**UWAGI METODYCZNE**

**1. Definicje i pojęcia spisowe**

Podstawą definicji i pojęć zastosowanych w Narodowym Spisie Powszechnym Ludności i Mieszkań przeprowadzonym w 2011 r. są Zalecenia międzynarodowe do spisów ludności i mieszkań około 2010 roku przyjęte przez Konferencję Statystyków Europejskich w czerwcu 2006 r. (ONZ, Genewa 2006)[[1]](#footnote-1).  Definicje dotyczące aktywności ekonomicznej ludności  opierają się na Rezolucji dotyczącej statystyki pracujących, bezrobotnych i niepełnozatrudnionych, przyjętej na XIII Międzynarodowej Konferencji Statystyków Pracy w październiku 1982 r. (z późniejszymi zmianami) i zalecane do stosowania przez Międzynarodową Organizację Pracy[[2]](#footnote-2)).

Wyodrębnienie pracujących, bezrobotnych i biernych zawodowo, czyli podstawowych kategorii ludności z punktu widzenia aktywności ekonomicznej (statusu osób na rynku pracy), jest możliwe dopiero po wcześniejszym określeniu całej zbiorowości, czyli precyzyjnym zdefiniowaniu pojęcia „ludność”. Ze względu na przepisy prawa obowiązujące w poszczególnych krajach – wyznaczające dolną granicę wieku umożliwiającą „wejście” na rynek pracy – w spisach ludności przyjęto określanie statusu na rynku pracy dla osób w wieku 15 lat i więcej.

* 1. **Ludność, gospodarstwa domowe i rodziny**

Na podstawie wyników NSP 2011 możliwe jest wyodrębnienie kilku kategorii ludności.[[3]](#footnote-3) W niniejszej publikacji dane z zakresu aktywności ekonomicznej dotyczą wyłącznie ludności faktycznie zamieszkałej – dalej określonej jako „ludność”.

**Ludność faktycznie zamieszkała (ludność faktyczna)** – kategoria ta, z punktu widzenia mieszkańców gminy, tj. najmniejszej jednostki podziału administracyjnego kraju – obejmuje następujące grupy:

1. Osoby mieszkające stale (stali mieszkańcy)– są to osoby, które w spisie zadeklarowały, że dana gminajest ich stałym miejscem zamieszkania (bez względu czy są tam zameldowane na pobyt stały czy nie)oraz – w momencie spisu:
   1. były obecne (mieszkały w gminie),
   2. były nieobecne, ale ich nieobecność trwała krócej niż 3 miesiące,
   3. były nieobecne dłużej niż 3 miesiące, ale ich nieobecność wynikała z powodu:

* przebywania w zakładzie karnym lub śledczym
* pobytu za granicą.

1. Osoby przebywające czasowo dłużej niż 3 miesiące – są to osoby, które w spisie zadeklarowały, że ich stałym miejscem zamieszkania jest inna gmina w kraju, natomiast w miejscu/gminie spisania przebywały czasowo (dłużej niż 3 miesiące) z następujących powodów: nauka, praca, warunki rodzinne lub mieszkaniowe, leczenie lub rehabilitacja, przebywanie w domu opieki.

Jako czas okresowej nieobecności lub przebywania przyjęty został czas zamierzony.

Do kategorii ludności faktycznej danej gminy nie były zaliczane osoby przybyłe z zagranicy na pobyt czasowy, tj. takie, które nie posiadały stałego pobytu w Polsce (pozwolenia na osiedlenie się). Kategoria nie ujmuje też stałych mieszkańców danej gminy, którzy w momencie spisu przebywali poza gminą dłużej niż 3 miesiące z przyczyn innych niż przebywanie w zakładzie karnym (śledczym) lub pobyt za granicą, tj. osób które przebywały czasowo w innej gminie kraju (zgodnie z przyjętą definicją zostali tam ujęci jako ludność faktyczna). Analogicznie wyodrębniana jest ludność faktyczna dla poszczególnych miejscowości.

**Gospodarstwa domowe[[4]](#footnote-4)**

W NSP 2011 gospodarstwa domowe ustalane były w mieszkaniach wylosowanych do badania reprezentacyjnego.

Gospodarstwa domowe były wyodrębniane **według definicji ekonomicznej**, tzn. przy zastosowaniu kryterium wspólnego zamieszkiwania i utrzymywania się osób.

**Gospodarstwo domowe** definiowane jest jako zespół osób (spokrewnionych lub niespokrewnionych) wspólnie zamieszkujących i utrzymujących się. Wspólnota utrzymywania członków danego zespołu decydowała o tym, czy w danym mieszkaniu występowało jedno, czy też więcej gospodarstw domowych. Osoba mieszkająca samotnie lub mieszkająca z innymi osobami, ale utrzymująca się oddzielnie tworzy odrębne – jednoosobowe – gospodarstwo domowe.

W każdym gospodarstwie domowym wyodrębniano **„reprezentanta”** gospodarstwa domowego wobec którego ustalano relacje pokrewieństwa/powinowactwa wszystkich pozostałych członków gospodarstwa domowego.

**Rodziny** w spisach wyodrębnia się spośród osób wchodzących w skład gospodarstwa domowego.

**Rodzina (biologiczna)** w spisie definiowana jest jako dwie lub większa liczba osób, które są związane jako mąż i żona, wspólnie żyjący partnerzy (kohabitanci) – osoby płci przeciwnej lub jako rodzic i dziecko. Tak więc, rodzina obejmuje parę bez dzieci lub parę z jednym lub większą liczbą dzieci, albo też samotnego rodzica z jednym bądź większą liczbą dzieci.

Jako **dziecko** określa się osobę w każdym wieku, która pozostaje w gospodarstwie domowym (lub przebywa w gospodarstwie zbiorowym) wraz z obojgiem lub jednym z rodziców. Do dzieci zalicza się również pasierbów oraz dzieci przysposobione.

**Dziecko do lat 24** pozostające na utrzymaniu jest to dziecko w wieku 0-24 lata, które nie posiada własnego źródła utrzymania i pozostaje na utrzymaniu innej osoby wchodzącej   
w skład gospodarstwa domowego.

Dane wynikowe dotyczące gospodarstw domowych i rodzin opracowane były tylko dla osób zamieszkałych w mieszkaniach, wylosowanych do badania reprezentacyjnego.   
Do ustalania struktury gospodarstw domowych i rodzin oraz ich typów nie były uwzględniane osoby bezdomne (bez dachu nad głową).

**1.2. Charakterystyka ekonomiczna**

W spisie zastosowano definicje **bieżącej** **aktywności ekonomicznej** ludności, zgodnie z którą podstawą podziału ludności na pracujących, bezrobotnych i biernych zawodowo jest praca, czyli fakt wykonywania, posiadania bądź poszukiwania pracy w określonym przedziale czasu. W przypadku NSP 2011 jako okres odniesienia przyjęto tydzień od 25 do 31 marca 2011 r. Zastosowana kolejność wyodrębniania poszczególnych zbiorowości gwarantuje zaklasyfikowanie każdej osoby tylko do jednej z trzech kategorii ludności wyróżnianych z punktu widzenia statusu na rynku pracy.

**Pracujący** – do kategorii tej zaliczono wszystkie osoby w wieku 15 lat i więcej, które w okresie badanego tygodnia (25 – 31 marca 2011 r.):

* wykonywały przez co najmniej 1 godzinę pracę przynoszącą zarobek lub dochód tzn. były zatrudnione w charakterze pracownika najemnego, pracowały we własnym (lub dzierżawionym) gospodarstwie rolnym lub prowadziły własną działalność gospodarczą poza rolnictwem, pomagały (bez wynagrodzenia) w prowadzeniu rodzinnego gospodarstwa rolnego lub rodzinnej działalności gospodarczej poza rolnictwem,
* nie wykonywały pracy (np. z powodu choroby, urlopu, przerwy w działalności zakładu, trudnych warunków atmosferycznych, strajku), ale formalnie miały pracę jako pracownicy najemni bądź pracujący na własny rachunek.

Zgodnie z „Zaleceniami międzynarodowymi do spisów ludności i mieszkań około 2010 roku” (ONZ, Genewa 2006) – pomagających członków rodziny, którzy nie wykonywali pracy w badanym tygodniu zaliczono do osób niepracujących.

**Bezrobotni** – są to osoby w wieku 15-74 lata, które spełniły jednocześnie trzy warunki:

* w okresie badanego tygodnia nie były osobami pracującymi,
* aktywnie poszukiwały pracy w okresie od 1 do 31 marca, tzn. podjęły przynajmniej jedno z następujących działań aby znaleźć pracę (niezależnie od uzyskanych efektów): odpowiadanie na ogłoszenia, poszukiwanie pracy przez krewnych i znajomych bądź bezpośrednio w zakładach pracy, podjęcie starań o zorganizowanie własnego miejsca pracy, zarejestrowanie się w biurze pośrednictwa pracy jako poszukujący pracy,
* były gotowe (zdolne) podjąć pracę w ciągu dwóch tygodni po tygodniu badanym, tj. w okresie od 1 do 14 kwietnia.

Do bezrobotnych zaliczono także osoby, które znalazły pracę i oczekiwały na jej rozpoczęcie w okresie 3 miesięcy oraz były gotowe tę pracę podjąć.

**Bierni zawodowo** – są to osoby w wieku 15 lat i więcej, które nie zostały zaklasyfikowane jako pracujące lub bezrobotne, tzn. osoby, które w badanym tygodniu:

* nie pracowały, nie miały pracy i jej nie poszukiwały,
* nie pracowały, poszukiwały pracy, ale nie były gotowe (zdolne) do jej podjęcia w ciągu dwóch tygodni po tygodniu badanym, tj. w okresie od 1 do 14 kwietnia,
* nie pracowały i nie poszukiwały pracy, ponieważ miały pracę załatwioną i oczekiwały na jej rozpoczęcie w okresie: dłuższym niż 3 miesiące lub do 3 miesięcy, ale nie były gotowe tej pracy podjąć.

**Aktywni zawodowo** – wszystkie osoby uznane za pracujące lub bezrobotne, zgodnie z definicjami podanymi powyżej.

***UWAGA.*** *W NSP 2011* ***charakterystyka ludności z punktu widzenia aktywności ekonomicznej jest możliwa jedynie w przypadku*** ***osób zamieszkujących w mieszkaniach****. Osobom przebywającym w obiektach zbiorowego zakwaterowania oraz bezdomnym nie zadawano pytań dotyczących ich sytuacji na rynku pracy. W związku z tym w tablicach dotyczących aktywności ekonomicznej ludności osoby te uwzględniono w rubryce „nieustalony status na rynku pracy”. Rubryka ta uwzględnia również osoby, które brały udział w spisie, ale nie udało się dla nich uzyskać informacji o statusie na rynku pracy, a także osoby, które nie wzięły udziału w spisie (np. emigranci, osoby, które zrezygnowały z udziału w spisie) i nie udało się dla nich pozyskać podstawowych informacji o ich statusie na rynku pracy na podstawie rejestrów administracyjnych.*

Do podstawowych wskaźników opisujących sytuację na rynku pracy należą:

**Współczynnik aktywności zawodowej** – procentowy udział aktywnych zawodowo danego grupowania w ogólnej liczbie ludności danego grupowania.

**Wskaźnik zatrudnienia** – procentowy udział pracujących danego grupowania w ogólnej liczbie ludności danego grupowania.

**Stopa bezrobocia** – procentowy udział bezrobotnych danego grupowania w liczbie aktywnych zawodowo danego grupowania.

***Uwaga.*** *W tablicach**powyższe**wskaźniki zostały obliczone bez uwzględniania osób o nieustalonym statusie na rynku pracy****.***

Moduł „**Dojazdy do pracy**” został opracowany i dostosowany do „Rozporządzenia Komisji (WE) nr 1201/2009 z dnia 30 listopada 2009 r. w sprawie wykonania rozporządzenia (WE) nr 763/2008 Parlamentu Europejskiego i Rady w sprawie spisów powszechnych ludności i mieszkań w zakresie specyfikacji technicznych tematów i dotyczących ich podziałów”.

W oparciu o wyniki badania reprezentacyjnego badano osoby pracujące w Polsce lub na terenach przygranicznych (mieszkające w Polsce i regularnie przekraczające granicę) bez względu na ich status zatrudnienia.

**Lokalizacja miejsca pracy** to dokładnie to miejsce, w którym osoba pracująca wykonuje swoją pracę główną. Jeżeli osoba pracuje w oddziale, filii lub innej składowej części przedsiębiorstwa (firmy), tj. w jednostce lokalnej, należało odnieść się do faktycznego miejsca pracy.

**Dojeżdżający do pracy** to osoby pracujące w Polsce lub na terenach przygranicznych (mieszkające w Polsce i regularnie przekraczające granicę) i przemieszczające się do pracy środkiem transportu, również w obrębie gminy czy miejscowości zamieszkania.

Osoby nieposiadające stałego miejsca pracy to te, dla których trudno jest ustalić stałe miejsce pracy, ze względu na jej specyficzny charakter, np. marynarze, geodeci, zawodowi kierowcy, akwizytorzy, itp.

Temat „Dojazdy do pracy” realizowany był w oparciu o dane z badania reprezentacyjnego (analizę tematu zamieszczono w niniejszej publikacji) jak i danych ze źródeł administracyjnych (publikacja „Dojazdy do pracy – NSP 2011” zamieszczona jest na stronie GUS – [www.stat.gov.pl](http://www.stat.gov.pl) ).

Należy zwrócić uwagę i podkreślić, że wyniki badania dojazdów do pracy na podstawie rejestrów administracyjnych nie są porównywalne z danymi uzyskanymi z badania reprezentacyjnego, głównie ze względu na odmienność źródeł danych i różnice metodologiczne.

**Źródła utrzymania**

W spisie powszechnym ustalano główne i dodatkowe źródło utrzymania dla poszczególnych osób oraz odrębnie dla gospodarstw domowych. Informacje o źródłach dochodów, z których pochodziły środki na finansowanie konsumpcyjnych i innych życiowych potrzeb osób dotyczyły całego roku (12 miesięcy) poprzedzającego moment spisu. Ustalano je niezależnie od stanu aktywności ekonomicznej osób.

**Główne źródło utrzymania osoby** jest to źródło przynoszące jej największy zarobek lub dochód. Jeżeli osoba miała w ciągu roku poprzedzającego spis tylko jedno źródło dochodów – było ono jednocześnie jej głównym, a zarazem wyłącznym źródłem utrzymania.

Za **dodatkowe źródło utrzymania** uważa się źródło najważniejsze spośród wszystkich pozostałych po wyodrębnieniu głównego źródła utrzymania, czyli przynoszące dochód drugi pod względem wysokości.

**1.3. Klasyfikacje i grupowania**

Wszystkie zastosowane klasyfikacje i grupowania wynikają z Rozporządzenia Komisji (WE) nr 1201/2009 z dnia 30 listopada 2009 r. w sprawie wykonania rozporządzenia (WE)   
nr 763/2008 Parlamentu Europejskiego i Rady w sprawie spisów powszechnych ludności i mieszkań w zakresie specyfikacji technicznych tematów i dotyczących ich podziałów.

**Wiek**

Wiek osób – określony liczbą lat ukończonych – ustalono poprzez porównanie pełnej daty urodzenia z datą przeprowadzenia spisu (tzw. dniem krytycznym, tj. 31 marca 2011 roku).

W tablicach informacje są podawane również wg ekonomicznych grup wieku:

* wiek przedprodukcyjny – mężczyźni i kobiety w wieku 15-17 lat,
* wiek produkcyjny – mężczyźni w wieku 18-64 lata, kobiety w wieku 18-59 lat,
* wiek mobilny (18-44 lata mężczyźni i kobiety),
* wiek niemobilny (45-64 lata – mężczyźni i 45-59 lat – kobiety),
* poprodukcyjny – mężczyźni w wieku 65 lat i więcej oraz kobiety w wieku 60 lat i więcej.

***UWAGA.*** *W tablicach dotyczących aktywności ekonomicznej ludności, o ile nie zaznaczono inaczej, informacje dotyczą ludności w wieku 15 lat i więcej (15-74 lata w przypadku tablic dla osób bezrobotnych).*

**Poziom wykształcenia**

Zgodnie z międzynarodową klasyfikacją kształcenia (*International Standard Classification of Education)* prezentowane poziomy opisują najwyższy ukończony cykl kształcenia w szkole lub szkolenia w innym trybie lub formie, uznany zgodnie z obowiązującym systemem szkolnictwa. Podstawą zaliczenia wykształcenia do określonego poziomu jest uzyskane świadectwo (dyplom) ukończenia odpowiedniej szkoły niezależnie od trybu jej ukończenia (dziennie, wieczorowo, zaocznie czy eksternistycznie).

W tablicach zastosowano następujące grupowania poziomu wykształcenia:

**Wykształcenie wyższe**

dotyczy osób:

* ze stopniem naukowym: doktora, doktora habilitowanego, profesora,
* z tytułem magistra, lekarza lub równorzędnym, uzyskanym po ukończeniu studiów drugiego stopnia (uzupełniających magisterskich) lub jednolitych studiów magisterskich,
* z tytułem zawodowym: inżyniera, licencjata, dyplomowanego ekonomisty, uzyskanym po ukończeniu studiów pierwszego stopnia (licencjackich, inżynierskich);

**Wykształcenie policealne**

dotyczy:

* absolwentów kolegiów nauczycielskich i nauczycielskich kolegiów języków obcych (z wyjątkiem organizowanych w ramach szkoły wyższej) oraz absolwentów kolegiów pracowników służb społecznych; do tej kategorii należy zaliczyć również osoby, które ukończyły studium nauczycielskie,
* osób, które posiadają dyplom (świadectwo) ukończenia szkoły pomaturalnej, do której przyjęcie było uwarunkowane posiadaniem świadectwa dojrzałości (z maturą, pomaturalne),
* osób, które posiadają dyplom/świadectwo ukończenia szkoły policealnej, do której przyjęcie było uwarunkowane posiadaniem świadectwa ukończenia szkoły średniej (bez matury).

**Wykształcenie średnie zawodowe:**

dotyczy osób, które:

* otrzymały świadectwo dojrzałości (maturę) w szkole średniej zawodowej – technikum, technikum uzupełniającym, liceum zawodowym, liceum technicznym, szkole artystycznej II stopnia (zawodowe z maturą),
* otrzymały świadectwo ukończenia szkoły średniej zawodowej – technikum, technikum uzupełniającego, liceum zawodowego, liceum technicznego, szkoły artystycznej II stopnia (zawodowe bez matury).

**Wykształcenie średnie ogólnokształcące:**

dotyczy osób, które:

* otrzymały świadectwo dojrzałości (maturę) uzyskane w liceum ogólnokształcącym lub w liceum profilowanym lub ukończyły gimnazjum przed 1932 rokiem (ogólnokształcące z maturą),
* otrzymały świadectwo ukończenia liceum ogólnokształcącego lub liceum profilowanego (ogólnokształcące bez matury).

**Wykształcenie zasadnicze zawodowe:**

dotyczy osób, które posiadają świadectwo ukończenia: szkoły zawodowej (zasadniczej szkoły zawodowej lub rolniczej, szkoły przemysłowej, gimnazjum zawodowego itp.), szkoły przysposobienia zawodowego lub rolniczego, korespondencyjnego kursu rolniczego wyłącznie o poziomie zasadniczej szkoły zawodowej oraz szkoły mistrzów.

**Wykształcenie podstawowe ukończone i gimnazjalne:**

dotyczy osób, które:

* uzyskały świadectwo ukończenia gimnazjum (gimnazja zaczęły funkcjonować w roku szkolnym 1999/2000); ponadto do tej kategorii zalicza się absolwentów specjalnych szkół przysposabiających do pracy (wykształcenie gimnazjalne),
* posiadających świadectwo ukończenia: szkoły podstawowej (przed wojną powszechnej) niezależnie od tego, ile klas liczyła (sześć, osiem, czy dawniej cztery), kursów dla pracujących w zakresie szkoły podstawowej, szkoły artystycznej I stopnia, realizującej jednocześnie program szkoły podstawowej (wykształcenie podstawowe ukończone).

**Wykształcenie podstawowe nieukończone i bez wykształcenia szkolnego:**

dotyczy osób, które uczą się w szkole podstawowej, uczyły się w szkole podstawowej, lecz jej nie ukończyły, bądź nigdy nie uczęszczały do szkoły.

**Klasyfikacja statusu zatrudnienia**

Zgodnie z Międzynarodową Klasyfikacją Statusu Zatrudnienia wyróżnia się następujące kategorie pracujących:

**Pracownicy najemni** – do tych osób zalicza się:

* osoby zatrudnione na podstawie stosunku pracy (umowa o pracę, umowa-zlecenie, umowa o dzieło, powołanie, wybór lub mianowanie) w jednostkach państwowych, spółdzielniach, jednostkach organizacji społecznych, politycznych i związków zawodowych, w jednostkach prywatnych, również u osób fizycznych i w indywidualnych gospodarstwach rolnych,
* osoby wykonujące pracę nakładczą,
* uczniów (oraz osoby przebywające na praktykach), z którymi zakłady pracy lub osoby fizyczne zawarły umowę o naukę zawodu lub przyuczenie do określonej pracy, jeżeli otrzymują wynagrodzenie,
* osoby zatrudnione w rodzinnym gospodarstwie rolnym bądź rodzinnej działalności gospodarczej i jednocześnie otrzymujące za wykonywaną pracę ustalony zarobek lub dochód (niekoniecznie w formie pieniężnej) – niezależnie od związku z gospodarstwem domowym użytkownika gospodarstwa rolnego czy właściciela rodzinnej działalności gospodarczej.

W kategorii pracowników najemnych wyróżniono osoby pracujące w pełnym i niepełnym wymiarze czasu pracy:

* **pracownicy najemni pełnozatrudnieni** – to osoby, które pracują w pełnym wymiarze godzin pracy obowiązującym w danej jednostce, na danym stanowisku pracy lub w danym zawodzie; do pracowników pełnozatrudnionych zalicza się również osoby, które zgodnie z obowiązującymi przepisami pracują w skróconym czasie pracy, osoby zatrudnione na podstawie umowy zlecenia, umowy o dzieło, wyboru lub wykonujących pracę nakładczą jeżeli przepracowali w badanym tygodniu co najmniej 36 godzin,
* **pracownicy najemni niepełnozatrudnieni** – to osoby, które zgodnie z zawartą umową pracują w niepełnym wymiarze godzin pracy obowiązującym w danym zakładzie pracy, na danym stanowisku lub w danym zawodzie.

**Pracujący na własny rachunek** – do tych osób zalicza się:

* właścicieli, współwłaścicieli i dzierżawców indywidualnych gospodarstw rolnych pracujących w tych gospodarstwach,
* członków rolniczych spółdzielni produkcyjnych,
* agentów we wszystkich systemach agencji,
* osoby prowadzące własną działalność gospodarczą poza rolnictwem (zarejestrowaną lub niezarejestrowaną), osoby którym narzucono zarejestrowanie własnej działalności jako 1 – osobowy podmiot gospodarczy.

Wśród pracujących na własny rachunek wyróżnia się:

* **pracodawców** (czyli zatrudniających przynajmniej 1 pracownika najemnego)
* **pracujących na własny rachunek** **niezatrudniających pracowników najemnych.**

**Pomagający członkowie rodzin** – osoby, które bez umownego – ustalonego wcześniej – wynagrodzenia pomagają w prowadzeniu rodzinnej działalności gospodarczej, w tym również w prowadzeniu gospodarstwa rolnego.

Informacje **o zawodach wykonywanych przez pracujących** prezentowane są w układzie Klasyfikacji Zawodów i Specjalności obowiązującej od 2010 roku, opracowanej przez Instytut Pracy i Spraw Socjalnych na podstawie ,,Międzynarodowej Klasyfikacji Zawodów ISCO 08” (,,International Standard Classification of Occupations ISCO 08”) zalecanej do stosowania przez Międzynarodową Organizację Pracy.

Informacje o **rodzaju działalności miejsca pracy** prezentowane są w układzie **Polskiej Klasyfikacji Działalności 2007** wprowadzonej rozporządzeniem Rady Ministrów z dnia 24 grudnia 2007 roku (Dz. U. Nr 251, poz. 1885) w sprawie Polskiej Klasyfikacji Działalności. PKD 2007 została opracowana na podstawie statystycznej klasyfikacji działalności gospodarczej NACE Rev 2.

Pracujący mogą należeć do jednego z trzech sektorów ekonomicznych, na które składają się określone sekcje PKD:

* rolniczy – Rolnictwo, leśnictwo, łowiectwo i rybactwo;
* przemysłowy - Górnictwo i wydobywanie; Przetwórstwo przemysłowe; Wytwarzanie i zaopatrywanie w energię elektryczną, gaz…; Dostawa wody; gospodarowanie ściekami i odpadami…; Budownictwo;
* usługowy – pozostałe sekcje.

**Źródła utrzymania**

Ze względu na posiadane źródła utrzymania ludność dzieli się na dwie kategorie:

* posiadający własne źródła utrzymania, do których zaliczono osoby utrzymujące się z dochodów uzyskiwanych w ciągu całego roku poprzedzającego moment spisu:

- z pracy osobiście wykonywanej, z niezarobkowych źródeł, a także z wynajmu lub z własności,

* utrzymywani, tj. osoby, dla których pozostawanie na utrzymaniu jest źródłem utrzymania.

**Klasyfikacja źródeł utrzymania**

**Praca:**

* **praca najemna w sektorze publicznym** – dla osób utrzymujących się z dochodów z pracy najemnej w jednostkach, stanowiących własność państwową (Skarbu Państwa lub państwowych osób prawnych) lub własność samorządu terytorialnego oraz własność mieszaną z przewagą własności Skarbu Państwa, osób prawnych, własności państwowej lub samorządowej;
* **praca najemna w sektorze prywatnym** – dla osób utrzymujących się z dochodów z pracy najemnej w jednostkach prowadzących działalność gospodarczą i będących własnością krajową, zagraniczną lub mieszaną, osób fizycznych (w zakładach, przedsiębiorstwach, również w indywidualnych gospodarstwach rolnych lub inną) oraz własnością np. spółek, przedsiębiorstw zagranicznych, spółdzielni, partii politycznych, fundacji, organizacji wyznaniowych;
* **praca na rachunek własny poza rolnictwem** – dla osób utrzymujących się z prowadzenia własnej działalności gospodarczej (zarejestrowanej lub niezarejestrowanej), tj. prowadzących zakłady (przedsiębiorstwa) przemysłowe, budowlane, handlowe, hotelowe, lecznice medyczne, lecznice weterynaryjne, zakłady usługowe (różnych branż), szkoły, teatry, przedszkola, żłobki, itp. a także osób wykonujących wolne zawody, np. lekarzy, również lekarzy weterynarii, prawników, artystów, pisarzy, a także osób, którym narzucono zarejestrowanie własnej działalności jako 1-osobowy podmiot gospodarczy. Do pracujących na rachunek własny poza rolnictwem należy zaliczyć agentów we wszystkich systemach agencji. Do utrzymujących się z pracy na rachunek własny należy zaliczyć także: pomagających – systematycznie i bez umownego wynagrodzenia w prowadzeniu działalności gospodarczej poza rolnictwem – członków rodziny oraz osoby uzyskujące dochody z wynajmu mieszkania, pokoi, budynków mieszkalnych i niemieszkalnych, garażu itp. Dochód z wynajmu wyodrębniony może być jako oddzielna pozycja lub grupowany jest razem z pracą na rachunek własny poza rolnictwem.
* **praca na rachunek własny w rolnictwie** – dla osób utrzymujących się z pracy w swoim gospodarstwie rolnym związanej z produkcją roślinną i zwierzęcą, tj. użytkowników (niezależnie czy są właścicielami, dzierżawcami czy też użytkują grunty z innego tytułu), a także właścicieli zwierząt gospodarskich; również osób utrzymujących się z prowadzenia działalności usługowej związanej z rolnictwem, np. nawożenie pól, opryski upraw, obsługa systemów irygacyjnych, sztuczne unasienianie zwierząt, usługi pasterskie, usługi w zakresie leśnictwa lub łowiectwa.   
  Do utrzymujących się pracy na rachunek własny w rolnictwie (w tym w rodzinnym gospodarstwie rolnym) należy zaliczyć także pomagających systematycznie i bez umownego wynagrodzenia członków rodziny. Do grupy tej zaliczono również członków rolniczych spółdzielni produkcyjnych, spółdzielni kółek rolniczych oraz członków spółek cywilnych w rolnictwie.

**Niezarobkowe źródło:**

* **emerytura (pracownicza, kombatancka, rolna)** – dla osób utrzymujących się ze wszystkich rodzajów emerytur przyznanych z tytułu pracy lub ubezpieczenia, w kraju lub za granicą;
* **renta strukturalna** – dla osób, które przekazały gospodarstwo rolne i zaprzestały prowadzenia działalności rolniczej. Renta strukturalna wypłacana jest nie dłużej niż do osiągnięcia 65 roku życia przez beneficjenta;
* **renta z tytułu niezdolności do pracy, renta inwalidzka** – dla osób otrzymujących świadczenia (stałe lub okresowe) uzyskane na podstawie orzeczenia komisji lekarskich (odpowiedniego organu orzekającego) o całkowitej lub częściowej utracie zdolności do wykonywania pracy, jak również dla osób utrzymujących się ze świadczenia rehabilitacyjnego, renty szkoleniowej, renty powypadkowej;
* **renta rodzinna (wdowia, sieroca)** – dla osób utrzymujących się z dochodów ze świadczenia przyznanego uprawnionym członkom rodziny po śmierci współmałżonka, rodzica lub innej osoby oraz zagranicznych rent rodzinnych. Rentę przyznaną   
  na dzieci poniżej 15 lat należy uznać jako źródło utrzymania osoby wychowującej   
  te dzieci;
* **renta socjalna** – dla osób utrzymujących się ze świadczenia przyznanego na skutek całkowitej niezdolności do pracy z powodu inwalidztwa powstałego przed ukończeniem 18 roku życia lub przed ukończeniem 25 roku życia, w trakcie nauki w szkole ponadpodstawowej lub wyższej;
* **zasiłek dla bezrobotnych** – dla osób utrzymujących się ze świadczenia przyznawanego osobom zarejestrowanym w powiatowych urzędach pracy na określony czas trwania bezrobocia;
* **świadczenia i zasiłki przedemerytalne** – dla osób utrzymujących się z takich świadczeń. Występować mogą o nie osoby, które łącznie spełniają określone warunki: utrata pracy z przyczyn dotyczących zakładu pracy, wiek, staż dotychczasowej pracy, rejestracja w urzędzie pracy. Świadczenie przyznane jest na czas określony tzn. do osiągnięcia przez beneficjenta wieku emerytalnego;
* **świadczenia pomocy społecznej** – dla osób utrzymujących się ze świadczenia przyznawanego ze względu na niskie dochody, często zły stan zdrowia, zaawansowany wiek, jak również trudną sytuację rodzinną oraz osobom, które nie są w stanie funkcjonować bez dodatkowej pomocy.

**Dochody z własności:** – dla osób utrzymujących się z dochodów z tytułu oddanych w dzierżawę gruntów rolnych, dochodów z lokat kapitałowych (z obrotu akcjami, obligacjami, zysków kapitałowych itp.), z odsetek z oszczędności i udziału w zyskach przedsiębiorstw (dywidend).

**Pozostałe źródła dochodów osobno nie wymienione** – dla osób utrzymujących się z dochodów ze źródeł innych niż wymienione w pozycjach powyżej np. alimenty od osób prywatnych, stypendium naukowe, sportowe.

**Na utrzymaniu** – dla osób pozostających na utrzymaniu innych osób posiadających własne źródła dochodów.

**Gospodarstwa domowe**

Ze względu na liczbę osób oraz ich stosunek pokrewieństwa wyróżniono następujące **typy gospodarstw domowych:**

* rodzinne
* jednorodzinne,
* dwurodzinne,
* trzy i więcej rodzinne,
* nierodzinne
* jednoosobowe,
* wieloosobowe.

**Rodziny**

Wyróżnia się następujące **typy rodzin biologicznych:**

* małżeństwo bez dzieci
* małżeństwo z dziećmi
* partnerzy bez dzieci
* partnerzy z dziećmi
* matka z dziećmi
* ojciec z dziećmi.

**2. Jakość wyników w badaniu reprezentacyjnym**

Badanie aktywności ekonomicznej w ramach NSP 2011 odbyło się metodą reprezentacyjną, stąd też poniższe punkty będą odnosiły się zasadniczo do tej części spisu.

***UWAGA! Dane w tablicach mogą się nie sumować ze względu na uogólnienia wyników z próby.***

Ważnym obszarem prac przygotowawczych do NSP 2011 były działania związane z tworzeniem elektronicznego wykazu adresowo-mieszkaniowego na potrzeby spisu. Ten etap prac miał istotne znaczenie dla spisu prowadzonego nowoczesnymi metodami. Wykaz adresowo-mieszkaniowy był źródłem danych do zasilania *Master Rekordu (MR)*, do zarządzania na wszystkich etapach prowadzenia spisu, do ustalenia kompletności spisu, a także był podstawą do utworzenia operatu losowania próby do badania reprezentacyjnego.

Wykaz adresowo-mieszkaniowy przygotowany został w oparciu o rejestr TERYT (Krajowy Rejestr Urzędowy Podziału Terytorialnego Kraju) z wykorzystaniem danych pochodzących z innych źródeł, m.in. z państwowego zasobu geodezyjnego i kartograficznego w zakresie lokalizacji przestrzennej budynków. W związku z podłączeniem tych dwóch źródeł do identyfikatorów adresowych budynków z rejestru TERYT dołączone zostały także współrzędne geograficzne x,y (punkty adresowe). W zakresie informacji o osobach na potrzeby wykazu adresowo-mieszkaniowego wykorzystane były gminne zbiory meldunkowe, które zostały połączone z systemem NOBC, tj. systemem identyfikacji adresowej ulic, nieruchomości, budynków i mieszkań, stanowiącym część rejestru TERYT. Przygotowane w tym trybie zestawienie budynków, mieszkań i osób było weryfikowane przez urzędy gmin podczas aktualizacji przedspisowej. Na tym etapie została ustalona zbiorowość budynków, mieszkań i obiektów zbiorowego zakwaterowania podlegających spisaniu, poprawność ich adresów oraz przyporządkowanie osób do poszczególnych mieszkań i obiektów.

W trakcie obchodu przedspisowego rachmistrzowie dokonali weryfikacji wszystkich punktów adresowych, które znalazły się w obszarze przydzielonych obwodów, porównując je z pozycjami ujętymi w wykazie adresowo-mieszkaniowym do obchodu. Przedmiotem obchodu było zatwierdzenie lub modyfikacja punktów adresowych wyszczególnionych w wykazie, a także usunięcie nieistniejących punktów adresowych lub ewentualne dodanie nowych nieodnotowanych w wykazie, na których znalazły się budynki mieszkalne, budynki niemieszkalne z mieszkaniami lub budynki obiektów zbiorowego zakwaterowania, np. dom studencki. W odróżnieniu do poprzednich spisów ludności i mieszkań rachmistrzowie w trakcie obchodu przedspisowego nie odwiedzali mieszkań. Zatem niemożliwa była weryfikacja informacji dotyczących głównego lokatora i liczby osób podlegających spisaniu.

**Operat losowania mieszkań do badania reprezentacyjnego**

Etapy tworzenia operatu losowania mieszkań obejmowały:

1. Utworzenie zintegrowanego wykazu adresowo-mieszkaniowego w oparciu o Zestawienie adresowo-mieszkaniowe;
2. Weryfikację wykazu w trakcie obchodu przedspisowego;
3. Uzupełnienie wykazu dodatkowymi zmiennymi pochodzącymi z rejestrów i systemów informacyjnych i utworzenie operatu losowania do badania reprezentacyjnego.

Dodatkowo do wykazu adresowo-mieszkaniowego (w Operacyjnej Bazie Mikrodanych) zostały wprowadzone informacje pochodzące z innych rejestrów i systemów informacyjnych lub baz statystycznych jak: zbiory Zakładu Ubezpieczeń Społecznych (Centralny Rejestr Ubezpieczonych – CRU, Emerytów i Rencistów – EiR oraz Centralny Rejestr Płatników Składek – CRPS), Kasy Rolniczego Ubezpieczenia Społecznego (KRUS), Krajowej Ewidencji Podatników (KEP), rejestru PESEL, spisu ludności i mieszkań 2002, bazy budynkowej oraz uaktualnionego Wykazu bazowego do PSR 2010. Celem włączania do wykazu dodatkowych informacji była poprawa jakości danych zgromadzonych w tym wykazie oraz zapewnienie informacji na potrzeby warstwowania operatu losowania. W trakcie weryfikacji danych w wykazie zostały ustalone kryteria (zmienne oraz ich wartości), które zostały wykorzystane do tworzenia bardziej jednorodnych grup (warstwowania) jednostek badania.

Przy tworzeniu operatu losowania zostały pominięte:

1. obiekty zbiorowego zamieszkania niezawierające mieszkań,
2. obiekty zamknięte oraz mieszkania na terenie obwodów zamkniętych,
3. mieszkania w obiektach zbiorowego zakwaterowania stanowiących obwody zamknięte,
4. pomieszczenia prowizoryczne,
5. mieszkania niezamieszkane, zniszczone na skutek klęsk żywiołowych w szczególności powodzi.

W konsekwencji w operacie zostały ujęte mieszkania zamieszkane, w których osoby były zameldowane lub przebywały bez meldunku oraz te mieszkania niezamieszkane, które były w trakcie remontu, zmiany lokatora lub nowo wybudowane. Budynki, które występowały pod jednym adresem były brane do losowania, jeżeli ich liczba nie przekroczyła dwóch. Operat losowania mieszkań został zaktualizowany w trakcie trwania obchodu przedspisowego, w wyniku którego do badania dołączono budynki, które nie znalazły się wcześniej w wykazie.

## Zmienne uwzględnione w operacie losowania

Dla potrzeb badania reprezentacyjnego w wykazie dopisano do każdego mieszkania zestaw informacji, który został przeniesiony do operatu losowania, spośród których następnie ostatecznie wybierano kryteria warstwowania:

1. symbol terytorialny,
2. położenie budynku/mieszkania na terenie objętym klęską żywiołową,
3. liczba mieszkań w budynku wielomieszkaniowym,
4. położenie mieszkania w budynku wielomieszkaniowym,
5. rok zakończenia budowy budynku,
6. liczba osób w mieszkaniu (0, 1, 2, 3, 4, 5, 6, 7 i więcej),
7. liczba osób zameldowanych na pobyt stały,
8. występowanie cudzoziemca w mieszkaniu,
9. występowanie osoby pracującej (posiadanie ubezpieczenia), emeryta lub rencisty w mieszkaniu oraz osoby bezrobotnej,
10. występowanie osoby niepełnosprawnej,
11. występowanie użytkownika gospodarstwa rolnego oraz powierzchni użytków rolnych,
12. położenie mieszkania w gminie o udziale społeczności mniejszościowej (narodowej lub etnicznej) co najmniej 10%.

Wartości zmiennych, jako kryteria warstwowania zostały zróżnicowane dla obszarów miejskich i wiejskich. Wybór kryteriów warstwowania został ostatecznie ustalony w trakcie losowania próbnego.

**Błędy nielosowe w spisie ludności i mieszkań 2011**

Podstawowym źródłem danych dla potrzeb oceny występowania oraz skali błędów nielosowych są wyniki spisu kontrolnego do NSP 2011. Dane zebrane w spisie kontrolnym powinny pozwolić na dokonanie oceny przede wszystkim błędów pokrycia (in plus oraz in minus), popełnianych z powodu niespisania mieszkań i osób, które mogą wynikać z nieujęcia jednostek losowania w operacie losowania, opuszczenia osób, które powinny być spisane lub podwójnego spisywania osób, ewentualnego dopisywania fikcyjnych osób, a także do określenia błędów treści, wynikających z braków odpowiedzi lub błędów w odpowiedziach. Przyjęte założenia oraz dane pozyskane w spisie kontrolnym, powinny pozwolić na ocenę wpływu czynników zewnętrznych, tj. rachmistrzów, ankieterów czy samych respondentów, którzy dokonali samospisu – na jakość wyników spisu ludności i mieszkań 2011.

Także dane pozyskiwane z rejestrów i systemów informacyjnych mogły być źródłem błędów nielosowych z uwagi na brak integracji pomiędzy tymi bazami danych. Dane te wymagały szczególnej walidacji, która obejmowała procedury normalizacji, konwersji oraz synchronizacji podmiotowej, przedmiotowej i przestrzennej. Szczególnie ważne jest gruntowne rozpoznanie definicji pojęć stosowanych w różnych rejestrach i systemach informacyjnych, które są konsekwencją zapisów w regulacjach prawnych, na bazie których te systemy funkcjonują.

Efekty pogłębionej analizy występowania ewentualnych błędów nielosowych tak w spisie pełnym jak i w badaniu reprezentacyjnym będą – w miarę możliwości – prezentowane w kolejnych ogólnopolskich publikacjach tematycznych. Należy jednak pamiętać, że błędy te są bardzo trudne do wykrycia, zaś oszacowanie ich wpływu na wyniki spisu jest jeszcze trudniejsze, jeśli nie niemożliwe. Szczegółowe zasady metodologiczne spisu kontrolnego oraz jego wyniki zostaną przedstawione w publikacji poświęconej metodologii spisu ludności i mieszkań 2011.

**Błędy losowe**

Błędy losowe występują wyłącznie w badaniach reprezentacyjnych i są one zawsze powiązane ze schematem losowania próby. Błędy losowe są łatwe do oszacowania, jeśli tylko przestrzegane są zasady metody reprezentacyjnej, a w szczególności dobre przygotowanie operatu losowania oraz właściwy dobór jednostek losowania do próby. Błąd losowy w badaniu reprezentacyjnym do spisu 2011 został wyznaczony poprzez wskaźniki precyzji estymacji.

## 2.1. Schemat losowania próby oraz alokacja jednostek w powiatach[[5]](#footnote-5)

Podstawowym celem badania reprezentacyjnego realizowanego w ramach spisu 2011 było uzyskanie informacji o sytuacji społeczno-demograficznej na poziomie powiatów. Ustalono, że aby uzyskać reprezentatywne wyniki na poziomie ok. 400 jednostek wyodrębnionych w przekroju terytorialnym (w tym 380 powiatów) należy wylosować próbę mieszkań liczącą ok. 20% w skali kraju. W celu wylosowania próby o ustalonej liczebności zastosowany został schemat jednostopniowego losowania warstwowego. Jednostki losowania – mieszkania zostały przed losowaniem pogrupowane w warstwy w celu zwiększenia efektywności losowania. Zastosowano zróżnicowane podejście do warstwowania w zależności od typu powiatu i gminy.

**W miastach – powiatach oraz w dzielnicach i delegaturach pięciu największych miast**, a także w innych wyróżnionych większych miastach niebędących powiatami – w pierwszym etapie dokonany został podział mieszkań na dwie kategorie:

1. mieszkania w „blokowiskach”,
2. pozostałe mieszkania.

Kryterium podziału na powyższe kategorie był przypisany do każdego mieszkania wskaźnik, wyrażający liczbę mieszkań w budynku, w którym znajduje się dane mieszkanie. Do „blokowisk” zaliczone zostały mieszkania o wartości tego wskaźnika powyżej jego mediany. Następnie w każdej z tych kategorii mieszkania zostały powarstwowane ze względu na liczbę osób w mieszkaniu, po czym w ramach tak utworzonych kategorii nastąpił dalszy podział na 4 grupy ze względu na:

1. występowanie osoby pracującej w mieszkaniu,
2. występowanie emeryta lub rencisty w przypadku braku pracującego,
3. występowanie bezrobotnego w mieszkaniu w przypadku braku w/w osób,
4. mieszkania z innymi osobami.

Ponadto, w mieszkaniach poza „blokowiskami”, dodano kategorię: mieszkanie z użytkownikiem gospodarstwa rolnego.

Warstwowanie według liczby osób zamieszkałych w danym mieszkaniu było istotnym czynnikiem mającym wpływ na precyzję wyników, zniwelowało bowiem negatywne efekty zróżnicowanej liczby osób zamieszkałych w mieszkaniach. Efektywność tego etapu warstwowania zależała od korelacji liczby osób w mieszkaniu zapisanej w operacie losowania ze stanem rzeczywistym. Warstwowanie ze względu na pozostałe zmienne miało z kolei pozytywny wpływ m.in. na wyniki spisu związane z aktywnością ekonomiczną.

**W pozostałych powiatach w pierwszym etapie założono warstwowanie według gmin, przy czym gminy miejsko – wiejskie traktowane były jako dwie oddzielne gminy.** W mniejszych miastach pominięty został etap podziału na mieszkania w tzw. blokowiskach i pozostałe mieszkania, a warstwowanie odbywało się analogicznie jak w dużych miastach poza blokowiskami.

**W gminach wiejskich** w pierwszym etapie mieszkania podzielono na dwie kategorie:

1. mieszkania z użytkownikiem gospodarstwa rolnego,
2. pozostałe mieszkania.

W kategorii pierwszej wykonano warstwowanie według liczby osób zamieszkałych w mieszkaniu, a następnie, w miarę możliwości, dalsze warstwowanie według powierzchni gospodarstwa rolnego (dwie ewentualnie trzy grupy obszarowe w zależności od liczby mieszkań z użytkownikiem gospodarstwa w gminie). W przypadku drugiej kategorii mieszkań wykonano warstwowanie analogiczne jak w małych miastach.

Przedstawiona wyżej koncepcja warstwowania mogła prowadzić do utworzenia w niektórych powiatach bardzo rozdrobnionych warstw. W związku z tym program losujący przy tworzeniu warstw sprawdzał jednocześnie czy tworzone w powyższy sposób warstwy spełniały wymóg minimalnej liczby mieszkań. Warstwy zawierające zbyt mało mieszkań zostały łączone z sąsiednimi. Założono odwrotną niż przy tworzeniu, hierarchię kryteriów tzw. „sklejania” warstw.

Po utworzeniu warstw ustalone zostały liczebności prób do wylosowania w poszczególnych warstwach. Przyjęto proporcjonalną alokację próby pomiędzy warstwy. Oznacza to m.in., że ustalona w wyniku alokacji pierwiastkowej frakcja losowania w danym powiecie obowiązywała we wszystkich gminach tego powiatu i w warstwach utworzonych wewnątrz gmin.

## Alokacja próby w powiatach

Podziału założonej 20% próby mieszkań dla Polski pomiędzy powiaty dokonano przy wykorzystaniu metody alokacji pierwiastkowej. Metoda ta stanowi kompromis pomiędzy alokacją proporcjonalną a alokacją zapewniającą jednakową precyzję dla subpopulacji. Przy założeniu, że zastosowanoby proporcjonalne losowanie próby, w każdym powiecie próba stanowiłaby 20 % populacji. Ponieważ precyzja wyników, tj. wielkość błędu losowego zależy od liczebności próby, uzyskanoby w efekcie niedostateczną precyzję dla wielu małych powiatów. Z kolei, w metodzie alternatywnej uzyskujemy, w przybliżeniu, jednakową precyzję wyników dla wszystkich powiatów, ale za cenę istotnego „spłaszczenia” liczebności próby. W efekcie liczebność próby, a tym samym pracochłonność przy realizacji spisu byłaby mało zróżnicowana pomiędzy dużymi i małymi powiatami. Z tych powodów jako metodę rozdziału próby przyjęto alokację pierwiastkową, w której np. liczba mieszkań losowanych w konkretnym powiecie jest proporcjonalna do pierwiastka kwadratowego z populacyjnej liczby mieszkań i wyrażona wzorem:



gdzie:

n\* – założona liczebność próby dla Polski,

Np – liczba mieszkań w p-tym powiecie,

Według powyższego wzoru obliczone zostały liczebności prób dla poszczególnych powiatów.Potraktowanie Warszawy, tak jak każdego innego powiatu, mogłoby w konsekwencji spowodować niemożność uzyskania w miarę precyzyjnych wyników dla poszczególnych dzielnic. Jako ostateczną próbę dla miasta Warszawy przyjęto wartość średnią 87,5 tys. mieszkań. Wartość ta została odjęta od założonej liczebności próby dla Polski tj. od 2 631 tys., po czym dokonano alokacji pierwiastkowej dla 378 powiatów. Próby w najmniejszych powiatach tj. beskidzkim i sejneńskim wynoszą po ok. 3200 mieszkań (49% populacji), zaś największa próba (poza Warszawą) jest w m. Łodzi – 23, tys. mieszkań, czyli 6,9% liczby mieszkań w tym mieście na prawach powiatu.

W Warszawie, Łodzi, Krakowie, Wrocławiu i w Poznaniu podział próby pomiędzy poszczególne dzielnice i delegatury dokonany został również metodą pierwiastkową. Poza Warszawą, w pozostałych miastach, ze względu na małą liczbę delegatur (4 lub 5) nie zachodziła potrzeba zwiększania liczebności próby, tak jak w przypadku Warszawy. Ustalone dla poszczególnych delegatur w tych miastach liczebności prób zapewniają uzyskanie precyzji nie gorszej niż w najmniejszych powiatach.

Efektem tych działań było wylosowanie, spośród prawie 13,7 mln mieszkań, próby liczącej ponad 2744 tys. mieszkań. Utworzono prawie 70,5 tys. warstw, zaś wielkość próby w poszczególnych warstwach wahała się od niemal 6% do ponad 49%.

**2.2. Łączenie danych pozyskanych z różnych źródeł**

Różnorodność metod badania i technik pozyskiwania informacji dla potrzeb spisu ludności i mieszkań 2011 spowodowała konieczność łączenia danych pochodzących z rejestrów i systemów informacyjnych z wynikami badania reprezentacyjnego oraz z wynikami spisu otrzymanymi poprzez Internet. W badaniu reprezentacyjnym zebrano szeroki zakres danych od ponad 8 mln osób, mieszkających lub przebywających   
w ok. 2 744 tys. wylosowanych mieszkaniach. Zdecydowana większość danych w tym badaniu została pozyskana poprzez bezpośrednie wywiady rachmistrzów z respondentami. Odsetek osób, które wypełniły formularz długi w samospisie internetowym wynosił niecałe 2%, natomiast udział osób, które dokonały samospisu na formularzu krótkim przekraczał 10%. Wywiady telefoniczne stosowano na ogół w sytuacji niekompletnie wypełnionych formularzy – do uzyskiwania odpowiedzi na pominięte pytania lub wyjaśniania nieścisłości i niekonsekwencji w zebranych danych.

Ze względu na przyjęte rozwiązania w badaniu reprezentacyjnym i badaniu pełnym osoba mogła zostać spisana więcej niż jeden raz. Na etapie opracowania wyników spisu zaistniała zatem konieczność wyboru źródła danych oraz wyboru „najlepszej” ankiety osobowej, w której pozyskano najszerszy zakres danych. Przyjęto przy tym założenie, że rozpatrywane ankiety są kompletne i zamknięte, co oznacza, że dla każdej ankiety osobowej uzyskano wszystkie informacje wynikające ze „ścieżek przejścia” w formularzu spisowym.

Procedury wyboru ankiety zostały przeprowadzone po etapie weryfikacji i ustaleniu kompletności osób w oparciu o informacje zawarte w wykazie adresowo-mieszkaniowym zaktualizowanym po przeprowadzeniu spisu oraz po deduplikacji mieszkań. Proces ten w przypadku badania reprezentacyjnego – polegał na wyborze zestawu „najlepszych” ankiet osobowych dla danego mieszkania tzw. „paczki osobowej” („paczka osobowa” rozumiana, jako zestaw ankiet osobowych powiązanych z danym mieszkaniem – spisanych w danym mieszkaniu). Natomiast formularze krótkie, wykorzystywane przy samospisie były poddawane deduplikacji indywidualnie na podstawie zdublowanego lub zwielokrotnionego numeru ewidencyjnego Pesel – w ramach tego samego mieszkania, jak również pomiędzy mieszkaniami. Procedura deduplikacji tych ankiet była bardziej skomplikowana w przypadku braku numeru Pesel jako unikalnego identyfikatora osoby.

Szczególnej walidacji, która obejmowała procedury normalizacji, konwersji oraz synchronizacji podmiotowej, przedmiotowej i przestrzennej – zostały poddane informacje pozyskiwane z rejestrów i systemów informacyjnych. Dane pozyskiwane z rejestrów i systemów informacyjnych mogły być źródłem błędów nielosowych z uwagi na brak integracji pomiędzy tymi bazami danych oraz nierozpoznane różnice definicyjne. Specjalnego podejścia wymagały sytuacje, kiedy występowały błędy w numerze Pesel lub niezgodności pomiędzy numerem Pesel, a datą urodzenia osoby. Należy także pamiętać, że rejestry i systemy informacyjne, które były wykorzystywane dla potrzeb spisu 2011 – zostały utworzone na podstawie przepisów prawnych dla zupełnie innych celów niż potrzeby statystyki. Każdy podmiot ujęty w takim rejestrze czy systemie informacyjnym nie może być usunięty z takiej bazy danych, dopóki nie zaistnieją okoliczności określone przepisami prawnymi i nie zostanie dostarczony stosowny dokument, potwierdzający zmianę statusu prawnego takiej jednostki.

Dane dla potrzeb spisu 2011 były pobierane ze źródeł administracyjnych po ich weryfikacji informacjami pozyskanymi od respondentów w samospisie internetowym (na formularzu krótkim) oraz w badaniu reprezentacyjnym (na formularzu długim). Łącznie dla potrzeb spisu pełnego zostały zaktualizowane informacje pobierane z rejestrów i systemów informacyjnych dla ok. 32% osób.

**Utworzona baza danych spisowych wymagała zastosowania dwóch podejść przy opracowywaniu wyników spisu 2011.**

**Pierwsze podejście** zostało zastosowane w odniesieniu do ustalenia stanów ludności oraz ich charakterystyki według cech demograficznych (płeć, wiek), a także miejsca zamieszkania. Kompletność tych danych dla całej populacji w kraju zapewniły zaktualizowane dane w zbiorach rejestrowych i systemach informacyjnych. Oznacza to, że stany ludności i struktury demograficzne takie jak płeć i wiek – mogą być opracowywane dla dowolnie małych jednostek podziału administracyjnego kraju, przekroju terytorialnego (miejscowości), geograficznego (określonego wymiarami x, y) oraz statystycznego (obwody spisowe, rejony statystyczne i miejscowości statystyczne).

**Drugie podejście** zastosowano w odniesieniu do większości tematów badanych w spisie w tym z zakresu aktywności ekonomicznej, dla których nie było odpowiadających danych w rejestrach i systemach informacyjnych. W obszarze tych tematów wyniki spisu zostały wyprowadzone poprzez uogólnienie na całą populację danych zebranych w 20% próbie (w skali kraju) osób mieszkających lub przebywających w wylosowanych mieszkaniach. Bezprecedensowa wielkość próby losowej mieszkań zastosowanej w spisie 2011, dobieranej z ponad 70 tys. warstw obejmujących wszystkie gminy w kraju, zagwarantowała reprezentatywne wyniki spisu począwszy od poziomu powiatu. Wyniki spisu 2011 wyprowadzone na podstawie badania reprezentacyjnego powinny być zawsze prezentowane wraz z błędami precyzji. Pozyskane dla 20%ludności informacje zawierają obszerny pod względem przedmiotowym zakres danych, które po uogólnieniu na całą populację pozwalają na szeroką charakterystykę demograficzną, społeczną i ekonomiczną ludności, gospodarstw domowych oraz rodzin.

**2.3. Uogólnianie wyników badania reprezentacyjnego**

Na etapie opracowania wyników badania reprezentacyjnego wyodrębnione zostały dwa etapy uogólniania (ważenia) wyników tego badania:

* etap I – ważenie przy zastosowaniu skorygowanych wag „z frakcji”, będących odwrotnością frakcji losowania, dla jednostek badania typu mieszkania,
* etap II – ważenie przy zastosowaniu skalibrowanych wag dla jednostek badania będących osobami.

Pierwotne wagi zostały wyznaczone jako odwrotność frakcji losowania dla prawie 70,5 tys. warstw. Należy przypomnieć, że celem warstwowania było wyodrębnienie możliwie jednorodnych grup jednostek losowania. W ramach każdej z warstw – wagi były identyczne. Pierwotne wagi z frakcji musiały zostać skorygowane z uwagi na 13,7% braki wypełnionych ankiet mieszkaniowych w badaniu reprezentacyjnym.

**Wagi skorygowane, wyznaczone w I etapie są stosowane do uogólniania wyników spisu w zakresie mieszkań.** Natomiast dla budynków jednorodzinnych, w ramach których występują 1 lub 2 mieszkania, wyprowadzone zostały odrębne wagi. Także odrębne wagi wyprowadzono dla gospodarstw domowych i rodzin. Wagi te zostały naliczone wtórnie – na bazie skalibrowanych wag dla osób tworzących gospodarstwa domowe.

Z uwagi na konieczność zintegrowania wyników badania reprezentacyjnego ze spisem pełnym (w zakresie podstawowych zmiennych dotyczących: płci, wieku oraz miejsca zamieszkania – poziom powiatu z wyodrębnieniem części miejskiej i wiejskiej) zaistniała potrzeba wyprowadzenia skalibrowanych wag dla poszczególnych osób.

**Kalibracja jest metodą polegającą na odpowiedniej korekcie wag**, tak aby zostały zrekompensowane straty informacji związane z występującymi brakami odpowiedzi.   
W NSP 2011 kalibracja miała na celu dostosowanie struktur płci i wieku pozyskanych w badaniu reprezentacyjnym na poziomie powiatów (z uwzględnieniem obszarów miejskich i wiejskich), do struktur ludności według płci i wieku ustalonych w spisie pełnym, którego wyniki stanowiły populację referencyjną, tj. populację odniesienia.

Kalibracja wag została wykonana przy wykorzystaniu programu CALMAR (*Calibration on Margins)* przez przedstawicieli środowiska naukowego Uniwersytetu Ekonomicznego w Poznaniu.[[6]](#footnote-6) **Wagi skalibrowane służą do uogólniania wyników badania reprezentacyjnego w zakresie stanów i struktur ludności.** Wagi kalibracyjne pozwalają odtwarzać struktury spisowe ludności na poziomie powiatów według płci, grup wieku (ostatnia grupa 85 lat i więcej). Ponadto wagi te pozwalają odtwarzać struktury według płci i pojedynczych roczników wieku dla 5 największych miast: Warszawa, Kraków, Łódź, Poznań i Wrocław. Dla Warszawy jest to możliwe w ramach płci i roczników wieku 0,1,…,99, 100 lat i więcej, a dla pozostałych 4 miast – tylko do wieku 85 lat i więcej.

Dla województw oraz dla całego kraju możliwe jest odtwarzanie struktur – oczywiście z uwzględnieniem płci i klasy miejscowości (w tym wyodrębnienie części miejskiej i wiejskiej) – w ramach roczników wieku, przy czym ostatnią grupę należy utworzyć jako 85 lat i więcej ze względu na brak reprezentacji w starszych rocznikach wieku dla niektórych województw w badaniu reprezentacyjnym, tj. powyżej 85 lat. Dzięki skalibrowaniu wag wyniki spisu wygenerowane z badania pełnego oraz uogólnione z badania reprezentacyjnego są tożsame – dla struktur ludności z uwzględnieniem płci, wieku (z pewnym ograniczeniem do grup wieku) oraz klasy miejscowości.

**Algorytm wyznaczania wag dla mieszkań w badaniu reprezentacyjnym**

W celu uzyskania poprawnych uogólnień konieczne było skorygowanie wag początkowych, wynikających z zastosowanego schematu losowania. Korekta uwzględniła zjawisko braku odpowiedzi oraz informacje o przyczynach nie uzyskania danych od niektórych badanych jednostek. Ponieważ jednostką losowania było mieszkanie, procedura korygowania wykorzystywała informacje przypisane do zmiennej „przyczyny niewypełnienia formularza”, gdzie zakodowano różne przyczyny dla mieszkań wylosowanych do próby.

W ramach każdej warstwy zdefiniowanej na etapie losowania próby mieszkań wyznaczono korektę wagi wynikającej z ustalonej frakcji losowania w danej warstwie. Korekta polegała na wykorzystaniu wskaźnika korygującego, wyliczanego jako iloraz liczby mieszkań, które powinny być zbadane do liczby mieszkań dla których udało się zrealizować badanie (odwrotność tego ilorazu można interpretować jako wskaźnik kompletności). Dla wylosowanej próby 2 684,2 tys. mieszkań udało się zebrać wypełnione formularze od 2 317,7 tys. mieszkań. Przypadki zaliczone jako będące poza badaną populacją, tzn. stwierdzono brak mieszkania lub mieszkanie niezamieszkane (pustostan), w trakcie postępowania spadkowego, przeznaczone do remontu, jeszcze nie zasiedlone itp. – stanowiły 140,4 tys., czyli około 5,2% wylosowanej próby. Występowanie takich jednostek w wylosowanej próbie zostało potraktowane jako błędy operatu losowania.

Waga skorygowana przypisana zbadanym mieszkaniom z danej warstwy była iloczynem tak wyznaczonej korekty oraz wagi z losowania (z frakcji). W przypadku warstw dla których obliczona korekta była zbyt duża i przekraczała umowny próg (przyjęty na poziomie 3) stosowano procedurę „sklejania” sąsiednich warstw i powtarzano obliczenia korekty aż do spełnienia założonego ograniczenia. Taka procedura stanowiła zabezpieczenie przed nadmiernym zróżnicowaniem wag, które zwiększa wariancję estymatorów i jest niekorzystne dla precyzji estymacji.

Oddzielnie traktowano część próby mieszkań dopisanych już po etapie losowania, w trakcie obchodu przedspisowego. W ramach badanych dodatkowo 105,4 tys. mieszkań zrealizowano badanie w 45,3 tys. mieszkaniach, poza badaną populacją okazało się być   
43,1 tys. mieszkań. W tej części próby założono, że wagi z losowania są równe 1 i zastosowano opisany algorytm korekty na poziomie umownych warstw zdefiniowanych jako powiaty.

Skorygowane w opisany sposób wagi mogą być wykorzystywane do wyprowadzania wartości globalnych (np. ogólnej liczby mieszkań, mieszkań zamieszkanych, mieszkań substandardowych oraz używane do estymacji dla cech mieszkaniowych (np. mieszkań według ich wielkości).

**Imputacja brakujących danych**

W spisie 2011 stosowano także imputację brakujących danych w odpowiedziach respondentów, tzw. imputacje dedukcyjną – na podstawie zależności pomiędzy zmiennymi. W większości przypadków była to imputacja pozycyjna, przeprowadzana z wykorzystaniem podejścia deterministycznego. Imputacja pozycyjna była stosowana także w przypadku brakujących odpowiedzi w formularzu do badania reprezentacyjnego – na podstawie wartości zmiennych dostępnych w rejestrach lub systemach informacyjnych. Procedury takie były możliwe do wykonania wyłącznie pod warunkiem pełnej zgodności definicyjnej zmiennych występujących w obu źródłach danych. W nielicznych przypadkach – z wykorzystaniem podejścia stochastycznego – została wprowadzona imputacja całych badanych jednostek.

W celu przedstawienia w niniejszej publikacji ostatecznych wyników z obszaru aktywności ekonomicznej ludności dokonano imputacji na podstawie rejestrów w zakresie podstawowych informacji o statusie na rynku pracy (osoba pracująca, bezrobotna czy bierna zawodowo) oraz pracy w indywidualnych gospodarstwach rolnych.

**2.4. Precyzja estymacji wartości globalnych**

Podstawowymi parametrami obliczanymi na podstawie danych zebranych w badaniu reprezentacyjnym w ramach NSP 2011 są wartości globalne dla poszczególnych zmiennych. Większość badanych zmiennych charakteryzuje populację osób, mniejsza część dotyczy mieszkań i budynków. Poniżej opisano sposób estymacji oraz oceny jakości estymacji dla cech dotyczących ludności.[[7]](#footnote-7)

Estymator wartości globalnej ustalonej cechy dla osób na poziomie agregacji typu: kraj ogółem, województwa, powiaty – jest sumą ważoną postaci:

gdzie *h* oznacza warstwy wynikające ze schematu losowania próby mieszkań, definiowane za pomocą kombinacji zmiennych: kod gminy przed obchodem spisowym i numer warstwy oraz

*i* – oznacza kod mieszkania, czyli jednostkę losowania próby,

*j* – oznacza osobę zaliczoną do składu danego mieszkania,

– oznacza wagę przypisaną do konkretnej osoby z danego mieszkania; jest to wartość zapisana w zmiennej *waga\_na\_lud\_pierwsza\_osoba*, tj. po procesie kalibracji,

– oznacza wartości analizowanej cechy dla osób np. 1 – jeżeli osoba posiada wykształcenie wyższe, 0 – w przeciwnym przypadku; dla konkretnych tablic wynikowych konieczne może być przygotowanie specjalnego zestawu nowych cech w oparciu o istniejące już w danych wynikowych np. do analizowania ludności w grupach wiekowych należy zdefiniować odpowiednie zmienne zero-jedynkowe dla ustalonych przedziałów wieku; jeżeli są potrzebne dodatkowo głębsze poziomy grupowania należy zdefiniować odpowiednio dodatkowe zmienne (np. uwzględniając podziały na miasta i wieś lub płeć).

**Do oceny jakości estymatora wartości globalnej wyznacza się tzw. wskaźniki precyzji estymacji.**

Umożliwiają one użytkownikom tablic wynikowych właściwą ocenę danych. Przyjęto, że prezentowane wskaźniki precyzji będą zdefiniowane jako względne błędy standardowe (inaczej współczynniki zmienności – oznaczane jako CV i wyrażone w procentach) za pomocą wzoru:

gdzie *Va*r oznacza oszacowanie wariancji estymatora, obliczane według formuły:

Wielkość oznacza liczbę mieszkań w danej warstwie zrealizowanej próby, natomiast jest liczbą mieszkań z danej warstwy w populacji; wielkości można odtworzyć odpowiednio sumując skorygowane wagi wynikające z losowania próby do badania reprezentacyjnego.

**Wskaźniki precyzji estymacji wartości globalnych dla podstawowych zmiennych zostały przedstawione w tablicach precyzji.**

Wskaźnik *CV* można wykorzystywać do wyznaczania przedziału ufności, który z określonym prawdopodobieństwem (tzw. poziomem ufności, np. 95%) pokrywa prawdziwą wartość estymowanego parametru. Na przykład dla poziomu ufności 95% granice takiego przedziału mają postać: ±CV. Jeżeli oszacowana liczba osób bezrobotnych w wieku 15-24  lata wynosiła 489,2 tys., a wskaźnik precyzji dla tej cechy wynosił *CV*=0,32%, to wyznaczając granice przedziału ufności można stwierdzić, że z prawdopodobieństwem 0,95 przedział wartości pomiędzy 486,2 tys. a 492,3 tys. obejmuje prawdziwą liczbę takich osób.

Kolejny przykład. Wskaźnik precyzji dla tej samej zmiennej, tj. liczby osób bezrobotnych w wieku 15-24 lata w woj. podlaskim wynosił 1,65%. Jeżeli oszacowana liczba osób bezrobotnych w wieku 15-24 lata wynosiła 15546 osób, a wskaźnik precyzji dla tej cechy wynosił *CV*=1,65%, to wyznaczając granice przedziału ufności można stwierdzić,   
że z prawdopodobieństwem 0,95 przedział wartości pomiędzy 15043 a 16049 pokrywa prawdziwą liczbę takich osób.

Z podanych przykładów widać, że uogólnienia dla województw są obciążone większym błędem losowym. Wynika to z tego, że liczebności prób w województwach były wielokrotnie mniejsze od liczebności próby ogólnopolskiej.

Interpretując wyniki badania reprezentacyjnego zwykle przyjmuje się umowne progi precyzji, powyżej których jakość uogólnień nie jest zadowalająca. Praktycznie wartość takiego progu można ustalić na poziomie np. CV=10%. Wtedy oszacowania z błędem względnym powyżej 10% powinny być traktowane jako mało precyzyjne; wskazuje to na konieczność prezentacji wyników na wyższych stopniach agregacji.

1. Conference of European statisticians recommendations for the 2010 censuses of population and housing, prepared in cooperation with the Statistical Office of the European Communities (EUROSTAT), (ONZ, Genewa 2006);

   [http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2006/zip.1.e.pdf](http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2006/zip.1.e.pdf ) [↑](#footnote-ref-1)
2. Current international recommendations on labour statistics, 2000 edition (Geneva, 2000);

   <http://www.ilo.org/public/english/bureau/stat/download/res/ecacpop.pdf> [↑](#footnote-ref-2)
3. Por.: publikacja *Ludność. Stan i struktura demograficzno-społeczna. Wyniki NSP 2011* – rozdział I Metodologia spisu ludności i mieszkań 2011 – wybrane aspekty; <http://www.stat.gov.pl/gus/5840_14076_PLK_HTML.htm> [↑](#footnote-ref-3)
4. ## Por.: publikacja „Gospodarstwa domowe i rodziny. Charakterystyka demograficzna – NSP 2011”, Warszawa 2014

   [↑](#footnote-ref-4)
5. Materiał roboczy opracowany przez zespół matematyków pod kierunkiem Bronisława Lednickiego. [↑](#footnote-ref-5)
6. Materiał roboczy: „Raport z opisem wyników z zakresu możliwości wykorzystania kalibracji na potrzeby korygowania wag w złotym rekordzie” pod kierunkiem prof. dr hab. Jana Paradysza. [↑](#footnote-ref-6)
7. Materiał roboczy opracowany przez dra Roberta Wieczorkowskiego. [↑](#footnote-ref-7)