

CENSUS_INS21ES_A_BE_2021_0000

National Reference Metadata in Euro SDMX Metadata Structure (ESMS)

Compiling agency: Statbel (Directorate-general Statistics – Statistics Belgium)



Eurostat metadata

Reference metadata

- [1. Contact](#)
 - [2. Metadata update](#)
 - [3. Statistical presentation](#)
 - [4. Unit of measure](#)
 - [5. Reference Period](#)
 - [6. Institutional Mandate](#)
 - [7. Confidentiality](#)
 - [8. Release policy](#)
 - [9. Frequency of dissemination](#)
 - [10. Accessibility and clarity](#)
 - [11. Quality management](#)
 - [12. Relevance](#)
 - [13. Accuracy](#)
 - [14. Timeliness and punctuality](#)
 - [15. Coherence and comparability](#)
 - [16. Cost and Burden](#)
 - [17. Data revision](#)
 - [18. Statistical processing](#)
 - [19. Comment](#)
- [Related Metadata](#)
[Annexes](#) (including footnotes)

For any question on data and metadata, please contact: [Eurostat user support](#)

1. Contact Top	
1.1. Contact organisation	Statbel (Directorate-general Statistics – Statistics Belgium)
1.1.1. Responsible party	Statbel (Directorate-general Statistics – Statistics Belgium)

1.1.2. Responsible party role	Point of Contact
1.2. Contact organisation unit	Databases Citizens
1.5. Contact mail address	Boulevard du Roi Albert II, 16 - 1000 Brussels

2. Metadata update Top	
2.1. Metadata last certified	23/12/2022
2.2. Metadata last posted	23/12/2022
2.3. Metadata last update	23/12/2022

3. Statistical presentation Top	
3.1. Data description	
See subcategories	
3.1.1. Resource title	
Population distribution on 1 km ² grid cell.	
3.1.2. Resource abstract	
Usual resident population by sex, broad age classes and usual residence expressed in km ² grid cells.	
3.1.3. Resource type	
Spatial data set	
3.1.3.1. Additional information if 'spatial data set' is marked	
https://data.europa.eu/en	
3.1.4. Resource locator	
https://www.geo.be/catalog/details/ea430cb8-2805-4c67-8392-9e2a47c1ef55?l=en	
3.1.5. Unique resource identifier	

https://data.europa.eu/en
3.1.6. Coupled resource
https://data.europa.eu/en
3.1.7. Resource language
English
3.1.8. Keyword value
Population Distribution
3.1.9. Originating controlled vocabulary
https://inspire.ec.europa.eu/theme/pd
3.1.10. Metadata language
English
3.2. Classification system
The cell code of the 1 km INSPIRE Statistical Units Grid for pan-European usage (ETRS89 Lambert Azimuthal Equal-Area (ETRS89-LAEA) coordinate reference system)
3.3. Coverage - sector
See subcategories
3.3.1. Topic category
Society
3.3.2. Spatial data service type
Download Service
3.4. Statistical concepts and definitions
See subcategories
3.4.1. Statistical concepts and definitions - Total population
For census purposes, the total population of the country consists of all the persons falling within the scope of the census in EU on usual residents.

Usual residence means the place where a person normally spends the daily period of rest, regardless of temporary absences for purposes of recreation, holidays, visits to friends and relatives, business, medical treatment or religious pilgrimage.

Persons usually resident in the place of enumeration but absent, or expected to be absent, at the time of the census for less than one year shall be considered as temporarily absent persons and thus included in the total usually population. In contrast, persons living or expected to live outside the place of enumeration for one year or more shall not be considered temporarily absent and shall therefore be excluded from the total population. This is regardless of the length of visits that they may pay to their families from time to time.

Persons who are enumerated but do not meet the criteria for usual residence in the place of enumeration, i.e. do not live or do not expect to live in the place of enumeration for a continuous period of at least 12 months, are considered temporarily present and are therefore not counted in the total usually resident population. Unless otherwise stated in this report.

3.4.2. Statistical concepts and definitions - Sex

'Sex' refers to the biological and physiological characteristics that define men and women.

3.4.3. Statistical concepts and definitions - Age

The age reached at the reference date (in completed years).

3.4.4. Statistical concepts and definitions - Employed person

Will be available later in 2023.

3.4.5. Statistical concepts and definitions - Place of birth

The place of birth for persons born within the country is the civil division in which the person was born; for those born in other countries, it is the country of birth. For persons born in the country (the native-born population), the concept of place of birth usually refers to the geographic unit where the mother of the individual resided at the time of the person's birth.

3.4.6. Statistical concepts and definitions - Place of usual residence one year prior to the census

The relationship between the current place of usual residence and the place of usual residence one year prior to the census.

For all persons that have changed their usual residence more than once within the year prior to the reference date, the previous place of usual residence is the last usual residence from which they moved to their current place of usual residence.

3.5. Statistical unit

grid cell

3.6. Statistical population
Usual resident population on 1st January 2021.
3.7. Reference area
See subcategories
3.7.1. Geographic bounding box
In WGS84: West 2.55° South 49.5° East 6.41° North 51.51°
3.7.2. Spatial resolution
1000 Unit:m
3.7.3. Coordinate Reference System
EPSG:3035 - ETRS89 / LAEA Europe
3.8. Coverage - Time
Data refer to the situation in the reporting country at the census reference date. 1st January 2021
3.9. Base period
Not applicable.

4. Unit of measure	Top
Counts of statistical units.	

5. Reference Period	Top
See subcategorie	

5.1. Temporal extent

1 January 2021

6. Institutional Mandate

[Top](#)

6.1. Institutional Mandate - legal acts and other agreements

The census in Belgium is regulated at the European level.

Regulation (EC) 763/2008; Regulation (EU) 2017/543; Regulation (EU) 2017/712 and Regulation (EU) 2017/881.

There isn't a specific law at national level for the 2021 census. Of course the Belgian statistical law of 4 July 1962 applies here.

Law of 4 July 1962 on official statistics (ammended by the Act of 1 August 1985 on tax and other measures, the Act of 21 December 1994 on social and other provisions, the Law of 2 January 2001 on social provisions, Budget and miscellaneous and by the program law of 22 December 2008).

6.2. Institutional Mandate - data sharing

not applicable

7. Confidentiality

[Top](#)

7.1. Confidentiality - policy

Confidentiality - Regulation (EU) 2017/712 Art 4 and 5

7.2. Confidentiality - data treatment

There was some research in order to decide which SDC-method to use.

First the microdata were aggregated without SDC. These results were checked for sensitive information. The first conclusion was that for the hypercubes as well for the GRID-data, there are not a lot of sensible frequency counts. So, a very light SDC-method should suffice.

Different SDC-methods were considered:

- Cell key method was never considered as a serious candidate because of the lack of additivity. Our dissemination team found it too difficult to explain to the general public that additivity is not preserved.

- We performed record swapping in 2011, but due to some inconsistency problems in 2011 for the household data, there was very little support to use this method again. There was also a lot of criticism from census users regarding the record swapping method.
- Also, cell suppression was considered as a bad choice due to reasons Eurostat mentioned in several meetings.

Statbel came up with a new method we could call "geographical perturbation". This method only works for a very light SDC approach when only a very small percentage of the frequency counts is considered as confidentially problematic. This method is now applied for the grid data. Briefly explained, this method works as follows:

It is a pretabulated method, which means some of the modifications is done in the microdata before aggregation.

1. Based on some confidentiality rules, we determine which grid cells have to be protected. These cells are determined on the basis of the aggregated data before a SDC method is applied.
2. Then we add some perturbations to the microdata. Households in a grid cell to protect are virtually moved to a point with different geographical coordinates. This is a point in another grid cell that should not be protected. This point is always located in the same municipality (same LAU2). We try to choose to the greatest extent possible a point in the same neighbourhood. In Belgium we have some statistical geographical areas that are very detailed. We call them the statistical sectors and these are more detailed than the LAU2 level. Statistical sectors can be considered as the neighbourhoods and are very important for our national dissemination. We choose, if possible, the nearest point (=inhabited building) in the same statistical sector. But (in some rare cases) if this point is too far away, we choose a point closer to the original point in another statistical sector, but always in the same LAU2 area.
3. Grid cells are recalculated based on the perturbed points. After that, the dataset is aggregated again.

Of course, the result of this method depends on the confidentiality rules used in step (1). We have tested this method where we started from different confidentiality rules, used in step 1 and we've compared the results with each other. These rules in step 1 are criteria based on frequency counts in grid cells such as number of people, number of households and the other frequency counts by grid cell we have to deliver to Eurostat in 2024. We compared relatively mild criteria with more severe criteria. After analysing the results, the conclusion was that one of the mildest criteria was sufficient for the protection of confidential data. Keeping the detailed rules secret is an important aspect to prevent anyone from being able to undo the protection.

8.1. Release calendar
23 December 2022
8.2. Release calendar access
23 December 2022
8.3. Release policy - user access
See subcategories
8.3.1. Conditions applying to access and use
https://inspire.ec.europa.eu/metadata-codelist/ConditionsApplyingToAccessAndUse/noConditionsApply
8.3.2. Limitations on public access
https://inspire.ec.europa.eu/metadata-codelist/ConditionsApplyingToAccessAndUse/noConditionsApply

9. Frequency of dissemination	Top
Decennial	

10. Accessibility and clarity	Top
10.1. Dissemination format - News release	
.gpkg .gml	
10.2. Dissemination format - Publications	
not yet available	
10.3. Dissemination format - online database	
not yet available	
10.4. Dissemination format - microdata access	

According to the Belgian statistical law of 4 July 1962, researchers have access to microdata. It concerns pseudonymised data. Researchers can submit an application to the DPO committee of statbel. Based on the principles of finality and proportionality (described in Belgian law), an assessment is made for researcher's application.

Finality = clear description of the purpose for which the data will be used. The researcher can only use the data for the purposes described in his application.

Proportionality = The researcher can only request the data that are really necessary for his research. (e.g. if the researcher only needs data on NUTS 3 level, he can't obtain data on LAU 2 level)

Annexes:

[access to microdata for research](#)

10.5. Dissemination format - other

not available

10.6. Documentation on methodology

See 18.5 data compilation

10.7. Quality management - documentation

See link.

Annexes:

[Information on quality](#)

11. Quality management

[Top](#)

11.1. Quality assurance

See subcategories

11.1.1. Lineage

See attached file in 11.2 Quality management - assessment. There was a longitudinal comparison between grid files.

11.1.2. Conformity, specification

The grid cells are conform with the definition of ETRS89-LAEA.

Population counts are conform with the definition of usual residence.

11.1.3. Conformity, degree

See 11.1.2

11.2. Quality management - assessment

For the grid data:

- Additivity of population figures and surface data were checked.
- Comparison was performed with the demographic statistics already published. The numbers match exactly. The total area is comparable with cadastre results.
- For each grid cell, the values of 2021 were compared with the grid in 2016. The differences are very acceptable.

Inside *Statistics Belgium*, a service which is independent of the service responsible for producing figures, is responsible for the final quality management of the data. Each year there is a survey that has to be filled in for all statistics that evaluates the different quality dimensions of the statistic.

Annexes:

[Quality check grid data 2021](#)

12. Relevance

[Top](#)

12.1. Relevance - User Needs

There are different working groups and task forces with different kind of users.

Each year the different users inside the IIS (Interfederal Institute for Statistics) can submit projects to improve statistics, develop new statistics or can ask to add new information in existing statistics.

There are also multiple task forces with researchers to discuss this.

In the high council of statistics, the most important stakeholders come together to follow up the statistical production. Specific working groups can be established if some new needs are expressed or if a reform of a particular statistic is necessary.

12.2. Relevance - User Satisfaction

See attached link.

Annexes:
[satisfaction survey](#)

12.3. Completeness

not available

13. Accuracy

[Top](#)

13.1. Accuracy - overall

Highly accurate.

Statbel uses the concept of usual residence and tries to deduce this from the data in the population register. The population register disposes of longitudinal information. Therefore, we have information of moving citizens. Based on that information, it's possible to see if condition (i) of the definition of usual resident population is satisfied:

"(i) those who have lived in their place of usual residence for a continuous period of at least 12 months before the reference date"

Indeed, we can check whether someone lives already for 12 months in Belgium.

It's more difficult to see if condition (ii) is satisfied:

"(ii) those who arrived in their place of usual residence during the 12 months before the reference date with the intention of staying there for at least one year"

In the population register, we have no information on the **intention** to stay. But, based on the records with information of moving citizens in the months after the reference date, it is possible to estimate this.

Remarks

- This differs from the figures statbel publishes on its own website, where the registered population is used at midnight 31 December 2020 (=1st January 2021). Moreover, asylum seekers are included in the figures we submit to Eurostat which isn't the case in our national publications.
- Both census statistics and demography use the same reference moment and data source (population register). As a consequence, the demographic part of the census is consistent with the demographic statistics published earlier.

Once the usual resident population is determined, we can precise with more geographical detail the usual residence. Municipality codes are available in the population register. Therefore, all levels NUTS1,2,3 and LAU2 can be derived from that information. To see in which km² grid cell the citizen has his/her usual residence, we use the address from the population register.

There is a geocoding of these addresses in what we call the CSAB file established by Statbel.

In a first step we require that members of the same household are attributed to the same geographical point. Therefore, it is important to determine the household id's before a grid cell id can be allocated. The composition of the households is also derived from information in the population register. Here, we distinguish between private households and collective living quarters. Private households are mainly determined based on information of the reference person in the household, whereas collective living quarters are determined by the information in the address.

To establish the CSAB file (geocode the addresses), the information in the population register has to be compared with geographical data coming from databases from the 3 regions (each NUTS 1 region has his own database). There is an automated step first. To compare addresses between the population register and the geographical data, intelligent algorithms are used as e.g. based on TF-IDF methods. Sometimes the exact address isn't available as such in one of the geographical registers. e.g. for a given street we have nr. 20 in the population register, but in the geographical databases we have nr. 18 and nr. 22. In those cases geographical interpolation can be used to estimate the geographical coordinates. In some other cases extrapolation can be used.

After that there is a manual step to geocode the rest of the addresses whereby geographical coordinates were not found in the automated step.

The CSAB file is updated each year with all new addresses. In the Belgian geographical databases, Lambert 72 (EPSG:31370) projection system is used. To attribute a grid ID, this is converted to the ETRS89 / LAEA Europe (EPSG:3035) system.

It is difficult to mention the exact accuracy of the geographical points, since this can change from address to address. Interpolated and extrapolated points are far less precise than addresses available in the original geographical databases. But the overall accuracy is very precise. Addresses in the original geographical databases are precise to the metre and sometimes even more detailed. Only 0,32% of the addresses is inter- or extrapolated. Exceptionally, it can be less precise for some extrapolated points. Mostly, only for points very near to the border of 2 different grid cells there is a very small risk the address is allocated to the wrong grid cell.

For the km² grid, not only population figures were supplied, but also land surface data. In most cases this is exactly 1 km². For border and coastal areas, the area was computed based on a GIS intersection of the GRID layer with the layer of the official Belgian borders both converted to the ETRS89 projection. Moreover, areas of water bodies such as lakes were subtracted from the initial surface. This information is estimated based on Corine Land Cover 2018 data.

13.2. Sampling error

not applicable

13.3. Non-sampling error

not applicable

14. Timeliness and punctuality

[Top](#)

14.1. Timeliness

See subcategories

14.1.1. Date of publication

23 December 2022

14.1.2. Date of last revision

23/12/2022

14.1.3. Date of creation

23/12/2022

14.2. Punctuality

Grid data for total population was delivered on time before the deadline of 31 December 2022.

15. Coherence and comparability

[Top](#)

15.1. Comparability - geographical

See 15.1.1

15.1.1. Geographic information - data quality

In all regions of Belgium, the data of the geographical databases are very precise. In most cases more precisely than one metre. See also 13.1

15.2. Comparability - over time

A comparison check was executed. See attached file in 11.2 **Quality management - assessment.**

15.3. Coherence - cross domain

not applicable

15.4. Coherence - internal

There is exact coherence with the already published population statistics.

16. Cost and Burden

[Top](#)

Not yet available.

17. Data revision

[Top](#)

17.1. Data revision - policy

See attached link.

Annexes:

[revision policy](#)

17.2. Data revision - practice

See 17.1.

The grid data will be each year a one shot. Census data and population statistics have the same reference date and uses the same administrative sources for the demographic part. As a consequence, no revisions are needed.

18. Statistical processing

[Top](#)

18.1. Source data

To establish the grid data, the Belgian population register was used.

18.2. Frequency of data collection

Decennial

18.3. Data collection

See 18.1

18.4. Data validation

Inside *Statistics Belgium*, a service which is independent of the service responsible for producing figures, is responsible for the final validation of the data. Internal validation procedures have been set up in order to ensure consistency between all hypercubes and grid data. An external evaluation was also carried out. The aim is, on the one hand, to ensure that the statistics produced are part of a time trend or, if interruptions are found, that these are the consequence of specific, cyclical socio-economic factors. Furthermore, the figures set out in the grid cells were compared with the already published population statistics.

So following checks were executed for the grid statistics

- Additivity of population figures and surface data were checked.
- Comparison was performed with the demographic statistics already published. The numbers match exactly. The total area is comparable with cadastre results.
- For each grid cell, the values of 2021 were compared with the grid in 2016. The differences are very acceptable.

18.5. Data compilation

To see in which km² grid cell the citizen has his/her usual residence, we use the address from the population register.

There is a geocoding of these addresses in what we call the CSAB file established by Statbel.

In a first step we require that members of the same household are attributed to the same geographical point. Therefore, the geographical point for each household is also derived from information in the population register. Here, we distinguish between private households and collective living quarters. For private households, the geographical point is the address of the person in the household, whereas collective living quarters are determined by the information in the address.

To establish the CSAB file (geocode the addresses), the information in the population register has to be compared with the geographical data. This is done in an automated step first. To compare addresses between the population register and the geographical data, intelligent algorithms are used. For example, for a given street we have nr. 20 in the population register, but in the geographical databases we have nr. 10. In some other cases extrapolation can be used.

After that there is a manual step to geocode the rest of the addresses whereby geographical coordinates were not found in the population register allocated one street, but the geographical registers the other street. The geographical coordinates are then manually assigned to the correct street.

The CSAB file is updated each year with all new addresses. In the Belgian geographical databases, Lambert 72 (EPSG:31466) is used as the coordinate system.

For the km² grid, not only population figures were supplied, but also land surface data. In most cases this is exactly 1 km². The official Belgian borders both converted to the ETRS89 projection. Moreover, areas of water bodies such as lakes were excluded from the grid.

18.6. Adjustment

not available

19. Comment

[Top](#)

3.3.2. We dispose download service and view service.

Is there a problem to indicate several answers?

Related metadata

[Top](#)

Annexes

[Top](#)