EVALUATING THE AGE SEX DISTRIBUTIONS FROM SOUTH AFRICA'S 2022 CENSUS BASED ON THE TEN PERCENT SAMPLE

By

Emeritus Professor Eric O Udjo (PhD)
Independent Consultant
Eric Udjo Consulting
www.ericudjoconsulting.co.za

ABSTRACT

Age and sex are cross-cutting variables in an economy. The quality of the age-sex distributions in a census is a general pointer to the overall quality of the census data. It is therefore important to provide an assessment of the quality of the age-sex distributions from the 2022 South Africa's census. This study undertook such assessment using primarily, the unweighted 10% sample of the 2022 census that is available in the public domain. The results indicate the following. The overall sex ratio of 91.5 nationally from Census 2022 10% sample is low and odd. The single-year age distributions were reasonably smooth nationally but less smooth provincially. The age-sex distributions nationally were of better quality than observed in some other African countries censuses. The shape of the five-year population pyramid from the Census 2022 10% sample is, however, inconsistent with the shape of the five-year population pyramid derived from the 2011 census and more importantly, inconsistent with the five-year population pyramid derived from the 2021 General Household Survey undertaken approximately one year before the 2022 census. As in other censuses elsewhere in Africa and other regions, there is evidence of digit preference and age shifting in the five-year age distributions in the 2022 census. The weights in the 2022 10% sample imposed a different age structure on the original age structure in the 10% sample. Users of the Census 2022 10% sample data therefore need to be sensitive to these issues and especially the weights since age-sex distributions are linked to all aspects of the life cycle and impact planning in all sectors of the economy.

BACKGROUND

I was appointed by Statistics South Africa on 10 January 2023 as a consultant to evaluate the quality of the 2022 census data and advice the Statistician General and Statistics Council on the data's fitness of use with specific focus on the demographic data.

In all, I produced seven technical reports based on my evaluation of the final cleaned full data. Five of the technical reports focused specifically on the quality of the age, sex, fertility, mortality and migration data in Census 2022. The seven technical reports are not in the public domain and since the individual records from the full Census 2022 data are also not in the public domain, this study was based on the released demographic data in the 10% sample of Census 2022 (hereafter referred to as, Census 2022 10% sample).

Evaluation of census data based on a 10% sample is not ideal because sampling may introduce new errors into the data. Though using same methods as in the evaluation of the full data, I made no comparisons in the results obtained from the current study (i.e. from the 10% sample) with the results obtained from the full data. However, where possible and necessary, comparison was made with results obtained from the 1996, 2001 and 2011 censuses. One should be cautious in interpreting the results of such comparison since the results obtained from the previous censuses were based on the full data.

OVERVIEW OF SOUTH AFRICA'S POST-APARTHEID CENSUSES

Censuses in South Africa during the apartheid era did not consistently cover the entire union of South Africa. The last census covering the entire country during the apartheid era was in 1970.The apartheid censuses of 1980, 1985 and 1991 excluded the former homelands – Transkei, Bophuthatswana, Venda, and Ciskei (TBVC). Four censuses have been undertaken in South Africa since 1994 (i.e., post-apartheid) and include the 1996, 2001, 2011 and 2022 (the most current census).

According to the official reports from Statistics South Africa, the overall undercounts in the 1996, 2001, and 2011 census were respectively 11%, 18% and 14.6%. (The estimated undercount in the 1970 census was 2%). According to the officially published figures on South Africa's censuses, the population of South Africa increased from 40,583,573 in 1996, to 44,819,778 in 2001, to 51,770,568 in 2011. Using the exponential growth formula of population growth, the above figures imply that the annual intercensal growth of South Africa's population was about 2.0% between 1996 and 2001, 1.4% between 2001 and 2011. Aside the varying degrees of coverage errors, the post-apartheid South Africa's censuses also had varying degrees of content errors. Several aspects of these errors have been reported on by different researchers, see for example, Dorrington (1999), Sadie (1999), Shell (1999), Phillips et al. (1999). Udjo (2005a, 2014).

PROBLEM STATEMENT

Age and sex are cross-cutting variables since planning in all sectors of a population is directly or indirectly related to age and sex. One of the sources of data on age and sex distributions is the census. However, the information on age and sex distributions from a census may not be taken at face value because there are often errors in the age and sex data. It is therefore important that the quality of the age and sex distributions from censuses be evaluated to identify age errors that may be present in the data and if possible, adjusted for such errors. Failure to adjust for such errors in planning, might lead to inefficient allocation of resources to different sectors since the age and sex foci differ from one sector to the other. For example, the focus on age by an education department may be different from that by a health department in planning for future vaccination of children. The quality of the age distribution in censuses (and surveys) is often a pointer to the general quality of the census data.

APPROACHES IN EVALUATING DEMOGRAPHIC DATA

Internal Consistency Evaluation

Demographic phenomena are intricately interwoven. For example, the age and sex composition of a population reflects fertility, mortality and migration operating in the population (Newell 1997). In this approach one seeks to answer the question for example: are the fertility reports consistent with the mortality reports as well as the reported age structure?

Use of Demographic Models

Demographic models have been developed based on present demographic knowledge. These models were developed for evaluating and estimating demographic parameters from limited and defective data. Using some of these models, error deviations can be identified and separated from the real features of the data. Fitting the Relational Gompertz model to fertility reports or comparison of the cumulated reported age distributions with a stable model are examples of use of demographic models are examples of using demographic models to evaluated demographic data.

Comparison with Other External Sources

This approach entails comparing one or several demographic aspects (including indicators) of the census data with one or several other independent sources of data. For example, one may compare trend in childhood mortality from the current census with corresponding trend in a previous census. It should however be borne in mind, that there is no ideal standard that can be used as a benchmark for assessing the 'correctness' of a census as the external sources themselves may contain biases. The approaches outlined above were utilised in the current study though some aspects were constrained by the available data, for example, fertility and mortality from household deaths are currently not available in the 10% sample of the Census 2022.

AIM AND OBJECTIVES

The overall aim of this study was to provide an assessment of the quality of the age sex distributions in Census 2022 10% sample. The specific objectives were:

1. Assess the sex ratios from Census 2022 10% unweighted sample data nationally and by province, and in comparison, with previous censuses.
2. Identify patterns of anomalies in the age sex distributions if present in the Census 2022 10% unweighted sample data nationally and by province.
3. Estimate corrected age distributions at national level through smoothing of the age distributions. Smoothing age distributions at provincial level is not advisable due to the complicating factor of migration.
4. Assess the effect of weighting on the age-sex distributions from Census 2022 10% unweighted sample nationally. Since the weighted number of persons in the age distribution in the Census 2022 10% sample sum up to 61,367,678 (valid cases) which is approximately the officially adjusted population size figure of 62,027,503 from Census 2022, the weights in Census 2022% sample are therefore a combination of Post Enumeration Survey (PES) adjustment weights for undercount in Census 2022 as provided by Statistics South Africa as well as expansion weights to blow the sample to the officially adjusted population size. Both weights are potential sources of error. The sampling procedure to obtain the 10% sample is also a potential source of error but is not the focus of this study.

DATA

As indicated above, the source of data for the evaluation of the age-sex distributions was primarily Census 2022 10% sample. This was published by Statistics South Africa in early September 2024. I downloaded the data from Statistics South Africa's website on 5 September 2024. Census 2022 10% sample in this study refers to this version of the data. The external sources of data used for comparison were the post 1994 South Africa's censuses: 1996, 2001 and 2011 censuses. The number of persons whose ages were provided in the unweighted census 2022 10% sample were 4,230,792. The corresponding figure for the weighted sample was 61,367,678. Note that fertility and household deaths data were not included in the Census 2022 10% sample that I downloaded at the time of this study, but they are important in a holistic evaluation of an age-sex distributions.

METHODS

One of the challenges in evaluating an age-sex distribution in a population is migration. An age-sex distribution reflects past fertility, mortality, and migration (Newell 1997). Whereas the impact of fertility and mortality on an age-sex distribution can be assessed through a careful analysis of past trends in fertility and mortality, the impact of migration is more difficult to assess, primarily due to lack of reliable data. It is therefore difficult to control for the impact of migration in the evaluation of age-sex distributions of South Africa's population.

There are often two forms of error in age-sex distributions: heaping (preferences for certain ages typically ending in even numbers, zeroes and fives) and shifting (transfer of one's age group to another). A commonly used approach in assessing the presence of these errors is to compute summary indices such as Whipple's and Myer's indices or United Nations joint scores. I find this approach unsatisfactory because such summary indices conceal much of the details in errors in age-sex distributions. To obtain deeper understanding of the pattern of errors in the data therefore, sex ratios were computed followed by the construction of single and five-year population pyramids, and finally, smoothing the age sex distributions.

Sex Ratios. A sex ratio is conventionally defined as the number of males per 100 females. The general formula for computing a sex ratio is:

$SexR_i = (nM_i/nF_i) * 100$ ……. 1

Where $SexR_i$ is the sex ratio for a specific group, *nMi* is the number of males in that group, and nF_i is the number of females in that group. Thus, sex ratio may be computed for the entire population (overall sex ratio) or for specific age groups (age specific sex ratios).

Population Pyramids: Population pyramids are two sets of histograms depicting the frequency distributions attributed to different age groups. The percentage in an age sex group in the histogram is conventionally computed as a proportion of the number of persons in that age sex group, of the total population (total males and females). A common mistake in the construction of population pyramids is the inclusion of the open-ended age group in the pyramid. This gives a distorted impression of the age sex distribution if the open-ended age group is included. Single year age distributions are suitable for detecting digit preference while five-year population pyramids may detect age shifting if present in the data.

Age Smoothing: This involved comparing the logit transformations of the cumulated age-sex distribution with an appropriate stable model. It should be noted that the use of a stable population in this context does not assume that the population under scrutiny is stable, rather it pushes the results in that direction (Brass, 1984). The logit transformations given by Brass (1971) are denoted as:

$Y_x = ½\ log_e(((1-P_x)/P_x))$ and ……(2)

$Y_{sx} = ½log_e(((1-P_{sx})/P_{sx}))$ ……(3)

Where Y_x and Y_{sx} are the logit transformations of the distributions in the data and stable age distributions respectively, P_x and P_{sx} are the cumulate proportions under age *x* in the data and stable model respectively. Analysis of the plotted values was undertaken to derive smoothed distributions at national level.

RESULTS

For brevity in presenting the results in what follows, Census 2022 is used hereafter to refer to Census 2022 10% sample data but within the context of Census 2022 in general.

SEX RATIOS

Table 1 shows the overall sex ratios from the unweighted Census 2022 in comparison with the overall sex ratios from previous censuses without PES adjustments.

Table 1: Overall Sex Ratios from Census 2022 10% Sample and in Previous Censuses

Region	2022	2011	2001	1996
National	91.5	**93.2**	**90.1**	**92.2**
Eastern Cape	91.1	88.9	85.8	85.7
Free State	88.2	92.2	90.5	97.0
Gauteng	95.2	100.9	98.4	103.1
KwaZulu-Natal	87.9	88.4	85.9	87.7
Limpopo	87.9	85.9	82.2	83.2
Mpumalanga	93.4	94.1	90.1	92.1
North West	96.4	98.7	95.0	97.5
Northern Cape	91.3	96.0	92.8	94.0
Western Cape	91.3	95.1	93.8	95.6

Source: Author's computation from South Africa's post 1994 censuses

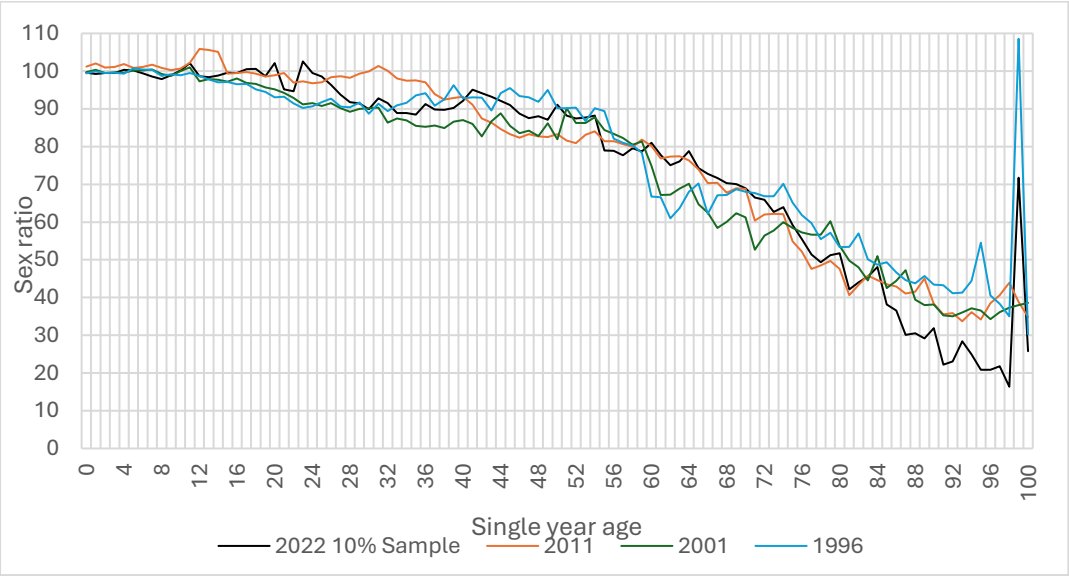
The table indicates an overall sex ratio of 91.5 nationally in 2022 based on Census 2022 unweighted 10% sample. This is seemingly low. Such low overall sex ratio is often observed in national populations with large scale emigration of males or in war situations i.e. substantial number of males leaving the population to fight in external war or large-scale mortality among males fighting internal or external wars. It is unlikely that the low overall sex ratio from Census 2022 unweighted 10% sample data is due to these factors. Low overall sex ratio is a consistent feature in South Africa's censuses and surveys (see Udjo 2005b for more details) and can be seen from the consistency of the overall sex ratio in 2022 compared with previous censuses (Table 1). As seen in the table, the overall sex ratio derived from the 2001 census (90.1) was even lower than that from the Census 2022 10% sample. Although there are usually excess females over males in human populations with a few exceptions (e.g., in some Muslim and Asian countries due to socio-cultural factors) resulting from differential sex ratio at birth, differential age specific mortality and to some extent by migration, the magnitude in the excess of males implied from the overall sex ratios from South Africa's censuses is improbable. Usually where there is no differential coverage of the census by sex and if the impact of international migration in the population is negligible, one would expect overall sex ratio in a population ranging between 95 – 99. The low overall sex ratio in Census 2022 and in previous censuses in South Africa is most probably due to lower coverage of males relative to females during the post 1994 censuses. It is noteworthy

that the overall sex ratio obtained from the 1970 census was 97. Other than under enumeration of males relative to females, I am unable to fathom the demographic processes that would have resulted in the overall sex ratio in South Africa dropping from 97 in 1970 to 91 in 2022 even after considering differential migration and mortality by sex. Sex ratios change very little over time.

The provincial pattern shown in Table 1 indicates low overall sex ratios in all the provinces in the census years except Gauteng and the North West. The overall sex ratio of 95.2 and 96.4 in Gauteng and the North West respectively are plausible. It may be argued that in the case of Gauteng and the North West, differential male migration relative to females from Limpopo may have contributed to the higher overall sex ratio in Gauteng and the North West. If that were the case, one would have expected a higher overall sex ratio in the Western Cape in 2022. The overall sex ratio of 91.3 in 2022 in the Western Cape is not plausible but the overall sex ratio of 95.1 and 95.6 in 2011 and 1996 in the Western Cape are plausible. The low overall sex ratios in most of the provinces in 2022 is probably primarily due to under-enumeration of males relative to females during Census 2022. Differential migration may have contributed also to the low overall sex ratios at provincial level, but it is difficult to establish this.

Further evaluation of the sex ratios was undertaken by examining the age specific sex ratios from Census 2022 10% sample data in comparison with previous censuses. As seen in Figure 1, in general, the pattern of the age-specific sex ratios from Census 2022 is consistent with what is expected: declining sex ratio with age resulting primarily from higher male than female mortality at any age. This is also consistent with the pattern in previous censuses. However, the curves in Figure 1 are not smooth and this could be due to differential age errors by sex in the data. Note the high spike in the sex ratio at age 99 years in Census 2022 as well as in the 1996 census. This is odd. It is not clear how much, editing of the data contributed to the pattern in Figure 1 but probably reflects age errors as well as administrative decision regarding the cleaning of the age data.

Figure 1: Age specific Single Year Sex Ratios from Census 2022 10% Sample and Previous Censuses



Source: Author's computation from Census 2022 10% Sample

To minimise the effect of differential age errors in the data, sex ratios were computed for broad age groups by population group from Census 2022 10% sample to identify the primary source of the overall low sex ratios. The results are shown in Table 2. As seen in table, the overall low sex ratios in Census 2022 were primarily driven by the African population. In a study focusing on age-sex distributions Udjo (2005b) observed that the birth history of the 1998 October Household Survey (now replaced by the General Household Survey) provided better indicator of sex ratios among the African population, though it was not clear why the mothers would report sex more accurately in the birth histories than sex in the household member listing.

Table 2: Sex Ratios in Broad Age groups by Population Group from Census 2022 10% Sample

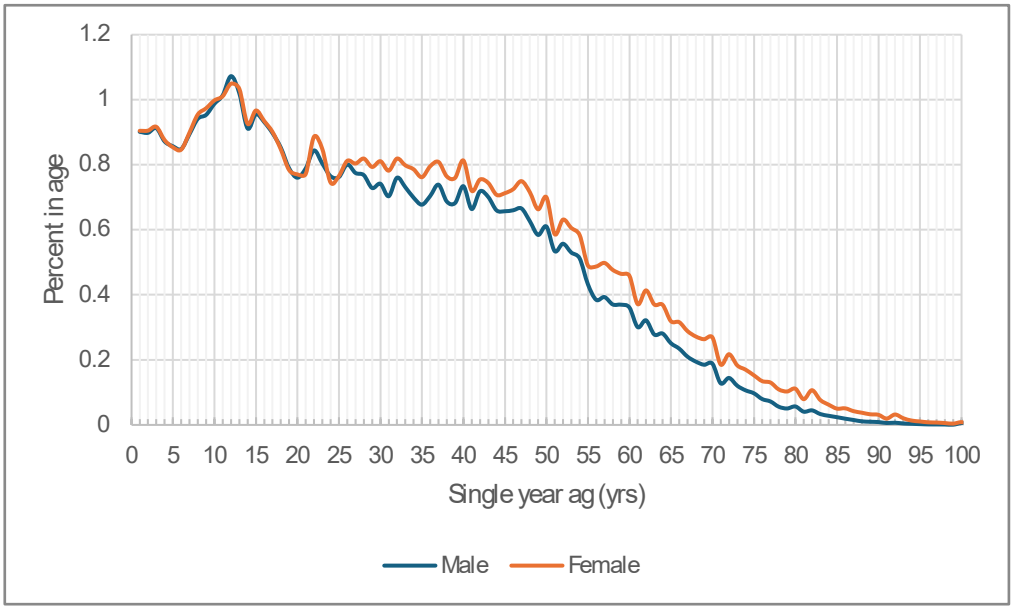
	Broad age group (years)		
	< 15	***15 - 64***	***65 & over***
National	**99.4**	**91.3**	**60.8**
African	99.0	89.8	54.6
Coloured	101.1	92.5	67.4
Indian/Asian	101.2	105.0	82.9
White	102.5	100.6	87.5

Source: Author's computation from Census 2022 10% Sample

Digit Preference

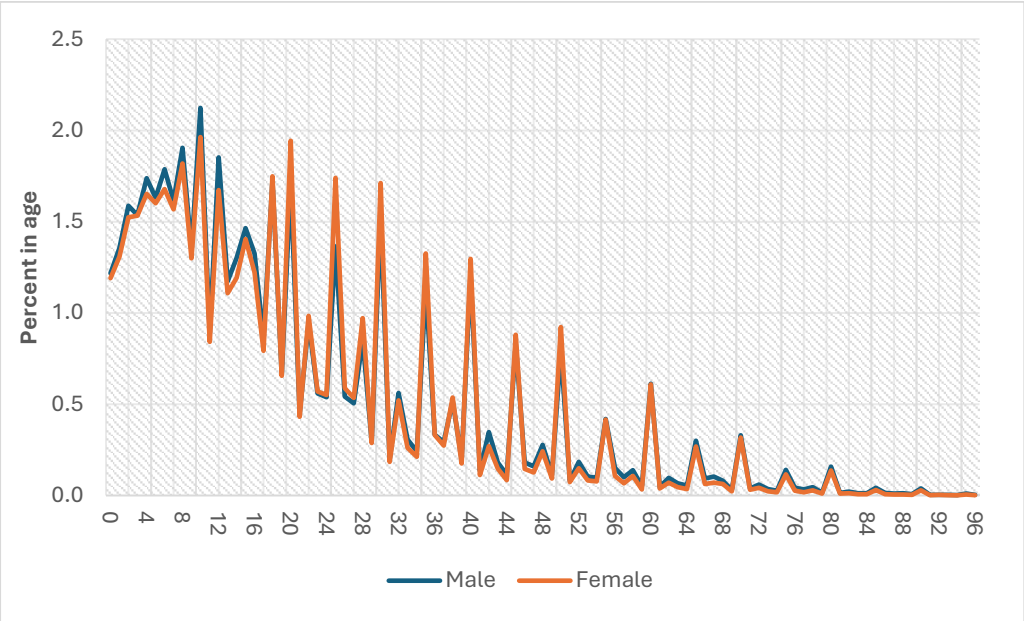
A useful tool in detecting digit preferences if present in age distributions is to examine single year age-sex distributions using line graphs though the conventional approach is to construct single year population pyramids. Single year age-sex distributions line graphs are easier to read off the ages where digit preferences are present than in single year population pyramids. Figure 2 shows the line graph of single year age-sex distributions from Census 2022 10% sample. As seen from the graph, the curves are reasonably smooth suggesting that the digit preference was not a huge problem in Census 2022. A comparison with age sex data from some other countries in the African continent suggests that the age-sex report from Census 2022 was of better quality than those from some other African countries. See for example Figure 3 for an African country in which the spikes and troughs are much more pronounced than those from South Africa's Census 2022. Figure 2 however, suggests that there was some degree of digit preference in the age distributions in Census 2022. The more pronounced digit preference was primarily in ages ending in even numbers (the spikes): 12, 22, 40, 50, This contrasts with the digit preference in the 2011 census in which the digit preference was more pronounced in ages ending in odd numbers: 21, 31, 51, 61,71 (Figure 4). The low proportions of males and females in ages below 12 years in Census 2022 are most probably due to under representations (partly due to under enumeration) of persons in these ages. There was a similar feature in Census 2011.

Figure 2: Single-year Age-Sex Distribution from Census 2022 10% Sample



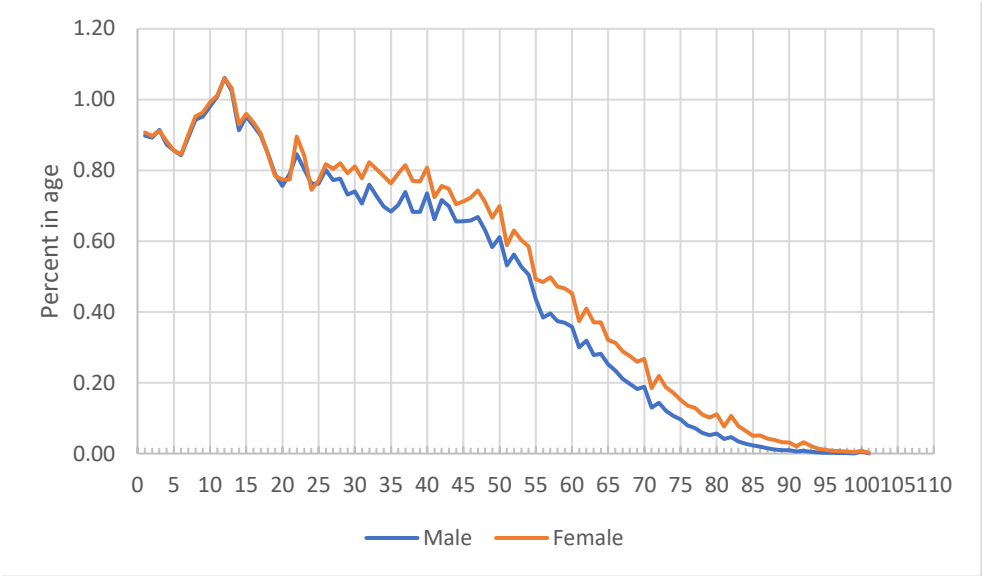
Source: Author's computation from Census 2022 10% Sample

Figure 3: Single Year Age Sex Distribution from and African Country's Census



Source: Author's computation from an African Country's census

Figure 4: Single-year Age-Sex Distribution from Census 2011

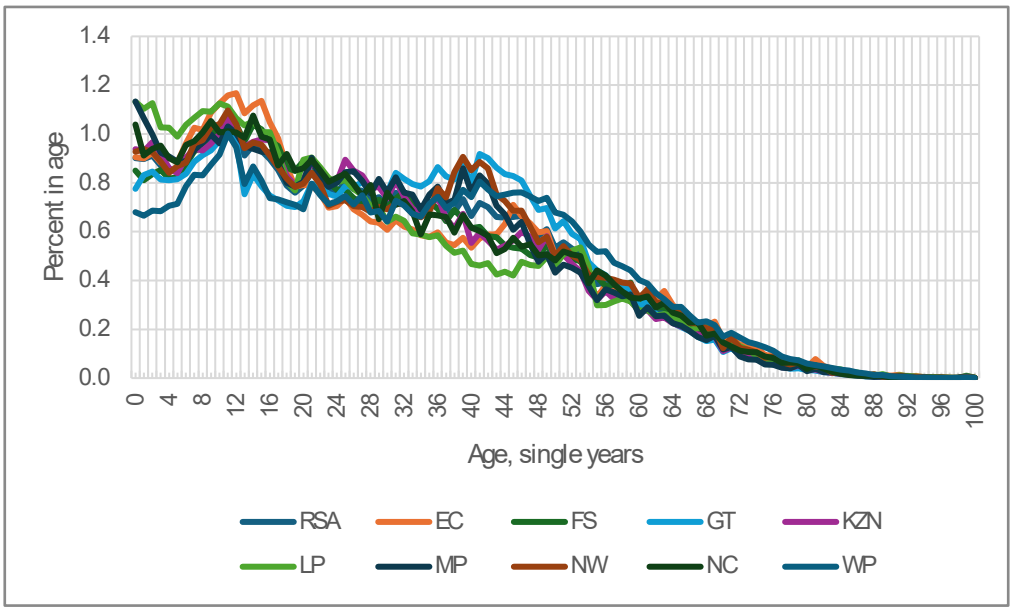


Source: Author's computation from South Africa's post 1994 censuses

The provincial pattern of digit preference in Census 2022 is illustrated in Figures 5 and 6. As seen in the graphs, the structure of the provincial single-year age distributions is not as smooth as the distributions in the national population for both males and females. Figure 5, depicting the provincial single year age distributions for males has pronounced humps in two sections due to spikes in even numbers around ages 12 and 40 indicating preference for even numbers relative to odd numbers in

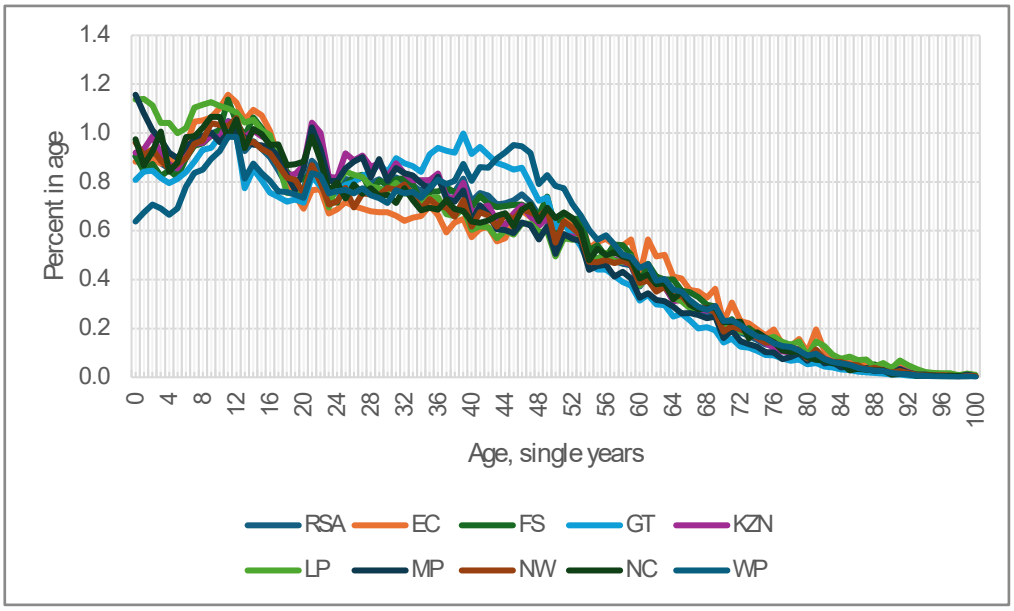
those ages. A similar pattern can be seen in Figure 6 depicting the pattern for females in the provinces but there is a third hump though less pronounced in age 21 years indicating the preference for this age in some of the provinces. Figure 5 appears to suggest an apparent deficit of males spanning ages 19 -39 in all the provinces as well as apparent deficit of males spanning ages 41 to 48 in all the provinces except the Western Cape.

Figure 5: Single-year Age-Sex Distribution from Census 2022 10% Sample, Males



Source: Author's computation from Census 2022 10% Sample

Figure 6: Single-year Age-Sex Distribution from Census 2022 10% Sample, Females



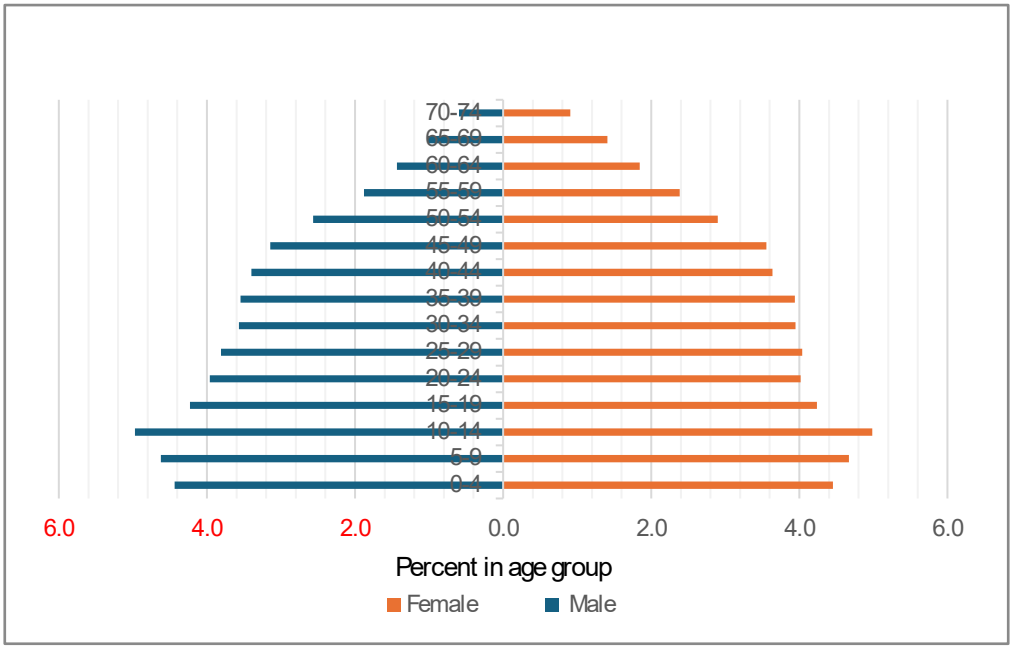
Source: Author's computation from Census 2022 10% Sample

The digit preference was more pronounced in ages ending in even numbers and less pronounced in ages ending in zeros and fives among males.

Age Shifting

Age shifting, though difficult to detect, is better assessed using five-year population pyramids. Figure 7 shows the five-year population pyramid constructed from Census 2022 10% sample. The pyramid is bulgy in certain age groups suggesting age shifting of person's true age among males aged 10-14, and 35-39, and among females aged 25-29,35-39 and 45-49. This is different from the pattern of age shifting in the 2011 census in which age shifting seemed to have been primarily in the age group 25-29. See the bulge in that age group in Figure 8. Thus, the pattern of age shifting in Census 2022 seemed different from the pattern in Census 2011.

Figure 7: Five-year Age Population Pyramid from Census 2022 10% Sample

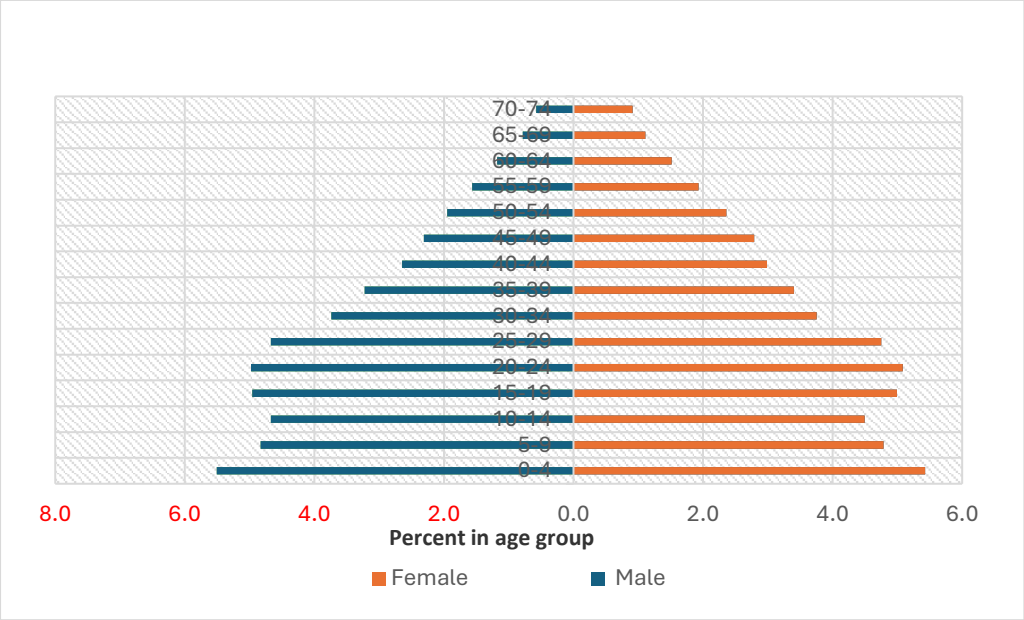


Source: Author's computation from Census 2022 10% Sample

Also, the structure of the population pyramid derived from Census 2022 10% sample (Figure 7) is different from the structure of the pyramid derived from Census 2011 (Figure 8). While the former (Figure 7) is more beehive in shape, the latter (Figure 8) is bell shaped. There would have had to be a steep decline in fertility and mortality in the last 50 – 60 years in South Africa to produce the shape of the population pyramid constructed from Census 2022 10% sample. It is unlikely that was the case because eleven years separate the last two census and sharp decline in fertility and mortality in the

last 11 years will not produce the beehive population pyramid seen in Census 2022. For example, the bulge in the population pyramid for persons aged 10-14 in the population from Census 2022 is inconsistent with an argument of steep decline in fertility in the population last 10-15 years. The bulge in the pyramid in that age group suggests the opposite argument – increase in fertility last 10-15 years – an unlikely situation. Further insight on this issue may be gained from fertility analysis from Census 2022. Unfortunately, the fertility data from Census 2022 are not (yet) available in the public domain.

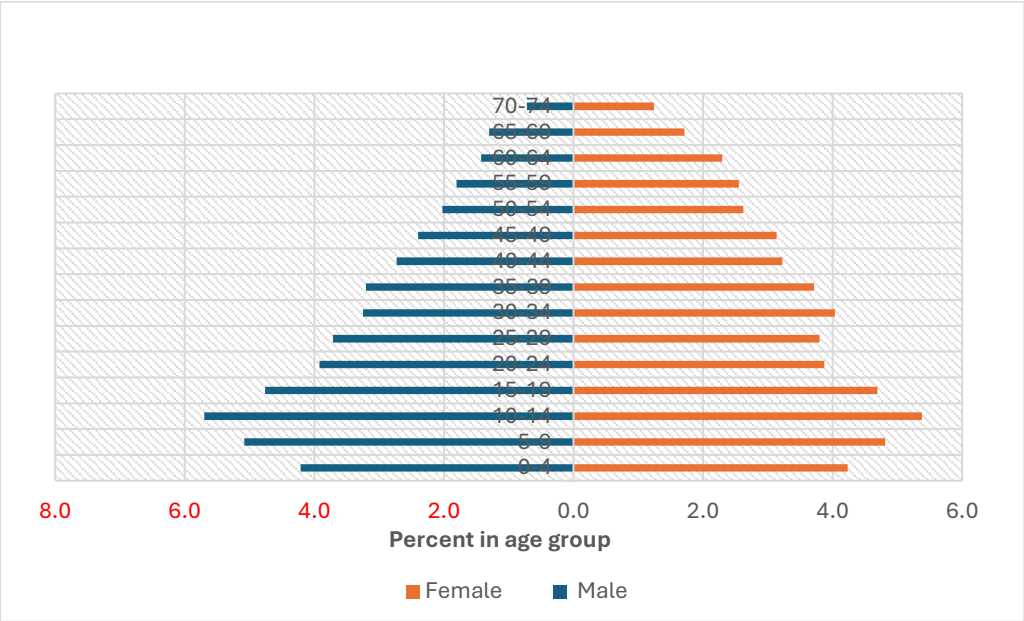
Figure 8: Five-year Age Population Pyramid from Census 2011



Source: Author's computation from Census 2011

It is clear from the above pyramids that the population pyramid be it single year (not shown) or five-year age sex pyramid from Census 2022 10% sample is inconsistent with the five-year pyramid from Census 2011. More importantly, the five-year age-sex pyramid from Census 2022 is not consistent with the five-year age sex pyramid from the 2021 General Household Survey (Figure 9). The 2021 General Household Survey was undertaken about year prior to 2022 Census, so one would expect close similarity between the population pyramids from Census 2022 and that from the 2021 General Household survey.

Figure 9: Five-year Population Pyramid from 2021 General Household Survey

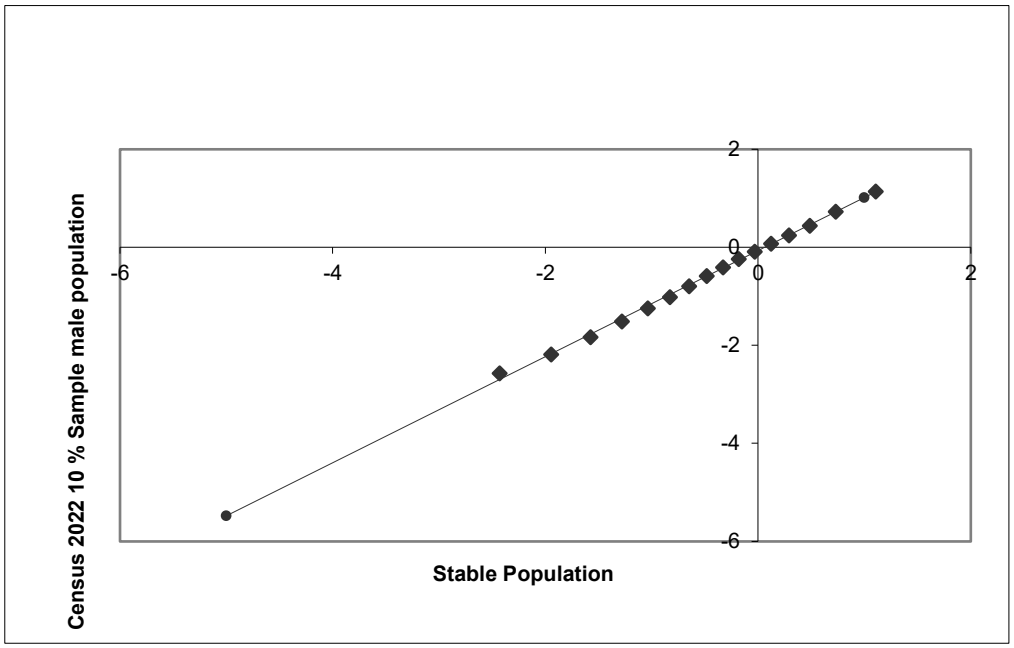


Source: Author's computation from 2021 General Household Survey

Age Smoothing

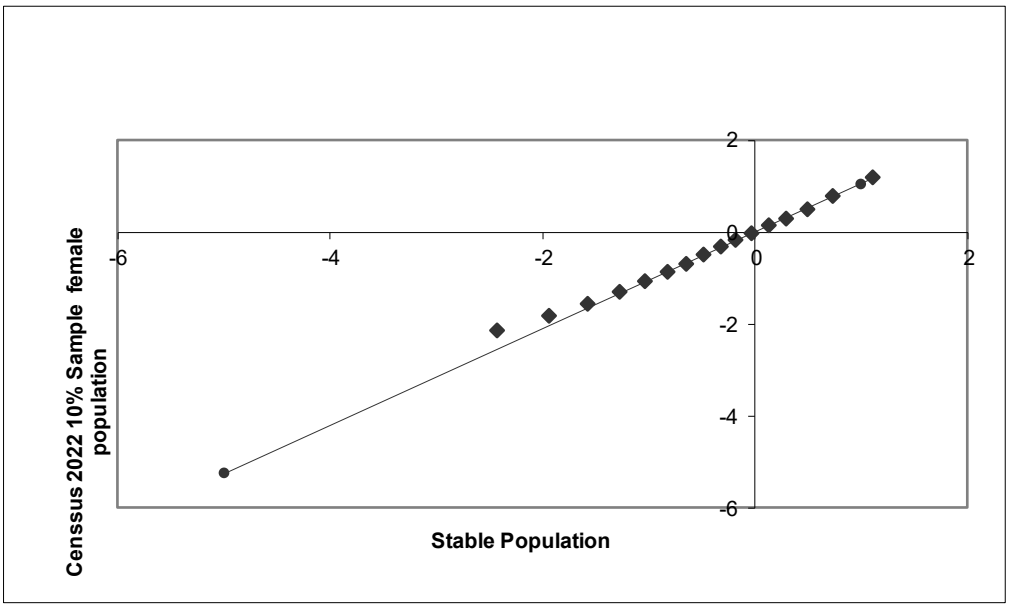
Age smoothing of the five-year distribution from Census 2022 10% sample was undertaken on the assumption that the Census 2022 age-sex distributions are partly attributable to content errors, The overall sex ratio was not tampered with in the smoothing process because that would have been tantamount to "turning" some females into males technically. To prevent that, the smoothing was done within each age-sex group. The results comparing the logit transformations of the cumulated age-sex distribution from Census 2022 10% sample with an appropriate stable model (see Carrier and Hobcraft 1971) are shown in Figures 10 and 11. As seen in the graphs, the stable age distribution fit the Census 2022 five-year age-sex distributions well (r = 1.00) for both males and females)). The resulting smoothed age-sex distributions from Figures 10 and 11 are shown in Figures 12-13. A negative value implies that the percentage in the age group in Census 2022 is lower than what the demographic modelling suggests. A positive value implies the opposite. Thus, Figure 10 suggests that the percentages in the age groups were higher among males aged 5-9, 10-14, 35-39, 40-44, 45-49 and 50-54 in Census 2022 than what the demographic modelling suggests. The pattern is similar for females though in varying degrees except for the age group 35-39. The corrected age distributions provide indications of the age groups that were affected by age shifting in Census 2022.

Figure 10: Logit Transformation of the Cumulated 2022 Census 10% Sample and Stable Age Distributions: Males



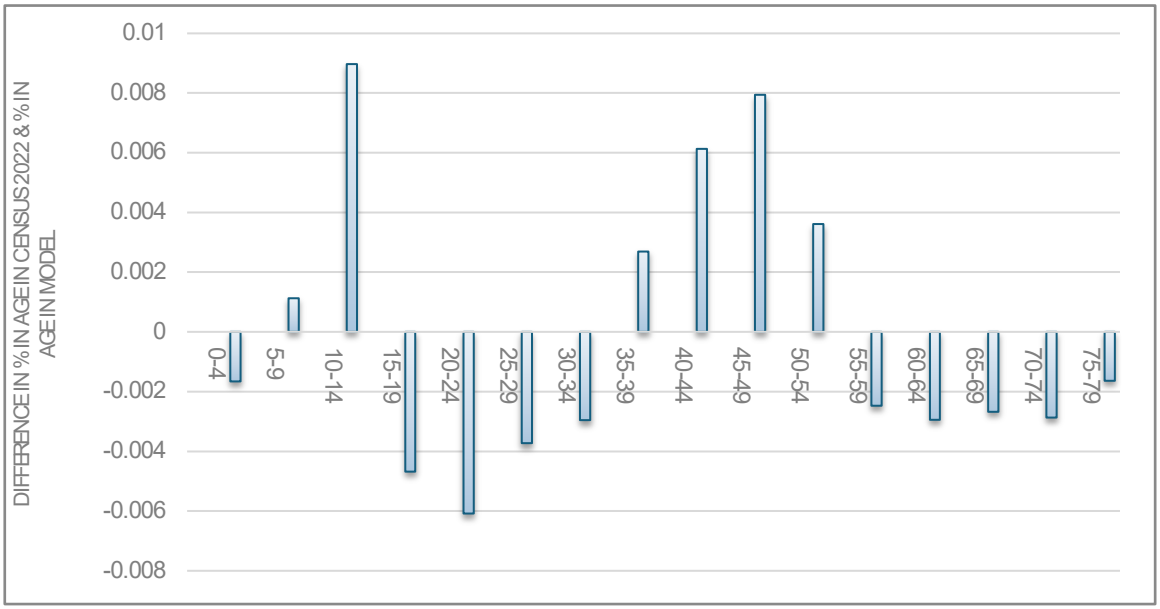
Source: Author's computation from Census 2022 10% Sample

Figure 11: Logit Transformation of the Cumulated 2022 Census 10% Sample and Stable Age Distributions: Females



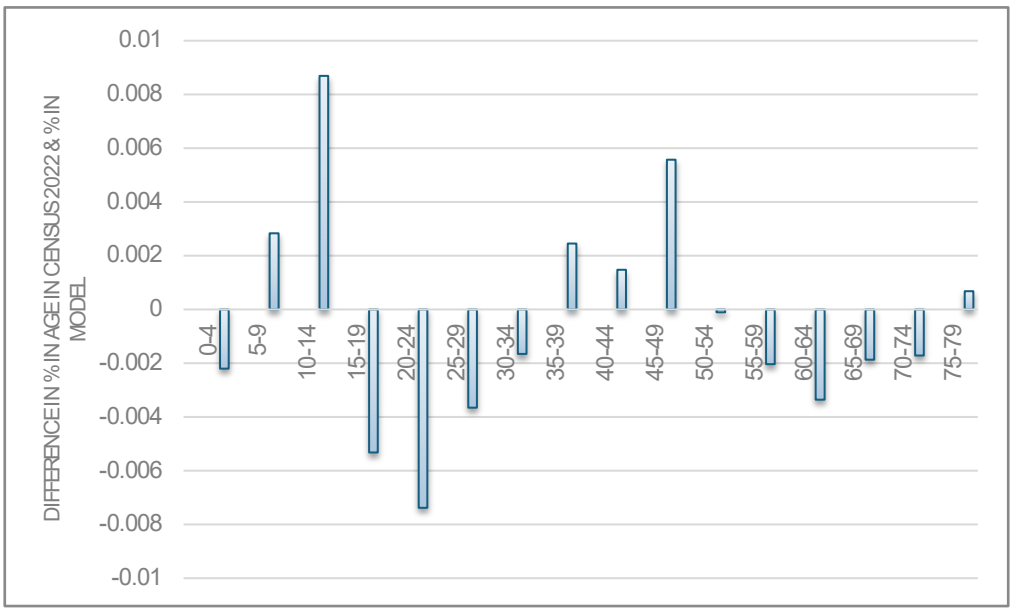
Source: Author's computation from Census 2022 10% Sample

Figure 12: Difference between the Percentage in Age in Census 2022 10% Sample and Percentage in Age in Model: Males



Source: Author's computation from Census 2022 10% Sample

Figure 13: Difference between the Percentage in Age in Census 2022 10% Sample and Percentage in Age in Model: Females



Source: Author's computation from Census 2022 10% Sample

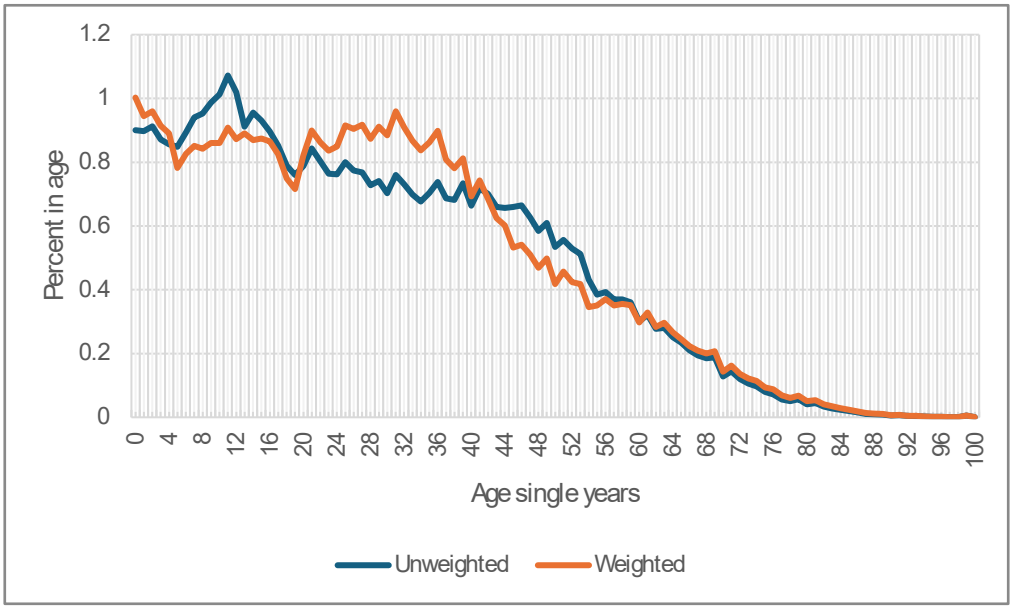
Effect of Weighting on the Age-Sex Distributions in Census 2022 10% Sample

The above analysis was performed on the unweighted 2022 10% sample. As noted above, the weights in the Census 2022% sample is a combination of PES and expansion weights. PES adjustments are

meant to adjust for undercount in a census while expansion weights are meant to raise the sample variable values to the officially adjusted population size for each variable value. Both actions should, however, maintain and not distort the structure of the original data.

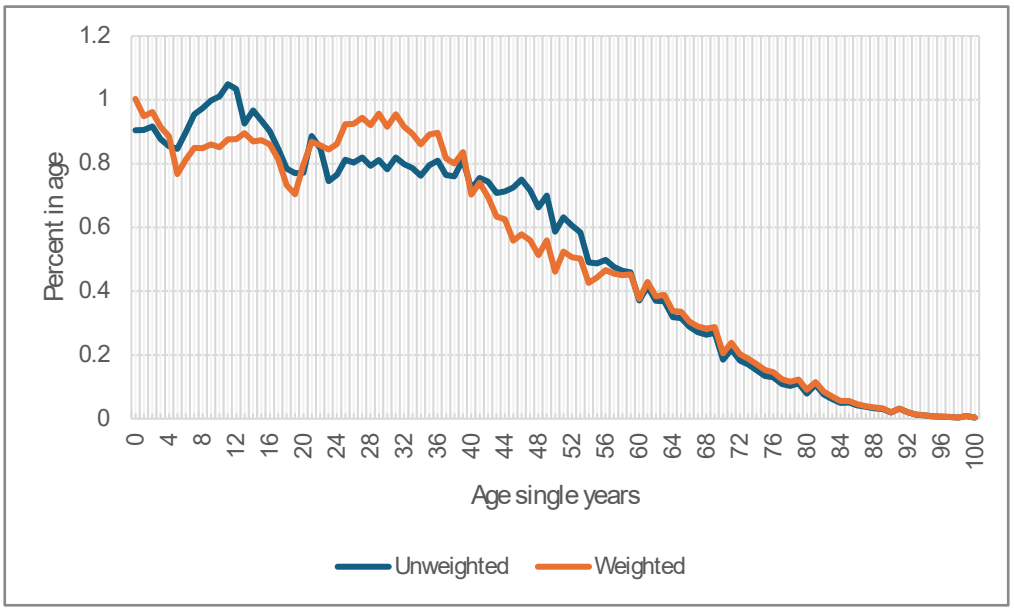
Figures 14-15 show the single year age distributions from the unweighted and weighted 2022 10% sample. As seen in the graphs, the weights in the 2022 10% sample distort the original single year age distributions up to age 56 among males and up to age 58 among females nationally. It is also evident comparing the five-year population pyramids from the weighted 2022 10% sample (Figure 16 below) with the unweighted five-year population pyramid (Figure 7 above) that weighting distorts the original five-year age distributions nationally. As in the single-year age distributions, weighting imposed a different age structure on the original five-year age structure nationally. Although a similar analysis comparison was not done at provincial level, but no doubt, weights in the 2022 10% sample would have also imposed different age structures on the original age structure at provincial levels in varying degrees.

Figure 14: Single-year Age Distribution from Census 2022 10% Sample, Unweighted and Weighted, Males



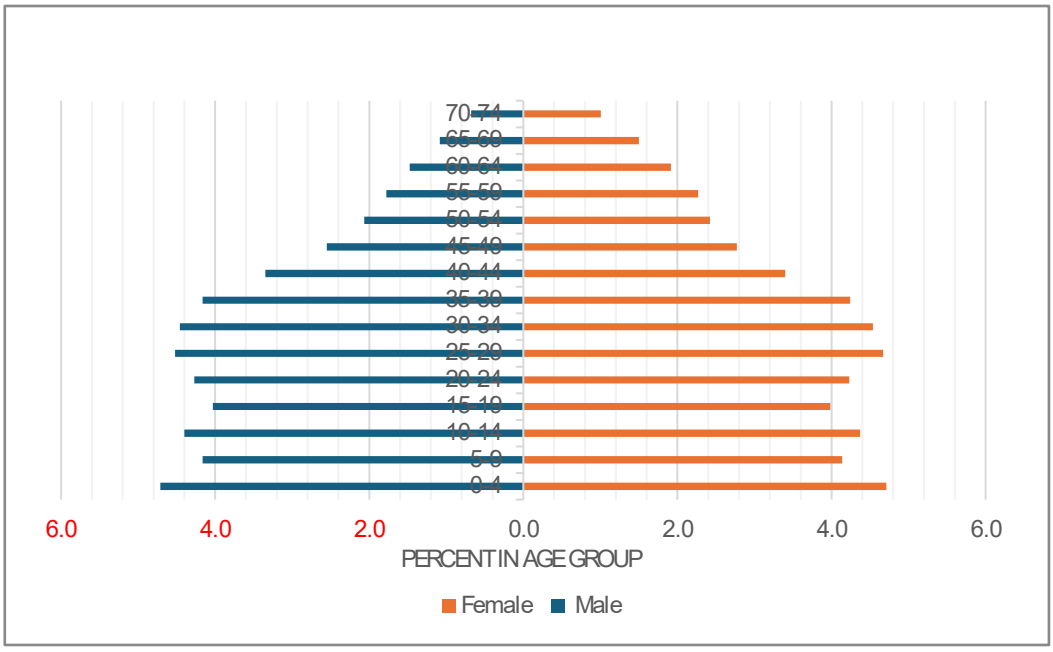
Source: Author's computation from Census 2022 10% Sample

Figure 15: Single-year Age Distribution from Census 2022 10% Sample, Unweighted and Weighted, Females



Source: Author's computation from Census 2022 10% Sample

Figure 16: Five-year Age Population Pyramid from Census 2022 10% Sample, weighted



Source: Author's computation from Census 2022 10% Sample

DISCUSSION AND CONCLUSION

Age and sex are linked to all aspects of the life cycle, including childhood, education, marriage, childbearing, entry into the labour market, retirement, ageing, morbidity, and mortality. It is therefore important to provide an assessment of the quality of the age-sex distributions from Censuses. The following were observed from such assessment.

The overall sex ratio of 91.5 nationally from Census 2022 10% sample is seemingly low and odd. The low overall sex ratio in Census 2022 is a consistent feature in South Africa's post 1994 censuses. For example, the unadjusted overall sex ratio from the 2011 census was 93.2. However, the officially published numbers from the 2011 census after the PES adjustment indicated an overall sex ratio of 95. It is not clear how that figure was derived but to boost a low overall sex ratio to a higher figure implies technically, turning some females into males in the process of adjustment.

The single-year age distributions from Census 2022 10% sample were reasonably smooth nationally but less smooth provincially. Comparing the single year age sex distribution nationally with age-sex distributions in some African countries, suggest that the age-sex distributions from Census 2022 were of better quality (i.e., less erratic, and less pronounced digit preference) than observed in some other African countries' censuses.

The five-year age population pyramid from Census 2022 10% sample exhibits a bee-hive shape that is characteristic of populations experiencing accelerated ageing of the population resulting from consistent sharp decline in fertility and consistent increase in life expectancies at any age. The shape is however inconsistent with the shape of the five-year population pyramid derived from the 2011 census and more importantly, inconsistent with the five-year population pyramid derived from the 2021 General Household Survey. Since the 2021 General Household Survey was undertaken about a year prior to Census 2022, one would have expected close similarities between the population pyramids from Census 2022 10% sample and 2021 General Household Survey.

The smoothing of the five-year age-sex distributions through demographic modelling appeared to suggest deficit of males and females in some age groups relative to what the demographic modelling suggests. The weights in the Census 2022 10% sample imposed a different age structure on the original age structure. Age errors are a common feature in any census or survey globally though in varying degrees. Users of Census 2022 10% sample therefore need to be sensitive to the issues highlighted

above especially the weights since age-sex distributions are linked to all aspects of the life cycle and impact planning in all sectors of the economy. A major limitation of age and sex structure analysis is that it is not possible to derive separate numerical estimates of the magnitude of coverage and content errors based on such analysis alone.

ACKNOWLEDGEMENT AND DISCLAIMER

It was a privilege that I was invited to evaluate the 2022 census data before their release to the public. I thank Statistics South Africa as well as South Africa's Statistics Council for that. The current study was not based on the full 2022 census data that I performed for Statistics South Africa and the Statistics Council but was based on the 10% sample of the data using the same methodologies that I employed in the evaluating the full data set. I wish to thank Dr Anne Khasakhala (Private Consultant, Nairobi) for her review comments on this study. The views expressed in this paper are, however, those of this author.

REFERENCES

Brass W. (1971). Methods for estimating fertility and mortality from limited and defective data. The University of North Carolina at Chapel Hill, 50-84, 1971.

Brass W. (1984). The use of stable populations in age-smoothing. Personal discussion with late Brass.

Carrier N. and Hobcraft J (1971). Demographic estimation for developing countries. Population Investigation Committee, London.

Dorrington R. (1999). To count or to model that is not the question: some possible deficiencies with the 1996 census results. Paper presented at the Arminel Roundtable workshop on the 1996 South Africa Census, Hogsback, 9-11 April.

Newell C. (1997). Methods and models in demography. John Wiley & Sons Ltd: Chichester.

Sadie JL (1999). The missing millions. Paper presented to NEDLAC meeting. Johannesburg.

Shell R (1999). An investigation into the reported sex composition of the South African population according to the census of 1996. Paper presented at the Workshop on Phase 2 of Census 1996 review, Wanderers Club, Johannesburg, 3-4 December.

Phillips HE, Anderson BA, Tsebe P (1999). Sex ratios in South African census data 1970 – 1996. Paper presented at the Workshop on Phase 2 of Census 1996 Review on Behalf of the Statistical Council, Johannesburg 3-4 December.

Udjo E. O. (2005a). An examination of recent census and survey data on mortality within the context of HIV/AIDS. In, Zuberi, Sibanda, and Udjo (eds). The demography of South Africa. New York: M. E. Sharp, 90-119.

Udjo E. O. (2005b). An evaluation of age-sex distributions of South Africa's population within the context of HIV/AIDS. Development Southern Africa, vol. 22, no. 3, pp 319-345.

Udjo E. O. (2014). Estimating demographic parameters from the 2011 South Africa population census. African Population Studies: Supplement on Population Issues in South Africa, 28(1): 564-578.