



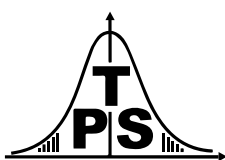
# III Kongres Statystyki Polskiej

Księga abstraktów

## The 3rd Congress of Polish Statistics

Abstracts

26-28 kwietnia 2022 r., Kraków





# **III Kongres Statystyki Polskiej**

Księga abstraktów

## **The 3rd Congress of Polish Statistics**

Abstracts

26-28 kwietnia 2022 r., Kraków



**Opracowanie graficzne i projekt okładki**

Graphic layout and cover design

Marek Wilczyński

**Druk i oprawa**

Printing and binding

Zakład Wydawnictw Statystycznych

Statistical Publishing Establishment

Publikacja dostępna na stronie: [stat.gov.pl](http://stat.gov.pl)

Publication available at: [stat.gov.pl](http://stat.gov.pl)

Warszawa 2022

© Copyright by Główny Urząd Statystyczny

ISBN 978-83-66466-91-3 (wersja elektroniczna / online version)

ISBN 978-83-66466-90-6 (wersja drukowana / printed version)

Szanowni Państwo,

III Kongres Statystyki Polskiej, organizowany w dniach 26–28 kwietnia 2022 r. w Krakowie przez Główny Urząd Statystyczny i Polskie Towarzystwo Statystyczne, to wyjątkowe spotkanie przedstawicieli statystyki publicznej i środowiska naukowego, zarówno z Polski, jak i z zagranicy. Biorą w nim udział pracownicy urzędów statystycznych, naukowcy, inni interesariusze statystyki publicznej oraz reprezentanci instytucji publicznych. Celem Kongresu jest przede wszystkim wzmocnienie wielopłaszczyznowej współpracy instytucji statystycznych i ośrodków naukowych w sprawach dotyczących rozwoju profesjonalnej i innowacyjnej statystyki publicznej oraz stałego unowocześniania metodologii badań statystycznych.

Program Kongresu obejmuje następujące obszary tematyczne: statystykę matematyczną, społeczną, gospodarczą i regionalną, statystykę ludności, statystykę polską i międzynarodową, a także metodę reprezentacyjną i big data oraz analizę i klasyfikację danych. Uczestnicy omówią wyzwania związane m.in. z gromadzeniem, przetwarzaniem i udostępnianiem danych w kontekście aktualnych uwarunkowań zewnętrznych, cyfryzacji i globalizacji.

Wydarzenie będzie upamiętniać 110. rocznicę powołania Polskiego Towarzystwa Statystycznego. Efektem działalności tej organizacji, zrzeszającej ludzi zainteresowanych teorią i praktyką badań statystycznych, jest wiele przedsięwzięć ważnych zarówno dla rozwoju statystyki publicznej, jak i statystyki jako dziedziny nauki.

Jesteśmy przekonani, że Kongres stanowi doskonałą okazję do twórczej wymiany poglądów, wiedzy i doświadczeń. Życzymy Państwu ciekawych dyskusji, inspirujących do dalszych badań oraz podejmowania inicjatyw na rzecz jak najlepszej jakości i dostępności danych statystycznych i ich efektywnego wykorzystywania.

Prezes  
Głównego Urzędu Statystycznego

dr Dominik Rozkrut

Prezes  
Polskiego Towarzystwa Statystycznego

prof. dr hab. Waldemar Tarczyński

Kraków, kwiecień 2022 r.

Dear Readers,

The 3rd Congress of Polish Statistics, organised on 26–28 April 2022 in Kraków by Statistics Poland and the Polish Statistical Association, is a unique meeting of the representatives of official statistics and the scientific community from Poland and abroad. It brings together employees of statistical offices, scientists, other stakeholders of official statistics, and representatives of public institutions. The primary goal of the Congress is to strengthen the multifaceted collaboration of statistical institutions and scientific centres in matters relating to the development of professional and innovative official statistics and in the continuous modernisation of the methodology of statistical research.

The programme of the Congress covers the following thematic areas: mathematical, social, economic and regional statistics, population statistics, Polish and international statistics, the representation method, big data, and data analysis and classification. The participants will discuss the challenges related to collecting, processing and sharing data, especially in the light of the current external circumstances, digitalisation and globalisation.

The event will also celebrate the 110th anniversary of the establishment of the Polish Statistical Society. The activity of this organisation, which brings together people interested in the theory and practice of statistical research, has resulted in many undertakings significant both to the development of official statistics and statistics as a field of science.

We are convinced that the Congress will provide an excellent opportunity for a creative exchange of opinions, knowledge and experience. We hope it will produce interesting and stimulating discussions that will inspire further research into the subject and spark new initiatives aimed at achieving the highest possible quality and the widest possible availability of statistical data, as well as enabling their effective use.

President  
Statistics Poland

Chairman  
Polish Statistical Association

Dominik Rozkrut, PhD

Waldemar Tarczyński, PhD, DSc, ProfTit

Kraków, April 2022 r.

# SPIS TREŚCI

## CONTENTS

01	STATYSTYKA MATEMATYCZNA .....	6
	MATHEMATICAL STATISTICS	
02	METODA REPREZENTACYJNA I BIG DATA .....	31
	REPRESENTATIVE METHOD AND BIG DATA	
03	STATYSTYKA LUDNOŚCI .....	48
	POPULATION STATISTICS	
04	STATYSTYKA SPOŁECZNA .....	63
	SOCIAL STATISTICS	
05	STATYSTYKA GOSPODARCZA .....	99
	ECONOMIC STATISTICS	
06	STATYSTYKA REGIONALNA .....	122
	REGIONAL STATISTICS	
07	ANALIZA I KLASYFIKACJA DANYCH .....	160
	DATA ANALYSIS AND CLASSIFICATION	
08	STATYSTYKA POLSKA I STATYSTYKA MIĘDZYNARODOWA .....	184
	POLISH AND INTERNATIONAL STATISTICS	

**Aleksandra Baszczyńska**

Uniwersytet Łódzki

University of Lodz

## **SZEROKOŚĆ PODPRZEDZIAŁÓW HISTOGRAMU W WYZNACZENIU WARTOŚCI PARAMETRU WYGŁADZANIA ESTYMACJI JĄDROWEJ FUNKCJI GĘSTOŚCI – PRZYPADEK DWUWYMIAROWY**

Celem pracy jest prezentacja oraz analiza własności metody wyboru parametru wygładzania w estymacji jądrowej funkcji gęstości wykorzystującej szerokość podprzedziałów w histogramie. Rozważania dotyczą zarówno przypadku jednowymiarowego, jak i dwuwymiarowego. Oceny własności zaproponowanej metody dokonano za pomocą analiz wyników badań symulacyjnych. Wykazują one, że metoda oparta na szerokości podprzedziałów dwuwymiarowego histogramu może być zastosowana w wyborze wartości parametru wygładzania estymacji jądrowej funkcji gęstości. Nowa metoda charakteryzuje się dobrymi własnościami i prostotą zastosowania.

Słowa kluczowe: estymacja jądrowa funkcji gęstości, parametr wygładzania, histogram

## **WIDTH OF THE HISTOGRAM SUB-INTERVALS IN DETERMINING THE VALUE OF THE SMOOTHING PARAMETER OF THE NUCLEAR ESTIMATION OF THE DENSITY FUNCTION – TWO-DIMENSIONAL CASE**

The aim of the work is to present and analyze the properties of the smoothing parameter selection method in the kernel estimation of the density function using the width of the sub-intervals in the histogram. The considerations apply to both the one-dimensional case and the two-dimensional case. The properties of the proposed method are assessed by analyzing the results of simulation methods. They show that the method based on the width of the sub-intervals of the two-dimensional histogram can be used in selecting the value of the smoothing parameter for the kernel estimation of the density function. The new method is characterized by both good properties and ease of use.

Keywords: kernel density of density function, smoothing parameter, histogram

**Małgorzata Bogdan**

Uniwersytet Wrocławski

University of Wrocław

## **ANALIZA DANYCH ZA POMOCĄ ESTYMATORA SLOPE**

Metoda SLOPE służy do redukcji wymiaru w dużych bazach danych. Zaprezentujemy nowe wyniki teoretyczne dotyczące własności tej metody i jej zastosowanie do analizy danych finansowych.

Słowa kluczowe: wybór modelu, regularyzacja, optymalizacja portfela

## **DATA ANALYSIS WITH THE SORTED L-ONE ESTIMATOR**

Sorted L-One Estimator is a method for dimensionality reduction in large data bases. We present new theoretical results on the properties of this method and its application in finance.

Keywords: model selection, regularization, portfolio optimization



**Bronisław Ceranka**

Uniwersytet Przyrodniczy w Poznaniu

Poznań University of Life Sciences

**Małgorzata Graczyk**

Uniwersytet Przyrodniczy w Poznaniu

Poznań University of Life Sciences

**D-OPTYMALNE SPRĘŻYNOWE UKŁADY WAGOWE**

Celem pracy jest podanie warunków określających istnienie D- optymalnych sprężynowych układów wagowych w wybranych klasach. Na podstawie własności macierzy wyznaczono górne oszacowanie wyznacznika macierzy informacji dla układu i uwzględniając to ograniczenie, wyprowadzono warunki określające istnienie układów optymalnych. Wynikiem analizy jest podanie warunków wyznaczających układy optymalne oraz metod ich konstrukcji. Nie w każdej klasie sprężynowych układów wagowych istnieje układ D- optymalny. Podane metody konstrukcji rozszerzają dotychczas znane klasy układów, w których istnieje ten najlepszy układ.

Słowa kluczowe: sprężynowe układy wagowe, układ zrównoważony o blokach niekompletnych

**D-OPTIMAL SPRING BALANCE WEIGHING DESIGNS**

The purpose of the paper is presenting the conditions determining the existence of the D-optimal spring balance weighing designs in selected classes. Based on the matrices properties, the upper bound of the determinant of the information matrix for the design was given. The conditions determining optimal design and construction methods were determined. In any class of the spring balance weighing designs it is not possible to determine optimal design. The presented method extends the class of optimal designs.

Keywords: spring balance weighing design, balanced incomplete block design

**Mariusz Czekala**

Wyższa Szkoła Bankowa we Wrocławiu

WSB University in Wrocław

**Agnieszka Matylda Schlichtinger**

Wyższa Szkoła Bankowa we Wrocławiu

WSB University in Wrocław

**ZASTOSOWANIE TESTU INWERSJI DO ANALIZY SKUTECZNOŚCI TERAPII MEDYCZNYCH**

Celem pracy jest przedstawienie nieparametrycznego testu inwersji wraz z ważną dziedziną zastosowań medycznych. Test ten, dzięki wprowadzeniu nowego parametru – prawdopodobieństwa inwersji – umożliwia testowanie hipotez prostych dotyczących korelacji z zadanymi wartościami różnymi od zera. W pracy przedstawiono związek proponowanego rozwiązania z testem korelacji rang Kendalla. Przedstawiono również dowód twierdzenia o wzajemnej odpowiedności prawdopodobieństwa inwersji z wartością statystyki tau Kendalla. Zaproponowano test statystyczny, dla którego istnieje możliwość precyzyjnej oceny mocy. Proponowane rozwiązanie nie wymaga dodatkowych założeń dotyczących liczebności próby czy też postaci rozkładu cechy. W części dotyczącej zastosowań przedstawiono możliwości wykorzystania proponowanego testu do oceny jakości terapii. Taka możliwość występuje wtedy, gdy oceniana cecha charakteryzująca jednostkę chorobową jest przedstawiona na skali porządkowej. Porównanie dwóch terapii polega wówczas na wykazaniu istotnej zmiany wielkości mierzącej korelację. W pracy podano przykłady zastosowań proponowanej metody w analizie terapii raka oraz ciśnienia tętniczego. Opisano również inne szerokie możliwości zastosowań testu.

Słowa kluczowe: test statystyczny, Kendall, inwersje, ranga, terapia medyczna

**APPLICATION OF THE INVERSION TEST TO ANALYZE THE EFFECTIVENESS OF MEDICAL THERAPIES**

The aim of the study is to present a non-parametric inversions test along with an important field of medical applications. Thanks to the introduction of a new parameter – the probability of inversion, this test makes it possible to test simple hypotheses regarding correlation with given values different

from zero. The paper presents the relations of the proposed solution with the Kendall rank correlation test. Proof of the theorem of mutual correspondence of the probability of inversion with the value of Kendall's tau statistics is also presented. A statistical test was proposed for which it is possible to accurately assess the power. The proposed solution does not require additional assumptions regarding the sample size or the form of the feature distribution. The application section presents the possibilities of using the proposed test to assess the quality of therapy. Such a possibility occurs when the assessed feature characterizing the disease entity is presented on an ordinal scale. The comparison of the two therapies then consists in demonstrating a significant change in the multiplicity measuring the correlation. The paper gives examples of applications of the proposed method. Other wide range of applications of the test are also described.

Keywords: statistical test, Kendall, inversions, rank, medical therapy

**Anna Dudek**

Akademia Górniczo-Hutnicza w Krakowie  
AGH University of Science and Technology

**Łukasz Lenart**

Uniwersytet Ekonomiczny w Krakowie  
Cracow University of Economics

**ESTYMACJA GĘSTOŚCI SPEKTRALNEJ DLA NIESTACJONARNYCH SZEREGÓW CZASOWYCH O NIEZEROWEJ FUNKCJI WARTOŚCI OCZEKIWANEJ**

Rozważamy problem estymacji gęstości spektralnej w przypadku niestacjonarnym. Klasyczne podejście do estymacji gęstości spektralnej wymaga usunięcia funkcji średniej z danych. Jednakże w przypadku niestacjonarnym może to być bardzo trudne. Dlatego istotne jest zaproponowanie nowej metody, która będzie zgodna i nie będzie zakładała zerowości funkcji wartości oczekiwanej. Nasza metoda estymacji gęstości spektralnej opiera się na metodzie subsamplingu, która pozwala aproksymować asymptotyczną wariancję pewnej przeskalowanej statystyki będącej przedmiotem zainteresowania. Ogólny wynik zgodności dotyczący estymatora subsamplingowego asymptotycznej wariancji, który przedstawiamy, nie wymaga żadnej szczególnej struktury danych (takiej jak np. okresowość lub prawie okresowość). Przedstawiamy nowe podejście do nieparametrycznej estymacji gęstości spektralnej oparte na technice subsamplingu, które następnie stosujemy do ważnej klasy niestacjonarnych szeregów czasowych. Są to szeregi prawie okresowo skorelowane. W przeciwieństwie do istniejących metod nasza technika nie wymaga usunięcia funkcji wartości oczekiwanej z danych. W przeprowadzonych badaniach symulacyjnych porównujemy nasz estymator funkcji gęstości spektralnej z estymatorem klasycznym. Dodatkowo proponujemy zmodyfikowany estymator, który pozwala zredukować efekt wycieku spektralnego. Pokazujemy również zastosowanie naszych wyników do danych rzeczywistych dotyczących ekonomii. Nasze wyniki pozwalają sformułować nowe problemy otwarte. Ponieważ klasyczne podejście do estymacji spektralnej wyższego rzędu wymaga również usunięcia funkcji wartości oczekiwanej z danych, uważamy, że nasze pomysły mogą być zastosowane w tym problemie.

Słowa kluczowe: subsampling, prawie okresowość, estymacja macierzy kowariancji, mieszanie

## SPECTRAL DENSITY ESTIMATION FOR NONSTATIONARY DATA WITH NONZERO MEAN FUNCTION

We focus on the spectral density estimation problem in a nonstationary setting. The classic approach for spectral density estimation requires removing the mean function from the data. However, this may be very challenging in the nonstationary case. Thus, it is crucial to propose a new method that provides consistent estimates and does not assume the nullity of the mean function. Our idea for spectral densities estimation is based on the subsampling method, which allows approximation of the asymptotic variance of the rescaled statistics of interest. The general consistency result concerning subsampling estimator of the asymptotic variance that we present does not require any particular data structure (such as periodicity or almost periodicity). We introduce a new approach for nonparametric spectral density estimation based on the subsampling technique, which we apply to the important class of nonstationary time series. These are almost periodically correlated sequences. In contrary to existing methods, our technique does not require demeaning of the data. On the simulated data examples we compare our estimator of spectral density function with the classical one. Additionally, we propose a modified estimator, which allows the reduction of the leakage effect. We apply our results to real data economic applications. Our results provide new open problems concerning subsampling estimation for the higher-order moments. Since the classic approach for the higher-order spectral estimation also requires removing the mean function from the data, we believe that our ideas can be adapted in higher-order spectral estimation problem under the non-zero mean assumption.

Keywords: subsampling, almost periodically correlated, covariance matrix estimation, mixing

**Przemysław Grzegorzewski**

Politechnika Warszawska

Warsaw University of Technology

## TESTY PERMUTACYJNE W ANALIZIE NIEPRECYZYJNYCH DANYCH

Załóżmy, że wynikiem eksperymentu jest próbka losowa złożona z nieprecyzyjnych danych, które daje się zadowalająco opisać za pomocą liczb rozmytych. Do analizy statystycznej takich danych potrzebny jest model matematyczny, który pozwoli uchwycić dwa aspekty niepewności występujące w tego typu próbkę, tj. losowość związaną z mechanizmem generowania danych oraz rozmycie odpowiadające brakowi precyzji. Aby poradzić sobie z tym problemem, Puri i Ralescu (1986) wprowadzili pojęcie rozmytej zmiennej losowej. Jednakże samo to pojęcie nie wystarcza do bezpośredniego uogólnienia klasycznych metod wnioskowania statystycznego na dane rozmyte. Próbuąc zastosować znane metody statystyczne do danych rozmytych, napotykamy pewne zasadnicze przeszkody związane z arytmetyką liczb rozmytych, z brakiem odpowiednich modeli rozkładów prawdopodobieństwa dla rozmytych zmiennych losowych czy też brakiem twierdzeń granicznych, które dałoby się zastosować wprost we wnioskowaniu statystycznym z danymi rozmytymi. Aby przezwyciężyć te przeszkody, stosuje się zwykle odpowiednio dobrane metryki w przestrzeni liczb rozmytych oraz metody bootstrap dla ogólnych elementów losowych.

W naszej prezentacji rozważymy zagadnienie testowania zgodności dla dwóch lub więcej próbek rozmytych. Zaproponujemy nowe testy statystyczne do weryfikacji hipotez o jednakowym rozkładzie, których konstrukcja łączy rozmaite techniki z metodologią testów permutacyjnych. Warto podkreślić, że zaproponowane procedury testowe są w pełni nieparametryczne, co ma ogromne znaczenie we wnioskowaniu statystycznym na podstawie danych rozmytych. Prezentację nowych testów uzupełnimy omówieniem ich podstawowych właściwości. Nową metodologię zilustrujemy również studium przypadku związanym z pewnym zagadnieniem statystycznej kontroli jakości.

Słowa kluczowe: liczby rozmyte, nieprecyzyjne dane, rozmyte zmienne losowe, testy permutacyjne, testy zgodności

## PERMUTATION TEST IN IMPRECISE DATA ANALYSIS

When the output of an experiment consists of a random sample of imprecise data that could be satisfactorily described by fuzzy numbers, we need a model which allows the grasping of both aspects

of uncertainty that appear in such data, i.e. randomness, associated with data generation mechanism and fuzziness, connected with data nature, i.e. their imprecision. To cope with this problem, Puri and Ralescu (1986) introduced the notion of a fuzzy random variable. However, this concept alone is not sufficient to generalize classic methods of statistical inference directly to fuzzy environment. Indeed, in analyzing fuzzy data from the statistical perspective we immediately come upon some key obstacles, like the nonlinearity associated with the fuzzy number arithmetic, the lack of suitable probability distribution models or no limit theorems for random mechanisms producing fuzzy data which could be directly applied in statistical inference. To overcome these drawbacks one usually applies an appropriate metric between fuzzy data and a bootstrapped central limit theorem for general space-valued random elements. In our presentation we consider the goodness-of-fit testing problem for two or more fuzzy samples. We suggest its solution based on another methodology and propose some new test constructions combining various techniques with permutation tests. All suggested testing procedures are completely distribution-free which is of extreme importance in statistical reasoning with fuzzy data. Besides deriving the aforementioned tests, we also compare their basic properties and discuss a case study related to quality assessment illustrating the proposed methodology.

Keywords: fuzzy numbers, fuzzy random variables, goodness-of-fit tests, imprecise data, permutation tests

**Mirosław Krzyśko**

Uniwersytet im. Adama Mickiewicza w Poznaniu

Adam Mickiewicz University Poznań

**Tomasz Górecki**

Uniwersytet im. Adama Mickiewicza w Poznaniu

Adam Mickiewicz University Poznań

**WIELOKROTNA ANALIZA KORELACJI KANONICZNYCH DLA WIELOWYMIAROWYCH DANYCH POWTARZANYCH POMIARÓW ORAZ WIELOWYMIAROWYCH DANYCH FUNKCJONALNYCH**

Klasyfikacja analiza korelacji kanonicznych Hotellinga (1936) została uogólniona na dwa sposoby. Po pierwsze wektory losowe zostały zastąpione macierzami losowymi zawierającymi obserwacje p zmiennych w T momentach czasowych (w tym przypadku mówimy o wielowymiarowych danych powtarzanych pomiarów, znanych także jako dane podwójnie wielowymiarowe) i badane były zależności między skończoną liczbą macierzy, a nie tylko dwiema. W tym przypadku skonstruowane zostały wielokrotne jądrowe zmienne kanoniczne. Po drugie wektory losowe zostały zastąpione wielozmiennymi procesami losowymi. W tym przypadku jednostki eksperymentalne charakteryzowane są za pomocą funkcji ciągłych, całkowalnych z kwadratem na pewnym przedziale czasowym (elementów przestrzeni Hilberta). Inaczej mówiąc, jednostki eksperymentalne charakteryzowane są za pomocą wielowymiarowych danych funkcjonalnych. Tak jak poprzednio, badane były zależności pomiędzy skończoną liczbą wielozmiennych procesów losowych, a nie tylko dwóch. W tym przypadku zostały skonstruowane wielokrotne funkcjonalne zmienne kanoniczne. Dla danych rzeczywistych, dotyczących plonów z hektara trzech różnych grup roślin uprawnych, lepszymi okazały się wielokrotne funkcjonalne zmienne kanoniczne.

Słowa kluczowe: wielokrotne jądrowe zmienne kanoniczne, wielokrotne funkcjonalne zmienne kanoniczne, wielowymiarowe dane powtarzanych pomiarów, wielowymiarowe dane funkcjonalne



## MULTIPLE CANONICAL CORRELATION ANALYSIS FOR MULTIVARIATE REPEATED MEASURES DATA AND MULTIVARIATE FUNCTIONAL DATA

Hotelling's (1936) classical canonical correlation analysis has been generalized in two ways. Firstly, random vectors were replaced by random matrices containing observations of  $p$  variables in  $T$  time moments on the same experimental unit (in this case we are talking about multivariate repeated measures data, also known as doubly multivariate data) and the relationships between a finite number of random matrices, not just two, were considered. In this case, multiple kernel canonical variables were constructed. Secondly, we replace random vectors with multidimensional random processes. In this case, the experimental units are characterized by continuous, square-integrable functions over a given interval (elements of Hilbert space). Otherwise, the experimental units are characterized by multivariate functional data. Moreover, as before, we consider the relationships between a finite number of multidimensional random processes, and not just two. In this case, multiple functional canonical variables were constructed. In the case of actual data on yields per hectare for three different groups of crops, multiple functional canonical variables turned out to be better.

Keywords: multiple kernel canonical variables, multiple functional canonical variables, multivariate repeated measures data, multivariate functional data

**Mirosław Krzyśko**

Akademia Kaliska im. Prezydenta Stanisława Wojciechowskiego

Calisia University

**Waldemar Wołyński**

Uniwersytet im. Adama Mickiewicza w Poznaniu

Adam Mickiewicz University Poznań

**ANALIZA SKŁADOWYCH GŁÓWNYCH DLA DANYCH CZASOWO-PRZESTRZENNYCH**

Jedną z głównych metod stosowanych w statystyce wielowymiarowej jest analiza składowych głównych (PCA). Ideą tej analizy jest redukcja wymiaru zbioru danych będących wynikiem obserwacji dużej liczby skorelowanych zmiennych przy zachowaniu jak największej występującej w nich zmienności. Osiąga się to poprzez przekształcenie zbioru zmiennych pierwotnych w nowy zbiór zmiennych, składowych głównych, które są nieskorelowane i uporządkowane tak, że kilka pierwszych z nich zachowuje większość zmienności obecnej we wszystkich zmiennych pierwotnych. Głównym celem referatu jest podanie konstrukcji oraz przetestowanie nowej metodologii właściwego analizowania zależności czasoprzestrzennych z wykorzystaniem techniki funkcjonalnych składowych głównych. W naszym podejściu maksymalizowany jest iloczyn wariancji funkcjonalnej składowej głównej oraz indeksu I Morana lub C Geary' ego. Użyteczność przedstawionego podejścia do analizy danych czasowo-przestrzennych zilustrowana zostanie rzeczywistym przykładem dotyczącym rozwoju polskich województw w latach 2002–2018 opisanym 12 cechami społeczno-ekonomicznymi.

Słowa kluczowe: składowe główne, dane czasowo-przestrzenne, dane funkcjonalne, indeks Morana, indeks Geary'ego

**SPATIO-TEMPORAL PRINCIPAL COMPONENT ANALYSIS**

One of the powerful methods in multivariate statistics is Principal Component Analysis (PCA). The idea of Principal Component Analysis is to reduce the dimensionality of a data set consisting of a large number of correlated variables, while retaining as much as possible of the variation present in the data. This is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in

all of the original variables. The present study aims to construct and test a new methodology for a proper mathematical statistical treatment of spatio-temporal dependencies among multivariate data in a PCA. In our approach, the product of the functional principal component variance and Moran's I index or the Geary's contiguity ratio C is optimized. The usefulness of spatio-temporal principal components is confirmed by the real example, which concerns the development of Polish regions in 2002-2018 characterized by 12 socioeconomic features.

Keywords: spatio-temporal principal components, spatio-temporal data, functional data, Moran's I index, Geary's contiguity ratio C

**Paweł Kurasiński**

Uniwersytet Przyrodniczy w Lublinie

University of Life Sciences in Lublin

**DOKŁADNE PRAWA WIELKICH LICZB I ICH ZASTOSOWANIE**

Celem prezentacji jest przedstawienie wyników dotyczących zbieżności z prawdopodobieństwem 1 ważonych sum zmiennych losowych do pewnej stałej różnej od 0. Twierdzenia dotyczące takiej zbieżności nazywamy dokładnymi prawami wielkich liczb. Pojawiają się one w szczególności w przypadku zmiennych losowych niezależnych, które mają jednakowe rozkłady z nieskończoną wartością oczekiwaną lub wartością oczekiwaną równą 0. W prezentacji będziemy rozważać sumy ilorazów niezależnych zmiennych losowych o tym samym rozkładzie ze zbieżnością prawie pewną. W dalszej części referatu skupimy się na badaniu zbieżności sum ważonych ilorazów statystyk porządkowych. W szczególności zaprezentowane będą wyniki dotyczące ilorazów najmniejszych statystyk porządkowych. Zostanie przedstawiony również wynik dotyczący maksimów ilorazów sąsiednich statystyk porządkowych.

Słowa kluczowe: sumy ważne, dokładne prawa wielkich liczb, ilorazy niezależnych zmiennych losowych, statystyki porządkowe

**EXACT STRONG LAWS OF LARGE NUMBERS AND THEIR APPLICATIONS**

The purpose of the presentation is to introduce the results concerning the convergence with probability of 1 of the weighted sums of random variables to a certain constant other than 0. We call theorems concerning such convergence the exact laws of large numbers. They appear, in particular, in the case of independent random variables that have equal distributions with an infinite expected value or an expected value equal to 0. In the presentation, we will consider the sums of ratios of independent random variables of the same distribution with almost certain convergence. In the further part of the presentation, we will focus on the study of the convergence of the sums of weighted ratios of order statistics. In particular, the results concerning the ratios of the smallest order statistics will be presented. The result of the ratios maxima of adjacent order statistics will also be presented.

Keywords: exact strong laws of large numbers, weighted sums of random variables, order statistics, ratios of independent random variables

**Marta Małecka**

Uniwersytet Łódzki

University of Lodz

**Radosław Pietrzyk**

Uniwersytet Ekonomiczny we Wrocławiu

Wroclaw University of Economics and Business

**ZASTOSOWANIE METODY BOOTSTRAPOWEJ DO ODPORNEJ ESTYMACJI PARAMETRÓW GARCH. PRZYKŁAD Z RYNKU KRYPTOWALUT**

Skomplikowane procesy finansowe obserwowane dziś na rynkach kryptowalut stawiają nowe wyzwania dotyczące rozwoju modeli szeregów czasowych. Jednym z głównych problemów utrudniających stosowanie klasycznych metod szeregów czasowych na rynku kryptowalut jest występowanie obserwacji nietypowych. Mają one m.in. wpływ na efektywność estymacji parametrów modeli GARCH, dlatego niezbędne jest stosowanie metod odpornych. Dotychczas w literaturze zaproponowano kilka metod odpornych na obserwacje nietypowe. Ich kluczowym narzędziem jest tzw. filtracja zmienności oparta na rozkładzie chi kwadrat. Przyjęcie parametrycznych założeń powoduje, że ograniczeniem tych metod jest możliwość stosowania jedynie do podstawowego modelu GARCH. Uniemożliwia to wykorzystanie m.in. dokładniejszych estymatorów zmienności, opartych na danych o wysokiej częstotliwości. Przez to metody te nie odpowiadają skomplikowanym procesom zachodzącym na rynkach finansowych, w szczególności na rynku kryptowalut.

Niniejsza praca rozszerza ten nurt badań w kierunku odpornej estymacji szerszej klasy modeli. Celem pracy jest opracowanie bootstrapowej metody filtracji zmienności pozbawionej parametrycznych założeń. Opracowanie tej metody umożliwia odporną estymację parametrów szerokiej klasy modeli GARCH. W szczególności pozwala to na wykorzystanie modeli z efektywnymi estymatorami zmienności. Praca przedstawia bootstrapowy algorytm filtracji zmienności oraz pokazuje jego porównanie z wcześniejszymi metodami odpornymi. Porównanie to przeprowadzone jest z wykorzystaniem danych z giełd kryptowalut. Badanie oparte jest na najbardziej płynnych instrumentach. Otrzymane wyniki potwierdzają skuteczność metody bootstrapowej jako narzędzia budowy metod estymacji odpornej na obserwacje nietypowe.

Słowa kluczowe: kryptowaluty, metoda bootstrapowa, odporna estymacja GARCH, obserwacje nietypowe

## APPLICATION OF BOOTSTRAP METHOD FOR ROBUST ESTIMATION OF GARCH PARAMETERS. EVIDENCE FROM CRYPTOCURRENCY MARKET

The complicated financial processes observed in the cryptocurrency markets today pose new challenges in the development of time series models. One of the main problems hindering the use of classic time series methods in the cryptocurrency market is the occurrence of outliers. They, among others, have an impact on the efficiency of estimation of GARCH parameters, which makes it necessary to use robust methods. So far, several methods robust to outliers have been proposed in the literature. Their key tool is the so-called volatility filtration based on the chi-square distribution. The adoption of parametric assumptions causes these methods to be relevant only for the basic GARCH model. This makes it impossible to use better estimators of variance, such as the ones based on high-frequency data. As a result, these methods do not correspond to the complex processes observed in the financial markets, in particular in the cryptocurrency market.

Our work extends previous studies towards robust estimation of a wider class of time series models. The aim of the work is to develop a bootstrap method of volatility filtration, free of any parametric assumptions. The development of this method enables a robust estimation of parameters of a wide class of GARCH models. In particular, it allows the use of models with effective estimators of variance. The work presents a bootstrap algorithm of filtering volatility and shows its comparison with earlier robust methods. This comparison is made using data from cryptocurrency markets. The study is based on the most liquid instruments. The obtained results confirm the effectiveness of the bootstrap method as an effective tool for building estimation methods robust to outliers.

Keywords: cryptocurrency, bootstrap method, robust GARCH estimation, outliers

**Łukasz Smaga**

Uniwersytet im. Adama Mickiewicza w Poznaniu  
Adam Mickiewicz University Poznań

**Marc Ditzhaus**

Uniwersytet Techniczny w Dortmundzie  
Technical University of Dortmund

**TESTOWANIE RÓWNOŚCI WIELOWYMIAROWEGO WSPÓŁCZYNNIKA ZMIENNOŚCI**

W referacie przedstawione zostaną nowe metody wnioskowania statystycznego dla wielowymiarowego współczynnika zmienności. Mimo że istnieje szereg różnych procedur testowych tego parametru w przypadku jednowymiarowym, problem ten jest w małym stopniu rozwiązany w przypadku wielowymiarowym. Istnieje tylko kilka testów, a ponadto opierają się one na restrykcyjnych założeniach dotyczących rozkładu danych. W celu rezygnacji z tego ograniczenia stosujemy statystykę typu Walda w ogólnym modelu czynnikowym, który nie zakłada heteroskedastyczności. Wykażemy, że skonstruowany test statystyczny asymptotycznie kontroluje prawdopodobieństwo błędu pierwszego rodzaju oraz jest zgodny. W celu poprawy własności tego testu w przypadku skończonego próbkowym zastosujemy metodę permutacyjną i pokażemy, że test permutacyjny zachowuje dobre własności testu asymptotycznego. Nowe i istniejące testy zostaną porównane w badaniach symulacyjnych. Przedstawiony zostanie również przykład ilustracyjny bazujący na danych rzeczywistych.

Słowa kluczowe: model czynnikowy, testowanie hipotez statystycznych, wielowymiarowy współczynnik zmienności

**TESTING THE EQUALITY OF THE MULTIVARIATE COEFFICIENT OF VARIATION**

New inference methods for the multivariate coefficient of variation are presented. While there are various testing procedures for this parameter in the univariate case, it is less known how to do inference in the multivariate setting appropriately. There are some existing procedures but they rely on restrictive assumptions on the underlying distributions. We tackle this problem by applying Wald-type statistic in the context of general, potentially heteroscedastic factorial designs. The resulting procedure is shown to be asymptotically valid

under the null hypothesis and consistent under general alternatives. To improve the finite sample performance, we suggest a permutation version of the test and show that the test's asymptotic properties can be transferred to it. An exhaustive simulation study compares the new tests and existing methods. To further analyse the differences between the tests, we conduct the illustrative real data example.

Keywords: factorial designs, statistical hypothesis testing, multivariate coefficient of variation



**Piotr Sulewski**

Akademia Pomorska w Słupsku

Pomeranian University in Słupsk

## **NOWI CZŁONKOWIE RODZINY ROZKŁADÓW JOHNSONA: WŁAŚCIWOŚCI I ZASTOSOWANIE**

W niniejszym artykule przedstawiamy nowe niemieszane rozkłady dwumodalne należące do rodziny rozkładów Johnsona (JFD) o nazwach SC i SD, gdzie SC jest szczególnym przypadkiem SD. SD jest jedyną dystrybucją z rodziny JFD, który może być zarówno unimodalny, jak i dwumodalny. Właściwości rozkładu SD zostały przedstawione. Rozkłady SC i SD porównano z siedmioma (nie licząc rozkładu normalnego) członkami JFD pod kątem elastyczności i możliwości zastosowania. Aby przetestować elastyczność, zdefiniowano specjalną miarę nazwaną skewness-kurtosis-square. Najlepsze rozproszenie punktów o współrzędnych (skośność, kurtoza) występuje dla SD oraz dla dobrze znanych SU i SB. Dwa zbiory danych rzeczywistych wykorzystano do testowania zastosowań nowego rozkładu. Rozkład EN okazał się lepszy od swoich konkurentów ze względu na kryteria informacyjne i wyniki trzech testów zgodności. Z treści artykułu wynika, że EN (w tym SC) jako nowy członek rodziny rozkładów Johnsona i jednocześnie jako nowy rozkład z kategorii rozkładów dwumodalnych niemieszanych jest modelem konkurencyjnym, który zasługuje na dodanie do istniejących rozkładów w modelowaniu danych.

Słowa kluczowe: rozkład normalny, elastyczność rozkładu, odejście od normalności

## **NEW MEMBERS OF THE JOHNSON FAMILY OF PROBABILITY DISTRIBUTIONS: PROPERTIES AND APPLICATION**

In this paper we introduce new non-mixed bimodal distributions belonging to the Johnson family of distributions (JFD) named SC and SD, where SC is a special case of SD. The SD is the only distribution among the JFD that can be both a unimodal and bimodal distribution. Properties of the SD are methodically studied. The SC and SD are compared with the seven (except the normal distribution) members of the JFD for flexibility and applicability. In order to test for flexibility, a special measure called skewness-kurtosis-square is defined. The best dispersion of points with coordinates (skewness, kurtosis) occurs for the SD and for the well-known SU and SB. Two real datasets were used to test the applicability. EN turned out to be better than its competitors in terms of information criteria and results of three goodness-of-fit tests.

The content of the paper shows that the EN (including SC) as a new member of the Johnson family of distributions and simultaneously as a new distribution from the non-mixed bimodal distribution category, is a competitive model that deserves to be added to the existing distributions in modeling data.

Keywords: normal distribution, flexibility of distribution, departure from normality

**Magdalena Szymkowiak**

Politechnika Poznańska

Poznan University of Technology

## **ANALIZA CZASU ŻYCIA ZA POMOCĄ FUNKCJI INTENSYWNOŚCI STARZENIA**

Celem pracy jest analiza danych dotyczących czasu życia badanych obiektów. Analizę tę można przeprowadzić za pomocą zdefiniowanej w pracy rodziny funkcji intensywności starzenia. Funkcje te mogą służyć do scharakteryzowania rozkładu czasu życia, zatem umożliwiają, na podstawie analizy danych dotyczących czasu działania obiektów, rozpoznanie rozkładu długości życia tych obiektów. Co więcej, za pomocą wprowadzonych funkcji można zdefiniować porządek intensywności starzenia. Oznacza to, że funkcje intensywności starzenia pozwalają zmierzyć i porównać tendencje starzenia badanych obiektów.

Słowa kluczowe: intensywność uszkodzeń, intensywność starzenia, charakteryzacja, estymacja

## **LIFETIME ANALYSIS BY AGING INTENSITY FUNCTIONS**

The goal of the paper is to analyze lifetime data. This analysis can be performed by the defined in the paper family of aging intensity functions. These functions can be used to characterize the lifetime distribution. Therefore, their analysis through some lifetime data allows us to recognize the data distribution. Moreover, the mentioned functions permit to define aging intensity order. It means that proposed aging intensity functions allow us to measure and compare the aging tendencies of the observed units.

Keywords: failure rate, aging intensity, characterization, estimation

**Sławomir Torbus**

Uniwersytet Kazimierza Wielkiego w Bydgoszczy  
Kazimierz Wielki University

**Karolina Mroczyńska**

Uniwersytet Kazimierza Wielkiego w Bydgoszczy  
Kazimierz Wielki University

**Bożena Piątkowska**

Politechnika Warszawska  
Warsaw University of Technology

**Edyta Gadomska**

Politechnika Warszawska  
Warsaw University of Technology

**FOTOWOLTAIKA I JEJ ROZUMIENIE W UJĘCIU STATYSTYCZNYM**

W opracowaniu zaprezentowano wyniki badania ankietowego mającego na celu ocenę poziomu świadomości młodzieży z różnych typów szkół średnich w zakresie znajomości idei fotowoltaiki i możliwości jej stosowania jako rozwiązania w szeroko rozumianym obszarze odnawialnych źródeł energii. Edukacja ekologiczna upowszechniająca wiedzę z obszaru odnawialnych źródeł energii oraz akcje informacyjne dotyczące możliwości uzyskania dofinansowania na zakup i montaż urządzeń fotowoltaicznych to kierunki działań, na które jest obecnie społeczne zapotrzebowanie. Stało się to asumptem do przygotowania niniejszego opracowania, w którym wykorzystano narzędzia statystyki matematycznej. Do badania ankietowego zaproszono uczniów z dwóch płockich szkół średnich o profilu technicznym kształcących w zupełnie odmiennych zawodach. Na podstawie wyników przeprowadzonej analizy statystycznej zaobserwowano istotną różnicę w zakresie wiedzy uczniów na temat fotowoltaiki zależnie od typu szkoły. Sformułowano wnioski, które wskazują kierunek realizacji kształcenia, aby rozumienie fotowoltaiki było na wysokim poziomie, niezależnie od grupy społecznej, z którą mamy do czynienia.

Słowa kluczowe: statystyka matematyczna, narzędzia statystyki matematycznej, świadomość ekologiczna, źródła energii odnawialnej

## PHOTOVOLTAICS AND ITS UNDERSTANDING IN STATISTICAL TERMS

The study presents the results of a survey aimed at assessing the level of awareness of young people from various types of secondary schools in the field of knowledge of the idea of photovoltaics and the possibility of using it as a solution in the broadly understood field of renewable energy sources. Ecological education disseminating knowledge in the field of renewable energy sources and information campaigns on the possibility of obtaining funding for the purchase and installation of photovoltaic devices are the directions of activities for which there is currently social demand. This prompted the preparation of this study, which uses the tools of mathematical statistics. Students from two technical secondary schools in Płock, educated in completely different professions, were invited to the survey. Based on the results of the statistical analysis, a significant difference was observed in the knowledge of students about photovoltaics, depending on the type of school. Conclusions were formulated that indicate the direction of the implementation of education so that the understanding of photovoltaics is at a high level, regardless of the social group with which we are dealing

Keywords: mathematical statistic, tools of mathematical statistics, ecological awareness, renewable energy sources

**Grzegorz Wyłupek**

Uniwersytet Wrocławski

University of Wrocław

## NIEPARAMETRYCZNY TEST W PROBLEMIE DWÓCH PRÓB DLA DANYCH SPAROWANYCH

Rozważamy klasyczny i bardzo ważny w zastosowaniach problem testowania. Dokładniej, niech  $(X_1, Y_1), \dots, (X_n, Y_n)$  będą niezależnymi wektorami losowymi o tym samym rozkładzie z dystrybuantą  $H(x, y)$  i ciągłymi dystrybuantami brzegowymi  $F(x)$  i  $G(y)$ , odpowiednio. Zmienne losowe  $X_i$  i  $Y_i$ ,  $i = 1, \dots, n$  są często interpretowane jako odpowiedzi pacjentów przed podaniem leku i po podaniu leku, odpowiednio. Ponieważ dane są sparowane,  $X_1$  i  $Y_1$  są zależne, przy czym struktura zależności może być wyrażona przez jednoznacznie wyznaczoną funkcję łącznikową  $C$ , tzn. dwuwymiarową dystrybuantę z jednostajnymi rozkładami brzegowymi, taką że  $H(x, y) = C(F(x), G(y))$ . Rozważamy problem testowania  $H: F(x) = G(x)$ , dla wszystkich  $x \in \mathbb{R}$ ,  $A: F(x) \neq G(x)$ , dla pewnego  $x \in \mathbb{R}$ . Hipoteza zerowa orzeka, że podany lek nie ma wpływu na stan zdrowia pacjentów, podczas gdy alternatywa jest jej zaprzeczeniem. W referacie podamy nowe rozwiązanie problemu testowania  $(H, A)$ , przedstawimy uzyskane wyniki teoretyczne oraz zaprezentujemy reprezentatywne rezultaty badań symulacyjnych.

Słowa kluczowe: bootstrap, copula, dane sparowane, proces empiryczny, problem dwóch prób

## A NONPARAMETRIC TEST FOR PAIRED DATA

The paper proposes a weighted Kolmogorov-Smirnov type test for the two-sample problem when the data is paired. The asymptotic distribution of the test statistic under the null model is derived. The dependence of both finite sample and asymptotic distribution of the test statistic from the dependence structure of the data forces the use of the wild bootstrap technique for the inference. The related wild bootstrap test turns out to be a consistent asymptotically level test. An application of the finite sample correction allows the test to keep the level well. An extensive simulation study demonstrates good finite sample behaviour of the test in comparison to the existing procedures.

Keywords: bootstrap, copula, empirical process, matched pairs, two-sample problem

**Wojciech Zieliński**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Warsaw University of Life Sciences

## PRZEDZIAŁ UFNOŚCI DLA ODSETKA PYTAŃ DRAŻLIWYCH

Jednym z problemów w badaniach ekonomiczno-społecznych jest oszacowanie odsetka pozytywnych odpowiedzi na pytanie drażliwe. Pytanie drażliwe jest takim pytaniem, na które respondent może nie chcieć udzielić rzetelnej odpowiedzi. W pracy podano konstrukcję przedziału ufności dla tego odsetka.

Słowa kluczowe: pytania drażliwe, model nierandomizowanych odpowiedzi, przedział ufności

## CONFIDENCE INTERVAL FOR THE FRACTION OF SENSITIVE QUESTIONS

In socioeconomic research estimation of the fraction of positive answers to a sensitive question. The sensitive question is such that the respondent would not give an honest answer. In the paper we present the construction of a confidence interval for such a fraction.

Keywords: sensitive questions, model of nonrandomized responses, confidence interval

**Adam Juszcak**

Uniwersytet Łódzki

University of Lodz

## WYKORZYSTANIE DANYCH SCRAPOWANYCH DO WYZNACZENIA ZMIANY CEN DLA ODZIEŻY I OBUWIA

Web scraping to technika pozwalająca automatycznie pozyskiwać informacje zamieszczone na stronach internetowych. Wraz ze wzrostem popularności zakupów online można dzięki niej wydobyć informacje o cenach dóbr sprzedawanych przez detalistów. Wykorzystanie danych scrapowanych nie tylko pozwala znacząco obniżyć koszty badania cen, lecz także poprawia precyzję szacunków inflacji i umożliwia śledzenie jej w czasie rzeczywistym. Z tego względu web-scraping jest dziś popularnym obiektem badań prowadzonych zarówno w ośrodkach statystycznych (Eurostat, brytyjski Office of National Statistics, belgijski Statbel), jak i na uniwersytetach (m.in. Billion Prices Project prowadzony na MIT). Zastosowanie danych scrapowanych do liczenia inflacji wiąże się jednak z wieloma wyzwaniami na poziomie ich zbierania, przetwarzania oraz agregacji. Celem badania jest porównanie metod liczenia indeksów cen wykorzystujących dane scrapowane. Wśród omawianych metod znajdują się m.in. indeks łańcuchowy Jevonsa, indeks GEKS-J oraz indeksy opierające się na metodach rozszerzania lub aktualizowania okna obliczeń (*movement splice method, half splice method, fixed base monthly expanding window, mean splice method*). Omówione zostaną także przykłady wykorzystania danych scrapowanych przez europejskie urzędy statystyczne, a także metody pozyskiwania informacji na temat cen ze stron sklepów internetowych oraz sposób agregacji danych do stanu, w jakim można wykorzystać je od obliczeń indeksów cen.

Słowa kluczowe: big data, web scraping, GEKS-J, CPI

## THE USE OF WEB-SCRAPED DATA TO ANALYSE THE DYNAMICS OF CLOTHING AND FOOTWEAR PRICES

Web scraping is a technique that allows obtaining information from websites automatically. As online shopping grows in popularity, it is an abundant source of information on the prices of goods sold by retailers. The use of scraped data allows not only for significantly reduced costs of price research, but also for the improvement of the precision of inflation estimates, and additionally allows real-time tracking. For this reason, web scraping is a popular research tool both for statistical centers (Eurostat,



British Office of National Statistics, Belgian Statbel) and universities (e.g. the Billion Prices Project conducted at MIT). However, the use of scraped data to calculate inflation is associated with many challenges at the level of their collection, processing and aggregation. The study aims to compare the methods of calculating price indices using scraped data. The discussed methods include Jevons chain index, GEKS-J index, and indexes based on the methods of extending or updating the calculation window (movement splice method, half splice method, fixed base monthly expanding window, mean splice method). Examples of the use of scraped data by the European statistical offices will also be discussed, as well as methods of obtaining information on prices from websites of online stores and the method of aggregating data to a level at which it can be used in the calculation of price indices.

Keywords: Big Data, web scraping, GEKS-J, CPI

**Arkadiusz Kozłowski**

Uniwersytet Gdański

University of Gdańsk

**Jerzy Auksztol**

Uniwersytet Gdański; Urząd Statystyczny w Gdańsku

University of Gdańsk; Statistical Office in Gdańsk

**WYKORZYSTANIE METODY REPREZENTACYJNEJ W BADANIU „WARUNKI PRACY”**

Celem artykułu jest omówienie uwarunkowań leżących u podstaw przygotowania strategii losowania, tj. planu losowania oraz metody estymacji, dla badania „Warunki pracy”, realizowanego dotychczas w Głównym Urzędzie Statystycznym jako badanie pełne. Zmiana metody badania wynika z implementacji podejścia oszczędnościowego (ang. lean approach), którego ideą jest zmniejszenie obciążenia zarówno respondentów, jak i służb statystyki publicznej, przy zapewnieniu jakości i zakresu publikowanych informacji akceptowanych przez ich głównych odbiorców. Proces badawczy obejmował analizę przebiegu oraz wyników dotychczasowych badań, zaproponowanie alternatywnych wariantów oraz przeprowadzenie symulacji komputerowej mającej na celu ocenę efektywności każdego z nich.

Analiza wyników badań przeszłych pozwoliła wyodrębnić następujące cechy charakterystyczne badanego zjawiska: 1) badane zmienne są rzadkie – dla większości jednostek wartości cech są równe zero, 2) rozkłady cech dla jednostek z niezerowymi wartościami są silnie asymetryczne, 3) wykazywanie niezerowej wartości jest silnie skorelowane w czasie. Przeprowadzone symulacje wykazały większą efektywność metod doboru próby, w których prawdopodobieństwo wyboru jednostki jest uzależnione od prawdopodobieństwa wykazywania wartości niezerowej na którejkolwiek z badanych cech. Ostatecznie zaproponowano metodę mieszaną doboru próby, tj. część jednostek jest włączana bez losowania, a część podlega losowaniu. Grupa losowana dobierana jest w sposób warstwowy nieproporcjonalny według dwóch niezależnych podziałów na warstwy. Dla zapewnienia stałych liczebności warstw stosowane jest losowanie zrównoważone. W celu zmniejszenia obciążenia wynikającego z braków odpowiedzi oraz zwiększenia precyzji szacunków wagi wynikające z planu losowania są kalibrowane. Przeprowadzone symulacje wskazują na wysoką efektywność zaproponowanej metody. Zdecydowana większość parametrów szacowana jest ze średnim błędem nieprzekraczającym 10% szacowanej wielkości.

Słowa kluczowe: warunki pracy, strategia losowania, nakładające się warstwy, losowanie zrównoważone, kalibracja

## THE USE OF SAMPLING TECHNIQUES IN THE WORKING CONDITIONS SURVEY

The aim of the article is to discuss the conditions underlying the preparation of the sampling strategy, i.e. the sampling design and the estimation method, for the Working Conditions survey, which was previously carried out by Statistics Poland as a complete enumeration. The change of the research method results from the implementation of the lean approach, the idea of which is to reduce the burden on both the respondents and employees of official statistics, while ensuring that the quality and scope of the published information are accepted by their main recipients. The research process included an analysis of the results of previous surveys, proposing alternative variants of strategy and conducting a computer simulation assessing the effectiveness of each of them.

The analysis of the previous surveys made it possible to distinguish the following features of the studied phenomenon: 1) the studied variables are rare – for most units, the values are zero, 2) the distributions for units with non-zero values are strongly asymmetric, 3) showing a non-zero value is strongly correlated in time. The simulations indicated that most effective are the sampling methods in which the probability of selecting a unit depends on the probability of showing a non-zero value on any of the variables of interest. Finally, a mixed sampling method was proposed, i.e. some units are included in the sample by definition (without being drawn), and some are randomly selected. The sampled part is selected in a stratified sampling with disproportionate allocation and overlapping strata (two-way stratification). In order to ensure a constant allocation, a balanced sampling is used. The sampling weights are calibrated to reduce non-responses bias and increase precision. The simulations results indicate that the proposed strategy meets its objectives. The vast majority of parameters are estimated with mean relative error far below 10%.

Keywords: working conditions, sampling strategy, two-way stratification, balanced sampling, calibration

**Agata Majkowska**

Uniwersytet Gdański

University of Gdańsk

**Kamila Migdał-Najman**

Uniwersytet Gdański

University of Gdańsk

**Krzysztof Najman**

Uniwersytet Gdański

University of Gdańsk

**Katarzyna Raca**

Uniwersytet Gdański

University of Gdańsk

**IDENTYFIKACJA SŁÓW I ZNAKÓW GRAFICZNYCH CHARAKTERYZUJĄCYCH WYRÓŻNIONE GRUPY WIEKOWE UŻYTKOWNIKÓW TWITTERA**

Wpisy użytkowników takich portali jak Twitter są cennym źródłem informacji, z których korzysta wiele przedsiębiorstw na całym świecie. Skuteczność tego kanału komunikacji wynika z prostoty jego użytkowania, formułowania krótkich i zwięzłych wiadomości, stosowania nieformalnego języka, który wpływa na różnorodność zawartości wpisów – od spontanicznych wypowiedzi do wulgaryzmów. Charakteryzuje się szerokim zasięgiem oddziaływania i możliwością nawiązywania bezpośrednich interakcji użytkownika z innymi w sieci. Tweety są źródłem danych wykorzystywanych w badaniach marketingowych, rynkowych, społecznych czy elementem wspierającym podejmowanie decyzji biznesowych. Pozwalają lepiej poznać użytkownika, jego potrzeby i preferencje. Jedną z kluczowych informacji, która jest podstawą tych zastosowań, jest wiek autorów wpisów. Jednak autorzy treści zamieszczanych w internecie często nie podają swojego wieku, co znacząco ogranicza możliwości wykorzystania danych o ich aktywności. Z badań językowych wynika, że przedstawiciele różnych grup wiekowych posługują się odmiennym słownictwem i innymi formami gramatycznymi. Ich wypowiedzi różnią się także pod względem stosowania znaków graficznych, takich jak emotikony, emoji, piktogramy itp. Celem prezentowanego badania jest próba wyróżnienia słów i znaków graficznych, które charakteryzują grupy wiekowe użytkowników Twittera. Do realizacji badania wykorzystano narzędzia text miningu i analizy skupień.

Słowa kluczowe: Twitter, text mining, analiza skupień, emoji, media społecznościowe

## IDENTIFICATION OF THE WORDS AND GRAPHICS SIGNS CHARACTERISING DISTINGUISHED AGE GROUPS OF TWITTER USERS

Messages in social media are a valuable source of information for many companies all over the world. For example, Twitter's value lies in its wide impact potential and the possibility for users to directly interact with others within the network. Hence, messages from this channel are a valuable source of data used in marketing, market and social research, or could support decision-making in business. Twitter also makes it possible to get to know the user better, including his/her needs and preferences. One of the most important pieces of information that is the basis of these uses is the age of the authors of the messages. However, they often do not specify their age, what significantly limits the possibilities of use. Linguistic research shows that representatives of different age groups use different vocabulary and grammatical forms. They also tend to use different graphic signs such as emoticons, emoji, pictograms and other graphic signs. The aim of the research presented in this paper is to attempt to distinguish between age groups by typical words and emoji used in messages on Twitter. Text mining methods and cluster analysis were used to carry out the research.

Keywords: Twitter, text mining, cluster analysis, emoji, social media

**Ewa Panek**

Instytut Geodezji i Kartografii  
Institute of Geodesy and Cartography

**Evangelos Gerasopoulos**

Narodowe Obserwatorium w Atenach  
National Observatory of Athens

**Katarzyna Dąbrowska-Zielińska**

Instytut Geodezji i Kartografii  
Institute of Geodesy and Cartography

**Ali Nadir Arslan**

Fiński Instytut Meteorologii  
Finnish Meteorological Institute

**Orestis Speyer**

Narodowe Obserwatorium w Atenach  
National Observatory of Athens

**Phillip Harwood**

Bruksela

**Anna Burzykowska**

Europejska Agencja Kosmiczna  
European Space Agency

**PROJEKT GAUSS: GENEROWANIE ZAAWANSOWANEGO WYKORZYSTANIA OBSERWACJI ZIEMI DO INTELIGENTNYCH STATYSTYK**

GAUSS (Generating Advanced Usage of Earth Observation for Smart Statistics) to 18-miesięczny projekt finansowany przez Europejską Agencję Kosmiczną (numer kontraktu: 4000136038/21/I/BG), prowadzony przez Narodowe Obserwatorium w Atenach (NOA) we współpracy z Fińskim Instytutem Meteorologii (FMI), Instytutem Geodezji i Kartografii w Warszawie (IGiK) oraz firmą Evenflow z siedzibą w Brukseli. Jego celem jest zaprezentowanie możliwości wykorzystania danych satelitarnych w statystyce publicznej w zakresie monitorowania: jakości powietrza, zasobów wodnych, pokrywy śnieżnej oraz zasięgu i kondycji terenów zielonych w poszczególnych jednostkach administracyjnych. Dzięki ścisłej współpracy konsorcjum z Głównym Urzędem Statystycznym możliwe jest przedstawienie perspektyw wykorzystania danych satelitarnych do

dalszego rozwoju krajowej statystyki w aspektach wykraczających poza zakres tego projektu. Projekt GAUSS, w odpowiedzi na potrzebę krajowych jednostek statystycznych, jaką jest dostarczanie odbiorcom rzetelnych danych dotyczących społeczeństwa i środowiska, ma na celu opracowanie metodyki wykorzystania danych satelitarnych w stworzeniu nowych wskaźników jakości środowiska przyrodniczego, które mają wpływ na dobrostan człowieka. Na podstawie poszczególnych studiów przypadków w ramach projektu zostaną opracowane zalecenia dotyczące wdrażania danych satelitarnych do statystyki publicznej. Uwzględnią one nie tylko kwestie techniczne, lecz także bariery operacyjne i regulacyjne w stosowaniu produktów satelitarnych w statystykach lokalnych. Projekt GAUSS wskaże możliwości operacyjnego wdrażania wykorzystania danych satelitarnych do wzbogacenia baz danych europejskich jednostek statystycznych (Eurostat) wskaźnikami umożliwiającymi monitorowanie stanu środowiska przyrodniczego.

Słowa kluczowe: smart statistics, dane satelitarne, jakość powietrza, susza, zielone gminy

## GENERATING ADVANCED USAGE OF EARTH OBSERVATION FOR SMART STATISTICS

GAUSS (Generating Advanced Usage of Earth Observation for Smart Statistics) is an 18-month project financed by the European Space Agency (Contract No.: 4000136038/21/I/BG), led by the National Observatory in Athens (NOA) in collaboration with the Finish Meteorological Institute (FMI), the Institute of Geodesy and Cartography in Warsaw (IGiK) and Evenflow company from Brussels. Its aim is to present the use of satellite data in official statistics in monitoring the air quality, water resources, snow cover, and the range and condition of green areas in individual administrative units. Furthermore, it will be possible to present prospects of satellite data for the further development of national statistics in the aspects beyond this project's scope. This will be possible thanks to the close collaboration between the consortium and Statistics Poland. In response to the needs of the national statistical units, the GAUSS project, whose purpose is to offer the most reliable data on the society and the environment, aims to develop a methodology for the use of satellite data to create new environmental quality indicators that affect human well-being. Based on individual case studies, the project will develop recommendations for implementing satellite data in official statistics. They will take into account technical issues and operational and regulatory barriers of satellite products usage in local statistics. The GAUSS project will indicate the possibilities of the operational implementation of satellite data to enrich the databases of European statistical units (Eurostat) with indicators enabling the monitoring of the state of the natural environment.

Keywords: smart statistics, EO, air quality, drought, green indicators

**Mirosław Szreder**

Uniwersytet Gdański

University of Gdańsk

**O NIEKTÓRYCH POZORNYM ZALETACH DUŻYCH PRÓB WE  
WNISKOWANIU STATYSTYCZNYM**

Duże liczebnie próby badawcze, do których dostęp staje się coraz łatwiejszy, zdają się stanowić dla niektórych badaczy usprawiedliwienie odrzucenia części założeń klasycznego modelu wnioskowania statystycznego. Dotyczy to w szczególności wymogu losowego generowania obserwacji w próbie. Czynnikiem wzmacniającym to zjawisko jest postęp, jaki dokonał się w doskonaleniu oprogramowania statystycznego. Jego użytkownik jest często osobą, która potrafi wykonać szereg obliczeń statystycznych, ale nie jest świadoma założeń, jakie powinny być spełnione, aby wynikiom tym nadać poprawną statystyczną i naukową interpretację. Celem opracowania jest przedstawienie konsekwencji nierespektowania lub złego rozumienia założeń wnioskowania statystycznego w sytuacjach, gdy badacz posiada dużą liczebnie próbę badawczą. Konsekwencje te są analizowane przy użyciu metod analitycznych i symulacyjnych. We wnioskach wskazano, że duża próba, obok wielu zalet, jakie posiada, rodzi niebezpieczeństwa w zakresie: ignorowania niektórych założeń modelu wnioskowania statystycznego, złego rozumienia jej wpływu na całkowity błąd badania oraz dyskusyjnego odwoływania się do rozstrzygnięć w testowaniu hipotez statystycznych na podstawie wskaźnika p-value.

Słowa kluczowe: wnioskowanie statystyczne, próba losowa, błędy losowe i nielosowe, weryfikacja hipotez, p-value

**ON SOME ILLUSIVE ADVANTAGES OF LARGE SAMPLE SIZES IN  
STATISTICAL INFERENCE**

Easily accessible large samples seem to be a good excuse for some researchers to ignore certain assumptions of the classical model of statistical inference. More specifically, this mainly concerns rejecting the requirement of having random sample observations. This tendency has been strengthened by a considerable progress in developing statistical software. A user of the software can quickly obtain various statistical results, and at the same time may be unaware of the assumptions that have to be met in order to properly interpret the results. The main objective of this paper is to present consequences of misunderstanding or ignoring by some researchers the assumptions of



statistical inference when large samples are at their disposal. Those consequences are analysed and discussed using analytical and simulation methods. The paper shows that a large sample, apart from its advantages, may encourage skipping some assumptions of statistical inference, the negligence of non-random errors and may contribute to false decision-making in testing hypotheses using p-value.

Keywords: statistical inference, random sample, random and non-random errors, testing hypotheses, p-value

**Katarzyna Kotańska**

Agencja Restrukturyzacji i Modernizacji Rolnictwa

Agency for Restructuring and Modernisation of Agriculture

**Jolanta Kossakowska**

Agencja Restrukturyzacji i Modernizacji Rolnictwa

Agency for Restructuring and Modernisation of Agriculture

**WKŁAD ARiMR W TWORZENIE STATYSTYKI PUBLICZNEJ Z ZAKRESU ROLNICTWA I ROZWOJU WSI**

Celem artykułu jest przedstawienie ewolucji zakresu danych raportowanych przez Agencję Restrukturyzacji i Modernizacji Rolnictwa (ARiMR) do systemu statystyki publicznej, źródeł ich pozyskiwania oraz sposobów ich przekazywania. Agencja jest kluczową instytucją realizującą politykę państwa na rzecz rozwoju rolnictwa i polskiej wsi. Wdraża instrumenty współfinansowane z budżetu Unii Europejskiej oraz udziela pomocy ze środków krajowych. Jej głównym celem jest zapewnienie prawidłowego, sprawnego i skutecznego wydatkowania środków pochodzących z funduszy unijnych i krajowych. W ramach realizowanych zadań ze wsparcia udzielanego przez ARiMR w ostatnich latach korzysta corocznie około 1,3 mln rolników. Beneficjentami pomocy wypłacanej przez ARiMR są także przedsiębiorcy, rybacy, samorządy oraz organizacje producentów, organizacje charytatywne, koła gospodyń wiejskich, uczniowie oraz samorządy. Agencja ustawowo prowadzi również systemy rejestracji danych, w tym System Identyfikacji i Rejestracji Zwierząt (IRZ), System Identyfikacji Działek Rolnych (LPIS), system danych przestrzennych GIS, Krajowy System Ewidencji Producentów (zarejestrowano około 2,3 mln podmiotów), a także wiele rejestrów, w tym Rejestr uznanych organizacji charytatywnych i uprawnionych innych jednostek oraz osób fizycznych, rejestry grup i organizacji producentów oraz ich zrzeszeń, Krajowy Rejestr Kół Gospodyń Wiejskich i Rejestr Upoważnionych Gwarantów. W ciągu ponad 28 lat działalności ARiMR rozszerzył się zakres przekazywanych danych, jak również zostały udoskonalone systemy ich przekazywania. Dane statystyczne przekazywane przez ARiMR do Głównego Urzędu Statystycznego są rzetelnym i wiarygodnym źródłem. Wtórne wykorzystanie danych z zasobów administracji publicznej pozwala na zmniejszenie kosztów zbierania danych oraz uciążliwości dla respondentów, jednocześnie zapewniając ich wysoką jakość.

Słowa kluczowe: informacja statystyczna, administracyjne źródła danych, rejestry, ARiMR, fundusze unijne

## CONTRIBUTION OF ARMA TO THE DEVELOPMENT OF PUBLIC STATISTICS ON AGRICULTURE AND RURAL DEVELOPMENT

The aim of the paper is to present the evolution of the scope of data reported by the Agency for Restructuring and Modernisation of Agriculture to the system of official statistics, the sources of their acquisition and the ways of their transmission. The Agency for Restructuring and Modernisation of Agriculture is a key institution implementing the state policy for the development of agriculture and the Polish rural areas. It implements instruments co-financed from the EU budget and provides assistance from domestic funds. The main objective of the Agency is to ensure the correct, efficient and effective spending of EU and national funds. Approximately 1.3 million farmers have benefitted annually from the support provided by the Agency in recent years. The beneficiaries of ARMA's aid are also entrepreneurs, fishermen, local governments and producer organisations, charity organisations, farmer housewives' circles, schoolchildren and local governments. Statutorily, ARMA also operates data registration systems, including the Animal Identification and Registration System (IRZ), the Land Parcel Identification System (LPIS) and the GIS spatial data System, the National Producers Registration System (approximately 2.3 million entities were registered), as well as numerous registers, including the Register of the Recognised Charity Organisations and Other Entitled Entities and Individuals, the Registers of Producer Groups and Organisations and their Associations, the National Register of Farmer Housewives' Circles and the Register of the Authorised Guarantors. Throughout over 28 years of the Agency's activity, the scope of transmitted data has been extended and the systems of their transmission improved. Statistical data transmitted by ARMA to Statistics Poland is a reliable and credible source. Secondary use of data from public administration resources allows the reduction of data collection costs and limits inconvenience for respondents.

Keywords: statistical information, administrative data sources, registers, ARMA, EU funds

**Marcin Szymkowiak**

Uniwersytet Ekonomiczny w Poznaniu; Urząd Statystyczny w Poznaniu  
Poznań University of Economics and Business, Statistical Office in Poznań

**Kamil Wilak**

Uniwersytet Ekonomiczny w Poznaniu; Urząd Statystyczny w Poznaniu  
Poznań University of Economics and Business, Statistical Office in Poznań

**ESTYMACJA WYBRANYCH CHARAKTERYSTYK RYNKU PRACY NA POZIOMIE PODREGIONÓW W ŚWIETLE PODEJŚCIA KALIBRACYJNEGO**

Głównym celem referatu jest przedstawienie wybranych wyników estymacji najważniejszych charakterystyk rynku pracy na poziomie podregionów ogółem oraz w bardziej szczegółowych domenach (według miejsca zamieszkania, płci i wieku). W przeprowadzonym badaniu wykorzystano dane jednostkowe pochodzące z Badania Aktywności Ekonomicznej Ludności (BAEL) za lata 2017–2019. Na potrzeby realizacji założonego celu rozwinięto podejście kalibracyjne, które obecnie stosowane jest w ramach BAEL. Estymacji podlegały wartości globalne pracujących, bezrobotnych i biernych zawodowo, a także współczynnik aktywności ekonomicznej, wskaźnik zatrudnienia i stopa bezrobocia. Takie charakterystyki, wyznaczone na podstawie BAEL, nie są publikowane przez Główny Urząd Statystyczny na poziomie podregionów. Problemem jest przede wszystkim szacowanie wskaźników dotyczących zjawiska bezrobocia, które w ostatnich latach cechuje się niskim poziomem. Związana z tym niewielka reprezentacja osób bezrobotnych w próbie skutkuje niską precyzją estymacji bezpośredniej wskaźników w obszarze tego zjawiska. W referacie zostanie poddana ocenie możliwość wykorzystania danych pochodzących z rejestru osób bezrobotnych, a także danych ludnościowych na poziomie podregionów w roli dodatkowych zmiennych pomocniczych. Autorzy przeprowadzą dyskusję na temat możliwości i ograniczeń wykorzystania podejścia kalibracyjnego w BAEL celem uogólniania wyników na niższe poziomy agregacji przestrzennej niż te, dla których obecnie GUS publikuje szacunki związane z rynkiem pracy w Polsce.

Słowa kluczowe: rynek pracy, podejście kalibracyjne, Badanie Aktywności Ekonomicznej Ludności

## ESTIMATION OF SELECTED CHARACTERISTICS OF THE LABOUR MARKET AT THE LEVEL OF SUB-REGIONS IN THE LIGHT OF THE CALIBRATION APPROACH

The main goal of the presentation is to propose a method for estimating key characteristics of the labour market at the level of sub-regions and for more detailed domains (by place of residence, sex and age) using data from the Polish Labour Force Survey (LFS) for 2017–2019. Methodology: To achieve this goal, the authors extended the calibration approach currently used in the LFS. The total counts of the employed, unemployed and economically inactive persons as well as the economic activity rate, employment rate and unemployment rate were estimated. These characteristics, calculated on the basis of LFS data, are not published by Statistics Poland at the level of sub-regions or for domains more precisely defined. The main challenge is to estimate the unemployment indicators. Since the unemployment in recent years has been rather low, the number of unemployed persons in the LFS sample has also been small, resulting in the low precision of directly estimated indicators of the unemployment. In the presentation, the authors will assess the possibility of using data from the unemployment register, additional auxiliary variables as well as population data at the sub-regional level. In the presentation, the authors will formulate conclusions about the possibilities and limitations of the calibration approach in the LFS to generalize the results at lower levels of spatial aggregation than those for which Statistics Poland currently publishes estimates related to the labour market in Poland.

Keywords: labour market, calibration approach, Labour Force Survey

**Jacek Wesołowski**

Główny Urząd Statystyczny; Politechnika Warszawska  
Statistics Poland; Warsaw University of Technology

**Robert Wieczorkowski**

Główny Urząd Statystyczny  
Statistics Poland

**Wojciech Wójciak**

Politechnika Warszawska  
Warsaw University of Technology

**EFFECTIVE ALGORITHMS OF OPTIMAL SAMPLE ALLOCATION IN STRATIFIED SCHEMES**

The optimal Tchuprov-Neyman (CN) allocation can give infeasible sample strata sizes, which are larger than the population strata sizes. A remedy is the iterative Neyman procedure, which in consecutive steps takes collections of take-all strata to the sample and reduces the original problem stepwise until the CN allocation gives feasible sample strata sizes in the remaining part of the population. In spite of the fact that this approach is popularly used, only recently it has been formally proved that the allocation it gives is optimal under the upper bounds for the sample strata sizes. Often a natural requirement in sample allocation is two-sided (lower and upper) bounds for the sample strata sizes. It appears that in such a situation the naive obvious modification of the iterative Neyman procedure does not work correctly. I will present a new algorithm, which while being more complicated, preserves the main idea of the iterative Neyman approach, i.e. step-wise choice of take-min or take-max strata to the sample and sufficiently careful method of the population reduction. The theoretical base of the algorithm are methods of convex analysis, and among them, the Karush-Kuhn-Tucker conditions in particular, which lead to an elegant and useful description of the optimal allocation vector.

Keywords: Tchuprov-Neyman optimal allocation, Neyman iterative procedure, box-constraints for sample strata sizes, Karush-Kuhn-Tucker conditions

**Katarzyna Widera**

Urząd Statystyczny w Opolu; Politechnika Opolska  
Statistical Office in Opole; Opole University of Technology

**Marcin Związek**

Urząd Statystyczny w Opolu  
Statistical Office in Opole

**Dastin Lipp**

Urząd Statystyczny w Opolu  
Statistical Office in Opole

**KLASTERYZACJA DANYCH SCRAPOWANYCH ZASTOSOWANYCH DO WYLICZENIA WSKAŹNIKÓW CEN TOWARÓW I USŁUG KONSUMPCYJNYCH – BADANIE EKSPERYMENTALNE**

Celem badania były eksperymentalne prace polegające na wykorzystaniu metody CLIP (ang. clustering large datasets into price indices) do obliczeń wskaźników cen towarów i usług konsumpcyjnych. Na podstawie danych uzyskanych metodą web scrapingu tworzono klastry homogenicznych towarów z pośrednim udziałem uczenia maszynowego. Klasteryzacja towarów zakłada, że wybór konsumenta ma charakter substytucyjny. Zastosowanie metody tworzenia klastrów uzasadniono praktyką wyliczania wskaźników cen towarów – reprezentantów. Często spotyka się sytuację, gdy dany towar przestaje być obserwowany i w konsekwencji następuje zerwanie szeregu czasowego. Dla utworzonych klastrów towarów obliczono wskaźniki cen. Rozwiązanie z zastosowaniem teorii klastrów pozwoliło na ciągłość w obliczaniu wskaźników. Tworzenie klastrów (grup towarów substytucyjnych z punktu widzenia konsumenta) i obliczanie na ich podstawie wskaźników cen odbywa się bez utraty ciągłości w przypadku braku danych pojedynczego towaru.

Słowa kluczowe: CLIP, web scraping, indeksy cen

## **CLUSTERING OF SCRAPED DATA USED TO CALCULATE CONSUMER PRICE INDICES AND CONSUMER SERVICES – EXPERIMENTAL STUDY**

The purpose of the study was experimental work involving the use of the CLIP method (clustering large datasets into price indices) to calculate consumer price indices. Based on the data obtained by web scraping, clusters of homogeneous goods were created with indirect machine learning. Clustering of goods assumes that the consumer's choice is of a substitutive nature. The application of the clustering method was justified by the practice of calculating price indices of goods – representatives. Often a situation is encountered where a given commodity ceases to be observed and, consequently, the time series is broken. Price indices were calculated for the created clusters of goods. The cluster theory solution allowed for smooth calculation of the indices, thanks to creating clusters (groups of goods that are substitutable from the consumer's point of view) and calculating price indices from them without losing continuity in the absence of single commodity data.

Keywords: CLIP, web scraping, price indexes



**Anita Abramowska-Kmon**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

**WHAT MAKES PEOPLE AGED 50+ IN POLAND HAPPY? THE ROLE OF LIFESTYLE: EVIDENCE FROM PANEL DATA**

The main aim of this paper is to investigate the determinants of subjective quality of life among Poles aged 50+, with a special focus on leisure activities, participation in social and religious activities and unhealthy behaviours. Data from four waves of the Social Diagnosis Polish panel survey carried out between 2009 and 2015 were used. Two types of ordered response panel models were employed: random effects and fixed effects ordered logit models. The models were estimated for the total population aged 50+ and two subpopulations: people aged 50–64 and people aged 65+. The modelling outcomes are mostly in line with those obtained in previous research. For example, participation in religious activities contributed positively to subjective well-being among Poles aged 50+, especially among those aged 50–64 who are more heterogenous with this respect than older people. Engagement in local community activities had no effect on happiness. Unhealthy behaviours (ever smoking and consuming too much alcohol) diminished subjective well-being. It holds true for people aged 50–64 who are probably more exposed to abuse alcohol than their older fellows. Time spent watching TV affected positively subjective well-being among both groups of respondents. A similar effect was found for participation in cultural events, especially among people aged 50–64. The results show the importance of lifestyle for subjective well-being of people 50+ in Poland as well as the moderation effects of age with respect to the associations with selected variables related to lifestyle among people aged 50–64 and aged 65+.

Keywords: happiness, cultural and physical activity, watching TV, alcohol abuse, smoking, people 50+

**Anita Abramowska-Kmon**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**Milena Chełchowska**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**Martin Piotrowski**

Uniwersytet w Oklahomie  
University of Oklahoma

**CONTACTS WITH OTHER PEOPLE AND MENTAL WELL-BEING DURING THE COVID-19 PANDEMIC**

Contact with other people (especially children, relatives or friends) may be a source of emotional support contributing to a higher level of subjective quality of life of individuals. The restrictions introduced during the Covid-19 pandemic caused a significant reduction in the frequency of social contact, especially among those aged 50+, which may have had a negative impact on their mental health. The aim of this paper is to analyze the relationship between the frequency of contact with other people and mental well-being among people aged 50+ utilizing longitudinal data available in the SHARE database. We used data from two rounds of the Survey of Health, Ageing and Retirement in Europe carried out during the pandemic (in 2020 and 2021) (SHARE-Covid). The sample was limited to respondents aged 50+ at wave 2 in 27 European countries and Israel. We estimated structural equation models (SEM) with fixed effects (FE-SEM) in which the dependent variable is a composite measure of mental health measures (depression, nervousness, trouble sleeping, and loneliness) regressed against measures of social contact during the pandemic and other variables that could influence simultaneously mental health and social contact (i.e. death of someone due to Covid-19, Covid symptoms, financial situation, and subjective health). Our results show that changes in frequency of contact with friends tended to be related to changes in mental health, while frequency of contact with relatives did not have statistically significant effects. Contact with children was especially important for women aged 50 or older. Specifically, there is a statistically significant difference between the reference category (“daily” contact) and contact “several times a week,” “about once a week,” and “less often.” Moreover, we observed a pattern of generally worsening mental health accompanying changes towards increasingly lower contact (i.e., a positive gradient of contact-mental health is evident).

Keywords: mental health, mental well-being, social contacts, SHARE, Covid-19

**Janusz Dygaszewicz**

Główny Urząd Statystyczny  
Statistics Poland

## **REALIZACJA NARODOWEGO SPISU POWSZECHNEGO LUDNOŚCI I MIESZKAŃ W 2021 ROKU**

Narodowy Spis Powszechny Ludności i Mieszkań roku 2021 został efektywnie przeprowadzony mimo trudnych warunków pandemicznych w kraju. Rozprzestrzenianie się pandemii COVID-19 miało wpływ na wiele różnych zjawisk społecznych i gospodarczych, jednak podjęto wszelkie możliwe działania zmierzające do zorganizowania spisów w taki sposób, aby zebrać dane najwyższej jakości. Spis charakteryzował się najnowocześniejszymi rozwiązaniami organizacyjnymi i metodologicznymi oraz wdrożeniem nowatorskich technik zbierania danych, łącznie z obowiązkiem spisu przez internet. Zgodnie z zapisami ustawy spis realizowany był w terminie od 1 kwietnia do 30 września 2021 r., według stanu na 31 marca 2021 r. o godz. 24.00. Przeprowadzenie spisu jest związane ze strategicznym znaczeniem tego badania dla funkcjonowania systemów informacyjnych państwa, lokalnych samorządów oraz biznesu.

Słowa kluczowe: spis, spis przez internet, Narodowy Spis Powszechny Ludności i Mieszkań, pandemia Covid-19

## **IMPLEMENTATION OF THE NATIONAL POPULATION AND HOUSING CENSUS IN 2021**

The National Population and Housing Census of 2021 was efficiently conducted despite the difficult pandemic conditions in the country. The spread of the COVID-19 pandemic affected a wide variety of social and economic phenomena, but every possible effort was made to organize the census so as to collect data of the highest quality. The census applied the most innovative organizational and methodological solutions and involved the implementation of state-of-the-art data collection techniques, including compulsory self-enumeration via the Internet. In accordance with the provisions of the relevant Act, the census was implemented between April 1 and September 30, 2021, as of midnight of March 31, 2021. Conducting the census is related to the strategic importance of this survey for the functioning of the state, local governments and business information systems.

Keywords: census, online census, National Population and Housing Census, Covid-19 pandemic

**Artur Mikulec**

Uniwersytet Łódzki

University of Lodz

**Iwona Rosset**

Uniwersytet Łódzki

University of Lodz

**Elżbieta Żądzińska**

Uniwersytet Łódzki

University of Lodz

**ZASTOSOWANIE MODELI ARMAX W BADANIACH NAD STRESEM EKONOMICZNYM WPŁYWAJĄCYM NA KONDYCJĘ BIOLOGICZNĄ POPULACJI LUDZKICH – PONOWNY TEST HIPOTEZY TRIVERSA-WILLARDA W POLSCE**

Hipoteza postawiona przez Triversa i Willarda (1973), zgodnie z którą naturalna selekcja faworyzuje matki rodzące synów w okresie „dobrej kondycji” (korzystne warunki środowiskowe, braku stresu ekonomicznego) i rodzące córki w okresie „słabej kondycji” (niekorzystne warunki środowiskowe, narażenie na stres ekonomiczny) zapoczątkowała badania nad wpływem stresu ekonomicznego i jego oddziaływaniem na liczbę urodzeń dzieci płci męskiej w relacji do liczby urodzeń dzieci płci żeńskiej. Zapoczątkowane przez Catalano i Brucknera (2005) badania nad sposobem weryfikacji tej hipotezy, tzw. hipotezy stresu ekonomicznego nie przyniosły dotychczas jednoznacznych wyników. Celem pracy była ponowna weryfikacja hipotezy Triversa-Willarda. Do analizy wykorzystano półroczne szeregi czasowe wskaźnika proporcji płci w populacji ludzkiej, liczonego jako stosunek urodzeń mężczyzn do kobiet (SSR – secondary sex ratio) lub jako odsetek urodzeń mężczyzn w stosunku do wszystkich urodzeń (SRB – sex ratio at birth) z lat 1995–2020 dla Polski oraz dla jednego z jej regionów – woj. łódzkiego. Materiał badawczy stanowiły dane statystyczne dotyczące ogólnej liczby żywo urodzonych dzieci płci męskiej i żeńskiej w latach 1995–2020 (obserwacje półroczne) dla Polski i woj. łódzkiego. Jako zmienną opisującą poziom stresu ekonomicznego ludności Polski oraz woj. łódzkiego przyjęto wartości konsumpcji prywatnej – spożycie indywidualne gospodarstw domowych w mln jednostek waluty krajowej (w cenach stałych z roku 2015) – dane wyrównane sezonowo i kalendarzowo. Obliczenia wykonano w programie GRETTL. Procedura analizy zaproponowana przez Catalano i Brucknera (2005) została przez autorów rozszerzona (m.in. koncepcja rozwiązania dualnego) i ponownie zastosowana do weryfikacji hipotezy stresu ekonomicznego w Polsce – pozwoliła na „poprawną” weryfikację hipotezy Triversa-Willarda. Wyniki wskazują, że okresy spadku konsumpcji faworyzowały matki rodzące córki, a okresy jej wzrostu faworyzowały matki

rodzące synów – udało się wykazać, że czynnik ekonomiczny ma pozytywny wpływ na kształtowanie się wskaźników proporcji płci w populacji ludzkiej. Jednocześnie uzyskano inną wartość opóźnienia w reakcji SSR i SRB na zmiany konsumpcji w przypadku województwa łódzkiego (1 rok) oraz w skali całego kraju (2 lata).

Wykazano również przydatność w analizie stresu ekonomicznego zarówno wskaźnika SSR, jak i wskaźnika SRB – rozwiązanie dualne.

Słowa kluczowe: hipoteza Triversa-Willarda, wskaźnik proporcji płci w populacji ludzkiej, stres ekonomiczny, spożycie indywidualne gospodarstw domowych, model ARMAX

## **APPLICATION OF ARMAX MODELS IN RESEARCH ON ECONOMIC STRESS INFLUENCING THE BIOLOGICAL CONDITION OF HUMAN POPULATIONS – A RE-TEST OF THE TRIVERS-WILLARD HYPOTHESIS IN POLAND**

The hypothesis formulated by Trivers and Willard (1973) where natural selection favours mothers giving birth to sons in the period of 'good conditions' (favourable environmental conditions with the absence of economic stress) and giving birth to daughters in the period of 'poor conditions' (unfavourable environmental conditions, involving exposure to economic stress) started the research on economic stress and its impact on the number of births of male babies in relation to the number of female babies. The research initiated by Catalano and Bruckner (2005) on the methods of testing the hypothesis viz. the economic stress hypothesis, has not provided unequivocal results. The aim of the study was to re-test the Trivers-Willard hypothesis. For the purpose of the analysis we used semi-annual time series of the human sex ratio, counted as the proportion of male to female births (SSR – secondary sex ratio) or as the percentage of male births (SRB – sex ratio at birth) for the years 1995-2020 for Poland and one of its regions – Łódzkie Voivodship. Research material comprised statistical data on the total number of live births of male and female babies in the years 1995-2020 (semi-annual observations) for Poland and Łódzkie Voivodship. The variable describing the level of economic stress of the population of Poland and Łódzkie Voivodship was assumed to be the value of private consumption – the final consumption expenditure of households in millions of the national currency units (at constant prices, with 2015 as the reference year) – seasonally and calendar adjusted data. Calculations were done with the use of the GRETTL program. The analysis procedure proposed by Catalano and Bruckner (2005) was expanded by the authors (e.g. the concept of dual solution) and re-applied for testing the hypothesis of economic stress in Poland, which allowed the 'correct' verification of the Trivers-Willard hypothesis. Research results revealed that periods of decrease in consumption favoured mothers giving birth to daughters, while periods of increase in consumption favoured mothers giving birth to sons – the economic factor proved to have a positive influence on the way human sex ratios were shaped. Simultaneously, a different value of lag in SSR and SRB reaction to consumption was obtained in case of the Łódzkie Voivodship (1 year) and the whole country (2 years).

Also, the usefulness of both secondary sex ratio (SSR) and sex ratio at birth (SRB) in analysis on economic stress – the dual solution, was demonstrated.

Keywords: Trivers-Willard hypothesis, human sex ratio, secondary sex ratio, sex ratio at birth, economic stress, final consumption expenditure of households, ARMAX model

**Dorota Szałtys**

Główny Urząd Statystyczny  
Statistics Poland

**Janusz Pappelbon**

Urząd Statystyczny w Olsztynie  
Statistical Office in Olsztyn

**Małgorzata Cierniak-Piotrowska**

Główny Urząd Statystyczny  
Statistics Poland

**Eliza Panfilów**

Urząd Statystyczny w Olsztynie  
Statistical Office in Olsztyn

**Wojciech Wasilewski**

Urząd Statystyczny w Olsztynie  
Statistical Office in Olsztyn

## NOWA JAKOŚĆ W BADANIU PRZYCZYN ZGONÓW

Celem prezentacji jest przedstawienie działań podjętych przez statystykę publiczną w odpowiedzi na rosnące potrzeby i oczekiwania odbiorców danych o zgonach i ich przyczynach, szczególnie w związku z pandemią COVID-19. Wyzwaniem w dążeniu do spełnienia tych potrzeb jest skrócenie czasu opracowywania i publikowania rzetelnych i porównywalnych statystyk dotyczących umieralności. Statystyka przyczyn zgonów jest najważniejszym i najbardziej złożonym zagadnieniem z zakresu badania umieralności oraz kluczowym elementem systemu monitorowania zdrowia publicznego. Długotrwały i niezautomatyzowany proces opracowywania danych, jak również nieprecyzyjne opisywanie przyczyn zgonów, stanowiły w ostatnich latach barierę przy bieżącym wykorzystywaniu tych statystyk oraz zdecydowanie utrudniały porównywalność w odniesieniu do innych krajów. W 2020 r. GUS oraz US w Olsztynie podjął działania na rzecz skrócenia procesu opracowywania danych o umieralności, w tym zgonów według przyczyn. Prace te zbiegły się w czasie z podjętą przez KPRM inicjatywą opracowania kompleksowej reformy regulacji administracyjnych związanych z ruchem naturalnym ludności i ochroną miejsc pamięci. Jednym z założeń Ustawy o cmentarzach i chowaniu zmarłych jest wprowadzenie elektronicznej postaci karty zgonu, co zdecydowanie ułatwi i przyspieszy etap gromadzenia przez GUS danych o zgonach. Wdrożenie e-karty stworzy możliwość wykorzystania oprogramowania do automatycznego kodowania przyczyn zgonu. Przyjęta w 2021 r. przez GUS nowa metodologia badania zgonów zakłada kodowanie wszystkich przyczyn zawartych na karcie zgonu. Podejmowane są również działania mające na celu poprawę jakości orzekania

o przyczynach zgonu. Zmiany organizacyjne i metodologiczne wprowadzone na przełomie lat 2020 i 2021 spowodowały znaczne skrócenie procesu opracowywania danych o zgonach i możliwość udostępniania tych danych w cyklach kwartalnych z przyczynami zgonów oraz w cyklach tygodniowych bez przyczyn.

Słowa kluczowe: statystyka zgonów, umieralność, przyczyny zgonów, epidemiologia, ruch naturalny

## A NEW QUALITY IN THE CAUSES OF DEATH SURVEY

The aim of the presentation is to discuss the activities undertaken by public statistics, responding to the growing needs and expectations of data users on the number and causes of deaths, especially in connection with the COVID-19 pandemic. The challenge in meeting these needs is to reduce the time taken to compile and publish reliable and comparable mortality statistics. Statistics of the causes of death are a key component of the public health monitoring system. Long-lasting and non-automated data processing, as well as imprecise descriptions of the causes of death have in the recent years been a barrier to the current use of these statistics and significantly hindered their comparability with other countries. In 2020, Statistics Poland took steps to shorten the process of compiling data on mortality, including causes of death. These undertakings coincided with an initiative of the Chancellery of the Prime Minister to develop a comprehensive reform of administrative regulations related to vital statistics and the protection of memorial sites. One of the assumptions of the Act on cemeteries and burying the deceased is the introduction of an electronic form of a death certificate, which will definitely facilitate and accelerate the stage of collecting data on deaths by Statistics Poland. The implementation of a death e-certificate will allow the use of software for an automatic coding of causes of death. The new methodology of death surveying adopted by Statistics Poland in 2021 assumes encoding all causes of death included in the death certificate. Actions are also taken to improve the quality of determining causes of death. The organizational and methodological changes introduced at the turn of 2020 and 2021 resulted in a significant shortening of the process of compiling data on deaths and the possibility of sharing this data on a quarterly basis with causes of death and on a weekly basis without the causes.

Keywords: death statistics, mortality, causes of death, epidemiology, vital statistics



**Maciej Potyra**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

**PROJEKCJA PROBABILISTYCZNA LUDNOŚCI W ALTERNATYWNYM UJĘCIU TERYTORIALNYM – PRZYKŁAD WOJEWÓDZTWA MAŁOPOLSKIEGO**

Referat przedstawia propozycję alternatywnego podziału terytorialnego Polski (na 30 jednostek), który zostanie wykorzystany do opracowania probabilistycznych projekcji ludności na poziomie regionalnym. Zarówno tradycyjny podział administracyjny (województwa, powiaty), jak i podział na regiony i podregiony nie dostarczają informacji o procesach demograficznych niezbędnych do opracowania założeń projekcji probabilistycznych. Jednostki wyższego szczebla (województwa, regiony) są zbyt duże, by pokazać w pełni zmienność terytorialną tych procesów, zaś jednostki niższego szczebla (powiaty, podregiony) są za małe. Alternatywny podział terytorialny powstał poprzez łączenie powiatów w większe, spójne pod względem charakterystyki demograficznej obszary. Do ich utworzenia wykorzystano analizę skupień opartą na podstawowych miernikach demograficznych: przekrojowym współczynniku dzietności, oczekiwanym trwaniu życia noworodka oraz strumieniach migracji. Taka klasyfikacja szczególnie podkreśla kontrast pomiędzy głównymi aglomeracjami miejskimi a obszarami peryferyjnymi, uwzględniając jednak również zróżnicowanie tych ostatnich. Zastosowanie tej metody wyznaczania projekcji regionalnych ludności jest zilustrowane wynikami przewidywań zmian ludności do 2040 r. dla woj. małopolskiego. Jego terytorium jest podzielone na aglomerację krakowską oraz część peryferyjną: południową („górską”) oraz północną.

Słowa kluczowe: podział terytorialny, analiza skupień, prognoza ludności, aglomeracja miejska

**PROBABILISTIC POPULATION PROJECTION USING ALTERNATIVE TERRITORIAL DIVISION - THE EXAMPLE OF MAŁOPOLSKIE VOIVODSHIP**

The presentation describes a proposal for an alternative territorial division of Poland (into 30 units) to develop a probabilistic projection of the population at regional level. The traditional administrative division (into voivodships and poviats) and the division into regions and sub-regions fail to provide information on the demographic processes necessary to develop assumptions for probabilistic projections. Higher-level units (voivodships, regions) are too large to fully show the territorial variability of these processes, while lower-level units (poviats, subregions) are too small. The alternative territorial division was created by merging poviats into larger areas, consistent in

terms of demographic characteristics. For this purpose, a cluster analysis based on basic demographic measures, i.e. the cross-sectional fertility rate, life expectancy of the newborn and migration flows was used. Such a classification emphasizes the contrast between major urban agglomerations and peripheral areas, but also takes into account the diversity of the latter. The application of this method of computing regional population projections is illustrated by the results of the predictions of population changes for Małopolskie Voivodship until 2040. Its territory is divided into the Kraków agglomeration and the peripheral southern ('highland') and northern parts.

Keywords: territorial division, cluster analysis, population forecast, urban agglomeration

**Marta Styrac**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

**ROZWOODY W POLSCE: INTENSYWNOŚĆ ZJAWISKA I ZMIANY STRUKTURY W CZASIE**

Rozwoody jako istotny element kształtujący życie rodzinne i wpływający na dobrostan zarówno partnerów, jak i pochodzących ze związku dzieci są ważnym zjawiskiem społecznym. Pomimo tego wydaje się, że wiedza na temat zjawiska w Polsce jest niewystarczająca. Składa się na to niewielka liczba badań prowadzonych na danych indywidualnych pochodzących z badań ankietowych oraz małe wykorzystanie danych zagregowanych udostępnianych przez GUS. Celem niniejszego referatu jest przedstawienie statystycznego obrazu zjawiska rozwodów w Polsce oraz jego zmian w czasie od lat 90. XX w. Analizie poddana zostanie intensywność rozwodów oraz ich struktura wg przyczyn rozpadu małżeństwa, liczby i wieku małoletnich dzieci, podziału władzy rodzicielskiej, długości trwania małżeństwa itd. Wybrane wskaźniki zostaną także zaprezentowane w porównaniu do innych krajów europejskich. Wykorzystane zostaną metody statystyki opisowej oraz dane zagregowane na temat rozwodów dostępne w ramach statystyki publicznej. W celu analizy zmian intensywności zjawiska w czasie zostanie zastosowany przekrojowy ogólny współczynnik rozwodów standaryzowany według długości trwania małżeństwa. Nie jest to wskaźnik udostępniany w ramach statystyki publicznej, ale możliwe jest jego obliczenie przy pomocy danych Eurostatu. Dane dotyczące zjawiska rozwodów zostaną zestawione z poglądami i postawami na temat małżeństwa i rozwodów pochodzącymi z badań ankietowych (np. GGS, EVS, ESS) i badań opinii publicznej. Szeregi czasowe danych opisujących intensywność rozwodów w czasie i zmiany ich struktury według cech raportowanych w statystyce publicznej pozwolą na uaktualnienie wiedzy o zjawisku rozwodów w Polsce oraz o jej ujęcie zarówno w perspektywie zmian zachodzących w czasie, jak i w perspektywie porównawczej z krajami europejskimi. Na podstawie przeprowadzonej analizy zostaną także sformułowane rekomendacje dotyczące udostępniania i zakresu danych o rozwodach w statystyce publicznej.

Słowa kluczowe: rozwód, małżeństwo, statystyka publiczna, Polska

## A STATISTICAL PICTURE OF DIVORCE IN POLAND

Divorce is a crucial factor for family life as well as for the well-being of partners and potential children from the union. Our knowledge about this phenomenon in Poland is insufficient because there have been few quantitative studies of divorce based on survey individual data, neither the aggregate data provided by the Central Statistical Office has been used to a satisfactory extent. With my study I intend to fill this gap and present the statistical picture of divorce in Poland and how it has changed over the last thirty years. I will analyse the frequency of divorces and their structure by causes of dissolution, number and age of children, care arrangement for minor children, duration of marriage, etc. Selected indicators will be presented in comparative European perspective. To this aim I will use descriptive statistics of aggregate data provided in the Polish and European social statistics. To follow the intensity of divorce, I will calculate the period total divorce rate standardized for the union duration. This indicator is not available in the official statistics but it can be computed with the data provided by Eurostat. The data on divorce will be supplemented with data on views and attitudes towards marriage and divorce from social surveys (e.g. Generations and Gender Survey, European Value Survey, European Social Survey) and opinion polls. With the time series on divorce intensity and structure by the characteristics published in the social statistics I will update the statistical picture of divorce in Poland. The investigation covers changes in time and comparative perspective against the European countries. Based on the analysis I will also formulate comments and recommendations regarding the scope and access to data on divorce.

Keywords: divorce, marriage, vital statistics, Poland

**Piotr Szukalski**

Uniwersytet Łódzki

University of Lodz

## **GDZIE POLKI I POLACY ZAWIERAJĄ ZWIĄZKI MAŁŻEŃSKIE Z OBCOKRAJOWCAMI?**

Wskutek masowej emigracji z Polski po przystąpieniu do UE oraz dużego napływu cudzoziemców do Polski w ostatnim dziesięcioleciu, w ostatnich latach wzrasta liczba małżeństw binacjonalnych zawieranych przez polskich obywateli. Małżeństwa takie mogą być rejestrowane zarówno w jednym z krajów pochodzenia nowożeńców, jak i w innym kraju będącym miejscem ich faktycznego zamieszkiwania. Celem referatu jest porównanie danych pochodzących z dziewięciu europejskich państw nt. liczby rejestrowanych w ostatnich latach małżeństw odmiennej płci obywateli RP z obywatelami owych państw z danymi pochodzącymi ze statystyk GUS. Porównanie takie wskazuje, że w przypadku małżeństw zawieranych z obywatelami Europy Zachodniej i Południowej dominują związki rejestrowane poza Polską, podczas gdy w przypadku państw Europy Wschodniej przeważają małżeństwa rejestrowane w Polsce. Wskazuje to na wybieranie jako miejsca rejestracji małżeństwa kraju faktycznego zamieszkiwania.

Słowa kluczowe: małżeństwa mieszane, małżeństwa binacjonalne, studia porównawcze, Polska, XXI w.

## **WHERE DO POLES REGISTER THEIR MARRIAGES WITH FOREIGNERS?**

The number of mixed marriages which Polish citizens enter into has significantly increased in the last dozen years due to the mass emigration of Poles following the EU accession and to mass immigration from Eastern European countries to Poland. The marriages could be entered into in one of the countries of the nupturients' citizenship or in a country where the pair actually lives in. The paper aims to compare data from 9 European countries to Polish ones to determine the preferences of the opposite sex nupturients in terms of the registration of the place of their unions. The comparisons reveal a tendency to register mixed marriages with citizens of Western and Southern European countries outside of Poland and in the case of marriages with Eastern European countries' citizens in Poland. The most important result of the study indicates the importance of the actual place of residence in choosing the place of marriage registration.

Keywords: mixed marriage, binational marriage, comparative studies, Poland, 21st century

**Krzysztof Tymicki**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

## DZIETNOŚĆ W UJĘCIU KOHORTOWYM I PRZEKROJOWYM. ANALIZY NA PODSTAWIE DANYCH Z REJESTRACJI URODZEŃ

Analizy porównują przemiany dzietności w Polsce w ujęciu kohortowym i przekrojowym do roku 2020 łącznie. Zestawienie tych ujęć analitycznych umożliwia pełniejsze przedstawienie i zrozumienie przemian wzorca dzietności w Polsce. Przy uwzględnieniu niedoskonałości pomiaru dzietności z wykorzystaniem miar przekrojowych wynikających z efektu odraczania, miary kohortowe pozwalają na wiarygodną ocenę zmian poziomu dzietności. Badanie wykorzystuje dane z rejestracji urodzeń na poziomie jednostkowym pozyskiwane cyklicznie z GUS i obejmujące lata 1985–2020. Dla okresu przed 1985 r. wykorzystano dane zebrane przez Holzera. W oparciu o ww. zbiory zrekonstruowano miary kohortowe, a także miary przekrojowe z wykorzystaniem dostosowania zaproponowanego przez Bongaartsa i Feeneya. Miary kohortowe dają podstawę do stwierdzenia, że dzietność jest wyższa niż ta wynikająca z zastosowania miar przekrojowych. Pierwszą kohortą, dla której zaobserwowano dzietność zrealizowaną poniżej wartości 2,1, jest kohorta urodzona w roku 1963. Kohorta ta miała dzietność zrealizowaną do wieku 35 lat na poziomie 1,97, podczas gdy analogiczny wskaźnik dla kohorty urodzonej w 1983 r. wyniósł już tylko 1,25. Różnica ta wynika głównie z procesu odraczania decyzji o posiadaniu potomstwa, a także „efektu ilości” czyli spadku odsetka kobiet rodzących dzieci drugie lub trzecie. Obserwacje te znajdują potwierdzenie również w analizie miar przekrojowych zmodyfikowanych z wykorzystaniem metody Bongaartsa-Feeneya. Z powodu uwzględnienia „efektu odraczania” klasyczne miary przekrojowe nie doszacowują dzietności przeciętnie o ok. 0,3 różnicy bezwzględnej. Przeprowadzone analizy dają podstawę do stwierdzenia, że miary przekrojowe nie doszacowują dzietności w Polsce głównie ze względu na proces odraczania decyzji o posiadaniu dzieci. W związku z tym miary kohortowe lub zmodyfikowane miary przekrojowe dostarczają bardziej wiarygodnego obrazu przemian dzietności w Polsce.

Słowa kluczowe: dzietność, kohorty, urodzenia, miary przekrojowe

## FERTILITY FROM COHORT AND PERIOD PERSPECTIVE. ANALYSIS WITH USE OF BIRTH REGISTRATION DATA

Analyses compare changes of fertility in Poland from a cohort and period perspective up to the year 2020. Those analytical approaches allow for a better understanding of fertility patterns in Poland. Taking into account the shortcomings in the application of period rates for the measurement of

fertility (due to the postponement effect), cohort measures allow for a more reliable measurement of fertility. The study uses data from the registration of births received from Statistics Poland. Databases cover the period of 1985 to 2020. For the period before 1985, the research applies data gathered by professor Jerzy Holzer. With the use of the aforementioned databases, cohort measures imply that the actual fertility is higher as compared to the period measures. The cohort of 1963 is the first one to reach sub-replacement fertility. This cohort had completed fertility up to the age of 35 at the level of 1.97. This same value for the 1983 cohort yielded 1.25. The observed difference results mainly from the postponement of fertility, but also from the quantum effect resulting in lower parity progression ratios. These observations are confirmed by analyses of period measures with the use of the Bongaarts-Feeney adjustment method. Taking into account the 'postponement effect', period measures seem to underestimate fertility in Poland by 0.3 of an absolute difference on average. Analyses seem to provide strong evidence that period measures underestimate fertility levels in Poland due to the 'postponement effect'. Therefore, cohort measures or at least adjusted period measures seem to provide a far more reliable picture of fertility changes in Poland.

Keywords: fertility, cohorts, births, period measures

**Anna Bieńkuńska**

Główny Urząd Statystyczny  
Statistics Poland

**Tomasz Piasecki**

Urząd Statystyczny w Łodzi  
Statistical Office in Łódź

**POMIAR I OCENA SPÓJNOŚCI SPOŁECZNEJ W POLSCE W KONTEKŚCIE STOSOWANYCH DEFINICJI ZJAWISKA**

Podobnie jak w przypadku innych pojęć stosowanych w naukach społecznych i powiązanych z polityką społeczną, również w przypadku spójności społecznej nie ma jednej ogólnie akceptowanej koncepcji i definicji, co przekłada się na stosowane metody pomiaru. Głównym celem referatu jest przedstawienie propozycji statystycznego pomiaru spójności społecznej w Polsce. Proponowany sposób pomiaru nawiązuje m.in. do koncepcji spójności społecznej przedstawionej przez Reginę Berger-Schmitt w 2000 r. Zgodnie z tą koncepcją osiągnięcie spójności społecznej wiąże się z dwoma celami: pierwszy zakłada redukcję nierówności i wykluczenia społecznego (w różnych sferach życia), drugi – wzmacnianie relacji, interakcji i więzi społecznych, czyli budowanie pozytywnego kapitału społecznego. Za kluczowe źródło danych umożliwiających kompleksowy pomiar spójności społecznej w Polsce (w kontekście przyjętej definicji) należy uznać realizowane cyklicznie przez GUS badanie spójności społecznej (BSS). W referacie przedstawione zostaną propozycje wskaźników mierzących różne wymiary spójności społecznej, jak również przykładowe wyniki analiz dokonanych na podstawie wyników tego badania. Pewien ekwiwalent dla danych z BSS, choć pokrywający obszar zainteresowania tylko w bardzo ograniczonym zakresie, mogą stanowić informacje pozyskiwane w corocznym Europejskim Badaniu Dochodów i Warunków Życia (EU-SILC). Na stan spójności społecznej w Polsce istotny wpływ mogą mieć zmiany będące skutkiem trwającej pandemii COVID-19. Pewne symptomy takich zmian zostaną pokazane na podstawie danych pochodzących z badania EU-SILC za 2020 r. Tymczasem ostatni kompleksowy obraz statystyczny tego zagadnienia pochodzący z BSS (edycja 2018) opisuje sytuację przedpandemiczną. Stanowi to mocny argument za realizacją kolejnej edycji BSS, co pozwoliłoby w pełniejszy sposób ocenić sytuację społeczną i uchwycić zmiany dotyczące jakości życia i spójności społecznej w okresie po pandemii.

Słowa kluczowe: spójność społeczna, nierówności, wykluczenie społeczne, kapitał społeczny, jakość życia



## MEASUREMENT AND ASSESSMENT OF SOCIAL COHESION IN POLAND IN THE CONTEXT OF THE APPLIED DEFINITIONS OF THIS PHENOMENON

As in the case of other terms used in the social sciences and social policy, social cohesion as well has no single commonly accepted concept and definition, which is reflected in the measurement methods used to assess it. The main purpose of the paper is to propose statistical measurement of social cohesion in Poland. The proposed approach refers, among other things, to the concept of social cohesion presented by Regina Berger-Schmitt in 2000. According to this concept, achieving social cohesion involves two goals: the first is to reduce inequality and social exclusion (in various spheres of life), the second is to strengthen relationships, interactions and social ties, i.e. to build positive social capital. The cyclical Social Cohesion Survey (BSS), carried out by Statistics Poland, should be considered as the key data source enabling a comprehensive measurement of social cohesion in Poland (in the context of the adopted definition). The paper will present proposals for indicators measuring the dimensions of social cohesion as well as selected examples of analysis based on the results of this survey. Information obtained from the annual European Survey of Income and Living Conditions (EU-SILC) may be some equivalent to data from the BSS, although covering the area of interest only to a very limited extent. The state of social cohesion in Poland may be significantly affected by changes resulting from the ongoing COVID-19 pandemic. Some symptoms of such changes will be shown on the basis of data from the EU-SILC for 2020. However, the last comprehensive statistical picture of this issue based on the Social Cohesion Survey (2018 edition) describes a pre-pandemic situation. This is a solid argument for carrying out the next edition of the Social Cohesion Survey, which would allow the assessment of the social situation in a more comprehensive way and to capture changes concerning the quality of life and social cohesion in the post-covid period in Poland.

Keywords: social cohesion, inequalities, social exclusion, social capital, quality of life

**Izabela Grabowska**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**Radosław Antczak**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**JAK MIERZYĆ WIELOWYMIAROWĄ JAKOŚĆ ŻYCIA OSÓB Z NIEPEŁNOSPRAWNOŚCIAMI NA POTRZEBY POLITYKI PUBLICZNEJ – PRZYPADEK POLSKI**

Koncepcje jakości życia (QoL) w przypadku osób z niepełnosprawnościami są zazwyczaj opracowywane specjalnie dla tej grupy. Konwencja ONZ o prawach osób z niepełnosprawnościami podkreśla znaczenie tej grupy dla społeczeństw, w których funkcjonują. Implikuje to wprost konieczność monitorowania jakości życia osób z niepełnosprawnościami także w porównaniu z osobami pełnosprawnymi. Celem artykułu jest skonstruowanie narzędzia, które może być wykorzystane do pomiaru wielowymiarowej jakości życia dla celów monitoringu Konwencji ONZ o prawach osób z niepełnosprawnościami. Zastosowano podejście porównawcze, które znalazło odzwierciedlenie w analizie różnic w jakości życia w różnych domenach życia między osobami z niepełnosprawnościami i bez niepełnosprawności. Do konceptualizacji pojęcia jakości życia zastosowano koncepcję możliwości Sena (ang. capability approach). Ponadto kierowano się wytycznymi pn. The Quality of Life Framework opracowanymi w ramach Europejskiego Systemu Statystycznego. Wartość zmiennej odzwierciedlającej QoL w poszczególnych wymiarach została skonstruowana przy użyciu modelu równań strukturalnych (MIMIC). Wszystkie analizy zostały wykonane na danych EU-SILC 2018 dla Polski. Zastosowane podejście umożliwia identyfikację wskaźników kluczowych dla domen QoL dla osób z niepełnosprawnościami i bez niepełnosprawności oraz determinant QoL w każdej domenie życia, a także ilościowe określenie różnicy między osobami z niepełnosprawnościami i bez niepełnosprawności w odniesieniu do QoL. Osoby pełnosprawne doświadczały wyższej jakości życia w porównaniu z osobami z niepełnosprawnościami. Wyniki potwierdzają także duże zróżnicowanie populacji osób z niepełnosprawnościami. Ponadto istnieją różnice we wskaźnikach odzwierciedlających QoL w każdej domenie życia między osobami z niepełnosprawnościami oraz bez niepełnosprawności, jak również można zidentyfikować różnice w oddziaływaniu czynników determinujących QoL w każdej domenie.

Słowa kluczowe: jakość życia, niepełnosprawność, domeny jakości życia, MIMIC

## HOW TO MEASURE MULTIDIMENSIONAL QUALITY OF LIFE OF PERSONS WITH DISABILITIES IN PUBLIC POLICIES – A CASE OF POLAND

The concepts of quality of life (QoL) of persons with disabilities are usually developed specifically for this group. The UN Convention on the Rights of Persons with Disabilities (UNCRPD) underlines the importance of persons with disabilities for societies they live in, which implies their quality of life should be monitored and compared with the quality of life of the population without disabilities. The aim of this article is to construct a tool with which it will be possible to measure the multidimensional quality of life for the purpose of monitoring of the UN Convention on the Rights of Persons with Disabilities. The comparative approach was used, involving tracing differences in the quality of life of persons with and without disabilities in different life domains. Sen's capability approach was applied to conceptualise the notion of quality of life. Moreover, we followed guidelines of The Quality of Life Framework developed within the European Statistical System. The QoL scores in dimensions were constructed using multiple indicators and a multiple causes model (MIMIC). All analyses were based on the 2018 EU-SILC data for Poland. This approach enables identifying indicators crucial for the QoL domains for the respective populations with and without disabilities, determinants of the QoL in each domain, and quantifying the differences between the QoLs of persons with and without disabilities. Persons without disability experienced higher quality of life than the population with disabilities. The results also confirm big internal diversity of population with disabilities. Moreover, there are differences in the indicators reflecting the QoL in each domain between persons with and without disabilities, as well as we can find some differences in the impacts of determinants of the QoL in each domain. The tool developed in this article enables comparisons between populations with and without disabilities in the dimensions covered by in the UNCRPD.

Keywords: quality of life, disability, life quality domains, MIMIC

**Wioletta Grzenda**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

**Jolanta Perek-Białas**

Uniwersytet Jagielloński w Krakowie; Szkoła Główna Handlowa w Warszawie

Jagiellonian University in Kraków; SGH Warsaw School of Economics

**AKTYWNOŚĆ ZAWODOWA OSÓB W WIEKU 55 LAT I WIĘCEJ  
W KONTEKŚCIE ZMIAN USTAWOWEGO WIEKU EMERYTALNEGO  
W POLSCE**

W porównaniu z innymi krajami europejskimi Polska charakteryzuje się niską aktywnością zawodową osób w starszym wieku. Jednocześnie w 2017 r. w Polsce weszła w życie ustawa o obniżeniu wieku emerytalnego do 60 lat dla kobiet oraz 65 lat dla mężczyzn, zmieniająca wcześniejszą ustawę z 2012 r. o zrównaniu wieku emerytalnego dla kobiet i mężczyzn. Celem badania była ocena zróżnicowania w aktywności zawodowej osób w wieku 55 lat i więcej w kontekście zmian wieku uprawniającego do przejścia na emeryturę. W analizach wykorzystano dane panelowe z badania Diagnostyki Społecznej z lat 2009–2015 dotyczące osób w wieku 55 lat i więcej (kobiet i mężczyzn). Ze względu na charakter zmiennej zależnej („Czy w ciągu ostatnich 7 dni osoba wykonywała pracę przynoszącą zarobek, dochód lub pomagała nieodpłatnie w rodzinnej działalności gospodarczej?”) zastosowano model logitowy w ujęciu wzdlużnym. Do oceny stopnia dopasowania modeli do empirycznych danych wykorzystano kryterium QIC (ang. quasi-likelihood under the independence model criterion).

Dzięki analizom udało się nie tylko ocenić wpływ zmiany ustawowego wieku emerytalnego na aktywność zawodową osób w wieku 55 lat i więcej w Polsce, ale także sprawdzono, jakie czynniki warunkowały ich decyzje zawodowe w badanym okresie. Wraz ze starzeniem się badanych osób szanse na wykonywanie pracy zmieniały się w większym stopniu w przypadku kobiet niż mężczyzn. Kobiety na ogół częściej niż mężczyźni wycofywały się z rynku pracy jeszcze przed osiągnięciem ustawowego wieku emerytalnego, ale po osiągnięciu tego wieku relacja ta uległa zmianie. Wnioskujemy, że większy wpływ na zakończenie lub kontynuowanie aktywności zawodowej ma wiek biologiczny niż ustawowy wiek emerytalny. Może to wynikać z pogarszania się stanu zdrowia wraz z wiekiem, które w większym stopniu determinowało wykonywanie pracy w przypadku mężczyzn niż kobiet (przy kontroli innych zmiennych).

Słowa kluczowe: rynek pracy, osoby w wieku 50 lat i więcej, wiek emerytalny, model logitowy, dane panelowe

## ECONOMIC ACTIVITY OF PERSONS AGED 55 AND OVER IN THE CONTEXT OF CHANGES OF ELIGIBLE RETIREMENT AGE IN POLAND

In comparison to other European countries Poland is characterized by low economic activity of older persons. Additionally in 2017 in Poland the law on decreasing the retirement age for women (60) and for men (65) entered into force, changing earlier law from 2012 which was aiming to increase and make equal retirement age for women and men. The aim of the analysis was to evaluate differences in economic activity of persons aged 55+ in the context of changes of eligible retirement age. In analysis the panel data from the Social Diagnosis Study of women and men in age of 55+ were used for years 2009-2015. Due to nominal character of dependent variable ("Does in last 7 days the person do any work which was paid or helped without payment in family business?") the logit model with longitudinal component was applied. In evaluation of models' goodness of fit Quasi-likelihood under the Independence model Criterion (QIC) was used. The analysis provided for the evaluation of the impact of retirement eligible age on economic activity of 55+ in Poland, which included checking which factors had an impact on decisions to work over the period of analysis. With the aging of the respondents, the chances of performing work changed to a greater extent for women than for men. Women more often than men withdraw from labour market, yet before the eligible retirement age but after reaching this age this relation was adverse. We conclude that the biological age has a greater impact on terminating or extending the working lives than the statutory retirement age. This may be due to the deterioration of health with age, which determined the employment to a greater extent in the case of men than of women (under control of other variables).

Keywords: labour market, persons 55+, retirement age, logit model, panel data

**Alina Jędrzejczak**

Uniwersytet Łódzki

University of Lodz

**Francesca Greselin**

Uniwersytet w Mediolanie-Bicocca

University of Milan-Bicocca

**ANALYSIS OF THE GENDER GAP IN POLAND AND ITALY**

For social and economic policies, it seems interesting to compare income inequalities across the countries and regions of the European Union. The focus of this article is income distribution in Poland and Italy, because these countries represent different economic backgrounds and thus differ in inequality patterns. The previous paper by the authors showed important differences between Poland and Italy in this respect; and especially so in the distribution of the discrepancies between men and women's incomes across regions. In this article, we would like to address the problem of the impact of selected covariates describing respondent or household characteristics of the samples on the existing gender gaps in both countries. Due to the fact that the concept of the 'gender gap' is related to the total distribution of income in a population, it is difficult to capture the full range of experiences by means of single metrics. In fact, quite the opposite is necessary — to uncover the factors contributing to gender discrepancies, one should adopt a variety of tools and move beyond the typical focus on the average or median income differences, focusing rather on how the entire distribution of women's earnings compares to that of men's earnings. To that end, income inequality curves were employed to detect changes at each income quantile. Afterwards, empirical distributions were compared through a relative approach. It was based on a transformation of the data from two distributions into a single distribution that contains all the information needed for scale-invariant comparison. The next step of the analysis was the search for socio-economic factors which could explain the observed differences, by means of a decomposition method. The methods applied in the study revealed important issues regarding the gender gaps in Poland and in Italy, showing both similarities and substantial differences between the countries.

Keywords: income inequality, gender gap, relative distribution method

**Stanisław Maciej Kot**

Politechnika Gdańska

Gdańsk University of Technology

## APPRAISING INEQUALITY: A POLICYMAKER ON THE OTHER SIDE OF THE VEIL OF IGNORANCE

The aim of the study is to develop the method of estimating Buhmann's et al. (1988, Rev. Inc. Wealth, 115-142) equivalence scale with the parameter  $\theta$ , assuming the Gini index,  $G$ , as the measure of inequality. The distributional theory hypothesis is a U-shaped relationship between  $\theta$  and  $G$ , with the minimum  $G_{\min}$ , at certain  $\theta_{\min}$ . We argue that the only convincing ex-post appraisal of the policy in question is that it enabled the reduction of inequality not less than  $G_{\min}$ . This appraisal determines the equivalence scale  $\theta_{\min}$  uniquely. We estimate the standards for 32 European countries in the years 2004-2018 using EU-SILC data. The data have corroborated the U-shape hypothesis in all 440 country-year cases. Economies of scale in the household size have turned out country- and year-specific. We propose recognising the pair  $(G_{\min}, \theta_{\min})$  as the standard of reporting income inequality in official statistics and assessing the poverty rate and other distributional issues using income distributions adjusted by the equivalence scale with  $\theta_{\min}$ .

Keywords: equivalence scale, inequality, U-test

**Ryszard Kowalski**

Uniwersytet Ekonomiczny w Krakowie

Cracow University of Economics

**Agnieszka Strzelecka**

Politechnika Koszalińska

Koszalin University of Technology

**Agnieszka Wałęga**

Uniwersytet Ekonomiczny w Krakowie

Cracow University of Economics

**Grzegorz Wałęga**

Uniwersytet Ekonomiczny w Krakowie

Cracow University of Economics

**ZADŁUŻENIE MŁODSZYCH I STARSZYCH GOSPODARSTW DOMOWYCH:  
CZY DZIECI ZWIĘKSZAJĄ OBCIĄŻENIE DŁUGIEM?**

Struktura osobowa gospodarstwa domowego jest jednym z czynników warunkujących poziom i strukturę ponoszonych wydatków. Szczególny wpływ na sytuację finansową gospodarstwa domowego mają dzieci, ponieważ ich obecność w gospodarstwie prowadzi do wzrostu kosztów utrzymania. Jednym z rozwiązań pozwalających na sfinansowanie rosnących wydatków gospodarstwa domowego stają się wówczas kredyty i pożyczki. Celem artykułu jest ocena skali nadmiernego zadłużenia wśród gospodarstw domowych z dziećmi. Przeprowadzono ją, wykorzystując obiektywne i subiektywne wskaźniki nadmiernego zadłużenia w podziale na gospodarstwa osób młodych (poniżej 35. roku życia) oraz starszych (35+). Zweryfikowano także zależność pomiędzy posiadaniem dzieci i innymi czynnikami społeczno-ekonomicznymi a ryzykiem wpadnięcia w nadmierne zadłużenie. Do oceny związku pomiędzy posiadaniem dzieci a nadmiernym zadłużeniem zastosowano model logitowy dla zmiennej binarnej. Wpływ posiadania dzieci i innych cech społeczno-ekonomicznych gospodarstw domowych na zadłużenie oceniono, wykorzystując regresję kwantylową. Obliczenia przeprowadzono z wykorzystaniem danych pochodzących z ogólnopolskich badań własnych przeprowadzonych na próbie zadłużonych gospodarstw domowych w 2018 r.

Gospodarstwa domowe z dziećmi charakteryzują się wyższym obciążeniem spłatami zadłużenia niż gospodarstwa bez dzieci. Nie wykazano jednakże większego ryzyka obiektywnego nadmiernego zadłużenia w tych gospodarstwach, co może wskazywać, że rodziny te rozważniej podchodzą do planowania swoich budżetów. Choć wśród młodych gospodarstw z dziećmi obserwowany jest



wyższy odsetek respondentów zgłaszających poczucie bycia nadmiernie zadłużonym w porównaniu z gospodarstwami bez dzieci, to wyłącznie w odniesieniu do grupy gospodarstw domowych w starszym wieku potwierdzono, że dzieci istotnie oddziałują na ryzyko psychologicznego nadmiernego zadłużenia. Co interesujące, obecność dzieci w gospodarstwach 35+ obniża ryzyko poczucia nadmiernego obciążenia długiem.

Słowa kluczowe: nadmierne zadłużenie, spłaty zadłużenia, dzieci, struktura gospodarstwa domowego

## **INDEBTEDNESS OF YOUNGER AND OLDER HOUSEHOLDS: DO CHILDREN REALLY INCREASE THE DEBT BURDEN?**

One of the factors determining the level and structure of expenses of a household is its age structure. The fact of having children impacts the financial situation of a household to a considerable extent, causing an increase in the cost of living. One of the solutions to increased household spending is taking out loans. The article aims to assess the scale of over-indebtedness of households with children. The study was conducted using objective and subjective indicators of over-indebtedness, broken down into households of younger people (age under 35) and older households (age 35+). The relationship between having children and other socio-economic factors and the risk of falling into over-indebtedness was also verified. A logit model for a binary variable was used to assess the relationship between having children and over-indebtedness. The impact of having children and other socio-economic characteristics of households on debt was assessed using quantile regression. The nationwide survey research of indebted households in Poland conducted in 2018 was used as a source of data.

Households with children have a higher debt burden than households without children. However, no greater risk of objective over-indebtedness was proven for these households, which might indicate that families with children are more prudent in planning their budgets. Although among 'young' households with children there was a higher percentage of respondents reporting a feeling of over-indebtedness than in households without children, only in the group of 'older' households it was confirmed that children significantly influenced the risk of psychological over-indebtedness. Interestingly, the presence of children in households aged 35+ reduces the risk of feeling over-indebted.

Keywords: over-indebtedness, debt repayments, children, household's structure

**Joanna Landmesser-Rusek**

Szkola Główna Gospodarstwa Wiejskiego w Warszawie

Warsaw University of Life Sciences

**Hanna Dudek**

Szkola Główna Gospodarstwa Wiejskiego w Warszawie

Warsaw University of Life Sciences

**DEKOMPOZYCJA ZMIAN W ZAKRESIE DEPRYWACJI MATERIALNEJ  
W POLSCE PRZED I W TRAKCIE PANDEMII COVID-19**

W ostatnich latach wzrasta zainteresowanie zagadnieniem deprivacji materialnej. Zagadnienie to odnosi się do braku możliwości zaspokojenia potrzeb uznanych w warunkach europejskich za podstawowe. Koncentrując się na braku możliwości zaspokojenia tych potrzeb ze względów finansowych, analiza deprivacji materialnej umożliwia bardziej bezpośredni pomiar standardu życia ludności niż wskaźniki dochodowe. Celem pracy jest identyfikacja zmian w deprivacji materialnej polskich gospodarstw domowych przed pandemią COVID-19 i w jej trakcie.

Niniejsza praca przedstawia wyniki ilościowej analizy deprivacji materialnej w Polsce. W badaniu wykorzystano dane pochodzące z Europejskiego Badania Warunków Życia Ludności (EU-SILC) za lata 2019 i 2020. Dokonano dekompozycji zaobserwowanych różnic w zakresie deprivacji materialnej w dwóch kolejnych latach – 2019 i 2020. Techniki dekompozycji stosowane są najczęściej w kontekście badania różnic w wynagrodzeniach kobiet i mężczyzn, wykorzystując przy tym modele regresji liniowej. Jednak metody te rozszerzono także na modele nieliniowe, w tym probitowe i logitowe. W pracy, podobnie jak w kilku poprzedzających ją publikacjach innych autorów, wykorzystano ideę dekompozycji do modeli z binarnymi zmiennymi zależnymi. Metoda ta dotychczas nie była stosowana w analizie deprivacji materialnej.

Oszacowane parametry modeli logitowych umożliwiły identyfikację czynników mających wpływ na doświadczanie deprivacji materialnej. Zastosowanie techniki dekompozycji do modeli logitowych pozwoliło na wyodrębnienie czynników wyjaśniających obserwowane różnice w deprivacji materialnej, które zaszły w okresie 2019–2020. Analiza ujawniła, że zarówno przed kryzysem COVID-19, jak i w jego trakcie znaczna część Polaków wykazywała symptomy deprivacji materialnej. Przeprowadzona dekompozycja zmian zachodzących w czasie dostarczyła pogłębionych informacji o badanym zjawisku.

Słowa kluczowe: deprivacja materialna, model logitowy, dekompozycja różnic

## DECOMPOSITION OF CHANGES IN MATERIAL DEPRIVATION IN POLAND BEFORE AND DURING THE COVID-19 PANDEMIC

In recent years, there has been growing interest in the issue of material deprivation. It refers to the inability to satisfy needs considered as basic by the European standards. By focusing on the financial inability to satisfy these needs, the analysis of material deprivation enables a more direct measurement of the standard of living of the population than income indicators. The aim of this paper is to identify changes in material deprivation of Polish households before and during the COVID-19 pandemic.

The paper presents the results of a quantitative analysis of material deprivation in Poland. The study uses data from the EU Statistics on Income and Living Conditions (EU-SILC) survey from 2019 and 2020, and decomposes the observed differences in material deprivation in two consecutive years: 2019 and 2020. Decomposition techniques are most commonly used in the context of studying gender pay gaps using linear regression models. However, these methods have been extended to non-linear models, including probit and logit models. The paper, like several others before it, applies the idea of decomposition to models with binary dependent variables. This method has not been used in material deprivation analysis before.

The estimated parameters of the logit models enabled the identification of factors influencing the experience of material deprivation. Applying the decomposition technique to the logit models enabled us to extract factors explaining the observed differences in material deprivation that occurred between 2019 and 2020. The analysis demonstrated that both before and during the COVID-19 crisis, a significant proportion of Poles exhibited symptoms of material deprivation. The decomposition of changes over time provided in-depth information about the phenomenon under study.

Keywords: material deprivation, logit model, decomposition of differences

**Karolina Goś-Wójcicka**

Główny Urząd Statystyczny  
Statistics Poland

**Edyta Makowska-Belta**

Główny Urząd Statystyczny  
Statistics Poland

**RACHUNEK SATELITARNY GOSPODARKI SPOŁECZNEJ JAKO KLUCZOWE NARZĘDZIE MONITOROWANIA ZNACZENIA TRZECIEGO SEKTORA W POLSKIEJ GOSPODARCE NARODOWEJ**

Celem wystąpienia jest zaprezentowanie rezultatów pierwszego rachunku satelitarnego gospodarki społecznej dla Polski. Stanowi on zwieńczenie ponad 10-letniego rozwoju badań trzeciego sektora w ramach statystyki publicznej. Zrealizowany został przy międzynarodowej współpracy pod auspicjami Eurostatu (grant „Rachunki satelitarne gospodarki społecznej” finansowany ze środków Komisji Europejskiej). Podstawę metodologiczną opracowania rachunku satelitarnego gospodarki społecznej stanowiły podręcznik ONZ „Rachunek satelitarny instytucji non-profit i pokrewnych oraz wolontariatu” oraz zasady określone w ESA 2010. Podobnie jak inne rachunki satelitarne łączy informacje ze wszystkich dostępnych źródeł, tj. z rachunków narodowych, danych administracyjnych oraz badań ankietowych GUS, w formę jednej kalkulacji zawierającej pełną sekwencję tablic rachunkowych.

Rachunek satelitarny gospodarki społecznej wyznaczył rozmiar i rzeczywistą monetarną wartość, jaką wnosi ten sektor w gospodarkę narodową w 2018 r. Wykracza poza klasyczny system rachunków narodowych, zarówno w wymiarze podmiotowym, jak i przedmiotowym. Obejmuje działalność organizacji non profit, spółdzielni i podmiotów pokrewnych, a także wolontariat instytucjonalny i świadczony bezpośrednio. Przedstawia produkcję globalną i inne kluczowe wskaźniki rachunkowe na tle gospodarki narodowej, uzupełnione analizami dotyczącymi zasobów pracy, kosztów pracy oraz źródeł przychodów. Wyodrębnienie działalności sektora gospodarki społecznej pozwoliło wygenerować szerszy i pełniejszy obraz tego sektora, zgodnie z ideą wyjścia poza PKB w opisie zjawisk społeczno-gospodarczych. Rachunek satelitarny gospodarki społecznej pozwala na monitorowanie udziału gospodarki społecznej w usługach publicznych, rozwoju społeczno-gospodarczym, a także na planowanie zrównoważonej działalności zarówno w sferze ekonomicznej, społecznej, jak i środowiskowej.

Słowa kluczowe: rachunek satelitarny, gospodarka społeczna, ekonomia społeczna, rachunki narodowe, polityka społeczna

## SOCIAL ECONOMY SATELLITE ACCOUNT AS A KEY TOOL FOR MONITORING THE IMPORTANCE OF THE THIRD SECTOR IN THE POLISH NATIONAL ECONOMY

The aim of this presentation is to show the results of the first social economy satellite account for Poland. It is the conclusion of an over-10-year-long research on the third sector under the auspices of official statistics, carried out with in the framework of international collaboration within Eurostat ('Social Economy Satellite Accounts' grant financed by the European Commission). The social economy satellite account methodology was based on an UN handbook entitled 'The Satellite Account on Non-profit and Related Institutions and Volunteer Work' and rules specified in the ESA 2010. Like other satellite accounts, it combines information from all accessible sources, i.e. national accounts, administrative data and Statistics Poland's social research, in a single calculation containing the entire sequence of accounting tables.

The social economy satellite account determined the 'size' of the contribution that this sector made to the national economy in 2018 and its direct monetary value. Social economy satellite accounts go beyond the basic system of national accounts, both in terms of the 'subject' and the 'object' dimension. This kind of an account features the activity of non-profit organizations, cooperatives and related entities, as well as institutional and direct volunteer work. It shows the global output and other key accounting indicators in relation to the national economy, complemented with analyses of labour resources, labour costs and sources of revenue. Separating the activity of the social economy sector from the whole economy made it possible to generate a broader and fuller picture of this sector, according to the idea of going beyond GDP in the description of socio-economic phenomena. The social economy satellite account allows monitoring the contribution of the social economy to public services and the socio-economic development, and planning sustainable activities in the economic, social and environmental spheres.

Keywords: satellite account, third sector, social economy, national accounts, social policy

**Barbara Pawełek**

Uniwersytet Ekonomiczny w Krakowie

Cracow University of Economics

**Tomasz Sekuła**

Urząd Statystyczny w Krakowie

Cracow University of Economics

**PORÓWNANIE PRZYDATNOŚCI WYBRANYCH METOD SZACOWANIA BRAKUJĄCYCH DANYCH W BADANIU FUNDACJI, STOWARZYSZEŃ I PODOBNYCH ORGANIZACJI SPOŁECZNYCH**

Celem referatu jest przedstawienie wyników badań nad przydatnością wybranych metod szacowania brakujących danych do uzupełniania brakujących informacji w sprawozdaniu z działalności fundacji, stowarzyszeń i podobnych organizacji społecznych (sprawozdanie SOF-1). Celem badania realizowanego na formularzu sprawozdawczym SOF-1 jest pomiar zasobów ekonomicznych i społecznych oraz opis działalności organizacji non profit. Zakres podmiotowy badania obejmuje: fundacje, stowarzyszenia, podmioty zrównane w prawach ze stowarzyszeniami oraz społeczne podmioty wyznaniowe. Badanie to jest realizowane w cyklu co 2 lata przez pracowników Urzędu Statystycznego w Krakowie. Do szacowania brakujących danych dotyczących liczby osób, które badane jednostki objęły swoim bezpośrednim działaniem w 2018 r., wykorzystano m.in. średnią arytmetyczną, medianę, metodę k najbliższych sąsiadów, drzewo klasyfikacyjne, metodę lasów losowych oraz imputację wielokrotną. W przeprowadzonych badaniach symulacyjnych rozważono różne mechanizmy generowania braków danych.

Wstępne wyniki badań potwierdzają hipotezę, że wybór metody służącej do uzupełniania brakujących informacji w sprawozdaniu z działalności fundacji, stowarzyszeń i podobnych organizacji społecznych należy poprzedzić badaniem zależności między występowaniem braków odpowiedzi w sprawozdaniu SOF-1 a charakterystykami jednostek uczestniczących w badaniu. Wnioski z takiego badania mogą być podstawą określenia mechanizmów generujących braki danych w poszczególnych pytaniach znajdujących się w formularzu SOF-1 i – w konsekwencji – wyboru metody szacowania brakujących danych. Zwiększenie kompletności odpowiedzi w sprawozdaniu SOF-1 przez wstawienie w miejsce brakujących informacji oszacowań uzyskanych w wyniku zastosowania dobrze dobranej metody szacowania braków danych wpływa nie tylko na zmniejszenie kosztów badania, lecz także na zwiększenie precyzji uogólnień.

Słowa kluczowe: organizacje non profit, brakujące dane, mechanizmy generujące braki danych, metody szacowania brakujących danych

## **A COMPARISON OF THE UTILITY OF SELECTED METHODS FOR ESTIMATING MISSING DATA IN A SURVEY OF FOUNDATIONS, ASSOCIATIONS, AND SIMILAR COMMUNITY ORGANIZATIONS**

The aim of the paper is to present the results of research on the usefulness of the selected methods of estimating missing data from the reports on the activity of foundations, associations and similar social organisations (SOF-1 report). The aim of the survey (delivered on the SOF-1 reporting form) is to measure the economic and social resources of non-profit organisations and to describe their activities. The survey examines foundations, associations, entities equal in rights to associations and social religious entities. It is conducted every 2 years by the Statistical Office in Krakow. In order to estimate the missing data on the number of people affected by or included in the activity of the surveyed units in 2018 several methods were used, including: the arithmetic mean, the median, the k nearest neighbours, the classification tree, the random forest method and the multiple imputation. Different mechanisms for generating missing data were considered in the simulation studies.

The preliminary results of the research support the hypothesis that the choice of the method to be used to find the information missing from reports on the activity of foundations, associations and similar social organisations should be preceded by a study of the potential relationships between the occurrence of missing responses and the characteristics of the units participating in the survey. The conclusions of this study may be the basis for the identification of mechanisms generating the originally-missing data for particular questions of the SOF-1 form and, consequently, for the choice of the method for the estimation of the missing data. Making the the SOF-1 complete by inserting estimates of the missing entries obtained by means of a fitting method for estimating missing data not only reduces the cost of the survey, but also increases the precision of generalisations.

Keywords: non-profit organisations, missing data, mechanisms generating the missing data, methods for estimating missing data

**Maria Penpeska**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**Anna Jasiówka**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**Michał Urbas**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

## **OPIEKA ZDROWOTNA NAD OSOBAMI STARSZYM W UJĘCIU BADAŃ STATYSTYCZNYCH**

Prognozy demograficzne wskazują, że postępujący proces starzenia się ludności doprowadzi do zasadniczej zmiany proporcji między osobami młodszymi i starszymi, stanowiąc jedno z najważniejszych wyzwań społecznych. Dla systemu ochrony zdrowia istotnego znaczenia nabiera zwiększenie obciążenia kosztami opieki nad najstarszą grupą ludności. Celem referatu jest przedstawienie statystycznego obrazu stopnia korzystania z usług opieki zdrowotnej przez osoby starsze na przestrzeni ostatnich lat. Dane obejmują również rok 2020 i mogą stanowić dobry punkt odniesienia dla badań wpływu pandemii COVID-19 na sytuację zdrowotną tej populacji. Informacje na temat świadczeń medycznych udzielonych grupie seniorów przez placówki lecznicze dostarczają prowadzone w cyklu rocznym badania statystyczne statystyki publicznej w zakresie ambulatoryjnej i stacjonarnej opieki zdrowotnej, pomocy doraźnej i ratownictwa medycznego oraz lecznictwa uzdrowiskowego. Prezentowane dane pochodzą z badań Głównego Urzędu Statystycznego, sprawozdawczości resortowej Ministerstwa Zdrowia, Ministerstwa Spraw Wewnętrznych i Administracji oraz Narodowego Instytutu Zdrowia Publicznego PZH – PIB. Uzupełnieniem są wyniki badania ankietowego „Ochrona zdrowia w gospodarstwach domowych w 2020 r.” na temat subiektywnego postrzegania stanu zdrowia przez osoby starsze. Jedną z konsekwencji procesu starzenia się społeczeństwa jest wzrost liczby świadczeń medycznych udzielanych osobom starszym, które znacznie częściej potrzebują różnorodnych usług opieki zdrowotnej. Tendencja ta dotyczy wszystkich form opieki zdrowotnej i – biorąc pod uwagę prognozowane trendy demograficzne – będzie się pogłębiać. Priorytetem staje się poprawa dostępności opieki zdrowotnej dla osób w podeszłym wieku, w tym w szczególności rozwój infrastruktury opieki geriatrycznej czy długoterminowej, oraz zapewnienie dostępu do odpowiednio przeszkolonego personelu medycznego w zakresie opieki nad seniorami.

Słowa kluczowe: starzenie się społeczeństwa, opieka zdrowotna, osoby starsze



## HEALTH CARE FOR ELDERLY PEOPLE FROM THE PERSPECTIVE OF STATISTICAL RESEARCH

Demographic forecasts indicate that progressing population ageing will lead to a fundamental change in the proportion of younger people to older ones, and this will be one of the most important social challenges. Therefore dealing with increasing financial burden resulting from the provision of healthcare services to a growing number of seniors is becoming a matter of vital importance for the healthcare system in Poland. The aim of the paper is to present the statistical picture of the healthcare services usage by the elderly people in recent years. The data also covers the year 2020 and may constitute a good reference point for research on the potential impact of the COVID-19 pandemic on the health situation of this sub-population. Information on the healthcare services provided to seniors is gathered in the framework of annual statistical surveys conducted by official statistical services in the field of outpatient and inpatient healthcare, emergency care and medical rescue as well as spa treatment. The presented data comes from surveys by Statistics Poland, reports of the Ministry of Health, the Ministry of Interior and Administration and the National Institute of Public Health NIH – the National Research Institute. They are complemented by the results of the 'Healthcare in households in 2020' survey on the elderly people's health self-assessment. One of the consequences of the population ageing is the increase in the number of medical services provided to the elderly, who much more often than younger people require a variety of healthcare services. This tendency applies to all forms of healthcare and, taking into account the demographic prognoses, it will only deepen. The priority is to improve the availability of healthcare services for the elderly people, including, in particular, the development of geriatric and long-term care infrastructure, and ensuring access to medical personnel trained in the field of senior care.

Keywords: population ageing, healthcare, elderly people

**Joanna Sanetra-Szeliga**

Urząd Statystyczny w Krakowie; Uniwersytet Ekonomiczny w Krakowie  
Statistical Office in Kraków; Cracow University of Economics

**WPŁYW PANDEMII COVID-19 NA SYTUACJĘ FINANSOWĄ INSTYTUCJI KULTURY**

Wybuch pandemii COVID-19 spowodował zmiany i czasowe ograniczenia w działalności instytucji kultury. Miał wpływ nie tylko na merytoryczną działalność, lecz także na sytuację finansową tych podmiotów. Celem prezentacji jest pokazanie zmian, jakie zaszły w sytuacji gospodarczej i finansowej instytucji kultury w okresie marzec 2020 – wrzesień 2021. Bazując na analizie wyników kwartalnych i rocznych sprawozdań o finansach instytucji kultury przekazywanych do GUS (formularze F-01/dk i F-02/dk) oraz raportów o sytuacji w kulturze przygotowanych na zlecenie wybranych podmiotów (m.in. Urząd Miasta Krakowa, Urząd Miejski w Gdańsku), przedstawiony zostanie obraz publicznych podmiotów z sektora kultury ze szczególnym uwzględnieniem przychodów, kosztów oraz poniesionych nakładów. Dane statystyczne zostaną opracowane według rodzajów działalności (PKD), form prawnych oraz województw.

Jednym z podstawowych problemów, z jakim muszą się mierzyć instytucje kultury, jest wynikające z restrykcji pandemicznych ograniczenie możliwości prowadzenia działalności gospodarczej (takiej jak sprzedaż biletów, wynajem powierzchni), która stanowi średnio ok. 20% ich przychodów. Drugim problemem mogą być zmiany w dotacjach do działalności podstawowej. O ile w 2020 r. w stosunku 2019 r. niższe dotacje otrzymały gminne i wojewódzkie samorządowe instytucje kultury (średnio odpowiednio o 2,25 i 2,95 p.proc.), o tyle państwowe instytucje kultury miały się dobrze, odnotowując ponad 14-procentowy wzrost dotacji. Samorządowe instytucje natomiast otrzymały relatywnie wysokie wsparcie z Funduszu Promocji Kultury, w ramach którego zorganizowano wsparcie ograniczające negatywne skutki pandemii. Sytuacja instytucji kultury, których budżety w przeważającej części zależą od wsparcia organizatora, pozostaje w bezpośredniej zależności możliwości finansowych samorządów i administracji centralnej. Ich przyszłość będzie więc zależeć od kondycji finansów publicznych, na które wpływa kryzys wywołany pandemią.

Słowa kluczowe: pandemia COVID-19, instytucje kultury, finanse kultury, działalność gospodarcza

## THE IMPACT OF COVID-19 PANDEMIC ON THE FINANCIAL SITUATION OF CULTURAL INSTITUTIONS

The outbreak of the COVID-19 pandemic caused changes and temporary limitations in the activity of cultural institutions. Not only their statutory activity was affected, but their financial situation as well. The aim of the presentation is to show changes that took place in the economic and financial situation of cultural institutions between March 2020 and September 2021. Based on the analysis of the results of quarterly and annual reports on the finances of cultural institutions, submitted to Statistics Poland (forms F-01 / dk and F-02 / dk), and reports on the situation in culture prepared at the request of selected entities (e.g. cities of Kraków, Gdańsk), the picture of public cultural entities will be presented, with a particular emphasis on revenues, costs and expenditures. Statistical data will be compiled according to types of activity (Classification of Business Activities), legal forms and voivodships.

One of the problems faced by cultural institutions is limiting their commercial activity (e.g. selling tickets, renting space) resulting from the pandemic restrictions, which on average accounts for about 20% of their revenues. The second problem may be changes in subsidies to the institutions' core activities. While in 2020, compared to 2019, lower subsidies were received by communal and voivodeship local government cultural institutions (on average lower by 2.25 and 2.95 percentage points, respectively), state cultural institutions recorded an over 14%-increase in their subsidies. To make the picture fuller, though, it has to be said that local government institutions received relatively high support from the Culture Promotion Fund, aimed at limiting the negative effects of the pandemic. The situation of cultural institutions, whose budgets mostly depend on the support of the organiser, is directly dependent on the financial possibilities of local governments and the central administration. Their future will, therefore, depend on the state of the public finances, affected by the crisis caused by the pandemic.

Keywords: COVID-19 pandemic, cultural institutions, finances of culture, commercial activity

**Tomasz Sekuła**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**Krzysztof Borysiak**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**Karolina Goś-Wójcicka**

Główny Urząd Statystyczny  
Statistics Poland

**Izabela Jajkiewicz**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**Mateusz Kolmaga**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**Agnieszka Pacut**

Uniwersytet Ekonomiczny w Krakowie  
Cracow University of Economics

**Urszula Stasiak-Jaśkiewicz**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

## PODMIOTY GOSPODARKI SPOŁECZNEJ NA RYNKU PRACY W POLSCE

Celem wystąpienia jest przedstawienie charakterystyki zatrudnienia w podmiotach gospodarki społecznej i jego znaczenia dla rynku pracy w Polsce. Szczególną uwagę zwrócono na reintegrację zawodową i społeczną osób zagrożonych wykluczeniem społecznym jako jedną z podstawowych funkcji realizowanych przez podmioty gospodarki społecznej. Podstawę referatu będą stanowić wyniki ogólnopolskich badań statystycznych dotyczących: fundacji, stowarzyszeń i podobnych organizacji społecznych, samorządu gospodarczego, spółdzielni i jednostek reintegracji społeczno-zawodowej. Prowadzone regularnie od ponad dekady badania w dziedzinie gospodarki społecznej

pozwołyły na zobrazowanie udziału podmiotów gospodarki społecznej w rynku pracy w Polsce, w tym tworzenia miejsc pracy oraz ich znaczenia w procesie reintegracji społeczno-zawodowej osób zagrożonych wykluczeniem społecznym. Zgromadzone dane umożliwiły prezentację wskaźników dotyczących zatrudnienia w podmiotach gospodarki społecznej z uwzględnieniem m.in. zmiennych demograficznych, rodzajów pracy płatnej i kategorii osób zagrożonych wykluczeniem społecznym. Dostępne dane pozwoliły na przedstawienie wewnętrznego zróżnicowania podmiotów gospodarki społecznej w Polsce oraz szerokiego zakresu ich aktywności. Badania podmiotów gospodarki społecznej stanowią istotne źródło informacji potrzebnych do monitorowania udziału tej grupy podmiotów w rynku pracy w Polsce, rozwoju społeczno-gospodarczego, a także pomiaru efektów reintegracji społeczno-zawodowej osób zagrożonych wykluczeniem społecznym.

Słowa kluczowe: gospodarka społeczna, ekonomia społeczna, reintegracja społeczno-zawodowa

## **PARTICIPATION OF THE SOCIAL ECONOMY ORGANIZATIONS ON THE LABOUR MARKET IN POLAND**

The presentation aims to describe the characteristics and specificity of the employment in social economy organisations as well as their impact on the labour market in Poland. Emphasis is put on social and professional reintegration services dedicated to persons at risk of social exclusion, which is one of the basic roles performed by the organizations of the social economy. The presentation is based on the results of the Polish nationwide surveys on foundations, associations and similar organisations, business associations, cooperatives and organisations providing social and professional reintegration services. Repeated surveys carried out in the field of social economy for over a decade makes it possible to measure the participation of the social economy organizations in the labour market in Poland, including job creation and their importance for the process of social and professional reintegration. The data collected in surveys makes it possible to present the crucial indicators of employment in social economy organizations, including demographic variables, type of paid work, and categories of persons at risk of social exclusion. The available data allowed some insight into the variability of the social economy organizations in Poland and their wide range of activities. The surveys on the social economy organisations constitute an important source of information necessary to monitor not only the participation of these organisations in the Polish labour market, but also the socio-economic development and the effects of the reintegration services offered by them on persons at risk of social exclusion.

Keywords: social economy, social and professional reintegration services

**Agnieszka Sompolska-Rzechuła**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**Agnieszka Kurdyś-Kujawska**

Politechnika Koszalińska

Koszalin University of Technology

**POKOLENIE GNIAZDOWNIKÓW W KRAJACH UNII EUROPEJSKIEJ**

Celem badania jest ocena i pomiar zjawiska gniazdownictwa w krajach Unii Europejskiej. Przedział czasowy studium obejmuje lata 2011–2019, wykorzystano dane z bazy Eurostatu. W badaniu dokonano dynamicznej oceny zmian odsetka gniazdowników z uwzględnieniem ich płci, wieku oraz formy aktywności zawodowej i statusu zatrudnienia. Wykorzystano liniowe porządkowanie krajów w ujęciu statycznym i dynamicznym. W 2019 r. odsetek gniazdowników w wieku 25–34 lat w badanych krajach wyniósł 30,5% i z roku na rok ulegał zwiększeniu. Odsetek gniazdowników płci żeńskiej był niższy w porównaniu z odsetkiem gniazdowników wśród mężczyzn. Poziom gniazdowników był najniższy w krajach skandynawskich, wyższy poziom badanego zjawiska zaobserwowano w krajach Europy Zachodniej i Południowej. Kraje UE są silnie zróżnicowane pod względem wielkości zjawiska gniazdownictwa i wieku opuszczenia domu rodzinnego przez młodych dorosłych. Głównym czynnikiem wpływającym na decyzję o opuszczeniu rodzinnego gniazda jest aspekt ekonomiczny odnoszący się zarówno do sytuacji młodego człowieka, jak i jego rodziny. Pozytywnym zjawiskiem jest podejmowanie aktywności zawodowej przez gniazdowników. Coraz więcej młodych dorosłych mieszka z rodzicami, a odsetek gniazdowników różni się w poszczególnych krajach Europy. Młodzi dorośli mieszkający z rodzicami to nie tylko bezrobotni czy osoby spoza rynku pracy, dużą grupę gniazdowników stanowią osoby aktywne zawodowo. Problem gniazdownictwa w poszczególnych krajach europejskich przebiegał z różnym natężeniem. Daje się zauważyć nasilenie tego zjawiska szczególnie wśród osób bez pracy oraz nieposiadających stałego zatrudnienia. Ponadto można zauważyć występowanie dużych dysproporcji w poziomie gniazdownictwa ze względu na aktywność zawodową i status zatrudnienia, co może wynikać z ogólnej koniunktury gospodarczej i dostępu do rynku pracy w poszczególnych krajach Europy.

Słowa kluczowe: gniazdownicy, podejście dynamiczne, porządkowanie krajów, aktywność zawodowa

## GENERATION OF YOUNG ADULTS LIVING WITH THEIR PARENTS IN EUROPEAN UNION

The aim of the study is to assess and measure the phenomenon of generation of young adults living with their parents in European Union countries. The time frame of the study covers the years 2011-2019, data from the Eurostat database were used. The study dynamically assessed the changes in the percentage of young adults living with their parents, taking into account their gender, age, and the form of activity and employment status. The linear ordering of countries in the static and dynamic approach was used. In 2019, the percentage of nests aged 25-34 in the studied countries amounted to 30.5% and was increasing year by year. The percentage of female was lower compared to the percentage of male young adults. The number of young adults living with their parents in the EU countries was the lowest in the Scandinavian countries, a higher level of the studied phenomenon was observed in the countries of western and southern Europe. The EU countries vary greatly in terms of the size of the nesting phenomenon and the age at which young adults leave their homes. The main factor influencing the decision to leave the family nest is the economic aspect, relating both to the situation of the young person and the family. A positive phenomenon is taking up professional activity by young adults. More and more young adults live with their parents. Young adults living with their parents are not only unemployed or people from outside the labor market, professionally active people constitute a large group. The problem of nesting in individual European countries was varied. It is noticeable that this phenomenon is intensifying, especially among the unemployed and those without permanent employment. In addition, large disproportions in the level of nesting can be noticed due to professional activity and employment status, which may result from the general economic situation and access to the labor market in individual European countries.

Keywords: young adults living with parents, dynamic approach, linear ordering of countries, activity and employment status

**Marta Styrac**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

**PRZESTRZEGANIE ZALECEŃ PRZECIWEPIDEMICZNYCH ZWIĄZANYCH Z COVID-19 PRZEZ OSOBY STARSZE W POLSCE**

W czasie epidemii COVID-19 zalecono wiele środków zapobiegających zakażeniu, które miały służyć ochronie zdrowia indywidualnego i ograniczyć rozprzestrzenianie się choroby w populacji. W niniejszym referacie koncentruję się następujących środkach bezpieczeństwa: utrzymywanie dystansu społecznego, noszenie maseczki, mycie rąk, używanie środków dezynfekujących, zasłanianie ust w czasie kaszlu i kichania oraz czynnikach, które wpływają na ich stosowanie wśród osób w wieku 50 lat i więcej. Wykorzystując dane z Badania Zdrowia, Starzenia się Populacji i Procesów Emerytalnych „SHARE 50+ w Europie” (Survey of Health, Ageing, and Retirement), szacuję modele regresji dla zachowań prozdrowotnych względem cech demograficznych i społeczno-ekonomicznych oraz stanu zdrowia. Wstępne wyniki wskazują, że osoby w bardziej zaawansowanym wieku oraz samotni mężczyźni są mniej skłonni do przestrzegania zaleceń przeciwepidemicznych.

Słowa kluczowe: zachowania prozdrowotne, zachowania zapobiegawcze, epidemia COVID-19, SHARE

**COMPLIANCE WITH COVID-19 PREVENTIVE BEHAVIOURS AMONG OLDER MEN AND WOMAN IN POLAND**

During the COVID-19 pandemic, several preventive behaviours were recommended to individuals to protect their health and reduce the spread of the disease. In this study I analyse the following preventive measures: (1) keeping distance (2) wearing face masks (3) washing hands, (4) using hand sanitisers and disinfection fluids, and (5) covering mouth and nose while coughing and sneezing, and the factors affecting their implementation among men and women aged 50+ in Poland. Using the data from the ‘Survey of Health, Ageing, and Retirement’, I estimate regression models for the implementation of the preventive behaviours on the set of demographic, health and socio-economic characteristics. Preliminary results indicate that older people and elderly single men are less inclined to comply with the safety measures.

Keywords: preventive behaviours, COVID-19, Health Belief Model, SHARE, population 50+



**Dominik Śliwicki**

Uniwersytet Mikołaja Kopernika w Toruniu

Nicolaus Copernicus University in Toruń

**Magdalena Osińska**

Uniwersytet Mikołaja Kopernika w Toruniu

Nicolaus Copernicus University in Toruń

**Małgorzata Szczepaniak**

Uniwersytet Mikołaja Kopernika w Toruniu

Nicolaus Copernicus University in Toruń

**Agnieszka Szulc-Obłóza**

Uniwersytet Mikołaja Kopernika w Toruniu

Nicolaus Copernicus University in Toruń

**SATYSFAKCJA Z PRACY I SATYSFAKCJA Z WYNAGRODZENIA OSÓB  
MŁODYCH W POLSCE**

Celem badania jest opracowanie narzędzia umożliwiającego zbadanie zróżnicowania wskaźnika satysfakcji z pracy oraz satysfakcji z wynagrodzenia osób młodych w Polsce ze względu na wykształcenie i płeć. Celem dodatkowym jest identyfikacja kluczowych czynników wpływających na poziom satysfakcji z pracy. Narzędziem pozwalającym na realizację celu badania jest kwestionariusz ankiety, przeprowadzonej na osobach w wieku 18–29 lat w grudniu 2020 r. Na podstawie uzyskanych wyników oszacowano wielomianowy model logitowy dla kategorii uporządkowanych. Oszacowanie modeli dla satysfakcji z pracy oraz satysfakcji z wynagrodzenia pozwala na wyodrębnienie istotnych czynników kreacji satysfakcji ze wskazaniem kierunku oraz siły ich wpływu. Ponadto oszacowane zostały skorygowane szacunki satysfakcji z pracy oraz satysfakcji z wynagrodzenia poprzez wykorzystanie (oszacowanych na podstawie modelu) prawdopodobieństw uzyskania określonych odpowiedzi w pytaniach o satysfakcję z pracy i satysfakcję z wynagrodzenia. Uzyskane wyniki zostały wykorzystane do analiz symulacyjnych satysfakcji z pracy i satysfakcji z wynagrodzenia. Symulacje opierają się na scenariuszach wzrostów i spadków satysfakcji z czynników szczegółowych. Z reguły spadki satysfakcji z czynników szczegółowych silniej wpływają na spadki satysfakcji z pracy niż ich wzrosty na wzrosty satysfakcji z pracy. Prowadzone symulacje pozwolą na walidację modelu, który będzie mógł być wykorzystany do dalszych badań, w tym oceny zmian poziomu satysfakcji z pracy oraz satysfakcji z wynagrodzenia.

Słowa kluczowe: młodzi, badanie ankietowe, satysfakcja z pracy, wielomianowy model logitowy

## **JOB SATISFACTION AND SATISFACTION WITH REMUNERATION OF YOUNG PEOPLE IN POLAND**

The aim of the study is to develop a tool that would check how the work satisfaction indicator and the indicator of satisfaction with remuneration received by young people in Poland vary according to the respondents' level of education and sex. An additional goal is to identify key factors influencing the level of job satisfaction. The tool enabling the achievement of the research objective is a questionnaire, conducted among people aged 18–29 in December 2020. Based on the obtained results, a polynomial logit model for the ordered categories was estimated. The estimation of models for job satisfaction and the satisfaction with remuneration led to the identification of factors critical to the generation of satisfaction, and additionally indicated the direction and strength of their impact. In addition, the 'adjusted estimates' of job satisfaction and salary satisfaction were estimated on the basis of the probabilities (estimated on the basis of the model) of obtaining specific answers in the questionnaire about job and salary satisfaction. The obtained results were used for simulation analyses of job satisfaction and satisfaction with salary. The simulations were based on scenarios of increases and decreases in satisfaction with specific factors. Usually, declines in satisfaction with specific factors have a stronger impact on declines in job satisfaction than their increases on the increases in job satisfaction. The conducted simulations will allow the validation of a model which can be used for further research, including the assessment of changes in the level of job satisfaction and the satisfaction with remuneration.

Keywords: young people, survey research, job satisfaction, multinomial logit model

**Kamila Trzcńska**

Uniwersytet Łódzki

University of Lodz

**Małgorzata Ćwiek**

Uniwersytet Ekonomiczny w Krakowie

Cracow University of Economics

**ESTYMACJA ROZKŁADU DOCHODÓW LUDNOŚCI CZECH, POLSKI I SŁOWACJI**

Badania rozkładów dochodów skupiają się głównie na próbach dopasowania rozkładów teoretycznych do empirycznego rozkładu dochodów w różnych przekrojach oraz na analizie tych rozkładów. Modele rozkładu dochodów mogą być wykorzystywane bezpośrednio do oceny poziomu życia lub do porównania poziomu życia pomiędzy poszczególnymi obszarami lub narodami. Analiza poziomu życia ludności jest bezpośrednio uzależniona od dochodów, konieczne jest zatem znalezienie odpowiedniego teoretycznego rozkładu dochodu ludności. W literaturze istnieją propozycje różnych typów funkcji matematycznych. Wyniki analiz wykazują, że model Daguma oraz model logarytmiczno-normalny bardzo dobrze aproksymują rozkłady dochodów wielu krajów. W 2010 r. Zenga skonstruował trzyparametrowy model rozkładu dochodów, który ma wiele interesujących własności statystycznych. Do estymacji parametrów funkcji gęstości rozważanych modeli zastosowano metodę największej wiarygodności oraz inwariantną metodę D'Addario. Głównym celem artykułu jest zastosowanie i porównanie modeli Daguma, log-normalnego oraz Zengi do badania rozkładu dochodów w Czechach, Polsce i na Słowacji. Wyniki przeprowadzonej analizy rozważanych modeli dla ludności Czech, Polski i Słowacji potwierdziły dobrą zgodność dopasowania rozkładów Daguma i Zengi do danych empirycznych.

Słowa kluczowe: rozkłady dochodów, model Daguma, model log-normalny, model Zengi, indeksy nierównomierności

## ESTIMATION OF THE INCOME DISTRIBUTION OF CZECH, POLAND AND SLOVAKIA HOUSEHOLDS

Studies on income distributions focus on describing empirical distributions by a theoretical model. This approach can be useful for many reasons. Income models may be used directly for the assessment of the living standard or for comparison of living standards among individual areas or nations. For statistical analysis of the living standard, the focus must be only on measurable elements of the living standard. To quantify this element of the living standard of a population which is directly dependent on income, it is necessary to elicit the correct level and structure of the population income to find suitable statistical models of income distribution for individual social classes as well as for the population as a whole, irrespective of social classes. The literature proposes various types of mathematical functions. Very high accuracy with empirical distribution is characteristic for the Dagum and log-normal models. In 2010, Zenga proposed a new three-parameter model for distributions by size that can be used to represent income, wealth financial and actuarial variables. The aim is to apply and compare the Dagum, log-normal, and Zenga models to income distributions in the Czech Republic, Poland and Slovakia. The parameter estimates of the Dagum and the log-normal distributions were obtained using the maximum likelihood method, in the case of the Zenga distribution, it was obtained using D'Addario's invariants methods. The results of the conducted approximations presented in the paper confirmed the consistency of the Dagum and Zenga distributions with the empirical household income distribution in the Czech Republic, Poland and Slovakia.

Keywords: income distribution, Dagum model, log-normal model, Zenga model, inequality index

**Paweł Ulman**

Uniwersytet Ekonomiczny w Krakowie  
Cracow University of Economics

**Agnieszka Wałęga**

Uniwersytet Ekonomiczny w Krakowie  
Cracow University of Economics

**Małgorzata Ćwiek**

Uniwersytet Ekonomiczny w Krakowie  
Cracow University of Economics

**DOCHODY I WYDATKI GOSPODARSTW DOMOWYCH Z OSOBAMI NIEPEŁNOSPRAWNYMI**

Celem referatu jest zaprezentowanie wyników badań własnych autorów dotyczących oceny sytuacji ekonomicznej gospodarstw domowych z osobami z niepełnosprawnościami w zakresie dochodów oraz zaspokojenia potrzeb tych gospodarstw (w tym ich wydatków) w różnych przekrojach społeczno-ekonomicznych. Dla realizacji celu badań wykorzystano indywidualne dane z badania budżetów gospodarstw domowych z 2019 r., na podstawie których oszacowano modele rozkładu dochodów (rozkład Daguma) oraz wyznaczono charakterystyki tych rozkładów. Zastosowano także mierniki odległości rozkładów oraz przeprowadzono analizę ubóstwa przy wykorzystaniu dochodowej i wydatkowej relatywnej linii ubóstwa. W zakresie wydatków oszacowano modele regresyjne pozwalające na określenie elastyczności dochodowej tych wydatków.

Zidentyfikowano zróżnicowanie rozkładów dochodów gospodarstw domowych z osobami z niepełnosprawnością i bez takich osób w zakresie przeciętnego poziomu tych dochodów oraz poziomu nierówności dochodowych. Rezultat ten potwierdziły mierniki odległości między rozkładami oraz analiza ubóstwa. W zakresie analizy wydatków pokazano, że w gospodarstwach domowych z osobami z niepełnosprawnościami wydatki charakteryzują się niższym poziomem i inną strukturą ze względu na ich rodzaj oraz inną elastycznością dochodową niż w przypadku gospodarstw jedynie z osobami sprawnymi.

Gospodarstwa domowe z osobami z niepełnosprawnościami muszą – oprócz innych trudności – zmagać się z przeciętnie gorszą sytuacją ekonomiczną, co jest spowodowane z jednej strony niską aktywnością ekonomiczną osób z niepełnosprawnościami oraz relatywnie niskimi ich dochodami, a z drugiej strony koniecznością finansowania świadczeń zdrowotnych. Przeprowadzone badania wskazują, że w Polsce wciąż jest jeszcze wiele do zrobienia w zakresie polityki społecznej kierowanej

do osób z niepełnosprawnościami i ich opiekunów, zwłaszcza w perspektywie starzenia się społeczeństwa, a więc potencjalnego wzrostu liczby osób z niepełnosprawnościami.

Słowa kluczowe: osoby z niepełnosprawnościami, dochody, wydatki, zaspokojenie potrzeb

## INCOME AND EXPENSES OF HOUSEHOLDS WITH DISABLED PEOPLE

The aim of the paper is to present the results of the authors' own research on the assessment of the economic situation of households with people with disabilities in terms of income and meeting the needs of these households (including their expenses) in various socio-economic dimensions. To achieve the aim of the research, individual data from the 2019 Household Budget Survey were used. Based on these data, models of income distribution (Dagum distribution) were estimated, and the characteristics of the distributions were determined. Distribution distance measures were also used and a poverty analysis was carried out using the income and expenditure relative poverty line. Regarding expenses, regression models allowing the determination of the income elasticity of these expenses were estimated.

A differentiation of the income distribution of households with and without disabled persons was identified in terms of the average level of this income and the level of income inequality. This result was confirmed by the measures of the distance between distributions and the poverty analysis. In terms of the analysis of expenditure, it was shown that in households with disabled persons, expenditure is characterized by a lower level and a different structure due to its type, and a different income elasticity than in the case of households with disabled persons only.

Households with people with disabilities must – among other difficulties – struggle with an average worse economic situation, which is caused, on the one hand, by low economic activity of people with disabilities and their relatively low incomes, and, on the other hand, by the need to finance health services. The conducted research shows that in Poland there is still a lot to be done in the field of social policy addressed to disabled people and their carers, especially in the perspective of the aging of the society, and thus the potential increase in the number of disabled people.

Keywords: disabled people, income, expenses, meeting needs

**Dorota Wyszowska**

Uniwersytet w Białymstoku

University of Białystok

## **SYTUACJA OSÓB STARSZYCH W POLSCE – STAN OBECNY, PROBLEMY I WYZWANIA**

Spadek liczby ludności razem z procesem starzenia się społeczeństwa w Polsce stanowi wyzwanie dla państwa, w tym w zakresie wdrażania polityki senioralnej, i dla podmiotów gospodarczych w zakresie zaspokajania specyficznych potrzeb osób starszych (w ramach gospodarki senioralnej). Celem opracowania jest przedstawienie wybranych zagadnień dotyczących sytuacji ekonomicznej, mieszkaniowej, ochrony zdrowia, opieki społecznej oraz aktywności osób starszych w Polsce wpływających na ich warunki życia, kierunki polityki senioralnej i zachowania seniorów na rynku. Do opisu wybranych aspektów życia osób starszych w prezentacji dla kraju i niekiedy także w układzie województw posłużą dane z zasobów statystyki publicznej oraz dane pozyskane przez Urząd Statystyczny w Białymstoku ze źródeł pozastatystycznych (NFZ i BIK). Na ich podstawie zostaną zaprezentowane spodziewane problemy i wyzwania polityki i gospodarki senioralnej.

Sytuacja osób starszych w Polsce jest bardzo zróżnicowana pod względem różnych jej aspektów, jak i w układzie terytorialnym. Wyższa przeciętna długość życia wpływa na stan zdrowia oraz singularyzację, które mają odzwierciedlenie w zdolności do zaspokajania potrzeb seniorów. Wszystko to sprawia, że państwo na szczeblu centralnym i lokalnym musi podejmować działania zmierzające do poprawy warunków życia osób starszych, a podmioty gospodarcze – dostosowywać swoją ofertę do nowych potrzeb klientów seniorów, stąd zarówno przed państwem, podmiotami gospodarczymi, ale także statystyką publiczną pojawiają się nowe wyzwania związane z rosnącym udziałem osób starszych w społeczeństwie, potrzebą monitorowania ich sytuacji i stopnia zaspokajania ich potrzeb.

Słowa kluczowe: osoby starsze, depopulacja, starzenie społeczeństwa, polityka senioralna, gospodarka senioralna

## **THE SITUATION OF OLDER PEOPLE IN POLAND – CURRENT STATE, PROBLEMS, AND CHALLENGES**

A gradual decrease in the population in Poland in the recent years together with the ageing of the population pose a challenge for the state, also within the implementation of the senior policy, as well as for economic entities while meeting the special needs of older people (within the silver

economy). The aim of this work is to present some issues concerning the situation of older people in Poland from the economic, housing, healthcare, social assistance and activity angles that affect seniors' living conditions, senior policy directions and the market behaviour of older people. Data from public statistics shown for the entire country and sometimes by voivodship are used along with figures acquired from the Statistical Office in Białystok and from non-statistical sources (NHF and CIB) to describe selected aspects of the life of older people. The data constitute the basis for the description of the expected problems and challenges for senior policy and the silver economy. There are visible differences in the situation of older people in Poland in terms of its different aspects and in territorial division. An increasing life expectancy has an influence on health and singularization, which are reflected in the ability to satisfy seniors' needs. All these make central and local levels of state authorities undertake activities to improve the living conditions of older people, while economic entities adjust to the changing needs of their senior clients. Thus, the state, economic entities and public statistics are faced with new challenges related to the growing share of seniors in the population, the need to monitor both the situation of older people and the extent to which their needs are met.

Keywords: older people, depopulation, population ageing, senior policy, silver economy



**Katarzyna Zielonka-Kołośńska**

Narodowy Instytut Muzealnictwa i Ochrony Zbiorów  
National Institute for Museums and Public Collections

**Katarzyna Figiel**

Narodowy Instytut Muzealnictwa i Ochrony Zbiorów  
National Institute for Museums and Public Collections

**FREKWENCJA JAKO WSKAŹNIK EFEKTYWNOŚCI MUZEÓW  
– RELEVANTNOŚĆ I METODOLOGIA**

Celem referatu jest omówienie relewantności frekwencji w kontekście metod parametryzacji muzeów oraz wskazanie na metodologiczne problemy z liczeniem frekwencji oraz frekwencji online w muzeach. Wnioski przedstawione zostaną na bazie zrealizowanych przez NIMOZ badań ankietowych wśród polskich muzeów. Mowa o badaniu „Statystyka muzeów” oraz ankiecie nt. skutków pandemii COVID-19 w środowisku muzealnym. Wyniki ankiet zostały pogłębione analizami literatury przedmiotu przeprowadzonymi na potrzeby projektu parametryzacyjnego realizowanego przez NIMOZ we współpracy z Ministerstwem Kultury i Dziedzictwa Narodowego oraz Uniwersytetem SWPS.

Muzea różnią się w metodach liczenia frekwencji stacjonarnej, np. często wliczają także spacerujących na terenach zielonych. Sprawia to, że dane stają się trudno porównywalne między instytucjami. Uczestnictwo wirtualne jest jeszcze trudniejsze do liczenia i muzea mają duży problem z wyliczeniem frekwencji w wymiarze online, a ten sposób uczestnictwa zyskuje na znaczeniu. Przewartościowania wymaga podejście do frekwencji jako najważniejszego wskaźnika efektywności muzeów. Muzea działają na różnych polach: jako instytucje edukacyjne, naukowo-badawcze, wydają publikacje, a przede wszystkim zarządzają zbiorami i dbają o to, by zachować je w należyтым stanie. W tym kontekście frekwencja wydaje się wskaźnikiem nieprzystającym i niewystarczającym. Co więcej, realizowane badania wskazują na to, że porównywanie danych frekwencyjnych między muzeami jest niemiernodajne, jako że w różnych instytucjach liczona jest ona na różne sposoby. Stąd badanie frekwencji stacjonarnej oraz wirtualnej wśród muzeów wymaga wyraźnego doprecyzowania na poziomie centralnym.

Słowa kluczowe: frekwencja, parametryzacja, muzeum, działalność online

## ATTENDANCE AS EFFICIENCY INDEX FOR MUSEUMS – RELEVANCY AND METHODOLOGY

The aim of the study is to discuss the relevancy of attendance as the method of museums' parametrisation and indicate the methodological problems with counting onsite and online visitors in museums. Conclusions are drawn based on two surveys conducted by the National Institute for Museums and Public Collections (NIMOZ): museum statistics and one about the effects of COVID-19 on museums in Poland. The surveys' results were deepened by additional research based on the literature and expert analysis. These were carried out as part of a research program on the parametrisation in museums conducted by NIMOZ in cooperation with the Ministry of Culture and National Heritage and the SWPS University.

The results of the study show that museums vary in terms of the methods of counting visitors onsite, i.e. some include amblers in their green areas, which makes data hard to compare between museums and thereby unreliable. Online attendance is even harder to measure and museums face problems with counting online visitors. In conclusion, a redefinition of attendance as the main efficiency index for museums is needed. Museums perform in different areas: educational, research, editorial, but what is most important is that they are crucial in storing and preserving collections. In that context museums' attendance as the main efficiency index seems to be insufficient. Furthermore, the results of the conducted research indicate that the collected data are incomparable, because museums vary in the methods of counting visitors. Thus, counting onsite and online attendance needs a clear methodology developed at the central level.

Keywords: attendance, parametrisation, museum, online activities

**Jan Zwierzchowski**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

**Tomasz Panek**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

**COMPARATIVE ANALYSIS OF SUBJECTIVE WELL-BEING IN EU MEMBER STATES**

The methodological objective of the study is to operationalize the measurement of subjective well-being (hereafter: SW-B). In the empirical part of the study we aim to measure and compare the SW-B in EU member states and to identify variables determining its values. The proposed method of measuring SW-B is based on the Sen's capabilities approach. The operationalization was carried out using the MIMIC model (multiple indicators and multiple causes model). This model incorporates both observable functionings (SW-B basic indicators) and socio-demographic characteristics of individuals (conversion factors). The two sets of variables are linked through the latent capabilities, which are interpreted as the SW-B measures. In the empirical part of the study, an analysis of the SW-B in European countries in 2018 was conducted using the EU-SILC data. Characteristics influencing the SW-B were identified across analysed countries. In most countries the SW-B depends mostly on age, self-perceived health and material deprivation. The proposed method of normalisation allowed for comparative analysis of SWB. The highest values of SW-B were observed in Scandinavian and Western European countries. Relatively lower values were observed in Bulgaria, Romania and Southern European countries. The proposed method allowed for conducting comparative analysis of SW-B across various countries and identification of its major determinants.

Keywords: subjective well-being, SEM, EU-SILC, quality of life, capabilities approach

**Iwona Bąk**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**Katarzyna Cheba**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**MODELOWANIE ZIELONEJ TRANSFORMACJI – ZASTOSOWANIE  
STATYSTYCZNEJ ANALIZY DANYCH W SYSTEMATYCZNYM  
PRZEGLĄDZIE LITERATURY**

Zielona transformacja to obecnie jeden z ważniejszych tematów badawczych poruszanych w publikacjach autorów reprezentujących różne dyscypliny naukowe. Prace z tego zakresu obejmują zarówno zagadnienia teoretyczne związane z definiowaniem tego terminu, jak również prace badawcze przedstawiające m.in. analizy aktualnego poziomu zielonej transformacji. Popularne są także opracowania, zawierające wyniki przeglądu literatury. Badania tego rodzaju ograniczają się często do ilościowej prezentacji liczby cytowań oraz liczby prac indeksowanych w bazach danych. Zdecydowanie rzadziej prezentowane są w nich zestawienia analizowanych tematów badawczych, koncentrujące się np. na statystycznym badaniu relacji występujących pomiędzy rozpatrywanymi zagadnieniami. Tego rodzaju podejście zostanie zaprezentowane w niniejszym opracowaniu. W pracy, z blisko 300 publikacji dostępnych w bazie Web of Science poruszających zagadnienia związane z zieloną transformacją, wyodrębniono prace, odwołujące się wprost do modeli zielonej transformacji. Główną uwagę skupiono na poszukiwaniu zmiennych, których można użyć do ich opisu. Wykorzystano w tym celu dwuetapową procedurę badawczą. Na pierwszym etapie dokonano badania współwystępowania słów kluczowych w programie VOSviewer. Na tej podstawie, na etapie drugim, wyodrębniono czynniki użyte do opisu modeli zielonej transformacji. Wykorzystanie analizy log-liniowej pozwoliło na ustalenie optymalnego zbioru zmiennych, charakteryzujących badane modele. Dodatkowym atutem analizy log-liniowej jest możliwość uwzględnienia wpływu interakcji pomiędzy zmiennymi. Z przeprowadzonych analiz wynika, że na obecnym etapie do budowy modeli zielonej transformacji stosuje się wiele różnych podejść badawczych, w których najczęściej wykorzystuje się czynniki środowiskowe i gospodarcze. Zdecydowanie rzadziej poruszane są w nich zagadnienia związane ze społecznymi uwarunkowaniami zielonej transformacji.

Słowa kluczowe: modele zielonej transformacji, analiza log-liniowa, systematyczny przegląd literatury

## MODELING THE GREEN TRANSFORMATION – APPLYING STATISTICAL DATA ANALYSIS IN A SYSTEMATIC REVIEW OF THE LITERATURE

The green transformation is currently one of the most important research topics discussed in the publications of authors representing various scientific disciplines. Works in this field include both theoretical issues related to the definition of this term, as well as research presenting an analysis of the current level of the green transformation. Studies containing the results of the literature review are also popular. This type of research is often limited to the quantitative presentation of the number of citations and the number of works indexed in various bases of publications. The comparisons of the analyzed research topics focusing, for example, on a statistical study of the relationships between the issues under consideration are much less frequently presented. Such an approach will be discussed in this study. In the work, from nearly 300 publications available in the Web of Science database dealing with issues related to the green transformation, papers referring directly to the models of green transformation have been identified. The main focus was to find variables that could be used to describe them. For this purpose, a two-stage research procedure was used. At the first stage, the coexistence of keywords in the VOSviewer program was tested. On this basis, in the second stage, the factors used to describe the green transformation models were distinguished. The use of log-linear analysis allowed the determination of an optimal set of variables characterizing the tested models. An additional advantage of log-linear analysis is the ability to take into account the influence of interactions between the variables. The analyses show that at the present stage, many different research approaches are applied to elaborate on green transformation models, which most often use environmental and economic factors. Issues related to the social determinants of green transformation are much less frequently discussed.

Keywords: green transformation models, log-linear analysis, systematic literature reviews

**Jacek Białek**

Uniwersytet Łódzki; Główny Urząd Statystyczny

University of Lodz; Statistics Poland

**DANE SKANOWANE I NOWE FORMUŁY INDEKSÓW CEN W POMIARZE CPI**

Dane skanowane są to informacje transakcyjne o cenach produktów i wielkości ich sprzedaży, pozyskiwane z informatycznych systemów sieci handlowych na podstawie zeskanowanych kodów kreskowych tych produktów (np. GTIN, EAN czy SKU). Dane skanowane są relatywnie nowym i tanim źródłem informacji w pomiarze inflacji, a ich główną zaletą jest to, iż dostarczają wiedzy zarówno o cenach produktów, jak i wartościach ich sprzedaży z ustalonego okresu obserwacji na poziomie elementarnym. Pomiar inflacji przy zastosowaniu danych skanowanych niesie jednak za sobą konieczność przeprowadzenia wielu zmian metodologicznych związanych z pozyskiwaniem, filtrowaniem oraz przetwarzaniem tego typu informacji. Jednym z towarzyszących temu pomiarowi wyzwań jest wybór formuły indeksu cenowego dla najniższego poziomu agregacji danych, przy czym najlepszym wyborem wydają się być tu indeksy multilateralne. Praca omawia dwie autorskie propozycje indeksów multilateralnych z punktu widzenia podejścia aksjomatycznego, podejścia stochastycznego oraz czasochłonności obliczeniowej. Zaprezentowano wyniki badania empirycznego, jak i symulacyjnego, które potwierdzają wysoką użyteczność proponowanych indeksów cen. W szczególności pokazano, iż autorskie indeksy spełniają najważniejsze, wymagane aksjomaty (włączając tzw. test identyczności) oraz są relatywnie szybkie do oszacowania.

Słowa kluczowe: dane skanowane, wskaźnik cen towarów i usług konsumpcyjnych, inflacja, indeksy multilateralne, łańcuchowy dryf

**SCANNER DATA AND NEW PRICE INDEX FORMULAS IN THE CPI MEASUREMENT**

Scanner data are transaction data that specify product prices and their expenditures obtained from supermarkets' IT systems by scanning bar codes (i.e. GTIN, EAN or SKU). Scanner data are a relatively new and cheap data source for the calculation of the Consumer Price Index (CPI) and the main advantage of using scanner data is the fact that they provide full information about products even on the lowest data aggregation level, i.e. scanner data provide information about product prices and their expenditures for the given time interval of observations at the elementary level. The inflation measurement based on scanner data requires lots of methodological changes connected

with obtaining, filtering and processing this kind of data. One of the main challenges concerning the above-mentioned measurement is the choice of the appropriate price index formula for the lowest (product) level where multilateral indices seem to be the best choice. The paper discusses two original propositions of multilateral indices with respect to axiomatic and stochastic approaches and their time-consumption. Results from empirical and simulation studies are presented and the high usefulness of these proposals is confirmed. In particular, it is proved that these two original price indices satisfy most of the important axioms (including the identity test) and that they can be computed relatively quickly.

Keywords: scanner data, Consumer Price Index, inflation, price indices, multilateral indices, chain drift

**Mirosław Błażej**

Główny Urząd Statystyczny  
Statistics Poland

**Mariusz Górajski**

Główny Urząd Statystyczny  
Statistics Poland

**Magdalena Ulrichs**

Główny Urząd Statystyczny  
Statistics Poland

## **POMIAR LUKI PRODUKCYJNEJ NA PODSTAWIE BADAŃ KONIUNKTURY GOSPODARCZEJ: CZY WYKORZYSTANIE MOCY PRODUKCYJNYCH POMAGA PRZEWIDZIEĆ LUKĘ PRODUKCYJNĄ W POLSCE?**

W artykule zaproponowano dwurównaniowy model mikroekonometryczny do pomiaru luki produkcyjnej na podstawie mikrodanych, pochodzących z badań koniunktury gospodarczej oraz rocznych sprawozdań z działalności przedsiębiorstw. W pierwszym etapie, wykorzystując metody funkcji kontrolnych, oszacowano parametry funkcji produkcji, na podstawie której wyznaczono wskaźniki łącznej produktywności czynników produkcji (TFP) przedsiębiorstw. Drugie równanie modelu opisuje wpływ stopnia wykorzystania mocy wytwórczych na TFP i jest oszacowane za pomocą dynamicznych modeli danych panelowych. We wskaźnikach TFP przedsiębiorstw wyodrębniono indywidualną składową cykliczną oraz średni poziom technologii w danym sektorze. Łącząc je, proponujemy nowe definicje produktu potencjalnego i luki produkcyjnej na poziomie firmy. Ponadto podejście oparte o mikrodane przedsiębiorstw pozwoliło dokonać dekompozycji luki na jej trzy główne składowe, tj. luki związane z kapitałem, pracą i stopniem wykorzystaniu mocy produkcyjnych, jak również pozwoliło wyznaczyć luki na poziomie sektorów i regionów. Przedstawiona metoda jest poddana empirycznej weryfikacji na danych panelowych z sektora przetwórstwa przemysłowego gospodarki polskiej w okresie 2009-2019. W artykule pokazano, że sektorowe i regionalne luki produktowe spełniają różne ścieżki czasowe w zależności od ich wewnętrznych charakterystyk. Zidentyfikowano sektory i regiony, które w danym roku były poniżej swoich potencjalnych wyników oraz te, w których wykorzystanie czynników produkcji generowało ponadprzeciętną wartość dodaną. Przeprowadzona analiza może więc pomóc w dostosowaniu narzędzi polityki makroekonomicznej do specyfiki danej branży lub regionu. Otrzymane nowe miary luki produkcyjnej oparte na mikrodanych są istotnie i dodatnio skorelowane ze wzrostem PKB i stopą inflacji w Polsce. Pokazano, iż miary te lepiej objaśniają inflację w krzywej Phillipsa niż luka produkcyjna publikowana przez OECD.



Słowa kluczowe: produkt potencjalny, luka produktowa na poziomie firm, estymacja funkcji produkcji, metoda Olleya Pakesa, badanie koniunktury gospodarczej

## **OUTPUT GAP ESTIMATION WITH BUSINESS TENDENCY SURVEYS DATA: DOES CAPACITY UTILIZATION HELP PREDICT OUTPUT GAP IN POLAND?**

This study employs a two-stage micro-econometric model to estimate an output gap based on business tendency surveys and annual reports on enterprises' activities. In the first stage, we applied control function methods to estimate the production function and determine individual enterprises' total factor productivity (TFP). The second model equation describes the positive relationship between firm-level TFP and capacity utilization; we estimated this using a dynamic panel-data model. We then decomposed firm-level unadjusted TFP indices into business cycle components and the actual technology levels in a given sector. In combining these, we obtained several new definitions of potential outputs and output gaps. Moreover, our firm-level approach decomposes the output gap into three main components (capital-, labor - and capacity utilization-based output gaps) and sectoral and regional output gaps. We tested our method with an illustration of the Polish manufacturing industry in 2009-2019 and assessed its advantages and limitations. We observed that sectoral and regional output gaps present different time paths depending on their intrinsic characteristics. Subsequently, we identified the sectors and regions below their potential outputs in a given year, and those in which production factors generated above-average value-added. Thus, our analysis may help policymakers adjust macroeconomic policy tools to the characteristics of a particular industry or region. Our results also reveal the economic significance of the new output-gap estimates. In particular, the micro-founded output gap measures significantly and positively correlate with GDP growth and inflation rates. We conclude that the micro-founded output gap measures in the Phillips curve are more important than the OECD output gap.

Keywords: potential output, firm-level output gap, production function estimation, Olley and Pakes method, business tendency survey

**Mirosław Błażej**

Główny Urząd Statystyczny  
Statistics Poland

**Hubert Stefaniak**

Główny Urząd Statystyczny  
Statistics Poland

**WPŁYW PANDEMII COVID-19 NA SYTUACJĘ POLSKICH PRZEDSIĘBIORSTW – METODOLOGIA ORAZ WYNIKI BADANIA**

W referacie zostaną zaprezentowane metodologiczne aspekty oraz efekty prac nad prowadzonym przez GUS badaniem wpływu pandemii COVID-19 na sytuację polskich przedsiębiorstw (dodatkowy moduł wprowadzony w ramach badania koniunktury gospodarczej). Wystąpienie obok zagadnień metodologicznych obejmie także dyskusję wyników tej pracy oraz wnioski z niej płynące (tak w obszarze wyzwań metodologicznych, jak i praktyki badawczej). Do standardowego badania koniunktury gospodarczej dołączany jest od kwietnia 2020 r. dodatkowy moduł pytań. W badaniu efektów pandemii uwzględniono zarówno pytania ilościowe, jak i jakościowe, odnoszące się do podstawowych wymiarów działania przedsiębiorstwa i zjawisk ekonomicznych. Większość pytań nie ulega zmianie, co pozwala na uzyskanie porównywalnego w czasie obrazu ewolucji sytuacji badanych przedsiębiorstw. Wprowadzane są jednak modyfikacje w celu dostosowania ich do aktualnych wyzwań, przed jakimi stoją przedsiębiorstwa. Dodatkowe pytania kierowane są do szerokiego grona przedsiębiorstw, ze znaczną częstotliwością (w ujęciu miesięcznym), a pierwsze wyniki są prezentowane w kilka dni po przeprowadzeniu badania. Informacje uzyskane z dodatkowego modułu dostarczają – równoległe z wynikami standardowej części badania koniunktury i innymi danymi publikowanymi przez statystykę publiczną – cennej wiedzy, pozwalającej na szybką ocenę wpływu pandemii na kondycję przedsiębiorstw. Badanie to swoim zakresem, jak i momentem publikacji, wyprzedziło inne opracowania o podobnym charakterze. Z uwagi na skalę przeprowadzanego badania, jego tematykę i szczegółowość, badanie jest bezprecedensowym źródłem szybkiej informacji o wpływie pandemii na przedsiębiorstwa. Doświadczenia wyniesione z pracy nad modulem COVID-19 mogą być wykorzystane w perspektywnym badaniu „szybkiej diagnostyki podstawowych procesów makroekonomicznych w średnim okresie”.

Słowa kluczowe: koniunktura gospodarcza, COVID-19, wskaźniki

## THE IMPACT OF COVID-19 PANDEMIC ON POLISH ENTERPRISES' SITUATION – METHODOLOGY AND RESULTS

The paper presents the methodological aspects and results of the work on a survey carried out by Statistics Poland regarding the impact of the COVID-19 pandemic on the Polish enterprises' situation (additional module introduced as part of business tendency surveys). The presentation, apart from the methodological issues, also presents the results of this survey as well as its conclusions (both in the field of methodological challenges and research practice). In April 2020 a supplementary set of questions has been added to the standard business tendency surveys. Both quantitative and qualitative questions relating to the basic dimensions of the company's activity as well as economic phenomena were taken into account. Most of the questions remain unchanged, which allows for a time-comparable picture of how enterprises' situations evolve. However, certain modifications are introduced to adapt them to the current challenges enterprises are facing. Additional questions are addressed to a wide range of companies with a significant frequency (on a monthly basis) and the first results are presented a few days after the survey was conducted. Information gathered from this set provide – along with the results of the standard part of the survey and other data published by official statistics – valuable knowledge that allows a quick evaluation of the impact the pandemic has on the economic situation of enterprises. This survey, regarding its scope as well as the time of its release, was ahead of other publications of a similar nature. Due to its extensiveness, its subject matter and detail, this survey is an unprecedented source of quick information on the impact of the pandemic on companies. The experience gained from working on the COVID-19 module can be used in a prospective study of 'quick diagnostics of basic macroeconomic processes in the medium term'.

Keywords: business tendency, COVID-19, indicators

**Mariusz Cembruch-Nowakowski**

Urząd Statystyczny w Krakowie; Uniwersytet Pedagogiczny im. KEN w Krakowie

Statistical Office in Kraków; Pedagogical University of Krakow

**METODA SZACOWANIA EFEKTÓW DZIAŁALNOŚCI KULTURALNEJ  
W POLSCE PRZY UŻYCIU RACHUNKU SATELITARNEGO KULTURY**

Niniejsza praca ma na celu zaprezentowanie metodologii satelitarnego rachunku kultury jako narzędzia oceny efektów działalności sektora kultury w Polsce przy wykorzystaniu dostępnych danych statystycznych. Przeprowadzona zostanie także analiza efektywności tego sektora gospodarki na tle innych wiodących branż gospodarki w Polsce. Dokonane zostanie porównanie efektywności kultury w Polsce i w kilku wybranych krajach Unii Europejskiej. Metodologia: Do opisu efektów działalności kulturalnej RSK wykorzystano dane GUS o podmiotach gospodarki narodowej, działających lub związanych z obszarem kultury oraz prowadzących obroty handlowe z zagranicą. Zastosowano także podstawowe mierniki ekonomiczne, takie jak produkcja globalna, zużycie pośrednie, wartość dodana brutto. Rezultaty: Przedmiotem analizy były aktualnie dostępne dane, dotyczące aktywności kulturalnej w Polsce zebrane przez GUS. Dzięki zastosowaniu metodologii Satelitarnego Rachunku Kultury dokonano poszerzenia obszaru obserwacji i bardziej kompleksowego ujęcia zagadnienia, wychodząc poza tradycyjnie analizowaną sekcję R i włączenie do analizy także innych działań Polskiej Klasyfikacji Działalności, stanowiących elementy aktywności kulturalnej. Ustalono, że działalność kulturalna partycypowała w tworzeniu PKB średnio na poziomie ok. 3,5%, a więc wносиła wkład większy niż łącznie sektory takie jak: rolnictwo, gospodarka leśna, łowiectwo i rybactwo. Wnioski: Kalkulacja efektów działalności kulturalnej i jej wkładu w tworzenie PKB w Polsce za pośrednictwem satelitarnego rachunku kultury stanowi znaczącą wartość dodaną. Podejście takie stwarza możliwość spojrzenia na funkcjonowanie kultury z innej perspektywy niż czyniono to dotychczas przy wykorzystaniu metodyki rachunków narodowych. Ustalenia te mogą mieć znaczenie w kontekście społecznego postrzegania roli i funkcji kultury w Polsce.

Słowa kluczowe: satelitarny rachunek kultury, rachunki narodowe, kultura, PKB, efektywność

**METHOD OF ESTIMATING THE EFFECTS OF CULTURAL ACTIVITY IN  
POLAND USING THE CULTURE SATELLITE ACCOUNT**

This paper aims to present the methodology of the satellite account of culture as a tool for assessing the effects of the activities of the cultural sector in Poland with the use of available statistical data. An

analysis of the effectiveness of this sector of the economy is also carried out in comparison to other leading sectors of the economy in Poland. The effectiveness of culture in Poland and in a few selected countries of the European Union are compared. To describe the effects of the cultural activity of the RSK, the study uses Statistics Poland data on entities of the national economy operating or related to the area of culture and trading with foreign countries. Basic economic measures such as global production, intermediate consumption and gross value added were also applied. The subject of the analysis was the currently available data on cultural activity in Poland collected by Statistics Poland. Through the methodology of the Satellite Cultural Account, the area of observation was extended and the issue was more comprehensive, going beyond the traditionally analyzed section R and including other divisions of the Polish Classification of Activities, which are the elements of cultural activity in the analysis. It was found that cultural activity contributed to the creation of GDP on average at the level of approx. 3.5%, and thus contributed more than the total of sectors such as agriculture, forestry, hunting and fishing. The calculation of the effects of cultural activity and its contribution to the creation of GDP in Poland via the satellite account of culture represents a significant added value. Such an approach makes it possible to look at the functioning of culture from a different perspective than it has been done so far with the use of the national accounts methodology. These findings may be important in the context of the social perception of the role and function of culture in Poland.

Keywords: satellite account of culture, national accounts, culture, GDP, efficiency

**Czesław Domański**

Uniwersytet Łódzki

University of Lodz

**Robert Kubacki**

## ZASTOSOWANIE MODELI DEEP LEARNING DO PROGNOZOWANIA ZAPOTRZEBOWANIA NA WODĘ DLA MIASTA ŁODZI

Celem pracy jest przedstawienie wyników prac nad modelem uczenia głębokiego, który może być wykorzystany do przygotowania prognozy zapotrzebowania na wodę zarówno w krótkim okresie (kilkunastu lub kilku dziesięciu dni), i w długim okresie (jednego roku). Do budowy modelu użyto danych dziennych o tłoczeniu wody w latach 2010-2020. W artykule skuteczność prognozy modelu została porównana z innymi klasycznymi metodami prognozowania szeregów czasowych. Metodologia: Model został zbudowany w oparciu o dane udostępnione przez Miejski Zakład Wodociągów i Kanalizacji w Łodzi. Na podstawie otrzymanego szeregu czasowego przygotowana została wstępna analiza zbioru danych. Zbiór został podzielony na próbki, które służyły do nauki, testowania i walidacji krzyżowej uzyskanych wyników. Sieć neuronowa składała się z dwóch warstw neuronów. Model LSTM był uczony z wykorzystaniem biblioteki Keras w środowisku R. Końcową miarą skuteczności prognozowanego modelu, porównywaną z innymi modelami, był pierwiastek ze średniokwadratowego błędu (RMSE). W trakcie nauki modelu porównywane były różne warianty parametrów i hiperparametrów modelu. Rezultaty: Szereg czasowy, zawierający dane o ilości tłoczzonej wody, charakteryzuje się sezonowością. Na krótkotrwałe zmiany ilości tłoczzonej wody w poszczególnych dniach tygodnia nakładają się różnice w ilości tłoczzonej wody w poszczególnych porach roku i ogólny trend spadkowy obserwowany w okresie kilku lat. W samym modelu najefektywniejsze okazało się użycie danych opóźnionych o 7 dni. Wyniki uzyskane w trakcie badań potwierdzają skuteczność modelu LSTM (Long\_short-term\_memory). Wnioski: Skuteczność modelu LSTM potwierdzona została najniższymi wskaźnikami RSME w obrębie porównywanych modeli na tym samym horyzoncie prognozy. Uniwersalność (każdy Zakład Wodociągowy zbiera tego typu dane) i dostępność oprogramowania (pakiet R jest bezpłatny) mogłaby poprawić zarządzanie systemem gospodarki wodnej w wielu miastach naszego kraju.

Słowa kluczowe: LSTM, uczenie głębokie, popyt na wodę, sieci neuronowe

## APPLICATION OF DEEP LEARNING MODELS TO FORECAST WATER DEMAND FOR THE CITY OF LODZ

The aim of the paper is to present the results of work on a deep learning model, which can be used to prepare forecasts of water demand both in the short term (several or ten days) and in the long term (one year). Daily water pumping data for the years 2010-2020 were used to build the model. In the paper, the forecasting effectiveness of the model was compared with other classic time series forecasting methods. The model was built based on data made available by the Municipal Water Supply and Sewage System Company in Lodz. Based on the obtained time series, a preliminary analysis of the dataset was prepared. The set was divided into samples which were used for learning, testing and the cross-validation of the obtained results. The neural network consisted of two layers of neurons. The Long\_short-term\_memory (LSTM) model was learned using the Keras library in the R environment. The final measure of the predictive performance of the model compared to other models was the root mean square error (RMSE). Different variants of model parameters and hyperparameters were compared during model learning. The time series containing pumped water volume data is characterised by seasonality. Short-term changes in pumped water volumes on particular days of the week are compounded by differences in pumped water volumes between seasons and a general downward trend observed over a period of several years. In the model itself, the use of data delayed by 7 days proved to be the most effective. The results obtained during the study confirm the effectiveness of the LSTM model. The effectiveness of the LSTM model was confirmed by the lowest RSME rates within the compared models at the same forecast horizon. The universality (every waterworks collects this type of data) and availability of the software (the R package is free) could improve the management of the water management system in many cities of our country.

Keywords: LSTM, deep learning, water demand, neural networks

**Eugeniusz Gatnar**

Uniwersytet Ekonomiczny w Katowicach

University of Economics in Katowice

**ŹRÓDŁA PRESJI INFLACYJNEJ W POLSCE**

Inflacja w Polsce mocno wzrosła i w styczniu osiągnęła poziom 9,2 proc. Zgodnie z prognozami pozostanie na podwyższonym poziomie co najmniej do końca przyszłego roku. W referacie zostaną przedstawione źródła procesów cenotwórczych w Polsce i przewidywany horyzont ich oddziaływania.

Słowa kluczowe: ceny, inflacja, CPI, polityka pieniężna, statystyka gospodarcza

**SOURCES OF INFLATIONARY PRESSURE IN POLAND**

Inflation in Poland has increased significantly and achieved the level of 9.2 percent. According to the current forecasts it will remain elevated for the 2 next years. In the presentation the sources of high inflation in Poland will be introduced.

Keywords: prices, inflation, CPI, monetary policy



**Aleksandra Hałka**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**Agnieszka Leszczyńska-Paczesna**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**POMIAR INFLACJI W CZASACH DUŻYCH ZMIAN W STRUKTURZE KONSUMPCJI. OBCIĄŻENIE WSKAŹNIKA CPI Z TYTUŁU SUBSTYTUCJI W POLSCE**

Artykuł porusza problem błędu w pomiarze inflacji, wynikającego z efektu substytucji produktu. Badanie opiera się na porównaniu oficjalnego wskaźnika cen towarów i usług konsumpcyjnych w Polsce ze wskaźnikami uznanymi za superlatywne. Wyniki wskazują na niewielkie niedoszacowanie CPI z tytułu substytucji, co jest wynikiem wyjątkowym na tle wyników dla innych krajów. Dodatkowo w 2020 r. w związku z szybkimi zmianami w strukturze konsumpcji gospodarstw domowych to niedoszacowanie się pogłębiło. Brak zawyżenia CPI może wynikać z częstych korekt wag wykorzystywanych do obliczenia CPI oraz szybszym niż CPI wzrostem cen tych towarów i usług, na które popyt jest stosunkowo nieelastyczny. Dodatkowo w artykule szacujemy wielkość luki plutokratycznej. Wyniki wskazują, że zastosowanie wag plutokratycznych w kalkulacji CPI prowadzi do wyższego wskaźnika cen niż jego demokratyczny odpowiednik

Słowa kluczowe: inflacja, obciążenie CPI, indeks Laspeyresa, indeks Fisher, indeks demokratyczny

**INFLATION MEASUREMENT IN TIMES OF LARGE CONSUMPTION SHIFT. EVIDENCE OF THE CPI BIAS FROM POLAND**

The paper addresses the problem of bias in the measurement of inflation as provided by the price index of consumer goods and services (CPI) in Poland. We estimate the size of the bias resulting from the product substitution effect. The study is based on a comparison of the official consumer price index in Poland with superlative indices. The survey does not identify an upward CPI bias and the findings indicate a slight understatement of the CPI, which is rather unusual. Additionally, in 2020, due to the rapid changes in the households' consumption pattern the underestimation increased. The lack of an overstatement in the CPI may be due to frequent adjustments in the weights used for

CPI calculation and a faster-than-CPI rise in the prices of those goods and services for which the demand is relatively inflexible. Additionally, we deliver the estimates of the plutocratic gap, which points that the use of the plutocratic weights in CPI calculation leads to a higher price index than its democratic equivalent.

Keywords: inflation, CPI bias, Laspeyres index, Fisher index, democratic index

**Maciej Jamiński**

Akademia Górniczo-Hutnicza w Krakowie

AGH University of Science and Technology

**ANALIZA WYDATKÓW NA ALKOHOL PODCZAS PANDEMII COVID-19**

Celem badania było ustalenie odpowiedzi na pytania:

Czy istnieje związek pomiędzy wiekiem, a zmianą modelu konsumpcji alkoholu?

Czy miejsce zamieszkania ma wpływ na poziom konsumpcji i wydatki z tym związane?

Czy wielkość spożywanego alkoholu i nakładów na to związanych różni się wśród kobiet i mężczyzn?

Czy brak lub zmiana sytuacji materialnej miała wpływ na konsumpcję alkoholu przez Polaków?

Jak zmieniły się wydatki ludzi na alkohol przez pandemię?

Badanie zostało przeprowadzone drogą elektroniczną na losowej grupie osób ( $n = 315$ ).

Wyniki:

Podział ze względu na wiek: Ponad 40% badanych osób w wieku od 18 do 25 lat zadeklarowało zmniejszenie spożycia wszystkich alkoholi, a także wydatków z tym związanych. 25% ankietowanych stwierdziło wzrost konsumpcji oraz wydatków związanych z zakupem alkoholu, natomiast 35% respondentów nie zauważyło zmiany swojego spożycia oraz wydatków na to ponoszonych. Największym wzrostem popularności wśród młodych osób podczas pandemii cieszy się wino (67 osób wskazało, że zwiększyło konsumpcję wina wobec 55 osób dla piwa i 57 dla mocniejszych alkoholi). Analogicznie przez to wzrosły także wydatki na konsumpcję wina, co odnotowało 60 osób. Największy spadek konsumpcji i wydatków przypada piwu, co ma swoją przyczynę w zamknięciu lokali oraz braku spotkań na świeżym powietrzu. W przeciwieństwie do młodszych osób badani w wieku od 26 do 30 lat w zdecydowanej większości pozostawili spożycie oraz wydatki na tym samym poziomie. Podobnie jak osoby w wieku 18-25 lat, także osoby od 26 do 30 roku życia chętniej sięgały po wino niż przed pandemią COVID-19, ale czyniły to chętniej (więcej osób odnotowało większą ilość spożywanego wina niż jego zmniejszoną ilość). Zauważalny jest także większy, w stosunku do innych alkoholi, spadek spożycia piwa i nakładów na jego picie. Połowa respondentów także stwierdziła brak zmian w spożyciu i nakładach na zakup alkoholi mocnych (wódki, whisky).

Słowa kluczowe: alkohol, wydatki, pandemia, spożycie, analiza

## THE ANALYSIS OF THE EXPENCE OF ALCOHOL DURING COVID-19 PANDEMIC

More than 40% of respondents aged 18 to 25 declared a decrease in the consumption of all alcohols and a drop in the related expenses. 25% of the respondents reported an increase in consumption and spending on alcohol, while 35% of the respondents did not notice any change in their alcohol consumption and spending on it. Wine noted the greatest increase in popularity among young people during the pandemic (67 persons indicated that they increased wine consumption compared to 55 persons for beer and 57 for stronger alcohols). Similarly, expenditure on wine consumption increased, as reported by 60 persons. The largest drop in consumption and expenditure is attributable to beer, which has its cause in the closure of pubs and lack of outdoor meetings. In contrast to younger people, respondents aged 26 to 30 overwhelmingly left their consumption and expenditure at the same level. Similarly as in the case of 18-25-year-olds, 26-30-year-olds were more likely to consume wine than before the COVID-19 pandemic, but they did so more readily (more people reported increased wine consumption than decreased wine consumption). There was also a noticeable decline in beer consumption and expenditure on beer compared to other alcohols. Half of the respondents also reported no change in consumption and expenditure on the purchase of spirits (vodka, whisky). Beer has also recorded significant drops in consumption (40% of respondents reported a decrease in consumption), as well as in expenditure on it (44% of respondents reported lower expenditure on beer). While according to the respondents aged 30 and above, their alcohol consumption remained at the previous level or it decreased, the value of this consumption increased. This is best illustrated by hard liquor – 71% of the respondents indicated that they have not decreased their spending on it, or even increased it, yet almost 80% indicated a similar or higher level of spending on hard liquor. The wine category also, although to a lesser extent, presents similar results.

Keywords: alcohol, expense, pandemic, consumption, analysis

**Jarosław Janecki**

Szkoła Główna Handlowa w Warszawie

SGH Warsaw School of Economics

## **OCZEKIWANIA INFLACYJNE KONSUMENTÓW – POMIAR I ICH ROLA W POLITYCE PIENIĘŻNEJ**

W celu oceny oczekiwań inflacyjnych wykorzystywanych jest wiele metod i miar. Różnią się one grupą podmiotów, która formułuje oczekiwania inflacyjne (ekonomiści z sektora finansowego, przedsiębiorcy, konsumenci) oraz sposobem kwantyfikacji oczekiwań inflacyjnych (badania ankietowe, zmiany w wycenie instrumentów finansowych). Celem opracowania jest przedstawienie stosowanych, podstawowych metod oceny oczekiwań inflacyjnych konsumentów, w tym szczególnie w przypadku Polski, gdzie zostanie omówiona metodologia stosowana przez GUS w ramach badań koniunktury konsumenckiej. W drugiej części referatu przedstawione zostaną możliwości wykorzystania oczekiwań inflacyjnych w procesie decyzyjnym w ramach prowadzonej polityki pieniężnej.

Słowa kluczowe: oczekiwania inflacyjne, polityka pieniężna

## **CONSUMER INFLATION EXPECTATIONS – MEASUREMENT AND THEIR ROLE IN MONETARY POLICY**

Many methods and measures are used to assess inflation expectations. They differ in the group of entities that formulate inflation expectations (economists from the financial sector, entrepreneurs, consumers) and in the method of quantifying inflation expectations (surveys, changes in the valuation of financial instruments). The aim of the study is to present the applied basic methods of assessing consumer inflation expectations, especially in the case of Poland, where the methodology used by Statistics Poland applied in the survey of the consumer sentiment will be discussed. The second part of the paper focuses on the possibilities of using inflation expectations in the decision-making process within the monetary policy.

Keywords: prices, inflation, CPI, monetary policy

**Magdalena Mojsiewicz**

Uniwersytet Szczeciński

University of Szczecin

## **CZAS UŻYTKOWANIA PRODUKTÓW WŁASNOŚCI INTELEKTUALNEJ**

W systemie rachunków narodowych ESA 2010 kładzie się nacisk na traktowanie nakładów na B+R jako nakładów kapitałowych. Amortyzacja tych nakładów wymaga przyjęcia określonych założeń, dotyczących okresu użytkowania środków trwałych. Przyjęte w polskim systemie rachunków narodowych założeni, odnoszące się do czasu wykorzystania wyników prac B+R, nie zostały do tej pory zweryfikowane. Na podstawie współpracy z Urzędem Patentowym Rzeczypospolitej Polskiej zbudowano bazę danych, dotyczącą czasu trwania ochrony patentowej. W artykule zaproponowano podejście, wykorzystujące czas użytkowania patentu jako czas skorelowany z czasem użytkowania wyników prac B+R z zastosowaniem metod badania czasu trwania w stanie.

Słowa kluczowe: modelowanie czasu trwania, okres użytkowania środków trwałych

## **INTELLECTUAL PROPERTY PROTECTION DURATION**

The ESA 2010 system of national accounts emphasizes the treatment of R&D expenditures as capital expenditures. The depreciation of these outlays requires the adoption of certain assumptions regarding the useful lifetime of fixed assets. The assumptions adopted in the Polish system of national accounts concerning the time of the utilization of R&D work results have not been verified so far. Based on the cooperation with the Patent Office of the Republic of Poland, a database on the duration of patent protection was built. This paper proposes an approach using patent duration as the time correlated with the time of the use of R&D results by means of state duration survey methods.

Keywords: lifetime modelling, useful lifetime of fixed assets

**Paweł Strzelecki**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**Robert Wyszyński**

Narodowy Bank Polski  
National Bank of Poland

**Jakub Growiec**

Szkoła Główna Handlowa w Warszawie  
SGH Warsaw School of Economics

**ROLA IMIGRANTÓW WE WZROŚCIE PKB W POLSCE**

Zjawiska zachodzące w ostatniej dekadzie w Polsce stawiają duże wyzwania dla pomiaru wkładu czynnika praca do wzrostu PKB. Jednym z nich w 2014 roku jest dynamiczny napływ imigrantów krótkookresowych, którzy w większości są „niewidzialni” dla statystyki publicznej. Celem niniejszego opracowania jest uwzględnienie w możliwie pełny sposób aktualnej i przeszłej roli imigrantów (w tym imigrantów krótkoterminowych) w wypracowywaniu wzrostu PKB w Polsce. W analizie wykraczamy poza proste uwzględnienie dodatkowej liczby pracowników w polskiej gospodarce, wynikające z autorskiego oszacowania liczby pracujących w Polsce imigrantów krótkoterminowych. Dzięki metodzie dekompozycji PKB, uwzględniającej heterogeniczność czynników produkcji oraz stopień wykorzystania mocy produkcyjnych w gospodarce, możliwa była ocen, w jakim stopniu zmiana struktury zatrudnienia według cech pracowników i ich miejsc pracy ma znaczenie dla zmian produktywności pracy. Ma to duże znaczenie dla oceny produktywności pracy imigrantów. Opracowanie oparte jest zarówno na danych GUS, jak i różnego typu danych administracyjnych oraz ankietach przeprowadzanych wśród imigrantów na zlecenie Narodowego Banku Polskiego. Wyniki scenariusza bazowego dla lat 2013-18 wskazują, że przeciętny wkład do wzrostu PKB wzrastającej liczby osób pracujących dzięki imigracji z Ukrainy wynosił 0,5pp. rocznie, czyli tylko nieco mniej niż wkład pracy obywateli polskich w tym samym czasie (0,8 pp.).

Słowa kluczowe: rachunkowość wzrostu gospodarczego, imigracja, wkład pracy do PKB, produktywność pracy, TFP

## THE ROLE OF IMMIGRANTS IN THE GDP ACCOUNTING IN POLAND

The phenomena taking place in the last decade in Poland pose challenges for the measurement of the contribution of the work factor to GDP growth. One of them is the dynamic influx of short-term economic immigrants after 2014. Most of these employees are 'invisible' to official statistics. This study aims to take into account as fully as possible the current and past role of immigrants (including short-term immigrants) in generating GDP growth in Poland. In our analysis, we go beyond simply taking into account the additional number of employees in the Polish economy resulting from the original estimate of the number of short-term immigrants working in Poland. Thanks to the GDP decomposition method, taking into account the heterogeneity of production factors and the degree of production capacity utilization in the economy, it was possible to assess the extent to which a change in the employment structure according to the characteristics of employees and their jobs is significant for changes in labour productivity. This is of great importance for assessing the labour productivity of immigrants. The study is based on data from Statistics Poland as well as various types of administrative data and surveys conducted among immigrants at the request of the National Bank of Poland. The results of the baseline scenario for the years 2013-2018 indicate that the average contribution of the increasing number of working people to the GDP growth due to immigration from Ukraine was 0.5 pp. per year, i.e. only slightly less than the labour input of Polish citizens at the same time (0.8 pp.).

Keywords: growth accounting, immigration, labour input, labour productivity, TFP



**Bohdan Wyżnikiewicz**

Główny Urząd Statystyczny

Statistics Poland

**MANIPULOWANIE DANymi STATYSTYKI GOSPODARCZEJ**

Celem referatu jest diagnoza, inaczej identyfikacja i opis zjawiska manipulowania danymi statystyki gospodarczej, na przykładzie produktu krajowego brutto, przez niektórych szeroko rozumianych interesariuszy, takich jak politycy, publicyści, dziennikarze, przedsiębiorcy, naukowcy i inni. Metodologia oparta jest o pogłębioną analizę zgromadzonych licznych udokumentowanych przypadków manipulowania danymi statystyki gospodarczej na przykładzie produktu krajowego brutto. Metodologia obejmuje także analizę domniemych przyczyn, motywów i celów manipulowania statystyką. Omówione zostają najczęściej spotykane techniki manipulacji danymi statystycznymi, jak również często spotykane podważanie wiarygodności statystyki poprzez krytykowanie stosowanych metod i metodologii badań statystycznych. Dotyczy to głównie systemu rachunków narodowych SNA i statystyki cen. Poruszona jest kwestia niewłaściwych interpretacji statystyk jako źródła manipulacji nimi. Rezultaty: Ilustrowana przykładami diagnoza zjawiska manipulowania danymi statystyki gospodarczej z przedstawieniem przyczyn i sposobów manipulowania zainteresowanych statystyką. Wnioski: Statystyka publiczna powinna bardziej przystępnie i zrozumiale dla społeczeństwa prezentować wyniki kluczowych i wrażliwych społecznie badań statystycznych. Większy nacisk powinien być położony na edukację statystyczną. Należy w odpowiedni sposób reagować na przypadki manipulacji danymi statystycznymi.

Słowa kluczowe: manipulacja danymi statystycznymi, produkt krajowy brutto, techniki manipulacji, edukacja statystyczna

**MANIPULATION OF ECONOMIC STATISTICAL DATA**

The purpose of the presentation is the diagnosis and identification of statistical data manipulation (GDP as an example) by various stakeholders such as politicians, journalists, entrepreneurs, scientists and others. The methodology is based on a profound analysis of numerous collected cases of economic statistical data with GDP as an example. The methodology covers the analysis of the alleged reasons and aims of data manipulation. The typical techniques of manipulation as well as cases of undermining statistics by questioning methodology are also discussed. The discussed criticisms concentrate on national accounts and price statistics. Data misinterpretation is considered as a source

of statistical data manipulation. The result of the presentation is a diagnosis of the economic statistical data manipulation with a presentation of the causes and techniques applied. Official statistics should pay increased attention to a more friendly presentation of statistics for the general public. Such an approach is important for socially sensitive statistical data. Statistical education is another task for official statistics. A reaction to the manipulation of statistical data is also necessary.

Keywords: statistical data manipulation, gross domestic product, techniques of data manipulation, statistical education

**Barbara Batóg**

Uniwersytet Szczeciński

University of Szczecin

**Katarzyna Wawrzyniak**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**WPŁYW PANDEMII NA POLARYZACJĘ POWIATÓW WOJEWÓDZTWA  
ZACHODNIOPOMORSKIEGO**

Pandemia COVID-19 ma bardzo duży wpływ prawie na wszystkie dziedziny gospodarki zarówno w skali mikro, jak i w skali makro. Jej skutki są widoczne w różnych obszarach życia społecznego i ekonomicznego. Pojawiają się liczne publikacje, w których analizowany jest wpływ pandemii na zjawiska demograficzne i gospodarcze, ze szczególnym uwzględnieniem rynku pracy. Przegląd wyników tych prac skłonił autorki do przeprowadzenia badań, mających na celu porównanie stopnia polaryzacji powiatów województwa zachodniopomorskiego w stosunku do centrów rozwoju – Szczecina jako stolicy województwa oraz Koszalina jako drugiego dużego ośrodka miejskiego w województwie – przed pandemią oraz w czasie jej trwania. Analiza będzie dotyczyła następujących trzech obszarów, charakteryzujących sytuację w regionie: demografia, podmioty gospodarcze oraz rynek pracy. Dane o wybranych do badania kategoriach z lat 2008–2021 będą pochodziły z Banku Danych Lokalnych GUS. W badaniu zostaną wykorzystane wybrane miary podobieństwa i odległości. Rezultatem przeprowadzonego badania będzie wskazanie tych powiatów, które w analizowanym okresie znalazły się w największej oraz najmniejszej odległości od centrów rozwoju w wybranych obszarach oraz określenie czy w kolejnych latach były to te same powiaty. Dzięki temu zostanie potwierdzona lub odrzucona hipoteza, że pandemia pogłębiła polaryzację powiatów województwa zachodniopomorskiego, czyli słabiej rozwinięte powiaty jeszcze bardziej oddaliły się od centrów rozwoju.

Słowa kluczowe: miary podobieństwa, miary odległości, centra rozwoju, polaryzacja, analizy regionalne

## THE IMPACT OF THE PANDEMIC ON THE POLARIZATION OF POWIATY IN THE ZACHODNIOPOMORSKIE VOIVODSHIP

The COVID-19 pandemic has a very large impact on almost all areas of the economy, both at the micro and macro level. Its effects are visible in various areas of social and economic life. Numerous publications are appearing which analyze the impact of the pandemic on demographic and economic phenomena, with particular emphasis on the labour market. A review of the results of these works led the Authors to conduct research to compare the degree of the polarization of the powiats of Zachodniopomorskie Voivodship in relation to the growth centers - Szczecin as the capital of the voivodship and Koszalin as the second largest urban center in the voivodship - before and during the pandemic. The analysis concerns the following three areas characterizing the situation in the region: the demography, enterprises and the labour market. Data on the variables selected for the study in 2008-2021 come from the Local Data Bank of Statistics Poland. Measures of similarity and distance will be used in the study. The result of the study is an indication of those powiats which were in the greatest and the smallest distance from the growth centers in selected areas in the analyzed period, and whether in consecutive years these were the same powiats. Thanks to this, the hypothesis that the pandemic deepened the polarization of powiats in Zachodniopomorskie Voivodship, i.e. less developed powiats moved further away from the growth centers, will be confirmed or rejected.

Keywords: similarity measures, distance measures, growth centers, polarization, regional analysis

**Agnieszka Borkowska**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Warsaw University of Life Sciences

**Alicja Stolarska**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Warsaw University of Life Sciences

**Robert Pietrzykowski**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Warsaw University of Life Sciences

**PRZESTRZENNE I CZASOWE ZMIANY ZASOBÓW LEŚNYCH W KONTEKŚCIE  
ZRÓWNOWAŻONEJ GOSPODARKI**

Celem pracy jest przedstawienie zmian zachodzących w czasie, jak i w przestrzeni w zasobach leśnych. W pracy badano konwergencję wybranych zasobów leśnych, wykorzystując metody statystyki przestrzennej oraz medianowe odchylenie absolutne. Wykorzystano standardowe narzędzia statystyki przestrzennej oraz medianowe odchylenie absolutne (MAD) do obliczenia odchylenia standardowego  $S$ , które jest niewrażliwe na wartości odstające nawet w przypadku ich 50% udziału. Obserwowane dane wykazywały duże zróżnicowanie i występowanie danych odstających i wpływowych. Zaproponowane działania pozwoliły na określenie kierunku zmian w zasobach leśnych, co wynika m.in. z uwarunkowań klimatyczno-środowiskowych i gospodarczych oraz z okresowego wsparcia zewnętrznego, np. PROW 2014-2020 (poddziałanie 8.5 - inwestycje zwiększające odporność ekosystemów leśnych i ich wartość dla środowiska: nabór do końca 2021 r.). Do określenia konwergencji badanego zjawiska wykorzystano sigma konwergencję w zupełnie nowy sposób, ujmując to zagadnienie w oparciu o medianowe odchylenie absolutne. Oprócz produktów drewnopochodnych lasy są źródłem nieдрzewnych produktów nieżywnościowych i żywnościowych (grzyby, jagody, dziczyzna). Analiza zasobów leśnych w Polsce jest ważna w kontekście ekonomicznym, społecznym i kulturowym. Odpowiednie zarządzanie gruntami leśnymi i ochrona ich bioróżnorodności umożliwia gospodarcze wykorzystanie naturalnych zasobów w zakresie gospodarki zrównoważonej.

Słowa kluczowe: MAD, zasoby leśne, statystyka przestrzenna

## SPATIAL AND TEMPORAL CHANGES OF FOREST RESOURCES IN THE CONTEXT OF SUSTAINABLE MANAGEMENT

The aim of the work is to present the changes of forest resources in time and space. The study investigated the convergence of selected forest resources using methods of spatial statistics and the absolute median deviation. The observed data showed great variability and the presence of outliers and influencers. The work used standard tools of spatial statistics and the median absolute deviation (MAD). MAD was used to calculate standard deviation  $S$ , which is insensitive to outliers even in their 50% share. The proposed actions allowed the determination of the direction of changes in forest resources, which results, among others, from climatic, environmental and economic conditions and periodic external support, e.g. RDP 2014-2020 (sub-measure 8.5 - investments increasing the resilience of forest ecosystems and their value for the environment: recruitment by the end of 2021). To determine the convergence of the studied phenomenon, sigma convergence was used in a completely new way, presenting this issue based on the absolute median deviation. The analysis of forest resources in Poland is important in the economic, social and cultural context. Apart from wood-based products, forests are a source of non-wood, non-food and food products (mushrooms, berries, game). The proper management of forest land and the protection of its biodiversity enables the economic use of natural resources in a sustainable economy.

Keywords: MAD, forest resources, spatial statistics

**Mariola Chrzanowska**

Urząd Statystyczny w Kielcach; Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
Statistical Office in Kielce; Warsaw University of Life Sciences

**Małgorzata Sobieraj**

Urząd Statystyczny w Kielcach  
Statistical Office in Kielce

**METODY SZACOWANIA SZAREJ STREFY – ANALIZA BIBLIOMETRYCZNA**

Szara strefa stanowi ważny i aktualny, wielowymiarowy problem badawczy. Stanowi przedmiot badań zarówno dla ekspertów zajmujących się analizą, oceną i pomiarem gospodarki, jak i dla organów państwa, których zadaniem jest przeciwdziałanie temu zjawisku. Szara strefa występuje od zawsze we wszystkich gospodarkach, niezależnie od poziomu ich rozwoju czy kontekstu geograficznego. Wciąż jednak problematyczne jest określanie wielkości tego zjawiska. Podłoża tego „sporu” można upatrywać już na samym początku, czyli przy wyborze szacowania metody szarej strefy. Każda z metod prezentowanych w literaturze ma zagorzałych zwolenników i przeciwników: uwzględnia nieco inny aspekt tego zjawiska. Powstaje zatem pytanie, które metody są najbardziej popularne i co przemawia za ich stosowaniem? Celem artykułu jest identyfikacja proponowanych w literaturze metod szacowania szarej gospodarki z wykorzystaniem analizy bibliometrycznej. W ramach badań ustalono najczęściej wybierane metody szacowania, wykorzystywane w opracowaniach indeksowanych w bazie Scopus. Dodatkowo w pracy zaprezentowano sieć powiązań badaczy, zajmujących się tą tematyką. Wyniki badań zostały zaprezentowane w postaci sieci powiązań.

Słowa kluczowe: szara strefa, estymacja, analiza bibliometryczna

**METHODS FOR ESTIMATING THE SHADOW ECONOMY – A BIBLIOMETRIC ANALYSIS**

The shadow economy (SE) is an important and current multidimensional research problem. It is a subject of research both for experts involved in the analysis, evaluation and measurement of the economy, and for state bodies whose task is to prevent this phenomenon. SE has always existed in all economies, regardless of their level of development or geographical context. However, it is still problematic to determine the size of this phenomenon. The basis of this ‘dispute’ can be found at

the very beginning, i.e. in the choice of the method for estimating the shadow economy. All the methods presented in the literature have their advantages and disadvantages because of the slightly different characteristics of this phenomenon. Therefore, a question arises of which methods are the most popular and what supports their use? The aim of this paper is to provide a bibliometric analysis of SE estimation methods proposed in the literature. The study indicates the most used estimation methods described in the Scopus database. In addition, the paper presents a network of researchers working on this topic. The results of the research are presented in the form of a network of links.

Keywords: shadow economy, estimation, bibliometric analysis



**Marek Cierpiął-Wolan**

Uniwersytet Rzeszowski

University of Rzeszów

**Dariusz Twaróg**

Politechnika Rzeszowska

Rzeszów University of Technology

**WYKORZYSTANIE REGIONALNEGO POTENCJAŁU W BUDOWANIU  
NIEZALEŻNOŚCI ENERGETYCZNEJ – STUDIUM PRZYPADKU DLA POLSKI  
I UKRAINY**

Produkcja biogazu w Polsce i na Ukrainie wydaje się dobrym sposobem zarówno na ograniczenie emisji gazów cieplarnianych, jak i zwiększenia samowystarczalności energetycznej poprzez uzupełnienie konwencjonalnych źródeł energii. Celem badań była ocena potencjału produkcji biogazu oraz możliwości jego zwiększenia na poziomie regionalnym obu badanych krajów w 2018 roku. Badając wpływ potencjału biogazu na pokrycie zużycia ciepła i energii elektrycznej, dokonano dezagregacji zapotrzebowania dla miesięcznych przedziałów czasowych z pomocą danych meteorologicznych. W ten sposób ustalono, że w miesiącach letnich, w szczególności w czerwcu i w lipcu, w każdym z badanych regionów wystąpi nadwyżka ciepła. Biorąc jednak pod uwagę rolę biogazu w pokryciu zapotrzebowania na gaz ziemny, stwierdzono, że pod tym względem przoduje region Podlaski (~261%). Zarówno w Polsce, jak i na Ukrainie potencjał ten powstaje w ponad 50% z nawozów naturalnych (głównie obornika bydlęcego). Analiza możliwości zwiększenia potencjału w oparciu o uprawę kukurydzy na nieużytkach wykazała, że pod tym względem regiony ukraińskie mają większe możliwości w porównaniu z polskimi. W przypadku 20 regionów maksymalne wykorzystanie potencjału biogazu powinno skutkować zwiększeniem udziału źródeł odnawialnych w miksie energetycznym do poziomu przekraczającego 25%. Polska i Ukraina mają porównywalne potencjały produkcji biogazu na poziomie ~10 mld m<sup>3</sup> rocznie, jednak polskie regiony wykazują znacznie większe zróżnicowanie potencjału produkcji biogazu (140~1094 mln m<sup>3</sup>) niż regiony ukraińskie (90~300 mln m<sup>3</sup>). Z drugiej strony, liczba biogazowni potrzebnych do wykorzystania potencjału biogazu, jak również liczba nowych miejsc pracy tworzonych w obu krajach pozostają na podobnym poziomie. Powyższe analizy przeprowadzono również na poziomie powiatów i rejonów, gdzie omówiono potencjał współpracy regionalnej dla czterech regionów przygranicznych.

Słowa kluczowe: odnawialne źródła energii, biogaz, potencjał regionalny, przedsiębiorczość

## THE USE OF THE REGIONAL POTENTIAL IN BUILDING ENERGY INDEPENDENCE – POLISH AND UKRAINE CASE STUDY

Biogas production in Poland and Ukraine seems to be a good way of both reducing greenhouse gas emissions and increasing energy self-sufficiency by supplementing conventional energy sources. The aim of the research was to assess the potential of biogas production and the possibility of increasing it at the regional level of both studied countries in 2018. When examining the impact of the biogas potential on covering heat and electricity consumption, the demand for monthly time intervals was disaggregated with the help of meteorological data. In this way, it was established that in the summer months, in particular in June and July, each of the studied regions will have a surplus of heat. However, taking into account the role of biogas in covering the demand for natural gas, it was found that the Podlasie region leads in this respect (~261%). Both in Poland and Ukraine, this potential is created by over 50% of natural fertilizers (mainly cattle manure). The analysis of the possibilities of increasing the potential based on the cultivation of maize in wastelands showed that in this respect the Ukrainian regions have better opportunities compared to Polish regions. In the case of 20 regions, the maximum use of the potential of biogas should result in an increase in the share of renewable sources in the energy mix to the level of above 25%. Poland and Ukraine have comparable biogas production potentials of ~10 billion m<sup>3</sup> annually, but Polish regions show a much greater diversity of the biogas production potential (140~1094 million m<sup>3</sup>) than Ukrainian regions (90~300 million m<sup>3</sup>). On the other hand, the number of biogas plants needed to consume the potential of biogas, as well as the number of new jobs created in both countries remains at a similar level. The above analyses were also carried out at the powiat and raion levels, where the potential of regional cooperation for four border regions was discussed.

Keywords: renewable energy sources, biogas, regional potential, entrepreneurship

**Dorota Ciołek**

Uniwersytet Gdański

University of Gdańsk

## **CHANGES IN THE LABOUR MARKET DURING THE COVID-19 PANDEMIC AND THEIR SPATIAL INTERACTIONS – EVIDENCE FROM MONTHLY DATA FOR POLISH LAU**

In addition to the direct negative effects in terms of morbidity and mortality, the pandemic caused by SARS-CoV-2 also has indirect negative effects that concerned, among others, the labour market. This study analysed changes in the unemployment rate that were observed at the level of Polish Local Administrative Units (LAU) during the ten months of the pandemic. Cross-sectional and panel econometric modelling with spatial interactions was applied. Both annual and monthly data were used. It was shown that the observed increase in unemployment was strongly influenced by the share of employment in services, especially in less knowledge-intensive services such as: trade, accommodation and gastronomy. Moreover, it turns out that a higher share of women working in services was associated with a higher increase in unemployment than in the case of men working in services. Significant positive spatial relationships between local labour markets in LAUs were also identified. It was also proven that both the timing and severity of the containment measures were significant. The strongest effect of the lockdown was observed three months after its introduction, while after six months the effect was significantly smaller. The study's findings may be important for post-pandemic recovery plans.

Keywords: COVID-19, labour market, unemployment, containment measures, spatial panel model

**Dawid Dawidowicz**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**Iwona Bąk**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**PORÓWNANIE SYTUACJI FINANSOWEJ JST PRZED I W OKRESIE COVID-19**

Jednostki samorządu terytorialnego (JST) jako samodzielne i autonomiczne podmioty realizują wiele zadań z zakresu infrastruktury społecznej, technicznej, a także porządku i bezpieczeństwa czym istotnie przyczyniają się do rozwoju lokalnych gospodarek, a także polepszenia dobrobytu społeczności lokalnej. Rozprzestrzenianie się pandemii koronawirusa w 2020 roku i w konsekwencji wprowadzane przez rząd zamykanie gospodarki, skutkujące ograniczeniem możliwości przemieszczania się ludności, wpłynęły na finanse nie tylko przedsiębiorstw, ale także i jednostek samorządu terytorialnego. Z kolei w okresach luzowania obostrzeń można było dostrzec szybkie odrabianie strat niektórych przedsiębiorców. Samorządy podjęły wiele dyskrecjonalnych decyzji celem zniwelowania negatywnych skutków spowodowanych koronawirusem. W konsekwencji pandemii zaburzona została przede wszystkim dochodowa część finansów jednostek samorządu terytorialnego. Wpłynęło to na sytuację finansową jednostek samorządu terytorialnego przez decyzje podejmowane nie tylko w zakresie dochodów i wydatków, ale także w zakresie realizowanych inwestycji. Celem artykułu jest określenie sytuacji finansowej poszczególnych szczebli jednostek samorządu terytorialnego w okresie COVID-19 w porównaniu do wybranych okresów z lat poprzednich. Do realizacji postawionego celu zastosowane zostały metody wielowymiarowe. Wykorzystano dane Ministerstwa Finansów. Badanie przeprowadzono w oparciu o najważniejsze wskaźniki, charakteryzujące sytuację finansową jednostek samorządu terytorialnego wszystkich szczebli (województw, powiatów i gmin). Wyniki badania dostarczają dowodów na to, że chociaż wartości zobowiązań per capita w większości JST wzrosły w roku covidowym, to jednak udział zobowiązań w dochodach ogółem zmniejszył się przy jednoczesnym spadku udziałów wydatków inwestycyjnych w wydatkach ogółem. Świadczy to o ograniczeniu realizacji inwestycji celem poprawy sytuacji finansowej JST.

Słowa kluczowe: JST, sytuacja finansowa, COVID-19, wielowymiarowa analiza statystyczna

## COMPARISON OF THE FINANCIAL SITUATION OF LOCAL GOVERNMENT UNITS BEFORE AND DURING THE COVID-19 PERIOD

Local government units, as independent and autonomous entities, perform many tasks in the field of social and technical infrastructure, as well as public order and safety, which significantly contribute to the development of local economies and improve the well-being of the local community. The spread of the coronavirus pandemic in 2020 and, in consequence, the lockdown of the economy introduced by the government, resulting in limiting the possibility of population movement, affected the finances of not only enterprises but also local government units. On the other hand, in the periods of easing the restrictions, it was possible to notice a quick increase in profits by some entrepreneurs. Local governments have taken many discretionary decisions to counter the negative effects caused by the coronavirus. As a result of the pandemic, the income part of the finances of local government units was disturbed. This influenced the financial situation of local government units, by decisions made not only in terms of income and expenses, but also in terms of the implemented investments. The aim of the article is to determine the financial situation of individual levels of local government units in the COVID-19 period in comparison to selected periods from the previous years. To achieve the goal, multi-dimensional methods were used. The data of the Ministry of Finance were used. The study was conducted on the basis of the most important indicators characterizing the financial situation of local government units at all levels (voivodships, powiats and gminas). The results of the study provide evidence that although the value of liabilities per capita in most of local government units increased in the covid year, the share of liabilities in the total income decreased with a simultaneous decline in the share of investment expenditure in the total expenditure. This proves that the implementation of investments is limited in order to improve the financial situation of local government units.

Keywords: local government units, financial situation, COVID-19, multivariate statistical analysis

**Sylwia Filas-Przybył**

Urząd Statystyczny w Poznaniu; Uniwersytet im. A. Mickiewicza w Poznaniu  
Statistical Office in Poznań; Adam Mickiewicz University Poznań

**MIGRACJE WAHADŁOWE ZWIĄZANE Z ZATRUDNIENIEM W UJĘCIU ANALIZY SIECIOWEJ**

Rosnące zapotrzebowanie na dane, pozwalające na opisanie rzeczywistości społeczno-gospodarczej za pomocą powiązań między różnymi obiektami i zjawiskami, inspirowane statystykę publiczną do poszukiwania źródeł umożliwiających pozyskanie takich informacji. Przykładem badania, prowadzonego przez statystykę publiczną, które dostarcza tego rodzaju danych, jest badanie przepływów ludności związanych z zatrudnieniem. Każdorazowo badanie, a doczekało się ono już trzech edycji, dostarcza informacji korelujących miejsca pracy i zamieszkania pracowników najemnych. Jedną z form prezentowania danych z badania dojazdów do pracy jest macierz przepływów, pokazująca skalę migracji wahadłowych pomiędzy gminami, z których osoby wyjeżdżają do pracy i gminami, w których pracują. Otrzymana macierz może być dla analityków bardzo wygodnym narzędziem danych, które w prosty sposób można z kolei przekształcić do postaci grafu, opisującego fragment rynku pracy za pomocą sieci powiązań między jednostkami terytorialnymi za pomocą wierzchołków (oczek siatki kilometrowej, gminy, powiaty, województwa) i krawędzi (strumień pracowników najemnych, odległość drogowa, czas przejazdu). Celem referatu jest pokazanie możliwości jakie niesie spojrzenie na zjawisko migracji wahadłowych z perspektywy teorii grafów na różnych etapach procesu realizacji badania statystycznego, zarówno na etapie opracowywania zebranego materiału statystycznego – danych z pozyskanych rejestrów administracyjnych, ale także podczas wizualizacji opracowanych wyników i prowadzenia analiz statystycznych.. Ponadto w referacie zostaną przedstawione rozwiązania, które mogą udoskonalić dotychczas stosowaną metodologię badań migracji wahadłowych związanych z pracą.

Słowa kluczowe: dojazdy do pracy, analiza sieciowa, analiza grafów, obserwacje nietypowe

**WORK-RELATED CIRCULAR MIGRATIONS – NETWORK ANALYSIS APPROACH**

The growing demand for data allowing the description of the socio-economic reality with the help of connections between various objects and phenomena, inspires public statistics to search for sources

enabling obtaining such information. An example of a survey conducted by official statistics that provides this type of data is the study of population flows related to employment. Each time, the study (which has already had its third edition) provides information correlating the places of work and residence of employees. One of the forms of presenting data from the study of commuting to work is the flow matrix, showing the scale of circular migrations between the gminas from which people leave for work and the gminas where they work. The obtained matrix can be a very convenient data tool for analysts and it can easily be transformed into a graph describing a part of the labor market by means of a network of connections between territorial units represented by vertices (1 km<sup>2</sup> grid squares, gminas, powiats, voivodships) and edges (employee streams, road distances, travel times). The aim of the paper is to show the possibilities of looking at the phenomenon of commuting to work from the perspective of the graph theory at various stages of the statistical research implementation process, both at the stage of developing the collected statistical material – data from the administrative registers – and also during the visualization of the developed results and conducting statistical analyses. In addition, the paper presents solutions that may improve the currently used methodology of work-related circular migration research.

Keywords: commuting to work, network analysis, graph analysis, atypical observations

**Sylvia Filas-Przybył**

Urząd Statystyczny w Poznaniu; Uniwersytet im. A. Mickiewicza w Poznaniu

Statistical Office in Poznań; Adam Mickiewicz University Poznań

**Tomasz Klimanek**

Urząd Statystyczny w Poznaniu; Uniwersytet Ekonomiczny w Poznaniu

Statistical Office in Poznań; Poznań University of Economics and Business

**KILKA UWAG O DYLEMATACH GEOPRZESTRZENNYCH W DOSKONALENIU  
BADAŃ NAD DOJAZDAMI DO PRACY**

Celem artykułu jest przedstawienie dylematów geoprzestrzennych, dotyczących uwzględniania dodatkowych informacji o dojazdach do pracy, tj. odległości drogowych w km oraz czasu potrzebnego na przejazd z jednej lokalizacji do drugiej. Oficjalne statystyki, dotyczące kierunków i wielkości dojazdów do pracy, są publikowane od 2006 r. w formie macierzy międzygminnej co pięć lat. Od 2019 r. zakres dostępnych informacji został rozszerzony o odległości drogowe i czasy podróży samochodem między gminami, które są przedstawiane w takiej samej formie macierzowej jak w przypadku przepływów ludności związanych z zatrudnieniem. W celu obliczenia odległości drogowych i czasu przejazdu pomiędzy gminami konieczne jest wyznaczenie współrzędnych jednego punktu, reprezentującego całą gminę. Dotychczas stosowanym rozwiązaniem były tzw. centroidy gminne. Ten sposób reprezentacji gmin w przestrzeni budzi jednak wiele zastrzeżeń. Autorzy przedstawiają różne sposoby definiowania punktów charakterystycznych gminy i dokonują krytycznej oceny każdego z nich. Z ich analizy wynika, że najlepszym rozwiązaniem jest stosowanie odrębnych medoid dla miejsc zamieszkania i miejsc pracy. Zaproponowano również metodę wyznaczania medoidy miejsca zamieszkania, która nie jest jednak oparta na odległości euklidesowej, lecz na algorytmie Dijkstry, który minimalizuje odległość drogową między wszystkimi współrzędnymi w zbiorze danych.

Słowa kluczowe: dojazdy do pracy, centroidy, medoidy, algorytm Dijkstry



## SOME REMARKS ON GEOSPATIAL DILEMMAS IN IMPROVING THE STUDY ON COMMUTING TO WORK

The aim of the article is to present geospatial dilemmas regarding the inclusion of additional information on commuting, i.e. road distances in km and the time needed to drive from one location to another. Official statistics on the directions and size of commuting flows have been published in the form of an inter-commune matrix every five years since 2006. Since 2019, the scope of the available information has been extended to include road distances and car travel times between communes, which are presented in the same matrix form as in the case of employment-related population flows. In order to calculate road distances and travel times between communes, it is necessary to determine the coordinates of one point representing the entire commune. Until now the solution has been to use commune centroids. However, there are many objections to this way of representing communes in space. The authors present various ways of defining characteristic points of a commune and provide a critical assessment of each approach. Their analysis indicates that the best solution is to use separate medoids for places of residence and workplaces. A method of determining medoids of places of residence is also proposed; it is not based on the Euclidean distance, but on the Dijkstra algorithm, which minimizes the road distance between all the coordinates in the dataset.

Keywords: commuting to work, centroids, medoids, Dijkstra algorithm

**Iwona Forys**

Uniwersytet Szczeciński

University of Szczecin

## **INDEKSY CEN LOKALI MIESZKALNYCH NA JEDNORODNYCH OBSZARACH MIASTA: ANALIZA NA PRZYKŁADZIE SZCZECINA**

Celem badania jest wyznaczenie indeksów cen mieszkań na małych obszarach - jednorodnych z uwagi na istotne dla rynku nieruchomości mieszkalnych czynniki cenotwórcze. Weryfikowana jest hipoteza badawcza, dotycząca rozbieżności wartości indeksów przy zastosowaniu różnych kryteriów wyboru zakresu badanego obszaru (indeksy małych obszarów, indeksy zagregowane). Wykorzystano modele hedoniczne cen dla pełnych danych. Indeksy liczone dla małych obszarów, a następnie agregowane dla miasta, istotnie różnią się od indeksu wyznaczanego dla całego obszaru ze zmienną lokalizacją. Uzyskane wyniki są przyczynkiem do dalszych badań na innych rynkach lokalnych oraz dla innych rodzajów nieruchomości. Jest to zgodne z koncepcją proponowaną przez Roberta Shillera, laureata Nagrody Nobla w tej dziedzinie.

Słowa kluczowe: rynek mieszkaniowy, indeksy cen, modele hedoniczne

## **HOUSING PRICE INDICES IN HOMOGENOUS CITY AREAS: ANALYSIS BASED ON THE EXAMPLE OF SZCZECIN**

The aim of the study is to determine housing price indices in small areas which are homogeneous in terms of the significant for the residential real estate market price drivers. The paper verifies the research hypothesis on the divergence of index values with the application of different criteria for the selection of the studied area scope (small area indices, aggregated indices). Hedonic price models for the complete data were used. Indices calculated for small areas and then aggregated for the city are significantly different from the index calculated for the whole area with the location variable. The results obtained are a contribution to the further research of other local markets and for other types of properties. This is in compliance with the concept proposed by Robert Shiller, a Nobel Prize winner in this field.

Keywords: housing market, price indices, hedonic models

**Romana Głowicka-Wołoszyn**

Uniwersytet Przyrodniczy w Poznaniu  
Poznań University of Life Sciences

**Andrzej Wołoszyn**

Uniwersytet Przyrodniczy w Poznaniu  
Poznań University of Life Sciences

**PRZESTRZENNY CHARAKTER NIERÓWNOŚCI DOCHODOWYCH  
SAMORZĄDÓW GMINNYCH W POLSCE**

Endogeniczne możliwości pozyskania dochodów przez gminy są bardzo zróżnicowane i skutkują powstawaniem nierówności dochodowych, a w konsekwencji również społecznych. Mogą one prowadzić do hamowania rozwoju lokalnego, rozwarstwienia społecznego i utrudniać realizację polityki spójności. Jednak pomiędzy jednostkami przestrzennymi, takimi jak gminy, występują wzajemne oddziaływania w postaci powiązań ekonomicznych (np. w sferze produkcji) czy demograficzno-społecznych (np. dojazdy do pracy). Powiązania te przenoszą się m.in. na sytuację dochodową gmin. Stąd też w licznych badaniach, dotyczących nierówności dochodowych gmin, coraz częściej pojawia się potrzeba wzięcia pod uwagę efektów przestrzennych, które uwzględniałyby znaczenia lokalizacji przestrzennej w kształtowaniu ogólnego poziomu nierówności. Celem badań była identyfikacja efektów przestrzennych w kształtowaniu poziomu nierówności w zakresie dochodów własnych gmin w Polsce. Analizowano znaczenie wzorców przestrzennych w kształtowaniu nierówności najważniejszych składowych dochodów własnych gmin. W badaniach do oceny wpływu wzorców przestrzennych na poziom ogólnej nierówności zastosowano proponowaną przez autorów modyfikację miernika nierówności przestrzennych, przedstawionego przez Reya i Smitha (2013). Zastosowany miernik nierówności pozwolił na ocenę nierówności, wynikających z oddziaływań sąsiedzkich, jako składnika ogólnego poziomu nierówności. Składnik nierówności sąsiedzkich wydzielony został z ogólnego poziomu nierówności na podstawie dekompozycji przestrzennej współczynnika Giniego. Wykonane badania wykazały relatywnie duże znaczenie nierówności, wynikających ze wzorców przestrzennych w kształtowaniu ogólnego poziomu nierówności dochodowych. Badania przeprowadzono na podstawie danych pochodzących z Banku Danych Lokalnych Głównego Urzędu Statystycznego w Warszawie.

Słowa kluczowe: nierówności przestrzenne, przestrzenna dekompozycja współczynnika Giniego, autokorelacja przestrzenna, dochody własne gmin

## SPATIAL NATURE OF INCOME INEQUALITY OF LOCAL GOVERNMENTS IN POLAND

The endogenous income sources of gminas are remarkably diverse, which results in economic and social inequalities. They may hinder local development, create social fragmentation and impede the implementation of the cohesion policy. Moreover, spatial units such as gminas often interact with each other demographically, socially or economically (e.g. in production or commuting). These interactions are reflected in, among others, the income status of gminas. Therefore, a growing body of local income inequality studies begins to consider spatial effects, which would account for the importance of the spatial location in shaping the overall inequality levels. The study aimed to identify the spatial effects that characterize inequality of the own income of Polish gminas, and of its most important components. It used the authors' proposed modification of the spatial inequality measure of Rey and Smith (2013) to assess the impact of spatial patterns on the level of the overall inequality. The measure provided an estimation of the share of inequality due to neighborhood interactions through a spatial decomposition of the Gini coefficient. The conducted research showed a relatively high relevance of spatial patterns in shaping the general level of income inequality. The study drew on data from the Local Data Bank of Statistics Poland.

Keywords: spatial inequality, spatial decomposition of the Gini coefficient, spatial autocorrelation, own income of gminas

**Agnieszka Jankowska**

Urząd Statystyczny w Zielonej Górze

Statistical Office in Zielona Góra

## **OCENA I ANALIZA SYTUACJI GOSPODARCZEJ W CZASIE PANDEMII COVID-19 NA PODSTAWIE WYNIKÓW BADAŃ KONIUNKTURY W WOJEWÓDZTWACH**

Celem referatu jest zaprezentowanie wyników badań koniunktury w wybranych województwach oraz oceny kondycji gospodarczej sektorów gospodarki, dla których realizowane jest badanie, w szczególnym okresie pandemii Covid-19. Wystąpienie, oprócz zagadnień metodologicznych badania, obejmie analizę wyników badania koniunktury oraz ocenę trafności prognoz w okresie pandemii. Badanie koniunktury gospodarczej dla województw realizowane jest od 2011 roku i stanowi uszczegółowienie wyników ogólnopolskich na poziomie terytorialnym. Ze względu na terminy badania, wskaźniki obliczane na podstawie wyników stanowią punkt wyjścia dla statystyki ilościowej. W referacie wykorzystano metody wielowymiarowej analizy porównawczej do oceny kondycji przedsiębiorstw w wybranych regionach. Informacje uzyskane z badania koniunktury na poziomie województw są źródłem wiedzy, która umożliwia przewidywanie zjawisk ekonomicznych w średnim okresie. Zawartość ankiet koniunktury gospodarczej pozwala na zbudowanie wskaźników ilościowych, jak i jakościowych, dzięki którym można ocenić bądź przewidzieć stan gospodarki na poziomie województw. Wyniki badania stanowią jeden z głównych elementów, na podstawie których budowana jest zarówno krótko-, jak i średniookresowa strategia rozwoju województw. Umiejętne przewidywanie przez przedsiębiorców przyszłości ma szczególne znaczenie dla analiz sytuacji gospodarczej i pozwala na podjęcie szybkich decyzji, dotyczących gospodarki w regionach.

Słowa kluczowe: koniunktura gospodarcza, COVID-19, wskaźniki

## **ASSESSMENT AND ANALYSIS OF THE ECONOMIC SITUATION DURING THE COVID-19 PANDEMIC BASED ON THE RESULTS OF BUSINESS TENDENCY SURVEYS IN VOIVODSHIPS**

The paper presents the research results of the business tendency in selected voivodships and the assessment of the economic condition of the sectors of the economy for which the research is carried out in the particular period of the Covid-19 pandemic. In addition to the methodological issues,

the study includes an analysis of the results of the business tendency survey and the assessment of the accuracy of forecasts carried out during the pandemic. The economic situation survey for voivodships has been carried out since 2011 and provides a more detailed description of the national results at a territorial level. Due to the timing of the survey, the performance indicators are the starting point for quantitative statistics. The paper uses methods of multidimensional comparative analysis to assess the condition of enterprises in selected voivodships. The information obtained from the survey of the economic situation at the voivodship level is a source of knowledge that enables forecasting economic phenomena in the medium term. The content of the economic situation surveys allows for the construction of quantitative and qualitative indicators, thanks to which it is possible to assess or predict the state of the economy at the voivodship level. The results of the study are one of the main elements on the basis of which both short- and medium-term development strategies for voivodships are built. A skilful forecasting prepared by entrepreneurs is of particular importance in the analyses of the economic situation and allows making quick decisions regarding the economy in the voivodships.

Keywords: business tendency, COVID-19, indicators

**Katarzyna Koczewska**

Uniwersytet Warszawski

University of Warsaw

## **PRZESTRZENNE UCZENIE MASZYNOWE – SZANSE DLA REGIONALISTYKI**

Celem referatu jest przedstawienie metod przestrzennego uczenia maszynowego oraz możliwych ścieżek jego rozwoju. Uczenie maszynowe przestrzenne stanowi alternatywę, jak i uzupełnienie dla dobrze ugruntowanej ekonometrii i statystyki przestrzennej. Istotą jest sposób uwzględniania informacji z sąsiedztwa i traktowania danych o różnej granulacji przestrzennej. Jest to nowy obszar analityczny i konieczna jest systematyka dostępnych metod. Zaprezentowane zostaną rozwiązania i wyzwania nowoczesnych analiz przestrzennych, uwzględniające jej potencjał i pułapki. Skatalogowane i skomentowane zostaną metody grupowania przestrzennego (dla lokalizacji i wartości, osobno i łącznie) do mapowania, bootstrap, walidacji krzyżowej, modelowania GWR etc. W odniesieniu do przestrzennych modeli nadzorowanego uczenia maszynowego omówione zostaną aspekty integracji danych przestrzennych, modelowania, trenowania modeli i predykcji w celu radzenia sobie z przestrzenną autokorelacją i dużymi zbiorami danych. Zostanie pokazane, że istnieje wiele konkurencyjnych rozwiązań, które dają bardzo zróżnicowane rezultaty. Można dostrzec, że trudno jest w prosty sposób zastąpić macierz wag przestrzennych oraz opóźnienie przestrzenne, typowe dla statystyki i ekonometrii przestrzennej. Z tego przeglądu metodologicznego wynika, że badacze łączą stare i nowe algorytmy i rozwiązania ilościowe, aby odpowiedzieć na wyzwania big data i skalowalności metod, przy uwzględnieniu specyfiki danych przestrzennych. To podejście jest kluczowe do lepszego wykorzystania danych satelitarnych, spektralnych, nocnych świateł, czy zdjęć, a także daje znaczące przewagi w predykcji. Jednak wciąż jest wiele do zrobienia w zakresie heterogeniczności przestrzennej i kierunkowości procesów, a także modelowania danych czasowo-przestrzennych. Do badań przyszłości należą metody przestrzenne 3D, modelowanie ruchu, przestępczości czy rozwoju, oparte na zdjęciach (np. z GoogleStreetView). Można także odpowiedzieć na zupełnie nowe wyzwania badawcze.

Słowa kluczowe: przestrzenne uczenie maszynowe, przegląd metodologiczny, autokorelacja przestrzenna, analizy regionalne

## SPATIAL MACHINE LEARNING – NEW OPPORTUNITIES FOR REGIONAL SCIENCE

The aim of the paper is to present the methods of spatial machine learning and to indicate its possible development paths. Spatial machine learning is an alternative and complement to the well-established spatial econometrics and statistics. The essence is how to take into account information from the neighbourhood and how to treat data with different spatial granularity. Spatial machine learning is a very new analytical area and it needs an overview of the available methods. The paper presents the solutions and challenges of modern spatial analyses, including their potential and pitfalls. The paper will catalogue and comment on spatial clustering methods (for locations and values, separately and jointly) for mapping, bootstrapping, cross-validation, GWR modeling etc. With regard to spatial supervised machine learning, aspects of spatial data integration, modeling, model training and prediction to deal with spatial autocorrelation and large datasets will be discussed. The paper shows that there are many competitive solutions that give very different results. One can see that it is difficult to easily replace the spatial weights matrix and the spatial lag typical of spatial statistics and econometrics. This methodological review shows that researchers combine old and new algorithms and quantitative solutions to meet the challenges of big data and method scalability, taking into account the specificity of spatial data. Spatial machine learning is also the key to a better use of satellite data, spectral data, night lights or photos, and gives significant advantages in prediction. However, much remains to be done in spatial heterogeneity, isotropy, and spatio-temporal data modeling. Future research includes 3D spatial methods, the modeling of traffic, crime, economic and social development based on photos (e.g. from Google StreetView). Modern methods, however, enable us to respond to completely new research challenges that were unattainable with the current quantitative instruments.

Keywords: spatial machine learning, methodological review, spatial autocorrelation, regional analysis



**Marcin Ćmielewski**

**Jakub Głowacki**

Uniwersytet Ekonomiczny w Krakowie  
Cracow University of Economics

**Michał Kudłacz**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

## **WATER CITY INDEX, RANKING EFEKTYWNOŚCI WYKORZYSTANIA ZASOBÓW WODY PRZEZ MIASTA W POLSCE**

Niniejsze badanie odnosi się do kwestii efektywności miejskich polityk wodnych w Polsce. Wynika to z faktu rosnącego znaczenia problematyki miejskiej polityki wodnej w Polsce oraz tego, że miasta są najchętniej wybieranym środowiskiem do życia człowieka. Celem badania było analiza i ocena efektywności gospodarowania zasobami wody przez 218 największych pod względem ludności miast w Polsce. Stanowi to punkt wyjścia do dalszych analiz nad remediami dla poszczególnych miast w zakresie miejskich polityk wodnych.

Water City Index (WCI2021) jest realizowanym od trzech lat rankingiem w zakresie gospodarowania zasobami wody przez miasta w Polsce, które podzielone są na trzy grupy: metropolie, pozostałe miasta na prawach powiatu i miasta średniej wielkości. Poprzez wskaźniki i odpowiednio dobrane wagi sklasyfikowano miasta, co pozwoliło na wyciągnięcie wniosków w odniesieniu do jakości zarządzania miastami oraz stworzenie rekomendacji w zakresie kierunków ich rozwoju.

Badanie zostało zrealizowane poprzez zastosowanie ok. 40 wskaźników ilościowych z różnych baz danych: BDL GUS, BDOT, dane IGWP oraz IMiGW PAN, a także na podstawie wyników ankiet skierowanych do urzędów miast i miejskich spółek wodociągowych wybranych miast w Polsce. Opracowano wskaźniki syntetyczne, które były podstawą do stworzenia rankingów miast w obszarach: „jakość życia”, „woda jako zagrożenie”, „woda dla energii i przedsiębiorczości”, a także „rekreacja i kultura”.

W rankingu WCI2021 najwyższymi sklasyfikowanymi metropoliami pod względem efektywności wykorzystywania zasobów wody są: Gdańsk, Kraków i Wrocław. Klasyfikacja dużych miast (miast na prawach powiatu nie będących metropoliami) kształtuje się następująco: Gliwice, Gorzów Wlkp., Świnoujście, natomiast w rankingu miast średniej wielkości najwyższymi sklasyfikowane zostały: Augustów, Lubliniec i Żywiec.

Autorzy WCI2021 prowadzą prace nad rankingiem, który będzie miał premierę w październiku 2022 r.

Słowa kluczowe: polityka rozwoju miast, miejska polityka wodna, efektywność wykorzystania zasobów, ranking miast w Polsce

## **WATER CITY INDEX, RANKING OF THE EFFECTIVENESS OF WATER RESOURCES USE BY CITIES IN POLAND**

This study addresses the issues of urban water policy. This is due to the growing importance of urban water policy issues in Poland and the fact that cities are the most-chosen environment for human life. In the context of climate change, firstly, we will be dealing with an increasing number of 'weather incidents' derived from climate change, and the consequences will be increasingly prolonged heat and droughts, and dangerous floods of all varieties. Secondly, the global water deficit increases with the growth of the population and wealth of societies, the escalation of climate extremes, the expansion of cities, and the exploitation and pollution of the environment. Thirdly, Poland's water resources are scarce. We are dealing with relatively small amounts of precipitation. Therefore, the aim of this study is to analyze the efficiency of cities in terms of water resource management. The Water City Index is a ranking in the field of water resources management which has been carried out for three years. The ranking includes 218 cities in Poland which are divided into three groups: metropolises, other cities with powiat rights and medium-sized cities. The study was carried out using the quantitative method. Water management was divided into four areas, for the measurement of which several dozen indicators obtained from various sources were used. By means of the indicators and appropriately selected weights, cities were ranked, which allowed for drawing conclusions as to the quality of city management and for creating recommendations regarding the directions of their development.

Keywords: urban development policy, urban water policy, resource efficiency, ranking of cities in Poland

**Małgorzata Markowska**

Uniwersytet Ekonomiczny we Wrocławiu  
Wroclaw University of Economics and Business

**Danuta Strahl**

Akademia WSB w Dąbrowie Górniczej  
WSB University, Dąbrowa Górnicza

**WSKAŹNIK AKTYWNOŚCI INNOWACYJNEJ PRZEDSIĘBIORSTW PRZEMYSŁOWYCH W WOJEWÓDZTWACH, W LATACH 2018–2020**

Celem pracy jest przedstawienie wyników oceny aktywności innowacyjnej przedsiębiorstw przemysłowych w polskich województwach. Aktywność ta jest mierzona wskaźnikiem agregatowym, który uwzględnia pięć obszarów: nakłady inwestycyjne na innowacje, wsparcie publiczne dla przedsięwzięć innowacyjnych, działalność innowacyjną, współpracę przedsiębiorstw w zakresie innowacji oraz efekty działalności innowacyjnej. Wagi dla poszczególnych obszarów i dla zmiennych je reprezentujących zostały ustalone metodą ekspertów. Pierwsza analiza przeprowadzona przy pomocy tego wskaźnika dotyczyła lat 2013–2015, a obecna obejmuje lata 2018–2020. Daje to możliwość zmierzenia stanu i oceny zmian jakie zaszły między tymi okresami. W pracy wykorzystano dane z badań przeprowadzonych i opublikowanych przez Główny Urząd Statystyczny. Lista zmiennych z ostatnich badań nie w pełni pokrywa się z listą badania pierwszego. Układ wag zaproponowany dla zmiennych w ramach poszczególnych obszarów może być nieznacznie zmieniony tak, aby zapewnić porównywalność dwóch okresów. Wskaźnik agregatowy liczony jest metodą klasyczną, z normalizacją na przedział (0;1), a wszystkie wykorzystane zmienne są stymulantami. Wyniki porównań wskazują na znaczną poprawę aktywności innowacyjnej przedsiębiorstw.

Słowa kluczowe: aktywność innowacyjna, przedsiębiorstwa przemysłowe, wskaźnik agregatowy

**INDUSTRIAL ENTERPRISES INNOVATION ACTIVITY INDEX FOR POLISH PROVINCES 2018–2020**

The aim of the paper is to present the result of Polish voivodships evaluation from the point of view of industrial enterprises activity. This activity is measured by a composite indicator, which covers five areas: investment in innovations, public support for innovations, innovation activity itself,

enterprise co-operation in the field of innovations and economic results of innovations. Weights for areas and variables were established by a panel of experts. The first analysis using the proposed index had been carried out for the 2013–2015 period, and the current one covers the years 2018–2020. It gives an opportunity for dynamic comparisons. The data for the research has been taken from Statistics Poland publications. The list of variables from the first research does not fully comply with the one from the last publication. The system of weights used is flexible enough to make the findings comparable. The Composite Index has been calculated according to a standard procedure with a normalization of variables onto a (0;1) interval. All variables were stimulants, satisfying (the bigger, the better) condition. The results of research show the improving innovation activity of Polish industrial enterprises.

Keywords: innovation activity, industrial enterprises, composite indicator

**Małgorzata Michalcewicz-Kaniowska**

Politechnika Bydgoska im. J.J. Śniadeckich  
Bydgoszcz University of Science and Technology

**Małgorzata Zajdel**

Politechnika Bydgoska im. J.J. Śniadeckich  
Bydgoszcz University of Science and Technology

**STATYSTYKA PUBLICZNA W OPINII JEDNOSTEK SAMORZĄDU  
TERYTORIALNEGO W WOJEWÓDZTWIE KUJAWSKO-POMORSKIM**

Ważność tematyki dotyczącej statystyki publicznej jako źródła danych, szczególnie jej wykorzystania w jednostkach samorządu terytorialnego, wynika z faktu odpowiedzialności władz za kreowanie właściwej polityki rozwoju społeczno-gospodarczego. Przygotowanie i wdrażanie m.in. spójnej strategii rozwoju wymaga dostępnych danych statystycznych Banku Danych Lokalnych (BDL), stanowiącego bazę danych o gospodarce, społeczeństwie i środowisku. Celem głównym badania była ocena poziomu wiedzy i stopnia wykorzystania BDL dla potrzeb planowania rozwoju regionalnego. Kolejny aspekt dotyczył oceny BDL w opinii badanych z punktu widzenia m.in. zawartości i przejrzystości prezentowanych danych, obsługi portalu, czy sposobu nawigacji serwisu. Pierwszy etap badań ankietowych przeprowadzono wśród wszystkich jednostek samorządu terytorialnego na terenie województwa kujawsko-pomorskiego, a odpowiedzi udzieliło 60% spośród wszystkich gmin objętych badaniem. Badanie miało charakter wyczerpujący. Analiza danych przeprowadzona została w oparciu o narzędzia statystyki opisowej. Badania m.in. wskazały, iż ponad 40% analizowanych gmin nie korzysta w ogóle z Banku Danych Lokalnych. Badani wskazywali, że dane zawarte w BDL wykorzystywane są głównie do opracowywania planów strategicznych i raportów. Wyniki przeprowadzonego badania pokazują, że zarówno poziom wiedzy o BDL, jak i o jego możliwościach, jest nadal bardzo niski. Oznacza to, że w woj. kujawsko-pomorskim ogólna dostępność BDL jest stosunkowo mało rozpowszechniona i wymaga dalszych działań edukacyjnych ze strony Głównego Urzędu Statystycznego.

Słowa kluczowe: Bank Danych Lokalnych, statystyka publiczna, rozwój regionalny, strategia rozwoju gminy

## THE VIEW OF THE LOCAL GOVERNMENT IN KUJAWSKO-POMORSKIE ON PUBLIC STATISTICS

Official statistics is an important source of information used especially in units of the local government, which is responsible for creating appropriate strategies for the socio-economic development of the region. The preparation and implementation of a coherent development strategy requires access to statistical data from the Local Data Bank (LDB), which contains information about the country's economy, society and the environment. The main aim of the research was to assess the awareness of the authorities of the information prepared by LDB for the benefit of local development planning. Another aim was to check how the respondents evaluate LDB with respect to its contents and the clarity of gathered data, the ease of the use and navigation of the LDB website. The first part of the survey was performed on all local government units in Kujawsko-Pomorskie. As many as 60% of all gminas in the voivodship responded to the survey. The study can be described as comprehensive. The authors used descriptive statistics tools to analyze the results. The study demonstrated that over 40% of the surveyed gminas do not make use of the information from the Local Data Bank. Moreover, the respondents claimed that data from LDB is used mainly to prepare strategic plans and reports. The results of the study prove that the awareness of LDB and its potential still remain very low. Therefore, it can be argued that the accessibility of the LDB remains scarce and requires further promotional initiatives by Statistics Poland.

Keywords: Local Data Bank, public statistics, local development, commune development strategy

**Dominika Rogalińska**

Główny Urząd Statystyczny  
Statistics Poland

**Marek Pieniążek**

Główny Urząd Statystyczny  
Statistics Poland

## **PRZESTRZENNE ZRÓŻNICOWANIE ŚWIADCZENIA WYBRANYCH USŁUG PUBLICZNYCH W KONTEKŚCIE NOWYCH PODZIAŁÓW TERENÓW WIEJSKICH W POLSCE**

Celem jest ukazanie przestrzennego zróżnicowania świadczenia usług publicznych w Polsce ze szczególnym uwzględnieniem terenów wiejskich przy użyciu danych z Systemu Monitorowania Usług Publicznych. Do analizy wybrano usługi obszarów: edukacja, lokalna polityka społeczna, kultura i rekreacja, ochrona środowiska, gospodarowanie nieruchomościami, inwestycje i budownictwo. Jako usługi publiczne traktowano wszelkie dobra materialne i niematerialne, za których dostarczenie odpowiada administracja publiczna. Zastosowano metody statystyki opisowej. Wyniki zaprezentowano w zestawieniach z wykorzystaniem: grup porównawczych utworzonych na podstawie kryterium administracyjno-rodzajowego, dochodowego i ludnościowego oraz w nowych podziałach terenów wiejskich, uwzględniających kryteria demograficzne i przestrzenne. Badanie wykonano na poziomie gmin na podstawie danych, opublikowanych w Systemie Monitorowania Usług Publicznych. Usługi publiczne prezentowane w SMUP są opisane wskaźnikami ilości, jakości, dostępności i efektywności. Dzięki temu dokonano analizy przestrzennej poszczególnych wymiarów opisu usług. Otrzymano wzorce zróżnicowania przestrzennego analizowanych usług. Zróżnicowanie ilości świadczonych usług odpowiada rozmieszczeniu ludności, natomiast zróżnicowanie jakości, dostępności i efektywności w większym stopniu związane są z poziomem dochodów gminy. Wyniki badania ukazują duże różnice pomiędzy miastami a obszarami wiejskimi. W obrębie obszarów wiejskich widać natomiast różnice pomiędzy gminami, znajdującymi się w obszarach oddziaływania dużych ośrodków miejskich, a gminami peryferyjnymi. Zastosowanie danych SMUP umożliwiło identyfikację gmin, uzyskujących ponadprzeciętne rezultaty w świadczeniu usług publicznych w poszczególnych grupach porównawczych.

Słowa kluczowe: usługi publiczne, statystyka regionalna, analiza przestrzenna, obszary wiejskie, SMUP

## **SPATIAL DIFFERENTIATION OF THE PROVISION OF SELECTED PUBLIC SERVICES IN THE CONTEXT OF NEW DIVISIONS OF RURAL AREAS IN POLAND**

The aim of the study is to show the spatial differentiation of the provision of public services in Poland, with particular emphasis on rural areas, using data from the Public Services Monitoring System (SMUP). The following services were selected for the analysis: education, local social policy, culture and recreation, environmental protection, real estate management, investments and construction. All tangible and intangible goods which the public administration is responsible for were treated as public services. Descriptive statistics methods were used. The results are presented in compilations with the use of comparative groups created on the basis of the administrative type, income and population criteria, and in new divisions of rural areas taking into account demographic and spatial criteria. The study was conducted at the level of gminas on the basis of data published in the SMUP. The public services presented in SMUP are described by indicators of quantity, quality, availability and efficiency. Thanks to this, a spatial analysis of individual dimensions of service description was performed. Patterns of spatial differentiation of the analyzed services were obtained. The differentiation in the number of the provided services corresponds to the distribution of the population, while the differentiation in quality, availability and efficiency are to a greater extent related to the level of the gmina's income. The results of the study show large differences between cities and rural areas. Within rural areas, there are differences between gminas located in the areas of influence of large urban centers and peripheral gminas. The use of SMUP data made it possible to identify gminas achieving above-average results in the provision of public services in particular comparative groups.

Keywords: public services, regional statistics, spatial analysis, rural areas, SMUP.



**Teresa Słaby**

Wyższa Szkoła Menedżerska w Warszawie

Warsaw Management University

**Artur Czech**

Wyższa Szkoła Menedżerska w Warszawie

Warsaw Management University

**TRANSPORT ZRÓWNOWAŻONY W CZASIE PANDEMII COVID-19 – UJĘCIE PRZESTRZENNE**

Transport jest podstawą dobrze funkcjonującej gospodarki i jej rozwoju. Działalność transportowa generuje zewnętrzne efekty pozytywne, ale i negatywne, które powodują problemy zdrowotne i zanieczyszczenie środowiska naturalnego. Zjawiska te nabrały szczególnego znaczenia w okresie pandemii Covid-19, która dodatkowo wzmogła odczuwanie negatywnych skutków transportu. Głównym celem pracy jest próba oceny w ujęciu geograficznym efektów transportu zrównoważonego przed i w czasie pandemii w poszczególnych województwach. W badaniu wykorzystano pozytywną metodę taksonomiczną. Ponadto, w procesie konstrukcji miary taksonomicznej zastosowano wielowymiarową medianę Webera. Przeprowadzoną analizę oparto na zbiorze cech wygenerowanych z Banku Danych Lokalnych Głównego Urzędu Statystycznego za lata 2018 i 2020 w Polsce. Wybór cech diagnostycznych był uwarunkowany dostępnością danych statystycznych. W wyniku przeprowadzonego badania otrzymano rankingi województw dla poszczególnych filarów zrównoważonego transportu, takich jak: środowiskowy, społeczny i ekonomiczny. W badaniu zastosowano autorski zbiór cech diagnostycznych. Zbiory te posłużyły do opisu poszczególnych łańdów. Następnie, syntetyczne mierniki cząstkowe posłużyły do konstrukcji miary zintegrowanej. Zastosowanie wielowymiarowej mediany Webera pozwoliło ponadto na uwzględnienie interakcji pomiędzy poszczególnymi cechami diagnostycznymi. Uwzględnione czynniki postrzegane jako reprezentantki poszczególnych filarów zrównoważonego systemu transportowego generują wzajemne zależności również w ujęciu ogólnym. Wykorzystanie mediany Webera jako miary statystycznej pozwoliło również na uodpornienie analizy na wpływ asymetrii rozkładów empirycznych wybranych cech diagnostycznych.

Słowa kluczowe: transport zrównoważony, filary/porządki, analiza taksonomiczna, pozycje województw, miernik syntetyczny

## SUSTAINABLE TRANSPORT DURING THE COVID-19 PANDEMIC – A SPATIAL APPROACH

Transport is the basis of a properly functioning economy and its development. Transport activity generates positive and negative external effects, which has implications for human health and environmental degradation. These phenomena gained particular importance during the Covid-19 pandemic, which additionally increased these negative effects of transport. The main aim of the study is to try to assess the effects of sustainable transport in a geographic approach before and during the Covid-19 pandemic in individual voivodships. For this purpose, the order taxonomic method was implemented. Moreover, the multidimensional Weber median was used in the process of the construction of a synthetic measure. The carried out analysis was based on a set of diagnostic variables drawn from the Local Data Bank of Statistics Poland for 2018 and 2020. As a result, rankings of voivodships were obtained for individual pillars of sustainable transport, such as environmental, social and economic. The study was based on an original set of diagnostic variables. These sets were used to describe individual orders of sustainable transport development. Then, synthetic partial measures were used to construct an integrated one. The application of the multidimensional Weber median allowed taking into account interactions among individual diagnostic variables. The factors taken into account, perceived as representatives of particular pillars of sustainable transport, generate mutual and not directly observable dependencies also in general terms. The implementation of the Weber median as a statistical measure also made the analysis more resistant to the influence of the asymmetry of the empirical distributions of selected diagnostic variables.

Keywords: transport sustainability, pillars/orders, taxonomic analysis, voivodships' positions, synthetic measure

**Agata Surówka**

Politechnika Rzeszowska

Rzeszow University of Technology

## **PROBLEMATYKA NAKŁADAJĄCYCH SIĘ OBSZARÓW WPŁYWU POLSKICH PORTÓW LOTNICZYCH W BADANIACH NAUKOWYCH**

W literaturze można znaleźć różne definicje obszarów ciężenia (izochrony, obszary wpływów). W prowadzonych badaniach przyjęto jedną z nich, zgodnie z którą jest to obszar w promieniu 100 km od portu. Przegląd literatury pokazuje również mnogość istniejących metod i narzędzi określania oraz prezentowania obszaru oddziaływania lotniska. Dostarczają one badaczowi różnorodnych informacji, począwszy od ogólnego zarysu na mapie, po sposoby identyfikacji jednostek statystycznych zlokalizowanych w zasięgu portów lotniczych. Brakuje jednak badań naukowych, dotyczących obszarów nakładania się izochron dwóch lub więcej lotnisk w Polsce. Postrzegana luka stała się bodźcem do podjęcia badań nad takimi zagadnieniami. Mając powyższe na uwadze, autorka postawiła sobie dwa cele. Pierwszym z nich jest identyfikacja powiatów położonych na obszarach oddziaływania dwóch lub więcej polskich portów. Drugim natomiast była próba odpowiedzi na pytanie, czy jednostki położone w nakładających się izochronach, charakteryzują się odmienną specyfiką? Jako narzędzia badawcze wykorzystano wybrane metody autokorelacji przestrzennej. Dzięki takiemu podejściu możliwe było określenie zależności przestrzennych badanych cech. Analizowaną kategorią był rynek pracy, który zdefiniowano za pomocą 27 wskaźników. Źródłem danych był Portal Geostatystyczny. Badanie przeprowadzono w ujęciu dynamicznym. Dzięki takiemu podejściu możliwe było przeprowadzenie analizy porównawczej dynamiki zmian badanych cech. Zastosowane podejście pozwoliło również na stwierdzenie występowania istotnych zależności przestrzennych w kształtowaniu wyróżnionych cech, a na podstawie tej wiedzy można m.in. zaprojektować określone wzorce reakcji w strefach oddziaływania badanych portów. Ze względu na założone cele szczególną uwagę zwrócono na jednostki statystyczne zlokalizowane w izochronach nakładających się obszarów wpływu.

Słowa kluczowe: rynek pracy, obszary oddziaływania transportu lotniczego, metody ilościowe, transport lotniczy, lotnisko

## THE ISSUE OF OVERLAPPING AREAS OF INFLUENCE OF POLISH AIRPORTS IN SCIENTIFIC RESEARCH

Various definitions of the areas of gravity (isochrones, areas of influence) can be found in the literature. One of them was adopted in the conducted research, according to which it is the area within 100 km from the port. The literature review also shows a multitude of methods and tools for determining and presenting the area of impact of an airport. They provide the researcher with a variety of information, ranging from a general outline on a map to ways of identifying statistical units located within the range of airports. However, there is a lack of scientific research on the areas of overlapping isochrones of two or more airports in Poland. The perceived gap has become a stimulus to undertake research on such issues. Bearing the above in mind, the author has set two goals. The first is the identification of powiats located in the areas of influence of two or more Polish ports. The second was an attempt to answer the question whether units located in overlapping isochrones are characterized by different specificity. Selected methods of spatial autocorrelation were used as research tools. Thanks to this approach, it was possible to determine the spatial relationships of the examined features. The analyzed category was the labor market, defined by means of 27 indicators. The source of data was the Geostatistics Portal. The study was carried out in a dynamic approach. Thanks to this approach, it was possible to carry out a comparative analysis of the dynamics of changes in the examined characteristics. The applied approach also allowed finding the presence of significant spatial relationships in shaping the distinguished features, and on the basis of this knowledge one can, inter alia, design specific patterns of reactions in the zones of influence of the studied ports. Due to the assumed objectives, special attention was paid to statistical units located in isochrones of overlapping areas of influence.

Keywords: labor market, areas of impact air transport, quantitative methods, air transport, airport

**Monika Wałaszek**

Urząd Statystyczny w Krakowie

Statistical Office in Kraków

**Michał Misygar**

Urząd Statystyczny w Krakowie

Statistical Office in Kraków

**DELIMITACJA POWIATÓW WOJEWÓDZTWA MAŁOPOLSKIEGO POD KĄTEM WPŁYWU PANDEMII NA RYNEK PRACY ORAZ WYKORZYSTANIE ŚRODKÓW POMOCY PUBLICZNEJ**

Głównym celem analizy jest wskazanie powiatów województwa małopolskiego najbardziej dotkniętych wpływem pandemii w obszarze rynku pracy, na podstawie wartości wskaźnika syntetycznego. Wybrane do modelu wskaźniki cząstkowe wykazały istotną statystycznie różnicę w wartościach dla lat trwania pandemii COVID-19 w porównaniu z latami poprzedzającymi wystąpienie pandemii. Dodatkowym celem opracowania jest przeanalizowanie dystrybucji środków pomocy publicznej na cele przeciwdziałania pozamedycznym skutkom pandemii w powiatach województwa małopolskiego. Pierwsza część artykułu zawiera ogólny opis działań władz lokalnych i centralnych, mających na celu zapobieganie negatywnemu wpływowi pandemii na gospodarkę krajową. W drugiej części artykułu autorzy skupiają się na analizie statystycznej wskaźników cząstkowych i budowie wskaźnika syntetycznego. Kolejna część publikacji poświęcona jest przeanalizowaniu dystrybucji środków pomocy publicznej w kontekście uzyskanego wskaźnika syntetycznego. Analizowane wskaźniki wykazały istotne statystycznie różnice pomiędzy latami przed wystąpieniem pandemii i w trakcie jej trwania. Wpływ pandemii na rynek pracy można było zaobserwować w zmianach wskaźników, dotyczących zwolnień z przyczyn zakładu pracy, wolnych miejsc pracy, podmiotów zawieszających działalność oraz wyników finansowych przedsiębiorstw w konkretnych sekcjach PKD. Zbudowany wskaźnik syntetyczny umożliwił delimitację powiatów najbardziej dotkniętych wpływem pandemii w obszarze rynku pracy. Analiza dystrybucji środków pomocowych nie wykazała istotnej korelacji Spearman z uzyskanym wskaźnikiem syntetycznym. Natomiast analiza korespondencji pokazała współwystępowanie utworzonych grup powiatów na bazie wskaźnika syntetycznego oraz grup powiatów utworzonych na bazie wielkości dystrybuowanych środków na jeden podmiot gospodarczy.

Słowa kluczowe: tarcza antykrzysowa, wskaźnik syntetyczny, delimitacja, rynek pracy, współwystępowanie

## **DELIMITATION OF POWIATS IN THE MAŁOPOLSKIE VOIVODSHIP IN TERMS OF THE IMPACT OF THE PANDEMIC ON THE LABOUR MARKET AND THE USE OF PUBLIC AID FUNDS**

The main aim of the analysis is to identify the powiats of Małopolskie Voivodship which were most affected by the pandemic in the area of the labour market on the basis of the synthetic indicator value. The partial indicators selected for the model showed a statistically significant difference in the values for the years of the COVID-19 pandemic compared to the years preceding the pandemic. An additional aim of the study is to analyse the distribution of public aid for the purposes of counteracting the non-medical effects of the pandemic in the Małopolskie Voivodship powiats. The first part of the article provides a general description of local and central authorities actions aimed at preventing the negative impact of the pandemic on the national economy. In the second part of the article, the authors focus on the statistical analysis of partial indices and the synthetic index construction. The next part of the publication is devoted to the analysis of the state aid funds distribution in the context of the obtained synthetic indicator. The analyzed indicators showed statistically significant differences between the years before and during the pandemic. The impact of the pandemic on the labour market could be observed in changes in the indicators related to layoffs due to the workplace, job vacancies, entities suspending their activity and financial results of enterprises in specific sections of the Polish Classification of Activities. The constructed synthetic index allowed the delimitation of powiats most affected by the pandemic in the area of the labour market. The analysis of aid funds distribution did not show a significant Spearman correlation with the obtained synthetic index. On the other hand, the analysis of the correspondence showed the coexistence of the created groups of powiats on the basis of the synthetic indicator and groups of powiats created on the basis of the amount of funds distributed per one economic entity.

Keywords: anti-crisis shield, synthetic indicator, delimitation, labour market, coexistence

**Iwona Bąk**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**Katarzyna Wawrzyniak**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**Maciej Oesterreich**

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

West Pomeranian University of Technology in Szczecin

**ANALIZA TEMPA PRZEMIAN SPOŁECZNO-GOSPODARCZYCH NA TERENACH WIEJSKICH W LATACH 2010–2020**

W publikacjach dotyczących sytuacji społeczno-gospodarczej na terenach wiejskich wskazywane są różnorodne czynniki, wpływające na problemy, z którymi muszą zmierzyć się mieszkańcy wsi. Najczęściej pogrupowane są one w czterech obszarach: demografia, odległość i infrastruktura oraz dostęp do podstawowych usług, edukacja i kapitał ludzki, rynek pracy. Analizując sytuację społeczno-gospodarczą na polskiej wsi nie można pominąć „obciążenia” dziedzictwem popegeerowskim, które wzmacnia problemy ludności wiejskiej. W artykule zaprezentowano wyniki badań nad procesem przemian na polskiej wsi w latach 2010-2020 z uwzględnieniem tych gmin wiejskich, w których występowały PGR-y. Badanie miało na celu uzyskanie odpowiedzi na kilka pytań – po pierwsze – jak silna jest zależność pomiędzy „obciążeniem” dziedzictwem popegeerowskim a poziomem wybranych wskaźników, charakteryzujących sytuację społeczno-gospodarczą ludności mieszkającej na tych terenach, po drugie – w których gminach, wraz z upływem czasu od momentu likwidacji PGR-ów, sytuacja społeczno-gospodarcza ulega poprawie, a w których nie i jakie mogą być przyczyny takiego stanu rzeczy, po trzecie – jak na szybkość przemian wpływa lokalizacja danej gminy w stosunku do centrów miejskich. W badaniu wykorzystano informacje udostępnione przez KOWR oraz wskaźniki, charakteryzujące sytuację społeczno-gospodarczą na terenach wiejskich z lat 2010-2020 r. Ponadto uwzględniono odległości geograficzne pomiędzy badanymi gminami a centrami miejskimi. Jako narzędzie badawcze wykorzystano wybrane metody wielowymiarowej analizy statystycznej. Wyniki potwierdziły, że na terenach popegeerowskich do dzisiaj istnieją problemy finansowe oraz trudności z odnalezieniem się ludności na rynku pracy. Również tempo przemian społeczno-gospodarczych na tych terenach jest znacznie wolniejsze niż na pozostałych terenach i są takie gminy, dla których nawet bliskość centrów miejskich nie przyczynia się do poprawy sytuacji społeczno-gospodarczej.

Słowa kluczowe: obszary wiejskie, przemiany społeczno-gospodarcze, metody wielowymiarowej analizy statystycznej

## ANALYSIS OF THE RATE OF SOCIO-ECONOMIC CHANGES IN RURAL AREAS IN 2010–2020

Publications on the socio-economic situation in rural areas indicate various factors influencing the problems faced by rural residents. Most often they are grouped in four areas: demography, distance and infrastructure and access to basic services, education and human capital, and the labor market. When analysing the socio-economic situation in the Polish countryside, one cannot ignore the 'burden' of the heritage of the former state-owned farms (PGR), which strengthens the problems of the rural population. The article presents the results of research on the process of changes in the Polish countryside in 2010–2020, including those rural gminas where PGRs existed. The limitation of the research to such gminas only was aimed at answering several questions – firstly, how strong is the relationship between the 'burden' of the former state-owned farm heritage and the level of selected socio-economic indicators, and secondly – in which gminas, over time since the liquidation of state-owned farms, the socio-economic situation improves, and in which not and what may be the reasons for such a state, thirdly - how the speed of changes is influenced by the location of a gmina in relation to city centers. The study used information provided by the National Support Centre for Agriculture (KOWR) and indicators characterizing the socio-economic situation in rural areas in the years 2010-2020. Moreover, geographical distances between the examined gminas and city centers were taken into account. Selected methods of multidimensional statistical analysis were used as a research tool. The results confirmed that there are still financial problems in the former state-owned farm areas and the difficulties of the population with finding their place on the labor market. Also, the rate of socio-economic changes in these areas is much slower than in other areas, and there are municipalities for which even the proximity of urban centers does not contribute to the improvement of their socio-economic situation.

Keywords: rural areas, socio-economic changes, methods of multidimensional statistical analysis



**Mariusz Cembruch-Nowakowski**

Uniwersytet Pedagogiczny w Krakowie  
Pedagogical University in Krakow

**Sławomir Dorocki**

Uniwersytet Pedagogiczny w Krakowie  
Pedagogical University in Krakow

**WYKORZYSTANIE METOD STATYSTYCZNYCH W GEOGRAFII  
SPOŁECZNO-EKONOMICZNEJ NA PODSTAWIE KWERENDY WYBRANYCH  
CZASOPISM NAUKOWYCH BAZY INTERNETOWEJ WEB OF SCIENCE**

Celem artykułu było sprawdzenie stopnia zróżnicowania wykorzystywanych metod i narzędzi statystycznych w artykułach naukowych z obszaru dyscypliny naukowej geografia społeczno-ekonomiczna. W celu weryfikacji założeń badawczych o intensywności i zróżnicowaniu stosowanych metod statystycznych w geografii społeczno-ekonomicznej ze względu na podejmowaną tematykę badawczą dokonano kwerendy użytych metod statystycznych w czasopismach o największej ilości artykułów z analizowanej tematyki w latach 2012–2021. W celu różnicowania podejmowanej tematyki wybrano trzy czasopisma reprezentujące odmienną tematykę badawczą. Są to czasopisma z obszaru geografii społecznej, geografii ekonomicznej i gospodarki przestrzennej.

W wyniku przeprowadzonej kwerendy zostało potwierdzone, że istnieje różnica w intensywności wykorzystania metod statystycznych pomiędzy analizowanymi czasopismami. Metody statystyczne najczęściej stosowane były w przypadku geografii ekonomicznej (dane statystyczne), natomiast rzadziej w geografii. Spośród wykorzystanych metod najczęściej stosowano wartości średnie oraz miary zależności. Powszechne było równocześnie stosowanie testów statystycznych oraz modeli opartych na regresji. Analizy uzupełnione były o wykresy, które powszechne były głównie w artykułach o tematyce gospodarczej, a rzadziej społecznej. W przypadku geografii ekonomicznej i gospodarki przestrzennej popularne były metody geostatystyczne (GIS). W przeprowadzonych badaniach wykorzystywano programy statystyczne, w tym głównie: SPSS, NVivo, GeoDa (geografia społeczna) oraz R, Arcgis, STATA (geografia ekonomiczna i gospodarka przestrzenna). Powszechne jest także wykorzystanie arkusza kalkulacyjnego Excel. Należy zwrócić uwagę, że metody statystyczne najczęściej stosowane są w przypadku badań prowadzonych przez zespoły badawcze złożone z przedstawicieli różnych dyscyplin naukowych realizujących duże projekty badawcze.

Słowa kluczowe: geografia, metoda statystyczna, regresja, wartości średnie, test statystyczny

## THE USE OF STATISTICAL METHODS IN SOCIO-ECONOMIC GEOGRAPHY ON THE BASIS OF A QUERY OF SELECTED SCIENTIFIC JOURNALS OF THE WEB OF SCIENCES DATABASE

The aim of the article is to check how varied statistical methods and tools used in scientific articles in the field of socio-economic geography are. In order to verify the research assumptions about the frequency of their use and about the diversity of statistical methods applied, and due to the specificity of the undertaken research subject, a query was made about the statistical methods applied in journals with the largest number of articles on the analysed subject in 2012–2021. In order to make the analysis more comprehensive, three journals representing different research topics (social, economic and spatial geography) were selected.

The conducted query demonstrated that the analysed journals used statistical methods with a different frequency. Methods of this kind were used most often in the case of economic geography (statistical data), and less frequently in social geography. Among the methods often used, mean values and measures of dependence were the most popular, as were statistical tests and models based mainly on regression. The above-mentioned analyses were often accompanied by charts, which were however more common in articles on economic, rather than social, issues. In the case of economic geography and spatial planning, geostatistical methods (GIS) were popular. Statistical tests most commonly used included SPSS, NVivo, GeoDa (social geography) and R, Arcgis, STATA (economic geography and spatial management), as well as Excel spreadsheets. Another conclusion of our research is that statistical methods are most often used in research conducted by teams composed of the representatives of various scientific disciplines who carry out large research projects.

Keywords: geography, statistical method, regression, mean values, statistical test

**Marek Walesiak**

Uniwersytet Ekonomiczny we Wrocławiu  
Wroclaw University of Economics and Business

**Grażyna Dehnel**

Uniwersytet Ekonomiczny w Poznaniu  
Poznań University of Economics and Business

**OCENA ZMIAN POZIOMU SPÓJNOŚCI SPOŁECZNEJ REGIONÓW POLSKI  
W LATACH 2010–2019 Z WYKORZYSTANIEM DYNAMICZNEJ TAKSONOMII  
RELATYWNEJ DLA DANYCH SYMBOLICZNYCH INTERWAŁOWYCH**

Celem referatu jest ocena zmian poziomu spójności społecznej regionów Polski w latach 2010–2019. Miary agregatowe odgrywają ważną rolę w analizie zjawisk społeczno-ekonomicznych. W literaturze opracowano różne koncepcje konstrukcji miar agregatowych. W zależności od stopnia kompensacji można je podzielić na: kompensacyjne, częściowo-kompensacyjne i niekompensacyjne. Przedmiotem zainteresowania jest metoda taksonomii relatywnej oraz jej modyfikacja dla ujęcia dynamicznego. Oba ujęcia dotyczą klasycznych zbiorów danych metrycznych. W artykule zaproponowano rozszerzenie zastosowania dynamicznej taksonomii relatywnej dla danych symbolicznych interwałowych. Zmienne symboliczne interwałowe opisują badane obiekty precyzyjniej niż dane metryczne. Dane metryczne mają charakter atomowy. Dane symboliczne interwałowe otrzymano w wyniku dwustopniowego gromadzenia danych. Najpierw zgromadzono dane klasyczne dotyczące spójności społecznej według powiatów, a następnie poddano je agregacji do poziomu regionów. Otrzymano dane symboliczne interwałowe obejmujące decyl 2 i 8 (60% obserwacji). Na podstawie danych symbolicznych interwałowych oraz dynamicznej taksonomii relatywnej przeanalizowano oraz zaprezentowano w formie graficznej zmiany poziomu spójności społecznej 17 regionów Polski w latach 2010–2019. W latach 2010–2019 odnotowano systematyczną poprawę w poziomie spójności społecznej 17 regionów Polski. Widoczne było wprawdzie zróżnicowanie w tym zakresie, ale dystans między regionami się zmniejszał. Dzięki zastosowaniu podejścia bazującego na danych symbolicznych pominięto obserwacje odstające.

Słowa kluczowe: spójność społeczna, taksonomia relatywna, miary agregatowe, ujęcie dynamiczne, dane symboliczne interwałowe

## ASSESSMENT OF CHANGES IN THE LEVEL OF SOCIAL COHESION OF POLISH REGIONS IN 2010–2019 WITH THE APPLICATION OF DYNAMIC RELATIVE TAXONOMY FOR INTERVAL-VALUED DATA

The aim of the paper is to assess changes in the level of social cohesion of Polish regions in 2010–2019. Aggregate measures play an important role in the analysis of socio-economic phenomena. Various concepts for the construction of aggregate measures have been developed in the source literature. Depending on the degree of compensation, three types of methods can be identified: compensatory methods, partially-compensatory methods, non-compensatory methods. The relative taxonomy method and its modification for the dynamic approach is the subject of this article. Both approaches relate to classic metric data sets. The article proposes an extension of the application of dynamic relative taxonomy for interval-valued data. Interval-valued variables describe the analysed objects more accurately than metric data. Classic data are of an atomic nature. In order to obtain interval-valued data, two-stage data collection proces was adopted. Firstly, the atomic metric data on social cohesion in Polish districts (LAU units) were collected (380 districts described by 22 variables), which were then aggregated to produce interval-valued data. The lower and upper limit of the interval for each province was determined on the basis of district-level data: 2nd and 8th deciles (60% of observations). Based on symbolic interval data and a dynamic relative taxonomy, changes in the level of social cohesion in 17 Polish regions in 2010–2019 were analysed and presented in a graphic form. In the years 2010–2019, there was a systematic improvement in the level of social cohesion in 17 regions of Poland. Despite the fact that the degree of this improvement was different in different regions, the distance between them was nevertheless narrowing. Due to the approach based on symbolic data, outliers were omitted.

Keywords: social cohesion, relative taxonomy, composite indicators, dynamic approach, interval-valued data

**Krzysztof Dmytrów**

Uniwersytet Szczeciński

University of Szczecin

**Beata Bieszk-Stolorz**

Uniwersytet Szczeciński

University of Szczecin

**METODA DYNAMIC TIME WARPING W OCENIE PODOBIENSTWA ZMIAN  
NA EUROPEJSKICH RYNKACH PRACY W KONTEKŚCIE ROZWOJU  
PANDEMII COVID-19**

Celem badania jest porównanie zmian wybranych wskaźników charakteryzujących europejski rynek pracy w okresie przed pandemią COVID-19 i w jej trakcie. W badaniu wykorzystano miesięczne i kwartalne dane z Eurostatu opisujące kształtowanie się podstawowych wskaźników dotyczących rynku pracy: stopy bezrobocia, wskaźnika zatrudnienia oraz wskaźnika aktywności zawodowej. Wszystkie wskaźniki analizowano dla populacji ogółem, kobiet, osób młodych oraz osób w wieku 55 lat i więcej. Podobieństwo szeregów zbadano za pomocą metody dynamicznego odwzorowania czasu (ang. dynamic time warping). Jednorodne grupy krajów ze względu na podobieństwo szeregów czasowych otrzymano za pomocą grupowania hierarchicznego. Analizę przeprowadzono w dwóch okresach: od początku 2018 r. do końca 2019 r. (okres przed pandemią) oraz od początku 2020 r. do chwili obecnej (okres pandemii).

Dla wszystkich wskaźników zarówno w okresie przed pandemią, jak i w okresie pandemii zidentyfikowano dwa, trzy lub cztery skupiska podobnych krajów pod względem dynamiki wskaźników. W większości przypadków podobieństwa w ramach każdego ze skupień były większe w okresie przed pandemią niż w okresie pandemii (wyjątek stanowiły wskaźnik aktywności zawodowej kobiet oraz wskaźnik zatrudnienia osób w wieku 55 lat i więcej). Skład skupień dla poszczególnych wskaźników w okresie przed pandemią i w okresie pandemii był na ogół odmienny. Można zauważyć wpływ pandemii COVID-19 na kształtowanie się podobieństw dynamiki wybranych wskaźników charakteryzujących rynek pracy. Przed pandemią skupienia wydają się być bardziej jednorodne niż w okresie jej trwania. Jest to wynikiem różnego stopnia rozwoju rynków pracy oraz ogólnie rozwoju gospodarczego, co wpływało na radzenie sobie ze skutkami kryzysu wywołanego pandemią.

Słowa kluczowe: europejski rynek pracy, pandemia COVID-19, Dynamic Time Warping, analiza skupień

## **DYNAMIC TIME WARPING METHOD IN ASSESSMENT OF SIMILARITY OF CHANGES IN EUROPEAN LABOUR MARKETS IN THE CONTEXT OF COVID-19 PANDEMIC DEVELOPMENT**

The aim of the study is the comparison of changes in the selected indicators used to characterise the European labour market before and during the COVID-19 pandemic. The study used monthly and quarterly data from Eurostat that reflected the shaping of the basic indicators which describe the labour market: the unemployment rate, the employment rate and the activity rate. All indicators were analysed from the point of view of the total population, women, young people and people aged 55+. The similarity of the series was analysed using the Dynamic Time Warping method. Homogeneous groups of countries based on the similarity of the time series were obtained using hierarchical clustering. The analysis was conducted over two periods: from the beginning of 2018 to the end of 2019 (pre-pandemic period) and from the beginning of 2020 to the present (pandemic period).

Two, three or four clusters of countries similar in terms of indicator dynamics were identified for each indicator both in the pre-pandemic and the pandemic period. In most cases, the similarities within each cluster were greater in the pre-pandemic period than in the pandemic period (the exceptions were: the female labour force participation rate and the employment rate of people aged 55+). The compositions of clusters for particular indicators in the pre-pandemic and pandemic periods were generally different. The impact of the COVID-19 pandemic on the shaping of the similarities of the dynamics of selected indicators characterising the labour market is clearly observable. Before the pandemic, the clusters seem to be more homogeneous than during the pandemic. This is the result of the European labour markets being at different stages of development and more generally, the European economies being developed to a different extent, which influenced particular countries' coping with the effects of the crisis caused by the pandemic.

Keywords: European labour market, COVID-19 pandemic, Dynamic Time Warping, cluster analysis

**Elżbieta Stańczyk**

Urząd Statystyczny we Wrocławiu; Uniwersytet Wrocławski  
Statistical Office in Wrocław; University of Wrocław

**Andrzej Dudek**

Uniwersytet Ekonomiczny we Wrocławiu  
Wrocław University of Economics and Business

**Halina Woźniak**

Urząd Statystyczny we Wrocławiu  
Statistical Office in Wrocław

**DRZEWY KLASYFIKACYJNE I REGRESYJNE W ANALIZIE JAKOŚCI BADAŃ  
STATYSTYCZNYCH**

Zmieniająca się rzeczywistość oraz potrzeby informacyjne odbiorców danych statystycznych, a także zakres wykorzystania statystyki w różnych dziedzinach stawiają przed statystyką publiczną nowe wyzwania i coraz to wyższe wymagania. Dotyczą one w dużej mierze utrzymania wysokiej jakości badań statystycznych. Jednym z komponentów jakości danych statystycznych jest dokładność. Analiza dokładności polega na rozpatrywaniu błędów losowych (błędów próby) i nielosowych (błędów niezwiązanych z próbą). Celem referatu jest zaprezentowanie zastosowania drzew klasyfikacyjnych i regresyjnych w ocenie jakości badań statystycznych w wymiarze dokładności. Model w postaci drzewa klasyfikacyjnego odwzorowuje przebieg procedury rekurencyjnego podziału zbioru badanych przedsiębiorstw na podzbiory maksymalnie jednorodnie z punktu widzenia wartości zmiennej objaśnianej utożsamianej z odmową wypełnienia kwestionariusza (tj. z podmiotowym błędem odpowiedzi jednostki). Jednocześnie model ten pozwoli na określenie, które ze zmiennych objaśniających (głównie wybranych charakterystyk przedsiębiorstw) w największym stopniu wpływały na zmienną objaśnianą. Dla zbiorowości przedsiębiorstw, które nie odmówiły udziału w badaniu kwestionariuszowym, skonstruowany zostanie model drzewa regresyjnego dla zmiennej objaśnianej utożsamianej z liczbą błędów popełnianych przez respondentów (tj. z przedmiotowymi błędami odpowiedzi jednostki). Model drzewa regresyjnego umożliwi dodatkowo predykcję zmiennej objaśnianej poprzez określenie relacji pomiędzy predyktorami (m.in. wybranymi charakterystykami przedsiębiorstw, pytaniami kwestionariusza, wybranymi cechami ankietera). Źródłem danych empirycznych będzie baza danych jednostkowych pozyskana w wyniku zrealizowanej w 2018 r. przez GUS w skali całego kraju, wśród 72,6 tys. przedsiębiorców, pracy badawczej „Uwarunkowania rozwoju przedsiębiorczości w sektorze MŚP”, finansowanej ze środków Programu Operacyjnego Pomoc Techniczna 2014–2022.

Słowa kluczowe: jakość badań statystycznych, przedsiębiorstwa, drzewa klasyfikacyjne i regresyjne, model predykcyjny

## CLASSIFICATION AND REGRESSION TREES IN THE QUALITY ANALYSIS OF STATISTICAL RESEARCH

The changing reality and information needs of the recipients of statistical data, likewise the degree to which statistics is used in various fields, all pose new challenges and ever-higher requirements for official statistics, which are most often related to maintaining high quality of statistical research. One of the indicators of statistical data quality is their accuracy. The analysis of accuracy is based on the consideration of random errors (sample errors) and non-random errors (non-sampling ones). The aim of the paper is to present the application of classification and regression trees to the assessment of the quality of statistical research, and more specifically, to the assessment of data accuracy. The model in the form of a classification tree will reflect the process of the recursive division of a set of the surveyed enterprises into sub-sets maximally homogeneous from the point of view of the value of the dependent variable, in this case identified with the refusal to fill in the questionnaire (i.e. with the 'subjective error' of the unit's answer). At the same time, this model will make it possible to determine which of the explanatory variables (selected characteristics of enterprises) had the greatest impact on the dependent variable. For the group of enterprises that did not refuse to participate in the questionnaire, a regression tree model will be constructed for the dependent variable, in this case identified with the number of errors made by respondents (i.e. with the 'subject errors' in the unit's answers). The regression tree model will additionally allow the prediction of the explained variable by determining the relationship between the predictors (e.g. selected characteristics of enterprises, selected characteristics of interviewers). The source of empirical data will be the unit database created as the result of the 'Determinants of the development of entrepreneurship in the SME sector' study carried out by Statistics Poland in 2018 among 72,600 entrepreneurs.

Keywords: quality of statistical research, enterprises, classification and regression trees, predictive model



**Jerzy Korzeniewski**

Uniwersytet Łódzki

University of Lodz

## **GRUPOWANIE TEKSTÓW POLSKOJĘZYCZNYCH POD WZGLĘDEM SENTYMENTU**

Celem referatu jest przedstawienie wyników badań dotyczących nienadzorowanej klasyfikacji tekstów polskojęzycznych pod względem opinii, czyli sentymentu tekstu. Badanie zostało przeprowadzone na 15 zbiorach dokumentów pochodzących ze stron internetowych podmiotów gospodarczych (hotele, firmy kurierskie, banki, apteki, księgarnie, inne sklepy), na których były one oceniane. Każdy zbiór składał się z kilku lub kilkunastu tysięcy dokumentów, które zawierały kilkadziesiąt tysięcy cech – terminów. Teksty były grupowane w dwie klasy: pozytywne i negatywne. Podstawową badaną grupą metod były metody partycjonujące, takie jak metoda  $k$ -średnich i jej modyfikacje oraz metoda PAM. Jakość uzyskanych grupowań była oceniana za pomocą precyzji oraz miary F1. Okazało się, że podstawowa metoda  $k$ -średnich przy zastosowaniu odległości cosinusowej zupełnie nie nadaje się do grupowania tekstów w korpusie tekstów z tak dużą liczbą cech. Zaproponowane zostały modyfikacje tej metody wykorzystujące skorelowanie częstościowe cech oraz dwie zdefiniowane listy terminów o ustalonym sentymencie. Listy terminów pomagają w ustaleniu sentymentu tekstu poprzez uśrednienie miary korelacyjnej sentymentu znalezionej dla wszystkich cech występujących w tekście ustalonego w na podstawie listy terminów. Ponadto miara siły korelacji z terminami z list może być wykorzystana do selekcji cech i uzyskania znacznie mniejszego zbioru cech pozwalającego na szybsze i lepsze grupowanie.

Słowa kluczowe: grupowanie tekstów, odległość cosinusowa, metoda  $k$ -średnich, współczynnik korelacji liniowej, sentyment tekstu

## **POLISH LANGUAGE TEXT SENTIMENT GROUPING**

The aim of the paper is to present the results of research concerning unsupervised text sentiment classification. The research was carried out on 15 sets of texts collected from internet sites of economic entities (hotels, delivery companies, banks, pharmacies, bookshops, and others), where they were evaluated. Each set consisted of several thousand documents containing several thousand feature-terms. The texts were classified into two classes: positive and negative. The basic group of

the examined methods were the k-means methods and their modifications as well as the PAM method. The quality of classifications was assessed by means of accuracy and the F1 measure. It turned out that the basic k-means method with the use of the cosine distance is inadequate for the classification of texts with such a big number of features. Some modifications of the basic method were proposed using a correlation measure of term frequencies and two lists of terms with strictly positive and negative sentiment. The two lists help to establish the text sentiment by averaging the correlation sentiment measure applied to all terms appearing in the text. Moreover, the measure of the correlation with the terms from the lists might be used for feature selection, thus reducing the number of features and allowing faster and more accurate classification.

Keywords: text grouping, cosine distance, k-means method, linear correlation coefficient, text sentiment

**Artur Markiel**

Politechnika Śląska

Silesian University of Technology

**Wojciech Kempa**

Politechnika Śląska

Silesian University of Technology

**ANALIZA SKUPIEŃ KRZYWYCH ZACHOROWAŃ III FALI COVID-19  
W POLSKICH POWIATACH**

Celem analizy był podział polskich powiatów na kategorie względem sposobu przejścia III fali COVID-19 przez ich populację oraz próba powiązania wyników klasyfikacji z innymi czynnikami. Na podstawie danych dotyczących dziennej liczby zachorowań na COVID-19 w powiatach w przeliczeniu na 10 000 mieszkańców obliczono średnie tygodniowe w każdym dniu. Krzywe średnich dla każdego powiatu przybliżono funkcją wielomianową. Następnie dla większości wyznaczonych funkcji wyznaczono pięć punktów charakterystycznych: maksimum funkcji oraz najbliższe mu maksimum i minimum pierwszej pochodnej i dwa maksima drugiej pochodnej. Na podstawie ich współrzędnych pogrupowano powiaty metodą  $k$ -średnich. Na końcu przeanalizowano średnie poziomy różnych czynników w otrzymanych kategoriach i wyznaczono różne modele klasyfikacji na ich podstawie. Zastosowanie pięciokrotnego sprawdzianu krzyżowego pozwoliło ustalić liczbę skupień na 5. Średnie wartości wspomnianych współrzędnych w każdym skupieniu pozwoliły wyznaczyć przybliżony kształt charakterystycznej dla danej grupy powiatów krzywej zachorowań. Jedynie pomiędzy częścią skupień wystąpiła znacząca różnica w gęstości zaludnienia i procencie ludności mieszkającej w miastach oraz przybliżonym procencie ludności zaszczepionej już po III fali COVID-19. Niemniej na podstawie tych czynników i struktury wiekowej ludności wyznaczono drzewa decyzyjne klasyfikujące większość kategorii fal z zadowalającą trafnością. Przeprowadzona analiza pozwoliła skonstruować modele predykcyjne przybliżonego kształtu krzywej zachorowań w danym powiecie. Modele te mogą stanowić obiekt dalszych rozważań w zakresie modelowania rozwoju epidemii.

Słowa kluczowe: analiza skupień, drzewo decyzyjne, krzywa zachorowań, metoda  $k$ -średnich, powiat

## CLUSTER ANALYSIS OF EPIDEMIC CURVES OF THE III COVID-19 WAVE IN POLISH POWIATS

The purpose of the analysis was to divide Polish poviats into categories according to the ways in which their populations dealt with the third wave of the COVID-19 pandemic and the attempt to link the resulting classifications with other factors. On the basis of the data about the new daily COVID-19 cases per 10,000 inhabitants in poviats, weekly averages were calculated every day. The curves of these averages for each powiat were approximated by a polynomial function. Then, for the majority of the generated functions, five characteristic 'points' were determined, i.e.: the maximum of the function together with maximum and minimum of its first derivative, and two maximums of its second derivative. On the basis of coordinates of these points, poviats were grouped using the k-means method. Finally, the mean levels of various factors in the obtained categories were analysed, and different classification models were determined on their basis. By using a V-fold cross-validation test, the number of clusters equal to five was determined. Mean values of the above-mentioned coordinates in each cluster made it possible to determine the approximate shape of the characteristic epidemic curve for a given group of poviats. Only a few clusters differed significantly from the rest in terms of the population density, the percentage of the population living in cities, and the approximate percentage of inhabitants vaccinated after the third wave of the COVID-19 pandemic. On the basis of these factors and the age structure of the population, decision trees which classify most of the wave categories with satisfactory accuracy were determined. The analysis allowed the construction of predictive models of the approximate shape of the epidemic curve in a given powiat. These models can be subject to further considerations in the modelling of the pandemic's development.

Keywords: cluster analysis, decision tree, epidemic curve, method of k-means, powiat

**Iwona Markowicz**Uniwersytet Szczeciński  
University of Szczecin**Paweł Baran**Uniwersytet Szczeciński  
University of Szczecin**IDENTYFIKACJA DETERMINANT CZASU TRWANIA RELACJI  
HANDLOWYCH PRZEDSIĘBIORSTW POLSKICH NA RYNKU UNIJNYM**

Handel międzynarodowy umożliwia poszerzenie oferty towarów i usług na rynku krajowym. Przyczynia się do rozwoju społeczno-gospodarczego i jest ważnym czynnikiem w podnoszeniu poziomu życia ludności. Celem pracy jest analiza czasu trwania relacji handlowych przedsiębiorstw polskich na rynku unijnym i identyfikacja czynników wpływających na długość tego czasu. Nasze badania dotyczą relacji handlowych na poziomie poszczególnych transakcji – określone: firma, produkt, typ transakcji, kraj wywozu/przywozu w ramach jednej, wybranej grupy towarowej (90 334 obserwacje). Dane, pochodzące z rejestru Intrastat i pozyskane z Krajowej Administracji Skarbowej, dotyczą wartości transakcji polskich przedsiębiorstw na rynku unijnym w okresie od stycznia 2008 r. do sierpnia 2020 r. Do szacowania funkcji przetrwania wykorzystano estymator Kaplana-Meiera, a do badania istotności różnic między krzywymi przetrwania – test Gehana-Wilcoxon. Relacje trwalsze w czasie dotyczą polskich eksporterów (WDT). Wśród wszystkich badanych transakcji wywozowych 1 rok przetrwało 38% przedsiębiorstw, 2 lata – 29%, a 5 lat – 18% (dla transakcji przywozowych udziały te były niższe: 25%, 17% i 9%, odpowiednio). Najtrwalsze relacje polskich przedsiębiorstw dotyczą firm z takich krajów, jak Wielka Brytania, Francja, Szwecja, Hiszpania, Portugalia i Czechy. Interesuje nas, czy relacje handlowe polskich przedsiębiorstw z kontrahentami unijnymi mają trwały charakter, czy są raczej sporadyczne. Obecne tendencje prezentowane w literaturze wskazują na odejście od badania specjalizacji krajów w handlu międzynarodowym na rzecz szerokiej, często spontanicznej wymiany dóbr. Badania skupiają się jednak zwykle na poziomie krajów i produktów. My badamy poszczególne transakcje dotyczące danej firmy i danego produktu łącznie, zatem prezentowane wyniki wypełniają lukę badawczą w tym zakresie.

Słowa kluczowe: wewnątrzspółnotowe relacje handlowe, analiza trwania, czas działalności gospodarczej na rynku UE

## IDENTIFYING DETERMINANTS OF THE DURATION OF TRADE RELATIONS OF POLISH ENTERPRISES ON THE EU MARKET

International trade makes it possible to expand the range of goods and services on the domestic market. It contributes to the socio-economic development and is an important factor in improving the standard of living of a population. The aim of this paper is to analyse time-spans in which Polish enterprises maintain trade relations within the EU market, and to identify factors that influence the length of these time-spans. Our study examines trade relations (at the level of individual transactions – i.e. the firm, the product, the transaction type, the country of exports/imports) within one selected group of products (90,334 observations). Data showing the values of transactions of Polish enterprises on the EU market in the period between January 2008 and August 2020 come from the Intrastat register and were obtained from the National Revenue Administration. Kaplan-Meier estimator was used to estimate the survival function, and Gehan-Wilcoxon test was used to test the significance of differences between the survival curves. Polish exporters tend to have more durable trade relations with the EU than importers (intra-Community suppliers ICS). 38% of exporters were present on the EU labour market for one year, 29% for 2 years, and 18% for 5 years (for importer, these figures were lower: 25%, 17% and 9%, respectively). Polish companies built most durable trade relations with partners from the following countries: Great Britain, France, Sweden, Spain, Portugal and the Czech Republic. We are interested in whether trade relations between Polish companies and the rest of the EU have a permanent character or are irregular. Current trends in the literature indicate a shift from studying countries' specialisations in international trade towards the focus on a broad, often spontaneous exchange of goods. At the same time, studies tend to focus on the country and the product level. We study individual transactions of a given firm and a given product jointly, thus filling the research gap in this area.

Keywords: intra-EU trade relationships, duration analysis, duration of economic activity on the EU market

**Andrzej Młodak**

Urząd Statystyczny w Poznaniu  
Statistical Office in Poznań

**Tomasz Józefowski**

Urząd Statystyczny w Poznaniu  
Statistical Office in Poznań

**Kamil Wilak**

Urząd Statystyczny w Poznaniu  
Statistical Office in Poznań

**Grzegorz Grygiel**

Urząd Statystyczny w Poznaniu  
Statistical Office in Poznań

**Tomasz Klimanek**

Urząd Statystyczny w Poznaniu  
Statistical Office in Poznań

## ZNACZENIE DANYCH SYNTETYCZNYCH W STATYSTYCE PUBLICZNEJ

Wystąpienie poświęcone zostanie rosnącej roli danych syntetycznych w statystyce publicznej na świecie, szczególnie w kontekście kontroli ujawniania danych. Stają się one bowiem coraz powszechniej stosowanym narzędziem ochrony tajemnicy statystycznej. Konstrukcja danych syntetycznych polega na generowaniu sztucznych mikrodanych z zakresu danego obszaru tematycznego lub badania statystycznego zachowujących w maksymalnym możliwym stopniu kształt zmiennych oraz rozkłady i związki występujące w odpowiednich danych oryginalnych. Mikrodane są następnie udostępniane użytkownikom zewnętrznym. Zaprezentowana zostanie istota takich danych, najważniejsze i coraz powszechniej stosowane metody ich konstrukcji (w tym pełna specyfikacja warunkowa – FCS, predykcyjne dopasowanie średniej – PMM, zaciemnianie danych zachowujące informację – IPSO czy różnicowanie prywatności – DP). Towarzyszyć temu będzie ukazanie ich zastosowań, a zwłaszcza użyteczności w minimalizacji ryzyka identyfikacji jednostki i straty informacji ponoszonej na skutek ochrony danych. Całość zilustrowana zostanie stosowną egzemplifikacją empiryczną, wskazującą najważniejsze korzyści i problemy tworzenia danych syntetycznych. Obejmie ona przykłady zastosowań danych syntetycznych w różnych krajach, jak również eksperymentalne wyniki syntetyzacji danych uzyskane przez zespół z Urzędu Statystycznego

w Poznaniu podczas HLG-MOS Synthetic Data Challenge, który odbył się w dniach 24–28 stycznia 2022 r., oraz dalszych własnych studiów w tym zakresie.

Słowa kluczowe: dane syntetyczne, metody syntezy, różnicowanie prywatności, kontrola ujawniania danych

## THE IMPORTANCE OF SYNTHETIC DATA IN OFFICIAL STATISTICS

The paper is devoted to the growing role of synthetic data in official statistics around the world, especially in the context of statistical disclosure control. Synthetic data are increasingly often being used as a way of protecting the confidentiality of statistical information. Synthetic datasets contain artificial microdata representing a given area of interest or a statistical survey, which have been generated in such a way as to resemble as closely as possible distributions and relationships observed in the original data. Such synthetic microdata are then made available to external users. The goal of the paper is to highlight the key features of synthetic data as well as the most common methods used for generating them (such as Fully Conditional Specification – FCS, Predictive Mean Matching – PMM, Information Preserving Statistical Obfuscation – IPSO or differential privacy – DP). We also present the potential areas of application, especially their usefulness in minimizing the risk of identifying individual units and information loss resulting from data protection. A number of empirical examples will be given to illustrate the main benefits of synthetic data as well as problems with generating them, including applications made in various countries and experimental results of a data synthesis exercise conducted by a team from the Statistical Office in Poznań during the HLG-MOS Synthetic Data Challenge, which took place on January 24–28, 2022.

Keywords: synthetic data, synthesizing methods, differential privacy, statistical disclosure control



**Michał Bernard Pietrzak**

Politechnika Gdańska  
Gdańsk University of Technology

**Marta Kuc-Czarnecka**

Politechnika Gdańska  
Gdańsk University of Technology

**PRZYDATNOŚĆ ANALIZY WRAŻLIWOŚCI W OCENIE KONSTRUKCJI  
TAKSONOMICZNEGO MIERNIKA ROZWOJU**

Mierniki syntetyczne są powszechnie używane w badaniach dotyczących oceny rozwoju zjawisk ekonomicznych, w szczególności oceny rozwoju społeczno-gospodarczego, poziomu życia, poziomu konkurencyjności czy stopnia innowacyjności. Wskaźniki te tworzone są przez międzynarodowe instytucje i są najczęściej oparte na zestawie kilkunastu lub nawet kilkudziesięciu zmiennych diagnostycznych. W takim przypadku istotnym problemem staje się kwestia oceny konstrukcji taksonomicznego miernika rozwoju. Problem ten sprowadza się do dwóch głównych aspektów: oceny doboru zmiennych diagnostycznych oraz wartości przyjętych wag. Narzędziem pozwalającym na rozważenie obydwu aspektów jest analiza wrażliwości dla miernika syntetycznego. Celem artykułu jest rozważenie przydatności analizy wrażliwości w ocenie konstrukcji taksonomicznego miernika rozwoju. Analiza wrażliwości przeprowadzona zostanie na przykładzie Global Competitiveness Index, który jest powszechnie stosowanym wskaźnikiem do klasyfikowania krajów pod względem ich konkurencyjności. Realizacja celu pozwoliła na ocenę stopnia poprawności konstrukcji wybranego miernika syntetycznego oraz propozycję jego ewentualnej korekty na podstawie analizy wrażliwości.

Słowa kluczowe: taksonomiczny miernik rozwoju, analiza wrażliwości, zmienne diagnostyczne, Global Competitiveness Index

**USEFULNESS OF THE SENSITIVITY ANALYSIS IN ASSESSING THE  
CONSTRUCTION OF A TAXONOMIC MEASURE OF DEVELOPMENT**

Synthetic indicators are commonly used in research concerning the evaluation of the development of economic phenomena, in particular the evaluation of socio-economic development, standard of living, the level of competitiveness or the degree of innovation. These indicators are created

by international institutions and are most often based on a set of several or even several dozen diagnostic variables. In such a case, the issue of assessing the construction of a taxonomic measure of development becomes a significant problem. This problem boils down to two main aspects: the assessment of the selection of diagnostic variables and the values of the adopted weights. A tool for considering both aspects is the sensitivity analysis for the synthetic measure. The aim of the article is to consider the usefulness of the sensitivity analysis in assessing the construction of a taxonomic measure of development. The sensitivity analysis will be carried out on the example of the Global Competitiveness Index, which is a commonly used index for classifying countries in terms of their competitiveness. The implementation of the goal allowed assessing the degree of the correctness of the structure of the selected synthetic indicator and a proposal for its possible correction based on the sensitivity analysis.

Keywords: taxonomic development measure, sensitivity analysis, diagnostic variables, Global Competitiveness Index

**Andrzej Sokołowski**

Uniwersytet Ekonomiczny w Krakowie  
Cracow University of Economics

**Małgorzata Markowska**

Uniwersytet Ekonomiczny we Wrocławiu  
Wroclaw University of Economics and Business

**Grzegorz Migut**

StatSoft Polska  
StatSoft Poland

**PRZEPLÝW FAL COVID-19 PRZEZ POLSKĘ**

Celem pracy jest przedstawienie wyników analizy przebiegu pięciu fal pandemii COVID-19 przez województwa Polski. Wykorzystano dane o liczbie zachorowań publikowane codziennie przez Ministerstwo Zdrowia, skorygowane dodatkowo o proporcjonalne przyporządkowanie osób o nieustalonym miejscu zamieszkania. Obliczono współczynnik zachorowalności na 100 000 ludności. W celu zniwelowania regularnych wahań okresowych o długości 7 dni, szeregi czasowe zostały wyrównane za pomocą średniej ruchomej. Następnie dla każdej pary województw szeregi te przesuwano względem siebie w zakresie od  $-20$  do  $+20$  dni i obliczano współczynniki korelacji. Maksymalny współczynnik korelacji wskazywał na opóźnienie (lub wyprzedzenie) przebiegu pandemii dla tych dwóch województw. Po przeanalizowaniu wszystkich par otrzymano macierz opóźnień/przyspieszeń. Stosunkowo łatwo było zidentyfikować województwa, w których zaczynały się poszczególne fale. Założono, że fala może przesuwać się do sąsiednich województw. Stwierdzono, że pierwsza fala rozpoczęła się w województwie dolnośląskim i dotarła do ostatniego województwa zajęło jej 20 dni. Druga fala rozpoczęła się w województwach warmińsko-mazurskim i zachodniopomorskim i ostatnie województwo osiągnęła w 16 dni. Trzecia fala zaczęła się w województwie opolskim i potrzebowała 15 dni na opanowanie całego kraju.

Słowa kluczowe: COVID-19, województwa, współczynnik zachorowalności, szeregi czasowe, opóźnienia

## COVID-19 WAVES FLOW OVER POLAND

The aim of the paper is to present the result of the analysis of five COVID-19 waves moving through the 16 voivodships of Poland. The data have been taken from official reports published daily by the Ministry of Health, adjusted for cases with unknown addresses. Daily crude incidence rates per 100,000 population were smoothed by a 7-day moving average. Then, for each pair of voivodships the time series were shifted backwards and forwards within the range of  $-20$  to  $+20$  days. The highest correlation coefficient identified the lag between given pairs of voivodships. Finally, the lag matrix allowed the identification of the initial voivodship and the movement of the pandemic, assuming that it can proceed to the neighboring geographical unit. The first wave started in Dolnośląskie and reached the last voivodship in 20 days. The second wave started in Warmińsko-Mazurskie and Zachodniopomorskie, and reached the last voivodship after 16 days. For the third wave it took 15 days, emerging from Opolskie Voivodship.

Keywords: COVID-19, voivodships, incidence rate, time series, lags

**Kateryna Zabarina**

Uniwersytet Warszawski

University of Warsaw

**ANALIZA ODLEGŁOŚCI DLA RÓŻNYCH TYPÓW DANYCH – CASE STUDY  
WOJEWÓDZTWA LUBELSKIEGO**

Kwestia odległości pozostaje ważna nie tylko w świetle pandemii, lecz także jako czynnik decydujący o lokalizacji firm, kupnie mieszkania, infrastrukturze transportowej itd. Obecnie dostępne dane wraz z rozwijającymi się metodami statystycznymi umożliwiają coraz bardziej zaawansowaną analizę odległości nie tylko na danych punktowych, lecz także liniowych i obszarowych. Celem niniejszego referatu jest pokazanie takiej analizy i omówienie jej potencjalnych zastosowań. Analiza odległości (w tym graficzna) na danych punktowych, liniowych (drogi krajowe i linie kolejowe) oraz obszarowych (województwa/powiaty) będzie wykonana w programie R. Dane pochodzą z Głównego Urzędu Statystycznego (dane punktowe lokalizacji firm i granice administracyjne województw i powiatów), Głównego Urzędu Geodezji i Kartografii (wzorzec liniowy dróg krajowych) i geofabrik.de (wzorzec liniowy linii kolejowych). Przedstawiona analiza pozwala na uzyskanie przejrzystych wyników w relatywnie krótkim czasie. Możliwa jest modyfikacja wyników (np. sortowanie, filtrowanie, mapowanie itd.) w celu dalszego zastosowania. Wyniki przedstawionych metod obliczania odległości na różnych typach danych nie tylko są użyteczne pod kątem podstawowej analizy, lecz także mogą być użyte do dalszej analizy, np. za pomocą modeli dla procesów punktowych czy też modeli przestrzennych.

Słowa kluczowe: dane punktowe, analiza odległości, wzorce liniowe, analiza przestrzenna

**DISTANCE ANALYSIS FOR VARIOUS TYPES OF DATA – CASE STUDY FOR  
LUBELSKIE VOIVODSHIP**

Distance-related issues remain important not only in the pandemic context, but also as a decisive factor in the location of firms, purchase of a flat, transport infrastructure, etc. Currently available data, along with developing statistical methods, allow a more and more advanced analysis not only on point but also line and areal data. The purpose of this study is to show this analysis and discuss its potential applications. The analysis of distances (including graphical) on point data, line data (national roads and railways) and areal data (voivodships/poviats) are performed in the R program. The data are from Statistics Poland (point data of firm locations and administrative region shapes),

the Head Office of Geodesy and Cartography (GUGIK) (roads line pattern) and geofabrik.de (railroad line pattern). The presented analysis allows for obtaining transparent results in a relatively short time. It is possible to modify the results (e.g. sorting, filtering, mapping, etc.) for further use. The results of the presented methods of calculating distances on various types of data are useful not only in terms of a basic analysis, but can also be used for further analysis, e.g. using models for point processes or spatial models.

Keywords: point data, distance analysis, line patterns, spatial analysis

**Elżbieta Zalewska**

Uniwersytet Łódzki

University of Lodz

## **ANALIZA SZKOLNICTWA WYŻSZEGO W PAŃSTWACH UNII EUROPEJSKIEJ**

Szkolnictwo wyższe odgrywa istotną rolę w Unii Europejskiej. Wysokiej jakości edukacja wpływa na różne aspekty życia, m.in. pracę, zarobki i innowacyjność. Poziom wykształcenia kraju zależy od wielu czynników, w tym od polityki edukacyjnej. Unia Europejska jest pod tym względem niejednorodną grupą krajów. Celem badań przedstawionych w referacie jest analiza szkolnictwa wyższego w krajach UE na podstawie charakterystyk dotyczących szkolnictwa wyższego oraz sytuacji absolwentów na rynku pracy. Wykorzystano dane Eurostat za 2020 r. Zidentyfikowano czynniki wpływające na konkurencyjność i jakość szkolnictwa wyższego, a także dokonano oceny podobieństw i różnic pomiędzy krajami, opierając się na utworzonym rankingu. W szczególności zwrócono uwagę na pozycję Polski w otrzymanych rankingach. W pracy zaprezentowano wyniki oraz porównanie klasyfikacji krajów UE pod względem szkolnictwa wyższego uzyskane za pomocą metod wielowymiarowych: analizy skupień oraz metod porządkowania liniowego Hellwiga. Dzięki zastosowaniu metody klasyfikacji możliwe było wyodrębnienie krajów najlepszych pod względem szkolnictwa wyższego. Należą do nich m.in. Wielka Brytania, Niemcy i Francja. Najśłabsze wyniki notują Grecja i Rumunia. W przeprowadzonych analizach Polska zajmuje średnią pozycję w rankingu, podobną do Słowenii i Czech.

Słowa kluczowe: szkolnictwo wyższe, ranking, Unia Europejska, analiza skupień, porządkowanie liniowe

## **ANALYSIS OF HIGHER EDUCATION IN THE EUROPEAN UNION COUNTRIES**

Higher education plays an important role in the European Union. High-quality education affects various aspects of life, including work, earnings, and innovation. A country's education level depends on many factors, including the education policy. In this respect, the European Union is a heterogeneous group of states. The aim of the research presented in the paper is to analyze higher education in the European Union countries on the basis of the characteristics of higher education and the situation of graduates in the labor market. Eurostat data from 2020 was used. Factors influencing the competitiveness and quality of higher education were identified, and the similarities and differences between countries were assessed, compiling a ranking of the EU countries. In particular, attention was paid to the position of Poland in the obtained rankings. The paper presents the results

and a comparison of the classification of EU countries in terms of higher education obtained with the use of multivariate methods: cluster analysis and Hellwig's linear ordering methods. The applied classification methods made it possible to distinguish the leading countries in terms of higher education. These countries include, among others Great Britain, Germany and France. The worst results are recorded by Greece and Romania. In the carried out analyzes, Poland occupies an average position in the ranking, similar to that of Slovenia and the Czech Republic.

Keywords: higher education, ranking, European Union, cluster analysis, linear ordering



**Robert Chrabąszcz**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**Jerzy Gruszka**

Urząd Statystyczny w Krakowie  
Statistical Office in Kraków

**EUROPEJSKI SYSTEM ZINTEGROWANYCH STATYSTYK NA TEMAT  
OCHRONY SOCJALNEJ (ESSPROS). POLSKA W PERSPEKTYWIE  
PORÓWNAWCZEJ**

Celem wystąpienia jest opisanie systemu ochrony socjalnej w Polsce w 2019 r. w szerszej perspektywie międzynarodowej statystyki społecznej, z wykorzystaniem metodologii Europejskiego systemu zintegrowanych statystyk o ochronie socjalnej (European System of Integrated Social Protection Statistics – ESSPROS). Do opisu danych o ochronie socjalnej wykorzystano definicje, klasyfikacje, jednostki statystyczne i zasady rachunkowości, zawarte w rozwijanej od lat 80. XX wieku metodologii ESSPROS. Zestaw przygotowanych w ten sposób informacji o wartości wydatków na ochronę socjalną jest zgodny z zagregowanymi danymi o ochronie socjalnej w systemie rachunków narodowych. Na potrzeby analizy, wydatki na ochronę socjalną w Polsce ujęto w ramach funkcji ESSPROS. Uzyskane informacje są porównywalne z danymi o ochronie socjalnej w innych państwach europejskich, cyklicznie publikowanymi w bazach Eurostatu. Oprócz informacji o wartości wydatków na świadczenia społeczne analizie poddano dane o liczbie beneficjentów rent i emerytur oraz informacje o wartości netto świadczeń społecznych, uwzględniających wpływ polityki fiskalnej państwa na ochronę socjalną. Dane o liczbie beneficjentów rent i emerytur oraz o wartości netto wydatków na ochronę socjalną stanowią istotne uzupełnienie opisu systemu ochrony socjalnej i są integralnym elementem metodologii ESSPROS. Zaprezentowana analiza porównawcza odmiennych systemów ochrony socjalnej opisuje specyfikę oraz cechy wspólne tych systemów w poszczególnych państwach. Umożliwia ona krajowym i międzynarodowym użytkownikom danych skuteczny monitoring aktywności państwa w tej sferze. Dane o ochronie socjalnej, w ujęciu zgodnym z metodologią ESSPROS, pozwalają decydom dokonywać identyfikacji określonych rodzajów potrzeb społecznych, ułatwiają kompleksową ocenę funkcjonowania systemu ochrony socjalnej oraz wpływają na prognozowanie przyszłych polityk społecznych, składających się na funkcję socjalną współczesnego państwa.

Słowa kluczowe: ochrona socjalna, systemy ochrony socjalnej, świadczenia społeczne, funkcje ochrony socjalnej, beneficjenci rent i emerytur

## EUROPEAN SYSTEM OF INTEGRATED SOCIAL PROTECTION STATISTICS (ESSPROS). POLAND IN COMPARATIVE PERSPECTIVE

The objective of the presentation is to describe the social protection system in Poland as of 2019 from a broader perspective of international social statistics, with the use of the European System of Integrated Social Protection Statistics (ESSPROS) methodology. Definitions, classifications, statistical units and accounting principles, included in the ESSPROS methodology, developed since the 1980s, were used in order to analyse the data on social protection. The produced set of information on the value of social protection expenditures corresponds with the aggregated data on social protection in the national accounts system. For the purpose of the analysis, social protection expenditures in Poland are assigned to the ESSPROS functions. The outcome data are comparable with the information on social protection in other European countries, periodically published in the Eurostat databases. Apart from the information on the value of expenditures on social benefits, the data on the number of pension beneficiaries and on the net value of social benefits, which consider the impact of the state's fiscal policy on social protection, were analysed. The data on the number of pension beneficiaries and on the net value of social protection expenditures are vital in the description of the social protection system and constitute an integral element of the ESSPROS methodology. The comparative analysis of different social protection systems emphasises the specific as well as common characteristics of those systems in individual countries. It facilitates national and international data users to effectively monitor the activity of the government in that area. The data on social protection from the perspective of the ESSPROS methodology allow the decision-makers to identify specific types of social needs, to facilitate a comprehensive evaluation of the social protection system and to impact the forecasting of social policy making which forms the social function of a contemporary state.

Keywords: social protection, social protection systems, social benefits, social protection functions, pension beneficiaries

**Czesław Domański**

Uniwersytet Łódzki

University of Lodz

## **POLSKIE TOWARZYSTWA STATYSTYCZNE FUNDAMENTEM STATYSTYKI PUBLICZNEJ**

Polskie Towarzystwo Statystyczne zostało utworzone przez grono krakowskich statystyków i ekonomistów 9 kwietnia 1912 r. w Krakowie. Powołane zostało w ściśle określonym celu: opracowania i wydania w jednym dziele rozproszonego materiału statystycznego, oświetlającego życie i rozwój polskiego społeczeństwa, rozdzielonego przez trzech zaborców. Dzieło pt. „Statystyka Polski” zostało wydane w Krakowie w 1915 r. Jest traktowane jako pierwszy polski rocznik statystyczny. Celem referatu jest przedstawienie osiągnięć autorów tego dzieła, które zostało przygotowane pod redakcją Adama Krzyżanowskiego i Kazimierza Władysława Kumanieckiego. Rocznik „Statystyka Polski” został porównany z pierwszym Rocznikiem Statystyki Rzeczypospolitej Polskiej wydanym - cz. I ukazała się w 1921 r. a cz. II w 1922 r. „Statystyka Polski” stanowi fundament polskiej statystyki publicznej, obejmujący informacje o Polsce od początku XIX wieku do I wojny światowej, i gwarantuje tym samym ciągłość narodowej statystyki.

Słowa kluczowe: rola statystyki, rozwój myśli statystycznej, prekursorzy statystyk

## **POLISH STATISTICAL SOCIETIES AS THE FOUNDATION OF OFFICIAL STATISTICS**

The Polish Statistical Society was established by a group of Kraków statisticians and economists in Kraków on April 9, 1912. This society was established for a strictly defined purpose: to prepare and publish, in one work, a dispersed statistical material illuminating the life and development of the Polish society divided by three partitioners. The work entitled 'Polish Statistics' was published in Kraków in 1915 and is treated as the first Polish statistical yearbook. The aim of the paper is to present the achievements of the authors of this work, which was prepared under the editorship of Adam Krzyżanowski and Kazimierz Władysław Kumaniecki. The Yearbook 'Statistics of Poland' has been compared with the first Yearbook of Statistics of the Republic of Poland, published in 1921 (part I) and part II in 1922. 'Polish Statistics' is the foundation of Polish official statistics, covering information about Poland from the beginning of the 19th century to World War I, and thus, guarantees the continuity of national statistics.

Keywords: the role of Polish statistics, development of statistical thought, precursors of statistics

**Agnieszka Pobłocka**

Uniwersytet Gdański

University of Gdańsk

## **POLSKI RYNEK UBEZPIECZEŃ W LATACH 2004–2020**

Celem pracy jest zbadanie sytuacji i ocena rozwoju polskiego rynku ubezpieczeń w latach 2004–2020 na tle innych krajów UE. W tym celu zdefiniowano pytania badawcze: Jak kształtuje się polski rynek ubezpieczeń na tle innych państw? Jak się on zmieniał i rozwijał od momentu wejścia Polski do Unii Europejskiej? I jak w przyszłości będzie się on kształtował? Jakie wyzwania czekają polskich ubezpieczycieli w najbliższym czasie? W badaniu wykonano analizę statystyczną (analizę szeregów czasowych – analizę dynamiki i analizę funkcji trendu oraz wielowymiarową analizę porównawczą (porządkowanie liniowe, analizę czynnikową) podstawowych charakterystyk ubezpieczeniowych oraz wskaźników rozwoju ubezpieczeń. Zakłady ubezpieczeń pełnią ważną rolę inwestorów w gospodarce ekonomicznej danego Państwa i świata. Ich łączna wartość portfela inwestycyjnego w Europie w 2019 r., jak podaje Insurance Europe, wynosiła aż 10,43 bln Euro, a łączna składka przypisana brutto na krajowych rynkach była równa 1,25 bln Euro. Największy udział w składce przypisanej brutto odnotowano w Wielkiej Brytanii (21 proc.), Francji (18 proc.) i Niemczech (17 proc.). Polskę notowano na 14 pozycji (1 proc.). Największy rozwój rynku według wskaźnika penetracji (składka przypisana brutto do PKB) na rynkach krajowych odnotowano w Finlandii (10,8 proc.), Danii (10,7 proc.) i Wielkiej Brytanii (10,4 proc.). A największy rozwój rynku według wskaźnika gęstości (składka przypisana brutto na 1 mieszkańca) na rynkach krajowych odnotowano w Szwajcarii (EUR 6,156), Danii (EUR 5,731), Luksemburgu (EUR 5,354). W Polsce wskaźnik penetracji był niski (2,7 proc.), a wskaźnik gęstości nieduży (EUR 373). Polski rynek ubezpieczeń w ciągu badanego okresu zmieniał się istotnie w czasie. Jednak, czy można Polskę zakwalifikować do kraju wysoko rozwiniętego względem ubezpieczeń, czy nadal jest to kraj rozwijający się?

Słowa kluczowe: ubezpieczenia osobowe i majątkowe, rynek ubezpieczeń, rozwój rynku, wskaźniki rozwoju rynku ubezpieczeń, składka przypisana brutto, gęstość ubezpieczeń, wskaźnik penetracji rynku

## **POLISH INSURANCE MARKET IN 2004–2020**

The aim of the study is to examine the situation and assess the development of the Polish insurance market in the years 2004–2020 compared to other EU countries. For this purpose, research questions were defined: How is the Polish insurance market shaped compared to other countries? How has it changed and developed since Poland joined the European Union? How will it be shaped in the future?

What challenges await Polish insurers in the near future? The study presents a statistical analysis (time series analysis – dynamics analysis and trend function analysis and multivariate comparative analysis (linear ordering, factor analysis)) of the basic insurance characteristics and insurance development indicators. Insurance companies play an important role of investors in the economic management of a given country and the world. Their total investment portfolio in Europe in 2019, according to Insurance Europe, was €10.43 trillion and their total gross direct premiums written on domestic markets was equal to €1.25 trillion. The largest share of the gross direct premiums written was recorded in the United Kingdom (21%), France (18%) and Germany (17%). Poland was listed on the 14th position (1%). The largest market development by penetration index (total gross written premiums to GDP) on domestic markets was recorded in Finland (10,8%), Denmark (10,7%) and the United Kingdom (10,4%). The largest market development by density indicator (total premiums per inhabitant) on domestic markets was recorded in Switzerland (EUR 6,156), Denmark (EUR 5,731) and Luxembourg (EUR 5,354). In Poland, the penetration indicator was low (2.7%) and the density index was low (EUR 373). The Polish insurance market changed significantly over time during the period under examination. However, can Poland be classified as a highly developed country in terms of insurance, or is it still a developing country?

Keywords: life and non-life insurance, insurance market, market development, indicators of the development of the insurance market, premium, density, penetration

**Agata Szczukocka**

Uniwersytet Łódzki

University of Lodz

**Elżbieta Zalewska**

Uniwersytet Łódzki

University of Lodz

**UDZIAŁ POLSKICH STATYSTYKÓW W KONGRESACH MIĘDZYNARODOWYCH**

Na początku lat 40. XIX stulecia przyjął się zwyczaj zbierania uczonych jednej gałęzi nauki lub kilku pokrewnych na sympozjach lub kongresach. Celem tych spotkań było ustalenie metod poszukiwania oraz rozwiązywania problemów tej gałęzi. Pierwszy Kongres Statystyczny odbył się w Brukseli w 1853 r. Następne odbyły się w Paryżu w 1855 r., w Wiedniu w 1857 r., w Londynie w 1860 r., w Berlinie w 1863 r., we Florencji w 1867 r., w Hadze 1869 r., w Petersburgu w 1872 r. i w Budapeszcie 1876 r. Celem referatu jest przedstawienie dorobku polskich statystyków, często reprezentujących inne kraje, na kongresach statystycznych w latach 1853-1876. Ich udział pozwolił zachować ciągłość polskiej myśli statystycznej w okresie rozbiorów Polski, z pewnością inspirował do tworzenia instytucji statystycznych – w tym Polskiego Towarzystwa Statystycznego w 1912 r., a także Głównego Urzędu Statystycznego. Przedstawimy osiągnięcia kilku wybitnych polskich uczonych, którzy wnieśli istotne postulaty, dotyczące metod i organizacji międzynarodowej statystyki. August Cieszkowski (1814–1894) na Kongresie w Paryżu w 1855 r. postulował m.in. ujednoczenie miar w statystyce międzynarodowej. Postulat ten został wprowadzony podczas kolejnych kongresów. Tadeusz Piłat był (1844–1923) m.in. współzałożycielem Międzynarodowego Instytutu Statystycznego i pierwszym profesorem wśród jego członków zwyczajnych. Witold Załęcki (1836–1908) był uczestnikiem Kongresu Statystycznego w 1876 r. w Budapeszcie. Józef Kleczyński (1841–1900) to m.in. autor obszernego artykułu pt. „Międzynarodowy Instytut Statystyczny”, oceniającego wszechstronnie kongresy Statystyczne oraz określającego dalszy rozwój Międzynarodowego Instytutu Statystycznego. Statystycy polscy w pierwszych organizacjach międzynarodowej zaznaczyli trwałe ślady swojego dorobku i umocnili pozycję polskiej statystyki na arenie światowej.

Słowa kluczowe: rola statystyki, rozwój myśli statystycznej, prekursorzy statystyki

## PARTICIPATION OF POLISH STATISTICIANS IN INTERNATIONAL CONGRESSES

At the beginning of the 1840s, it was customary to gather scholars of one or more related branches of science at symposia or congresses. The purpose of these meetings was to establish methods of searching and solving problems in this branch. The first Statistical Congress was held in Brussels in 1853. The next ones were held in Paris in 1855, in Vienna in 1857, in London in 1860, in Berlin in 1863, in Florence in 1867, in the Hague in 1869, in St. Petersburg in 1872 and in Budapest in 1876. The aim of the paper is to present the achievements of Polish statisticians, often representing other countries, at statistical congresses in the years 1853–1876. Their participation allowed the continuity of the Polish statistical thought during the Partitions of Poland; it certainly inspired the creation of statistical institutions, including the Polish Statistical Association in 1912 and the Central Statistical Office. We present the achievements of several outstanding Polish scientists who made important postulates on the methods and organization of international statistics. August Cieszkowski (1814–1894) at the Congress in Paris in 1855, postulated, the standardization of measures in international statistics. This postulate was introduced during subsequent congresses. Tadeusz Piłat (1844–1923) was the co-founder of the International Statistical Institute and the first professor among its ordinary members. Witold Załęcki (1836–1908) was a participant of the Statistical Congress in 1876 in Budapest. Józef Kleczyński (1841–1900) was the author of an extensive article entitled ‘The International Statistical Institute’, comprehensively evaluated statistical congresses and determined the further development of the International Statistical Institute. Polish statisticians in the first international organizations made a solid mark of their achievements and strengthened the position of Polish statistics internationally.

Keywords: the role of Polish statistics, development of statistical thought, precursors of statistics