Age Reporting Behaviour : A Case Study of 1991 and 2000 Population and Housing Censuses, Malaysia

Abd. Latib Talib, Mohd Sofi Ali, Marlini Sahul Hamid
and Khamsiah Mohd Zin[1]

Abstract

This paper examines the quality of age data reporting in the two Malaysian Population and Housing Censuses, 1991 and 2000. Various combinations of methods were used to analyse single ages of population at various levels of aggregation: total, stratum, gender and ethnic groups. The pyramid chart, Whipple-type Index and test differences of terminal digits of single age seem to suggest that misstatements in age reporting do exist due to digit preferred and digit avoided in both censuses.

Introduction

The quality of Malaysia Population and Housing Censuses data was evaluated through Post Enumeration Survey (PES) or Census Coverage Evaluation Survey (CCES). The purpose of the PES is to determine the extent of under-enumeration or coverage error, to evaluate the content error as well as benchmark for the planning and implementation of future censuses or surveys. Generally, the rate of under-enumeration of population is relatively low at 4.4 percent in 1991 and 4.6 percent in 2000 which is acceptable under international standard. However, the content error of the surveys shows that the level of inconsistencies and unmatched responses of age variable is relatively higher than the other two demographic variables, gender and ethnic groups (Department of Statistics, Malaysia, 1995 and Department of Statistics, Malaysia, 2003)[2].

Age analysis study for Peninsular Malaysia was initiated by Saw Swee Hock (1967). The study using 1957 Census data found that the quality of age data for Malays was relatively poor compared to that of Indians and Chinese. As the PES 1991 or CCES 2000 only show the level of inconsistency or unmatched in age reporting, this paper further examines the quality of data collected by analysing single ages population data in both censuses. Various combinations of methods were adopted and thoroughly examined to analyse the data on population age by stratum, ethnic group and gender.

[1] Abd. Latib Talib is currently the Deputy Director of Industrial Production and Construction Statistics Division and Mohd Sofi Ali, Marlini Sahul Hamid and Khamsiah Mohd Zin are Assistant Directors of Terengganu State Office.

[2] Coverage error refers to living quarters, households or population that were omitted, erroneously included or duplicated during census enumeration. Content error on the other hand, refers to the error due to the difference in responses on specific characteristics of matched persons obtained from the census and CCES.

© 2007 Department of Statistics Malaysia

Irregularities in age reporting

Most researchers agreed that the age data compiled by national population censuses may have some irregularities in age reporting. These errors are quite common in many developing countries as compared to developed countries. Irregularities that refer to digit preference and digit avoidance will normally distort the age population distribution. In the absence of irregularities in age reporting, the count of adjacent ages should be virtually smooth.

Mason and Cope (1987) concluded that there are four sources that could be attributed to age misreporting in any censuses or surveys. Firstly, ignorance of actual ages among respondents; secondly, miscommunication between interviewers and informants; thirdly, the distortion of age to meet preconceptions or social norms about the relationship of age to other social characteristics or events; and finally, errors in recording or processing.

Ueda (1980) stated that one of the major types of errors most commonly found in the sex age data derived from censuses or similar surveys is the false reporting of age. In many cases, the erroneous reporting of age is attributable to the ignorance of respondent. In most cases, ages are being reported on some particular digits, “0” and “5”.

Mukherjee and Mukhopadhyay (1988) in their study using Turkish Census data found that age heaping occurs in terminal digit “0” and “5”. Kabir and Chowdhury (1981) in their analysis for census data of Bangladesh found errors in age reporting due to digit preference and there were strong tendency to report age ending with “0” and “5”, with subsidiary heaping at ages ending with “8” and “2” respectively.

Some researchers found that higher tendency of age heaping occurs in the older age category of population. Hill, Preston and et. al. (1997) noted that the age misreporting remains substantially high for older African American. Nagi, Stockwell and Snavley (1973) revealed that age heaping is a major source of inaccuracy in the age statistics in many of the developing nations on the African continent, particularly among Islamic populations. This phenomenon was found to be more pronounced among women than men, and it tends to increase with age. However, Bairagi, Aziz, et.al. (1982) pointed that misreporting also occurs in the early age of population especially in the rural area. The misstatement for young children in rural Bangladesh increase monotonically with age and systematic errors in age misstatement display modest overstatement for the first years of life and more pronounced understatement for ages 4, 5 and 6.

Aimee and Samuel (1991) concluded that misreporting is most severe at older age. They found evidence of very pervasive overstatement of age at advanced ages. The evidence of increasing age misstatement with old age is consistent with the observation that literacy rates also decline with age, since age misstatement is associated with literacy and low educational attainment.

Tradition and beliefs can also be part of age misstatement especially in the method of age-reckoning for traditional Chinese or Muslim community. You Poh Seng (1959), Saw Swee Hock (1967) pointed that the Chinese traditional method of age-reckoning differs from the international method in a systematic manner. On the day a child is born he or she is already considered as one year old. If a child was born one week before Chinese New Year, after two weeks the child will be two years old.

Most countries in Asia like Korea and China, there is sometimes preference for age ending in “3” because the numeral 3 sounds like the word or character for life. However, they would avoid the number 4 because it has the same sound as the word or character meaning death.

Age at first marriage especially in the developing countries also contributes to age misreporting. In some developing countries marriage at very young ages still exist. Indonesia is one example of a country characterised by relatively young age at marriage for females (Savitridina, 1997). Interviewers have some motivation to shift the ages of women who are within the boundaries of the 15 to 59 interval to be below or above the minimum age of respondent eligibility. There may also be some shifting of birth to be outside the maximum age of eligibility for the health questions (Pullum, 2005).

Generally researchers agreed that the irregularities in age reporting for censuses or surveys do exist. Unfortunately, little is known why age heaping or age misreporting occurs. Roberts and Brewer (2001) pointed out that although heaping represents a common type of measurement error, it was apparently due to no prior general measures being applied.

Age reporting error in Malaysian Censuses

Study of age reporting in Malaysia was conducted by Saw Swee Hock using the 1957 Malaya census data. Whipple index shows that the degree of heaping is by far the most severe among the Malay, slightly less among the Indians and very much less among the Chinese. Under United Nations grading (United Nations, 1990), the age accuracy of Malay may be classified as “rough” for males and “very rough” for females; that of the Chinese as “fairly accurate” for males and “highly accurate” for females; and that of Indians as “rough” for both males and females (Saw Swee Hock, 1967).

Myers index that measures the degree of preference or dislike of terminal digits tracked similar pattern of age misreporting seems to exist among the Malay and Indians. There is no clear preference between even or odd numbers, but “0”, “5”, and “8” are preferred and other terminal digits are disliked (Saw Swee Hock, 1967).

In the CCES 2000, it was reported that the consistency of coverage for ages recorded 87.83 percent as compared to the consistency for ethnic group and gender of 98.11 percent and 99.32 percent respectively[3]. This is consistent with the observation by Shyrock and Siegel (1971) that the quality of data collected in a survey or census, if examined more intensively, is due more to errors in age reporting rather than any other data items. Thus, the need to explore the details of behavioural aspects of Malaysian in age reporting is very essential.

[3] Malaysia, Census Coverage Evaluation Survey 2000 Report, Department of Statistics, Putrajaya, 2003.

The methodology and data sources

The methodology

There are several methods of evaluating age reporting errors. Whipple index, Myers blended method, Bachi index, Carrier index and Ramachandran index are examples of methods used in evaluating goodness of fit of census or surveys data. Whipple index is a classic measurement for evaluating the age heaping. (Zhenglian, Yi, Jeune & Vaupel, 1998).

Other than indices of evaluating age data, pyramid chart is one of the most elegant ways of presenting age and sex distribution data graphically. Irregularities can be easily detected using this method by observing the age distribution. Spike in bar chart for any terminal digits in relation to the adjacent ages will show that there is digit preferences. In contrast, if bar chart suddenly dropped down for any terminal digits relatively to the adjacent ages, that would represent the digit avoidance. However, this method is incapable to prove any irregularities in age distribution statistically.

This paper combines several methodologies in the analysis of population at single age using whipple-type index and pyramid chart with additional digits differential test.

a. Whipple-type Index

One of the most popular age-heaping analyses is Whipple's Method (WM), an index designed to reflect preference for the terminal digits of "0" and "5", usually in the age range 23-62 denoted as,

$$WM = \frac{\Sigma\left(P25 + P30 + P35 + P40 + P45 + P50 + P55 + P60\right)}{\frac{1}{5}\Sigma\left(P23 + P24 + P25...P60 + P61 + P62\right)} * 100 \qquad (1)$$

WM varies from 0 to 500. WM is 0 when the digits "0" and "5" are not reported in the census data, 100 is when there is no preference for "0" or "5" reported in the census data and maximum of 500 when only the digits "0" and "5" are reported in the census data.

The judgment in the age distribution data as noted by the United Nations (1990) are as follows;

- i) "highly accurate" if the values of WM are less than 105
- ii) "fairly accurate" if the WM values are between 105 and 109.9
- iii) "approximate" if WM values are between 110 and 124.9
- iv) "rough" if WM values are between 125 and 174.9
- v) "very rough" if WM values are 175 or more

Although Whipple index measures only the effects of preferences for ages ending in 5 and 0, it can be assumed that such digit preference is usually connected with other sources of inaccuracy in age statements and the indices can be accepted as a fair measure of the general reliability of the age distribution (United Nations, 1990).

The Whipple index which calculated age preference for ages 23 to 62 is an arbitrary one. The ages of early childhood and old age are excluded because they are more frequently influenced by other types of errors and issues than digit preference (Shyrock and Siegel, 1971).

The "Whipple-type Index" developed by Poston and his students (Poston, Chu et al. 2000) reflects only the degree of heaping on age 3 for the ages between 23 and 53 as follows,

$$W-3=\frac{\Sigma(P23+P33+P43+P53)}{\frac{1}{10}\Sigma(P23+P24+P25...P60+P61+P62)}*100 \qquad (2)$$

Applying the same mechanisms of calculating the index, this paper will extend the age range from 20 to 79 since it is expected that the irregularities in age distribution exists within that range in the case of Malaysia. Thus, the formula being extended as follows,

$$Wi_{t}^{j,k}=\frac{\left(P2i_{t}^{j,k}+P3i_{t}^{j,k}+P4i_{t}^{j,k}+P5i_{t}^{j,k}+P6i_{t}^{j,k}+P7i_{t}^{j,k}\right)}{\frac{1}{10}.\sum_{n=20}^{79}P_{nt}^{j,k}}*100 \qquad (3)$$

Where,

i = 0,1,2,…..,9
j = 1 and 2
k = 1,2,3
t = 1991, 2000

the i values from 0 to 9 refer to terminal digit of ages 20 to 79 years. Thus $Wi^{j,k}$ indicates the Whipple Index of terminal digit of i for j's stratum (urban and rural) and k's ethnic groups (Malay, Chinese, and Indians).

Judgement of age distribution data of each terminal digit is based on United Nations recommendation as stated earlier.

b. Statistical test of age heaping

There is no specific model to prove any terminal digit preferred or avoided statistically. However, in this paper every terminal digit of ages (from 0 to 9) will be tested with the adjacent terminal digit by comparing the means differences. For example for digit "0", will be compared together with digits "9" and "1". Sets of digit to be tested are as follows,

"0" and "9", "0" and "1", "1" and "2", "2" and "3", "3" and "4", "4" and "5", "5" and "6", "6" and "7", "7" and "8" and "8" and "9". The hypotheses of the terminal digit avoided are generally denoted as,

$H_o : \mu_{p-1} - \mu_P = 0$ and $\mu_p - \mu_{P+1} = 0$

$H_1 : \mu_{p-1} - \mu_P > 0$ and $\mu_p - \mu_{P+1} < 0$

The hypotheses of the terminal digit preferred or heaping are generally denoted as,

$H_o : \mu_{p-1} - \mu_P = 0$ and $\mu_p - \mu_{P+1} = 0$

$H_1 : \mu_{p-1} - \mu_P < 0$ and $\mu_p - \mu_{P+1} > 0$

where μ_{p-1} and μ_{p+1} are the means of adjacent digit of (ρ) and μ_P is means for digit tested. If the test failed to reject the null hypothesis, there is no digit preference or digit avoidance. Conversely, if the test rejects the null hypothesis, one can conclude that there is digit preferred or digit avoided in the ages reporting.

Hypotheses

The paper tries to examine the behaviour of Malaysians in age reporting in census. Is there any significant change in age reporting among Malaysians? The hypotheses constructed for this analysis are as follows:

a. There are misstatements in age reporting among Malaysian in both censuses;
b. The degree of age misstatements reduce as literacy rate improved;
c. Digit “0” and “5” are the most preferred compared to other terminal digit “1”, “2”, “3”, “4”, “6”, “7”, “8” and “9”;
d. Digit preference in age reporting occurs at every level of population, stratum, sex, and ethnic groups; and
e. Digit avoidance does not exist in the case of Malaysia at any level of population.

The above hypotheses were set-up with the assumptions that the enumerators were well trained for the census and information on ages were given by the respondents directly.

Findings

a. Pyramid charts and whipple’s indices

Total single age distribution

Pyramid charts of single age data distribution for 1991 and 2000 censuses at total level are shown in chart 1(a) and 1(b) in Appendix 1. Based on the charts, it seems that heaping occurred in both censuses 1991 and 2000 for terminal digits of “0” and “5”. The bar charts clearly spiked at both digits from ages 20 to 80. However, there is no clear evidence that digit avoidance exists at total level of single age distribution data in the case of Malaysia for both censuses.

Whipple’s indices at total level for both censuses shows that digits “1”, “2”, “6”, “7”, “8”, and “9” fall under category “highly accurate”. This indicates that there are no digits preferred or digits avoided for those terminal digits. On the other hand, digit “4” is "fairly accurate" and "approximate" for digits "0", "3" and "5". This shows that there is preference or avoidance for those terminal digits (Table 1).

Single age distribution by stratum

Pyramid charts of single age data distribution for 1991 and 2000 censuses at stratum level are shown in charts 2a, 2b, 2c and 2d.

Based on the charts, it seems that population were under enumerated among urban-females population between 5-19 years old and rural-males as well as rural-females population age between 20-34 years. Age heaping occurred at both urban and rural stratum for 1991 and 2000 censuses. Terminal digits of "0" and "5" are the most preferred one. The bar charts clearly spike at both digits from ages 20 to 80.

By examining Whipple index as in equation (3) for stratum, digits "0" and "5" are the most preferred for males and females on both censuses (see Table 2a and 2b). Digit "0" falls under category of "rough" for urban and rural stratum in both censuses. Digit "5", however is "fairly accurate" for urban stratum and is "approximate" for rural population. Digits "3", "4", "6", "7", "8" and "9" based on the Whipple-type Index seems to suggest that the distribution are "highly accurate" for both censuses.

Single age distribution by ethnic groups

Pyramid charts of single age data distribution for Malay for 1991 and 2000 censuses are shown in chart 3a, 3b, Chinese in charts 3c and 3d and Indians in Charts 3e and 3f respectively.

Based on the charts, it seems that the age distribution are contains more "noise" for Chinese especially for population age between 15 to 25 years, age between 10 to 25 years and 40 to 50 years for Indians and relatively well distributed for Malay. It seems that "the gaps" are contributed by under enumeration for those age groups (charts 3a, 3c and 3e).

Generally, digit "0" is considered very rough while digits "2" and "5" is fairly accurate for 1991 and 2000 censuses for all ethnic groups. Digit "1" is also fairly accurate for 1991 census and highly accurate for 2000 census. The remaining digits recorded the WM index less than 100 (Table 3a, 3b and 3c).

b. Differential tests

Terminal digits test for total population provides clear evidence of digits "0" and "5" are statistically different from its adjacent digits at 5 percent significant level. Both digits are preferred to be reported in the census as shown by the negative sign of mean differences of digits ("9" and "0", "4" and "5") and positive sign of digits ("0" and "1", "5" and "6")[4]. Digit "8" is also statistically difference and preferred by female population in the 2000 census at 5 percent significant level. Digit "4" is avoided to be reported in the 1991 census for both males and females but not in the 2000 census (Table 4a).

[4] Digit preferred or digit avoided is determined by the following example. If mean differences of digits of 4 and 5 is positive and mean differences of digit 5 and 6 is negative and statistically significance, this indicates that digit 5 is preferred. On the other hand, if mean differences of digits 4 and 5 is negative and mean differences of digit 5 and 6 is also negative at any significant level, thus indicates that digit 5 is avoided.

Analysing population age at stratum level, shows that heaping at terminal digit of “0” and “5” for urban and rural for both censuses. Given that the literacy rate of Malaysian is improving from 90 percent for urban and 80 percent for rural in 1991 to 94.3 percent and 85.4 percent respectively in 2000, it is also expected that the degree of awareness among Malaysians in reporting age should also increased. With this assumption, one should expect that the level of age reporting is better for urban than rural. However, the results shows that there is no correlation between literacy rate and age reporting. The age reported were not consistent between the two censuses. In 1991, digit “4” is avoided by males and females at urban stratum but there is no clear evidence for the 2000 census. Digit “0”, “5”, “6” and “8” is preferred by females in 2000 but no clear evidence of digit avoidance at urban level. Rural-stratum population also have similar pattern of age reporting as urban stratum population where, digit “0”, “5” are the most preferable digits with additional digit of “6” and “8” for female population in 2000 (Table 4b and 4c).

The result on age reporting by ethnic group shows that Malay community preferred digit “0”, “5” for both censuses with additional of digit “6” for male and digit “8” female in the 2000 census. Digit “4” was avoided in 1991 census and is not statistically significant for 2000 census. Chinese and Indians, on the other hand, show significant heaping for digit “0” for both censuses and there is no strong evidence for the remaining terminal digits (Table 4d, 4e and 4f).

To sum up, all the three methods that were used in this analysis were in unanimity with the issues highlighted that irregularities does exist in age reporting for Malaysians. Digits “0” and “5” are the most preferred, as shown by the bar–charts that clearly spike at both digits. The Whipple index falls under category “approximate” for those digits compared to “highly accurate” and “fairly accurate” for the remaining digits. The t-test for differences of those digits also in accordance to heaping for digit “0” and “5”.

The findings also suggest that Malay contributes more than Chinese and Indian for misstatement of ages and preference of digit “0” and “5”. Digit “4” was avoided for both males and females in 1991 census. In the 2000 census, digit “6” and “8” female were the additional digits preferred by Malay males and females respectively. On the other hand, significant heaping occurs at digit “0” for Chinese and Indians in both censuses and there is no strong evidence for the remaining terminal digits. Through this finding, it can be concluded that there is no significant change in trend in age reporting for Malay since it was reported by Saw Swee Hock (1967), that the age accuracy of Malays may be classified as “rough” for males and “very rough” for females; that of the Chinese as “fairly accurate” for males and “highly accurate” for females; and that of Indians as “rough” for both males and females.

Conclusion and recommendations

In general, there is digit preference in age reporting in both censuses for gender and ethnic group. Digit avoidance, however, occurred only for 1991 census. The findings are also consistent with an earlier study that Malay contributed more to the quality of age misreporting compared to Chinese and Indians. If age heaping occurs at the start of age groups, age misreporting may result in a slight bias in parity calculations. Since age heaping seems to have occurred in all categories, the net effect is expected to be negligible.

Field enumerators of Kelantan and Sabah in the 1991 Census and Kelantan and Terengganu in the 2000 Census benefitted from the training conducted directly by DOSM. The field enumerators of other states, however, received their training from District Officer. It is suggested that further analysis can be carried out to determine whether insufficient training of field enumerators contributes to age misreporting. The outcome from the suggested analysis will be useful to make concrete conclusion on the misreporting that is attributed to insufficient training.

References

Amiee R. D, Samuel H. Preston. (1991). Age misreporting and its effects on adult mortality estimated in Latin America. *Population Bulletin of the United Nations*, 31/32

Bairagi R, Aziz K.M, Chowdhury M.K and Edmonton B. (1982, November). Age misstatement for young children in rural Bangladesh*. Demography*, vol 19, no. 4

Department of Statistics, Malaysia. (1995). *Post enumeration surveys 1991 report, (restricted edition)*. Kuala Lumpur

Department of Statistics, Malaysia. (2003). *Malaysia, census coverage evaluation survey 2000 report (restricted edition)*. Putrajaya

Hill, M.E, Preston, S.H and et. al. (1997, March). Age-linked institutions and age reporting among older African Americans*. Social Forces*, vol. 75, no. 3

Ibrahim Naseem, Bakhta Ghubhaju and Hussain Niyaaz. (2004). Rapid fertility decline in the Maldives: An assessment*. Asia-Pasific Population Journal*, vol. 19, no. 3

Kabir M and Chowdhury K. (1981). The pattern of age reporting errors in the districts of Bangladesh*. Rural Demography Journal*, vol 8, no. 2

Mason, K.O and Cope, L.G. (1987, November). Sources of age and date of birth misreporting in 1900 U.S. census*. Journal of Demography*, vol 24, no. 4

Mukherjee, B.N and Mukhopadhyay, B.K. (1988, Jan.-June). A study of digit preference and quality of age data in Turkish censuses*. Genus*, vol. 44, no. 1-2

Nagi M.H., Stockwell, E.G and Snavley, L.M. (1973, August). Digit preference and avoidance in the age statistics of some recent African censuses: some patterns and correlates*. International Statistical Review*, vol. 41, no. 2

Poston, D. L. Jr, Chu, I.H.J and et. al. (2006, February 23). *An analysis of the quality of the age and sex data of the Republic of Korea and its province, 1970 and 1995, 2000 (draft)*. Available from: http://sociweb.tamu.edu/Faculty/POSTON/Postonweb/research/SKE/text.pdf

Pullum T. W. (2005, April 22). *Assessing the quality of age and date reporting in the demographic and health surveys.* Paper prepared for UT Population Research Center brownbag

Rini Savitridina. (1997, June). Determinants and consequences of early marriage in Java, Indonesia. *Asia – Pasific Population Journal* , vol.12, no. 2

Roberts, J. M and Brewer, D.D. (2001, September). Measures and test of heaping in discrete quantative distributions. *Journal of Applied Statistics*, vol, 28, no. 7

Saw, Swee Hock, (1967). Error in Chinese age statistics. *Demography 4*

Shyrock, Henry S., Jacob S. Siegel, and Associates. (1971). *The methods and materials of demography*. U.S. Bureau of Census, Washington

Ueda, K. (1980). *Basic demographic data, their accuracy and use for subnational population projection in Asian and Pacific countries*. Tokyo: South East Asian Medical Information Center

United Nations. (1990). *1988 Demographic Yearbook*. New York

Wang Zhenglian, Zeng Yi, Jeune Bernard, James W. Vaupel. (1998, Jan.-June). Age validation of Han Chinese centenarian. *Genus*, vol. 54, no 1-2

You, Poh Seng. (1959, November). Errors in Age Reporting in Statistically Underdevelop countries. *Population studies*, vol. 13, no. 2

Appendix 1

Chart 1a: Pyramid chart of single age distribution of total population by sex group, Malaysia 1991

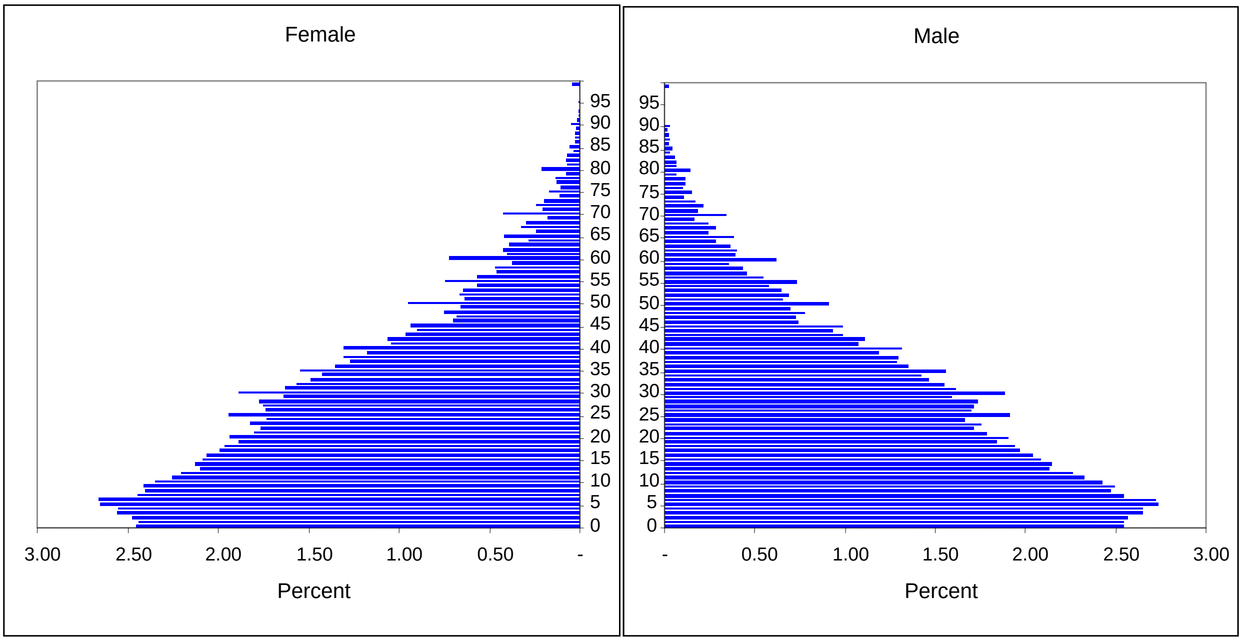


Chart 1b: Pyramid chart of single age distribution of total population by sex group, Malaysia 2000

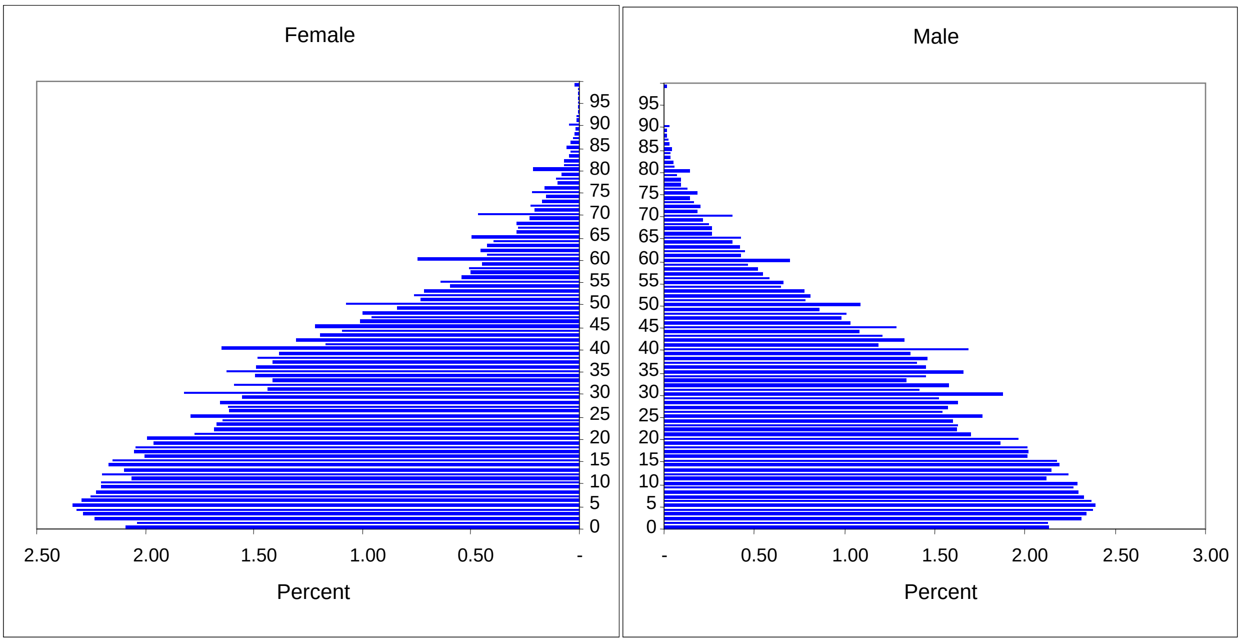


Chart 2a: Pyramid chart of single age distribution of urban-area population by sex group, Malaysia 1991

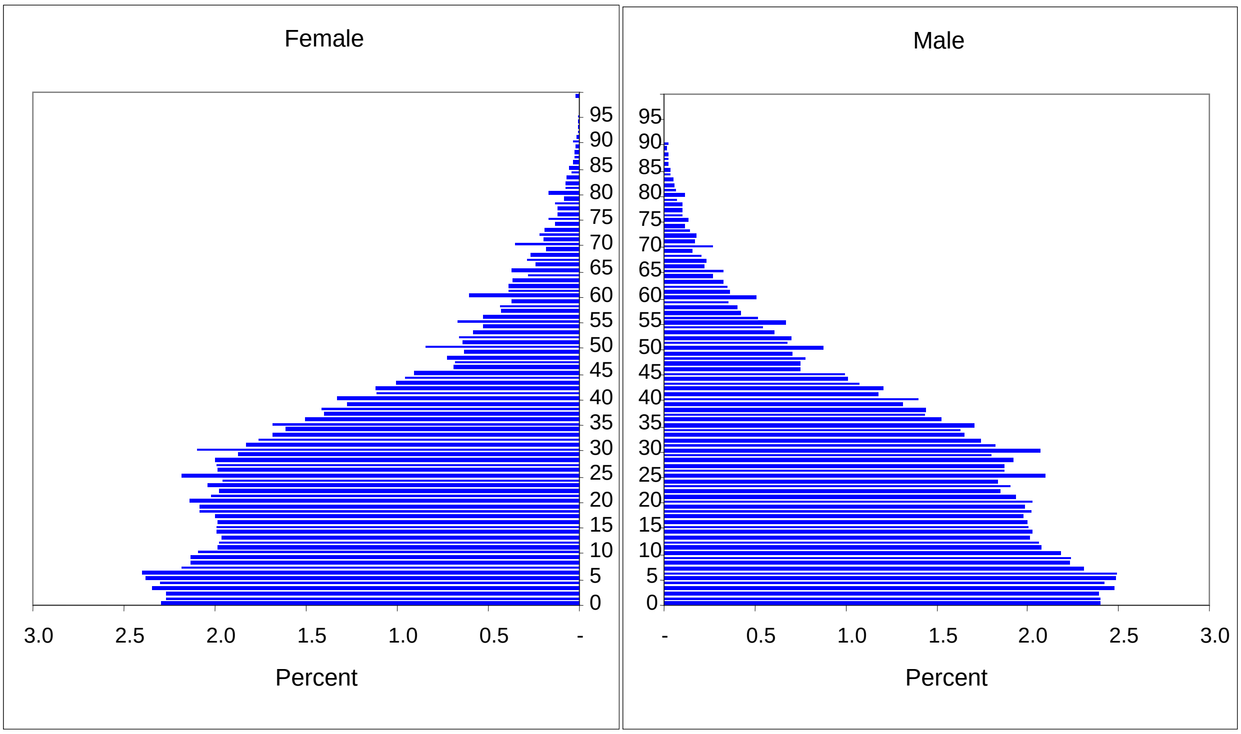


Chart 2b: Pyramid chart of single ages distribution of urban-area population by sex group, Malaysia 2000

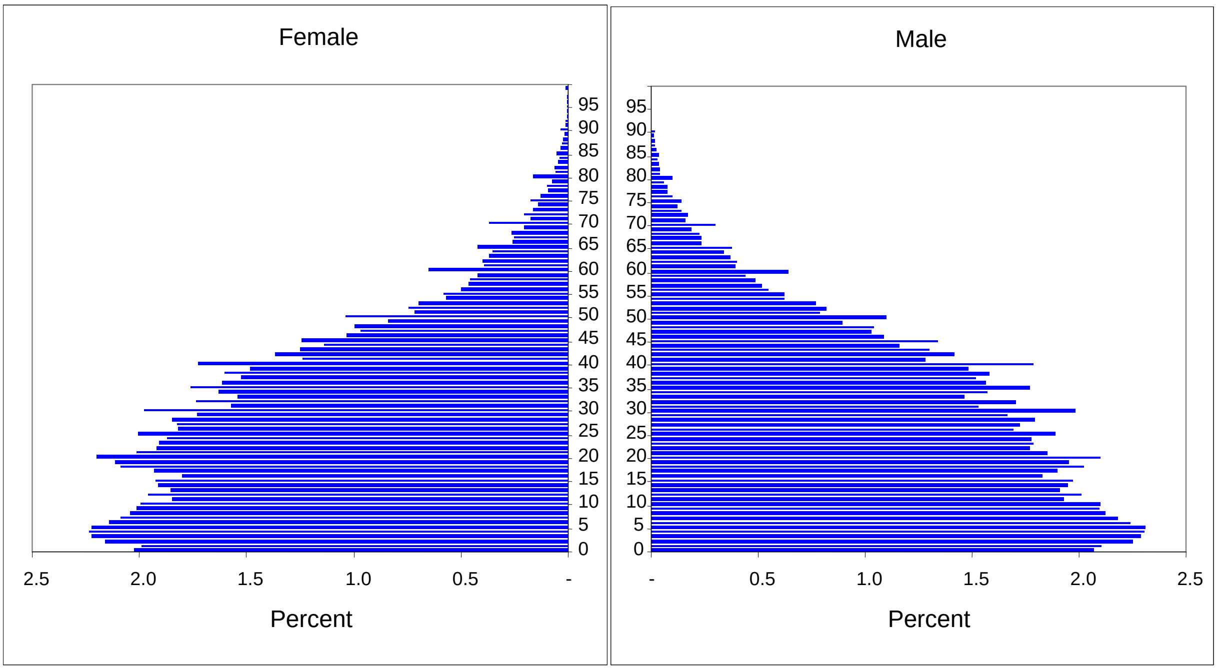


Chart 2c: Pyramid chart of single age distribution of rural-area population by sex group, Malaysia 1991

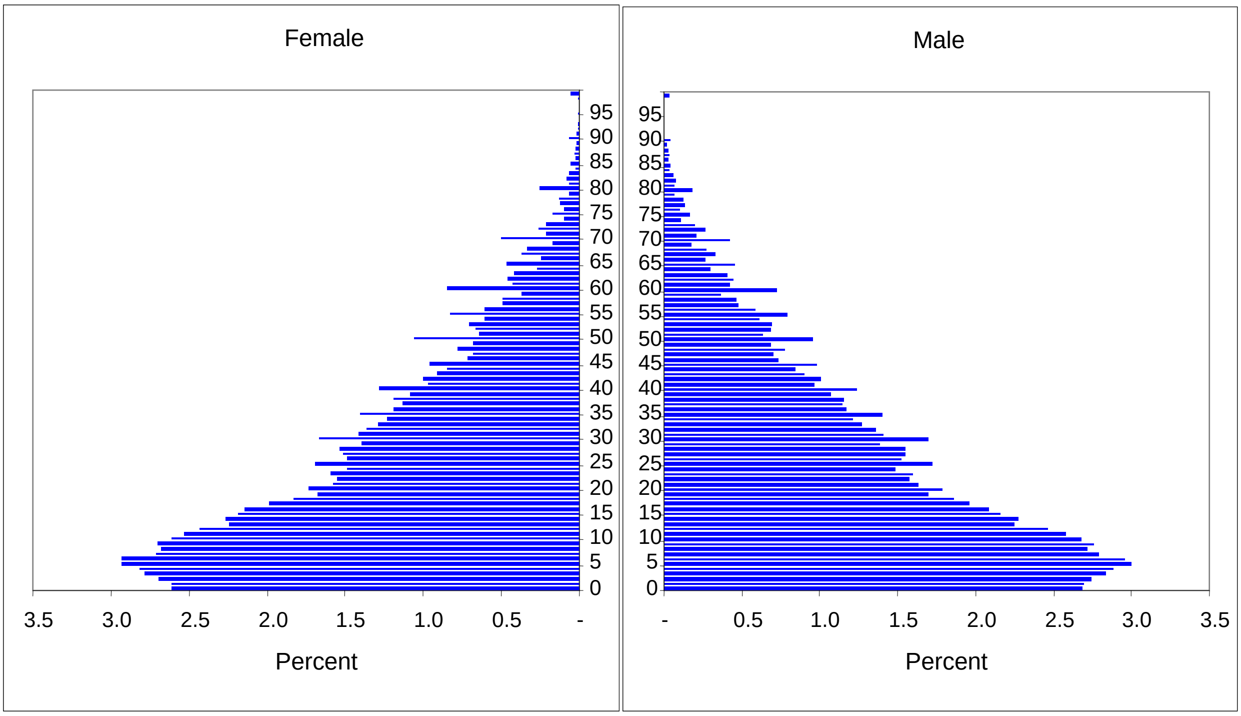


Chart 2d: Pyramid chart of single age distribution of rural-area population by sex group, Malaysia 2000

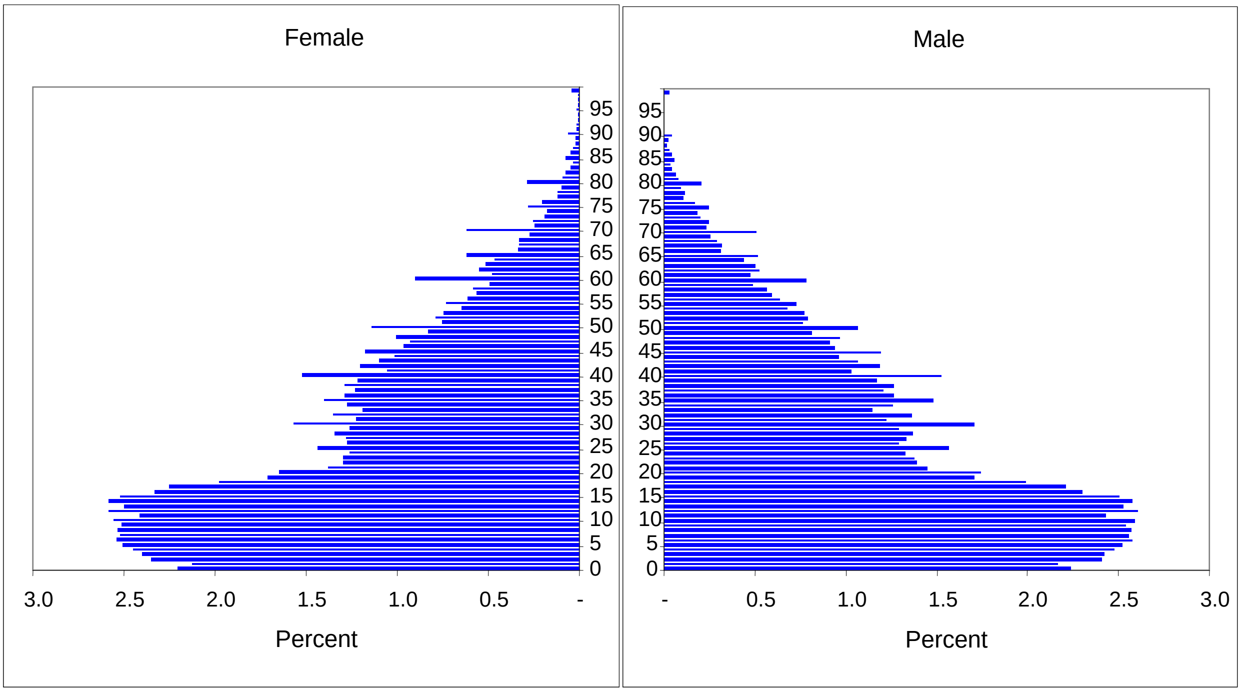


Chart 3a: Pyramid chart of single age distribution of Malay population by sex group, Malaysia 1991

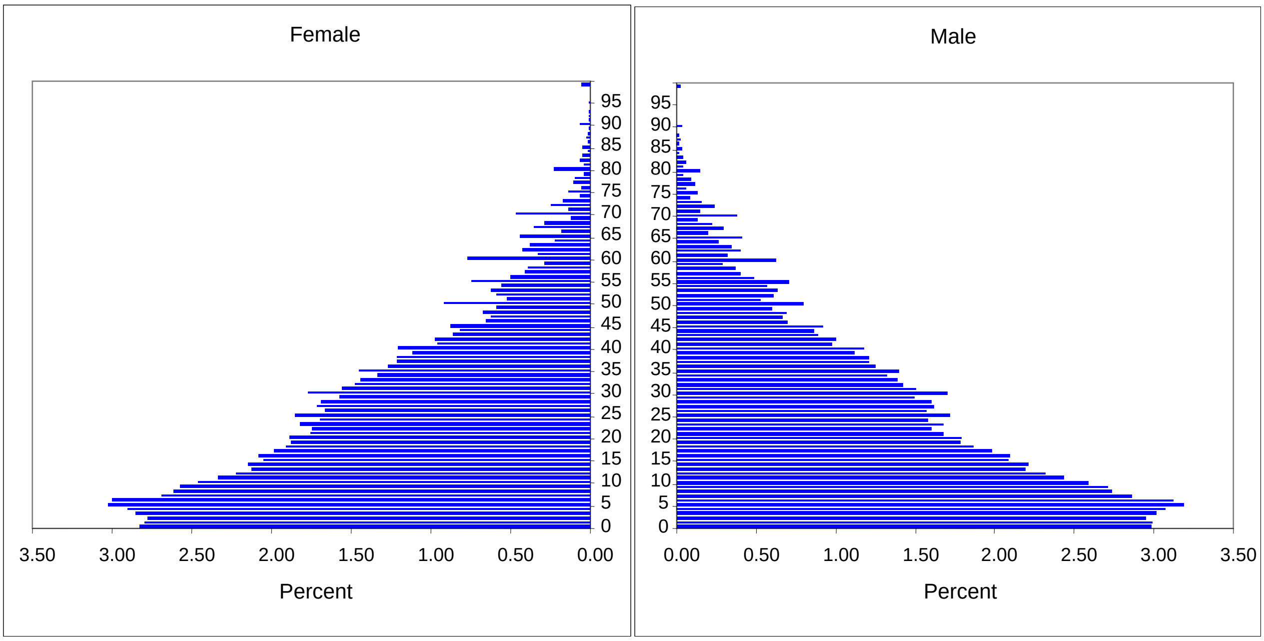


Chart 3b: Pyramid chart of single age distribution of Malay population by sex group, Malaysia 2000

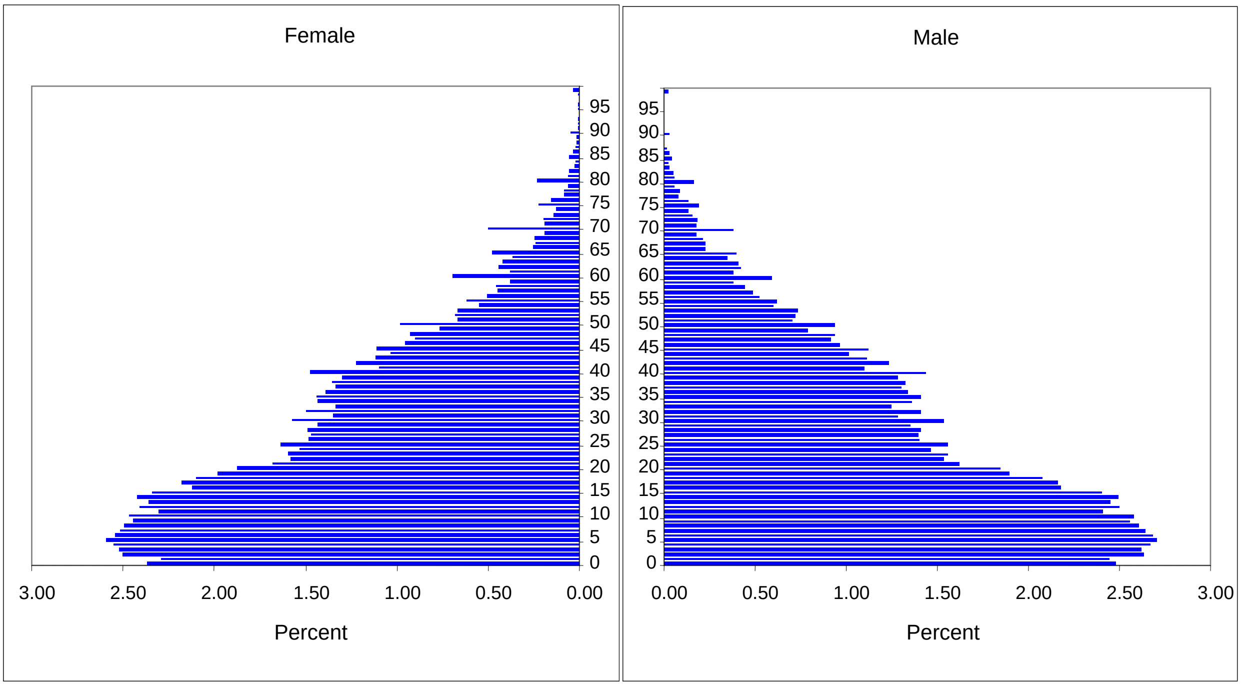
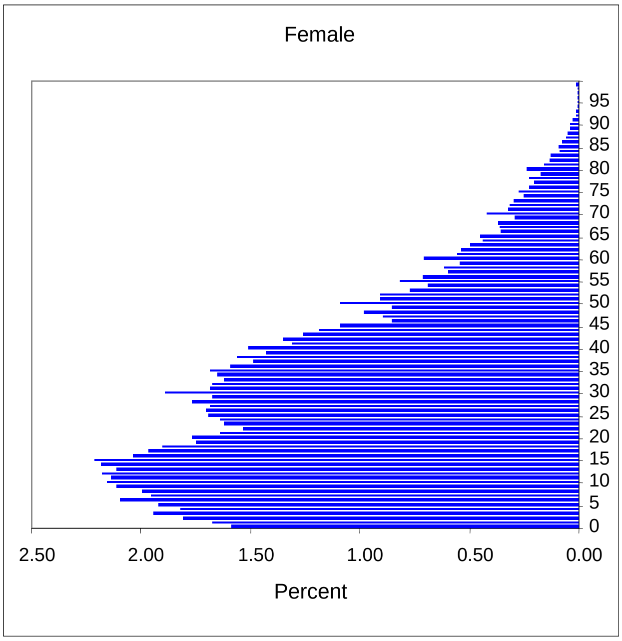


Chart 3c: Pyramid chart of single age distribution of Chinese population by sex group, Malaysia 1991



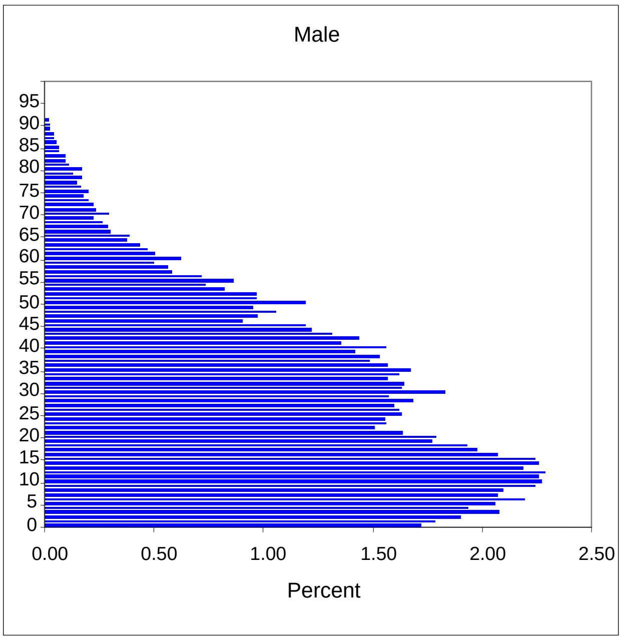
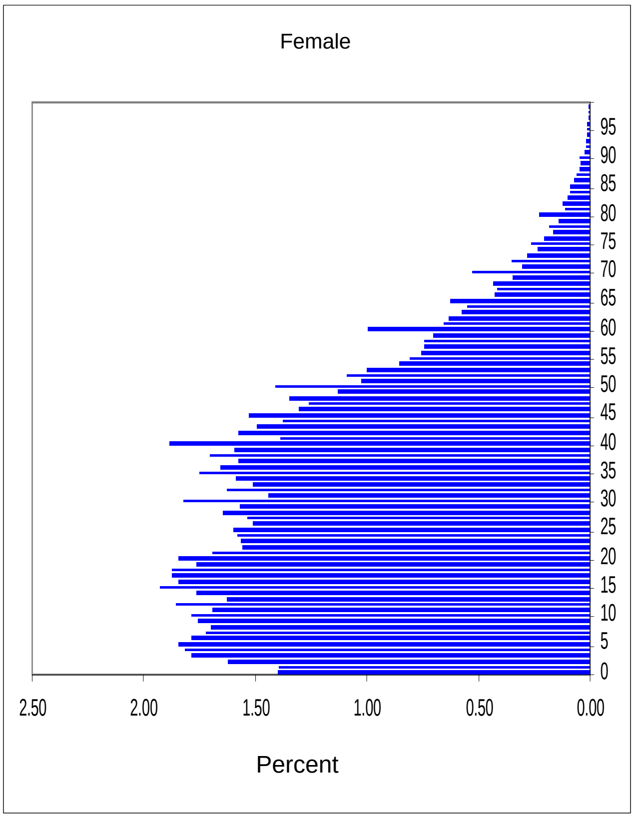


Chart 3d: Pyramid chart of single age distribution of Chinese population by sex group, Malaysia 2000



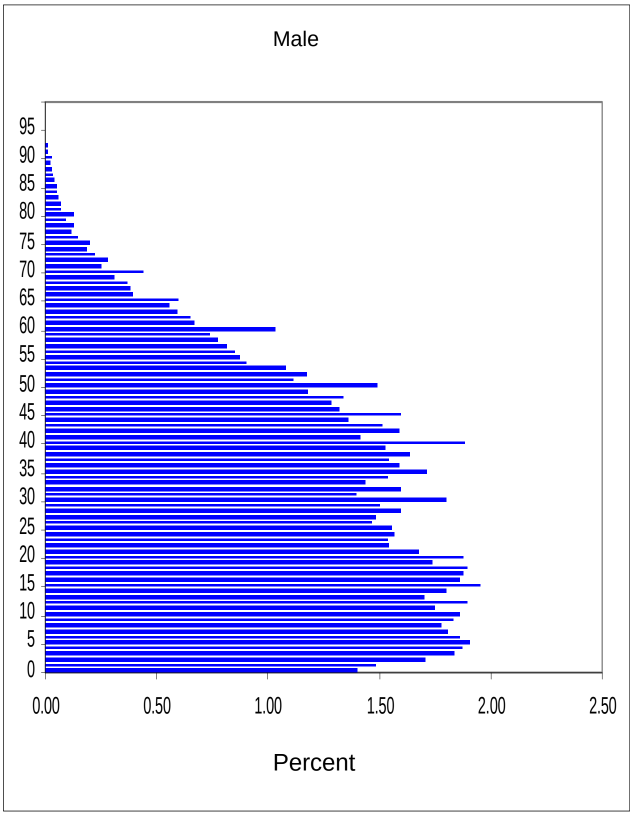


Chart 3e: Pyramid chart of single age distribution of Indian population by sex group, Malaysia 1991

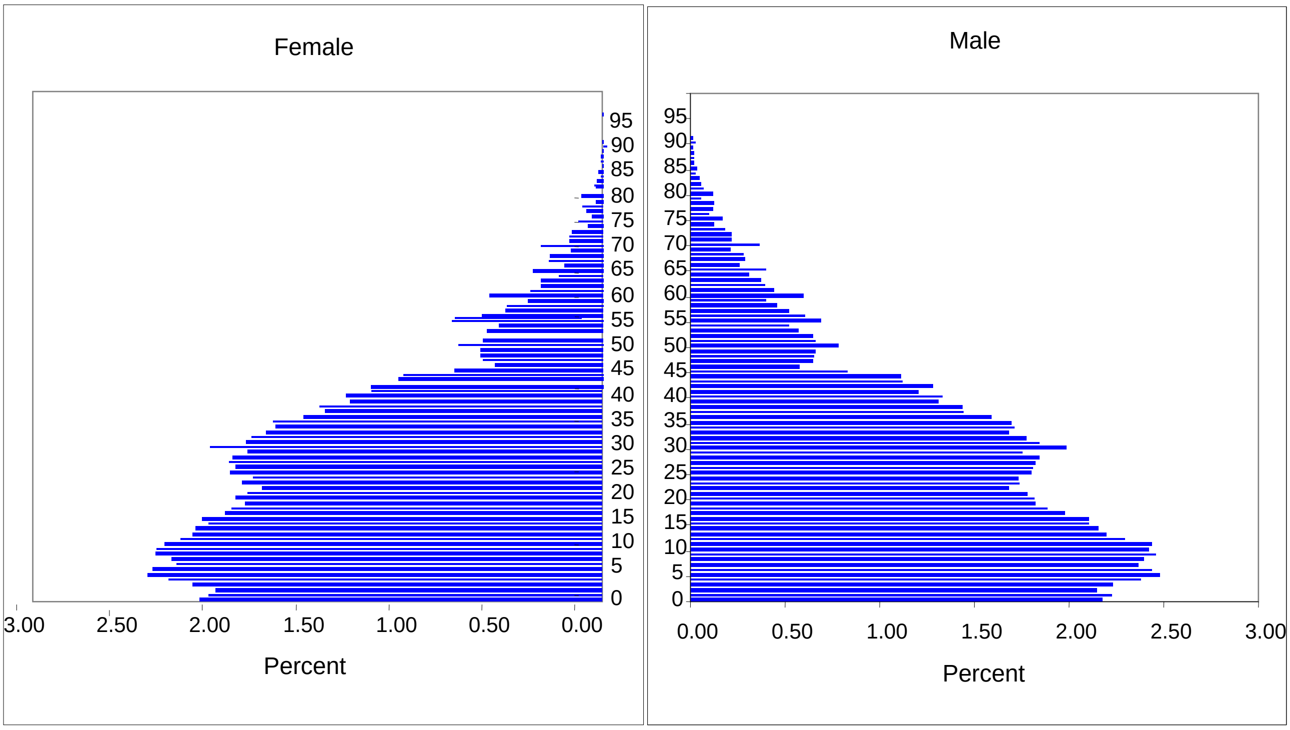


Chart 3f: Pyramid chart of single age distribution of Indian population by sex group, Malaysia 2000

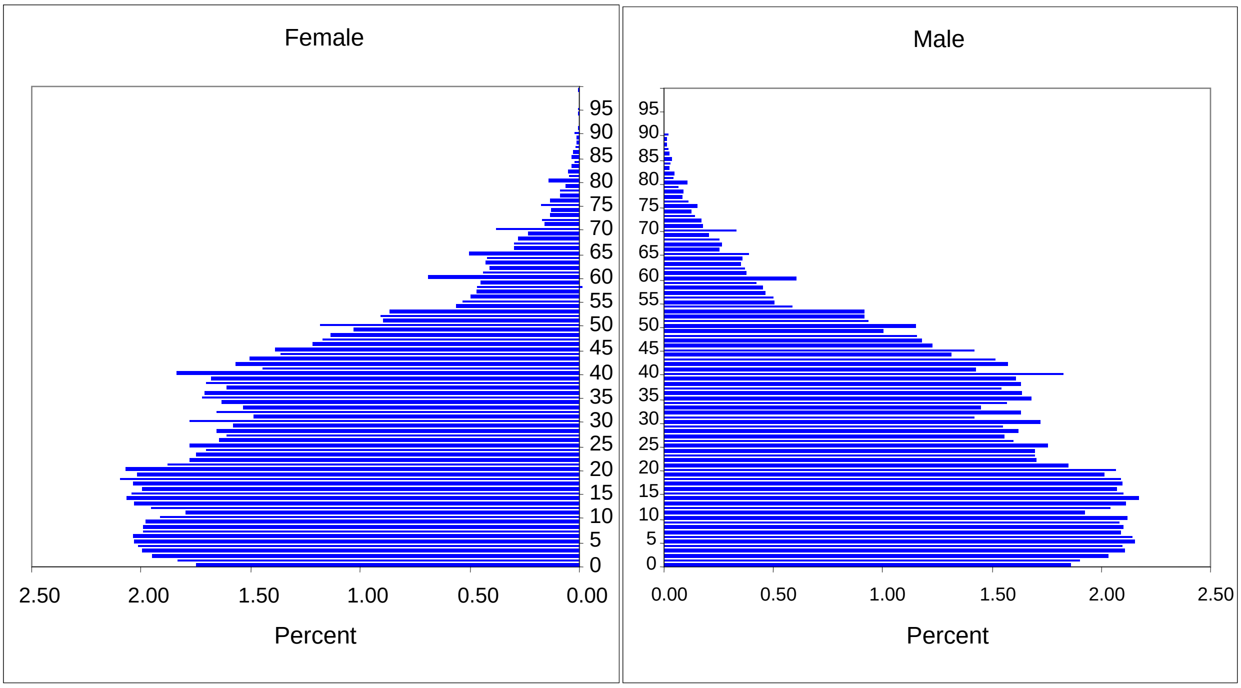


Table 1: Whipple-type indices of terminal digit of ages of total population by sex group, Malaysia, 1991 and 2000

Age Terminal Digit	Single ages	1991		2000	
		Male	Female	Male	Female
0	20,30,40,50,60,70	133.4	136.1	138.6	138.4
1	21,31,41,51,61,71	108.9	107.8	102.5	102.6
2	22,32,42,52,62,72	108.4	107.7	107.9	107.4
3	23,33,43,53,63,73	102.8	103.7	99.8	99.8
4	24,34,44,54,64,74	95.2	94.4	95.6	95.9
5	25,35,45,55,65,75	109.3	108.4	107.8	106.8
6	26,36,46,56,66,76	89.4	88.6	89.9	91.0
7	27,37,47,57,67,77	87.3	86.9	87.5	87.0
8	28,38,48,58,68,78	87.7	89.1	89.4	90.1
9	29,39,49,59,69,79	77.7	77.3	81.0	81.0

Table 2a: Whipple-type indices of terminal digit of ages of urban population by sex group, Malaysia, 1991 and 2000

Age Terminal Digit	Single ages	1991		2000	
		Male	Female	Male	Female
0	20,30,40,50,60,70	129.1	130.7	136.3	135.5
1	21,31,41,51,61,71	110.8	109.7	103.5	104.0
2	22,32,42,52,62,72	108.8	108.5	108.4	108.4
3	23,33,43,53,63,73	103.1	104.1	100.6	101.0
4	24,34,44,54,64,74	97.7	97.1	96.4	97.0
5	25,35,45,55,65,75	107.2	106.1	105.7	105.3
6	26,36,46,56,66,76	89.7	89.7	90.1	91.1
7	27,37,47,57,67,77	86.8	87.1	87.8	87.2
8	28,38,48,58,68,78	87.6	88.4	89.7	89.7
9	29,39,49,59,69,79	79.3	78.5	81.6	80.9

Table 2b: Whipple-type indices of terminal digit of ages of rural population by sex group, Malaysia, 1991 and 2000

Age Terminal Digit	Single ages	1991		2000	
		Male	Female	Male	Female
0	20,30,40,50,60,70	138.4	142.5	142.7	143.8
1	21,31,41,51,61,71	106.7	105.4	100.8	99.9
2	22,32,42,52,62,72	107.9	106.8	107.1	105.5
3	23,33,43,53,63,73	102.5	103.1	98.3	97.7
4	24,34,44,54,64,74	92.3	91.3	94.1	94.0
5	25,35,45,55,65,75	111.6	111.1	111.6	109.6
6	26,36,46,56,66,76	88.9	87.3	89.6	90.8
7	27,37,47,57,67,77	87.9	86.8	87.0	86.7
8	28,38,48,58,68,78	88.0	89.8	88.9	90.9
9	29,39,49,59,69,79	75.8	75.9	79.9	81.2

Table 3a: Whipple-type indices of terminal digit of ages of Malay population by sex group, Malaysia, 1991 and 2000

Age Terminal Digit	Single ages	1991		2000	
		Male	Female	Male	Female
0	20,30,40,50,60,70	133.7	140.1	129.0	132.1
1	21,31,41,51,61,71	106.6	105.3	100.8	99.8
2	22,32,42,52,62,72	109.0	109.2	105.5	104.5
3	23,33,43,53,63,73	105.3	105.8	99.7	98.2
4	24,34,44,54,64,74	96.6	93.7	94.3	93.8
5	25,35,45,55,65,75	109.3	110.0	101.2	102.4
6	26,36,46,56,66,76	88.2	86.4	87.8	88.2
7	27,37,47,57,67,77	89.0	88.1	84.5	83.3
8	28,38,48,58,68,78	86.4	86.9	84.5	84.9
9	29,39,49,59,69,79	76.0	74.5	77.4	77.0

Table 3b: Whipple-type indices of terminal digit of ages of Chinese population by sex group, Malaysia, 1991 and 2000

Age Terminal Digit	Single ages	1991		2000	
		Male	Female	Male	Female
0	20,30,40,50,60,70	126.1	124.8	130.8	129.1
1	21,31,41,51,61,71	109.5	108.4	100.3	99.1
2	22,32,42,52,62,72	108.1	106.6	104.9	104.0
3	23,33,43,53,63,73	102.0	102.5	97.9	97.8
4	24,34,44,54,64,74	98.3	98.9	93.8	94.1
5	25,35,45,55,65,75	102.8	101.6	100.4	100.0
6	26,36,46,56,66,76	91.1	92.0	88.5	89.2
7	27,37,47,57,67,77	87.7	88.3	86.4	86.7
8	28,38,48,58,68,78	91.3	93.1	89.6	92.1
9	29,39,49,59,69,79	83.0	83.8	82.2	83.3

Table 3c: Whipple-type indices of terminal digit of ages of Indian population by sex group, Malaysia, 1991 and 2000

Age Terminal Digit	Single ages	1991		2000	
		Male	Female	Male	Female
0	20,30,40,50,60,70	125.4	126.2	127.6	128.3
1	21,31,41,51,61,71	112.3	110.1	102.5	101.8
2	22,32,42,52,62,72	109.5	108.6	105.5	104.8
3	23,33,43,53,63,73	103.4	105.2	100.6	100.3
4	24,34,44,54,64,74	100.5	98.5	93.7	93.8
5	25,35,45,55,65,75	102.1	103.9	98.0	98.5
6	26,36,46,56,66,76	90.3	89.4	88.3	88.7
7	27,37,47,57,67,77	88.6	88.8	84.4	84.6
8	28,38,48,58,68,78	87.8	89.4	86.4	86.3
9	29,39,49,59,69,79	80.2	79.9	80.5	81.4

Table 4a: Paired sample test (paired differences) of terminal digit of total population ages by sex, Malaysia

Pair digits	1991			2000		
	Male	Female	Total	Male	Female	Total
0 – 1	18991,2* (8.787)	21847.2* (9.027)	40838.3* (9.292)	37508.3* (6.687)	36545.8* (8.914)	78089.6* (7.503)
1 – 2	393.5 (0.219)	22.2 (0.016)	415.7 (0.131)	-5582.7 (-1.361)	-4924.0 (-1.243)	-16264** (-2.360)
2 - 3	4328.5** (2.129)	3133.7 (1.615)	7462.2 (1.881)	8469.8** (2.078)	7757.7* (2.708)	(19502.8)** (2.605)
3 - 4	5889.8* (8.852)	7085.2* (9.849)	12975.0* (10.053)	4344.0 (1.094)	3945.5 (1.233)	8554.0 (0.975)
4 - 5	-10898.2* (-4.047)	-10745.2* (-4.317)	-21643.3* (-4.270)	-12690.3* (-3.173)	-11106.2* (-5.865)	-21619.4* (-3.358)
5 - 6	15414.5* (6.174)	15245.3* (7.432)	30659.8* (11072.8)	18594.0* (5.072)	16145.2* (5.842)	32707.2* 4.577
6 - 7	1577.2 (0.811)	1271.7 (0.502)	2848.8 (0.640)	2508.0 (1.669)	4096.7* (2.912)	8774.4* (3.775)
7 - 8	-336.0 (-0.271)	-1624.3 (-1.412)	-1960.3 (-0.831)	-1980.2 (-1.103)	-3173.7* (-2.677)	-4003.8 (-1.223)
8 - 9	7806.5* (6.389)	9071.2* (8.660)	16877.7 (7.813)	8765.5* (3.937)	9311.7* (4.278)	16947.0* (3.272)
9 - 0	-43167* (-7.110)	-45306.8* (-7.807)	-88473.8* (-7.509)	-59936.5* (-7.412)	-58598.7* (-8.223)	-122687.4* (-6.918)

Note: Figures in bracket are t statistics

* Significant at 5 percent significant level
** Significant at 10 percent significant level

Table 4b: Paired sample test (paired differences) of terminal digit of urban population by sex, Malaysia

Pair digits	1991		2000	
	Male	Female	Male	Female
0 - 1	7540.5* (6.470)	8736.2* (8.879)	21955.5* (5.646)	20900.8* (6.339)
1 - 2	818.2 (0.927)	517.0 (0.732)	-3249.8 (-1.213)	-2911.7 (-1.123)
2 - 3	2359.3 (1.952)	1808.5 (1.587)	5224.8 (1.977)	4963.8* (2.582)
3 - 4	2228.0* (6.921)	2933.0* (11.461)	2794.3 (1.065)	2645.2 (1.248)
4 - 5	-3920.3** (-2.263)	-3753.7** (-2.227)	-6206.5** (-2.540)	-5533.2* (-3.985)
5 - 6	7202.7* (5.036)	6814.7* (6.380)	10438.2* (4.541)	9442.3* (5.381)
6 - 7	1206.2 (1.307)	1100.8 (0.980)	1539.3 (1.746)	2605.3* (2.574)
7 - 8	-311.0 (-0.567)	-537.5 (-1.632)	-1272.3 (-1.073)	-1672.7** (-2.107)
8 - 9	3392.3* (4.420)	4129.5* (6.169)	5422.3* (3.668)	5875.8* (3.936)
9 - 0	-20515.8* (-4.770)	-21748.5* (-5.240)	-36645.8* (-5.834)	-36315.8* (-6.530)

Note: Figures in bracket are t statistics

* Significant at 5 percent significant level
** Significant at 10 percent significant level

Table 4c: Paired sample test (paired differences) of terminal digit of rural population by sex, Malaysia

Pair digits	1991		2000	
	Male	Female	Male	Female
0 - 1	11450.7* (9.870)	13111.0* (7.341)	15552.8* (8.702)	15645.0* (13.576)
1 - 2	-424.7 (-0.458)	-494.8 (-0.680)	-2332.8 (-1.597)	-2012.3 (-1.402)
2 - 3	1969.2 (1.964)	1325.2 (1.306)	3245.0** (2.226)	2793.8* (2.908)
3 - 4	3661.8 (7.836)*	4152.2* (7.694)	1549.7 (1.103)	1300.3 (1.175)
4 - 5	-6977.8* (-6.281)	-6991.5* (-7.198)	-6483.8* (-4.002)	-5573.0* (-9.031)
5 - 6	8211.8* (7.803)	8430.7* (7.979)	8155.8* (5.716)	6702.8* (5.395)
6 - 7	371.0 (0.327)	170.8 (0.114)	968.7 (1.393)	1491.3* (2.771)
7 - 8	-25.0 (-0.031)	-1086.8 (-1.313)	-707.8 (-0.997)	-1501.0* (-2.884)
8 - 9	4414.2* (7.192)	4941.7* (8.067)	3343.2* (4.079)	3435.8* (4.142)
9 - 0	-22651.2* (-11.584)	-23558.3* (-9.655)	-23290.7* (-12.566)	-22282.8* (-9.249)

Note: Figures in bracket are t statistics

* Significant at 5 percent significant level
** Significant at 10 percent significant level

Table 4d: Paired sample test (paired differences) of terminal digit of Malay population by sex, Malaysia

Pair digits	1991		2000	
	Male	Female	Male	Female
0 - 1	9296.3* (8.160)	12385.0* (6.073)	13925.3* (11.901)	16280.0* (10.546)
1 - 2	-833.7 (-0.619)	-1385.7 (-1.106)	-2301.5 (-1.207)	-2358.8 (-1.151)
2 - 3	1255.8 (1.041)	1225.7 (0.987)	2872.2 (1.662)	3166.2** (2.223)
3 - 4	2995.5* (6.628)	4298.8* (6.149)	2640.3 (1.282)	2253.3 (1.295)
4 - 5	-4360.7* (-5.385)	-5805.0* (-5.372)	-3387.0* (-4.613)	-4346.5* (-4.676)
5 - 6	7224.5* (7.001)	8412.7* (7.728)	6621.3* (5.931)	7146.8* (4.653)
6 - 7	-259.3 (-0.215)	-628.8 (-0.376)	1654.3* (3.106)	2476.3* (4.164)
7 - 8	900.2 (1.485)	439.2 (0.699)	-19.0 (-0.033)	-817.5* (-3.043)
8 - 9	3536.7* (10.930)	4424.8* (7.654)	3487.7* (3.372)	4020.8* (3.804)
9 - 0	-19755.3* (-9.514)	-23367.0* (-9.181)	-25493.7* (-7.469)	-27820.7* (-8.118)

Note: Figures in bracket are t statistics

* Significant at 5 percent significant level
** Significant at 10 percent significant level

Table 4e: Paired sample test (paired differences) of terminal digit of Chinese population by sex, Malaysia

Pair digits	1991		Chinese	
	Male	Female	Male	Female
0 - 1	3729.2* (6.246)	3713.0* (9.307)	9074.3* (7.099)	8645.3* (6.553)
1 - 2	312.7 (0.463)	402.8 (0.867)	-1388.2 (-.993)	-1411.3 (-1.059)
2 - 3	1364.5 (1.955)	930.5 (1.308)	2081.2* (3.555)	1791.5* (4.153)
3 - 4	833.2 (1.619)	809.0 (1.840)	1239.7 (1.052)	1043.2 (1.154)
4 - 5	-1004.7 (-1.962)	-594.2 (-0.866)	-1958.8 (-1.615)	-1693.5 (-1.930)
5 - 6	2615.5* (2.783)	2169.8* (2.796)	3526.3* (3.253)	3105.8* (3.943)
6 - 7	770.7 (1.205)	834.0 (1.402)	623.2** (2.321)	733.0 (1.847)
7 - 8	-796.3 (-1.74)	-1087.8* (-3.142)	-967.3 (-1.446)	-1552.5* (-2.711)
8 - 9	1858.8* (5.530)	2102.3* (6.878)	2216.5* (3.988)	2541.2* (3.695)
9 - 0	-9683.5* (-4.867)	-9279.5* (-4.858)	-14446.8* (-5.914)	-13202.7* (-5.329)

Note: Figures in bracket are t statistics

* Significant at 5 percent significant level
** Significant at 10 percent significant level

Table 4f: Paired sample test (paired differences) of terminal digit of Indian population by sex, Malaysia

Pair digits	1991		2000	
	Male	Female	Male	Female
0 - 1	786.8* (6.706)	980.8* (7.148)	1989.8* (7.402)	2161.7* (9.604)
1 - 2	173.8 (1.048)	97.3 (0.812)	-238.7 (-0.586)	-243.0 (-0.757)
2 - 3	366.8 (1.866)	204.5 (0.911)	388.2 (1.774)	362.7** (2.426)
3 - 4	176.1 (1.770)	412.8* (5.979)	551.2 (1.046)	536.2 (1.188)
4 - 5	-98.2 (-0.233)	-328.83 (-0.685)	-337.7 (-1.507)	-389.0* (-2.767)
5 - 6	711.0* (3.037)	882.7* (4.692)	766.2* (3.167)	799.3* (3.008)
6 - 7	105.0 (0.480)	37.2 (0.152)	304.8* (2.879)	338.7* (3.170)
7 - 8	49.0 (0.634)	-34.2 (-0.647)	-153.8 (-1.041)	-141.3 (-0.934)
8 - 9	454.5* (3.801)	581.0* (3.930)	462.5* (2.860)	403.0* (3.603)
9 - 0	-2725.0* (-4.320)	-2833.3* (-4.188)	-3732.5* (-4.283)	-3828.2* (-4.512)

Note: Figures in bracket are t statistics

* Significant at 5 percent significant level
** Significant at 10 percent significant level