Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

Assessing Digit Preference and Age Heaping in Nepal: Evidence from the Census 2021

Sudesh Pandit
Tribhuvan University
Prithvi Narayan Campus, Pokhara, Nepal
sudesh.pandit@prnc.tu.edu.np

Om Chandra Thasineku*
Tribhuvan University
Research Centre for Educational Innovation and Development (CERID), Kathmandu, Nepal
om.thasineku@cerid.tu.edu.np

Ramesh Adhikari
Tribhuvan University
Mahendra Ratna Campus, Kathmandu, Nepal
rameshipsr@gmail.com

Sushil Sharma
Tribhuvan University
Prithvi Narayan Campus, Pokhara, Nepal
sushil.bhattarai@prnc.tu.edu.np

Pitambar Paudel
Tribhuvan University
Prithvi Narayan Campus, Pokhara, Nepal
pitambar.paudel@prnc.tu.edu.np

Corresponding Author*
Received: November 02, 2025; Revised & Accepted: December 28, 2025

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

Abstract

Background: Accurate information regarding age is necessary for demographic studies, policy formulation, and planning for development. Age data in the census suffers from digit preference and age heaping. This study aims to assess the extent and patterns of digit preference and age heaping in Nepal.

Methods: This study used secondary data from the National Population and Housing Census 2021 by descriptive analytical approach. Whipple's, Myer's Blended Index and UN Age-Sex Accuracy Index were used to evaluate the quality of age data. These indices can be measured to examine the degree of age heaping in census data.

Results: Age-heaping analysis indicates a moderate level of data quality concern. Whipple's index reveals substantial digit preference for ages ending in 0 and 5 by the value 149, while Myer's Blended index recorded the deviation score of 9.89 across terminal digits. Consistently, the UN age-sex accuracy score of 21.90 classifies the data as moderately accurate. Together these findings suggest persistent inaccuracies in self- or proxy reported age data.

Conclusion: Despite methodological and technological advancements in the 2021 census, age misreporting persists due to limited literacy and numeracy, and administrative thresholds. That may influence both respondents and enumerators for digit preference on 0 and 5. Enhanced enumerator training, and developing a proving questionnaire for the census are essential to improve the accuracy and reliability of age data.

Novelty: This paper provides a national-level assessment of age data quality using the 2021 census, providing information for future assessments beyond Nepal's demographic data.

Keywords: accuracy index, age heaping, census, digit preference, Myer's Blended index, Whipple's index

Introduction

Precise and reliable age data constitute the foundation of demographic analysis, providing the basis for essential estimation of mortality, fertility, population structure, and human capital assessments. Yet, in many developing countries, including Nepal, age is misreported due to tendency towards digit preference. Digit preference describes the systematic bias in reporting ages ending in specific digits, most often 0 and 5, resulting in age heaping. Such a phenomenon destructs census and survey data, misrepresents the actual age structure, and compromises the reliability of demographic estimates used for policy making, planning, research, and governance (Preston et al., 2003; Pullum, 2005; Shryock et al., 1976) .

Age heaping emerges from an interplay of cognitive, cultural, and institutional factors. Limited literacy and numeracy, absence of official birth documentation, the symbolic salience of certain digits, and the routinized use of round numbers in administrative and legal frameworks contribute to the problem (Frank B. Hobbs, 2004; Robinson, 2013; Shryock et al., 1976). Both respondents and enumerators may be responsible for these biases; respondents may do so because of intentional misreporting or real confusion, while enumerators may do so because of

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

time restrictions or default rounding procedures, particularly in rural or marginalized areas. (Mason & Cope, 1987; Niyonsenga et al., 2008).

Nepal has conducted decennial censuses for over a century, providing a continuous record of population enumeration since 1911 (Central Bureau of Statistics, 2021). Compared to other South Asian countries, Nepal has historically displayed moderate levels of age heaping, as indicated by Whipple's Index and Myers' Blended Index (Singh et al., 2022). Although the overall data quality ranks of Nepal relatively higher than the South Asian countries Bangladesh, Afghanistan, and India. (Pardeshi, 2010; Singh et al., 2022),

Age is a central variable for demographic analysis, underpinning population projections, fertility and mortality estimates, and the formulation of age-targeted policies (Basannar et al., 2022; Preston et al., 2003). Although age data across global censuses frequently exhibit patterns of digit preference and age heaping, and census data often show patterns of digit preference and age heaping, and Nepal is not an exception to this pervasive reporting bias (Central Bureau of Statistics, 2014; West et al., 2005). These falsifications often arise from a combination of the education level of respondents, cultural norms, religious values, cognitive biases, interviewer performance and systemic enumeration challenges (Mason & Cope, 1987; Pullum, 2005). In some societies, symbolic meanings associated with particular digit preference influence age reporting, while in others, logistical limitations during fieldwork may lead interviewers to round ages to save time and effort (Niyonsenga et al., 2008; Robinson, 2013).

The literature shows that there is a significant preference for terminal digits 0 and 5, leading to systematic deviations from real age distribution (Pullum, 2005; Singh et al., 2022). These errors could bias age data and hamper the quality of demographic parameters, especially for ages in studies involving education status, labour force analysis, and health planning (Pardeshi, 2010; Preston et al., 2003). It is therefore, important to have a mechanism that can detect and quantify falsifications for the purpose of improving data quality and methodological rigor in demographic research. The error in age reporting can be precisely assessed by observing the census records.

This study aims to assess the extent and patterns of digit preference and age heaping in Nepal using single-year age data from the 2021 National Population Census. It employs established demographic indices- Whipple's Index, Myers' Blended Index, and the United Nations Age-Sex Accuracy Index to quantify the accuracy of age reporting and assess the quality of data. The findings generate empirical evidence on the magnitude of age misreporting, insights and policy recommendations for improving census and statistical precision in Nepal.

Methods

Research Design

This study adopted a quantitative descriptive research design to assess the quality of age data in Nepal, based on the 2021 population and housing census, specifically, focusing on the detection of digit preference and age heaping. The design was appropriate for evaluating patterns and irregularities in age data without altering the underlying dataset. Standardized

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

demographic indices - Whipple's index, Myer's Blended Index and UN Age-Sex Accuracy Index were used to compute and measure the magnitude and direction of digit preference, assess the degree of age heaping and determine the overall accuracy of age reporting in the census data.

Data Source

The dataset used in this study is extracted from the National Population and Housing Census of Nepal, conducted by the National Statistics Office. The census employed a modified de-jure enumeration method, collecting demographic and socio-economic data from individuals present in households during the enumeration period. Age information was self-reported or provided by a proxy respondent; no supporting documents were verified during data collection. The study focuses specifically on age data from individuals aged above 10 years, with detailed emphasis on the 23-62 years age group for certain index calculations. Age data can be obtained by asking about age directly or by asking for date of birth (Frank B. Hobbs, 2004). Age data are collected and tabulated in single-year intervals (1, 2, 3, 4, 5,…) of age as well as five-year intervals (0-4, 5-9, 10-14,…) of age groups or broader categories (0-14, 15-59, 60+) (Central Bureau of Statistics, 2014).

Variables

The main variable analyzed in this study is reported age in completed years. From this variable, the terminal digit of the reported age, focused on 0 and 5, and sex are created as derived variables for the purpose of index computation of age reporting quality indices.

Demographic techniques for assessing age heaping

To assess the quality of age data reporting in the Census 2021, Nepal, this study employed three well-established demographic indices: Whipple's Index, Myers' Blended Index, and the United Nations Age-Sex Accuracy Index. Each of these indices offers a complementary approach to identifying and quantifying errors in age reporting, particularly related to age heaping and digit preference. These three indices were triangulated to ensure the methodological precision to evaluate census data quality, enabling cross-validation of findings and improving reliability regarding age reporting accuracy.

Whipple's Index is a widely used measure of age heaping that evaluates the extent to which ages ending in 0 and 5 are overreported. It is calculated using the proportion of individuals aged 23-62 (Rasul et al., 2025; Shryock et al., 1976; United Nations, 1955), whose reported ages end in 0 or 5, relative to the expected number in a uniform distribution, which is generally considered to be less likely to misreport age in children and the elderly (Shryock et al., 1976, p. 122). The index value ranges from 100, indicating no preference, to 500, extreme heaping for terminal digits 0 or 5, measured by the formula as outlined by George Chandler Whipple, separately for both 5-year and 10-year age groupings (Pardeshi, 2010; Shryock et al., 1976).

$$\text{Whipple's index} = \frac{\sum (P_{25}+P_{30}+\cdots+P_{55}+P_{60})}{\frac{1}{5}\sum(P_{23}+P_{24}+P_{25}+\cdots+P_{60}+P_{61}+P_{62})} \times 100$$

Whipple's index values less than 105 indicate highly accurate, 105 to 109.9 fairly accurate, 110 to 124.9 approximate, 125 to 174.9 rough and 175 and over indicates very rough (Central Bureau of Statistics, 2014). This index is particularly effective in detecting systematic rounding

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

behavior among adults a population segment where ages are generally known but are often approximated to rounding figures due to a preference for 0 and 5 digits

Myers' Blended Index was used on age data ranging from 10-99 years in order to identify the digital preference patterns across all terminal digits (0 through 9). Using age frequencies ending in each terminal digit to weight cumulative deviations from ten percent theoretical expectation for each digit, this method combines population distributions (West et al., 2005). A score of 90 suggests strong preference on a single terminal digit, indicating significant age heaping, while a score of 0 suggests perfect accuracy and no digit preference. The blended version, which allows for flexible selection of age ranges, allows for more reliable evaluation of age-heaping tendencies across all digits than the original methods presented by Myers (1954) concept. This method is superior in this regard; the index is sensitive to variation in the underlying age distribution, Interpretation of the index is usually aided by including either a graphical or statistical assessment that demonstrates the clarity and usefulness of the data in representing age, enhancing the validity of graphical representation.

The procedure for computation comprises the following steps.

Step I. Sum of populations ending in each digit over the whole range starting with the lower limit of the range (e.g., 10, 20, 30, …. 80; 11, 21, 31, …. 81)

Step II. Ascertain sum excluding the first population combined in step 1 (e.g., 20, 30, 40, …. 80; 21, 31, 41, …. 81)

Step III. Weight the sums in steps 1 and 2 and add the results to obtain a blended population (e.g., weights 1 and 9 for 0 digit, weights 2 and 8 for 1 digit.)

Step IV. Convert distribution in step 3 into percentages

Step V. Take the deviation of each percentage in step 4 from 10.0, which is the expected value for each percentage

Step VI. A summary index of preference for all terminal digits is derived as one half of the sum of the deviations from 10.0 percent, each regardless of signs.

The resulting summary index is expressed as half the sum of absolute deviations from the expected 10 percent, offering a composite measure of digit preference. A value of zero denotes no digit preference and thus high accuracy, while a maximum score of 90 would reflect complete heaping on a single digit. This index is particularly effective in detecting detailed and multidimensional patterns of age misreporting that are not adequately captured by Whipple's index alone (Dahiru & Dikko, 2013; United Nations, 1955).

The internal consistency and reliability of age reporting across male and female populations were evaluated using the United Nations Age-Sex Accuracy Index. The Age Ratio Score for Male (ARSM), the Age Ratio Score for Female (ARSF) and Sex Ratio Score (SRS), which collectively identify the variances, are the three diagnostic components that are integrated into this composite index. The index computed using the formula UN Index = ARSM + ARSF + 3 × SRS, assigning greater weight to inconsistencies in sex ratios across all age groups.

It provides a systematic way to assess the accuracy of age-sex data, particularly for five intervals up to age 70 years, when reporting biases and enumeration errors are most likely to occur. Thus, according to United Nations standards, values below 20 denote high accuracy of

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

data score; between 20 to 40 indicate moderately inconsistency; and above 40, serious reporting error (Central Bureau of Statistics, 2014; United Nations, 1955).

A comprehensive and methodologically appropriate and multidimensional assessment of age data quality is provided by the integrated use of Whipple's index, Myers' Blended index, and the UN Age-Sex Accuracy Index. Whipple's index is well regarded for its simplicity and for being widely used across countries, measuring the concentration in ages ending in 0 and 5, which captures the underlying rounding tendency to heaping. Myers' Blended Index extends this calculation to all possible terminal digits, providing a deeper assessment of systematic reporting bias. Likewise, the UN age-sex accuracy index similarly estimates internal consistency of the age and sex distribution as an overall measure of quality for the age data.

While both Whipple's and Myers' indices are moderately affected by the underlying age structure that exists, and the UN index conveys a wider reporting bias, the examining of all three indices provides a more rounded and comprehensive study of the quality of the age data. If taken together, these indices are valuable in enhancing data quality assessment, illuminating the various segments of census data.

Results

This analysis utilizes data from the National Population and Housing Census 2021 of Nepal (National Statistics Office, 2023), which was conducted under a modified de-jure method that enumerated individuals at their place of usual residence. Data collection employed a hybrid methodology: digital enumeration through CSPro-enabled tablets in selected areas and traditional paper-based schedules elsewhere. The census mobilized 35,657 enumerators and 8,545 supervisors, recruited through a merit-based national selection process managed via the human resource management and monitoring system. Educational requirements and a bachelor's degree for supervisors. Additionally, 1,800 schoolteachers were trained as master trainers, while census officers at district and local levels were deployed from government offices. Training was provided at both central and local levels emphasizing procedural consistency and data accuracy. Nonetheless, despite these improvements in personal qualification and technological application, the training framework did not specially address the issue of age misreporting or methods to minimize the age heaping. Consequently, the continued dependence on self-reported or proxy-reported ages, often without documentary verification, likely sustained rounding tendencies and residual age reporting inaccuracies in the 2021 census dataset. (National Statistics Office, 2023).

The final cleaned dataset included 16,283,500 individuals aged 23-62, used for calculating Whipple's index, while broader age ranges were applied for Myers' and the UN Age-Sex Accuracy Index. An age histogram is presented in Figure 1, clearly showing visible spikes at ages ending in '0' and '5', suggesting systematic digit preference. The full single-year age distribution is available in Supplementary Table S1.

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

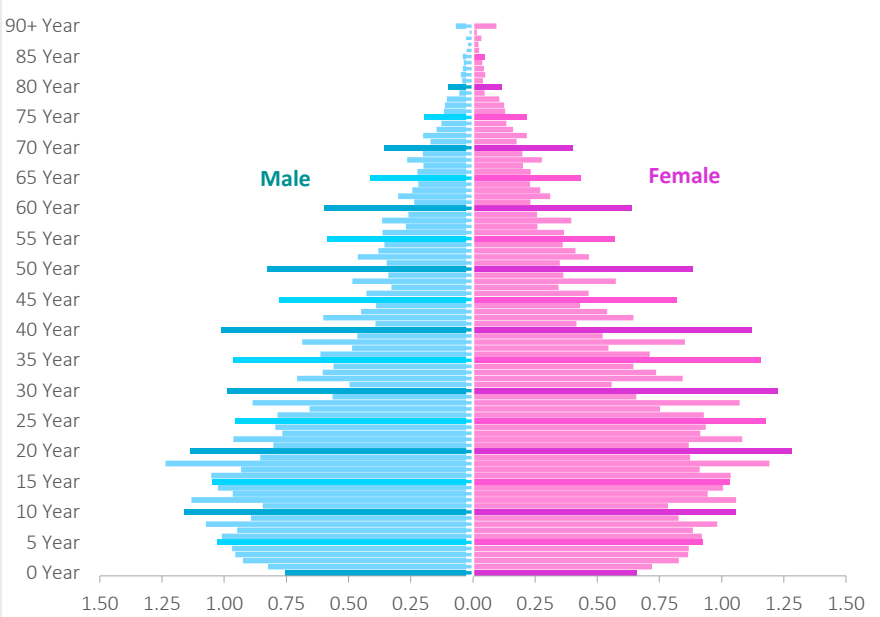

Figure 1: Single year age and sex distribution of population (% in total), 2021 Nepal census (National Statistics Office, 2024b)

The application of Whipple's index provides a valuable tool for assessing the accuracy and potential digit preference in reported age distributions within a population. In this study, the focus is on a substantial cohort comprising a total population of 16,283,500 individuals aged 23 to 62 years. The primary aim is to assess the prevalence of digit preference, specifically examining ages terminating with the digits "5" and "0" within this age range. Among the 16,283,500 individuals, 4,167,794 reported ages ending with the digit "5." Additionally, 2,125,760 individuals reported ages ending with the digit 0.

The analysis of digit preference and age heaping for age reporting in 5-year intervals revealed a Whipple's index value of 149. This index quantifies the extent of digit preference and age heaping or deviation from a uniform distribution of ages ending in 0 and 5. A lower index value suggests a lesser degree of digit preference, while a higher value indicates a more pronounced preference for certain digits.

Whipple's Index values of 149 for ages ending in '0' and '5' indicate that the age data for the 23–62-year cohort is of "rough" accuracy. The results reveal an evident tendency to digit preferences, particularly for ages ending in the digits of 0 and 5. Whipple's index is quite valuable for its methodological simplicity and focused approach to the most commonly heaped digits, making it a useful and interpretable tool for establishing systematic heaping in the middle-age ranges where age reporting is usually presumed to be accurate. But it is limited in that it does not capture younger or older cohorts and does not capture age heaping on digits other than 0 and 5. Thus, while the standard Whipple's index captures rounding behaviour over all ages, the separate disaggregated calculation of the digit 0 allows for the identification of disproportionate logging of decade-ending ages such as, ages 30 or 40, compared to ages of 25 or 35.

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

Age heaping analysis was conducted for ages 10 to 99 years, using Myer's Blended method to examine deviation from the expected uniform distribution of 10 percent for each terminal digit. Myer's method aggregates and evaluates age frequencies to determine the extent of digit preference across all digits. The analysis resulted in a summary deviation score of 9.90, which indicates there was a moderate level of digit preference present among respondents in the 2021 census of Nepal. The greatest deviations of digit preference occurred for terminal digits 0 (+4.98) and 5 (+3.15), which was consistent with their predisposition at the reported ages. The results indicate that it appears there is a systematic rounding pattern to these terminal digits, and is consistent with the findings from similar developing contexts characterized by low age awareness and limited vital registration systems. Myer's Blended Index is better because, in addition to being diagnostic, it measures deviation across the entire spectrum of digit and can recognize the subtle reporting bias that may not be sensed with Whipple's index. Myer's index is also sensitive to the irregularities in the underlying age distribution, which might expose structural population parameters along with reporting inaccuracies, and should not be interpreted earlier out of context.

The UN age-sex accuracy index shows the age ratio score for male (ARSM) is 4.00, the age ratio score for female (ARSF) is 4.88, and the sex ratio score (SRS) is 4.37. The overall composite score of the UN age-sex accuracy is 21.90 in the 2021census which, according to United Nations guidelines, is categorized as moderately inaccurately classified. This index provides assessments of age and sex distribution inconsistencies while accounting for age ratio inconsistencies for males and females as well as sex ratio inconsistencies in the distributions. However, the limitation is that it is influenced not only by digit preference but also by sex-selective underreporting or enumeration errors, thus complicating the interpretation of age biases.

Through the triangulation of these aggregate summaries, the study is able to demonstrate a convergent pattern of moderate, though persistent, age misreporting in the 2021 census of Nepal. The three indices converge and, although differ in their methodological approaches and analytical focus, are indicating on almost a unified basis substantial irregularity in age data quality. All three indices independently confirm the presence of digit preference, particularly for ages ending in '0' and '5', but with varying sensitivity and specificity. While Whipple’s Index separates the core heaping issue, Myers’ Index expands the scope to subtler digit preferences, and the UN Index contextualizes age errors within the broader demographic structure. Triangulation enhances the validity of the findings by cross-verifying results through multiple analytical lenses. Despite their methodological differences, the convergence of these indices underscores the need for ongoing improvements in enumeration practices and age data reporting in Nepal’s national statistical system.

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

Table 1: *Accuracy of Age Reporting in Nepal: 2011 and 2021 Census*

Index	2011			2021		
	Male	Female	Both	Male	Female	Both
Whipple's Index	191	186	189	149	149	149
Myer's Blended Index	15.7	15.6	15.6	9.6	10.1	9.9
UN Age-Sex Accuracy Index	-	-	23.3	-	-	21.2

(National Statistics Office, 2024b)

Table 1 provides an overview of three major demographic indices, Whipple's Index, Myers' Blended Index, and the UN Age-Sex Accuracy Index, that examine the quality of age reporting in the 2011 and 2021 censuses of Nepal. Together these indicators provide some triangulated evidence of digit preference patterns, numerical accuracy, and structural consistency of census data.

Whipple's Index, which tracks the preference of terminal digit '0' and '5', shows a marked decline from 189 (both sexes) in 2011 to 149 in 2021, reduction by Whipple's Index value 40. This decline indicates that a measurable improvement in the quality of age reporting over the intercensal period. Such improvement likely to reflects cumulative effect of structural and procedural development, including increased awareness, increased literacy, improved questionnaire design, or enhanced rigor in practices of enumeration. Although, the index value remains within the range indicative of rough, suggesting that digit preference has been declining in trend but still remain roughly age reporting.

The Myers' Blended Index also shows an important movement in accurate age reporting accuracy, by measuring heaping across all terminal digit 0 to 9, during the same period. The theoretical range of Myers' index is from 0 to 90, where 0 indicates no age heaping and 90 indicates (Central Bureau of Statistics, 2014) the extreme case where all recorded ages end in the same digit. Table 1 presents the inaccuracy index of age reporting by Myers' Blended Index, decreased from 15.6 census 2011 to 9.9 of census 2021 with minimal differences between male and female (10.1 and 9.9).

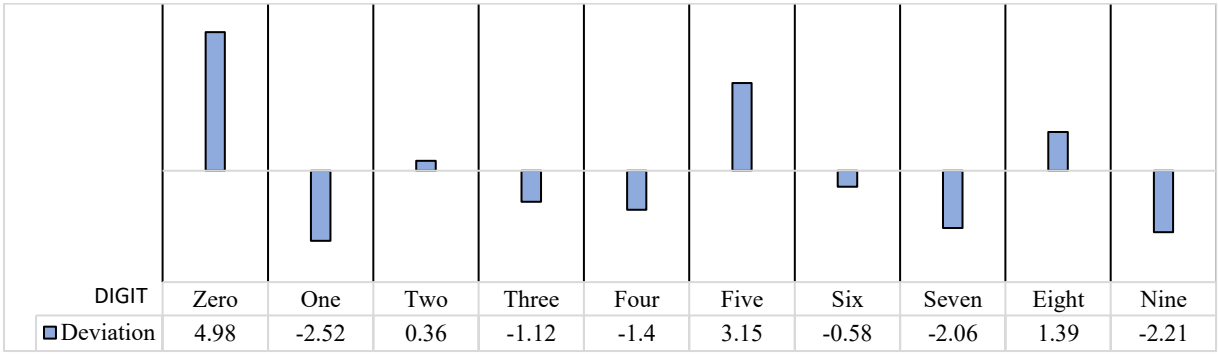

Figure 2: Myers index of digit preference for censuses 2021

Figure 2 presents a clear pattern of digit preference and avoidance in reported age by the Myer's Blended index for the census 2021. Positive deviations indicate over-reporting, while negative values denote under-reporting of specific digits. Ages ending in 0 show the strongest preference (+4.98), followed by 5 (+3.15), confirming pronounced rounding to culturally salient digits. A

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

modest preference is also observed for age 8 (+1.39) and 2 (+0.36). In contrast, ages ending in 1, 3, 4, 6, 7, and 9 are systematically under-reported, with the greatest avoidance for 1 (−2.52) and 9 (−2.21). Overall, the pattern reflects persistent digit heaping centered on 0 and 5, consistent with moderate age misreporting in the 2021 census.

The UN Age-Sex Accuracy Index, which measures consistency in the reported distribution of age and sex, slightly declined from 23.2 in 2011 to 21.2 in 2021. Although this might still be regarded as considered stable at the structural level of reporting, the small change would reflect potential structural issues associated with age-sex misalignment with assumptions such as proxy reporting or enumeration bias. Thus, while age misreporting remains an issue, significant progress has been made in reducing age heaping in Nepal in the last decade. Notable improvements in accuracy of age were evident in both Whipple's and Myer's scores, which directly reflect reported digit preference from respondents. Although the limited degree of changes in the UN index indicates that monitoring of enumeration quality, consideration of ongoing monitoring, especially as the indicators reflect significant challenges for rural and marginalized populations.

Discussion

This research studied if the reporting of age has improved improve between the 2011 and 2021 censuses of Nepal, using three common demographic indices of age accuracy - Whipple's Index, Myers' Blended Index, and the UN Age-Sex Accuracy Index. In general, we found moderate levels of age heaping and varied improvement between the 2011 census of Nepal, which suggesting a pattern of convergence. While a descriptive level examines the progress, this discussion focuses on interpreting the results, explaining the potential of socio-ethno-cultural trends, and contextualizing by models.

Although previous censuses in Nepal have been criticized for having poor quality of age data, the 2021 census data reveals that some moderate improvements have been made, although there are substantial inconsistencies. Based on Whipple's Index of 149 in a 5-year range and 152 in a 10-year range, the quality of age reporting in Nepal falls in the "rough" category (Shryock et al., 1976). Moreover, the Myers' Blended Index shows 9.9 suggesting moderate digit preference, while the UN Age-Sex Accuracy Index is 21.2 suggests moderately inaccurate age-sex data (United Nations, 1955). These represent marginal improvement over the 2011 census, but still nominal improvements over the census, but still indicate that with the age data, substantial issues remain in respect to quality to enhance precision regarding demographic estimation.

To comprehend the drivers of these trends, it is necessary to analyze the context even more deeply. Globally, age misreporting is distributed mostly among populations with low educational attainment and low numerical skill (Rosenzweig, 2021). In Nepal, many older populations and even some rural residents do not know their exact age because of inadequate civil registration records, especially before the 1990s. This is consistent with patterns observed in other countries, where reliance on memory and the absence of formal birth records have led to widespread age heaping (Moultrie et al., 2013; Spoorenberg, 2007).

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

Cultural symbolism, particularly the ritual significance of the number five in Nepalese Hindu traditions, likely contributes to persistent age heaping on digit 5. Ethnographic literature suggests that religious practices such as *Panchayan Deuta* (worship of five deities), Pancha Bali (five forms of animal sacrifice), *Panchamrit* (a sacred mixture of five ingredients) (Udit, 2023), *Panchagavya* (a combination of five cow-derived substances) (Ramachandran, 2016), which is compulsory for in any holy functions and *Pancha Kanya* (veneration of five virgin girls), all of which represent symbolic completeness in Hindu customs (Wonder Nepal, 2025), subconsciously influence age reporting patterns.

Additionally, *Gadhimai Mela*, a major Hindu festival held every five years in Bara District (Pant, 2024; Sparke, 2024), epitomizes this numerical symbolism and attracts devotees from across Hindu society, reinforcing the cultural salience of quinquennial cycles . These longstanding traditions may subconsciously influence individuals, especially in rural and less literate communities, to prefer ages ending in 5, potentially contributing to digit preference and systematic age heaping in national demographic data.

Nepal's institutional frameworks, such as the marriageable age of 20 (Government of Nepal, 2017), pension eligibility after 20 years of service (Government of Nepal, 2049), and the old-age allowance program, which provides monthly benefits to citizens aged 70 years and above (Government of Nepal, 2025), emphasize rounded age limits. Furthermore, *Dirgha Sewa Padak*, one of the long-service awards awarded after 25 years of service in institutions like Tribhuvan University (Tribhuvan University, 2050), further illustrates how formal systems institutionalize round-number 0 and 5 milestones. Other milestones formalized in the literature include five- or ten-year cycles in censuses, surveys, and national planning. Currency in different formats, in denominations that reflect a broad cognitive bias towards cognitive ease are issued in increments of 0 and 5 digits, such as Rs.5, Rs.10, and Rs.50. (Dehaene, 2011). In contexts, particularly with weak birth registration, criteria that stimulate or prescribe take attribution of 0 or 5, which refers back to round number heaping in specific preferred digits 0 and 5 in Nepal's census data. Age-based policies in law construct these hierarchical or granular age-based categories based on age; they create internal-external distinctions that, by taking age as a hegemonic axis of social organization (Sanghi, 2022), or invoked for public health and social science, which condition people to report an age on an age peg, as opposed to reporting their actual age.

The Whipple's Index indicated a marked reduction from 189 in 2011 to 149 in 2021, or approximately a 21 percent improvement in heaping, showing a significant reduction of age heaping to terminal digits '0' and '5'. The upward departure in the Myers' Blended Index was indicative of that reduction for all digits from 15.6 to 9.9. These improvements are likely attributable to several key factors: rising levels of literacy and numeracy, particularly among younger cohorts (CBS, 2021b); enhanced educational qualifications of the population; and the incorporation of digital tools and improved training for enumerators during the 2021 census (National Statistics Office, 2023).

However, the UN Age-Sex Accuracy Index remained relatively stable, declining only slightly from 23.2 to 21.2, indicating persistent inconsistencies in age-sex structures. This suggests that

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

while respondents may be increasingly capable of providing accurate ages, structural and procedural weaknesses in census administration, such as proxy reporting, and lack of probing, continue to affect the quality of demographic data (Rozelle et al., 2023).

Digit preference is defined as the tendency of response to report their age ending in certain digits, 0 or 5, more frequently than others instead of a more precise digit like one, nine or four. This bias is believed to stem from social and cultural preferences (Frank B Hobbs, 2004), prevailing social norms (Nagi et al., 1973, p. 167), or cognitive biases that affect memory around age reporting. Related to this, the Population Monograph of Nepal has also highlighted the instance of digit preference as well (Central Bureau of Statistics, 2014; Pantha & Sharma, 2003). Age heaping is an associated phenomenon of digit preference and refers to a clustering of age reporting at certain digits or age groupings. The Whipple's index is commonly used to measure age heaping, with values above a certain threshold indicating significant heaping (Dahiru & Dikko, 2013; Singh et al., 2022). In Nepal, age heaping has been observed in various censuses (Central Bureau of Statistics, 2014; Pantha & Sharma, 2003), raising concerns about the accuracy of age data.

Several factors contribute to digit preference and age heaping in Nepal. Cultural factors (Frank B Hobbs, 2004), such as the importance of certain milestone ages in societal norms and rituals (Nagi et al., 1973), may influence individuals to round their ages to culturally significant digits. Additionally, a low literacy rate, limited numeracy skills, and unavailability of calendars, especially in rural areas, can contribute to digit preference, as individuals may find it easier to remember and report ages that end in 0 or 5.

This study has assessed the quality of data of the 2021 census 2021 of Nepal, particularly in terms of age reporting. From the analyses of Nepal census data of 2021, the quality of age data is poor with respect to Whipple's index, Myers' index, and the UN age-sex accuracy index or joint ratio. Whipple's index excludes early childhood and old age from the formula due to their susceptibility to different errors and issues, distinct from digit preference (Pardeshi, 2010). The quality of age data in the 2021 census, Whipple's index for Nepal is between 125 to 174.9; that means data is rough (Shryock et al., 1976). The data indicate the concentration on terminal digits 0 and 5 and digit avoidance for terminal digits 1 and 9. As it has been argued that misreporting of age is most common in areas with the least developed countries (Frank B Hobbs, 2004; Seng, 1959). Although there was noticeable improvement in the Whipple's index, Myer's index and UN age-sex accuracy index between the censuses, decreasing from 189, 15.6 and 23.2, respectively, in the 2011 (Central Bureau of Statistics, 2014) census to 148.94, 9.89 and 21.90 in the 2021 census.

This observation may be ascribed to various factors, including the literacy level of respondents, the effectiveness of conscientious enumerators, and the availability of calendars for reference. The literacy rate of Nepal has increased from 65.9 percent in 2011 (Central Bureau of Statistics, 2012) to 76.2 percent in 2021 (CBS, 2021b). The observed digit preferences may be influenced by cultural practices or beliefs associated with specific numbers (Frank B Hobbs, 2004). As literacy is increasing in Nepal, it is seen that the inaccuracy of age reporting is also decreasing (Central Bureau of Statistics, 2014). Individuals lacking educational background exhibited a

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

higher incidence of age misreporting compared to those with some level of education (Singh et al., 2022). In addition, public awareness programs related to census, publicity, training of enumerators, and use of technology in census may have helped to reduce the misreporting of age in Nepal (CBS, 2021a). Even though the overall pattern of age enumeration in Nepal is inconsistent.

The effect of digit preference and age heaping has important implications for demographic research, policy formulation, and program planning. Age data that is biased will produce inaccurate estimates of population structure, cause errors in age-specific mortality rates, and lead to erroneous estimates of other core demographic measures, all of which influence the success of public health and education planning, as well as social welfare programs. Addressing digit preference and age heaping in Nepal should include targeted interventions like improving educational and awareness campaigns. Encouraging communities to provide actual ages in identification documentation in addition to training census enumerators to more accurately collect age data could also help reduce these biases as well. Accurate age reporting in census and survey data collection depends not only on direct questioning but also on the systematic use of probe-type questions to enhance reliability and validity. Beyond asking respondents their current age, trained enumerators are encouraged to elicit supplementary information such as month and year of birth or the timing of life event associated with birth, which enables cross-verification of reported age data (Pullum & Staveteig, 2017). These additional probes qualify the stated age, improved accuracy for this age variable, reduce age misreporting, and minimize age heaping and digit preference. Besides, high-level statistical techniques to adjust for population biases and age groupings in demographic analysis are essential to provide more reliable population estimates.

Digit preference and heaping of ages are serious challenges that need to be accounted for when evaluating census data from Nepal. Investigating the issues of the factors creating bias will be critical to developing effective approaches to improve the quality of age reporting and, therefore, the reliability of demographic information, ensuring its utility for policymaking and strategic planning purposes.

Conclusion

While there have been improvements since 2011, age reporting is still a major challenge in the 2021census Nepal, reflecting persistent structural and sociocultural constraints. Limited literacy, cultural beliefs of numeracy, legal and administrative thresholds, and institutional reinforcement of rounded numbers all continue to propagate digit preference and age heaping. All of these biases distort demographic estimates, undermining policy, planning, and research. Additionally, improved literacy in age reporting, conducting counting check techniques, establishing rigorous training procedures, and conducting educational campaigns would contribute to better data quality. All of these issues are essential to generating reliable age data that is foundational to evidence-based decision-making in health, education, and social welfare program in Nepal. To improve data quality, the findings underscore the need for strengthened census enumeration practices, particularly enumerator training and the systematic use of probe-

type questions, such as calendar year of birth and association with life events, to validate reported ages and reduce misreporting.

Funding: There is no funding received.
Conflicts of Interest: The authors have no conflict of interests to this article.
Data availability statement
Data are available in a public, open access repository. All data relevant to the study are included in the article or uploaded as supplementary information. Data is available in https://censusnepal.cbs.gov.np/results/files/national/Formated_NR_Indv_Table03_SingleYearAge_edited_17.xlsx

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

References

Basannar, D., Singh, S., Yadav, J., & Yadav, A. K. (2022). Quantifying age heaping and age misreporting in a multicentric survey. *Indian Journal of Community Medicine 47*(1), 104. https://doi.org/10.4103/ijcm.ijcm_1179_21

CBS. (2021a). *Census Advertise Videos* Thapathali, Kathmandu, Nepal, Central Bureau of statistics. https://censusnepal.cbs.gov.np/Home/Details?tpid=10&tfsid=85

CBS. (2021b). *Preliminary Report of Census 2021*. CBS.

Central Bureau of Statistics. (2012). *National Population and Housing Census 2011 (National Report)*

Central Bureau of Statistics. (2014). *Population Monograph of Nepal*. Central Bureau of Statistics.

Central Bureau of Statistics. (2021). *Hundred Ten Years of Census (1968-2078)*. Central Bureau of Statistics, Government of Nepal.

Dahiru, T., & Dikko, H. G. (2013). Digit preference in Nigerian censuses data of 1991 and 2006. *Epidemiology, Biostatistics Public Health*, *10*(2). https://doi.org/10.2427/8843

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. OUP USA. https://tinyurl.com/5b2d2w57

The National Civil (Code) Act, 2017 (2074), (2017).

Government of Nepal. (2025). *Budget statement for the fiscal year 2082/83*. Kathmandu, Nepal: Government of Nepal, Retrieved from https://giwmscdntwo.gov.np/media/pdf_upload/Webside%20Final%20budget%20Speech%202_15_qbn3jyt.pdf

Civil service act 2049, (2049).

Hobbs, F. B. (2004). Age and sex composition. In J. S. Siegel & D. A. Swanson (Eds.), *The Methods and Materials of Demography, 2nd Edition* (pp. 125-173). Elsevier Academic Press. https://www.researchgate.net/publication/304822739

Hobbs, F. B. (2004). Age and Sex Composition. In D. Swanson & J. S. Siegel (Eds.), *The methods and materials of demography* (Second ed., pp. 125-174). Elsevier Academic Press California, USA.

Mason, K. O., & Cope, L. G. (1987). Sources of age and date-of-birth misreporting in the 1900 US census. *Demography*, *24*(4), 563-573. https://doi.org/10.2307/2061392

Moultrie, T. A., Dorrington, R. E., Hill, A. G., Hill, K., Timæus, I. M., & Zaba, B. (2013). *Tools for demographic estimation*. International Union for the Scientific Study of Population.

Myers, R. J. (1954). Accuracy of age reporting in the 1950 United States census. *Journal of the American Statistical Association*, *49*(268), 826-831. https://doi.org/10.1080/01621459.1954.10501237

Nagi, M. H., Stockwell, E. G., & Snavley, L. (1973). Digit preference and avoidance in the age statistics of some recent African censuses: Some patterns and correlates. *International Statistical Review/Revue Internationale de Statistique 41*(2), 165-174. https://doi.org/10.2307/1402833

National Statistics Office. (2023). *National Population and Housing Census 2021 (National Report)*. N. S. Office. file:///C:/Users/Acer/Downloads/National%20Report_English%20(1).pdf

National Statistics Office. (2024a). *National Population and Housing Census 2021* National Statistics Office. https://censusnepal.cbs.gov.np/results/files/national/Formated_NR_Indv_Table03_SingleYearAge_edited_17.xlsx

National Statistics Office. (2024b). *National Population and Housing Census 2021, Population Composition of Nepal* (Thematic report-1, Issue.

https://censusnepal.cbs.gov.np/results/files/result-folder/Final_Population_compostion_12_2.pdf

Niyonsenga, T., Vanasse, A., Courteau, J., & Cloutier, L. (2008). Impact of terminal digit preference by family physicians and sphygmomanometer calibration errors on blood pressure value: implication for hypertension screening. *The Journal of Clinical Hypertension*, *10*(5), 341-347. https://doi.org/10.1111/j.1751-7176.2008.06620.x

Pant, B. B. (2024, fri, 6December 2024). Gadhimai Festival: A Celebration Of Devotion. *The Rising Nepal*. https://risingnepaldaily.com/news/53230?utm_source=chatgpt.com

Pantha, R., & Sharma, B. R. (2003). POPULATION SIZE, GROWTH AND DISTRIBUTION. In *Population monograph of Nepal* (Vol. 1). Government of Nepal, National Planning Commission Secretariat, Central Bureau of Statistics. https://censusnepal.cbs.gov.np/Home/Details?tpid=5&dcid=5db0e1c6-5c69-4183-88cc-6f18ac5ec838&tfsid=26

Pardeshi, G. S. (2010). Age heaping and accuracy of age data collected during a community survey in the Yavatmal district, Maharashtra. *Indian Journal of Community Medicine*, *35*(3), 391. https://doi.org/10.4103/0970-0218.69256

Preston, S. H., Elo, I. T., Hill, M. E., & Rosenwaike, I. (2003). Understanding the Sources of Age Misreporting. In *The Demography of African Americans* (pp. 63-73). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-0325-3

Pullum, T. W. (2005). A statistical reformulation of demographic methods to assess the quality of age and date reporting, with application to the Demographic and Health Surveys. Annual Meeting of the Population Association of America, Philadelphia,

Pullum, T. W., & Staveteig, S. (2017). *An assessment of the quality and consistency of age and date reporting in DHS surveys, 2000-2015* (DHS Methodological Reports 19, Issue. ICF, T. D. Program, & A. Health.

Rasul, A., Nasir, J. A., Jasilionis, D., & Jdanov, D. A. (2025). Quality of reported ages: A robust re-modification in Total Modified Whipple's Index. *PLoS One*, *20*(8), e0330519. https://doi.org/10.1371/journal.pone.0330519

Robinson, O. C. (2013). Values and adult age: Findings from two cohorts of the European Social Survey. *European journal of ageing*, *10*(1), 11-23. https://doi.org/10.1007/s10433-012-0247-3

Rosenzweig, S. C. (2021). Age is measured with systematic measurement error in developing country surveys: A diagnosis and analysis of consequences. *Research Politics*, *8*(3), 20531680211044068. https://doi.org/10.1177/20531680211044068

Rozelle, J. W., Meyer, M. J., McKenna, A. H., Obaje, H., & Kraemer, J. D. (2023). The effect of interviewer-respondent age difference on the reporting of sexual activity in the Demographic and Health Surveys: Analysis of data from 21 countries. *Journal of Global Health*, *13*, 04002. https://doi.org/10.7189/jogh.13.04002

Sanghi, S. (2022). Age as a site of law's meaning-making practices. *Jindal Global Law Review*, *13*(2), 221-229. https://doi.org/0.1007/s41020-022-00185-6

Seng, Y. P. (1959). Errors in age reporting in statistically underdeveloped countries. *Population Studies*, *13*(2), 164-182. https://doi.org/10.2307/2172409

Shryock, H. S., Siegel, J. S., & Stockwell, E. G. (1976). *The Methods and Materials of Demography* (Nachdr ed.). New York : Academic Press.

Singh, M., Kashyap, G. C., & Bango, M. (2022). Age heaping among individuals in selected South Asian countries: evidence from Demographic and Health Surveys. *Journal of Biosocial Science*, *54*(4), 725-734. https://doi.org/10.1017/S0021932021000249

Sparke, T. (2024, 11 December 2024). Tradition and Controversy at Nepal's Gadhimai Festival With Water Buffalo, Rats, and Pigeons Sacri. *Focusing on Wildlife*.

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

https://www.onegreenplanet.org/animals/tradition-and-controversy-at-nepals-gadhimai-festival-with-water-buffalo

Spoorenberg, T. (2007). Quality of age reporting: extension and application of the modified Whipple's index. *Population and Development Review*, *62*(4), 729-741. https://doi.org/10.3917/popu.704.0847

Tribhuvan University teacher staff service rule 2050, (2050).

Udit. (2023). *Panchamrit: Ingredients Used, History & Benefits of Panchamrit in Detail*. Darshan Enlighten your soul. Retrieved 6-1-2025 from file:///C:/Users/Acer/Desktop/Panchamrit%20Ingredients%20Used,%20History%20&%20Benefits%20of%20Panchamrit%20in%20Detail%20-%20Darashan.html

United Nations. (1955). *Methods of appraisal of quality of basic data for population estimates*. New York: United Nations.

West, K. K., Robinson, J. G., & Bentley, M. (2005). Did proxy respondents cause age heaping in the census 2000. *ASA Section on Survey Research Methods*, 3658-3665.

Wonder Nepal. (2025). *Panchakanya: The ritual offering of five virgen girls in Nepali tratition*. Wonder Nepal. Retrieved 6-1-2025 from https://thewondernepal.com/articles/panchakanya-the-ritual-offering-of-five-virgin-girls-in-nepali-tradition/?utm_source=chatgpt.com

Nepal Journal of Multidisciplinary Research (NJMR)
Vol. 9, No. 1, March 2026. Pages: 1-18
ISSN: 2645-8470 (Print), ISSN: 2705-4691 (Online)
DOI: 10.3126/njmr.v9i1.86274

Supplementary Table 1 *Population by single age group of Nepal, 2021*

Age	Population	34 Year	351,877	69 Year	116,571
0 Year	411,159	35 Year	618,195	70 Year	220,685
1 Year	450,469	36 Year	385,964	71 Year	101,019
2 Year	510,855	37 Year	300,741	72 Year	121,570
3 Year	531,295	38 Year	448,956	73 Year	89,967
4 Year	535,505	39 Year	287,636	74 Year	76,129
5 Year	568,365	40 Year	621,284	75 Year	120,603
6 Year	562,745	41 Year	235,550	76 Year	71,787
7 Year	534,278	42 Year	363,738	77 Year	69,471
8 Year	599,728	43 Year	288,324	78 Year	61,504
9 Year	501,311	44 Year	238,936	79 Year	29,838
10 Year	646,772	45 Year	465,773	80 Year	62,234
11 Year	475,016	46 Year	260,568	81 Year	24,549
12 Year	638,817	47 Year	195,764	82 Year	28,369
13 Year	556,890	48 Year	308,758	83 Year	24,653
14 Year	592,370	49 Year	205,177	84 Year	21,751
15 Year	607,128	50 Year	499,602	85 Year	25,043
16 Year	608,979	51 Year	202,968	86 Year	14,508
17 Year	537,901	52 Year	270,951	87 Year	12,336
18 Year	708,386	53 Year	231,221	88 Year	17,925
19 Year	504,010	54 Year	209,110	89 Year	8,515
20 Year	704,630	55 Year	336,733	90 Year	13,580
21 Year	487,099	56 Year	212,738	91 Year	4,716
22 Year	596,573	57 Year	154,189	92 Year	4,958
23 Year	490,146	58 Year	221,728	93 Year	4,119
24 Year	504,612	59 Year	150,556	94 Year	3,033
25 Year	621,333	60 Year	359,659	95 Year	4,361
26 Year	499,807	61 Year	136,482	96 Year	2,249
27 Year	410,888	62 Year	178,068	97 Year	1,536
28 Year	571,114	63 Year	150,089	98 Year	2,580
29 Year	356,207	64 Year	131,306	99 Year	1,579
30 Year	645,215	65 Year	246,944	≥100 Year	4,929
31 Year	307,222	66 Year	133,073	**Total**	**29,164,578**
32 Year	452,204	67 Year	116,934		
33 Year	391,194	68 Year	158,096		

(National Statistics Office, 2024a)