie (by 11.0%). The discrepancies are larger for age groups and regions with greater mobility, showing the need to take into account the resident population when measuring the intensity of demographic processes.

Keywords: fertility, population, resident population, population census

1. Wprowadzenie

Szacunek liczby ludności ma zasadnicze znaczenie dla badania i oceny wszystkich procesów demograficznych, a także ekonomicznych i społecznych, szczególnie gdy poziom zjawiska mierzony jest per capita. Określenie właściwego stanu i struktury ludności, które powinny stanowić podstawę odniesienia w konstrukcji współczynników demograficznych, jest jednak złożone i dyskusyjne.

Najważniejszym źródłem danych o mieszkańcach kraju jest Narodowy Spis Powszechny Ludności i Mieszkań (NSP) przeprowadzany przez Główny Urząd Statystyczny. W niniejszym opracowaniu wyniki NSP 2021 posłużyły do oszacowania populacji Polski według dwóch definicji: krajowej definicji *ludności* stosowanej przez GUS na potrzeby polskiej statystyki oraz definicji *ludności rezydującej* (ang. *usually resident population*) stosowanej przez GUS na potrzeby statystyki międzynarodowej i przez Eurostat. W kategorii *ludności* określanej według definicji krajowej ujmowani są stali mieszkańcy, w tym osoby, które przebywają czasowo za granicą (bez względu na to, jak długo), ale zachowały stałe zameldowanie w Polsce, nie są natomiast do niej zaliczani imigranci czasowo przebywający w Polsce (GUS, 2022a, s. 19). Z kolei *ludność rezydująca* – zgodnie z Rozporządzeniem Parlamentu Europejskiego i Rady (UE) nr 1260/2013 z dnia 20 listopada 2013 r. w sprawie statystyk europejskich w dziedzinie demografii (dalej: rozporządzenie nr 1260/2013) – to ludność zamieszkująca nieprzerwanie przez okres co najmniej 12 miesięcy przed czasem odniesienia¹. Różnica polega więc na tym, że w definicji krajowej są uwzględnione osoby, które wyjechały z kraju, i nie są uwzględnieni cudzoziemcy mieszkający w Polsce, jeśli nie mają pozwolenia na osiedlenie się (karty stałego pobytu), nawet gdy czas ich zamieszkania w Polsce jest dłuższy niż 12 miesięcy. Każde z tych podejść do definiowania pojęcia *ludności* ma swoich zwolenników i przeciwników, co uwidoczniło się podczas dyskusji, którą wywołała publikacja Unii Metropolii i Fundacji im. Stefana Batorego (Abramowska-Kmon i in., 2023)².

¹ Bądź deklarująca zamiar przebywania w miejscu zamieszkania przez co najmniej rok.

² Panel dyskusyjny poświęcony tej tematyce zorganizował prezes GUS Dominik Rozkrut podczas Konferencji Metodologia Badań Statystycznych MET2023 5 lipca 2023 r. w Warszawie, a 12 lipca 2023 r. odbyła się

Rozważane zagadnienie wiąże się przede wszystkim z problemem szacowania migracji, który nie jest ani nowy, ani prosty do rozwiązania. Dyskusja dotycząca definiowania pojęcia *ludności* była podejmowana w wielu opracowaniach (m.in.: Bijak i in., 2007; Fihel, 2015; Gołata, 2012, 2013, 2014, 2016a, 2016b, 2018; Jończy, 2015, 2023; Kędelski, 1990; Paradysz, 2007a, 2007b; Sakson, 2002; Śleszyński, 2004a, 2004b, 2005). Wzrost mobilności międzynarodowej, spowodowany przez rosnące nierówności ekonomiczne i społeczne, kryzys klimatyczny oraz konflikty lokalne i międzynarodowe, nie tylko komplikuje pomiar liczby ludności, lecz także potęguje jego skalę i znaczenie (Komitet Nauk Demograficznych Polskiej Akademii Nauk, 2022). Do 2015 r. rozbieżności w pomiarze populacji według definicji krajowej i według definicji ludności rezydującej tłumaczono skalą emigracji z Polski, szczególnie dużą liczbą osób, które wyjechały za granicę, ale nie wymeldowały się z pobytu stałego (zob. Fihel, 2015; Gołata, 2012, 2013, 2018; Jończy, 2015)³. Od 2016 r. Polska stała się również jednym z najważniejszych krajów docelowych migracji w Unii Europejskiej, a masowa imigracja i fala uchodźców z Ukrainy po 24 lutego 2022 r. przyniosły dodatkowe wyzwania. Zadaniem statystyki publicznej jest dokonanie pomiaru migracji i szacunku liczby ludności jak najlepiej odzwierciedlającego stan rzeczywisty, a przyjęte rozwiązania powinny umożliwiać porównywalność międzynarodową.

Problematyka definiowania pojęcia *ludności* w badaniach demograficznych jest bardzo obszerna. W niniejszym opracowaniu ograniczono się do ukazania konsekwencji przyjętej definicji ludności dla pomiaru jednego z najczęściej dyskutowanych zjawisk demograficznych – płodności kobiet. Pierwsze pytania dotyczące natężenia płodności pojawiają się już przy porównaniu wartości powszechnie znanej miary syntetycznej, jaką jest współczynnik dzietności teoretycznej (ang. *total fertility rate* – TFR). Okazuje się, że wartości TFR wyznaczone na podstawie danych Eurostatu różnią się od wartości obliczanych na podstawie danych GUS – te pierwsze są nieco wyższe od drugich, co może dziwić, zwłaszcza że bazy danych Eurostatu są zasilane przez GUS⁴.

Rozważania podjęte w niniejszym opracowaniu zmierzają do wyjaśnienia przyczyn powyższych rozbieżności. Omówiono konsekwencje niespójności definicji zdarzeń ujmowanych w liczniku i ludności stanowiącej podstawę odniesienia w mianowniku współczynników demograficznych. Zasadniczym celem artykułu jest przedstawienie różnic w pomiarze płodności kobiet w Polsce w zależności od metody szacunku liczby ludności. Dodatkowym zamierzeniem jest określenie zróżnicowania płodności według wieku i w ujęciu terytorialnym jako konsekwencji przyjętych definicji ludności.

debata zorganizowana przez prezesa Fundacji im. Stefana Batorego Edwina Bendyka i Marcina Wojdata z Unii Metropolii im. Pawła Adamowicza.

³ Zdaniem Pawła Kaczmarczyka (Gołata, 2022) nie powinniśmy o tym zapominać, ponieważ Polska nadal jest jednym z najważniejszych w Europie krajów emigracji.

⁴ Źródłem większości danych bezwzględnych w bazach Eurostatu są statystyki krajowe, a dane względne, w tym współczynniki odnoszące liczby zdarzeń do liczby ludności, są opracowywane bezpośrednio przez Eurostat – jednolicie dla wszystkich krajów (GUS, 2016).

2. Metoda badania

Badanie dotyczyło oceny wpływu szacunku liczby ludności na pomiar płodności kobiet w Polsce, w tym na jej zróżnicowanie według wieku i w ujęciu terytorialnym. Najpierw przedstawiono rezultaty szacowania populacji Polski według dwóch definicji: krajowej definicji *ludności* i definicji *ludności rezydującej*, opublikowane przez GUS na podstawie wyników NSP 2021. Konsekwencje szacunku liczby ludności przedstawiono w ujęciu – najczęściej stosowanej – analizy poprzecznej (przekrojowej) dla kohort hipotetycznych z uwzględnieniem zróżnicowania terytorialnego oraz analizy wzdłużnej (kohortowej) dla generacji rzeczywistych. Ponieważ dane z NSP 2021 stanowią podstawę korekty bilansów ludności (według definicji krajowej) dla lat wcześniejszych oraz szacunku liczby ludności dla lat kolejnych, do badania wybrano dane za 2021 r. W celach porównawczych w analizie uwzględniono także lata wcześniejsze, co zostało każdorazowo wskazane w komentarzach. Dane GUS zaczerpnięto z roczników demograficznych, Bazy Demografia GUS⁵ i Banku Danych Lokalnych (BDL) GUS⁶. Korzystano także z danych Eurostatu i Human Fertility Database (HFD)⁷.

Zastosowano metody demograficznej analizy płodności, intensywności i kalendara w ujęciu poprzecznym i wzdłużnym. Obliczono podstawowe miary płodności dla ludności rezydującej, które porównano z publikowanymi przez GUS współczynnikami dla ludności określanej według definicji krajowej. Wprowadzono odpowiednie indeksy dla rozróżnienia liczby kobiet i współczynników obliczanych w odniesieniu do ludności według definicji krajowej (K) i ludności rezydującej (R).

Szczegółowe informacje dotyczące systemu współczynników i kryteriów ich klasyfikacji (m.in. surowe i cząstkowe, pierwszej i drugiej kategorii oraz przekrojowe i wzdłużne) zostały wyczerpująco przedstawione w podręcznikach Holzera (2003), Kędelskiego i Paradysza (2006) oraz Kurkiewicz (2010). Podstawowa zasada konstrukcji współczynników polega na odniesieniu liczby zdarzeń demograficznych do liczby osób mogących doświadczyć tych zdarzeń (Kędelski i Paradysz, 1984, 2006). Aby nie naruszać zasady spójności, liczbę urodzeń żywych w Polsce odnoszono do liczby kobiet rezydujących w Polsce. Cząstkowe współczynniki płodności według wieku i miejsca zamieszkania w przekroju województw obliczono według następującego wzoru:

⁵ <https://demografia.stat.gov.pl/bazademografia>.

⁶ <https://bdl.stat.gov.pl/bdl>.

⁷ Human Fertility Database (<https://www.humanfertility.org/>) powstała w 2007 r. jako rezultat międzynarodowego projektu z udziałem zespołów badawczych z Instytutu Badań Demograficznych Maxa Plancka i Wiedeńskiego Instytutu Demografii.

$$F_R(x) = \frac{U(x)}{K_R(x)} \cdot C, \quad (1)$$

gdzie:

$F_R(x)$ – cząstkowy współczynnik płodności kobiet rezydujących w wieku x . Jeśli $C = 1000$, to współczynnik ten przedstawia liczbę urodzeń żywych na 1000 kobiet rezydujących w wieku x w roku $t = 2021$,

$U(x)$ – liczba dzieci żywo urodzonych przez kobiety w wieku x w roku $t = 2021$,

$\overline{K_R(x)}$ – średnia liczba kobiet rezydujących w 2021 r. w wieku x , którą z uwagi na dostępność danych obliczono jako średnią dla stanu na 31 grudnia 2020 r. i 2021 r.,

x – wiek w latach ukończonych $x = 15, \dots, 49$,

C – stała (1 lub 1000).

W analizie wykorzystano także TFR, nazywany *sumą zdarzeń zredukowanych*. Wskazuje on przeciętną liczbę żywo urodzonych dzieci, które urodziłaby kobieta w ciągu całego okresu zdolności rozrodczej, przy cząstkowych współczynnikach płodności według wieku takich jak w roku badania (np. 2021). Obliczono go jako sumę cząstkowych współczynników płodności według wieku:

$$TFR_R = \sum_x F_R(x). \quad (2)$$

W pierwszej kolejności przedstawiono wyniki szacunku miar płodności w skali kraju, a następnie w przekroju województw.

Analiza wzdłużna pozwoliła uzupełnić wnioski z badania w dłuższej perspektywie. Porównano płodność najmłodszych generacji kobiet oraz tych, które są już po okresie prokreacyjnym. Zastosowano tutaj analogiczne miary intensywności płodności, ale obliczone dla generacji rzeczywistych G , tj. zgodnie z zasadami analizy wzdłużnej (kohortowej). Z uwagi na wymagany długi okres obserwacji i zasady analizy porównawczej, w tej części ograniczono się do ludności według definicji krajowej, ale na tyle, na ile pozwalała na to dostępność danych o ludności rezydującej, dokonano odpowiednich szacunków i przedstawiono relacje pomiędzy analizowanymi współczynnikami. W tym celu wykorzystano dane GUS, dane z HFD oraz wyniki wcześniejszych badań Tymickiego i in. (2018), które uzupełniono własnymi szacunkami wzdłużnych cząstkowych współczynników płodności (ang. *Cohort age-specific fertility rates*):

$$F_G(x) = \frac{U_G(x)}{K_G(x)} \cdot C, \quad (3)$$

gdzie:

$F_G(x)$ – wzdłużny cząstkowy współczynnik płodności kobiet generacji G , np. urodzonych w roku $G = 1980$ w wieku x w roku t . Jeśli $C = 1000$, to współczynnik ten przedstawia liczbę urodzeń żywych na 1000 kobiet urodzonych np. w roku $G = 1980$, będących w wieku x w roku obserwacji t ,

$U_G(x)$ – liczba dzieci żywo urodzonych przez kobiety generacji G będące w wieku x w roku obserwacji t ,

$\overline{K_G(x)}$ – średnia liczba kobiet generacji G będących w wieku x w roku obserwacji t .⁸

Wzdłużny współczynnik dzietności kobiet generacji G (ang. *cohort total fertility rate* – CTFR) jest miarą rzeczywistej dzietności kobiet urodzonych w danym roku. Przedstawia on przeciętną liczbę żywo urodzonych dzieci przez kobietę generacji G , np. urodzoną w roku $G = 1980$, w ciągu całego okresu zdolności rozrodczej, przy wzdłużnych cząstkowych współczynnikach płodności dla poszczególnych lat wieku prokreacyjnego $x = 15, \dots, 49$ rzeczywiście obserwowanych dla tej generacji kobiet. Obliczono go jako sumę wzdłużnych cząstkowych współczynników płodności według wieku, zgodnie ze wzorem:

$$CTFR(G) = \sum_x F_G(x). \quad (4)$$

3. Wyniki badania

Wartości TFR publikowane przez GUS i Eurostat są różne. W danych za lata 1990–2022 (wykr. 1) zwracają uwagę dwa okresy, w których uwidoczniły się szczególnie duże różnice: 1990–1999 i 2009–2021. Pierwszy okres to czas przed akcesją Polski do UE, a drugi obejmuje lata, dla których dostępny jest szacunek ludności rezydującej. Przyczyn rozbieżności można szukać w odmiennym definiowaniu *urodzenia żywego*⁹ i szacowaniu liczby kobiet według wieku.

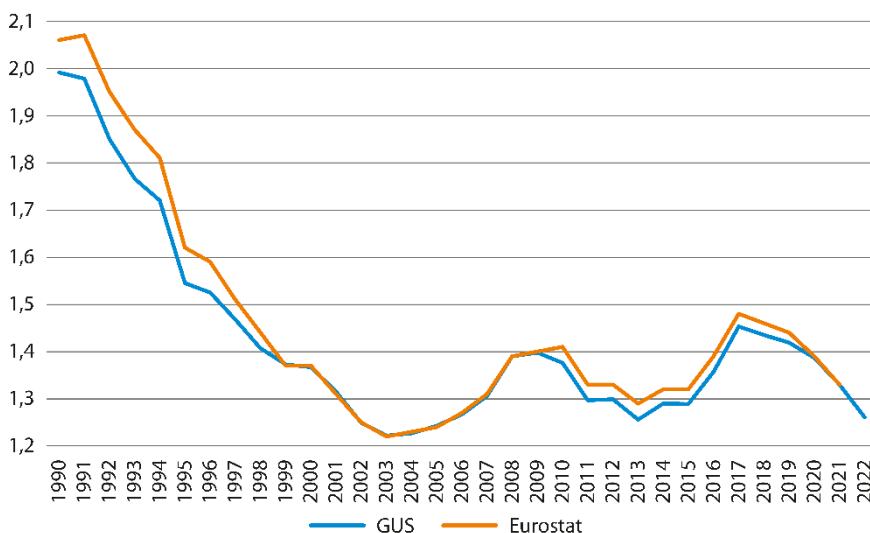
Tymicki i Zeman (2022) w opracowaniu przedstawiającym wyniki szacunków w HFD zastrzegają, że ze względu na dużą emigrację z Polski, która nie jest w pełni ujmowana w oficjalnych statystykach ludności, wskaźniki dzietności od 2000 r. mogą być zaniżone nawet o 10%. Omawiając jakość danych, autorzy zwracają uwagę na

⁸ Zwyczajowo przyjmuje się stan na 30 czerwca danego roku albo wyznacza się średnią ze stanów na 31 grudnia roku poprzedniego (1 stycznia roku badanego) i 31 grudnia roku badanego.

⁹ W 1994 r. wprowadzono definicję *urodzenia żywego* opartą na trzech elementach: czasie trwania ciąży, urodzeniowej masie ciała i wzroście urodzeniowym. Według zaleceń Światowej Organizacji Zdrowia noworodka uważa się za żywego, jeśli waży co najmniej 501 g lub – w przypadku nieznaney masy urodzeniowej – został wydobyty z organizmu matki po co najmniej 22. tygodniu ciąży lub ma co najmniej 25 cm długości.

definicję *urodzenia*¹⁰, zgodność roku urodzenia i rejestracji zdarzenia, urodzenia poza Polską i szacunek liczby ludności, zwłaszcza w zakresie uwzględniania emigracji. Zauważają, że według danych GUS (2015) liczba obywateli polskich, którzy byli zaliczani do populacji rezydentów, ale w rzeczywistości mieszkali za granicą, wzrosła z 0,4 mln w 2002 r. do 2,3 mln w 2014 r. i w większości były to osoby w wieku prokreacyjnym.

Wykr. 1. Współczynnik dzietności teoretycznej (TFR) dla Polski według danych GUS i Eurostatu



Źródło: opracowanie własne na podstawie Eurostatu (b.r.), roczników demograficznych i danych z BDL.

Dyskusja na temat szacowania liczby ludności, która toczyła się po NSP 2011, skłoniła GUS (2016) do opublikowania informacji o zakresie danych dotyczących stanu i struktury ludności i ruchu naturalnego oraz ich dostępności w świetle rozporządzenia nr 1260/2013. Opisano mechanizm gromadzenia danych i źródła możliwych rozbieżności. Wyjaśniono, że we wskazanym akcie prawnym zostały przyjęte rozwiązania odmienne od rozwiązań stosowanych przez niektóre kraje na własne potrzeby, ale Eurostat dopuszcza różnice szacunków w stosunku do podawanych przez poszczególne kraje. W odniesieniu do Polski oznacza to uznanie podwójnego szacunku liczby ludności: według definicji krajowej i według definicji ludności rezydującej, tj. zgodnie ze wspomnianym rozporządzeniem. Chociaż omawiane rozporządzenie obowiązuje od 2013 r., to GUS przekazał do Eurostatu dane o liczbie i strukturze ludności rezydującej

¹⁰ Również GUS (2016) zwraca uwagę na różnice definicyjne. Podkreśla, że baza Eurostatu zawiera informacje o dzieciach martwo urodzonych po okresie trwania ciąży wynoszącym co najmniej 28 tygodni, natomiast dane publikowane przez GUS dotyczą dzieci martwo urodzonych o ciężarze urodzeniowym wynoszącym co najmniej 500 g (co odpowiada ok. 22 tygodniom trwania ciąży).

również za lata 2009–2012. Jednak do tej pory wszystkie współczynniki demograficzne szacowane są w odniesieniu do definicji krajowej.

3.1. Różnice w szacunku liczby ludności według definicji krajowej i ludności rezydującej

W 2021 r. urodziło się w Polsce 331 511 dzieci (liczba urodzeń żywych). Określenie miary intensywności, jaką jest współczynnik płodności, wymaga odniesienia liczby tych dzieci do liczby kobiet w wieku prokreacyjnym w połowie roku¹¹. Z NSP 2021 wynika, że ludność Polski według definicji krajowej liczyła 38 036 118 osób, w tym 8 790 600 kobiet w wieku prokreacyjnym (15–49 lat), a według definicji ludności rezydującej Polskę zamieszkiwało 37 019 327 osób, w tym 8 425 384 kobiety w wieku prokreacyjnym. Różnica dla całej populacji wynosi ponad milion osób, co stanowi 2,7% ludności według definicji krajowej. Największe różnice liczebności w zakresie grup wieku pomiędzy obiema kategoriami ludności dotyczą grupy wieku 20–55 lat (wykr. 2). Są to osoby w wieku produkcyjnym, w większości mobilnym, w tym kobiety w wieku prokreacyjnym. W danych wyraźnie uwidacznia się pokłosie migracji poakcesyjnych w postaci dużo mniej licznych roczników dzieci w wieku do 18 lat. Liczba rezydujących dzieci i młodzieży w wieku do 18 lat okazała się niższa o ponad 230 tys. od wielkości obliczonej dla ludności według definicji krajowej. Z kolei liczba osób w wieku produkcyjnym jest niższa o ponad 800 tys., a liczba kobiet w wieku prokreacyjnym – o 365 216. Największa różnica pomiędzy omawianymi kategoriami ludności dotyczy kobiet w grupie wieku 35–39 lat i wynosi blisko 93 tys., co stanowi prawie 6%. Współczynnik płodności dla 2021 r. oszacowany na podstawie liczby ludności spełniającej definicję krajową wynosi 37,77 urodzenia na 1000 kobiet, a na podstawie liczby ludności rezydującej¹² – 39,41.

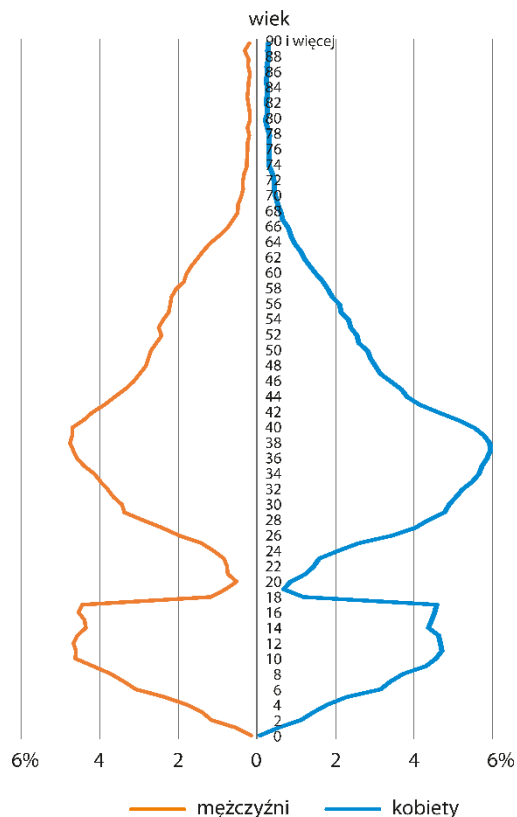
Zróznicowanie relacji liczby ludności określanej według definicji krajowej i liczby ludności rezydującej uwidacznia się szczególnie w ujęciu terytorialnym. We wszystkich województwach liczba kobiet w wieku prokreacyjnym mieszczących się w definicji ludności rezydującej jest niższa niż kobiet w populacji Polski określanej według definicji krajowej. W woj. podkarpackim w wymiarze względnym jest to 8,8%. Podobnie wysokie różnice dotyczą województw podlaskiego (8,0%) i opolskiego (7,8%). Najmniejsza różnica dotyczy woj. mazowieckiego (ok. 1%). Wyraźniejsze rozbieżności w rozpatrywanym zakresie uwidaczniają się na niższym poziomie podziału administracyjnego¹³.

¹¹ Zwyczajowo przyjmuje się stan na 30 czerwca danego roku albo wyznacza średnią ze stanów na 31 grudnia roku poprzedniego (1 stycznia roku badanego) i 31 grudnia roku badanego.

¹² Obydwa współczynniki uwzględniają stany skorygowane z wykorzystaniem danych z NSP 2021.

¹³ W gminie Trzcianne (woj. podlaskie) różnica między liczbą ludności rezydującej i liczbą ludności określanej według definicji krajowej wynosi 14,4%, a w Gorzowie Śląskim (woj. opolskie) – 14,0%. Odwrotna relacja między szacunkami opartymi na tych definicjach dotyczy jedynie 2,8% jednostek podziału administracyjnego. W mieście Wronki (woj. wielkopolskie) liczba ludności rezydującej jest o 6,8% wyższa niż liczba

Wykr. 2. Względna różnica między liczbą ludności według definicji krajowej i liczbą ludności rezydującej w podziale na wiek i płeć w 2021 r.



Źródło: obliczenia własne na podstawie: GUS (b.r.).

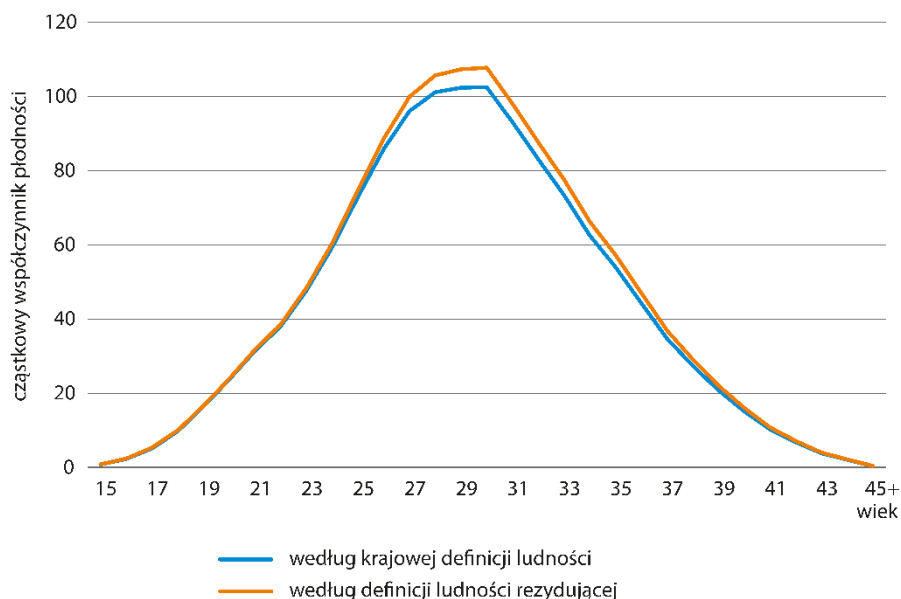
3.2. Pomiar płodności w analizie poprzecznej

Powyższe różnice w szacunku liczby kobiet zmieniają wartości cząstkowych współczynników płodności. Zmiany są szczególnie wyraźne w grupie kobiet w wieku o najwyższej intensywności płodności. Dla kobiet w wieku 37 i 38 lat cząstkowe współczynniki płodności są o ponad 6,3% wyższe w przypadku ludności rezydującej niż w przypadku ludności określanej według definicji krajowej (wykr. 3 i 4). Ta zmiana wielkości odniesienia nie wpływa istotnie na zmianę kalendarza płodności¹⁴.

ludności określane według definicji krajowej. Podobnie jest w mazowieckich gminach Raszyn i Lesznowola (odpowiednio 6,7% i 5,8%).

¹⁴ Kalendarze płodności to statystyczne charakterystyki rozkładu urodzeń według wieku kobiet, np. średni wiek kobiet w momencie rodzenia, wiek najczęstszy, miary dyspersji i asymetrii. Analiza uwzględniająca kolejność urodzeń pozwala także na wskazanie średniego wieku w momencie rodzenia pierwszego dziecka i kolejnych dzieci oraz odstępów intergenetycznych.

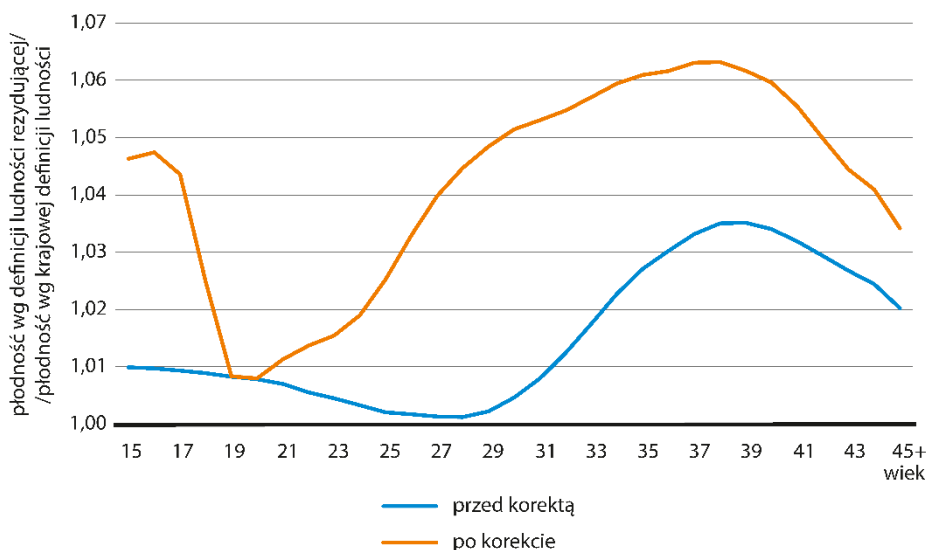
Wykr. 3. Krzywa płodności w 2021 r. dla ludności określonej według definicji krajowej i ludności rezydującej na podstawie danych po korekcie wynikającej z oszacowań NSP 2021



Źródło: opracowanie własne na podstawie GUS (2022b, 2022c) i danych z Bazy Demografia.

Suma cząstkowych współczynników płodności dla kolejnych roczników wieku prokreacyjnego stanowi TFR, który w 2021 r. dla kobiet rezydujących wynosił 1,386 w porównaniu z 1,328 dla kobiet ujętych w krajowej definicji ludności. Różnica wartości $TFR_R - TFR_K = 0,058$ w odniesieniu do przeciętnej liczby żywo urodzonych dzieci jest znaczna. Ten wynik dotyczy porównania współczynników obliczonych na podstawie danych po korekcie wynikającej z oszacowań NSP 2021. Szacunki ludności rezydującej sprzed tej korekty również wskazywały na zniżenie cząstkowych współczynników płodności i w efekcie – współczynnika dzietności (1,333 w stosunku do 1,319), jednak nie było ono tak wyraźne, jak po uwzględnieniu danych z NSP 2021 (wykr. 4).

Wykr. 4. Relacja między cząstkowymi współczynnikami płodności obliczonymi na podstawie liczby kobiet oszacowanej zgodnie z krajową definicją ludności i zgodnie z definicją ludności rezydującej w 2021 r. przed korektą wynikającą z oszacowań NSP 2021 i po korekcie



Źródło: opracowanie własne na podstawie GUS (2022b, 2022c) i danych z Bazy Demografia.

Relacje przedstawione na wyk. 4 unaoczniają, jak złożony i trudny jest szacunek liczby ludności, w tym ludności rezydującej. Gdyby w poszczególnych rocznikach liczba kobiet rezydujących i liczba kobiet w populacji określanej według krajowej definicji ludności były takie same, to cząstkowe współczynniki płodności byłyby identyczne, co odzwierciedlałaby linia równoległa do osi wieku. Warto w tym miejscu zauważyć, jak niewielka była różnica liczby kobiet w wieku 25–29 lat reprezentujących dwójako definiowaną ludność według danych przed korektą wynikającą z oszacowań NSP 2021 (niebieska linia na wyk. 4). Gdyby wyniki spisu były zgodne z wcześniejszymi bilansami, to na wykresie widniałaby tylko jedna krzywa przedstawiająca relację współczynników dla ludności rezydującej i ludności według definicji krajowej. Jest tak tylko w odniesieniu do kobiet w wieku 19 lat, co wyraźnie uwidoczniło się na wyk. 2 i co można określić mianem efektu akcesji (wzrost emigracji po 2004 r.). Poza tym jednym rocznikiem relacja szacunku liczby kobiet rezydujących w 2021 r. przed korektą wynikającą z oszacowań NSP 2021 i po korekcie jest różna co do poziomu oraz zróżnicowana według wieku. Szczególnie duże różnice dotyczą grupy wieku 30–34 lata (różnica w liczbie rezydentek wynosi ponad 62,5 tys. kobiet). W efekcie współczynnik płodności dla rezydentek tej grupy wieku wynosił 83 urodzenia żywe na 1000

kobiet przed korektą opartą na wynikach NSP 2021 w porównaniu z 87 po uwzględnieniu wyników NSP 2021.

3.3. Pomiar płodności w ujęciu terytorialnym

Wyraźne różnice szacunku liczby ludności obserwowane w skali całego kraju oraz przestrzenne zróżnicowanie rozmieszczenia ludności, jej struktur i natężenia procesów demograficznych sugerują potrzebę zwrócenia uwagi na rozbieżności szacunku płodności w ujęciu terytorialnym. W rocznikach demograficznych GUS zamieszczane są szacunki w odniesieniu do ludności według definicji krajowej. W tabl. 1 przedstawiono wartości współczynnika płodności dla ludności rezydującej. Z uwagi na dostępność danych i dla celów porównawczych wartości współczynnika oszacowano dla 5-letnich grup wieku.

Różnice szacunku liczby kobiet w wieku prokreacyjnym rozpatrywanych w populacji według krajowej definicji ludności i w zbiorowości ludności rezydującej w przekroju województw wynoszą w wymiarze względnym 1–9%. Przekłada się to bezpośrednio na oszacowania wartości współczynnika płodności kobiet w wieku prokreacyjnym. Dla wszystkich województw wyniki dotyczące kobiet rezydujących są wyższe niż dla kobiet ujmowanych w ramach ludności według definicji krajowej. Różnice między wartościami współczynnika w odniesieniu do tych definicji wynoszą od 0,8% dla woj. mazowieckiego do 9,6% dla woj. podkarpackiego. Stosunkowo niewielkie różnice dotyczą dwóch najmłodszych grup wieku: 15–19 lat i 20–24 lata. W przypadku województw łódzkiego i mazowieckiego wartości współczynnika płodności dla rezydentek i kobiet w ujęciu krajowej definicji ludności są zbliżone, a dla województw podlaskiego, warmińsko-mazurskiego czy zachodniopomorskiego – o blisko 6% wyższe w przypadku ludności rezydującej. W kolejnych grupach wieku rozbieżności są zdecydowanie wyraźniejsze. W grupach wieku o największej płodności, od 25 do 44 lat, w województwach podkarpackim, opolskim, podlaskim i warmińsko-mazurskim współczynniki płodności dla rezydentek są o 10–13% wyższe niż te określone dla kobiet w ujęciu krajowej definicji ludności.

Bardzo zróżnicowane są relacje między szacunkiem ludności według definicji krajowej i ludności rezydującej w przekroju województw dla najstarszej grupy wieku prokreacyjnego – 44–49 lat. Dodatkowo niewielka liczba urodzeń przez kobiety w tym wieku powoduje, że względna różnica pomiędzy wartościami współczynnika płodności dla rezydentek i dla kobiet w ujęciu krajowej definicji ludności waha się od –13% (woj. lubuskie) do 18% (woj. wielkopolskie).

Tabl. 1. Płodność kobiet i współczynnik dzietności teoretycznej w 2021 r. według województw na podstawie szacunku ludności według definicji krajowej i ludności rezydującej

Województwa		Ogółem	Urodzenia żywe na 1000 kobiet w wieku							TFR
			15–19 lat	20–24	25–29	30–34	35–39	40–44	45–49 lat	
<i>K</i> – ludność według definicji krajowej										
<i>R</i> – ludność rezydująca										
Ogółem	<i>K</i>	37,8	6,9	40,6	92,4	82,5	36,1	7,7	0,4	1,330
	<i>R</i>	39,4	7,1	41,1	95,9	87,0	38,3	8,1	0,4	1,390
Dolnośląskie	<i>K</i>	36,0	8,2	38,8	80,3	80,4	35,6	6,7	0,4	1,249
	<i>R</i>	37,5	8,5	39,5	82,8	84,2	37,8	7,1	0,4	1,301
Kujawsko-pomorskie	<i>K</i>	36,1	9,2	48,6	92,2	74,6	30,9	7,0	0,4	1,314
	<i>R</i>	37,9	9,6	49,7	96,9	79,6	33,1	7,4	0,4	1,383
Lubelskie	<i>K</i>	36,5	6,2	35,9	91,0	82,2	36,4	8,1	0,4	1,300
	<i>R</i>	38,7	6,4	36,5	96,5	89,2	39,6	8,7	0,4	1,386
Lubuskie	<i>K</i>	35,0	9,4	51,3	90,0	71,3	31,9	6,4	0,2	1,304
	<i>R</i>	36,6	9,8	51,7	93,3	76,0	34,1	6,7	0,2	1,359
Łódzkie	<i>K</i>	36,8	6,6	38,9	94,5	84,1	35,3	7,0	0,4	1,333
	<i>R</i>	37,7	6,7	38,6	96,1	87,1	36,7	7,2	0,4	1,365
Małopolskie	<i>K</i>	40,5	4,4	34,3	93,8	88,7	40,7	9,0	0,4	1,349
	<i>R</i>	42,6	4,5	35,1	98,1	94,1	43,5	9,6	0,5	1,427
Mazowieckie	<i>K</i>	40,8	5,7	34,8	90,5	93,8	43,7	9,8	0,6	1,390
	<i>R</i>	41,1	5,7	34,3	90,5	94,9	44,6	10,0	0,6	1,403
Opolskie	<i>K</i>	33,8	7,6	40,6	90,4	71,3	29,0	5,7	0,6	1,226
	<i>R</i>	36,7	7,9	41,5	97,2	79,3	32,6	6,3	0,6	1,327
Podkarpackie	<i>K</i>	37,1	5,0	34,4	94,6	81,5	36,3	7,2	0,3	1,294
	<i>R</i>	40,6	5,2	35,9	104,8	91,9	40,9	8,0	0,4	1,436
Podlaskie	<i>K</i>	39,2	5,2	36,8	101,2	86,8	35,3	8,0	0,5	1,364
	<i>R</i>	42,6	5,5	38,0	109,0	96,3	39,4	8,8	0,5	1,488
Pomorskie	<i>K</i>	41,5	9,1	50,9	96,3	84,9	37,2	8,6	0,4	1,432
	<i>R</i>	43,2	9,5	52,4	100,1	88,7	39,3	9,0	0,5	1,497
Śląskie	<i>K</i>	35,4	7,1	39,3	93,5	76,7	33,2	7,0	0,4	1,284
	<i>R</i>	36,7	7,3	39,8	97,0	80,6	35,1	7,3	0,5	1,337
Świętokrzyskie	<i>K</i>	33,6	4,9	38,3	90,8	75,0	30,7	5,8	0,2	1,228
	<i>R</i>	35,7	5,1	39,0	96,7	81,6	33,6	6,2	0,2	1,312
Warmińsko-mazurskie	<i>K</i>	33,7	10,0	48,1	85,1	69,4	28,6	6,3	0,2	1,240
	<i>R</i>	36,2	10,6	49,6	92,1	76,4	31,5	6,7	0,2	1,335
Wielkopolskie	<i>K</i>	40,5	6,9	47,6	102,5	84,8	35,9	7,8	0,3	1,425
	<i>R</i>	41,5	7,0	47,9	104,5	87,5	37,2	8,0	0,4	1,462
Zachodniopomorskie	<i>K</i>	34,2	9,5	48,5	84,8	71,5	30,6	6,6	0,4	1,260
	<i>R</i>	36,0	10,0	49,2	88,5	76,0	32,8	7,0	0,4	1,320

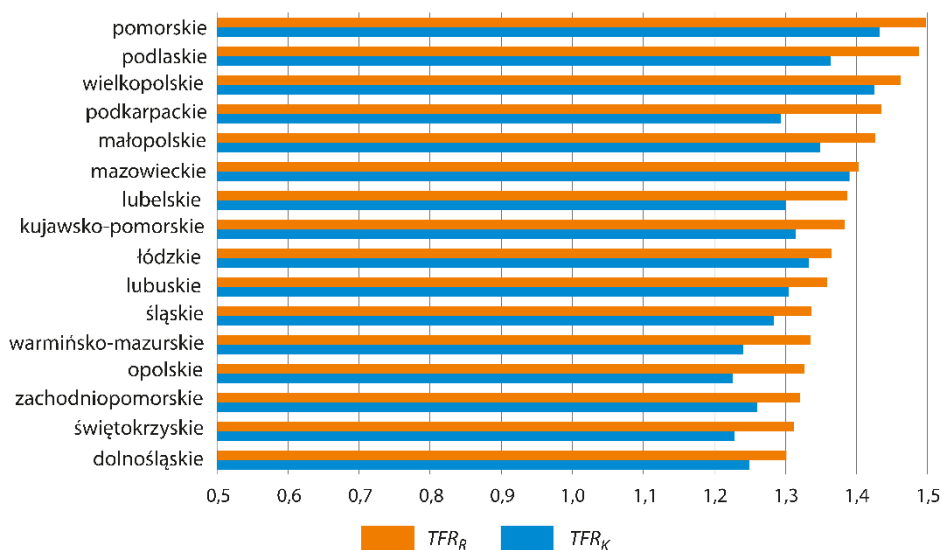
Źródło: opracowanie własne na podstawie GUS (2022b, 2022c) i danych z Bazy Demografia.

Synteza powyższych rozbieżności uwidacznia się w wartościach TFR dla kobiet rozpatrywanych w dwóch odmiennie zdefiniowanych populacjach kraju (tabl. 1). W woj. podkarpackim szacunek TFR_R jest o 11,0% wyższy niż TFR_K . Podobna relacja dotyczy województw podlaskiego (9,1%), opolskiego (8,2%) czy warmińsko-mazurskiego (7,7%). Wyniki te potwierdzają relację zaobserwowaną w przypadku miar płodności dla ludności według definicji krajowej i dla ludności rezydującej już na podstawie danych z NSP 2011. Dla województw opolskiego, podkarpackiego i podlaskiego wartości współczynnika płodności były zaniżone o 10–17% (Gołata, 2013).

Dane NSP 2021 wskazują, że w woj. mazowieckim TFR_R jest wyższy od TFR_K o niespełna 1%, a w woj. łódzkim – o 2,4%. Stosunkowo niewielka różnica wartości tych współczynników w woj. mazowieckim jest rezultatem dużego zróżnicowania terytorialnego tego obszaru. Wobec tego warto byłoby podjąć badania w przekroju mniejszych jednostek podziału administracyjnego. Jest to często utrudnione nie tylko z powodu braku danych, lecz także małej liczby zdarzeń demograficznych, co może potęgować oddziaływanie czynników losowych. Trzeba podkreślić, że w samym m.st. Warszawie liczba mieszkańców odpowiadających definicji ludności rezydującej jest o 4398 osób wyższa niż liczba mieszkańców według definicji krajowej (GUS, b.r.). Uwzględniając różne relacje według wieku, zauważono, że w 2021 r. średnia liczba rezydentek w wieku prokreacyjnym była o 807 osób wyższa niż kobiet w ujęciu krajowej definicji ludności (dane z BDL i Bazy Demografia). Częstkowe współczynniki płodności kobiet ujmowanych według krajowej definicji ludności okazały się nieco wyższe niż w przypadku rezydentek, poza grupą wieku 35–44 lata. W rezultacie wartość $TFR_R = 1,222$ była wyższa od $TFR_K = 1,213$ o 0,7%. Można przypuszczać, że sytuacja podobna do obserwowanej w Warszawie dotyczy także innych gmin woj. mazowieckiego, ponieważ wśród dziesięciu gmin, w których liczba ludności rezydującej była większa niż liczba ludności według definicji krajowej, trzy znajdują się w woj. mazowieckim. Są to gminy Raszyn, Lesznowola i Tarczyn.

Ważna refleksja pojawia się w odniesieniu do klasyfikacji województw według poziomu płodności. Przestrzenne zróżnicowanie migracji powoduje zniekształcenie wyników analizy terytorialnego zróżnicowania dzietności dla ludności według definicji krajowej (wykr. 5).

Wykr. 5. Współczynnik dzietności teoretycznej dla ludności rezydującej i dla ludności według definicji krajowej w przekroju województw w 2021 r.



Źródło: opracowanie własne na podstawie GUS (2022b, 2022c) i danych z Bazy Demografia.

Jeżeli szacunki wielkości licznika i mianownika współczynników płodności są zgodne, to można zauważyć, że najwyższe wartości TFR_R (tabl. 1) dotyczą województw: pomorskiego (1,5), podlaskiego (1,49), wielkopolskiego (1,46), podkarpackiego (1,44) i małopolskiego (1,43). W przypadku ludności według definicji krajowej uporządkowanie województw ze względu na TFR_K jest następujące: pomorskie (1,43), wielkopolskie (1,43), podlaskie (1,39), mazowieckie (1,36) i małopolskie (1,35). Jest to efekt odniesienia liczby urodzeń do liczby kobiet uwzględniającej te osoby, które w rzeczywistości nie mieszkają na danym obszarze. Wykorzystanie TFR_K jako miary dzietności prowadzi do konfuzji, ponieważ przykładowo z porównania TFR_K dla województw wielkopolskiego i podlaskiego można by wnioskować, że dzietność kobiet w tym pierwszym (1,425) jest wyższa niż w tym drugim (1,364). Tymczasem TFR_R dla woj. podlaskiego (1,488) jest wyższy niż dla woj. wielkopolskiego (1,462), więc nieprawdziwe jest twierdzenie o wyższej dzietności kobiet w Wielkopolsce. Podobna relacja dotyczy także innych województw, np. dzietność według TFR_K jest wyższa w woj. mazowieckim (1,390) niż w woj. podkarpackim (1,294), podczas gdy TFR_R wskazuje na odmienne uporządkowanie i znaczną różnicę wartości (odpowiednio 1,403 i 1,463).

3.4. Pomiar płodności w analizie wzdłużnej

Przedstawiając problemy pomiaru płodności, nie sposób pominąć analizy kohortowej. We wprowadzeniu wskazano na odmienne wartości współczynnika dzietności teoretycznej szacowane przez GUS i przedstawiane przez Eurostat (wykr. 1). W HFD, w której uwzględniane są długie szeregi czasowe w ujęciu przekrojowym i wzdłużnym, widnieje ostrzeżenie o zaniżeniu dzietności dla kohort urodzonych po 1965 r. wynoszącym nawet 10%. Powoduje to konsternację wśród odbiorców danych w środowisku międzynarodowym, ale także refleksję dotyczącą szacowania liczby ludności. Wyjściowy stan ludności w sposób oczywisty decyduje o rezultatach projekcji, a wyniki analiz kohortowych powinny stanowić podstawę założeń dotyczących płodności. Z tych względów dyskusji poddano także kohortowe współczynniki płodności według wieku i wzdłużne współczynniki dzietności dla generacji urodzonych przed 2000 r. Jest to syntetyczne uzupełnienie analizy dla Polski dostępnej w HFD, ale – zgodnie z celem artykułu – tam, gdzie to możliwe, wskazano konsekwencje stosowanej metody szacunku liczby ludności.

Wzdłużne ujęcie płodności kobiet stosunkowo rzadko bywa przedmiotem analiz, ponieważ wymaga obserwacji trwającej co najmniej 35 lat. Jednak analiza zmian zachowań prokreacyjnych znajduje uzasadnienie właśnie dla rzeczywistych generacji kobiet. Ujęcie przekrojowe dotyczące kohort hipotetycznych obserwowanych w jednym roku kalendarzowym jest rozwiązaniem teoretycznym. Jest ono uwarunkowane spełnieniem założeń dotyczących natężenia płodności w poszczególnych rocznikach wieku, takich jak w roku obserwacji dla kobiet różnych generacji urodzonych w okresie 35 lat. W celu rozszerzenia dotychczasowego ujęcia przekrojowego poniżej przedstawione zostaną syntetyczne wyniki analizy dla rzeczywistych generacji kobiet. Jest to kontynuacja (po 2015 r.) szacunków i analiz omówionych w artykule Tymickiego i in. (2018).

W tabl. 2 i na wykr. 6 przedstawiono szacunki wzdłużnych cząstkowych współczynników płodności dla generacji rzeczywistych 1945–2005. Tablicę 2 uzupełniono o wartości kohortowych współczynników dzietności, które przedstawiono także na wykr. 7. Dla generacji 1945–1970 współczynniki dzietności dotyczą płodności zakończonej¹⁵. Podobnie można traktować wartości tych współczynników dla generacji 1975–1980, choć w wyniku urodzeń przez kobiety w wieku 43 lat i więcej mogą one wzrosnąć. Nie ma podstaw, by przypuszczać, aby była to istotna zmiana, ponieważ w generacji 1970 r. wskutek urodzeń po ukończeniu 42 lat wzdłużny współczynnik dzietności wzrósł o 0,006. Z kolei wzdłużne współczynniki dzietności dla generacji 1985 r. i młodszych dotyczą płodności niezakończonych. Zaznaczono je kolorami: brązowym, pomarań-

¹⁵ Współczynniki dzietności obliczone dla wskazanych generacji kobiet uwzględniają cały – zakończony już – okres zdolności prokreacyjnej.

czowym, czerwonym i fioletowym i wprowadzono jedynie w celu analizy porównawczej. Wszystkie powyższe miary płodności obliczono dla ludności według definicji krajowej. Uwzględnienie ludności rezydującej jest możliwe dopiero od 2009 r. Przy ograniczonej migracji w latach wcześniejszych różnice są stosunkowo niewielkie.

Tabl. 2. Wzdłużne cząstkowe współczynniki płodności dla generacji rzeczywistych

Wiek	Urodzenia żywe na 1000 kobiet generacji												
	1945	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005
15 lat	0,0	1,3	1,1	1,3	1,4	1,5	1,8	1,3	1,0	0,9	1,4	1,5	0,9
16	7,3	4,9	4,1	5,4	5,5	6,4	6,7	4,7	3,9	3,5	4,7	4,1	2,3
17	24,1	16,4	14,4	17,8	19,5	19,9	18,5	11,6	9,8	11,0	11,5	8,1	4,3
18	59,5	43,5	41,1	43,8	48,8	46,1	37,9	26,9	20,2	23,1	19,7	14,7	.
19	118,3	90,3	86,4	90,3	96,4	90,0	64,0	44,0	31,4	35,3	27,2	21,9	.
20	163,8	140,6	133,8	136,5	140,8	131,7	81,3	55,5	40,8	43,4	35,1	27,1	.
21	190,7	174,0	168,7	176,4	170,8	156,0	96,4	60,8	50,1	46,7	42,7	31,2	.
22	201,9	182,8	183,6	203,3	185,5	156,3	101,8	66,6	58,4	49,2	49,8	35,5	.
23	196,9	192,1	191,0	206,2	184,5	147,1	101,1	72,1	63,4	57,3	57,7	.	.
24	185,5	177,9	185,2	193,4	171,6	138,4	100,9	78,7	71,7	62,8	67,6	.	.
25	164,5	165,7	168,2	173,4	155,9	122,0	99,2	83,9	78,0	73,1	79,0	.	.
26	147,6	149,2	151,1	148,1	135,7	112,1	93,9	91,3	84,6	85,8	87,9	.	.
27	128,6	128,9	141,8	131,7	117,2	101,4	90,2	97,7	91,0	102,2	94,6	.	.
28	121,1	116,1	125,7	110,6	100,6	86,0	88,7	99,5	92,5	107,1	.	.	.
29	106,3	102,7	106,4	93,6	87,2	75,6	83,9	98,1	93,5	106,5	.	.	.
30	91,9	88,3	90,0	79,3	71,1	67,4	79,2	90,7	92,9	102,5	.	.	.
31	79,7	76,1	72,9	68,4	62,6	57,8	74,7	78,7	87,7	93,0	.	.	.
32	66,4	69,9	62,6	57,2	54,3	50,2	70,5	69,7	89,2	80,0	.	.	.
33	56,6	60,9	50,7	49,6	44,1	44,2	61,9	61,4	77,7
34	49,6	51,1	41,7	42,0	37,5	38,8	54,9	54,4	67,5
35	40,5	41,2	35,7	33,4	30,5	33,5	45,9	46,7	56,4
36	35,7	32,0	28,9	27,5	25,4	29,3	36,1	37,2	44,7
37	29,9	26,2	23,8	22,2	20,4	26,4	28,9	35,7	33,9
38	23,7	19,5	20,5	16,8	16,8	20,7	22,6	28,0
39	19,8	15,0	15,6	13,0	13,2	16,4	17,5	21,8
40	13,3	11,3	10,4	9,1	9,6	12,0	12,8	15,6
41	8,9	8,1	7,2	6,5	6,8	8,1	8,0	10,2
42	5,8	5,3	4,8	3,9	5,1	5,2	5,9	6,2
43	3,5	3,6	2,5	2,4	2,8	3,1	3,7
44	1,8	1,8	1,4	1,4	1,6	1,6	2,0
45	0,9	0,7	0,7	0,7	0,8	0,8	1,0
46	0,4	0,3	0,3	0,3	0,4	0,3	0,5
47	0,2	0,1	0,1	0,2	0,1	0,2	0,3
48	0,1	0,0	0,1	0,1	0,1	0,0
49 lat	0,0	0,0	0,0	0,0	0,0	0,1
<i>CTFR(G)</i>	2,345	2,198	2,173	2,166	2,025	1,807	1,593	1,449	1,340	1,083	0,579	0,144	0,007

Źródło: opracowanie własne na podstawie danych z badań Tymickiego i in. (2018) oraz danych z Bazy Demografia.

Tablica 3 zawiera wartości wzdłużnych cząstkowych współczynników płodności dla generacji rzeczywistych 1965–2005 z uwzględnieniem szacunku dla ludności rezydującej od 2009 r., co było możliwe tylko dla niektórych lat wieku (zaznaczonych szarym

kolorem). Z uwagi na zbyt krótki okres obserwacji nie jest możliwe obliczenie $CTFR_R(G)$, czyli współczynnika dzietności bazującego na szacunku liczby kobiet rezydujących, w odniesieniu do żadnej z analizowanych kohort. Konsekwencje wynikające z przyjętej definicji ludności uwiadcniają się najwyraźniej w generacjach 1980 i 1985 r. W poniższej analizie, w miarę możliwości, starano się wprowadzić porównania uwzględniające sposób szacunku liczby ludności. Syntetycznie zostały one ujęte w tabl. 4.

Tabl. 3. Wzdłużne cząstkowe współczynniki płodności dla generacji rzeczywistych z uwzględnieniem szacunku ludności rezydującej od 2009 r.

Wiek	Urodzenia żywe na 1000 kobiet generacji								
	1965	1970	1975	1980	1985	1990	1995	2000	2005
15 lat	1,4	1,5	1,8	1,3	1,0	0,9	1,6	1,5	1,0
16	5,5	6,4	6,7	4,7	3,9	3,5	4,5	4,1	2,4
17	19,5	19,9	18,5	11,6	9,8	11,0	11,6	8,2	4,5
18	48,8	46,1	37,9	26,9	20,2	23,1	20,1	14,8	.
19	96,4	90,0	64,0	44,0	31,4	35,3	28,6	22,1	.
20	140,8	131,7	81,3	55,5	40,8	42,0	35,3	27,5	.
21	170,8	156,0	96,4	60,8	50,1	43,8	42,9	31,7	.
22	185,5	156,3	101,8	66,6	58,4	49,6	50,0	35,8	.
23	184,5	147,1	101,1	72,1	63,4	55,3	57,9	.	.
24	171,6	138,4	100,9	78,7	71,7	62,8	67,8	.	.
25	155,9	122,0	99,2	83,9	86,6	73,2	80,5	.	.
26	135,7	112,1	93,9	91,3	91,0	86,4	89,6	.	.
27	117,2	101,4	90,2	97,7	97,4	103,0	97,3	.	.
28	100,6	86,0	88,7	99,5	95,9	107,9	.	.	.
29	87,2	75,6	83,9	98,1	98,9	107,2	.	.	.
30	71,1	67,4	79,2	95,2	95,8	105,7	.	.	.
31	62,6	57,8	74,7	82,9	90,4	95,9	.	.	.
32	54,3	50,2	70,5	73,6	91,9	84,1	.	.	.
33	44,1	44,2	61,9	62,0	80,0
34	37,5	38,8	54,9	56,5	69,4
35	30,5	33,5	45,9	48,4	59,0
36	25,4	29,3	36,0	38,5	46,8
37	20,4	26,4	28,8	37,0	35,9
38	16,8	20,7	21,7	28,9
39	13,2	16,4	17,6	22,5
40	9,6	12,2	13,1	16,3
41	6,8	8,2	8,2	10,7
42	5,1	5,1	6,1	6,6
43	2,8	3,1	3,8
44	1,6	1,6	2,0
45	0,8	0,8	1,0
46	0,4	0,3	0,5
47	0,1	0,2	0,3
48	0,1	0,0
49 lat	0,1	0,1

Uwaga. Wartości współczynników, do obliczenia których wykorzystano szacunki ludności rezydującej, umieszczono na szarym tle.

Źródło: opracowanie własne na podstawie danych z badań Tymickiego i in. (2018) oraz danych z Bazy Demografia.

Tabl. 4. Wzdłużne współczynniki dzietności dla generacji rzeczywistych według krajowej definicji ludności i z uwzględnieniem szacunku ludności rezydującej od 2009 r.

Wyszczególnienie	Generacje rzeczywiste kobiet urodzonych w roku (G)								
	1965	1970	1975	1980	1985	1990	1995	2000	2005
$CTFR_K(G)$	2,025	1,807	1,593	1,449	1,340	1,083	0,579	0,144	0,007
$CTFR_R(G)$	2,025	1,807	1,593	1,472	1,390	1,091	0,588	0,146	0,008

Uwaga. Wytłuszczoną czcionką zaznaczono wartości $CTFR$, dla których różnice są największe.

Źródło: opracowanie własne na podstawie danych z badań Tymickiego i in. (2018) oraz danych z Bazy Demografia.

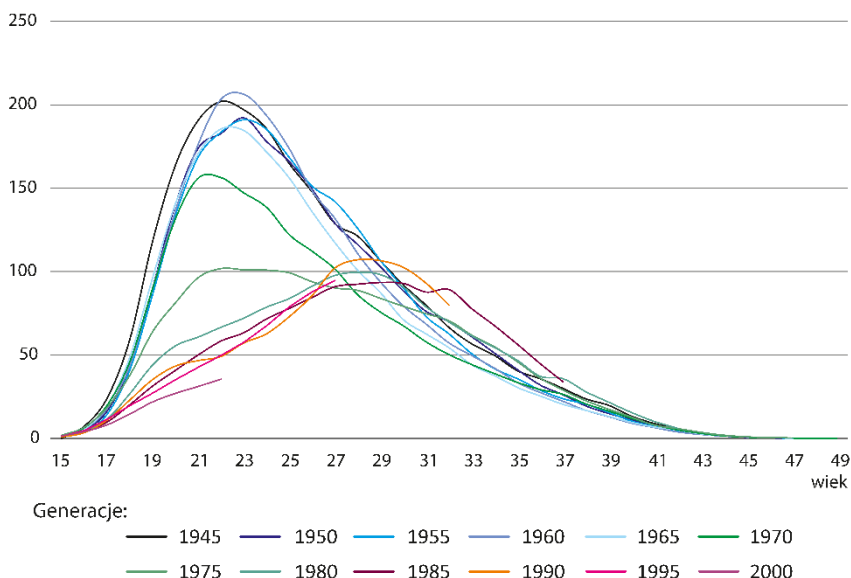
Krzywe płodności dla kobiet generacji 1945–1965 są silnie prawostronnie asymetryczne (wykr. 6). Wyraźne maksimum natężenia urodzeń dotyczy wieku 22–23 lata, a kohortowe współczynniki płodności przyjmują wartości powyżej 2, jednak obniżają się dla każdej kolejnej generacji (Tymicki i in., 2018). Dla kobiet urodzonych między 1945 r. a 1965 r. spadek dzietności wynosi $CTFR(1945) - CTFR(1965) = 2,345 - 2,025 = 0,32$.

Zmiany dotyczące kolejnych generacji są jeszcze bardziej znaczące. Wśród urodzonych w 1970 r. następuje wyraźny spadek płodności, chociaż kalendarz nie ulega zmianie. W tej generacji dla wieku 23 lata zanotowano spadek płodności o 37,4 urodzenia na 1000 kobiet w porównaniu z generacją 1965 r. Podobny, ok. 20-procentowy, spadek płodności dotyczy kolejnych rocznych grup wieku kobiet urodzonych w 1970 r. W rezultacie dzietność całkowita generacji 1970 r. jest o 0,22 niższa niż kobiet urodzonych w 1965 r.

Dla generacji 1975 r. zaobserwowano dalsze znaczne obniżenie płodności i pierwsze symptomy zmiany kalendarza. Polegały one na zniwelowaniu wyraźnego maksimum rozkładu wieku rodzenia dzieci i większym jego zróżnicowaniu. W grupie wieku 21–28 lat intensywność urodzeń była podobna i wynosiła od 89 do 102 urodzeń na 1000 kobiet.

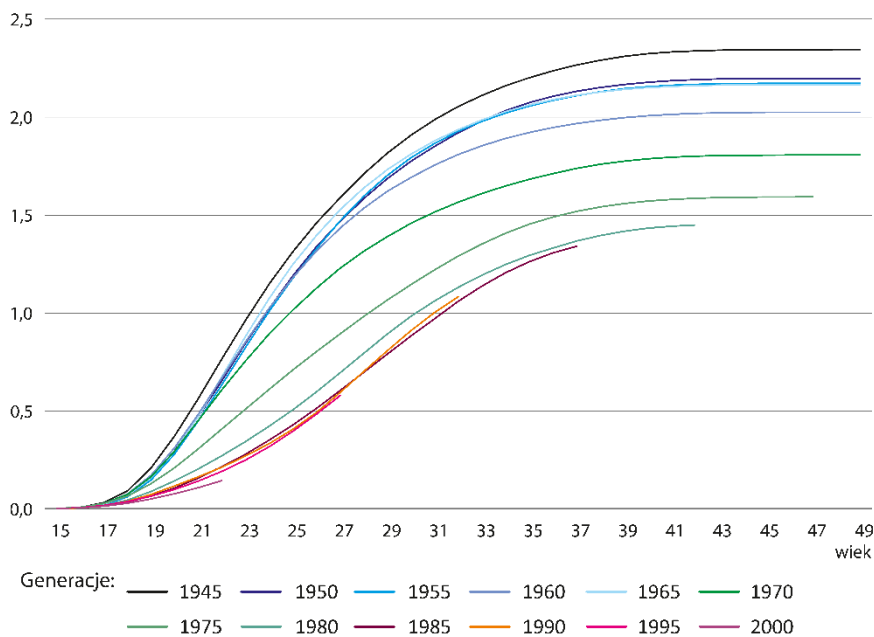
Generacja 1980 r. doświadczyła kolejnej znacznej zmiany natężenia i kalendarza płodności. Polegała ona przede wszystkim na wyraźnym obniżeniu płodności w grupie wieku 25–29 lat, a w efekcie – dalszym obniżeniu kohortowego współczynnika dzietności ($CTFR(1975) - CTFR(1980) = 1,593 - 1,449 = 0,144$). Od tej generacji krzywa płodności jest niemal symetryczna, a maksimum częstości urodzeń na 1000 kobiet przypada na wiek ok. 29–30 lat.

Wykr. 6. Kohortowe cząstkowe współczynniki płodności według wieku dla generacji rzeczywistych



Źródło: opracowanie własne na podstawie danych z badań Tymickiego i in. (2018) oraz danych z BazyDemografia.

W odniesieniu do generacji kobiet 1980 r. uwidacznia się też efekt sposobu szacowania liczby ludności. Uwzględnienie kobiet rezydujących w mianownikach kohortowych cząstkowych współczynników płodności dotyczy roczników wieku o intensywnej płodności. Dla wieku 30–32 lata oznacza to zwiększenie wartości współczynników o ponad 4 urodzenia na 1000 kobiet (powyżej 5%). W rezultacie współczynnik dzietności z uwzględnieniem szacunku liczby kobiet rezydujących wynosi 1,472 i jest o 0,023 wyższy niż w przypadku ludności według definicji krajowej. Nadal jednak – chociaż nieco mniej – natężenie płodności jest zaniżone w stosunku do kobiet urodzonych 5 lat wcześniej.

Wykr. 7. Kohortowe współczynniki dzietności dla generacji rzeczywistych

Źródło: opracowanie własne na podstawie danych z badań Tymickiego i in. (2018) oraz danych z Bazy Demografia.

Generacja 1985 r. i młodsze nie zakończyły jeszcze okresu prokreacyjnego, więc poniższe uwagi mają charakter porównawczy w stosunku do starszych generacji. W przypadku generacji 1985 r. zaobserwowano nieznaczne obniżenie płodności dla wieku do 30 lat, ale nieco wyższą płodność dla wieku do 38 lat – w porównaniu z generacją 1980 r. Do 2022 r. dzietność tej generacji kobiet wynosi 1,344 (dla kobiet rezydujących – 1,39). W generacji 1990 r. zaobserwowano wyższą płodność w wieku 27–32 lata w porównaniu z generacją 1985 r., ale po osiągnięciu maksimum w wieku 29 lat nastąpił gwałtowny spadek wartości współczynnika płodności. Dzietność w wieku do 30 lat dla generacji 1990 r. jest nieco niższa niż w starszych generacjach. Wartości współczynników dzietności z uwzględnieniem rezydentek są o blisko 0,01 wyższe.

Również kobiety urodzone w 1995 r. jeszcze nie zakończyły okresu prokreacyjnego. Dzietność w tej generacji do wieku 27 lat jest na poziomie 0,579, tj. o 0,015 niższym niż dla generacji 1990 r. i o 0,116 niż dla generacji 1980 r. Kobiety urodzone w 1995 r. właśnie rozpoczynają okres największego natężenia płodności, ale przytoczone dane nie dają podstaw, by przewidywać płodność wyższą niż obserwowana w młodszych generacjach. Podobnie trudno ocenić płodność wśród kobiet urodzonych w 2000 r.,

które w momencie analizy miały 22 lata. Jednak cząstkowe współczynniki płodności dla wieku 20–22 lata w tej generacji są o 23–29% niższe niż dla kobiet o 5 lat starszych. Dla wieku 22 lata cząstkowy współczynnik płodności wynosi 35,5 urodzenia na 1000 kobiet w porównaniu z 49,8 dla generacji 1995 r. Uwzględnienie ludności rezydującej w mianowniku kohortowych cząstkowych współczynników płodności nieznacznie łagodzi to obniżenie płodności.

Wpływ przyjętej definicji ludności na uzyskane rezultaty najwyraźniej uwidocznił się w przypadku kohorty 1985 r. (tabl. 4). Wynika to z tego, że w kohortowych cząstkowych współczynnikach płodności ludność rezydującą uwzględniono dla wieku o dużej intensywności urodzeń, tj. 25–37 lat. Tylko dla kobiet generacji 1985 r. będących w wieku 25 lat kohortowy cząstkowy współczynnik płodności rezydentek jest o 11% wyższy niż w ujęciu ludności według definicji krajowej.

7. Podsumowanie

W artykule przedstawiono wyniki analizy płodności kobiet w Polsce z uwzględnieniem szacunku liczby ludności według definicji krajowej i liczby ludności rezydującej. Analizowano szacunki płodności kobiet według wieku i w ujęciu terytorialnym, w przekroju województw. Płodność kobiet poddano analizie przekrojowej i wzdłużnej.

Zwrócono uwagę na potrzebę przestrzegania zasady konstrukcji współczynników, czyli odnoszenia zdarzeń demograficznych do średniej liczby ludności mogącej doświadczyć tych zdarzeń. Zgodnie z tym podejściem uznano, że ludność rezydująca stanowi lepszą podstawę nie tylko porównań międzynarodowych, lecz także badania procesów demograficznych i społecznych w Polsce, w szczególności w ujęciu regionalnym i lokalnym.

W ramach badania przeliczono wszystkie miary płodności w odniesieniu do ludności rezydującej. Wykazano, że współczynnik dzietności kobiet w Polsce w 2021 r. obliczany dla ludności według definicji krajowej był o 4,3% niższy niż dla ludności rezydującej. Największe zaniżenie szacunku dotyczy kobiet w wieku o największej intensywności urodzeń i wiąże się z uwarunkowaniami historycznymi dotyczącymi migracji. Płodność kobiet w wieku 35–39 lat oszacowana dla ludności według definicji krajowej jest zaniżona o 6,3%. Analiza zróżnicowania terytorialnego w przekroju województw wykazała największe zaniżenie współczynnika płodności kobiet zamieszkujących woj. podkarpackie (11,0%).

Omówione wyniki analizy wzdłużnej wskazują na znaczny spadek płodności wśród kobiet młodszych generacji. Uwzględnienie ludności rezydującej nieznacznie zmniejsza szacunek spadku płodności w kolejnych generacjach kobiet, ale go nie niweluje. Dostępne od 2009 r. szacunki ludności rezydującej dotyczą zbyt krótkiego okresu, by

móc przedstawić różnice w pomiarze dzietności dla którejkolwiek z badanych generacji rzeczywistych.

Analiza wzdłużna została przeprowadzona dla ludności według definicji krajowej. W 2004 r., w momencie wejścia Polski do UE, kobiety urodzone w 1980 r. miały 24 lata. Wówczas też zwiększyła się intensywność emigracji. Dlatego dla generacji 1980 i 1985 r., pomimo niezakończonego okresu prokreacji, różnice dotyczące wartości kohortowych cząstkowych współczynników płodności są najwyraźniejsze. Od 2016 r. Polska stała się także ważnym krajem docelowym migracji. Uwzględnienie nie tylko emigracji, lecz także imigracji stanowi poważne wyzwanie dla szacunku liczby ludności rezydującej.

Ustalenie właściwego stanu ludności, który powinien stanowić podstawę odwołań w analizie procesów demograficznych, jest dyskusyjne. Przede wszystkim pojawia się pytanie o osoby, które wyjechały za granicę, ale jeszcze się nie wymeldowały, nie podjęły decyzji o powrocie i nie wiadomo, czy podejmą ją w przyszłości. Zadaniem statystyki publicznej jest zaś dokonanie pomiaru jak najbliższego stanu rzeczywistego w danym momencie.

Bibliografia

- Abramowska-Kmon, A., Batorski, D., Cześniak, M., Marczewski, P., Rakocy, K., Salamońska, J., Turlejski, M., Wojdat, M. (2023). *Czy wiemy, ile nas jest? O potrzebie reformy sposobu pomiaru liczebności i struktury populacji*. Fundacja im. Stefana Batorego, Unia Metropolii Polskich im. Pawła Adamowicza. https://www.batory.org.pl/wp-content/uploads/2023/06/Czy_wiemy_ile_nas_jest_raport.pdf.
- Bijak, J., Kicinger, A., Kupiszewski, M. (2007). *Studium metodologiczne oszacowania rzeczywistej liczby ludności Warszawy* (CEFMR Working Paper No. 2/2007). http://www.cefmr.pan.pl/docs/cefmr_wp_2007-02.pdf.
- Eurostat. (b.r.). *Fertility indicators*. https://ec.europa.eu/eurostat/databrowser/view/DEMO_FIND__custom_7013627/default/table?lang=en.
- Fihel, A. (2015). Wpływ czasowych migracji zagranicznych na perspektywy demograficzne Polski. *Wiadomości Statystyczne*, 60(7), 74–89. <https://doi.org/10.5604/01.3001.0015.5550>.
- Gołata, E. (2012). Spis ludności i prawda. *Studia Demograficzne*, (1), 23–55. <https://econjournals.sgh.waw.pl/SD/article/view/2556/2215>.
- Gołata, E. (2013). Ludność rezydująca a faktyczna. W: P. Ulman (red.), *Zjawiska społeczne w badaniach statystycznych. Księga jubileuszowa dedykowana Profesor Barbarze Podolec* (s. 69–84). Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Gołata, E. (2014). Jakość spisów nowej generacji. *Wiadomości Statystyczne*, 59(5), 26–38. <https://doi.org/10.59139/ws.2014.05.3>.
- Gołata, E. (2016a). Estimation of fertility in Poland and of Polish born women in the United Kingdom. *Studia Demograficzne*, (1), 13–38. <https://doi.org/10.33119/SD.2016.1.1>.
- Gołata, E. (2016b). Shift in methodology and population census quality. *Statistics in Transition new series*, 17(4), 631–658. <https://doi.org/10.59170/stattrans-2016-035>.

- Gołata, E. (2018). *Koniec ery tradycyjnych spisów ludności*. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- Gołata, E. (red.). (2022). *Członkowie Komitetu Nauk Demograficznych Polskiej Akademii Nauk wyjaśniają zmianę demograficzną w Polsce*. https://knd.pan.pl/images/Panel12.12.2022_1.pdf.
- Główny Urząd Statystyczny. (b.r.). *Informacje o wynikach Narodowego Spisu Powszechnego Ludności i Mieszkań 2021*. <https://stat.gov.pl/aktualnosci/informacje-o-wynikach-narodowego-spisu-powszechnego-ludnosci-i-mieszkan-2021,422,1.html>.
- Główny Urząd Statystyczny. (2015). *Informacja o rozmiarach i kierunkach czasowej emigracji z Polski w latach 2004–2014*. <https://stat.gov.pl/obszary-tematyczne/ludnosc/migracje-zagraniczne-ludnosci/informacja-o-rozmiarach-i-kierunkach-emigracji-z-polski-w-latach-20042014,2,8.html>.
- Główny Urząd Statystyczny. (2016). *Zakres danych dotyczących stanu i struktury ludności oraz ruchu naturalnego, a także ich dostępność w świetle Rozporządzenia Parlamentu Europejskiego i Rady (UE) nr 1260/2013 w sprawie statystyk europejskich w dziedzinie demografii*. <https://stat.gov.pl/obszary-tematyczne/ludnosc/ludnosc/baza-eurostat-u-w-zakresie-statystyki-demografii,25,1.html>.
- Główny Urząd Statystyczny. (2022a). *Narodowy Spis Powszechny Ludności i Mieszkań 2021. Metodologia i organizacja badania*. <https://stat.gov.pl/spisy-powszechne/nsp-2021/nsp-2021-wyniki-wstepne/narodowy-spis-powszechny-ludnosci-i-mieszkan-2021-metodologia-i-organizacja-badania-,5,2.html>.
- Główny Urząd Statystyczny. (2022b). *Rezydenci (ludność rezydująca)*. Pobrano 1 sierpnia 2023 r. z <https://stat.gov.pl/obszary-tematyczne/ludnosc/ludnosc/rezydenci-ludnosc-rezydujaca,19,1.html>.
- Główny Urząd Statystyczny. (2022c). *Rocznik Demograficzny 2022*. <https://stat.gov.pl/obszary-tematyczne/roczniki-statystyczne/roczniki-statystyczne/rocznik-demograficzny-2022,3,16.html>.
- Główny Urząd Statystyczny. (2023). *Narodowy Spis Powszechny Ludności i Mieszkań 2021. Ocena jakości danych*. <https://stat.gov.pl/spisy-powszechne/nsp-2021/nsp-2021-wyniki-ostateczne/narodowy-spis-powszechny-ludnosci-i-mieszkan-2021-ocena-jakosci-danych,2,1.html>.
- Holzer, J. Z. (2003). *Demografia* (wyd. 6 zmienione). Polskie Wydawnictwo Ekonomiczne.
- Jończy, R. (2015). *Ekonomiczno-społeczne skutki współczesnych migracji w wymiarze regionalnym – na przykładzie regionu opolskiego*. Wydawnictwo Instytut Śląski.
- Jończy, R. (2023). *Migracje jako przyczyna wyludniania w skali regionalnej*. W: B. Solga (red.), *Migracje i rozwój regionu. Materiały z III Kongresu Demograficznego. Część 7* (s. 53–66). Główny Urząd Statystyczny, Rządowa Rada Ludnościowa.
- Kędełski, M. (1990). *Fikcja demograficzna w Polsce i RFN (ze studiów nad migracjami zagranicznymi)*. *Studia Demograficzne*, (1), 21–55.
- Kędełski, M., Paradysz, J. (1984). *Uwagi o kolejnych edycjach Rocznika Demograficznego*. *Wiadomości Statystyczne*, 29(5), 41–43.
- Kędełski, M., Paradysz, J. (2006). *Demografia*. Wydawnictwo Akademii Ekonomicznej w Poznaniu.
- Komitet Nauk Demograficznych Polskiej Akademii Nauk. (2022). *Przesłanie Komitetu Naukowego konferencji „Zrozumieć zmianę demograficzną”*. https://knd.pan.pl/images/Przeslanie_Komitetu_Naukowego_konferencji_Zrozumiec_zmiane_demograficzna_ost.pdf.
- Kurkiewicz, J. (red.). (2010). *Procesy demograficzne i metody ich analizy*. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Paradysz, J. (2007a). *Recenzja opracowania „Studium metodologiczne oszacowania rzeczywistej liczby ludności Warszawy”, przygotowanego w Środkowoeuropejskim Forum Badań Migracyjnych przez zespół w składzie: Jakub Bijak, Anna Kicingier, Marek Kupiszewski (maszynopis)*.

- Paradysz, J. (2007b). Rejestry administracyjne jako źródło zasilania w statystyce regionalnej. W: J. Paradysz (red.), *Statystyka regionalna w jednoczącej się Europie* (s. 193–206). Internetowa Oficyna Wydawnicza Centrum Statystyki Regionalnej.
- Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 1260/2013 z dnia 20 listopada 2013 r. w sprawie statystyk europejskich w dziedzinie demografii (Dz.U. UE L330/39).
- Sakson, B. (2002). *Wpływ „niewidzialnych” migracji zagranicznych lat osiemdziesiątych na struktury demograficzne Polski*. Szkoła Główna Handlowa.
- Śleszyński, P. (2004a). Regionalne różnice pomiędzy liczbą ludności według Narodowego Spisu Powszechnego w 2002 r. i rejestrowaną na podstawie ewidencji bieżącej. *Studia Demograficzne*, (1), 93–103.
- Śleszyński, P. (2004b). Różnice liczby ludności wykazane w NSP 2002 – suplement. *Studia Demograficzne*, (2), 104–109.
- Śleszyński, P. (2005). Różnice w spisie ludności ujawnione w Narodowym Spisie Powszechnym 2002. *Przegląd Geograficzny*, 77(2), 193–212.
- Tymicki, K., Zeman, K., Holzer-Żelazewska, D. (2018). Cohort fertility of Polish women, 1945–2015: the context of postponement and recuperation. *Studia Demograficzne*, (2), 5–23. <https://doi.org/10.33119/SD.2018.2.1>.
- Tymicki, K., Zeman, K. (2022). *Human Fertility Database Documentation: Poland*. <https://www.humanfertility.org/File/GetDocumentFree/Docs/POL/POLcom.pdf>.

Ocena zależności między rozwojem branży leasingowej a wzrostem gospodarczym w Polsce

Tomasz Banasik^a

Streszczenie. Z danych opublikowanych w 2022 r. przez Polską Agencję Rozwoju Przedsiębiorczości wynika, że małe i średnie przedsiębiorstwa (MŚP) wytwarzają 49,6% PKB Polski. Jednocześnie sektor MŚP jest największym odbiorcą usług leasingowych. Badanie omawiane w artykule ma na celu określenie, czy występuje długookresowa zależność między rozwojem branży leasingowej a wzrostem gospodarczym w Polsce. W analizie wykorzystano szeregi czasowe wartości PKB per capita i umów leasingowych netto w przeliczeniu na mieszkańca – odpowiednio jako miary wzrostu gospodarczego i aktywności branży leasingowej. Posłużono się danymi kwartalnymi za lata 2010–2021 uzyskanymi z publikacji Głównego Urzędu Statystycznego i statystykami zamieszczonymi na stronie Związku Polskiego Leasingu. Wyniki analizy kointegracji i modelu korekty błędem wskazały na występowanie długookresowej zależności między rozwojem branży leasingowej a wzrostem gospodarczym w Polsce.

Słowa kluczowe: wzrost gospodarczy, małe i średnie przedsiębiorstwa, MŚP, modelowanie szeregów czasowych, model korekty błędem ECM, analiza kointegracji

JEL: O47, D02, C22, C51

Assessment of the relationship between the development of the leasing industry and economic growth in Poland

Abstract. According to the data published in 2022 by the Polish Agency of Enterprise Development (Polska Agencja Rozwoju Przedsiębiorczości – PARP), small and medium-sized enterprises (SMEs) generate 49.6% of the Polish GDP. At the same time, the SME sector is the largest recipient of services offered by the leasing industry. The aim of the research described in the article was to determine whether a long-run relationship exists between the development of the leasing industry and the economic growth in Poland. The analysis was based on GDP *per capita* time series and the actual net value of leasing contracts *per capita* as measures of economic growth and the activity of the leasing industry, respectively. The research used quarterly data for the years 2010–2021, obtained from Statistics Poland publications and the dataset available on the Polish Leasing Association (Związek Polskiego Leasingu – ZPL) website. The results of cointegration analysis and the error correction mechanism confirmed the existence of a long-run relationship between the activity of the leasing industry and the economic growth in Poland.

Keywords: economic growth, small and medium enterprise, SME, modelling time series, error correction mechanism, ECM, cointegration analysis

^a Politechnika Świętokrzyska, Wydział Zarządzania i Modelowania Komputerowego, Polska / Kielce University of Technology, Faculty of Management and Computer Modelling, Poland.
ORCID: <https://orcid.org/0000-0002-5646-829X>. E-mail: skarb@tu.kielce.pl.

1. Wprowadzenie

Rozważania na temat determinantów wzrostu gospodarczego, prowadzone z uwzględnieniem teoretycznych modeli wzrostu gospodarczego, nasuwają wniosek, że jego głównym czynnikiem są inwestycje (Roszkowska, 2004). W przypadku Polski stanowią one kluczowy czynnik wzrostu PKB. Są najskuteczniejszym mechanizmem przygotowania przedsiębiorstw na wyzwania czekające krajową gospodarkę w najbliższej przyszłości, takie jak wzrost kosztów pracy, deficyt wykwalifikowanej kadry, szeroko rozumiana transformacja energetyczna i wykorzystanie sztucznej inteligencji do usprawnienia procesu produkcji (EY-Parthenon i Związek Polskiego Leasingu [ZPL], 2024). Aby im sprostać, firmy będą musiały korzystać z zewnętrznych źródeł finansowania działalności inwestycyjnej.

Jedną z kluczowych form finansowania inwestycji jest leasing – alternatywa dla kredytów bankowych i zakupów gotówkowych. W Polsce branża leasingowa rozwija się od ponad 30 lat, a całkowita wartość aktywnego portfela firm leasingowych przewyższyła w ostatnich 6 latach wartość kredytów inwestycyjnych udzielonych przedsiębiorstwom (Radło i in., 2022). Szacowany na koniec 2022 r. udział leasingu w nakładach inwestycyjnych podmiotów gospodarczych w Polsce kształtował się na poziomie 26,8% (EY-Parthenon i ZPL, 2024). Głównymi odbiorcami usług leasingowych były następujące działy gospodarki: transport, handel hurtowy, przetwórstwo przemysłowe i budownictwo. Ważną rolę odgrywa w nich sektor małych i średnich przedsiębiorstw (MŚP). Z danych opublikowanych przez Polską Agencję Rozwoju Przedsiębiorczości (2022) wynika, że MŚP wytwarzają 49,6% PKB Polski. Warto podkreślić, że w Polsce leasing jest – w porównaniu z innymi krajami europejskimi – usługą bardzo popularną, o czym świadczy wartość wskaźnika penetracji rynku leasingu w odniesieniu do wartości PKB¹. W 2022 r. poziom tego wskaźnika dla naszego kraju wyniósł 3,4% PKB, co było czwartym wynikiem w Europie (Główny Urząd Statystyczny [GUS], 2023)². Warto zatem ustalić, jak kształtowała się zależność między rozwojem branży leasingowej a zmianami PKB.

Należy zaznaczyć, że w artykule nie zawarto informacji o historii i rodzajach leasingu, jego aspektach podatkowych i bilansowych oraz wadach i zaletach tego narzędzia dla każdej ze stron umowy, natomiast przywołano aktualną literaturę przedmiotu dotyczącą tych zagadnień. Podkreślono znaczenie branży leasingowej w finansowaniu polskich przedsiębiorstw, a także jej udziału w europejskim rynku usług leasingowych. Celem badania omawianego w artykule jest określenie, czy występuje długookresowa zależność między rozwojem branży leasingowej a wzrostem gospodarczym w Polsce.

¹ Wartość nowych umów leasingowych podpisanych w ciągu roku jako procent PKB kraju.

² Wyższy poziom wskaźnika zanotowano w Estonii (3,9% PKB), na Litwie (3,7% PKB) i w Wielkiej Brytanii (3,5% PKB).

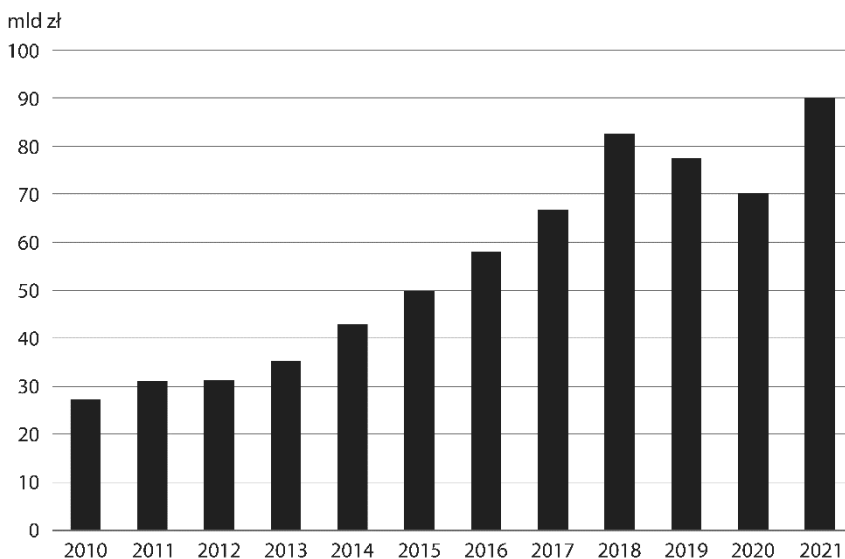
Analiza obejmuje lata 2010–2021, czyli okres względnie szybkiego wzrostu gospodarczego i jednocześnie dynamicznego rozwoju branży leasingowej.

2. Rola branży leasingowej w finansowaniu działalności gospodarczej w Polsce

Leasing wspiera działalność gospodarczą, ponieważ umożliwia nabycie nieruchomości i ruchomych składników majątku trwałego niezbędnych w procesie rozwoju firmy i budowania konkurencyjnej pozycji na rynku. W efekcie może się przyczyniać do unowocześnień gospodarki dzięki wprowadzeniu na rynek najnowszych technologii.

Pierwszą umowę leasingową podpisano w Polsce na początku lat 90. XX w (Price-waterhouseCoopers [PwC] i ZPL, 2019). Do 2000 r. leasing realizowano w formie umów cywilnoprawnych nieuregulowanych w kodeksie cywilnym. W przepisach podatkowych był traktowany jako działalność usługowa prowadzona na podstawie umowy o charakterze podobnym do umowy najmu lub dzierżawy. Dopiero wraz z wejściem w życie Ustawy z dnia 26 lipca 2000 r. o zmianie ustawy – Kodeks cywilny umowa leasingu stała się umową nazwaną w polskim prawie cywilnym, co miało fundamentalne znaczenie dla rozwoju całej branży leasingowej i finansowania działalności gospodarczej. Udział branży leasingowej w finansowaniu inwestycji gospodarczych w Polsce ilustrują dane przedstawione na wyk. 1.

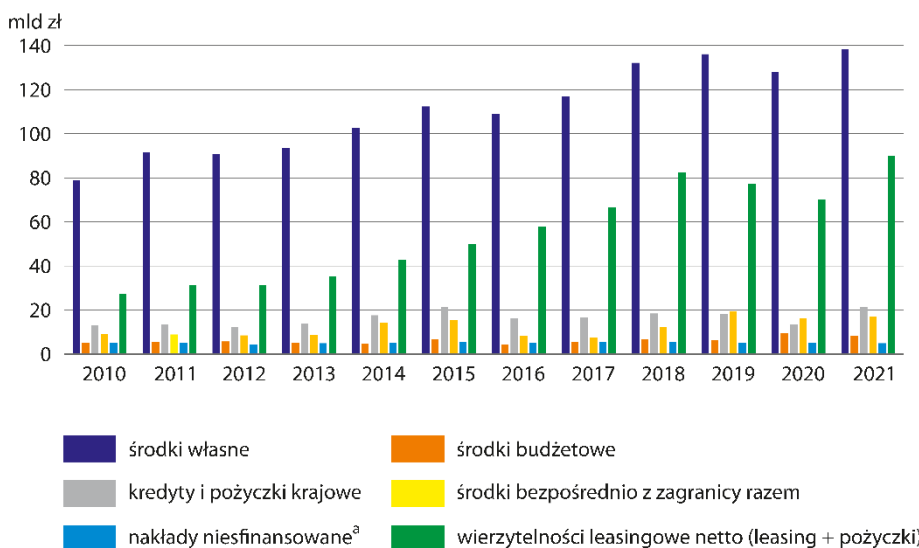
Wykr. 1. Wartość finansowania udzielonego przez branżę leasingową w Polsce



Źródło: opracowanie własne na podstawie: ZPL (2011–2020, 2021a, 2022a).

Do 2018 r. następował systematyczny wzrost wartości finansowania udzielanego przez branżę leasingową. Ten trend odwrócił się w latach 2019–2020. Przyczyną spadku finansowania była przede wszystkim zmiana przepisów o podatku dochodowym, która weszła w życie 1 stycznia 2019 r. Mniej korzystny sposób zaliczania rat leasingowych do kosztów uzyskania przychodu implikował zmniejszenie zainteresowania leasingiem operacyjnym samochodów osobowych o wartości powyżej 150 tys. zł (Krupiński, 2020). Kolejny spadek wartości finansowania w 2020 r. (o blisko 10% w stosunku do 2019 r.) był związany z wybuchem pandemii COVID-19 (ZPL, 2021b). Ograniczenie możliwości prowadzenia działalności gospodarczej spowodowało, że leasingobiorcy rezygnowali z inwestycji i wycofywali wnioski o finansowanie przedmiotu leasingu, co skutkowało spadkiem wartości i liczby nowych umów leasingowych. Koniunktura na rynku usług leasingowych poprawiła się w 2021 r. – wartość zawartych wówczas umów tego typu kształtowała się na rekordowo wysokim poziomie 90 mld zł (wynik o 28% wyższy niż w poprzednim roku). Rok 2021 okazał się szczególny dla branży leasingowej, ponieważ – jak wynika z raportu ZPL (2022b) – całkowita wartość aktywnego portfela wierzytelności leasingowych wyniosła 165,8 mld zł, a wartość kredytów inwestycyjnych udzielanych firmom przez banki ukształtowała się na poziomie 159 mld zł. Zwiększenie roli leasingu jako popularnego źródła finansowania nakładów inwestycyjnych w przedsiębiorstwach potwierdzają dane przedstawione na wyk. 2.

Wykr. 2. Źródła finansowania nakładów inwestycyjnych przez polskie przedsiębiorstwa



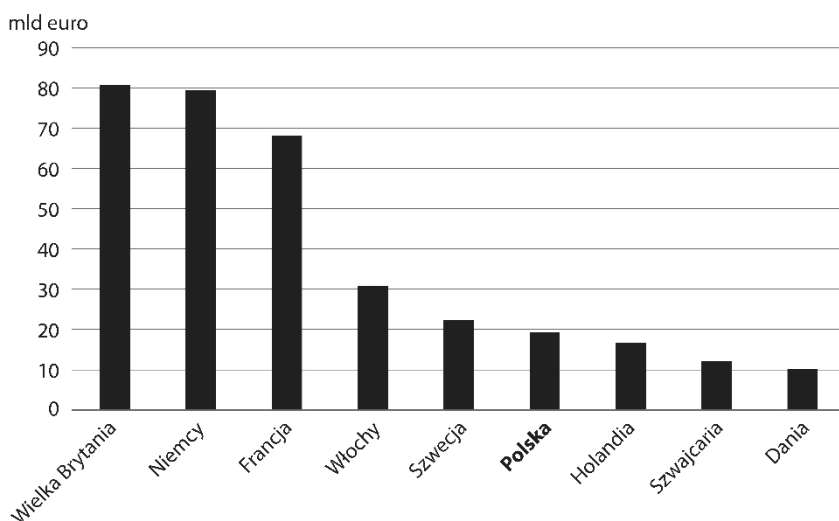
a Nakłady, w przypadku których w danym roku przedsiębiorstwa nie otrzymały finansowania (np. niezapłacone faktury, nieotrzymane dotacje lub transze kredytów).

Źródło: opracowanie własne na podstawie danych z Banku Danych Lokalnych GUS i ZPL (2011–2020, 2021a, 2022a).

Z danych przedstawionych na wyk. 2 wynika, że głównym źródłem finansowania nakładów inwestycyjnych przedsiębiorstw w latach 2010–2021 były środki własne inwestorów. Na koniec 2021 r. ich udział w finansowaniu ogółem stanowił 49,4%. Na drugim miejscu uplasował się leasing. Udział wierzytelności leasingowych netto w finansowaniu ogółem stanowił 32,2% i był o 24,6 p.proc. wyższy od udziału kolejnego źródła finansowania, czyli kredytu bankowego. Można na tej podstawie stwierdzić, że leasing jest obecnie najpopularniejszym źródłem kapitału obcego wykorzystywanego w celu finansowania działalności gospodarczej.

Omawiając znaczenie rynku usług leasingowych dla polskiej gospodarki, warto się przyjrzeć, jak rynek krajowy prezentuje się na tle rynków europejskich (wykr. 3).

Wykr. 3. Wartość portfela wierzytelności obsługiwanych przez branżę leasingową w Polsce w 2021 r. na tle wybranych krajów europejskich



Źródło: opracowanie własne na podstawie: Leaseurope (2022).

Według informacji Leaseurope, instytucji zrzeszającej 46 organizacji reprezentujących firmy leasingowe z 35 krajów, na koniec 2021 r. wartość wierzytelności obsługiwanych przez działające w Europie firmy leasingowe wyniosła 404,44 mld euro³ (Leaseurope, 2022). Z danych przedstawionych na wyk. 3 wynika, że największymi udziałami w europejskim rynku usług leasingowych wykazali się leasingodawcy z Wielkiej

³ W cytowanym raporcie sklasyfikowano 25 krajów oraz utworzone przez Bank Al-Maghrib z Maroka stowarzyszenie Association Professionnelle des Sociétés de Financement (APSF). Autorzy raportu zastrzegają, że dane dla krajów, w których funkcjonuje więcej niż jedna organizacja zrzeszająca firmy leasingowe, mogą się w pewnym stopniu różnić (np. Wielka Brytania, Francja i Holandia).

Brytanii, Niemiec i Francji – odpowiednio 19,97%, 19,65% i 16,83%. Polskie firmy leasingowe obsługiwały wierzytelności w kwocie 19,28 mld euro, co stanowiło 4,77% udziału w rynku (szóste miejsce w zestawieniu Leaseurope, 2022)⁴. Uprawnia to do stwierdzenia, że Polska jest ważnym uczestnikiem europejskiego rynku usług leasingowych.

3. Przegląd literatury

W literaturze ekonomicznej istnieje niewiele opracowań poświęconych zależnościom występującym między leasingiem a wzrostem gospodarczym. W jednym z najnowszych (Lleshaj i Kripa, 2021) omówiono wyniki badań z zastosowaniem modelu regresji progowej, których przedmiotem była ocena oddziaływania sektora leasingu na rozwój systemu finansowego i wzrost gospodarczy Albanii w latach 2008–2020. Za zmienną objaśnianą w modelu przyjęto poziom PKB, a zmiennymi objaśniającymi były: leasing finansowy (zmienna progowa), kredyty bankowe udzielone przedsiębiorstwom i osobom indywidualnym, udział finansowania typu prywatnego w ogólnej sumie kredytów udzielonych przez krajowe instytucje finansowe, agregat monetarny M2 oraz depozyty (terminowe i bezterminowe) w systemie bankowym. Oszacowano dynamiczny model progowy, w którym dla zmiennej leasing finansowy (mierzony jako procent PKB) przyjęto próg 0,16%. Autorzy dowodzą, że wzrost znaczenia leasingu jako źródła finansowania działalności gospodarczej wpływa na zwiększenie konkurencji w systemie pośrednictwa finansowego Albanii, co z kolei prowadzi do spadku ryzyka nieuregulowania zobowiązań i obniżki kosztów transakcyjnych związanych z funkcjonowaniem systemu finansowego. Co ważne, wymienione zależności przyczyniają się do wzrostu gospodarczego.

Vakhitov i Zamaletdinov (2015) przeprowadzili badanie, którego celem była weryfikacja hipotezy zakładającej, że leasing jest narzędziem stymulującym wzrost PKB. Stwierdzili, że leasing – finansujący wymianę wyposażenia używanego w procesie produkcji – oddziałuje na materiałochłonność i energochłonność na wszystkich etapach tego procesu. Wymiana wyposażenia na nowsze wiąże się często z oszczędzaniem zasobów materiałowych i energii, co zwiększa wydajność produkcji w przedsiębiorstwie, a w ujęciu makroekonomicznym – pozytywnie wpływa na wzrost gospodarczy. Autorzy wybrali trzy wskaźniki: obrotu środkami trwałymi, energochłonności i materiałochłonności jako propozycje miar oddziaływania leasingu na poziom produkcji w przemyśle wytwórczym, rolnictwie, przemyśle budowlanym i branży remontowej, w których działalność gospodarcza jest w znacznym stopniu finansowana leasingiem.

⁴ Za 100% przyjęto wartość wierzytelności obsługiwanych przez firmy leasingowe działające w Europie (404,44 mld euro).

Haiss i Kichler (2009) wykonali badania obejmujące 10 krajów Europy Środkowej i Wschodniej określanych mianem *emerging economies*⁵. Wybrane do analizy dane statystyczne (w ujęciu rocznym) dotyczyły lat 1999–2006. Autorzy zastosowali funkcję produkcji, w której zmienną modelowaną (efektem produkcji) był PKB per capita, a zmienną objaśniającą – wartość obsługiwanych wierzytelności leasingowych w relacji do wartości PKB⁶. Wyniki analiz wskazują, że leasing przyczynił się do wzrostu PKB każdego badanego kraju. Zaobserwowano pozytywny – statystycznie istotny – wpływ wzrostu wartości wierzytelności leasingowych, który nastąpił dwa lata lub rok wcześniej, na wzrost wartości PKB per capita w roku bieżącym. Stąd wniosek, że pojawienie się pozytywnych efektów oddziaływania leasingu na gospodarkę wymaga czasu, ale są one niezaprzeczalne.

W polskiej literaturze ekonomicznej poruszano różne zagadnienia dotyczące leasingu, w szczególności regulacje umowy leasingu w polskim prawie na tle praktyki sądów powszechnych (Grochowski, 2016), a także wpływ uregulowań Międzynarodowego Standardu Sprawozdawczości Finansowej na strukturę aktywów i pasywów w sprawozdaniach finansowych spółek giełdowych (Hońko, 2016). Problematykę wykorzystania leasingu jako alternatywnego źródła finansowania działalności podjęto w pracach Nesterowicz i Nesterowicz (2020), Olkowskiej (2009) i Politowskiej-Iwaszko (2020). Ciekawą dyskusję na temat korzyści podatkowych płynących z wykorzystania leasingu do finansowania inwestycji rzeczowych przedsiębiorstw przedstawiono w artykule Białek-Jaworskiej i in. (2014).

Ilościowy opis relacji między rozwojem branży leasingowej a wzrostem PKB w Polsce można znaleźć jedynie w raporcie przygotowanym przez PwC we współpracy z ZPL. Przytoczono w nim wyniki analizy ekonometrycznej wykonanej w 2017 r. przez ekonomistów ze Szkoły Głównej Handlowej w Warszawie. Wynika z nich, że wzrost wartości nowych umów w branży leasingowej (z uwzględnieniem udzielonych pożyczek leasingowych) o 1% w danym kwartale przekłada się na wyższą nakładów inwestycyjnych na środki trwałe ogółem o 0,17% i wzrost kwartalnego PKB o 0,19% (PwC i ZPL, 2019).

⁵ Rynki wschodzące to kraje, których gospodarki charakteryzują się szybszym niż w krajach rozwiniętych tempem wzrostu gospodarczego napędzanego konsumpcją, postępującą industrializacją, niższymi kosztami pracy, a także słabiej rozwiniętą gospodarką.

⁶ Należy podkreślić, że badania empiryczne miały szerszy zakres i dotyczyły wpływu systemu pośrednictwa finansowego na wzrost gospodarczy. Z tego powodu wykorzystano również inne zmienne objaśniające, np. kredyt bankowy, finansowanie *private debt*, kapitalizację giełdową czy wierzytelności powstałe w wyniku emisji obligacji.

4. Metoda badania

4.1. Analiza kointegracyjna

Należy postawić pytanie o rodzaj zależności występującej między zmianami zachodzącymi w branży leasingowej i gospodarce, a zwłaszcza zastanowić się, czy ma ona charakter równowagi długookresowej.

W podjętym badaniu za miarę wzrostu gospodarczego przyjęto PKB w przeliczeniu na mieszkańca w mln zł (PKB_{pc}), a za miernik rozwoju branży leasingowej – sumę wartości zawartych umów leasingowych i udzielonych pożyczek leasingowych (wartość netto w mln zł), również w przeliczeniu na mieszkańca (WLN_{pc}). Dane dotyczące powyższych wskaźników zaczerpnięto z zasobów ZPL (2011–2020, 2021a, 2022a) i GUS (b.r.). Analizie poddano dane kwartalne za okres od pierwszego kwartału 2010 r. do czwartego kwartału 2021 r. Wybór 2010 r. jako początkowego pozwolił na pominięcie niestabilności i negatywnego wpływu kryzysu finansowego z lat 2008–2009 na wyniki przedsiębiorstw leasingowych w Polsce. Badanie polegało na weryfikacji hipotezy zakładającej występowanie długookresowej zależności między rozwojem branży leasingowej a wzrostem gospodarczym.

W sytuacji gdy pożądane jest uwzględnienie teorii ekonomicznej dotyczącej zarówno długookresowej relacji zmiennych, jak i krótkookresowej nierównowagi (Bednarczyk, 2012; Charemza i Deadman, 1997), zalecane jest stosowanie podejścia opartego na analizie kointegracji i mechanizmie korekty błędem (ang. *error correction mechanism* – ECM). Uważa się, że procesy ekonomiczne są skointegrowane, gdy zachodzi między nimi relacja równowagi długookresowej. Równowaga jest stanem systemu (określającym jego stabilność), w którym zmienna modelowana i zmienna objaśniająca nie oddalają się od siebie znacząco wraz z upływem czasu. Stabilność zależności między zmiennymi oznacza, że są one niezmiennie ze względu na charakter, kierunek i siłę. Jeżeli system zostanie wytrącony ze stanu równowagi w wyniku działania sił zewnętrznych (zaburzeń losowych), to powraca do stanu równowagi. Wartości znajdujące się poza stanem równowagi stanowią krótkookresowe – zależne od czasu – odchylenia od tego stanu (Osińska, 2006; Welfe, 2003).

4.2. Model korekty błędem

W celu określenia siły zależności długookresowej zastosowano jednorównaniowy model korekty błędem należący do klasy modeli dynamicznych opisujących proces dostosowania się zmiennej objaśnianej do relacji długookresowej. Analizę przeprowadzono zgodnie z twierdzeniem Grangera, które brzmi następująco (Charemza i Deadman, 1997):

Jeśli (y_t, x_t) są skointegrowane oraz y_t i x_t są $I(1)$, to y_t można przedstawić w postaci ECM:

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta x_t + \gamma \underbrace{(y_{t-1} - \beta x_{t-1})}_{ECM_{t-1}} + e_t, \quad (1)$$

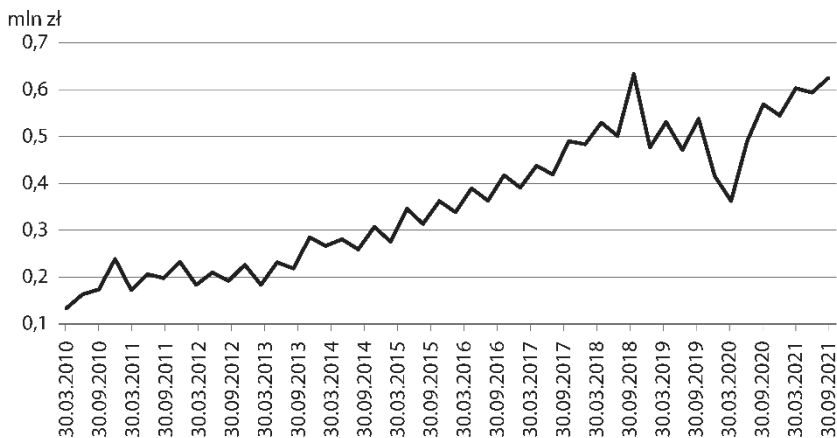
gdzie e_t oznacza składnik losowy.

Wyrażenie $y_{t-1} = \beta x_{t-1}$ należy rozumieć jako równowagę długookresową, a $(y_{t-1} - \beta x_{t-1})$ – jako odchylenie od równowagi. Współczynnik γ jest związany z szybkością dostosowania zmiennej y_t do poziomu równowagi, a współczynnik α_1 – z dynamiką krótkookresową.

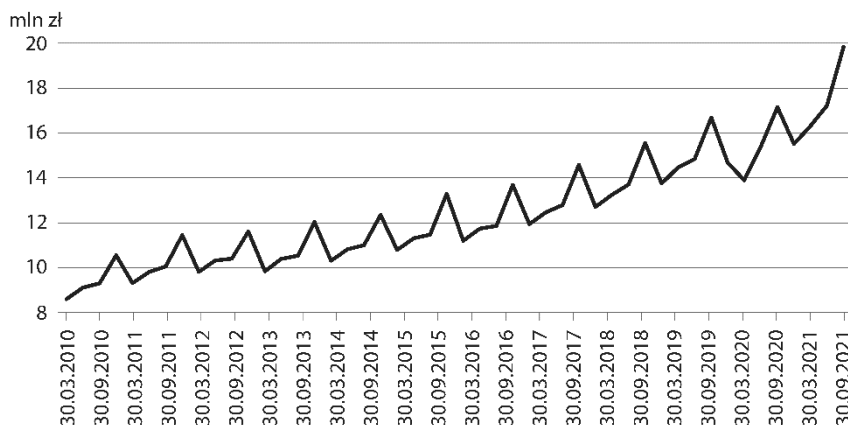
5. Wyniki badania

Analizę rozpoczęto od graficznego przedstawienia badanych szeregów WLN_{pc} i PKB_{pc} (wykr. 4 i 5).

Wykr. 4. Wartość netto zawartych umów leasingowych i udzielonych pożyczek leasingowych w przeliczeniu na mieszkańca



Źródło: opracowanie własne na podstawie: GUS (b.r.) i ZPL (2011, 2020, 2021a, 2022a).

Wykr. 5. Wartość PKB w przeliczeniu na mieszkańca

Źródło: opracowanie własne na podstawie: GUS (b.r.).

W związku z tym, że analizowane szeregi WLN_{pc_t} i PKB_{pc_t} podlegają sezonowym wahaniom kwartalnym oraz są niestacjonarne ze względu na trend, na kolejnym etapie badania podjęto próbę sprowadzenia ich do stacjonarności poprzez różnicowanie sezonowe – wyznaczając czwarte przyrosty analizowanych zmiennych: $\Delta_4 WLN_{pc_t}$, $\Delta_4 PKB_{pc_t}$ i ich pierwsze różnice: $\Delta \Delta_4 WLN_{pc_t}$, $\Delta \Delta_4 PKB_{pc_t}$.

Do badania stacjonarności szeregów zastosowano rozszerzony test Dickeya-Fullera (ADF) i test KPSS (Charemza i Deadman, 1997, s. 113). Wyniki uzyskane na poziomie istotności 0,05 zebrano w tabl. 1.

Tabl. 1. Analiza stacjonarności badanych szeregów

Zmienne	Test KPSS	Test ADF
WLN_{pc}	0,082 (>0,1)	-3,252 (0,070)
PKB_{pc}	0,283 (<0,01)	0,394 (0,999)
$\Delta_4 WLN_{pc}$	0,072 (>0,1)	-0,840 (0,961)
$\Delta_4 PKB_{pc}$	0,096 (>0,1)	-2,603 (0,279)
$\Delta \Delta_4 WLN_{pc}$	0,037 (>0,1)	-5,757 (0,000)
$\Delta \Delta_4 PKB_{pc}$	0,061 (>0,1)	-3,523 (0,037)

Uwaga. W nawiasach podano wartości p .

Źródło: obliczenia własne na podstawie: GUS (b.r.) i ZPL (2011–2020, 2021a, 2022a).

W przypadku badania pierwszych różnic czwartych przyrostów, $\Delta \Delta_4 WLN_{pc_t}$ i $\Delta \Delta_4 PKB_{pc_t}$ wyniki testu ADF zostały potwierdzone przez wyniki testu KPSS. Można zatem przyjąć, że obydwa badane szeregi, czyli WLN_{pc_t} i PKB_{pc_t} , są sezonowo zintegrowane: $WLN_{pc_t} \sim SI_4(1, 1)$ i $PKB_{pc_t} \sim SI_4(1, 1)$.

Zgodnie z dwustopniową procedurą Engla-Grangera dwa szeregi są skointegrowane, jeżeli można znaleźć taką ich kombinację liniową, która będzie stacjonarna. Ponieważ szeregi czwartych przyrostów wartości PKB per capita i czwartych przyrostów sumy wartości zawartych umów leasingowych i udzielonych pożyczek leasingowych są zintegrowane na poziomie 1, oszacowano relację długookresową (tabl. 2A).

Tabl. 2. Wyniki estymacji klasyczną metodą najmniejszych kwadratów równania kointegrującego oraz modeli korekty błędem

Wyszczególnienie	Współczynnik	Błąd standardowy	Statystyka t-Studenta	Wartość p
A. Równanie kointegrujące; N = 44; Y = $\Delta_4PKB_pc_t$				
Δ_4WLN_pc	8,112	1,431	5,671	<0,0001 ***
B. Jednorównaniowy model ECM (model 1); N = 43; Y = $\Delta\Delta_4PKB_pc_t$				
Const	0,129	0,068	1,889	0,066 *
$\Delta\Delta_4WLN_pc$	4,408	0,976	4,518	<0,0001 ***
ECM_{t-1}	-0,233	0,114	-2,050	0,047 **
C. Zmodyfikowany model ECM (model 2); N = 40; Y = $\Delta\Delta_4PKB_pc_t$				
Const	0,10	0,06	1,77	0,09 *
$\Delta\Delta_4WLN_pc$	4,06	0,94	4,35	0,00 ***
ECM_{t-1}	-0,19	0,10	-2,01	0,05 ***
$\Delta\Delta_4PKB_pc_{t-2}$	0,37	0,11	3,40	0,00 ***
$\Delta\Delta_4PKB_pc_{t-3}$	0,36	0,15	2,40	0,02 ***
$\Delta\Delta_4PKB_pc_{t-4}$	-0,35	0,17	-2,02	0,05 ***

Uwaga. N – wykorzystane obserwacje, Y – zmienna zależna. Poziom istotności: * – 0,1; ** – 0,05; *** – 0,01.

Źródło: obliczenia własne na podstawie: GUS (b.r.) i ZPL (2011–2020, 2021a, 2022a).

Wynik badania przy użyciu testu KPSS wskazuje na brak podstaw do odrzucenia hipotezy zakładającej stacjonarność reszt z równania kointegrującego na poziomie istotności 0,05. Na podstawie przeprowadzonej analizy można więc stwierdzić, że między badanymi szeregami (wartości umów leasingowych netto per capita i PKB per capita) występuje kointegracja.

Wyniki oszacowanego jednorównaniowego modelu korekty błędem opartego na czwartych przyrostach logarytmów WLN_pc_t przedstawiono w tabl. 2B. Ze względu na autokorelację reszt występującą w modelu 1 zmodyfikowano go, wprowadzając opóźnione wartości zmiennej modelowanej. Wyniki estymacji przedstawiono w tabl. 2C.

Wszystkie parametry modelu 2 – oprócz stałej (istotność na poziomie 0,1) – są statystycznie istotne na poziomie 0,05, a wartość statystyki ADF wynosi 5,22, z asymptotyczną wartością $p < 0,001$, co wskazuje na stacjonarność reszt modelu. Badanie testami: chi-kwadrat (statystyka testu: $\chi^2 = 4,55$, $p = 0,10$), Breuscha-Godfrey

$LMF = 0,77, p = 0,55$) i White'a na heteroskedastyczność reszt $LM = 31,23, p = 0,05$) wskazuje, że reszty modelu mają charakter Gaussowskiego białego szumu. Dodatkowo na podstawie testu CUSUM (ang. *cumulated sum of residuals*) przeprowadzono badanie stabilności parametrów modelu. Wyniki testu Harveya-Colliera ($t(32) = 1,01, p = 0,32$) wskazują na brak podstaw do odrzucenia hipotezy zerowej. Można zatem przyjąć, że oszacowany model jest stabilny.

Wartość parametru stojącego przy zmiennej ECM_{t-1} ($-0,19$) określa tempo powrotu do równowagi długookresowej. Ujemny znak tego parametru zapewnia powrót do stanu równowagi, czyli oznacza, że odchylenia od stanu równowagi były korygowane. Wartość parametru wskazuje, że 19% odchylenia jest korygowane po upływie jednego kwartału.

Uzyskane wyniki są zbieżne z prezentowanymi w raporcie PwC i ZPL (2019). Wpływają z nich następujące wnioski. Po pierwsze leasing stymuluje inwestycje, ponieważ zapewnia dostęp do kapitału podmiotom, które nie mają dostatecznej zdolności kredytowej lub z różnych powodów nie są w stanie uzyskać środków finansowych z innych źródeł. Inwestycje, jako jeden z czterech filarów kształtujących PKB, są natomiast ważnym czynnikiem wzrostu gospodarczego. Po drugie leasing pozytywnie wpływa na rozwój innowacyjności, ponieważ znacznie ułatwia dostęp do zaawansowanych rozwiązań technologicznych (na ogół stosunkowo drogich), a ze względu na swój charakter sprzyja ciągłej odnowie i unowocześnianiu wykorzystywanych środków. Co więcej, wpływając na stopę inwestycji, przyczynia się do wzrostu zatrudnienia – większy kapitał, będący następstwem wzrostu inwestycji, podnosi krańcową produktywność pracy, co prowadzi do zwiększenia zapotrzebowania na pracowników. Wzrost zatrudnienia może mieć z kolei wpływ na poprawę koniunktury w gospodarce, ponieważ podnosi popyt konsumpcyjny. Po trzecie zwiększający się wpływ leasingu na polską gospodarkę można zaobserwować, analizując przede wszystkim relacje wartości nowych umów do poziomu inwestycji i PKB.

6. Podsumowanie

W 2021 r. leasing okazał się najpopularniejszym źródłem kapitału obcego wykorzystywanym do finansowania działalności gospodarczej. Największym odbiorcą usług leasingowych był sektor MŚP. Polscy leasingodawcy utrzymywali wysoką pozycję na europejskim rynku usług leasingowych, plasując się zaraz za najbardziej rozwiniętymi gospodarczo krajami Europy. Rekordowa wartość finansowania udzielonego przez branżę leasingową w 2021 r. dowodzi, że branża ta szybko poradziła sobie ze skutkami pandemii COVID-19, co należy uznać za duży sukces.

Analiza kointegracji wartości umów leasingowych netto per capita i wartości PKB per capita w Polsce wykazała, że między rynkiem leasingu a wzrostem gospodarczym kraju występuje istotna długookresowa zależność. Innymi słowy, wyniki analizy świadczą o tym, że rozwój branży leasingowej odgrywa istotną rolę w procesie wzrostu gospodarczego.

Badanie przedstawione w niniejszym artykule nie wyczerpuje podjętego zagadnienia. Zmiany zachodzące w polskiej gospodarce w 2022 r. – bardzo wysoki poziom inflacji, wzrost stóp procentowych, wysokie ceny surowców spowodowane wybuchem wojny w Ukrainie, spadek dynamiki inwestycji czy wreszcie zmiany w obowiązujących przepisach podatkowych – sprawiają, że potrzebne są dalsze badania uwzględniające wymienione zjawiska.

Bibliografia

- Bednarczyk, T. H. (2012). *Wpływ działalności sektora ubezpieczeniowego na wzrost gospodarczy*. Wydawnictwo UMCS.
- Białek-Jaworska, A., Dzik-Walczak, A., Nehrebecka, N. (2014). Determinanty finansowania leasingiem inwestycji rzeczowych przedsiębiorstw. *Bank i Kredyt*, 45(6), 557–588. https://bankikredyt.nbp.pl/home.aspx?f=/content/2014/06/bik_06_2014_pl.html.
- Charemza, W. W., Deadman, D. F. (1997). *Nowa ekonometria*. Polskie Wydawnictwo Ekonomiczne.
- EY-Parthenon, Związek Polskiego Leasingu. (2024). *Leasing na ścieżce transformacji. 30 lat napędzamy polską gospodarkę. Raport przygotowany na 30-lecie Związku Polskiego Leasingu*. https://leasing.org.pl/uploads/Raport_ZPL_druk_1403.pdf.
- Główny Urząd Statystyczny. (b.r.). *Wskaźniki makroekonomiczne*. Pobrane 8 czerwca 2022 r. z <https://stat.gov.pl/wskazniki-makroekonomiczne/>.
- Główny Urząd Statystyczny. (2023). *Wyspecjalizowane segmenty rynku finansowego 2022*. <https://stat.gov.pl/obszary-tematyczne/podmioty-gospodarcze-wyniki-finansowe/przedsiębiorstwa-finansowe/wyspecjalizowane-segmenty-ryнку-finansowego-2022,18,8.html>.
- Grochowski, M. (2016). *Umowa leasingu w orzecznictwie sądów powszechnych*. Instytut Wymiaru Sprawiedliwości. https://iws.gov.pl/wp-content/uploads/2018/08/Grochowski-M._Umowa-leasingu.pdf.
- Haiss, P., Kichler, E. (2009). *Leasing and Economic Growth – Evidence for Central and South Eastern Europe*. https://www.efmaefm.org/0efmameetings/efma%20annual%20meetings/2009-Milan/papers/EFMA2009_0342_fullpaper.pdf.
- Hońko, S. (2016). Ocena wpływu zasad ujmowania umów leasingowych według MSSF 16 „Leasing” na sprawozdania finansowe wybranych spółek notowanych na GPW. *Finanse, Rynki Finansowe, Ubezpieczenia*, (2, cz. 1), 41–51. <https://doi.org/10.18276/frfu.2016.2.80/1-05>.
- Krupiński, M. (2020, 21 lutego). *Rynek leasingu w 2019 roku. Miał być wzrost, jest spadek*. <https://www.fmleasing.pl/aktualnosci/rynek-leasingu-2019>.
- Leaseurope. (2022). *Annual Statistical Enquiry 2021*. https://www.leaseurope.org/_flysystem/s3?file=Statistics/Annual%20statistics/Leaseurope%20Annual%20Statistical%20Enquiry%202021_Summary%20Report_V2.pdf.

- Lleshaj, L., Kripa, D. (2021). The Effect of Financial Leasing Threshold in the Financial Development and Economic Growth: Evidence from Albania. *International Journal of Economics and Business Administration*, 9(2), 165–177. <https://doi.org/10.35808/ijeba/695>.
- Nesterowicz, R., Nesterowicz, A. (2020). Znaczenie leasingu jako źródła finansowania działalności i rozwoju przedsiębiorstw. *Przedsiębiorczość – Edukacja*, 16(2), 250–260. <https://doi.org/10.24917/20833296.162.20>.
- Olkowska, W. (2009). Leasing w finansowaniu rozwoju makro i małych przedsiębiorstw. *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Ekonomiczne Problemy Usług*, (34), 81–89.
- Osińska, M. (2006). *Ekonometria finansowa*. Polskie Wydawnictwo Ekonomiczne.
- Politowska-Iwaszko, M. (2020). Leasing jako źródło finansowania samodzielnych publicznych zakładów opieki zdrowotnej w Polsce w latach 2013–2019. *Journal of Finance and Financial Law*, 2(26), 83–97. <https://doi.org/10.18778/2391-6478.2.26.06>.
- Polska Agencja Rozwoju Przedsiębiorczości. (2022). *Raport o stanie małych i średnich przedsiębiorstw w Polsce*. <https://www.parp.gov.pl/component/publications/publication/raport-o-stanie-sektora-malych-i-srednich-przedsiębiorstw-w-polsce-2022>.
- PricewaterhouseCoopers, Związek Polskiego Leasingu. (2019). *25 lat leasingu w Polsce. Od finansowej innowacji do finansowania polskiej innowacyjności*.
- Radło, M. J., Szczech-Pietkiewicz, E., Szarek-Piaskowska, A., Sagan, M. (2022). Znaczenie leasingu w finansowaniu inwestycji przedsiębiorstw w krajach Europy Środkowo-Wschodniej – stan obecny i trendy przyszłości. W: A. Chłoń-Domińczak, R. Sobiecki, M. Strojny, B. Majewski (red.), *Raport SGH i Forum Ekonomicznego 2022* (s. 261–282). Oficyna Wydawnicza SGH. <https://doi.org/10.33119/978-83-8030-563-2.2022.9.261.281>.
- Roszkowska, S. (2004). Inwestycje a wzrost gospodarczy w krajach OECD. *Gospodarka Narodowa*, 195(10), 1–21. <https://doi.org/10.33119/GN/113701>.
- Ustawa z dnia 26 lipca 2000 r. o zmianie ustawy – Kodeks cywilny (Dz.U. nr 74, poz. 857).
- Vakhitov, D. R., Zamaletdinov, A. (2015). Leasing as A Factor of Economic Growth. *Procedia Economics and Finance*, 23, 839–845. [https://doi.org/10.1016/S2212-5671\(15\)00365-2](https://doi.org/10.1016/S2212-5671(15)00365-2).
- Welfe, A. (2003). *Ekonometria*. Polskie Wydawnictwo Ekonomiczne.
- Związek Polskiego Leasingu. (2011). *Statystyki za rok 2010*. http://www.leasing.org.pl/page.php?entry_details&id=98.
- Związek Polskiego Leasingu. (2012). *Statystyki za rok 2011*. http://www.leasing.org.pl/page.php?entry_details&id=99.
- Związek Polskiego Leasingu. (2013). *Statystyki za rok 2012*. http://www.leasing.org.pl/page.php?entry_details&id=100.
- Związek Polskiego Leasingu. (2014). *Statystyki za rok 2013*. http://www.leasing.org.pl/page.php?entry_details&id=101.
- Związek Polskiego Leasingu. (2015). *Statystyki za rok 2014*. http://www.leasing.org.pl/page.php?entry_details&id=102.
- Związek Polskiego Leasingu. (2016). *Statystyki za rok 2015*. http://www.leasing.org.pl/page.php?entry_details&id=103.
- Związek Polskiego Leasingu. (2017). *Statystyki za rok 2016*. http://www.leasing.org.pl/page.php?entry_details&id=104.
- Związek Polskiego Leasingu. (2018). *Statystyki za rok 2017*. http://www.leasing.org.pl/page.php?entry_details&id=105.

- Związek Polskiego Leasingu. (2019). *Statystyki za rok 2018*. http://www.leasing.org.pl/page.php?entry_details&id=106.
- Związek Polskiego Leasingu. (2020). *Statystyki za rok 2019*. http://www.leasing.org.pl/page.php?entry_details&id=107.
- Związek Polskiego Leasingu. (2021a). *Statystyki za rok 2020*. http://www.leasing.org.pl/page.php?entry_details&id=108.
- Związek Polskiego Leasingu. (2021b). *W 2021 r. rynek leasingu wróci do poziomu finansowania sprzed pandemii*. Pobrane 13 września 2022 r. z <http://leasing.org.pl/pl/aktualnosci/2021/w-2021r-rynek-leasingu-wroci-do-poziomu-finansowania-sprzed-pandemii>.
- Związek Polskiego Leasingu. (2022a). *Statystyki za rok 2021*. http://www.leasing.org.pl/page.php?entry_details&id=109.
- Związek Polskiego Leasingu. (2022b). *Wyniki branży leasingowej w 2021 r.* Pobrane 15 września 2022 r. z http://leasing.org.pl/files/uploaded/ZPL_Prezentacja_Wyniki%20bran%C5%BCy%20leasingowej_2021_14.02.2022_1.pdf.

The application of multidimensional scaling of data to determining changes in retailer customers' preferences

Marcin Pełka,^a Antonio Irpino^b

Abstract. The COVID-19 pandemic has significantly affected several aspects of human activity, including the functioning of different markets, therefore it is important to research its impact. The aim of the study discussed in this paper is to determine which of the three approaches within the method of multidimensional scaling (i.e. multidimensional scaling of classical, symbolic interval-valued or symbolic histogram data) is most adequate for capturing the shifts in retailer customers' preferences that took place during the pandemic. The research concerned the health and beauty market. It was based on the data on orders for beauty products from a Polish producer of cosmetics of a global reach placed by 18 small, mainly family-managed, health and beauty retailers from Lower Silesia. The shops were selected through convenient sampling. Such shops are not a part of large health and beauty retailer chains, therefore they are more vulnerable to all the fluctuations and shifts on the market. The results of this study indicate that in 2020 and 2021, important changes took place on the Lower Silesian health and beauty market as compared to 2019. These changes involved cosmetics for eyes and eyebrows gaining popularity at the expense of cosmetics for lips and cheeks. Multidimensional scaling of symbolic histogram data turned out to be the most effective method (in the sense of the measure of fit and the Pearson correlation coefficient) of capturing and analysing changes happening on a market over a period of time.

Keywords: multidimensional scaling, preferences of customers, symbolic data, pandemic, COVID-19, beauty industry

JEL: C87, C30, L19, N84

Zastosowanie wielowymiarowego skalowania danych do określenia zmian preferencji klientów sklepów

Streszczenie. Pandemia COVID-19 miała istotny wpływ na wiele obszarów ludzkiej aktywności, w tym na funkcjonowanie różnych rynków, dlatego ważne jest badanie jej oddziaływania. Celem badania omawianego w artykule jest wskazanie, które podejście w zakresie metody skalowania wielowymiarowego: danych klasycznych, danych symbolicznych interwałowych czy danych

^a Uniwersytet Ekonomiczny we Wrocławiu, Wydział Ekonomii i Finansów, Polska / Wrocław University of Economics and Business, Department of Economy and Finance, Poland. ORCID: <https://orcid.org/0000-0002-2225-5229>. Autor korespondencyjny / Corresponding author, e-mail: marcin.pelka@ue.wroc.pl.

^b Second University of Naples, Department of Mathematics and Physics, Italy. ORCID: <https://orcid.org/0000-0001-9293-7180>. E-mail: antonio.irpino@unicampania.it.

symbolicznych histogramowych jest najodpowiedniejsze do rozpoznania zmian preferencji klientów sklepów, które nastąpiły w okresie pandemii. Badanie dotyczyło rynku kosmetyków. Wykorzystano dane o zamówieniach na produkty do makijażu jednej z polskich firm o globalnym zasięgu złożonych przez 18 niewielkich, głównie rodzinnych, drogerii z woj. dolnośląskiego (zastosowano technikę doboru według wygody). Tego typu sklepy nie należą do sieci wielkich drogerii, dlatego są znacznie bardziej narażone na wszelkie wahania i zmiany na rynku. Uzyskane rezultaty wskazują, że w latach 2020 i 2021 na dolnośląskim rynku kosmetyków nastąpiły istotne zmiany w porównaniu z 2019 r. polegające na tym, że zmniejszyła się popularność kosmetyków do upiększania ust i policzków, a zwiększyła popularność kosmetyków do makijażu oczu i brwi. Najefektywniejszą metodą wychwycenia i analizy zmian zachodzących w czasie na rynku (w sensie miary dopasowania i współczynnika korelacji Pearsona) okazało się skalowanie wielowymiarowe danych symbolicznych histogramowych.

Słowa kluczowe: skalowanie wielowymiarowe, preferencje klientów, dane symboliczne, pandemia, COVID-19, rynek kosmetyków

1. Introduction

The COVID-19 pandemic has had a significant impact on different aspects of human life and the economy. It affected 213 countries and territories all over the world. Over 23 million people were infected and forced to stay at home (Chanda & Kaul, 2022, p. 111). Many different papers show the ways in which the pandemic disrupted the world's economy and the beauty industry (see for example Chanda & Kaul, 2022; Gardner, 2021; Gerstell et al., 2020; Lee & Kwon, 2022; Pikoos et al., 2020; Unger, 2022).

Chandra and Kaul (2022) analysed the Indian beauty industry. They investigated the possibility of a paradigm shift in this industry due to the COVID-19 pandemic and the small chances of its recovery at the granular level. The authors also discussed the impact of the pandemic on the attributes of the Indian market connected to its size and stage of development, the competitive landscape, and future strategies. Pikoos et al. (2020) conducted a survey on a sample of 216 participants in Australia in order to analyse the relationship between the engagement in appearance-focused behaviours and distress regarding the beauty service closure.

Gerstell et al. (2020, p. 4) from McKinsey and Company checked the ways in which COVID-19 was changing the beauty industry (opinion survey). They showed that consumers in China, Japan, the UK and the US intended to spend less on beauty products. The authors also observed a decline by 55% and 75% in the purchases of cosmetics and fragrances, respectively, in 2020 compared to 2019, with the exception of above-the-mask products.

A doctoral thesis by Gardner (2021) and a master's thesis by Unger (2022) analysed the impact of the COVID-19 pandemic on the health and beauty market. Their results show that people limited the use of all the types of cosmetics (face, eye, lip, and skin)

compared to the pre-COVID-19 times. Eye, lip, and skin cosmetics were expected to return to the previous level of popularity after the pandemic.

Werner-Lewandowska et al. (2021) presented a preliminary research that focused on the effects of COVID-19 on the consumer behaviour in Poland in terms of daily shopping. In the second half of 2020, they questioned 214 respondents using the CAWI technique. The questions concerned fast-moving consumer goods (FMCG) and on-line shopping. The results showed that the 2020 lockdown contributed to consumers stockpiling food products, which was due to the fact that most of them purchased FMCG less frequently than before, thus buying larger quantities at a time and stockpiling.

Ratajczak et al. (2023) researched into Polish customer preferences regarding natural cosmetics. The authors observed that significant changes in consumer attitudes towards the cosmetics market occurred after the outbreak of the COVID-19 pandemic. Mościcka (2023), Mościcka et al. (2020) and Ścieszko et al. (2021) focused more on the impact of the pandemic on body care in general. In addition to other aspects, these articles point to considerable changes in the use of cosmetics during the pandemic. Similarly to Ratajczak et al. (2023), they indicated that DIY (do-it-yourself) cosmetics gained much popularity during COVID-19.

However, all these authors use classical data, where objects are usually described by single-valued variables. This makes it possible to represent them as a vector of quantitative or qualitative measurements, where each column represents a variable. Nevertheless, this kind of data representation is too restrictive to represent more complex data. If we want to take into account the variability and uncertainty of data, variables must assume sets of categories or intervals, even with weights or frequencies. Such kind of data has been mainly studied in Symbolic Data Analysis (SDA). The main aim of SDA is to provide suitable methods and techniques for managing aggregated or complex data sets described by multi-valued variables, where the cells of the data table contain sets of categories, intervals, weight (probability) and distribution (see for example Billard & Diday, 2006; Bock & Diday, 2000; Diday & Noirhomme-Fraiture, 2008).

The aim of the study discussed in this paper is to determine which of the three approaches within the method of multidimensional scaling (i.e. multidimensional scaling of classical, symbolic interval-valued or symbolic histogram data) is most adequate for capturing the shifts in retailer customers' preferences that took place during the pandemic. When dealing with the problem of how to analyse its impact on the economy, many different methods and approaches can be used. Among those, multidimensional scaling stands out as very popular (see, for example, its applications presented by Braithwaite et al., 2021; Coco et al., 2021; Güre et al., 2021; Kutlar et al., 2021). The authors decided to use symbolic data analysis and combine individual,

monthly information from Polish health and beauty retailers to obtain symbolic interval-valued variables that would make it possible to capture variability in the data (see for example Billard & Diday, 2006).

2. Research method

The empirical part of this paper uses data on orders for 2019, 2020 and 2021¹ collected from health and beauty retailers that offer cosmetics products. The sample consists of small companies (usually family-managed) that are not a part of large retail chains (e.g. Rossmann, Hebe, etc.). They do not have large storing capacity and they make orders as they sell products.

The company selected for the purpose of this study is one of the biggest and best-known brands of cosmetics in Poland. It was established in 1988. Its products are offered in health and beauty retailers in Poland (over 1,500 cooperating shops all over Poland plus an online store) and additionally in some shops in Saudi Arabia, Bahrain, China, Croatia, Cyprus, the Czech Republic, Georgia, Estonia, India, Italy, Kazakhstan, Latvia, Moldavia, Portugal, Pakistan, the Seychelles, Slovenia, Turkey, Vietnam and the United Arab Emirates.² This brand is known for a wide range of beauty products for different skin types.

The data set contains 3,080 monthly orders made by a sample of 18 health and beauty retailers (selected by convenient sampling) located in the Lower Silesia region in 12 Polish cities (Dzierżoniów, Góra, Jawor, Jelenia Góra, Jelenia Góra (outskirts), Kąty Wrocławskie, Mirsk, Rawicz, Świdnica, Wałbrzych, Wrocław and Zgorzelec). These retailers were ordering the following products: nail polishes, lipsticks, lip glosses, eye shadows, fluids, mascaras, powders, blushes and lip pencils.

In the empirical part of the paper, symbolic multidimensional scaling for symbolic histogram data was applied. Unlike classical data, where each object is described by a set of qualitative or quantitative variables, symbolic data analysis enables each object to be described not only by nominal, ordinal, interval, or ratio variables, but also by symbolic interval-valued variables, symbolic multivalued variables, symbolic multivalued variables with weights and symbolic histogram variables (see e.g. Billard & Diday, 2006, pp. 7–30; Bock & Diday, 2000, pp. 2–3; Brito & Dias, 2022; Diday & Noirhomme-Fraiture, 2008, pp. 10–19). More details about symbolic variables and symbolic objects can be found in Billard and Diday (2006, pp. 7–66), Bock and Diday (2000, pp. 2–8), Brito and Dias (2022, pp. 6–35) and Diday and Noirhomme-

¹ The data were collected by Michał Swachta, the student of Wrocław University of Economics, for his bachelor's thesis *Analiza unfolding jako narzędzie identyfikacji zmian na rynku kosmetycznym przed i w trakcie pandemii Covid-19* (2022).

² Information obtained from the official company's website.

Fraiture (2008, pp. 3-30). Table 1 presents some symbolic variables as well as their examples.

Table 1. Symbolic variables and their examples

Variable name	Example	Variable type
Preferred car price (in EUR)	(10,000, 30,000); (15,000, 45,000); (12,000, 22,000)	symbolic interval-valued (non-disjoint intervals)
Engine capacity for insurance purposes (in ccm)	(up to 1,000), (1,000, 2,000), (2,000, 3,000), (over 3,000)	symbolic interval-valued (disjoint variables)
Preferred car colour	orange, yellow, blue, red	symbolic multi-valued
Preferred car brand	{Toyota (0.6), VW (0.4)}, {Audi (1.0)} {Skoda (0.5), Renault (0.4), other (0.1)}	symbolic multi-valued (with weights in brackets)
Home-work travel time (minutes)	{[40, 60] (0.6), [60, 85] (0.4)}	symbolic histogram variable (with weights in brackets)
Age	25, 45, ...	ratio
Gender	F, M	nominal

Source: authors' work.

In this study, we applied multidimensional scaling methods to the classical data and two types of symbolic data. Multidimensional scaling (regardless of whether we apply it to classical or symbolic data) is the method that makes it possible to represent efficiently (i.e. as close to the original as possible) the distance matrix between objects in m -dimensional space $[\delta_{ik}]$ (usually $m > 3$) in the distance matrix between the objects in q -dimensional space (where $q < m$). For the purposes of the graphical presentation of the results, we usually consider $q=2$ or $q=3$. Lower q -dimensional space allows the analysis of relations between the objects and a clear identification (interpretation) of each dimension. The dimensions cannot be observed directly. They represent latent variables which enable the explanation of the similarities and differences between the analysed objects.

2.1. Multidimensional scaling for classical data

As mentioned before, multidimensional scaling aims to present the relationships between objects from a high-dimensional (usually over three-dimensional) space in a lower-dimensional space (usually two- or three-dimensional). A similar technique that is capable of representing the relationships between multidimensional data in a lower-dimensional space is the popular principal component analysis (PCA). The PCA, which is probably the oldest technique of multivariate analysis, was introduced by Karl Pearson in 1901 and developed by Harold Hotelling in 1933 (see Jolliffe, 2002, pp. 6-8).

The general idea of the PCA is to reduce the dimensionality of the dataset in which there are a large number of interrelated variables, while retaining as much as possible

of the variation present in the data. This is achieved by transforming principal components into a new set of variables which are uncorrelated and ordered in such a way that the first few retain most of the variation present in all of the initial (original) variables (Jolliffe, 2002, p. 1).

In order to maximise the variance, the first principal component is as follows (Jolliffe, 2002):

$$Z_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p, \quad (1)$$

where a_{11}, \dots, a_{1p} are the coefficients for initial (original) variables.

In a similar way, the subsequent i -th principal components are defined as follows: $Z_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p$. Each subsequent component is selected so as to maximise the variance that was not captured by the previous one, and the subsequent components are orthogonal (they are not correlated).

As mentioned before, the multidimensional scaling (MDS) is a method that represents the measurements of similarities (or dissimilarities) among pairs of objects as distances between points of a low-dimensional multidimensional space (Borg & Groenen, 2005, p. 3).

The MDS can be seen as an explanatory technique. It makes it possible to study data that are not linked to an explicit theory and to predict their patterns. The MDS enables researchers to find patterns and structure in the data (Borg & Groenen, 2005, p. 4). It can also be seen as a tool for testing structural hypotheses. In addition, the MDS makes it possible to find out how the dimensions (features, factors) by which the objects are described are reflected in corresponding distances between observations of these items (Borg & Groenen, 2005, p. 6). Another application of the MDS is the exploration of psychological structures (Borg & Groenen, 2005, pp. 9–11). Finally, the technique can be used as a model of similarity judgements (Borg & Groenen, 2005, p. 11).

We applied the iterative-algorithm procedure of multidimensional scaling performed by means of smacof algorithm (Borg & Groenen, 2005, pp. 204–205). The name smacof stands for Stress Majorisation of a Complicated Function. The majorisation was introduced to multidimensional scaling by De Leeuw (1977), and subsequently developed by De Leeuw and Heiser (1977, 1980). The general stress function for the symmetric data case is defined as follows (Mair et al., 2015, p. 4):

$$\sigma_r(X) = \sum_{i < j} w_{ij} (\hat{d}_{ij}^2 - d_{ij}(x))^2, \quad (2)$$

where w_{ij} are weights, \hat{d}_{ij}^2 are d-hats, disparities, target distances or pseudo-distances (see Borg & Groenen, 2005, p. 199), $\hat{d}_{ij}^2 = f(\delta_{ij})$, X are coordinates of objects in high-dimensional space, and x are final coordinates for objects in the lower space.

The general smacof algorithm, where we minimise the stress function over the configuration matrix \mathbf{V} , can be described using pseudo-code in the following way (Borg & Groenen, 2005, pp. 191, 204–205):

- a) set $\mathbf{V} = \mathbf{V}^{[0]}$, where $\mathbf{V}^{[0]}$ is some random or non-random start configuration. Set iteration counter (k) to 0 and convergence criterion ε to a small positive number (i.e. $\varepsilon = 1 \cdot 10^{-6}$),
- b) find optimal disparities \hat{d}_{ij}^2 for fixed distances $d_{ij}(\mathbf{V}^{[0]})$,
- c) standardise \hat{d}_{ij}^2 so that $\eta_{\hat{d}}^2 = n(n-1)/2$, to avoid degenerated solution,
- d) compute Stress function – see equation (2),
- e) set $\sigma_r^{[-1]} = \sigma_r^{[0]}$,
- f) while $k = 0$ or $\sigma_r^{[k-1]} - \sigma_r^{[k]} > \varepsilon$ and $k \leq$ maximum iterations do,
- g) increase number of iterations by one ($k = k + 1$),
- h) compute Guttman transform $\mathbf{V}^{[k]}$,
- i) find optimal disparities \hat{d}_{ij}^2 for fixed distances $d_{ij}(\mathbf{V}^{[k]})$,
- j) standardise \hat{d}_{ij}^2 so that $\eta_{\hat{d}}^2 = n(n-1)/2$,
- k) compute updated stress function for new values of \hat{d}_{ij}^2 and $\mathbf{V}^{[k]}$,
- l) end while.

As the final result, we obtain points that represent data rows (usually objects) in q dimensions. The details about the majorisation in the case of the smacof algorithm can be found in Borg and Groenen (2005).

2.2. Multidimensional scaling for symbolic interval-valued data

We applied I-Scal algorithm proposed by Groenen et al. (2006) to symbolic interval-valued data. In the case of symbolic interval-valued data, we assumed that we either have symbolic interval-valued variables and then we calculate a symbolic interval-valued dissimilarities $[\delta_{ij}^{(L)}, \delta_{ij}^{(U)}]$, where $\delta_{ij}^{(L)}$ is the lower bound of the dissimilarity between i -th and j -th object, and $\delta_{ij}^{(U)}$ is the upper bound of the dissimilarity, or we have the symbolic interval-valued dissimilarities as input data.

In order to develop the MDS for interval-valued data, we have to assume that the ranges of dissimilarities are represented by the ranges of distances. Not only are the distances represented by ranges, but the coordinates of objects themselves are ranges. The maximum Euclidean distance between rectangles i and j is defined as follows (Groenen et al., 2006, p. 362):

$$d_{ij}^{(U)}(\mathbf{X}, \mathbf{R}) = \left(\sum_{s=1}^p [|x_{is} - x_{js}| + (r_{is} + r_{js})]^2 \right)^{1/2}, \quad (3)$$

and the minimum Euclidean distance is defined as:

$$d_{ij}^{(L)}(\mathbf{X}, \mathbf{R}) = \left(\sum_{s=1}^p \max [0, |x_{is} - x_{js}| - (r_{is} + r_{js})]^2 \right)^{1/2}, \quad (4)$$

where x_{is}, x_{js} are the centers of the i -th and the j -th rectangle that are used to build matrix \mathbf{X} , r_{is}, r_{js} are the radiuses of the i -th and the j -th rectangle that are used to build matrix \mathbf{R} , and $s = 1, \dots, p$ is the variable number.

The definition of $d_{ij}^{(L)}(\mathbf{X}, \mathbf{R})$ guarantees that if two hyperboxes overlap, their minimum distance is zero.

The I-Stress loss function that is used in this case is defined as follows (Groenen et al., 2006, p. 363):

$$\sigma_I^2(\mathbf{X}, \mathbf{R}) = \sum_{i < j}^n w_{ij} [\delta_{ij}^{(U)} - d_{ij}^{(U)}(\mathbf{X}, \mathbf{R})]^2 + w_{ij} [\delta_{ij}^{(L)} - d_{ij}^{(L)}(\mathbf{X}, \mathbf{R})]^2, \quad (5)$$

where $\delta_{ij}^{(U)}$ is the upper bound of the dissimilarity of objects i and j , $\delta_{ij}^{(L)}$ is the lower bound, w_{ij} is a given non-negative weight, and $d_{ij}^{(U)}(\mathbf{X}, \mathbf{R})$ and $d_{ij}^{(L)}(\mathbf{X}, \mathbf{R})$ are upper and lower bound dissimilarities based on the centres and ranges (spreads), respectively.

The normalised version of the I-Stress will be applied in the empirical part of the paper. The normalized I-Stress is defined as follows (Groenen et al., 2006, p. 363):

$$\frac{\sigma_I^2(\mathbf{X}, \mathbf{R})}{\sum_{i < j}^n w_{ij} [\delta_{ij}^{(U)}]^2 + \sum_{i < j}^n w_{ij} [\delta_{ij}^{(L)}]^2}, \quad (6)$$

where all elements are the same as in equation (5).

As a result, we obtained rectangles (in two-dimensional space) or cuboids (in three-dimensional space).

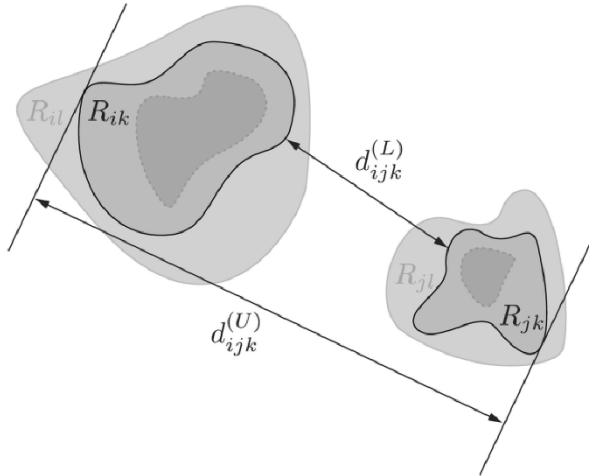
The details about the majorisation of the I-Stress function can be found in the paper by Groenen et al. (2006).

2.3. Multidimensional scaling for symbolic histogram data

We applied rectangle non-concentric multidimensional scaling to the symbolic histogram data. To use this approach, a histogram-valued dissimilarity data (also known as percentile-valued dissimilarity data) were needed. Let us consider set A with elements A_k containing the ranges of the percentile values $[\alpha_k^{(L)}, \alpha_k^{(U)}]$ such that the ranges become increasingly large, and the smaller one contains the large one, that is $A_k \subset A_l$ (A_l is a subset of A_k) if $k < l$. A simple example can be $A_1 = [0.4, 0.6]$, $A_2 = [0.3, 0.7]$, $A_3 = [0.2, 0.8]$. We can see that intervals in A_k do not have to be symmetric around any value, the only requirement is that $A_k \subset A_l$ if $k < l$.

We applied a constrained version of the standard interval MDS to symbolic histogram-valued dissimilarity data. Let R_{ki} be a region of object i corresponding to percentile range A_k . The regions are satisfying $R_{ki} \subset R_{li}$ for $A_{ki} \subset A_{li}$. Figure 1 shows the general idea of a region (Brito & Dias, 2022, p. 284).

Figure 1. Histogram representation of general regions, where region R_{ik} is contained in R_{il} if percentile range $A_{ik} \subset A_{il}$



Source: Brito and Dias (2022, p. 285).

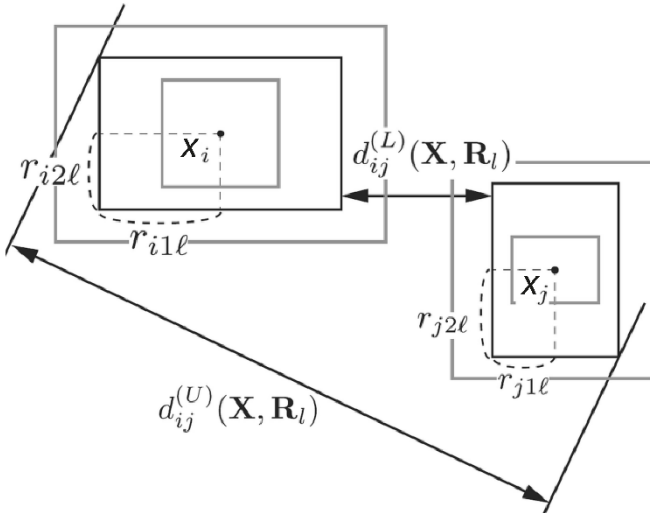
The stress function for symbolic histogram-valued dissimilarity data was applied (Brito & Dias, 2022, p. 285):

$$\begin{aligned} \min \sigma_{\text{HIST}}^2(R_1, \dots, R_K) := & \sum_{k=1}^K \sum_{i=1}^{n-1} \sum_{j=i+1}^n w_{ijk} \left(\delta_{ijk}^{(L)} - d_{ij}^{(L)}(R_k) \right)^2 + \\ & + \sum_{k=1}^K \sum_{i=1}^{n-1} \sum_{j=i+1}^n w_{ijk} \left(\delta_{ijk}^{(U)} - d_{ij}^{(U)}(R_k) \right)^2 \end{aligned} \tag{7}$$

subject to $R_{ki} \subset R_{li}$ for $A_k \subset A_l$.

The rectangle non-concentric model is shown in Figure 2.

Figure 2. Rectangle non-concentric model



Source: Brito and Dias (2022, p. 286).

In the non-concentric model, each object is represented by nested rectangles that do not necessarily have the same centre point. The basic stress function for histogram-valued dissimilarities has changed. The elements $d_{ij}^{(L)}(R_k)$ and $d_{ij}^{(U)}(R_k)$ are replaced by $d_{ij}^{(L)}(\mathbf{X}_k, \mathbf{R}_k)$ and $d_{ij}^{(U)}(\mathbf{X}_k, \mathbf{R}_k)$.

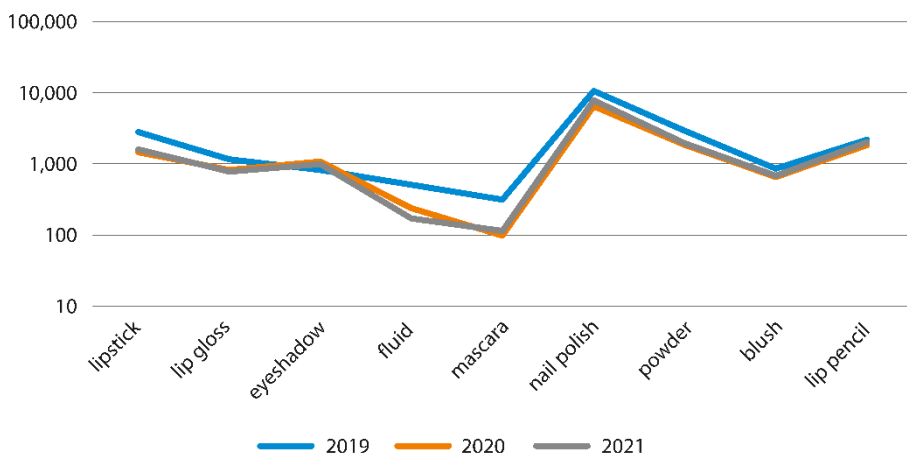
Details of other MDS types for symbolic histogram-valued dissimilarities and the majorisation steps for this data case are presented by Brito and Dias (2022, pp. 275–294).

3. Results

This part presents results of multidimensional scaling for the dataset in three cases. Firstly, we assume that the data will be classical data, secondly, that it will be interval-valued data, and thirdly, that it will be symbolic histogram data. When looking at the whole data set, we can see some changes in customers' behaviour when comparing three analysed years: 2019, 2020 and 2021.

Figure 3 presents the ordered quantities of each product (with a logscale) in 2019, 2020 and 2021.

Figure 3. Ordered quantities of beauty products

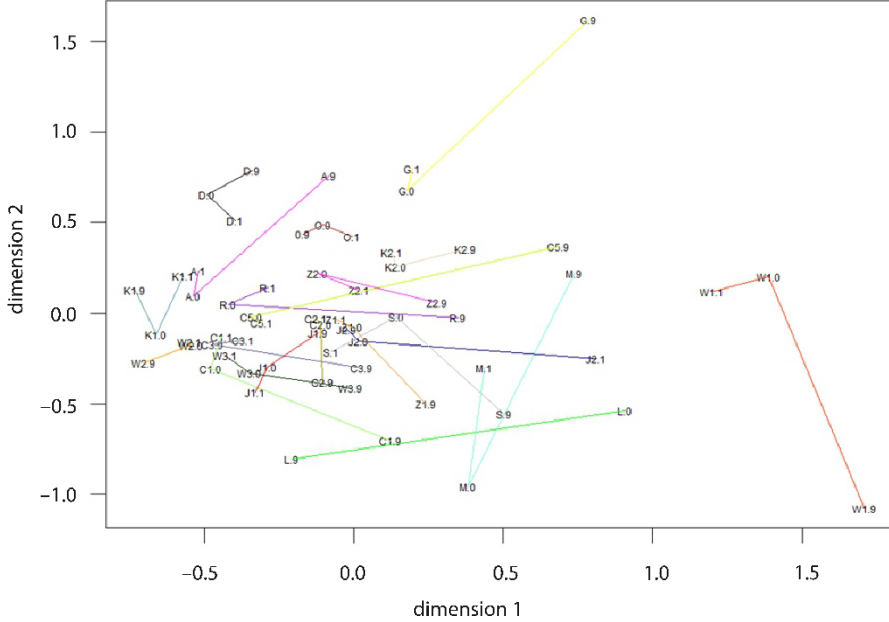


Source: authors' work based on the dataset for selected health and beauty retailers.

If we compare the volumes of products that were ordered in 2019, 2020 and 2021 to the corresponding purchase volumes for the studied company from the selected small health and beauty retailers in Lower Silesia, we can see that starting in 2019, all these retailers were ordering and selling fewer products, with just one exception – mascaras.

In the first attempt to see if the MDS would be able to capture the changes that happened in 2019, 2020 and 2021, we used the classical MDS, where the values were averaged. Figure 4 presents the final configuration plot for all the years, where the averages for each year were calculated, as were the lines that connect the same products over time.

Figure 4. The MDS final configuration plot where all the years are presented simultaneously (as one dataset)



Note. The letter denotes the location of a health and beauty retailer, the first digit (in front of the dot), if present, stands for the studied company's separate sales area, and the second digit indicates the year (9 for 2019, 0 for 2020 and 1 for 2021), so we have: A – Jawor (A.9, A.0, A.1), C – Wrocław (C1.9, C1.0, C1.1, C2.9, C2.0, C2.1, C3.9, C3.0, C4, C4.9, C4.0, C4.1), D – Dzierżoniów (D.9, D.0, D.1), G – Góra (G.9, G.0, G.1), J – Jelenia Góra (J1.9, J1.0, J1.1, J2.9, J2.0, J2.1), K – Kąty Wrocławskie (K1.9, K1.0, K1.1, K2.9, K2.0, K2.1), L – Jelenia Góra (outskirts; L.9, L.0, there is no L.1 as no transactions were done), M – Mirsk (M.9, M.0, M.1), O – Jawor (O.9, O.0, O.1), R – Rawicz (R.9, R.0, R.1), S – Świdnica (S.9, S.0, S.1), W – Wałbrzych (W1.9, W1.0, W1.1, W2.9, W2.0, W2.1, W3.9, W3.0, W3.1), Z – Zgorzelec (Z.9, Z.0, Z.1, Z2.9, Z2.0) Z2.1).

Source: authors' work based on the company's data with application of R software.

The HHI (Hirschman-Herfindahl Index, with $HHI = \sum_{i=1}^n spp_{pi}^2$) approach was used to select the optimal MDS procedure (see Walesiak & Dudek, 2017 for details). The HHI is a sum of squared stress per point values (stress contribution in percentages). The HHI Index takes values from the interval $\left[\frac{10000}{n}; 10000\right]$. The value $\frac{10000}{n}$ means that the distribution of errors for individual objects is uniform. The maximum value is obtained if the summary stress is assigned to one object, and for all others it is zero.

The standardisation and the Euclidean distance were selected as the best steps by `mdsOpt` package for R software (Walesiak & Dudek, 2022).

Figure 4 presents the results of the MDS when we take all health and beauty retailers and years in one dataset, and the values for each product are a simple sum of values for all the invoices in a year – see Table 2 for reference.

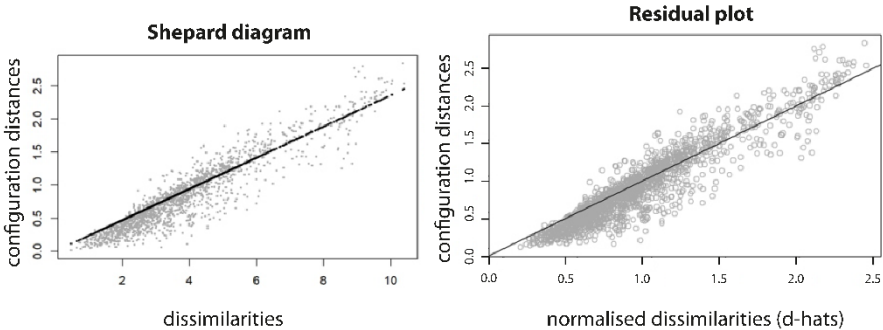
Table 2. Classical data – example of a data creation step

Health and beauty retailer's code	Quantities in 2019			Sum
	lipstick 1	...	lipstick 20	
PL-WAL002	4	...	7	$\sum n_i = 2,024$
...	
PL-WAL002	5	...	12	

Source: authors' work based on the data on the studied company.

The same colours in Figure 4 connect the same companies. Some changes can be clearly seen, while others are harder to find, as many points from different years can be found in the lower bottom side of Figure 4. The Shepard diagram and residual plot for classical MDS results are presented in Figure 5.

Figure 5. Shepard diagram and residual plot for classical MDS results



Source: authors' work based on data on the studied company with the application of R software.

The Shepard diagram and the residual plot confirm that the chosen MDS model was adequate. The final STRESS-1 value is 0.3985992. The Pearson's correlation coefficient for delta and configuration distances is 0.7358537. It confirms that the selected MDS model fits the data.

However, if we look at the points that represent each year, we see virtually no changes between the years. It means that although the model fits the data, the classical MDS model was not able to capture all the changes that happened in 2020 and 2021 compared to 2019.

When considering symbolic interval-valued data, the I-Scal method was selected (Groenen et al., 2006) with minimal and maximal values for each product in each year. This is a temporal aggregation, where data are recorded at different moments in time (see Brito, 2014, p. 282 for details). Table 3 presents details of the data creation step.

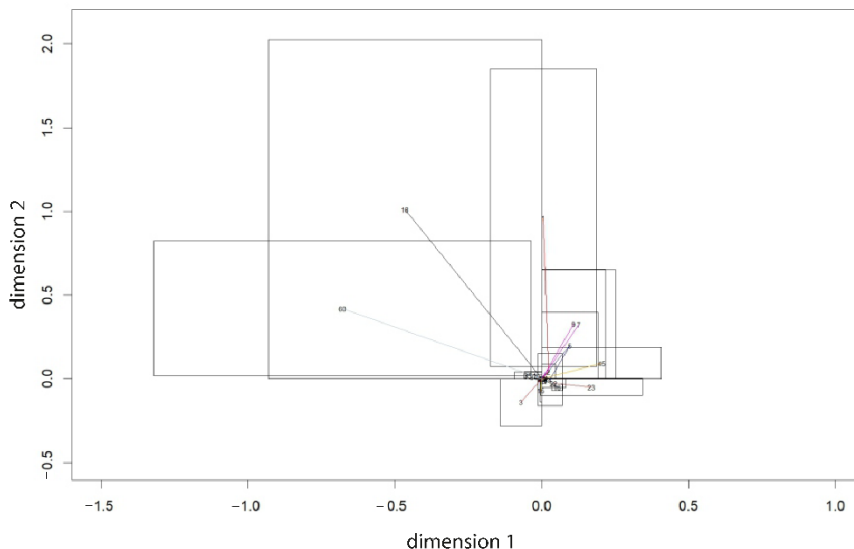
Table 3. Interval-valued symbolic data for the data on the studied company – example of a data creation step

Health and beauty retailer's code	Quantities in 2019			Interval-valued data
	lipstick 1	...	lipstick 20	
PL-WAL002	4	...	7	[min; max] = [8; 31]
...	
PL-WAL002	5	...	12	

Source: authors' work based on the data on the studied company.

Figure 6 presents the MDS results for products (interval-valued data).

Figure 6. Results of the I-Scal MDS when data are represented as interval-valued data



Note. 1, 2, 3 – Jelenia Góra (2019, 2020 and 2021); 4, 5, 6 – Dzierżoniów (2019, 2020 and 2021); 7, 8, 9 – Góra (2019, 2020 and 2021); 10, 11, 12 – Jawor (2019, 2020 and 2021); 13, 14, 15 – Jelenia Góra (outskirts); 16, 17, 18 – Jordanów (2019, 2020 and 2021); 19, 20, 21 – Kąty Wrocławskie (1st shop; 2019, 2020 and 2021) and 22, 23, 24 (2nd shop; 2019, 2020 and 2021); 25, 26, 27 – Mirsk (2019, 2020 and 2021); 28, 29, 30 – Rawicz (2019, 2020 and 2021); 31, 32, 33 – Świdnica (2019, 2020 and 2021); 34, 35, 36, 37, 38, 39, 40, 41, 42 – Wałbrzych (three shops and three years for each); 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54 – Wrocław (four shops and three years); 55, 56, 57 – Zgorzelec (1st shop; 2019, 2020 and 2021) and 58, 59, 60 (2nd shop; 2019, 2020 and 2021).

Source: authors' work using R software.

When we look at Figure 6, we can clearly see that there were significant changes in 2020 and 2021 compared to 2019.

As the I-Scal literature does not propose any guidelines as how to prepare Shepard’s diagram or residual plot, we will present only I-Stress value as well as the Pearson’s correlation coefficient for delta and configuration distances.

The final I-Stress value is 0.1170164 (for symbolic interval-valued data with no data normalisation, interval-valued dissimilarities and 1,000 iterations). The Pearson’s correlation coefficient for delta and configuration distances is 0.9954633. We can say that this model fits the data better than the classical MDS, and thus it was able to capture changes that took place in 2020 and 2021 compared to 2019.

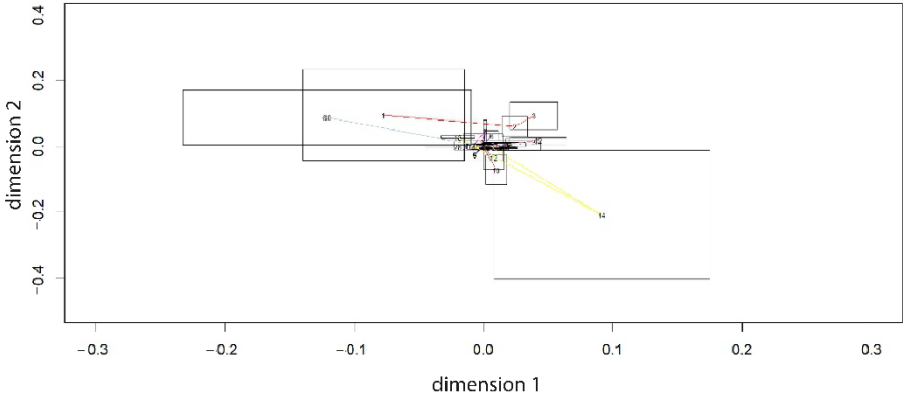
For the symbolic histogram data, non-concentric rectangle MDS model was used, where the 5th and 95th percentiles were used as the lower and upper bounds for histograms. Figure 7 presents the results of the MDS for symbolic histogram data. See Table 4 for details of the data-creation step.

Table 4. Histogram-valued symbolic data for data on the studied company – example of the data-creation step

Health and beauty retailer’s code	Quantities in 2019			Histogram-valued data
	lipstick 1	...	lipstick 20	
PL-WAL002	4	...	7	a) build histograms b) histogram-valued data:
...	[4; 6] (0.2), [6; 8] (0.5),
PL-WAL002	5	...	12	[6; 10] (0.2), [10; 12] (0.1)

Source: authors’ work based on the data on the studied company.

Figure 7. Results of the non-concentric rectangle MDS model for symbolic histogram data



Note. Numbers have the same meaning as in Figure 6.
Source: authors’ work using R software.

As no Shepard's diagram or residual plot was proposed for the non-concentric rectangle MDS, the authors provided information about the final stress value and the Pearson's correlation coefficient for delta and configuration distances.

The final stress value is 0.065558, and the Pearson's correlation coefficient for delta and configuration distances is 0.9994597. These results suggest that this model fits the data best, and was able to capture all the changes that took place in 2020 and 2021 as compared to 2019.

4. Conclusions

This paper presents three different approaches to analysing customers' preferences. The best results in terms of the stress measure was obtained for the MDS applied to histogram-valued symbolic data. The study presented in this paper shows also how to apply three different multidimensional scaling approaches when dealing with data changing over time. The first one is the popular classical multidimensional scaling, where average values are used. Although this approach fits the data, it was not able to capture all the changes that happened in the beauty industry in Poland in 2020 and 2021 as compared to 2019. The second approach is the interval-valued multidimensional scaling for symbolic data (I-Scal). This approach yields better results in terms of the stress value and Pearson's correlation coefficient for delta and configuration distances. The last approach is the non-concentric rectangle multidimensional scaling for symbolic histogram data.

Data prior to 2019 showed that customers buying products of the analysed company had fairly stable preferences. Preferences play the key role in the theory of economy, hence knowing what customers are likely to buy enables a company to make profits, acquire new customers and retain the existing ones.

Products that had been most popular before 2019 were also very popular in 2019 (as far as customers' purchases are considered), with only slight changes occurring. However, post 2019, face beauty products that were located 'below and near the mask line' (e.g. lipsticks, blushers) were bought less often than before, whereas eye beauty products (e.g. mascaras, eye shadows) became more popular than before. It seems that customers tried to boost their beauty but with application of other tools.

The obtained results show that this approach fits the data best (according to the Stress value and Pearson's correlation coefficient for delta and configuration distances), and managed to capture all the changes that took place in 2020 and 2021 compared to 2019. These changes were confirmed by invoices and volumes of ordered products.

Symbolic histogram data and the MDS for this type of data seem to be the best way to capture, compare and analyse changes that happen on markets over time.

References

- Billard, L., & Diday, E. (2006). *Symbolic Data Analysis. Conceptual Statistics and Data Mining*. John Wiley & Sons. <https://doi.org/10.1002/9780470090183>.
- Bock, H.-H., & Diday, E. (Eds.). (2000). *Analysis of Symbolic Data. Explanatory Methods for Extracting Statistical Information from Complex Data*. Springer Verlag. <https://doi.org/10.1007/978-3-642-57155-8>.
- Borg, I., & Groenen, P. J. F. (2005). *Modern Multidimensional Scaling. Theory and Applications* (2nd edition). Springer Science+Business Media.
- Braithwaite, J., Tran, Y., Ellis, L. A., & Westbrook, J. (2021). The 40 health systems, COVID-19 (40HS, C-19) study. *International Journal for Quality in Health Care*, 33(1), 1–7. <https://doi.org/10.1093/intqhc/mzaa113>.
- Brito, P. (2014). Symbolic data analysis: another look at the interaction of data mining and statistics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(4), 281–295. <https://doi.org/10.1002/widm.1133>.
- Brito, P., & Dias, S. (Eds.). (2022). *Analysis of distributional data*. CRC Press. <https://doi.org/10.1201/9781315370545>.
- Coco, M., Guerrero, C. S., Santisi, G., Riggio, F., Grasso, R., Di Corrado, D., Di Nuovo, S., & Ramaci, T. (2021). Psychosocial Impact and Role of Resilience on Healthcare Workers during COVID-19 Pandemic. *Sustainability*, 13(13), 1–9. <https://doi.org/10.3390/su13137096>.
- Chanda, R., & Kaul, G. (2022). COVID-19: Effect on Indian Beauty Industry. In A. Chandani, R. Divekar, J. K. Nayak, & K. Chopra (Eds.), *Pandemic, New Normal and Implications on Business* (pp. 111–127). Springer. https://doi.org/10.1007/978-981-19-4892-3_8.
- De Leeuw, J. (1977). Applications of Convex Analysis to Multidimensional Scaling. In J. R. Barra, F. Brodeau, G. Romier, & B. van Cutsem (Eds.), *Recent Developments in Statistics* (pp. 133–145). North Holland Publishing Company.
- De Leeuw, J., & Heiser, W. J. (1977). Convergence of Correction Matrix Algorithms for Multidimensional Scaling. In J. Lingoes (Ed.), *Geometric Representations of Relational Data* (pp. 735–752). Mathesis Press.
- De Leeuw, J., & Heiser, W. J. (1980). Multidimensional Scaling with Restrictions on the Configuration. In P. R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 501–522). North Holland Publishing Company.
- Diday, E., & Noirhomme-Fraiture, M. (Eds.). (2008). *Symbolic Data Analysis and the SODAS Software*. John Wiley & Sons. <https://doi.org/10.1002/9780470723562>.
- Gardner, K. (2021). *Beauty During a Pandemic: The Impact of COVID-19 on the Cosmetic Industry* (Doctoral dissertation, Honors College Middle Tennessee State University). <https://jewelscholar.mtsu.edu/server/api/core/bitstreams/0e2d09ce-5a9d-4e84-b1da-29c9223d6d66/content>.
- Gerstell, E., Marchessou, S., Schmidt, J., & Spagnuolo, E. (2020). *How COVID-19 is changing the world of beauty*. McKinsey & Company. <https://www.mckinsey.com/~media/McKinsey/Industries/Consumer%20Packaged%20Goods/Our%20Insights/How%20COVID%2019%20is%20changing%20the%20world%20of%20beauty/How-Covid-19-is-changing-the-world-of-beauty-vF.pdf>.
- Groenen, P. J. F., Winsberg, S., Rodríguez, O., & Diday, E. (2006). I-Scal: Multidimensional Scaling of Interval Dissimilarities. *Computational Statistics & Data Analysis*, 51(1), 360–378. <https://doi.org/10.1016/j.csda.2006.04.003>.

- Güre, Ö. B., Kayri, M., & Şevgin, H. (2021). Investigation of Coronavirus Pandemic Indicators of the Countries with Hierarchical Clustering and Multidimensional Scaling. *Eastern Journal of Medicine*, 26(2), 308–315. <https://doi.org/10.5505/ejm.2021.72681>.
- Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd edition). Springer-Verlag.
- Kutlar, A., Gülmez, A., & Öncel, A. (2021). *The Analysis of the Effect of Covid 19 on Macroeconomic Indicators via MDS and Clustering Methods* [Preprint]. <https://doi.org/10.21203/rs.3.rs-194221/v1>.
- Lee, J., & Kwon, K. H. (2022). Sustainable changes in beauty market trends focused on the perspective of safety in the post-coronavirus disease-19 period. *Journal of Cosmetic Dermatology*, 21(7), 2700–2707. <https://doi.org/10.1111/jocd.14493>.
- Mair, P., de Leeuw, J., & Groenen, P. J. F. (2015). *More on Multidimensional Scaling and Unfolding in R: smacof Version 2*. <https://cran.r-project.org/web/packages/smacof/vignettes/smacof.pdf>.
- Mościcka, P. (2023). Body care activities and its consequences related to COVID-19 pandemic. *Journal of Cosmetic Dermatology*, 22(1), 16–20. <https://doi.org/10.1111/jocd.15314>.
- Mościcka, P., Chróst, N., Terlikowski, R., Przyłipiak, M., Wołosik, K., & Przyłipiak, A. (2020). Hygienic and cosmetic care habits in Polish women during COVID-19 pandemic. *Journal of Cosmetic Dermatology*, 19(8), 1840–1845. <https://doi.org/10.1111/jocd.13539>.
- Pikoos, T. D., Buzwell, S., Sharp, G., & Rossell, S. L. (2020). The COVID-19 pandemic: Psychological and behavioral responses to the shutdown of the beauty industry. *International Journal of Eating Disorders*, 53(12), 1993–2002. <https://doi.org/10.1002/eat.23385>.
- Ratajczak, P., Landowska, W., Kopciuch, D., Paczkowska, A., Zaprutko, T., & Kus, K. (2023). The Growing Market for Natural Cosmetics in Poland: Consumer Preferences and Industry Trends. *Clinical, Cosmetic and Investigational Dermatology*, 16, 1877–1892. <https://doi.org/10.2147/CCID.S411032>.
- Swachta, M. (2022). *Analiza unfolding jako narzędzie identyfikacji zmian na rynku kosmetycznym przed i w trakcie pandemii Covid-19* [Unpublished bachelor's thesis]. Wrocław University of Business and Economy.
- Ścieszko, E., Budny, E., Rotsztein, H., & Erkiert-Polguj, A. (2021). How has the pandemic lockdown changed our daily facial skincare habits?. *Journal of Cosmetic Dermatology*, 20(12), 3722–3726. <https://doi.org/10.1111/jocd.14059>.
- Unger, J. K. (2022). *The impact of Covid-19 on buying decision-making: changes in beauty consumer behavior during lockdown* [Master's thesis, LUT University]. <https://lutpub.lut.fi/handle/10024/164690>.
- Walesiak, M., & Dudek, A. (2017). Selecting the Optimal Multidimensional Scaling Procedure for Metric Data with R Environment. *Statistics in Transition new series*, 18(3), 521–540. <https://doi.org/10.59170/stattrans-2017-027>.
- Walesiak, M., & Dudek, A. (2022). *mdsOpt – Searching for Optimal MDS Procedure for Metric and Interval-Valued Data*. <https://cran.r-project.org/web/packages/mdsOpt/vignettes/mdsOpt.pdf>.
- Werner-Lewandowska, K., Lubiński, P., & Słoniec, J. (2021). The Effect of Covid-19 on Consumer Behavior in Poland – Preliminary Research Results. *European Research Studies Journal*, 24(2), 405–416. <https://doi.org/10.35808/ersj/2272>.

Elektroniczne zbiory danych jako źródło informacji o stanie bezpieczeństwa i higieny w miejscu pracy

Electronic datasets as a source of information on the state of health and safety in the workplace

1. Wprowadzenie

System zbierania, gromadzenia i udostępniania informacji o stanie bezpieczeństwa i higieny pracy (bhp) w miejscu jej wykonywania powinien być przede wszystkim użyteczny dla osób, które z niego korzystają. We współczesnym świecie, w którym kluczową rolę odgrywają digitalizacja danych i koncentrowanie wszystkich procesów w jednym narzędziu informatycznym, tradycyjne podejście do prowadzenia rejestrów, wykazów i ewidencji (dalej pod wspólną nazwą: rejestry) okazuje się nieefektywne. Dotychczasowe rozwiązania powodują, że każdy użytkownik systemu ustala zakres informacji o bhp na podstawie własnych potrzeb. Prowadzi to do niespójności systemu, powielania gromadzonych i udostępnianych danych i znacznego obciążenia pracodawców zadaniami administracyjnymi. Zmiany w tym obszarze są niewątpliwie możliwe, lecz wymagają podjęcia działań legislacyjnych oraz stworzenia cyfrowej platformy służącej gromadzeniu danych o bhp i wymianie informacji pomiędzy użytkownikami systemu.

2. Rejestry

W Polsce na pracodawcy spoczywa obowiązek sporządzenia szeregu dokumentów z zakresu bhp. Do podstawowych należy zaliczyć:

- ocenę ryzyka zawodowego (Rozporządzenie Ministra Pracy i Polityki Socjalnej z dnia 26 września 1997 r. w sprawie ogólnych przepisów bezpieczeństwa i higieny pracy);
- rejestr badań i pomiarów czynników szkodliwych (Rozporządzenie Ministra Zdrowia z dnia 2 lutego 2011 r. w sprawie badań i pomiarów czynników szkodliwych dla zdrowia w środowisku pracy);
- karty badań i pomiarów czynników szkodliwych (Rozporządzenie Ministra Zdrowia z dnia 2 lutego 2011 r. w sprawie badań i pomiarów czynników szkodliwych dla zdrowia w środowisku pracy);
- rejestr prac, których wykonywanie powoduje konieczność pozostawania w kontakcie z substancjami chemicznymi, ich mieszaninami, czynnikami lub procesami technologicznymi o działaniu rakotwórczym lub mutagennym, oraz rejestr pra-

owników zatrudnionych przy tych pracach (Rozporządzenie Ministra Zdrowia z dnia 24 lipca 2012 r. w sprawie substancji chemicznych, ich mieszanin, czynników lub procesów technologicznych o działaniu rakotwórczym lub mutagennym w środowisku pracy);

- rejestr wypadków przy pracy (Rozporządzenie Rady Ministrów z dnia 1 lipca 2009 r. w sprawie ustalania okoliczności i przyczyn wypadków przy pracy);
- rejestr obejmujący przypadki stwierdzonych chorób zawodowych i podejrzeń o takie choroby (Ustawa z dnia 26 czerwca 1974 r. – Kodeks pracy);
- wykaz stanowisk pracy, na których są wykonywane prace w szczególnych warunkach lub o szczególnym charakterze (Ustawa z dnia 26 czerwca 1974 r. – Kodeks pracy);
- rejestr prac narażających pracowników na działanie szkodliwych czynników biologicznych i rejestr pracowników zatrudnionych przy tych pracach (Rozporządzenie Ministra Zdrowia z dnia 22 kwietnia 2005 r. w sprawie szkodliwych czynników biologicznych dla zdrowia w środowisku pracy oraz ochrony zdrowia pracowników zawodowo narażonych na te czynniki);
- ewidencję pracowników wykonujących prace w szczególnych warunkach lub o szczególnym charakterze, w przypadku których istnieje obowiązek opłacania składek na Fundusz Emerytur Pomostowych (Ustawa z dnia 26 czerwca 1974 r. – Kodeks pracy).

Wymienione dokumenty są kluczowe, jeśli chodzi o informowanie instytucji zewnętrznych o warunkach pracy, ale na pracodawcy spoczywają też inne obowiązki informacyjne w tym zakresie. Dotyczy to w szczególności sprawozdań i informacji przekazywanych Zakładowi Ubezpieczeń Społecznych (ZUS), Głównemu Urzędowi Statystycznemu (GUS), Państwowej Inspekcji Pracy (PIP) i Państwowej Inspekcji Sanitarnej (PIS):

- do ZUS pracodawcy są zobowiązani corocznie przekazywać informację o danych do ustalenia składki na ubezpieczenie wypadkowe – ZUS IWA (Rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 29 listopada 2002 r. w sprawie różnicowania stopy procentowej składki na ubezpieczenie społeczne z tytułu wypadków przy pracy i chorób zawodowych) oraz ewidencję pracowników wykonujących prace w szczególnych warunkach lub o szczególnym charakterze, za których jest przewidziany obowiązek opłacania składek na Fundusz Emerytur Pomostowych – ZUS ZSWA (Rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 20 grudnia 2020 r. w sprawie określenia wzorów zgłoszeń do ubezpieczeń społecznych i ubezpieczenia zdrowotnego, imiennych raportów miesięcznych i imiennych raportów miesięcznych korygujących, zgłoszeń płatnika składek, deklaracji rozliczeniowych i deklaracji rozliczeniowych korygujących, zgłoszeń danych o pracy

w szczególnych warunkach lub o szczególnym charakterze, raportów informacyjnych, oświadczeń o zamiarze przekazania raportów informacyjnych, informacji o zawartych umowach o dzieło oraz innych dokumentów);

- do GUS należy corocznie przekazywać sprawozdanie o warunkach pracy – Z-10 (Ustawa z dnia 29 czerwca 1995 r. o statystyce publicznej), a jeśli dojdzie do wypadku przy pracy – dodatkowo statystyczną kartę wypadku – Z-KW (Rozporządzenie Ministra Rodziny i Polityki Społecznej z dnia 9 grudnia 2022 r. w sprawie statystycznej karty wypadku przy pracy);
- do właściwego państwowego wojewódzkiego inspektora sanitarnego i okręgowego inspektora pracy pracodawca zobowiązany jest corocznie przekazywać informację o substancjach chemicznych, ich mieszaninach, czynnikach lub procesach technologicznych o działaniu rakotwórczym lub mutagennym (Rozporządzenie Ministra Zdrowia z dnia 24 lipca 2012 r. w sprawie substancji chemicznych, ich mieszanin, czynników lub procesów technologicznych o działaniu rakotwórczym lub mutagennym w środowisku pracy).

Obecny system należy uznać za archaiczny, ponieważ obciąża pracodawców obowiązkiem prowadzenia licznych papierowych lub elektronicznych rejestrów, a także wypełniania określonych formularzy. Warto zauważyć, że dane przekazywane przez pracodawców do instytucji zewnętrznych mają jedynie charakter deklaracyjny, a ich weryfikacja jest bardzo trudna, co prowadzi do rozbieżności między danymi prezentowanymi przez różne instytucje (Pawłowska, 2020).

3. Przykład ZUS IWA

Niedoskonałość systemu przekazywania instytucjom zewnętrznym danych dotyczących warunków pracy można przedstawić na przykładzie informacji o danych do ustalenia składki na ubezpieczenie wypadkowe. Na początku każdego roku pracodawca zatrudniający więcej niż dziewięciu pracowników jest zobowiązany przekazać ZUS – na formularzu papierowym lub elektronicznym – ZUS IWA zawierającą następujące dane:

- swoją nazwę, numer identyfikacji podatkowej (NIP), numer identyfikacyjny REGON lub PESEL i adres korespondencyjny;
- rodzaj działalności gospodarczej według Polskiej Klasyfikacji Działalności (PKD);
- liczbę ubezpieczonych w roku sprawozdawczym;
- liczbę poszkodowanych w wypadkach przy pracy za okres składania informacji, w tym liczbę poszkodowanych w wypadkach śmiertelnych i ciężkich;
- liczbę pracujących w warunkach zagrożenia w roku objętym informacją.

Większość danych zawartych w ZUS IWA to powielone informacje, które pracodawca już wcześniej przekazał ZUS. Dotyczy to przede wszystkim szczegółowych danych identyfikujących płatnika składek na ubezpieczenie społeczne, które zostały podane podczas zgłaszania pierwszej osoby do ubezpieczenia społecznego. Rodzaj działalności gospodarczej według PKD to z kolei informacja ogólnodostępna w bazie REGON. Liczba osób ubezpieczonych w danym okresie jest znana ZUS ze względu na obowiązek zgłaszania wszystkich zatrudnionych do ubezpieczenia społecznego (niezależnie od rodzaju umowy), a poza tym wynika bezpośrednio ze składek na ubezpieczenie społeczne odprowadzanych przez pracodawców. W tych okolicznościach żądanie przekazywania danych po raz kolejny wydaje się nieuzasadnione. Sytuacja wygląda podobnie w przypadku informacji o poszkodowanych w wypadkach przy pracy. Liczba wypadków wynika bezpośrednio z liczby świadczeń wypłacanych z tego tytułu (np. świadczenia za czas niezdolności do pracy z powodu wypadku), natomiast źródłem danych o liczbie wypadków o skutkach ciężkich i śmiertelnych mogłyby stać się dla ZUS zakładowe dokumentacje powypadkowe (protokoły powypadkowe i karty wypadków). ZUS nie ma informacji o liczbie osób pracujących w warunkach zagrożenia. Należy jednak wskazać, że w ostatnim okresie występowały istotne problemy interpretacyjne w zakresie zasad ustalania wartości tego wskaźnika (Smoliński, 2023).

4. Centralna platforma danych z zakresu bhp

Ekspansja nowoczesnych technologii cyfrowych stanowi wyzwanie dla organów administracji publicznej. Tradycyjne formy gromadzenia, przetwarzania i udostępniania danych dotyczących bhp wymagają zmian. Obligowanie pracodawców do prowadzenia wielu rejestrów, a następnie do przekazywania zgromadzonych informacji do różnych instytucji uznano za praktykę przestarzałą i odbiegającą od planu zintegrowanej informatyzacji państwa (Uchwała nr 232/2023 Rady Ministrów z dnia 23 listopada 2023 r. zmieniająca uchwałę w sprawie przyjęcia programu rozwoju „Program Zintegrowanej Informatyzacji Państwa”). Co ważne, na potrzebę doskonalenia systemu rejestracji, analizy i wykorzystania wiarygodnych informacji o wypadkach przy pracy i chorobach zawodowych wskazuje Międzynarodowa Organizacja Pracy (International Labour Organization – ILO). W czasie kampanii społecznej zorganizowanej z okazji Światowego Dnia Bezpieczeństwa i Ochrony Zdrowia w Pracy 2017 koncentrowano się na pilnej potrzebie poprawy zdolności do gromadzenia i wykorzystywania wiarygodnych danych dotyczących bhp (ILO, 2017). Działanie zostało wpisane jako cel Agendy na rzecz Zrównoważonego Rozwoju 2030 (Organizacja Narodów Zjednoczonych, 2015).

Niezbędne wydaje się zatem stworzenie platformy umożliwiającej pracodawcom łatwe gromadzenie wymaganych prawem informacji z zakresu bhp. Nie ma merytorycznych ani ekonomicznych przeszkód, aby zbierać te informacje w centralnym systemie informatycznym. Dane o warunkach pracy, które pracodawca jest z mocy przepisów zobowiązany gromadzić i przechowywać, a następnie archiwizować, są de facto danymi publicznymi, więc wykorzystanie do tego centralnej platformy nie stanowiłoby naruszenia tajemnicy pracodawcy; prowadziłyby za to do poprawy komunikacji między pracodawcami i innymi interesariuszami a państwem. Zdjęłoby także z pracodawców obowiązek archiwizowania tych dokumentów, co pozwoliłoby zredukować obciążenia administracyjne i koszty z tym związane.

Na centralnej platformie – która zastąpiłaby dotychczasowe rejestry prowadzone przez pracodawców – podmioty, osoby i instytucje zamieszczałyby dane z zakresu warunków pracy. Dane te byłyby następnie automatycznie pobierane przez inne uprawnione instytucje. Takie rozwiązanie ułatwiłoby współpracę między pracodawcami a administracją publiczną, w tym ZUS, i innymi podmiotami. Za przykład tego typu platformy mogą posłużyć elektroniczne bazy wypadków przy pracy, np. hiszpański elektroniczny system zgłaszania wypadków przy pracy Delt@ czy aplikacja Injury Tracking Application (ITA) stosowana w Stanach Zjednoczonych.

Z jednej strony dane dotyczące bhp są niezbędne do identyfikowania nowych zagrożeń i niebezpiecznych sektorów, jak również ustalania priorytetów działań, w tym programów prewencyjnych dotyczących zdrowia i bezpieczeństwa pracowników. Z drugiej strony gromadzenie tych danych nie może być zbędnym obciążeniem dla pracodawców i organów administracji państwowej (Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów w sprawie strategicznych ram UE dotyczących bezpieczeństwa i higieny pracy na lata 2014–2020).

Dzięki działaniu platformy zostałyby zrealizowane zalecenia ILO i Unii Europejskiej, a także projekt – zgodny z założeniami statystyki publicznej opartymi na koncepcji *lean* – pełnej automatyzacji procesu zbierania danych wykorzystywanych do badań dotyczących warunków pracy (Auksztol, 2023). Aby zaistniała szansa realizacji projektu, niezbędne jest wprowadzenie licznych zmian legislacyjnych znoszących obowiązki prowadzenia dotychczasowych rejestrów zakładowych i przekazywania informacji o warunkach pracy w tradycyjny sposób. Należałoby także powołać instytucję, która administrowałaby zasobami danych.

Centralna platforma danych z zakresu bhp stanowiłaby także odpowiedź na wzrost znaczenia cyfryzacji komunikacji w życiu społecznym i gospodarczym, poszerzenie zakresu danych gromadzonych przez sektory publiczny i prywatny, rosnące ocze-

kiwania użytkowników danych statystycznych i potrzebę ograniczenia obowiązków sprawozdawczych (Auksztol, 2021).

Stworzenie platformy prowadziłyby do zmniejszenia obciążeń wynikających z konieczności weryfikacji rejestrów. Poza tym ułatwiłyby prace instytucjom kontrolnym, czyli PIS i PIP, ponieważ wiele czynności można byłoby wykonywać zdalnie (Smoliński, 2019).

5. Gromadzenie danych z wykorzystaniem centralnej platformy

Kluczowych informacji o warunkach pracy dostarczają rejestry wypadków przy pracy i chorób zawodowych. Obecnie to pracodawca ma obowiązek prowadzenia własnego rejestru wypadków przy pracy. Na podstawie rejestru powinno być możliwe przeprowadzenie – bez sięgania do dokumentów źródłowych – analizy wypadkowości; ma on także służyć udostępnianiu informacji o wypadkach uprawnionym instytucjom zewnętrznym. Gdyby udało się utworzyć centralną bazę wypadków przy pracy, prowadzenie indywidualnych rejestrów stałoby się niecelowe. Baza zawierałaby dane wymagane w Z-KW. Warto byłoby także rozważyć możliwość zgłaszania wypadków przez samych poszkodowanych i uprawnione organy (np. policję), co zapobiegałoby ukrywaniu wypadków. Ze zgromadzonych danych mogłoby korzystać zarówno pracodawca, jak i uprawnione instytucje, przede wszystkim kontrolne, ZUS i GUS.

Centralna baza wypadków przy pracy pozwalałaby nie tylko na zbieranie danych wynikających z ustaleń dokonanych bezpośrednio po zdarzeniu, lecz także na ich ewentualną modyfikację. Ustalenia zawarte w protokole powypadkowym lub karcie wypadku mogą być poddawane korektom w wyniku postępowań administracyjnych lub sądowych. Każda zmiana mogłaby być odnotowywana w centralnej bazie przez uprawnioną instytucję lub organ w momencie uprawomocnienia się decyzji administracyjnej lub wyroku. W wielu przypadkach początkowe ustalenia ulegają istotnym zmianom w wyniku dalszych postępowań dotyczących przyczyny, skutków i klasyfikacji prawnej zdarzenia. Zgromadzone dane obrazowałyby wypadek nie tylko z punktu widzenia pracodawcy, lecz także z perspektywy związanych z nim postępowań administracyjnych i sądowych. Możliwe byłoby także bezpośrednie porównywanie ustaleń audytora zewnętrznego, czyli inspektora PIP, z danymi raportowanymi przez pracodawcę, a w dalszej kolejności – z decyzjami administracyjnymi i wyrokami sądów. Tak prowadzona baza pozwalałaby na wyeliminowanie rozbieżności występujących obecnie w statystyce GUS, sprawozdaniach PIP i raportach ZUS (Centralny Instytut Ochrony Pracy – Państwowy Instytut Badawczy, 2013). Dane byłyby automatycznie przetwarzane, analizowane i udostępniane uprawnionym osobom, instytucjom i organom. GUS miałby więc możliwość przeprowadzania analiz wypadkowości

na podstawie nie tylko subiektywnych ustaleń pracodawców, lecz także danych zweryfikowanych przez kompetentne instytucje. Obraz wypadkowości byłby wtedy niewątpliwie pełniejszy.

Podobną bazę można byłoby stworzyć w celu rejestracji chorób zawodowych – w miejsce centralnego rejestru prowadzonego przez Instytut Medycyny Pracy (IMP; Rozporządzenie Rady Ministrów z dnia 30 czerwca 2009 r. w sprawie chorób zawodowych). Co ważne, zastępowałaby ona także prowadzone przez pracodawców rejestry, których wymaga kodeks pracy, i umożliwiałaby bieżące monitorowanie sprawy – od zgłoszenia podejrzenia aż do ewentualnej prawomocnej decyzji o uznaniu choroby zawodowej. Tym samym liczba użytkowników bazy chorób zawodowych znacznie by się zwiększyła. Pierwszym z nich byłby lekarz, ponieważ zazwyczaj to on dokonuje rozpoznania podejrzenia choroby zawodowej i zgłasza ją do PIS, która z kolei przeprowadza postępowanie administracyjne i wydaje decyzję o stwierdzeniu choroby zawodowej (możliwe są także postępowania sądowe). Wszystkie te zdarzenia byłyby zarejestrowane w bazie, dzięki czemu nie byłoby potrzeby sięgania do jednostkowych rejestrów dostępnych u pracodawcy i w PIS oraz wypełniania tradycyjnych kart stwierdzenia choroby zawodowej (Rozporządzenie Ministra Zdrowia z dnia 1 sierpnia 2002 r. w sprawie sposobu dokumentowania chorób zawodowych i skutków tych chorób). Zgromadzone w ten sposób dane byłyby udostępniane zainteresowanym instytucjom, w tym IMP i GUS.

Warto byłoby również rozważyć stworzenie kolejnych baz: czynników szkodliwych, prac i pracowników narażonych na czynniki rakotwórcze lub mutagenne, prac i pracowników zatrudnionych w szczególnych warunkach, prac narażających pracowników na działanie szkodliwych czynników biologicznych i pracowników zatrudnionych przy tych pracach. Z jednej strony pozwoliłoby to zwolnić pracodawców z obowiązku prowadzenia wielu tradycyjnych rejestrów, a z drugiej uporządkowałoby sposób raportowania oraz gromadzenia i udostępniania danych instytucjom zewnętrznym.

Ważną rolę w analizie warunków pracy odgrywa informacja dotycząca występowania czynników szkodliwych na poszczególnych stanowiskach pracy. Obecnie informacje te są odnotowywane w kartach oraz rejestrach badań i pomiarów czynników szkodliwych prowadzonych przez poszczególnych pracodawców (najczęściej w formie papierowej). Dane z tych kart i rejestrów są przekazywane do ZUS (w ZUS IWA, jako liczba pracujących w warunkach zagrożenia) i GUS (Z-10). Kontrolę i analizę danych zawartych w rejestrze i kartach przeprowadzają także PIP i PIS. Stworzenie centralnej bazy czynników szkodliwych zwolniłoby pracodawców z obowiązku prowadzenia rejestrów kart badań i pomiarów czynników szkodliwych. Dane z pomiarów laboratoria przekazywałyby automatycznie do centralnej bazy, do której dostęp miałiby zarówno

pracodawca, jak i uprawnione instytucje. Tym samym takie instytucje, jak PIP i PIS mogłyby analizować warunki pracy bez konieczności przeprowadzania kontroli u pracodawcy, a GUS i IMP mieliby bezpośredni dostęp do rzetelnych danych o czynnikach szkodliwych występujących na poszczególnych stanowiskach pracy.

Stworzenie bazy prac i pracowników pracujących w narażeniu na czynniki rakotwórcze i mutagenne pozwoliłoby uniknąć corocznego raportowania do PIP i PIS. Przekazywanie tych danych w sposób tradycyjny znacząco utrudnia ich analizę, co powoduje, że GUS zbiera je dodatkowo w ramach informacji Z-10. Podobnie jest w przypadku prac narażających pracowników na działanie szkodliwych czynników biologicznych i pracowników zatrudnionych przy tych pracach – centralna baza zastąpiłaby rejestry zakładowe, a jednocześnie umożliwiła PIS przeprowadzanie szczegółowych analiz zgromadzonych w niej informacji. GUS i inne zainteresowane instytucje mogłyby wykorzystywać te dane do badań naukowych, analiz i planowania ewentualnych kontroli.

Ważnym elementem systemowego zbioru danych o warunkach pracy jest też informacja w zakresie prac i pracowników zatrudnionych w szczególnych warunkach. Centralna baza umożliwiłaby pogłębioną analizę tych danych.

6. Podsumowanie

Gromadzenie w systemie teleinformatycznym wszystkich danych o bhp powinno się stać powszechną praktyką. Digitalizacja dokumentów i danych z zakresu bhp pozwoli na zwiększenie produktywności zarówno po stronie pracodawców, jak i instytucji uprawnionych do analizowania warunków pracy. Można domniemywać, że z biegiem czasu zakres zbieranych informacji będzie się zwiększać, w tym o informacje z dokumentów dotyczących oceny ryzyka zawodowego. Należy się także spodziewać stopniowej integracji centralnej platformy bhp z oprogramowaniem stosowanym przez zakłady pracy w zakresie zarządzania bezpieczeństwem pracy (Smoliński, 2020). W konsekwencji tak zaplanowanych procesów powstanie zbiór big data – cenny zasób ułatwiający podejmowanie progresywnych działań zmierzających do poprawy warunków pracy.

Bibliografia

Auksztol, J. (2021). Lean official statistics concept based on the working conditions survey. *Wiadomości Statystyczne. The Polish Statistician*, 66(12), 75–97. <https://doi.org/10.5604/01.3001.0015.5595>.

- Auksztol, J. (2023). Koncepcja systemu informacji o bezpieczeństwie i zdrowiu w miejscu pracy. *Wiadomości Statystyczne. The Polish Statistician*, 68(3), 44–60. <https://doi.org/10.59139/ws.2023.03.3>.
- Centralny Instytut Ochrony Pracy – Państwowy Instytut Badawczy. (2013). *Wypadki przy pracy*. https://www.ciop.pl/CIOPPortalWAR/appmanager/ciop/pl?_nfpb=true&_pageLabel=P32000168841444824271997&html_tresc_root_id=553&html_tresc_id=300003722&html_klucz=553#:~:text=Podstawowe%20r%C3%B3C5%BCnice%20w%20publikacjach%20wynikaj%C4%85%20z%20r%C3%B3C5%BC.
- International Labour Organization. (2017). *World Day for Safety and Health at Work 2017 – Optimize the collection and use of OSH data*. https://www.ilo.org/global/topics/safety-and-health-at-work/events-training/events-meetings/world-day-safety-health-at-work/WCMS_617323/lang-en/index.htm.
- Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów w sprawie strategicznych ram UE dotyczących bezpieczeństwa i higieny pracy na lata 2014–2020 ((COM)2014, 322).
- Organizacja Narodów Zjednoczonych. (2015). *Rezolucja przyjęta przez Zgromadzenie Ogólne w dniu 25 września 2015 r. 70/1. Przekształcamy nasz świat: Agenda na rzecz Zrównoważonego Rozwoju 2030*. https://www.unic.un.org.pl/files/164/Agenda%202030_pl_2016_ostateczna.pdf.
- Pawłowska, Z. (2020). Źródła danych o warunkach środowiska pracy w Polsce. *Bezpieczeństwo Pracy. Nauka i Praktyka*, (11), 14–17. <https://doi.org/10.5604/01.3001.0014.4996>.
- Rozporządzenie Ministra Pracy i Polityki Socjalnej z dnia 26 września 1997 r. w sprawie ogólnych przepisów bezpieczeństwa i higieny pracy (t.j. Dz.U. 2003 nr 169, poz. 1650 ze zm.).
- Rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 29 listopada 2002 r. w sprawie różnicowania stopy procentowej składki na ubezpieczenie społeczne z tytułu wypadków przy pracy i chorób zawodowych (t.j. Dz.U. 2022 poz. 2189 ze zm.).
- Rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 20 grudnia 2020 r. w sprawie określenia wzorów zgłoszeń do ubezpieczeń społecznych i ubezpieczenia zdrowotnego, imiennych raportów miesięcznych i imiennych raportów miesięcznych korygujących, zgłoszeń płatnika składek, deklaracji rozliczeniowych i deklaracji rozliczeniowych korygujących, zgłoszeń danych o pracy w szczególnych warunkach lub o szczególnym charakterze, raportów informacyjnych, oświadczeń o zamiarze przekazania raportów informacyjnych, informacji o zawartych umowach o dzieło oraz innych dokumentów (t.j. Dz.U. 2023 poz. 2032).
- Rozporządzenie Ministra Rodziny i Polityki Społecznej z dnia 9 grudnia 2022 r. w sprawie statystycznej karty wypadku przy pracy (Dz.U. 2022 poz. 2750).
- Rozporządzenie Ministra Zdrowia z dnia 1 sierpnia 2002 r. w sprawie sposobu dokumentowania chorób zawodowych i skutków tych chorób (t.j. Dz.U. 2013 poz. 1379 ze zm.).
- Rozporządzenie Ministra Zdrowia z dnia 22 kwietnia 2005 r. w sprawie szkodliwych czynników biologicznych dla zdrowia w środowisku pracy oraz ochrony zdrowia pracowników zawodowo narażonych na te czynniki (Dz.U. 2005 nr 81, poz. 716 ze zm.).
- Rozporządzenie Ministra Zdrowia z dnia 2 lutego 2011 r. w sprawie badań i pomiarów czynników szkodliwych dla zdrowia w środowisku pracy (t.j. Dz.U. 2023 poz. 419).

- Rozporządzenie Ministra Zdrowia z dnia 24 lipca 2012 r. w sprawie substancji chemicznych, ich mieszanin, czynników lub procesów technologicznych o działaniu rakotwórczym lub mutagenym w środowisku pracy (t.j. Dz.U. 2021 poz. 2235 ze zm.).
- Rozporządzenie Rady Ministrów z dnia 30 czerwca 2009 r. w sprawie chorób zawodowych (t.j. Dz.U. 2022 poz. 1836).
- Rozporządzenie Rady Ministrów z dnia 1 lipca 2009 r. w sprawie ustalania okoliczności i przyczyn wypadków przy pracy (Dz.U. 2009 nr 105, poz. 870).
- Smoliński, D. (2019). Inspekcja 4.0. *Inspektor Pracy*, (1), 39–41. <https://www.pip.gov.pl/ua/dla-sluzb-bhp/magazyn-inspektor-pracy>.
- Smoliński, D. (2020). E-bezpieczeństwo i ochrona zdrowia. *Atest – Ochrona pracy*, (6), 13–16.
- Smoliński, D. (2023). Glosa krytyczna do wyroku Sądu Najwyższego z 8 grudnia 2020 r. w zakresie sposobu ustalania liczby zatrudnionych w warunkach zagrożenia w rozumieniu art. 31 ust. 3 pkt 3 ustawy z 30 października 2002 r. o ubezpieczeniu społecznym z tytułu wypadków przy pracy i chorób zawodowych. *Zabezpieczenie Społeczne. Teoria, Prawo, Praktyka*, (17), 49–59. <https://doi.org/10.31971/2299-2332.2023.17.4>.
- Uchwała nr 232/2023 Rady Ministrów z dnia 23 listopada 2023 r. zmieniająca uchwałę w sprawie przyjęcia programu rozwoju „Program Zintegrowanej Informatyzacji Państwa”.
- Ustawa z dnia 26 czerwca 1974 r. – Kodeks pracy (t.j. Dz.U. 2023 poz. 1465).
- Ustawa z dnia 29 czerwca 1995 r. o statystyce publicznej (t.j. Dz.U. 2023 poz. 773).

Dariusz Smoliński

Państwowa Inspekcja Pracy, Okręgowy Inspektorat Pracy w Gdańsku,
Oddział w Słupsku, Polska / State Labour Inspection, District Labour Inspectorate in Gdańsk,
Branch in Słupsk, Poland

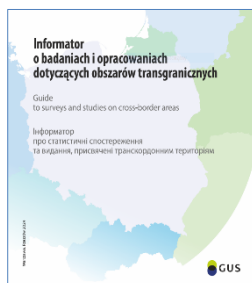
WYDAWNICTWA GUS. MARZEC 2024 PUBLICATIONS OF STATISTICS POLAND. MARCH 2024

W ofercie wydawniczej Głównego Urzędu Statystycznego z ubiegłego miesiąca warto zwrócić uwagę na następującą publikację:

Among Statistics Poland's publications from the previous month, we would like to recommend:

Informator o badaniach i opracowaniach dotyczących obszarów transgranicznych **Guide to surveys and studies on cross-border areas**

Piąta edycja informatora zawierającego ogólne dane o prowadzonych badaniach obszarów przygranicznych oraz o euroregionach i prezentującego wydane w ciągu ostatnich dwóch lat publikacje poświęcone tej tematyce. Za opracowanie merytoryczne tego wydania odpowiadał Urząd Statystyczny w Rzeszowie we współpracy z urzędami statystycznymi w Białymstoku, Lublinie, Opolu, Wrocławiu i Zielonej Górze.



Język: polski, angielski, ukraiński

Language: Polish, English, Ukrainian

Seria: Foldery i publikacje okolicznościowe

Series: Brochures and celebratory publications

Dostępne wersje: drukowana i elektroniczna

Available in: printed and electronic form

Nawiązywanie i podtrzymywanie dobrych relacji między sąsiadującymi ze sobą krajami pozytywnie wpływa na rozwój gospodarczy regionów przygranicznych. Wiąże się z tym duże zapotrzebowanie na informacje na temat tych obszarów. Wychodząc mu naprzeciw, statystyka publiczna przeprowadza wielowymiarowe i ponadnarodowe badania regionów przygranicznych, obejmujące różne aspekty zjawisk społeczno-gospodarczych. Co ważne, przy tej okazji zacieśnia się współpraca między służbami statystyki publicznej w Polsce i innych krajach. Z *Informatora* można dowiedzieć się m.in., jak rozwijają się badania obszarów transgranicznych oraz jakie euroregiony znajdują się w granicach Polski, a także zapoznać się z ich krótką charakterystyką. W publikacji zamieszczono też informacje o źródłach danych dotyczących obszarów przygranicznych.

W marcu br. ukazały się ponadto:

- *Bezrobocie rejestrowane 1–4 kwartał 2023 r.*;
- „Biuletyn statystyczny” nr 2/2024;

- *Ceny robót budowlano-montażowych i obiektów budowlanych (styczeń 2024 r.);*
- *Jak korzystamy z Internetu? 2023;*
- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2024 (marzec 2024);*
- *Nauka i technika w 2022 r.;*
- *Obszary wiejskie w Polsce w 2022 r.;*
- *Produkcja ważniejszych wyrobów przemysłowych w lutym 2024 r.;*
- *Rachunek podaży i wykorzystania wyrobów i usług w 2020 r.;*
- *„Statistics in Transition new series” nr 1/2024;*
- *Sytuacja społeczno-gospodarcza kraju w lutym 2024 r.;*
- *Zeszyt metodologiczny. Warunki pracy – zatrudnieni w warunkach zagrożenia;*
- *Zmiany strukturalne grup podmiotów gospodarki narodowej w rejestrze REGON, 2023 r.*

Joanna Sadowy

Główny Urząd Statystyczny, Departament Opracowań Statystycznych, Polska
Statistics Poland, Statistical Products Department, Poland

Wszystkie publikacje GUS w wersji elektronicznej są dostępne na stronie stat.gov.pl/publikacje/publikacje-a-z. Wersje drukowane (jeśli zostały wydane) można zamawiać pod adresem: zws-sprzedaz@stat.gov.pl.

All the publications of Statistics Poland available in electronic form can be accessed at stat.gov.pl/en/publications. Printed versions (if available) may be ordered at: zws-sprzedaz@stat.gov.pl.

DLA AUTORÓW FOR THE AUTHORS

(for the English translation of the information given below, please visit ws.stat.gov.pl/ForAuthors)

W „Wiadomościach Statystycznych. The Polish Statistician” („WS”) zamieszczane są artykuły o charakterze naukowym poświęcone teorii i praktyce statystycznej, które prezentują wyniki oryginalnych badań teoretycznych lub analitycznych wykorzystujących metody statystyki matematycznej, opisowej bądź ekonometrii. Ukazują się również artykuły przeglądowe, recenzje publikacji naukowych oraz inne opracowania informacyjne. W czasopiśmie publikowane są prace w języku polskim i angielskim.

Od 2007 r. „WS” znajdują się na liście czasopism naukowych MEiN. Zgodnie z komunikatem Ministra Edukacji i Nauki z dnia 1 grudnia 2021 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych „WS” otrzymały 70 punktów.

„Wiadomości Statystyczne. The Polish Statistician” są udostępniane w następujących bazach, repozytoriach, katalogach i wyszukiwarkach: Agro, BazEkon, Biblioteka Nauki, Central and Eastern European Academic Source (CEEAS), Central and Eastern European Online Library (CEEOL), Central European Journal of Social Sciences and Humanities (CEJSH), Directory of Open Access Journals (DOAJ), EBSCO Discovery Service, European Reference Index for the Humanities and Social Sciences (ERIH Plus), Exlibris Primo, Google Scholar, ICI Journals Master List, ICI World of Journals, Norwegian Register for Scientific Journals and Publishers (The Nordic List) oraz Summon.

Za publikację artykułów na łamach „WS” autorzy nie otrzymują honorariów ani nie wnoszą opłat.

1. Zgłaszanie artykułów

Prace przeznaczone do opublikowania w „WS” należy przysyłać za pośrednictwem platformy Editorial System: www.editorialsystem.com/ws.

Zgłaszany artykuł powinien być zanonimizowany, tj. pozbawiony informacji o autorze/autorach (również we właściwościach pliku), podziękowań i informacji o źródłach finansowania, a także innych informacji wskazujących na afiliację lub umożliwiających zidentyfikowanie autora. Jeżeli w pracy występują tablice, wykresy lub mapy, powinny być umieszczone w treści artykułu. Materiały graficzne, razem z danymi do nich, należy ponadto załączyć jako osobny plik / osobne pliki, najlepiej w formacie Excel. **Prosimy o niestosowanie stylów i ograniczenie formatowania do wymogów redakcyjnych.** Więcej informacji w pkt 4 *Wymogi redakcyjne*.

Razem z artykułem należy przesłać skan/zdjęcie oświadczenia o oryginalności pracy i nielożeniu jej w innym wydawnictwie. **Załączenie oświadczenia jest warunkiem poddania pracy ocenie wstępnej i skierowania do recenzji.**

Zgłoszenie artykułu do opublikowania w „WS” oznacza zgodę na jego udostępnienie na licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0).

Autorzy mają prawo do samodzielnego umieszczania w wybranych przez siebie repozytoriach artykułu w wersji zarówno zgłoszonej do „WS”, jak i zaakceptowanej do opublikowania

oraz opublikowanej, z zastrzeżeniem wymogu niezwłocznego podania w repozytorium informacji o numerze „WS”, w którym praca się ukazała, wraz z linkiem do niej (DOI).

2. Przebieg prac redakcyjnych

Zgłoszony artykuł jest oceniany i opracowywany w czteroetapowym procesie:

1. **Ocena wstępna**, dokonywana przez redakcję. Polega na weryfikacji naukowego charakteru artykułu oraz jego struktury i zawartości pod kątem wymogów redakcyjnych, a także zgodności tematyki z profilem czasopisma. Autor uzupełnia i poprawia artykuł stosownie do uwag redakcji, a w przypadku nieuwzględnienia danej uwagi uzasadnia swoje stanowisko. Warunkiem skierowania pracy do recenzji jest potwierdzenie oryginalności tekstu uzyskane za pomocą systemu antyplagiatowego. W przypadku wykrycia znacznego podobieństwa do innych prac artykuł zostanie odrzucony.
2. **Ocena recenzentów**, dokonywana przez specjalistów w danej dziedzinie. Artykuł oceniają dwaj recenzenci spoza jednostki naukowej, przy której afiliowany jest autor; w przypadku pracy w języku angielskim co najmniej jeden recenzent jest afiliowany przy jednostce zagranicznej. W razie sprzecznych opinii dwóch recenzentów powoływany jest trzeci recenzent. Recenzenci kierują się kryteriami oryginalności i jakości opracowania zarówno w odniesieniu do treści, jak i formy artykułu.

Autorzy artykułów, które otrzymały pozytywne oceny, wprowadzają poprawki zalecane przez recenzentów i przesyłają zmodyfikowaną wersję pracy. Jeśli pojawi się różnica zdań dotycząca zasadności proponowanych zmian, autorzy są zobligowani do uzasadnienia swojego stanowiska.

3. **Ocena Kolegium Redakcyjnego (KR)**, decydująca o przyjęciu pracy do publikacji. Jest dokonywana na podstawie recenzji, z uwzględnieniem opinii redaktorów tematycznego i merytorycznego. Polega m.in. na weryfikacji dokonania przez autora zmian w artykule stosownie do uwag recenzentów. KR ocenia artykuł pod względem poprawności i spójności merytorycznej oraz zaleca autorowi wprowadzenie poprawek, jeśli są one konieczne, aby praca spełniała wymogi czasopisma. Autorowi przysługuje prawo do odwołania od decyzji o niepublikowaniu artykułu. W takim przypadku powinien on skontaktować się z redakcją „WS” i przedstawić uzasadnienie. Ostateczna decyzja w tej sprawie należy do redaktora naczelnego.

W „WS” publikowane są wyłącznie te artykuły, które otrzymają pozytywną ocenę na każdym z wymienionych etapów i zostaną poprawione przez autora zgodnie z otrzymanymi uwagami (chyba że autor przedstawi argumenty uzasadniające nieuwzględnienie danej uwagi).

Artykuły przyjęte przez KR do publikacji są zamieszczane na stronie internetowej czasopisma w zakładce Early View, gdzie znajdują się do czasu opublikowania w konkretnym wydaniu „WS”.

4. **Opracowanie redakcyjne, autoryzacja i korekta**. Artykuł zakwalifikowany do druku jest poddawany opracowaniu merytorycznemu i językowemu. Redakcja zastrzega sobie prawo do zmiany tytułu i śródtytułów, modyfikowania tablic, wykresów i innych elementów graficznych oraz przededagowania treści bez naruszenia zasadniczej myśli autora.

Po opracowaniu redakcyjnym artykuł jest przesyłany do autoryzacji. Tekst zatwierdzony przez autora, po składzie i łamaniu, jest poddawany korekcie i rewizji (II korekcje).

Autor dokonuje korekty autorskiej tekstu na etapie rewizji. Wykresy i inne materiały graficzne są opracowywane na podstawie plików i danych przekazanych przez autora i poddawane korekcie i rewizji. Autor dokonuje ich akceptacji na etapie rewizji.

W przypadku odkrycia błędów w opublikowanym artykule zamieszcza się na łamach „WS” sprostowanie, a artykuł w wersji elektronicznej jest poprawiany i umieszczany na stronie internetowej „WS” ze stosownym wyjaśnieniem.

3. Zasady etyki publikacyjnej COPE

Redakcja „WS” podejmuje wszelkie starania w celu utrzymania najwyższych standardów etycznych zgodnie z wytycznymi Komitetu ds. Etyki Publikacyjnej (COPE), dostępnymi na stronie internetowej www.publicationethics.org, oraz wykorzystuje wszystkie możliwe środki mające na celu zapobieżenie nadużyciom i nierzetelności autorskiej. Przyjęte zasady postępowania obowiązują autorów, Radę Naukową, Kolegium Redakcyjne, redakcję, pracowników Wydziału Czasopism Naukowych GUS, recenzentów i wydawcę.

3.1. Odpowiedzialność autorów

1. Artykuły naukowe kierowane do opublikowania w „WS” powinny zawierać precyzyjny opis badanych zjawisk i stosowanych metod oraz autorskie wnioski i sugestie dotyczące rozwoju badań i analiz statystycznych. Autorzy powinni wyraźnie określić cel artykułu oraz jasno przedstawić wyniki przeprowadzonej analizy. Prezentacja efektów badań statystycznych zaprojektowanych i przeprowadzonych przez autorów wymaga opisanie zastosowanej w nich metodologii. W przypadku nowatorskich metod analizy pożądanym jest podanie przykładu ilustrującego ich zastosowanie w praktyce statystycznej. Autorzy ponoszą odpowiedzialność za treści prezentowane w artykułach. W razie zgłaszania przez czytelników zastrzeżeń odnoszących się do tych treści autorzy są zobligowani do udzielenia odpowiedzi za pośrednictwem redakcji.
2. Na autorach spoczywa obowiązek zapewnienia pełnej oryginalności przedłożonych prac. Redakcja nie toleruje przejawów nierzetelności naukowej autorów, takich jak:
 - duplikowanie publikacji – ponowne publikowanie własnego utworu lub jego części;
 - plagiat – przywłaszczenie cudzego utworu lub jego fragmentu bez podania informacji o źródle;
 - fabrykowanie danych – oparcie pracy naukowej na nieprawdziwych wynikach badań;
 - autorstwo widmo (*ghost authorship*) – nieujawnianie współautorów, mimo że wnieśli oni istotny wkład w powstanie artykułu;
 - autorstwo gościnne (*guest authorship*) – podawanie jako współautorów osób o znikomym udziale lub niebiorących udziału w opracowywaniu artykułu;
 - autorstwo grzecznościowe (*gift authorship*) – podawanie jako współautorów osób, których wkład jest oparty jedynie na słabym powiązaniu z badaniem.

Autorzy deklarują w stosownym oświadczeniu, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i nie został złożony w innym wydawnictwie oraz że jest ich oryginalnym dziełem, i określają swój wkład w opracowanie artykułu. Jeżeli doszło do zaprezentowania podobnych materiałów podczas konferencji lub

symposium naukowe, to podczas składania tekstu do publikacji w „WS” autorzy są zobowiązani poinformować o tym redakcję.

3. Autorzy są zobowiązani do podania w treści artykułu wszelkich źródeł finansowania badań będących podstawą pracy.
4. Główną odpowiedzialność za rzetelność przekazanych informacji, łącznie z informacją na temat wkładu poszczególnych współautorów w powstanie artykułu, ponosi zgłaszający artykuł.
5. Autorzy zgłaszający artykuły do publikacji w „WS” biorą udział w procesie recenzji double-blind peer review, dokonywanej przez co najmniej dwóch niezależnych ekspertów z danej dziedziny. Po otrzymaniu pozytywnych recenzji autorzy wprowadzają zalecane przez recenzentów poprawki i dostarczają redakcji zaktualizowaną wersję opracowania wraz z pisemnym poświadczeniem uwzględnienia poprawek. Jeśli pojawi się różnica zdań co do zasadności proponowanych zmian, należy wyjaśnić, które poprawki zostały uwzględnione, a w przypadku ich nieuwzględnienia – uzasadnić swoje stanowisko.
6. Jeżeli autorzy odkryją w swoim maszynopisie lub tekście już opublikowanym błędy, nieścisłości bądź niewłaściwe dane, powinni niezwłocznie poinformować o tym redakcję w celu dokonania korekty, wycofania tekstu lub zamieszczenia sprostowania. W przypadku korekty artykułu już opublikowanego jego nowa wersja jest zamieszczana na stronie internetowej „WS” wraz ze stosownym wyjaśnieniem.

3.2. Odpowiedzialność Rady Naukowej, Kolegium Redakcyjnego i Wydziału Czasopism Naukowych GUS

1. Rada Naukowa (RN) kształtuje profil programowy czasopisma, określa kierunki jego rozwoju i konsultuje jego zakres merytoryczny.
2. Kolegium Redakcyjne (KR) podejmuje decyzję o publikacji danego artykułu z uwzględnieniem ocen recenzentów oraz opinii zespołu redakcyjnego. W swoich rozstrzygnięciach członkowie KR kierują się kryteriami merytorycznej oceny wartości artykułu, jego oryginalności i jasności przekazu, a także ścisłego związku z celem i zakresem tematycznym „WS”. Oceniają artykuły niezależnie od płci, rasy, pochodzenia etnicznego, narodowości, religii, wyznania, światopoglądu, niepełnosprawności, wieku lub orientacji seksualnej ich autorów.
3. Zespół redakcyjny, wyodrębniony z KR, tworzą redaktor naczelny i jego zastępca, redaktorzy tematyczni i redaktor merytoryczny. Członkowie zespołu redakcyjnego weryfikują nadsyłane artykuły pod względem merytorycznym, oceniają ich zgodność z celem i zakresem tematycznym „WS” oraz sprawdzają spełnienie wymogów redakcyjnych i przestrzeganie zasad rzetelności naukowej. Ponadto wybierają recenzentów w taki sposób, aby nie wystąpił konflikt interesów, i dbają o zapewnienie uczciwego, bezstronnego i terminowego procesu recenzowania.
4. Za sprawny przebieg procesu wydawniczego, poinformowanie wszystkich jego uczestników o konieczności przestrzegania obowiązujących zasad i przygotowanie artykułów do publikacji odpowiadają pracownicy Wydziału Czasopism Naukowych (WCN) GUS. W celu uzyskania obiektywnej oceny oryginalności nadsyłanych artykułów przed skierowaniem ich do recenzji WCN wykorzystuje system antyplagiatowy. Informacje dotyczące

artykułu mogą być przekazywane przez WCN wyłącznie autorom, recenzentom, członkom RN i KR oraz wydawcy.

5. Zmiany dokonane w tekście na etapie przygotowania artykułu do publikacji nie mogą naruszać zasadniczej myśli autorów. Wszelkie modyfikacje o charakterze merytorycznym są z nimi konsultowane.
6. W przypadku podjęcia decyzji o niepublikowaniu artykułu nie może on zostać w żaden sposób wykorzystany przez wydawcę lub uczestników procesu wydawniczego bez pisemnej zgody autorów. Autorzy mogą się odwołać od decyzji o niepublikowaniu artykułu. W tym celu powinni się skontaktować z redaktorem naczelnym lub sekretarzem redakcji „WS” i przedstawić stosowną argumentację. Odwołania autorów są rozpatrywane przez redaktora naczelnego.
7. Członkowie RN i KR ani pracownicy WCN nie mogą pozostawać w jakimkolwiek konflikcie interesów w odniesieniu do artykułów zgłaszanych do publikacji. Przez konflikt interesów należy rozumieć sytuację, w której jakiegokolwiek interesy lub zależności (służbowe, finansowe lub inne) mogą mieć wpływ na ocenę artykułu lub decyzję o jego publikacji.
8. W celu przeciwdziałania nierzetelności naukowej wymagane jest złożenie przez autorów oświadczenia, w którym deklarują, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i jest ich oryginalnym dziełem, a także określają swój wkład w opracowanie artykułu.
9. W celu zapewnienia wysokiej jakości recenzji wymagane jest złożenie przez recenzentów oświadczenia o przestrzeganiu zasad etyki recenzowania COPE i niewystępowaniu konfliktu interesów.
10. W przypadku uzasadnionego podejrzenia na jakimkolwiek etapie procesu wydawniczego, że autorzy dopuścili się nierzetelności naukowej (zob. pkt 3.1. Odpowiedzialność autorów), zespół redakcyjny skrupulatnie zbada sprawę ewentualnego nadużycia. Jeśli nierzetelność autorów zostanie udowodniona, to zgłoszony przez nich artykuł zostanie odrzucony przez KR, a autorzy otrzymają informację o podjętej decyzji wraz z jej uzasadnieniem.
11. Czytelnicy, którzy mają wobec autorów opublikowanego artykułu uzasadnione podejrzenia o nierzetelność naukową, powinni powiadomić o tym redaktora naczelnego lub sekretarza redakcji. Po zbadaniu sprawy ewentualnego nadużycia czytelnicy zostaną poinformowani o rezultacie przeprowadzonego postępowania. W przypadku potwierdzenia nadużycia, na łamach czasopisma zostanie zamieszczona stosowna informacja.

3.3. Odpowiedzialność recenzentów

1. Recenzenci przyjmują artykuł do recenzji tylko wtedy, gdy uznają, że:
 - posiadają odpowiednią wiedzę w określonej dziedzinie, aby rzetelnie ocenić pracę;
 - zgodnie z ich stanem wiedzy nie istnieje konflikt interesów w odniesieniu do autorów, przedstawionych w artykule badań i instytucji je finansujących, co potwierdzają w oświadczeniu;
 - mogą wywiązać się z terminu ustalonego przez redakcję, aby nie opóźnić publikacji.
2. Recenzenci są zobligowani do zachowania obiektywności i poufności oraz powstrzymania się od osobistej krytyki. Zawsze powinni uzasadnić swoją ocenę, przedstawiając stosowną argumentację.

3. Recenzenci powinni wskazać ważne dla wyników badań opublikowane prace, które w ich ocenie powinny zostać przywołane w ocenianym artykule.
4. W razie stwierdzenia wysokiego poziomu zbieżności treści recenzowanej pracy z innymi opublikowanymi materiałami lub podejrzenia innych przejawów nierzetelności naukowej recenzenci są zobowiązani poinformować o tym redakcję.
5. Po ukończeniu recenzji przechowywanie przesłanych przez redakcję materiałów (w jakiegokolwiek formie) oraz posługiwanie się nimi przez recenzentów jest niedozwolone.

3.4. Odpowiedzialność wydawcy

1. Materiały opublikowane w „WS” są chronione prawem autorskim.
2. Wydawca udostępnia pełną treść wszystkich artykułów w internecie w trybie otwartego dostępu, tj. bezpłatnie i bez technicznych ograniczeń, od 1 stycznia 2022 r. na licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0). W przypadku artykułów zgłoszonych do „WS” od 2022 r. dozwolone jest dzielenie się artykułem (kopiowanie i rozpowszechnianie go w dowolnym medium i formie) oraz adaptowanie go (w dowolnym celu, także komercyjnym) na warunkach określonych w tej licencji. Z pozostałych artykułów zamieszczonych w czasopiśmie można korzystać w ramach otwartego dostępu, zgodnie z ustawą o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego.
3. Wydawca deklaruje gotowość do opublikowania poprawek, wyjaśnień oraz przeprosin.

4. Wymogi redakcyjne

Zgodnie z wymogami czasopisma omawiany w artykule problem badawczy powinien być jednoznacznie zdefiniowany oraz istotny dla oceny zjawisk społecznych lub gospodarczych. Artykuł powinien zawierać wyraźnie określony cel badania, precyzyjny opis badanych zjawisk i stosowanych metod, uzyskane wyniki przeprowadzonej analizy oraz autorskie wnioski.

4.1. Struktura i zawartość artykułu

Wymagane elementy artykułu:

1. Tytuł.
2. Dane autora: imię/imiona i nazwisko, afiliacja w języku polskim i angielskim, ORCID, wkład w powstanie artykułu, adres e-mail. Wśród autorów artykułu wieloautorskiego należy wskazać autora korespondencyjnego.
3. Streszczenie (zalecana objętość – do 1200 znaków ze spacjami, forma bezosobowa). W przypadku artykułu opisującego badanie empiryczne powinno zawierać: cel, przedmiot, okres i metodę badania, źródła danych i najważniejsze wnioski z badania. W przypadku artykułów o innym charakterze należy podać co najmniej cel pracy, jej przedmiot i najważniejsze wnioski.

Streszczenie to podstawowe źródło informacji o artykule, warunkujące też decyzję czytelnika o zapoznaniu się z całą pracą. Dlatego powinno być przygotowane ze szczególną starannością i dbałością o umieszczenie w nim wszystkich wymaganych elementów.

4. Słowa kluczowe – najistotniejsze pojęcia lub wyrażenia użyte w pracy (nie mniej niż trzy). Powinny być zawarte w streszczeniu i/lub tytule.
5. Kod/kody z klasyfikacji Journal of Economic Literature (JEL).
6. Tłumaczenie tytułu, streszczenia i słów kluczowych (na język angielski w przypadku artykułu napisanego w języku polskim, a na język polski w przypadku artykułu napisanego w języku angielskim).
7. W artykule opisującym badanie empiryczne wymagane są następujące części:
 - wprowadzenie, zawierające syntetyczne przedstawienie zagadnień teoretycznych, uzasadnienie podjęcia danego problemu badawczego, cel badania i krytyczne odniesienie do literatury przedmiotu. W wyjątkowych przypadkach, kiedy istotne dla podjętego tematu jest obszerniejsze przedstawienie dyskusji toczącej się w literaturze, przegląd literatury może stanowić odrębną część artykułu;
 - metoda badania, uwzględniająca przedmiot i okres badania, źródła danych i zastosowane metody badawcze, w tym uzasadnienie ich wyboru;
 - wyniki badania – analiza danych oraz interpretacja wyników i odniesienie ich do rezultatów wcześniejszych badań (dyskusja). W uzasadnionych przypadkach dyskusja może stanowić odrębną część artykułu;
 - podsumowanie, które powinno być zwarte i odzwierciedlać istotę problemu badawczego przedstawionego w artykule, bez podawania danych liczbowych; końcowe wnioski powinny odnosić się do treści artykułu, a w szczególności do celu badania.Wszystkie części powinny być opatrzone numerami.
8. Bibliografia, zawierająca pełny wykaz prac i materiałów przywołanych w artykule, przygotowana zgodnie z wymogami czasopisma.

4.2. Przygotowanie artykułu

1. Artykuł powinien być utrzymany w formie bezosobowej.
2. Tekst należy zapisać alfabetem łańciskim. Nazwy własne, tytuły itp. oryginalnie zapisane innym alfabetem powinny być poddane transliteracji.
3. Nie należy stosować stylów; formatowanie należy ograniczyć do wymogów redakcyjnych.
4. Objętość artykułu łącznie ze streszczeniem, słowami kluczowymi, bibliografią, tablicami, wykresami i innymi materiałami graficznymi nie powinna być mniejsza niż 10 stron maszynopisu ani przekraczać 20 stron.
5. Edytor tekstu: Microsoft Word, format *.doc lub *.docx.
6. Krój czcionki:
 - Arial – tytuł, autor, streszczenie, słowa kluczowe, kody JEL, śródtytuły, elementy graficzne (tablice, zestawienia, wykresy, schematy), przypisy;
 - Times New Roman – tekst główny, bibliografia.
7. Wielkość czcionki:
 - 14 pkt – tytuł, autor, śródtytuły wyższego rzędu;
 - 12 pkt – tekst główny, śródtytuły niższego rzędu;
 - 10 pkt – pozostałe elementy.
8. Marginesy – 2,5 cm z każdej strony.

9. Interlinia – 1,5 wiersza; tablice i przypisy – 1 wiersz; przed tytułami rozdziałów i podrozdziałów oraz po nich – pusty wiersz.
10. Wcięcie akapitowe – 0,4 cm; bibliografia – bez wcięcia, wysunięcie 0,4 cm.
11. Przy wycienieniach należy posłużyć się listą punktowaną z punktorami w postaci kropek (wysunięcie 0,4 cm, wcięcie 0 cm); wiersze (oprócz ostatniego) zakończone średnikiem.
12. Strony ponumerowane automatycznie.
13. Tablice i elementy graficzne (wykresy, mapy, schematy) muszą być przywołane w tekście.
14. Wykresy, mapy i schematy należy zamieścić w tekście głównym. Wykresy powinny być edytowalne (optymalnie wykonane w programie Excel; w przypadku wykonania w programie graficznym powinny mieć postać wektorową). Wykresy i inne materiały graficzne należy przekazać osobno, najlepiej w pliku programu Excel lub innym edytowalnym w pakiecie Microsoft Office.
15. Tablice muszą być edytowalne. Nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
16. Wskazówki dotyczące opracowywania map znajdują się w publikacji *Mapy statystyczne. Opracowanie i prezentacja danych*, dostępnej na stronie internetowej GUS.
17. Pod tablicami i każdym elementem graficznym należy podać źródło opracowania, a także objaśnić użyte w nich skróty i symbole.
18. Literowe symbole liczb i innych wielkości niezłożonych należy zapisywać małą lub dużą literą i pismem pochyłym (np. a , A , $y(x)$, a_i); wektorów – pismem pochyłym i pogrubionym (np. \mathbf{a} , \mathbf{A} , \mathbf{w} , $\mathbf{y}(x)$, \mathbf{w}_i); macierzy – pismem prostym i pogrubionym (np. \mathbf{A} , \mathbf{a} , \mathbf{M} , $\mathbf{Y}(x)$, \mathbf{M}_i).
19. Objasnienia znaków umownych i zapisów w tablicach: kreska (–) – zjawisko nie wystąpiło; zero (0) – zjawisko istniało w wielkości mniejszej od 0,5; (0,0) – zjawisko istniało w wielkości mniejszej od 0,05; kropka (.) – brak informacji, konieczność zachowania tajemnicy statystycznej, wypełnienie pozycji jest niemożliwe lub niecelowe; „w tym” – oznacza, że nie podaje się wszystkich składników sumy.
20. Stosowane są następujące skróty: tablica – tabl., wykres – wykr.
21. Wszystkie zawarte w artykule informacje, dane i stwierdzenia wykraczające poza wiedzę powszechną – np. wyniki badań innych autorów, zarówno o charakterze empirycznym, jak i koncepcyjnym – muszą być opatrzone przypisem bibliograficznym. Przez wiedzę powszechną należy rozumieć informacje ogólnie znane i niebudzące wątpliwości ani kontrowersji w danej grupie społecznej, np. utworzenie GUS w 1918 r. lub powstanie UE w 1993 r. na podstawie traktatu z Maastricht. Natomiast dane statystyczne udostępniane lub publikowane np. przez GUS lub Eurostat nie należą do takich informacji. Charakteru wiedzy powszechnej nie mają również stwierdzenia odnoszące się do idei, zjawisk i procesów społecznych, politycznych czy gospodarczych. Nawet pozornie zdroworozsądkowe idee zmieniają bowiem swój sens w zależności od kultury, języka lub dyscypliny naukowej, a także bywają w rozmaity sposób konceptualizowane, jak np. pojęcie poznania w naukach społecznych.

Podanie źródła jest konieczne niezależnie od tego, czy informacje lub stwierdzenia są ujęte w ramy cytatu, czy przedstawione bez dosłownego przytoczenia, np. w formie parafrazy. Jeżeli stwierdzenie może budzić jakiegokolwiek wątpliwości odbiorców, autor powinien wskazać stosowne źródło podawanej informacji.

22. Przypisy rzeczowe, słownikowe lub informacyjne należy umieszczać na dole strony. Przypisy bibliograficzne, zgodnie ze standardem APA (American Psychological Association), należy podawać w tekście głównym.
23. Bibliografię należy przygotować zgodnie ze standardem APA.

4.3. Zasady przywoływania publikacji w treści artykułu

Wyszczególnienie	Przykład przywołania	
	w odsyłaczu	w treści zdania
Autor indywidualny		
Jeden autor	(Iksiński, 2001)	Iksiński (2001)
Dwóch autorów	(Iksiński i Nowak, 1999)	Iksiński i Nowak (1999)
Trzech autorów lub więcej	(Jankiewicz i in., 2003)	Jankiewicz i in. (2003)
Autor instytucjonalny		
Nazwa funkcjonuje jako powszechnie znany skrótowiec: pierwsze przywołanie w tekście	(International Labour Organization [ILO], 2020)	International Labour Organization (ILO, 2020)
kolejne przywołanie	(ILO, 2020)	ILO (2020)
Pełna nazwa	(Stanford University, 1995)	Stanford University (1995)
Typ publikacji		
Publikacja bez ustalonego autorstwa	(<i>Skrócony tytuł ...</i> , 2015)	<i>Pełny tytuł</i> (2015)
Publikacja bez roku wydania	(Iksiński, b.r.)	Iksiński (b.r.)
Akt prawny	(Pełny tytuł)	Pełny tytuł
Strona internetowa / Zbiór danych: znana data publikacji	(Iksiński, 2020) / (Nazwa instytucji, 2020)	Iksiński (2020) / Nazwa instytucji (2020)
nieznana data publikacji	(Iksiński, b.r.) / (Nazwa instytucji, b.r.)	Iksiński (b.r.) / Nazwa instytucji (b.r.)
Rodzaj przywołania		
Przywoływanie kilku prac (porządek prac – chronologiczny, porządek autorów – alfabetyczny)	(Iksiński, 1997, 1999, 2004a, 2004b; Nowak, 2002)	Iksiński (1997, 1999, 2004a, 2004b) i Nowak (2002)
Przywoływanie publikacji za innym autorem (uwaga: w bibliografii należy wymienić tylko pracę czytaną)	(Nowakowski, 1990, za: Zienniecka, 2007)	Nowakowski (1990, za: Zienniecka, 2007)

Źródło: opracowanie na podstawie: American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th edition). <https://doi.org/10.1037/0000165-000>.

4.4. Przykłady opisu bibliograficznego

Bibliografia powinna być zamieszczona na końcu opracowania. Prace należy uszeregować alfabetycznie według nazwiska pierwszego autora. W przypadku dwóch lub więcej prac tego samego autora / tych samych autorów trzeba je uporządkować chronologicznie według roku publikacji. Jeśli kilka prac tego samego autora / tych samych autorów zostało opublikowanych w tym samym roku, należy podać je w kolejności alfabetycznej według tytułu i odpowiednio oznaczyć literami a, b, c itd.

Typ publikacji	Przykład opisu bibliograficznego
Artykuł w czasopiśmie	
W wersji drukowanej	Nazwisko, X. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik (zeszyt)</i> , strona początku–strona końca.
Dostępny w internecie, z DOI	Nazwisko, X., Nazwisko 2, Y. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik(zeszyt)</i> , strona początku–strona końca. https://doi.org/xxx .
Dostępny w internecie, bez DOI	Nazwisko, X., Nazwisko 2, Y., Nazwisko 3, Z. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik(zeszyt)</i> , strona początku–strona końca. https://xxx .
Maszynopis	
Niepublikowany / przygotowywany przez autora / zgłoszony do publikacji, ale jeszcze niezaakceptowany	Nazwisko, X. (rok). <i>Tytuł [maszynopis niepublikowany / w przygotowaniu / zgłoszony do publikacji]</i> .
Zaakceptowany do publikacji	Nazwisko, X. (w druku). Tytuł artykułu. <i>Tytuł czasopisma</i> .
Opublikowany nieformalnie (np. na stronie internetowej autora)	Nazwisko, X., Nazwisko 2, Y. (rok). <i>Tytuł artykułu</i> . https://xxx .
Opublikowany w trybie online first (przed włączeniem do zeszytu)	Nazwisko, X. (rok). Tytuł artykułu. <i>Tytuł czasopisma</i> . Online first. https://xxx .
Książka	
W wersji drukowanej	Nazwisko, X. (rok). <i>Tytuł książki</i> . Wydawnictwo.
Dostępna w internecie, z DOI	Nazwisko, X. (rok). <i>Tytuł książki</i> . Wydawnictwo. https://doi.org/xxx .
Dostępna w internecie, bez DOI	Nazwisko, X. (rok). <i>Tytuł książki</i> . Wydawnictwo. https://xxx .
W przekładzie	Nazwisko, X. (rok). <i>Tytuł książki</i> (tłum. Y. Nazwisko). Wydawnictwo.
Wydanie wielotomowe: tom zatytułowany	Nazwisko, X. (rok). <i>Tytuł książki: nr tomu. Tytuł tomu</i> . Wydawnictwo.
tom niezatytułowany	Nazwisko, X. (rok). <i>Tytuł książki (nr tomu)</i> . Wydawnictwo.
Kolejne wydanie	Nazwisko, X. (rok). <i>Tytuł książki (nr wydania)</i> . Wydawnictwo.
Pod redakcją: w języku polskim	Nazwisko, X. (red.). (rok). <i>Tytuł książki</i> . Wydawnictwo.
w języku angielskim	Nazwisko, X. (Ed.). (rok). <i>Tytuł książki</i> . Wydawnictwo.
Rozdział w pracy zbiorowej	Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, Z. Nazwisko 2 (red.), <i>Tytuł książki (s. strona początku–strona końca)</i> . Wydawnictwo. https://doi.org/xxx lub https://xxx .
Inne prace	
Raport: autor indywidualny	Nazwisko, X. (rok). <i>Tytuł raportu</i> . Wydawnictwo.
autor instytucjonalny	Nazwa instytucji. (rok). <i>Tytuł raportu</i> . Wydawnictwo.
Working Papers	Nazwisko, X. (rok). <i>Tytuł pracy</i> (nazwa serii i numer). https://doi.org/xxx lub https://xxx .
Sesja konferencyjna / prezentacja / referat	Nazwisko, X. (rok, dzień i miesiąc). <i>Tytuł pracy</i> [typ wystąpienia, np. referat]. Nazwa konferencji, miejsce konferencji.
Rozprawa doktorska: nieopublikowana	Nazwisko, X. (rok). <i>Tytuł pracy</i> [nieopublikowana rozprawa doktorska]. Nazwa instytucji nadającej tytuł doktorski.
opublikowana	Nazwisko, X. (rok). <i>Tytuł pracy</i> [rozprawa doktorska, nazwa instytucji nadającej tytuł doktorski]. https://xxx .
Akt prawny	Pełny tytuł aktu prawnego wraz z datą publikacji w dzienniku urzędowym.

Typ publikacji	Przykład opisu bibliograficznego
Strona internetowa	
Znana data publikacji, zawartość strony się nie zmienia	Nazwisko, X. (rok, dzień i miesiąc). <i>Tytuł</i> . https://xxx .
Nieznana data publikacji, zawartość strony się zmienia	Nazwa instytucji. (b.r.). <i>Tytuł</i> . Pobrane dzień, miesiąc i rok pobrania z https://xxx .
Zbiór danych	
Surowe dane nieopublikowane	Nazwisko, X. (rok wydania pracy, w której dane są wykorzystywane) [opis danych, np. surowe dane nieopublikowane dotyczące...]. Źródło danych (np. nazwa uniwersytetu).
Dane opublikowane: znana data publikacji, zawartość zbioru się nie zmienia	Nazwisko, X. (rok). <i>Nazwa zbioru danych</i> [zbiór danych]. Wydawca. https://xxx .
nieznana data publikacji, zawartość zbioru się zmienia	Nazwa instytucji. (b.r.). <i>Nazwa zbioru danych</i> [zbiór danych]. Wydawca. Pobrane dzień, miesiąc i rok pobrania z https://xxx .

Źródło: opracowanie na podstawie: American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th edition). <https://doi.org/10.1037/0000165-000>.

Praca przygotowana w sposób niezgodny z powyższymi wskazówkami będzie odesłana do autora z prośbą o dostosowanie formy artykułu do wymogów redakcyjnych.

DZIAŁY „WS” – TEMATYKA ARTYKUŁÓW WS SECTIONS – TOPICS OF THE ARTICLES

Pełny opis zakresu tematycznego działów: ws.stat.gov.pl/AimScope

Description of the topics covered in each section: ws.stat.gov.pl/AimScope

Studia metodologiczne / Methodological studies

- Oryginalne teoretyczne rozwiązania metodologiczne ze wskazaniem ich praktycznej użyteczności
- Prace przeglądowe i porównawcze oraz dotyczące etyki w statystyce, które wnoszą pionierski wkład poznawczy do obecnego stanu wiedzy

Statystyka w praktyce / Statistics in practice

- Nowatorskie zastosowania narzędzi i modeli statystycznych oraz analiza i ocena statystyczna zjawisk społeczno-ekonomicznych i innych, prowadzona w szczególności na danych z zasobów statystyki publicznej
- Wykorzystanie narzędzi informatycznych do uzyskiwania i przetwarzania informacji statystycznych, naliczania danych wynikowych, ich prezentacji i rozpowszechniania
- Projektowanie badań statystycznych, uzyskiwanie, integracja i przetwarzanie danych oraz generowanie wynikowych informacji statystycznych i kontrola ich ujawniania

Studia interdyscyplinarne. Wyzwania badawcze / Interdisciplinary studies. Research challenges

- Wyzwania badawcze wynikające z rosnących potrzeb użytkowników danych statystycznych i wymagające zaangażowania znacznych środków oraz rozwiązań z różnych dziedzin nauki i techniki
- Wykorzystanie technologii informacyjnych i komunikacyjnych, innowacyjność, przetwarzanie i analiza zagadnień związanych z data science i big data
- Wyniki badań prowadzonych przez przedstawicieli dyscyplin innych niż statystyka z wykorzystaniem metod statystycznych

Spisy powszechne – problemy i wyzwania / Issues and challenges in census taking

- Propozycje rozwiązań – zarówno organizacyjnych, jak i metodologicznych – możliwych do zastosowania w spisach oraz rezultaty analiz danych spisowych
- Praktyczne aspekty związane z gromadzeniem i udostępnianiem danych ze spisów, w tym dotyczące obciążenia odpowiedzi i ochrony tajemnicy statystycznej

Edukacja statystyczna / Statistical education

- Metody i efekty nauczania statystyki oraz popularyzacja myślenia statystycznego i rzetelnego posługiwania się informacjami statystycznymi
- Problemy związane z kształceniem w zakresie umiejętności stosowania statystyki na wszystkich poziomach edukacji, a także dotyczące wykorzystywania nowoczesnych koncepcji i metod dydaktycznych oraz pomocy naukowych w nauczaniu statystyki

Z dziejów statystyki / From the history of statistics

- Historia prowadzenia obserwacji statystycznych oraz rozwoju ich metodologii i narzędzi
- Życie i osiągnięcia zawodowe wybitnych statystyków, jak również działalność najważniejszych instytucji i organizacji statystycznych w Polsce i za granicą

In memoriam

- Nekrologi i artykuły wspomnieniowe

Informacje. Recenzje. Dyskusje / Discussions. Reviews. Information

- Teksty nierecenzowane i niemające charakteru artykułów naukowych: sprawozdania z konferencji naukowych i innych wydarzeń dotyczących statystyki, recenzje książek, omówienia nowości wydawniczych GUS, polemiki i dyskusje