Appendix

Sample Design and Sampling Variability

Sample Design and Selection

The sample for the Census of Population 2020 was selected from a sampling frame comprising all residential dwelling units in Singapore. As the sample survey of the Census 2020 covered only households in residential dwellings, institutions such as military camps, hostels and hotels were excluded from the frame.

The sample was selected based on a stratified design. Dwelling units in the sampling frame were divided into different groups. The groups are defined based on the planning areas demarcated by the Urban Redevelopment Authority and broad dwelling type groups. A random sample was then selected from the different groups by systematic sampling with a random start. The samples selected from each group were combined to form the required sample of about 150,000 dwelling units.

Sampling Variability

As the survey estimates are based on information obtained from a fraction of the population instead of the whole population, the precision of estimates derived from the sample survey are affected by sampling errors. Sampling errors refer to the difference between the estimate based on a sample and its 'true' population value that would result if the whole population has been surveyed.

The extent of sampling error of an estimate under a particular sample design is assessed by the variability of the estimate across all possible samples under the design. One common measure of this variability is given by the standard error (SE), which is the standard deviation of the sampling distribution of the estimate. Another measure is the relative standard error (RSE), which is obtained by expressing the standard error as a percentage to the estimate. The smaller the RSE, the more precise is the estimate.

Generalised Sampling Errors Tables

From Table A1, the DEFT[1] for most of the selected attributes (T_Y) is about 1. It is impractical to compute and display the sampling error for each and every of the possible estimates such as the total number of elements in the population with a given attribute Y from the Census 2020. Thus, generalized sampling errors tables are provided instead as a guide to data users for estimating the errors of any estimates.

Since most of the attributes in Table A1 have DEFT about 1, data users should generally use Table A2 (DEFT value of 1) to determine sampling errors for the attribute of interest. For attributes of individuals with larger DEFT, such as language most frequently spoken at home, data users can refer to Table A3 (DEFT value of 2).

The smaller the estimate, the larger is the RSE. This implies that sample estimates of a rare characteristic would have high RSEs and users would have to be careful in drawing inferences based on the sample estimates.

Table A1 Sampling Errors and DEFT of T_Y for Selected Attributes, Census 2020

	Sample Estimate	Standard Error	Relative Standard Error	95% Confidence Interval ('000)		DEFT
	('000)	(T_Y)	(T_Y)	Lower	Upper	(T_Y)
Residents Aged 15 Years & Over						
Single	1,090.7	3,885	0.4	1,083.1	1,098.3	1.22
Married	2,035.4	4,296	0.2	2,027.0	2,043.8	1.14
Widowed	183.2	1,429	0.8	180.4	186.0	0.98
Divorced/Separated	149.8	1,375	0.9	147.1	152.5	1.04
Ever-Married Resident Females Aged 15 Years & Over						
With No Children Born	164.3	1,387	0.8	161.6	167.0	1.00
With 1 Child Born	256.0	1,674	0.7	252.7	259.3	0.98
With 2 Children Born	462.6	2,034	0.4	458.7	466.6	0.90
With 3 Children Born	238.6	1,604	0.7	235.5	241.8	0.97
With 4 or More Children Born	136.5	1,266	0.9	134.0	139.0	1.00
Residents Aged 15 Years & Over*						
With Below Secondary Qualifications	765.9	3,162	0.4	759.7	772.1	1.13
With Secondary Qualifications	505.6	2,599	0.5	500.5	510.7	1.11
With Post-Secondary (Non-Tertiary) Qualifications	341.9	2,193	0.6	337.6	346.2	1.12
With Diploma and Professional Qualifications	519.9	2,623	0.5	514.8	525.1	1.11
With University Qualifications	1,007.2	3,493	0.3	1,000.3	1,014.0	1.12

  * Data pertain to residents who were not attending educational institutions as full-time students and include those who were upgrading their qualifications through part-time courses.

---

[1] The DEFT is the ratio of the standard error of the estimate, under the sample design used, to that of a simple random sample. This ratio measures the effect of the complexity of the sample design on the standard error.

Table A1 Sampling Errors and DEFT of T_Y for Selected Attributes, Census 2020 (cont'd)

	Sample Estimate	Standard Error	Relative Standard Error	95% Confidence Interval ('000)		DEFT
	('000)	(T_Y)	(T_Y)	Lower	Upper	(T_Y)
Residents Aged 15 Years & Over						
Literate in English	2,852.0	5,428	0.2	2,841.3	2,862.6	1.44
Literate in Two or More Languages	2,497.8	5,504	0.2	2,487.0	2,508.6	1.44
Residents Aged 5 Years & Over						
Spoke English Most Frequently at Home	1,735.2	6,195	0.4	1,723.1	1,747.4	1.69
Spoke Mandarin Most Frequently at Home	1,075.2	5,270	0.5	1,064.8	1,085.5	1.66
Spoke Malay Most Frequently at Home	332.3	3,837	1.2	324.7	339.8	1.98
Spoke Tamil Most Frequently at Home	89.9	1,916	2.1	86.2	93.7	1.86
Residents Aged 15 Years & Over						
Buddhism	1,074.2	5,130	0.5	1,064.1	1,084.2	1.61
Taoism^	304.0	2,938	1.0	298.2	309.7	1.58
Christianity	654.4	3,948	0.6	646.6	662.1	1.51
Islam	539.3	4,669	0.9	530.1	548.4	1.94
Hinduism	173.0	2,474	1.4	168.1	177.8	1.74
Other Religions	21.9	833	3.8	20.2	23.5	1.62
No Religion	692.5	3,861	0.6	685.0	700.1	1.44
Resident Households						
With 1 Person	220.3	1,591	0.7	217.2	223.4	1.05
With 2 Persons	309.8	1,747	0.6	306.3	313.2	1.00
With 3 Persons	280.5	1,674	0.6	277.3	283.8	1.00
With 4 Persons	275.6	1,653	0.6	272.4	278.8	0.99
With 5 Persons	163.5	1,334	0.8	160.9	166.1	1.00
With 6 or more Persons	122.9	1,171	1.0	120.6	125.2	1.00
Employed Residents Aged 15 Years & Over						
Travelled to Work by Public Bus Only	325.8	2142	0.7	321.6	330.0	1.12
Travelled to Work by MRT/LRT Only	287.4	2055	0.7	283.4	291.5	1.14
Travelled to Work by MRT/LRT and Public Bus Only	559.3	2861	0.5	553.7	564.9	1.17
Travelled to Work by Car Only	459.8	2461	0.5	455.0	464.6	1.10
Residents Aged 5 Years & Over						
Unable to Perform/ with A Lot of Difficulty in At Least One Basic Activity	97.6	1145	1.2	95.4	99.9	1.06

^ 'Taoism' includes Chinese Traditional Beliefs.

Table A2 Sampling Errors for Square Root of Design Effect (DEFT) Equals 1

Size of Estimates	Proportion of Total Population (%)	Standard Error	Relative Standard Error (%)	95% Confidence Interval	
				Lower	Upper
PERSONS					
4,000,000	82.65	2,774	0.1	3,994,563	4,005,437
3,500,000	72.32	3,278	0.1	3,493,576	3,506,424
3,000,000	61.99	3,556	0.1	2,993,030	3,006,970
2,500,000	51.66	3,661	0.1	2,492,825	2,507,175
2,000,000	41.33	3,607	0.2	1,992,930	2,007,070
1,500,000	31.00	3,388	0.2	1,493,360	1,506,640
1,000,000	20.66	2,966	0.3	994,186	1,005,814
750,000	15.50	2,651	0.4	744,804	755,196
500,000	10.33	2,230	0.4	495,630	504,370
250,000	5.17	1,621	0.6	246,822	253,178
100,000	2.07	1,042	1.0	97,957	102,043
75,000	1.55	905	1.2	73,226	76,774
50,000	1.03	741	1.5	48,548	51,452
25,000	0.52	525	2.1	23,971	26,029
10,000	0.21	333	3.3	9,348	10,652
7,500	0.15	288	3.8	6,935	8,065
5,000	0.10	235	4.7	4,539	5,461
2,500	0.05	166	6.7	2,174	2,826
1,000	0.02	105	10.5	794	1,206
500	0.01	74	14.9	354	646
HOUSEHOLDS					
1,150,000	69.71	1,962	0.2	1,146,154	1,153,846
850,000	51.52	2,134	0.3	845,817	854,183
550,000	33.34	2,013	0.4	546,054	553,946
250,000	15.15	1,531	0.6	246,999	253,001
100,000	6.06	1,019	1.0	98,003	101,997
75,000	4.55	890	1.2	73,257	76,743
50,000	3.03	732	1.5	48,565	51,435
25,000	1.52	522	2.1	23,978	26,022
10,000	0.61	331	3.3	9,350	10,650
7,500	0.45	287	3.8	6,937	8,063
5,000	0.30	235	4.7	4,540	5,460
2,500	0.15	166	6.6	2,174	2,826
1,000	0.06	105	10.5	794	1,206
500	0.03	74	14.9	354	646

Table A3 Sampling Errors for Square Root of Design Effect (DEFT) Equals 2

Size of Estimates	Proportion of Total Population (%)	Standard Error	Relative Standard Error (%)	95% Confidence Interval	
				Lower	Upper
PERSONS					
4,000,000	82.65	5,548	0.1	3,989,126	4,010,874
3,500,000	72.32	6,555	0.2	3,487,152	3,512,848
3,000,000	61.99	7,112	0.2	2,986,060	3,013,940
2,500,000	51.66	7,322	0.3	2,485,649	2,514,351
2,000,000	41.33	7,215	0.4	1,985,859	2,014,141
1,500,000	31.00	6,776	0.5	1,486,719	1,513,281
1,000,000	20.66	5,932	0.6	988,373	1,011,627
750,000	15.50	5,302	0.7	739,608	760,392
500,000	10.33	4,460	0.9	491,259	508,741
250,000	5.17	3,243	1.3	243,644	256,356
100,000	2.07	2,084	2.1	95,915	104,085
75,000	1.55	1,810	2.4	71,453	78,547
50,000	1.03	1,482	3.0	47,096	52,904
25,000	0.52	1,050	4.2	22,941	27,059
10,000	0.21	665	6.7	8,696	11,304
7,500	0.15	576	7.7	6,370	8,630
5,000	0.10	471	9.4	4,077	5,923
2,500	0.05	333	13.3	1,847	3,153
1,000	0.02	211	21.1	587	1,413
500	0.01	149	29.8	208	792
HOUSEHOLDS					
1,150,000	69.71	3,924	0.3	1,142,308	1,157,692
850,000	51.52	4,268	0.5	841,634	858,366
550,000	33.34	4,026	0.7	542,109	557,891
250,000	15.15	3,062	1.2	243,998	256,002
100,000	6.06	2,038	2.0	96,006	103,994
75,000	4.55	1,779	2.4	71,513	78,487
50,000	3.03	1,464	2.9	47,130	52,870
25,000	1.52	1,043	4.2	22,955	27,045
10,000	0.61	663	6.6	8,701	11,299
7,500	0.45	575	7.7	6,374	8,626
5,000	0.30	469	9.4	4,080	5,920
2,500	0.15	332	13.3	1,849	3,151
1,000	0.06	210	21.0	588	1,412
500	0.03	149	29.7	209	791

Simple Guide on Using Relative Standard Error and Confidence Interval

To compute the 95% confidence interval of an estimate of persons with size 1,100,000 as below.

Step 1:   From Table A2, the RSE of this estimate is close to 0.3%

Step 2:   To compute 95% confidence interval of the estimate

Lower Confidence Interval = 1,100,000 – 1.96 x 0.3% x 1,100,000 = 1,093,532

Upper Confidence Interval = 1,100,000 + 1.96 x 0.3% x 1,100,000 = 1,106,468

95% Confidence Interval = (1,093,532 , 1,106,468)

There is a 95% chance that the 'true' population value is between 1,093,532 and 1,106,468.