



**Opracowanie metodologii i oszacowanie
liczby pracujących w gospodarce narodowej
według miejsca zamieszkania i głównego miejsca
pracy na poziomie NTS 4,
stopy bezrobocia rejestrowanego na poziomie NTS 5
oraz miar wynagrodzeń brutto na poziomie NTS 4**

*Development of a methodology
and estimation of the number of employed persons
in the national economy by place of residence and main
workplace at NTS level 4,
the registered unemployment rate at NTS level 5
and measures of gross wages and salaries at NTS level 4*

Prace studialne

Study works

Warszawa 2018

Warsaw 2018

ISSN

ISBN



Główny Urząd Statystyczny
Statistics Poland

Prace studialne

Study works

**Opracowanie metodologii i oszacowanie
liczby pracujących w gospodarce narodowej
według miejsca zamieszkania i głównego miejsca
pracy na poziomie NTS 4,
stopy bezrobocia rejestrowanego na poziomie NTS 5
oraz miar wynagrodzeń brutto na poziomie NTS 4**

***Development of a methodology
and estimation of the number of employed persons
in the national economy by place of residence and main
workplace at NTS level 4,
the registered unemployment rate at NTS level 5
and measures of gross wages and salaries at NTS level 4***

Raport końcowy

Final report

Warszawa 2018

Warsaw 2018

Nazwa jednostki opracowującej raport

The name of the unit developing the report

Centrum Badań i Edukacji Statystycznej GUS

Centre for Statistical Research and Education

Kierownik projektu

Project leader

Dominik Śliwicki

Zespół autorski

Editorial team

Joanna Wierzińska, Wiesława Gierańczyk, Bożena Białasik, Agata Cendrowska, Aleksandra Czapska, Małgorzata Garnuszek, Klaudia Gietka-Cichańska, Ewa Jankowska, Katarzyna Jaszczuk, Katarzyna Klamrowska, Ewelina Kodan, Leszek Kozłowski, Aneta Kuchta, Olga Lewandowska, Dorota Malicka, Bożena Matecka, Zbigniew Osypiuk, Michał Pacuski, Iwona Pasierowska, Agnieszka Piórek, Małgorzata Potocka, Jacek Pruski, Agnieszka Ratajska, Alicja Redelbach, Cezary Rogoś, Ewa Rupniewska, Maciej Ryczkowski, Małgorzata Saroska, Aleksandra Schreiber, Ewa Seweryn, Małgorzata Soból, Paweł Stopiński, Renata Strużyńska, Ewa Tuszyńska, Małgorzata Winnicka, Bartosz Wilk, Marzenna Anna Wiśniewska, Anna Wolska, Dominik Ziółkowski, Wioletta Zwara, Teresa Żelazny

ISBN

ISSN

Raport dostępny na stronie

Report available on website

<http://www.stat.gov.pl>

Przy publikowaniu danych GUS prosimy o podanie źródła

When publishing Statistics Poland data – please indicate the source

Druk publikacji współfinansowany ze środków Unii Europejskiej w ramach projektu „Statystyka dla polityki spójności. Wsparcie systemu monitorowania polityki spójności w perspektywie finansowej 2014-2020 oraz programowania i monitorowania polityki spójności po 2020 roku”

Przedmowa

Pozyskiwanie danych z rejestrów administracyjnych jest ogólnosięciowym trendem w podejściu do realizacji badań statystycznych. Wszelkie zagadnienia związane z pracami nad wykorzystaniem administracyjnych zbiorów danych należy traktować jako wyzwanie zarówno w zakresie metodologicznym, jak i programistycznym. Wyzwanie w obszarze metodologii dotyczy przede wszystkim wypracowania takich metod, w wyniku zastosowania których dane przechowywane w rejestrach administracyjnych na ogół do celów ewidencyjnych, zostaną przekształcone w dane statystyczne. Wyzwaniem jest również oprogramowanie wypracowanych metod w narzędziach do automatycznego przetwarzania danych.

Niewątpliwie wartością dodaną z wykorzystania danych rejestrowych w badaniach statystycznych jest możliwość prezentowania danych na niskich poziomach agregacji przestrzennej - niższych niż obecnie prezentowane dane jako wynikowe informacje statystyczne. Ponadto daje to możliwość ograniczenia sprawozdawczości podmiotów gospodarki narodowej, co z pewnością przełoży się na koszty badań. Podkreślić należy również, że pewne dane, które dotychczas były pozyskiwane w ramach badań reprezentacyjnych, mogą stanowić rezultat badania pełnego opartego na rejestrze administracyjnym.

Celem niniejszej pracy badawczej było opracowanie metodologii i oszacowanie danych dot. liczby pracujących w gospodarce narodowej według miejsca pracy oraz miejsca zamieszkania na poziomie powiatów, stopy bezrobocia rejestrowanego na poziomie gmin oraz miar wynagrodzeń brutto na poziomie powiatów. Uzyskane dane stanowią cenne uzupełnienie danych obecnie publikowanych przez służby statystyki publicznej. Przejawia się to w możliwości prezentowania informacji na niższych niż dotąd poziomach agregacji terytorialnej, według nowych wymiarów (np. według grup wiekowych, płci, miejsca zamieszkania), czy oszacowaniu dla całej gospodarki narodowej nowych mierników (np. miar pozycyjnych wynagrodzeń brutto). W pewnej części uzyskane z rejestrów administracyjnych dane z powodzeniem mogą zastąpić te, które obecnie pozyskiwane są w ramach badań statystycznych. W tym zakresie praca badawcza stanowi innowacyjne podejście do metodyki prowadzenia badań statystycznych. Niemniej jednak wyniki prezentowane w niniejszym raporcie należy traktować jako eksperymentalne i niemające charakteru oficjalnych statystyk publikowanych przez służby statystyki publicznej.

Przekazując Państwu niniejszy raport zwracamy się z prośbą o uwagi i sugestie co do jego treści, jak również do uzyskanych wyników i wniosków z nich płynących. Każda wskazówka będzie stanowić istotny wkład w rozwój wypracowanej metody badawczej oraz zakresu danych pozyskiwanych z rejestrów administracyjnych.

Preface

Acquiring data from administrative registers is a global tendency when it comes to approaches to conducting statistical surveys. All matters related to works on the use of administrative datasets should be regarded as a challenge both in methodological and programming aspects. A challenge in the methodological area relates primarily to developing such methods, by means of which data stored in administrative registers, usually for record-keeping purposes, will be transformed into statistical data. Designing software for the methods developed, that is automatic data processing tools is also a challenging task.

Undoubtedly, the added value resulting from the use of administrative register data in statistical surveys is the possibility to present data at low spatial aggregation levels – lower than data currently presented as statistical information. Moreover, it makes it possible to reduce statistical data reporting obligations for the national economy entities, which certainly will translate into lower costs of surveys. It should be also emphasised that some types of data, which have so far been collected from representative surveys may be gathered through census surveys based on administrative registers.

The objective of this research work was to develop a methodology and estimate data concerning the number of employed persons in the national economy by workplace and place of residence at the powiat level, the registered unemployment rate at the gmina level as well as measures of gross wages and salaries at the powiat level. The data obtained constitutes a valuable complement to the data presently being published by official statistical services. This is particularly reflected in the possibility to present information at territorial aggregation levels lower than hitherto and broken down by new criteria (for instance age groups, sex or place of residence) or estimating new statistical measures for the whole national economy (for example measures of positions for gross wages and salaries). To some extent data retrieved from administrative registers may successfully replace that collected from statistical surveys. In this respect the research work may be regarded as an innovative approach to the methodology of conducting statistical surveys. However, the results presented in this report should be regarded as experimental and not as official statistics published by the official statistical services.

Dear Readers, we are pleased to present you with this report and we would be grateful for any comments and suggestions regarding its content as well as the results obtained and conclusions derived from them. Any feedback will contribute to the development of the research method applied and the scope of data acquired from administrative registers.

Spis treści

Table of Contents

Przedmowa	7
<i>Preface</i>	8
Wykaz skrótów	11
<i>Abbreviations</i>	11
Wstęp	12
<i>Introduction</i>	13
Synteza	14
<i>Executive summary</i>	16
1. Wyniki badania	18
<i>Survey results</i>	
1.1. Wyniki oszacowania liczby pracujących	18
<i>Results of the estimation of the number of employed persons</i>	
1.2. Wyniki oszacowania stopy bezrobocia rejestrowanego	23
<i>Results of the estimation of the registered unemployment rate</i>	
1.3. Wyniki oszacowania miar wynagrodzeń brutto	25
<i>Results for the estimation of measures of gross wages and salaries</i>	
2. Metodologia badania	33
<i>Survey methodology</i>	
2.1. Podstawowe pojęcia dotyczące rynku pracy	34
<i>Basic concepts concerning the labour market</i>	
2.2. Charakterystyka administracyjnych zbiorów danych wykorzystanych do wyznaczenia i naliczenia wskaźników	40
<i>Characteristics of administrative data sets used for defining and calculating indicators</i>	
2.2.1. Ocena źródłowych zbiorów danych	45
<i>Evaluation of source data sets</i>	
2.2.2. Opis źródłowych zbiorów danych	47
<i>Description of source data sets</i>	
2.3. Opis rozwiązań metodologicznych i metodycznych zastosowanych do generowania wskaźników i danych wyników	49
<i>Description of methodological and methodical solutions used for the generation of indicators and result data</i>	
2.3.1. Opis rozwiązań metodologicznych przyjętych w celu wyznaczenia zbiorowości pracujących ..	49
<i>Description of methodological solutions adopted for counting the population of employed persons</i>	
2.3.2. Opis rozwiązań metodycznych przyjętych w celu wyznaczenia liczby pracujących w głównym miejscu pracy	53
<i>Description of methodical solutions adopted for counting the number of employed persons working in the main workplace</i>	
2.3.3. Opis rozwiązań metodologicznych przyjętych w celu wyznaczenia zbiorowości bezrobotnych zarejestrowanych	54

	<i>Description of methodological solutions adopted for counting the population of registered unemployed persons</i>	
2.3.4.	Opis rozwiązań metodycznych przyjętych w celu wyznaczenia stopy bezrobocia rejestrowanego	55
	<i>Description of methodical solutions adopted for the calculation of the registered unemployment rate</i>	
2.3.5.	Opis rozwiązań metodologicznych przyjętych w celu wyznaczenia zbiorowości zatrudnionych i oszacowania miar wynagrodzeń	56
	<i>Description of methodological solutions adopted for counting the population of employees and the estimation of their wages and salaries</i>	
2.3.6.	Opis rozwiązań metodycznych przyjętych w celu oszacowania miar wynagrodzeń i miar rozkładu wynagrodzeń	57
	<i>Description of methodical solutions adopted for the estimation of the measures of wages and salaries and measures for distribution of wages and salaries</i>	
2.3.7.	Procedury przygotowujące zbiory KRUS i ZUS	59
	<i>Procedures for the preparation of data sets of the Polish Social Insurance Institution (ZUS) and the Agricultural Social Insurance Fund (KRUS)</i>	
2.3.8.	Opis rozwiązań metodologicznych i metodycznych przyjętych w celu uzupełnienia braków danych	60
	<i>Description of methodological and methodical solutions adopted for filling data gaps</i>	
2.4.	Lista oszacowanych wskaźników i danych wynikowych	63
	<i>List of estimated indicators and result data</i>	
2.5.	Ocena jakości badania	65
	<i>Assessment of survey quality</i>	
	Bibliografia	67
	<i>Bibliography</i>	
	Spis załączników	68
	<i>List of appendices</i>	

Wykaz skrótów Abbreviations

Skrót <i>Abbreviation</i>	Znaczenie <i>Meaning</i>
AZD	Administracyjne Źródła Danych <i>Administrative Data Sources</i>
BJS	Baza Jednostek Statystycznych <i>The Base of Statistical Units</i>
CRP KEP	Centralny Rejestr Podmiotów Krajowej Ewidencji Podatników <i>Central Register of National Taxpayer Records</i>
GUS	Główny Urząd Statystyczny <i>Statistics Poland</i>
KRUS	Kasa Rolniczego Ubezpieczenia Społecznego <i>The Agricultural Social Insurance Fund</i>
MF	Ministerstwo Finansów <i>Ministry of Finance</i>
MICE	Metoda imputacji wielokrotnej <i>Multivariate Imputation by Chained Equations</i>
MRPiPS	Ministerstwo Rodziny, Pracy i Polityki Społecznej <i>Ministry of Family, Labour and Social Policy</i>
NIP	Numer Identyfikacji Podatkowej <i>Taxpayer Identification Number</i>
NTS	Nomenklatura Jednostek Terytorialnych do Celów Statystycznych <i>Nomenclature of Territorial Units for Statistical Purposes</i>
p.proc.	Punkt procentowy <i>Percentage point</i>
PNA	Pocztowy Numer Adresowy <i>Postal Address Number</i>
PBSSP	Program Badań Statystycznych Statystyki Publicznej <i>Programme of Statistical Surveys of Official Statistics</i>
PKD 2007	Polska Klasyfikacja Działalności 2007 <i>Polish Classification of Activities</i>
REGON	Krajowy Rejestr Urzędowy Podmiotów Gospodarki Narodowej <i>National Official Business Register</i>
TERYT	Krajowy Rejestr Urzędowy Podziału Terytorialnego Kraju <i>National Official Register of the Territorial Division of the Country</i>
UR	Użytki rolne <i>Agricultural Lands</i>
ZUS	Zakład Ubezpieczeń Społecznych <i>Social Insurance Institution</i>

Wstęp

Niniejszy dokument jest raportem końcowym stanowiącym podsumowanie pracy badawczej pt. „Opracowanie metodologii i oszacowanie liczby pracujących w gospodarce narodowej według miejsca zamieszkania i głównego miejsca pracy na poziomie NTS 4, stopy bezrobocia rejestrowanego na poziomie NTS 5 oraz miar wynagrodzeń brutto na poziomie NTS 4”, realizowanej w ramach projektu „Wsparcie systemu monitorowania polityki spójności w perspektywie finansowej 2014-2020 oraz programowania i monitorowania polityki spójności po 2020” współfinansowanego przez Unię Europejską ze środków Programu Operacyjnego Pomoc Techniczna 2014-2020.

W dobie nasilającego się zapotrzebowania na informacje statystyczne coraz częściej odczuwalna jest konieczność pozyskiwania danych szybciej i na niższych poziomach agregacji. Duże znaczenie ma również dążenie do obniżenia kosztów realizacji badań statystycznych. Realizacja badań statystycznych poprzez podmioty gospodarki narodowej czy gospodarstwa domowe jest czasochłonna i najczęściej dokonywana na próbach losowych, co dodatkowo ogranicza możliwości prezentacji danych wynikowych na niskich poziomach agregacji. Wychodząc naprzeciw tym wyzwaniom podjęto próby wykorzystania rejestrów administracyjnych w badaniach statystycznych. Uzyskano w ten sposób nową wiedzę dotyczącą liczby pracujących na poziomie powiatów, stopy bezrobocia rejestrowanego na poziomie gmin oraz miar wynagrodzeń brutto na poziomie powiatów. Wymagało to opracowania odpowiedniego aparatu definicyjnego uwzględniającego potrzeby statystyki publicznej z jednej strony oraz możliwości rejestrów administracyjnych z drugiej strony. Podkreślić należy, że rejestry administracyjne na ogół nie są tworzone i prowadzone w celach statystycznych. Przyjęcie nowych rozwiązań definicyjnych skutkuje niepełną porównywalnością nowych danych z danymi uzyskanymi w ramach oficjalnych badań statystycznych statystyki publicznej. Jest to ten koszt, który należy bezwzględnie ponieść w związku z planowaną zmianą źródła danych w pewnych obszarach badań statystycznych rynku pracy i wynagrodzeń. Zyskiem z wykorzystania danych rejestrowych są przede wszystkim możliwości generowania danych na niższych poziomach agregacji terytorialnej, a w pewnych przypadkach z większą częstotliwością (np. dane o pracujących w rolnictwie indywidualnym pozyskiwane są w ramach powszechnych spisów rolnych i w okresach pomiędzy spisami przyjmują stałe wartości). W przypadku szacunków stopy bezrobocia wartością dodaną jest uwzględnienie miejsca zamieszkania zarówno dla bezrobotnych, jak i pracujących. W przypadku wynagrodzeń brutto nową wartość tworzą mierniki pozycyjne na poziomie powiatów. Uzyskane wyniki nie są porównywalne z żadnymi innymi danymi pochodzącymi z różnych badań.

W ramach pierwszego etapu pracy badawczej przeprowadzono inwentaryzację wybranych rejestrów administracyjnych oraz zasobów informacyjnych statystyki publicznej. Ponadto dokonano szczegółowej analizy definicji zmiennych stosowanych w badaniach statystycznych oraz wykorzystywanych w rejestrach administracyjnych. Wyniki tych analiz pozwoliły na opracowanie metodologii i oszacowanie liczby pracujących według miejsca pracy głównej i miejsca zamieszkania na poziomie powiatów.

W etapie drugim przygotowano metodologię i oszacowano stopę bezrobocia rejestrowanego na poziomie gmin.

Etap trzeci pracy badawczej dotyczył opracowania metodologii i oszacowania miar wynagrodzeń brutto na poziomie powiatów. Do osiągnięcia tego celu przeanalizowano definicje wynagrodzeń stosowane w statystyce publicznej oraz wybranych zmiennych z rejestrów administracyjnych.

Etap czwarty pracy badawczej został poświęcony doskonaleniu wypracowanych w ramach trzech etapów metod, ponownym naliczeniu danych oraz dokonaniu ich analizy.

Wyniki zaprezentowane w niniejszym raporcie dotyczą trzeciego i czwartego kwartału 2016 roku oraz kwartałów 2017 roku.

Raport składa się z: przedmowy, wstępu, syntezy, dwóch rozdziałów oraz aneksu. Cele pracy badawczej, podsumowanie uzyskanych wyników, rekomendacje oraz ocenę możliwości zastąpienia danych z badań statystycznych danymi rejestrowymi zamieszczono w syntezie. Rozdział pt. *Wyniki badania* zawiera analizę wybranych informacji wynikowych dotyczących pracujących, stopy bezrobocia oraz wynagrodzeń brutto. Z kolei rozdział pt. *Metodologia badania* zawiera opis przyjętych rozwiązań metodologicznych w zakresie pozyskiwania danych statystycznych z rejestrów administracyjnych.

Introduction

This document is a final report summarising a research work entitled: *'Development of a methodology and estimation of the number of employed persons in the national economy by place of residence and main workplace at NTS level 4, the registered unemployment rate at NTS level 5 and measures of gross wages and salaries at NTS level 4'* carried out within the framework of the project 'Support for the cohesion policy monitoring system in the 2014–2020 financial perspective and for programming and monitoring cohesion policy after 2020', co-financed by the European Union through the Operational Programme 'Technical Assistance 2014-2020'.

In the times characterised by a growing demand for statistical information there is an increasing recognition for the need of acquiring data faster and at lower aggregation levels. Striving for the reduction of statistical survey costs is also very important. Conducting statistical surveys on the national economy entities or households is time-consuming and usually is carried out on samples, which additionally limit the possibility to present statistical information at low aggregation levels. In order to address these challenges attempts have been made to use administrative registers in statistical surveys. As a result a new knowledge was gained regarding the number of employed persons at the powiat level, the registered unemployment rate at the gmina level and measures of gross wages and salaries at the powiat level. It required to prepare an appropriate definitional apparatus, which takes into account the needs of the official statistics system on the one hand and possibilities offered by administrative registers on the other hand. It should be stressed that administrative registers are not usually created and run for statistical purposes. The adoption of new definitions results in only partial comparability between new statistics and data obtained from official statistical surveys conducted by official statistical services. This is a disadvantageous consequence that must be absolutely accepted in connection with a planned replacement of a data source in some areas of statistical surveys concerning the labour market, including wages and salaries. The advantage of using administrative register data is above all the possibility to generate statistics at lower territorial aggregation levels and in some cases also on a more frequent basis (for instance data on employed persons in individual farming are collected from agricultural censuses and in the periods between censuses are considered to remain constant). As regards the estimation of the unemployment rate the inclusion of information about place of residence both for unemployed and employed persons is an additional advantage. In the case of gross wages and salaries measures of position at the powiat level constitute a new added value. The results obtained are not comparable to any other data derived from various surveys.

During the first stage of the research work an inventory of selected administrative registers and information resources of official statistics system was made. Moreover, definitions of variables used in statistical surveys and also in administrative registers were comprehensively analysed. The results of this analysis made it possible to develop a methodology and estimate the number of employed persons by workplace and place of residence at the powiat level.

In the second stage a methodology was developed and the registered unemployment rate at the gmina level was estimated.

The third phase of the research work was focused on developing a methodology and estimating measures of gross wages and salaries at the powiat level. In order to fulfil this aim definitions of wages and salaries used in official statistics and selected variables from administrative registers were reviewed.

The fourth phase of the research work was devoted to improving the methods worked out during the first three stages, recalculating the data and making an analysis of them.

The results presented in this report relate to the third and fourth quarters of 2016 and the quarters of 2017.

The report comprises: a preface, an introduction, an executive summary, two chapters and an appendix. The aims of the research study, the summary of the results obtained, recommendations and assessment of the possibility of replacing data collected from statistical surveys with administrative register data are presented in the executive summary. The chapter entitled Survey results contains an analysis of selected results information concerning employed persons, the unemployment rate and gross wages and salaries. The chapter entitled Survey methodology describes the adopted methodological solutions in relation to acquiring statistical data from administrative registers.

Synteza

Zasadniczym celem niniejszej pracy badawczej było uzyskanie informacji dotyczących liczby pracujących w gospodarce narodowej według głównego miejsca pracy i miejsca zamieszkania na poziomie NTS 4 w podziale na płeć i grupy sekcji PKD, stopy bezrobocia rejestrowanego na poziomie NTS 5 w podziale na płeć oraz miar wynagrodzeń brutto w gospodarce narodowej (miary klasyczne oraz pozycyjne) i miar koncentracji rozkładu wynagrodzeń na poziomie NTS 4 w podziale na płeć i grupy sekcji PKD. Praca badawcza miała na celu również rozpoznanie możliwości opracowania metodologii i naliczenia wymienionych wyżej wskaźników w innych agregacjach, tj. w podziale na grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej, w podziale na miasto i wieś oraz w sektorze przemysłowym i usługowym.

Osiągnięcie założonych celów wymagało przeprowadzenia szeregu działań dotyczących między innymi zbadania możliwości pozyskania wiedzy z istniejących źródeł, co zrodziło potrzebę przeprowadzenia inwentaryzacji istniejących źródeł danych. W związku z powyższym dokonano szerokiej analizy zasobów informacyjnych Głównego Urzędu Statystycznego, a przede wszystkim zasobów innych jednostek administracji publicznej stanowiących potencjalne źródło danych dla oszacowania omawianych wskaźników.

Na podstawie przeprowadzonej analizy dokumentacji i przepisów prawnych określających zasady funkcjonowania i organizacji poszczególnych systemów informacyjnych administracji publicznej i tworzonych w ich ramach baz danych zdecydowano, że najlepszymi i najbardziej aktualnymi źródłami danych dla potrzeb oszacowania badanych wskaźników według założonych kryteriów ze względu na stopień pokrycia zakresu podmiotowego i przedmiotowego są zbiory danych z systemów informacyjnych Zakładu Ubezpieczeń Społecznych (systemów KSI ZUS – Centralny Rejestr Płatników Składek, Centralny Rejestr Ubezpieczonych) oraz Kasy Rolniczego Ubezpieczenia Społecznego (systemu nSIU – nowy System Informatyczny Ubezpieczeń dawniej KRUSna), a dane zebrane w tych rejestrach pozwolą na uzyskanie wyników w podziale na płeć, grupy wieku, miejsce zamieszkania ubezpieczonych i rodzaj działalności.

Wybrane zbiory wymagały przeprowadzenia szeregu prac umożliwiających ich wykorzystanie do realizacji założonych celów. Przede wszystkim sprawdzono poprawność danych w wybranych zbiorach, w tym zidentyfikowano błędy logiczne, wartości spoza dopuszczalnego zakresu oraz wartości nietypowe (tzw. anomalie), a także braki wartości poszczególnych zmiennych. Proces generowania bazy danych do naliczeń oczekiwanych wskaźników i danych wynikowych polegał na przeprowadzeniu korekty błędnych wartości, wskazaniu i wdrożeniu rozwiązań metodologicznych i metodycznych pozwalających na uzupełnienie brakujących wartości. Rozwiązania te pozwoliły na uzupełnienie braków na poziomie zarówno danych jednostkowych, jak i agregatów, dzięki wykorzystaniu innych zmiennych znajdujących się w danej bazie lub zmiennych dostępnych w innych źródłach danych. Następnie przeprowadzono integrację zbiorów, co poprzedzono podjęciem szeregu decyzji metodologicznych dotyczących zakresu podmiotowego i przedmiotowego (w tym w zakresie powtórzonych lub pustych obserwacji) generowanej bazy danych.

Tak przygotowana baza stanowiła źródło danych dla naliczenia wskaźników i danych wynikowych określonych celem niniejszej pracy badawczej.

Badanie wykazało, iż liczba pracujących w gospodarce narodowej wahała się od 14,1 mln osób do 14,4 mln osób. Liczba pracujących w sektorze rolniczym systematycznie malała i oscylowała wokół wartości 1,5 mln osób. Spośród wszystkich pracujących najwięcej pracowało w sektorze usługowym – od 8,6 mln osób w III kw. 2016 r. do 8,8 mln osób w II kw. 2017 r.

Stopa bezrobocia rejestrowanego w czerwcu 2017 r. wyniosła 7,5%, trzy miesiące wcześniej była o 1,1 p. proc. wyższa. Dla kobiet stopa bezrobocia rejestrowanego była wyższa o 2,1 p. proc. w porównaniu z mężczyznami (6,5%). Najwyższa stopa bezrobocia w obu analizowanych okresach wystąpiła w województwie warmińsko-mazurskim – 13,4% w marcu 2017 r. i 11,4% w czerwcu 2017 r. Najniższe wartości wskaźnik ten osiągnął w województwie wielkopolskim – odpowiednio 5,2% oraz 4,4%. Gminą o najniższej stopie bezrobocia było Kleszczewo (woj. wielkopolskie, pow. poznański) – odpowiednio 1,6% i 1,4%. Najwyższą wartość wskaźnik ten osiągnął w gminie Lelkowo (woj. warmińsko-mazurskie, pow. braniewski) – odpowiednio 35,1% oraz 31,5%.

Wynagrodzenia brutto zaprezentowano w postaci miar klasycznych (średnia arytmetyczna) oraz miar pozycyjnych (kwantyle, dominanta). Każda z tych miar zachowuje tendencję wzrostową w ciągu badanego okresu czasu. W czerwcu 2017 r. przeciętne wynagrodzenie przyjmuje wartości z przedziału od 3000 do 4000 zł. Najwyższe wartości przekraczają nawet 10000 zł (mężczyźni w pow. łęczyńskim

w woj. lubelskim). Wartości dominanty nie są istotnie zróżnicowane w żadnym z badanych okresów. Najczęściej przekraczają one minimalne wynagrodzenie o co najwyżej kilkanaście złotych. Mediana wynagrodzeń ogółem w czerwcu 2017 r. mieściła się przedziale od prawie 2366 zł w powiecie kępińskim (woj. wielkopolskie) do prawie 4725 zł w m.st. Warszawa.

Wyniki przeprowadzonej pracy badawczej potwierdzają użyteczność wypracowanej metodyki szacowania danych o pracujących, stopie bezrobocia oraz miarach wynagrodzeń brutto z wykorzystaniem rejestrów administracyjnych. W związku z tym uzasadniony jest wniosek, że w pewnych obszarach badań statystycznych statystyki oficjalnej możliwe jest zastąpienie źródła danych – ze zbieranych od podmiotów gospodarki narodowej na te, które gromadzone są i przetwarzane w rejestrach przede wszystkim systemu zabezpieczenia społecznego. Dane głównie Zakładu Ubezpieczeń Społecznych i Kasy Rolniczego Ubezpieczenia Społecznego pozwoliły na uzyskanie nowych informacji. Uzyskane dane o pracujących mają charakter innowacyjny z uwagi na to, że m.in. dotyczą miejsca zamieszkania, są prezentowane w podziale na grupy wieku, w zakresie pracujących w rolnictwie opracowana metodyka pozwala na uzyskiwanie danych z częstotliwością roczną (dotychczas dane o tej grupie pracujących pochodziły z Powszechnego Spisu Rolnego). Dane dotyczące stopy bezrobocia rejestrowanego również mają innowacyjny charakter z uwagi na to, że przede wszystkim są dostępne na poziomie gminy, w podziale na grupy wieku. Ponadto wyznaczony wskaźnik jest spójny metodologicznie z uwagi na to, że zarówno pracujący, jak i bezrobotni liczeni są według miejsca zamieszkania. Ze względu na to, dane te nie mają w polskiej statystyce odpowiednika i powinny zostać wprowadzone do praktyki badawczej jako dodatkowe informacje o rynku pracy. Dostępność informacji o stopie bezrobocia na poziomie gmin w podziale na płeć i grupy wieku będzie cenną wskazówką do podejmowania decyzji o strategicznym znaczeniu dla rozwoju lokalnych rynków pracy. Również oszacowane miary wynagrodzeń brutto nie mają odpowiednika w statystyce publicznej ze względu na to, że dotyczą poziomu powiatu i obejmują wszystkie podmioty gospodarki narodowej. Podkreślić należy, że szczególnego znaczenia nabiera kwestia prezentowania miar pozycyjnych wynagrodzeń do poziomu powiatu.

Z przedstawionych względów rekomenduje się:

- kontynuowanie prac w zakresie doskonalenia metodyki pozyskiwania danych o pracujących i bezrobociu rejestrowanym na poziomie gmin oraz miar wynagrodzeń brutto na poziomie powiatów,
- doskonalenie wypracowanej metodyki również z wykorzystaniem innych niż dotychczas stosowane rejestrów administracyjnych,
- w ramach trwałości projektu zastąpienie istniejących źródeł danych danymi rejestrowymi (w zakresie możliwym do zastąpienia) i wprowadzenie wypracowanej metodyki do praktyki badawczej oraz stworzenie narzędzia dedykowanego do prezentowania i wizualizacji eksperymentalnych danych.

Tablice wynikowe stanowiące rezultat badania zamieszczone zostaną na stronie internetowej Głównego Urzędu Statystycznego. Mogą być one jednym z elementów monitorowania sytuacji na rynku pracy i wspomagać decyzje w sferze gospodarczej i społecznej.

Executive summary

The principal objective of this research work was to obtain information on the number of employed persons in the national economy by main workplace and place of residence at NTS level 4 broken down by sex and groups of PKD sections, the registered unemployment rate at NTS level 5 broken down by sex and measures of gross wages and salaries (classical and positional measures) as well as measures of concentration for the distribution of wages and salaries at NTS level 4 broken down by sex and groups of PKD sections. The research work was also aimed at exploring the possibilities of developing a methodology and calculating the above mentioned indicators for other aggregations, broken down by the following age groups: 15-24, 25-54, 55-64, 65 years or more, broken down by urban and rural areas as well as for the industrial sector and the service sector. In addition, it was planned to examine the possibility of calculating the registered unemployment rate at NTS level 5 in the scope of long-term unemployment (according to the national definition and according to the definition used by Eurostat).

Achieving this objective required to perform a range of actions concerning among others exploring the possibilities of acquiring knowledge from existing sources, which gave rise to the need for making an inventory of existing data sources. In connection with the foregoing, a comprehensive analysis of the information resources of Statistics Poland and, above all, the resources of other public administration units constituting a potential data sources for the estimation of the said indicators was made.

On the basis of the analysis of documentation and legal provisions defining the principles of functioning and organisation of particular public administration information systems and databases created within them, it was concluded that the best and most current data sources for the purpose of estimating the said indicators according to the set criteria, that is the coverage degree of subjective and objective scopes are data sets from the information systems of the Polish Social Insurance Institution (ZUS) – (the KSI ZUS - Central Register of Contribution Payers, Central Register of Insured Persons) and the Agricultural Social Insurance Fund (KRUS) – (the nSIU system – the New Social Insurance IT System, formerly the KRUSnal), and also that data collected in these registers would enable information broken down by sex, age groups, place of residence of the insured and type of activity to be obtained.

The selected data sets required to carry out a series of works enabling their use to fulfil the established objectives. Before anything else, the correctness of data in selected sets was checked, including logical errors, values outside the acceptable range and non-standard values (the so-called anomalies) as well as missing values of particular variables. The process of generating a database for the calculation of the mentioned indicators and result data involved correction of erroneous values, identification and implementation of methodological and methodical solutions enabling missing values to be filled. These solutions made it possible to fill gaps at the level of both microdata and aggregates, thanks to the use of other variables stored in a given database or variables available from other data sources. Then, the integration of sets was carried out, which was preceded by taking a number of methodological decisions regarding the subjective and objective scopes (also in reference to blank or duplicate observations) of the generated database.

The database prepared in such a way was the source of data for the calculation of indicators and result data specified as the purpose of this research work.

The study showed that the number of employed persons in the national economy oscillated between 14.1 and 14.4 million people. The number of employed persons in the agricultural sector was systematically decreasing and oscillated around the level of 1.5 million people. Most of the employed persons worked in the service sector – from 8.6 million people in the third quarter of 2016 do 8.8 million people in the second quarter of 2017.

The registered unemployment rate in June 2017 was 7.5 %. Three months earlier it was 1.1 percentage points higher. Women's registered unemployment rate was 2.1 percentage points higher than men's (6.5 %). The highest unemployment rate in both analysed periods occurred in the Warmińsko-Mazurskie Voivodship – 13.4 % in March 2017 and 11.4 % in June 2017. The lowest values of this indicator were observed in the Wielkopolskie Voivodship – 5.2 % and 4,4 %, respectively. The gmina with the lowest unemployment rate was Kleszczewo (Wielkopolskie Voivodship, poznański powiat) – 1.6 % and 1.4 %, respectively. The highest value of this measure was observed in Gmina Lelkowo (Warmińsko-Mazurskie Voivodship, braniewski powiat) – 35.1 % and 31.5 %, respectively.

Gross wages and salaries were presented in the form of classical measures (the arithmetic mean) and positional measures (quantiles, the mode). Each of these measures showed an increasing tendency over

the analysed period of time. In June 2017, the average wage and salary ranged from PLN 3,000 to PLN 4,000. The highest values exceeded even PLN 10,000 (men in the łęczyński powiat in the Lubelskie Voivodship). The values of the mode did not significantly differ in any of the periods studied. Most often, they exceeded the minimum wage by at most a dozen or so zlotys. The median wage and salary in June 2017 was in the range between nearly PLN 2366 in the kępiński powiat (Wielkopolskie Voivodship) to nearly PLN 4725 in the Capital City of Warsaw.

The results of the research work confirm the usefulness of the developed methodology for estimating data on employed persons, the unemployment rate and measures of gross wages and salaries with the use of administrative registers. Consequently, it is reasonable to conclude that in some areas of statistical surveys conducted by official statistical services, it is possible to change the data source – replacing data collected from entities of the national economy with those collected and processed in registers, primarily those used in the social security system. The data, received mainly from the Polish Social Insurance Institution (ZUS) and the Agricultural Social Insurance Fund (KRUS) enabled new information to be obtained. The acquired data on employed persons are innovative due to the fact that among others they relate to the place of residence, and are presented as broken down by age groups. As regards the employed in agriculture, the developed methodology enables data to be obtained on an annual basis (hitherto, data on this group of employed persons came from the National Agricultural Census). Data on the registered unemployment rate are also innovative due to the fact that above all they are available at the gmina level. They are broken down by age groups. Moreover, the calculated indicator is methodologically consistent as both the employed and the unemployed are counted according to the place of residence. Therefore, these data do not have an equivalent in Polish official statistics and should be included to the scope of its research activity as additional information on the labour market. The availability of information on the unemployment rate at the gmina level, broken down by sex and age groups, will constitute a valuable support in making decisions of strategic importance for the development of local labour markets. The estimated measures of gross wages and salaries also do not have an equivalent in official statistics due to the fact that they concern the powiat level and include all entities of the national economy. It should be emphasised that the question of presenting positional measures of wages and salaries for the powiat level is particularly important.

For the reasons presented, it is recommended to:

- continue the works in the field of improving the methodology of obtaining data on employed persons and registered unemployment at the gmina level and measures of gross wages and salaries at the powiat level,
- improve the developed methodology also with the use of administrative registers other than those used so far,
- as part of maintaining project sustainability replace existing data sources with administrative register data (to the extent possible to replace) and introduce the developed methodology into research practices and create a tool dedicated to presenting and visualising experimental data.

The result tables being the outcome of the research work will be published on the website of Statistics Poland. They may be one of the elements of monitoring the situation in the labour market and support making decisions in the economic and social spheres.

1. Wyniki badania Survey results

Dane uzyskane na podstawie rejestrów administracyjnych odnoszą się do miejsca zamieszkania pracujących, co niewątpliwie jest wartością dodaną w porównaniu do badań statystyki publicznej. W statystyce publicznej grupowania terytorialne wyznaczane są na podstawie siedziby jednostki lub faktycznego miejsca pracy. Innymi atutami pozyskiwania danych o liczbie pracujących, stopie bezrobocia rejestrowanego i miarach wynagrodzeń¹ pozyskiwanych na podstawie rejestrów administracyjnych są: dostępność danych na niższych niż dotychczas poziomach agregacji terytorialnych, według nowych przekrojów (grupy wieku, płeć, sektory działalności, podział na miasto i wieś) oraz fakt, że dane te dotyczą pełnej zbiorowości.

Poniżej zaprezentowano wyniki oszacowania liczby pracujących w gospodarce narodowej według głównego miejsca pracy i miejsca zamieszkania na poziomie NTS 4, stopy bezrobocia rejestrowanego na poziomie NTS 5 oraz miar wynagrodzeń brutto na poziomie NTS 4.

Liczba pracujących została scharakteryzowana w podziale na grupy wieku, grupy sekcji PKD, sektory działalności, płeć i w podziale na miasto i wieś.

Stopa bezrobocia rejestrowanego na poziomie NTS 5 została zaprezentowana według grup wieku, płci i w podziale na miasto i wieś.

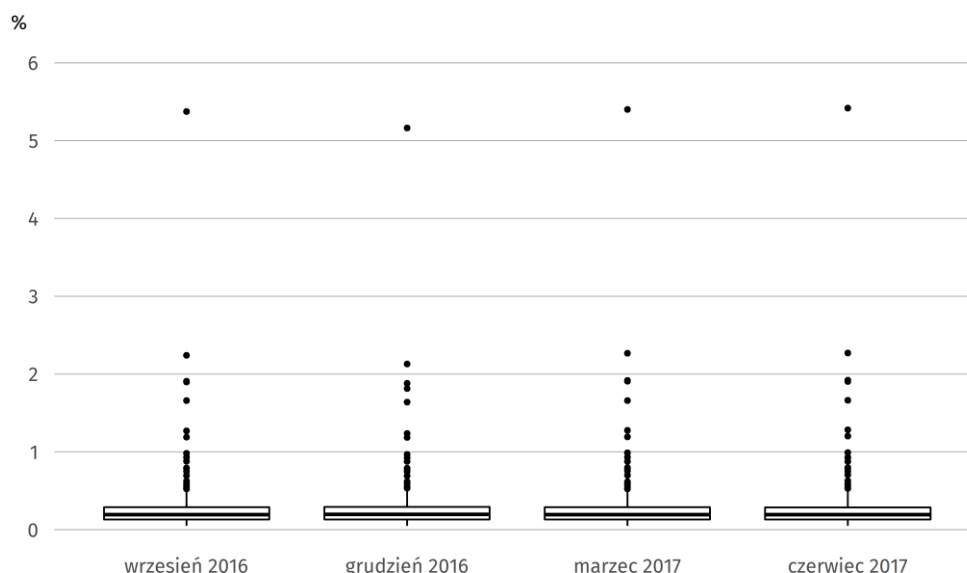
Miary wynagrodzeń brutto (miary klasyczne oraz pozycyjne) oraz miary koncentracji rozkładu wynagrodzeń na poziomie NTS 4 zostały zaprezentowane w podziale na grupy wieku, grupy sekcji PKD, sektory działalności, płeć i w podziale na miasto i wieś.

1.1. Wyniki oszacowania liczby pracujących

Results of the estimation of the number of employed persons

Ogólna liczba pracujących w gospodarce narodowej w głównym miejscu pracy w Polsce we wrześniu i w grudniu 2016 roku oraz w marcu i w czerwcu 2017 roku kształtowała się na poziomie pomiędzy 14,1 mln a 14,4 mln osób. Poniższy wykres świadczy o stabilności danych w przekroju powiatów.

Wykres 1. Udział pracujących w powiatach w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy w Polsce według miejsca zamieszkania

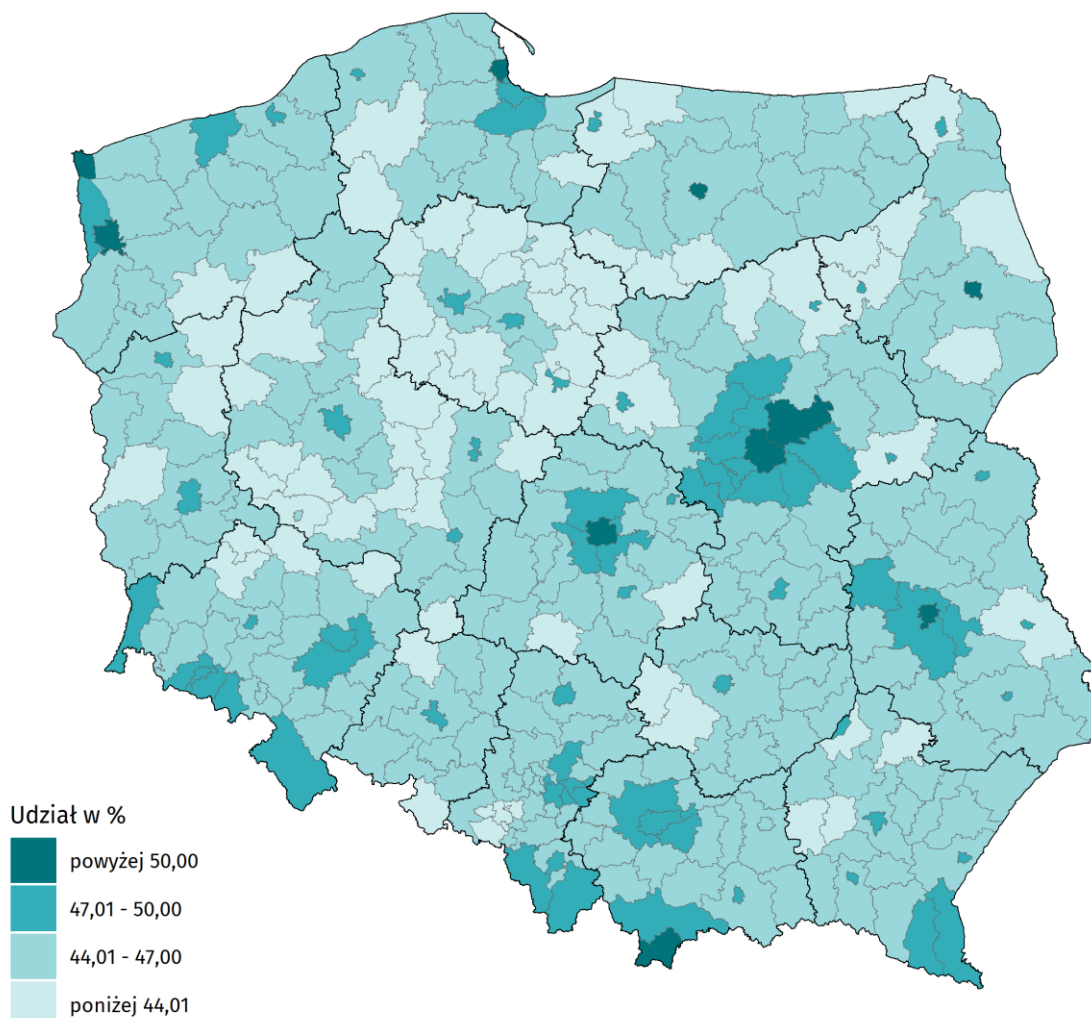


Źródło: Opracowanie własne.

¹ Ilekroć w niniejszym rozdziale pojawia się pojęcie pracujących, zatrudnionych, bezrobotnych rejestrowanych, stopy bezrobocia rejestrowanego lub wynagrodzeń, rozumie się pod tymi pojęciami pracujących, zatrudnionych, bezrobotnych rejestrowanych, stopę bezrobocia rejestrowanego oraz wynagrodzenia według AZD.

Udział pracujących kobiet w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy dla wyżej wymienionych okresów oscylował na poziomie 47%. W przekroju powiatów dla wymienionych okresów można zaobserwować zależność, że udział ten był najwyższy dla miast.

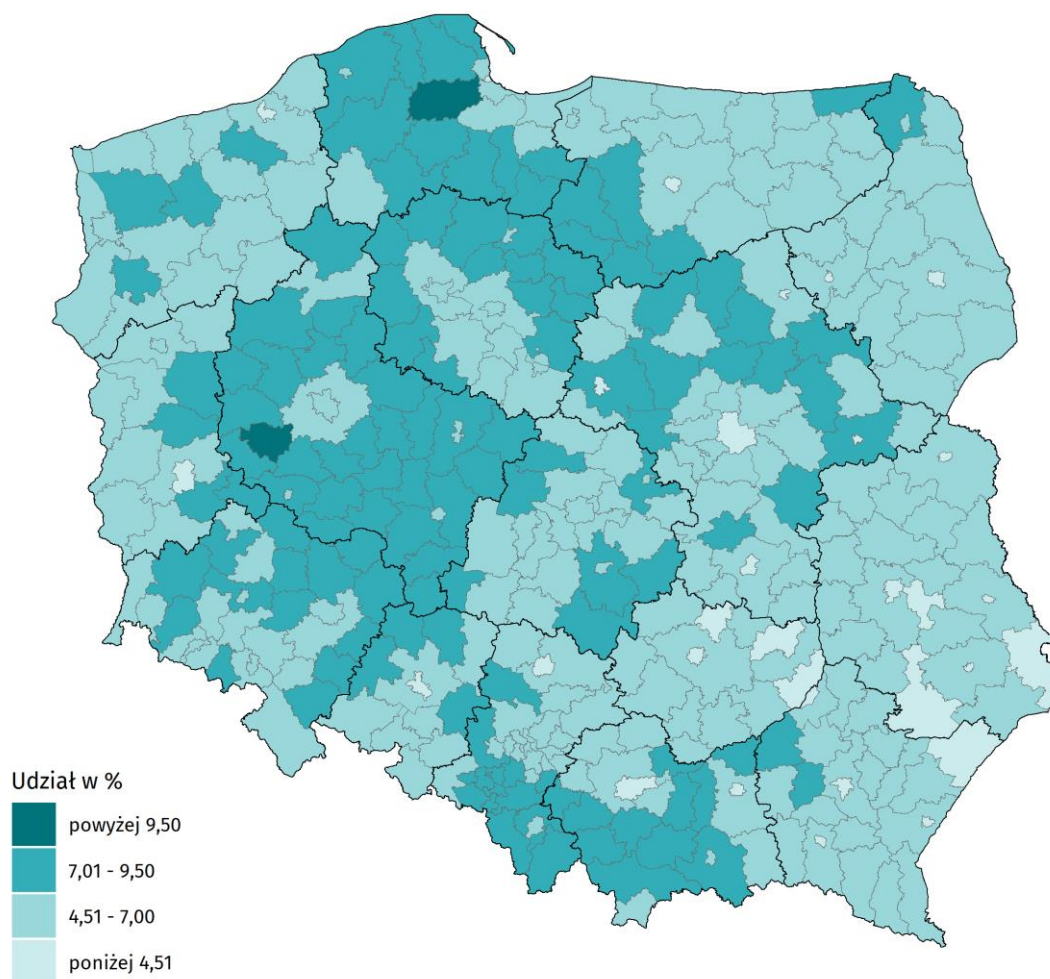
Kartogram 1. Udział pracujących kobiet w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

Udział osób pracujących w wieku 15-24 lata w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy dla Polski we wrześniu i w grudniu 2016 roku oraz w marcu i w czerwcu 2017 roku kształtował się na poziomie 6%. W przekroju powiatów dla wymienionych okresów można zaobserwować zależność, że udział ten był najwyższy dla powiatów Polski centralnej, północnej i południowej. Ponadto warto zauważyć, że udział ten był w wielu przypadkach wyższy w powiatach otaczających największe miasta niż w samych miastach.

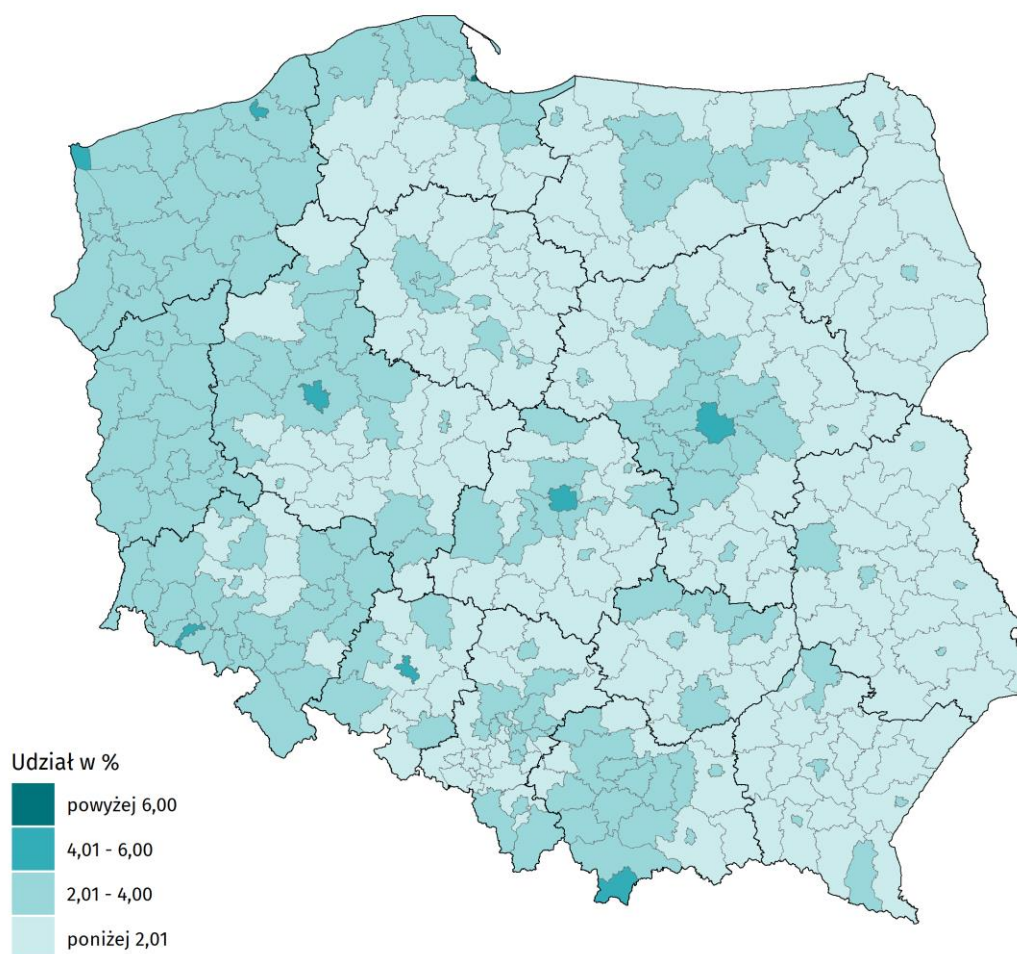
Kartogram 2. Udział osób pracujących w wieku 15-24 lata w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

Udział osób pracujących w wieku 65 lat i więcej w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy dla Polski we wrześniu i w grudniu 2016 roku oraz w marcu i w czerwcu 2017 roku kształtował się na poziomie 2,5%. W przekroju powiatów dla wymienionych okresów można zaobserwować zależność, że udział ten był najwyższy dla powiatów Polski zachodniej. Ponadto warto zauważyć, że udział ten, w przeciwieństwie do udziału pracujących w wieku 15-24 lata w ogólnej liczbie pracujących, był wyższy w miastach. Na szczególną uwagę zasługuje powiat miasto Sopot (woj. pomorskie), w którym udział osób pracujących w wieku 65 lat i więcej w ogóle pracujących był najwyższy i wynosił blisko 7%.

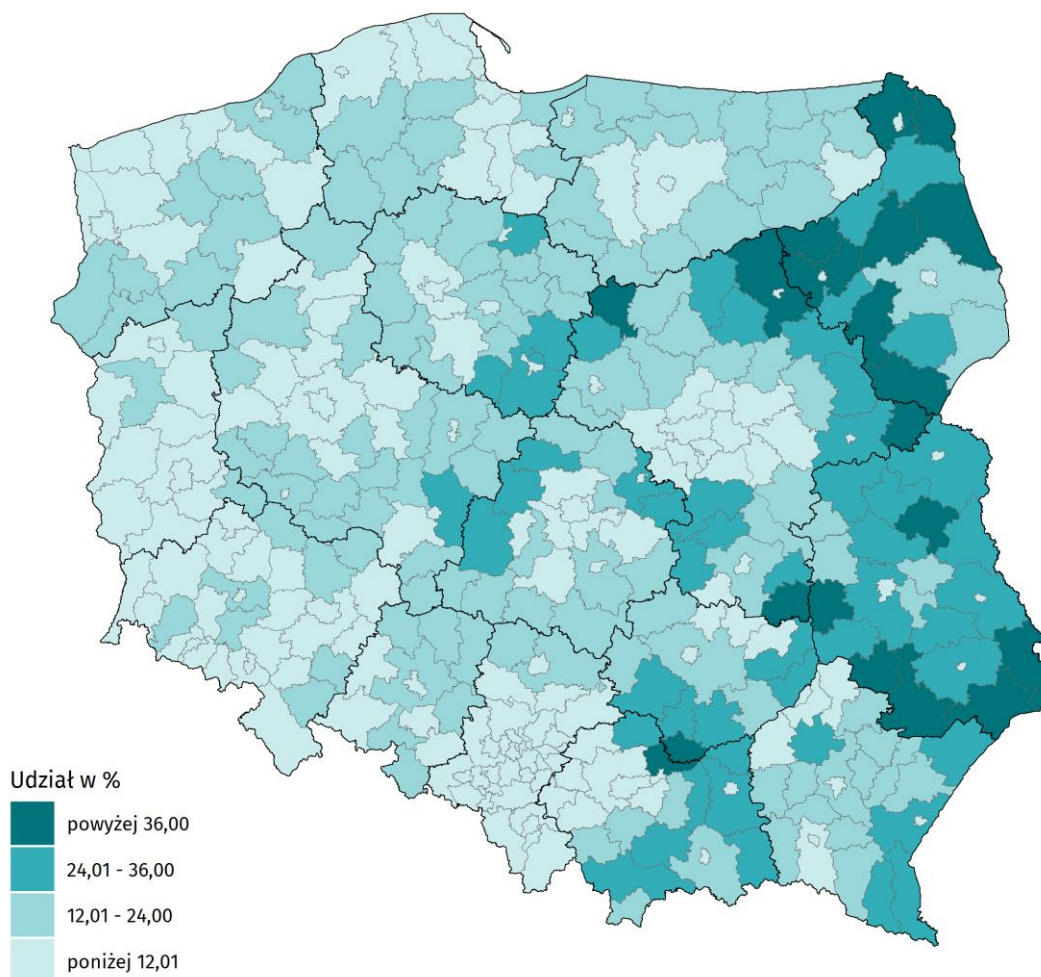
Kartogram 3. Udział osób pracujących w wieku 65 lat i więcej w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

Udział osób pracujących w rolnictwie w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy dla Polski we wrześniu i w grudniu 2016 roku oraz w marcu i w czerwcu 2017 roku kształtował się na poziomie 10,5%. W przekroju powiatów dla wymienionych okresów można zaobserwować wyraźną zależność, że udział ten był najwyższy dla powiatów Polski wschodniej. Szczególnie wysoki udział pracujących w rolnictwie, nawet na poziomie czterdziestu kilku procent, można zaobserwować w powiatach województwa lubelskiego i województwa podlaskiego.

Kartogram 4. Udział pracujących w rolnictwie w ogólnej liczbie pracujących w gospodarce narodowej w głównym miejscu pracy w czerwcu 2017 r. według miejsca zamieszkania



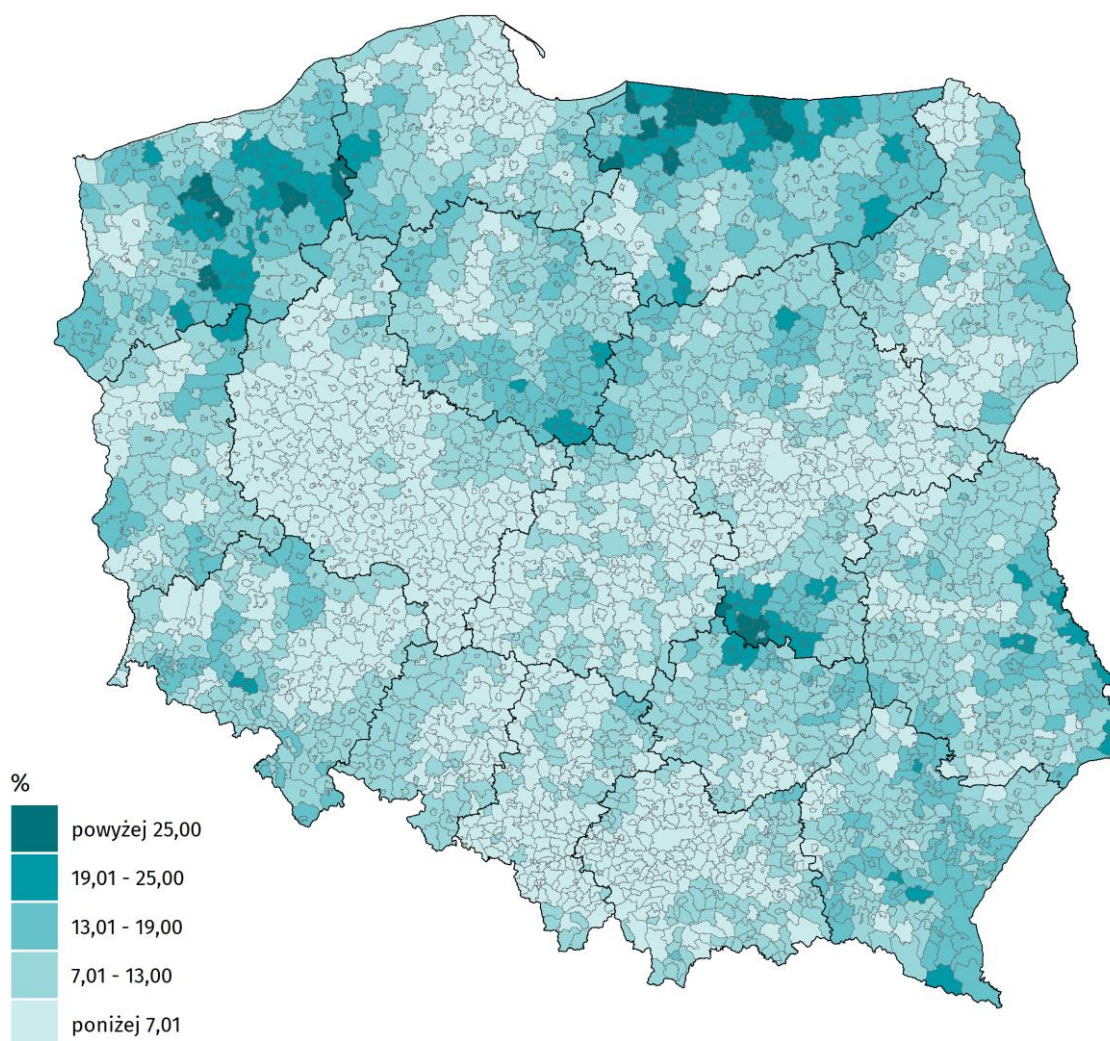
Źródło: Opracowanie własne.

1.2. Wyniki oszacowania stopy bezrobocia rejestrowanego

Results of the estimation of the registered unemployment rate

Stopa bezrobocia rejestrowanego ogółem dla Polski według stanu na czerwiec 2017 roku wyniosła 7,5%. Była ona o 2,1 p.proc. wyższa w przypadku kobiet (8,6%) niż w przypadku mężczyzn (6,5%). Biorąc pod uwagę podział wg grup wieku (15-24 lata, 25-54 lata, 55-64 lata oraz 65 lat i więcej) najniższa stopa bezrobocia rejestrowanego dotyczyła osób w wieku 65 lat i więcej (2,9%), co mogło być związane z osiągnięciem wieku emerytalnego i możliwością przechodzenia na emeryturę. Stopa bezrobocia rejestrowanego osób w grupie wieku 55-64 lata wyniosła 9,0%, podczas gdy dla kobiet z tej grupy wieku równała się zaledwie 7,5% (mężczyźni – 10,0%).

Kartogram 5. Stopa bezrobocia rejestrowanego według gmin w czerwcu 2017 r.



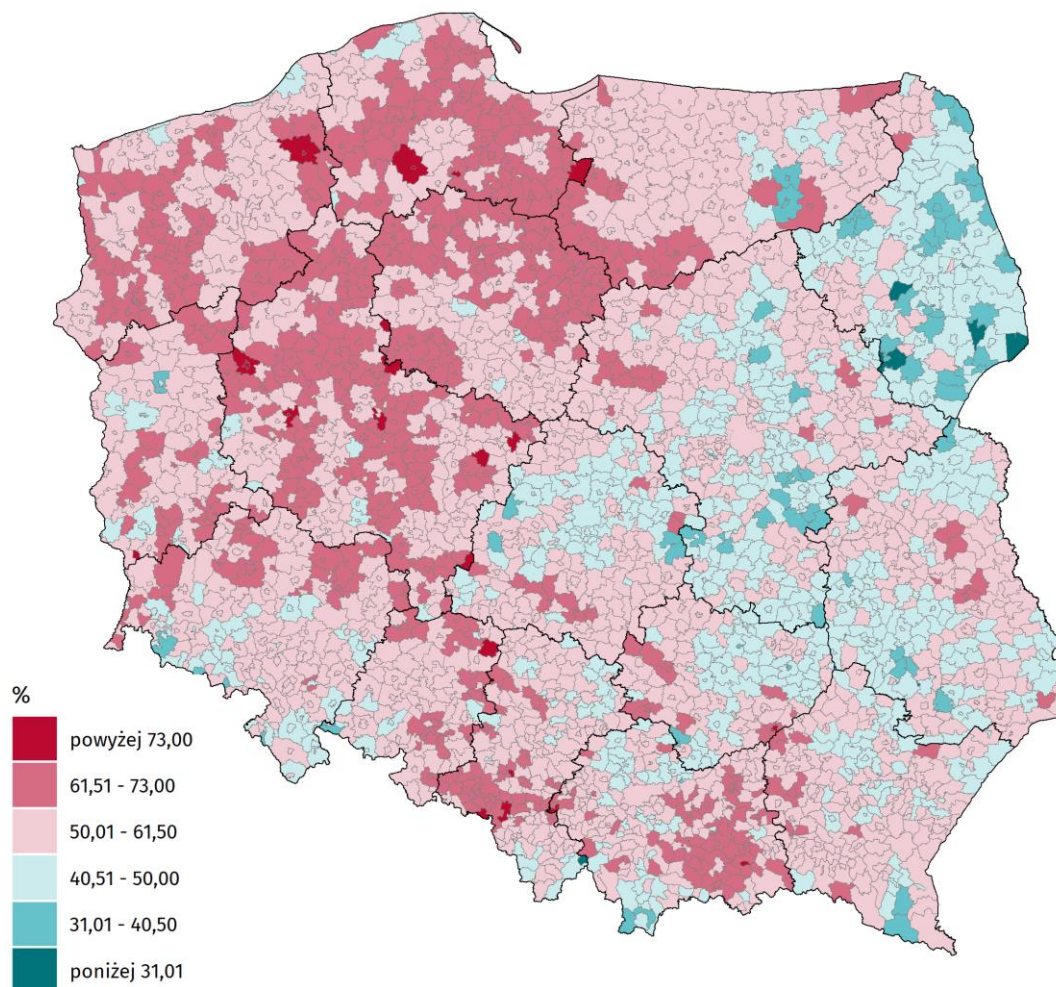
Źródło: Opracowanie własne.

Grupą wieku, w której stopa bezrobocia była najniższa, była grupa osób od 25 do 54 lat (6,7%). Najwyższa stopa bezrobocia dotyczyła osób do 24 lat i wyniosła w tej grupie wieku 14,0%, przy czym 10,1% dla mężczyzn i aż 19% dla kobiet.

Województwem o najwyższej stopie bezrobocia rejestrowanego ogółem było województwo warmińsko-mazurskie (11,4%). Najniższa stopa bezrobocia rejestrowanego ogółem wystąpiła w województwie wielkopolskim (4,4%). Gminą o najwyższej stopie bezrobocia rejestrowanego wśród mężczyzn była gmina

Chlewiska (27,6%; woj. mazowieckie, pow. szydłowiecki), natomiast wśród kobiet była to gmina Frombork (40,7%; woj. warmińsko-mazurskie, pow. braniewski).

Kartogram 6. Udział bezrobotnych kobiet w ogólnej liczbie bezrobotnych według gmin w czerwcu 2017 r.



Źródło: Opracowanie własne.

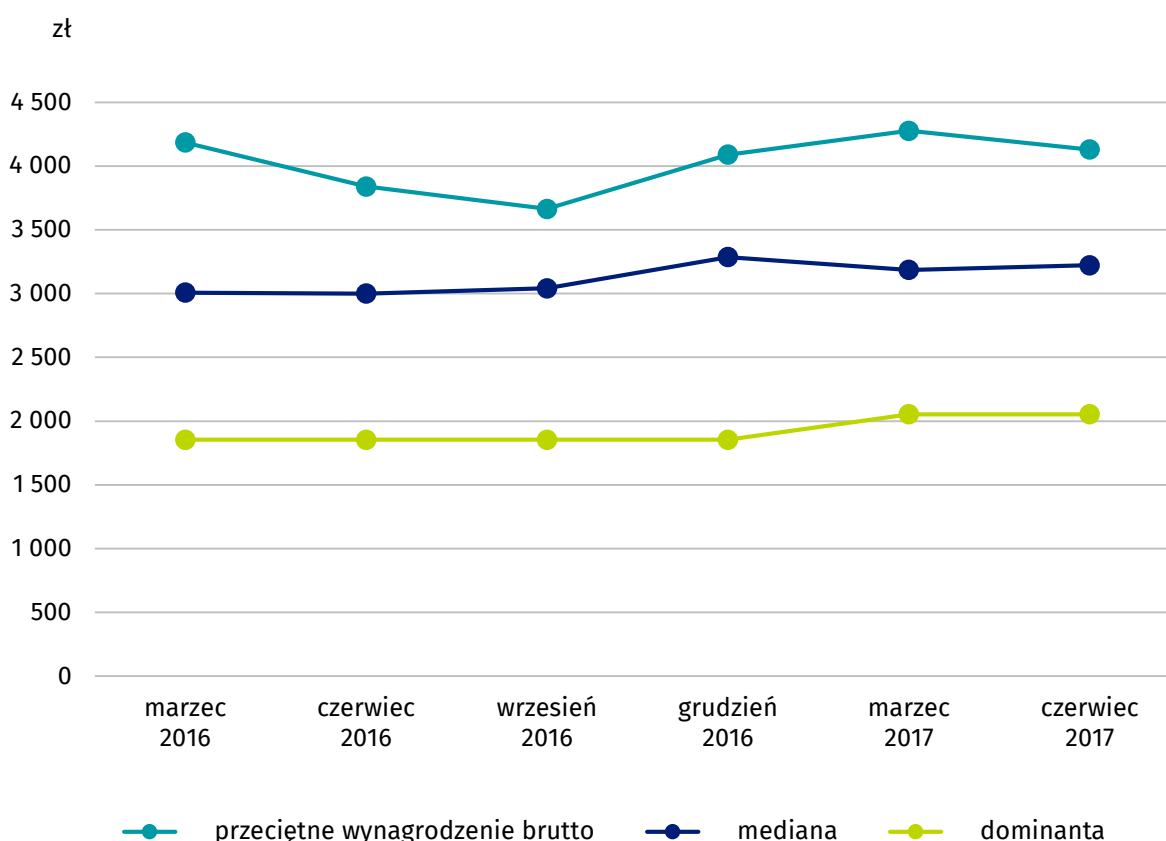
Dane administracyjne umożliwiają prezentowanie wyników w różnych agregacjach i podziałach. Jedną z takich możliwości jest przedstawienie udziału bezrobotnych kobiet w bezrobotnych ogółem w podziale na gminy. Najwyższy udział bezrobotnych kobiet w bezrobotnych ogółem (powyżej 80%) wystąpił w gminach Czajków (woj. wielkopolskie, pow. ostrzeszowski) i Dobrzany (woj. zachodniopomorskie, pow. stargardzki). Pod względem województw najniższy udział bezrobotnych kobiet w bezrobotnych zarejestrowanych ogółem był w województwie podlaskim. Województwa wschodnie charakteryzowały się niższym udziałem bezrobotnych kobiet w bezrobotnych zarejestrowanych ogółem.

1.3. Wyniki oszacowania miar wynagrodzeń brutto

Results for the estimation of measures of gross wages and salaries

Wartości średniej arytmetycznej wynagrodzeń brutto, mediany i dominanty wyraźnie wskazywały na typową cechę rozkładów wynagrodzeń, tj. na prawostronną asymetrię. Relacja miar w rozkładach prawostronnie asymetrycznych jest taka, że najniższe wartości przyjmuje dominanta, druga w kolejności jest mediana, a najwyższą wartość spośród tych trzech miar ma średnia arytmetyczna.

Wykres 2. Wybrane miary wynagrodzeń brutto wypłaconych w okresie od marca 2016 r. do czerwca 2017 r.



Źródło: Opracowanie własne.

Wartości dominanty wynagrodzeń brutto wykazywały niewielką zmienność. Dla każdego powiatu dominanta wynosiła kilka złotych (2016 rok) oraz kilkadziesiąt złotych (2017 rok) powyżej ustawowej płacy minimalnej. W 2016 roku płaca minimalna wynosiła 1850 zł, a w 2017 r. 2000 zł. Wartości dominanty w 2017 roku były o ponad 50 zł wyższe od płacy minimalnej. Z ekonomicznego punktu widzenia oznacza to, że najczęściej osiągnięty poziom wynagrodzenia przekraczał o ponad 50 zł płacę minimalną.

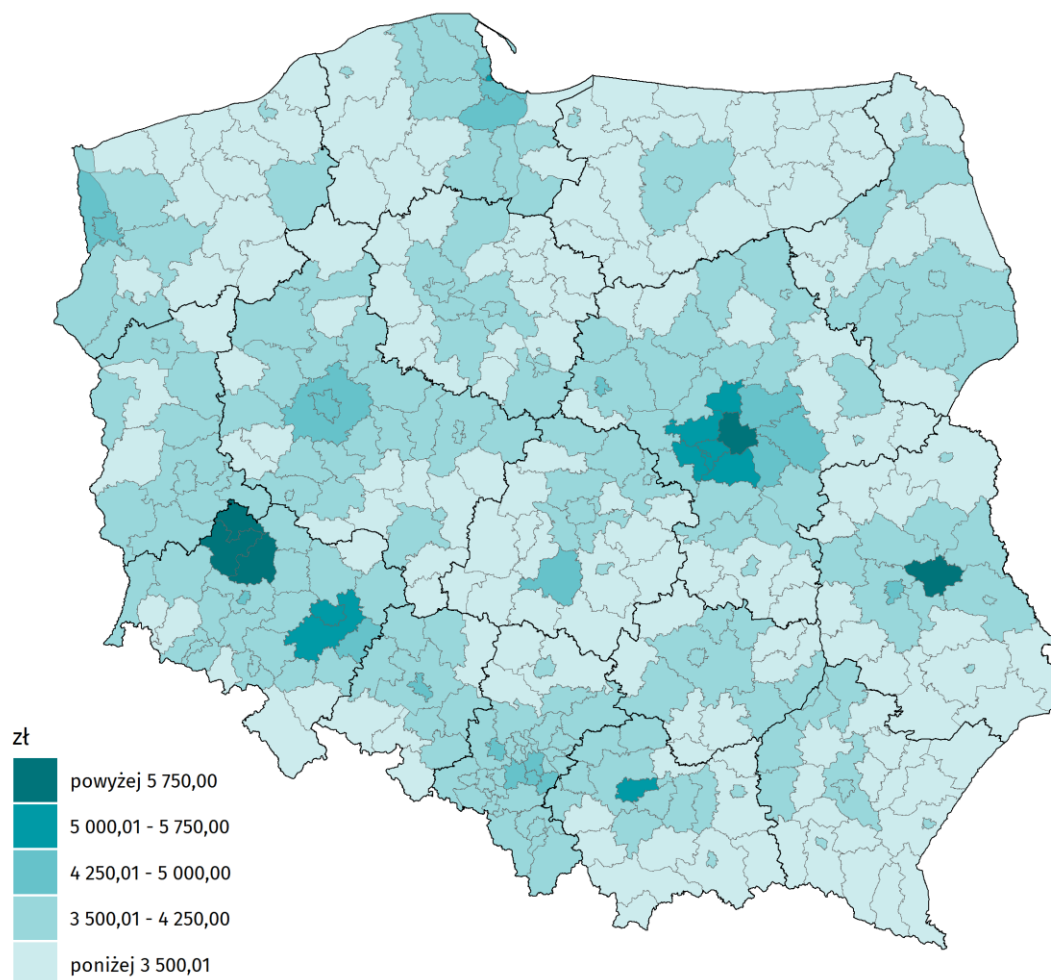
Tablica 1. Wybrane miary wynagrodzeń brutto zatrudnionych w czerwcu 2016 r. i czerwcu 2017 r. według województw

Wyszczególnienie	Czerwiec 2016			Czerwiec 2017		
	Przeciętne wynagrodzenie brutto	Mediana wynagrodzeń brutto	Dominanta wynagrodzeń brutto	Przeciętne wynagrodzenie brutto	Mediana wynagrodzeń brutto	Dominanta wynagrodzeń brutto
Polska	3839,25	3000,00	1854,36	4129,59	3221,84	2053,58
woj. dolnośląskie	3876,82	3143,01	1854,28	4588,52	3461,80	2053,94
woj. kujawsko-pomorskie	3329,32	2700,65	1854,77	3698,62	3000,00	2054,05
woj. lubelskie	3481,33	2726,20	1855,30	3888,21	3004,37	2054,30
woj. lubuskie	3432,39	2836,00	1854,56	3760,29	3076,16	2054,31
woj. łódzkie	3544,49	2793,08	1854,35	3802,18	3045,00	2053,42
woj. małopolskie	3586,37	2842,67	1853,73	4090,65	3209,30	2052,89
woj. mazowieckie	4848,06	3701,32	1853,99	5127,75	3708,00	2052,92
woj. opolskie	3461,60	2881,65	1855,07	3770,82	3168,86	2054,32
woj. podkarpackie	3330,04	2753,38	1855,41	3575,30	2984,00	2055,00
woj. podlaskie	3415,99	2830,40	1855,42	3704,29	3103,81	2054,86
woj. pomorskie	3529,62	2830,00	1853,60	4116,86	3253,21	2053,13
woj. śląskie	3780,66	3000,00	1854,57	4044,49	3343,60	2053,72
woj. świętokrzyskie	3333,60	2712,08	1854,69	3663,85	2987,04	2053,95
woj. warmińsko-mazurskie	3275,42	2686,11	1854,37	3546,55	2911,90	2053,59
woj. wielkopolskie	3749,81	2987,78	1854,14	3954,80	3150,77	2053,17
woj. zachodniopomorskie	3457,14	2830,00	1854,04	3795,77	3108,50	2053,38

Źródło: Opracowanie własne.

Wartości średniej arytmetycznej wynagrodzeń brutto były zróżnicowane. Najczęściej mieściły się w przedziale do 4000 zł brutto, ale występowały również wartości osiągające prawie lub ponad 10000 zł brutto (powiat łączyński w województwie lubelskim, mężczyźni). Powiaty, dla których najczęściej uzyskano wysokie wartości średniej arytmetycznej charakteryzują się wysokim stopniem uprzemysłowienia i zurbanizowania. Minimalne wartości średniej arytmetycznej wynagrodzeń brutto oscylowały wokół kwoty 3000 zł.

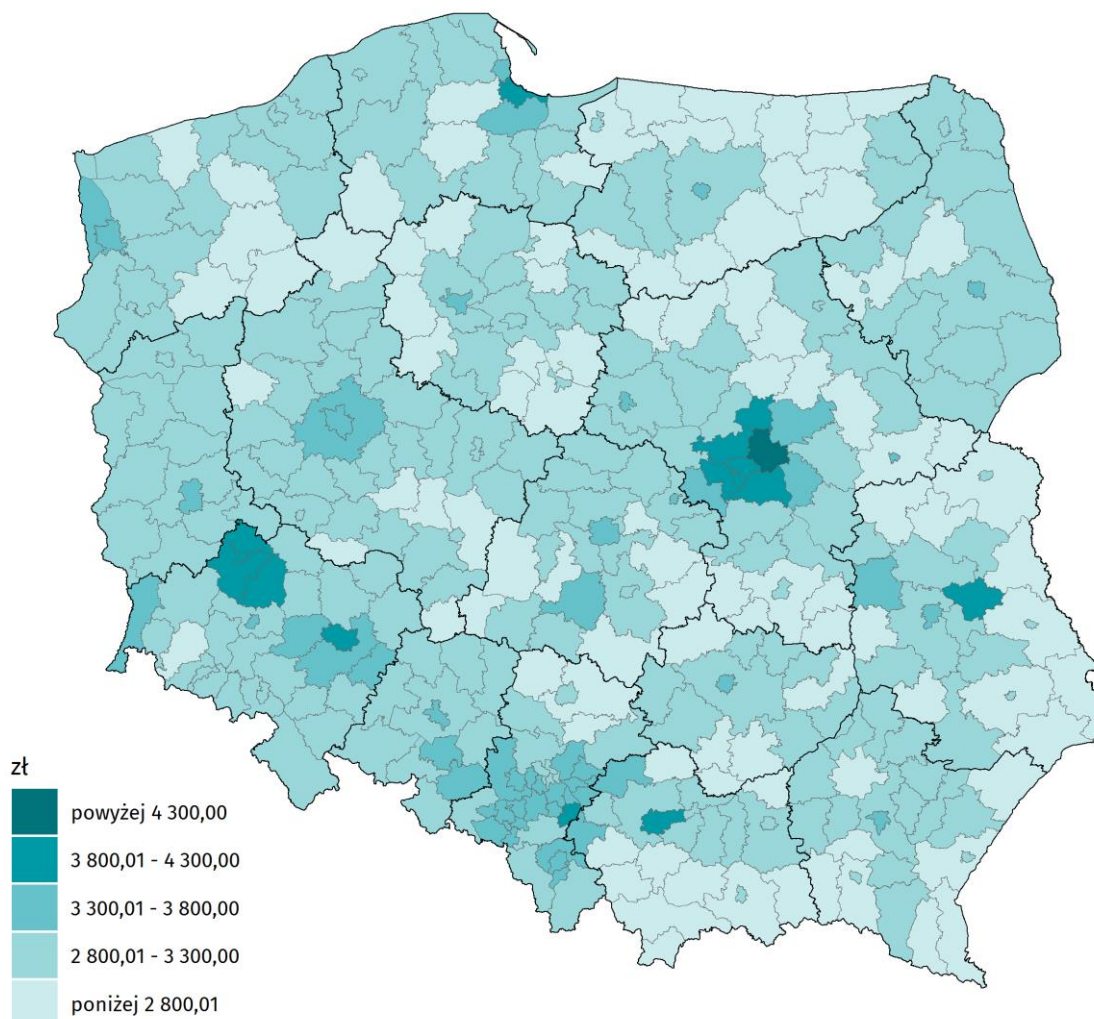
Kartogram 7. Przeciętne wynagrodzenie brutto zatrudnionych wypłacone w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

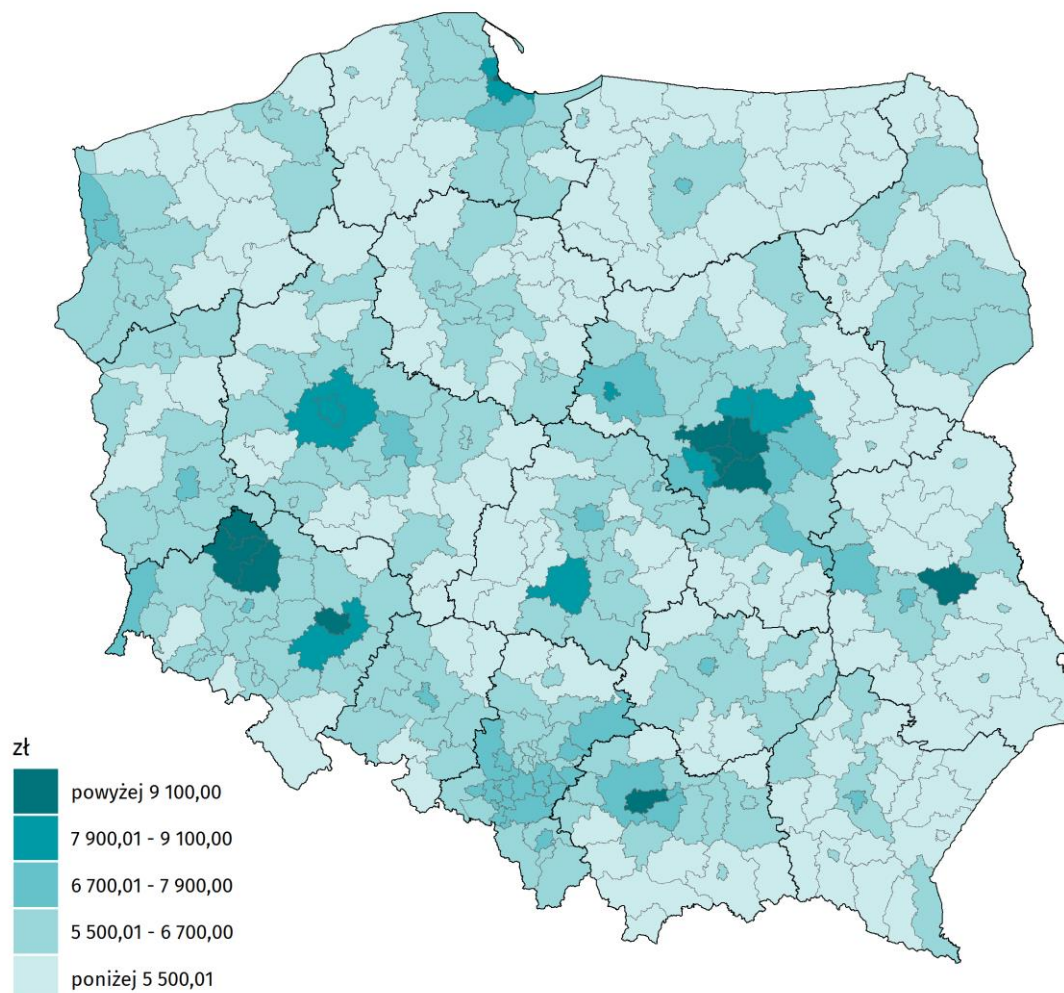
Na potrzeby pracy badawczej oszacowano decyl 1, decyl 5 (mediana, kwartyl 2) oraz decyl 9. Wartości decyla pierwszego były praktycznie stałe i wynosiły tyle, ile wynosiła płaca minimalna w danym roku kalendarzowym. Oznacza to, że 10% zatrudnionych osiągało wynagrodzenie co najwyżej równe minimalnemu. W każdym przypadku mediana była niższa od wartości średniej i jednocześnie znacznie wyższa od dominanty wynagrodzeń. Decyl 9 z kolei opisuje tę wysokość wynagrodzenia, która jest osiągnięta przez 10% najlepiej zarabiających.

Kartogram 8. Mediana wynagrodzeń brutto zatrudnionych wypłaconych w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

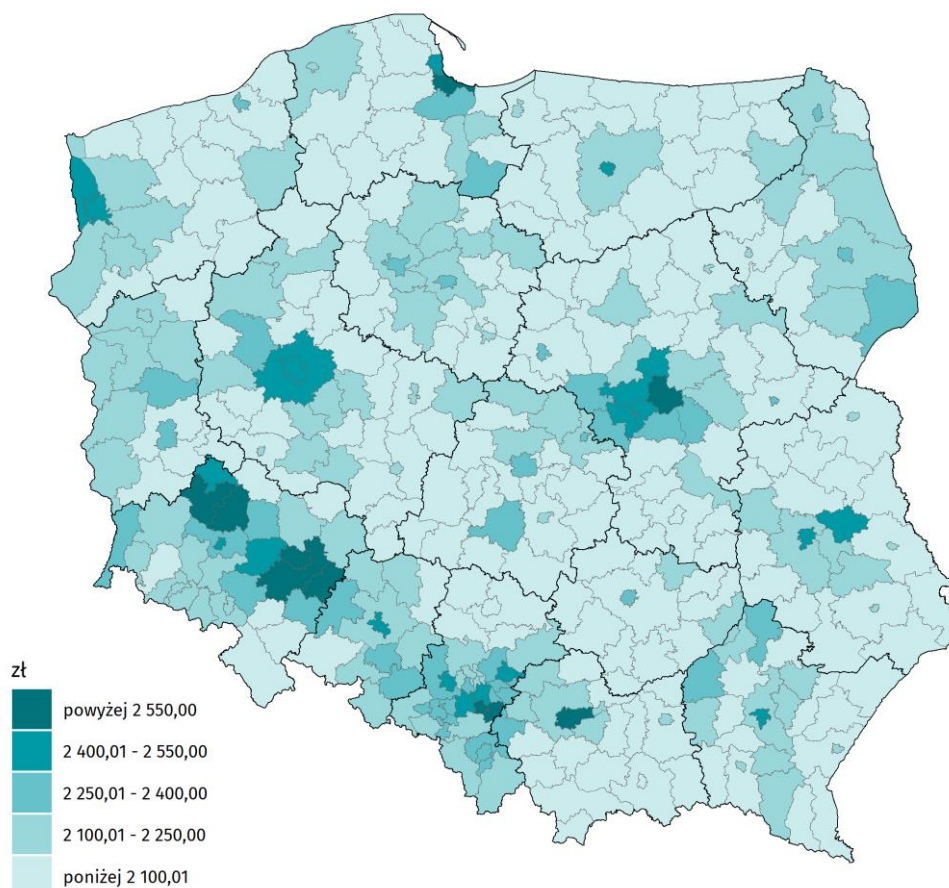
Kartogram 9. Decyl 9 wynagrodzeń brutto zatrudnionych wypłaconych w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

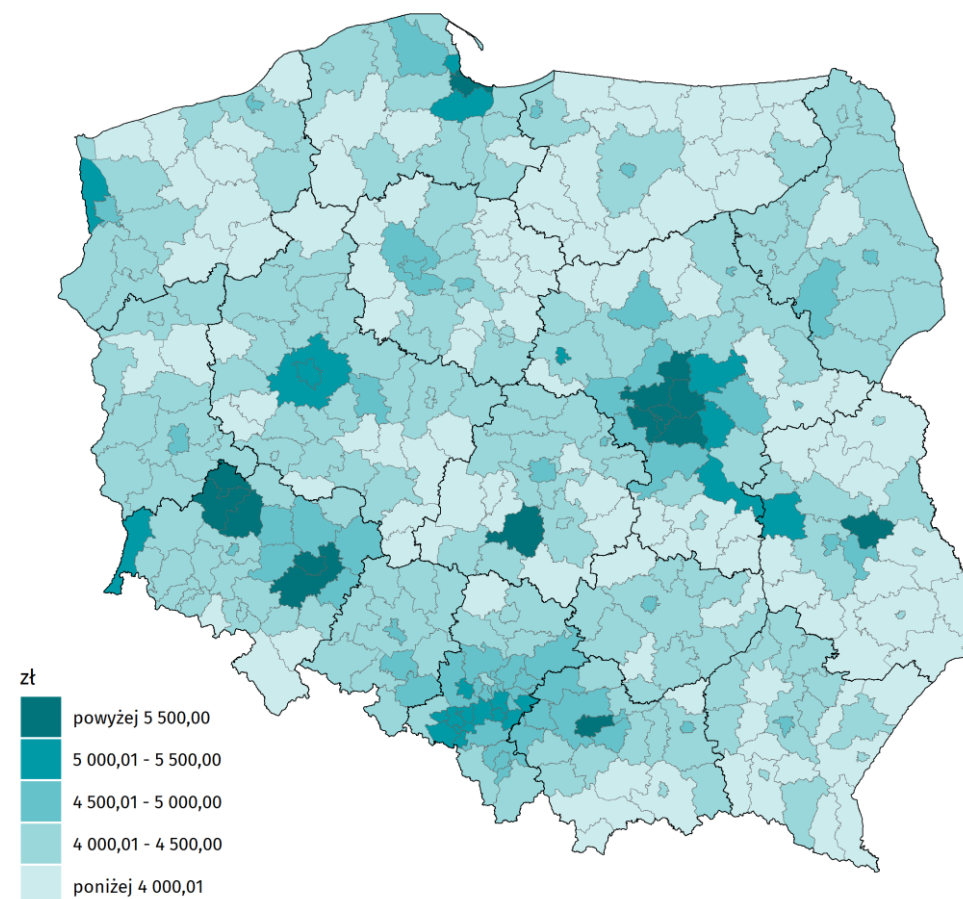
Kwartył 1 oraz kwartył 3 wynagrodzeń brutto opisują takie poziomy wynagrodzeń, które są osiągnięte przez co najwyżej odpowiednio 25% i 75% zatrudnionych. Kwartył 1 w niektórych powiatach nie różnił się znacznie od wynagrodzenia minimalnego, co wyraźnie świadczyło o występowaniu na terenie danego powiatu relatywnie niskiego poziomu wynagrodzeń. Kwartył trzeci natomiast wynosił co najmniej 3200 zł.

Kartogram 10. Kwartył 1 wynagrodzeń brutto zatrudnionych wypłaconych w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

Kartogram 11. Kwartył 3 wynagrodzeń brutto zatrudnionych wypłaconych w czerwcu 2017 r. według miejsca zamieszkania

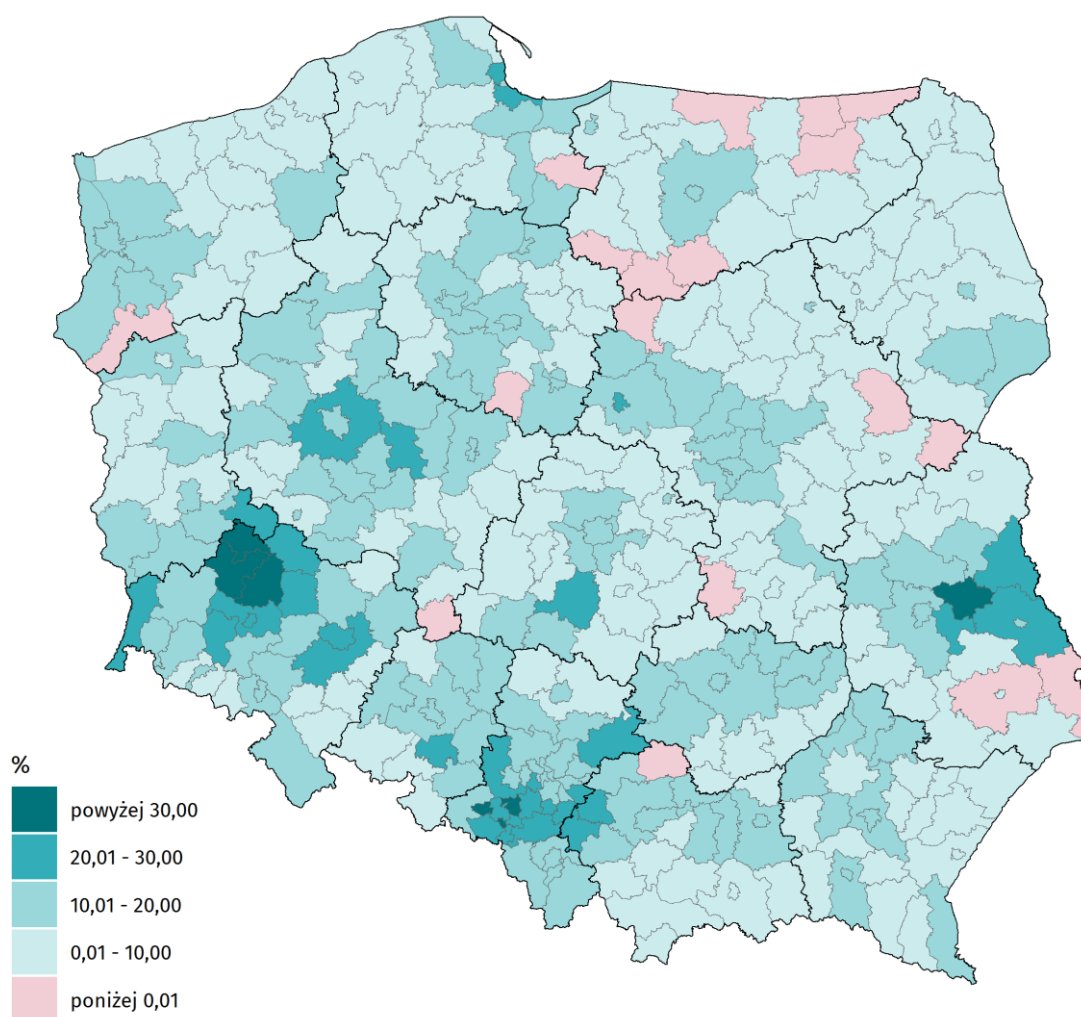


Źródło: Opracowanie własne.

Interesujące z ekonomicznego punktu widzenia są oszacowania współczynnika Giniego oraz pola powierzchni pod krzywą Lorenza. Obie te miary opisują nierównomierności w rozkładach wynagrodzeń brutto. Wyższe wartości wskaźnika Giniego informują o większych nierównościach w dochodach zatrudnionych. Wyznaczone wartości tego wskaźnika mieściły się w szerokim przedziale zmienności od ponad 0,2 do ponad 0,6.

Ponadto istotne z ekonomicznego punktu widzenia jest również badanie zróżnicowania wynagrodzeń ze względu na płeć. Na poniższych mapach przedstawiono zróżnicowanie wynagrodzeń mężczyzn i kobiet w przekroju powiatów. Zaprezentowano je jako udział różnicy przeciętnego wynagrodzenia brutto mężczyzn i kobiet w przeciętnym wynagrodzeniu brutto mężczyzn.

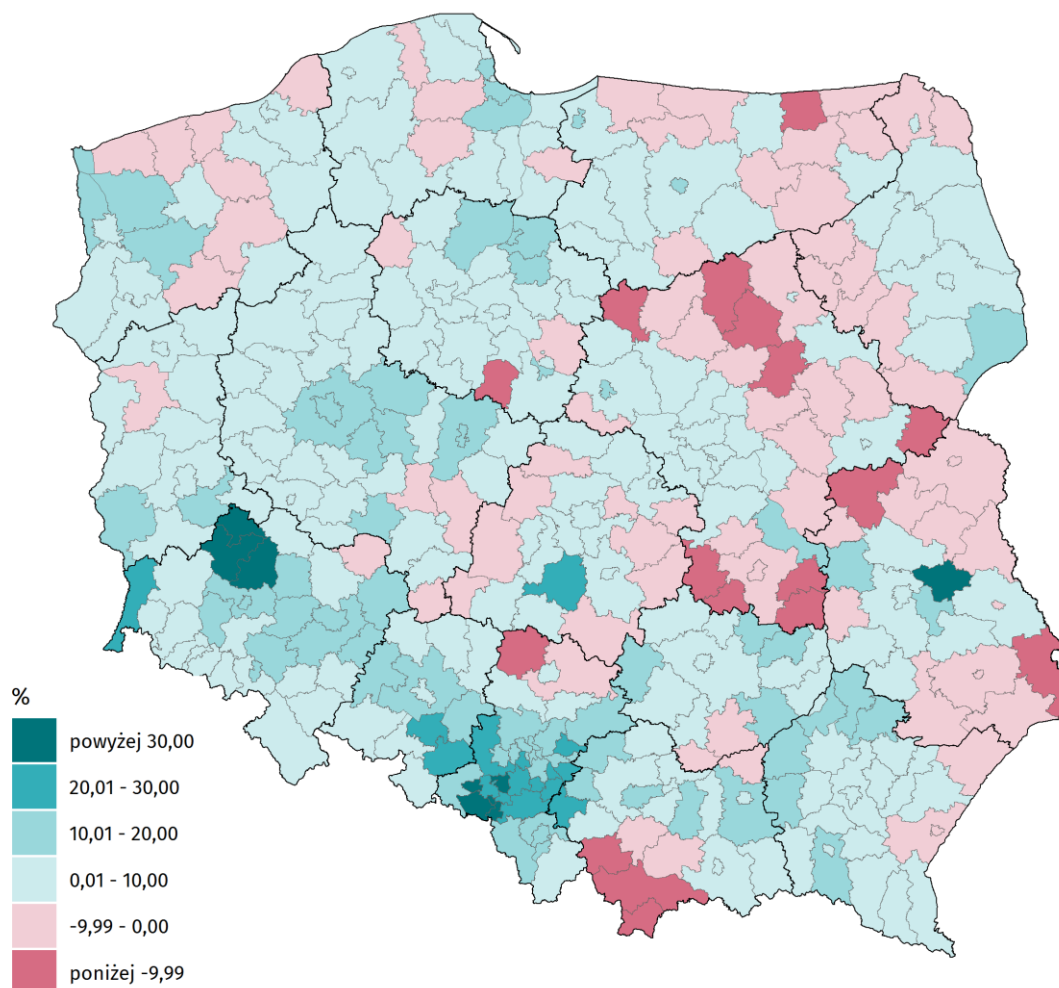
Kartogram 12. Procentowa różnica przeciętnych wynagrodzeń brutto zatrudnionych mężczyzn i kobiet wypłacanych w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.



Kartogram 13. Procentowa różnica mediany wynagrodzeń brutto zatrudnionych mężczyzn i kobiet wypłaconych w czerwcu 2017 r. według miejsca zamieszkania



Źródło: Opracowanie własne.

2. Metodologia badania *Survey methodology*

Ze względu na złożoność i szeroki zakres praca badawcza została zrealizowana w czterech etapach. Każdy z etapów 1-3 realizował konkretny cel szczegółowy – na potrzeby niniejszej pracy badawczej nazwano je modułami. W pierwszym etapie zrealizowano moduł I, czyli opracowano metodologię i naliczono liczbę pracujących w gospodarce narodowej według miejsca zamieszkania oraz zweryfikowano metodologię i naliczono **liczbę pracujących w gospodarce narodowej według głównego miejsca pracy do poziomu NTS 4** w podziale na płeć i grupy sekcji PKD. Moduł II, mający na celu opracowanie metodologii i naliczenie **stopy bezrobocia rejestrowanego do poziomu NTS 5** w podziale na płeć, zrealizowano w etapie drugim. W trzecim etapie opracowano metodologię i naliczono **miary wynagrodzeń oraz miary koncentracji rozkładu wynagrodzeń do poziomu NTS 4** w podziale na płeć i grupy sekcji PKD, czyli wykonano zadania przewidziane w module III. Każdy z etapów został zakończony utworzeniem pośredniego raportu metodologicznego z wykonanych prac. W czwartym i ostatnim etapie pracy badawczej dokonano weryfikacji i ewentualnych korekt metodologii i naliczeń wykonanych w poprzednich etapach. Było to konieczne ze względu na wzajemne powiązania poszczególnych modułów. Wnioski wyciągnięte z realizacji każdego z modułów mogły mieć zastosowanie także w pracach realizowanych w pozostałych modułach. Ponadto etap ten obejmował również dokonanie analizy otrzymanych wyników oraz opracowanie końcowego raportu metodologiczno-analitycznego i jego streszczenia w języku nietechnicznym.

Aby osiągnąć cel pracy badawczej przyjęto następujące postępowanie:

- 1) Jako badaną populację przyjęto zbiór osób ubezpieczonych w KRUS lub w ZUS.
- 2) Zidentyfikowano w zbiorach KRUS i ZUS osoby uznane wg stanu na ostatni dzień każdego z okresów referencyjnych badania za:
 - a) pracujące, zgodnie z wprowadzoną w niniejszej pracy definicją pracujących wg AZD
 - b) zarejestrowane jako bezrobotne.
- 3) Dla wszystkich osób uznanych za pracujące ustalono pracę główną.
- 4) Dokonano naliczeń liczby pracujących w gospodarce narodowej w głównym miejscu pracy wg AZD według m.in.: płci, grup wieku, sekcji PKD i miejsca zamieszkania.
- 5) Wyliczono stopę bezrobocia rejestrowanego (jako wypadkową liczby pracujących wg AZD i liczby bezrobotnych zarejestrowanych w każdej z agregacji) według m.in.: płci, grup wieku i miejsca zamieszkania.
- 6) Zidentyfikowano w zbiorach ZUS osoby uznane za zatrudnione w okresie referencyjnym badania, zgodnie z wprowadzoną w niniejszej pracy definicją zatrudnionych wg AZD.
- 7) Dla każdego z zatrudnionych ustalono rzeczywisty wymiar etatu w okresie referencyjnym badania zgodnie z zawartymi w publikacji „Zasady metodyczne statystyki rynku pracy i wynagrodzeń”² wytycznymi do obliczania przeciętnej liczby zatrudnionych.
- 8) Oszacowano rzeczywiste wynagrodzenie brutto wg AZD wypłacone osobom zatrudnionym wg AZD w każdym z okresów referencyjnych badania.
- 9) Naliczono miary wynagrodzeń wg AZD oraz miary koncentracji rozkładu wynagrodzeń wg AZD według m.in.: płci, grup wieku, sekcji PKD i miejsca zamieszkania.

²<http://stat.gov.pl/obszary-tematyczne/rynek-pracy/>

2.1. Podstawowe pojęcia dotyczące rynku pracy

Basic concepts concerning the labour market

Przez pojęcie **pracujących w gospodarce narodowej**³ rozumie się osoby wykonujące pracę przynoszącą im zarobek (w formie wynagrodzenia za pracę) lub dochód.

Do pracujących zalicza się:

- 1) osoby zatrudnione na podstawie stosunku pracy (umowa o pracę, powołanie, mianowanie, wybór lub stosunek służbowy);
- 2) pracodawców i pracujących na własny rachunek, a mianowicie:
 - a) właścicieli, współwłaścicieli i dzierżawców gospodarstw indywidualnych w rolnictwie (łącznie z pomagającymi członkami ich rodzin), z pewnymi wyłączeniami w latach 2002–2009;
 - b) właścicieli i współwłaścicieli (łącznie z pomagającymi członkami ich rodzin; z wyłączeniem wspólników spółek, którzy nie pracują w spółce) podmiotów prowadzących działalność gospodarczą poza gospodarstwami indywidualnymi w rolnictwie,
 - c) inne osoby pracujące na własny rachunek, np. osoby wykonujące wolne zawody;
- 3) osoby wykonujące pracę nakładczą;
- 4) agentów (łącznie z pomagającymi członkami ich rodzin oraz osobami zatrudnionymi przez agentów);
- 5) członków spółdzielni produkcji rolniczej (rolniczych spółdzielni produkcyjnych i spółdzielni powstałych na ich bazie oraz spółdzielni kółek rolniczych);
- 6) duchownych pełniących obowiązki duszpasterskie.

Przez pojęcie **gospodarstwo indywidualne**⁴ rozumie się gospodarstwo rolne użytkowane przez osobę fizyczną. Gospodarstwa indywidualne obejmują:

- 1) gospodarstwa o powierzchni 1 ha i więcej użytków rolnych,
- 2) gospodarstwa o powierzchni poniżej 1 ha użytków rolnych (w tym nieposiadające użytków rolnych) prowadzące działy specjalne produkcji rolnej lub produkcję o znaczącej (określonej odpowiednimi progami) skali: drzew owocowych, krzewów owocowych, szkótek sadowniczych i ozdobnych, warzyw i truskawek gruntowych, chmielu, tytoniu, kóz oraz dzikich zwierząt utrzymywanych w gospodarstwie dla produkcji mięsa (np. dziki, sarny, daniele).

Pod pojęciem **ludność cywilna aktywna zawodowo**⁵ należy rozumieć osoby pracujące w jednostkach sektora publicznego i prywatnego, jak również osoby bezrobotne.

Przez pojęcie **bezrobotni zarejestrowani**⁶ rozumie się osoby, które ukończyły 18 lat i nie osiągnęły wieku emerytalnego, niezatrudnione i niewykonywujące innej pracy zarobkowej, zdolne i gotowe do podjęcia zatrudnienia w pełnym wymiarze czasu pracy i zarejestrowane we właściwym dla miejsca zameldowania (stałego lub czasowego) powiatowym urzędzie pracy oraz poszukujące zatrudnienia lub innej pracy zarobkowej.

Do osób bezrobotnych należy zaliczyć osoby niepełnosprawne zdolne i gotowe do podjęcia zatrudnienia co najmniej w połowie tego wymiaru czasu pracy, nieuczące się w szkole, z wyjątkiem szkół dla dorosłych (lub przystępujące do egzaminu eksternistycznego z zakresu tej szkoły) lub szkół wyższych w systemie studiów niestacjonarnych, oraz poszukujące zatrudnienia, z dodatkowymi wyłączeniami dotyczącymi źródeł, zapisanymi w ustawie.

Pod pojęciem **stopy bezrobocia rejestrowanego**⁷ należy rozumieć udział zarejestrowanych bezrobotnych w cywilnej ludności aktywnej zawodowo, tj. bez pracowników jednostek budżetowych prowadzących działalność w zakresie obrony narodowej i bezpieczeństwa publicznego.

Pod pojęciem **zatrudnieni**⁸ należy rozumieć osoby zatrudnione na podstawie stosunku pracy lub stosunku służbowego na czas określony (w tym zatrudnione sezonowo i dorywczo) i nieokreślony, w pełnym i niepełnym wymiarze czasu pracy (inaczej: pracownicy najemni). Przy ustalaniu stanu zatrudnienia na dany dzień w osobach fizycznych brane są pod uwagę osoby pełnozatrudnione oraz osoby niepełnozatrudnione.

³ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/3399,pojecie.html>

⁴ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/3138,pojecie.html>

⁵ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/2392,pojecie.html>

⁶ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/2958,pojecie.html>

⁷ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/2390,pojecie.html>

⁸ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/825,pojecie.html>

Pod pojęciem **przeciętne zatrudnienie**⁹ należy rozumieć średnią wielkość zatrudnienia obliczoną dla badanego okresu (np. miesiąca, kwartału, roku) na podstawie ewidencyjnego stanu zatrudnienia. Przeciętne zatrudnienie w badanym okresie uwzględnia pracowników pełnozatrudnionych oraz niepełnozatrudnionych w przeliczeniu na pełne etaty.

Pod pojęciem **wynagrodzenie brutto**¹⁰ należy rozumieć całkowitą wartość ujęcia wynagrodzeń, tj. łącznie z zaliczkami na poczet podatku dochodowego od osób fizycznych oraz od 1999 r. ze składkami na obowiązkowe ubezpieczenia społeczne (emerytalne, rentowe, chorobowe) płatnymi przez ubezpieczonego pracownika.

Pod pojęciem **wynagrodzenie ogółem**¹¹ należy rozumieć wypłaty pieniężne wypłacane pracownikom lub innym osobom fizycznym, stanowiące wydatki ponoszone przez pracodawców na opłacenie wykonywanej pracy, niezależnie od źródeł ich finansowania oraz bez względu na podstawę stosunku pracy bądź innego stosunku prawnego lub czynności prawnej, na podstawie których jest świadczona praca lub pełniona służba.

Pod pojęciem **wynagrodzenia za czas choroby**¹² należy rozumieć wynagrodzenia za czas niezdolności pracownika do pracy wskutek:

- 1) choroby lub odosobnienia w związku z chorobą zakaźną trwającą łącznie do 33 dni w ciągu roku kalendarzowego,
- 2) wypadku przy pracy, wypadku w drodze do pracy lub z pracy albo choroby zawodowej lub choroby przypadającej w czasie ciąży - trwającej łącznie do 33 dni w ciągu roku kalendarzowego.

Pod pojęciem **zasitek chorobowy**¹³ należy rozumieć świadczenie pieniężne wypłacane za okresy niezdolności do pracy spowodowanej chorobą lub wypadkiem oraz odosobnieniem z powodu choroby zakaźnej, przysługujący począwszy od 15 lub 34 dnia skumulowanej w roku niezdolności do pracy. Za okres pierwszych 14 lub 33 dni niezdolności do pracy ubezpieczony zachowuje prawo do wynagrodzenia finansowanego ze środków pracodawcy lub w pewnych okolicznościach z Funduszu Gwarantowanych Świadczeń Pracowniczych. Zasiłek ten przysługuje przez okres trwania niezdolności do pracy, nie dłużej jednak niż przez 182 dni, a jeśli niezdolność do pracy została spowodowana gruźlicą lub występuje w trakcie ciąży - nie dłużej niż przez 270 dni.

Pod pojęciem **przeciętne miesięczne wynagrodzenie brutto**¹⁴ należy rozumieć stosunek sumy wynagrodzeń osobowych brutto, honorariów wypłaconych niektórym grupom pracowników za prace wynikające z umowy o pracę, wypłat z tytułu udziału w zysku lub w nadwyżce bilansowej w spółdzielniach oraz dodatkowych wynagrodzeń rocznych dla pracowników jednostek sfery budżetowej do przeciętnej liczby zatrudnionych w danym okresie; po wyeliminowaniu osób wykonujących pracę nakładczą oraz zatrudnionych za granicą.

Pod pojęciem **średnia arytmetyczna**¹⁵ - należy rozumieć sumę wartości zmiennej wszystkich jednostek badanej zbiorowości podzielona przez liczbę tych jednostek.

Pod pojęciem **dominanta (modalna, wartość najczęstsza)**¹⁶ - należy rozumieć taką wartość zmiennej, która w danym rozkładzie empirycznym występuje najczęściej.

Pod pojęciem **kwantyle**¹⁷ - należy rozumieć zbiór wartości cechy badanej zbiorowości dzielący ogólną liczebność próby na n równych części.

Pod pojęciem **decyl**¹⁸ - należy rozumieć wartość cechy, która dzieli uporządkowaną zbiorowość statystyczną na 10 równych części co 10%; np. decyl pierwszy oznacza, że 10% jednostek ma wartości cechy mniejsze bądź równe od decyla pierwszego, a 90% jednostek wartości cechy równe lub większe od decyla pierwszego.

⁹ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/376,pojecie.html>

¹⁰ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/3390,pojecie.html>

¹¹ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/607,pojecie.html>

¹² <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/743,pojecie.html>

¹³ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/648,pojecie.html>

¹⁴ <http://stat.gov.pl/metainformacje/sloownik-pojec/pojecia-stosowane-w-statystyce-publicznej/693,pojecie.html>

¹⁵ M. Sobczyk, *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2018.

¹⁶ Ibidem.

¹⁷ Ibidem.

¹⁸ Ibidem.

Pod pojęciem **kwartyle**¹⁹ – należy rozumieć trzy wartości cechy badanej zbiorowości, które dzielą uporządkowany zbiór danych na cztery równe pod względem liczebności części.

Pod pojęciem **kwartyl pierwszy**²⁰ – należy rozumieć wartość cechy, która dzieli uporządkowaną zbiorowość statystyczną na dwie części w taki sposób, że 25% jednostek zbiorowości ma wartości cechy mniejsze lub równe wartości kwartyla pierwszego, a 75% jednostek badanej zbiorowości ma wartości cechy większe lub równe wartości kwartyla pierwszego.

Pod pojęciem **kwartyl trzeci**²¹ – należy rozumieć wartość cechy, która dzieli uporządkowaną zbiorowość statystyczną na dwie części w taki sposób, że 75% jednostek zbiorowości ma wartości cechy mniejsze lub równe wartości kwartyla trzeciego, a 25% jednostek badanej zbiorowości ma wartości cechy większe lub równe wartości kwartyla trzeciego.

Pod pojęciem **mediana (kwartyl drugi Q2, wartość środkowa)**²² – należy rozumieć wartość cechy, która dzieli uporządkowaną zbiorowość statystyczną na dwie równoliczne części w ten sposób, że poniżej i powyżej tej wartości znajduje się jednakowa liczba obserwacji. Mediana jest kwartylem drugim. Warto podkreślić, iż mediana jako miara pozycyjna jest niewrażliwa na wartości nietypowe (odstające, ekstremalne) rozkładu, jej wartość nie zależy tym samym od wartości krańcowych rozkładu badanej cechy.

Pod pojęciem **kurtoza (współczynnik koncentracji, współczynnik skupienia)**²³ – należy rozumieć stopień względnego nachylenia wzniesienia krzywej rozkładu w sąsiedztwie wartości dominanty. Jest to względna miara spłaszczenia.

Pod pojęciem **krzywa Lorenza**²⁴ – należy rozumieć metodę graficzną polegającą na wykreśleniu tzw. wieloboku koncentracji Lorenza. W tym celu na osi odciętych odmierza się skumulowane częstości względne (w %), natomiast na osi rzędnych – procentowe skumulowane częstości względne łącznego funduszu cechy. Łącząc punkty o tych współrzędnych otrzymujemy tzw. krzywą Lorenza.

Pod pojęciem **współczynnik Giniego**²⁵ – należy rozumieć miarę nierówności rozkładu dochodów; przybiera wartość między 0 a 1. Wskaźnik ten osiągnąłby wartość 0 (rozkład jednorodny), gdyby wszystkie osoby miały ten sam dochód, natomiast wartość 1, gdyby wszystkie osoby poza jedną miały dochód zerowy. Zatem, im wyższa jest wartość wskaźnika, tym większy jest stopień koncentracji dochodów i większe jest ich zróżnicowanie.

Ponadto ilekroć w pracy badawczej jest mowa o:

- 1) **podmiocie gospodarki narodowej (jednostce lub przedsiębiorcy)** należy przez to rozumieć jednostkę prawną, jednostkę organizacyjną nieposiadającą osobowości prawnej lub osobę fizyczną prowadzącą działalność gospodarczą,
- 2) **obserwacjach** należy przez to rozumieć poszczególnych pracujących i bezrobotnych stanowiących rekordy z administracyjnych baz danych; w ten sposób, z nielicznymi wyjątkami, każda obserwacja będzie posiadała odpowiadający jej numer PESEL,
- 3) **zmiennych** należy przez to rozumieć właściwości opisujące poszczególne obserwacje, które są mierzone, kontrolowane lub ustalone, pod względem których poszczególne obserwacje różnią się między sobą, np.: płeć, rok urodzenia, miejsce zamieszkania lub miejsce pracy; zmienne mogą przyjmować różne wartości, np.: dla zmiennej płeć może to być mężczyzna i kobieta; w ten sposób zmienne opisane są przez „wartości zmiennych” (w pracy nie będzie wykorzystywane natomiast pojęcie „cech informacyjnych” ze względu na brak jednoznaczności tego pojęcia. Nie jest bowiem intuicyjnie możliwe do rozpoznania, czy pojęcie cechy informacyjnej dotyczy zmiennej czy jej wartości),

¹⁹ Ibidem.

²⁰ Ibidem.

²¹ Ibidem.

²² Ibidem.

²³ Ibidem.

²⁴ Ibidem.

²⁵ Ibidem.

- 4) **ubezpieczonym** należy przez to rozumieć osobę fizyczną podlegającą co najmniej jednemu z ubezpieczeń społecznych, tj. ubezpieczeniom emerytalnemu, rentowemu, chorobowemu, wypadkowemu lub ubezpieczeniu zdrowotnemu,
- 5) **podstawie wymiaru składek na ubezpieczenia społeczne** należy przez to rozumieć podstawę wymiaru składek na ubezpieczenia społeczne pracownika, którą stanowią przychody ze stosunku pracy i stosunków pokrewnych, czyli wszelkiego rodzaju wypłaty pieniężne oraz pieniężna wartość świadczeń w naturze bądź ich ekwiwalenty (bez względu na źródło ich finansowania), a w szczególności wynagrodzenia zasadnicze, wynagrodzenia za godziny nadliczbowe, różnego rodzaju dodatki, nagrody, ekwiwalenty za niewykorzystany urlop, a także świadczenia pieniężne ponoszone za pracownika, jak również wartość innych nieodpłatnych świadczeń lub świadczeń częściowo odpłatnych,
- 6) **wymiarze czasu pracy** należy przez to rozumieć wymiar czasu pracy, który zgodnie z *Zasadami wypełniania dokumentów ubezpieczeniowych*²⁶ określony w umowie o pracę powinien być zapisany w dokumentach ubezpieczeniowych w postaci ułamka zwykłego np.:
- 1/1 - dla pełnego wymiaru czasu pracy,
 - 1/2 - dla połowy wymiaru czasu pracy,
 - 3/4 - dla trzech czwartych wymiaru czasu pracy,
- 7) **kodzie tytułu ubezpieczenia** należy przez to rozumieć sześciocyfrowy kod tytułu ubezpieczenia (podawany w zgłoszeniu do ubezpieczeń), którym posługuje się Zakład Ubezpieczeń Społecznych dla rozróżnienia poszczególnych grup ubezpieczonych. Znaki te oznaczają:
- pierwsze cztery znaki określają podmiot podstawowy wraz z rozszerzeniem,
 - znak piąty oznacza ustalone bądź nieustalone prawo do emerytury lub renty,
 - szósty znak określa stopień niepełnosprawności,
- Szczegółowy wykaz kodów znajduje się w załączniku nr 33.
- 8) **kodzie świadczenia/przerwy (ostatnim w miesiącu rozliczeniowym w kontekście płatnika)** należy przez to rozumieć kod, który zawiera informacje o przerwach w opłacaniu składek. Przerwy w opłacaniu składek mogą być spowodowane: chorobą, urlopem bezpłatnym, macierzyństwem itp.,
Szczegółowy wykaz kodów znajduje się w załączniku nr 33.
- 9) **okresie „od” za jaki nastąpiła wypłata świadczenia lub przerwa w opłacaniu składki** należy przez to rozumieć początkową datę okresu przerwy w opłacaniu składek lub okresu wypłaconych świadczeń/wynagrodzeń za czas absencji chorobowej. Okres ten nie może wykraczać poza ostatni dzień kalendarzowy miesiąca, za który jest sporządzany raport ZUS RSA – oznacza to, iż „data od” (dzień/ miesiąc/ rok) może być wcześniejsza niż pierwszy dzień kalendarzowy miesiąca, za który jest sporządzany raport, i nie może być późniejsza niż ostatni dzień miesiąca kalendarzowego, za który jest sporządzany raport,
- 10) **okresie „do” za jaki nastąpiła wypłata świadczenia lub przerwa w opłacaniu składki** należy przez to rozumieć końcową datę okresu przerwy w opłacaniu składek lub okresu wypłaconych świadczeń/wynagrodzeń za czas absencji chorobowej. Okres ten nie może wykraczać poza ostatni dzień kalendarzowy miesiąca, za który jest sporządzany raport ZUS RSA – oznacza to, iż „data do” (dzień/ miesiąc/ rok) może być wcześniejsza niż pierwszy dzień kalendarzowy miesiąca, za który jest sporządzany raport, i nie może być późniejsza niż ostatni dzień miesiąca kalendarzowego, za który jest sporządzany raport,
- 11) **kodzie informacji o przekroczeniu rocznej podstawy wymiaru składek na ubezpieczenie emerytalne i rentowe** należy przez to rozumieć kod, który zawiera informację, czy w danym miesiącu przychód ubezpieczonego przekroczył kwotę rocznej podstawy wymiaru składek na ubezpieczenia emerytalne i rentowe.

²⁶ Dostępne na stronie internetowej ZUS: www.zus.pl

- Kod informacji o przekroczeniu rocznej podstawy wymiaru składek na ubezpieczenia emerytalne i rentowe przyjmuje następujące wartości:
- a) 1 – w przypadku, gdy informację o przekroczeniu rocznej podstawy wymiaru składek przekazał ubezpieczony;
 - b) 2 - w przypadku, gdy informację o przekroczeniu rocznej podstawy wymiaru składek przekazał płatnik składek;
 - c) 3 - w przypadku, gdy informację o przekroczeniu rocznej podstawy wymiaru składek przekazał Zakład Ubezpieczeń Społecznych.
- 12) **podatniku** należy przez to rozumieć osobę fizyczną, osobę prawną lub jednostkę organizacyjną niemającą osobowości prawnej, podlegającą na mocy ustaw podatkowych obowiązkowi podatkowemu,
- 13) **płatniku** należy przez to rozumieć:
- a) jednostkę opłacającą składki na ubezpieczenie społeczne lub zdrowotne za co najmniej jednego ubezpieczonego (płatnik czynny),
 - b) jednostkę nie będącą aktualnie płatnikiem czynnym, ale posiadającą zadłużenie z tytułu nieopłaconych składek na ubezpieczenie,
 - c) osobę fizyczną, osobę prawną lub jednostkę organizacyjną niemającą osobowości prawnej, obowiązującą na podstawie przepisów prawa podatkowego do obliczenia i pobrania od podatnika podatku i wpłacenia go we właściwym terminie organowi podatkowemu,
- 14) **ustawie o promocji zatrudnienia i instytucjach rynku pracy** należy przez to rozumieć ustawę z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy (Dz. U. z 2017 r. poz. 1065, z późn. zm.),
- 15) **ustawie o ubezpieczeniu społecznym rolników** należy przez to rozumieć ustawę z dnia 20 grudnia 1990 r. o ubezpieczeniu społecznym rolników (Dz.U. z 2016 r. poz. 277 z późn. zm.),
- 16) **ustawie o systemie ubezpieczeń społecznych** należy przez to rozumieć ustawę z dnia 13 października 1998 r. o systemie ubezpieczeń społecznych (Dz.U. z 2017 r. poz. 1778),
- 17) **umowie uaktywniającej** należy przez to rozumieć rodzaj umowy o świadczenie usług, do której zgodnie z przepisami Kodeksu cywilnego stosuje się przepisy dotyczące zlecenia, na mocy której można zatrudnić nianię, która sprawuje opiekę nad dzieckiem w wieku od 20 tygodnia życia do trzech lat,
- 18) **cudzoziemcu** należy przez to rozumieć osobę nieposiadającą obywatelstwa polskiego,
- 19) **członku rodziny** należy przez to rozumieć:
- a) osobę pozostającą z obywatelem polskim lub cudzoziemcem w związku małżeńskim uznawanym przez prawo Rzeczypospolitej Polskiej,
 - b) zstępного, obywatela polskiego lub cudzoziemca, w wieku do 21 lat lub pozostającego na jego utrzymaniu,
- 20) **poszukującym pracy** oznacza to osobę, o której mowa w art. 1 ust. 3 pkt 1–3 (Dz. U. z 2017 r. poz. 1065, z późn. zm.), lub cudzoziemca – członka rodziny obywatela polskiego, poszukujących zatrudnienia, innej pracy zarobkowej lub innej formy pomocy określonej w ustawie, zarejestrowanych w powiatowym urzędzie pracy,
- 21) **pozarolniczej działalności** oznacza to pozarolniczą działalność w rozumieniu przepisów o systemie ubezpieczeń społecznych,
- 22) **głównym miejscu pracy**, zwanym dalej **miejscem siedziby pracy głównej** oznacza to lokalizację (np. miejscowość, adres, współrzędne geograficzne) siedziby jednostki, na rzecz której osoba wykonuje pracę główną,
- 23) **miejscu zamieszkania** oznacza to:
- a) dla bezrobotnych w sprawozdaniach MRPIPS informacje o miejscu zamieszkania ustalone na podstawie czasowego lub stałego miejsca zameldowania lub pobytu osoby bezrobotnej, która

- nie jest zameldowana, zgodnie z wykazem identyfikatorów i nazw jednostek podziału terytorialnego kraju, określonych w rozporządzeniu Rady Ministrów z dnia 15 grudnia 1998 r. w sprawie szczegółowych zasad prowadzenia, stosowania i udostępniania krajowego rejestru urzędowego podziału terytorialnego kraju oraz związanych z tym obowiązków organów administracji rządowej i jednostek samorządu terytorialnego (Dz. U. Nr 157, poz.1031, z późn. zm.),
- b) dla danych o pracujących (w tym zatrudnionych) z ZUS lokalizację (np. miejscowość, adres, współrzędne geograficzne), którą osoba uznaje za miejsce swojego pobytu,
- c) w podziale na:
- miasto, czyli gminy miejskie, miasta w gminach miejsko-wiejskich oraz miasta na prawach powiatu,
 - wieś, czyli gminy wiejskie oraz obszary wiejskie w gminach miejsko-wiejskich.
- 24) **poziomie NTS 4** należy rozumieć, że oznacza to czwarty poziom podziału terytorialnego kraju obejmujący powiaty,
- 25) **poziomie NTS 5** należy rozumieć, że oznacza to piąty poziom podziału terytorialnego kraju obejmujący gminy,
- 26) **rodzaju ubezpieczenia** należy przez to rozumieć:
- a) jedno z ubezpieczeń społecznych, tj. emerytalne, rentowe, chorobowe, wypadkowe,
 - b) ubezpieczenie zdrowotne,
- 27) **typie ubezpieczenia** należy przez to rozumieć obowiązkowe lub dobrowolne podleganie jednemu z rodzajów ubezpieczeń,
- 28) **użytkach rolnych** należy przez to rozumieć: zasiewy; łąki i pastwiska trwałe; uprawy trwałe, w tym sady; ogrody przydomowe (bez powierzchni przeznaczonej na rekreację); grunty ugorowane oraz pozostałe użytki,
- 29) **Polskiej Klasyfikacji Działalności 2007 (PKD 2007)**, oznacza to klasyfikację utworzoną na podstawie klasyfikacji macierzystej NACE Rev. 2, w której najwyższy poziom obejmuje dwadzieścia jeden sekcji, tj.:
- sekcja A - rolnictwo, leśnictwo, łowiectwo i rybactwo
 - sekcja B - górnictwo i wydobywanie
 - sekcja C - przetwórstwo przemysłowe
 - sekcja D - wytwarzanie i zaopatrywanie w energię elektryczną, gaz, parę wodną, gorącą wodę i powietrze do układów klimatyzacyjnych
 - sekcja E - dostawa wody; gospodarowanie ściekami i odpadami oraz działalność związana z rekultywacją
 - sekcja F - budownictwo
 - sekcja G - handel hurtowy i detaliczny; naprawa pojazdów samochodowych, włączając motocykle
 - sekcja H - transport i gospodarka magazynowa
 - sekcja I - działalność związana z zakwaterowaniem i usługami gastronomicznymi
 - sekcja J - informacja i komunikacja
 - sekcja K - działalność finansowa i ubezpieczeniowa
 - sekcja L - działalność związana z obsługą rynku nieruchomości
 - sekcja M - działalność profesjonalna, naukowa i techniczna
 - sekcja N - działalność w zakresie usług administrowania i działalność wspierająca
 - sekcja O - administracja publiczna i obrona narodowa; obowiązkowe zabezpieczenia społeczne
 - sekcja P - edukacja
 - sekcja Q - opieka zdrowotna i pomoc społeczna
 - sekcja R - działalność związana z kulturą, rozrywką i rekreacją
 - sekcja S - pozostała działalność usługowa
 - sekcja T - gospodarstwa domowe zatrudniające pracowników; gospodarstwa domowe produkujące wyroby i świadczące usługi na własne potrzeby
 - sekcja U - organizacje i zespoły eksterytorialne;

- 30) **grupowaniu sekcji PKD 2007-** w niniejszej pracy badawczej oznacza prezentowanie danych według sekcji rodzaju działalności jednostek w pięciu grupach:
- Rolnictwo, leśnictwo, łowiectwo i rybactwo (sekcja A).
 - Przemysł i budownictwo (sekcje B, C, D, E, F).
 - Handel, naprawa pojazdów samochodowych, transport i gospodarka magazynowa, zakwaterowanie i gastronomia, informacja i komunikacja (sekcje G, H, I, J).
 - Działalność finansowa i ubezpieczeniowa oraz obsługa rynku nieruchomości (sekcje K, L).
 - Pozostałe usługi (sekcje M, N, O, P, Q, R, S, T, U).

Jednym z istotnych z punktu widzenia gospodarki narodowej sposobem prezentacji danych według sekcji PKD 2007 jest klasyfikowanie według **sektorów gospodarki narodowej**:

- Sektor rolniczy (sekcja A).
 - Sektor przemysłowy (sekcje B, C, D, E, F).
 - Sektor usługowy (sekcje G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U);
- 31) **ubezpieczeniach społecznych**, należy przez to rozumieć ubezpieczenia społeczne, które obejmują: ubezpieczenie emerytalne, ubezpieczenia rentowe, ubezpieczenie w razie choroby i macierzyństwa (chorobowe), ubezpieczenie z tytułu wypadków przy pracy i chorób zawodowych (wypadkowe);
- 32) **ubezpieczeniach zdrowotnych**, należy przez to rozumieć ubezpieczenie zdrowotne obejmujące osoby posiadające obywatelstwo polskie i zamieszkujące na terytorium Polski oraz cudzoziemców (posiadających kartę stałego pobytu lub kartę czasowego pobytu wydaną w związku z udzieleniem statusu uchodźcy), jeśli podlegają obowiązkowi ubezpieczenia zdrowotnego, ubezpieczają się dobrowolnie lub są członkami rodziny osób obowiązkowo lub dobrowolnie objętych ubezpieczeniem zdrowotnym.

2.2. Charakterystyka administracyjnych zbiorów danych wykorzystanych do wyznaczenia i naliczenia wskaźników

Characteristics of administrative data sets used for defining and calculating indicators

Administracyjne źródła danych będące podstawą do naliczenia wskaźników na potrzeby pracy badawczej pochodzą z:

- Kasy Rolniczego Ubezpieczenia Społecznego²⁷ - nowy System Informatyczny Ubezpieczeń (nSIU, dawniej KRUSna),
- Zakładu Ubezpieczeń Społecznych²⁸ - Kompleksowy System Informatyczny (KSI ZUS):

²⁷ Podstawa prawna funkcjonowania systemu: ustawa z dnia 20 grudnia 1990 r. o ubezpieczeniu społecznym rolników (tekst jednolity (dalej tj.) Dz. U. z 2016 r., poz. 277 z późn. zm.); ustawa z dnia 13 października 1998 r. o systemie ubezpieczeń społecznych (tj. Dz. U. z 2016r., poz. 963 z późn. zm.); ustawa z dnia 29 sierpnia 1997 r. Ordynacja podatkowa (tj. Dz. U. z 2017r., poz. 201 z późn. zm.); przepisy o koordynacji dotyczące zabezpieczenia społecznego: rozporządzenie Parlamentu Europejskiego i Rady (WE) Nr 883/2004 z dnia 29 kwietnia 2004 r. w sprawie koordynacji systemów zabezpieczenia społecznego (Dz. U. UE L 166 z dnia 30 kwietnia 2004 r. z późn. zm.; Dz. Urz. UE Polskie wydanie specjalne, rozdz. 5, t. 5 z późn. zm.) i rozporządzenie Parlamentu Europejskiego i Rady (WE) Nr 987/2009 z dnia 16 września 2009 r. dotyczącym wykonywania rozporządzenia (WE) Nr 883/2004 z dnia 29 kwietnia 2004 r. w sprawie koordynacji systemów zabezpieczenia społecznego (Dz. U. UE L 284 z dnia 30 października 2009 r.); ustawa z dnia 29 sierpnia 1997 r. o ochronie danych osobowych (tj. Dz. U. z 2016 r., poz. 922); inne akty prawne mające zastosowanie w sprawach z zakresu ubezpieczenia społecznego rolników; inne akty prawne nakładające na Kasę obowiązek udostępniania danych ze zbiorów danych.

²⁸ Podstawa prawna funkcjonowania systemu: ustawa z dnia 13 października 1998 r. o systemie ubezpieczeń społecznych (tj. Dz. U. 2016, poz. 963 z późn. zm.); ustawa z dnia 25 czerwca 1999 r. o świadczeniach pieniężnych z ubezpieczenia społecznego w razie choroby i macierzyństwa (tj. Dz. U. 2017, poz. 1368), rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 21 grudnia 2009 r. w sprawie szczegółowego zakresu danych zawartych w centralnych rejestrach prowadzonych przez Zakład Ubezpieczeń Społecznych (tj. Dz. U. 2013, poz. 219), rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 23 października 2009 r. w sprawie określenia wzorów zgłoszeń do ubezpieczeń społecznych i ubezpieczenia zdrowotnego, imiennych raportów miesięcznych i imiennych raportów miesięcznych korygujących, zgłoszeń płatnika, deklaracji rozliczeniowych i deklaracji rozliczeniowych korygujących, zgłoszeń danych o pracy w szczególnych warunkach lub o szczególnym charakterze oraz innych dokumentów (tj. Dz. U. 2016, poz. 222 z późn. zm.).

- a) Centralny Rejestr Płatników Składek – CRPS,
 - b) Centralny Rejestr Ubezpieczonych – CRU;
- 3) Ministerstwa Finansów - Centralny Rejestr Podmiotów Krajowa Ewidencja Podatników (CRP KEP).

Tablica 2. Charakterystyka systemów informacyjnych

Kasa Rolniczego Ubezpieczenia Społecznego – nSIU - nowy System Informatyczny Ubezpieczeń (dawniej KRUSna)

Cel systemu	System w merytorycznym zakresie ubezpieczeń służy do ewidencjonowania osób objętych ubezpieczeniem społecznym rolników, okresów podlegania ubezpieczeniu i opłacania składek za każdego ubezpieczonego oraz spraw związanych z ich dochodzeniem i udzielaniem ulg w spłacie należności z tytułu składek.
Zakres podmiotowy	Płatnicy składek oraz osoby objęte ubezpieczeniem społecznym rolników.
Zakres przedmiotowy	Dane dotyczące m.in. okresów podlegania ubezpieczeniu społecznemu i zdrowotnemu rolników i domowników, wymiaru i okresu poboru składek na ubezpieczenie społeczne i zdrowotne rolników oraz świadczeń pieniężnych wypłacanych z ubezpieczenia emerytalno-rentowego, ubezpieczenia wypadkowego, chorobowego i macierzyńskiego oraz świadczeń pozaubezpieczeniowych.
Źródła danych	Dokumenty przedkładane przez rolników i osoby objęte ubezpieczeniem społecznym rolników oraz inne dokumenty źródłowe uzyskiwane przez jednostki organizacyjne KRUS bądź nadsyłane przez inne jednostki organizacyjne. Aktualizacja gromadzonych w systemie danych dokonywana jest na bieżąco w oparciu o ww. dokumenty.

Zakład Ubezpieczeń Społecznych - Kompleksowy System Informatyczny KSI ZUS

Cel systemu	System obejmuje swoją funkcjonalnością procesy zachodzące w ZUS oraz procesy, jakie mają miejsce w relacji ZUS z podmiotami zewnętrznymi, z jego otoczenia, w zakresie niezbędnym do realizacji zadań statutowych Zakładu.
Zakres podmiotowy	Płatnicy składek oraz osoby objęte ubezpieczeniem społecznym, z wyłączeniem osób podlegających ubezpieczeniom społecznym rolników. Zgodnie z art. 4 ust. 2 ustawy o systemie ubezpieczeń społecznych płatnikami składek na ubezpieczenia społeczne są: 1) pracodawca - w stosunku do pracowników i osób odbywających służbę zastępczą, oraz jednostka organizacyjna lub osoba fizyczna pozostająca z inną osobą fizyczną w stosunku prawnym uzasadniającym objęcie tej osoby ubezpieczeniami społecznymi, w tym z tytułu przebywania na urlopie wychowawczym albo pobierania zasiłku macierzyńskiego, z wyłączeniem osób, którym zasiłek macierzyński wypłaca ZUS; 2) jednostka wypłacająca świadczenie socjalne, zasiłki socjalne oraz wynagrodzenia przysługujące w okresie korzystania ze świadczenia górniczego lub w okresie korzystania ze stypendium na przekwalifikowanie - w stosunku do osób pobierających świadczenia socjalne wypłacane w okresie urlopu, osób pobierających zasiłek socjalny wypłacany na czas przekwalifikowania zawodowego i poszukiwania nowego zatrudnienia oraz osób pobierających

- wynagrodzenie przysługujące w okresie korzystania ze świadczenia górniczego lub w okresie korzystania ze stypendium na przekwalifikowanie;
- 3) podmiot, na rzecz którego wykonywana jest odpłatnie praca w czasie odbywania kary pozbawienia wolności lub tymczasowego aresztowania - w stosunku do osób, które ją wykonują, na podstawie skierowania do pracy, lub osób podlegających ubezpieczeniom społecznym z tytułu pobierania zasiłku macierzyńskiego, jeżeli zasiłek wypłaca ten podmiot;
 - 4) ubezpieczony zobowiązany do opłacenia składek na własne ubezpieczenia społeczne;
 - 5) Kancelaria Sejmu - w stosunku do posłów i posłów do Parlamentu Europejskiego, o których mowa w art. 1 ust. 1 ustawy z dnia 30 lipca 2004 r. o uposażeniu posłów do Parlamentu Europejskiego wybranych w Rzeczypospolitej Polskiej (Dz.U. nr 187, poz. 1925 ze zm.);
 - 6) Kancelaria Senatu - w stosunku do senatorów;
 - 7) duchowny niebędący członkiem zakonu albo przełożony domu zakonnego lub klasztoru - w stosunku do członków swych zakonów lub, za zgodą Zakładu Ubezpieczeń Społecznych, inna zwierzchnia instytucja diecezjalna lub zakonna w odniesieniu do duchownych objętych tą zgodą;
 - 8) jednostka organizacyjna podległa Ministrowi Obrony Narodowej - w stosunku do żołnierzy niezawodowych pełniących czynną służbę wojskową, z wyłączeniem żołnierzy pełniących służbę kandydacką;
 - 9) ośrodek pomocy społecznej - w stosunku do osób rezygnujących z zatrudnienia w związku z koniecznością sprawowania bezpośredniej, osobistej opieki nad długotrwale lub ciężko chorym członkiem rodziny oraz wspólnie niezamieszkującymi matką, ojcem lub rodzeństwem, jak również w stosunku do osób pobierających świadczenie pielęgnacyjne, jeżeli zostanie uprawniony do wypłaty tych świadczeń przez wójta, burmistrza lub prezydenta miasta;
 - 10) powiatowy urząd pracy - w stosunku do osób pobierających zasiłek dla bezrobotnych, świadczenie integracyjne lub stypendium. Od 1 lutego 2009 r. płatnikiem składek w stosunku do osób pobierających świadczenie integracyjne jest centrum integracji społecznej;
 - 11) powiatowy urząd pracy - w stosunku do osób pobierających stypendium na podstawie przepisów o promocji zatrudnienia i instytucjach rynku pracy w okresie odbywania studiów podyplomowych i niepozostawiania w zatrudnieniu;
 - 12) inne niż powiatowe urzędy pracy podmioty wypłacające stypendium w okresie szkolenia, stażu lub przygotowania zawodowego w ramach projektów lub programów finansowanych z udziałem środków funduszy strukturalnych Unii Europejskiej - w stosunku do osób pobierających stypendium w okresie szkolenia, stażu lub przygotowania zawodowego;
 - 13) Zakład Ubezpieczeń Społecznych - w stosunku do osób podlegających ubezpieczeniom społecznym z tytułu pobierania zasiłku macierzyńskiego albo zasiłku w wysokości zasiłku macierzyńskiego, jeżeli zasiłki te wypłaca ZUS, oraz w stosunku do osób określonych w art. 6a ust. 1 i art. 6b ust. 1;
 - 14) podmiot wypłacający stypendium sportowe - w stosunku do osób pobierających te stypendia;
 - 15) minister właściwy do spraw finansów publicznych oraz dyrektor izby

celnej - w stosunku do funkcjonariuszy celnych;

- 16) Krajowa Szkoła Administracji Publicznej - w stosunku do słuchaczy pobierających stypendium;
- 17) osoba prowadząca pozarolniczą działalność - w stosunku do osób współpracujących przy prowadzeniu tej działalności;
- 18) wojewódzki urząd pracy - w stosunku do osób, których świadczenia pracownicze finansowane są ze środków Funduszu Gwarantowanych Świadczeń Pracowniczych, jeżeli świadczenia te wypłacane są przez ten urząd;
- 19) jednostka obsługi ekonomiczno-administracyjnej utworzona przez jednostkę samorządu terytorialnego - jeżeli rozlicza i opłaca składki za ubezpieczonych wykonujących pracę w podlegających jej szkołach, przedszkolach i innych jednostkach organizacyjnych systemu oświaty;
- 20) podmiot, w którym pełniona jest służba - w odniesieniu do żołnierzy zawodowych i funkcjonariuszy oddelegowanych do pełnienia w nim służby, jeżeli podmiot ten wypłaca im uposażenie;
- 21) wójt, burmistrz lub prezydent miasta - w odniesieniu do osób otrzymujących świadczenie pielęgnacyjne, specjalny zasiłek opiekuńczy lub zasiłek dla opiekuna;
- 22) podmiot, który wypłaca świadczenie szkoleniowe po ustaniu zatrudnienia - w stosunku do osób, którym wypłaca to świadczenie;
- 23) pracownik, który na podstawie art. 21 rozporządzenia 987/2009 przejął obowiązki pracodawcy (podmiotu zagranicznego, który nie ma w Polsce siedziby ani przedstawicielstwa) w zakresie opłacania składek na własne ubezpieczenia społeczne;
- 24) podmiot, w którym działa rada nadzorcza - w stosunku do członków rad nadzorczych.

Podstawowym obowiązkiem płatnika składek jest dokonanie zgłoszenia osoby ubezpieczonej do ubezpieczeń społecznych, a w przypadku ustania tytułu ubezpieczeń - do wyrejestrowania na odpowiednich formularzach, a także - comiesięczne rozliczanie i opłacanie należnych składek w odpowiednich terminach.

Zasady podlegania ubezpieczeniom nie są jednolite dla wszystkich ubezpieczonych w ZUS i są zależne od tytułu ubezpieczenia. Nie wszyscy ubezpieczeni w ZUS podlegają obowiązkowym ubezpieczeniom społecznym. Obowiązkowe ubezpieczenia: emerytalne, rentowe, chorobowe, wypadkowe obowiązują wszystkich pracowników od dnia nawiązania do dnia ustania stosunku pracy.

Obowiązkowe ubezpieczenie emerytalne, rentowe i wypadkowe (bez obowiązkowego ubezpieczenia chorobowego) dotyczy osób:

- 1) wykonujących umowę agencyjną, umowę zlecenia lub umowę o świadczenie usług albo współpracującą przy wykonywaniu tych umów - od dnia oznaczonego w umowie jako dzień rozpoczęcia pracy do dnia rozwiązania lub wygaśnięcia umowy;
- 2) prowadzących pozarolniczą działalność lub współpracujących przy jej prowadzeniu - od dnia rozpoczęcia działalności do dnia jej zaprzestania (oprócz okresu, na który jest zawieszona działalność), a osoby współpracujące od dnia rozpoczęcia współpracy do dnia jej zakończenia;
- 3) duchowne - od dnia przyjęcia do stanu duchownego do dnia

	<p>wystąpienia z niego.</p> <p>Dla w/w osób ubezpieczenie chorobowe jest ubezpieczeniem dobrowolnym.</p> <p>Natomiast tylko do obowiązkowych ubezpieczeń emerytalnych i rentowych (bez obowiązkowych ubezpieczeń chorobowych i wypadkowych) zobowiązane są osoby:</p> <ol style="list-style-type: none"> 1) pobierające zasiłek macierzyński albo zasiłek w wysokości zasiłku macierzyńskiego - od dnia spełnienia warunków do objęcia tymi ubezpieczeniami do dnia zaprzestania spełniania warunków; 2) przebywające na urlopie wychowawczym i niemające prawa do emerytury lub renty czy innych tytułów do obowiązkowych ubezpieczeń społecznych od dnia spełnienia warunków do objęcia tymi ubezpieczeniami do dnia zaprzestania spełniania warunków; 3) pobierają świadczenie pielęgnacyjne, specjalny zasiłek opiekuńczy albo zasiłek dla opiekuna - przez okres niezbędny do uzyskania 25-letniego okresu ubezpieczenia; 4) członkowie rady nadzorczej i pobierające wynagrodzenie za pełnienie tej funkcji - od dnia powołania do dnia, kiedy przestaniesz pełnić tę funkcję. <p>Osoby nieobjęte obowiązkowymi ubezpieczeniami emerytalnym i rentowymi mogą przystąpić do nich dobrowolnie. Natomiast dobrowolne ubezpieczenie chorobowe można mieć tylko wtedy, gdy jest się objętym obowiązkowym ubezpieczeniem emerytalnym i rentowymi.</p>
Zakres przedmiotowy	Dane ubezpieczeniowe i finansowe na poziomie kont funduszy, kont płatników składek oraz kont ubezpieczonych dotyczące m.in. okresów podlegania ubezpieczeniu społecznemu i zdrowotnemu, wymiaru i okresu poboru składek na ubezpieczenia społeczne i zdrowotne oraz świadczeń pieniężnych wypłacanych z ubezpieczenia emerytalnego, rentowego, wypadkowego, chorobowego i macierzyńskiego.
Źródła danych	Dokumenty zgłoszeniowe ZUS przekazywane przez płatników składek oraz przez banki w ustalonych terminach: 5, 10, 15 każdego miesiąca. Dane z dokumentów zgłoszeniowych są przetwarzane i wprowadzane do systemu na bieżąco. Źródłem aktualizacji danych są również stanowiska pracy pracowników merytorycznych ZUS.

Ministerstwo Finansów – Centralny Rejestr Podmiotów – Krajowa Ewidencja Podatników

Cel systemu	Usprawnienie i spełnianie obowiązków wynikających z przepisów prawa oraz wykonywania przez administrację podatkową określonych prawem zadań realizowanych dla dobra publicznego.
Zakres podmiotowy	Podatnicy, płatnicy podatków, płatnicy składek na ubezpieczenia społeczne oraz na ubezpieczenia zdrowotne.
Zakres przedmiotowy	<p>Dla osób, dla których identyfikatorem podatkowym jest numer PESEL system obejmuje dane jednostkowe, m.in. w zakresie: nazwisko, imiona, data urodzenia, data zgonu, płeć, adres miejsca zamieszkania.</p> <p>Dla podmiotów, dla których identyfikatorem podatkowym jest NIP system obejmuje dane jednostkowe, m.in. w zakresie: nazwisko, imiona, nazwa, data urodzenia, płeć, adres miejsca zamieszkania, numer PESEL, numer REGON, adresy prowadzenia działalności gospodarczej i przedmiot działalności.</p>
Źródła danych	Źródłem danych dla ewidencji podatników jest rejestr PESEL oraz zgłoszenia

identyfikacyjne i aktualizacyjne zgodnie z ustawą z dnia 13 października 1995 roku o zasadach ewidencji i identyfikacji podatników i płatników.

Źródło: Opracowanie własne.

Dane występujące w zbiorach KRUS i ZUS pokrywają populację pracujących, bezrobotnych zarejestrowanych oraz zatrudnionych, a występujące w nich zmienne dostarczają informacji niezbędnych do prezentowania liczby pracujących w gospodarce narodowej w głównym miejscu pracy, stopy bezrobocia rejestrowanego oraz miar wynagrodzeń brutto w wymaganych agregacjach.

2.2.1. Ocena źródłowych zbiorów danych *Evaluation of source data sets*

Warunkiem wykorzystania zewnętrznych systemów informacyjnych administracji publicznej jako źródeł danych dla statystyki było spełnienie przez nie wymogów jakości przyjętych w statystyce. Dlatego na podstawie obowiązujących kryteriów oceny jakości rejestrów administracyjnych oraz oceny jakości badań statystycznych przeprowadzono ocenę jakości źródeł danych wykorzystanych w niniejszej pracy badawczej. Dla każdego ze źródeł danych analizie poddano następujące komponenty jakości: przydatność danych, dokładność danych, terminowość i punktualność, dostępność i przejrzystość, porównywalność i spójność danych.

Przydatność

Zgodnie z obowiązującymi przepisami dotyczącymi ubezpieczeń społecznych i zdrowotnych wykonywanie pracy w rozumieniu definicji pracującego w gospodarce narodowej rodzi obowiązek ubezpieczenia społecznego lub zdrowotnego z tytułu tej pracy w rejestrze KRUS lub rejestrze ZUS. Także rejestracja w Urzędzie Pracy pociąga za sobą konieczność zgłoszenia osoby do ubezpieczeń. Implikuje to wysoką przydatność tych rejestrów, ponieważ zawierają informacje o wszystkich osobach objętych badaniem. Ponadto wykorzystanie przez statystykę publiczną danych z systemów administracyjnych częściowo ograniczy obciążenie respondentów m.in. w zakresie: ograniczenia liczby zmiennych na formularzach statystycznych, ograniczenia częstotliwości pozyskiwania danych na formularzach statystycznych, zmniejszenia liczby obowiązków sprawozdawczych, zmniejszenia kosztów (prowadzenie odpowiednich ewidencji i związany z tym zakup i aktualizacja odpowiednich systemów informatycznych, a także koszty obsługi sprawozdawczości wynikające z wykorzystania pracowników własnych, jak i jednostek zewnętrznych).

Dostępność

1) dostępność metainformacji

Na potrzeby niniejszej pracy badawczej zarówno zakres przedmiotowy w strukturze zbiorów danych, jak i warunki wypełniania wartości zmiennych dla poszczególnych obserwacji przekazywane były bez szczegółowych informacji ze strony gestorów.

Pozyskanie metainformacji było możliwe z następujących, publicznie dostępnych źródeł:

- a) *ustawa o ubezpieczeniu społecznym rolników*,
- b) dostępne na stronie internetowej www.krus.gov.pl poradniki, m. in.: broszura *Kasa Rolniczego Ubezpieczenia Społecznego. Informacje podstawowe 2016*,
- c) *ustawa o systemie ubezpieczeń społecznych*,
- d) dostępny na stronie internetowej www.zus.pl słownik pojęć oraz poradniki, m. in.: *Ogólne Zasady Wypełniania i Korygowania Dokumentów Ubezpieceniowych*.

Kontakt z pracownikami KRUS i ZUS odpowiadającymi merytorycznie i technicznie za wygenerowanie i przekazanie baz Głównemu Urzędowi Statystycznemu pozwalał na wyjaśnienie podstawowych kwestii niezbędnych do prawidłowego przeprowadzenia badania.

2) dostępność baz i danych

W trakcie badania dostęp do bazy, której gestorem był KRUS był zapewniony w systemie rocznym ze stanem na dzień 31 grudnia. Biorąc pod uwagę niewielką dynamikę zmian zachodzących wśród pracujących w rolnictwie na potrzeby badania taka częstotliwość była wystarczająca.

Dostęp do baz ZUS zapewniony był w systemie kwartalnym (cztery razy w roku za ostatni miesiąc każdego kwartału (marzec, czerwiec, wrzesień, grudzień)). Była to częstotliwość wystarczająca dla naliczania liczby pracujących w gospodarce narodowej ogółem oraz liczby bezrobotnych zarejestrowanych (jako składowej do wyznaczenia stopy bezrobocia rejestrowanego) w agregacjach określonych w celu niniejszej pracy badawczej. Do naliczenia liczby zatrudnionych (w etatach) i naliczenia miar wynagrodzeń brutto dostęp do baz był zapewniony częściowo. Z uwagi na brak zbiorów dotyczących wszystkich miesięcy z danego kwartału, oszacowanie danych o zatrudnionych i wynagrodzeniach brutto dokonano za miesiące kończące dany kwartał.

Kompletność danych

Dane pochodzące ze zbiorów administracyjnych KRUS i ZUS w pełni pokrywają **zakres podmiotowy** badania:

- 1) zgodnie z *ustawą o ubezpieczeniu społecznym rolników* udostępniony zbiór KRUS w pełni pokrywa zakres podmiotowy (zakres obserwacji) dla potrzeb naliczania pracujących w gospodarce narodowej (ogółem oraz w agregacjach określonych w celu niniejszej pracy badawczej) w gospodarstwach indywidualnych w rolnictwie
- 2) zgodnie z *ustawą o systemie ubezpieczeń społecznych* udostępnione zbiory ZUS pokrywają w pełni zakres podmiotowy (zakres obserwacji) dla potrzeb naliczania liczby pracujących w gospodarce narodowej (poza pracującymi w gospodarstwach indywidualnych w rolnictwie) i liczby bezrobotnych zarejestrowanych (ogółem oraz w agregacjach określonych w celu niniejszej pracy badawczej), a także liczby zatrudnionych (w etatach) i ich wynagrodzeń brutto.

Zakres podmiotowy obu źródeł danych znacznie wykracza poza zakres podmiotowy niniejszego badania. Jednakże informacje zawarte w obu rejestrach pozwalają na eliminację obserwacji spoza zakresu podmiotowego badania.

Dane pochodzące ze zbiorów administracyjnych KRUS i ZUS częściowo pokrywają **zakres przedmiotowy** badania:

- 1) Zakres informacji przekazywanych z Kasy Rolniczego Ubezpieczenia Społecznego wynika z Programu Badań Statystycznych Statystyki Publicznej. Jest to obszar częściowo pokrywający zakres przedmiotowy (zakres zmiennych) stosowany w dotychczasowych badaniach statystyki publicznej (niezbędne jest pozyskanie informacji dotyczących wielkości użytków rolnych gospodarstwa w hektarach), natomiast w pełni pokrywa on zakres przedmiotowy niniejszego badania.
- 2) Zakres informacji przekazywanych z Zakładu Ubezpieczeń Społecznych wynika z Programu Badań Statystycznych Statystyki Publicznej. Przekazywane dane częściowo pokrywają zakres przedmiotowy (zakres zmiennych) zarówno dotychczasowych badań statystyki publicznej, jak i niniejszego badania. Do pełnego pokrycia zakresu przedmiotowego niezbędne jest pozyskanie informacji dotyczącej sekcji PKD 2007 płatnika.

Dokładność danych

Dane pozyskiwane z KRUS i ZUS cechują się wysoką dokładnością. Skala występowania braków wartości zmiennych w pojedynczych obserwacjach jest niewielka:

- 1) W zbiorze KRUS występują braki wartości w pojedynczych obserwacjach dla poszczególnych zmiennych, takich jak: numer PESEL, cechy adresowe. Istnieje możliwość uzupełnienia braków tych zmiennych na podstawie pozostałych zmiennych lub innych baz danych. Nie występują braki w zmiennych istotnych do naliczeń.
- 2) W zbiorze ZUS występują braki wartości w pojedynczych obserwacjach dla poszczególnych zmiennych, takich jak: REGON płatnika, numer PESEL oraz płeć ubezpieczonego. Istnieje możliwość uzupełnienia braków na podstawie pozostałych zmiennych, zależności między zmiennymi oraz innych baz danych.

Terminowość i punktualność

Zbiory KRUS i ZUS zostały przekazane przez gestorów na potrzeby badań ujętych w PBSSP i wykorzystane wtórnie w niniejszej pracy badawczej. Były one dostępne w momencie rozpoczęcia pracy badawczej.

2.2.2. Opis źródłowych zbiorów danych

Description of source data sets

Do realizacji pracy badawczej konieczne było pozyskanie danych z rejestrów KRUS i ZUS, a pomocniczo również z rejestru MF. W tym celu wtórnie wykorzystano dane pozyskiwane od gestorów na potrzeby badań ujętych w Programie Badań Statystycznych Statystyki Publicznej.

W ramach danych Kasy Rolniczego Ubezpieczenia Społecznego w zasobach Głównego Urzędu Statystycznego dostępne były zbiory:

- 1) *krusubezp_krl3_xxxx_t2* zawierający dane o ubezpieczonych,
- 2) *krusubezp_kr11_xxxx_t2* zawierający dane o ubezpieczonych i płatnikach.

W ramach danych Zakładu Ubezpieczeń Społecznych w zasobach Głównego Urzędu Statystycznego dostępne były zbiory:

- 1) *zusplatnik_fizpu01_xxxx_t2* zawierający dane o płatnikach będących osobami fizycznymi prowadzącymi działalność gospodarczą oraz osobach ubezpieczonych przez tych płatników,
- 2) *zusplatnik_prapu01_xxxx_t2* zawierający dane o płatnikach będących osobami prawnymi lub jednostkami organizacyjnymi nieposiadającymi osobowości prawnej oraz osobach ubezpieczonych przez tych płatników.

Struktury wymienionych wyżej zbiorów zostały przedstawione w załączniku nr 34.

Ponadto pomocniczo skorzystano z danych Ministerstwa Finansów, dostępnych w zasobach Głównego Urzędu Statystycznego w postaci zbiorów:

- 1) *mfosoba_kepoa02_xxxx_t2* zawierający dane o podatnikach będących osobami fizycznymi,
- 2) *mffirma_kepopjo_xxxx_t2* zawierający dane o podatnikach będących osobami prawnymi i jednostkami organizacyjnymi niemającymi osobowości prawnej,
- 3) *mffirma_keposfi_xxxx_t2* zawierający dane o podatnikach będących osobami fizycznymi prowadzącymi działalność gospodarczą.

Tablica 3. Przybliżona liczba rekordów w zbiorach KRUS, ZUS i MF

Gestor danych	Nazwa zbioru	Przybliżona liczba rekordów
KRUS	<i>krusubezp_krl3_xxxx_t2</i>	1,3 mln
	<i>krusubezp_kr11_xxxx_t2</i>	1,8 mln
ZUS	<i>zusplatnik_fzpu01_xxxx_t2</i>	4,7 mln
	<i>zusplatnik_prapu01_xxxx_t2</i>	11,5 mln
MF	<i>mfosoba_kepoa02_xxxx_t2</i>	29,9 mln
	<i>mffirma_kepopjo_xxxx_t2</i>	1,1 mln
	<i>mffirma_keposfi_xxxx_t2</i>	2,6 mln

Źródło: Opracowanie własne.

Zawarte w zbiorach informacje dotyczą:

- 1) w przypadku KRUS - stanu na ostatni dzień roku oraz stanu w roku; w zbiorach znalazły się również informacje wprowadzone w pierwszych 10 dniach następnego roku (data zgonu ubezpieczonego, data rozpoczęcia działalności rolniczej); wynika to z cyklu przetwarzania danych w systemie informatycznym KRUS,
- 2) w przypadku ZUS - stanu z ostatniego miesiąca kwartału; w zbiorach znalazły się również informacje wprowadzone po miesiącu kończącym kwartał, ale dotyczące tego okresu; wynika to z możliwości wprowadzania korekt przez płatników oraz z cyklu przetwarzania danych w systemie informatycznym ZUS.

Stany zbiorów opisane powyżej nazywane są **okresami referencyjnymi** stosowanymi w niniejszej pracy badawczej.

Przygotowanie bazy danych było podstawą oszacowania wskaźników i naliczenia danych, których uzyskanie było celem niniejszej pracy badawczej. Wymagało to zintegrowania zharmonizowanych zbiorów KRUS i ZUS.

Harmonizacja jest procesem niezbędnym do przeprowadzenia prawidłowej integracji różnych źródeł danych. Wyróżnić można następujące etapy harmonizacji (van der Laan 2000; Roszka 2013):

- 1) harmonizacja definicji jednostek,
- 2) harmonizacja okresów referencyjnych,
- 3) badanie kompletności populacji,
- 4) harmonizacja zmiennych,
- 5) harmonizacja wariantów cech,
- 6) korekta błędów pomiaru,
- 7) korekta związana z brakami danych (imputacja),
- 8) tworzenie zmiennych pochodnych.

Pierwszy i trzeci z powyższych warunków został zapewniony poprzez właściwy dobór źródeł danych, spójnych pod względem definicyjnym wynikającym z przepisów prawa oraz w znacznej mierze względem siebie komplementarnych pod kątem uzyskania pełnego pokrycia badanej populacji.

Okresy referencyjne obu integrowanych zbiorów nie były jednakowe ale, na potrzeby realizacji pracy badawczej, wystarczające ze względu na:

- 1) wyżej wspomnianą komplementarność zbiorów,
- 2) możliwość przeprowadzenia działań znacząco zwiększających spójność czasową zawartych w nich zmiennych,
- 3) niewielką dynamikę zmian zakresu podmiotowego obserwowanego w KRUS.

Spełnienie pozostałych warunków zapewniły rozwiązania opisane w kolejnych punktach niniejszego rozdziału. Przy czym, w celu zwiększenia poziomu harmonizacji wariantów cech dla poszczególnych grup zmiennych przyjęto następujące dane referencyjne. Dla:

- 1) adresu zamieszkania – zbiory ZUS i KRUS,
- 2) adresu miejsca siedziby pracy głównej – zbiory ZUS i KRUS,
- 3) grup sekcji PKD 2007 i sektorów gospodarki narodowej – zbiór BJS,
- 4) płci – zbiory ZUS i KRUS,
- 5) wieku – zbiory ZUS i KRUS.

W przypadku wystąpienia w zbiorach referencyjnych braków w wartościach zmiennych jako rejestr pomocniczy przyjęto CRP KEP.

Źródłowe zbiory danych otrzymane przez GUS zostały przetworzone w Ośrodku Administracyjnych Źródeł Danych w Warszawie. Analiza poprawności danych dotyczy zbiorów już przekształconych.

Zgodnie z założeniami, pierwszym etapem przygotowania bazy danych do wyznaczenia wskaźników i danych wynikowych zgodnie z przedmiotem pracy badawczej było sprawdzenie poprawności danych w podstawowych zbiorach źródłowych, tj. zbiorach ZUS i KRUS, w tym zidentyfikowanie:

- 1) błędów logicznych,
- 2) wartości spoza dopuszczalnego zakresu,
- 3) wartości nietypowych.

Przeprowadzona eksploracja błędów i braków pozwoliła na uzyskanie informacji o zmiennych i wartościach wymagających korekty, a procedury opisane w podrozdziale 2.3.8. wyznaczyły te z nich, które wymagały zastosowania imputacji danych, dla których korekta nie była możliwa.

Wyniki analizy są zaprezentowane w załączniku nr 35.

2.3. Opis rozwiązań metodologicznych i metodycznych zastosowanych do generowania wskaźników i danych wynikowych

Description of methodological and methodical solutions used for the generation of indicators and result data

Oszacowanie poszczególnych zmiennych na podstawie danych administracyjnych wymagało zmian w stosunku do definicji dotychczas stosowanych przez statystykę publiczną. Ponadto specyfika zmiennych wymagała, aby prace nad dostosowaniem metodologii i naliczeniem zmiennych przebiegały w odpowiedniej kolejności. Najpierw zrealizowano, objęte modułem I niniejszej pracy badawczej, prace metodologiczne i metodyczne dotyczące zdefiniowania i wyodrębnienia zbiorowości pracujących (w dniu referencyjnym badania), wyboru pracy głównej dla każdego z pracujących i naliczenia liczby pracujących w pracy głównej w poszczególnych agregacjach. Następnie wykonano prace metodologiczne i metodyczne umożliwiające wybranie zbiorowości bezrobotnych zarejestrowanych (w dniu referencyjnym badania), co wraz z wynikami uzyskanymi w module I pozwoliło na realizację celu przewidzianego w module II, czyli naliczenie stopy bezrobocia rejestrowanego w poszczególnych agregacjach. Wynikiem prac przeprowadzonych w module III było oszacowanie miar wynagrodzeń i miar rozkładu wynagrodzeń w poszczególnych agregacjach. Prace metodologiczne i metodyczne wykonane w ramach tego modułu obejmowały: wyodrębnienie zbiorowości zatrudnionych (w okresie referencyjnym badania) i ustalenie dla każdej obserwacji z tej zbiorowości rzeczywistego wymiaru etatu w okresie referencyjnym, a następnie wydzielenie zbiorowości, dla której możliwe było oszacowanie wynagrodzeń i naliczenie dla tej zbiorowości miar wynagrodzeń i miar rozkładu wynagrodzeń.

Zmiana źródła danych wymusza zmianę podejścia do oszacowania wskaźników i danych wynikowych dotyczących pracujących, bezrobotnych zarejestrowanych, zatrudnionych oraz wynagrodzeń w gospodarce narodowej, tj. przejścia z obserwowania osób pracujących, bezrobotnych zarejestrowanych oraz zatrudnionych poprzez podmioty gospodarki narodowej do bezpośredniej obserwacji tych osób.

2.3.1. Opis rozwiązań metodologicznych przyjętych w celu wyznaczenia zbiorowości pracujących

Description of methodological solutions adopted for counting the population of employed persons

Zakres podmiotowy wskazany w definicji pracujących w gospodarce narodowej jest zasadniczo zgodny z zakresem podmiotowym możliwym do uzyskania na podstawie danych administracyjnych. Prawidłowe wyznaczenie liczby pracujących wymagało jednak dostosowania metodologii w zakresie młodocianych niewykonujących pracy w celu przygotowania zawodowego oraz właścicieli, współwłaścicieli i dzierżawców gospodarstw indywidualnych w rolnictwie niewytwarzających towarów i usług wyłącznie lub głównie na swoje potrzeby. Wynika to ze specyfiki źródeł administracyjnych, które nie dają możliwości prawidłowego wyodrębnienia tych grup. W rejestrze Zakładu Ubezpieczeń Społecznych młodociani, niezależnie od tego czy podejmują pracę w ramach przygotowania zawodowego czy nie, klasyfikowani są do jednej grupy, uniemożliwiając wyodrębnienie tych, którzy nie wykonują pracy w celu

przygotowania zawodowego. Z kolei w ramach danych gromadzonych przez Kasę Rolniczego Ubezpieczenia Społecznego ubezpieczeniom podlegają osoby, których gospodarstwo rolne obejmuje obszar użytków rolnych powyżej 1 ha przeliczeniowego, zamiast stosowanego w statystyce 1 ha użytków rolnych. Także progi rodzajów i rozmiarów działów specjalnych ustalone dla ubezpieczenia społecznego rolników różnią się od progów ustalonych na potrzeby badań gospodarstw rolnych.

W ramach systemów informacyjnych administracji publicznej funkcjonuje wiele definicji gospodarstwa indywidualnego w rolnictwie, dlatego w zależności od zastosowanego źródła konieczne będzie przyjęcie:

- 1) innego areалу kwalifikującego gospodarstwo nieposiadające działu specjalnego,
- 2) innych progów działów specjalnych produkcji rolnej.

W przypadku progów działów specjalnych, sytuację prezentuje poniższa tablica.

Tablica 4. Rodzaje i rozmiary działów specjalnych produkcji rolnej

Lp.	Rodzaje upraw i produkcji	Ustalony próg	
		dla badań gospodarstw rolnych	dla ubezpieczenia społecznego rolników
		GUS ²⁹	KRUS
1.	Uprawy w szklarniach ogrzewanych:		100 m ²
2.	Uprawy w szklarniach nieogrzewanych:	ogółem 1000 m ²³⁰	100 m ²
3.	Uprawy w tunelach foliowych ogrzewanych:		200 m ²
4.	Uprawy grzybów	25 m ²	100 m ²
5.	Plantacje drzew owocowych, krzewów owocowych, uprawy warzyw gruntowych, truskawek gruntowych, chmielu	0,5 ha	-
6.	Szkółki sadownicze i ozdobne	0,3 ha	-
7.	Drób rzeźny	100 szt.	X
	- kurczęta		1000 szt.
	- gęsi		500 szt.
	- kaczki		500 szt.
	- indyki		500 szt.
	- strusie		50 szt.
8.	Drób nieśny		X
	- kury nieśne		2000 szt.
	- kury mięsne		-
	- kury (produkcja jaj konsumpcyjnych)		1000 szt.

²⁹ Działy specjalne (progi) wg definicji Badania Struktury Gospodarstw Rolnych 2016 i PSR 2020.

³⁰ W prowadzonych przez GUS badaniach gospodarstw rolnych stosuje się określenie „uprawy pod ostonami”, gdzie „ostony” definiowane są jako ogrzewane lub nieogrzewane szklarnie i tunele foliowe (minimum 1,5 m w szczycie) oraz inspekty.

Lp.	Rodzaje upraw i produkcji	Ustalony próg	
		dla badań gospodarstw rolnych	dla ubezpieczenia społecznego rolników
		GUS ²⁹	KRUS
	- gęsi		200 szt.
	- kaczki		500 szt.
	- indyki		500 szt.
	- strusie		6 szt.
9.	Wylęgarnie drobiu (zdolność produkcyjna – liczba jaj)		X
	- kurczęta		3000 szt.
	- gęsi		3000 szt.
	- kaczki		3000 szt.
	- indyki		3000 szt.
	- strusie		50 szt.
10.	Króliki	50 szt. samic	50 szt. samic
11.	Pozostałe zwierzęta futerkowe		X
	- lisy i jenoty		50 szt. samic
	- norki	5 szt. Samic	50 szt. samic
	- tchórze		50 szt. samic
	- szynszyle		50 szt. samic
	- nutrie		50 szt. samic
12.	Pasieki	20 pni pszczelich	80 rodzin pszczelich
13.	Bydło ogółem:	10 szt. ³¹	-
	- krowy	5 szt.	-
	- cielęta	-	-
	- bydło rzeźne	-	-
14.	Świnie ogółem:	50 szt. ³²	-
	- lochy	10 szt.	100 szt. Świń ³³

³¹ Stosuje się określenie „bydło ogółem”, wliczając krowy.

³² Stosuje się określenie „świnie ogółem”, wliczając lochy.

³³ Nie stosuje się rozróżnienia na tuczniaki, warchlaki, prosięta.

Lp.	Rodzaje upraw i produkcji	Ustalony próg	
		dla badań gospodarstw rolnych	dla ubezpieczenia społecznych rolników
		GUS ²⁹	KRUS
	- tuczniaki	-	
	- prosięta i warchlaki	-	
15.	Owce		x
	- hodowla	20 szt. ogółem	-
	- tucz		-
16.	Konie rzeźne		x
	- rzeźne	5 szt. ogółem	-
	- hodowlane		-
17.	Kozy	20 szt. ogółem	-
18.	Zwierzęta „dzikie” (dla produkcji mięsa)	10 szt.	-

Źródło: Opracowanie własne.

W związku z powyższym na potrzeby pracy badawczej do pracujących zaliczono osoby:

- 1) ubezpieczone w ZUS z kodem tytułu ubezpieczenia świadczącym o wykonywaniu pracy, bez uwzględnienia pracowników młodocianych,
- 2) będące płatnikami KRUS lub osobami ubezpieczonymi przez tych płatników, których gospodarstwo rolne obejmuje obszar użytków rolnych powyżej 1 ha przeliczeniowego lub prowadzącymi dział specjalny spełniający progi ustalone dla ubezpieczenia społecznych rolników.

Zgodnie z definicją praca główna to praca, która zwykle zajmuje więcej czasu a jeżeli dwie prace (lub więcej) zajmują taką samą ilość czasu – to ta, która przynosi wyższy dochód.

Na potrzeby pracy badawczej dla osób pracujących wykonujących więcej niż jedną pracę, pracę główną ustalono w następujący sposób:

- 1) w przypadku pracujących w rolnictwie indywidualnym, dla których źródłem danych jest KRUS przyjęto, że czas poświęcony na pracę w gospodarstwie jest wprost proporcjonalny do wielkości gospodarstwa. W przypadku posiadania działu specjalnego uznano, że praca na rzecz gospodarstwa posiadającego dział specjalny zajmuje więcej czasu niż praca na rzecz gospodarstwa rolnego nie posiadającego działu specjalnego.
- 2) dla pracujących pozyskanych ze zbiorów ZUS przyjęto, że ich pracą główną była praca o najwyższym wymiarze czasu pracy, a w przypadku wystąpienia więcej niż jednej pracy o tym samym wymiarze czasu pracy, wyboru dokonano w oparciu o wysokość podstawy składki emerytalno-rentowej. Ponadto dla ubezpieczonych takich jak: funkcjonariusze, sędziowie przyjęto, że ich wymiar czasu pracy jest zawsze maksymalny.

2.3.2. Opis rozwiązań metodycznych przyjętych w celu wyznaczenia liczby pracujących w głównym miejscu pracy

Description of methodical solutions adopted for counting the number of employed persons working in the main workplace

W celu wyodrębnienia osób pracujących w pracy głównej (zgodnie z definicją przyjętą na potrzeby niniejszego badania) według stanu na wskazany dzień (ostatni dzień okresu referencyjnego badania) w zbiorze KRUS należało:

- 1) zidentyfikować osoby, które spełniały poniższe warunki:
 - a) podlegali ubezpieczeniu w KRUS na wskazany dzień lub byli płatnikami, którzy nie podlegali ubezpieczeniu na wskazany dzień, ale na ich koncie była zarejestrowana osoba, która podlegała ubezpieczeniu na wskazany dzień,
 - b) jeżeli nie byli płatnikami, to podlegali ubezpieczeniu emerytalno-rentowemu na wskazany dzień i jednocześnie podlegali ubezpieczeniu wypadkowemu, chorobowemu i macierzyńskiemu na wskazany dzień,
 - c) jeżeli byli płatnikami, to posiadali co najmniej jedno gospodarstwo lub co najmniej jeden dział specjalny; jeżeli nie byli płatnikami, to ich płatnicy posiadali co najmniej jedno gospodarstwo lub co najmniej jeden dział specjalny,
 - d) jeżeli byli płatnikami i miały wypełnioną zmienną informującą o dacie rozpoczęcia działalności rolniczej, to powinna ona być równa lub wcześniejsza niż wskazany dzień (osób, które rozpoczęły działalność rolniczą po zakończeniu okresu referencyjnego nie zaliczono do pracujących); jeśli nie byli płatnikami i ich płatnicy mieli wypełnioną zmienną informującą o dacie rozpoczęcia działalności rolniczej, to powinna ona być równa lub wcześniejsza niż wskazany dzień,
 - e) we wskazanym dniu miały ukończone 16 lat (osób, które ukończyły 16 lat po zakończeniu okresu referencyjnego nie zaliczono do pracujących),
 - f) jeżeli miały wypełnioną zmienną informującą o dacie zgonu, to data zgonu powinna być późniejsza niż wskazany dzień (osoby zmarłe ostatniego dnia okresu referencyjnego zaliczono do pracujących),
 - g) we wskazanym dniu nie pobierały renty strukturalnej,
- 2) dla osób wyznaczonych w punkcie 1. wykonujących tylko jedną pracę na rzecz gospodarstwa rolnego lub działu specjalnego, pracę tę uznać za pracę główną,
- 3) dla osób wyznaczonych w punkcie 1. wykonujących więcej niż jedną pracę na rzecz gospodarstw rolnych i/lub działów specjalnych dokonać wyboru pracy głównej na podstawie:
 - a) faktu świadczenia pracy na rzecz działu specjalnego - biorąc pod uwagę, że specyfika działu specjalnego wymaga codziennego wkładu pracy, w niniejszym opracowaniu przyjęto, że praca na rzecz gospodarstwa posiadającego dział specjalny w KRUS zajmuje więcej czasu niż praca na rzecz gospodarstwa rolnego nie posiadającego działu specjalnego,
 - b) powierzchni gospodarstwa rolnego w hektarach przeliczeniowych,
 - c) numeru konta płatnika
 - d) numeru rekordu w zbiorze.

Z kolei w zbiorze ZUS należało:

- 1) zidentyfikować ubezpieczonych z kodem tytułu ubezpieczenia świadczącym o fakcie wykonywania pracy (szczegółowy wykaz w załączniku nr 33),
- 2) dla wszystkich osób ubezpieczonych występujących z tylko jednym kodem tytułu ubezpieczenia świadczącym o wykonywaniu pracy uznać, że jest to ich praca główna,
- 3) dla wszystkich osób ubezpieczonych z więcej niż jednym kodem tytułu ubezpieczenia świadczącym o wykonywaniu pracy wybrać pracę główną na podstawie:

- a) wymiaru czasu pracy,
- b) wysokości podstawy wymiaru składki emerytalno-rentowej,
- c) hierarchii kodów tytułu ubezpieczenia,
- d) daty początku obowiązywania składek.

W ostatnim etapie dokonano wyboru pracy głównej między KRUS i ZUS w następujący sposób:

- 1) dla osób, które wystąpiły tylko wśród pracujących w KRUS lub tylko wśród pracujących w ZUS praca główna pozostała bez zmian,
- 2) dla osób, które wystąpiły zarówno wśród pracujących w KRUS jak i wśród pracujących w ZUS pracę główną ustalono na podstawie:
 - a) wymiaru czasu pracy wykazanego w ZUS,
 - b) posiadania statusu ubezpieczonego w KRUS,
 - c) posiadania statusu płatnika w KRUS.

2.3.3. Opis rozwiązań metodologicznych przyjętych w celu wyznaczenia zbiorowości bezrobotnych zarejestrowanych

Description of methodological solutions adopted for counting the population of registered unemployed persons

Aktualnie głównym źródłem informacji o osobach bezrobotnych są dane gromadzone przez poszczególne powiatowe urzędy pracy. Przy wyborze najlepszego źródła danych administracyjnych do pozyskania informacji o tych osobach kierowano się możliwością ich uzyskania z jednego rejestru. Powiatowe urzędy pracy mają ustawowy obowiązek zgłaszać zarejestrowane osoby bezrobotne do ubezpieczeń w ZUS³⁴. Dane występujące w zbiorze ZUS dość dobrze pokrywają się z populacją bezrobotnych dotychczas wykazywaną w sprawozdawczości MRPIPS, a występujące w nim zmienne dostarczają informacji niezbędnych do prezentowania stopy bezrobocia rejestrowanego w potrzebnych agregacjach. ZUS jest również głównym źródłem informacji o osobach pracujących poza indywidualnymi gospodarstwami rolnymi.

Ważnym atutem wyżej wymienionego zbioru jest możliwość prezentowania stopy bezrobocia rejestrowanego według miejsca zamieszkania na poziomie gminy (NTS 5). Dotychczas prezentowane dane o bezrobotnych oraz pracujących wykorzystywane do obliczenia stopy bezrobocia pochodzą ze źródeł różniących się metodą zbierania, co może powodować pewne nieścisłości, zwłaszcza w przypadku stopy bezrobocia ustalonej dla powiatów. Dla bezrobotnych miejscem lokalizacji jest ich miejsce zamieszkania lub pobytu, zaś dla pracujących lokalizacja miejsca pracy.

Podstawą wyznaczania osób zarejestrowanych jako bezrobotne ze zbioru ZUS było ustalenie czy osoby te podlegają ubezpieczeniu w ZUS.

Zgodnie z ustawą o systemie ubezpieczeń społecznych obowiązkowo ubezpieczeniom emerytalnemu, rentowym, chorobowemu i wypadkowemu podlegają osoby bezrobotne od dnia nabycia prawa do zasiłku, świadczenia integracyjnego lub stypendium do dnia utraty prawa do nich. Natomiast bezrobotni nie pobierający zasiłku lub stypendium zgodnie z ustawą z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych (t.j. Dz. U. z 2017 r. poz. 1938 z późn. zm.) podlegają obowiązkowo ubezpieczeniu zdrowotnemu.

Na podstawie ww. aktów prawnych ustalono, że do **bezrobotnych zarejestrowanych według ZUS zaliczane będą osoby spełniające poniższe warunki:**

- 1) ubezpieczone z kodem tytułu ubezpieczenia:
 - a) „0910” – osoba bezrobotna pobierająca zasiłek dla bezrobotnych oraz osoba pobierająca świadczenie integracyjne,

³⁴ Ustawa z dnia 13 października 1998 r. o systemie ubezpieczeń społecznych (t.j. Dz. U. z 2017 r. poz. 1778); ustawa z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych (t.j. Dz. U. z 2017 r. poz. 1938 z późn. zm.)

- b) „0911” – osoba bezrobotna niepobierająca zasiłku dla bezrobotnych lub stypendium,
- 2) nie są właścicielami lub posiadaczami (samoistnym lub zależnym) nieruchomości rolnej w rozumieniu przepisów ustawy z dnia 23 kwietnia 1964 r. – Kodeks cywilny (Dz. U. z 2017 r. poz. 459 i 933³⁵), o powierzchni użytków rolnych powyżej 1 ha przeliczeniowego lub nie podlegają ubezpieczeniom emerytalnemu i rentowym z tytułu stałej pracy jako współmałżonek lub domownik w gospodarstwie rolnym o powierzchni użytków rolnych przekraczającej 1 ha przeliczeniowy,

Ponadto do bezrobotnych zarejestrowanych zaliczone zostały osoby, które w danym miesiącu utraciły status osoby bezrobotnej i fakt ten do dnia referencyjnego badania nie został zgłoszony do ZUS przez powiatowy urząd pracy.

Zgodnie z ustawą o promocji zatrudnienia i instytucjach rynku pracy, jako osoba bezrobotna może być zarejestrowana tylko osoba, która ukończyła 18 lat i nie osiągnęła wieku emerytalnego. Zatem uznano, że status osoby bezrobotnej mogły mieć tylko osoby spełniające kryterium wieku osoby bezrobotnej.

Biorąc pod uwagę możliwość zgłoszenia do ZUS danej osoby przez kilku płatników, należało ustalić, czy osoba ta została zgłoszona do ubezpieczeń w ZUS wielokrotnie.

W sytuacji, gdy dana osoba została zgłoszona do ZUS tylko raz z jednym tytułem ubezpieczenia świadczącym o statusie osoby bezrobotnej, uznano tę osobę za bezrobotną, jeśli jej płatnikiem był powiatowy urząd pracy.

Dla przypadków, gdy osoba została zgłoszona do ZUS więcej niż raz, z tytułami ubezpieczenia świadczącymi o statusie osoby bezrobotnej, niezbędne było ustalenie, które z tych zgłoszeń najlepiej opisywało jej sytuację w dniu referencyjnym badania.

Natomiast, gdy dana osoba została zgłoszona do ZUS z tytułem ubezpieczenia świadczącym o statusie osoby bezrobotnej oraz innym tytułem ubezpieczenia, trzeba było dokonać oceny, czy należy tę osobę uznać za bezrobotną. W takim przypadku konieczne było przeanalizowanie kolejności zgłoszeń tej osoby do ubezpieczeń w ZUS.

2.3.4. Opis rozwiązań metodycznych przyjętych w celu wyznaczenia stopy bezrobocia rejestrowanego

Description of methodical solutions adopted for the calculation of the registered unemployment rate

W celu wyznaczenia osób bezrobotnych należało kolejno zidentyfikować osoby:

- 1) które zostały zgłoszone do ubezpieczenia z kodem tytułu ubezpieczenia równym „0910” („osoba bezrobotna pobierająca zasiłek dla bezrobotnych oraz osoba pobierająca świadczenie integracyjne”) lub „0911” („osoba bezrobotna niepobierająca zasiłku dla bezrobotnych lub stypendium”);
- 2) których płatnikiem był powiatowy urząd pracy;
- 3) które spełniały kryterium wieku dla osoby bezrobotnej;
- 4) których data początku obowiązywania składki była wcześniejsza lub równa dacie referencyjnej zbioru.

Następnie, dla tak przygotowanego zbioru należało:

- 1) dla osób, które spełniając powyższe kryteria, wystąpiły w zbiorze ZUS tylko raz przyjąć, że osoby te posiadały status osoby bezrobotnej na dzień referencyjny zbioru;
- 2) dla osób, które spełniając powyższe kryteria, wystąpiły w zbiorze ZUS więcej niż raz, wyłącznie z kodami tytułu ubezpieczenia osoby bezrobotnej, ustalić, które wystąpienie będzie opisywało daną osobę jako bezrobotną. Wyboru dokonano w oparciu o datę początku obowiązywania składki;
- 3) dla osób, które spełniając powyższe kryteria, wystąpiły w zbiorze ZUS więcej niż raz, zarówno z kodem tytułu ubezpieczenia osoby bezrobotnej, jak i z kodami tytułu ubezpieczenia wykluczającymi posiadanie statusu osoby bezrobotnej, ustalić, na podstawie daty początku obowiązywania składki ich status na dzień referencyjny badania;

³⁵ Do 02.03.2017 r. Dz. U. z 2016 r. poz. 380 z późn. zm.

- 4) zidentyfikować osoby, które spełniając powyższe kryteria, wystąpiły również w zbiorze KRUS i zgodnie z metodologią przyjętą w I etapie pracy badawczej zostały uznane za pracujące. Takie osoby należało wyeliminować ze zbioru osób bezrobotnych.

Po uzyskaniu zbioru osób bezrobotnych zarejestrowanych oraz dysponując uzyskanym we wcześniejszym etapie prac zbiorem osób pracujących można było naliczyć stopę bezrobocia rejestrowanego zgodnie ze wzorem:

$$\text{stopa bezrobocia rejestrowanego} = \frac{\text{liczba bezrobotnych zarejestrowanych}}{\text{liczba aktywnych zawodowo}},$$

gdzie $\text{liczba aktywnych zawodowo} = \text{liczba pracujących} + \text{liczba bezrobotnych zarejestrowanych}$.

2.3.5. Opis rozwiązań metodologicznych przyjętych w celu wyznaczenia zbiorowości zatrudnionych i oszacowania miar wynagrodzeń

Description of methodological solutions adopted for counting the population of employees and the estimation of their wages and salaries

W zbiorach ZUS nie występuje informacja o wynagrodzeniach, dlatego oszacowania miar wynagrodzeń dokonano w oparciu o wysokość podstawy wymiaru składki emerytalno-rentowej. Nie wszystkie składniki wynagrodzeń wchodzące w zakres wynagrodzeń w gospodarce narodowej miały swoje odzwierciedlenie w podstawie wymiaru składki emerytalno-rentowej.

Podstawowe różnice:

- 1) w przypadku wystąpienia zbiegu tytułów ubezpieczeń nie było możliwości wyodrębnienia z wynagrodzeń naliczonych w oparciu o postawę wymiaru składki emerytalno-rentowej kwot, które zostały wypłacone na podstawie umowy zlecenia, umowy o świadczenie usług lub umowy o dzieło. Jeśli pracownik ubezpieczony z kodem tytułu ubezpieczenia 011 zawarł ze swoim pracodawcą umowę zlecenia, umowę o świadczenie usług lub umowę o dzieło, to podstawa składki emerytalno-rentowej dla kodu tytułu ubezpieczenia 011 odnosi się zarówno do wynagrodzenia z tytułu umowy o pracę, jak i wynagrodzenia z tytułu umowy zlecenia, umowy o świadczenie usług lub umowy o dzieło;
- 2) w przypadku wystąpienia niezdolności pracownika do pracy z powodu absencji chorobowej, wynagrodzenie za okres absencji nie jest uwzględniane w podstawie wymiaru składki emerytalno-rentowej;
- 3) odprawy rentowe i emerytalne nie stanowią podstawy wymiaru składki emerytalno-rentowej;
- 4) nagrody jubileuszowe, które według zasad określających warunki ich przyznawania przysługują pracownikowi nie częściej niż co 5 lat, nie stanowią podstawy wymiaru składki emerytalno-rentowej.

Ponadto nie było możliwości wyodrębnienia z podstawy wymiaru składki emerytalno-rentowej kwot, które:

- 1) zostały wypłacone z tytułu udziału w zysku lub nadwyżce bilansowej;
- 2) zostały wypłacone z tytułu dodatkowego wynagrodzenia rocznego dla pracowników jednostek sfery budżetowej.

Szczegółowe porównanie składników wchodzących w zakres wynagrodzeń w gospodarce narodowej ze składnikami wchodzącymi w zakres podstawy wymiaru składki emerytalno-rentowej przedstawione zostało w załączniku nr 36.

Oszacowanie wynagrodzeń możliwe było dla osób posiadających status zatrudnionych w gospodarce narodowej. Wśród osób zatrudnionych w gospodarce narodowej znalazły się osoby objęte odrębnym systemem zaopatrzenia emerytalnego, dla których w zbiorze ZUS nie było informacji o wysokości podstawy wymiaru składek na ubezpieczenia emerytalne i rentowe.

Na potrzeby niniejszej pracy badawczej zbiorowość zatrudnionych w gospodarce narodowej ograniczono do osób ubezpieczonych w ZUS, dla których była informacja o wysokości podstawy wymiaru składek na ubezpieczenia emerytalne i rentowe. Osoby te w dalszej części pracy będą nazywane jako **zatrudnieni według AZD**.

W związku z powyższym ustalenie wynagrodzeń w niniejszej pracy badawczej zostało dokonane dla osób zatrudnionych według AZD.

Wynagrodzenia naliczone w oparciu o wysokość podstawy wymiaru składki emerytalno-rentowej w niniejszej pracy badawczej są nazywane **wynagrodzeniami według AZD**.

Wyznaczenie miar wynagrodzeń brutto w gospodarce narodowej oraz miar koncentracji rozkładu wynagrodzeń na poziomie NTS 4 w podziale na płeć i grupy sekcji PKD 2007 wymaga wykonania na zbiorach ZUS określonych czynności następujących po sobie.

2.3.6. Opis rozwiązań metodycznych przyjętych w celu oszacowania miar wynagrodzeń i miar rozkładu wynagrodzeń

Description of methodical solutions adopted for the estimation of the measures of wages and salaries and measures for distribution of wages and salaries

W celu wyznaczenia miar wynagrodzeń brutto w gospodarce narodowej oraz miar koncentracji rozkładu wynagrodzeń należało:

- 1) zidentyfikować osoby, które w okresie referencyjnym posiadały status zatrudnionego
Do zatrudnionych zaliczono osoby, które ukończyły 16 lat i więcej i ubezpieczone były z kodem tytułu ubezpieczenia świadczącym o fakcie bycia zatrudnionym.
- 2) ustalić nominalny wymiar etatu dla osób zatrudnionych
Nominalny wymiar etatu ustalono na podstawie informacji o wymiarze czasu pracy osoby ubezpieczonej w następujący sposób:
 - a) w przypadku osób ubezpieczonych jako poseł lub senator pobierający uposażenie albo funkcjonariusz Służby Celno-Skarbowej przyjęto, że były one zatrudnione w pełnym nominalnym wymiarze etatu;
 - b) w przypadku osób ubezpieczonych jako pracownik lub osoba wykonująca odpłatnie pracę na podstawie skierowania do pracy w czasie odbywania kary pozbawienia wolności lub tymczasowego aresztowania przyjęto, że nominalny wymiar etatu jest równy ilorazowi licznika i mianownika.
- 3) ustalić rzeczywisty wymiar etatu w okresie referencyjnym dla osób zatrudnionych
Rzeczywisty wymiar etatu został utworzony na podstawie nominalnego wymiaru etatu. Do jego wyznaczenia wzięto pod uwagę okresy przerw w opłacaniu składek, okresy pobierania świadczeń oraz inne czynniki mające wpływ na podstawę wymiaru składek emerytalno-rentowych.
- 4) oszacować rzeczywiste wynagrodzenie brutto wypłacone osobie zatrudnionej w okresie referencyjnym
Rzeczywiste wynagrodzenie brutto ustalono na podstawie informacji o wysokości podstawy wymiaru składek na ubezpieczenia emerytalne i rentowe.
- 5) oszacować wynagrodzenie brutto przypadające na pełen etat
Rzeczywiste wynagrodzenie brutto wypłacone osobie zatrudnionej w okresie referencyjnym odniesiono do rzeczywistego wymiaru etatu w okresie referencyjnym.
- 6) uzupełnić informacje o wynagrodzeniu brutto dla osób, które przekroczyły roczną podstawę wymiaru składek na ubezpieczenia emerytalne i rentowe.
Informacja o wynagrodzeniu brutto jest możliwa do uzupełnienia jedynie dla zbiorów, dla których istniała możliwość porównania ze zbiorem za wcześniejszy okres referencyjny.

Oszacowane w powyższym postępowaniu wynagrodzenie brutto przypadające na pełen etat dały możliwość wyliczenia miar wynagrodzeń (klasycznych i pozycyjnych) oraz miar koncentracji rozkładu wynagrodzeń.

Wyliczona w toku prac średnia arytmetyczna wynagrodzeń brutto w danym miesiącu referencyjnym stanowi **przeciętne wynagrodzenie brutto** w gospodarce narodowej. Wyliczenia średniej arytmetycznej dokonano na podstawie wynagrodzeń przeliczonych na pełen etat.

Z uwagi, że wynagrodzenie jest cechą ciągłą, wartość modalną (dominantę) naliczono w szeregu rozdzielczym przedziałowym. W szeregach rozdzielczych przedziałowych można wyznaczyć przedział, w którym znajduje się dominanta – przedział o największej liczebności. W celu wyznaczenia konkretnej wartości dominanty wykorzystano następujący wzór (Sobczyk, 2018):

$$D = x_0 + \frac{n_0 - n_{-1}}{(n_0 - n_{-1}) + (n_0 - n_{+1})} \cdot h_0,$$

gdzie:

x_0 – dolna wartość przedziału dominanty,

n_0 – liczebność przedziału dominanty,

n_{-1} – liczebność przedziału poprzedzającego przedział dominanty,

n_{+1} – liczebność przedziału następującego po przedziale dominanty,

h_0 – rozpiętość przedziału dominanty.

Na potrzeby niniejszej pracy badawczej wartość modalna została wyznaczona na podstawie wynagrodzeń przeliczonych na pełen etat. Przy naliczaniu najczęstszego wynagrodzenia brutto przyjęto założenie, że przedziały są lewostronnie domknięte, a ich rozpiętość ustalono na poziomie 100 zł, tj. [0, 100), [100, 200), itd. Natomiast za miarę liczebności przedziałów przyjęto sumę etatów.

Aby zachować czytelność i przejrzystość raportu, przedstawiono jedynie wartości wybranych decyli (I, V i IX) oraz kwartyli (I i III). Miary te zostały naliczone na podstawie wynagrodzeń w przeliczeniu na pełen etat.

Minimalne i maksymalne wynagrodzenie brutto naliczono na podstawie wynagrodzeń w przeliczeniu na pełen etat.

Miarę spłaszczenia (kurtozę) rozkładu wynagrodzeń naliczono ze wzoru:

$$K = \frac{\left((n + 1) * \left(\frac{m_4}{m_2^2} - 3 \right) + 6 \right) * (n - 1)}{((n - 2) * (n - 3))},$$

gdzie

n – liczebność grupy, dla której liczymy kurtozę,

m_4 – czwarty moment centralny,

m_2 – drugi moment centralny.

Wartość współczynnika Giniego rozkładu wynagrodzeń, po uporządkowaniu wynagrodzeń w kolejności rosnącej, naliczono ze wzoru:

$$G(x) = \frac{\sum_{i=1}^n (2i - n - 1) * x_i}{n * \sum_{i=1}^n x_i},$$

gdzie:

n – liczebność grupy, dla której liczymy współczynnik Giniego,

x_i – wartość i -tej obserwacji (wynagrodzenie i -tej osoby zatrudnionej).

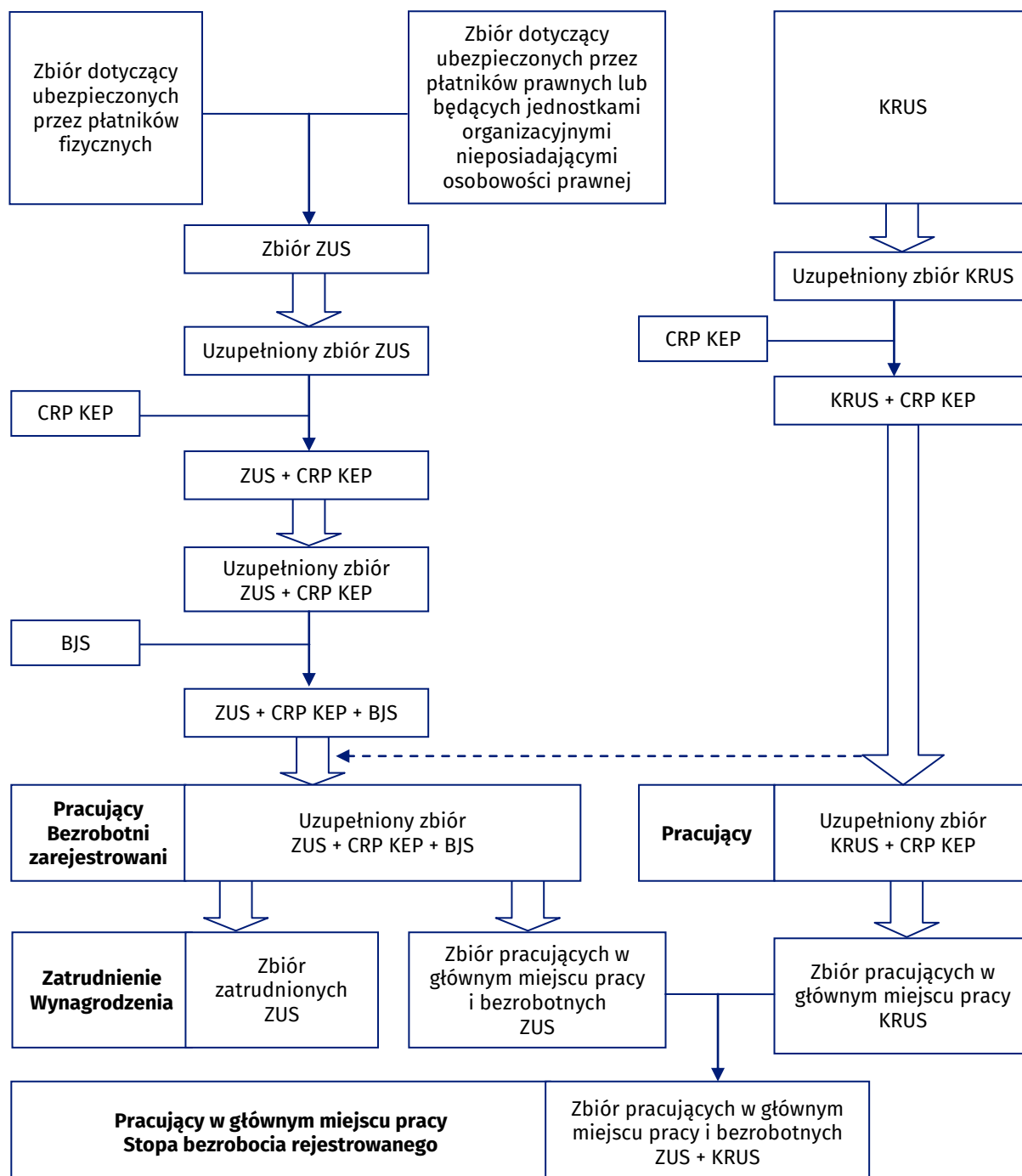
Krzywą Lorenza rozkładu wynagrodzeń w postaci graficznej zaprezentowano tylko w ogólnych agregacjach. Natomiast w pozostałych podziałach w tablicach przedstawiono wartości pola powierzchni pod krzywą.

2.3.7. Procedury przygotowujące zbiory KRUS i ZUS

Procedures for the preparation of data sets of the Polish Social Insurance Institution (ZUS) and the Agricultural Social Insurance Fund (KRUS)

Poniższy schemat przedstawia w formie graficznej procedury przygotowujące zbiory AZD do naliczania wskaźników i danych wynikowych.

Schemat 1. Procedury przygotowujące zbiory AZD



Źródło: Opracowanie własne.

Szczegółowy opis procedur przygotowujących zbiory KRUS i ZUS do wyznaczenia wskaźników i danych wynikowych przedstawiono w załączniku nr 37.

2.3.8. Opis rozwiązań metodologicznych i metodycznych przyjętych w celu uzupełnienia braków danych

Description of methodological and methodical solutions adopted for filling data gaps

W zbiorach danych pochodzących ze źródeł administracyjnych ZUS oraz KRUS występowały braki wartości dla niektórych zmiennych. Niezbędnym działaniem podejmowanym podczas całego procesu badawczego było zatem uzupełnienie brakujących wartości zmiennych istotnych z punktu widzenia przeprowadzanej analizy. W tym celu w pierwszej kolejności zastosowano metody imputacji prostej, a w przypadkach wymagających bardziej zaawansowanego działania skorzystano z technik imputacji bazującej na rozkładzie prawdopodobieństwa oraz imputacji wielokrotnej³⁶.

Opis stosowanych metod imputacji

1) Metoda dedukcyjna

Spośród dostępnych metod imputacji prostej³⁷ zdecydowano się na zastosowanie metody dedukcyjnej. Próba uzupełniania braków danych tym sposobem została podjęta w odniesieniu do wszystkich zmiennych wymagających imputacji. Metoda ta opiera się na możliwości pozyskania informacji zawartych w pozostałych zmiennych dla danego rekordu, pozostałych rekordach dla danej zmiennej lub innych, dostępnych bazach danych.

2) Metoda bazująca na rozkładzie prawdopodobieństwa

W przypadku, gdzie metoda dedukcyjna okazała się niewystarczająca, kolejnym etapem uzupełniania niekompletnych rekordów było wykorzystanie metody opartej na analizie rozkładu prawdopodobieństwa danej cechy. W tym celu użyto funkcji *sample()* dostępnej w środowisku programistycznym R. Przy jej pomocy generowana była próbka odwzorowująca odpowiedni rozkład prawdopodobieństwa w ustalonym rozmiarze. Składnia funkcji wygląda w następujący sposób:

$$\text{sample}(x, \text{size}, \text{replace} = \text{FALSE}, \text{prob} = \text{NULL}),$$

gdzie:

x – wektor występujących wartości dla imputowanej zmiennej

size – rozmiar wektora z wygenerowanymi wartościami zachowującymi rozkład prawdopodobieństwa

replace – określa, czy wybór próby ma być wykonywany z zastępowaniem

prob – wektor prawdopodobieństw.

3) Metoda imputacji wielokrotnej MICE

Ostatecznym rozwiązaniem dla uzupełnienia braków danych było wykorzystanie metody imputacji wielokrotnej: MICE – *Multivariate Imputation by Chained Equations*³⁸. W tym celu wykorzystano funkcję *mice()* dostępną w pakiecie programu R o takiej samej nazwie. Metoda MICE opiera się na stosowaniu bayesowskich technik estymacji. Poprzez tworzenie wielu zestawów imputacyjnych w odróżnieniu od pojedynczego uzupełniania pozwala na zredukowanie losowości wartości imputacyjnej. Szczegółowy opis składni funkcji *mice()* dostępny jest na stronie internetowej³⁹.

Funkcja *mice()* generuje wiele imputacji dla niekompletnych, wielowymiarowych danych poprzez zastosowanie próbkowania Gibbasa. Brakujące dane mogą występować w dowolnym miejscu w bazie danych. Algorytm imputuje kolumny zawierające braki (kolumny docelowe), generując dopuszczalne sztuczne wartości uzyskane z innych kolumn w zbiorze. Każda niekompletna kolumna musi wystąpić jako kolumna docelowa oraz ma swój własny szczególny zbiór predyktorów. Domyślny zestaw predyktorów dla danej kolumny docelowej zawiera wszystkie inne kolumny ze zbioru danych. W przypadku predyktorów, które są niekompletne, ostatnie generowane imputacje są używane do ich uzupełnienia przed przypisaniem wartości w kolumnie docelowej. Osobny model imputacji może być określony dla każdej kolumny. Pod używanym tutaj pojęciem „kolumny” rozumie się zmienne występujące w zbiorze.

³⁶ Korczyński A. (2014). Review Of Methods For Data Sets With Missing Values And Practical Applications. “Silesian Statistical Review” 12(18).

³⁷ Ostasiewicz W. (2004). Ocena i analiza jakości życia. Wydawnictwo Akademii Ekonomicznej, Wrocław.

³⁸ Little R.J.A., Rubin D.B. (2002), Statistical Analysis with Missing Data, 2nd Edition, A John Wiley & Sons, Inc., Hoboken, New Jersey.

³⁹ Van Buuren, S., Groothuis-Oudshoorn, K. (2011). MICE: Multivariate Imputation by Chained Equations in R. “Journal of Statistical Software” 45(3), 1-67.

Założmy, że pełna zmienna Y jest częściowo obserwowaną losową próbką z wielowymiarowego rozkładu $P(Y|\theta)$. Zauważmy, że ten rozkład zmiennej Y całkowicie zależy od wektora nieznanymi parametrów θ , należy więc znaleźć jego rozkład. Algorytm MICE uzyskuje rozkład *a posteriori* wektora θ poprzez próbkowanie iteracyjne z rozkładów warunkowych postaci:

$$P(Y_1|Y_{-1}, \theta_1)$$

⋮

$$P(Y_p|Y_{-p}, \theta_p)$$

Parametry $\theta_1, \dots, \theta_p$ są określone dla odpowiednich gęstości warunkowych. Począwszy od prostej próbki z obserwowanych rozkładów brzegowych, t -ta iteracja (dla $t \in N_+$) równań łańcuchowych jest próbkowaniem Gibbsa, gdzie:

$$\theta_1^{*(t)} \sim P(\theta_1|Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)})$$

$$Y_1^{*(t)} \sim P(Y_1|Y_1^{obs}, Y_2^{(t-1)}, \dots, Y_p^{(t-1)}, \theta_1^{*(t)})$$

⋮

$$\theta_p^{*(t)} \sim P(\theta_p|Y_p^{obs}, Y_1^{(t)}, \dots, Y_{p-1}^{(t)})$$

$$Y_p^{*(t)} \sim P(Y_p|Y_p^{obs}, Y_1^{(t)}, \dots, Y_{p-1}^{(t)}, \theta_p^{*(t)}),$$

natomiast $Y_j^{(t)} = (Y_j^{obs}, Y_j^{*(t)})$ jest j -tą imputowaną zmienną w iteracji t .

Opis procesu imputacji zmiennych, dla których występowały braki danych

Dedukcyjną metodę imputacji w odniesieniu do populacji pracujących i zatrudnionych stosowano dla następujących zmiennych występujących w ZUS:

- 1) płeć i wiek ubezpieczonego,
- 2) PESEL ubezpieczonego,
- 3) sekcja PKD 2007 przeważającej działalności płatnika,
- 4) kod TERYT gminy adresu zamieszkania ubezpieczonego,
- 5) kod TERYT gminy adresu siedziby płatnika

oraz dla zmiennych występujących w KRUS:

- 1) płeć i data urodzenia ubezpieczonego,
- 2) PESEL ubezpieczonego,
- 3) kod TERYT gminy adresu zamieszkania ubezpieczonego,
- 4) kod TERYT gminy adresu działalności płatnika

dla nich bowiem mogła istnieć możliwość pozyskania informacji na podstawie innych zmiennych występujących w bazie bądź po połączeniu z odrębnymi zbiorami danych.

Jeżeli PESEL ubezpieczonego nie był brakiem danych, to zmienną płeć uzupełniono na podstawie informacji o płci z 10. (przedostatniej) cyfry numeru PESEL ubezpieczonego. Podobny sposób postępowania przyjęto dla braków w zmiennej dotyczącej wieku ubezpieczonego. Braki w zmiennej dotyczącej sekcji PKD 2007 płatnika osoby ubezpieczonej uzupełniono przy wykorzystaniu zmiennej zawierającej informację o sekcji PKD 2007 pochodzącej ze zbioru CRP KEP (zbiory zostały połączone za pomocą NIP lub za pomocą numeru REGON płatnika) oraz w ramach zbioru ZUS za pomocą kodu tytułu ubezpieczenia (tylko dla wybranych kodów tytułu ubezpieczenia). Adres zamieszkania ubezpieczonego uzupełniono adresem zameldowania ubezpieczonego (zgodnie z zasadami przyjętymi w ZUS adres zamieszkania jest wypełniony, jedynie gdy jest on różny od adresu zameldowania). Braki w kodzie gminy adresu siedziby płatnika uzupełniono za pomocą zmiennych z BJS oraz za pomocą zmiennych z CRP KEP.

W odniesieniu do zmiennych płeć i data urodzenia ubezpieczonych w KRUS, mimo niewystępowania braków wartości, nastąpiła ich weryfikacja numerem PESEL. Z wykorzystaniem metody dedukcyjnej dokonano również uzupełnienia numeru identyfikacyjnego PESEL. Odbyło się to przy pomocy identyfikatora utworzonego na bazie nazwiska, imienia, daty urodzenia i adresu zamieszkania ubezpieczonego wewnątrz zbioru KRUS oraz z wykorzystaniem zbioru CRP KEP. Uzupełnienia adresu zamieszkania ubezpieczonego dokonano również w oparciu o dedukcję wewnątrz zbioru KRUS i przy pomocy zbioru CRP KEP, natomiast imputację adresu działalności płatnika wykonano poprzez metodę dedukcyjną ograniczając się jedynie do bazy KRUS.

W przypadkach, w których zastosowanie metody dedukcyjnej nie przyniosło wymaganych rezultatów, zarówno dla zbiorów ZUS jak i KRUS konieczne było wykorzystanie dwóch pozostałych metod opisanych w poprzednim podrozdziale. W odniesieniu do niektórych zmiennych skutecznym działaniem okazało się dopiero połączenie wybranych metod w celu osiągnięcia jak najlepszego efektu estymacyjnego⁴⁰.

Imputacja brakujących kodów TERYT gminy zamieszkania została przeprowadzona na podstawie zmiennych zawierających informację o numerach oddziału terenowego KRUS i placówki terenowej KRUS. Na podstawie tych informacji można za pomocą wektora prawdopodobieństwa wystąpienia gminy w danej placówce terenowej ustalić brakujące kody TERYT gminy zamieszkania.

Ze względu na niewielki udział braków gminy adresu zamieszkania wśród pracujących płatników (0,00035%) podjęto decyzję o zastosowaniu uproszczonej metody imputacji braków. W zbiorze wszystkich pracujących w KRUS posiadających poprawny adres zamieszkania, dla każdej placówki terenowej wyznaczono gminę, dla której w danej placówce występuje najwięcej rekordów i zgodnie z tym wykazem uzupełniono brakujące kody TERYT gminy zamieszkania dla pracujących płatników w zbiorze KRUS. Jeśli dla danej placówki KRUS wystąpiła więcej niż jedna gmina z maksymalną liczbą rekordów występujących w danej placówce, to do uzupełnienia braków wykorzystano gminę, która w klasyfikacji TERYT występuje najwcześniej (przy uszeregowaniu województw od 02 do 16).

Imputacja dotycząca płci oraz wieku ubezpieczonych występujących w bazie ZUS (w przypadku, gdy inne metody okazały się niewystarczające) została przeprowadzona z wykorzystaniem całego zbioru ubezpieczonych. W celu uzupełnienia braków danych w zmiennych zawierających informacje o płci i wieku ubezpieczonego konieczne było przeprowadzenie procesu z wykorzystaniem metody imputacji wielokrotnej MICE. Do imputacji zmiennych opisujących płeć i wiek ubezpieczonego wybrano ten sam zestaw predyktorów, do którego należały:

- 1) LP – unikalny numer rekordu w zbiorze (zmienna nie brała udziału w imputacji, była potrzebna do identyfikacji rekordu),
- 2) płeć ubezpieczonego,
- 3) wiek ubezpieczonego,
- 4) kod terytorialny powiatu zamieszkania ubezpieczonego,
- 5) kod terytorialny powiatu siedziby płatnika,
- 6) sekcja PKD,
- 7) podstawa składki emerytalno-rentowej,
- 8) wymiar etatu,
- 9) kod tytułu ubezpieczenia (4 znaki),
- 10) kod tytułu ubezpieczenia (5. znak),
- 11) kod tytułu ubezpieczenia (6. znak),

tj. zmienne, które mogły mieć największy wpływ na zmienne imputowane. W zbiorach za marzec i czerwiec 2016 r. jako predyktora zamiast zmiennej opisującej kod TERYT gminy zamieszkania ubezpieczonego wykorzystano kod TERYT powiatu zamieszkania.

⁴⁰Wesołowski J., Tarczyński J. (2016). Podstawy matematyczne technik imputacyjnych. "Wiadomości Statystyczne" nr 9, 7-54.

Kolejną imputację przeprowadzono w celu uzupełnienia braków danych w kodzie TERYT gminy miejsca zamieszkania ubezpieczonego. Dla zbiorów za marzec i czerwiec 2016 imputowano kod TERYT powiatu miejsca zamieszkania ubezpieczonego. Podczas tych procesów krokiem pośrednim było wykorzystanie metody imputacji wielokrotnej MICE. Przy jej pomocy przypisano brakom w kodzie województw dwucyfrowy kod terytorialny TERYT. Jako zestaw zmiennych predykcyjnych wybrano:

- 1) LP – unikalny numer rekordu w zbiorze (zmienna nie brała udziału w imputacji, była potrzebna do identyfikacji rekordu),
- 2) województwo zamieszkania ubezpieczonego,
- 3) województwo siedziby płatnika,
- 4) identyfikator płatnika,
- 5) sekcja PKD,
- 6) kod kraju.

Następnie imputację na poziomie gminy (powiatu) przeprowadzono poprzez zastosowanie rozkładu prawdopodobieństwa. Wektor prawdopodobieństw określono postępując się udziałem liczby ubezpieczonych w poszczególnych gminach w stosunku do liczby ubezpieczonych w danym województwie (po wcześniejszym wykluczeniu osób zamieszkałych za granicą).

W celu imputacji braków w kodzie TERYT gminy adresu siedziby płatnika w pierwszej kolejności przeprowadzono imputację dedukcyjną kodu TERYT województwa siedziby płatnika w oparciu o numer oddziału ZUS. Następnie dla wszystkich województw utworzono wektory prawdopodobieństw postępując się udziałem liczby płatników w poszczególnych gminach w stosunku do liczby płatników w danym województwie.

W ostatnim kroku dokonano imputacji zmiennej informującej o sekcji PKD 2007 przeważającej działalności płatnika. W tym celu skorzystano z metody rozkładu prawdopodobieństwa wystąpienia płatników w sekcjach PKD 2007 w ramach województw.

Zastosowanie powyżej opisanych metod imputacji pozwoliło na uzupełnienie wszystkich brakujących wartości w kluczowych zmiennych ze zbiorów ZUS i KRUS.

Zestawienie liczby braków danych w źródłach administracyjnych przeznaczonych do imputacji znajdują się w załączniku nr 38.

2.4. Lista oszacowanych wskaźników i danych wynikowych

List of estimated indicators and result data

Opierając się na możliwościach technicznych dotyczących zasobu informacyjnego i narzędzi analitycznych oraz możliwościach merytorycznych dotyczących eksperckiej wiedzy i kompetencji zespołu realizującego niniejszą pracę badawczą, przy zastosowaniu każdej z zaproponowanych metod także wynikach przeprowadzonych symulacji, naliczono wskaźniki i dane będące celem pracy badawczej oszacowane ze stanem za dany okres referencyjny:

- 1) liczba pracujących na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 2) liczba pracujących na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i miasto/wieś,
- 3) liczba pracujących na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć w sektorze przemysłowym i usługowym,
- 4) liczba pracujących na poziomie NTS 4 według miejsca siedziby pracy głównej w podziale na grupy sekcji PKD 2007 i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 5) liczba pracujących na poziomie NTS 4 według miejsca siedziby pracy głównej w podziale na grupy sekcji PKD 2007 i miasto/wieś,
- 6) liczba pracujących na poziomie NTS 4 według miejsca zamieszkania w podziale na grupy sekcji PKD 2007 i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 7) liczba pracujących na poziomie NTS 4 według miejsca zamieszkania w podziale na grupy sekcji PKD 2007 i miasto/wieś,
- 8) stopa bezrobocia rejestrowanego na poziomie NTS 5 według miejsca zamieszkania i płci,

- 9) stopa bezrobocia rejestrowanego na poziomie NTS 5 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 10) stopa bezrobocia rejestrowanego na poziomie NTS 5 według miejsca zamieszkania, płci i miasto/wieś,
- 11) przeciętne wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania, grupy sekcji PKD 2007 i płci,
- 12) przeciętne wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 13) przeciętne wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na sektory działalności i miasto/wieś,
- 14) przeciętne wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności,
- 15) przeciętne wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej i miasto/wieś,
- 16) przeciętne wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na grupy sekcji PKD 2007 i miasto/wieś,
- 17) najniższe i najwyższe wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania, grup sekcji PKD 2007 i płci,
- 18) najniższe i najwyższe wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 19) najniższe i najwyższe miesięczne wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności,
- 20) kwartyle (I i III) wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania, grup sekcji PKD 2007 i płci,
- 21) kwartyle (I i III) wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 22) kwartyle (I i III) wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności,
- 23) wybrane decyle (I, V i IX) wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania, grup sekcji PKD 2007 i płci,
- 24) wybrane decyle (I, V i IX) wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 25) wybrane decyle (I, V i IX) wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności,
- 26) najczęstsze wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania, grup sekcji PKD 2007 i płci,
- 27) najczęstsze wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 28) najczęstsze wynagrodzenie brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności,
- 29) kurtoza rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania, grup sekcji PKD 2007 i płci,
- 30) kurtoza rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 31) kurtoza rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności,
- 32) współczynnik Giniego rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania, grup sekcji PKD 2007 i płci,
- 33) współczynnik Giniego rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 34) współczynnik Giniego rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności,
- 35) krzywa Lorenza rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania, grup sekcji PKD 2007 i płci,
- 36) krzywa Lorenza rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i grupy wieku 15-24, 25-54, 55-64, 65 lat i więcej,
- 37) krzywa Lorenza rozkładu wynagrodzeń brutto na poziomie NTS 4 według miejsca zamieszkania w podziale na płeć i sektory działalności.

2.5. Ocena jakości badania

Assessment of survey quality

Uzyskane wyniki badania poddano analizie i ocenie jakości pod względem przydatności, dokładności, terminowości, dostępności, porównywalności, spójności oraz bezpieczeństwa danych.

Przydatność

Odbiorcy informacji z zakresu rynku pracy oczekują danych na jak najniższych poziomach agregacji, co było celem niniejszej pracy badawczej. Do szerokiego grona odbiorców tych informacji należą: centralne i terenowe organy administracji publicznej (administracja rządowa i samorządowa), głównie: Ministerstwo Rodziny, Pracy i Polityki Społecznej, Publiczne służby zatrudnienia, instytucje naukowe (pracownicy naukowci, studenci), organizacje pracodawców, organizacje związkowe, instytucje finansowe, media oraz inne osoby i podmioty zainteresowane zagadnieniami związanymi z rynkiem pracy.

Z uwagi na konieczność monitorowania zatrudnienia, zwalczanie bezrobocia, rozwijania badań naukowych, wprowadzania innowacji, jak również wzmacniania konkurencyjności przedsiębiorstw głównie małych i średnich jest potrzeba dostarczania informacji o lokalnych rynkach pracy.

Wyniki pracy badawczej mogą być wykorzystywane w analizach ekonomicznych i społecznych w celu prowadzenia efektywnej polityki społecznej i gospodarczej na lokalnych rynkach pracy. Uzyskane dane statystyczne o pracujących w głównym miejscu pracy według miejsca zamieszkania oraz stopie bezrobocia rejestrowanego w gminach mogą pomóc w podejmowaniu działań ukierunkowanych na monitorowanie zatrudnienia czy przeciwdziałanie bezrobociu. Dotychczas niedostępne dane mogą również stanowić podstawę do prowadzenia badań naukowych, których celem może być ustalenie nowych niezbadanych wcześniej zależności.

Wyniki pracy badawczej pozwolą wprowadzić dodatkowe informacje do badań ujętych w programie badań statystycznych statystyki publicznej: *Pracujący w gospodarce narodowej, Bezrobotni i poszukujący pracy zarejestrowani w urzędach pracy oraz Wynagrodzenia w gospodarce narodowej*.

Wyników pracy badawczej nie należy traktować jako ostateczne z uwagi na trwające prace mające na celu doskonalenie dotychczasowej metody badawczej.

Dokładność

Badaną populacją są osoby uznane za pracujące, zatrudnione i bezrobotne, które stanowiły ogólną populację badania. Nie występują więc rozbieżności pomiędzy populacją docelową a populacją badaną.

Wykaz błędów związanych z przetwarzaniem i brakiem danych został szczegółowo przedstawiony w załączniku nr 35.

Terminowość i punktualność

Praca badawcza realizowana była w 2017 i 2018 roku i dotyczyła danych za wybrane miesiące 2016 i 2017 roku. Dane o pracujących dotyczyły okresów: wrzesień i grudzień 2016 roku oraz marzec i czerwiec 2017 roku. Dane na temat stopy bezrobocia rejestrowanego zostały naliczone za marzec i czerwiec 2017 roku, natomiast miary wynagrodzeń wyznaczono za miesiące: marzec, czerwiec, wrzesień i grudzień 2016 roku oraz marzec i czerwiec 2017 roku.

Poszczególne etapy pracy badawczej realizowane były zgodnie z harmonogramem badania.

Dostępność i przejrzystość

Wyniki pracy badawczej zostały przedstawione w niniejszym raporcie zawierającym analizę danych oraz wyjaśnienia metodologiczne. Zestawienie otrzymanych danych zawarte jest w formie tabelarycznej w załącznikach nr 1-32 w postaci tablic. W raporcie znalazły się również wizualizacje wybranych danych za pomocą wykresów i kartogramów. Wyniki badania zostaną również zaprezentowane na spotkaniu podsumowującym projekt.

Porównywalność

W aspekcie pojęciowym i metodologicznym nie ma różnic, które miałyby wpływ na porównywalność danych z wyjątkiem dwóch okresów tj. marzec i czerwiec 2016 roku dotyczących miar wynagrodzeń.

W tych dwóch okresach wystąpiła konieczność nadania kodów TERYT na podstawie kodów pocztowych, nazw miejscowości i nazw gmin. W pozostałych zbiorach kod TERYT uzupełniany był w Ośrodku Przetwarzania Danych US Warszawa. W związku z tym między okresami w 2016 i 2017 roku może występować utrudnienie w porównywalności czasowej i przestrzennej danych.

Poza wymienionymi wyżej wyjątkami metodologia zastosowana w celu wyznaczenia liczby pracujących, stopy bezrobocia oraz miar wynagrodzeń była identyczna dla wszystkich badanych okresów, co zapewniło wysoki stopień porównywalności otrzymanych wyników zarówno w aspekcie czasowym jak i przestrzennym.

Spójność

Badanie przeprowadzone w pracy badawczej jest spójne z badaniami z tej samej dziedziny prowadzonymi zgodnie z programem badań statystyki publicznej, tj. *Pracujący w gospodarce narodowej, Bezrobotni i poszukujący pracy zarejestrowani w urzędach pracy oraz Wynagrodzenia w gospodarce narodowej*. Podobieństwa i różnice metodologiczne między niniejszą pracą badawczą a wymienionymi badaniami statystyki publicznej pod względem definicji zmiennych, zakresu podmiotowego, okresu referencyjnego, częstotliwości, stosowanych przekrojów zostały przedstawione w rozdziale 2.

Bezpieczeństwo danych

Dostęp do danych odbywał się poprzez Operacyjną Bazę Mikrodanych (OBM), która jest środowiskiem chronionym z ograniczonym dostępem. Dostęp do danych miały wyłącznie osoby posiadające odpowiednie upoważnienie.

Podsumowanie

Istotna z punktu widzenia pracy badawczej jest ocena okresów od września 2016 r. do czerwca 2017 r. z uwagi na fakt, iż od tego momentu zbiory przekazywane przez ZUS zawierają unikalny numer osoby ubezpieczonej PESEL, co pozwoli w przyszłości na polepszenie poziomu komponentów oceny jakości: dokładność i porównywalność. Wyniki badania w tych okresach charakteryzują się wysokim poziomem porównywalności, co przy kontynuacji badania na następne okresy gwarantuje wysoką akceptowalność danych. Ponadto możliwość prezentowania wyników na niskim poziomie agregacji zadowoli szerokie grono odbiorców, które takich danych o rynku pracy oczekuje.

Bibliografia

Bibliography

- Korczyński A. (2014). Review Of Methods For Data Sets With Missing Values And Practical Applications. "Silesian Statistical Review" 12(18).
- Little R.J.A., Rubin D.B. (2002), Statistical Analysis with Missing Data, 2nd Edition, A John Wiley & Sons, Inc., Hoboken, New Jersey.
- Ostasiewicz W. (2004). Ocena i analiza jakości życia. Wydawnictwo Akademii Ekonomicznej, Wrocław.
- Roszka W. (2013), Statystyczna integracja danych w badaniach społeczno-ekonomicznych, Uniwersytet Ekonomiczny w Poznaniu.
- Sobczyk M. (2018), Statystyka, Wydawnictwo Naukowe PWN, Warszawa.
- Van Buuren, S., Groothuis-Oudshoorn, K. (2011). MICE: Multivariate Imputation by Chained Equations in R. "Journal of Statistical Software" 45(3).
- Van der Laan P. (2000), Integrating administrative registers and household surveys, „Netherlands Official Statistics”, vol. 15, Summer 2000, Special issue: Integrating administrative registers and household surveys, Statistics Netherlands, Voor-burg/Heerlen.
- Wesołowski J., Tarczyński J. (2016). Podstawy matematyczne technik imputacyjnych. "Wiadomości Statystyczne" nr 9.
- Kodeks cywilny (Dz. U. z 2017 r. poz. 459 i 933).
- Ustawa z dnia 13 października 1998 r. o systemie ubezpieczeń społecznych (t.j. Dz. U. z 2017 r. poz. 1778).
- Ustawa z dnia 27 sierpnia 2004 r. o świadczeniach opieki zdrowotnej finansowanych ze środków publicznych (t.j. Dz. U. z 2017 r. poz. 1938 z późn. zm.).
- Ustawa z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy (t.j. Dz.U.2017,poz.1065 z późn.zm.).
- Ustawa z dnia 20 grudnia 1990 r. o ubezpieczeniu społecznym rolników (t.j. Dz. U. z 2016 r., poz. 277 z późn. zm.).
- Ustawa z dnia 29 sierpnia 1997 r. Ordynacja podatkowa (t.j. Dz. U. z 2017r., poz. 201 z późn. zm.).
- Ustawa z dnia 29 sierpnia 1997 r. o ochronie danych osobowych (t.j. Dz. U. z 2016 r., poz. 922).
- Ustawa z dnia 25 czerwca 1999 r. o świadczeniach pieniężnych z ubezpieczenia społecznego w razie choroby i macierzyństwa (t.j. Dz. U. 2017, poz. 1368).
- Rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 21 grudnia 2009 r. w sprawie szczegółowego zakresu danych zawartych w centralnych rejestrach prowadzonych przez Zakład Ubezpieczeń Społecznych (t.j. Dz. U. 2013, poz. 219).
- Rozporządzenie Ministra Pracy i Polityki Społecznej z dnia 23 października 2009 r. w sprawie określenia wzorów zgłoszeń do ubezpieczeń społecznych i ubezpieczenia zdrowotnego, imiennych raportów miesięcznych i imiennych raportów miesięcznych korygujących, zgłoszeń płatnika, deklaracji rozliczeniowych i deklaracji rozliczeniowych korygujących, zgłoszeń danych o pracy w szczególnych warunkach lub o szczególnym charakterze oraz innych dokumentów (t.j. Dz. U. 2016, poz. 222 z późn. zm.).
- <http://stat.gov.pl/obszary-tematyczne/rynek-pracy/>
- <http://stat.gov.pl/metainformacje/sownik-pojec/pojecia-stosowane-w-statystyce-publicznej/>
- <http://sms.stat.gov.pl/>
- <https://www.krus.gov.pl/>
- <http://www.zus.pl/sownik>

Spis załączników *List of appendices*

1. Załącznik nr 1 – Pracujący w gospodarce narodowej w głównym miejscu pracy
2. Załącznik nr 2 – Stopa bezrobocia rejestrowanego
3. Załącznik nr 3 – Przeciętne wynagrodzenie brutto według grup sekcji PKD, płci i miejsca zamieszkania
4. Załącznik nr 4 – Wartości skrajne wynagrodzeń brutto według grup sekcji PKD, płci i miejsca zamieszkania
5. Załącznik nr 5 – Kwartyle wynagrodzeń brutto według grup sekcji PKD, płci i miejsca zamieszkania.
6. Załącznik nr 6 – Decyle wynagrodzeń brutto według grup sekcji PKD, płci i miejsca zamieszkania
7. Załącznik nr 7 – Dominanta wynagrodzeń brutto według grup sekcji PKD, płci i miejsca zamieszkania
8. Załącznik nr 8 – Kurtoza wynagrodzeń brutto według grup sekcji PKD, płci i miejsca zamieszkania
9. Załącznik nr 9 – Współczynnik Giniego wynagrodzeń brutto według grup sekcji PKD, płci i miejsca zamieszkania
10. Załącznik nr 10 – Przeciętne wynagrodzenie brutto według grup wieku, płci i miejsca zamieszkania
11. Załącznik nr 11 – Wartości skrajne wynagrodzeń brutto według grup wieku, płci i miejsca zamieszkania
12. Załącznik nr 12 – Kwartyle wynagrodzeń brutto według grup wieku, płci i miejsca zamieszkania
13. Załącznik nr 13 – Decyle wynagrodzeń brutto według grup wieku, płci i miejsca zamieszkania
14. Załącznik nr 14 – Dominanta wynagrodzeń brutto według grup wieku, płci i miejsca zamieszkania
15. Załącznik nr 15 – Kurtoza wynagrodzeń brutto według grup wieku, płci i miejsca zamieszkania
16. Załącznik nr 16 – Współczynnik Giniego wynagrodzeń brutto według grup wieku, płci i miejsca zamieszkania
17. Załącznik nr 17 – Przeciętne wynagrodzenie brutto według sektorów działalności, płci i miejsca zamieszkania
18. Załącznik nr 18 – Wartości skrajne wynagrodzeń brutto według sektorów działalności, płci i miejsca zamieszkania
19. Załącznik nr 19 – Kwartyle wynagrodzeń brutto według sektorów działalności, płci i miejsca zamieszkania
20. Załącznik nr 20 – Decyle wynagrodzeń brutto według sektorów działalności, płci i miejsca zamieszkania
21. Załącznik nr 21 – Dominanta wynagrodzeń brutto według sektorów działalności, płci i miejsca zamieszkania
22. Załącznik nr 22 – Kurtoza wynagrodzeń brutto według sektorów działalności, płci i miejsca zamieszkania
23. Załącznik nr 23 – Współczynnik Giniego wynagrodzeń brutto według sektorów działalności, płci i miejsca zamieszkania
24. Załącznik nr 24 – Przeciętne wynagrodzenie brutto według grup sekcji PKD i miejsca zamieszkania (miasto, wieś)
25. Załącznik nr 25 – Przeciętne wynagrodzenie brutto według grup wieku i miejsca zamieszkania (miasto, wieś)
26. Załącznik nr 26 – Przeciętne wynagrodzenie brutto według sektorów działalności i miejsca zamieszkania (miasto, wieś)
27. Załącznik nr 27 – Krzywa Lorenza rozkładu wynagrodzeń brutto w marcu 2016 r.
28. Załącznik nr 28 – Krzywa Lorenza rozkładu wynagrodzeń brutto w czerwcu 2016 r.
29. Załącznik nr 29 – Krzywa Lorenza rozkładu wynagrodzeń brutto we wrześniu 2016 r.
30. Załącznik nr 30 – Krzywa Lorenza rozkładu wynagrodzeń brutto w grudniu 2016 r.
31. Załącznik nr 31 – Krzywa Lorenza rozkładu wynagrodzeń brutto w marcu 2017 r.
32. Załącznik nr 32 – Krzywa Lorenza rozkładu wynagrodzeń brutto w czerwcu 2017 r.
33. Załącznik nr 33 – Wykaz kodów stosowanych w ZUS
34. Załącznik nr 34 – Struktury zbiorów źródłowych
35. Załącznik nr 35 – Zestawienie błędów i braków w zbiorach KRUS i ZUS
36. Załącznik nr 36 – Zakres składników wynagrodzeń
37. Załącznik nr 37 – Wykaz szczegółowych procedur przygotowujących zbiory KRUS i ZUS
38. Załącznik nr 38 – Liczba braków przeznaczonych do imputacji