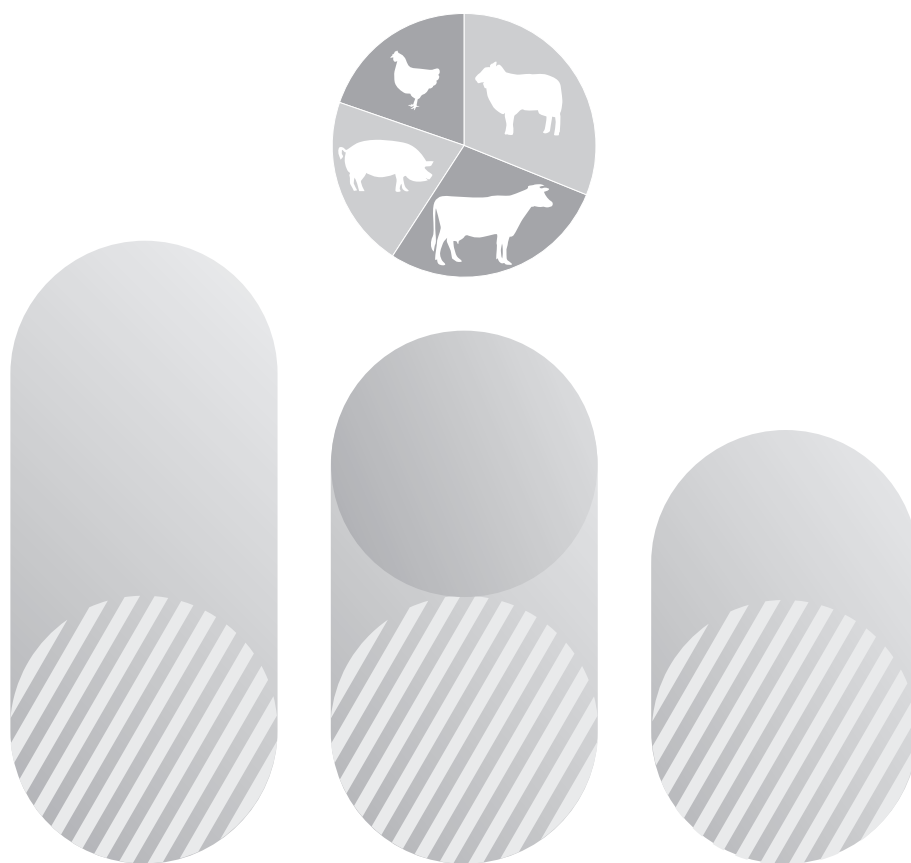


Zwierzęta gospodarskie w 2021 r.

Farm animals in 2021



Zwierzęta gospodarskie w 2021 r.

Farm animals in 2021

Opracowanie merytoryczne

Content-related works

Główny Urząd Statystyczny, Departament Rolnictwa
Statistics Poland, Agriculture Department

Zespół autorski

Editorial team

Małgorzata Kuliś, Jolanta Przypaśniak, Robert Wieczorkowski, Izabela Dach-Oleszek, Anna Tylkowska-Siek.

Kierujący

Supervisor

Artur Łączyński

Prace redakcyjne

Editorial work

Izabela Dach-Oleszek, Anna Tylkowska-Siek

Skład i opracowanie graficzne

Typesetting and graphics

Anna Tylkowska-Siek, Izabela Dach-Oleszek

Publikacja dostępna na stronie internetowej

Publications available on website

stat.gov.pl

Przy publikowaniu danych GUS prosimy o podanie źródła

When publishing Statistics Poland data — please indicate the source

Przedmowa

Publikacja zawiera podstawowe dane z przeprowadzonych w 2021 r. badań pogłowia bydła, świń, owiec, drobiu i produkcji zwierzęcej związanej z tymi gatunkami zwierząt, sprawozdawczości miesięcznej z ubojni zwierząt i wylęgarni drobiu.

Dla zilustrowania przemian i tendencji w produkcji zwierzęcej, dane z pogłowia poszczególnych gatunków zwierząt podano w porównaniu z analogicznymi danymi roku poprzedniego.

Publikacja składa się z uwag metodycznych oraz działów zawierających część tabelaryczną.

W uwagach metodycznych, oprócz wyjaśnienia podstawowych kwestii terminologicznych i zakresowych podano informacje o badaniach reprezentacyjnych pogłowia bydła, świń, owiec, drobiu i produkcji zwierzęcej.

Aneks tabelaryczny zawiera informacje o pogłowie poszczególnych gatunków zwierząt gospodarskich, podstawowe dane o ubojach zwierząt w rzeźniach przemysłowych i lokalnych, a także o wylęgach drobiu według kierunku wykorzystania piskląt.

Wyniki badania pogłowia zwierząt gospodarskich przedstawiono w przekroju krajowym, regionalnym i wojewódzkim dla rolnictwa ogółem.

Publikacja została opracowana w Wydziale Produkcji Roślinnej i Zwierzęcej.

Dyrektor Departamentu Rolnictwa
Artur Łączyński

Preface

The publication contains basic data of the surveys of cattle, pigs, sheep, poultry and animal output related to these species of animals, monthly reporting from slaughterhouses and hatcheries.

In order to illustrate changes and tendencies in animal production, the data on particular animal species are presented in comparison with the analogous data from previous year. The publication consists of methodological notes and sections containing the tabular part.

Beside the explanation of basic terminology and scope-related issues, the methodological notes contain information on sample surveys of cattle, pigs, sheep, poultry and animal production.

Statistical tables includes information on the number of individual livestock species, basic data on animal slaughters carried out in industrial and local slaughterhouses, as well as on poultry hatching according to the directions of nestlings' use.

Results of livestock survey are presented in cross national, regional, voivodship in total for agriculture.

The publication was prepared in the Crop and Animal Production Section.

Director of the Agriculture Department
Artur Łączyński

Uwagi metodologiczne

1. Źródła danych

Dane zawarte w niniejszej publikacji opracowano na podstawie:

- uogólnionych wyników reprezentacyjnych badań pogłowia¹ bydła, owiec, drobiu i świń oraz produkcji zwierzęcej w gospodarstwach indywidualnych,
- sprawozdań statystycznych z zakresu pogłowia zwierząt gospodarskich oraz produkcji zwierzęcej w gospodarstwach państwowych, spółdzielczych i spółkach z udziałem mienia sektora publicznego i prywatnego,
- sprawozdań statystycznych z rzeźni i ubojni zwierząt gospodarskich,
- sprawozdań statystycznych z wylęgarni drobiu,
- informacji rzeczoznawców wojewódzkich o pogłowie drobiu,
- szacunków własnych.

Badania pogłowia bydła, owiec i drobiu przeprowadzono na próbie gospodarstw indywidualnych utrzymujących wymienione gatunki zwierząt; próba ta liczyła 30 tys. gospodarstw.

Badania pogłowia świń przeprowadzono na próbie gospodarstw indywidualnych utrzymujących świnię; próba ta liczyła 30 tys. gospodarstw.

Wyniki badań pogłowia zwierząt gospodarskich i produkcji zwierzęcej opracowane zostały w układzie wojewódzkim oraz regionów NUTS 2, według siedziby użytkownika gospodarstwa, tj. dla gospodarstw indywidualnych – według miejsca siedziby (zamieszkania) użytkownika, a dla gospodarstw państwowych, spółdzielczych i spółek – według miejsca siedziby przedsiębiorstwa (gospodarstwa).

2. Ważniejsze definicje, pojęcia i zasady spisywania

Gospodarstwo rolne – jednostka wyodrębniona pod względem technicznym i ekonomicznym, posiadająca odrębne kierownictwo (użytkownik lub zarządzający) i prowadząca działalność rolniczą.

Do **działalności rolniczej** zaliczamy działalność związaną z uprawą roślin, która obejmuje: wszystkie uprawy rolne (w tym również uprawę grzybów), warzywnictwo i ogrodnictwo, szkółkarstwo, hodowlę i nasiennictwo roślin rolniczych i ogrodniczych oraz chów i hodowlę zwierząt gospodarskich, tj. bydła, owiec, kóz, koni, świń, drobiu, królików, pozostałych zwierząt futerkowych, dzikich zwierząt utrzymywanych w gospodarstwie dla produkcji mięsa (np. dziki, sarny, daniele) i pszczoł, a także działalność polegającą na utrzymaniu gruntów rolnych, już niewykorzystywanych do celów produkcyjnych, według zasad dobrej kultury rolnej przy zachowaniu wymogów ochrony środowiska (zgodnie z normami).

Gospodarstwo rolne osoby fizycznej (gospodarstwo indywidualne) – to gospodarstwo użytkowane przez osobę fizyczną o powierzchni 1 ha i więcej użytków rolnych (UR) lub gospodarstwo poniżej 1 ha UR (w tym nieposiadające użytków rolnych), jeżeli spełnia co najmniej jeden z poniższych progów:

- 0,5 ha dla: plantacji drzew owocowych,
- 0,5 plantacji krzewów owocowych, warzyw gruntowych, truskawek gruntowych, chmielu,
- 0,3 ha szkółek sadowniczych i ozdobnych,
- 0,1 ha dla: warzyw pod osłonami,
- 0,1 ha dla: truskawek pod osłonami, kwiatów i roślin ozdobnych pod osłonami, tytoniu,
- 25 m² grzybów jadalnych,
- 10 szt. bydła ogółem,
- 5 szt. krów,
- 50 szt. świń ogółem,

¹Stale badania pogłowia świń, bydła, owiec i drobiu prowadzone są 2-krotnie w roku, tj. w czerwcu i w grudniu.

- 10 szt. loch,
- 20 szt. dla: owiec ogótem, kóz ogótem,
- 100 szt. drobiu ogótem,
- 5 szt. koni ogótem,
- 50 szt. samic królików,
- 5 szt. samic pozostałych zwierząt futerkowych,
- 10 szt. dzikich zwierząt (np. dziki, sarny, daniela) utrzymywanych w gospodarstwie dla produkcji mięsa,
- 20 pni pszczelich
- lub niezależnie od ww. progów jest gospodarstwem ekologicznym.

Gospodarstwo rolne osoby prawnej lub jednostki organizacyjnej niemającej osobowości prawnej to gospodarstwo rolne prowadzone przez osobę prawną lub jednostkę organizacyjną niemającą osobowości prawnej, którego podstawowa działalność jest zaliczana według Polskiej Klasyfikacji Działalności do sekcji A, dział 01, grupy:

- 01.1– uprawy rolne inne niż wieloletnie,
- 01.2– uprawy roślin wieloletnich,
- 01.3 – rozmnażanie roślin,
- 01.4 – chów i hodowla zwierząt,
- 01.5 – uprawy rolne połączone z chowem i hodowla zwierząt (działalność mieszana),

01.6, klasa 01.61 – działalność usługowa wspomagająca produkcję roślinną (utrzymywanie gruntów w dobrej kulturze rolnej przy zachowaniu wymogów ochrony środowiska), a także niezależnie od zaklasyfikowania działalności podstawowej, gdy w gruntach użytkowanych przez jednostkę powierzchnia użytków rolnych wynosi 1 ha i więcej lub prowadzony jest chów/hodowla zwierząt gospodarskich.

Za **użytkownika gospodarstwa rolnego** uważa się osobę fizyczną, osobę prawną oraz jednostkę organizacyjną niemającą osobowości prawnej, faktycznie użytkującą gospodarstwo rolne, niezależnie od tego, czy jest właścicielem, dzierżawcą tego gospodarstwa czy też użytkuje je z innego tytułu i niezależnie od tego, czy grunty wchodzące w skład gospodarstwa rolnego są położone na terenie jednej czy kilku gmin.

Zwierzęta gospodarskie

Badaniu podlegały zwierzęta gospodarskie znajdujące się w czasie badania w gospodarstwie rolnym oraz zwierzęta wystane na redyki, wypasy i do bacówek. Spisywano wszystkie zwierzęta, tj. stanowiące własność użytkownika gospodarstwa lub członków jego gospodarstwa domowego, jak również zwierzęta przetrzymywane czasowo lub stale w gospodarstwie, tj. przyjęte na wychów, opas itp. niezależnie od tego, czy przyjęte je od gospodarstw indywidualnych, czy od jednostek państwowych, spółdzielczych, spółek.

Za **krowy mleczne** uważa się krowy, które ze względu na rasę lub odmianę lub szczególne właściwości utrzymywane są w gospodarstwie wyłącznie lub głównie do produkcji mleka przeznaczonego do konsumpcji lub przetworzenia na produkty mleczne. Zalicza się tu również krowy mleczne wybrakowane już z chowu, które pozostają jeszcze w gospodarstwie na tzw. dopasie, po czym skierowane zostaną do uboju.

Za **krowy pozostałe** uważa się krowy, które ze względu na rasę lub odmianę (krowy ras mięsnych i urodzone z krzyżówek z rasami mięsnymi) lub szczególne właściwości utrzymywane są w gospodarstwie wyłącznie lub głównie do produkcji cieląt rzeźnych, a których mleko wykorzystywane jest do odchowu cieląt lub przeznaczone na paszę dla innych zwierząt. Zalicza się tu również krowy "mamki" wybrakowane już z chowu, które pozostają jeszcze w gospodarstwie na tzw. dopasie, po czym skierowane zostaną do uboju.

Dla gospodarstw o dużej skali chowu drobiu (np. ferma wielkotowarowa produkująca brojlery lub jaja konsumpcyjne), w których w dniu badania nie było na stanie drobiu w związku z trwającą właśnie przerwą technologiczną w produkcji, a przerwa ta nie przekraczała 8 tygodni, przyjmowano stany drobiu z okresu przed opróżnieniem pomieszczeń (kurników).

Informacje o pogłowie świń, bydła, owiec i drobiu zawarte w niniejszej publikacji dotyczą stanów w czerwcu i grudniu 2021 roku.

Informacje liczbowe w ujęciu odsetkowym prezentowane z jednym znakiem po przecinku z uwagi na elektroniczną technikę zaokrągleń mogą nie sumować się (również na 100%). Liczby te są merytorycznie poprawne.

Schemat losowania prób

Badanie pogłowia bydła, owiec i drobiu

1. Uwagi wstępne

Celem przeprowadzanych przez GUS dwa razy w roku (tj. w czerwcu i w grudniu) badań jest uzyskanie szczegółowych informacji o pogłowie bydła i drobiu według regionów, województw i dla Polski, zaś pogłowia owiec tylko dla Polski. Badaną populację stanowią indywidualne gospodarstwa rolne, które według danych dostępnych w operacie posiadały bydło lub drób lub owce. Badana populacja liczyła w 2021 roku 602 tys. gospodarstw. Założono, że próba do badania liczyć będzie około 30 tys. gospodarstw. Taka wielkość próby była konsekwencją uwzględnienia kilku czynników m.in.: kosztów realizacji badania, analizy danych z lat poprzednich oraz konieczności zapewnienia odpowiedniej jakości wyników na poziomie kraju i regionów.

2. Operat losowania

Przy tworzeniu operatu losowania wykorzystano Operat do Badań Rolniczych. Rejestr ten stworzony został w GUS na bazie indywidualnych wyników Powszechnego Spisu Rolnego 2020. Informacje o gospodarstwach rolnych zapisane w rejestrze aktualizowane są danymi ze źródeł administracyjnych oraz wynikami kolejnych rolniczych badań reprezentacyjnych. Jednostką losowania było indywidualne gospodarstwo rolne. Dla każdego gospodarstwa zapisane zostały następujące informacje:

- symbol województwa oraz regionu NUTS 2,
- nr gospodarstwa,
- powierzchnia ogólna gospodarstwa,
- powierzchnia użytków rolnych w gospodarstwie,
- liczba sztuk bydła,
- liczba sztuk drobiu,
- liczba owiec.

3. Schemat losowania

Populacja gospodarstw została przed losowaniem podzielona na trzy części. Do części pierwszej zaliczono gospodarstwa spełniające przynajmniej jeden z warunków tj. posiadało przynajmniej 1 sztukę bydła lub powyżej 50 sztuk drobiu oraz nie posiadające owiec. Ta część populacji liczyła 395 108 gospodarstw. Do części drugiej zaliczono gospodarstwa nie posiadające bydła i owiec, zaś drobiu co najwyżej 50 sztuk. Część druga liczyła 194 845 gospodarstw. Do części trzeciej zaliczono gospodarstwa posiadające owce. Liczyła ona 11 661 gospodarstw.

W celu wylosowania próby zastosowany został schemat losowania warstwowego - optymalnego. W części pierwszej populacji jako kryteria tworzenia warstw i alokacji próby pomiędzy warstwy przyjęto liczbę sztuk bydła oraz liczbę sztuk drobiu. W części drugiej warstwy tworzone były w oparciu o powierzchnię użytków rolnych, zaś w trzeciej ze względu na liczbę sztuk owiec. W każdym regionie utworzono 8 warstw, z czego 5 w odniesieniu do gospodarstw części pierwszej, 3 w części drugiej. W części trzeciej utworzono 6 warstw

ogólnopolskich, tzn. obejmowały one gospodarstwa ze wszystkich regionów.

Pośród gospodarstw **części pierwszej** postanowiono wylosować próbę liczącą około 24,7 tys. gospodarstw.

Przystępując do losowania próby z tej kategorii gospodarstw przyjęto następujące założenia:

- (1) liczebność próby n ustalona jest dla populacji gospodarstw w Polsce, a nie dla poszczególnych regionów, przy czym n liczy ok. 24 700 gospodarstw,
- (2) w poszczególnych regionach próba losowana jest według schematu losowania warstwowego - optymalnego przy wykorzystaniu metody Neymana,
- (3) w każdym regionie dokonywany jest jednocześnie podział populacji na 5 warstw ($h = 1, 2, \dots, 5$), oraz dokonuje się alokacji próby pomiędzy warstwy,
- (4) w każdym regionie do warstwy nr 5 (tj. $h = 5$) wydzielane są jednostki losowania o wartościach zmiennych przyjętych za podstawę warstwowania powyżej określonego progu. Utworzona w ten sposób tzw. górna warstwa zawiera jednostki, które nie są losowane, lecz wszystkie zaliczane są do próby,
- (5) przyjęto, że oczekiwana precyzja wyników badania, mierzona współczynnikiem zmienności pogłowia bydła albo drobiu będzie jednakowa dla każdego regionu i równa 1,0% (konkretny parametr liczbowy był ustalony empirycznie tak, aby spełnić narzucone ograniczenia na liczebność próby).

Powyższy problem rozwiązany został przy wykorzystaniu metod optymalizacji numerycznej¹. Populacja podzielona została na warstwy, których granice (górne) wyrażone w liczbie sztuk bydła oraz liczb sztuk drobiu przedstawione zostały w poniższej tablicy.

Tabl.1. Granice warstw według regionów w badaniu pogłowia bydła, owiec i drobiu w 2021 roku.

Kod regionu	B – bydło D - drób	b_1	b_2	b_3	b_4
021	B	3	15	32	75
	D	729	746	752	5985
041	B	12	33	71	278
	D	815	897	903	14990
061	B	7	23	64	831
	D	711	1051	1597	16067
081	B	6	23	49	130
	D	1727	2032	2080	18050
101	B	6	22	48	154
	D	807	847	863	10400
121	B	3	11	37	797
	D	261	473	498	12512
141	B	2	4	14	37
	D	78	172	184	248
142	B	8	29	70	520
	D	1322	1385	1494	20927
161	B	8	29	62	132
	D	810	811	872	15871
181	B	0	3	18	250
	D	143	263	293	74998

¹ Opis rozwiązania tego zadania przedstawiony został w artykule: Lednicki B., Wieczorkowski R. (2003). Optymalna alokacja próby pomiędzy subpopulacje i warstwy. *Wiadomości Statystyczne*, nr 10, 18-26.

201	B	21	57	174	1584
	D	754	755	771	28950
221	B	9	30	80	1004
	D	1091	2197	4007	19039
241	B	4	20	52	1585
	D	1004	1007	6705	29990
261	B	3	8	21	55
	D	221	289	298	299
281	B	11	32	65	184
	D	617	620	639	6970
301	B	13	37	81	314
	D	1256	1393	1451	18454
321	B	7	27	64	1303
	D	2250	5243	16099	50039

Przedstawione w tabl.1. granice warstwy 4 tj. **b₄** stanowią jednocześnie **próg** powyżej którego gospodarstwa zaliczane są do warstwy 5 czyli nie podlegają losowaniu, lecz wszystkie zaliczane są do próby. Dla pozostałych warstw tj. $h = 1, 2, \dots, 4$ ustalone zostały, zgodnie z metodą optymalnej alokacji Neymana, wartości n_{wh} tj. liczebności prób losowanych z h -tej warstwy w w -tym regionie.

Do próby wylosowano następnie, zgodnie z przyjętą alokacją, 24 748 gospodarstw, w tym 7 420 gospodarstw z warstwy nr 5.

W **części drugiej**, w każdym regionie, utworzono przed losowaniem po 3 warstwy ($h = 6, 7, 8$). Warstwy te tworzone ze względu na powierzchnię użytków rolnych tj.:

$h = 6$: gospodarstwa poniżej 15 ha, $h = 7$: od 15 ha do 49.99 ha, $h = 8$: gospodarstwa o powierzchni 50 ha i więcej. Jako kryterium alokacji próby pomiędzy regiony przyjęto jednakową precyzję liczby sztuk drobiu dla każdego z regionów, zaś wewnątrz regionów próba alokowana była przy wykorzystaniu metody optymalnej Neymana. Z tej części populacji wylosowano 3 121 gospodarstw.

W **części trzeciej**, w której utworzono 6 warstw ogólnopolskich ($h = 9, 10, \dots, 14$), wylosowano do próby 2 131 gospodarstw. Z warstwy 14 do próby zaliczono wszystkie gospodarstwa (tzn. 203). Były to bowiem gospodarstwa posiadające owce, a jednocześnie posiadające 50 i więcej sztuk bydła lub co najmniej 400 sztuk drobiu. Granice pozostałych warstw oraz alokację założonej liczebności próby pomiędzy warstwy wyznaczono przy wykorzystaniu wspomnianej wyżej metody optymalizacji numerycznej. Wydzielona została warstwa górna ($h = 13$), z której nie losowano gospodarstw. Do warstwy tej zaliczono gospodarstwa, nie zaliczone wcześniej do warstwy 14 i posiadające powyżej 45 sztuk owiec. Górne granice pozostałych warstw były następujące: $b_9 = 3$, $b_{10} = 8$, $b_{11} = 20$, $b_{12} = 45$. Celem wydzielenia tej kategorii gospodarstw oraz optymalizacji podziału na warstwy było precyzyjne oszacowanie danych dotyczących pogłowia owiec w skali kraju, bez przekrojów regionalnych.

Ostatecznie cała próba do badania bydła, drobiu i owiec liczyła 30 000 gospodarstw.

2. Uogólnianie wyników i metoda oceny precyzji

Podstawowym parametrem szacowanym w badaniu pogłowia bydła, owiec i drobiu jest suma wartości zmiennej X np. pogłowie bydła ogółem.

Parametr ten dla w -tego regionu jest postaci:

$$(1) \hat{x}_w = \sum_h \sum_i W 1_{whi} * x_{whi}, \quad (i = 1, 2, \dots, n_{wh}; h = 1, 2, \dots, 14)$$

gdzie:

x_{whi} – wartość zmiennej X w i -tym gospodarstwie (jednostce losowania) wylosowanym z h -tej warstwy w w -tym regionie,

$W1_{whi}$ – waga przypisana i -temu gospodarstwu wylosowanemu z h -tej warstwy w w -tym regionie, przy czym waga ta obliczana jest ze wzoru:

$$(2) W1_{whi} = \frac{N_{wh}}{n_{wh}},$$

N_{wh} – liczba jednostek losowania w operacji w h -tej warstwie w -tego regionu,

n_{wh} – liczba jednostek losowania wylosowanych do próby z h -tej warstwy w -tego regionu,

Waga $W1_{whi}$ może być stosowana do estymacji wyników badania tylko wtedy, gdy badanie jest kompletne. Waga ta musi być korygowana, jeżeli część gospodarstw wylosowanych do próby odmawia udziału w badaniu. W tym celu wylosowaną próbę na podstawie informacji o realizacji badania dzielimy na 4 grupy:

- (1) gospodarstwa zbadane,
- (2) gospodarstwa, które odmówiły udziału w badaniu,
- (3) gospodarstwa zlikwidowane,
- (4) gospodarstwa, z którymi nie nawiązano kontaktu podczas realizacji badania.

Dla każdej warstwy, oddzielnie w każdym regionie, ustalamy liczebności powyższych grup tj.: $n1_{wh}$, $n2_{wh}$, $n3_{wh}$ i $n4_{wh}$, po czym ustalamy frakcję gospodarstw zbadanych i gospodarstw niezbadanych wśród gospodarstw o ustalonym statusie czyli:

$$(3) c_{wh} = \frac{n1_{wh} + n2_{wh}}{n_{wh} - n4_{wh}}.$$

Następnie obliczamy, dla wylosowanej próby, liczbę gospodarstw aktywnych n_{awh} w h -tej warstwie w -tego regionu:

$$(4) n_{awh} = n1_{wh} + n2_{wh} + c_{wh} * n4_{wh}$$

Na tej podstawie, dla danej warstwy, obliczany jest mnożnik korygujący R_{wh} :

$$(5) R_{wh} = \frac{n_{awh}}{n1_{wh}}$$

Mnożnik ten służy do korekty wagi $W1_{whi}$ w celu uzyskania wagi końcowej W_{whi} :

$$(6) W_{whi} = R_{wh} * W1_{whi}$$

Ocena sumy zmiennej X dla Polski jest sumą wartości uzyskanych dla regionów tj.:

$$(7) \hat{x} = \sum_w \hat{x}_w, \quad (w = 1, 2, \dots, 17)$$

Wagi pierwotne wynikające z losowania próby, korygowane są nie tylko ze względu na niepełną kompletność badania, ale również ze względu na występowanie tzw. gospodarstw nietypowych (*ang. outlier*). Dotyczy to gospodarstw, którym przypisano dużą wagę (losowanych z dużą frakcją), a jednocześnie ze stosunkowo dużymi wartościami dla niektórych badanych zmiennych. Korekta wagi ma zapobiegać w tym wypadku znacznemu przeszacowaniu wartości badanej zmiennej.

W ostatnim kroku wagi były kalibrowane uwzględniając dane z Agencji Restrukturyzacji i Modernizacji Rolnictwa (ARiMR) o liczbie gospodarstw, liczbie sztuk bydła oraz liczbie sztuk drobiu na poziomie województw.

Dla wybranych ważniejszych ocen parametrów oszacowane zostały, jako miary precyzji, ich współczynniki zmienności (względne błędy standardowe). W przypadku estymatora wyrażonego wzorem (1), tj. dla w -tego regionu oszacowanie jego współczynnika zmienności przedstawia się następująco:

$$(8) v(\hat{x}_w) = \frac{\sqrt{d^2(\hat{x}_w)}}{\hat{x}_w},$$

gdzie $d^2(x_w)$ jest oszacowaniem tzw. wariancji estymatora \hat{x}_w i oblicza się według wzoru (formuła wynika z klasycznej teorii metody reprezentacyjnej dla schematu losowania warstwowego):

$$(9) d^2(\hat{x}_w) = \sum_h \left(1 - \frac{n_{1wh}}{\sum_i W_{whi}}\right) \frac{n_{1wh}}{n_{1wh}-1} \sum_i \left(W_{whi} x_{whi} - \frac{\sum_j W_{whj} x_{whj}}{n_{1wh}}\right)^2.$$

Dla Polski współczynnik zmienności sumy X szacowanej wg wzoru (7) wyrażony jest wzorem:

$$(10) v(\hat{x}) = \frac{\sqrt{d^2(\hat{x})}}{\hat{x}},$$

zaś

$$(11) d^2(\hat{x}) = \sum_w d^2(\hat{x}_w).$$

W tablicach prezentujących precyzje oszacowań współczynniki zmienności podane są w procentach.

Badanie pogłowia świń

1. Uwagi wstępne

Celem przeprowadzanych przez GUS dwa razy w roku (tzn. w czerwcu oraz w grudniu) badań pogłowia trzody jest uzyskanie szczegółowych informacji o pogłowie trzody według regionów, województw oraz dla Polski. Badaną populację stanowią indywidualne gospodarstwa rolne, które według danych dostępnych w operacie posiadały trzodę. Badana populacja liczyła 94 627 gospodarstw. Założono, że próba do badania liczyć będzie około 30 tys. gospodarstw. Taka wielkość próby była konsekwencją uwzględnienia kilku czynników m.in.: kosztów realizacji badania, analizy danych z lat poprzednich oraz konieczności zapewnienia odpowiedniej jakości wyników na poziomie kraju i województw.

2. Operat losowania

Przy tworzeniu operatu losowania wykorzystano Operat do Badań Rolniczych. Rejestr ten stworzony został w GUS na bazie indywidualnych wyników Powszechnego Spisu Rolnego 2020. Informacje o gospodarstwach rolnych zapisane w rejestrze aktualizowane są danymi ze źródeł administracyjnych oraz wynikami kolejnych rolniczych badań reprezentacyjnych. Jednostką losowania było indywidualne gospodarstwo rolne. Dla każdego gospodarstwa zapisane zostały następujące informacje:

- symbol województwa oraz regionu NUTS 2,
- nr gospodarstwa,
- powierzchnia ogólna gospodarstwa,
- powierzchnia użytków rolnych,
- liczba sztuk trzody.

3. Schemat losowania

W celu wylosowania próby zastosowany został schemat losowania warstwowo- optymalnego. Spośród gospodarstw prowadzących hodowlę trzody postanowiono

wylosować próbę liczącą około 30 tys. gospodarstw.

Przystępując do losowania próby z tej kategorii gospodarstw przyjęto następujące założenia:

- (1) liczebność próby n ustalona jest dla populacji gospodarstw w Polsce, a nie dla poszczególnych regionów, przy czym n liczy ok. 30 tys. gospodarstw,
- (2) w poszczególnych regionach próba losowana jest według schematu losowania warstwowego - optymalnego wg metody Neymana,
- (3) w każdym regionie dokonywany jest najpierw podział populacji na 5 warstw ($h = 1, 2, \dots, 5$), po czym dokonuje się alokacji próby pomiędzy warstwy,
- (4) w każdym regionie do warstwy nr 5 (tj. $h = 5$) wydzielane są jednostki losowania o wartościach zmiennej przyjętej za podstawę warstwowania powyżej określonego progu. Utworzona w ten sposób tzw. górna warstwa zawiera jednostki, które nie są losowane, lecz wszystkie zaliczane są do próby,
- (5) przyjęto, że oczekiwana precyzja wyników badania, mierzona współczynnikiem zmienności pogłowia trzody będzie jednakowa dla każdego regionu i równa 0,13 % (konkretny parametr liczbowy jest ustalony empirycznie tak, aby spełnić narzucone ograniczenia na liczebność próby).

Powyższy problem rozwiązany został przy wykorzystaniu metod optymalizacji numerycznej. Populacja podzielona została na warstwy, których granice (górne) wyrażone w liczbie sztuk trzody przedstawione zostały w poniższej tablicy.

Tabl. 2. Granice warstw według regionów w badaniu pogłowia trzody w 2021 roku.

Kod regionu	b_1	b_2	b_3	b_4
021	7	15	25	43
041	16	34	61	104
061	7	16	38	204
081	7	15	26	202
101	15	38	83	338
121	4	8	15	25
141	5	10	74	221
142	15	46	109	493
161	13	28	52	188
181	3	6	12	17
201	11	29	65	557
221	12	26	45	75
241	7	14	24	113
261	6	15	25	119
281	30	73	462	1043
301	18	43	83	153
321	7	18	68	255

Przedstawiona w tabl.2. granica warstwy 4 tj. b_4 stanowi jednocześnie próg, powyżej którego jednostki losowania zaliczane są do warstwy 5 czyli nie podlegają losowaniu, lecz badane są w 100%. Dla pozostałych warstw tj. $h = 1, 2, \dots, 4$ ustalone zostały, zgodnie z metodą optymalnej alokacji Neymana, wartości n_{wh} tj. liczebności prób losowanych z h -tej warstwy w w -tym regionie.

Do próby wylosowano następnie, zgodnie z przyjętą alokacją, 30 000 gospodarstw, w tym 12 529 gospodarstw z warstwy nr 5.

4. Uogólnianie wyników i metoda oceny precyzji

Podstawowym parametrem szacowanym w badaniu pogłowia trzody jest suma wartości zmiennej X np. pogłowie trzody ogółem.

Parametr ten dla w -tego regionu jest postaci:

$$(1) \hat{x}_w = \sum_h \sum_i W1_{whi} * x_{whi}, \quad (i = 1, 2, \dots, n_{wh}; h = 1, 2, \dots, 5)$$

gdzie:

x_{whi} – wartość zmiennej X w i -tym gospodarstwie (jednostce losowania) wylosowanym z h -tej warstwy w w -tym regionie,

$W1_{whi}$ – waga przypisana i -temu gospodarstwu wylosowanemu z h -tej warstwy w w -tym regionie, przy czym waga ta obliczana jest ze wzoru:

$$(2) W1_{whi} = \frac{N_{wh}}{n_{wh}},$$

N_{wh} – liczba jednostek losowania w operacji w h -tej warstwie w -tego regionu,

n_{wh} – liczba jednostek losowania wylosowanych do próby z h -tej warstwy w -tego regionu,

Waga $W1_{whi}$ może być stosowana do estymacji wyników badania tylko wtedy, gdy badanie jest kompletne. Waga ta musi być korygowana, jeżeli część gospodarstw wylosowanych do próby odmawia udziału w badaniu. W tym celu wylosowaną próbę na podstawie informacji o realizacji badania dzielimy na 4 grupy:

- (1) gospodarstwa zbadane,
- (2) gospodarstwa, które odmówiły udziału w badaniu,
- (3) gospodarstwa zlikwidowane,
- (4) gospodarstwa, z którymi nie nawiązano kontaktu podczas realizacji badania.

Dla każdej warstwy, oddzielnie w każdym województwie, ustalamy liczebności powyższych grup tj.: $n1_{wh}$, $n2_{wh}$, $n3_{wh}$ i $n4_{wh}$, po czym ustalamy frakcję gospodarstw zbadanych i gospodarstw niezbadanych wśród gospodarstw o ustalonym statusie czyli:

$$(3) c_{wh} = \frac{n1_{wh} + n2_{wh}}{n_{wh} - n4_{wh}}.$$

Następnie obliczamy, dla wylosowanej próby, liczbę gospodarstw aktywnych n_{awh} w h -tej warstwie w -tego województwa:

$$(4) n_{awh} = n1_{wh} + n2_{wh} + c_{wh} * n4_{wh}$$

Na tej podstawie, dla danej warstwy, obliczany jest mnożnik korygujący R_{wh} :

$$(5) R_{wh} = \frac{n_{awh}}{n1_{wh}}$$

Mnożnik ten służy do korekty wagi $W1_{whi}$ w celu uzyskania wagi końcowej W_{whi} :

$$(6) W_{whi} = R_{wh} * W1_{whi}$$

Ocena sumy zmiennej X dla Polski jest sumą wartości uzyskanych dla regionów tj.:

$$(7) \hat{x} = \sum_w \hat{x}_w, \quad (w = 1, 2, \dots, 17)$$

Wagi pierwotne wynikające z losowania próby, korygowane są nie tylko ze względu na niepełną kompletność badania, ale również ze względu na występowanie tzw. gospodarstw nietypowych (*ang. outlier*). Dotyczy to gospodarstw, którym przypisano dużą wagę (losowanych z dużą frakcją), a jednocześnie ze stosunkowo dużymi wartościami dla niektórych badanych zmiennych. Korekta wagi ma zapobiegać w tym wypadku znacznemu przeszacowaniu wartości badanej zmiennej.

Do oszacowania precyzji uogólnień z badania pogłowia świń zastosowano formuły opisane w części dotyczącej badania pogłowia bydła, owiec i drobiu.

Tabl. 3. Wartości względnego błędu standardowego dla wybranych cech dla Polski – na podstawie wyników reprezentacyjnego badania pogłowia bydła, owiec i drobiu oraz badania pogłowia świń zrealizowanych w grudniu 2021 r.

Nr kolejny cechy	Nazwa cechy	Względny błąd standardowy
1.	Bydło ogółem	0,65
2.	Krowy	1,00
3.	Świnie ogółem	0,68
4.	Lochy ogółem	1,41
5.	Drób kurzy	1,68
6.	Nioski kurze	1,11

Methodological notes

1. Sources of data

The data in this publication were compiled on the basis of:

- generalized results of sample surveys¹ on cattle, sheep, poultry and pigs, as well as, the animal output in private farms,
- statistical reports in the scope of livestock in state and cooperative farms and companies with public and private property share,
- statistical reports from slaughterhouses of farm animals,
- statistical reports from poultry hatcheries,
- information on the livestock of poultry from voivodship experts,
- own estimates.

Surveys on cattle, sheep, poultry and animal output were conducted in private farms breeding the above-listed species of animals; this sample amounted to 30 thousand farms.

Surveys on pigs and production of pigs for slaughter were carried out in a sample of private farms breeding pigs; this sample amounted to 30 thousand farms.

The results of the survey of farm animal stocks and animal output were compiled by voivodship and NUTS 2 regions, according to the residence of the farm user, i.e. for private farms - according to the official residence (place of residence) of the farm user, while for state owned farms, cooperative farms and companies – according to the official residence of the holding.

2. Major definitions, terms and enumeration rules

Agricultural holding is understood as an organised economic and technical unit with separate management (a holder or a manager), conducting agricultural activity.

Agricultural activity includes activity associated with crop production, which covers: all agricultural crops (including mushroom production), vegetable production and horticulture, arboriculture, agricultural and horticultural crop and seed production, as well as animal production in a holding, i.e. cattle, sheep, goats, horses, pigs, poultry, rabbits, other fur animals, game kept for slaughter, bees, and activity consisting of maintaining unused agricultural land for production purposes according to the rules of a good agricultural conditions in compliance with environmental protection requirements.

Natural person's holding (private farm) is understood as a holding used by a natural person with an area of 1.0 ha or more of agricultural land, or a holding of the area of less than 1.0 ha of agricultural land (including holdings without agricultural land), which meets at least one of the thresholds mentioned below:

0.5 ha of fruit trees plantation,

0.5 ha of fruit shrubs plantation, soil-grown vegetables, soil-grown strawberries, hop,

0.3 ha of fruit and ornamental nurseries,

0.1 ha of vegetables under cover,

0.1 ha of strawberries under cover, flowers and ornamental plants under cover, tobacco,

¹The surveys on pigs, cattle, sheep and poultry stock are conducted twice a year, i.e. in June and in December.

25 m² of edible mushrooms,
 10 head of cattle in total,
 5 head of cows in total,
 50 head of pigs in total,
 10 head of sows,
 20 head of sheep in total, goats in total,
 100 head of poultry in total,
 5 head of horses in total,
 50 head of female rabbits,
 5 head of other female fur animals,
 10 head of game (e.g. wild boars, roe deer, fallow deer) kept for slaughter,
 20 beehives
 or, regardless of the above thresholds, is an organic farm.

The holding of a legal person or an organisational unit without a legal personality is an agricultural holding run by legal person or an organisational unit without legal personality, the basic activity of which is classified, according to the Polish Classification of Activities, to Section A, division 01, group:

01.1 – growing of non perennial crops,

01.2 – growing of perennial crops,

01.3 – plant propagation,

01.4 – animal production,

01.5 – mixed farming,

01.6, class 01.61 – support activities for crop production (maintaining the land in accordance with cultivation principles with respect for environment protection requirements), and also, irrespective of the basic activity classification, when the area of agricultural land used by the holding is 1 ha or more or when livestock is reared or bred.

Holder of an agricultural farm is understood as a natural person or a legal person or an organisational unit without a legal personality actually using the farm, regardless of whether he or she is an owner or a leaseholder, or uses the farm in any other respect, and regardless of whether land constituting the farm is situated in one or in several gminas.

Livestocks

The survey covered the livestock staying in the agricultural holding during the survey, as well as animals sent to herding, grazing and shepherd's huts. All animals were registered, i.e. the ones owned by a holder or members of his/her household, as well as animals temporarily or permanently kept in the holding, i.e. taken for rearing, fattening, etc., irrespective whether they were taken from private holdings, state-owned holdings, cooperative entities, or companies.

Dairy cows are understood as cows which, due to their breed, variety or particular qualities, are kept in a holding exclusively or mainly for the production of milk to be consumed or to

be processed into dairy products. Dairy cows culled from breeding herds which at the moment of performing the survey were kept in the holding as pre-slaughter pasturing, after which they are sent to slaughter, are also included in this group.

Other cows are understood as cows which, due to their breed (beef breed cows and cows born from a cross-breed with beef breeds) or particular qualities, are kept in a holding exclusively or mainly for calves for slaughter, and whose milk is used to feed calves or other animals. Suckling cows culled from breeding herds which at the moment of performing the survey were kept in the holding as pre-slaughter pasturing, after which they are sent to slaughter, are also included in this group.

In the case of holdings engaged in the production of poultry on a large scale (such as a large-scale holding producing broilers or hen eggs), in which no poultry has been recorded on the survey day due to the current technological break in production, whenever such break does not exceed 8 weeks, the poultry stocks from the period before emptying the rooms (poultry houses) have been adopted.

Information on the number of pigs, cattle, sheep and poultry contained in this publication refers to the stock in June and December 2021.

The percentages are presented with one decimal point and due to the electronic technique of rounding may not sum up into 100%. These figures are substantially correct.

Sampling scheme

Survey on cattle, sheep and poultry stock

1. Introductory notes

The purpose of the surveys conducted by the Statistics Poland twice a year (i.e. in June and in December) is to obtain detailed information on the number of cattle and poultry, both by regions, voivodships and for Poland, and on the number of sheep for Poland only. The surveyed population consists of private agricultural farms which, according to the sampling frame were keeping cattle, or poultry, or sheep. The surveyed population in 2021 consisted of 602 thousand farms. It was assumed that the sample for the survey would consist of about 30,000 farms. This sample size was a consequence of taking into account several factors: the costs of the study, the analysis of data from previous years and the need to ensure the quality of results at the national and regional level.

2. Sampling frame

Sampling frame for all agricultural research as the registry was created by the Statistics Poland on the basis of individual results of the Agricultural Census 2020.

Information on farms in this frame are updated using data from administrative sources and from the results of the following agricultural sample surveys. An individual agricultural farm constituted a sampling unit. The following information was recorded for each farm:

- voivodship code and NUTS 2 region code,
- farm number,
- total farm area,
- agricultural land in the farm ,

- number of cattle,
- number of poultry,
- number of sheep.

3. Sampling scheme

Before sampling, the population of farms was divided into three parts. **The first part** included farms fulfilling at least one of the following criteria, i.e. farms with at least one head of cattle or farms with more than 50 head of poultry and without sheep. This part of population included 395 108 farms. **The second part** consisted of farms with no cattle or sheep, and with no more than 50 head of poultry. The second part amounted to 194 845 farms. Finally, **the third part** included farms keeping sheep, and it amounted to 11 661 farms.

Sample drawing was done with a stratified and optimal sampling scheme. The number of cattle and poultry was used in the first part of the population as the criteria for stratification and allocation of the sample between the strata. In the second part, the strata were established on the basis of the agricultural land, whereas in the third part – on the basis of the number of head of sheep. There were created 8 strata in each region, of which 5 related to farms from the first part, and 3 related to farms from the second part. In the third part, 6 national strata were established, i.e. strata that covered farms from all regions.

It was decided that from the first part farms a sample consisting of approx. 24.7 thousand farms will be drawn.

The following assumptions were made while drawing the sample from this category of farms:

- (1) the size of sample n established for the population of farms in Poland, and not for individual regions, where n consists of approx. 24 700 farms,
- (2) the sample is drawn in individual regions according to the stratified and optimal sampling scheme, with the use of the Neyman method,
- (3) the population in each region is first divided into 5 strata ($h = 1, 2, \dots, 5$), and the sample is then allocated between these strata,
- (4) stratum no. 5 (i.e. $h = 5$) in each region consists of such sampling units for which the value of variables adopted as the stratification basis is above the specified threshold. The stratum created in this way, so called the upper stratum, includes the units which are not drawn, but which are all included in the sample,
- (5) it has been assumed that the expected accuracy of the survey results, measured with the variation coefficient of the livestock of cattle or poultry, will be identical for each region and will be approximately equal to 1.0% (a specific numerical parameter was set empirically to meet the imposed limits on the sample size).

The above problem was solved with the use of the numerical optimization method. The population was divided into strata whose boundaries (upper) expressed in the number of cattle and poultry are presented in Table 1.

Table 1. Boundaries of strata by region in the survey on cattle, sheep, and poultry stocks in 2021

NUTS 2	B – cattle	b_1	b_2	b_3	b_4
--------	------------	-------	-------	-------	-------

	D - poultry				
PL51	B	3	15	32	75
	D	729	746	752	5985
PL61	B	12	33	71	278
	D	815	897	903	14990
PL81	B	7	23	64	831
	D	711	1051	1597	16067
PL43	B	6	23	49	130
	D	1727	2032	2080	18050
PL71	B	6	22	48	154
	D	807	847	863	10400
PL21	B	3	11	37	797
	D	261	473	498	12512
PL91	B	2	4	14	37
	D	78	172	184	248
PL92	B	8	29	70	520
	D	1322	1385	1494	20927
PL52	B	8	29	62	132
	D	810	811	872	15871
PL82	B	0	3	18	250
	D	143	263	293	74998
PL84	B	21	57	174	1584
	D	754	755	771	28950
PL63	B	9	30	80	1004
	D	1091	2197	4007	19039
PL22	B	4	20	52	1585
	D	1004	1007	6705	29990
PL72	B	3	8	21	55
	D	221	289	298	299
PL62	B	11	32	65	184
	D	617	620	639	6970
PL41	B	13	37	81	314
	D	1256	1393	1451	18454
PL42	B	7	27	64	1303
	D	2250	5243	16099	50039

The boundaries of stratum 4, i.e. \mathbf{b}_4 , presented in Table 1, constitute also a threshold above

which the farms are included in stratum 5, which means that they are not subject to sampling, but are all included in the sample. For other strata, i.e. $h = 1, 2, \dots, 4$, the Neyman optimal allocation method was applied for establishing the values of n_{wh} i.e. size of the samples drawn from the h -stratum in w -region.

After that, 24 748 farms were drawn to the sample, based on the assumed allocation, including 7 420 farms from stratum no. 5.

Before sampling, 3 strata were established in each region in the second part ($h = 6, 7, 8$). These strata were created in respect to agricultural land, i.e.: $h = 6$: farms of less than 15 ha; $h = 7$: farms of 15 ha to 49,99; $h = 8$: farms of 50 ha or more. Identical accuracy of the number of poultry in each region was adopted as the criterion for allocation of the sample between regions, while within regions the sample was allocated by means of the Neyman optimal method. From this part of the population 3 121 farms were drawn.

In the third part, in which 6 national strata were established ($h = 9, 10, \dots, 14$), 2 131 farms were drawn for the sample. All farms from stratum 14 were included in the sample. These were farms keeping sheep and simultaneously 50 or more head of cattle, or at least 400 head of poultry (i.e. 203 farms). The boundaries of other strata, and the assumed number of the sample allocated between these strata, were established with the above mentioned numerical optimization method. The upper stratum ($h = 13$) was also established, from which no farms were drawn. This stratum included farms which had not been previously included in stratum 14, and which kept more than 45 head of sheep. The upper boundaries of the remaining strata were the following: $b_9 = 3$, $b_{10} = 8$, $b_{11} = 20$, $b_{12} = 45$. The aim of establishing this category of farms as a separate one, as well as optimizing the division into strata, was to accurately estimate the data on the livestock of sheep in country terms, with no regional breakdown.

Eventually, the entire sample for the survey on cattle, poultry, and sheep consisted of 30 000 farms.

4. Results generalization and the accuracy assessment method

The sum of X variable value, such as cattle stock in total, is the basic parameter estimated in the survey of livestock of cattle, sheep and poultry.

This parameter for w -region is calculated according to the formula:

$$(1) \hat{x}_w = \sum_h \sum_i W1_{whi} * x_{whi}, \quad (i = 1, 2, \dots, n_{wh}; h = 1, 2, \dots, 14)$$

where:

x_{whi} – the value of X variable in i -farm (sampling unit) drawn from h -stratum in w -region,

$W1_{whi}$ – the weight assigned to i -farm drawn from h -stratum of w -region, whereas this weight is calculated according to this formula:

$$(2) W1_{whi} = \frac{N_{wh}}{n_{wh}},$$

N_{wh} – the number of sampling units in h -stratum of w -region,

n_{wh} – the number of sampling units drawn for the sample from h -stratum of w -region.

The $W1_{whi}$ weight might be used to estimate the survey results only if the survey is complete. This weight must be corrected if some of the sampled farms refuse to participate in the survey. For this purpose, the drawn sample is divided into 4 groups based on information on the survey performance:

- (1) the surveyed farms,
- (2) farms that refused to participate in the survey,
- (3) closed down farms,

(4) farms with which the contact was not established during the survey performance.

For each stratum separately in each region, the size of the above groups, namely $n1_{wh}$, $n2_{wh}$, $n3_{wh}$, and $n4_{wh}$ is established, and then the likelihood function of surveyed and not surveyed among the farms with a determined status is established, that is:

$$(3) c_{wh} = \frac{n1_{wh} + n2_{wh}}{n_{wh} - n4_{wh}}.$$

Then the number of the n_{awh} active farms in h -stratum of w -region is calculated for the drawn sample:

$$(4) n_{awh} = n1_{wh} + n2_{wh} + c_{wh} * n4_{wh}$$

On this basis, the R_{wh} correction factor is calculated for a given stratum:

$$(5) R_{wh} = \frac{n_{awh}}{n1_{wh}}$$

The purpose of this factor is to correct the $W1_{whi}$ weight in order to obtain final W_{whi} weight:

$$(6) W_{whi} = R_{wh} * W1_{whi}$$

The sum of X variable value for Poland is the sum of values obtained for particular regions, i.e.:

$$(7) \hat{x} = \sum_w \hat{x}_w, \quad (w = 1, 2, \dots, 17)$$

Original weights resulting from sampling are corrected not only due to incompleteness of the survey but also due to the occurrence of so called unusual farms (outliers). This pertains to farms with high assigned weight (drawn with a high likelihood function) and, at the same time, with relatively high values for some of the analyzed variables. In this case, the weight correction is to prevent significant overestimation of the value of the surveyed variable.

In the last step weights were calibrated taking into account the data from the Agency for Restructuring and Modernisation of Agriculture (ARMA) for the number of farms, number of cattle, and number of poultry in voivodships.

For the selected major assessments of the parameters, their variation coefficients were calculated as the accuracy measures. For an estimator expressed by formula (1) i.e. for w -region, its variation coefficient estimation is expressed in the following formula:

$$(8) v(\hat{x}_w) = \frac{\sqrt{d^2(\hat{x}_w)}}{\hat{x}_w},$$

where $d^2(x_w)$ is an estimate of the so-called variance of the estimator \hat{x}_w and is calculated according to the formula (which follows from the classical survey sampling theory for stratified sampling scheme):

$$(9) d^2(\hat{x}_w) = \sum_h \left(1 - \frac{n1_{wh}}{\sum_i W_{whi}}\right) \frac{n1_{wh}}{n1_{wh}-1} \sum_i \left(W_{whi} x_{whi} - \frac{\sum_j W_{whj} x_{whj}}{n1_{wh}}\right)^2.$$

For Poland the variation coefficient of the sum X estimated with the formula (7) is expressed by the following formula:

$$(10) v(\hat{x}) = \frac{\sqrt{d^2(\hat{x})}}{\hat{x}},$$

whereas:

$$(11) d^2(\hat{x}) = \sum_w d^2(\hat{x}_w).$$

In the tables presenting the precision of the estimates, the coefficients of variation are given as a percentage.

Survey on pigs

1. Introductory notes

The purpose of the survey on pigs stocks, conducted by the Statistics Poland two times a year in June and December is to obtain detailed information on the number of pigs by regions, voivodships and for Poland. The surveyed population consisted of individual farms which, according to the data available in the sampling frame, were keeping pigs. The surveyed population in 2021 consisted of 94 627 farms. It was assumed that the sample for the survey would consist of about 30,000 farms. This sample size was a consequence of taking into account several factors: the costs of the study, the analysis of data from previous years and the need to ensure the quality of results at the national and regional level.

2. Sampling frame

Sampling frame for all agricultural surveys as the frame was created by the Statistics Poland on the basis of individual results of the Agricultural Census 2020.

Information on farms in this frame are updated using data from administrative sources and from the results of the following agricultural sample surveys. An individual agricultural farm constituted a sampling unit. The following information was recorded for each farm:

- voivodship code and NUTS 2 region code,
- farm number,
- total farm area,
- agricultural land,
- number of pigs.

3. Sampling scheme

In order to draw sample, a stratified sampling and optimal scheme was used with respect to farms which, according to the sampling frame, reared pigs. It was decided that a sample consisting of approx. 30 thousand farms will be drawn from all farms breeding and rearing pigs.

The following assumptions were made while drawing the sample from this category of farms:

- (1) the size of n sample is established for the population of farms in Poland, and not for individual regions, where n consists of approx. 30 thousand farms,
- (2) the sample is drawn in individual regions according to the stratified and optimal sampling scheme, by means of the Neyman method,
- (3) the population in each region is first divided into 5 strata ($h = 1, 2, \dots, 5$), and then the sample is allocated between these strata,
- (4) stratum no. 5 (i.e. $h = 5$) in each region consists of such sampling units, for which the value of at least one of the variables adopted as the stratification basis is above the specified threshold. The stratum created this way, regarded as the upper stratum, includes the units which are not drawn, but which are all included in the sample,
- (5) it has been assumed that the expected accuracy of the survey results, measured with the variation coefficient of the livestock of pigs, will be identical for each region and will be equal approximately to 0.13 % (a specific numerical parameter was set empirically to meet the imposed limits on the sample size).

The above problem was solved with the use of the numerical optimization method. The population was divided into strata whose boundaries (upper) expressed in the number of pigs were presented in Table 2 below.

Table 2. Boundaries of strata by region in the survey on pigs stock in 2021.

NUTS 2	b_1	b_2	b_3	b_4
PL51	7	15	25	43
PL61	16	34	61	104
PL81	7	16	38	204
PL43	7	15	26	202
PL71	15	38	83	338
PL21	4	8	15	25
PL91	5	10	74	221
PL92	15	46	109	493
PL52	13	28	52	188
PL82	3	6	12	17
PL84	11	29	65	557
PL63	12	26	45	75
PL22	7	14	24	113
PL72	6	15	25	119
PL62	30	73	462	1043
PL41	18	43	83	153
PL42	7	18	68	255

The boundary of stratum 4, i.e. b_4 , presented in Table 2, also constitutes a threshold, above which the sampling units are included in stratum 5, which means they are not subject to sampling, but are all included in the sample. For other strata, i.e. $h = 1, 2, \dots, 4$, the Neyman optimal allocation method was employed for establishing the values of n_{wh} , i.e. size of the samples drawn from the h -stratum in w -region. After that, 30 000 farms were drawn to the sample, based on the assumed allocation, including 12 529 farms from stratum no. 5.

2. Results generalization and the accuracy assessment method

The sum of X variable value, such as pigs stock in total, is the basic parameter estimated during the survey on the livestock of pigs.

This parameter for w -region is calculated according to the formula:

$$(1) \hat{x}_w = \sum_h \sum_i W1_{whi} * x_{whi}, (i = 1, 2, \dots, n_{wh}; h = 1, 2, \dots, 5)$$

where:

x_{whi} – the value of X variable in i -farm (sampling unit) drawn from h -stratum in w -region,

$W1_{whi}$ – the weight assigned to i -farm drawn from h -stratum of w -region, whereas this weight is calculated according to this formula:

$$(2) W1_{whi} = \frac{N_{wh}}{n_{wh}},$$

N_{wh} – the number of sampling units in h -stratum of w -region,

n_{wh} – the number of sampling units drawn for the sample from h -stratum of w -region.

The $W1_{whi}$ weight might be used to estimate the survey results only if the survey is complete. This weight must be corrected if some of the sampled farms refuse to participate in the survey. For this purpose, the drawn sample is divided into 4 groups based on information on the survey performance:

- (1) the surveyed farms,
- (2) farms that refused to participate in the survey,
- (3) closed down farms,
- (4) farms with which the contact was not established during the survey performance.

For each stratum separately in each region, the size of the above groups, namely $n1_{wh}$, $n2_{wh}$, $n3_{wh}$, and $n4_{wh}$ is established, and then the likelihood function of surveyed and not surveyed among the farms with a determined status is established, that is:

$$(3) c_{wh} = \frac{n1_{wh} + n2_{wh}}{n_{wh} - n4_{wh}}.$$

Then the number of the n_{awh} active farms in h -stratum of w -region is calculated for the drawn sample:

$$(4) n_{awh} = n1_{wh} + n2_{wh} + c_{wh} * n4_{wh}$$

On this basis, the R_{wh} correction factor is calculated for a given stratum:

$$(5) R_{wh} = \frac{n_{awh}}{n1_{wh}}$$

The purpose of this factor is to correct the $W1_{whi}$ weight in order to obtain final W_{whi} weight:

$$(6) W_{whi} = R_{wh} * W1_{whi}$$

The sum of X variable value for Poland is the sum of values obtained for particular regions, i.e.:

$$(7) \hat{x} = \sum_w \hat{x}_w, \quad (w = 1, 2, \dots, 17)$$

Original weights resulting from sampling are corrected not only due to incompleteness of the survey but also due to the occurrence of so called unusual farms (outliers). This pertains to farms with high assigned weight (drawn with a high likelihood function) and, at the same time, with relatively high values for some of the analyzed variables. In this case, the weight correction is to prevent significant overestimation of the value of the surveyed variable.

To estimate the precision of the estimation from survey of pig the formulas described in the section on sample survey of the livestock of cattle, sheep and poultry were used.

Table 3. The values of the relative standard error of selected characteristics for Poland – based on the results of a sample survey of the livestock of cattle, sheep and poultry as well as the results of a survey of pigs – conducted in December 2021.

No. of the attribute	Name of characteristics	Relative standard error
1.	Cattle total	0.65
2.	Cows	1.00
3.	Pigs total	0.68
4.	Sows total	1.41
5.	Hens	1.68
6.	Laying hens	1.11

Objaśnienia znaków umownych i ważniejsze skróty

Symbols and main abbreviations

Symbol <i>Symbol</i>	Opis <i>Description</i>
p. proc.	punkt procentowy percentage point
tys.	tysiąc thousand
mln	milion million
ha	hektar hectare
r.	rok year
cd.	ciąg dalszy continued
dok.	dokończenie completion
tabl.	tablica table
Znak #	dane ukryte ze względu na tajemnicę statystyczną necessity of maintaining statistical confidentiality
Kreska (-)	zjawisko nie wystąpiło magnitude zero
Zero (0,0)	zjawisko istniało w wielkości mniejszej od 0,5 magnitude not zero but less than 0.05 of a unit
Kropka (.)	zupełny brak informacji albo brak informacji wiarygodnych data not available or not reliable
Znak x	wypełnienie pozycji jest niemożliwe lub niecelowe not applicable or not reliable
„W tym”	oznacza, że nie podaje się wszystkich składników sumy indicates that not all elements of the sum are given