

GŁÓWNY URZĄD STATYSTYCZNY
Departament Badań Demograficznych

**Zasady opracowywania wyników
Narodowego Spisu Powszechnego Ludności i Mieszkań 2011
w zakresie mniejszości narodowych i etnicznych
oraz języka regionalnego**

Notatka na XXXI posiedzenie Komisji Wspólnej
Rządu i Mniejszości Narodowych i Etnicznych

Białystok, 22-23 lutego 2012r.

1. Metody i techniki pozyskiwania danych w spisie ludności i mieszkań 2011

Powszechny spis ludności i mieszkań 2011 został przeprowadzony przy zastosowaniu zróżnicowanej metodologii, tj. poprzez pozyskiwanie danych ze źródeł administracyjnych (rejestrów i systemów informacyjnych) oraz zbieranie danych bezpośrednio od ludności w ramach badania reprezentacyjnego oraz tzw. badania pełnego. Oprócz tego prowadzone były dwa pełne, aczkolwiek niewielkie badania, obejmujące osoby przebywające w obiektach zbiorowego zakwaterowania oraz osoby bezdomne.

Badania spisowe były realizowane każdą z trzech dostępnych technik badawczych, tj.: poprzez samospis internetowy (technika CAII), bezpośredni wywiad przeprowadzany przez rachmistrza spisowego (CAPI) oraz wywiad telefoniczny przeprowadzany przez teleankietera (CATI). W każdym przypadku były wykorzystywane formularze w postaci elektronicznej, jako aplikacja działająca bezpośrednio na stronie internetowej (on-line) lub aplikacja przystosowana do pobrania ze strony, zainstalowania na komputerze osobistym respondenta i odesłania danych (off-line), a także aplikacja zaprojektowana na przenośne urządzenia rejestrujące (typu handheld), używane przez rachmistrzów spisowych.

Każde badanie spisowe mogło być realizowane dowolną techniką, przy czym, ze względu na przyjętą koncepcję spisu i wynikającą z niej konieczność uzyskania możliwie największej kompletności danych w ramach badania reprezentacyjnego, technikę wywiadu bezpośredniego przeprowadzanego przez rachmistrzów spisowych, zastosowano głównie w badaniu reprezentacyjnym, chociaż i w tym przypadku uzyskano część formularzy wypełnionych w drodze samospisu internetowego, a także w ramach wywiadów telefonicznych.

Dane dla osób mieszkających w niewylosowanych mieszkaniach pochodziły ze źródeł administracyjnych, a ponadto kierowany był do nich formularz krótki. Osoby takie mogły uczestniczyć w spisie przede wszystkim poprzez samospis internetowy, wspierany w miarę możliwości lokalnych biur spisowych - bezpośrednimi wywiadami rachmistrzów oraz wywiadami przez telefon. Wywiady telefoniczne stosowano na ogół w sytuacji niekompletnie wypełnionych formularzy - do uzyskiwania odpowiedzi na pominięte pytania lub wyjaśniania nieścisłości i niekonsekwencji w zebranych danych.

W ramach spisu 2011 przeprowadzono także pełne badanie ludności w 86 gminach, wstępnie wytypowanych na podstawie wyników spisu 2002. Kryterium do wyróżnienia tych gmin był co najmniej 10% udział osób należących do mniejszości w liczbie mieszkańców gminy w 2002 roku. Dla osób zamieszkałych lub przebywających w tych gminach w wylosowanych mieszkaniach wypełniany był

formularz długi, zaś w pozostałych mieszkaniach - formularz krótki. Dane z tego badania będą miały duże znaczenie dla opracowania wyników spisu w obszarze narodowości i języka, zwłaszcza w zakresie ustalania tzw. gmin mniejszościowych.

W spisie 2011 wykorzystywano dwa rodzaje formularzy, dostępnych wyłącznie w formie elektronicznej. Formularz długi, o szerokim zakresie tematycznym z dużą liczbą pytań, był stosowany w badaniu reprezentacyjnym, natomiast formularz krótki znalazł zastosowanie w badaniu pełnym, przede wszystkim w celu aktualizacji danych pozyskiwanych do spisu z rejestrów. Pytania o przynależność narodowo-etniczną oraz o język używany w kontaktach domowych zostały włączone do obu formularzy, m.in. także z tego powodu, że w żadnym z systemów administracyjnych nie występowała informacja o narodowości, możliwa do wykorzystania w spisie ludności.

2. Opracowywanie wyników spisu

Dane dotyczące osób i mieszkań pozyskane w spisie 2011 na etapie opracowywania wyników tworzą dwa podstawowe zbiory informacji:

1. Pierwszy zbiór informacji, to zbiór z **badania pełnego**, obejmujący dane odnoszące się do wszystkich mieszkańców kraju. Zbiór ten zawiera zweryfikowane i ujednolicone informacje z różnych rejestrów i systemów informacyjnych, także pochodzące z formularza krótkiego oraz części danych z formularza długiego – z badania reprezentacyjnego. Zakres informacji w tym zbiorze jest stosunkowo wąski pod względem tematycznym, co jest konsekwencją dostępności danych w systemach administracyjnych;
2. Drugi zbiór danych stanowi **zbiór z badania reprezentacyjnego**. Zbiór ten zawiera szeroki pod względem przedmiotowym zakres danych uzyskanych na podstawie formularza długiego. Dane te z założenia dotyczą tylko części ludności kraju, tj. tych osób, które zamieszkiwały lub przebywały w mieszkaniach wylosowanych do badania reprezentacyjnego. Dane z badania reprezentacyjnego zostaną uogólnione na całą populację.

Zbiór z badania pełnego umożliwia w sposób bezpośredni prezentowanie i analizę wyników spisu, na odpowiednio szczegółowym poziomie podziału terytorialnego kraju, tj. gmin, ich części miejskich i wiejskich, a także wybranych miejscowości. Jego ograniczeniem jest natomiast stosunkowo wąski zakres przedmiotowy, który wyznacza w zasadzie ta część informacji, która była dostępna w rejestrach i systemach informacyjnych.

W badaniu reprezentacyjnym zebrano szeroki zakres danych od ponad 8 mln osób, mieszkających lub przebywających w ok. 2744 tys. wylosowanych mieszkaniach.

Zdecydowana większość danych w badaniu została pozyskana poprzez bezpośrednie wywiady rachmistrzów z respondentami. Odsetek osób, które wypełniły formularz długi w samospisie internetowym wynosił niecałe 2%, podczas gdy udział osób które dokonały samospisu na formularzu krótkim przekraczał 10%.

Dane pozyskane różnymi metodami w spisie ludności i mieszkań 2011 zostały ostatecznie zebrane w bazie informatycznej, zwanej operacyjną bazą mikrodanych. Szczególne miejsce w tej bazie zajmuje połączony zbiór danych jednostkowych, zwany złotym rekordem, który zawiera dane pochodzące wprost z formularzy czy z rejestrów, jak również część zmiennych pochodnych, wyprowadzanych wtórnie. Opracowanie wyników spisu na podstawie tak złożonych i rozproszonych źródeł oraz obszernego zakresu informacyjnego (w formularzu do spisu reprezentacyjnego znajdowało się ponad 100 pytań) wymaga dużego zaangażowania sił ekspertów, metodologów, matematyków i informatyków oraz odpowiedniego czasu potrzebnego do uzyskania rzetelnych i dobrych jakościowo wyników. Aktualnie dane zebrane w spisie są poddawane procedurom walidacji.

Wyniki badania reprezentacyjnego zostaną uogólnione na całą populację (poprzez odpowiednie ważenie) zgodnie z zasadami metody reprezentacyjnej. Przygotowany starannie operat losowania mieszkań z uwzględnieniem głębokiego warstwowania, bardzo duża próba losowa (ok. 20% mieszkań w skali kraju), jak również schemat doboru próby do badania oraz jej alokacja - zapewniły pozyskanie co najmniej dobrych jakościowo wyników (z dobrą precyzją) na poziomie powiatów (z rozróżnieniem obszarów miejskich i wiejskich oraz dzielnic i delegatur w dużych miastach). Zatem na poziomie ogólnopolskim, województw, większych miast, powiatów oraz części miejskich i wiejskich powiatów - dostępny będzie szeroki zakres przedmiotowy danych z badania reprezentacyjnego, z możliwością szczegółowych zestawień i analiz.

Jednocześnie istnieje duże zapotrzebowanie użytkowników danych statystycznych na wyniki spisu dla jednostek lokalnych, tj. gmin i miejscowości. Dlatego w ramach wyników spisu na poziomie gmin lub niższym będą w pierwszym rzędzie opracowane dane ze spisu pełnego. Ponadto w ramach statystyki małych obszarów, są podejmowane działania mające na celu dezagregację wyników badania reprezentacyjnego na gminy (w tym miejscowości miejskie) oraz wybrane miejscowości wiejskie.

Należy podkreślić, że mimo zastosowania w spisie 2011 kilku źródeł danych opracowany będzie jeden zestaw spójnych wyników, ponieważ rezultaty badania pełnego i reprezentacyjnego zostaną połączone poprzez zastosowanie odpowiednich rozwiązań z zakresu metody reprezentacyjnej oraz statystyki małych obszarów.

Należy zwrócić uwagę, że na podstawie wyników spisu ludności i mieszkań 2011 zostaną wyodrębnione trzy kategorie ludności:

- 1) **Ludność stała** (zameldowana na pobyt stały);
- 2) **Ludność faktycznie zamieszkała**, aktualnie nazywana ludnością faktyczną;
- 3) **Ludność rezydująca**; wyprowadzana zgodnie z wymogami Eurostatu.

Kategorie ludności stałej i faktycznej wyodrębniane w spisie są tożsame ze stosowanymi dotychczas w bilansach opracowywanych w ramach bieżących badań demograficznych, prowadzonych przez GUS. W skali kraju obie kategorie dają tę samą liczbę ludności. Różnica pomiędzy ludnością stałą i faktyczną wynika z przemieszczeń osób na pobyt czasowy powyżej trzech miesięcy - pomiędzy gminami oraz obszarami miejskimi i wiejskimi gmin w ramach województw, a także między województwami.

Kategoria ludności faktycznej jest szeroko wykorzystywana przez statystykę, jak i przez odbiorców zewnętrznych. Dane dotyczące ludności faktycznej służą do opisu procesów demograficznych oraz do oceny perspektyw demograficznego rozwoju, w tym w prognozowaniu demograficznym. Ludność faktyczna stanowi punkt odniesienia dla wielu wskaźników społecznych i niektórych ekonomicznych, wykorzystywanych w analizach społecznych i gospodarczych. Z kolei ludność stała jest podstawą do wyprowadzania innych kategorii ludności, a ponadto jest wykorzystywana do wyliczeń współczynników demograficznych.

Kategoria ludności rezydującej oparta na definicji miejsca przebywania (*usual residence*) została zdefiniowana w Rozporządzeniu PE i Rady nr 763/2008 w sprawie spisów ludności, a także w Rozporządzeniu PE i Rady nr 862/2007 w sprawie wspólnotowej statystyki migracji i ochrony międzynarodowej - jako obowiązująca w krajach członkowskich UE. Ludność rezydująca jest wyprowadzana z uwzględnieniem kryterium co najmniej 12 miesięcy czasowego pobytu lub czasowej nieobecności w danym miejscu (miejscowości) w kraju lub zagranicą.

Liczba ludności Polski wyprowadzona w spisie 2011 dla tej kategorii będzie znacząco niższa od ludności faktycznej z uwagi na duże rozmiary długookresowych zagranicznych migracji ludności. Ludność rezydująca będzie przekazywana do baz Eurostatu począwszy od wyników spisu dla 2011 roku.

3. Opracowywanie wyników spisu w zakresie narodowości i języka

3.1. Struktura narodowo-etniczna ludności

W spisie ludności 2011 zostały wprowadzone dwa pytania o przynależność narodowo-etniczną, a także pytanie o język używany w kontaktach domowych, które

zamieszczono w obydwu rodzajach formularza. Ponadto w formularzu długim znalazło się pytanie o język ojczysty (*p. załącznik*).

Struktura narodowo-etniczna ludności będzie ustalana na podstawie danych pozyskanych w badaniu reprezentacyjnym oraz w samospisie internetowym, a także z uwzględnieniem wyników badania pełnego przeprowadzonego w 86 gminach, w których przynależność do określonej grupy narodowej lub etnicznej zadeklarowało minimum 10% mieszkańców - na podstawie wyników spisu ludności z 2002 roku. W tych gminach przeprowadzone zostało badanie pełne, czyli wszyscy mieszkańcy zostali objęci spisem i mogli zadeklarować swoją tożsamość narodowo-etniczną. Z wykorzystaniem wszystkich trzech metod pozyskiwania informacji zostanie ustalona liczebność grup narodowo-etnicznych dla całego kraju oraz jednostek administracyjnych.

Zatem podobnie jak zdecydowana większość informacji pozyskanych w spisie - odpowiedzi na pytania etniczne i językowe będą opracowywane przede wszystkim, jako wyniki uzyskane metodą reprezentacyjną. W ramach tej metody zostanie zaprezentowana zarówno pełna struktura identyfikacji narodowo-etnicznej ludności Polski, jak i liczebności poszczególnych mniejszości narodowych, etnicznych oraz społeczności posługującej się językiem regionalnym.

Niezależnie od powyższego wszystkie odpowiedzi na pytania narodowo-etniczne uzyskane z każdego źródła mają cenny walor poznawczy i metodologiczny, także jako kryterium weryfikacji i podwyższania jakości odpowiadających im danych z badania reprezentacyjnego.

Dane o deklaracjach narodowo-etnicznych, podobnie jak wszystkie wyniki spisu będą publikowane zgodnie z najlepszą wiedzą przedmiotową i metodologiczną oraz doświadczeniem pracowników statystyki publicznej. Odpowiedzi na obydwie pytania o przynależność narodowo-etniczną, jak również dane dotyczące języka używanego w domu i języka ojczystego zostaną uwzględnione w prezentowanych wynikach spisu. W miarę postępowania analiz danych spisowych w ramach metody reprezentacyjnej, pozyskana zostanie odpowiednia wiedza na temat jakości i reprezentatywności (precyzji i wiarygodności) danych z badania reprezentacyjnego, w tym także danych zebranych na poziomie niższym niż powiaty oraz ocena możliwości uogólnienia czy oszacowania pewnych rodzajów wyników dla mniejszych jednostek terytorialnych.

3.2. Mniejszości narodowe i etniczne oraz kaszubska społeczność regionalna

Zaliczanie osób do poszczególnych mniejszości w rozumieniu ustawy *o mniejszościach narodowych i etnicznych oraz o języku regionalnym* (Dz. U. Nr 17, poz. 141, z późn. zm.), będzie odbywało się w oparciu o odpowiedzi na dwa pytania o przynależność narodowo-etniczną oraz - w przypadku społeczności posługującej się

językiem regionalnym - na pytanie o język używany w kontaktach domowych, w sposób zgodny ze stanowiskiem Komisji Wspólnej Rządu i Mniejszości Narodowych i Etnicznych, wyrażonym na konferencji w Lublinie 16 września 2010r.

We wspomnianym stanowisku znalazł się następujący tekst: „*W opinii Komisji Wspólnej w przypadku zadeklarowania przynależności do narodowości polskiej i równocześnie przynależności do mniejszości narodowej lub etnicznej, przy ustalaniu liczby osób należących do mniejszości narodowych i etnicznych powinna być wzięta pod uwagę deklaracja przynależności do mniejszości.*”

W przypadku zadeklarowania przynależności do dwóch różnych mniejszości narodowych lub etnicznych powinna zostać wzięta pod uwagę odpowiedź na pierwsze z pytań.”

Liczba obywateli polskich należących do poszczególnych mniejszości narodowych i etnicznych w rozumieniu ustawy o mniejszościach będzie ustalana na podstawie liczby odpowiednich deklaracji narodowościowych wyrażonych jako odpowiedź na pierwsze bądź drugie pytanie o przynależność narodowo-etniczną. W sytuacjach, gdy osoba zadeklaruje przynależność do dwóch społeczności mniejszościowych w rozumieniu ustawy, jako kryterium przypisania jej do jednej z nich posłuży rodzaj odpowiedzi udzielonej na pierwsze z pytań narodowościowych. Liczebność społeczności posługującej się językiem regionalnym (kaszubskim) będzie ustalana na podstawie odpowiedzi na pytania o język używany w kontaktach domowych.

Jednocześnie na podstawie badania przeprowadzonego w 86 gminach dobranych celowo, zostanie ustalona imienna lista gmin spełniających 20% kryterium ustawowe (na bazie wyników spisu z 2002 roku wyróżniono 51 gmin spełniających to kryterium). Na podstawie wyników spisu w obszarze narodowości i języka będzie możliwe włączenie do wspomnianej listy gmin mniejszościowych – także gminy (gmin) spoza grupy 86 zbadanych w sposób pełny.

Opracowali:

*Grzegorz Gudaszewski – główny specjalista
oraz Lucyna Nowak – zastępca dyrektora
w Departamencie Badań Demograficznych*

Akceptacja

Prezes Głównego Urzędu Statystycznego

Janusz Witkowski

Załącznik

Pytania spisowe w zakresie badania przynależności narodowo-etnicznej i języka, ujęte w formularzach spisowych do NSP 2011 miały następującą postać:

1) Jaka jest Pana(i) narodowość? Przez narodowość należy rozumieć przynależność narodową lub etniczną - nie należy jej mylić z obywatelstwem.

2) Czy odczuwa Pan(i) przynależność także do innego narodu lub wspólnoty etnicznej?

3) Jakim językiem(ami) zazwyczaj posługuje się Pan(i) w domu? (*język używany w codziennych kontaktach domowych z osobami najbliższymi – nie należy mylić ze znajomością języka obcego*)

4) Jaki jest Pana(i) język ojczysty? (*język, którego nauczył (a) się Pan (i) jako pierwszego we wczesnym dzieciństwie – np. od matki*) – pytanie występowało wyłącznie w badaniu reprezentacyjnym

Obydwa pytania o narodowość są równoważne, w każdym z nich chodzi o przynależność narodowo-etniczną. Poprzez udzielanie odpowiedzi na dwa pytania badane osoby mają możliwość przedstawienia swoich pełnych tożsamości, określanych jako narodowość złożona - bez konieczności rozstrzygania, czy jest to identyfikacja z grupą narodową czy etniczną. Respondenci mieli zapewnioną możliwość wyboru dobrowolnej deklaracji w zakresie przynależności narodowo-etnicznej.