

# Population Census Microdata Availability

Jiří Novák<sup>1</sup> | Prague University of Economics and Business, Prague, Czech Republic

Received 12.12.2021 (revision received 30.3.2022), Accepted (reviewed) 12.4.2022, Published 16.9.2022

## Abstract

In this paper, the author aims to describe the dissemination of microdata from the population census by National Statistical Offices. This type of data is highly confidential, and approaches to protection vary across the world. National Statistical Offices mostly strive to publish their data as much as possible, but they are bounded by national and international laws to protect the personal data of respondents. The primary goal is mapping the differences between countries and their categorization. Different approaches to microdata availability are described, and various data access approaches are depicted. The information was obtained from publicly available documentation and a survey in which selected statistical offices were contacted. Discovered were that of the 223 countries (including dependent territories), 100 countries have made microdata available for the scientific community, with 30 countries also providing microdata access to the public. This paper presents a mapped overview and aggregated information on the publication of microdata of the population census from around the world.

## Keywords

*Microdata, population census, statistical disclosure control, confidentiality, public use files, scientific use files*

## DOI

<https://doi.org/10.54694/stat.2021.44>

## JEL code

C80

## INTRODUCTION

A growing pressure from the research community, policymakers and citizens has been observed for publishing more data in increasing detail. Producers of the statistic, specifically National Statistical Offices (NSO) or other statistical agencies, stand on the other side. They are bound by national and international laws protecting personal data and keeping the trust of their respondents. The confidentiality of the data is essential to all statistical offices and agencies because respondents and NSO's can trust each other only if the trust is maintained, which increases the quality of the data obtained from individual respondents.

However, the dissemination of census data has many positives, and their subsequent analysis has many benefits that aim to improve the social well-being and everyday lives of ordinary people, who mostly do not even know that it was the census data used to improve their surroundings. The population census is a unique statistical survey that makes it possible to obtain essential information that cannot usually be obtained in any other way. Census microdata can be used for research and planning in areas such as health, fertility, housing, transportation, education, employment, migration and regional development.

<sup>1</sup> Department of Economic Statistics, Prague University of Economics and Business, W. Churchill Sq. 1938/4, 130 67, Prague 3, Czech Republic. E-mail: jiri.novak.kest@vse.cz, phone: (+420)224095435.

For example, as stated in Černý (2021a), NSO's obtain information about population distribution, which in European Union censuses is based on usual residence (the place where the person actually resides, regardless of where he is domiciled). Not only that, NSO's further obtain information about education with other socio-demographic indicators, such as the place of residence, size of the dwelling, number of household members, etc.

In the following article by Černý (2021b), several projects were presented, which have used census data to improve the state of affairs. To give an idea of the benefits of census data for the city, one of the projects was the Transport model of the city of České Budějovice. In the project, the authors created a mathematical model of automobile transport in České Budějovice and its surroundings. The model then enabled them to plan municipal transit/transport, assess changes in the organization of transit/transport and forecast car transport or evaluate the impact of new investments. This model included data that cannot be obtained from a source other than the census, such as information concerning commuting to work or the actual number of people in the surrounding area.

There are many approaches how to make microdata accessible by NSO. It can be accessed via Data Laboratories, Remote Access Facilities or are provided as products for use outside the NSO. Disseminated microdata can be categorized into two groups based on intended. The first group is microdata that the general public can access, and the second one is microdata that can access only by approved researchers. These groups do not have a fixed nomenclature, as it varies by region. In the European region is used term Public Use Files (PUF) for the first group and the Scientific use files (SUF) for the second. Other regions are described further in the article.

In this research paper, the author aims to examine the approaches taken by the individual European National/Central Statistical Offices and other statistical agencies around the world in relation to the publication of microdata. The main goal is to compare specific approaches between the individual countries (European and non-European). The primary interest of this work is whether the statistical offices publish microdata at all and, if so, what approach and models they chose to protect the personal data of their respondents. The information has been obtained from publicly available documentation and from a survey sent to selected statistical offices.

## 1 LEGISLATION AND LITERATURE

Microdata are data that contain information about an individual person, household, business, or other entity (Templ et al., 2014). Usually, they are collected directly by NSO, or they can be obtained from the administrative sources or surveys. Raw data of this character are highly sensitive on disclosure of confidential information, and in case of release to the scientific community or sometimes even to public, there have to be special measures applied to secure data and protect personal information. In the case of a population census, microdata is the data that contains information about individual households and housing units (United Nations 2017), who are located in a specific territory.

According to United Nations (2007) every official statistical system should support any research based on microdata. The following benefits are defined, which result from more accessible publishing of micro-data:

- i. microdata permits policy makers to pose and analyse complex questions. In economics, for example, analysis of aggregate statistics does not give a sufficiently accurate view of the functioning of the economy to allow analysis of the components of productivity growth;*
- ii. access to microdata permits analysts to calculate marginal rather than just average effects. For example, microdata enable analysts to do multivariate regressions whereby the marginal impact of specific variables can be isolated;*
- iii. broadly speaking, widely available access to microdata enables replication of important research;*
- iv. access to microdata for research purposes, and the resulting feedback, can facilitate improvements in data quality. For example, the US Bureau of the Census has formalised the documentation it requires from researchers to assist it in improving the quality of its surveys;*

v. *it increases the range of outputs derived from statistical collections and hence the overall value for money obtained from these collections.*

International organizations such as United Nations (UN) and European Union (EU) support the publication of microdata in compliance with strict privacy of personal data.

United Nations defined this approach in their Fundamental Principles of Official Statistics (United Nations, 2015a) as a sixth principle – Confidentiality, which states that: individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes. This principle is followed by United Nations (2007), which sets out further principles for the handling of confidential microdata. The mentioned principles are as follow:

*Principle 1: It is appropriate for microdata collected for official statistical purposes to be used for statistical analysis to support research as long as confidentiality is protected.*

*Principle 2: Microdata should only be made available for statistical purposes.*

*Principle 3: Provision of microdata should be consistent with legal and other necessary arrangements that ensure that confidentiality of the released microdata is protected.*

*Principle 4: The procedures for researcher access to microdata, as well as the uses and users of microdata, should be transparent and publicly available.*

European Union defined this approach in their European Statistics Code of Practice (Eurostat, 2018) as a fifth principle – Statistical Confidentiality and Data Protection, which states that: *the privacy of data providers, the confidentiality of the information they provide, its use only for statistical purposes and the security of the data are absolutely guaranteed.* This is more specified in Eurostat (2018) by the following subpoints:

*5.1 Statistical confidentiality is guaranteed in law.*

*5.2 Staff sign legal confidentiality commitments on appointment.*

*5.3 Penalties are prescribed for any wilful breaches of statistical confidentiality.*

*5.4 Guidelines and instructions are provided to staff on the protection of statistical confidentiality throughout the statistical processes. The confidentiality policy is made known to the public.*

*5.5 The necessary regulatory, administrative, technical and organisational measures are in place to protect the security and integrity of statistical data and their transmission, in accordance with best practices, international standards, as well as European and national legislation.*

*5.6 Strict protocols apply to external users accessing statistical microdata for research purposes.*

Efforts to make microdata available are not in conflict with UN principles and the EU Code as long as the data are used for statistical purposes (what is the statistical purpose is described below) and it is prevented that individual data on respondents can be identified from the published data.

## **1.1 European legislation**

Microdata publications have to be supported by corresponding legislation as is set out in the principles above. The protection of personal data is usually already covered by national law, and in this paper, when examining the approach of national offices to publishing microdata, the author also recorded the relevant laws that cover this issue. The relevant laws of individual countries can be found in the results section, if information about the laws was available.

In European Union, the national statistical confidentiality is provided for in the EU legislation. The principal regulations on the EU level that cover statistical confidentiality and protection of personal information are the following regulations.

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 also known as General Data Protection Regulation (GDPR) represents an EU-wide

legal framework for the protection of personal data, which protects the rights of its citizens against the unauthorized treatment of their data and personal data.

In Recital 162 of GDPR (Regulation (EU) 2016/679), there are defined statistical purposes as *"any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results"* and that *"the statistical purpose implies that the result of processing for statistical purposes is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person"*. This is a crucial sentence that guarantees to all persons whose data is used in the microdata analysis that the results of the analyses will not be used against them. The result of microdata analysis should not be information that could lead to measures against certain individuals but rather aggregated information from which can a population benefit as a whole. The researcher should not be interested in information about the persons contained in the data, but in the structure and links hidden in the data, which is information that may be lost by pre-processing of the data when the researcher has available only an aggregated dataset. In the recital is also stated that *"where personal data are processed for statistical purposes, this Regulation should apply to that processing"* and that *"Union or member state law should, within the limits of this Regulation, determine statistical content, control of access, specifications for the processing of personal data for statistical purposes and appropriate measures to safeguard the rights and freedoms of the data subject and for ensuring statistical confidentiality"*.

GDPR Regulation (EU) 2016/679 lays down rules for the publication in Article 6 Lawfulness of processing, where is stated that *Processing shall be lawful only if and to the extent that at least one of the following applies:*

- a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes;*
- b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;*
- c) processing is necessary for compliance with a legal obligation to which the controller is subject;*
- d) processing is necessary in order to protect the vital interests of the data subject or of another natural person;*
- e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;*
- f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.*

Sub-point c), which refer to compliance with a legal obligation, is a crucial legal instrument for publishing microdata. The legal obligation is represented by other legislation, that has specified the rights and obligations for the handling of personal data in the given case specified by law. The legislation that specifies this is a regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics as amended by Regulation (EU) No 2015/759, which then describes in detail the handling of data for statistical purposes. This is also confirmed in GDPR by Recital 163.

In Recital 26 of (Regulation (EC) No 223/2009) is acknowledged that *the research community should enjoy wider access to confidential data used for the development, production and dissemination of European statistics, for analysis in the interest of scientific progress in Europe. Access to confidential data by researchers for scientific purposes should therefore be improved without compromising the high level of protection that confidential statistical data require.* In Article 19 Public use files (PUFs) is then further specified that data on individual statistical units, disseminated in the form of a public use file consisting of anonymised records which have been prepared in such a way that the statistical unit cannot be identified,

either directly or indirectly, when account is taken of all relevant means that might reasonably be used by a third party. In Article 23 Access to confidential data for scientific purposes, which describes Scientific use files (SUFs) is stated that access to confidential data which only allow for indirect identification of the statistical units may be granted to researchers carrying out statistical analyses for scientific purposes by the Commission (Eurostat) or by the NSIs or other national authorities, within their respective spheres of competence. (Regulation (EC) No 223/2009).

EU legislation not only provides legal tools for publishing microdata, in its regulations, it also encourages data to be made available to the scientific community, while maintaining strict confidentiality in the data and under the conditions of examining the data as a whole. Each member country has its own approach to the protection of microdata. There is no regulation at the European level that prescribes protection procedures. However, there is inter-European coordination, where experiences are exchanged within a working group and expert group on statistical disclosure control.

## 1.2 Statistical disclosure control

The set of methods focusing on the protection of disseminated data are called Statistical Disclosure Control (SDC). SDC is the term most used in the European region, but in other regions are used slightly different terms like Statistical disclosure limitation or Disclosure avoidance. Every statistic released from NSO is controlled by these methods to protect statistical confidentiality.

There are many approaches to publishing data and according to United Nations (2007) they can be categorized as follows:

(i) Statistical products for use outside the NSO.

– Statistical Tables which can be standardized or specially generated on specific request of the researcher. Data Cubes which are flexible multi-dimensional matrices that allow to researcher to create required table on their own.

– Public Use Files (PUF) are microdata that can be accessed by general public, usually after registration, and because of that the highest level of security against disclosure of respondents' personal data is necessary. PUF is the most used expression in the European region; its definition is codified by Regulation EC No 223/2009 2015. In the USA, PUF's are called Anonymised Microdata Files.

– Scientific use files (SUF) are microdata that can access only approved researchers, usually is signed contract. This type of microdata needs lower level of security than PUFs, but the personal information of respondents is still preserved against disclosure. SUF is term most used in the European region, its definition is codified by Regulation EC No 223/2009 2015. In USA is SUF called Licensed Anonymised Microdata Files.

(ii) A service window through which researchers can submit data requests.

– Remote Access Facilities which means that researchers could create their statistical tasks remotely from microdata through computer networks.

(iii) Arrangements for allowing researchers to work on the premises of the National Statistical Office.

– Data Laboratories also known as Data Centers or Safe Centers, which allows researchers to access microdata in secure place at NSO and because of that can allow to researchers work with more detailed data.

The methods designed for protecting microdata are in detail described in Hundepool et al. (2010) and Hundepool (2012). Microdata protection can be categorized into three main groups: Non-perturbative microdata masking, Perturbative microdata masking, Synthetic and hybrid data. A general overview of these methods is as follows:

(i) Non-perturbative microdata masking is based on reduction of detail in data. Examples of t methods in this category are Sampling, Global recoding, Top and bottom coding, Local suppression. In Sampling is published only a sample of the original dataset. In Global recoding are generalized categories, which means creating new more general categories. Top and bottom coding is global recoding for ranked data.

Top values or bottom values are merged together, which means creating new more general top or bottom category. Local suppression means replacing risk values with missing values.

(ii) Perturbative microdata masking is based on perturbing data, which means distortion of the original data. Examples of the methods in this category are: Noise masking, Microaggregation, Data swapping and Rounding. Noise masking is based on adding noise into the data. It is a broad category with many approaches to adding noise. For example, normally distributed errors are added into the original data. Microaggregation is a broad category of methods for continuous microdata. Those methods are based on replacing individual values with values computed on small aggregates. This approach creates groups of  $k$  or more individuals, where no individual dominates the group and  $k$  is a threshold value. Data swapping is based on exchanging values of confidential variables between individual records. In Rounding are original values replaced with rounded values.

(iii) Synthetic and hybrid data are methods based on generating new data with the preservation of certain statistics or internal relationships of the original data set. Methods in this category are Fully synthetic data, Partially synthetic data and Hybrid data. In Fully synthetic data a completely new dataset which does not contain the original data is released. In Partially synthetic data only the most sensitive data are generated, and the rest of the original data are kept in the dataset. Hybrid data are the original data and synthetic data combined together.

### **1.3 Census recommendations**

This paper is focused on publication of microdata outputs from the population census. A comprehensive description of the methodology of the population census is provided in the United Nations (2017) and United Nations (2015a). Papers describing the content, methodology and preparation of the Census 2021 in the Czech Republic are following: Čtrnáct (2016), Sudková (2016), Škrabal et al. (2016), Škrabal (2017), Báčová (2018), Moravec (2018), and Kozelek (2019).

The recommendations for confidentiality and security of census data are described in United Nations (2006) and United Nations (2015a). The general aim of the census is to collect information on each person, household, and dwelling. The purpose is to provide information about the population, which can then be used to improve life in the country. In the use of these data, NSO are not so interested in information about each individual, but they gain valuable information about the structure of the population. However, in order to obtain reliable results, the confidentiality of the information collected about the respondents must be guaranteed. The security rules around the census carry out the entire operation from the actual data collection to its processing and subsequent publication to both the scientific community and the general public. The census is a unique project that allows the state to obtain detailed data down to the level of very little geographical detail.

In order to release the data obtained from the census, SDC methods have to be applied to all outputs, which prevent the disclosure of sensitive information about specific respondents. The goal of statistical protection of data confidentiality is to ensure maximum data utility while minimizing data information loss and maximizing the protection of published data. In case of the publishing of census microdata, there must be always removed direct identifiers such as name, addresses and personal identification numbers. Any unique variables, which could cause re-identification of any potential respondent, have to be perturbed by SDC methods, in order to minimize the risk of their disclosure. In the case of the census, the public confidence in the security and confidentiality of the information is primary and all operations are carried out to respect this precondition (United Nations, 2006; United Nations, 2015a).

### **1.4 International organizations**

Several international organizations are interested in the field of microdata dissemination. Mostly it is a recommendation of methods and procedures how microdata would be published. Important



documents of United Nations and European Union are described above. Another relevant organization is OECD, which created expert group for international collaboration on microdata access. In OECD (2014) is their summary of the methods, practices, and recommendations. Two important organizations are further described, which strive for the widest possible dissemination of microdata from population census, IPUMS and the Pacific Community.

#### 1.4.1 IPUMS

IPUMS (Integrated Public Use Microdata Series) is an organization which is mainly focusing on inventory, preservation, harmonization, and disseminating of census microdata from countries around the world. IPUMS is a part of the Institute for Social Research and Data Innovation at the University of Minnesota. According to their website,<sup>2</sup> they collaborate with 105 national statistical agencies, 9 national archives, and 3 genealogical organizations, which according to them created the world's largest database of census microdata. This database contains U.S. censuses from 1790 to the present and other international censuses of over 100 countries. IPUMS claims that the library has over a billion records. They stored in their library microdata describing 1.4 billion individuals which come from over 450 censuses and surveys.

Countries listed in Table 1 to Table 5 are the countries that can be found in their library and for clarity they were categorized geographically by continent.

**Table 1** Datasets of IPUMS in Europe

Austria	Belarus	Finland	France	Germany
Greece	Hungary	Ireland	Italy	Netherlands
Poland	Portugal	Romania	Russia	Slovakia
Slovenia	Spain	Switzerland	Ukraine	United Kingdom

Source: Own construction

**Table 2** Datasets of IPUMS in America

Argentina	Bolivia	Brazil	Canada	Chile
Colombia	Costa Rica	Cuba	Dominican Republic	Ecuador
El Salvador	Guatemala	Haiti	Honduras	Jamaica
Mexico	Nicaragua	Panama	Paraguay	Peru
Puerto Rico	Saint Lucia	Suriname	Trinidad and Tobago	United States of America
Uruguay	Venezuela			

Source: Own construction

**Table 3** Datasets of IPUMS in Asia

Armenia	Bangladesh	Cambodia	China	Indonesia
Iran	Iraq	Israel	Jordan	Kyrgyzstan
Laos	Malaysia	Mongolia	Myanmar	Nepal
Pakistan	Palestine	Philippines	Thailand	Turkey
Vietnam				

Source: Own construction

<sup>2</sup> <<https://ipums.org>>.

**Table 4** Datasets of IPUMS in Africa

Benin	Botswana	Burkina Faso	Cameroon	Egypt
Ethiopia	Ghana	Guinea	Kenya	Lesotho
Liberia	Malawi	Mali	Mauritius	Morocco
Mozambique	Rwanda	Senegal	Sierra Leone	South Africa
South Sudan	Sudan	Tanzania	Togo	Uganda
Zambia	Zimbabwe			

Source: Own construction

**Table 5** Datasets of IPUMS in Oceania

Fiji	Papua New Guinea
------	------------------

Source: Own construction

Table 30, which was placed in the Annex due to its size, then shows the complete range of microdata available to IPUMS library<sup>3</sup> on 1/3/2022. Picture 1 shows on a world map the countries that are represented in their database. The full-size map can be downloaded from the author's GitHub.<sup>4</sup> The differences between the table and the map are the countries Iceland, Norway, Sweden and Denmark, because these countries provide only historical values from the years before 1960.

**Picture 1** Map of member states of IPUMS

Source: Own construction

According to Sobek and Cleveland (2020) confidentiality of the data is ensured by application of suppression to very small categories and also application of top- and bottom-code thin tails of continuous variables. At geographical level they swap a small number of households to add an additional degree of uncertainty and for areas with fewer than 20 000 population in recent censuses are combined together with neighbouring units until the threshold is achieved.

<sup>3</sup> <<https://international.ipums.org/international-action/samples>>.

<sup>4</sup> <[https://github.com/Kyoshido/Population-census-microdata-availability/blob/main/map\\_IPUMS\\_03-2022.png](https://github.com/Kyoshido/Population-census-microdata-availability/blob/main/map_IPUMS_03-2022.png)>.



Microdata are available at their website,<sup>5</sup> but the World Bank library of microdata offers a better search environment at their website.<sup>6</sup> The microdata are free of charge, but the researcher must submit an application to use restricted microdata via an electronic authorization form identifying the user by name, electronic address, and institutional affiliation.

#### 1.4.2 The Pacific Community

The Pacific Community (SPC) is the scientific and technical organisation in the Pacific region. SPC division, which is responsible for the dissemination of the statistics data, is Statistics for Development Division<sup>7</sup> (SDD). SDD has provided the census support for its members since the 1990's and their technical support covers all aspects of the census cycle, from the questionnaire design and preparation of census cartography, to training of field staff, data processing, tabulation, analysis, and reporting and dissemination of results.

Member States of the Pacific Community, which are geographically located in Oceania are listed in Table 6.

**Table 6** Member states of The Pacific Community

Australia	American Samoa	Cook Islands	Federated States of Micronesia	Fiji
French Polynesia	Guam	New Zealand	Kiribati	Mariana Islands
Marshall Islands	Nauru	New Caledonia	Niue	Palau
Papua New Guinea	Pitcairn Islands	Samoa	Solomon Islands	Tokelau
Tonga	Tuvalu	Vanuatu	Wallis and Futuna	

Source: Own construction

The microdata are available on their website only if the country signs a Memorandum of Understanding and gives SPC the right to publish their data sets. In the case of census microdata, all datasets in the microdata library are set as No Access. Because SPC helped with the collection of the microdata, they include them in the list in the library and also hold the dataset, but SPC don't disseminate them. Researchers have to contact the country National Statistics office to get those data.

## 2 SURVEY ON THE PRACTICES

There are plenty of approaches on how to make microdata available. A comprehensive list of recommended ways can be found at United Nations (2007). Generally, the publication of microdata can be divided based on the audience for which it is intended, into PUF and SUF microdata. Public Use Files (PUFs) in the USA called Anonymised microdata files, are microdata that can be accessed by the general public. That is the reason why the highest level of security against disclosure of individual respondent personal data is necessary. Scientific use files (SUF) in the USA called Licensed anonymised microdata files, are microdata that can be accessed only by approved researchers, usually with a signed contract. This type of microdata needs a lower level of security than PUFs, but the personal information of respondents is still preserved against disclosure.

In order to find out the approaches of individual statistical offices the survey and following research were made, in which author was focused on the current situation and the plans following the data

<sup>5</sup> <<https://international.ipums.org/international>>.

<sup>6</sup> <<https://microdata.worldbank.org/index.php/catalog/ipums>>.

<sup>7</sup> <<https://sdd.spc.int>>.

protection in the European and non-European countries. Mentioned information is based on author's research and the information from the official websites. A survey (questions are in the Annex) was created and then sent to the countries, which are shown in Table 7 to Table 10 and are categorized by continent. The following countries were chosen to represent the European approach with the addition of important world statistical offices.

**Table 7** Surveyed countries of Europe

Albania	Austria	Belgium	Bulgaria	Croatia
Cyprus	Czech Republic	Denmark	Estonia	Finland
France	Germany	Greece	Hungary	Ireland
Israel	Italy	Latvia	Lithuania	Luxembourg
Malta	Montenegro	Netherlands	Poland	Portugal
North Macedonia	Romania	Russia	Serbia	Slovakia
Slovenia	Spain	Sweden	Switzerland	United Kingdom

Source: Own construction

**Table 8** Surveyed countries of America

Brazil	Canada	Mexico	USA
--------	--------	--------	-----

Source: Own construction

**Table 9** Surveyed countries of Asia

China	Indie	Japan	Turkey
-------	-------	-------	--------

Source: Own construction

**Table 10** Surveyed countries of Oceania

Australia	New Zealand
-----------	-------------

Source: Own construction

Surveyed countries were grouped by their approach to publishing of microdata. Created clusters with the information obtained were: 1) Publishing SUF, 2) Publishing SUF and PUF, 3) Not publishing, and 4) No data.

Author purposely distinguishes between not publishing and no data, because he wanted to provide complete information and with this distinction, he purposefully differentiates situations where author was sure that a given country intentionally does not publish data and situations where author does not have the opportunity to make an exact decision.

Initially, 45 countries were addressed in the questionnaire; however, while gathering information from the survey and researching literature for the paper, author found that it was an initial mistake to consider only a limited range of countries as he found out that other countries also had something to offer in terms of microdata publication and that the original intention would capture only limited information about dissemination of microdata from the population census.

Thus, author decided to extend the research from the limited number of statistical offices to the whole world. Therefore, the results give general information about all the census microdata that can be obtained.

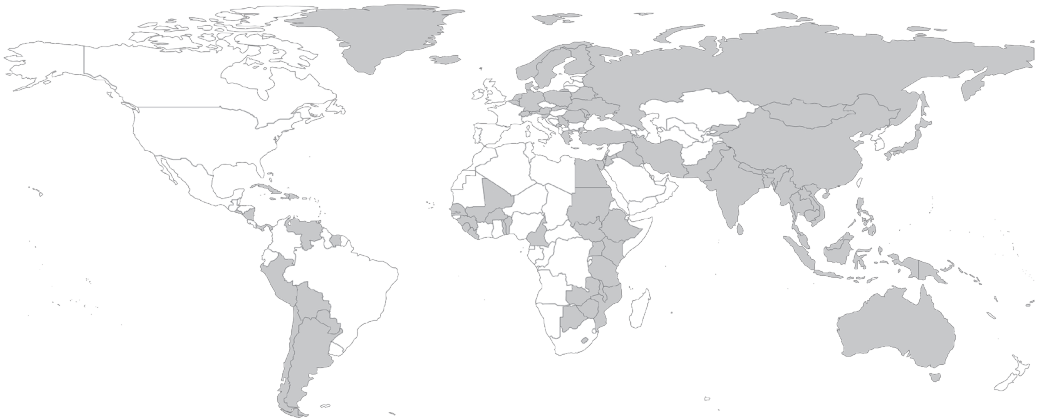
## 2.1 Publishing SUF

As the first one is presented countries that publish only census microdata for scientific use purposes (SUF). Countries that were categorized it this cluster are listed in Table 11 to Table 15 and are categorized by continent. Not all countries are described in the text below. This is because if a country provides its microdata via IPUMS and does not provide additional useful information, it has been omitted from the detailed description.

Further, the individual countries and their approaches to publishing SUF microdata are described. For clarity, a form was chosen where the links to individual offices and to microdata are listed in the Annex (Tables 31 to 33).

Picture 2 shows on a world map the countries that publish SUF microdata. Full size map can be downloaded from the author's GitHub.<sup>8</sup>

**Picture 2** Map of countries publishing SUF



Source: Own construction

### 2.1.1 Countries of Europe publishing SUF microdata

The part further below describes countries only publishing SUF microdata in Europe. Not all countries in Table 11 are described in the text below. This is because if a country provides its microdata via IPUMS or does not provide additional useful information, it is omitted from the detailed description.

**Table 11** Countries of Europe publishing SUF microdata

Austria	Belarus	Belgium	Denmark	Finland
France	Faroe Islands	Germany	Greece	Iceland
Luxembourg	Malta	Montenegro	Moldova	Netherlands
Norway	Poland	Romania	Russia	Serbia
Slovakia	Slovenia	Sweden	Switzerland	Ukraine

Source: Own construction

<sup>8</sup> <[https://github.com/Kyoshido/Population-census-microdata-availability/blob/main/map\\_SUF\\_03-2022.png](https://github.com/Kyoshido/Population-census-microdata-availability/blob/main/map_SUF_03-2022.png)>.

**Belgium**

The scientific use of microdata is subject to the prior approval of the Commission for the Protection of Privacy. Researchers interested in census microdata have to submit their data request to the data protection officer. The data request must comply with the principles of finality and proportionality as set out in Belgian legislation. However, Belgium NSO provides only aggregated data with the lowest territorial level being a building. To ensure the confidentiality of the personal information of the enumerated persons, the Cell Suppression Method is applied.

**Denmark**

Since 1981 Denmark has had annual censuses based on their register data. Access to microdata can only be granted to researchers and analysts in Danish research environments on the research servers after approval from Denmark NSO.

**Germany**

The legislation that covers the confidentiality of data is Federal Statistics Law (BStatG). Germany NSO do not publish a sample of microdata from the census, instead they have a different approach when they publish the so-called Microcensus, which is a census executed on the representative sample (one percent) of the population, about 370 000 households with 810 000 household members. Microdata are available to researchers on site through Safe Centre (GWAP) or Remote Execution (KDFV) at Research Data Centre of the Federal Statistical Office. Available SUF is anonymised 70% subsample of the households from the Microcensus.

**Luxembourg**

To access microdata, researchers have to sign a formal convention between the researcher, the institution and Luxembourg NSO. After the convention is signed, Luxembourg NSO would send a sample (5% of the data) to the researcher to prepare his work, especially the syntax, which is sent to Luxembourg NSO, who will then check the results if the confidentiality of the data is respected. Luxembourg NSO does not have the SDC model to protect the data confidentiality, but for detailed data they do not provide data with less than 3 observations.

**Montenegro**

Statistics to researchers are provided under the Law on Official Statistics and Official Statistical System. To obtain access to individual data without identifiers, researchers need to send the license/proof document that your institution is scientific, and the research institution issued by a licensed institution. After that a commission will decide on the provision of individual data.

**Norway**

Researchers who are affiliated with an approved research institution, or a public authority, can apply to access microdata. The census is not specifically mentioned in their available datasets, but Norway NSO provides data on population or families and households, immigration, and other demographic information, which are updated annually from registers.

**Slovakia**

Slovakia NSO disseminates microdata with 28 variables, and more detail can be found at their website, which is listed in Table 31, but only in the Slovakian language. An English version of their websites offers much less data, and the information about microdata access is not visible. The Slovakian version

informs that the data are available only for scientific and research purposes. Those interested can request the information service of the Statistical Office of the Slovak Republic for selected microdata.

### *Sweden*

Sweden NSO's platform for access to microdata is called MONA (Microdata Online Access). Users can log in through Security card or Smart phone. Data are delivered in SQL format from STATA, SAS, SPSS or R.

### **2.1.2 Countries of Americas publishing SUF microdata**

Below are listed countries only publishing SUF microdata in the Americas. Countries in Table 12 are not further described in the text below. This is because if a country provides its microdata via IPUMS or does not provide additional useful information, it is omitted from the detailed description.

Argentina	Bolivia	Brazil	Canada	Chile
Colombia	Costa Rica	Cuba	Dominican Republic	Ecuador
El Salvador	Greenland	Haiti	Jamaica	Nicaragua
Panama	Paraguay	Peru	Saint Lucia	Suriname
Trinidad and Tobago	Venezuela			

Source: Own construction

### **2.1.3 Countries of Asia publishing SUF microdata**

In the part further below are described countries only publishing SUF microdata in Asia. Not all countries in Table 13 are described in the text below. This is because if a country provides its microdata via IPUMS or does not provide additional useful information, it is omitted from the detailed description.

Armenia	Bangladesh	Bhutan	Cambodia	China
India	Indonesia	Iran	Iraq	Israel
Japan	Jordan	Kyrgyzstan	Laos	Malaysia
Maldives	Mongolia	Myanmar	Nepal	Pakistan
Palestine	Philippines	Sri Lanka	Thailand	Turkey
Vietnam				

Source: Own construction

### *India*

Researchers can access the data under secure environment of the Workstation for Research on Sample Micro-Data from Census. Confidentiality of microdata is secured by the law in The Census Act 1948. 1% and 5% sample of micro-data from the 2001 Census and 2011 Census are available for research to Universities or Institutes. Available are SPSS and Stata software.

### *Japan*

For the dissemination of the microdata, they established the Micro Data Usage Portal Site (Miripo). Miripo website is listed in Table 32. Microdata from the years 1960, 1980 and 1990 are provided

on a magnetic medium and are in TXT format. Microdata from the years 2005, 2010 and 2015 are available only on-site.

### ***Maldives***

Request to access the Maldives census 2014 micro dataset can be found at their website, which is listed in Table 32. The dataset will be made available to users after submitting the form and once the application has been processed a Memorandum of Understanding (MoU) will be signed between NBS and the user.

### ***Sri Lanka***

Microdata are disseminated in compliance with the Statistics Law of Sri Lanka. Researchers must submit an application which can be found at their website, which is listed in Table 32.

## **2.1.4 Countries of Africa publishing SUF microdata**

In following part are described countries only publishing SUF microdata in Africa. Countries in Table 14 are not further described in the text below. This is because if a country provides its microdata via IPUMS or does not provide additional useful information, it is omitted from the detailed description.

**Table 14** Countries of Africa publishing SUF microdata

Benin	Botswana	Burkina Faso	Cameroon	Egypt
Ethiopia	Ghana	Guinea	Kenya	Lesotho
Liberia	Malawi	Mali	Mauritius	Mozambique
Senegal	Sierra Leone	South Sudan	Sudan	Tanzania
Togo	Uganda	Zambia	Zimbabwe	

Source: Own construction

## **2.1.5 Countries of Oceania publishing SUF microdata**

In following part are described countries only publishing SUF microdata in Oceania. Not all countries in Table 15 are described in the text below. This is because if a country provides its microdata via IPUMS or does not provide additional useful information, it is omitted from the detailed description.

**Table 15** Countries of Oceania publishing SUF microdata

Australia	Fiji	Papua New Guinea
-----------	------	------------------

Source: Own construction

### ***Australia***

Confidentiality of microdata is secured in Australia by their law under the Census and Statistics Act 1905, Privacy Act 1988 and Census and Statistics (Information Release and Access) Determination 2018. From the 2016 Census are available the following microdata: 1% sample Basic CURF (Confidentialised Unit Record File), 5% Expanded CURF. CURF is the regional name for SUF.

The 1% Basic CURF contains data on 87 798 dwellings, 93 002 families and 215 597 persons. It provides a sample of one private dwelling record in every hundred from the census, and the associated family and person records. Dwellings with more than six usual residents were removed from the sample to ensure confidentiality of large dwellings. For non-private dwellings the sampling is applied to persons present,

where one person in every hundred is selected and the associated dwelling records included on the file. Data are available on CD-ROM and through the Remote Access Data Laboratory or the Data Laboratory.

The 5% Expanded CURF contains data on 422 725 dwellings, 450 038 families and 1 083 585 persons. It provides a sample of one private dwelling in every twenty from the Census, and the associated family and person records. Dwellings with more than eight usual residents were also removed from the sample to ensure confidentiality of large dwellings. For non-private dwellings the sampling is applied to persons present, where five persons in every hundred are selected and the associated dwelling records included on the file. Data are available through the Remote Access Data Laboratory or the Data Laboratory. Both the 1% and 5% CURFs are available in SAS, SPSS and STATA formats. Both CURF files are not available on CD-ROM to overseas customers.

## 2.2 Publishing SUF and PUF

Here are presented countries that publish census microdata not only for scientific purpose (SUF) but also for public (PUF). Countries that were categorized in this cluster are listed in Table 16 to Table 20 and are categorized geographically by continent. Picture 3 shows on a world map the countries that publish SUF and PUF microdata. A full size map can be downloaded from the author's GitHub.<sup>9</sup> Picture 3 shows on a world map the countries that publish PUF microdata.

Picture 3 Map of countries publishing SUF and PUF



Source: Own construction

Further, the individual countries and their approaches to publishing SUF and PUF microdata are described. For clarity, a form was chosen where the links to individual offices and to microdata are listed in the Annex (Tables 34 to 37).

### 2.2.1 Countries of Europe publishing SUF and PUF microdata

In following part are described countries publishing SUF and PUF microdata in Europe. Countries in Table 16 are described in the text below.

<sup>9</sup> <[https://github.com/Kyoshido/Population-census-microdata-availability/blob/main/map\\_PUF\\_03-2022.png](https://github.com/Kyoshido/Population-census-microdata-availability/blob/main/map_PUF_03-2022.png)>.



**Table 16** Countries of Europe publishing SUF and PUF microdata

Albania	France	Hungary	Ireland	Italy
Portugal	Spain	United Kingdom		

Source: Own construction

### ***Albania***

The legislation that covers the confidentiality of data is Law No 9888, dated 1.3.2008 "On the Protection of Personal Data" and Law No 17/2018 "On official statistics". The PUF microdata from the Population and Housing Census 2011 can be found at their website, which is listed in Table 33. Data can be downloaded directly without the need for registration or any control.

In the explanatory note to the PUF micro data there is written that they provide 3% of households and the sample technique used to select the sample is a simple random sample. The sample from Population and Housing Census 2011, which was randomly selected, was 21 665 households from 722 226 households in the whole country. In their opinion this method ensures that the sample is representative not just at national level, but also at prefecture level. The data are provided in SPSS format. The methods used to protect confidentiality of the census microdata are resampling and perturbative methods.

### ***France***

The processing of personal data for statistical purposes complies with the law Informatique et Libertés. The PUF harmonized microdata from the Population and Housing Census from 1968 to 2017 can be found at their website, which is listed in Table 33. The file contains 18 variables and 52 448 313 records. This dataset has no explanations and documentation in English, it is only available in French. However, it is not the only one, they have 7 126 datasets with English explanations in their database, while they have 23 465 datasets in French. Data can be downloaded directly without the need for registration or any control. The data are provided in dBase and CSV format. The file is created as a file of cumulative individuals, so the observations having the same modalities for all the variables have been grouped, therefore it's essential to use a weighting (variable POND) to recreate population.

### ***Hungary***

Hungary NSO is providing PUF files in form of Test files. Test files have the same structure as the microdata sets, which are available internally, but do not reflect the relationships between the variables. That means the files are not determined for analysis. Their purpose is to provide preparatory support for the researchers in accessing data in the Safe Centre and remote execution. Thus, the researcher can pre-prepare his code in advance of the test data at rest and then apply this code in a secure environment to microdata already suitable for the given analyses. Each record is fictitious and the logical correlations between the variables are not fulfilled either. The Test files PUF microdata from the Population and Housing Census 2011 can be found at their website, which is listed in Table 33.

### ***Ireland***

The legislation that covers the confidentiality of data is the Statistics Act, 1993. Ireland NSU provides access to two types of microdata files: Anonymised Microdata Files (AMFs), which are PUFs and Research Microdata Files (RMFs), which are SUFs.

Access to AMFs must be approved in advance by Ireland NSU. 5% of anonymised samples of the population census from 1996, 2002, 2006 are available. More information here can be found at their website, which is listed in Table 33.

Access to RMF's is strictly controlled and can be accessed remotely via CSO Researcher Data Portal (RDP) or on-site in a Ireland NSU via the Researcher Data Portal (RDP). Recently in June 2021 Ireland NSU has allowed access to Researcher Microdata Files from home offices. They achieved it by adding an extra layer of security provided by two-factor authentication (2FA).

### ***Italy***

Data are protected by the Legislative Decree No. 322 of 6.9.1989. The PUF microdata from the Population and Housing Census 2011 can be found at their website, which is listed in Table 33. The data represent a 1% sample and are provided in TXT format. Data can be downloaded after User Authentication through registration to Italy NSO single sign-on system.

### ***Portugal***

The PUF microdata from the Population and Housing Census 1981, 1991, 2011, 2011 can be found at their website, which is listed in Table 33. The microdata represents a 5% sample on individuals and dwellings. The data are provided in Microsoft Access Database. Data can be downloaded directly without the need for registration or any control, just acceptance of use conditions is necessary.

### ***Spain***

The data are provided in ASCII format. Data can be downloaded directly without the need for registration or any control. The PUF microdata from the Population and Housing Census 1991, 2001 and 2011 can be found at their website, which is listed in Table 33. The following SDC methods were used to secure confidentiality. Variables of municipalities with fewer than 20 000 inhabitants have to be recorded: place of residence, place of birth, previous place of residence, place of residence 1 year ago, place of residence 10 years ago, place of second home, place of work/study. To protect people who work in occupations related to the Armed Forces, they were moved to other categories.

### ***United Kingdom***

Data confidentiality is protected by the law Statistics and Registration Service Act 2007 and Data Protection Act. ONS has a facility called the Secure Research Service which is providing access for approved researchers to restricted microdata. England and Wales microdata samples, two Northern Ireland microdata samples, and two Scottish microdata samples, which are created from 10% of people or households in the 2011 Census are available. From statistical disclosure control methods applied to 2011 Census data were targeted record swapping and restriction of detail.

United Kingdom NSO also publishes Microdata Teaching File, which is the PUF microdata from the Population and Housing Census 2011 and can be found at their website, which is listed in Table 33. The Microdata Teaching File contains a 1% sample of people with just a small number of characteristics. Its purpose is to serve as an educational tool to assist with the teaching.

## **2.2.2 Countries of Americas publishing SUF and PUF microdata**

In following part are described countries publishing SUF and PUF microdata in America. Countries in Table 17 are described in the following text.

### ***Brazil***

The PUF microdata from the Population and Housing Census 2010 can be found at their website, which is listed in Table 34. The microdata are in ASCII format. Data can be downloaded directly without the need for registration or any control.

**Table 17** Countries of Americas publishing SUF and PUF microdata

Brazil	Canada	Chile	Colombia	Costa Rica
Ecuador	Guatemala	Honduras	Mexico	Puerto Rico
Uruguay	USA			

Source: Own construction

### ***Canada***

Data are protected by laws – the Statistics Act and the Privacy Act. The PUF microdata from the Population and Housing Census 2016 can be found as Hierarchical File and as Individuals File at their website, which is listed in Table 34. Hierarchical File provides access to non-aggregated data covering a sample of 1% of the Canadian households, 140 705 household records, which are representing 343 330 persons that have been anonymized. The Individuals File contains a 2.7%, sample of anonymous responses, 930 421 individual records, to the 2016 Census questionnaire. The researcher must complete an order, where he describes his intentions, to get the data. The data are provided in ASCII format and SAS, SPSS or Stata format program source codes.

### ***Chile***

The PUF microdata from the Population and Housing Census 2017 can be found at their website, which is listed in Table 34. The data are provided in CSV format. Data can be downloaded directly without the need for registration or any control.

### ***Colombia***

Data are protected by Law 79 of 1993 Article 5: the PUF microdata from the Population and Housing Census 2018 can be found at their website, which is listed in Table 34. Data can be downloaded directly without the need for registration or any control.

### ***Costa Rica***

The PUF microdata from the Population and Housing Census 2011 can be found at their website, which is listed in Table 34. Microdata is covering a sample of 10%, 427 972 records. To get the microdata you need to register to their website and then just fill in the reason and the intended purpose of what you want to do with the microdata. After that, you can download the microdata without any control of your text admission. The data are provided in SPSS format.

### ***Ecuador***

The PUF microdata from the Population and Housing Census 2010 can be found at their website, which is listed in Table 34. Data can be downloaded directly without the need for registration or any control. The data are provided in SPSS format.

### ***Guatemala***

The PUF microdata from the Population and Housing Census 2018 can be found at their website, which is listed in Table 34. The microdata are in CSV and SPSS format. Data can be downloaded directly without the need for registration, only control is reCaptcha.

### ***Honduras***

The PUF microdata from the Population and Housing Census 2013 can be found at their website, which is listed in Table 34. Unfortunately, the link to download the data does not work because the data are stored on external storage. The link to these databases will transfer researcher to the main page.

### ***Mexico***

The PUF microdata from the Population and Housing Census 2017 can be found at their website, which is listed in Table 34. Mexico NSO provides various files of microdata: files of Census (basic questionnaire) and files of Sample (extended questionnaire). Census (basic questionnaire) files provides examples of the Basic Questionnaire database, while ensuring that any type of inference can't be made. Their task is to show characteristics and for users to test their syntax before sending it to be processed through the sections of: Microdata Laboratory, Remote Processing and Processing Service. Sample (extended questionnaire) files provides the results derived from the Extended Questionnaire on the characteristics of inhabited private housing units and their occupants. On these data researchers can perform their analyses.

The data are provided in CSV format. Data can be downloaded directly without the need for registration or any control.

### ***United States of America***

By United States of America (USA) law, more precisely by Title 13 of the United States Code, they are obligated to ensure that private information about any specific individual, household, or business is never published and revealed. The results of their census are of great political importance for the USA as they are used to determine the number of seats in their House of Representatives and further determine the size of the legislative districts from the congress to the city councils.

The USA NSO allows researchers to access microdata through a nationwide network of secure Research Data Centres which they created by partnership with various universities, non-profit research institutions, and government agencies. In these Data Centres, researchers can access microdata in the form of restricted use files. As the USA NSO states on its data centre website, research based on micro-data is crucial for them as it also provides them with feedback on the quality of the data and provides them with feedback on the strengths and weaknesses of the microdata records.

In dissemination of the Public Use Microdata (PUM) USA's NSO distinguish between Stateside and Island Areas. Island Areas are Guam and Virgin Islands. The PUM for Island Areas from the Population and Housing Census 2010 can be found at their website, which is listed in Table 35. Stateside areas are 51 states plus Puerto Rico. The PUM for Stateside areas from the Population and Housing Census 2010 can be found at their website, which is listed in Table 34. Each country has its own PUM file. Data can be downloaded directly without the need for registration or any control, The data are provided in TXT format. The disclosure avoidance methods which they used to protect PUM were Data swapping, Synthetic data, Top-coding and bottom-coding, Age perturbation, Reduced detail for categorical variables. Minimum population threshold is set on 100 000. They provide a 10% sample of the population.

In terms of future of statistical disclosure control applied on census data, their researchers are fully aware of the increasing possibilities of growing computer power in combination with progress which was done in the field of mathematic and its possible misuse to compromise and disclose private data. Because of this, they moved in Census 2020 from classical methods to the new concept called "differential privacy", which was firstly introduced in Dwork et al. (2006) and protect every record in their database with this new approach. The Census Bureau was the first one which used this method to disseminate its data. Differential privacy was also implemented by private industries such as Google in their browser Chrome, Uber in their app, Microsoft in their operating system Windows 10 and Apple in their product iPhone.

### ***Uruguay***

The PUF microdata from the Population and Housing Census 2011, 2004, 1996, 1985, 1975, 1963 can be found at their website, which is listed in Table 34. The data are provided in SPSS and DBF format. Data can be downloaded directly without the need for registration or any control.

### 2.2.3 Countries of Asia publishing SUF and PUF microdata

In following part are described countries publishing SUF and PUF microdata in Asia. Countries in Table 18 are described in the text below.

**Table 18** Countries of Asia publishing SUF and PUF microdata

Armenia	South Korea
---------	-------------

Source: Own construction

#### Armenia

The PUF microdata from the Population and Housing Census 2014 can be found at their website, which is listed in Table 35. Data can be downloaded directly without the need for registration or any control.

To maintain confidentiality, they made the following changes to the data. The sample from the census contains every tenth household, and they have chosen only households consisting of not more than 20 members and at least with one of them with a status of permanent inhabitant.

Because of small numbers in some variables, they had altered the data as follows. Variable Age had the maximum limit changed to 90 years, so the ages above 90 are equated to 90. In variable Country of Citizenship, they changed countries with less than 5 cases to the group "other". In variable Nationality, they changed the cases of less than 20 to the group "other". In variable Mother tongue, they changed the cases of less than 5 to the group "other". In variable Religion, they changed the cases of less than 10 to the group "other". Variable Born and alive children to mothers had changed the cases of more than 7 to 7. Variable Economic activity type and occupation are presented in two-digit and one-digit codes, and the occupation "servicemen" is recorded as "not applicable". For variable household living conditions, they had changed the cases of more than 5 rooms to 5, and the cases of more than 200 sq/m dwelling space were changed to 200. Their 2011 census sample data file is a text file, which is created by software CSPro.

#### South Korea

South Korea NSO run a service called Microdata Integrated Service (MSI), which allows Download, Remote Access service, Microdata Research Center Service, On-Demand Service. The English version of MSI page contains only general information. In order to download data or obtain information, it is necessary to switch to Korean. The PUF microdata from the Population and Housing Census 2015 can be found at their website, which is listed in Table 35. They provide a 2% sample of the population. To download data it's necessary to subscribe to Statistics Korea ONE-ID, which is only for Koreans.

### 2.2.4 Countries of Africa publishing SUF and PUF microdata

In following part are described countries publishing SUF and PUF microdata in Africa. Countries in Table 19 are described in the text below.

**Table 19** Countries of Africa publishing SUF and PUF microdata

Angola	Burundi	Ghana	Morocco	Namibia
Rwanda	Somalia	South Africa		

Source: Own construction

#### Angola

The PUF microdata from the Population and Housing Census 2014 can be found at their website, which is listed in Table 36. Data can be downloaded directly without the need for registration or any control.

***Burundi***

The PUF microdata from the Population and Housing Census 2008 can be found at their website, which is listed in Table 36. To get the microdata you need to register to their website and then just fill in the reason and the intended purpose of what you want to do with the microdata. After that, you can download the microdata without any control of your text admission. The data are provided in SPSS format.

***Ghana***

The PUF microdata from the Population and Housing Census 2010 can be found at their website, which is listed in Table 36. To get the microdata researcher need to register to their website and then just fill in the reason and the intended purpose of what researcher want to do with the microdata. After that, anybody can download the microdata without any control of their text admission. The data are provided in SPSS format.

***Morocco***

The PUF microdata from the Population and Housing Census 2014 can be found at their website, which is listed in Table 36. The data are provided in STATA, SPSS and TXT format. Data can be downloaded directly without the need for registration or any control.

***Namibia***

The PUF microdata from the Population and Housing Census 2011 can be found at their website, which is listed in Table 36. The dataset is available to everyone after registering in the system. Microdata were sampled based on a stratified random sample. The stratum was at the constituency and urban/rural levels, the threshold for sampling within stratum in file is 250 households. In other cases a simple random sampling was used for each stratum at a 20% sample rate. The microdata includes 93 674 housing units and 418 362 people.

***Rwanda***

The PUF microdata from the Population and Housing Census 2012 can be found at their website, which is listed in Table 36. To download the data researchers, have to register at their website. Microdata are created as 10% equal probability sample from the population.

***Somalia***

In Somalia is available Population Estimation Survey which is a first milestone reached towards implementing a full and comprehensive population and housing census. The aim of the analysis was to count Somali Population to know size of population living with cities. To get the microdata researcher need to register to their website and then just fill in the reason and the intended purpose of what researcher want to do with the microdata. After that, anybody can download the microdata without any control of their text admission. The data are provided in SPSS format. Population Estimation Survey can be found at website, which is listed in Table 36.

***South Africa***

The PUF microdata from the Population and Housing Census 2011 can be found at their website, which is listed in Table 36. The microdata represents a 10% sample of the population.

***2.2.5 Countries of Oceania publishing SUF and PUF microdata***

In following part are described countries publishing SUF and PUF microdata in Oceania. Countries in Table 20 are described in the text below.

**Table 20** Countries of Oceania publishing SUF and PUF microdata

New Zealand

Source: Own construction

***New Zealand***

New Zealand NSO have a large research database called Integrated Data Infrastructure (IDI) which contains microdata of people and households. Available are Census data from 2013 and 2018. Data can be accessed only in their secure virtual environment, in approved facilities (the Data Lab). Outside of the Data Lab environment researchers can apply for confidentialised unit record files (CURFs). The application is listed in a Table 37. CURFs are microdata that were created to protect confidentiality and maintain integrity of the data. They can therefore be classified as SUF microdata. SDC methods which they are using are top-coding, data swapping, and collapsing categorical variables to the unit records. Available are Census data from 2001 and 2013. After the approval, the microdata can be downloaded.

**2.3 Not publishing**

Countries that do not publish microdata from the population and housing census are listed here. These are countries that have openly stated in the survey that their country does not publish this data, it is clear from their website that they do not publish census microdata, or their datasets are not part of the IPUMS microdata library. These countries are listed in Tables 21 to 25.

**Table 21** Countries of Europe not publishing microdata

Bosnia and Herzegovina	Bulgaria	Croatia	Cyprus	Czech Republic
Estonia	Gibraltar	Guernsey	Kosovo	Latvia
Liechtenstein	Monaco	North Macedonia	San Marino	Vatican

Source: Own construction

**Table 22** Countries of Americas not publishing microdata

Anguilla	Antigua and Barbuda	Aruba	Bahamas	Barbados
Belize	Bermuda	Cayman Islands	Guyana	Montserrat
Saint Kitts and Nevis	Saint Martin	Saint Vincent and the Grenadines	Sint Maarten	

Source: Own construction

**Table 23** Countries of Asia not publishing microdata

Afghanistan	Azerbaijan	Bahrain	Brunei	Georgia
Hong Kong	Lebanon	Macau	Nigeria	Oman
Qatar	Singapore	Syria	Taiwan	Tajikistan
Turkmenistan	United Arab Emirates	Uzbekistan	Yemen	

Source: Own construction



**Table 24** Countries of Africa not publishing microdata

Algeria	Cape Verde	Comoros	Eritrea	Gambia
Guinea-Bissau	Libya	Mauritania	Mayotte	Saint Helena
Sao Tome and Principe	Swaziland			

Source: Own construction

**Table 25** Countries of Oceania not publishing microdata

American Samoa	Cocos Islands	Christmas Island
----------------	---------------	------------------

Source: Own construction

## 2.4 No data

A special table was designed for those countries in respect of which author does not have enough information about their approach to access to statistical data. These are the countries that have not responded to the survey or the author could not find any information about their approach to the publication of microdata on their websites or their datasets are not part of the IPUMS microdata library. These countries are listed in Tables 26 to 29.

**Table 26** Countries of Europe with no data

Andorra	Lithuania
---------	-----------

Source: Own construction

**Table 27** Countries of America with no data

Dominica	Falkland Islands	Grenada	Haiti
----------	------------------	---------	-------

Source: Own construction

**Table 28** Countries of Asia with no data

East Timor	Kazakhstan	Kuwait	North Korea	Saudi Arabia
------------	------------	--------	-------------	--------------

Source: Own construction

**Table 29** Countries of Africa with no data

Central African Republic	Chad	Democratic Republic of the Congo	Djibouti	Equatorial Guinea
Gabon	Ivory Coast	Madagascar	Niger	Republic of the Congo
Seychelles	Tunisia			

Source: Own construction

## CONCLUSION

Aim of in this paper was to examine the approaches of individual European statistical offices and other statistical offices around the world to the publication of microdata. The main goal was to compare individual approaches between countries. The primary interest of this work was to summarize whether

NSO publish microdata at all and if so what approach and models did they choose to protect the personal data of their respondents.

At the beginning, author was focused only on a selected sample of statistical offices. During the collection of information on available microdata, author decided to expand the field of processing to the whole world and provide complete information on all possible published microdata from population censuses. The most common problem, which was encountered during processing, was non-response by statistical offices and also the language barrier. Fortunately, the language barrier was overcome with the help of Google Translator, which was an invaluable helper during the writing of this paper. Many NSO do not have their pages and documents translated into English. Complete translated webpages are a rare phenomenon and usually researchers will find only a part of websites in English.

Countries were divided according to their approach to microdata publishing into the following categories: 1) countries that publish only for the scientific community, 2) countries that also provide microdata to the public and 3) countries that do not publish microdata at all. The third group is further divided into two sub-groups. Author distinguishes here between countries for which he was sure that NSO do not publish microdata and NSO for which author did not have enough information to make a final decision on categorization of the approach. This sub-group can be considered as a non-publishing group, but for the sake of clarity, it was decided to create a separate subgroup.

International organizations play an important role in the dissemination of microdata. These international organizations help their members with the preparation of surveys and censuses, as well as with the processing. The most important organization is IPUMS from the United States, which dedicates its existence to the harmonization, collection, and dissemination of microdata to the scientific community. In their microdata library, they made available census data from 96 countries. Thanks to the IPUMS organization, there is a large number of countries that publishes SUF microdata, but they do not have all countries that publish in their library. This is because population census microdata is a very sensitive topic and not all countries are willing to pass this data to a foreign institution located in another country. In total, there are 100 countries that publish SUF microdata.

Countries that provide PUF microdata are much more uncommon. In total, there are 30 countries, but there are large differences between them. The differences consist in the degree of control of microdata access when on the one side there are completely free census microdata that can be downloaded by anyone without any control and on the other side, there is maximum protection, and microdata can be downloaded only if you are a citizen of the state and have the appropriate identifiers.

This paper furthermore provides three maps that were created for countries in IPUMS library, countries that publish SUF and countries that publish PUF. Author hopes that these maps will provide a better overview of the differences between countries. In future research, the aim is to learn more from the findings on the publication of microdata from other statistical offices and enable access to microdata from the population census for the scientific community in the Czech Republic.

## ACKNOWLEDGMENTS

This paper has been prepared with the support of a project of the Prague University of Economics and Business – Internal Grant Agency, project No. F4/50/2021.

## References

- 
- BÁČOVÁ, P. (2018) Příští sčítání bude provázet mnoho novinek [online]. *Statistika&My*, 8(6). [cit. 1.7.2021]. <<https://www.statistikaamy.cz/wp-content/uploads/2018/06/18041806.pdf>>.
- ČERNÝ, P. (2021a). K čemu je dobré sčítání [online]. *Statistika&My*, 11(1). [cit. 1.7.2021]. <[https://www.statistikaamy.cz/wp-content/uploads/2021/02/01\\_21\\_SaM.pdf](https://www.statistikaamy.cz/wp-content/uploads/2021/02/01_21_SaM.pdf)>.

- ČERNÝ, P. (2021b). Kde pomohla data ze sčítání [online]. *Statistika&My*, 11(1). [cit. 1.7.2021]. <[https://www.statistikaamy.cz/wp-content/uploads/2021/02/01\\_21\\_SaM.pdf](https://www.statistikaamy.cz/wp-content/uploads/2021/02/01_21_SaM.pdf)>.
- ČTRNÁCT, P. (2016). Příprava světových sčítání kolem roku 2020 pokračuje [online]. *Demografie*, 58(2). [cit. 1.7.2021]. <<https://www.czso.cz/documents/10180/33199357/SLDB.pdf/6d6e4dca-fce8-4d1e-b09c-f464905bf4aa?version=1.0>>.
- DWORK, C., MCSHERRY, F., NISSIM, K., SMITH, A. (2006). *Calibrating Noise to Sensitivity in Private Data Analysis* [online]. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer Berlin Heidelberg. [cit. 1.7.2021]. <[https://doi.org/10.1007/11681878\\_14](https://doi.org/10.1007/11681878_14)>.
- EUROSTAT. (2018). *European Statistics Code of Practice: for the National Statistical Authorities and Eurostat (EU Statistical Authority)* [online]. Luxembourg: Publications Office of the European Union. [cit. 1.7.2021]. <<https://ec.europa.eu/eurostat/documents/4031688/8971242/KS-02-18-142-EN-N.pdf/e7f85f07-91db-4312-8118-f729c75878c7?t=1528447068000>>.
- HUNDEPOOL, A., DOMINGO-FERRER, J., FRANCONI, L., GIESSING, S., LENZ, R., LONGHURST, J., NORDHOLT, E. S., SERI, G., DE WOLF, P. P. (2010). *Handbook on Statistical Disclosure Control* [online]. ESSNet. [cit. 1.7.2021]. <[https://ec.europa.eu/eurostat/cros/system/files/SDC\\_Handbook.pdf](https://ec.europa.eu/eurostat/cros/system/files/SDC_Handbook.pdf)>.
- HUNDEPOOL, A. (2012). *Statistical disclosure control*. Wiley series in survey methodology, Chichester, West Sussex, United Kingdom: Wiley.
- KOZELEK, V. (2019). Územní příprava sčítání se neobejde bez spolupráce s obcemi [online]. *Statistika&My*, 9(9). [cit. 1.7.2021]. <<https://www.statistikaamy.cz/wp-content/uploads/2019/03/18041903.pdf>>.
- OECD (2014). *Final report* [online]. OECD expert group for international collaboration on microdata access, Paris: OECD. [cit. 1.7.2021]. <<https://www.oecd.org/sdd/microdata-access-final-report-OECD-2014.pdf>>.
- MORAVEC, Š. (2018). Dvoustanné pracovní jednání ČSÚ a ŠÚSR k přípravě sčítání lidu, domů a bytů v roce 2021 [online]. *Demografie*, 60(1). [cit. 1.7.2021]. <<https://www.czso.cz/documents/10180/61449042/scitani+lidu.pdf/cd6f2eee-615a-47cb-bd2b-138a5faa8034?version=1.0>>.
- Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics and repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom establishing a Committee on the Statistical Programmes of the European Communities [online]. European Parliament, Council of the European Union, OJ L 87, 31.3.2009: 164–173. [cit. 1.7.2021]. <<http://data.europa.eu/eli/reg/2009/223/oj>>.
- Regulation (EU) No 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 (General Data Protection Regulation) [online]. European Parliament, Council of the European Union, OJ L 119, 4.5.2016: 1–88. [cit. 1.7.2021]. <<https://eur-lex.europa.eu/eli/reg/2016/679/oj>>.
- SOBEK, M., CLEVELAND, L. (2020). *Building Research Infrastructure for Harmonized International Census Microdata* [online]. IPUMS Working Paper Series. [cit. 1.7.2021]. <<https://doi.org/10.18128/IPUMS2020-01>>.
- SUDKOVÁ, E. (2016). Konzultace s uživateli dat o obsahu Sčítání lidu, domů a bytů v roce 2021 [online]. *Demografie*, 58(3). [cit. 1.7.2021]. <<https://www.czso.cz/documents/10180/33199355/scitani+lidu.pdf/88b2d15c-1387-4a20-b630-3ca726696d57?version=1.0>>.
- ŠKRABAL, J., ŠANDA, R., HABARTOVÁ, P. (2016). Současný stav přípravy sčítání lidu, domů a bytů v roce 2021 [online]. *Demografie*, 58(1). [cit. 1.7.2021]. <<https://www.czso.cz/documents/10180/33199359/SLDB.pdf/d4061f0c-0123-4d88-896c-8e5677f12ba4?version=1.1>>.
- ŠKRABAL, J. (2017). Příprava sčítání lidu, domů a bytů v roce 2021 [online]. *Demografie*, 59(2). [cit. 1.7.2021]. <<https://www.czso.cz/documents/10180/46203818/SLDB.pdf/77b9c078-cd8f-439f-b5b9-1367cd31cecl?version=1.0>>.
- TEMPL, M., MEINDL, B., KOWARIK, A., CHEN, S. (2014). *Introduction to Statistical Disclosure Control (SDC)* [online]. International Household Survey Network. [cit. 1.7.2021] <<https://www.ihsn.org/sites/default/files/resources/ihsn-working-paper-007-Oct27.pdf>>.
- UNITED NATIONS. (2006). *Conference of european statisticians recommendations for the 2010 censuses of population and housing* [online]. United Nations Publication. [cit. 1.7.2021]. <[https://unece.org/fileadmin/DAM/stats/publications/CES\\_2010\\_Census\\_Recommendations\\_English.pdf](https://unece.org/fileadmin/DAM/stats/publications/CES_2010_Census_Recommendations_English.pdf)>.
- UNITED NATIONS. (2007). *Managing Statistical Confidentiality & Microdata Access: Principles and Guidelines of Good Practice* [online]. New York and Geneva: UNECE. [cit. 1.7.2021]. <[https://unece.org/fileadmin/DAM/stats/publications/Managing\\_statistical\\_confidentiality\\_and\\_microdata\\_access.pdf](https://unece.org/fileadmin/DAM/stats/publications/Managing_statistical_confidentiality_and_microdata_access.pdf)>.
- UNITED NATIONS. (2015). *Conference of european statisticians recommendations for the 2010 censuses of population and housing* [online]. United Nations Publication. [cit. 1.7.2021]. <[https://unece.org/DAM/stats/publications/2015/ECECES41\\_EN.pdf](https://unece.org/DAM/stats/publications/2015/ECECES41_EN.pdf)>.
- UNITED NATION. (2017). *Principles and Recommendations for Population and Housing Censuses* [online]. Economic & Social Affairs, New York: United Nations. [cit. 1.7.2021] <[https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles\\_and\\_Recommendations/Population-and-Housing-Censuses/Series\\_M67rev3-E.pdf](https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf)>.

## ANNEX 1: SURVEY QUESTIONS

Author have asked the following questions:

- (i) If you do not currently provide data to respondents, are you considering this? If so, in what form?
- (ii) If you are already publishing data, in what form is the presentation taking place? What models do you use to ensure the protection of respondents' data?
- (iii) If you are considering publishing microdata, what will be your next step, or what has hindered your implementation?
- (iv) If you have considered the possibility of presenting microdata but have withdrawn from it, please send this message as well.

## ANNEX 2: TABLES

**Table 30** Library of IPUMS

Countries	2010s	2000s	1990s	1980s	1970s	1960s	Pre-1960
Argentina	2010	2001	1991	1980	1970		
Armenia	2011	2001					
Austria	2011	2001	1991	1981	1971		
Bangladesh	2011	2001	1991				
Belarus		2009	1999				
Benin	2013	2002	1992		1979		
Bolivia	2012	2001	1992		1976		
Botswana	2011	2001	1991	1981			
Brazil	2010	2000	1991	1980	1970	1960	
Burkina Faso		2006	1996	1985			
Cambodia	2013	2008	1998				
		2004					
Cameroon		2005		1987	1976		
Canada	2011	2001	1991	1981	1971		1911
							1901
							1891
							1881
							1871
							1852
Chile		2002	1992	1982	1970	1960	
China		2000	1990	1982			
Colombia		2005	1993	1985	1973	1964	
Costa Rica	2011	2000		1984	1973	1963	
Cuba	2012	2002					
Denmark							1801
							1787

Table 30

(continuation)

Countries	2010s	2000s	1990s	1980s	1970s	1960s	Pre-1960
Dominican Republic	2010	2002		1981	1970	1960	
Ecuador	2010	2001	1990	1982	1974	1962	
Egypt		2006	1996	1986			
El Salvador		2007	1992				
Ethiopia		2007	1994	1984			
Fiji	2014	2007	1996	1986	1976	1966	
Finland	2010						
France	2011	2006	1999	1982	1975	1968	
			1990			1962	
Germany				1987	1971		1819
				1981	1970		
Ghana	2010	2000		1984			
Greece	2011	2001	1991	1981	1971		
Guatemala		2002	1994	1981	1973	1964	
Guinea	2014		1996	1983			
Haiti		2003		1982	1971		
Honduras		2001		1988	1974	1961	
Hungary	2011	2001	1990	1980	1970		
Iceland							1910
							1901
							1801
							1729
							1703
Indonesia	2010	2005	1995	1985	1976		
		2000	1990	1980	1971		
Iran	2011	2006					
Iraq			1997				
Ireland	2016	2006	1996	1986	1979		1911
	2011	2002	1991	1981	1971		1901
Israel		2008	1995	1983	1972		
Italy	2011	2001					
Jamaica		2001	1991	1982			
Jordan		2004					
Kenya		2009	1999	1989	1979	1969	
Kyrgyz Republic		2009	1999				
Laos		2005					
Lesotho		2006	1996				

Table 30

(continuation)

Countries	2010s	2000s	1990s	1980s	1970s	1960s	Pre-1960
Liberia		2008			1974		
Malawi		2008	1998	1987			
Malaysia		2000	1991	1980	1970		
Mali		2009	1998	1987			
Mauritius	2011	2000	1990				
Mexico	2015	2005	1995		1970	1960	
	2010	2000	1990				
Mongolia		2000		1989			
Morocco	2014	2004	1994	1982			
Mozambique		2007	1997				
Myanmar	2014						
Nepal	2011	2001					
Netherlands	2011	2001			1971	1960	
Nicaragua		2005	1995		1971		
Norway							1910
							1900
							1875
							1865
							1801
Pakistan			1998	1981	1973		
Palestine	2017	2007	1997				
Panama	2010	2000	1990	1980	1970	1960	
Papua New Guinea		2000	1990	1980			
Paraguay		2002	1992	1982	1972	1962	
Peru		2007	1993				
Philippines	2010	2000	1995				
			1990				
Poland	2011	2002		1988	1978		
Portugal	2011	2001	1991	1981			
Puerto Rico	2010	2005	1990	1980	1970		
		2000					
Romania	2011	2002	1992		1977		
Russia	2010	2002					
Rwanda	2012	2002	1991				
Saint Lucia			1991	1980			
Senegal	2013	2002		1988			
Sierra Leone		2004					

Table 30								(continuation)
Countries	2010s	2000s	1990s	1980s	1970s	1960s	Pre-1960	
Slovenia		2002						
South Africa	2016	2007	1996					
	2011	2001						
South Sudan		2008						
Spain	2011	2001	1991	1981				
Sudan		2008						
Suriname	2012	2004						
Sweden							1910	
							1900	
							1890	
							1880	
Switzerland		2000	1990	1980	1970			
Tanzania	2012	2002		1988				
Thailand		2000	1990	1980	1970			
Togo	2010				1970	1960		
Trinidad and Tobago	2011	2000	1990	1980	1970			
Turkey		2000	1990	1985				
Uganda	2014	2002	1991					
Ukraine		2001						
United Kingdom		2001	1991				1911	
							1901b	
							1901a	
							1891b	
							1891a	
							1881b	
							1881a	
							1871b	
							1861b	
							1861a	
							1851c	
							1851b	
							1851a	
United States	2015	2005	1990	1980	1970	1960	1910	
	2010	2000					1900	
							1880b	
							1880a	
							1870	



Table 30

(continuation)

Countries	2010s	2000s	1990s	1980s	1970s	1960s	Pre-1960
							1860
							1850b
							1850a
Uruguay	2011	2006	1996	1985	1975	1963	
Venezuela		2001	1990	1981	1971		
Vietnam		2009	1999	1989			
Zambia	2010	2000	1990				
Zimbabwe	2012						

Source: IPUMS Microdata library<sup>10</sup>

Table 31 Countries of Europe publishing SUF with links to microdata

Country	Web pages of the office	Web pages of the microdata
Belgium	< <a href="https://statbel.fgov.be">https://statbel.fgov.be</a> >	-
Denmark	< <a href="https://www.dst.dk">https://www.dst.dk</a> >	-
Faroe Islands	< <a href="https://hagstova.fo">https://hagstova.fo</a> >	< <a href="https://hagstova.fo/fo/atgongd-til-avnevndar-mikrodatur">https://hagstova.fo/fo/atgongd-til-avnevndar-mikrodatur</a> >
Germany	< <a href="https://www.destatis.de">https://www.destatis.de</a> >	-
Luxembourg	< <a href="https://statistiques.public.lu">https://statistiques.public.lu</a> >	-
Malta	< <a href="https://nso.gov.mt">https://nso.gov.mt</a> >	< <a href="https://nso.gov.mt/en/Services/Microdata/Pages/Access-to-Microdata.aspx">https://nso.gov.mt/en/Services/Microdata/Pages/Access-to-Microdata.aspx</a> >
Moldova	< <a href="https://statistica.gov.md">https://statistica.gov.md</a> >	< <a href="https://statistica.gov.md/pageview.php?l=en&amp;idc=636">https://statistica.gov.md/pageview.php?l=en&amp;idc=636</a> >
Montenegro	< <a href="http://www.monstat.org">http://www.monstat.org</a> >	-
Norway	< <a href="https://www.ssb.no">https://www.ssb.no</a> >	-
Serbia	< <a href="https://www.stat.gov.rs">https://www.stat.gov.rs</a> >	-
Slovakia	< <a href="https://slovak.statistics.sk">https://slovak.statistics.sk</a> >	< <a href="https://slovak.statistics.sk/wps/portal/ext/themes/demography/census/indicators">https://slovak.statistics.sk/wps/portal/ext/themes/demography/census/indicators</a> >
Sweden	< <a href="https://www.scb.se">https://www.scb.se</a> >	-

Source: Own construction

Table 32 Countries of Asia publishing SUF with links to microdata

Country	Web pages of the office	Web pages of the microdata
Bhutan	< <a href="https://www.nsb.gov.bt">https://www.nsb.gov.bt</a> >	< <a href="https://www.nsb.gov.bt/services/statistical-data-request">https://www.nsb.gov.bt/services/statistical-data-request</a> >
India	< <a href="https://www.censusindia.gov.in">https://www.censusindia.gov.in</a> >	< <a href="https://censusindia.gov.in/2011census/workstation.html">https://censusindia.gov.in/2011census/workstation.html</a> >
Japan	< <a href="http://www.stat.go.jp">http://www.stat.go.jp</a> >	< <a href="https://www.e-stat.go.jp/microdata">https://www.e-stat.go.jp/microdata</a> >
Maldives	< <a href="http://statisticsmaldives.gov.mv">http://statisticsmaldives.gov.mv</a> >	< <a href="http://statisticsmaldives.gov.mv/census-dataset">http://statisticsmaldives.gov.mv/census-dataset</a> >
Sri Lanka	< <a href="http://www.statistics.gov.lk">http://www.statistics.gov.lk</a> >	< <a href="http://www.statistics.gov.lk/Datadessimination">http://www.statistics.gov.lk/Datadessimination</a> >

Source: Own construction

<sup>10</sup> <<https://international.ipums.org/international-action/samples>>.

**Table 33** Countries of Europe publishing PUF and SUF with links to microdata

Country	Web pages of the office	Web pages of the microdata
Albania	< <a href="http://www.instat.gov.al">http://www.instat.gov.al</a> >	< <a href="http://www.instat.gov.al/en/figures/micro-data">http://www.instat.gov.al/en/figures/micro-data</a> >
France	< <a href="https://www.insee.fr">https://www.insee.fr</a> >	< <a href="https://www.insee.fr/fr/statistiques/4995124?sommaire=2414232">https://www.insee.fr/fr/statistiques/4995124?sommaire=2414232</a> >
Hungary	< <a href="http://www.ksh.hu">http://www.ksh.hu</a> >	< <a href="http://www.ksh.hu/nepszamlalas/tesztallomanyok">http://www.ksh.hu/nepszamlalas/tesztallomanyok</a> >
Ireland	< <a href="https://www.cso.ie">https://www.cso.ie</a> >	< <a href="https://www.cso.ie/en/census/censusreports1821-2006">https://www.cso.ie/en/census/censusreports1821-2006</a> >
Italy	< <a href="https://www.istat.it">https://www.istat.it</a> >	< <a href="https://www.istat.it/en/archivio/196131">https://www.istat.it/en/archivio/196131</a> >
Portugal	< <a href="https://www.ine.pt">https://www.ine.pt</a> >	< <a href="https://www.ine.pt/xportal/xmain?xpid=INE&amp;xpgid=ine_pufs&amp;xlang=en">https://www.ine.pt/xportal/xmain?xpid=INE&amp;xpgid=ine_pufs&amp;xlang=en</a> >
Spain	< <a href="https://www.ine.es">https://www.ine.es</a> >	< <a href="https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&amp;cid=1254736176992&amp;menu=resultados&amp;idp=1254735572981#tbs-1254736195714">https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&amp;cid=1254736176992&amp;menu=resultados&amp;idp=1254735572981#tbs-1254736195714</a> >
United Kingdom	< <a href="https://www.ons.gov.uk">https://www.ons.gov.uk</a> >	< <a href="https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/datasets/2011censusteachingfile">https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/datasets/2011censusteachingfile</a> >

Source: Own construction

**Table 34** Countries of Americas publishing PUF and SUF with links to microdata

Country	Web pages of the office	Web pages of the microdata
Brazil	< <a href="https://www.ibge.gov.br">https://www.ibge.gov.br</a> >	< <a href="https://www.ibge.gov.br/en/statistics/social/population/18521-2000-population-census.html?edicao=18531&amp;t=microdados">https://www.ibge.gov.br/en/statistics/social/population/18521-2000-population-census.html?edicao=18531&amp;t=microdados</a> >
Canada	< <a href="https://www.statcan.gc.ca">https://www.statcan.gc.ca</a> >	Hierarchical file < <a href="https://www150.statcan.gc.ca/n1/en/catalogue/98M0002X">https://www150.statcan.gc.ca/n1/en/catalogue/98M0002X</a> > Individuals file < <a href="https://www150.statcan.gc.ca/n1/en/catalogue/98M0001X">https://www150.statcan.gc.ca/n1/en/catalogue/98M0001X</a> >
Colombia	< <a href="https://www.dane.gov.co">https://www.dane.gov.co</a> >	< <a href="http://microdatos.dane.gov.co/index.php/catalog/643/get_microdata">http://microdatos.dane.gov.co/index.php/catalog/643/get_microdata</a> >
Costa Rica	< <a href="https://www.inec.cr">https://www.inec.cr</a> >	< <a href="http://sistemas.inec.cr/pad5/index.php/catalog/113/get-microdata">http://sistemas.inec.cr/pad5/index.php/catalog/113/get-microdata</a> >
Ecuador	< <a href="https://www.ecuadorencifras.gob.ec">https://www.ecuadorencifras.gob.ec</a> >	< <a href="https://anda.inec.gob.ec/anda/index.php/catalog/659/get_microdata">https://anda.inec.gob.ec/anda/index.php/catalog/659/get_microdata</a> >
Guatemala	< <a href="https://www.ine.gob.gt">https://www.ine.gob.gt</a> >	< <a href="https://www.censopoblacion.gt/descarga">https://www.censopoblacion.gt/descarga</a> >
Honduras	< <a href="http://www.ine.gob.hn">www.ine.gob.hn</a> >	< <a href="http://170.238.108.229/index.php/catalog/69/get_microdata">http://170.238.108.229/index.php/catalog/69/get_microdata</a> >
Chile	< <a href="https://www.ine.cl">https://www.ine.cl</a> >	< <a href="http://www.censo2017.cl/microdatos">http://www.censo2017.cl/microdatos</a> >
Mexico	< <a href="https://www.inegi.org.mx">https://www.inegi.org.mx</a> >	< <a href="https://www.inegi.org.mx/programas/ccpv/2020/?ps=microdatos">https://www.inegi.org.mx/programas/ccpv/2020/?ps=microdatos</a> >
United States of America	< <a href="https://www.census.gov">https://www.census.gov</a> >	Stateside areas: < <a href="https://www2.census.gov/census_2010/12-Stateside_PUMS">https://www2.census.gov/census_2010/12-Stateside_PUMS</a> > Island areas: < <a href="https://www2.census.gov/census_2010/11-Island_Areas_PUMS">https://www2.census.gov/census_2010/11-Island_Areas_PUMS</a> >
Uruguay	< <a href="https://www.ine.gub.uy">https://www.ine.gub.uy</a> >	< <a href="https://www.ine.gub.uy/web/guest/censos1">https://www.ine.gub.uy/web/guest/censos1</a> >

Source: Own construction

**Table 35** Countries of Asia publishing PUF and SUF with links to microdata

Country	Web pages of the office	Web pages of the microdata
Armenia	< <a href="https://www.armstat.am">https://www.armstat.am</a> >	< <a href="https://www.armstat.am/en/?nid=210">https://www.armstat.am/en/?nid=210</a> >
South Korea	< <a href="http://kostat.go.kr">http://kostat.go.kr</a> >	< <a href="https://mdis.kostat.go.kr/extract/extYearsSurvSearchNew.do?curMenuNo=UI_POR_P9012">https://mdis.kostat.go.kr/extract/extYearsSurvSearchNew.do?curMenuNo=UI_POR_P9012</a> >

Source: Own construction

**Table 36** Countries of Africa publishing PUF and SUF with links to microdata

Country	Web pages of the office	Web pages of the microdata
Angola	< <a href="https://www.ine.gov.ao">https://www.ine.gov.ao</a> >	< <a href="https://andine.ine.gov.ao/nada/index.php/catalog/3">https://andine.ine.gov.ao/nada/index.php/catalog/3</a> >
Burundi	< <a href="http://www.isteebu.bi">http://www.isteebu.bi</a> >	< <a href="http://www.isteebu.bi/nada/index.php/catalog/3">http://www.isteebu.bi/nada/index.php/catalog/3</a> >
Ghana	< <a href="https://statsghana.gov.gh">https://statsghana.gov.gh</a> >	< <a href="https://www2.statsghana.gov.gh/nada/index.php/catalog/51/get_microdata">https://www2.statsghana.gov.gh/nada/index.php/catalog/51/get_microdata</a> >
Morocco	< <a href="https://www.hcp.ma">https://www.hcp.ma</a> >	< <a href="https://www.hcp.ma/downloads/RGPH-2014-Microdonnees-anonymisees-Open-Data_t21400.html">https://www.hcp.ma/downloads/RGPH-2014-Microdonnees-anonymisees-Open-Data_t21400.html</a> >
Namibia	< <a href="https://nsa.org.na">https://nsa.org.na</a> >	< <a href="https://nsa.org.na/microdata1/index.php/catalog/19">https://nsa.org.na/microdata1/index.php/catalog/19</a> >
Rwanda	< <a href="https://www.ine.rw">https://www.ine.rw</a> >	< <a href="https://microdata.statistics.gov.rw/index.php/catalog/65/related_materials">https://microdata.statistics.gov.rw/index.php/catalog/65/related_materials</a> >
South Africa	< <a href="http://www.statssa.gov.za">http://www.statssa.gov.za</a> >	< <a href="https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/485/get_microdata">https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/485/get_microdata</a> >
Somalia	< <a href="https://www.nbs.gov.so">https://www.nbs.gov.so</a> >	< <a href="http://microdata.nbs.gov.so/index.php/catalog/4/get_microdata">http://microdata.nbs.gov.so/index.php/catalog/4/get_microdata</a> >

Source: Own construction

**Table 37** Countries of Oceania publishing PUF and SUF with links to microdata

Country	Web pages of the office	Web pages of the microdata
New Zealand	< <a href="https://www.stats.govt.nz">https://www.stats.govt.nz</a> >	< <a href="https://www.stats.govt.nz/integrated-data/apply-to-use-microdata-for-research">https://www.stats.govt.nz/integrated-data/apply-to-use-microdata-for-research</a> > < <a href="https://www.stats.govt.nz/integrated-data/apply-to-use-microdata-for-research/confidentialised-unit-record-files-curfs">https://www.stats.govt.nz/integrated-data/apply-to-use-microdata-for-research/confidentialised-unit-record-files-curfs</a> >

Source: Own construction