Thematic Report
Evaluation of Age and Sex Data
Timor-Leste
Population and Housing Census 2022

Thematic Report
Evaluation of Age and Sex Data

May, 2024
Title:
Timor-Leste Population and Housing Census 2022, Thematic Report – Evaluation of Age and Sex Data

Design, Typesetting
Timor-Leste National Institute of Statistics (INETL)

Copyright © Timor-Leste National Institute of Statistics 2024

All right reserved

INETL/06

The Timor-Leste National Institute of Statistics (INETL)

Dili, May 2024
Preface

The population and Housing Census 2022, conducted by the Timor-Leste National Institute of Statistics (INETL) from 5 September to 5 October 2022, underscored our continuing commitment to decision-making grounded in robust data analysis. Utilizing modern technology in the form of tablets for data collection, the census laid the foundation for a comprehensive understanding of our nation’s demographic landscape. Initial census basic tablets were released in the latter part of 2023, enriching our understanding of the demographic intricacies captured during the 2022 Population and Housing Census.

This publication, is one of several detailed thematic census reports, addressing crucial areas such as fertility and nuptiality, mortality, migration, population projection, household and housing conditions, labour force and economic activity, disability, youth, gender, and a comprehensive census atlas, that will provide nuanced insights into specific demographic dimensions.

This volume provides a detailed analysis of mortality in Timor-Leste. Encompassing critical areas such as infant mortality, under five mortality, life expectancy, age specific mortality rates and maternal mortality, we expect this report to serve as a cornerstone for informed policymaking.

At this juncture, I would like to acknowledge the collective dedication of the INETL staff and the multi-stakeholder Census Technical Committee, led by Mr. Elias dos Santos Ferreira, President of INETL, I.P., in the successful completion of this Census. Their tireless efforts steered the census process from its initiation to the release of this report. Gratitude is also extended to our development partners from the UN System, among them, the UNFPA, UNICEF, UNDP, UN Women, and WFP, whose invaluable contributions have significantly enriched the census endeavour. Similarly, the support that we received from the Australian Bureau of Statistics (ABS) throughout the census process is also duly recognized.

On behalf of the Ministry of Finance, I would like to invite to all individuals with a keen interest in demographics and statistical data to leverage the insights presented in this publication. It is our hope that this compilation shall serve as reliable resource for informed discourse and evidence-based decision-making in the realm of our country’s national development.

Thank you.

Santina JRF Viegas Cardoso
Minister of Finance, RDTL
Acknowledgements

The Timor-Leste National Institute of Statistics (INETL), formerly the General Directorate of Statistics, implemented the Population and Housing Census 2022 under the slogan ‘Our census, our future be part of it’.

The census was largely financed by the Government of Timor-Leste through the Ministry of Finance. Additional financial and material support in form of tablets and power banks, was provided by UNFPA, UN Women, UNDP, UNICEF and WFP. I would like to express my sincere gratitude for the unwavering support towards the census.

I would like to acknowledge the invaluable technical support provided by UNFPA throughout the census process. The support provided during field staff training by UN Women is also acknowledged.

Let me also thank a team of independent international monitors from the Australian Bureau of Statistics, who worked with us throughout the enumeration period, their feedback from the field was valuable and timely.

I would like to further express my appreciation to all members of the Census Technical Committee for their advice and guidance in the census. The Census Publication Commission led a successful implementation of the census publicity campaign to ensure that the stakeholders, including the general public, were informed about the census.

Special thanks go to each and every enumerator and supervisor who worked tirelessly, visiting households across the country to collect the census information. They worked for long hours and also at odd hours in trying to collect the information.

Finally, I would like to commend the work of all staff from INETL headquarters and municipality offices, who continue to work for the success of the 2022 census project. I am grateful to the people of Timor-Leste for their cooperation, without which a successful census would not have been possible.

Eliar dos Santos Ferreira, L.Ec, MM
President, Timor-Leste National Institute of Statistics
Contents

Acknowledgements ............................................................................................................................ ii
List of tables ...................................................................................................................................... v
List of figures ..................................................................................................................................... v
1 Timor-Leste Population and Housing Census ........................................................................... 1
  1.1 Background and Context ........................................................................................................ 1
  1.2 Objectives of the Census ........................................................................................................ 1
  1.3 Census methodology .............................................................................................................. 2
    1.3.1 Census planning ................................................................................................................ 2
    1.3.2 Census mapping ................................................................................................................ 2
    1.3.3 Census Publicity ............................................................................................................... 2
    1.3.4 Field staff training ............................................................................................................ 3
    1.3.5 Census enumeration ....................................................................................................... 3
    1.3.6 Post-Enumeration Survey ............................................................................................. 3
    1.3.7 Data processing and analysis ......................................................................................... 4
    1.3.8 Census dissemination .................................................................................................... 5
2 Methodology and data .................................................................................................................. 6
  2.1 Rationale ................................................................................................................................... 6
  2.2 Sources of errors .................................................................................................................... 7
  2.3 Concepts of age and sex data evaluation ............................................................................... 7
  2.4 Data and methodology .......................................................................................................... 8
    2.4.1 Data .................................................................................................................................... 8
    2.4.2 Method of analysis ......................................................................................................... 9
3 Analysis of age and sex data ...................................................................................................... 10
  3.1 Graphical analysis ................................................................................................................ 10
    3.1.1 Population distribution ................................................................................................. 10
    3.1.2 Population Pyramid ...................................................................................................... 11
    3.1.3 Graphical cohort analysis ............................................................................................ 13
  3.2 Age and Sex ratios ................................................................................................................ 16
    3.2.1 Sex ratio ...................................................................................................................... 16
    3.2.2 Age ratio .................................................................................................................... 17
  3.3 Whipples Index ..................................................................................................................... 18
  3.4 Myers Blended Index .......................................................................................................... 20
  3.5 United Nations (UN) Age-Sex Accuracy Index .................................................................. 22
4 Conclusion and recommendations ........................................................................................... 24
References ......................................................................................................................................... 26
Annex ............................................................................................................................................... 27
Report contributors .......................................................................................................................... 29
List of tables

Table 3.1: Whipple’s Index by sex at national and municipality, 2022................................................. 20
Table 3.2: Myers Blended Index by municipality and sex, 2022.............................................................. 22

List of figures

Figure 3.1: Single year age population distribution by sex, 2022............................................................. 10
Figure 3.2: Population distribution by five year age group sex, Timor-Leste, 2022 ................................. 11
Figure 3.3: Single age population pyramid, Timor-Leste, 2022................................................................. 12
Figure 3.4: Single age population pyramid, Timor-Leste, 2015................................................................. 13
Figure 3.5: Graphical cohort analysis, Timor-Leste, 2015 and 2022......................................................... 14
Figure 3.6: Graphical cohort analysis of males, Timor-Leste, 2015 and 2022........................................... 15
Figure 3.7: Graphical cohort analysis of females, Timor-Leste, 2015 and 2022 ..................................... 15
Figure 3.8: Sex ratio, Timor-Leste, 2022..................................................................................................... 16
Figure 3.9: Sex ratio by Municipality, Timor-Leste, 2022................................................................. 17
Figure 3.10: Age ratio by sex and five year age-group, Timor-Leste, 2022 ............................................. 18
Figure 3.11: Myers preference by terminal digit index and sex, Timor-Leste, 2022 ............................ 21
Figure 3.12: Myers Blended Index at national and municipality, 2022 ................................................... 21
Figure 3.13: UN Joint Age Sex Accuracy Index by municipality, 2022 .................................................. 23
1 Timor-Leste Population and Housing Census

1.1 Background and Context

The evaluation of age and sex data in this report uses data from the 2022 Population and Housing census that was conducted by the then General Directorate of Statistics (GDS), later known as National Statistical Institute of Timor-Leste (INETL). The census was conducted from September to October 2022. Timor-Leste regained its independence in 2002 from Indonesian rule after 27 years. Previously the country was a Portuguese colony until 1975 when it got its independence. The 2022 census was the fourth to be conducted in Timor-Leste since the restoration of independence. Previous censuses were conducted in 2004, 2010 and 2015, that is approximately every five years.

1.2 Objectives of the Census

The objectives of the Timor-Leste Population and Housing Census 2022 were the following:

- Produce reliable statistical information on the size, composition, characteristics and spatial distribution of the resident population of Timor-Leste, as well as on the size, characteristics and spatial distribution of the country’s housing stock, and more specifically on
  - levels of fertility, mortality and migration
  - the educational attainment level of the population
  - the size and characteristics of the labour force
  - the participation of the population in agricultural production
  - the prevalence of disability by type
  - the rate and patterns of urbanization
  - housing conditions and availability of amenities and assets.
- Produce a sampling frame for household- and housing surveys in the ten years following the 2022 census.
- Produce the baseline for population projections.
- Providing the base for Timor-Leste’s Geographical Information System.
1.3 Census methodology

1.3.1 Census planning

Preparations for the census started in 2018 with development of a census project document which contained the complete census plan and the estimated cost of the whole census project. The document was crucial for resource mobilization from both the Government and development partners. Work on questionnaire development also started early with stakeholder consultation on topics and questions to be included in the census. This work built on topics that are recommended internationally to be included in the census.

1.3.2 Census mapping

Update of census maps was one of the key activities in the pre-enumeration period, starting from 2018. The mapping programme for the 2022 census adopted a digital approach, based on an integrated system composed of Geographic Information System (GIS) tools, Global Navigation Satellite Systems, high-resolution Earth Observation satellite images and tablets. The mapping field work was aimed at updating the census frame with new buildings that did not exist in 2015 census. The process was aimed at producing accurate maps to be used in the census enumeration process. The EA maps were delineated based on the number of households, a criteria that was set to determine a reasonable workload for enumerators during the enumeration process. Each building was geo-tagged, building on the existing geo-coded census frame. The mapping data was verified and office EA delineation was completed in 2022, resulting in 2,384 EAs.

1.3.3 Census Publicity

The success of the census depends on the quality of the information provided by the public. INETL updated the 2022 census publicity plan in February 2022 after the government approved the establishment of the census administrative structure. The campaign strategy identified four main target groups: government ministries and departments, traditional leaders, the media and the general public. A series of key messages were developed, including information about the census date, duration and importance. There were several census materials developed to promote census awareness. These included radio and TV adverts, a census jingle, stickers, leaflets, posters, banners, T-shirts, bags, umbrellas, hand fans, pens and notepads.

The INETL communication and public relations unit designed a census logo as well as developing a slogan and a census mascot. The slogan for the census was ‘Our census, our future be part of it’.
1.3.4 Field staff training

Due to their large numbers, the training of census field staff was implemented with a cascade approach, involving three levels of training:

From August 2022, census training started with 26 master trainers in Dili who later trained 180 field staff trainers in six parallel classes, also in Dili. Finally, decentralized trainings of 2,345 enumerators, 623 census supervisors and other local census staff were conducted across all 14 municipalities. Each of the trainings took at least 7 days. Apart from the master trainers who were all permanent civil servants, the majority of the other field staff were recruited for as temporary staff for the census enumeration phase. The training was standardized using a power point version of the instruction manuals for enumerators and supervisors that were developed as guidelines for the census fieldwork.

1.3.5 Census enumeration

The 2022 census data was collected through a traditional approach where households were visited by enumerators across the whole country to collect information on all individuals by asking pre-defined questions to the respondents. Enumeration was conducted in regular households as well as collective living quarter. The information collected was on a de jure population meaning that the people enumerated were usual household members as of the night of 4 September, 2023 (census night). The census enumeration period was from 5 September to 5 October, 2023, and was followed by a sweeping exercise to enumerate households that were missed out during the official enumeration period.

The information from the regular household population was collected through a long questionnaire that had all the census questions while a short questionnaire with only a few selected questions was used for the collective living quarters.

1.3.6 Post-Enumeration Survey

In accordance with international recommendation, a Post-Enumeration Survey was conducted to evaluate the quality of the census, mainly under-coverage or over-coverage of the population in 2022.

The Timor-Leste 2022 PES tools that included the questionnaire and the tablet computer application were developed as part of the census pilot phase. A full pilot PES was conducted
in April 2002 immediately following the pilot census to test these tools but also the training and enumeration procedures and initial matching procedures.

The PES sample a nationally representative sample stratified by municipality and urban-rural location. There were 16 thousand households in 149 EAs selected for a complete re-enumeration. The actual sample was drawn with probability-proportionate-to-size in late September 2022, strategically scheduled not to affect or influence completion of the main census enumeration. Other measures to preserve the independence between the census and the PES included the condition that PES field staff did not work in EAs where they were engaged in the census.

The training of the PES enumerators by INETL staff and a member of the technical assistance team was conducted from 4 to 5 November 2022. The three-week PES data collection period started on 7 November, which implies a relatively small period between the census and PES enumeration.

Though the PES data collection was conducted timely towards the end of 2023, the commencement of matching and analysis of the PES was postponed from mid-2023 to mid-2024.

1.3.7 Data processing and analysis

Computer-assisted personal interviewing (CAPI) using tablets was used in the census, which was for the first time to use such electronic data collection in a census. The CAPI application, developed in CSPro integrated digital maps of EAs that were used as a base for geo-referencing both the households and the housing units and to organise the work of the enumerators. Daily data synchronisation to the server managed at Census head office, enabled effective monitoring of the enumeration process including generating reports for the census management team.

This technological innovation enabled the data to be captured electronically during the enumeration phase but also to be transferred to the census head office therefore making it possible to monitor the data collection progress in real time. A census dashboard developed before enumeration, was used to produce daily reports and visualization of the progress by municipality and Administrative- Post level

Data editing to correct inconsistencies in the data started took place from November 2022 to April 2023. A significant improvement in the quality of the data collected was observed due
to the use of CAPI data collection. There was a set of editing rules developed, which showed that less than 5 percent of the records violated at least one skipping rule, that no implausible values were recorded for gender and that just few cases had missing or implausible values in the age.

The 2022 census collected some information that required coding after the enumeration – information such as field of study, occupation and industry. For the information on industry, the International Standard Industrial Classification of All Economic Activities (ISIC, rev 4) (United Nations, 2008) was used for response coding. The information on occupation was coded based on the International Standard Classification of Occupations (ISCO-08) (International Labour Office, 2012). The information on field of study was be coded on the basis of the International Standard Classification of Education, Fields of education and training (ISCED-F 2013) (UNESCO, 2015).

The census coding process started in December 2022 with manual coding by a team of ten coders and three supervisors from the pool of census secretariat staff. For the manual coding there was an existing code list in Tetum that was developed for the 2019 Timor-Leste Labour Force Survey. The manual coding process was time consuming, therefore later from February 2023, automated coding was attempted to match census information on occupation and industry using the partially coded census data. More efforts were made to finalise the coding later in 2023.

1.3.8 Census dissemination

Preliminary census results were released in November 2022 which was within six weeks of completion of census enumeration. There was a national dissemination seminar held in Dili with several Government officials and Development partners, where the report was shared.

The census main report was released in May 2023, at a national dissemination workshop, also held in Dili. In a similar manner like the preliminary report release, the participants were from both the Government and Development agencies.

In October 2023, selected indicators were compiled into a wall chart and disseminated. The indicators were presented at national, urban -rural, municipality and administrative post level.

All the census reports and the wall chart were also published on the INETL for easy access by the public.
2 Methodology and data

2.1 Rationale

Demographic changes affect the size and structure of the population, therefore understanding such information is important for development planning in countries. Age and sex are important variables in understanding the demographic composition of any population. Population censuses remain the major and reliable source of age and sex data in developing countries. The other source of such data is from national registration systems on vital events in the population such as births, deaths and marriages. For most developing countries including Timor-Leste, such systems are still being developed, therefore it is not possible to get complete information on such events that are expected to be reported by age and sex. Analysis of the most important indicators used in population analysis hinges on these two important characteristics of the population. The indicators that use age and sex in the calculations cover a wide range of topics such as fertility, mortality, migration but also projections. Inaccuracies in age and sex data have an effect on the quality of indicators produced. For example, studies in fertility have shown that a woman’s fertility declines with age. Fertility indicators are usually computed based on information collected on women only, but also focusses on women in childbearing ages which could be from the young age of 12 to older women of age 50. Therefore, any significant inaccuracies in quality of the age data can affect the estimates on the fertility indicators. The same can apply in mortality analysis, for instance where children’s age is not accurate, the estimates on child mortality rates can also be affected.

Population censuses as a source of demographic data are not exceptional to data quality issues arising from inaccuracies. It is important that census data is examined to assess the quality of age and sex of the population, as it will be useful information in understanding the series of thematic analysis for the 2022 census such as fertility, mortality, migration and most importantly population projections.

The basic objectives of this report are the following:

- To assess the accuracy of age and sex reporting from the 2022 Population and Housing Census
- To assess the pattern of digit preference in 2022
- To provide the variation in age and sex reporting at municipality level
2.2 Sources of errors

There are several reasons why the census data may have errors, and these are grouped into two main categories namely coverage errors and content errors as outlined in the international guidelines of censuses.

Coverage errors in census enumerations are of two types: inclusion of people in the enumeration who should not have been included known as overcount, and omission of people who should have been included known as undercount but also duplications where the person is enumerated more than once.

Content errors are those that arise from the incorrect reporting or recording of the characteristics of persons. The most common errors on age mis-reporting are age heaping or digit preferences.

As noted by Arriaga (1994) that during census enumeration, the respondent may provide age of a person that is purposely misrepresented or the age may be approximated where the actual age is not known. In other cases, the enumerator may estimate the age of the respondent or even the other household members, in the process providing age estimates that are not even close enough to the actual ages of the persons.

Such misreporting of age may result in under-reporting or over-reporting of certain ages. In this case, digit preference arises, and previous censuses have shown that population tends to report ages ending in 0s and 5s while avoiding reporting ages ending in digits the other digits such as 3 or 7. There may also be age-selective under-reporting such as leaving out infants, or older people.

2.3 Concepts of age and sex data evaluation

The graphical cohort analysis is aimed at tracking a population cohort at particular ages.

Population pyramid is a graphical representation of the age and sex composition of the population, representing the age and sex structure of the population.

Age ratio is the ratio of the population in a particular age group divided by half of the population in the adjacent ages.

Sex ratios is defined as the proportion of males to females in a population.
**Whipple's index** is a summary measure of age heaping on ages ending usually in the digits 0 or 5, used to determine variability in the quality of age reporting.

**Myers Blended Index** is another measure of age reporting on digit preference at ages ending in each digit from 0 to 9.

**UN Joint Age Sex Accuracy Index** is a combined or composite indicator of the overall age displacement in five-year age groups and the differential errors in age misreporting between males and females.

### 2.4 Data and methodology

#### 2.4.1 Data

The 2022 census questionnaire had questions related to age and sex for each person that was enumerated during the census. The question that was in the census questionnaire to establish sex was “What is <NAME>’s” gender? and the response options were male, female and other. During analysis, it was found that out of the 1.34 million people enumerated in Timor-Leste, only 50 persons were reported in the category ‘Other’. Since these were too few for any meaningful analysis, there were re-classified into either male or female based on the name listed in the questionnaire.

The other question asked on was the date of birth. The question had three separate questions as follows:

- What was <NAME> day of birth?
- What was <NAME> month of birth?
- What was <NAME> year of birth?

From this information, the age of the person was calculated by the in-built checks in the data capture system on the tablet. If the respondent did not know any of these three pieces of birthdate information, the response was recorded as unknown during the enumeration. For example, if only year was given, and the respondent could not remember the day and month, then those response were recoded under day and month of birth questions respectively.

A question was also asked to confirm the current age of a person that was calculated after providing the birth date.
For those who did not provide both the month and year of birth, there was a specific question asked on the age of that person on the census night of 4 to 5 September, 2022.

With these set of questions, information on age and sex was available for each person enumerated in 2022 census.

2.4.2 Method of analysis

The accuracy of the age and sex data from 2022 Population and Housing census will be evaluated by using the reported age distribution classified by sex.

Graphical analysis of the age and sex of the population using a population pyramid will depict the structure of the population. The population at the bottom of the pyramid can give an indication of the fertility in the population, while the rapid declining of population at older ages gives an indication of mortality experiences for the older population. The pyramid will be produced for single years. In addition, the graphical cohort analysis will be conducted for the 2022 and 2015 censuses.

The measures that will be used for evaluating the accuracy of age and sex data will include age and sex ratios, Whipple's Index, Myer's Blended Index, and the United Nations (UN) Age-Sex Accuracy Index.

The calculation for these indices will be done using Microsoft Excel spreadsheets. Other spreadsheets developed by US Bureau of Census such as follows:

PYRAMID – to construct five-year population pyramid
AGESEX for calculation of age and sex ratios, and the U Age Sex Accuracy Index
SINGAGE for assessment of digit preference using Whipples and Myers Indices.

Analysis will also use Data evaluation Excel spreadsheets from the IUSSP Tools of Demographic Estimation.
3 Analysis of age and sex data

3.1 Graphical analysis

3.1.1 Population distribution

The population enumerated during the 2022 census is depicted in Figure 3.1 below showing the distribution of the population by age and sex. There are more people at the younger ages than at the older ages and the gap is wider between ages 3 to 7 years but also between ages 49 to 50. Females are a lot more with a wider margin at older ages 72-74. The pattern of the population distribution by single age for both males and females is generally fluctuating showing a zig-zag pattern at both younger and older ages.

Figure 3.1: Population distribution by single age sex, 2022
Grouping the single age distribution into five-year age groups smoothest distortions in the population shown by the single ages.

The pattern of the population distribution shown in Figure 3.2 on the five-year age group distribution of the population looks much smoother with a few dips at age groups 10-14, 40-44 and 65-69. There is a slight spike of the population in the age group 15-19, probably due to the dip in the population in the age group 10-14 for both males and females.

Figure 3.2: Population distribution by five-year age group sex, Timor-Leste, 2022

3.1.2 Population Pyramid

The population distribution by age and sex takes the shape of a pyramid, unless mortality and fertility levels are unusual, or if the population has experienced significant levels of migration. In recent times, low fertility rates and decline in mortality rates have resulted in emerging population shapes that do not resemble a pyramid. This is a common feature of most developed countries. For countries like Timor-Leste where fertility and mortality remain high, the
population still shows a normal pyramid shape. When the population is displayed by single
years of age, it may also show age misstatement.

Graphical analysis of the age and sex of the population using a population pyramid depicts the
structure of the population. The population at the lower part of the pyramid is a good indication
of the fertility in the population, and the small proportions of population at older ages is an
ideal indication of mortality in the population.

The single age population pyramid from 2022 census data is depicted in Figure 3.3
summarizing the single age distribution of the population by sex. The data shows a pyramid
shape with more people at the younger ages that confirms higher fertility in the population. The
fewer people at the older ages show high mortality in the population where most people do not
survive to older ages. The 2022 pyramid also shows some deficit in the population at very
young ages but also between ages 40-45 for both males and females. There is also a deficit in
the population between ages 6 and 11 for both sexes.

**Figure 3.3: Single age population pyramid, Timor-Leste, 2022**

![Figure 3.3](image)

Figure 3.4 shows the 2015 population pyramid for single ages that has a similar shape to the
2022 pyramid. In terms of consistency in age reporting, the 2022 pyramid has cleanly trimmed
edges which shows better age reporting. The 2015 population pyramid shows some spikes at several ages such as 12, 20, 25 and 40 among others, showing over-reporting at those ages. The dip in the population can be seen at ages between ages 35 and 37 for both males and females. It is the same ages that are now appearing at older ages in early 40s in the 2022 data. The smooth edges of the population pyramid in 2022 give confidence that the quality of the age and sex data was very good.

**Figure 3.4: Single age population pyramid, Timor-Leste, 2015**

![Population pyramid](image)

### 3.1.3 Graphical cohort analysis

The graphical cohort analysis is important in census analysis as it tracks a population cohort at particular ages. The expectation is that the size of each cohort should decline over each census due to mortality, that is in the absence of notable international migration. The age pattern of the population for the two censuses in 2015 and 2022 should be also be similar in the absence of problems with quality of the census data.

In the 7-year period between 2015 and 2022 censuses, the population can be tracked using population cohorts in five-year age groups. The population enumerated at age 0 in the 2015 census would have reached the age 7 in the 2022 census. In order to track a complete five-year age group cohort with population that was enumerated in 2015, the analysis will compare from age group 10-14 up to 80-84.
The analysis presented in Figure 3.5 shows that more persons were enumerated in 2015 for each age cohort except for those in the cohort that was born between 1972 and 1977. Comparison of the two censuses by sex generally shows a decline of the population in this cohort in the 2015 census. Generally, the size of the population at the majority of the ages follows the expected pattern, as some people in the cohort from 2015 were lost through deaths in 2022.

**Figure 3.5: Graphical cohort analysis, Timor-Leste, 2015 and 2022**

Figure 3.6 presents the graphical cohort analysis for male population using 2015 and 2022 census data. The graph shows that there are slightly more males enumerated in 2022 for the cohorts born between 1966-1971 and 1972-1977. The results also show insignificant less males in the 1960-1965 birth cohort that were enumerated in 2022. This implies that in 2022, there were excess males enumerated in the age cohort 1960-1977 which could be a result of migration of males into the country during the intercensal period or an indication of errors in reporting of males by age in the 2015 or 2022 censuses.
The analysis for the female population shows that there is the expected pattern of fewer females for each age cohort as shown in Figure 3.7. The expected pattern deviates among females in the age cohort born between 1948 and 1953 where the enumerated females in 2022 were more than in 2015. There are also notably very few women less in 2022 census from the 1960-1965 and 1972-1977 age cohorts of the women.
3.2 Age and Sex ratios

3.2.1 Sex ratio

The sex ratio is the number of males per 100 females. At birth, this ratio generally ranges between 102 and 107, as biologically more males are born than females. Studies have also documented that a lot more females with increase in age compared to males, as a result the sex ratio tends to decline gradually with age.

For most middle age groups, and depending on migration level in a population, the larger the departure from 100, the larger the possibility that the deviation is from errors in the data.

The sex ratio of the Timor-Leste population stands at 102 males per 100 females. The sex ratio by five-year age groups at national level confirms variation in the sex composition of the population by age as shown in Figure 3.8. There are more males than females at ages below 15 years of age, averaging 105 males per 100 females but also at ages 45-59 years averaging 110 males per 100 females. For the population aged 70 years and over, the sex ratio is low at 82 males per 100 females.

Figure 3.8: Sex ratio, Timor-Leste, 2022
The results presented in Figure 3.9 show that in all municipalities, there are more males than females except in Lautem which has a sex ratio of 99 males per 100 females. Compared to 2015, Lautem continues to be the municipality with the lowest sex ratio. Manufahi is ranking as the municipality with the highest sex ratio at 109 males per 100 females, a shift from Dili which had a sex ratio of 108 males per 100 females in 2015. The sex ratios for 2015 and 2022 are presented in Annex Table 1.

**Figure 3.9: Sex ratio by Municipality, Timor-Leste, 2022**

![Figure 3.9: Sex ratio by Municipality, Timor-Leste, 2022](image)

### 3.2.2 Age ratio

Age ratio uses population data for 5-year age groups to detect age misreporting in populations where fertility has not fluctuated previously and where international migration is minimal. The procedure is simple as it involves dividing the population in a particular 5-year age group by the average population of the two adjacent 5-year age groups, expressed as a percent. The larger the fluctuations of these ratios, the larger their deviation from 1, therefore the higher the likelihood of data errors.

Figure 3.10 presents the age ratio by five-year age groups shows a similar pattern for males and females. The pattern is generally fluctuating, but it is moderate at the younger ages than at older ages for both sexes. Even though the results seem to point to some inaccuracies, however the gap is quite small.
Figure 3.10: Age ratio by sex and five year age-group, Timor-Leste, 2022

![Graph showing age ratio by sex and five year age-group](image)

### 3.3 Whipple's Index

The Whipple’s index was developed to measure preference for or avoidance of terminal digits 0 and 5. The index is calculated as the number of individuals between the ages of 23 and 62 whose reported age ends in 0 or 5, over the expected number of individuals whose ages should end in 0 or 5 in the 23-62 age group, multiplied by 100. The choice of 23 and 62 as the limits of the age band to be examined in the classic Whipple’s index calculation is arbitrary however it is known to be the most suitable for measuring age heaping in general in the population. Ages during childhood and old age are often excluded because they are more strongly affected by other types of errors of reporting than by preference for specific terminal digit.
Whipple’s index is known to vary from 0 to 500. The values of the index are classified as below:

<table>
<thead>
<tr>
<th>Index</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>digits ‘0’ and ‘5’ are not reported,</td>
</tr>
<tr>
<td>100</td>
<td>no preference for ‘0’ or ‘5’,</td>
</tr>
<tr>
<td>500</td>
<td>only the digits ‘0’ and ‘5’ are reported in the age data.</td>
</tr>
<tr>
<td>Less than 105</td>
<td>highly accurate age data</td>
</tr>
<tr>
<td>105-109.9</td>
<td>fairly accurate</td>
</tr>
<tr>
<td>110-124.9</td>
<td>approximate</td>
</tr>
<tr>
<td>125-174.5</td>
<td>Rough</td>
</tr>
<tr>
<td>At least 175</td>
<td>very rough</td>
</tr>
</tbody>
</table>

The total population within the ages 23-62 was 550.3 thousand of which 51.1 thousand persons reported ages ending in 0 and 57.6 thousand persons reported ages ending on 5. The Whipples index for the ages reporting in digit 0 is 93.0 that is lower than the Whipples Index ending in 5 that is calculated to be 104.8.

The analysis summarized in Table 3.1 reveals that the overall Whipples index, detecting a preference for ages 0 or 5 is 98.9, showing that the data is highly accurate. At municipality level the Whipples index ranges from 92.6 in Atauro to 103.4 in Manufahi. It has to be noted that a notably lower index below 100, shows preference of other digits other than 0 and 5.
Table 3.1: Whipple’s Index by sex at national and municipality, 2022

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aileu</td>
<td>98.3</td>
<td>97.2</td>
<td>99.5</td>
</tr>
<tr>
<td>Ainaro</td>
<td>100.3</td>
<td>99.3</td>
<td>101.4</td>
</tr>
<tr>
<td>Atauro</td>
<td>92.6</td>
<td>98.2</td>
<td>87.1</td>
</tr>
<tr>
<td>Baucau</td>
<td>98.2</td>
<td>98.7</td>
<td>97.6</td>
</tr>
<tr>
<td>Bobonaro</td>
<td>98.7</td>
<td>96.8</td>
<td>100.6</td>
</tr>
<tr>
<td>Covalima</td>
<td>99.0</td>
<td>101.6</td>
<td>96.4</td>
</tr>
<tr>
<td>Dili</td>
<td>98.7</td>
<td>99.5</td>
<td>97.9</td>
</tr>
<tr>
<td>Ermera</td>
<td>97.0</td>
<td>97.0</td>
<td>97.0</td>
</tr>
<tr>
<td>Lautem</td>
<td>101.5</td>
<td>101.2</td>
<td>101.7</td>
</tr>
<tr>
<td>Liquiça</td>
<td>100.9</td>
<td>101.1</td>
<td>100.7</td>
</tr>
<tr>
<td>Manatuto</td>
<td>98.5</td>
<td>100.3</td>
<td>96.7</td>
</tr>
<tr>
<td>Manufahi</td>
<td>103.4</td>
<td>103.2</td>
<td>103.6</td>
</tr>
<tr>
<td>Oecusse</td>
<td>97.2</td>
<td>96.7</td>
<td>97.8</td>
</tr>
<tr>
<td>Viqueque</td>
<td>98.8</td>
<td>99.4</td>
<td>98.2</td>
</tr>
<tr>
<td>Timor-Leste</td>
<td>98.9</td>
<td>99.2</td>
<td>98.6</td>
</tr>
</tbody>
</table>

3.4 Myers Blended Index

The Myers Blended index was developed to measure preference or avoidance of ages ending in each digit from 0 to 9. The theoretical range of the Myers index is from 0 to 90, where 0 indicates no age heaping and 90 indicates the extreme case where all recorded ages end in the same digit. Terminal digits with positive values mean are overreporting, while digits with negative scores show underreporting, compared to expected digit frequencies.

The results presented in Figure 3.11 on Myers preference on each of the terminal digits shows quite a small deviation from expected frequencies, with slightly higher under-reporting for age ending with the digit 1.
The total population age 10 years and older in 2022 census was 1.0 million persons. The Myers index for the country is calculated to be 3.3, showing that the data was very accurate. The index presented at municipality level presented in Figure 3.12 shows that the value ranges from 2.7 in Dili to 6.5 in Viqueque. The Myers index is very low, an indication of very good age reporting at sub-national level.

Figure 3.12: Myers Blended Index at national and municipality, 2022

![Figure 3.11: Myers preference by terminal digit index and sex, Timor-Leste, 2022](image)
Table 3.2 presents Myers Blended Index by sex at national and municipality. It is worth noting that the quality of reporting between males and females is equally good as variation in reporting between males and females at national level is 3.3 and 2.8 respectively. The index for females is highest in Ainaro (7.6) while that for males is highest in Aileu (6.3).

Table 3.2: Myers Blended Index by municipality and sex, 2022

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aileu</td>
<td>6.3</td>
<td>5.1</td>
</tr>
<tr>
<td>Ainaro</td>
<td>5.0</td>
<td>7.6</td>
</tr>
<tr>
<td>Atauro</td>
<td>5.6</td>
<td>5.5</td>
</tr>
<tr>
<td>Baucau</td>
<td>5.6</td>
<td>4.5</td>
</tr>
<tr>
<td>Bobonaro</td>
<td>3.1</td>
<td>3</td>
</tr>
<tr>
<td>Covalima</td>
<td>5.9</td>
<td>4.8</td>
</tr>
<tr>
<td>Dili</td>
<td>2.8</td>
<td>4.1</td>
</tr>
<tr>
<td>Ermera</td>
<td>4.8</td>
<td>5.1</td>
</tr>
<tr>
<td>Lautem</td>
<td>3.1</td>
<td>3.9</td>
</tr>
<tr>
<td>Liquiça</td>
<td>4.4</td>
<td>3.3</td>
</tr>
<tr>
<td>Manatuto</td>
<td>4.4</td>
<td>4.6</td>
</tr>
<tr>
<td>Manufahi</td>
<td>4.1</td>
<td>4.5</td>
</tr>
<tr>
<td>Oecusse</td>
<td>3.4</td>
<td>3.3</td>
</tr>
<tr>
<td>Viqueque</td>
<td>5.3</td>
<td>4.8</td>
</tr>
<tr>
<td>Timor-Leste</td>
<td>3.3</td>
<td>2.8</td>
</tr>
</tbody>
</table>

3.5 United Nations (UN) Age-Sex Accuracy Index

The United Nations (UN) Age-Sex Accuracy Index is based on sex ratios and age ratios, both of which are calculated in five-year age groups for populations below age 70. The age-sex accuracy index developed by the United Nations in 1952 and 1955 is the sum of the male and female age ratio scores plus three times the sex ratio scores all calculated using data for ages 0-14 through 65-69.

Scores obtained from UN Joint Age Sex Score are interpreted as follows:

- **Less than 20** Accurate data
- **Between 20 and 40** Inaccurate data
- **Over 40** Highly inaccurate data

The UN Age Sex Accuracy Index for the 2022 census is 28.1, which points to data that is classified as fairly inaccurate.
Figure 3.13 shows the UN Joint Age Sex Index at municipality level. The results show that the age and sex data from nine municipalities are not reliable with Viqueque showing the highest Index value. It should be noted, however, that the index does not take into account the normal decline of sex ratios by age, and distortions in the age structure due to large labour migration or displacements caused by wars or natural disasters.

**Figure 3.13: UN Joint Age Sex Accuracy Index by municipality, 2022**
4 Conclusion and recommendations

Assessment of age and sex data is the first step in evaluation of census data, as this is the key information in analysis of fertility, mortality, migration and population projections. It is important to understand the challenges with the data in order to understand the findings of further analysis of the 2022 census. There are several methodologies available for appraising age and sex data. In this report, the three indices used in the analysis were Whipple’s Index, Myers Blended Index and UN Joint Age Sex Score but also graphical analysis using population pyramid and graphical cohort analysis.

Even though it not possible to attribute the exact sources of errors in the data, the assessment can inform the census processes in the future. The results can also serve as a basis for further analysis of other aspects of the census process especially during data collection stage.

One important finding from this analysis of age and sex data assessment in 2022 census shows that the data is of very good quality. At municipality level, using the three indices, the data show that Bobonaro and Oecussi are consistently showing good quality age and sex data.

The graphical analysis showed that the 2022 census had minimal age mis-reporting. The population pyramid by single year ages in 2022 did not show significant outliers at each of the ages. This was a contrast to population enumerated in 2015 that had several spikes at various ages showing notable age misreporting.

Taking into account the fact that it was the first digital census to be conducted in the history of census taking in Timor-Leste, it is clear that the innovation improved the census enumeration. Both coverage in terms of complete enumeration and improved coverage of the population. Improvements in technology available at different stages of the census will have to be fully maximized for best census. The benefits of use of tablets during enumeration can further be reaped by taking advantage of real time data transfers that will enable timely checking of inconsistencies in the data. Timely feedback to enumerators during enumeration will go a long way to improve data quality.

For the future, it is recommended that population at Administrative Post level should be further analyzed to evaluate the age and sex data. This can possibly reveal the problematic localities, so that further analysis can also be targeted.
There should also be efforts to explore other data sources such as Civil Registration and Vital Statistics System, for further evaluation of data in Timor-Leste. The census and good quality CRVS data can complement to provide reliable demographic data sources.
References


Annex

Annex 1: Sex ratio at national and municipality, 2022

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Census Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2022</td>
</tr>
<tr>
<td>Aileu</td>
<td>107.1</td>
</tr>
<tr>
<td>Ainaro</td>
<td>104.7</td>
</tr>
<tr>
<td>Atauro</td>
<td>101.0</td>
</tr>
<tr>
<td>Baucau</td>
<td>102.1</td>
</tr>
<tr>
<td>Bobonaro</td>
<td>101.5</td>
</tr>
<tr>
<td>Covalima</td>
<td>103.5</td>
</tr>
<tr>
<td>Dili</td>
<td>103.0</td>
</tr>
<tr>
<td>Ermera</td>
<td>104.1</td>
</tr>
<tr>
<td>Lautem</td>
<td>99.1</td>
</tr>
<tr>
<td>Liquiça</td>
<td>102.7</td>
</tr>
<tr>
<td>Manatuto</td>
<td>103.9</td>
</tr>
<tr>
<td>Manufahi</td>
<td>108.7</td>
</tr>
<tr>
<td>Oecusse</td>
<td>103.3</td>
</tr>
<tr>
<td>Viqueque</td>
<td>101.4</td>
</tr>
<tr>
<td>Timor-Leste</td>
<td>103.1</td>
</tr>
</tbody>
</table>
Annex 2: Myers Blended Index by sex at national and municipality, 2022

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aileu</td>
<td>5.7</td>
<td>6.3</td>
<td>5.5</td>
</tr>
<tr>
<td>Ainaro</td>
<td>5.7</td>
<td>6.3</td>
<td>5.5</td>
</tr>
<tr>
<td>Atauro</td>
<td>5.7</td>
<td>6.3</td>
<td>5.5</td>
</tr>
<tr>
<td>Baucau</td>
<td>6.1</td>
<td>5.3</td>
<td>5.7</td>
</tr>
<tr>
<td>Bobonaro</td>
<td>3.6</td>
<td>3.4</td>
<td>3.3</td>
</tr>
<tr>
<td>Covalima</td>
<td>5.9</td>
<td>5.3</td>
<td>5.4</td>
</tr>
<tr>
<td>Dili</td>
<td>2.7</td>
<td>3.9</td>
<td>3.1</td>
</tr>
<tr>
<td>Ermera</td>
<td>4.8</td>
<td>5.2</td>
<td>4.9</td>
</tr>
<tr>
<td>Lautem</td>
<td>3.8</td>
<td>4.7</td>
<td>4.1</td>
</tr>
<tr>
<td>Liquiça</td>
<td>4.5</td>
<td>3.7</td>
<td>3.9</td>
</tr>
<tr>
<td>Manatuto</td>
<td>4.9</td>
<td>5.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Manufahi</td>
<td>4.0</td>
<td>4.3</td>
<td>4.0</td>
</tr>
<tr>
<td>Oecusse</td>
<td>3.6</td>
<td>3.8</td>
<td>3.3</td>
</tr>
<tr>
<td>Viqueque</td>
<td>6.5</td>
<td>5.6</td>
<td>6.0</td>
</tr>
<tr>
<td>Timor-Leste</td>
<td>3.4</td>
<td>3.1</td>
<td>3.3</td>
</tr>
</tbody>
</table>
# Contributors to Thematic Report – Evaluation of Age and Sex Data

<table>
<thead>
<tr>
<th>Name and Institution</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Elias dos Santos Ferreira, The President of the National Institute of Statistics of Timor-Leste (INETL, I.P.,) Ministry of Finance</td>
<td>Project Administration and Coordination</td>
</tr>
<tr>
<td>Mr. Silvino Lopes Director National Statistics Directorate</td>
<td>National Census Coordinator</td>
</tr>
<tr>
<td>Mr. Ricardo dos Santos Cruz Director National Statistics Directorate</td>
<td>Deputy Census Coordinator</td>
</tr>
<tr>
<td>Mr. Lourenco Soares Chief Data Processing</td>
<td>Data Collection Manager</td>
</tr>
<tr>
<td>Mr. Helder Henriques Mendes Chief Publicity and Advocacy</td>
<td>Chief of Publicity and Advocacy Campaign</td>
</tr>
<tr>
<td>Mrs. Paulina Rita Viegas INETL, I.P.,</td>
<td>Information Technology Services</td>
</tr>
<tr>
<td>Mr. Francelino Leão Freitas INETL, I.P.,</td>
<td>Information Technology Services</td>
</tr>
<tr>
<td>Mr. Alipio Cardoso Moniz INETL, I.P.,</td>
<td>Produced maps for monographs</td>
</tr>
<tr>
<td>Mr. João Soares Gusmão INETL, I.P.,</td>
<td>Produced maps for monographs</td>
</tr>
<tr>
<td>Mr. Silveiro Martins Pinto INETL, I.P.,</td>
<td>Administration and Finance</td>
</tr>
<tr>
<td>Mr. Domingos Sebastião Freitas INETL, I.P.,</td>
<td>Design and Typesetting Thematic Report</td>
</tr>
<tr>
<td>Name and Institution</td>
<td>Contribution</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Mrs. Pressia Arifin-Cabo, UNFPA Representative</td>
<td>Report Reviewer</td>
</tr>
<tr>
<td>Mrs. Domingas Bernardo, Assistant Representative UNFPA</td>
<td>Report Reviewer</td>
</tr>
<tr>
<td>Mr. Egidio da Costa, Population and Development Program Analyst, UNFPA</td>
<td>Report Reviewer</td>
</tr>
<tr>
<td>Secondinho Salsinha, Monitoring and Evaluation Program Officer UNFPA</td>
<td>Report Reviewer</td>
</tr>
<tr>
<td>Mrs. Angela Msosa, UNFPA Timor-Leste Technical Adviser Census Project</td>
<td>Report Reviewer</td>
</tr>
<tr>
<td>Mrs. Angela Msosa, UNFPA Timor-Leste Technical Adviser Census Project</td>
<td>Author of Evaluation of Age and Sex Thematic Report</td>
</tr>
</tbody>
</table>
INETL, I.P. = Instituto Nacional de Estatística de Timor-Leste, Instituto Publico

**Contact**

Telefone: +670 3311348
E-mail: dge@mof.gov.tl
Website: https://inetl-ip.gov.tl