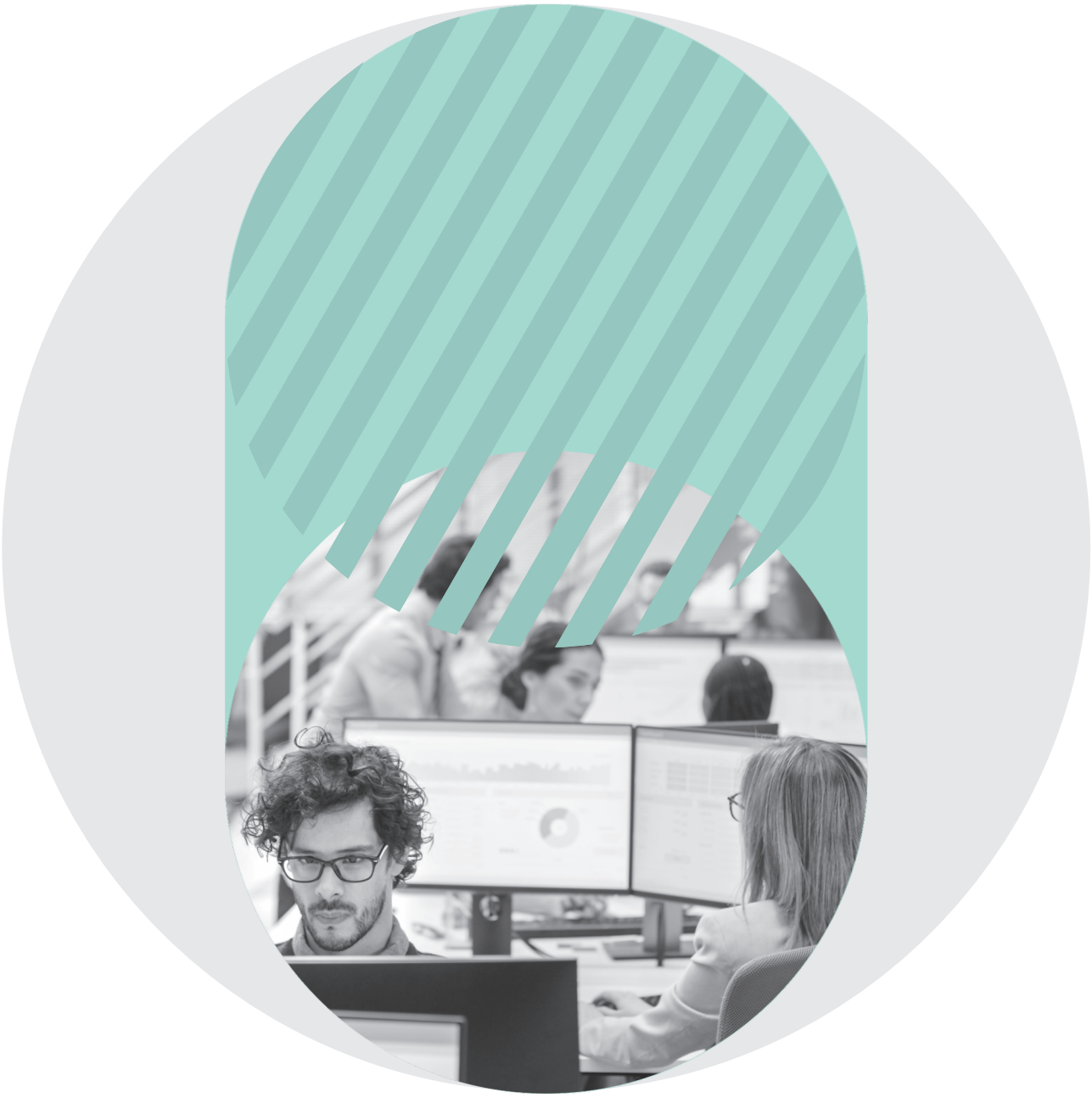
Statistical research methodology

Statistics Poland

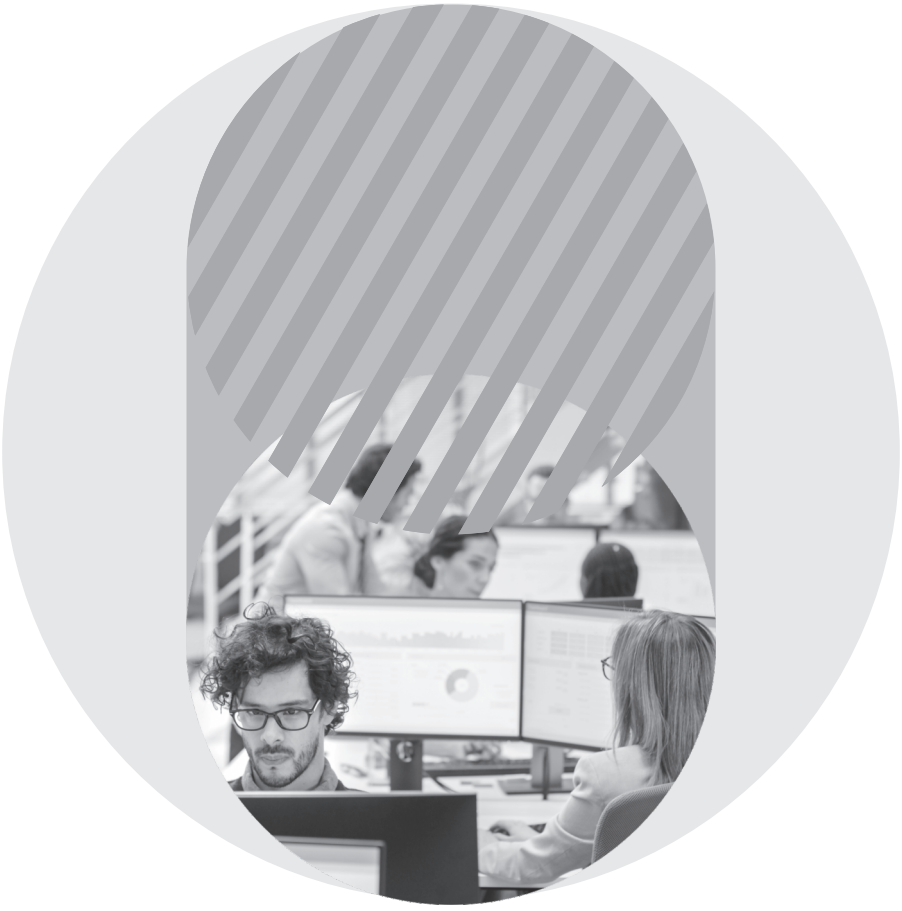


Methodological report
Adult Education Survey

Gdańsk 2026

Statistical research methodology

Statistics Poland



Methodological report
Adult Education Survey

Statistical Office in Gdańsk

Gdańsk 2026

Content-related works

Statistical Office in Gdańsk, Centre for Education and Human Capital Statistics

Supervisor

Jerzy Auksztol

Editorial team

Ewelina Grochowska-Subotowicz, Piotr Jaworski, Anna Mistrzak, Aleksandra Sarnowska, Joanna Wyzina, Anna Żochowska

Cooperation

Robert Wieczorkowski	Department for Innovation, Statistics Poland
Janusz Partyka	Centre for Data Engineering, Statistical Office in Gdańsk

Editorial work

Magdalena Poleszuk	Pomeranian Centre for Regional Surveys, Statistical Office in Gdańsk

Typesetting and graphics

Malgorzata Kloz	Pomeranian Centre for Regional Surveys, Statistical Office in Gdańsk

Methodological report, reviewed by the Methodological Commission of Statistics Poland

Publication available on website

stat.gov.pl

When publishing please indicate the source

Metadata

Title of the methodological report	Adult Education Survey
Author	Statistical Office in Gdańsk
Statistical units in the survey	The observation unit are persons who are members of households living in sampled dwellings and meeting the criteria of the resident population. The units of analysis in the AES are: • households, • persons aged 18-69.
Scope of data in the survey	Educational activity of household members, i.e. participation in formal education, non-formal education and informal learning, as well as access to information on learning possibilities, obstacles to participation in education, learning methods used, time spent on learning, expenditure related to acquiring knowledge and skills, and the level and distribution of foreign language proficiency (during the 12 months preceding the interview).
Type and method of the survey	The Adult Education Survey is a voluntary sample survey conducted every 6 years. The sample is selected using a two-stage sample design. The respondents are persons living in sampled dwellings.
Data collection tools and data sources	Data collection tools: the AES household questionnaire (intended for household) and the AES individual questionnaire (addressed to persons aged 18-69 who are members of a given household) – in both an electronic and paper versions. Data collection procedure: information from respondents is collected through face-to-face or telephone interviews conducted by interviewers using the above questionnaires.
Dissemination formats of the survey results	Publications: Adult Education 2022 Human capital in Poland in the years 2018–2022 News release: Adult Education in 2022
Classifications used	1. Polish Classification of Activities (PKD 2007): Polska Klasyfikacja Działalności PKD 2007 2. Classification of Occupations and Specialisations: Klasyfikacja Zawodów i Specjalności 3. International Standard Classification of Education (ISCED 2011) and International Standard Classification of Education – Fields of Education and Training (ISCED-F 2013): International Standard Classification of Education (ISCED) 4. Classification of Learning Activities 2016 (CLA 2016): Classification of Learning Activities (2016 edition) 5. Classification of Territorial Units for Statistics (NUTS): Klasyfikacja NUTS w Polsce 6. ISO 3166 — Country Codes 7. ISO 639 - Language Codes 8. Common European Framework of Reference for Languages (CEFR)
Date of preparation of the methodological report	May 2026

Contents

Metadata 2
Main abbreviations 4
Introduction 5
Historical overview 7
1. Statistical units and scope of data in the survey 9
2. Type and method of the survey 10
2.1. Survey method and reference period 10
2.2. Sample design 10
2.3. Weighting procedure 12
3. Data collection tools and data sources 16
3.1. Data collection tools 16
3.2. Data collection methods 16
3.3. Direct and proxy interviews 17
4. Variables covered in the survey 18
4.1. Description of the main variables 18
4.2. Definitions of the basic terms used in the AES 19
4.3. Main indicators calculated based on AES results and their definitions 23
4.4. Classifications 25
5. Organisation and management of survey implementation 26
5.1. Official statistics units conducting the survey 26
5.2. Organisation of data collection 26
6. Presentation of the survey results 27
6.1. Main results obtained from the AES 27
6.2. Restrictions regarding the publication of AES results 27
6.3. Principles of access to AES data 28
6.4. Structure of the microdata files 29
6.5. Use of AES results and users of AES data 29
7. Assessment of the survey quality 30
7.1. Unit non-response and item non-response (response rates) 30
7.2. Precision of results and limitations of data interpretation 32
7.3. Measurement and processing errors 33
7.4. Quality reports 34
Bibliography 35
Appendices 36

Main abbreviations

AES	– Adult Education Survey
AUC	– Area under curve – area under the ROC curve – measure of classification performance
CAPI	– Computer Assisted Personal Interviewing
CATI	– Computer Assisted Telephone Interviewing
CIS	– Statistical Computing Centre
CEDEFOP	– European Centre for the Development of Vocational
CORstat	– Survey Management System
CV	– coefficient of variation; also referred to as RSE – relative standard error
CVTS	– Continuing Vocational Training Survey
DG EAC	– Directorate General Education and Culture
EFTA	– European Free Trade Association
ESS	– European Statistical System
EU	– European Union
Eurostat	– Statistical Office of European Union
IESS FR	– Integrated European Social Statistics Framework
ISCED 2011	– International Standard Classification of Education (ISCED) 2011
ISCED-F 2013	– International Standard Classification of Education. Fields of education and training
LFS	– Labour Force Survey
NUTS	– Classification of Territorial Units for Statistics
OBS	– Sampling Frame for Social Surveys
OEK	– Centre for Education and Human Capital Statistics at the Statistical Office in Gdańsk
OECD	– The Organisation for Economic Co-operation and Development
PAPI	– Paper and Pencil Interview
PK	– Programming and Survey Coordination Department of Statistics Poland
PSU	– Primary sampling unit
PWA	– Progressive Web App
ROC	– Receiver Operating Characteristics
SDG	– Sustainable Development Goals
SE	– standard error
SIMS	– Single Integrated Metadata Structure
SSU	– Secondary sampling unit
TERYT	– The National Register of Territorial Division of the Country
UN	– The United Nations

Introduction

Adult Education Survey (AES) is one of the components of the system of European statistical system on adult learning. The other two components are the *Labour Force Survey* (LFS) and the *Continuing Vocational Training Survey* (CVTS).

The main objective of the *Labour Force Survey* (LFS) is to provide data on the labour market; however, it also provides a limited set of indicators concerning adult participation in education and training (formal and non-formal) during the 12 months preceding the interview. These indicators are compiled every two years and currently constitute the primary source of information enabling regular monitoring of participation in lifelong learning in European Union countries. In the LFS, these data are collected and published quarterly. The *Labour Force Survey* is also a source of information on participation in education and training during the four weeks preceding the survey.

The *Adult Education Survey* (AES) is conducted every six years and provides more detailed data on various aspects related to adult participation in formal and non-formal education. It is also the only source of information on informal learning among persons aged 18-69. Formal education refers to learning through participation in education and training programmes leading to the acquisition of a qualification that has been formally confirmed by an authorised institution and entered into the Integrated Qualifications Register. Non-formal education can be defined as institutionalised, organised learning that takes place outside programmes of education and training leading to qualifications formally confirmed by an authorised institution. Informal learning refers to the acquisition of knowledge, skills and social competences through various activities outside organised forms of learning. Detailed definitions of these concepts are presented in Subchapter 4.2.

Unlike the AES, the LFS does not cover informal learning or certain forms of non-formal education. All methodological differences between the two surveys are described in the Eurostat information note available at: Participation in education and training during the last 12 months – differences between data available from two sources. The AES and LFS are household surveys. The European system of adult learning statistics in Europe is complemented by the Continuing Vocational Training Survey (CVTS), which is an enterprise survey. It provides information on enterprises' investment in continuing vocational training for their employees and delivers data on enterprises' strategies regarding continuing vocational training and skills development of employees.

The primary objective of the *Adult Education Survey* is to provide data comparable across European Union countries on learning among persons aged 18-69, which serve as a tool for monitoring adult participation in lifelong learning at both international and national levels. The set of mandatory variables in the survey includes participation of persons in this age group in formal education, non-formal education and informal learning, in relation to socio-demographic characteristics. Additionally, the survey provides information on access to information about learning possibilities, obstacles to participation in education, learning methods, time spent on learning, expenditure related to acquiring knowledge and skills, and the level and distribution of foreign language skills.

The pilot *Adult Education Survey* was conducted in Poland for the first time in the fourth quarter of 2006. It was carried out under a gentlemen's agreement with Eurostat and was partially financed by the European Commission. Subsequent editions of the survey (conducted in 2012, 2017 and 2023) were implemented under EU legal regulations.

The legal basis for conducting the survey in 2023 in Poland was the Act of 29 June 1995 on Official Statistics (Journal of Laws 1995, No. 88, item 439, as amended), as well as the Regulation of the Council of Ministers on the Programme of Statistical Surveys of Official Statistics for the year 2022 of 19 November 2021 (Journal of Laws 2021, item 2303, as amended). Given that the survey also fulfils obligations arising from Poland's membership in the European Union, the legal acts governing the conduct of the European Adult Education Survey are also applicable. In 2021, changes were introduced to the legal framework regulating this survey at the EU level – the Integrated European Social Statistics Framework (IESS FR), i.e. Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019, along with its implementing acts, was put into effect. Compliance with the guidelines set out in EU regulations ensures that the survey is consistent with international methodology.

The presented methodological report consists of seven chapters. It outlines the historical background of the survey, the statistical units (the target population), and the scope of data (the range of information covered by the survey). The subsequent chapters describe: the type of survey and the methods used to conduct it, data collection tools and data sources, the characteristics of the variables included in the survey, the organisation of the survey, the method of presenting the results, and the assessment of survey quality.

The description of the *Adult Education Survey* presented in this report concerns the 2022 edition of the survey. It refers to the methodology introduced in 2022, while the fieldwork was carried out in 2023. The electronic version of the report is also expected to be updated systematically.

Historical overview

The first pilot surveys of AES were conducted by EU countries (including Poland), EFTA and candidate countries between 2005 and 2008. The AES 2007 covered the resident population aged 25-64. Its aim was to collect information on participation in education and training, i.e. participation in formal, non-formal and informal learning during the 12 months preceding the interview. The information collected focused on the purpose (job--related or not) and other characteristics of the learning activities. In addition, AES 2007 covered self-reported language skills, social and cultural participation, computer skills and variables related to the main socio--demographic characteristics of respondents.

In April 2008, Regulation (EC) No 452/2008 of the European Parliament and of the Council concerning the production and development of statistics on education and lifelong learning was adopted. This regulation set the European legal basis for AES data collection. The entry into force of the implementing regulation (Commission Regulation (EU) No 823/2010) allowed for the data collection within a fully harmonised AES survey, covering the resident population aged 25-64. According to this regulation, data collection could take place between 1 July 2011 and 30 June 2012. In Poland, the *Adult Education Survey* 2011 was conducted from 2 January to 29 February 2012 and covered the population aged 18-69.

The next survey (AES 2016) was carried out based on Commission Implementing Regulation (EU) No 1175/2014. Under this regulation, data collection could be conducted between 1 July 2016 and 31 March 2017. The methodological changes introduced (including those concerning the target population and scope of data in the survey) were based on the experience of countries from the previous edition of the survey (AES 2011). They also reflected the evolving demand for data from stakeholders involved in European education and training policies, such as DG EAC and CEDEFOP. The requirements for the AES 2016 were defined so as to take into account changes and progress in the field of lifelong learning, while maintaining consistency with the AES 2011 wherever possible. The scope of the survey was reduced by removing modules on participation in social and cultural activities and on computer skills. Significant changes were also made to the core international classifications used in the survey and their national equivalents. In particular, this involved the introduction into the survey of the new International Standard Classification of Education (ISCED 2011) and the International Standard Classification of Fields of Education and Training (ISCED-F 2013). In Poland, the *Adult Education Survey* 2016 was conducted from 2 January to 28 February 2017 and covered persons aged 18-69. The entry into force of Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019 establishing a common framework for European social statistics resulted in two major changes concerning the target population and the organisation of the survey. It made it mandatory to cover persons aged 18-69 instead of the previously applicable target population of 25-64 across all participating countries. It was also specified that the survey should be conducted every six years, rather than every five years as previously.

The methodological assumptions for the AES 2022 survey did not change significantly, and the overall consistency with the previous edition of the survey was maintained. Only a small number of questions were removed or modified – those that were no longer needed from the perspective of data demand or that raised the greatest substantive doubts. Another important reason for modifying or introducing additional variables in the AES survey was the need to capture new phenomena related to adult education and learning.

The most significant changes to the AES 2022 questionnaire involved modifications to questions concerning e-learning and educational guidance and counselling. For all multiple-choice variables, an additional question was introduced to identify the main response when two or more categories were selected. Variables were also added concerning skills acquired as a result of participation in non-formal education and the initiator of job--related non-formal education. A question was also introduced regarding the motivation for participation in informal learning.

In Commission Implementing Regulation (EU) 2021/861 (Appendix 2), new or more detailed guidelines were specified regarding data collection periods and methods, weighting procedures, and rules for selecting survey respondents and non-formal education activities. The previously allowed nine-month time period (between 1 July 2022 and 31 March 2023) during which the survey should be conducted was maintained (as in AES 2016). However, it was specified that data collection could not last longer than six months within this period. It was also required that data be collected using computer-assisted methods. It was explicitly specified that, in the weighting procedures applied to the survey results, reliable external data on the distribution of the target population (demographic data) by sex, age groups, and level of educational attainment should be taken into account. A rule was introduced that no more than two members of the same household may be interviewed. A requirement was also established that non-formal education activities selected for further data collection in the questionnaire are to be selected at random. This requirement results from the assumption that information is collected for a maximum of five non-formal education activities, of which two are covered in more detail.

The Polish name of the survey was changed from "Kształcenie dorosłych" (questionnaire symbol: KD) to "Uczenie się osób dorosłych" (questionnaire symbol: UD). The change aimed to more precisely reflect the scope of the survey, which covers all forms of intentional and planned activities undertaken to acquire knowledge and skills. In Poland, the *Adult Education Survey* 2022 was conducted from 1 February to 31 March 2023.

1. Statistical units and scope of data in the survey

The unit of observation in the *Adult Education Survey* consists of persons aged 18-69 who are members of households residing in randomly selected dwellings and who meet the criteria of the resident population:

- present in the household for 12 months or more[1] – registered for permanent or temporary residence, and currently residing there or intending to reside there without registration;
- absent for up to 12 months (the total duration of actual and planned absence is considered), e.g. persons temporarily staying abroad, in collective households, or in another household within the country.

The survey does not cover children and youth in pre-primary or school education (aged 3-17 years). It also excludes homeless persons and members of households residing abroad for 12 months or more[2], as well as those living in collective households such as workers' hostels, student dormitories, boarding schools, nursing homes.

The target population to which the estimation of AES results applies consists of persons residing in private households in Poland. Due to this restriction, the population size obtained after weighting is lower than the figure based on current population estimates[3].

The subject of the survey is the educational activity of household members, i.e. participation in educational activities within formal and non-formal education and informal learning during the 12 months preceding the interview.

In addition, the survey provides information on:

- the characteristics of undertaken educational activities;
- the number of hours spent on learning;
- the reasons for and barriers to participation in formal and non-formal education;
- access to information on learning possibilities;
- the financing and the costs of learning;
- self-reported foreign language skills.

Detailed information on the scope of the data collected in the survey is provided in the description of variables in the survey, presented in Subchapter 4.1.

[1] The period of actual presence, including the planned duration, is taken into account.
[2] The actual period, including the planned period, is taken into account.
[3] Data on the population of Poland are based on the results of the National Population and Housing Census of 31 March 2011, with component data (births, deaths and migration) derived from the registers of civil status offices, population registers of individual gminas and the PESEL register (see: *Population. Size and Structure and Vital Statistics in Territorial Breakdown in 2022*. As of 30 June, Statistics Poland, 2022).

2. Type and method of the survey

2.1. Survey method and reference period

The *Adult Education Survey* is a voluntary sample survey, which involves observing a random sample of units of the target population and using statistical inference to generalise the results to the entire population. All households in the selected dwellings are included in the survey. Dwellings in mobile units, collective accommodation establishments, and temporary dwellings are excluded from the sampling frame.

In the selected dwellings, interviews are conducted with persons aged 18-69. In AES 2022, one of the requirements specified in the implementing regulations (Appendix 2) is that interviews should not be conducted with more than two members of the same household. If a household consists of three or more persons aged 18-69, the interviewer must randomly select two persons for participation in the survey. To avoid arbitrary decisions by the interviewer, the selection procedure within the sampled household should use a random mechanism (e.g. the Kish method). Interviews are conducted first with persons aged 18-24. These assumptions aim to ensure objective results and to provide adequate representation of young people in the survey. The rules limiting the number of respondents from a single household are intended to mitigate the homogeneity effect, i.e. the occurrence of similarities in surveyed characteristics among persons living in the same household. Ensuring adequate representation of young participants is intended to meet the precision requirement for the "participation in formal education and training (age 18-24)" indicator specified in the Integrated European Social Statistics Framework (IESS FR).

In exceptional cases, the implementing regulations allow for the possibility of conducting interviews with more than two persons in a single household. Therefore, in the Polish version of the survey, the interviewers' manual includes appropriate guidelines specifying the situations in which this is justified. This applies primarily to persons who spend long periods away from their permanent place of residence (household). It is assumed that factors such as limited personal contact or other causes weakening emotional ties between household members may reduce their mutual influence on individual life decisions, including choices related to education and attitudes toward learning. In AES 2022, a total of 34 households nationwide were identified in which more than two persons were interviewed (which accounted for 0.3% of all households included in the survey).

The reference period adopted in the survey is the 12 months preceding the date of the interview. It applies to all variables relating to participation in lifelong learning (in formal and non-formal education as well as informal learning), a detailed description of educational activities, access to information on possibilities to participate in education and the use of guidance and counselling, as well as obstacles to participation in education. For the remaining variables, the required information refers to the situation at the time of the interview. For questions concerning financing and the costs incurred for educational activities within formal and non-formal education (variables: FEDPAID FEDPAIDBY, NFEPAID, NFEPAIDBY, NFEPAIDVAL)[4], no specific reference period is defined. These refer to all direct expenditures related to education or training, including those incurred earlier than the 12-month reference period.

2.2. Sample design

The sample size and sample design were determined based on experience from the previous edition of the survey (carried out in 2017) and other social surveys conducted by Statistics Poland. The sample was drawn using the Sampling Frame for Social Surveys (OBS), i.e. a database primarily based on TERYT – the National Register of Territorial Division of the Country, including a list of territorial statistical units and an address list of dwellings. The OBS database also includes additional information from key administrative registers.

[4] Detailed information on the survey variables can be found in Appendices 4–6 to the Methodological Report.

A two-stage sample design with unequal selection probabilities at the first stage was applied. Enumeration districts were primary sampling units (PSUs). The sampling frame includes approximately 186,000 districts (PSUs), with an average of 83 dwellings per district. The number of dwellings in a PSU varies depending on the size of the locality and averages 94 in urban areas and 66 in rural areas. Adjacent PSUs were merged if they contained too few dwellings; information on the (x, y) coordinates assigned to dwelling addresses in the sampling frame was used in the merging algorithm. At the second stage, dwellings were selected by simple random sampling without replacement. In the selected dwellings, all households and persons aged 18-69 in the selected dwellings were covered by the survey, with no more than two persons per household (permitted exceptions to this rule are described in Subchapter 2.1).

Primary sampling units were selected independently within each stratum with selection probabilities proportional to the number of dwellings in a given PSU (using the Hartley–Rao method). Prior to sampling, PSUs were stratified by region (according to the NUTS 2 classification[5]).

In constructing the strata, particular attention was paid to dwellings with persons aged 18-24. This was due to Eurostat requirements specifying a precision constraint for the key indicator "participation in formal education and training for persons (age 18-24)".

The strata were defined as cross-classifications of NUTS 2 region, locality size class (6 classes, detailed parameters are presented in Table 2 in subchapter 2.3), and three groups defined on the basis of the share of dwellings with persons aged 18-24 in a PSU. The method used to determine the three PSU groups, based on data available in the sampling frame, was as follows:

- for each PSU, the share of dwellings with persons aged 18-24 relative to the total number of dwellings in that PSU was calculated, and the set of PSUs was then sorted according to this characteristic;
- the sorted set of PSUs was divided into three parts using the 1/3 and 2/3 quantiles. In this way, three groups differing with respect to the selected characteristic were obtained. One of these groups contains substantially more addresses with persons aged 18-24 (i.e. the group with share values above the 2/3 quantile).

The final set of strata was obtained by dividing PSUs in each stratum into two parts of similar size. This division was carried out based on the median number of employed persons in dwellings within each PSU. Subsequently, two new strata were created: one including PSU segments with dwellings where the number of employed persons was above the median, and the other containing PSU segments with dwellings where the number of employed persons was below the median. This division was applied only in cases where the resulting strata contained more than 10,000 addresses. As a result, 385 strata were formed. The additional stratification using the number of employed persons was intended to increase the efficiency of the sample design and was justified by the results of the analysis of data from the previous wave of the survey. This analysis confirmed significant correlations between the number of employed persons and most of the estimated parameters.

The sample allocation was determined in accordance with the algorithm described by Wesołowski and Wieczorkowski (2017). The algorithm uses the eigenvalues of a specially constructed matrix to determine the optimal allocation of the sample in a two-stage design across subpopulations and strata. The optimisation criterion was the minimisation of the relative standard error (coefficient of variation) of the estimator of a given variable (i.e. the number of employed persons available in the sampling frame) at the level of defined subpopulations (i.e. NUTS 2 regions), subject to sample size constraints at each stage of sampling. It was assumed that, on average, four dwellings would be selected from each PSU. The allocation algorithm also took into account factors related to expected non-response at the second stage of sampling. In addition, to improve the efficiency of the sample design with respect to achieving an adequate number of dwellings with persons aged 18-24, oversampling of 10% was applied in strata with a significantly higher share of such dwellings.

[5] According to the revision of the Nomenclature of Territorial Units for Statistics – NUTS 2016, in force since 2018.

Secondary sampling units (SSUs) were dwellings. From each selected PSU, addresses were drawn by simple random sampling without replacement, in accordance with the pre-determined allocation. As a result, the final sample for the survey consisted of 6,646 PSUs and 26,148 dwellings.

The distribution of PSUs and dwellings across regions, as well as the total number of dwellings in the population and the number of strata created, is presented in Table 1.

Table 1. Distribution of the sample and population by region

Specification	Number of strata	Number of dwellings		Number of primary sampling units PSUs
		in the population	in the selected sample	
Poland	**385**	**15310345**	**26148**	**6646**
Dolnośląskie	28	1275612	1957	550
Kujawsko-Pomorskie	22	781378	1358	324
Lubelskie	23	788508	1256	321
Lubuskie	20	394808	899	293
Łódzkie	23	1040008	1617	392
Małopolskie	26	1295569	1671	436
Warszawski Stołeczny	24	1595460	2417	612
Mazowiecki Regionalny	22	837144	1371	317
Opolskie	20	360929	934	290
Podkarpackie	21	702149	1040	253
Podlaskie	22	460724	1002	296
Pomorskie	23	951638	1747	424
Śląskie	24	1813957	3415	732
Świętokrzyskie	19	455007	1097	322
Warmińsko-Mazurskie	22	534482	1074	289
Wielkopolskie	23	1326999	1889	435
Zachodniopomorskie	23	695973	1404	360

All households and persons aged 18-69 living in the dwellings selected for the sample should be surveyed (details on the rules for selecting persons within a household are provided in Subchapter 2.1).

2.3. Weighting procedure

In sample surveys, only a small part of the population is observed, and the information obtained from households and individuals is generalised to the entire population. Appropriate weights are used to generalise the results, allowing estimates for the entire population to be produced. The design weight (base weight) is defined for each selected dwelling as the inverse of the probability of selection. It is equal to the ratio of the number of dwellings in a given stratum in the sampling frame to the product of the number of PSUs selected from that stratum and the number of dwellings selected per PSU.

The weighting procedure took into account the following factors:

1) probability of selection of dwellings;
2) household response rate;
3) individual response rate.

These factors were applied in the initial adjustment of the design weights.

The design weights for dwellings (and households) were first adjusted to account for non-response due to refusals, temporary absence of household members, inability to contact the selected dwelling, etc.

In the Sampling Frame for Social Surveys, additional information from administrative registers is available for the sampled dwellings. This information was used to refine the algorithm previously applied at this stage of the weighting process. In the classical approach, the adjustment for non-response would be carried out using household and individual response rates, computed separately for locality size classes within voivodships or regions. An extension of this classical approach involved the use of machine learning algorithms to construct a mathematical model of the relationship between a predefined set of auxiliary variables and the probability of participation in the survey, referred to as the propensity score. The predicted participation probabilities generated by the model for the surveyed dwellings were then used to adjust the design weights. The simplest form of adjustment consists of multiplying the design weights by the inverse of the predicted participation probabilities. A commonly used practical solution is to apply the adjustment within groups defined on the basis of the distribution of the predicted probabilities. Such stratification reduces excessive variability of weights after adjustment. In the case of the AES, once the model was estimated, the dwellings were divided into five groups defined by propensity score quintiles.

Using adjustment factors, the design weights were adjusted for each NUTS 2 region, separately within each of the five propensity score groups. The adjustment factors were calculated as the ratio of the sum of design weights for eligible dwellings to the sum of design weights for dwellings in which an interview was completed. The propensity scores used to create the five groups were obtained from a machine learning model (using the h2o_automl function from the lares package in R). In this model, the following characteristics available in the sampling frame were used as predictor variables: the NUTS 2 region, locality size class, whether persons aged 18-69 (including 18-24) were present in the dwelling, the number of employed persons, the income decile group (based on tax records), etc. The final model resulted from an automated procedure for testing multiple candidate models with optimal parameter selection. The candidate models included well-established machine--learning classification algorithms, such as ensemble decision-tree models (including Gradient Boosting Machines as well as Random Forests and Extremely Randomised Forests) and Generalised Linear Models. The final model was a Gradient Boosting Machine and was selected due to the highest value of the AUC, used as a measure of classification performance. The AUC is defined as the area under the ROC curve, which illustrates the relationship between the true positive rate and the false positive rate. In these calculations, all possible propensity score thresholds (ranging from 0 to 1) are taken into account.

The values presented in Table 2 show substantial variation in participation propensity scores across locality size classes, which justifies the use of this classification in the construction of strata. Similar variation within locality size classes is observed for response rates; therefore, in the classical adjustment algorithm, response rates are used to adjust the design weights separately for each locality size class (with possible additional stratification by voivodship or region). This approach was applied in the previous edition of the survey (AES 2016).

Table 2. Household participation propensity scores by locality size class

Locality size class	Poland	Participation propensity at the regional level	
		minimum	maximum
Total	**0.434**	**0.354**	**0.504**
Warsaw	0.349	•	•
Cities above 500 thousand inhabitants [a]	0.392	0.365	0.423
Cities from 100 thousand to 500 thousand inhabitants	0.389	0.286	0.411
Cities from 20 thousand to 100 thousand inhabitants	0.418	0.339	0.465
Cities up to 20 thousand inhabitants	0.436	0.340	0.523
Rural areas	0.494	0.380	0.553

a The Tri-city agglomeration is also included in this class. Data for this class do not include Warsaw.

Aggregate indicator values presented in Table 3 result from the adopted method of dividing the data into five groups according to the quintiles of the distribution of propensity scores. The largest interval widths defining the participation propensity groups occur for the extreme groups – Group 1 and Group 5.

Table 3. Household participation propensity scores by five propensity groups

Group	Participation propensity	Poland	Participation propensity at the regional level	
			Minimum	Maximum
1	[0.045; 0.358]	0.219	0.182	0.258
2	(0.358; 0.443]	0.409	0.397	0.428
3	(0.443; 0.482]	0.463	0.458	0.469
4	(0.482; 0.532]	0.504	0.489	0.510
5	(0.532; 0.623]	0.575	0.538	0.591

At the next stage of the weighting adjustment process, missing individual interviews were taken into account. The individual response rate was 65%, with the lowest rate recorded for persons aged 18-24 (46%). The adjustment of weights (from the previous stage of adjustment for households) for missing individual interviews was carried out in the following cross-classifications: urban/rural * sex * (persons aged 18-24 and nine 5-year age groups). After adjustment for non-response, adjusted weights for individual interviews were obtained.

These weights were then calibrated[6] using data from current population estimates for sex, age group coded as "1" (the 18-24 age group and nine 5-year age groups), age group coded as "2" (persons aged 18-24 and four 10-year age groups), place of residence (urban or rural areas), and region (NUTS 2). The weights were calibrated according to the marginal distributions in the following breakdowns:

- urban/rural * sex * age groups "1";
- NUTS 2 regions * urban/rural * age groups "2";
- NUTS 2 regions * sex * age groups "2";
- NUTS 2 regions * urban/rural * age groups "1".

[6] A detailed description of this method can be found i.a. in the following publications: Health Status of the Population of Poland in 2009, Statistics Poland, Warsaw 2011; M. Szymkowiak, T. Józefowski, Calibration Estimators in Statistical Surveys, Statistical News (1), pp. 31–43, Statistics Poland, Warsaw 2012.

The calibrated weights were calculated using a dedicated programme in R. One of the calibration algorithms described by Williams and Savitsky (2019) was used. A variant of the calibration algorithm providing range limits for the weights was used, i.e. the range of the resulting weights relative to the pre-calibration weights was maintained within the range from 1/3 to 3.

In addition to the standard weights, two additional types of weights are calculated and used in the estimation of certain indicators related to non-formal education. This results from the methodological framework adopted for this form of learning. The collected information may relate to a maximum of five educational activities within non-formal education, of which two are covered in greater detail. When a respondent reports participation in a larger number of educational activities, the survey methodology requires that the activities selected for further data collection be selected at random. Therefore, in the analysis of the survey results, the overall educational activity of respondents is taken into account, expressed as the number of educational activities undertaken. The European Commission's implementing regulation specifies two additional types of weights used in the calculation of certain indicators related to non-formal education:

1) NFEACTWEIGHT_5 is used to calculate indicators based on five randomly selected non-formal education activities;
2) NFEACTWEIGHT_2 is used to calculate indicators based on two randomly selected non-formal education activities.

These additional weights are obtained by multiplying the individual weights by the total number of non-formal educational activities declared by the respondent and dividing by the number of selected activities relevant to a given indicator (i.e. from 1 to 5 or from 1 to 2).

3. Data collection tools and data sources

3.1. Data collection tools

The *Adult Education Survey* (AES) was conducted using two questionnaires:

- The *household questionnaire*, intended for each household in the sampled dwelling;
- The *individual questionnaire*, completed for the selected respondent (aged 18-69).

These questionnaires were available both in paper form, used when conducting interviews using the PAPI method, and in electronic form, as a mobile application designed for data entry on mobile devices (smartphones, tablets, or laptops) during interviews conducted using the CAPI method, as well as on desktop computers. The application was also used to conduct telephone interviews.

More information on the scope of the questionnaires can be found in Subchapter 4.1.

The household and individual questionnaires used in the 2022 wave of the AES are presented in Appendices 3 and 4.

To ensure the correct completion of the questionnaires – from both technical and methodological perspectives (i.e. proper interpretation of the questions) – guidelines for completing the questionnaires were prepared. These were accompanied by a set of appendices including a list of symbols used in the survey, particularly for variables based on classifications and used in questions requiring coded responses. The paper version of the questionnaire also included a dedicated appendix containing definitions, a list of example educational activities, and explanations of the main forms of learning within non-formal education.

3.2. Data collection methods

In accordance with Eurostat methodological guidelines, the preferred approach is face-to-face interviewing, particularly computer-assisted personal interviewing (CAPI), in order to eliminate errors and inconsistencies already at the fieldwork stage. The individual questionnaire of the Adult Education Survey contains complex and detailed questions. Face-to-face interviews help reduce measurement errors – i.e. errors resulting from misunderstanding a question, concept, or definition – because the interviewer assists the respondent in completing the questionnaire. In line with guidelines on interviewer work organisation, the survey in Poland was conducted using a mixed-mode approach comprising three data collection methods: computer-assisted telephone interviewing (CATI), computer-assisted personal interviewing (CAPI), and paper-and-pencil personal interviewing (PAPI). Most interviews were carried out using CATI (60.6%) and CAPI (35.6%). The introduction of telephone interviews (CATI) in the AES 2022 survey was a response to unforeseen external factors (including the COVID-19 pandemic caused by the SARS-CoV-2 virus) and was also intended to facilitate access to specific groups of respondents (young, mobile, and economically active persons). An additional benefit was the reduction of survey costs and interviewer workload. In exceptional cases resulting from circumstances beyond the interviewer's control (e.g. inability to use the tablet due to technical or weather-related reasons), it was possible to complete a paper questionnaire (PAPI) and subsequently enter the data into the CAPI application (3.8%). Interviewers were equipped with electronic devices designed for conducting surveys using an electronic questionnaire, performing data checks and transmitting validated data to the server while ensuring full compliance with data security requirements. The data collection software was developed using Progressive Web Application (PWA) technology. This made it possible to run the application on any type of device (desktop computer, laptop, tablet, or smartphone) with any operating system (e.g. Android, Windows, Linux, iOS). Data entry was performed via a mobile application. The software allowed operation in CATI and CAPI modes (entering data directly during the interview) or in PAPI mode (entering data from paper questionnaires). All electronic devices used with the mobile application were synchronised with the "Survey Management System (CORStat)", enabling continuous data transmission to the server.

3.3. Direct and proxy interviews

In its methodological guidelines, Eurostat strongly recommends that answers be provided directly by the selected person rather than by another member of the household. The AES focuses on issues related to participation in formal and non-formal education as well as informal learning. These topics are closely related to the respondent's personal experiences and attitudes toward learning. A relative or another household member may have difficulty providing reliable information about the respondent's educational activities (particularly in the case of education and training that took place at the respondent's workplace). Moreover, recalling all educational activities undertaken during the last 12 months may already be challenging for the respondent themselves, and for another member of the household it is virtually impossible. However, in the Polish version of the survey, the use of proxy interviews was allowed under certain conditions, whereby another household member could provide answers on behalf of an absent respondent. The introduction of this solution was intended to reduce non-response. On the other hand, proxy interviews may negatively affect data quality for certain variables. Therefore, the interviewer instructions restricted the use of proxy interviews for the individual questionnaire. Proxy interviews could be conducted only in exceptional cases, for example if household members were able to provide reliable information that the respondent had not participated in any form of learning for objective reasons, such as health-related reasons. The share of proxy interviews in the AES survey amounted to 9.6%.

4. Variables covered in the survey

4.1. Description of the main variables

The overall scope, content, and structure of the questions included in the *Adult Education Survey* questionnaires are based on the set of variables defined in the European *Adult Education Survey* (AES). The AES 2022 questionnaires did not include additional variables collected solely for national purposes. The optional set of questions concerning COVID-19 was not included. The Polish questionnaire differed slightly from the model questionnaire proposed by Eurostat in terms of the order of questions in some modules, coding, and certain response categories.

Therefore, at the data processing stage it was necessary to recode certain variables so that the dataset transmitted to Eurostat met the formal requirements concerning response codes and permitted value ranges (assumptions for variable conversion are provided in Appendix 5). The main differences between the Eurostat model questionnaire and the national survey questionnaires include:

- in the national survey, two separate questionnaires are used: a household questionnaire and an individual questionnaire. These were designed to adapt the Eurostat model questionnaire to the specific characteristics of surveys based on a sample of dwellings. Information on the number and age structure of household members (HHNBPERS) is derived from the household composition reported in the household questionnaire;
- the variable on self-declared labour market status (MAINSTAT) was introduced into the questionnaire as two separate questions to enable respondents to specify their labour market status more precisely;
- the order of questions on parents' education and country of birth was modified – these questions were placed before the module on the respondent's labour market status;
- the order of questions in the "Formal Education" module was modified. This concerned the variables specifying the month and year of starting education and the labour market status at that time (FEDSTARTYEAR, FEDSTARTMONTH, FEDMAINSTAT), which were placed after the variable on education level (FEDLEVEL), while the question on completion of education (FEDCOMP) was placed after the variable on labour market status at the start of education (FEDMAINSTAT).

Variables Included in the household questionnaire

Apart from technical variables and information on the reason for non-interview, included at the beginning of the questionnaire, the main part of the household questionnaire is divided into three sections:

- section 1 – collects information on household members, both present and absent, whose total period of absence (actual and intended) is less than one year. The variables describe the household composition and include demographic and social characteristics of individual household members (including date of birth, sex, citizenship, and basic migration-related information), as well as additional information about participation in the survey;
- section 2 – information on the type of household, including the variable HHTYPE describing the household composition;
- section 3 – information on the monthly net income of the household received in the month preceding the survey month. The variable HHINCOME represents the total income of all household members, whether received individually or jointly.

At the end of the questionnaire, supplementary information is recorded, including the duration of the household interview and contact details for conducting a CATI interview.

Variables Included in the individual questionnaire

Apart from the technical variables included at the beginning of the questionnaire (general information in Section 1), which are necessary for conducting the survey and linking the individual questionnaire with the information collected in the household questionnaire, the main part of the individual questionnaire is divided into the following sections:

- de facto marital status – the variable MARSTADEFACTO indicating whether the respondent lives with a partner based on the actual living arrangement resulting from living in the same household;
- education – information on level of educational attainment and characteristics of the respondent's education (including the field of education and the year the qualification was obtained, or information on incomplete education);
- parents – variables describing the educational attainment and the country of birth of the respondent's mother and father;
- labour market situation – self-declared labour market status and characteristics of employment, occupation, and the main job for employed persons (including the economic activity and the size of the enterprise or institution);
- access to information on opportunities to participate in formal and non-formal education and the use of educational guidance services – information on the level of interest among adults in learning possibilities in formal and non-formal education, as well as institutional support received in the field of educational guidance. This includes activities related to the provision of information, counselling, skills assessment, support, and training in decision-making and career management. The module also contains questions on the institution providing such support and the way in which it was obtained;
- formal education – information on the respondent's participation in formal education during the 12 months preceding the interview, including the level and field of education, organisation and financing of the education, reasons for participation and benefits obtained, as well as time spent on learning;
- non-formal education and training – variables concerning the respondent's participation in non-formal education and training during the 12 months preceding the interview, including the level and field of learning activities, reasons for participation, benefits and skills acquired, as well as the organisation, financing arrangements, costs incurred, and time spent on learning. The module also includes questions on the relationship between non-formal education and employment or personal interests, who initiated participation, who provided the training, and how the acquired qualifications were certified or recognised;
- obstacles to participation in education and training – information on the reasons for non-participation in formal or non-formal education and training, as well as the barriers encountered in accessing these forms of learning;
- informal learning – variables concerning the respondent's participation in informal learning activities, including the methods of informal learning and the purpose of undertaking such learning (whether related to employment or personal interests);
- knowledge of foreign languages – information on the respondent's mother tongue as well as their proficiency in foreign languages;
- general health status – variables concerning the respondent's self-perceived health status, as well as variables enabling the assessment of whether and to what extent health problems limit the performance of usual daily activities.

4.2. Definitions of the basic terms used in the AES

A *household* is a group of people related by blood or marriage or not related, living together in the same dwelling and jointly maintaining a household. If any of the persons living together maintains a separate budget, he/she constitutes a separate one-person household. The members of a household include persons who are:

- present in the household (registered for permanent or temporary residence, staying or intending to stay without registration for a period of 12 months or more);
- absent (taking into account the total duration of actual and planned absence) for a period of up to 12 months (e.g. persons temporarily staying abroad, in collective accommodation establishments or in another household within the country for less than 12 months).

The survey also covers foreign nationals who are members of households in the sampled dwellings, provided that they meet the above criteria.

The *main labour market status* was determined according to the respondent's self-assessment, identifying the actual rather than the formal employment situation of the surveyed persons. The definitions adopted in this respect were as follows:

Employed persons are those performing work generating pay or income or contributing family workers. The category comprises:

1**.** *Self-employed persons,* including:

- owners, co-owners and leaseholders of private farms in agriculture working on these farms. Work performed by the respondent on a farm was included if it generated income that was the respondent's main or additional source of income;
- members of agricultural production cooperatives;
- agents working under any agency system;
- persons who conducted their own economic activity (registered or unregistered) outside agriculture, i.e. persons who ran their own business or production or service establishment, shop or earned money by operating their own taxi, liberal professionals (e.g. lawyers, writers, artists, doctors and nurses with private practice), as well as persons running private schools, nursery schools, persons giving private lessons directly paid by the students or their parents, etc.; married couples or related persons (e.g. fathers and sons, siblings) working together (on equal terms) on their own account were treated as co-owners and all these people were classified as "self-employed";
- persons who were persuaded by their employer to register their own business as a sole proprietorship;

1a. *Self-employed persons with employees* are persons working in their own business, professional practice or farm for the purpose of earning a profit, who employ at least one person. If the people employed in that business, practice or farm did not receive a salary, then the worker is referred to as a self-employed person without employees;

1b. S*elf-employed persons without employees* are persons working in their own business, professional practice or farm for the purpose of earning a profit, who do not employ any other person. These people may involve family members or apprentices who do not receive remuneration for their work. This category includes farmers who run their own farms with the help of family members;

2. *Employed persons (employees*), including:

- persons employed under an employment relationship (employment contract, contract of mandate, specific--task contract, posting, appointment, election) in public sector units, cooperatives, social organisation units, political units and trade unions, in private units, at businesses run by natural persons and private farms in agriculture;
- outworkers;
- apprentices and trainees receiving remuneration, who have a contract for apprenticeship or on-the-job training;
- persons employed in a family farm or family business receiving a regular remuneration or income for the work performed (in cash or in kind) - regardless of the connection with the household of the farm user or the owner of the family business;

3. *Contributing family workers (unpaid)* are persons who help another family member in running a farm or other business without being classified as employed persons. Persons working in a company or on a family farm without remuneration should share the same place of residence with the business or farm owner, or, in a slightly broader sense, should live in a building located on the same plot and be a common household member. These persons often receive remuneration in the form of additional benefits and payments in kind. This applies where the owner or manager is the person concerned or a family member. This category excludes work performed on a voluntary basis for charitable organisations.

Occupation in the main job is defined in accordance with the Classification of Occupations and Specialisations developed on the basis of the International Standard Classification of Occupations ISCO-08. It concerns the main job currently performed by the employed person. For persons who had several jobs, the main job is the one that usually occupies the greatest amount of time. If the jobs take up the same amount of time, the main job is the one that brought the higher income. Self-employed persons as owners of small, often single-person service establishments: shops, bakeries, tailors, etc. were classified not as managers but according to their profession, i.e. as salespersons, bakers or tailors.

Persons not in employment, or those who do not consider employment to be their main labour market status, include the following categories:

- *Unemployed persons*, defined as persons without employment who are seeking work and available to start work. This category could also include persons who retired early for economic reasons (according to the respondent's perception), if the early retirement was caused by company closure, downsizing, etc. Respondents may classify themselves as unemployed irrespective of their official status or registration at the Labour Office. Unemployed persons could also have minor jobs while looking for a main job;
- Economically i*nactive persons*, i.e. respondents who are neither classified as employed nor as unemployed. The following categories of inactive persons were distinguished:
  - *Persons in education or training, or participating in unpaid traineeships or internships* – persons in education or training, including students participating in unpaid internships. Respondents who participate in various forms of vocational education or training (partly) held at the workplace may classify themselves as employed. This category also includes apprentices and paid trainees;
  - *Retired persons or persons in early retirement* – persons receiving pensions, early retirement pensions or structural pensions. If early retirement was due to economic reasons (company closure, downsizing, etc.), the status is classified according to the respondent's self-perception (retired or unemployed);
  - *Persons receiving disability benefits/persons permanently unable to work* – persons who have not reached retirement age but are unable to work due to health reasons or disabilities for a long or indefinite period. Respondents may select this category even if they did not receive any benefits (including pension). Persons receiving social pensions, training benefits or rehabilitation benefits are included in the category of "economically inactive for other reasons";
  - *Economically* i*nactive persons due to family responsibilities* – persons engaged in domestic tasks, persons on parental leave, providing unpaid care for family members. Respondent may select this category even if they have a job, receive a pension, are retired or are engaged in voluntary work;
  - *Economically inactive for other reasons* – a category referring to persons not working for reasons other than those listed above. This category includes persons receiving survivors' pension, social pensions, training benefits or rehabilitation benefits, as well as persons engaged in voluntary work.

Formal education, according to the definition presented in the Act of 22 December 2015 on the Integrated Qualifications System, is the education provided by public and non-public schools and other entities of the education system, higher education institutions and other higher education and research institutions, within the framework of programmes which lead to obtaining full qualifications awarded on completion of specific stages of education, qualifications awarded on completion of postgraduate studies, or qualifications in a profession, confirmed by a certificate of professional competence or a professional diploma. In the *Adult Education Survey*, it covers learning in the school system (at primary, secondary and post-secondary levels), as well as education at tertiary, postgraduate and doctoral levels and in the form of Vocational Qualification Courses. A Vocational Qualification Course should be understood as a course based on a programme which includes the core curriculum of vocational education in the scope of one qualification, the completion of which leads to vocational examination in the scope of this qualification. Formal education is conducted in accordance with approved curricula. The teaching is based on the standards adopted in the education system and higher education, in compliance with applicable legislation. Formal education leads to obtaining the qualifications listed in the Integrated Qualifications Register and validated by a certificate or diploma from an authorised institution. Formal education is characterised by the presence of a person conducting classes (teacher/lecturer/trainer).

Non-formal education, pursuant to the Act of 22 December 2015 on the Integrated Qualifications System, can be defined as education and training carried out as part of programmes that do not lead to obtaining full qualifications, qualifications awarded on completing postgraduate studies, or qualifications in a profession, validated by a vocational qualification certificate or a diploma. In the *Adult Education Survey*, non-formal educational programmes lasting less than one semester or half a year of full-time teaching, such as "vocational skills course", "general competences course" or "theoretical training course for young employees", were also included in non-formal education. Non-formal education comprises all organised/institutionalised educational activities that do not correspond to the definition of formal education. Such activities are usually conducted in the form of courses or instruction (at the workplace or externally), seminars or lectures to which the respondent registered and in which he/she participated. This type of education also includes private lessons (e.g. foreign languages) as well as "distance" learning which takes place via postal correspondence or electronic means (e.g. computer, video, DVD). Non-formal education does not change the level of educational attainment, but usually leads to the development, expansion and acquisition of skills in various areas of professional, social and cultural life. Unlike informal learning, non-formal education should involve the participation of a lecturer, instructor, teacher or tutor. Classes may take the form of courses or training to improve professional qualifications, they may lead to obtaining skills needed in social life (e.g. self-defence course) or for personal benefits, one's own intellectual development or hobby (e.g. cooking course, art history lectures, modelling course). Participation in such an educational activity may or may not result in obtaining a certificate.

Courses at the workplace or during free time - a planned series of educational activities in a specific field, offered by an educational provider. They are typically subject-oriented and taught by one or more instructors who specialise in a given field. Courses can be conducted in a classroom (on-site) or using distance learning techniques (online). Online courses, like on-site courses, may include such elements as curriculum, enrolment, teaching and tests, but are conducted by correspondence or via electronic media connecting instructors, teachers or trainers with learners who are not physically present in the same classroom (e-learning).

Workshops and seminars at the workplace or during free time – a type of training combining elements of practice and instruction with theory. These may have the form of a course combining theoretical instruction (classroom instruction) with practice in real or simulated situations. *A workshop* is a form of classes during which participants actively take part in a meeting. This type of training is based on the exchange of experience and confronting it with the trainer's expertise. Workshops are intended to teach a specific skill at a specific time. The training is scenario-based, taking into account the tasks and goals to be achieved. *A seminar* involves active participation of trainees who independently develop some of the issues discussed at the seminar and then present their work in the form of a presentation, lecture or in another way, as well as contribute to discussions on a given issue, demonstrating their knowledge.

Guided on-the-job training – a form of training with the presence of a tutor, organised by the employer for one person or a small group. This form of non-formal education is characterised by planned periods of training, instruction or practical experience using typical tools of work, directly in the workplace or in a work-related situation, with the presence of a tutor. The training takes place with the participation and supervision of a person with appropriate qualifications (employer, manager, instructor). The training is individual and practical. It is usually organised by the employer to support the adaptation of (new) employees, including transferred, re--employed and seasonal or temporary workers, to new or existing jobs. It may include general training about the company (organisation, operating procedures, etc.) and specific job-related instructions (safety and health risks, work practices).

Private lessons taught by a teacher or tutor. The survey did not include lessons given by colleagues, friends or family members, except when they were paid lessons.

Informal learning, according to the definition contained in the Act of 22 December 2015 on the Integrated Qualifications System, is the acquisition of knowledge, skills and social competences through various types of activities outside formal and non-formal education. It includes activities undertaken to learn something independently and the learning process that accompanies other activities (work, participation in various events or recreational activities). Unlike formal and non-formal education, informal learning should take place without the participation of a tutor or instructor. It is not institutionalised, i.e. it does not involve courses, workshops, seminars, private lessons, etc. Lessons taught free of charge by family members, household members, neighbours or friends are considered a type of informal learning. The scope of the survey covered only educational activities undertaken with the intention of acquiring or expanding knowledge and skills. Learning could not be incidental ("by the way"). They had to be planned in advance, the objectives set, as well as expected effects of the undertaken educational activities. Questions about informal learning did not concern additional activities that usually accompany formal or non-formal education, such as self-study or doing homework.

Educational activity is an activity undertaken with the aim of improving or acquiring knowledge, skills or competences. A single learning activity is characterised by a combination of a specific method and one field of education. *Method* is understood as the organisational framework used for learning or teaching (i.e. acquiring or transmitting ideas, information, knowledge, skills and competences). It may take the form of an educational programme (school, field of study), a course or another more or less organised undertaking. *A field of education* is a subject area to which the learning activity relates, defined on the basis of the classification of fields of education. The classification of fields of education used in the survey was developed on the basis of ISCED-F 2013 (International Standard Classification of Fields of Education and Training), primarily using the broad field groupings defined in this classification.

Language proficiency levels were defined in accordance with the Common European Framework of Reference for Languages (CEFR). These descriptors enable respondents to self-assess their ability to communicate in a given language. The lowest level of language competence – "almost no ability to use the language" – is not covered by the CEFR classification. It was introduced in the survey in order to distinguish a category of persons with very limited language skills. The following language proficiency levels were distinguished:

1. Below A1 -almost no ability to use the language – the respondent understood and was able to use only a few isolated words and expressions;
2. Basic User (A1–A2) – the respondent could communicate in a simple way provided that the interlocutor was willing to repeat what is said slowly, simplify them or provide assistance. The respondent was able to ask and answer simple questions on familiar topics or matters of immediate need;
3. Independent User (B1–B2) – the respondent was able to deal with most situations likely to arise while travelling in an area where the language is spoken. The respondent could take part without preparation in conversations on topics related to personal interests or everyday life, such as family, leisure, work, travel and current events;
4. Proficient User (C1–C2) – the respondent was able to communicate fluently and spontaneously without much obvious searching for expressions. The respondent could use the language effectively and flexibly for social and professional purposes and express ideas and opinions clearly, relating them to those of other speakers.

4.3. Main indicators calculated based on AES results and their definitions

The key indicators for Poland and other countries participating in the AES are presented in the Eurostat database:

1. *Adult participation rate in education and training* – the percentage of persons participating in formal education, non-formal education, formal and/or non-formal education in a given category, in the total population of that category (distinguished i.a. by age, sex, level of educational attainment, labour market status and occupation);
2. *Adult participation rate in job-related non-formal education* – the percentage of persons participating in job--related non-formal education (overall and broken down into employer-sponsored and not employer--sponsored non-formal education) within a given category, in the total population of that category (distinguished i.a. by age, sex, level of educational attainment, labour market status and occupation);

3. *Educational activities in which adults participate*, by *field of education* – the percentage of educational activities in a given field of education in the total number of educational activities. The indicator is calculated for formal education and non-formal education (overall and broken down into job-related and non-job-related education), by age group. In the calculation of the indicator, NFEACTWEIGHT_2 weights were used (see explanations in Subchapter 2.3);
4. *Non-formal educational activities in which adults participate, by type of provider* – the percentage of educational activities provided by a given education and training provider in the total number of educational activities. The indicator is presented by age group. In the calculation of the indicator, NFEACTWEIGHT_2 weights were used (see explanations in Subchapter 2.3);
5. *Non-formal educational activities in which adults participate, by motivation and employer contribution to costs* – the percentage of educational activities that are job-related (overall and broken down into employer-sponsored and not employer-sponsored), job-related educational activities employer-sponsored and undertaken by employed persons, and non-job-related educational activities in the total number of non-formal education activities. The indicator is presented by sex, age group and level of educational attainment. In the calculation of the indicator, NFEACTWEIGHT_5 weights were used (see explanations in Subchapter 2.3);
6. *Adults seeking information on learning possibilities, by participation in learning* – the percentage of persons (overall, including those participating in formal and/or non-formal education, those participating only in informal learning, and those not participating in any form of learning) who seek information about opportunities to participate in formal and/or non-formal education, in the total number of persons. The indicator is presented by sex, age group and level of educational attainment;
7. *Average number of instruction hours spent by adults on education and training* – the average number of hours spent by participants on formal and/or non-formal education, formal education, and non-formal education during the 12 months preceding the interview. The indicator is presented by sex, age group and level of educational attainment. In the calculation of the indicator for non-formal education, NFEACTWEIGHT_2 weights were used (see explanations in Subchapter 2.3);
8. *Distribution of instruction hours spent by adults on education and training by field of education* – the share of hours spent on formal and/or non-formal education (including separately formal and non-formal education) in a given field of education to the total number of learning hours. The indicator is calculated for each of these forms of education and presented by age group. In the calculation of the indicator for non-formal education, NFEACTWEIGHT_2 weights were used (see explanations in Subchapter 2.3);
9. *Adults by willingness to participate in education or training* – the percentage of persons expressing a given attitude towards their education (willingness or unwillingness to participate in education or training) in the total number of persons. Data are presented by age group;
10. *Adults not participating in education or training, by main reason* – the percentage of persons not participating in formal or non-formal education for a specified reason (obstacles encountered by type or lack of interest) in the total number of persons not participating in these forms of learning. Data are presented by sex, age group and level of educational attainment;
11. *Adults willing to participate in education or training, by obstacles encountered, including the obstacle indicated as the main one* – the percentage of persons willing to participate in formal or non-formal education who declare having encountered a specific obstacle (including the one indicated as the main obstacle) in the total number of persons expressing willingness to participate in education. Data are presented by sex, age group and level of educational attainment;
12. *Adults not willing to participate in education or training, by the main obstacle to undertaking education* – the percentage of persons not willing to participate in formal or non-formal education who declare having encountered a specific obstacle to such participation, in the total number of persons not expressing willingness to participate in these forms of learning. Data are presented by sex, age group and level of educational attainment;
13. *Adults knowing foreign languages* – the percentage of persons of a given category declaring knowledge of foreign languages in the total population of that category (distinguished i.a. by age, sex, level of educational attainment, labour market status and occupation), by number of languages known;

14. *Adults by level of knowledge of the foreign language* – the percentage of persons of a given category declaring knowledge of foreign languages at a given level (including the language selected as the best known) in the total population of that category (distinguished i.a. by age, sex, level of educational attainment, labour market status and occupation) declaring knowledge of foreign languages.

4.4. Classifications

The classifications used in the *Adult Education Survey* (AES) in 2022 are as follows:

1. AES data on level of education and educational attainment are presented according to the *International Standard Classification of Education (ISCED 2011)*;
2. Information on the field of education and training is presented according to the International Standard Classification of Education – Fields of Education and Training (ISCED-F 2013);
3. Forms of learning are defined in accordance with *the Classification of Learning Activities (CLA 2016);*
4. AES data on occupation are compiled on the basis of *the Classification of Occupations and Specialisations* (KZiS 2014). This classification was introduced by the Regulation of the Minister of Labour and Social Policy of 7 August 2014 (consolidated text: Journal of Laws 2018, item 227, as amended);
5. Information on the economic activity of the local unit is presented according to the *Polish Classification of Activities* (PKD 2007), introduced by the Regulation of the Council of Ministers of 24 December 2007 (Journal of Laws No. 251, item 1885, as amended) on the Polish Classification of Activities. PKD 2007 was developed on the basis of the *Statistical Classification of Economic Activities in the European Community (NACE Rev. 2);*
6. AES data in territorial breakdowns are presented in accordance with the statistical territorial division of Poland, based on the *Nomenclature of Territorial Units for Statistics* (NUTS 2016);
7. *ISO 3166-1* provides country codes used to identify data on the citizenship and country of birth of the respondent and their parents;
8. The mother tongue and foreign languages known by the respondent are coded according to *ISO 639 – language codes*;
9. The basis for determining levels of language proficiency is provided by the guidelines set out in the *Common European Framework of Reference for Languages (CEFR).*

Information on the methodology of the Adult *Education Survey* (AES) for 2022, including the definitions used and references to international classifications applied in the European *Adult Education Survey* (AES), is available at: Adult Education Survey (AES) methodology – Statistics Explained (europa.eu)

5. Organisation and management of survey implementation

5.1. Official statistics units conducting the survey

The preparation and implementation of the *Adult Education Survey* (AES) is coordinated by the Statistical Office in Gdańsk – Centre for Education and Human Capital Statistics. The survey is prepared in cooperation with the following units of Statistics Poland: the Programming and Survey Coordination Department, the Education and Communication Department, the Labour Market Department, and the Social Surveys Department, as well as the Ministry of Education and Science.

At the regional level, the survey is carried out by the relevant statistical offices. IT support for the survey is provided by the Section for Data Collection Technologies and Systems of the Statistical Computing Centre – Radom Branch and the Centre for Data Engineering at the Statistical Office in Gdańsk.

5.2. Organisation of data collection

Data collection within the survey is preceded by the preparation of appropriate survey instruments (described in Chapter 3) and by the creation of a register containing the addresses of dwellings in which the households residing there are to be surveyed. The preparation of the survey instruments was carried out by the Centre for Education and Human Capital Statistics (OEK) at the Statistical Office in Gdańsk in cooperation with the Social Surveys Department and the Labour Market Department of Statistics Poland. The responsible unit (OEK) also developed the specifications for the Section for Data Collection Technologies and Systems of the Statistical Computing Centre – Radom Branch to prepare the data collection application using the CAPI and CATI methods, as well as the organisational guidelines for the survey. The survey register was created by CIS on the basis of the list of dwellings selected for the survey (sample) prepared in accordance with the assumptions of the responsible unit by the Programming and Survey Coordination Department (PK) of Statistics Poland. The AES sample is drawn from the Social Surveys Sampling Frame, maintained by Statistics Poland in accordance with the provisions of the Act of 29 June 1995 on Official Statistics and used as the basis for drawing samples for household survey programmes.

Data collection in the selected dwellings is carried out by interviewers – employees of statistical offices. Contact between the interviewer and the respondent (either face-to-face or by telephone) is preceded by a letter from the President of Statistics Poland informing about the purpose of the AES and about the fact that the dwelling has been selected for the survey. In addition, in order to encourage participation in the survey, the advance letter is accompanied by a brochure presenting the results of the previous survey round. Information on the launch of the survey, together with a short description, is also published on the Information Portal of Statistics Poland.

Data collection is carried out using the data collection modes described in Subchapter 3.2 of the AES methodology.

6. Presentation of the survey results

6.1. Main results obtained from the AES

The AES enables a comprehensive assessment of participation of persons aged 18-69 in Poland in various forms of learning. It provides information on participation of adults in formal education, non-formal education, and informal learning, as well as on access to information on learning possibilities, and on the level and distribution of foreign language skills. The survey covers the entire target population, which is further characterised by demographic, social and economic characteristics, such as:

- sex;
- age;
- place of residence (urban/rural area, voivodship);
- the respondent's level and field of educational attainment, as well as the educational attainment of their parents;
- type and income of the household;
- health status and the impact of health problems on limitations in performing daily activities;
- labour market status;
- employment status, economic activity, number of employees, and occupation in the main job (for the population group identified in the AES as employed persons).

AES results also make it possible to characterise the population of adult participants in formal and non-formal education according to the following characteristics:

- reasons and motivation for undertaking education;
- time and place of learning;
- sources of funding for education;
- level (in the case of formal education) and field of education and training;
- learning methods and forms;
- the initiator of participation in non-formal education;
- benefits obtained from participation in education;
- costs (in the case of non-formal education) and time spent on learning;
- the provider of education and training services (in the case of non-formal education);
- the form of recognition or certification of acquired qualifications (in the case of non-formal education).

Presentation of results

Survey results are disseminated through publications issued in the form of news releases and statistical publications. Titles from the most recent edition include:

- News release: *Adult Education in 2022*;
- Publication: *Adult Education 2022*;
- Publication: *Human capital in Poland in the years 2020–2024.*

AES results are also available in the SDG database – *Sustainable Development Goals* maintained by Statistics Poland.

6.2. Restrictions regarding the publication of AES results

When analysing and disseminating data from the AES, it is necessary to take into account the limitations resulting from the sample-based survey design. Due to this methodology, the survey cannot provide reliable data for highly detailed population breakdowns. A small number of observations may lead to large sampling errors.

Accordingly, the following restrictions apply to the publication of AES results:

- the lowest administrative level at which core data on the number of persons participating in different forms of learning (formal education, non-formal education and informal learning) are disseminated is the voivodship. In terms of the Nomenclature of Territorial Units for Statistics (NUTS), the lowest level for data dissemination is NUTS 2 regions;
- data on employed persons by occupation are published at the level of major groups (1-digit level according to the Classification of Occupations and Specialisations);
- data on economic activity according to PKD (NACE Rev. 2) are presented at the level of sections (1-digit level).

Due to the limitations inherent in the sample survey design, AES results are presented rounded to the nearest thousand persons. In addition, caution is recommended when using data for more detailed breakdowns where the phenomenon is based on at least 20 but fewer than 51 observations in the sample. For data at the voivodship level, an additional precision criterion has been applied. Values for which the relative standard error (coefficient of variation, CV) for the estimated totals ranges between 10% and 20% should also be interpreted with caution. In assessing the precision of AES results, the item non-response rate for individual variables was also taken into account. All estimates based on variables with a non-response rate between 20% and 50% have been flagged as potentially subject to high sampling error. This applies in particular to household income data (HHINCOME variable) and costs related to participation in non-formal education (NFEPAIDVAL1 and NFEPAIDVAL2 variables). All estimates based on fewer than 20 observations have not been published. Furthermore, in the presentation of results at the voivodship level, data for which the relative standard error exceeds 20% have also been excluded from publication due to very high sampling variability. Limitations in the interpretation of data, together with the associated sampling errors calculated for key survey indicators, are described in detail in Subchapter 7.2.

6.3. Principles of access to AES data

The individual data and personal data collected and compiled within the survey [7] are confidential and subject to special protection. These data may be used exclusively for the preparation of statistical outputs, tabulations and analyses, and their disclosure or use for purposes other than those specified above is prohibited (statistical confidentiality). Personal data in the AES are processed only at the data collection stage. They are necessary for organisational purposes, enabling contact with respondents and, consequently, the implementation of the survey. After the data have been collected by interviewers using the survey questionnaires, the dataset is stripped of variables allowing the identification of individual persons and is subsequently processed as a set of non-identifiable microdata.

AES data are disseminated primarily in the form of aggregated statistical results. However, the provision of microdata is permitted upon individual request. Non-identifiable microdata sets are made available by Statistics Poland, on a fee basis, for scientific purposes to higher education institutions and other organisational units of the higher education system, as well as to research institutes conducting scientific studies. Data sets prepared for users are subject to confidentiality checks to prevent the disclosure of any confidential information and, in order to ensure compliance with statistical confidentiality requirements, may be subject to anonymisation.

[7] Information on the processing of personal data by Statistics Poland in accordance with Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, is available on the website: https://bip.stat.gov.pl/rodo.

6.4. Structure of the microdata files

Microdata from the *Adult Education Survey* are stored in six data files (Appendix 5). The data files listed below contain information on:

1) dwellings selected for the survey;
2) households, collected using the household questionnaire;
3) individual persons, including basic demographic and socio-economic characteristics (such as age, citizenship, country of birth and length of residence in Poland) from the household questionnaire;
4) individual persons included in data file3, with additional variables corresponding to Sections 1-7 of the individual questionnaire;
5) individual persons included in data file 3, with additional variables corresponding to Sections 8-12 of the individual questionnaire;
6) individual persons included in data file 3, with additional sets of basic and additional weights calculated for each respondent. More information on the types of weights and their calculation is provided in Subchapter 2.3.

The structure of these data files consists of coded responses to successive questions/variables included in the respective questionnaires. In addition, each file contains certain technical variables, such as the dwelling identification number, the household sequence number within the dwelling, and the person identifier within the household. These variables enable linking information on a given person collected across the six different data files.

As the AES also fulfils obligations arising from European Union legislation, non-identifiable microdata from the survey must be transmitted to Eurostat. Therefore, it is necessary to prepare a dataset compliant with the record structure required by Eurostat. This structure is defined by a European Commission implementing regulation--currently, *Commission Implementing Regulation (EU) 2021/861 of 21 May 2021 laying down technical specifications of the data set and establishing the technical formats for transmitting information on the organisation of the sample survey in the field of "education and training", pursuant to Regulation (EU) 2019/1700 of the European Parliament and of the Council* (Appendix 1). As the record structure for the EU dataset differs from the national structure, it is necessary to prepare an appropriate correspondence (mapping) key between the national data file structure and the EU record structure (Appendix 6).

6.5. Use of AES results and users of AES data

Statistics based on the *Adult Education Survey* support the monitoring of adult participation in lifelong learning at the EU level. Adult participation in lifelong learning plays an important role in the European policy agenda, particularly in the context of the European Pillar of Social Rights, the European Skills Agenda, and the Strategic framework for European cooperation in education and training towards the European Education Area and beyond. At the national level, data on participation in formal and non-formal education are also of key importance for supporting initiatives aimed at improving the skills of the adult population and increasing their employability, such as the Long-term National Development Strategy Poland 2030, the Human Capital Development Strategy, and the Social Capital Development Strategy. The results of the AES constitute a key data source used by the Interministerial Team for Lifelong Learning and the Integrated Qualifications System to monitor the implementation of national policy objectives in the field of adult formal and non-formal education.

User needs are systematically monitored and taken into account when preparing the Programme of Official Statistics for the following year. AES results are used by key national and international institutions, including ministries and central government bodies, the United Nations (UN) and its agencies, the Organisation for Economic Co-operation and Development (OECD), as well as national and international higher education institutions, research institutions, and the media.

7. Assessment of the survey quality

7.1. Unit non-response and item non-response (response rates)

Survey response rates, understood as the degree to which households in the selected dwellings were surveyed, is continuously monitored during fieldwork at both the regional and central levels. The following basic indicators are used: response rates, non-response rates, and refusal rates. Continuous monitoring and ongoing actions to increase response rates are crucial for enhancing survey quality. A low response rate may negatively affect the precision of the results and limit their interpretability (results may not be statistically significant). Furthermore, higher levels of non-response are associated with an increased risk of non-sampling errors (estimator bias). The reasons for non-completion of interviews at both the household and individual (respondent) levels are subsequently analysed in more detail during the weighting procedure, at the stage of generalising the survey results to the target population.

European regulations on household surveys specify precision requirements. For the AES, these relate to two key indicators: the participation rate in formal education and training (for persons aged 18-24) and the participation rate in non-formal education and training (for persons aged 25-69). Precision requirements are defined in Annex II to Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019. A particularly high non-response rate among young people (aged 18-24) negatively affected the magnitude of sampling error for the participation rate in formal education and training. The individual response rate for the youngest age group was 46%, compared with 65% for the overall surveyed population. Consequently, the sampling error calculated for this indicator exceeded the minimum standard error threshold specified in the EU regulation.

Despite a systematic increase in non-response rates (especially among young people) compared with previous AES editions, survey results still demonstrate a high degree of accuracy. Efforts were made to reach the youngest respondents, but low willingness to participate was particularly evident in this group. According to the interviewer instructions, persons aged 18-24 were interviewed first in the selected dwellings. Nevertheless, young people were often away from their usual residence, mainly for educational reasons, which significantly hindered establishing contact for conducting interviews, even after introducing the option of telephone interviews.

Response rate indicators for the AES are calculated according to the following formulas:

Response rate	R=Z*100/(Z+B^)
Refusal rate[8]	W=O*100/(Z+B^)
Non-response rate	I=B^*100/(Z+B^)

where:

Z – number of responding households;
B^ – number of households (registered during the survey in selected dwellings) and dwellings not surveyed but eligible for survey. This includes dwellings and households not surveyed due to: lack of access to the dwelling (e.g. inability to reach the dwelling, security issues), no contact with residents (residents not at home/no answer by phone, illness, advanced age, language barrier), short-term absence of residents (up to 12 months), or refusal to participate in the survey[9];
O – number of households (registered during the survey in selected dwellings) and dwellings not surveyed due to refusal by residents to participate.

[8] This indicator forms part of the non-response rate – refusal by residents to participate in the survey (denoted as O) is one of the reasons for non-response at the household (dwelling) level, and therefore is included in B^.
[9] The listed reasons for non-surveying have been in effect since 2017 (following the introduction of a new classification of non--response reasons). Consequently, response rate indicators for 2011 and 2016 are not fully comparable with those for 2022.

When calculating the above indicators, the denominator is based on the set of eligible units (Z+B^), i.e. households (dwellings) whose non-surveying can be considered justified are excluded (e.g. no dwelling selected, unoccupied dwelling, long-term absence of residents exceeding 12 months, or absence of persons aged 18-69). The selective inclusion of non-surveyed households (dwellings) in the denominator is related to the occurrence of coverage errors in representative surveys, i.e. errors in the coverage of the target population by the sampling frame – which are a type of non-sampling error. In the AES, overcoverage occurs, meaning that sampling frame includes dwellings that fall outside the scope of the survey (e.g. due to the absence of persons aged 18-69) or that do not exist in practice.

A consistent decline in response rates has been observed, regardless of the surveyed population or the size of the locality.

Response rate indicators for Poland are presented in the table below.

Table 4. Response rate indicators for the AES

Specification	2011	2016	2022
	in %		
Response rate (R)	54.7	54.1	51.2
Refusal rate (W)	16.3	29.9	32.9
Non-response rate (I)	45.3	45.9	48.8

In addition to unit non-response, item non-response also occurs, meaning that information is missing for at least one, but not all, questionnaire variables. The vast majority of questions in the household and individual questionnaires allow respondents to select the option "don't know/no answer." Such non-responses are not subject to imputation. According to Eurostat guidelines, data imputation is permitted only for three variables: SEX, AGE, and HHINCOME (household net monthly income category). The most sensitive questions, and therefore those most prone to high non-response rates, concern financial matters. These include household income (HHINCOME) and expenses related to participation in non-formal education (NFEPAIDVAL1 and NFEPAIDVAL2). Furthermore, the complexity of these concepts and difficulties in determining household income and expenditures contributed to a relatively high item non-response rate, and, consequently, lower precision of the resulting estimates based on these variables.

Information on variables with the highest item non-response rates (over 10%) is presented in Table 5. The proportion of non-response for a given question is calculated relative to the relevant target population, depending on the variable: the total population, the population participating in formal education, non-formal education, etc. (i.e. relative to the number of respondents to whom the question should have been asked during the interview).

Table 5. Item non-response rates for questions in AES 2022

Questionnaire	Question or section number	Variable name	Variable description	Item non-response rate (%)
individual	Q 60, col. 2	NFEPAIDVAL2	Expenditure for the 2nd randomly-selected non-formal learning activity (in PLN)	47.9
individual	Q 60, col. 1	NFEPAIDVAL1	Expenditure for the 1st randomly-selected non-formal learning activity (in PLN)	47.4
household	Sec. 3	HHINCOME	Net current monthly household income	27.7
individual	Q 67a or 67b	NONEED	No need for (further) education and training	13.2
individual	Q 19	JOBTIME	Year in which the person started working for current employer or as self-employed in main job	10.5
individual	Q 04	HATYEAR	Year when the highest level of education was successfully completed	10.3

7.2. Precision of results and limitations of data interpretation

The assessment of the precision of AES 2022 results is based on criteria defined in accordance with Eurostat guidelines. Thresholds were established to determine data reliability depending on the sample size used for estimation. These are as follows:

- unreliable data (should not be published): estimated results are based on fewer than 20 sample observations, or the item non-response rate for a given question exceeds 50%;
- low reliability data (may be published, but users are advised to exercise caution when using, interpreting, and drawing conclusions from them): estimated results are based on 20–50 sample observations, or the item non--response rate for a given question is between 20% and 50%;
- reliable data (published without reservations): estimated results are based on 51 or more sample observations, and the item non-response rate for a given question is below 20%.

For selected key statistics at the voivodship level (expressed as aggregate values), sampling errors were calculated in the form of relative standard errors (RSE), also referred to as coefficients of variation (CV). These are presented in the analytical publication *Adult Education Survey 2022*[10]. The values of relative errors were calculated taking into account the applied calibration of weights to demographic data. Their magnitude, expressed as a percentage, constitutes an additional criterion for assessing the reliability of the presented data:

- unreliable data (should not be published): the relative standard error for estimated aggregate values exceeds 20%;
- low reliability data (may be published, but users are advised to exercise caution when using, interpreting and drawing conclusions from them): the relative standard error for estimated aggregate values is between 10% and 20%;
- reliable data (published without reservations): the relative standard error for estimated aggregate values is below 10%.

The coefficient of variation (CV) is calculated by dividing the standard error of an estimator (SE) by the estimated value of that estimator (i.e. the value of a given parameter, such as an indicator or an aggregate value relating to a specific characteristic in a given population). The standard error of an estimator is the square root of its variance[11]:

$$(1)\quad SE(\hat{x}) = \sqrt{V(\hat{x})}$$

where:

$\hat{x}$ – estimator of the parameter x;

$SE(\hat{x})$ – standard error of the estimator $\hat{x}$;

$V(\hat{x})$ – variance of the estimator $\hat{x}$.

The basic formula for the coefficient of variation is as follows:

$$(2)\quad CV(\hat{x}) = \frac{\sqrt{V(\hat{x})}}{\hat{x}}$$

where:

$CV(\hat{x})$ – coefficient of variation of the estimator $\hat{x}$.

[10] Table 8.1. Persons aged 18–69 by participation in formal and non-formal education and informal learning, place of residence, labour market status, sex, age and voivodships; estimated totals (in thousands) and corresponding relative standard errors (%).
[11] Variance estimation was carried out using the linearisation method, implemented in the *vardpoor* package in R.

The coefficients of variation (CV), alongside standard errors (SE), are also calculated for key indicators in the field of adult learning, namely: participation in formal education and training by sex and age group; participation rate in non-formal education and training by sex, age group, labour market status and educational attainment; share of job-related activities among non-formal education and training activities by age group; participation in informal learning; average expenditure on non-formal education and training; and the average number of hours spent on formal and non-formal education and training. In addition, information on 95% confidence intervals for these indicators is transmitted to Eurostat in the quality report. This is another measure used to express the magnitude of sampling error. The general formula defining the confidence interval is as follows:

$$(3) \quad [\hat{x} - Z * SE(\hat{x}),\ \hat{x} + Z * SE(\hat{x})]$$

where:

Z – coefficient depending on the assumed confidence level (e.g. for a 95% confidence level, Z=1.96).

The interpretation of the 95% confidence interval is as follows: if the survey were repeated many times, then according to probability theory, on average 95% of the calculated intervals would contain the true value of the estimated parameter.

7.3. Measurement and processing errors

The main sources of measurement errors that may occur in the AES survey include:

- data entry errors;
- inadequately designed and developed survey questionnaires;
- inappropriate selection and training of survey staff (e.g. interviewers);
- unclear instructions for interviewers.

To reduce such errors, the following measures were taken:

- as part of the preparatory work, methodological and linguistic consultations (both external and internal) were conducted to better reflect the specific characteristics of non-formal education and training in Poland and its forms in the individual questionnaire, and to improve consistency with the Labour Force Survey (LFS) questionnaire;
- a series of training sessions was organised for coordinators and interviewers in regional statistical offices;
- during data collection, the survey authors and trained coordinators provided methodological support via telephone and e-mail;
- a mobile application with an electronic questionnaire was used for data entry; it included built-in plausibility and consistency checks, as well as alerts in cases of unlikely or invalid values;
- all statistical classifications used in the survey were embedded directly in the electronic questionnaire to support interviewers in coding certain responses;
- pilot testing of the mobile data collection application was carried out;
- the Survey Information System (SIB) enabled the control, analysis and validation of control tables, as well as data editing at both national and regional (voivodship) level.

The data entry process was carried out using an electronic questionnaire implemented in a mobile application. This questionnaire enabled real-time validation of data during the interview, which reduced errors at the data processing stage. All electronic devices used with the mobile application were synchronised with the Survey Management System (CORStat), enabling continuous data transmission to the central server. Raw data from the AES were subsequently imported into the Survey Information System (SIB). This system is used for data processing and enabled re-checking, correction, completion and coding of missing responses based on textual descriptions entered by interviewers in the mobile application during the interview. It also served as a tool for analysing aggregated data and validating datasets at both regional and national levels.

The structure of the national AES data files differs from the model dataset that must be transmitted to Eurostat. Therefore, it is necessary to prepare an appropriate conversion key to transform variables from the national questionnaire into a dataset compliant with the harmonised structure applied across all participating countries. The variable conversion process, carried out at the final stage of data processing, may be a source of hard errors (requiring correction). However, these were relatively few and could be easily eliminated during dataset validation using a dedicated programme developed by Eurostat. They mainly concerned missing or incorrectly coded categories of certain variables, as well as an incorrect record structure. Another category of errors includes soft (plausibility) errors, related to potentially inconsistent responses to questionnaire items. Such errors may be accepted following detailed data analysis. Some soft errors indicated the need for corrections where the identified inconsistencies in responses could not be explained. In such cases, data was corrected based on logical inference derived from the analysis of responses to all questionnaire items, without the need to obtain additional clarification from respondents.

7.4. Quality reports

For the purpose of assessing the quality of the survey, a national quality report and a quality report for Eurostat were prepared. The reports were compiled after the completion of the survey, within the deadlines specified by Eurostat (for European reports) and by the Statistical Office in Łódź (for national reports). The national report consists of several sections, each devoted to the analysis of a selected quality component. Each report includes: a description of the component, quality measurement (quality assessment indicators), and conclusions aimed at improving quality within that component. The quality dimensions subject to evaluation are as follows: relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability, coherence, costs and respondent burden, confidentiality, transparency, and data security. The report also includes a summary presenting proposals for strengthening quality in selected areas requiring improvement. The structure of the quality report is based on the general standards applicable within the European Statistical System (ESS).

The report prepared for Eurostat was also developed in accordance with the ESS standard structure for quality reports (Single Integrated Metadata Structure – SIMS), and quality assessment was carried out across dimensions analogous to those used in the national report. Nevertheless, reports prepared for Eurostat, within the same quality components, may include additional information specific to the survey. This is primarily achieved through a dedicated appendix developed for the AES quality report. It is intended to provide more detailed and extensive information, including numerical data on sampling errors calculated for key survey indicators, coverage errors, unit and item non-response, deviations from the methodology adopted by Eurostat, the comparability of AES results with other household surveys (information on the total population), etc. Quality reports for individual countries, including Poland, are available on the Adult Education Survey website.

Bibliography

Adult Education Survey (AES) methodology. Statistics Explained, Eurostat (b.r.).
Adult Education Survey (AES) methodology - Statistics Explained - Eurostat.

Kształcenie Dorosłych, Główny Urząd Statystyczny, Warszawa 2009.
https://stat.gov.pl/cps/rde/xbcr/gus/WZ_ksztalcenie_doroslych.pdf.

Stan zdrowia ludności Polski w 2009 r., Główny Urząd Statystyczny, Warszawa 2011.
Główny Urząd Statystyczny / Obszary tematyczne / Zdrowie / Zdrowie / Zdrowie i ochrona zdrowia w 2024 r.

Kompendium wiedzy nt. zapewnienia wysokiej jakości badań w polskiej statystyce publicznej, Główny Urząd Statystyczny, Warszawa 2024.
https://bip.stat.gov.pl/dzialalnosc-statystyki-publicznej/jakosc-w-statystyce/ocena-i-monitorowanie-jakosci-procesow-statystycznych/?fbclid=IwY2xjawJkeD5leHRuA2FlbQIxMAABHs9rmsD4Haviei9p_D4ru9Cnz4M6gh-s1g33y0r-gl_egaR1fwUUc1A6Iz2p_aem_oJoIM0XkjUHN_CqGp3kU_A.

Kształcenie dorosłych 2011, Główny Urząd Statystyczny, Warszawa 2013.
https://stat.gov.pl/obszary-tematyczne/edukacja/edukacja/ksztalcenie-doroslych-2011,3,2.html.

Kształcenie dorosłych 2016, Główny Urząd Statystyczny, Warszawa 2018.
https://stat.gov.pl/obszary-tematyczne/edukacja/edukacja/ksztalcenie-doroslych-2016,3,3.html.

Adult Education 2022, Statistics Poland, Warsaw 2024.
Statistics Poland / Topics / Education / Education / Adult Education 2022.

Adult Education in 2022, News release, Statistics Poland, Warsaw 2024.
Statistics Poland / Topics / Education / Education / Adult Education in 2022.

Szymkowiak M., Józefowski T., *Calibration estimators in statistical surveys, The Polish Statistician* (1), 31-42, Statistics Poland, Warsaw 2012.
Wiadomości Statystyczne The Polish Statistician Calibration estimators in staistical surveys.

Wesołowski J., Wieczorkowski R., *An eigenproblem approach to optimal equal-precision sample allocation in subpopulations, Communications in Statistics – Theory and Methods* 46(5), s. 2212-2231, 2017.
https://doi.org/10.1080/03610926.2015.1040501.

Williams M. R., Savitsky T. D., *Optimization for Calibration of Survey Weights under a Large Number of Conflicting Constraints, Journal of Computational and Graphical Statistics* 33(3), s. 1047–1060, 2024.
https://doi.org/10.1080/10618600.2024.2304071.

Appendices

Appendix 1

Legislation at EU level concerning *Adult Education Survey*

Appendix 2

Commission Implementing Regulation (EU) 2019/2240 2021/861 of 21 May 2021 specifying the technical items of the data set and establishing the technical formats for transmission of information on the organisation of a sample survey in the education and training domain pursuant to Regulation (EU) 2019/1700 of the European Parliament and of the Council

Appendix 3

Questionnaires for the AES – household questionnaire

Appendix 4

Questionnaires for the AES – individual questionnaire

Appendix 5

Description of the structure of the AES 2022 data files

Appendix 6

Assumptions for conversion of variables for Eurostat

Appendix 7

Code book and classification guidelines for the AES 2022