STATISTICAL RELEASE

P0301.5

Post-enumeration Survey (PES) 2022

Embargoed until:
10 October 2023
15:30

ENQUIRIES**:**
User Information Services
012 310 8600/8390

Dipalopalo tsa Aforikaborwa • Dipalopalo tsa Aforika Borwa • Ezazibalo zaseNingizimu Afrika • Tshitatistika Afrika Tshipembe • Tinhlayo Afrika-Dzonga

Statistieke Suid-Afrika • Dipalopalo tša Aforika Borwa • Telubalo zaseNingizimu Afrika • EzeeNkcukacha maNani zoMzantsi Afrika • Iimbalobalo zeSewula Afrika

IMPROVING LIVES THROUGH DATA ECOSYSTEMS

Contents

List of tables

List of figures

KEY FINDINGS

A Post-enumeration Survey (PES) is an independent sample survey that is conducted immediately after the completion of census enumeration in order to evaluate the coverage and content errors of the census, in this case Census 2022. The PES 2022 was undertaken shortly after the completion of census enumeration, from June 2022 to August 2022 with a *de facto* reference night on the midnight of 22–23 June 2022. A sample of 840 sub-enumeration areas was selected across South Africa's nine provinces. Coverage errors are a measure of how many persons or households were missed or counted more than once in the census. Content errors indicate the quality of key characteristics in the census. The output of the PES provides a statistically informed basis for adjusting the census count. The key findings of the PES 2022 are:

i. The estimated South African population, including those in collective living quarters, transients and the homeless, is 62,0 million persons on the Census 2022 reference night, midnight of 02–03 February 2022. The estimated final true population from private dwelling units within the scope of the PES 2022 is 61,4 million persons. This estimate is based on the adjustment of Census 2022 data using the adjustment factors derived under the PES 2022. The Census 2022 count of those in-scope for enumeration is 42,3 million persons. The uncorrected census count estimated by PES 2022 is 43,9 million persons. The final net coverage error rate relative to the final true population of 61,4 million persons is thus 31,1%.

ii. The estimated final true population of households within the scope of the PES 2022 is 19,3 million households. This estimate is based on the adjustment of Census 2022 data using the adjustment factors derived under the PES 2022. The Census 2022 population of households that were in-scope for enumeration is 13,4 million. The uncorrected census population for households is 13,8 million. The final net coverage error rate relative to the final true population of 19,3 million households is 30,5%.

iii. With respect to content errors, six variables were tested for consistency in terms of the responses that were recorded in the Census and the PES. The aggregated index of inconsistency was 7,5% for population group, 8,2% for sex, and 13,6% for age group, indicating a high level of agreement. The aggregated index of inconsistency for marital status was 23,0%, relationship to head of household was 34,8%, and country of birth was 42,3%, indicating moderate rates of agreement.

A mixed-mode data collection methodology was implemented to counteract the effects of the COVID-19 pandemic. This was made possible by having integrated, digitally enabled survey processes with a geo-spatial information frame as a base.

Risenga Maluleke
Statistician-General

DETAILED RESULTS

Estimation of preliminary true population – Persons

Figure 1: Estimation of preliminary true population for persons (in millions rounded to two decimals)
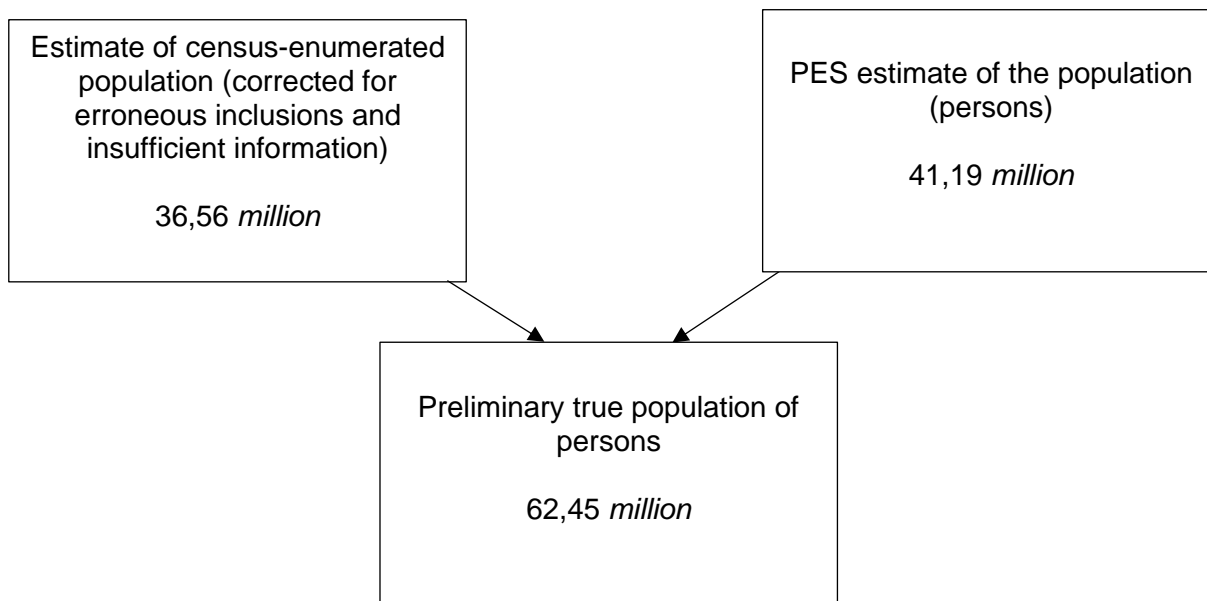
The in-scope Census 2022 and PES 2022 enumerations produced 43,96 million (including erroneous inclusions and insufficient information) and 41,19 million persons, respectively. The Corrected census population (corrected for erroneous inclusions and insufficient information) of 36,56 million persons was used to derive the preliminary true population. The matched population between the census and PES was 24,11 million persons. Using the dual-system estimation method, the in-scope preliminary true population of South Africa was estimated at 62,45 million persons. Four components together make up the dual-system estimate of the preliminary true population and these are shown in Figure 2 below.

Figure 2: Components of preliminary true population for persons (in millions rounded to two decimals)

Components (a), (b), (c) and (d) are obtained through the probabilities of inclusion and omission of a person based on the assumption of independence between the census and PES (Table 10). Component (a), the population included in both the census and the PES, was estimated at 24,11 million persons; component (b), the population included in the census but missed in the PES, was estimated at 12,45 million persons; component (c), the population included in the PES but missed in the census, was estimated at 17,07 million persons; and component (d), the population missed in both the census and the PES, was estimated at 8,81 million persons.

Estimation of preliminary true population – Households

Figure 3: Estimation of preliminary true population for households (in millions rounded to two decimals)
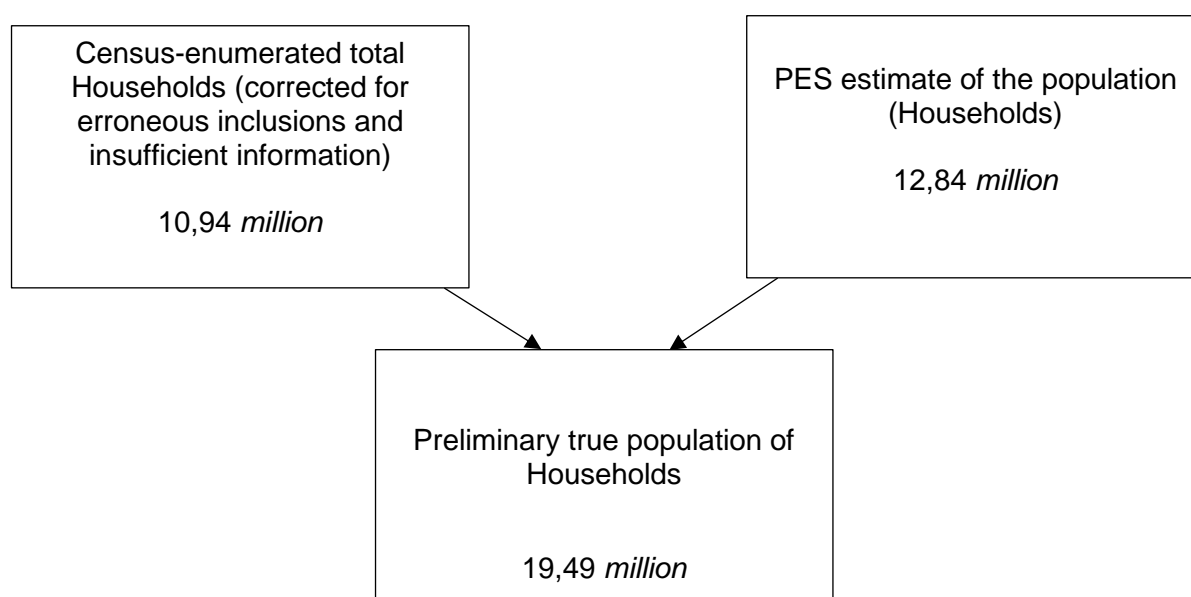
The in-scope Census 2022 and PES 2022 enumerations were estimated at 10,94 million and 12,84 million households, respectively. The matched population was 7,20 million households. Using the dual-system estimation method, the preliminary true population of households was estimated at 19,49 million. Four components together make up the dual-system estimate of the preliminary true population and these are shown in Figure 4 below.

Figure 4: Components of preliminary true population for households (in millions rounded to two decimals)

Preliminary true population of households 19,49 *million*

(a) Households included in both census and PES 7,20 *million*	(b) Households included in census, missed in PES 3,73 *million*	(c) Households included in PES, missed in census 5,63 *million*	(d) Households missed in both census and PES 2,92 *million*

Component (a), the households included in both the census and the PES, was estimated at 7,20 million households; component (b), the households included in the census but missed in the PES, was estimated at 3,73 million households; component (c), the households included in the PES but missed in the census, was estimated at 5,63 million households; and component (d), the households missed in both the census and the PES, was estimated at 2,92 million households.

Estimation of the preliminary net coverage error

The net undercount or overcount is the difference between the estimated true population and the estimated census population. The preliminary net undercount rate is the net undercount expressed as a percentage of the estimated preliminary true population.

Table 1: Preliminary net undercount rate for persons by province and geography type (values expressed in percentage points rounded to two decimals)

| Province | Undercount rate (%) | Standard error | 95% Confidence interval limits | |
			Lower	Upper
Western Cape	35,56	1,55	32,52	38,61
Eastern Cape	30,56	0,59	29,40	31,72
Northern Cape	29,11	2,27	24,67	33,55
Free State	18,52	2,62	13,37	23,66
KwaZulu-Natal	31,19	1,08	29,07	33,31
North West*	17,92	5,39	7,36	28,48
Gauteng	30,89	2,07	26,85	34,94
Mpumalanga	34,00	4,51	25,15	42,85
Limpopo	23,77	1,02	21,77	25,77
South Africa	**29,60**	**0,82**	**28,00**	**31,20**

| Geography type | Undercount rate (%) | Standard error | 95% Confidence interval limits | |
			Lower	Upper
Urban	31,23	1,03	29,21	33,25
Traditional	24,10	0,96	22,21	25,99
Farm*	46,86	12,92	21,53	72,18
South Africa	**29,60**	**0,82**	**28,00**	**31,20**

*The standard errors and resultant coefficients of variation for the estimated undercount rate for North West and Farm geography type indicate high levels of variation associated with the estimates.

In Table 1, it can be observed that the preliminary net undercount rate at the national level was estimated at 29,60%, with confidence intervals ranging from 28,00% to 31,20%. Among the provinces, the highest preliminary net undercount rate was observed in Western Cape with 35,56%, followed by Mpumalanga at 34,00%. North West had the lowest preliminary net undercount rate at 17,92%, followed by the Free State at 18,52%. Farms had the highest preliminary net undercount rate at 46,86% while Traditional areas had the lowest preliminary net undercount rate at 24,10%.

Table 2: Preliminary net undercount rate for households by province and geography type (values expressed in percentage points rounded to two decimals)

Province	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Western Cape	35,89	1,47	33,01	38,76
Eastern Cape	25,60	0,94	23,76	27,44
Northern Cape	33,59	2,08	29,50	37,67
Free State	16,65	2,87	11,03	22,27
KwaZulu-Natal	26,80	1,65	23,56	30,04
North West*	16,84	5,98	5,11	28,56
Gauteng	35,19	2,00	31,26	39,11
Mpumalanga*	34,05	9,31	15,79	52,31
Limpopo	19,23	1,56	16,17	22,29
South Africa	**29,37**	**1,17**	**27,08**	**31,67**

Geography type	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Urban	32,85	1,11	30,66	35,03
Traditional	14,78	1,24	12,34	17,21
Farm*	49,31	22,36	5,49	93,13
South Africa	**29,37**	**1,17**	**27,08**	**31,67**

*The standard errors and resultant coefficients of variation for the estimated undercount rate for North West and Farm geography type indicate high levels of variation associated with the estimates.

In Table 2, it can be observed that the preliminary net undercount rate for households at the national level was estimated at 29,37%, with confidence intervals ranging from 27,08% to 31,67%. Among the provinces, the highest preliminary net undercount rate was observed in Western Cape with 35,89%, followed by Gauteng at 35,19%. Free State had the lowest preliminary net undercount rate at 16,65%. Farms had the highest preliminary net undercount rate at 49,31% while Traditional areas had the lowest preliminary net undercount rate at 14,78%.

Estimation of final true population – Persons

Figure 5: Estimation of final true population for persons (in millions rounded to two decimals)

```
┌─────────────────────────┐            ┌─────────────────────────┐
│  PES adjustment factors  │            │ Census-enumerated persons │
│                          │            │                          │
│    By adjustment class   │            │      42,31 million       │
│                          │            │                          │
└─────────────────────────┘            └─────────────────────────┘
                    ┌─────────────────────────┐
                    │                          │
                    │ Final true population of │
                    │        persons           │
                    │                          │
                    │      61,37 million       │
                    │                          │
                    └─────────────────────────┘
```
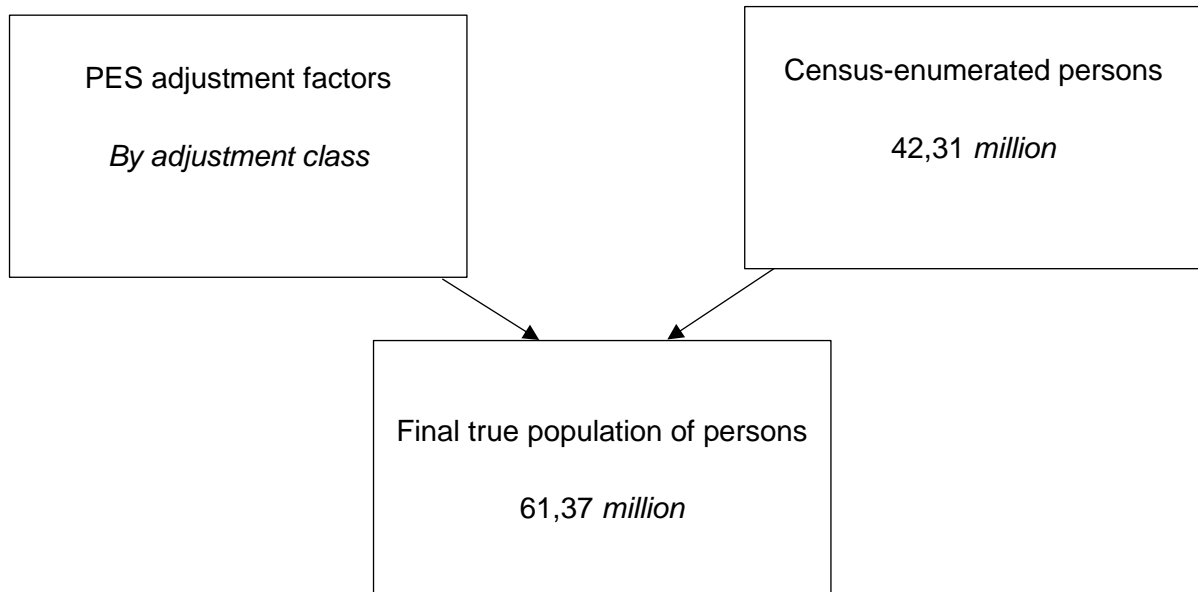
The in-scope final true population for persons was calculated using the PES adjustment factors and the full census-enumerated dataset for persons. Census 2022 enumeration produced 42,31 million persons. The in-scope final true population was estimated at 61,37 million persons. Four components together make up the estimate of the final true population and these are shown in Figure 6.

Figure 6: Components of final true population for persons (in millions rounded to two decimals)
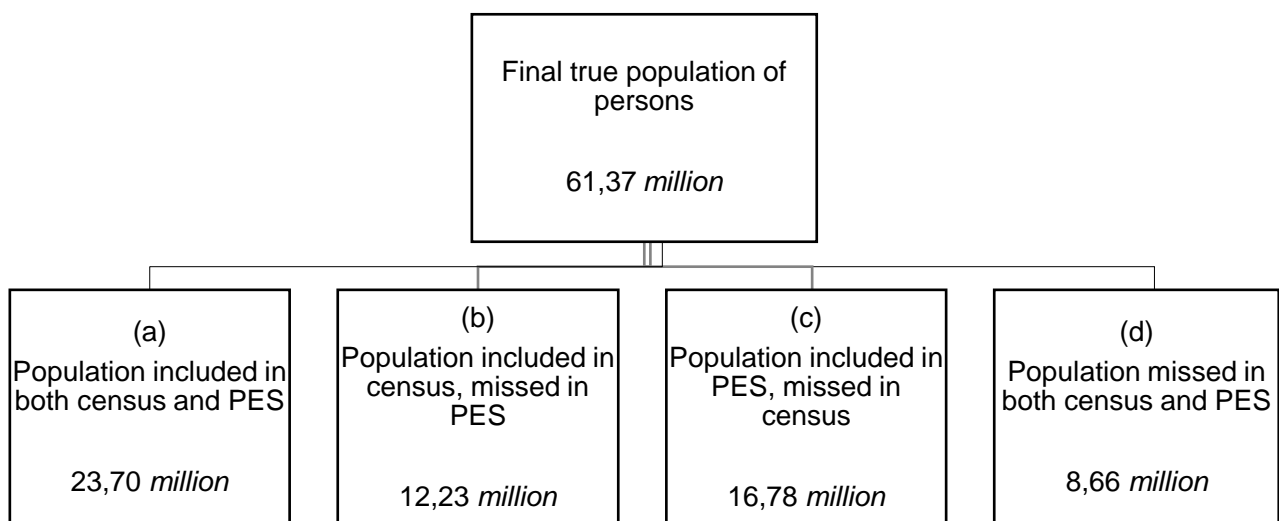
```
                    ┌─────────────────────────┐
                    │ Final true population of │
                    │        persons           │
                    │                          │
                    │      61,37 million       │
                    └─────────────────────────┘
```
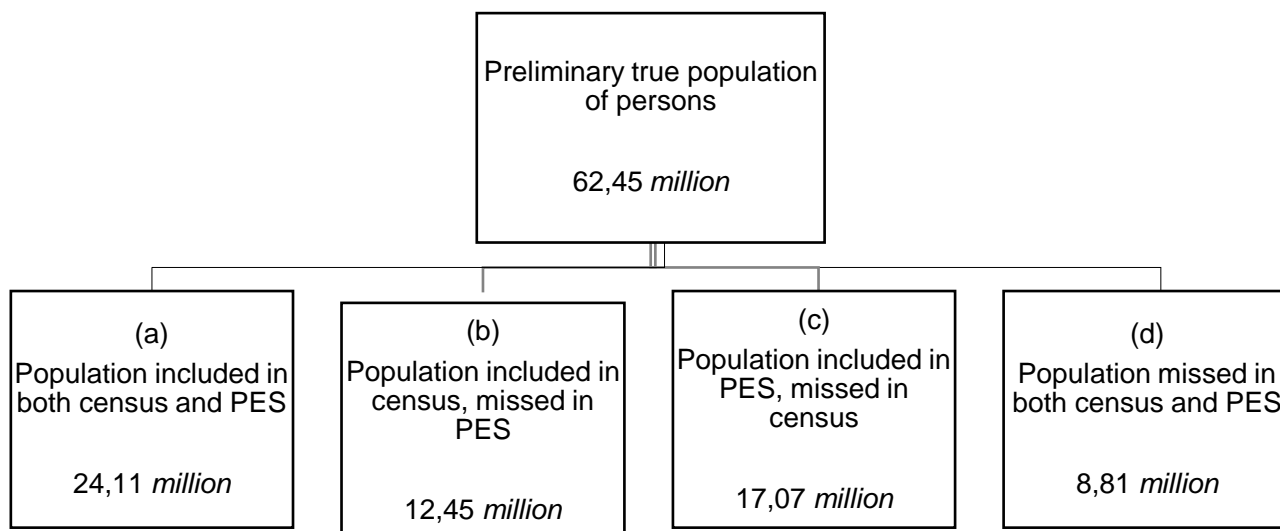
(a) Population included in both census and PES	(b) Population included in census, missed in PES	(c) Population included in PES, missed in census	(d) Population missed in both census and PES
23,70 million	12,23 million	16,78 million	8,66 million

Components (a), (b), (c) and (d) are obtained through the probabilities of inclusion and omission of a person based on the assumption of independence between the Census and PES (Table 10). Component (a), the population included in both the census and the PES, was estimated at 23,70 million persons; component (b), the population included in the census but missed in the PES, was estimated at 12,23 million persons; component (c), the population included in the PES but missed in the census, was estimated at 16,78 million persons; and component (d), the population missed in both the census and the PES, was estimated at 8,66 million persons.

Estimation of final true population – Households

Figure 7: Estimation of final true population for households (in millions rounded to two decimals)
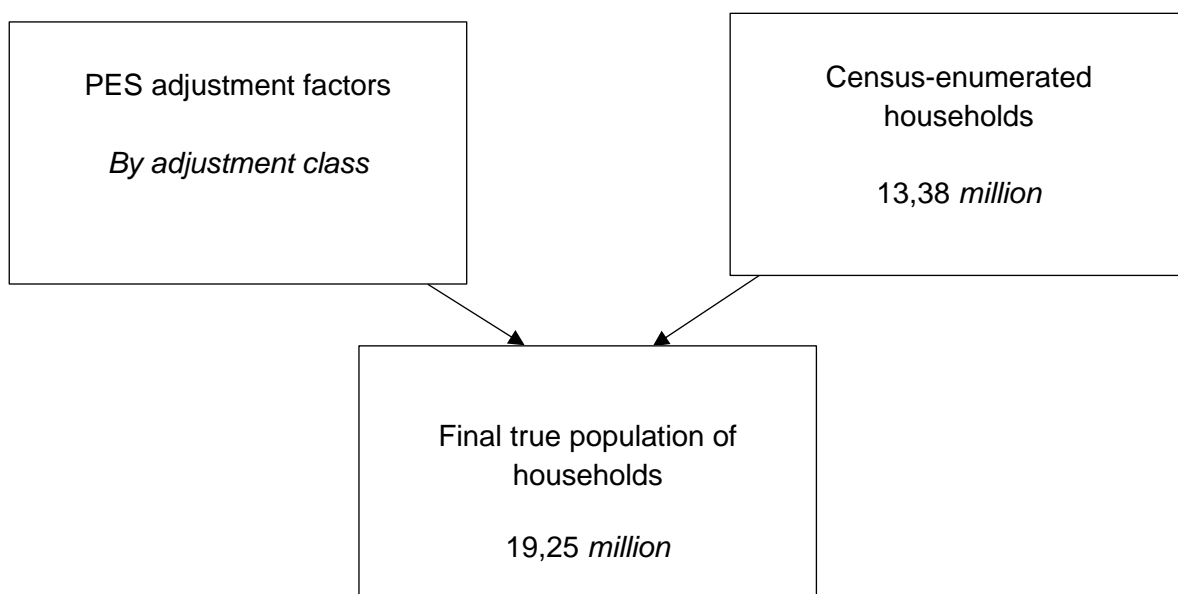
The in-scope final true population was calculated using the PES adjustment factors and the full census-enumerated dataset for households. Census 2022 enumeration produced 13,38 million households. The in-scope final true population was estimated at 19,25 million households. Four components together make up the estimate of the final true population of households and these are shown in Figure 8 below.

Figure 8: Components of final true population for households (in millions rounded to two decimals)

```
                    ┌─────────────────────┐
                    │ Final true population of │
                    │     Households      │
                    │                     │
                    │    19,25 million    │
                    └─────────────────────┘
        ┌──────────────┬──────────┴──────────┬──────────────┐
┌───────────────┐ ┌───────────────┐ ┌───────────────┐ ┌───────────────┐
│      (a)      │ │      (b)      │ │      (c)      │ │      (d)      │
│Households included│ │Households included│ │Households included│ │Households missed│
│in both census and│ │in census, missed in│ │in PES, missed in│ │in both census and│
│      PES      │ │      PES      │ │    census     │ │      PES      │
│               │ │               │ │               │ │               │
│  7,12 million │ │  3,69 million │ │  5,57 million │ │  2,88 million │
└───────────────┘ └───────────────┘ └───────────────┘ └───────────────┘
```
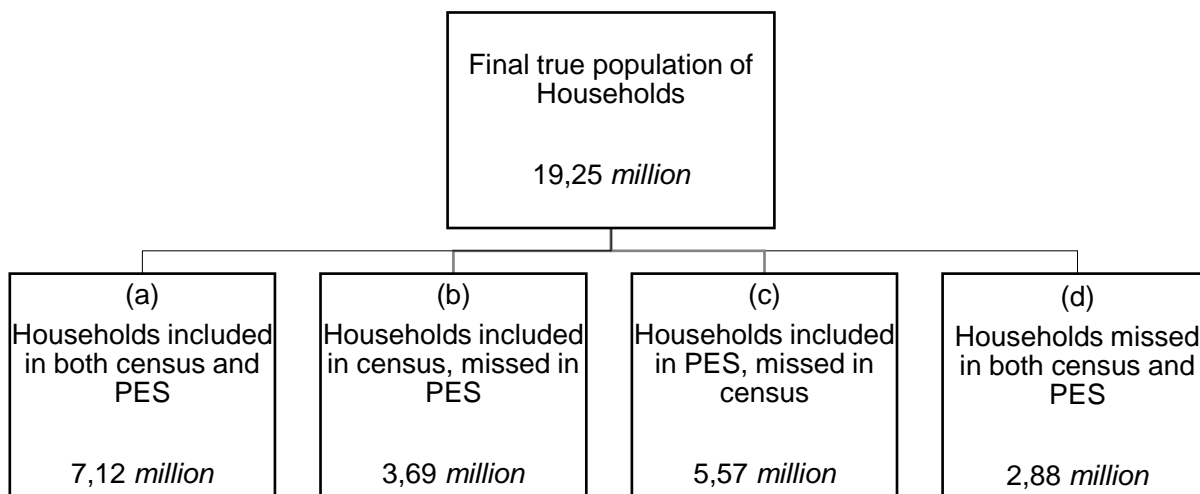
Component (a), the households included in both the census and the PES, was estimated at 7,12 million households; component (b), the households included in the census but missed in the PES, was estimated at 3,69 million households; component (c), the households included in the PES but missed in the census, was estimated at 5,57 million households; and component (d), the households missed in both the census and the PES, was estimated at 2,88 million households.
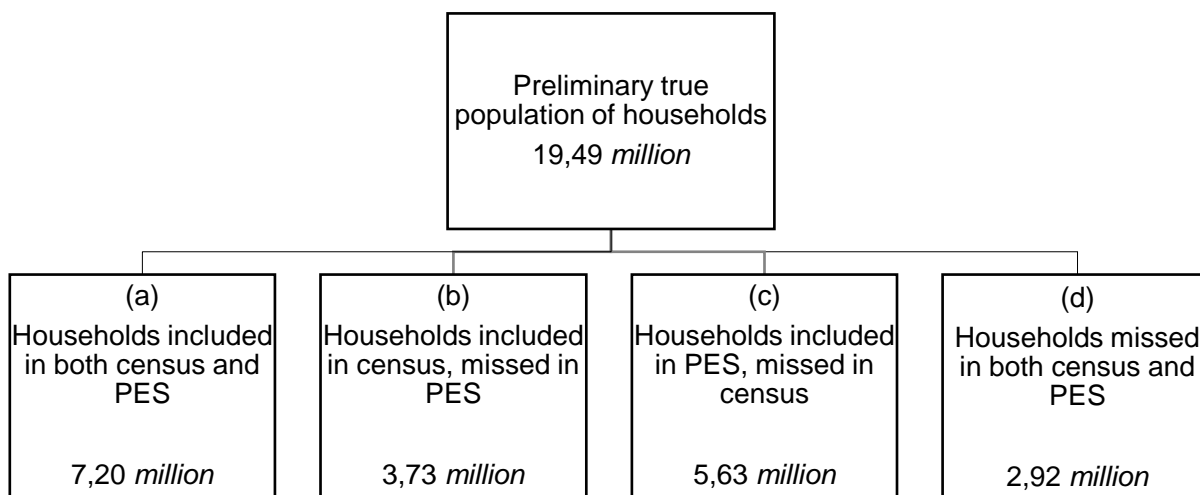
Estimation of the final net coverage error

Table 3: Final net undercount rate for persons, by province and geography type

Province	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Western Cape	35,58	0,02	35,56	35,61
Eastern Cape	32,40	0,01	32,37	32,42
Northern Cape	33,41	0,03	33,36	33,46
Free State	20,95	0,02	20,92	20,99
KwaZulu-Natal	34,67	0,01	34,65	34,69
North West	23,33	0,02	23,30	23,36
Gauteng	30,79	0,01	30,77	30,81
Mpumalanga	35,26	0,02	35,24	35,29
Limpopo	23,53	0,01	23,50	23,55
South Africa	**31,06**	**0,01**	**31,05**	**31,07**

Geography type	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Urban	33,18	0,01	33,17	33,19
Traditional	25,72	0,01	25,71	25,73
Farm	41,78	0,03	41,73	41,83
South Africa	**31,06**	**0,01**	**31,05**	**31,07**

In Table 3, the final net undercount rate at the national level was estimated at 31,06%, with confidence intervals ranging from 31,05% to 31,07%. Among the provinces, the highest undercount rate was observed in Western Cape with 35,58%, followed by Mpumalanga at 35,26%. Free State had the lowest final net undercount rate at 20,95%. Farms had the highest final net undercount rate at 41,78% while Traditional areas had the lowest final net undercount rate at 25,72%.

Table 4: Final net undercount rate for households, by province and geography type

Province	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Western Cape	36,30	0,00	36,30	36,30
Eastern Cape	27,24	0,02	27,20	27,27
Northern Cape	27,14	0,02	27,10	27,18
Free State	17,93	0,01	17,91	17,95
KwaZulu-Natal	27,93	0,01	27,92	27,94
North West	23,05	0,01	23,04	23,07
Gauteng	35,82	0,00	35,81	35,82
Mpumalanga	32,00	0,01	31,98	32,02
Limpopo	24,83	0,03	24,79	24,87
South Africa	**30,49**	**0,00**	**30,49**	**30,50**
Geography type	**Undercount rate (%)**	**Standard error**	**95% Confidence interval limits**	
			Lower	Upper
Urban	33,96	0,00	33,95	33,96
Traditional	15,27	0,00	15,27	15,28
Farm	55,63	0,03	55,59	55,68
South Africa	**30,49**	**0,00**	**30,49**	**30,50**

In Table 4, the final net undercount rate for households at the national level was estimated at 30,49%, with confidence intervals ranging from 30,49% to 30,50%. Among the provinces, the highest undercount was observed in Western Cape with 36,30%, followed by Gauteng at 35,82%. Free State had the lowest undercount rate at 17,93%. Farms had the highest undercount rate at 55,63% while Traditional areas had the lowest undercount rate at 15,27%.

Estimating the total South African population

Figure 9: Total South African population for persons (in millions rounded to two decimals)

```
┌─────────────────────────────┐        ┌─────────────────────────────┐
│ Final true population of    │        │ Census count of persons in  │
│ persons                     │        │ Other collective living     │
│                             │        │ quarters                    │
│ 61,37 million               │        │                             │
│                             │        │ 0,66 million                │
└─────────────────────────────┘        └─────────────────────────────┘
              ↘                                    ↙
              ┌─────────────────────────────┐
              │ Total South African         │
              │ population                  │
              │                             │
              │ 62,03 million               │
              └─────────────────────────────┘
```
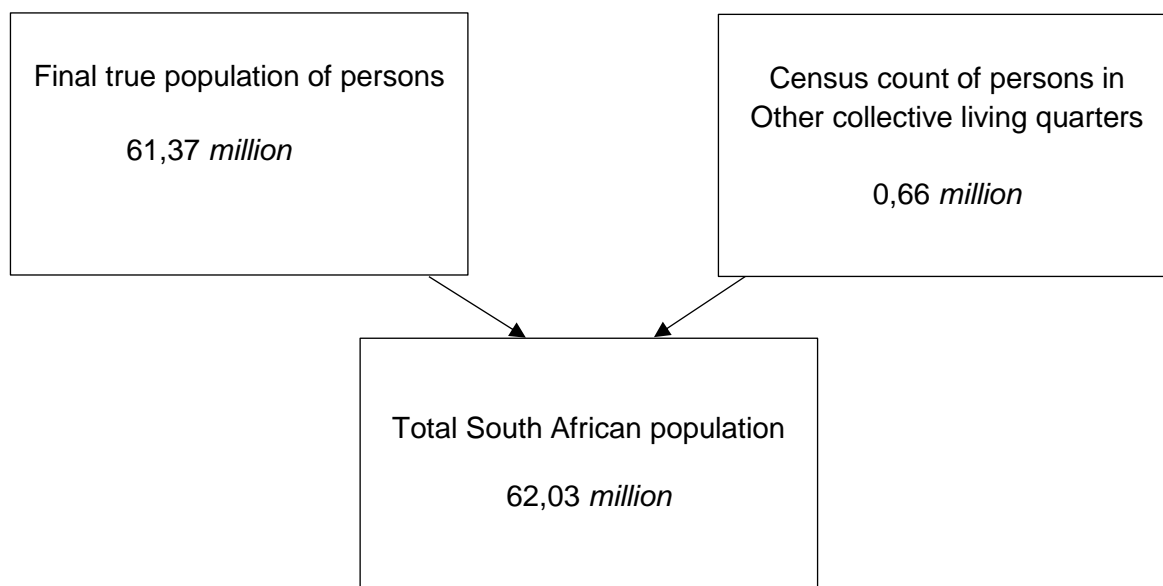
Figure 9 shows that the total South African population of 62,03 million persons was calculated by adding the number of census-enumerated persons in the "Other collective living quarters" to the in-scope final true population of 61,37 million persons.

Table 5: Total South African population of persons by province (in millions rounded to two decimals)

Province	Estimate (million)	Standard error	95% Confidence Interval	
			Lower	Upper
Western Cape	7,43	0,07	7,31	7,55
Eastern Cape	7,23	0,06	7,13	7,33
Northern Cape	1,35	0,03	1,30	1,40
Free State	2,96	0,04	2,89	3,02
KwaZulu-Natal	12,39	0,10	12,24	12,55
North West	3,80	0,04	3,73	3,87
Gauteng	15,12	0,09	14,97	15,27
Mpumalanga	5,16	0,06	5,06	5,26
Limpopo	6,57	0,06	6,47	6,67
South Africa	**62,03**	**0,12**	**61,84**	**62,22**

Table 5 shows the adjusted total population by province and the corresponding confidence intervals, which reflect the sampling error around the estimate. The confidence interval was obtained by adding the census-enumerated persons in "Other collective living quarters" to the lower and upper limits of the confidence interval for the adjusted in-scope population.

Figure 10: Total South African population of households (in millions rounded to two decimals)
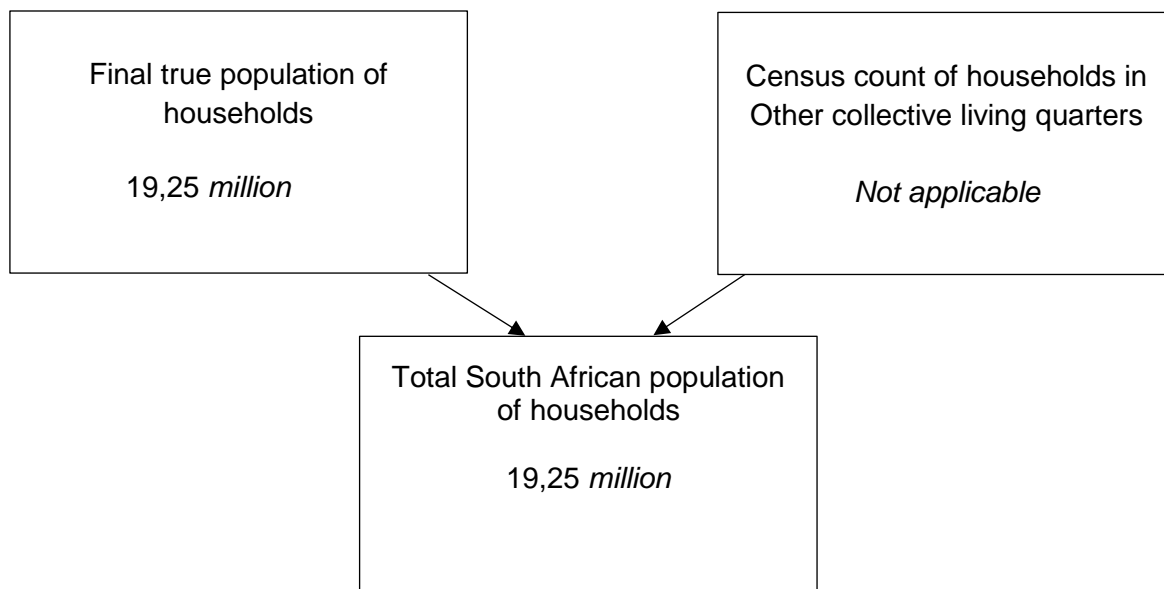
Figure 10 shows that the adjusted household total is the same as the final true population, which was estimated at 19,25 million. The population in "Other collective living quarters" is not linked to households and dwelling units.

Table 6: Total South African population of households by province and geography type (in millions rounded to nearest thousand)

Province	Estimate (million)	Standard error	95% Confidence interval limits	
			Lower	Upper
Western Cape	2,46	0,03	2,42	2,51
Eastern Cape	2,08	0,02	2,05	2,12
Northern Cape	0,36	0,01	0,34	0,37
Free State	0,89	0,01	0,87	0,91
KwaZulu-Natal	3,09	0,03	3,05	3,13
North West	1,22	0,01	1,19	1,24
Gauteng	5,74	0,04	5,67	5,81
Mpumalanga	1,49	0,02	1,46	1,52
Limpopo	1,92	0,02	1,89	1,96
South Africa	**19,25**	**0,05**	**19,17**	**19,33**

Geography type	Estimate (million)	Standard error	95% Confidence interval limits	
			Lower	Upper
Urban	13,27	0,05	13,19	13,36
Traditional	4,87	0,03	4,82	4,91
Farm	1,12	0,02	1,08	1,15
South Africa	**19,25**	**0,05**	**19,17**	**19,33**

Table 6 shows the adjusted total number of households by province and geography type, and the corresponding confidence intervals, which reflect the sampling error around the estimate. This estimate includes households in dwelling units only in the in-scope EAs and the out-of-scope EAs; however, "Other collective living quarters" (i.e. excludes hostels and other collective living quarters that do not contain any households). The confidence interval for this estimate was obtained by adding the unadjusted households in out-of-scope housing units to the lower and upper limits of the confidence interval for the adjusted household total in in-scope housing units.

The adjusted total number of households was estimated at 19,25 million. The highest adjusted number of households was observed in Gauteng at 5,74 million while the lowest was observed in Northern Cape at 0,36 million. Urban areas had the highest adjusted number of households at 13,27 million and Farms had the lowest at 1,12 million.

Content error analysis

Table 7 shows that Population group, Sex, and Age group have a low level of variability between the census and the PES, indicating that they were measured reliably between the census and PES. The variables Marital status, Relationship to head of household, and Country of birth show a moderate level of variability. Moderate and high levels of variability might be indicative of a need for clearer concepts and definitions, wording, and more probing for the questionnaire.

Table 7: Content error indices for demographic characteristics

Demographic characteristic	Aggregated index of inconsistency (%)	Interpretation
Population group	7,52	Low
Sex	8,19	Low
Age group	13,61	Low
Marital status	22,99	Moderate
Relationship to head of household	34,79	Moderate
Country of birth	42,27	Moderate

Detailed results – Evaluation of coverage error for persons

Estimation of preliminary true population for persons

Table 8: Distribution of uncorrected census population for persons (in millions rounded to two decimals)

Component of Uncorrected Census Population	Census Enumeration (million)
Included in PES	24,11
Omitted from PES	12,45
Erroneous inclusions	1,65
Insufficient information	5,75
Total of Uncorrected Census Population	**43,96**

Table 8 shows that, of the estimated 43,96 million in-scope persons who were counted in the in-census, 36,56 million persons were estimated to be correctly enumerated (the PES included 24,11 million persons and missed 12,45 million persons). The census erroneous inclusions (false enumeration, duplications, and geographic misallocations) were 1,65 million persons, which is approximately 3,75% of the estimated uncorrected census population. The census insufficient information total was estimated to be 5,75 million persons, which is approximately 13,08% of the estimated uncorrected census population.

Table 9: Distribution of preliminary true population for persons (in millions rounded to two decimals)

PES Population	Census Enumeration (million)		
	Included	Omitted	Total
Included	24,11	17,07	41,19
Omitted	12,45	8,81	21,26
Total excluding erroneous inclusions and insufficient information	36,56	25,89	62,45

Table 9 shows that the census omitted an estimated 25,89 million persons in total, of which 17,07 million were correctly enumerated in the PES. There were 8,81 million persons who were missed in both the census and the PES. While the PES estimated the total population at 41,19 million persons, it omitted 12,45 million persons who were correctly enumerated in the census. The PES omitted a total of 21,26 million persons.

Table 10: Probabilities of inclusion and omission of an in-scope person

Probability	Value
Probability of inclusion in census	0,5854
Probability of inclusion in PES	0,6596
Probability of inclusion in both census and PES	0,3861
Probability of inclusion in census, missed in PES	0,1993
Probability of inclusion in PES, missed in census	0,2734
Probability of missed in both census and PES	0,1411

In Table 10, the overall empirical probabilities of inclusion and omission of a person in the census or in the PES are calculated using the outcomes from the matching exercise. According to the results from the matching process, a member of the in-scope population had an approximately 58,54% chance of being enumerated in the census, 65,96% chance of being enumerated in the PES, and 38,61% chance of being enumerated in both. Conversely, the person had approximately 19,93% chance of being included in the census but missed in the PES, 27,34% chance of being included in the PES but missed in the census, and 14,11% chance of being missed in both.

Estimation of the preliminary net coverage error for persons

In Table 1, it can be observed that the preliminary net undercount rate at the national level was estimated at 29,60%, with confidence intervals ranging from 28,00% to 31,20%. When comparing rates for different sets of persons, the confidence intervals must be considered. An overlap in the intervals indicates that, except for a 5% chance of erring in the conclusion, the difference observed is not statistically significant due to random error; in other words, there is no evidence of a real difference. A 'floating bars' chart is useful for visualising the intervals (see Figure 11 below).

Figure 11: Graphic representation of confidence intervals for persons preliminary net undercount rate (%), by province (Lower and Upper limits)
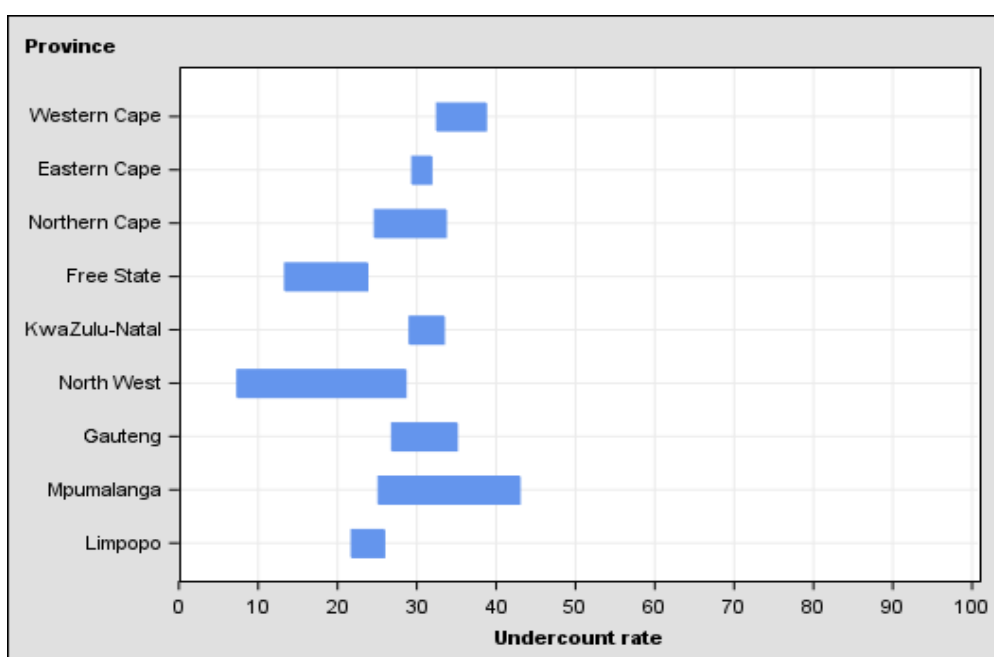
Figure 11 shows that undercount rates in the provinces of Free State (18,52%), North West (17,92%) and Limpopo (23,77%) overlap, i.e. there is no statistically significant difference between them.

Table 11: Preliminary net undercount rate for persons, by demographic groups – single-variable classifications (values expressed in percentage points rounded to two decimals)

Population group	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Black African	36,74	0,17	36,40	37,07
Coloured	44,93	0,43	44,09	45,78
Indian or Asian	72,30	0,82	70,69	73,91
White	61,64	0,45	60,76	62,52
Other	65,05	0,79	63,50	66,60
South Africa	**29,60**	**0,82**	**28,00**	**31,20**

Sex	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Male	33,12	0,91	31,34	34,91
Female	29,80	0,74	28,35	31,26
South Africa	**29,60**	**0,82**	**28,00**	**31,20**

Age group	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Under 5 years	33,24	1,00	31,28	35,19
5–14 years	22,89	1,12	20,70	25,08
15–19 years	28,15	0,64	26,89	29,40
20–29 years	37,79	0,78	36,26	39,33
30–44 years	37,04	1,31	34,48	39,61
45–64 years	19,94	1,50	17,01	22,88
65 or more	28,37	1,07	26,26	30,47
South Africa	**29,60**	**0,82**	**28,00**	**31,20**

Note: The undercount rates for undetermined categories form part of the calculations but are excluded from these tables.

Table 11 shows that the population group Indian or Asian has the highest undercount rate at 72,30%, with black African having the lowest at 36,74%. The undercount rate for males (33,12%) is higher than that for females (29,80%). The age group 20–29 years has the highest undercount rate at 37,79% while the age group 45–64 years had the lowest undercount rate at 19,94%.

Figure 12: Graphic representation of confidence intervals for persons preliminary net undercount rate (%), by Population group (Lower and Upper limits)
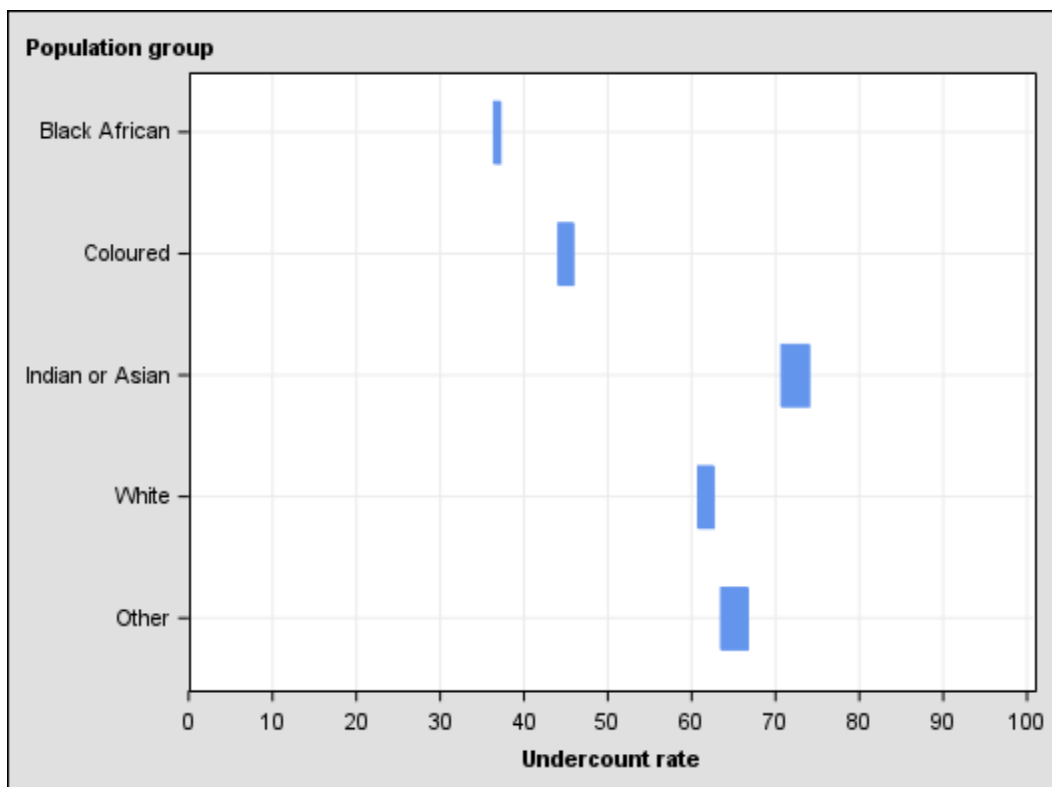
When population groups are compared (see Figure 12 above), the population group Indian or Asian had the highest undercount rate (72,30%), which was statistically different from all other population groups. The population group black African had the lowest undercount rate (36,74%), which was statistically different from all other population groups. The differences in undercount rates between all population groups are statistically significant.

Figure 13: Graphic representation of confidence intervals for persons preliminary net undercount rate (%), by Sex (Lower and Upper limits)
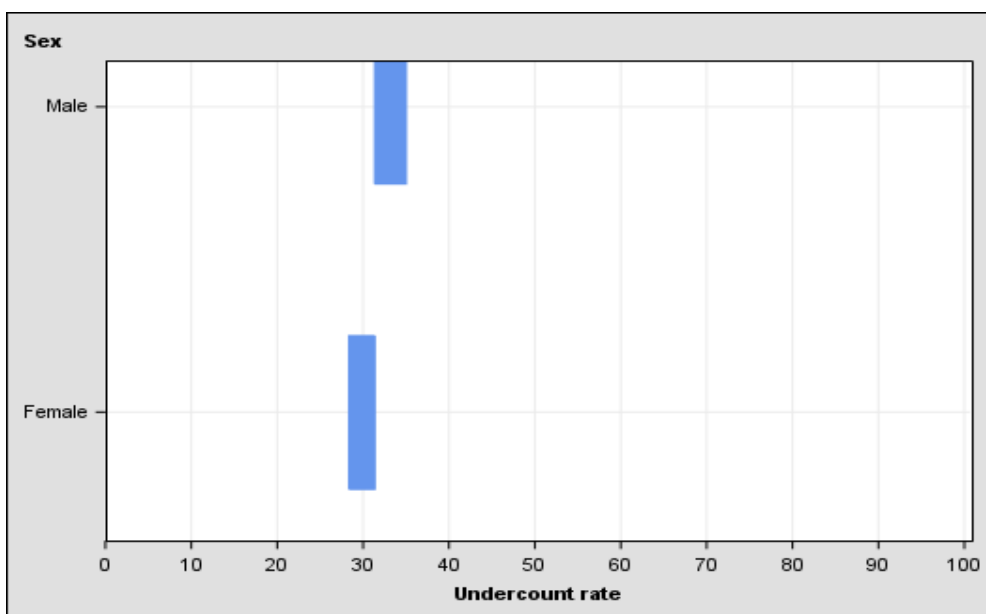
Figure 13 shows that males have a higher undercount rate than females at 33,12% and 29,80% respectively. The difference in undercount rates between males and females is statistically significant.

Figure 14: Graphic representation of confidence intervals for persons preliminary net undercount rate (%), by Age group (Lower and Upper limits)

Figure 14 shows that the difference in net undercount rates for age groups 20–29 years (37,79%) and 30–44 years (37,04%) was not statistically significant. The difference in net undercount rates for age groups 15–19 years (28,15%) and 65 or more (28,37%) was also not statistically significant.

Table 12: Preliminary net undercount rate for in-scope persons by demographic group; two-variable classifications (values expressed in percentage points rounded to two decimals)

Population group and Sex		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
Population group	Sex				
Black African	Male	39,10	0,17	38,76	39,44
	Female	34,60	0,17	34,27	34,94
Coloured	Male	46,41	0,46	45,51	47,32
	Female	43,54	0,41	42,73	44,35
Indian or Asian	Male	73,62	0,79	72,07	75,18
	Female	70,85	0,87	69,15	72,56
White	Male	62,78	0,47	61,87	63,70
	Female	60,56	0,44	59,70	61,42
Other	Male	66,27	0,81	64,68	67,87
	Female	63,08	0,93	61,27	64,90

Population group and Age group		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
Population group	Age group				
Black African	Under 5 years	37,45	0,20	37,06	37,84
	5–14 years	30,47	0,17	30,13	30,81
	15–19 years	32,31	0,19	31,93	32,68
	20–29 years	41,45	0,19	41,07	41,83
	30–44 years	42,46	0,18	42,11	42,82
	45–64 years	32,12	0,18	31,76	32,48
	65 or more	23,53	0,26	23,03	24,04
Coloured	Under 5 years	46,93	0,52	45,92	47,94
	5–14 years	42,91	0,54	41,86	43,96
	15–19 years	44,09	0,53	43,06	45,12
	20–29 years	48,51	0,43	47,66	49,35
	30–44 years	45,38	0,41	44,57	46,19
	45–64 years	41,36	0,44	40,51	42,21
	65 or more	38,75	0,50	37,78	39,72
Indian or Asian	Under 5 years	78,04	1,07	75,95	80,13
	5–14 years	74,95	0,73	73,52	76,37
	15–19 years	70,75	1,03	68,73	72,77
	20–29 years	72,86	1,02	70,86	74,85
	30–44 years	77,52	0,78	75,99	79,04
	45–64 years	67,71	0,85	66,05	69,37
	65 or more	61,85	0,99	59,90	63,80

		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
White	Under 5 years	65,25	0,66	63,95	66,55
	5–14 years	61,03	0,49	60,06	61,99
	15–19 years	59,02	0,77	57,52	60,53
	20–29 years	67,34	0,54	66,28	68,40
	30–44 years	63,45	0,45	62,57	64,34
	45–64 years	60,54	0,45	59,65	61,43
	65 or more	55,81	0,57	54,70	56,93
Other	Under 5 years	67,37	0,99	65,43	69,32
	5–14 years	53,68	1,14	51,45	55,91
	15–19 years	55,32	1,77	51,86	58,78
	20–29 years	76,35	1,14	74,12	78,57
	30–44 years	65,59	1,05	63,53	67,65
	45–64 years	45,91	0,97	44,00	47,82
	65 or more	60,96	0,72	59,55	62,37

Sex and Age group		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
Sex	**Age group**				
Male	Under 5 years	32,70	1,06	30,63	34,78
	5–14 years	22,78	1,16	20,50	25,06
	15–19 years	28,28	0,64	27,02	29,53
	20–29 years	39,19	0,65	37,91	40,46
	30–44 years	41,36	1,45	38,51	44,21
	45–64 years	22,05	1,90	18,33	25,77
	65 or more	30,46	1,25	28,01	32,91
Female	Under 5 years	33,58	0,97	31,68	35,47
	5–14 years	22,79	1,09	20,65	24,93
	15–19 years	27,82	0,70	26,44	29,20
	20–29 years	36,39	1,08	34,28	38,50
	30–44 years	33,10	1,54	30,09	36,12
	45–64 years	18,55	1,34	15,92	21,18
	65 or more	27,17	0,99	25,22	29,12

Note: The undercount rates for undetermined categories form part of the calculations but are excluded from these tables.

Table 12 and Figures 15 to 17 show different comparisons based on population group by sex, population group by age group, or sex by age group. For example, Figure 15 shows that the difference in net undercount rates between males and females for population groups Indian or Asian and Other is not statistically significant. Table 12 shows the highest net undercount rate was observed among the Indian or Asian population group in the age group under 5 years at 78,04%, while the lowest net undercount rate was 18,55% for females in the age group 45–64 years.

Figure 15: Graphic representation of confidence intervals for persons preliminary net undercount rate, by Population group and Sex (Lower and Upper limits)
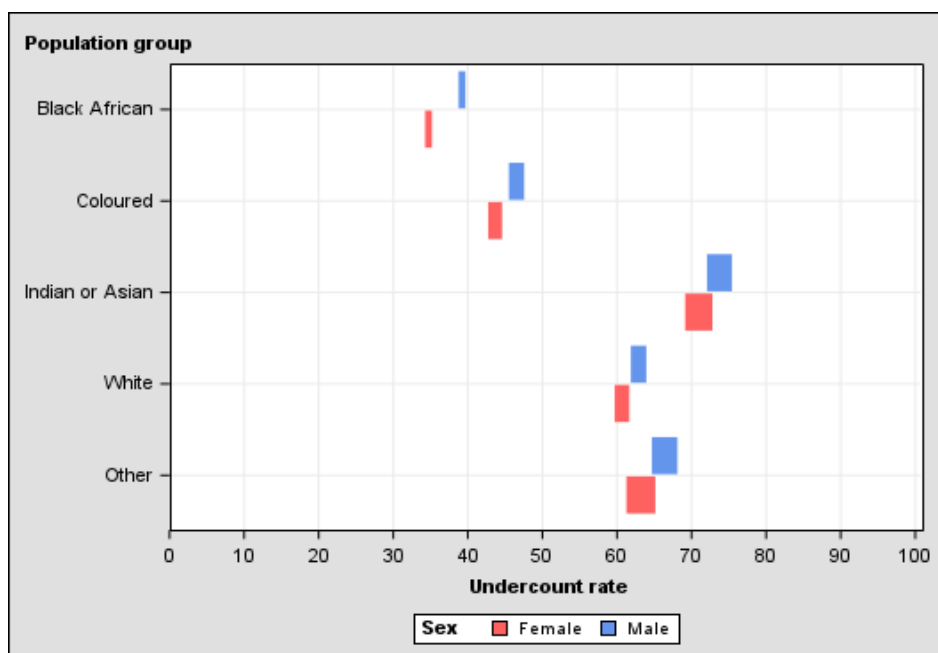
Figure 15 shows that females in the population group black African had the lowest net undercount rate (34,60%), which was statistically different from all other population groups by sex. There is no statistically significant difference in the net undercount rates for males and females in the population group Indian or Asian and for Other. Males (73,62%) and females (70,85%) in the population group Indian or Asian had the two highest net undercount rates, which were statistically different from all other population groups by sex.

Figure 16: Graphic representation of confidence intervals for persons preliminary net undercount rate (%), by Population group and Age group (Lower and Upper limits)
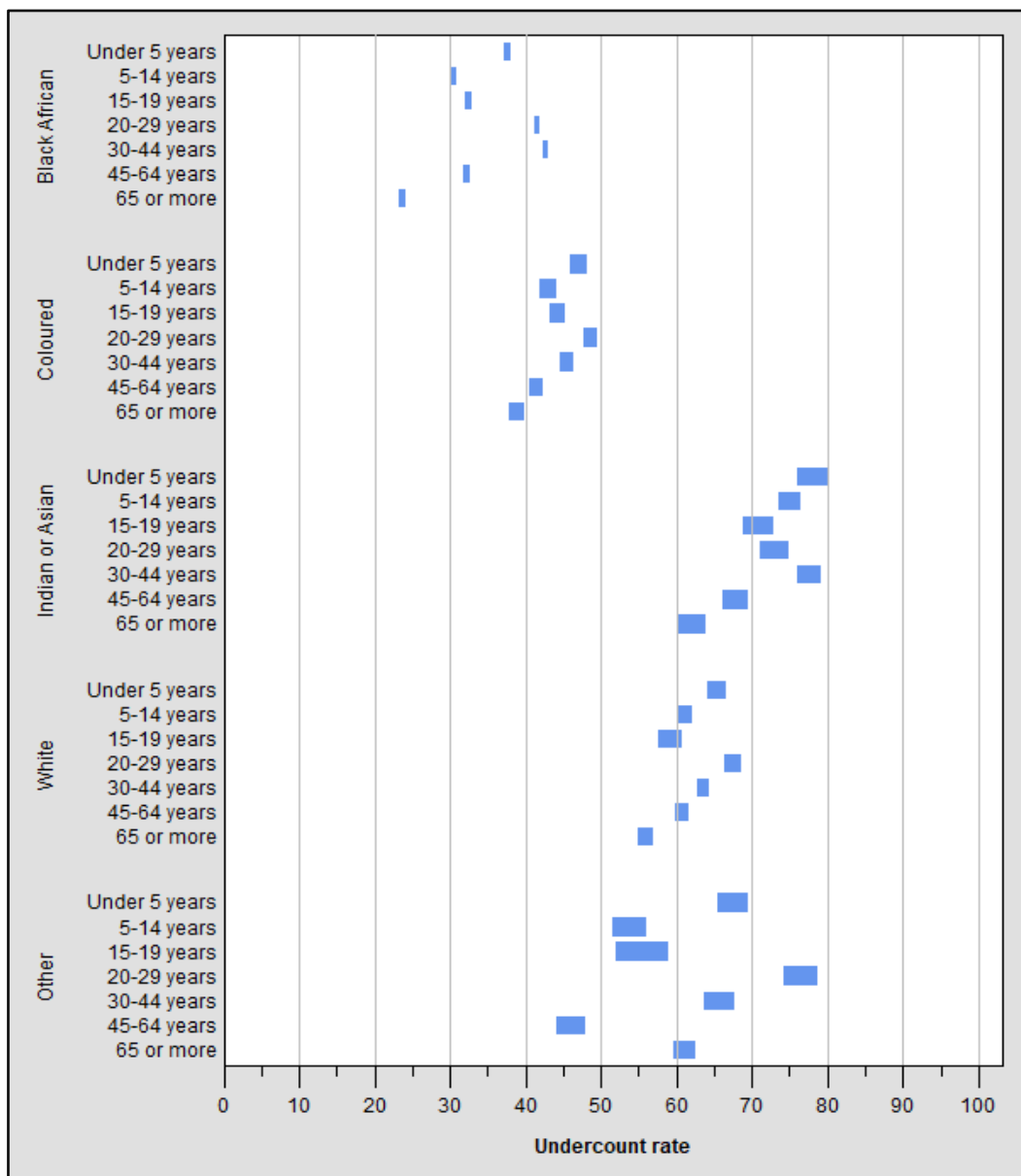
Figure 16 shows that the net undercount rate for the age group 65 or more years for the black African population group (23,53%) is lower than all other age groups by population group and age group, and the difference is statistically significant. There is no statistically significant difference between Indian or Asian within the age group under 5 years (78,04%), Indian or Asian within the age group 30–44 years (77,52%), and Other within the age group 20–29 years (76,35%).

Figure 17: Graphic representation of confidence intervals for persons preliminary net undercount rate (%), by Sex and Age group
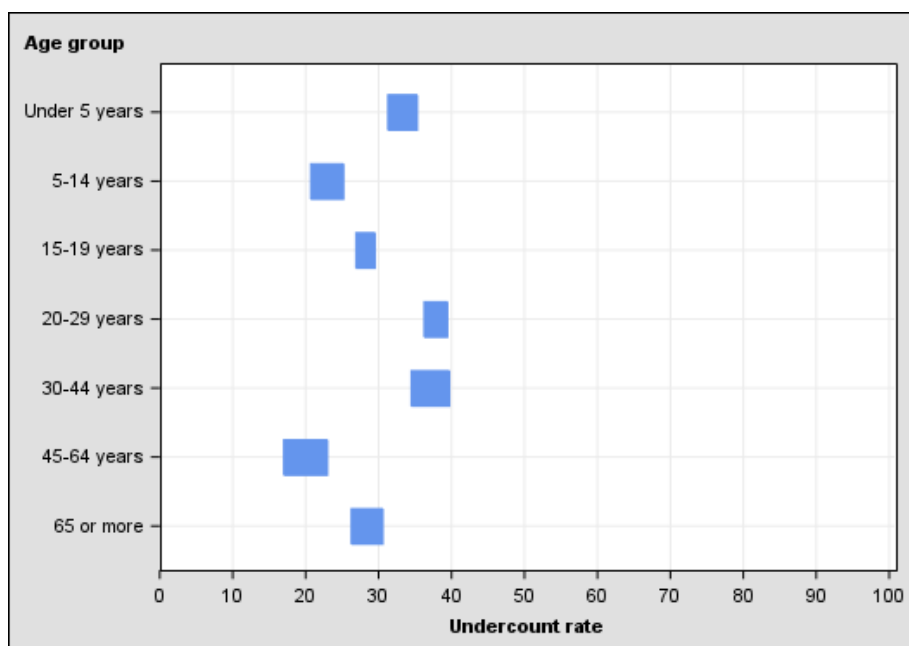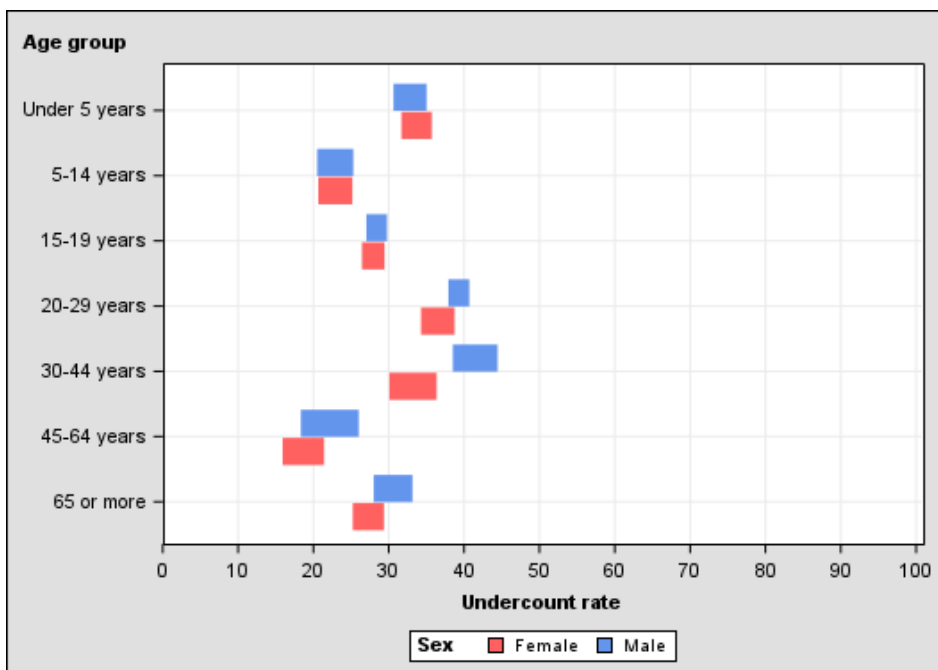

Figure 17 shows that the net undercount rate for females in the age group 30–40 years is lower than males in the same age group and the difference is statistically significant. There is no statistically significant difference in net undercount rates for males and females in the age groups under 5 years, 5–14 years, and 15–19 years, 20–29 years, 45–64 years, and 65 or more.

Table 13: Preliminary net undercount rate for in-scope persons by demographic; three-variable classifications (values expressed in percentage points rounded to two decimals)

Population group, Sex and Age group			Undercount rate (%)	Standard error	95% Confidence interval limits	
Population group	**Sex**	**Age group**			**Lower**	**Upper**
Black African	Male	Under 5 years	37,05	0,21	36,63	37,46
		5–14 years	30,41	0,18	30,05	30,77
		15–19 years	32,09	0,21	31,69	32,50
		20–29 years	42,92	0,20	42,53	43,30
		30–44 years	47,12	0,20	46,73	47,51
		45–64 years	37,74	0,20	37,36	38,13
		65 or more	26,87	0,29	26,30	27,43
	Female	Under 5 years	37,66	0,21	37,24	38,08
		5–14 years	30,31	0,18	29,96	30,66
		15–19 years	32,31	0,21	31,91	32,72
		20–29 years	39,97	0,21	39,56	40,38
		30–44 years	38,20	0,18	37,85	38,55
		45–64 years	27,74	0,19	27,37	28,12
		65 or more	21,69	0,27	21,17	22,22
Coloured	Male	Under 5 years	47,02	0,54	45,96	48,09
		5–14 years	43,16	0,62	41,95	44,37
		15–19 years	43,97	0,54	42,91	45,02
		20–29 years	50,20	0,49	49,24	51,16
		30–44 years	45,98	0,45	45,10	46,86
		45–64 years	44,53	0,46	43,63	45,44
		65 or more	40,59	0,47	39,68	41,51
	Female	Under 5 years	46,68	0,62	45,45	47,90
		5–14 years	42,47	0,50	41,49	43,44
		15–19 years	44,05	0,58	42,91	45,18
		20–29 years	46,74	0,42	45,93	47,56
		30–44 years	44,71	0,41	43,91	45,51
		45–64 years	38,59	0,43	37,75	39,42
		65 or more	37,47	0,55	36,40	38,55
Indian or Asian	Male	Under 5 years	81,69	1,61	78,53	84,86
		5–14 years	76,41	0,65	75,14	77,67
		15–19 years	73,88	1,24	71,45	76,31
		20–29 years	73,99	0,94	72,15	75,83
		30–44 years	78,83	0,84	77,19	80,48
		45–64 years	68,52	0,88	66,79	70,25
		65 or more	60,63	0,99	58,70	62,57

Population group, Sex and Age group			Undercount rate (%)	Standard error	95% Confidence interval limits	
					Lower	Upper
White	Female	Under 5 years	73,95	0,82	72,33	75,56
		5–14 years	73,39	0,96	71,51	75,27
		15–19 years	67,00	1,12	64,80	69,21
		20–29 years	71,33	1,42	68,56	74,11
		30–44 years	75,93	0,80	74,35	77,51
		45–64 years	66,79	0,85	65,12	68,47
		65 or more	62,69	1,02	60,69	64,68
	Male	Under 5 years	65,85	0,73	64,42	67,28
		5–14 years	64,23	0,48	63,29	65,18
		15–19 years	57,61	0,92	55,80	59,41
		20–29 years	68,86	0,63	67,62	70,10
		30–44 years	64,17	0,44	63,31	65,04
		45–64 years	61,33	0,48	60,38	62,28
		65 or more	55,73	0,58	54,59	56,88
	Female	Under 5 years	64,59	0,68	63,25	65,93
		5–14 years	57,62	0,55	56,53	58,70
		15–19 years	60,39	0,79	58,84	61,94
		20–29 years	65,81	0,50	64,82	66,80
		30–44 years	62,64	0,48	61,70	63,58
		45–64 years	59,69	0,44	58,83	60,56
		65 or more	55,76	0,58	54,62	56,90
Other	Male	Under 5 years	60,11	1,26	57,65	62,57
		5–14 years	43,05	1,16	40,77	45,33
		15–19 years	49,70	2,30	45,19	54,22
		20–29 years	77,88	0,96	75,99	79,76
		30–44 years	67,64	1,27	65,14	70,14
		45–64 years	52,34	1,28	49,83	54,85
		65 or more	67,73	0,81	66,14	69,33
	Female	Under 5 years	74,72	1,35	72,08	77,36
		5–14 years	64,51	1,59	61,40	67,61
		15–19 years	58,79	2,54	53,80	63,77
		20–29 years	73,71	2,14	69,52	77,91
		30–44 years	60,75	0,96	58,87	62,62
		45–64 years	33,81	1,05	31,75	35,86
		65 or more	52,82	0,91	51,03	54,60

Note: The undercount rates for undetermined categories form part of the calculations but are excluded from these tables.

In Table 13, the highest preliminary net undercount rate was observed among Indian or Asian males in the age group under 5 years at 81,69%, while the lowest preliminary net undercount rate was 21,69% in the population group black African for females in the age group 65 years or more.

Estimation of final true population for persons

Table 14: Distribution of final true population for persons (in millions rounded to two decimals)

PES Population	Census Enumeration (million)		
	Included	**Omitted**	**Total**
Included	23,70	16,78	**40,48**
Omitted	12,23	8,66	**20,89**
Total excluding erroneous inclusions and insufficient information	**35,93**	**25,44**	**61,37**

Table 14 shows that the census omitted an estimated 25,44 million persons in total, of which 16,78 million were correctly enumerated in the PES. There were 8,66 million persons who were missed in both the census and the PES. While the PES estimated the total population at 40,48 million persons, it omitted 12,23 million persons who were correctly enumerated in the census. The PES omitted a total of 20,89 million persons.

Estimation of the final net coverage error for persons

Table 15: Final net undercount rate for persons, by demographic groups – single-variable classifications (values expressed in percentage points rounded to two decimals)

Population group	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Black African	31,33	0,00	31,32	31,34
Coloured	28,73	0,02	28,69	28,76
Indian or Asian	42,10	0,06	42,00	42,20
White	24,86	0,05	24,79	24,94
Other	59,52	0,07	59,41	59,63
South Africa	**31,06**	**0,01**	**31,05**	**31,07**

Sex	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Male	32,00	0,01	31,99	32,01
Female	30,17	0,01	30,16	30,18
South Africa	**31,06**	**0,01**	**31,05**	**31,07**

Age group	Undercount rate (%)	Standard error	95% Confidence interval limits	
			Lower	Upper
Under 5 years	34,97	0,01	34,95	34,99
5–14 years	22,25	0,01	22,23	22,27
15–19 years	27,16	0,02	27,13	27,19
20–29 years	38,30	0,01	38,28	38,31
30–44 years	37,04	0,01	37,02	37,06
45–64 years	21,23	0,02	21,20	21,26
65 or more	37,45	0,02	37,42	37,48
South Africa	**31,06**	**0,01**	**31,05**	**31,07**

Table 15 shows that the population group Other has the highest final net undercount rate at 59,52%, with population group White having the lowest at 24,86%. The undercount rate for males (32,00%) is higher than that for females (30,17%). The age group 20–29 years has the highest undercount rate at 38,30% while the age group 45–64 years had the lowest undercount rate at 21,23%.

Table 16: Final net undercount rate for in-scope persons by demographic group; two-variable classifications (values expressed in percentage points rounded to two decimals)

Population group and Sex		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
Population group	Sex				
Black African	Male	32,73	0,01	32,72	32,74
	Female	30,02	0,01	30,01	30,03
Coloured	Male	27,70	0,03	27,65	27,75
	Female	29,66	0,03	29,61	29,70
Indian or Asian	Male	41,80	0,09	41,66	41,95
	Female	42,40	0,09	42,26	42,54
White	Male	22,74	0,07	22,63	22,85
	Female	26,86	0,06	26,76	26,96
Other	Male	60,02	0,09	59,88	60,17
	Female	58,76	0,11	58,59	58,94

Population group and Age group		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
Population group	Age group				
Black African	Under 5 years	35,95	0,01	35,93	35,97
	5–14 years	24,15	0,01	24,13	24,16
	15–19 years	27,72	0,01	27,70	27,74
	20–29 years	39,02	0,01	39,01	39,03
	30–44 years	37,14	0,01	37,13	37,16
	45–64 years	21,31	0,01	21,29	21,33
	65 or more	27,68	0,01	27,66	27,70
Coloured	Under 5 years	32,64	0,07	32,53	32,76
	5–14 years	20,16	0,06	20,06	20,26
	15–19 years	28,51	0,07	28,40	28,62
	20–29 years	37,33	0,04	37,27	37,40
	30–44 years	32,40	0,04	32,33	32,47
	45–64 years	19,95	0,06	19,86	20,04
	65 or more	40,15	0,04	40,09	40,21
Indian or Asian	Under 5 years	19,07	0,46	18,32	19,82
	5–14 years	16,77	0,31	16,26	17,28
	15–19 years	40,58	0,28	40,13	41,04
	20–29 years	45,87	0,14	45,64	46,10
	30–44 years	53,27	0,09	53,12	53,42
	45–64 years	32,95	0,14	32,73	33,17
	65 or more	65,33	0,05	65,24	65,41

		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
White	Under 5 years	14,08	0,28	13,62	14,53
	5–14 years	-12,23	0,25	-12,64	-11,83
	15–19 years	6,30	0,28	5,83	6,76
	20–29 years	18,50	0,16	18,24	18,76
	30–44 years	30,64	0,09	30,48	30,79
	45–64 years	17,07	0,10	16,91	17,23
	65 or more	57,02	0,04	56,96	57,08
Other	Under 5 years	63,54	0,21	63,19	63,88
	5–14 years	51,95	0,27	51,51	52,39
	15–19 years	51,78	0,40	51,12	52,44
	20–29 years	64,23	0,12	64,03	64,43
	30–44 years	64,60	0,10	64,43	64,77
	45–64 years	50,87	0,18	50,57	51,17
	65 or more	50,54	0,26	50,11	50,97

Sex and Age group		Undercount rate (%)	Standard error	95% Confidence interval limits	
				Lower	Upper
Sex	**Age group**				
Male	Under 5 years	35,13	0,02	35,10	35,16
	5–14 years	34,80	0,02	34,77	34,84
	15–19 years	22,81	0,02	22,78	22,84
	20–29 years	21,69	0,02	21,66	21,72
	30–44 years	27,76	0,02	27,72	27,80
	45–64 years	26,56	0,02	26,52	26,60
	65 or more	39,05	0,01	39,03	39,08
Female	Under 5 years	37,55	0,01	37,53	37,57
	5–14 years	39,51	0,01	39,48	39,53
	15–19 years	34,61	0,01	34,59	34,64
	20–29 years	20,89	0,03	20,85	20,94
	30–44 years	21,51	0,02	21,48	21,55
	45–64 years	38,58	0,03	38,53	38,63
	65 or more	36,75	0,02	36,71	36,78

Table 16 shows the highest final net undercount rate was observed among the Indian or Asian population group in the age group under 65 or more years at 65,33%, while the lowest final net undercount rate was 6,30% for the White population group in the age group 15–19 years. A final net overcount rate at 12,23% was observed for the White population group in the age group 5–14 years.

Table 17: Final net undercount rate for in-scope persons by demographic; three-variable classifications (values expressed in percentage points rounded to two decimals)

Population group, Sex and Age group			Undercount rate (%)	Standard error	95% Confidence interval limits	
Population group	**Sex**	**Age group**			**Lower**	**Upper**
Black African	Male	Under 5 years	36,12	0,01	36,10	36,14
		5–14 years	24,74	0,01	24,71	24,76
		15–19 years	28,25	0,02	28,23	28,28
		20–29 years	40,06	0,01	40,04	40,08
		30–44 years	39,87	0,01	39,85	39,89
		45–64 years	22,33	0,02	22,29	22,36
		65 or more	28,51	0,02	28,47	28,54
	Female	Under 5 years	35,78	0,01	35,76	35,80
		5–14 years	23,55	0,01	23,53	23,58
		15–19 years	27,17	0,02	27,15	27,20
		20–29 years	38,00	0,01	37,98	38,02
		30–44 years	34,47	0,01	34,45	34,49
		45–64 years	20,47	0,02	20,44	20,50
		65 or more	27,22	0,01	27,20	27,24
Coloured	Male	Under 5 years	32,53	0,10	32,37	32,69
		5–14 years	20,21	0,09	20,06	20,35
		15–19 years	28,72	0,09	28,56	28,87
		20–29 years	36,24	0,06	36,13	36,34
		30–44 years	33,08	0,06	32,97	33,18
		45–64 years	15,78	0,09	15,63	15,93
		65 or more	37,98	0,07	37,86	38,10
	Female	Under 5 years	32,76	0,10	32,59	32,92
		5–14 years	20,11	0,09	19,97	20,26
		15–19 years	28,30	0,09	28,15	28,46
		20–29 years	38,39	0,05	38,31	38,48
		30–44 years	31,79	0,06	31,69	31,89
		45–64 years	23,33	0,07	23,22	23,44
		65 or more	41,51	0,04	41,45	41,57
Indian or Asian	Male	Under 5 years	18,74	0,65	17,68	19,81

Population group, Sex and Age group			Undercount rate (%)	Standard error	95% Confidence interval limits	
					Lower	Upper
		5–14 years	18,26	0,42	17,56	18,95
		15–19 years	41,91	0,37	41,29	42,52
		20–29 years	45,26	0,20	44,93	45,58
		30–44 years	55,51	0,12	55,32	55,71
		45–64 years	29,75	0,21	29,41	30,10
		65 or more	63,48	0,10	63,32	63,64
	Female	Under 5 years	19,40	0,64	18,34	20,45
		5–14 years	15,17	0,46	14,42	15,92
		15–19 years	39,17	0,41	38,49	39,85
		20–29 years	46,59	0,20	46,26	46,92
		30–44 years	50,75	0,15	50,51	50,99
		45–64 years	35,96	0,18	35,67	36,26
		65 or more	66,72	0,06	66,63	66,82
White	Male	Under 5 years	15,67	0,38	15,05	16,30
		5–14 years	-10,91	0,34	-11,47	-10,35
		15–19 years	8,74	0,38	8,12	9,36
		20–29 years	17,32	0,23	16,95	17,69
		30–44 years	30,62	0,13	30,40	30,83
		45–64 years	12,66	0,15	12,41	12,90
		65 or more	55,40	0,06	55,30	55,50
	Female	Under 5 years	12,36	0,41	11,69	13,03
		5–14 years	-13,62	0,36	-14,22	-13,03
		15–19 years	3,69	0,42	3,01	4,38
		20–29 years	19,75	0,22	19,39	20,11
		30–44 years	30,66	0,13	30,44	30,87
		45–64 years	21,02	0,12	20,82	21,23
		65 or more	58,35	0,04	58,28	58,42
Other	Male	Under 5 years	63,36	0,30	62,87	63,85
		5–14 years	52,12	0,37	51,50	52,73
		15–19 years	52,02	0,56	51,10	52,93
		20–29 years	63,86	0,16	63,60	64,12
		30–44 years	64,43	0,13	64,21	64,64
		45–64 years	51,25	0,25	50,84	51,66
		65 or more	50,81	0,38	50,19	51,44
	Female	Under 5 years	63,71	0,29	63,23	64,19
		5–14 years	51,77	0,38	51,14	52,40
		15–19 years	51,51	0,58	50,57	52,46
		20–29 years	64,91	0,19	64,60	65,22

Population group, Sex and Age group			Undercount rate (%)	Standard error	95% Confidence interval limits	
					Lower	Upper
		30–44 years	64,95	0,17	64,67	65,23
		45–64 years	50,38	0,26	49,94	50,81
		65 or more	50,29	0,36	49,69	50,89

In Table 17, the highest final net undercount rate was observed among Indian or Asian females in the age group 65 years or more at 66,72%, while the lowest final net undercount rate was 3,69% in the population group White for females in the age group 15–19 years or more. Final net overcount rates were observed among the White population group for females and males in the age group 5–14 years at 13,62% and 10,91% respectively.

Detailed results – Evaluation of coverage error for households

Estimation of preliminary true population for households

Table 18: Distribution of uncorrected census population for households (in millions rounded to two decimals)

Component of Uncorrected Census Population	Census Enumeration (million)
Included in PES	7,20
Omitted from PES	3,73
Erroneous inclusions	0,50
Insufficient information	2,33
Total of Uncorrected Population	**13,77**

Table 18 shows that, of the 13,77 million households counted in the in-scope census, 10,94 million are estimated to be correctly enumerated, i.e. the PES enumerated 7,20 million and missed 3,73 million. The census erroneous inclusions and census insufficient information are estimated to be 0,50 million and 2,33 million respectively.

Table 19: Distribution of preliminary true population for households (in millions rounded to two decimals)

PES Population	Census Enumeration (million)		
	Included	**Omitted**	**Total**
Included	7,20	5,63	**12,84**
Omitted	3,73	2,92	**6,65**
Total excluding erroneous inclusions and insufficient information	**10,94**	**8,55**	**19,49**

Table 19 shows that the census omitted an estimated 8,55 million households in total, 5,63 million which were correctly enumerated in the PES, and another 2,92 million which were missed in the PES as well as in the census. While the PES estimated the in-scope household total at 12,84 million households, it omitted 3,73 million households that were correctly enumerated in the census. There were 2,92 million households that were missed in both the census and the PES, for a PES total omission of 6,65 million households.

Table 20: Probabilities of inclusion and omission of an in-scope household

Probability	Value
Probability of inclusion in census	0,5611
Probability of inclusion in PES	0,6586
Probability of inclusion in both census and PES	0,3696
Probability of inclusion in census, missed in PES	0,1915
Probability of inclusion in PES, missed in census	0,2891
Probability of missed in both census and PES	0,1498

Table 20 shows the overall empirical probabilities of inclusion and omission of a household in the census or in the PES. A household in the in-scope universe had an approximately 56,11% chance of being enumerated in the census, 65,86% chance of being enumerated in the PES, and 36,96% chance of being enumerated in both. Conversely, the household had approximately 19,15% chance of being included in the census but missed in the PES, 28,91% chance of being included in the PES but missed in the census, and 14,98% chance of being missed in both.

Estimation of the preliminary net coverage rate for households

Figure 18: Graphic representation of confidence intervals for households undercount rate (%), by province (Lower and Upper limits)
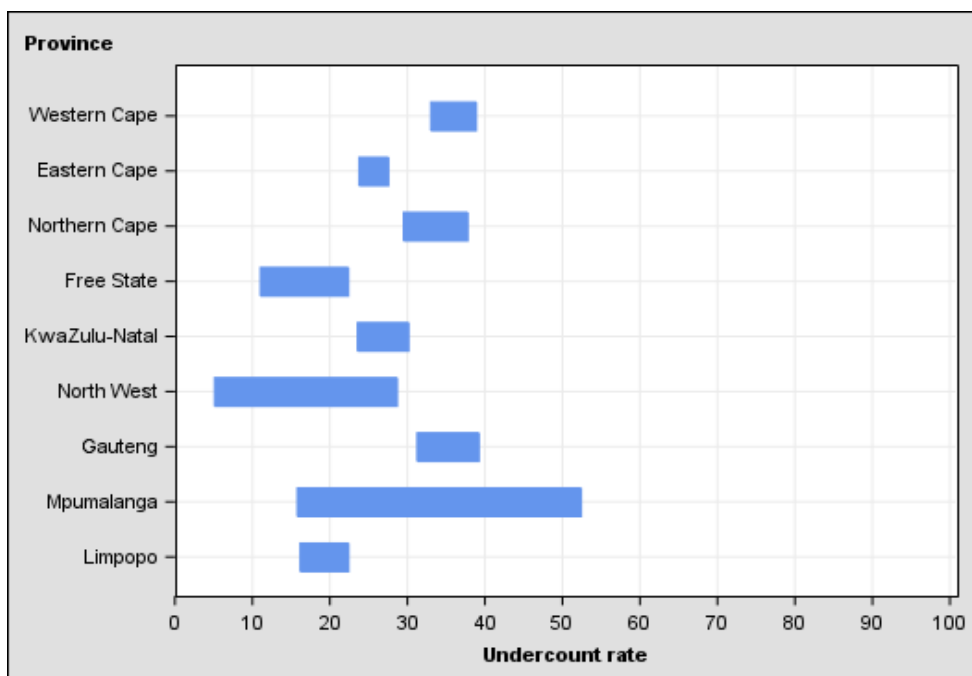
In Figure 18, it can be observed that among the provinces, there is no statistically significant difference in the undercount rates for Free State (16,65%), North West (16,84%), Mpumalanga (34,05%) and Limpopo (19,23%). There is also no statistically significant difference in the undercount rates for Western Cape (35,89%), Northern Cape (33,59%), and Gauteng (35,19%).

Estimation of final true population for households

Table 21: Distribution of final true population for households (in millions rounded to two decimals)

	Census Enumeration (million)		
PES Population	**Included**	**Omitted**	**Total**
Included	7,12	5,57	**12,68**
Omitted	3,69	2,88	**6,57**
Total excluding erroneous inclusions and insufficient information	**10,80**	**8,45**	**19,25**

Table 21 shows that the census omitted an estimated 8,45 million households in total, 5,57 million of which were correctly enumerated in the PES, and another 2,88 million that were missed in the PES as well as in the census. While the PES estimated the in-scope household total at 12,68 million households, it omitted 3,69 million households that were correctly enumerated in the census. There were 2,88 million households that were missed in both the census and the PES, for a PES total omission of 6,57 million households.

Detailed results – Content evaluation (person only)

Nature of content analysis

Content analysis uses unweighted "matched" persons identified in the census and PES. The variables under consideration were Population group, Sex, Age group, Relationship to head of household, Marital status, and Country of birth. The sole purpose of these totals is to compare the census responses with the PES responses and to calculate the measures of consistency; they are not for socio-demographic analytical purposes. Content error is measured using four indicators: Rate of agreement, Index of inconsistency, Aggregate index of inconsistency, and absolute value of Net difference rate (NDR) relative to mean or proportion (NDR/P).

Table 22: Standards for the interpretation of the different content error measures

Measure	Low	Moderate	High
Index of inconsistency	<20%	20%–50%	>50%
Aggregate index of inconsistency	<20%	20%–50%	>50%
Absolute value of NDR relative to mean or proportion (NDR/P)	<0,01	0,01–0,05	>0,05

Content analysis for Sex

Table 23: Sex as reported in the census and the PES for matched persons

PES population	Census population			
	Male	**Female**	**Undetermined**	**Total**
Male	73 307	3 169	0	**76 476**
Female	3 574	88 846	0	**92 420**
Undetermined	69	54	0	**123**
Total	**76 950**	**92 069**	**0**	**169 019**

Table 24: Net difference rate and index of inconsistency for Sex

Category	Consistent cases	Census	PES	Net difference rate			Index of inconsistency		
				Rate (%)	95% Confidence interval limits		Index (%)	95% Confidence interval limits	
					Lower	Upper		Lower	Upper
Male	73 307	76 950	76 476	0,28	0,18	0,38	8,13	7,94	8,33
Female	88 846	92 069	92 420	-0,21	-0,31	-0,11	8,11	7,92	8,31
Undetermined			123	.	.	.	.	.	.
Total	**162 153**	**169 019**	**169 019**	**0,00**	**0,00**	**0,00**	**.**	**.**	**.**

Table 25: Aggregated index of inconsistency, Gross difference rate and Rate of agreement for Sex

Category	Aggregated index of inconsistency (%)	Gross difference rate (%)	Rate of agreement (%)
Total	8,19	4,23	95,76

The characteristic Sex has an aggregated index of inconsistency of 8,19%, which shows a **low** level of inconsistency or variability (index < 20%) and was reported reliably and consistently between the census and PES.

Content analysis for Age group

Table 26: Age group as reported in the census and the PES for matched persons

PES population	Census population								Total
	Under 5 years	5–14 years	15–19 years	20–29 years	30–44 years	45–64 years	65 or more	Undetermined	
Under 5 years	14 450	446	31	40	26	7	2	444	**15 446**
5–14 years	399	32 392	372	114	28	10	5	1 220	**34 540**
15–19 years	25	372	14 007	293	27	1	5	585	**15 315**
20–29 years	35	133	258	21 907	678	45	11	728	**23 795**
30–44 years	12	30	31	696	29 398	566	54	1 074	**31 861**
45–64 years	8	4	3	51	548	26 227	413	950	**28 204**
65 or more	1	1		8	30	324	10 165	387	**10 916**
Undetermined	560	1 455	610	1 214	1 890	1 544	750	919	**8 942**
Total	**15 490**	**34 833**	**15 312**	**24 323**	**32 625**	**28 724**	**11 405**	**6 307**	**169 019**

Table 27: Net difference rate, index of inconsistency, and gross difference rate for Age group

Category	Consistent cases	Census	PES	Net difference rate			Index of inconsistency		
				Rate (%)	95% Confidence interval limits		Index (%)	95% Confidence interval limits	
					Lower	Upper		Lower	Upper
Under 5 years	14 450	15 490	15 446	0,03	-0,03	0,08	7,24	6,93	7,57
5–14 years	32 392	34 833	34 540	0,17	0,09	0,25	8,32	8,08	8,57
15–19 years	14 007	15 312	15 315	-0,00	-0,06	0,06	9,38	9,02	9,76
20–29 years	21 907	24 323	23 795	0,31	0,23	0,39	10,43	10,12	10,75
30–44 years	29 398	32 625	31 861	0,45	0,36	0,54	10,90	10,62	11,20
45–64 years	26 227	28 724	28 204	0,31	0,23	0,39	9,45	9,17	9,74
65 or more	10 165	11 405	10 916	0,29	0,24	0,34	9,55	9,13	9,99
Undetermined	919	6 307	8 942	-1,56	-1,70	-1,42	91,97	90,40	93,57
Total	**149 465**	**169 019**	**169 019**	**0,00**	**0,00**	**0,00**	**.**	**.**	**.**

Table 28: Aggregated index of inconsistency, Gross difference rate and Rate of agreement for Age group

Category	Aggregated index of inconsistency (%)	Gross difference rate (%)	Rate of agreement (%)
Total	13,61	13,08	86,92

In both the census and PES, Age was derived from the date of birth. Age derived from the date of birth was preferred over reported Age when the two were inconsistent. The characteristic Age (as derived) has an aggregated index of inconsistency of 13,61%, which shows a low level of inconsistency or variability (index < 20%) and was reported reliably and consistently between the census and PES.

Content analysis for Population group

Table 29: Population group as reported in PES and in census for matched persons

PES population	Census population						Total
	Black African	**Coloured**	**Indian or Asian**	**White**	**Other**	**Undetermined**	
Black African	146 325	673	50	57	51	305	**147 461**
Coloured	614	13 352	54	68	141	111	**14 340**
Indian or Asian	43	50	1 237	27	5	7	**1 369**
White	57	114	19	5 139	72	55	**5 456**
Other	121	54	16	21	24	1	**237**
Undetermined	142	11		3			**156**
Total	**147 302**	**14 254**	**1 376**	**5 315**	**293**	**479**	**169 019**

Table 30: Net difference rate, index of inconsistency, and gross difference rate for Population group

Category	Consistent cases	Census	PES	Net difference rate			Index of inconsistency		
				Rate (%)	95% Confidence interval limits		Index (%)	95% Confidence interval limits	
					Lower	Upper		Lower	Upper
Black African	146 325	147 302	147 461	-0,09	-0,15	-0,04	5,60	5,36	5,85
Coloured	13 352	14 254	14 340	-0,05	-0,10	0,00	7,22	6,90	7,56
Indian or Asian	1 237	1 376	1 369	0,00	-0,02	0,02	9,95	8,82	11,24
White	5 139	5 315	5 456	-0,08	-0,11	-0,06	4,73	4,32	5,17
Other	24	293	237	0,03	0,01	0,06	91,08	83,16	99,77
Undetermined		479	156	0,19	.	.	.	.	.
Total	**166 077**	**169 019**	**169 019**	**0,00**	**0,00**	**0,00**	**.**	**.**	**.**

Table 31: Aggregated index of inconsistency, Gross difference rate and Rate of agreement for Population group

Category	Aggregated index of inconsistency (%)	Gross difference rate (%)	Rate of agreement (%)
Total	7,52	1,77	98,23

The characteristic 'Population group' has an aggregated index of inconsistency of 7,52%, which also exhibits a **low** degree of inconsistency and variability (index < 20%) among the characteristics measured. It is quite robust and reliable from one measurement to another.

Content analysis for Relationship to head of household

Table 32: Relationship to head of household as reported in PES and in census for matched persons

PES population	Census population																
	Head or acting head	Husband or wife or partner	Son or daughter	Adopted son or daughter	Foster child	Stepchild	Brother or sister	Parent (mother or father)	Parent-in-law	Grand or great grandchild	Son or daughter-in-law	Brother or sister-in-law	Grandmother or father	Other relative	Non-related person	Undetermined	Total
Head or acting head	36 920	5 218	1 617	21	4	18	1 138	804	38	255	65	51	103	606	217	139	**47 214**
Husband or wife or partner	5 094	9 033	222	4		4	132	176	9	23	66	33	10	112	160	69	**15 147**
Son or daughter	1 518	154	45 920	543	45	399	1 552	52	3	2 923	149	43	4	1 556	80	160	**55 101**
Adopted son or daughter	28	7	565	35	14	20	23	2	1	78	5			70	8	4	**860**
Foster child	5	1	35	16	27	4	9			37	1			49	8		**192**
Stepchild	14	9	484	17	3	108	18			53	4	2		68	6	4	**790**
Brother or sister	1 111	107	1 556	24	5	21	2 674	9	2	319	25	76	7	501	42	15	**6 494**
Parent (mother or father)	882	173	148	1		1	39	283	27	28	6	3	29	36	7	11	**1 674**
Parent-in-law	38	15	19				2	45	34	9	60	8	6	15	6	1	**258**
Grand or great grandchild	164	22	3 075	80	32	56	369	22	1	25 038	111	22	15	1 406	48	46	**30 507**
Son or daughter-in-law	78	83	122	9	2	3	26	7	3	93	317	43		44	37	1	**868**
Brother or sister-in-law	74	30	55		3	2	63	3	1	23	32	63	1	53	15		**418**
Grandmother or father	119	9	28				9	18	7	157	3	1	10	13	1		**375**

PES population	Census population																
	Head or acting head	Husband or wife or partner	Son or daughter	Adopted son or daughter	Foster child	Stepchild	Brother or sister	Parent (mother or father)	Parent-in-law	Grand or great grandchild	Son or daughter-in-law	Brother or sister-in-law	Grandmother or father	Other relative	Non-related person	Undetermined	Total
Other relative	470	131	1 549	62	33	60	564	20	6	1 629	69	53	4	3 007	91	9	**7 757**
Non-related person	301	277	95	5	5	3	61	8	2	43	30	8	3	115	230	4	**1 190**
Undetermined	24	7	44	2		1	7	3		70	1	1		14			**174**
Total	**46 840**	**15 276**	**55 534**	**819**	**173**	**700**	**6 686**	**1 452**	**134**	**30 778**	**944**	**407**	**192**	**7 665**	**956**	**463**	**169 019**

Table 33: Net difference rate, index of inconsistency, and gross difference rate for Relationship to head of household

| Category | Consistent cases | Census | PES | Net difference rate | | | Index of inconsistency | | |
| | | | | Rate (%) | 95% Confidence interval limits | | Index (%) | 95% Confidence interval limits | |
					Lower	Upper		Lower	Upper
Head or acting head	36 920	46 840	47 214	-0,22	-0,39	-0,05	29,78	29,36	30,20
Husband or wife or partner	9 033	15 276	15 147	0,08	-0,06	0,21	44,63	43,84	45,44
Son or daughter	45 920	55 534	55 101	0,26	0,09	0,42	25,25	24,89	25,62
Adopted son or daughter	35	819	860	-0,02	-0,07	0,02	96,31	91,63	101,23
Foster child	27	173	192	-0,01	-0,03	0,01	85,30	76,16	95,53
Stepchild	108	700	790	-0,05	-0,10	-0,01	85,88	81,20	90,83
Brother or sister	2 674	6 686	6 494	0,11	0,01	0,22	61,83	60,45	63,25
Parent (mother or father)	283	1 452	1 674	-0,13	-0,19	-0,07	82,65	79,45	85,99
Parent-in-law	34	134	258	-0,07	-0,09	-0,05	82,74	74,04	92,46
Grand or great grandchild	25 038	30 778	30 507	0,16	0,04	0,29	22,34	21,92	22,77
Son or daughter-in-law	317	944	868	0,04	0,00	0,09	65,36	61,66	69,28
Brother or sister-in-law	63	407	418	-0,01	-0,04	0,02	84,93	78,75	91,61
Grandmother or father	10	192	375	-0,11	-0,14	-0,08	96,62	88,70	105,24
Other relative	3 007	7 665	7 757	-0,05	-0,17	0,06	63,92	62,62	65,25
Non-related person	230	956	1 190	-0,14	-0,19	-0,09	79,06	75,30	83,01
Undetermined		463	174	0,17	.	.	.	.	.
Total	**123 699**	**169 019**	**169 019**	**0,00**	**0,00**	**0,00**	**.**	**.**	**.**

Table 34: Aggregated index of inconsistency, Gross difference rate and Rate of agreement for Relationship to head of household

Category	Aggregated index of inconsistency (%)	Gross difference rate (%)	Rate of agreement (%)
Total	34,79	36,64	63,36

The characteristic 'Relationship to head of household' has an aggregated index of inconsistency of 34,79%, which shows a **moderate** level of inconsistency or variability (20% < index < 50%). It may not be possible to ensure more consistent responses for this variable in future surveys unless a *de jure* rule (usual residence) is used. With a *de jure* rule, the head of household generally remains the same, even when temporarily absent from the household.

Content analysis for Marital status

Table 35: Marital status as reported in PES and in census for matched persons

PES population	Census population								
	Married (include customary, traditional, religious, etc.)	Living together like husband and wife or life partners	Divorced	Separated, but still legally married	Widowed	Never married	Do not know	Undeter mined	Total
Married (include customary, traditional, religious, etc.)	23 423	1 673	138	138	1 305	2 079	89	204	**29 049**
Living together like husband and wife or life partners	1 547	4 200	47	26	91	2 088	95	40	**8 134**
Divorced	171	73	730	48	108	446	21	17	**1 614**
Separated, but still legally married	162	31	34	72	42	142	4	2	**489**
Widowed	1 606	76	132	29	3 282	907	51	23	**6 106**
Never married	2 526	2 548	450	154	884	74 332	533	2 384	**83 811**
Do not know	104	32	34	4	55	440	8	49	**726**
Undetermined	48	25	5	3	14	2 023	34	36 938	**39 090**
Total	**29 587**	**8 658**	**1 570**	**474**	**5 781**	**82 457**	**835**	**39 657**	**169 019**

Post-enumeration Survey, PES 2022

Table 36: Net difference rate, index of inconsistency, and gross difference rate for Marital status

Category	Consistent cases	Census	PES	Net difference rate			Index of inconsistency		
				Rate (%)	95% Confidence interval limits		Index (%)	95% Confidence interval limits	
					Lower	Upper		Lower	Upper
Married (include customary, traditional, religious, etc.)	23 423	29 587	29 049	0,32	0,19	0,45	24,33	23,88	24,78
Living together like husband and wife or life partners	4 200	8 658	8 134	0,31	0,20	0,42	52,59	51,45	53,75
Divorced	730	1 570	1 614	-0,03	-0,08	0,02	54,66	52,09	57,36
Separated, but still legally married	72	474	489	-0,01	-0,04	0,02	85,29	79,53	91,46
Widowed	3 282	5 781	6 106	-0,19	-0,28	-0,11	46,41	45,16	47,70
Never married	74 332	82 457	83 811	-0,80	-0,96	-0,64	20,83	20,52	21,15
Do not know	8	835	726	0,06	0,02	0,11	99,43	94,50	104,62
Undetermined	36 938	39 657	39 090	0,34	0,25	0,42	8,06	7,84	8,30
Total	**142 985**	**169 019**	**169 019**	**0,00**	**0,00**	**0,00**	**.**	**.**	**.**

Table 37: Aggregated index of inconsistency, Gross difference rate and Rate of agreement for Marital status

Category	Aggregated index of inconsistency (%)	Gross difference rate (%)	Rate of agreement (%)
Total	22,99	18,21	81,79

The characteristic 'Marital status' has an aggregated index of inconsistency of 22,99%, which shows a **moderate** level of inconsistency or variability (20% < index < 50%). It was not reported consistently between census and PES. At the level of each response category, the Undetermined shows an even greater degree of inconsistency.

Content analysis for Country of birth

Table 38: Country of birth as reported in PES and in census for matched persons

PES population	Census population			
	Outside South Africa	South Africa	Undetermined	Total
Outside South Africa	2 520	735	134	**3 389**
South Africa	455	162 730	1 777	**164 962**
Stateless		1		**1**
Undetermined	64	574	29	**667**
Total	**3 039**	**164 040**	**1 940**	**169 019**

Table 39: Net difference rate, index of inconsistency and gross difference rate for Country of birth

Category	Consistent cases	Census	PES	Net difference rate			Index of inconsistency		
				Rate (%)	95% Confidence interval limits		Index (%)	95% Confidence interval limits	
					Lower	Upper		Lower	Upper
Outside South Africa	2 520	3 039	3 389	-0,21	-0,25	-0,16	22,01	20,86	23,22
South Africa	162 730	164 040	164 962	-0,55	-0,62	-0,48	40,26	38,93	41,64
Stateless			1	.	.	.	.	.	.
Undetermined	29	1 940	667	0,75	0,69	0,81	98,35	94,53	102,33
Total	**165 279**	**169 019**	**169 019**	**0,00**	**0,00**	**0,00**	**.**	**.**	**.**

Table 40: Aggregated index of inconsistency, Gross difference rate and Rate of agreement for Country of birth

Category	Aggregated index of inconsistency (%)	Gross difference rate (%)	Rate of agreement (%)
Total	42,27	2,26	97,74

The characteristic Country of birth shows a **moderate** level of inconsistency or variability at 42,27% (20% < index < 50%).

EXPLANATORY NOTES

Background **1** South Africa conducts a de facto population census to get an official count of the country's population decennially. Stats SA conducted data collection for Census 2022 between February 2022 and May 2022, with a Census reference night on the midnight of 02–03 February 2022. A *de facto* count is based on a person's presence on the reference night at a particular dwelling unit, while a de jure count is based on a person's long-term presence at their usual place of residence. Stats SA also uses an evaluation instrument called the Post-enumeration Survey (PES) to measure the accuracy in terms of content error of the information collected during census, and reach in terms of the coverage error of the census. PES 2022 data collection was conducted immediately after Census 2022 data collection, between June 2022 and August 2022, with a *de facto* reference night on the midnight of 22–23 June 2022.

Purpose of the survey **2** The Post-enumeration Survey (PES) is an independent survey used to measure the accuracy and reach of the Census through content and coverage errors respectively. Coverage error is a measure of how many households and persons were erroneously included, missed, or counted more than once, while Content error indicates the quality of the key characteristics in the census by examining the concordance of responses between the PES and Census for the same characteristic. The output of the PES provides a statistically informed basis for adjusting the census count.

Dual system estimation (DSE) **3** The methodology used to derive the estimate of the true population based on the assumption that data collection in the PES is independent from the Census. One enumerates the persons and/or households independently in the two, to identify how many are found in each and in both as a basis for estimating the size of the true population. The first attempt at measuring the true population comes from the census-enumerated population, based on the enumeration in the Census. The second attempt comes from the enumeration of the PES sample to yield the corresponding PES estimate. Instead of assuming that these two estimates are better than each other, the methodology combines the two to derive a third, composite estimate of the true population referred as to as the Dual-system Estimate (DSE) of the true population. Resultantly, this estimate is more complete than either that from the Census or the PES. The difference between the estimated true population and the census-enumerated population is called the net coverage error of the Census. It denotes the net undercount or overcount of the Census as applicable.

| **DSE assumptions** | **4** | The DSE methodology assumes the following: |

- Closed population: migration between the census and PES is insignificant and the composition of the population remains relatively unchanged.
- Independence: there is independence between census and PES in the sense that the organisation of the census and PES, especially fieldwork operations, is separately managed.
- Error-free enumeration: there is absence of erroneous inclusions in either the census or PES.
- Perfect linking: there are no incomplete matches in the sense that any failure to match census and PES records is due to actual omission and not the inability to match.

Summary of dual system estimation

Enumeration status	Found in Census	Not found in Census	TOTAL
Found in PES	Matched	Excluded erroneously	Counted
Not found in PES	Included erroneously	Missed	Omitted
TOTAL	Counted	Omitted	True population

Classifications   **5**   The system of classifications used in Census 2022 is followed.

Statistical unit   **6**   The statistical units for which information is compiled and published are persons and households who spent the Census and/or PES reference nights in private dwelling units (DUs).

Preliminary estimate   **7**   An estimate computed with no adjustment factor applied.

Final estimate   **8**   An estimate computed with an adjustment factor applied.

Symbols and abbreviations   **9**   Figures are rounded off to the fourth decimal. Where figures have been rounded off, discrepancies may occur between sums of the component items and the totals. Additionally:

Symbol	Meaning
..	data is not available
0 or 0,0	more than nothing but less than half the final digit shown
blank space	data is not yet available
-	Nil or not applicable

Previously published information   **9**   Stats SA has conducted a PES for all post-1994 Censuses, including 1996, 2001, and 2011. Past PES publications and data are available on the Stats SA website and/or through Stats SA.

TECHNICAL NOTES

Scope of the survey

The scope of the PES 2022 is to estimate the total number of persons and households in private dwelling units (DUs) on Census 2022 reference night. The scope of the PES includes the following sub-EA types:

- Urban;
- Traditional; and
- Farms.

The scope of the PES excludes the following:

- Homes for the aged unless they are structured into separate households;
- Student residences;
- Tourist hotels/motels/inns;
- Institutions, e.g. prisons and hospitals; and
- The homeless.

Survey methodology and design

A single-stage stratified sample design is used. The sample of sub-EAs was selected using fractional interval systematic random sampling within each design stratum based on the required number of sub-EAs to be selected per design stratum. The sub-EAs for the PES 2022 sampling frame were stratified to improve the efficiency of the PES 2022 sample design. The sampling frame was first stratified by province, and within each province, the sub-EAs were explicitly stratified by sub-EA geographic type such as Urban, Traditional, and Farms. Three design strata were defined per province, with the exception of the Western Cape which had two design strata. Due to a low number of sub-EAs classified as traditional in the Western Cape, the geographic area types Traditional and Farms were combined to form one design strata. As a result, a total of 26 design strata were defined for the PES 2022. Within each of these strata, the sub-EAs were further ordered by sub-EA type and sub-EA code to provide additional implicit stratification gains.

Sample size

To determine the required samples size of PES 2022 the achieved precision levels at provincial level for the key indicator of the preliminary net undercount rate from PES 2011 was considered, allowing for a 15% reduction in the coefficients of variation for that indicator for PES 2022. This resulted in a total sample size of 840 sub-EAs allocated to the formed design strata. The provincial sample was allocated proportionally to the design strata based on the DU counts within each of the design strata per province. The systematic random sampling method with fractional intervals was used to select the sub-EAs within each design stratum. Of the 840 sub-EAs nationally, the sample size per province is:

Province	Sample size
Western Cape	113
Eastern Cape	144
Northern Cape	28
Free State	49
KwaZulu-Natal	165
North West	54
Gauteng	168
Mpumalanga	48
Limpopo	71
South Africa	**840**

Response rate at household level

The overall National response rate is 95,2%, and that per province is:

Province	Response rate (%)
Western Cape	93,41
Eastern Cape	95,25
Northern Cape	94,14
Free State	95,22
KwaZulu-Natal	95,70
North West	95,49
Gauteng	95,45
Mpumalanga	96,35
Limpopo	94,80
South Africa	**95,15**

Collection rate at household level

The overall National collection rate is 98,3%, and that per province is:

Province	Collection rate (%)
Western Cape	95,1
Eastern Cape	97,7
Northern Cape	100,0
Free State	99,9
KwaZulu-Natal	98,8
North West	98,2
Gauteng	97,7
Mpumalanga	99,1
Limpopo	99,8
South Africa	**98,3**

Weighting methodology

The base weight is the inverse sampling rate at the sub-EA level, accounting for the selection of sub-EAs within each design stratum. The base weight at household and person level is the same as the base weight at sub-EA level since all households and persons within the selected sub-EAs are included with certainty in the sample. Based on the sample requirements and specifications for the PES 2022, and the analysis of the Census 2022 sub-EA frame, the PES 2022 sampling frame excluded in-scope sub-EAs with a missing, zero, or DU count below 25. These sub-EAs are excluded from the sampling frame because they are often very remote and sparsely populated, representing only a small portion of the population. For the purpose of weighting, only the sub-EAs with a DU count below 25 were eligible for adjustment. If a sampled sub-EA experienced total non-response in a case such as access being denied to a sampled sub-EA, a non-response adjustment factor is applied to the base weights of responding sub-EAs to compensate for the sub-EAs with total non-response. The sub-EA non-response adjustment is applied within the design strata. In the PES 2022 there were no non-responding sub-EAs, and non-response adjustment was therefore not computed. Three sampled sub-EAs found to be out-of-scope because of having a zero household count were excluded from the weighting process, with no adjustment to the weights of in-scope sub-EAs.

Base weight

The inverse of the probability of inclusion, defined as follows:

$$W_{Bh} = \frac{N_h}{n_h}$$

where:
W_{Bh} = base weight for selected Sub-EAs in stratum "h",
N_h = total number of Sub-EAs in stratum "h", and
n_h = number of sampled Sub-EAs selected in stratum "h".

An adjustment is done to the base weight within the design strata to account for the excluded Sub-EAs with a dwelling unit count below 25 using the below adjustment:

$$W_{Eh} = \frac{N_h + N_{25E}}{N_h}$$

where:
W_{Eh} = adjustment for excluded Sub-EAs with less than 25 DUs in stratum "h", and
N_{25E} = total number of Sub-EAs with less than 25 DUs excluded in stratum "h".

An adjustment for Sub-EA non-response is also provided for. The Sub-EA non-response adjustment is applied within the design strata, and defined as follows:

$$W_{NRh} = \frac{n_{*h}}{n_{rh}}$$

where:
W_{NRh} = non-response adjustment in stratum "h",
n_{*h} = total number of eligible Sub-EAs in stratum "h", and
n_{rh} = total number of responding Sub-EAs in stratum "h".

Adjusted base weight

Therefore, the adjusted base weight (W_{ABh}) for strata that experience non-response at Sub-EA level is the product between the base weight and the adjustments:

$$W_{ABh} = W_{Bh} \times W_{Eh} \times W_{NRh}$$

Since there are no non-responding Sub-EAs in PES 2022, the final adjusted base weight is the product of the first two terms.

**Dual-system estimation
sequence**

In the dual system estimation methodology, the matched (M) population is the sum of the estimated:
- Total number of matched non-movers plus
- Total number of matched in-movers

The E-sample estimate of the population enumerated in the Census (CP) is the sum of the estimated:
- Total number of matched non-movers plus
- Total number of matched out-movers plus
- Total number of erroneously included in the Census plus
- Total number of correctly enumerated in the Census but missed in the PES

The P-sample estimate of the total population referred to as the PES population (PESP) is the sum of the estimated:
- Total number of non-movers in the population plus
- Total number of in-movers in the population

Subtracting the matched population (M) from the PES estimate of the total population (PESP) gives Census omissions (CO).

The census omission rate (OR) is the ratio of Census omissions to the PES population:

$$OR = \frac{CO}{PESP}$$

The coverage rate (CR) is the complement of the omission rate, and may be computed that way or as the ratio of the matched population to the PES population:

$$CR = 1 - OR = \frac{M}{PESP}$$

The cross count between the P-sample and the E-sample gives the total number of census erroneous inclusions (CEI). The census erroneous inclusion rate (CER) is the ratio between the total number of persons erroneously included in the Census relative to the E-sample estimate of the Census population:

$$CER = \frac{CEI}{CP}$$

The preliminary dual-system estimate of the true population (TP_p) is the product between the inverse of the coverage rate and the difference between the census population and those erroneously included in the Census:

$$TP_p = \frac{1}{CR} \times (CP - CEI)$$

The preliminary net coverage error (NCE_p) follows as the difference between the preliminary estimate of the true population and the census population estimated from the E-sample:

$$NCE_p = TP_p - CP$$

The preliminary net error rate (NERp) is the ratio between the preliminary net coverage error and the preliminary estimate of the true population:

$$NER_p = \frac{NCE_p}{TP_p}$$

The ratio of the preliminary true population estimate to the Census population estimate is referred to as the first estimate of the adjustment factor, otherwise also known as the first estimate of the coverage correction factor:

$$AF = \frac{TP_p}{CP}$$

An improvement to the first estimate of the adjustment factor is achieved by post-stratification. After the creation of homogenous adjustment classes within which the coverage rates are the same, a second refined estimate of the adjustment factor (AF_R) is calculated independently within each class. At household level these classes are obtained by splitting the matched household population by province followed by geography type of the Sub-EA. At person level these classes are obtained by splitting the matched person population provincially and nationally by Sub-EA geography type, followed by sex, followed by age group, and lastly followed by population group. The splitting is carried out by a regression classification tree referred to as Chi-squared Automatic Interaction Detection (CHAID).

The final estimate of the true population (TP_F) is the product between the refined adjustment factor and the actual Census count (ACC):

$$TP_F = AF_R \times ACC$$

Precision of estimates

The estimates in this publication are based on a sample drawn from units in the surveyed population, making them subject to sampling error. This error is measured by the standard error (SE) of an estimate (E) either in absolute or relative terms where it is referred to as its coefficient of variation, and describes its precision. An estimate's precision may also be computed from its confidence interval (CI) usually set at a 95% confidence level assuming a normal distribution:

$$CI = E \pm 1.96 \times SE$$

For a known confidence interval, the standard error of an estimate is:

$$SE = \frac{UL_{CI} - LL_{CI}}{2 \times 1.96}$$

where:
UL_{CI} = Upper limit of the confidence interval; and
LL_{CI} = Lower limit of the confidence interval.

Its coefficient of variation (CV) is:

$$CV = \frac{SE}{E} \times 100$$

Content error indices

Content error is synonymous with response error, and refers to error in the characteristics as reported for those persons that were enumerated in the Census. It is only estimated for the matched persons for a number of marker variables such as Population group, Sex, Age group, Relationship to head of household, Marital status, and Country of birth, in order to measure the inconsistency between the captured answers from Census and PES to the same questions. The content error indices compare the Census responses with the PES responses in order to calculate the consistency or lack thereof between the two. In this way they indicate the extent of response bias, which is the systematic tendency to respond differently to the same question between Census and the PES. Content error is measured by means of four indices:
- Net difference rate, NDR;
- Index of inconsistency, I;
- Gross difference rate, GDR; and
- Rate of agreement, RA.

Net difference rate

The difference between the number of cases in the Census and the number of cases in the PES that fall under each response category, relative to the total number of matched persons in all response categories.

$$NDR = \frac{1}{n}(Y_{Ci} - Y_{PESi}) \times 100$$

where:
Y_{Ci} = unweighted census number of cases in i-th category;
Y_{PESi} = unweighted PES number of cases in i-th category;
n = unweighted number of matched cases; and
i = the i-th response category for characteristic "Y".

Index of inconsistency

The relative number of cases for which the response varied between the Census and the PES.

$$I = n\frac{(Y_{Ci} + Y_{PESi} - 2Y_{ii})}{[Y_{Ci}(n - Y_{PESi}) + Y_{PESi}(n - Y_{Ci})]} \times 100$$

where:
Y_{ii} = number of cases where the i-th category is given as a response in both Census and the PES.

For all the response categories of the characteristic as a whole, the aggregate index of inconsistency is:

$$I_{AG} = \frac{(n - \sum_{i=1} Y_{ii})}{\left(n - \frac{1}{n}\sum_{i=1} Y_{Ci} \cdot Y_{PESi}\right)} \times 100$$

Gross difference rate

The number of discrepancies between the Census responses and the PES responses relative to the total number of persons matched.

$$GDR = \frac{(n - \sum_{i=1} Y_{ii})}{n}$$

Rate of agreement

The complement of the gross difference rate, indicating the level at which the information given in the Census agrees with that given during the PES.

$$RA = \frac{\sum_{i=1} Y_{ii}}{n} \times 100$$

Interpretation of content errors

Criteria for the interpretation of the different content error indices:

Measure	Low	Moderate	High
NDR absolute value	<0,01	0,01–0,05	>0,05
I / I_{AG}	<20%	20%–50%	>50%

A low rate of agreement indicates a high degree of variability in the collected responses between Census and the PES, and vice-versa. The rate of agreement is therefore a good measure of the gross error for an item, which is the error due to carelessness or lack of experience.

GLOSSARY

Adjustment factor	An under- or overcount to increase and/or decrease the Census population as applicable.
Born after Census	Person who is present in the household on the PES reference night but who was not born on the Census reference night.
Census reference night	Night before the census day, for Census 2022 set as midnight between 02 and 03 February 2022.
Collection rate	Proportion of contacted units to the total number of units in the sample.
Confidence interval	The probable range of an estimate by adding and/or subtracting its absolute standard error calculated from sample values.
Coverage correction factor	Another name for adjustment factor.
De facto census	Census in which population units are enumerated according to where they were on census night.
De jure census	Census in which population units are enumerated according to where they usually live.
Dwelling	Structure intended or used for human habitation.
Dwelling unit (DU)	Structure or part of a structure or group of structures occupied or meant to be occupied by one or more than one household.
Enumeration	Process of counting members of a given population and collecting demographic and other information about each person. This counting takes place by means of administering a questionnaire to households and institutions in the country.
Enumeration area (EA)	Smallest geographical unit (piece of land) into which the country is divided for census or survey purposes.
EA number	Unique 8-digit ID number given to an EA for purposes of record-keeping and coding. The first digit indicates the Province. The next two digits indicate the District municipality and the last five distinguish among the different EAs within a Local municipality.
Final match	The match status assigned after a reconciliation visit.
E-sample	The evaluation sample drawn from the census corresponding to the P-sample.
Head/Acting head of household	Member of the household identified by the household as their head and/or any member of the household acting in the absence of the head of the household at the time of the interview.
Household	Person or group of persons who live/stay together and provide themselves jointly with food or other essentials for living.
Initial match	The match status assigned before a reconciliation visit.
In-mover	Person who was present in the household on the post-enumeration survey reference night, but who was absent on the census reference night.

Match	Record-by-record comparison of persons, households, and living quarters found in both the P-sample and E-sample, within predetermined tolerance.
Matching	Operation whereby households and persons enumerated during a census and a post-enumeration survey are compared for similarities and differences.
Moving status	An indication of a household or a person's whereabouts on the Census and PES reference nights.
Net coverage error	The difference between the estimated true population and the estimated census population, which may be an undercount or overcount.
Non-mover	Person who was present in the household on the post-enumeration survey reference night, and who was also present on the census reference night.
Out-mover	Person who was absent from the household on the post-enumeration survey reference night, but who was present on the census reference night.
Out-of-scope	Persons and/or households who do not belong to the target population as of the Census reference night.
P-sample	Shortened name for the PES sample.
PES reference night	Night before post-enumeration survey day, for PES 2022 set as midnight between 22 and 23 June 2022.
Primary matching variables	The main variables used to determine the match status of a person, which are first name and surname, date of birth, age, sex, population group, and country of birth.
Reconciliation visit	Follow-up visit carried out in the PES segments following initial matching in order to resolve doubtful or non-matched cases and identify erroneous Census enumerations.
Response rate	The proportion of eligible respondents who completed a questionnaire with usable information to the total number of eligible respondents.
Secondary matching variables	Supplementary variables used to determine the match status of a person if required, which are marital status and relationship to head/acting head of household.
Sub-EA	A smaller geographical area split from a larger EA for easier fieldwork management.
Target population	Set of elements about which information is sought and estimates are required. Also known as the scope of a survey.

REFERENCES

1   United Nations. (2017). *Principles and recommendations for population and housing censuses, revision 3.* New York: United Nations Statistics Division.

2   United Nations. (2010). *Post-enumeration surveys operational guidelines, technical report.* New York: United Nations Statistics Division.

3   United States Bureau of the Census. (1985). *Evaluating censuses of population and housing.* Washington DC: United States Department of Commerce.

GENERAL INFORMATION

Stats SA publishes approximately 300 different statistical releases each year. It is not economically viable to produce them in more than one of South Africa's 12 official languages. Since the releases are used extensively, not only locally but also by international economic and social-scientific communities, Stats SA releases are published in English only.

Stats SA has copyright on this publication. Users may apply the information as they wish, provided that they acknowledge Stats SA as the source of the basic data wherever they process, apply, utilise, publish or distribute the data; and also that they specify that the relevant application and analysis (where applicable) result from their own processing of the data.

Advance release calendar
An advance release calendar is disseminated on www.statssa.gov.za.

Stats SA products
A complete set of Stats SA publications is available at the Stats SA Library and the following libraries:

National Library of South Africa, Pretoria Division
National Library of South Africa, Cape Town Division
Natal Society Library, Pietermaritzburg
Library of Parliament, Cape Town
Bloemfontein Public Library
Johannesburg Public Library
Eastern Cape Library Services, King William's Town
Central Regional Library, Polokwane
Central Reference Library, Mbombela
Central Reference Collection, Kimberley
Central Reference Library, Mmabatho

Stats SA also provides a subscription service.

Electronic services
A large range of data are available via on-line services, diskette and computer printouts. For more details about our electronic data services, contact 012 310 8600.

You can visit us on the internet at www.statssa.gov.za.

ENQUIRIES

User information services	Telephone number: 012 310 8600 Email address: info@statssa.gov.za
Orders/subscription services	Telephone number: 012 310 8619 Email address: millies@statssa.gov.za
Technical enquiries	Mandla Masimula 012 310 8939 MandlaM@statssa.gov.za
	Riaan Grobler 012 310 3474 RiaanGr@statssa.gov.za
Postal address	Private Bag X44, Pretoria, 0001

Produced by Stats SA