

Census, Revision, Projections

Experiences from the Republic of North Macedonia

Jane Krsteski

Advisor for Censuses, population estimation and projections, SSO of Republic of North Macedonia

jane.krsteski@stat.gov.mk

Abstract

The purpose of this paper is to share the knowledge and experience acquired by the SSO in terms of strengthening the capacities for utilizing its opportunities and the role of the institutions in providing an innovative approach to the use of statistical data for the population, which is:

1. The use of administrative data as a source for assessing the quality of census data,
2. Using data sources in the SSO's own, and
3. Methodological solutions and approaches for:
 - a) Activities related to data processing for the 2021 Census,
 - b) Revision of the series of population estimates up to the last Census 2002,
 - c) The population projections made in the period 2022-2070.

Based on the above, the focus of this paper will be directed to the following four aspects:

1. The processing of data for the Census. The SSO conducted Census 2021 based on the combined method. For this purpose, a Pre-Census Database was developed, with data from six institutions. In the processing of field data from the Census, this database was used as a source for quality assessment. Through this database, the SSO identified that of the 2 192 778 persons enumerated, 258,932 persons live abroad and are not part of the total population, while 132 725 persons were not enumerated.
2. Revision of the population estimates series. The SSO made a revision of the series of population estimates using PINs for all enumerated persons in the 2021 Census up to the 2002 Census using the data on births, deaths and migrations and thus managed to identify unregistered emigrants according to their PIN.
3. Development of population projections. Based on the data for the revision of the population series, SSO prepared population projections for the period 2022 - 2070. For this purpose, seven variants were developed, which are based on four hypotheses about fertility, two hypotheses about mortality and three hypotheses about migrations. A brief overview of the obtained results is given in the attachment.
4. Conclusion. As a conclusion, this paper will give a brief comment on all the challenges faced in each of the phases previously mentioned.

Keywords: Republic of North Macedonia, Census 2021, Population estimation revision, Population projection 2022 – 2070.

1. Introduction

The Republic of North Macedonia conducted the 2021 Census using a combined method through the use of CAPI and a specially designed application that captured and returned data in real time to an established IT Census Data System. (The last official Census before the 2021 Census was conducted in 2002.

The success of the 2021 Census is due to the use of the combined census method and established an IT Census Data System which managed to methodologically circumvent the challenges that proved to be insurmountable in the unsuccessful Census 2011.

SSO developed a PCD (pre-census database) as part of the established IT Census Data System, which used the PIN (Personal Identification Number) as a way to recognize persons in six state institutions, which together with the data available in the SSO as the seventh (institution) provided quality data for the use of the Usual Resident Status (URS) indicator.

In addition, the SSO provided an opportunity for self-registration of its citizens living abroad through a web application in which access was enabled only through an IP address from a foreign country and registration of persons through a PIN.

2. Overview of activities

2.1. Preparations for developing the PCD (pre-census database)

Preparations for the 2021 Census began in 2016. when a Working Group was formed in the SSO that visited all State institutions that have certain sub-categories for the population in order to make a list of available data that can be used as potential official sources in the preparation of the PCD (pre-census database).

Based on these activities SSO has divided the holders of relevant data into two basic groups:

1. Primary Institutions that are directly responsible for collecting this data and maintaining it; and
2. Secondary Institutions that take over the data from the primary Institutions.

In the next evaluation step, the primary institutions were subjected to additional data analysis based on the following additional criteria:

1. Basic analysis of an administrative source
2. Data accuracy (de facto versus de jure registered status);
3. Timeliness
4. Availability (data availability frequency);
5. Completeness of data

6. Comparability (comparability of the same features in different registers);
7. Legal basis for obtaining data.

Additional information about this phase can be found at the References.

2.2. Test Census 2019

The Test Census was successfully conducted in the spring of 2019. It was conducted in all eight statistical regions with one census area in an urban settlement and one census area in a rural settlement. During the selection of the areas, care was taken that they fully have the characteristics of the region they represented in terms of the demographic, economic, educational and ethno-cultural characteristics of the region.

The objectives of the Test Census were:

- Testing the functionality of the application in the two basic parts: the GIS part and the Methodological part.
- Testing of the IT Census Data System in terms of two-way data transfer and Internet availability in different terrain and its geographical characteristics.
- Testing the quality of the data in PCD and changing certain data that proved to be outdated or incorrect in the field.
- Testing the methodological approach in the development of the application in terms of the quality and volume of acquired knowledge of the enumerators during the training and minimizing the subjective methodological decisions of the enumerators in the enumeration process.
- Testing of personal data protection and built-in solutions that prevented the possibility of possible data manipulation after the enumerators left the household, by any of the participants in the Census and in any part of the process.
- Testing the acceptability of this way of enumeration by citizens and the general public.

Based on the results of the Test Census in the IT Census Data System including the census application, additional solutions were implemented that made this system fully functional.

The application incorporated a completely autonomous enumeration for each person individually based solely on the answers given, in which the enumerator's role as a technical person who should follow the methodology was reduced to a technical person who only had to ask the questions that appeared to him on the screen, give additional explanation about the question if necessary and write down the answer.

With this solution, the application recognized 44 different groups of questions (like 44 different questionnaires), and most importantly, it recognized automatically which of the enumerated persons was part of the total population, which was not the case in the 2002 Census when that decision was made by enumerators in the field, which is why tensions and mistrust were created and which was the main obstacle due to which the Census 2011 was stopped.

This methodological approach in the application for Census in the country was incorporated in the development of the application for self-enumeration from abroad.

2.3 Census 2021

The 2021 census was conducted in the period from September 5 to 30. In addition to CAPI, CAWI was also enabled for citizens residing abroad, through a special Web application. These persons were able to self-enumerate between March 1 and September 30, 2021.

As a result of the previous activities, in the Pre-census database (PCD), 2,608,921 active records (person) were recognized through PIN.

The use of PIN to connect persons in the specified data sources enabled the use of 33 variables in the enumeration and processing process.

Of these, 13 were used to facilitate the census in the field or required their confirmation by the persons enumerated.

The greatest help during the field activity for the enumerated persons and for the enumerators were the following 8 variables (in brackets is the number of persons for whom the data was available in the Pre-census database):

- 1) Educational attainment (776 000)
- 2) Field of educational attainment (273 000)
- 3) School attendance – level of the educational program (291 000)
- 4) Field of the educational program (45 000)
- 5) Occupation (460 000)
- 6) Activity (642 000)
- 7) Source of livelihood – pension (320 000)
- 8) Source of livelihood – social transfers (307 000)

Thanks to the significant amount of data in PCD for each PIN, and the algorithmic results for Usual Resident Status (URS) as a combination of data from different institutions, the following three categories of population were identified:

1. Enumerated (CAPI) and self-enumerated (CAWI) persons who are part of the total population;
2. Enumerated (CAPI) and self-enumerated (CAWI) persons who are not part of the total population;
3. Unenumerated persons who, according to the algorithm for Usual Resident Status (URS), should be included in the Usual Resident Population.

See Appendix, Table 1. Total enumerated population, total resident population, and total non-resident population. Table 2. Enumerated population and Population stocks

Additional information about this phase can be found at the References.

2.4 Revision of the series of estimates of the population for the period January 2021 - January 2003

The methodological approach to revise the total population between the census of 2021 and 2002 took place in two stages. The period January 2021 - January 2003 was taken as the dates on which the audit was carried out.

2.4.1 First stage

In the first stage, the SSO conducted an activity based on PIN (sex, age and official address) to recognize all persons enumerated in the 2002 Census who are part of the Total Population enumerated in the 2021 Census, including persons who were born, died or moved away during this period by using data from their own sources, in order to:

1. To identify all persons who have been enumerated in the two successive censuses according to their usual and official address of residence;
2. To identify those persons who died during that period according to the official place of residence;
3. To identify persons who were born during that period according to the mother's official place of residence;
4. To identify internal migrants according to the official area of residence from which they moved and to which they settled;
5. To identify the immigrated foreigners according to the official area of registration; and
6. To identify all invalid entries.

In the entire data processing process from this stage, SSO was aware that two types of problems would arise:

1. Differences between the official place of residence and the place of habitual residence of persons. This difference was due to the different definitions of residence used in the Censuses and those used in other sources.
2. The unavailability of proof of the date when the person moved to another settlement, or when he left the country to stay abroad for more than 12 months.

Based on these calculations, the number of the total population on January 1 2021 was 1 843 210 people. For comparison, the total population according to the 2021 Census is 1 836 713 people, (with a critical moment of September 5, 2021). That is, the difference was about 6 500 persons, or in percentage about 0.3%.

2.4.2. Second stage

An expert hired with the support of UNPFA was involved in the second stage of the data audit.

Based on the data prepared by the SSO in the first stage, he made a mathematical calculation based on an adjusted interpolation model of migrants by age and gender to align the peaks of the trends and adjust the population movement according to the expected trends through mirroring statistics on displaced persons from the country.

See Appendix Table 3. Total estimated resident population of the Republic of North Macedonia on 1 January, revised data for the period 2003-2021, based on the Census 2021

Additional information about results from this phase can be found at the References.

2.5. Population projections to 2070

2.5.1 Setting the hypotheses

Calculations for population projections until 2070 were performed using the cohort-component method, by using the computer program DemProj, which is one of the Spectrum System of Policy Models, Version 6.29 modules developed by Avenir Health.

(<https://www.avenirhealth.org/software-spectrum.php>).

The calculations represent results obtained on the basis of the set hypotheses without any influence on the factors of demographic change in the number and geographical distribution of the population as a result of the socio-economic circumstances and the security situation, not only in the country but also in the region of the Western Balkans, in the entire period for which they were calculated.

Additional information about hypothesis can be found at the References

In the process of setting up the hypotheses, the historical trends in the country were taken into account and an analysis of the possible movements of the demographic factors was carried out based on the internal policies and regional experiences that influenced the demographic changes in the countries of the region of South-east Europe and Western Balkan in the period from 2000 until now.

Hypotheses are consisted of four variants for fertility: low, medium, high and constant fertility; two for mortality: declining and constant mortality, and three for migration: variable (moderate), with zero-migration balance and with constant migrations.

In the calculation of population projections, it has been selected 7 projection variants. Those are as follows, according to their names: 1) low fertility variant, 2) medium fertility variant, 3) high fertility variant, 4) constant fertility variant, 5) constant mortality variant, 6) constant variant, and finally, 7) zero-migration variant.

Although the SSO did not aim at predicting the number of the population in 2070, it still considers variants 1 to 5 as realistically feasible, unlike variants 6 and 7, which are included as realistically unfeasible hypotheses and serve only for analytical purposes and those will not be comment. See Appendix Table 4. Combination of hypothesis.

2.5.2 Fertility hypotheses

For all four variants of the fertility hypotheses, it is assumed that in the starting year of the projection period, the fertility of the female population by age would be identical, calculated on the basis of data from the demographic statistics on the volume of births by age of the mother and the estimated number of the fertile female population for 2022. For the constant fertility variant, the same assumption has been adopted for the entire projection period (until the middle of 2070). The remaining three variants include an assumption about the variability of the fertility level, both on the total, as well as on its distribution by age. See Appendix Table 5. Total fertility rate for the hypothesis for fertility

2.5.3 Mortality hypotheses

In the projections of the population of the Republic of North Macedonia, two variants of hypotheses for the future movement of mortality, by age and by sex, have been adopted: a hypothesis for constant mortality and a hypothesis for variable, that is, decreasing mortality. The variant with constant mortality has an analytical character, in the second variant, the movements of mortality by age and gender, i.e. the life expectancy in North Macedonia in the last two decades, was analyzed, but with the abstraction of changes in mortality caused by Covid-19. See Appendix Table 6. Life expectancy at birth for second hypothesis for mortality

2.5.4 Hypothesis for migrations

Several variants of hypotheses have also been set in projecting future migration flows: variable migrations, zero-migration balance and constant migrations. The variable migration hypothesis variant is based on the adopted assumptions about changes in the external migration balance and its distribution by age, for both sexes individually. For the zero-migration balance variant, the absence of external migrations is assumed, while for the constant migrations variant, an identical migration balance as estimated for 2022, is assumed for each year of the projection period. Both the zero-migration balance hypothesis and the constant migrations hypothesis are theoretical hypotheses, whereby the former one aims to show the consequences only of the action of the natural factor (fertility and mortality), while the latter aims to show the consequences for the population when migrations have a constant character. See Appendix Table 7. Net migration

3.5.5 Results

1. None of the variants show stabilization or increase in population size.
2. The current conditions of the age structure, the low level of fertility, the mortality rates and the factors that determine migration movements are solely responsible for the calculated values.
3. The results of the analyzed data from the population projections indicate that in the near future it is difficult to expect an improvement in the demographic situation if appropriate solutions are not implemented.

In 2022 as the start year, population of North the Republic of North Macedonia was 1 831 712 persons. The most favorable population number in 2070 is obtained with the High Fertility scenario according to which the population decreases by 23.1% compared to the starting year, while the most unfavorable result for the final year is obtained with the Constant Mortality scenario where the population decreases by 43.8% compared to 2022.

At population structure by age groups the most favorable situation of the number of the population aged from 0 to 19 years is in the scenario with high fertility, where it decreased by 94,044 people, i.e. by 22.8% compared to the situation in the starting year, while the most unfavorable situation is in the scenario with low fertility, where this population group decreased by 250,591 people, i.e. by 60.8% compared to 2022.

For the population aged 20 to 64, the most favorable situation is also in the scenario with high fertility, where it decreases by 34%, i.e. by 374,195 people compared to 2022, while the most unfavorable situation is the same as before in the scenario with low fertility, where this population group decreased by 506,282 persons, that is, by 46% compared to the starting year.

In the population older than 64 years, there is an increase in the number in the scenarios with low, medium, high and constant fertility, by 45,433 people, that is, by 14.2%. This is due to the fact that this population is not affected by fertility because these individuals were born in 2022. Only, a decrease is observed in the scenario with constant mortalities, where a decrease of this group of people is observed by 101,262 people, that is, by 31.7% compared to the starting year.

Regarding the ratio of the total dependent population, the most favorable situation is in the scenario with constant mortality, where for every 100 persons aged 20 to 64, there is 72.1 persons aged 0 to 19 and persons older than 64. While the most unfavorable situation is in the scenarios with medium and high fertility where it amounts between 94.3 and 94.1 persons aged 0 to 19 years and persons older than 64 years in relation to 100 persons aged 20 to 64 years.

Regarding the ratio of young dependent population, the most favorable situation is in the scenario with high fertility, where for 100 people aged 20 to 64, we have 43.9 people aged 0 to 19, while most unfavorable situation is in the scenario with low fertility, where we have 27.3 persons aged 0 to 19 per 100 persons aged 20 to 64.

Table and graphs with results are available in Appendix Table 8. Population projections of the Republic of North Macedonia. And population Pyramids.

4. Over all conclusions

1. All activities about Census 2021 were successful.
2. During this activities SSO gain huge experience and knowledge.
3. SSO is confident that can share experience and knowledge with other SSO's as a direct support or consultation about all aspects of this process.
4. SSO is confident that can conduct Census 2031 only from registers and all focus is in that direction.
5. More detail information can be obtained from the References

References

Krsteski, J. (2020). Development of a pre-census database and computer-assisted methodology for conducting the 2021 census in the Republic of North Macedonia - HUNGARIAN STATISTICAL REVIEW, VOLUME 3, NUMBER 1, PP. 46–59. DOI: 10.35618/hsr2020.01.en046. https://real.mtak.hu/112019/1/2020_01_046.pdf

Simovski, A., & Krstevska, D. (2022). Combined Census of population, households and dwellings in the Republic of North Macedonia, 2021 - Group of Experts on Population and Housing Censuses, Twenty-fourth Meeting, Lessons learned from censuses of the 2020 round, Geneva, Switzerland, 21-23 September 2022. https://unece.org/sites/default/files/2022-09/WP11_NorthMacedonia_SimovskiKrstevska_ENG_2.pdf

Census data

https://makstat.stat.gov.mk/PXWeb/pxweb/en/MakStat/MakStat_Popisi_Popis2021_NaselenieSet/T1008P21a.px/table/tableViewLayout2/?rxid=46ee0f64-2992-4b45-a2d9-cb4e5f7ec5ef

Population revision data

https://makstat.stat.gov.mk/PXWeb/pxweb/en/MakStat/MakStat_Naselenie_ProcenkiNaselenie_ProcenkiPopis2021_Proceni2003_2020Revidirani/125_ProcNasMK01.ml.px/?rxid=46ee0f64-2992-4b45-a2d9-cb4e5f7ec5ef

Population projection 2022 – 2070

https://makstat.stat.gov.mk/PXWeb/pxweb/en/MakStat/MakStat_Naselenie_Proekcii/125_Popis_Procenki2070.ml.px/table/tableViewLayout2/?rxid=46ee0f64-2992-4b45-a2d9-cb4e5f7ec5ef

APPENDIX

	Enumerated population			Resident population			Non-resident population		
	TOTAL	Men	Women	TOTAL	Men	Women	TOTAL	Men	Women
Total enumerated population	2 097 319	1 054 178	1 043 141	1 836 713	911 087	925 626	260 606	143 091	117 515

Enumerated population	Number	Share (%)	Population Stock		
				Number	Share (%)
Total	2 192 778.00	100	Total	2 097 319.00	100
Field data	1 855 249.00	84.61	Usual resident population (URP)	1 836 713.00	87.57
Web data	204 804.00	9.34	Temporary present population	1 674.00	0.08
Imputed records	132 725.00	6.05	Population living abroad	258 932.00	12.35

Year	Total	Males	Females	Year	Total	Males	Females
2003	2 023 496	1 015 790	1 007 706	2013	1 925 283	955 911	969 372
2004	2 021 170	1 014 679	1 006 491	2014	1 919 380	953 477	965 903
2005	2 010 409	1 007 450	1 002 959	2015	1 914 887	951 671	963 216
2006	1 999 450	1 000 110	999 340	2016	1 909 155	949 521	959 634
2007	1 988 318	992 794	995 524	2017	1 902 668	946 719	955 949
2008	1 976 743	985 497	991 246	2018	1 893 802	942 586	951 216
2009	1 965 453	978 110	987 343	2019	1 883 505	937 771	945 734
2010	1 951 316	968 805	982 511	2020	1 868 205	929 327	938 878
2011	1 940 488	962 642	977 846	2021	1 843 210	914 653	928 557
2012	1 933 522	959 677	973 845				

Projection variant	Hypotesis		
	Fertility	Mortality	Migration
1 – Low fertility	Low	Decreasing	Variable
2 – Medium fertility	Medium	Decreasing	Variable
3 – High fertility	High	Decreasing	Variable
4 – Constant fertility	Constant	Decreasing	Variable
5 – Constant mortality	Medium	Constant	Variable
6 - Constant variant	Constant	Constant	Constant
7 – Zero-migration	Medium	Decreasing	Zero-outmigration

Year	2022	2023	2024	2025	2030	2040	2050	2060	2070
Low fertility	1.57	1.55	1.53	1.51	1.4	1.31	1.31	1.31	1.31
Medium fertility	1.57	1.55	1.53	1.51	1.44	1.42	1.52	1.71	1.81
High fertility	1.57	1.61	1.65	1.69	1.88	2.03	2.03	2.03	2.03

Sex	2022	2025	2030	2040	2050	2060	2070
Male	73.9	75.3	76.5	79	81	82.7	84.3
Female	78.4	80	81.2	84.1	86	87.2	88.3

Year	2022	2025	2030	2040	2050	2060	2070
Total	-14347	-16154	-20439	-3714	-889	2543	6376
Male	-7773	-8752	-11074	-2012	-482	1052	2585
Female	-6574	-7402	-9365	-1702	-407	1491	3791

Year	Variant 1: Low fertility	Variant 2: Medium fertility	Variant 3: High fertility	Variant 4: Constant fertility	Variant 5: Constant mortality	Variant 6: Constant values	Variant 7: Zero-migration balance
2030	1 642 299	1 643 450	1 663 333	1 649 656	1 621 395	1 654 901	1 792 479
2040	1 427 885	1 436 152	1 502 590	1 456 117	1 365 794	1 382 585	1 724 052
2050	1 302 646	1 323 009	1 435 168	1 350 650	1 203 922	1 091 777	1 655 462
2060	1 193 458	1 235 458	1 394 596	1 262 899	1 085 711	791 878	1 573 924
2070	1 120 279	1 192 963	1 408 911	1 217 438	1 029 243	501 506	1 485 738

Population pyramids – Projections 2070



