askafrika
LIBERATING TRUTH FOR HUMANITY

OMC TRAVEL SURVEY
TECHNICAL DOCUMENT

OUT OF HOME MEASUREMENT COUNCIL

ROADSIDE OUTDOOR AUDIENCE DATA

OMC TRAVEL SURVEY
TECHNICAL DOCUMENT
2017-2024

askafrika
LIBERATING TRUTH FOR HUMANITY

Tel: +27 12 428 7400

www.askafrika.co.za

Table of content

1. Introduction and objectives

Introduction and objectives

Out-of-home (OOH) advertising is advertising that reaches consumers while they are outside their homes. Out-of-home advertising is focused on marketing to consumers when they are "on the go" in public places, in transit, waiting areas (such as in a medical office), and/or in specific commercial locations (such as in a retail venue, shopping malls). OOH advertising formats fall into four main categories of which this research is referring to billboard advertising.

The Out of Home Measurement Council (OMC), registered in early 2015, is a council who is representative of OOH media owners who commissioned Ask Afrika in 2014 to supply travel, demographic and branded product data as part of a solution to produce a common trading currency for roadside static media in South Africa. The data is provided to MGE (new modelling partners since 2022) who uses the travel survey as input in its travel models.

The main objectives of the survey are:

- To provide regular trip data for a representative sample of South Africans.
- To collect individual and household demographics that will be included in the final currency data set.
- To collect travel data that will be included in the MGE travel models.

The purpose of the document is to provide the technical detail behind the OMC travel survey across the research value chain i.e., methodology, questionnaire design, fieldwork, universe, sampling and data releases.

The information provided is based on the 2023A surveys.

Implication of COVID-19

Due to lockdown as a result of the Coronavirus epidemic data collection for OMC2020A was stopped and a decision was made by OMC that data collection for 2020 will be on hold for the duration for the 2020 year as well as for 2021. This was due to the major shift in travelling behaviour as a result of lockdown.

In 20A, n=2 828 interviews were conducted before lockdown and for the year 2021 no interviews were conducted. Data collection resumed in 2022 of which n=7 648 interviews were conducted. In 2023 n=7 583 interviews were conducted and in 2024 n=7 499 interviews were conducted.

Waves	Interviews completed
2020A	2 828
2020B	On hold due to COVID-19
2021A	On hold due to COVID-19
2021B	On hold due to COVID-19
2022A	7 648
2023A	7 583
2024A	7 499

Travel Survey sample design change

Changes for the year 2017 - 2020

The travel survey's sample design changed since 2017. Community size, which was an old AMPS area classification was replaced with Statistics South Africa's area classification, namely the Geo-type. The Statistics SA area classification is more in-line with current currency projects in the industry.

A decision was made to exclude economic uncommercial areas due to the small sample allocation given to these areas annually. Thus, the third change was the exclusion of 14 of the 52 District Municipalities as of 2017.

Similar to previous years a stratified, multistage sample design is employed. Also, similar to previous years, is the use of the Enumeration Area database, which currently enables the lowest level of sampling in the Country. It is regularly updated by one of South Africa's most recognise sampling expert, Dr Ariane Neethling together with two GIS specialists and Geo Terra Image.

Changes for the year 2022 and future

A decision was made by Ask Afrika to migrate high income area interviews with a Land use Classifications being Retirement Village, Townhouses, Estates and Security Villages to an online panel methodology. High income areas with a land use classification being a Free Hold Formal/Informal dwelling was still conducted face-to-face. Face-to-face interviews was also done for Land use Classifications being Retirement Village, Townhouses, Estates and Security Villages in Mid to Low-income areas.

The decision was based due to the difficulty in interviewing respondents in high income areas, gated communities and complex dwellings. COVID-19 also resulted in an increased refusal rate of these areas. Entering these areas became even more difficult and previous measures taken to assist with these areas became ineffective.

The biggest challenges with a mixed methodology approach are the sample design as well as the weighting of data afterwards. In order to ensure that a representative and probability sample could be applied for the online proportion, Ask Afrika had to ensure that a big enough online panel could be utilised as well as the linking of these panel members to the sampling frame, namely the EA sampling frame.

Ask Afrika had numerous discussions with Dr Neethling, Ask Afrika's sampling expert as well as numerous online panel providers. After many considerations, CINT was selected as the online panel research partner.

The sample design was done similar to the previous years, where Dr Neethling designed a stratified, multistage sample from the EA sampling frame including high income areas. The sample was designed as per normal. After the EA's were drawn, the sampling points were selected. These selected EAs with sampling points were then linked to the land use classifications of GeoTerra Image (GTI), Ask Afrika's GIS partner as well as the co-owner with Dr Neething of the EA frame. Dwelling classifications were linked to each sampling point.

High income areas were identified by making use for the Neighbourhood Lifestyle Index (NLI) which is an economic income indicator designed by GTI. Sampling points falling within NLI8-10 were identified. These sampling points together with the Land use Classification of sampling points being Retirement Village, Townhouses, Estates and Security Villages were excluded from the face-to-face sample design and was used for the online portion.

The linking of these online sampling points to the online panel were done on a Provincial, District Municipality and Main-Place level as these information points are available in the CINT data file as well as within the EA frame. Strata as designed during the initial stage was created within the CINT data file, thus a quota sampling approach was followed. Online invitations were sent to respondents falling within these strata to full the strata quotas.

The randomness of the online interviews can however just be ensured on a Main-Place level. Online invites are submitted to the full panel base and completed interviews are controlled only by the set quotas. This being said, with an active panel base of over one million, the representivity of interviews are very high.

Face to face interviews were conducted for:

- High income areas with a Land Use Classification being a Free Hold Formal/Informal dwelling was still conducted face-to-face.
- Mid to Low-income areas with a Land Use Classifications being Retirement Village, Townhouses, Estates and Security Villages in.
- All mid to low income.

The split between the face-to-face and online interviews are as follow:

Methodology	2017	2018	2019	2020/2022	2023	2024	17-24
Face-to-face	15 152	15 114	15 195	9 908	7 077	6 991	**69 437**
Online				568	506	508	**1 582**
Total sample per year	15 152	15 114	15 195	10 476	7 583	7 499	**71 019**

Sample differences 2016 and before versus 2017-2019 and 2022 and future

Due to 2020 not being a completed wave due to lockdown, 2020 is not covered in the table below.

	2016 and before	2017 - 2019	2024
Waves per year	2	2	1
Annual sample size	15 000	15 000	7 500
Sample size per wave	7 500	7 500	7 500 only 1 wave
Data collection methodology	Face-to-Face	Face-to-Face	Face-to-Face: 93% Online: 7%
OMC interviews	Number of EAs per wave: 1,875 Sampling points per EA per wave: 4 OMC interviews per wave: 4 per EA	Number of EAs per wave: 2,500 OMC interviews per wave: 3 per EA	Number of EAs drawn for the face-to-face portion: 1,793 EA's Online interviews: 508
Data collection period	2 waves: February - June & July - November	2 waves: February - June & July - November	1 wave: February - May
Data releases	Bi-Annual	Bi-Annual	Annual
Sampling frame	EA data base updated to previous year's Statistics SA population estimates	EA data base updated to previous year's Statistics SA population estimates	Face-to-Face: EA data base updated to previous year's Statistics SA population estimates Online: Online panel, active members accounting for R1.5 million individuals
Sampling frame updates and maintenance	**1.** Dr Ariane Neethling (Modelling and Benchmarking of population estimates to EA frame based on Statistics SA's mid-year population estimates) **2.** Geo Terra Image (Specialist geospatial mapping, Remote sensing services (focused on earth observation, satellite imagery and aerial photography, updated imagery and aerial photography of dwelling unit counts and land use information, information on areas development,	**1.** Dr Ariane Neethling (Modelling and Benchmarking of population estimates to EA frame based on Statistics SA's mid-year population estimates) **2.** Geo Terra Image (Specialist geospatial mapping, Remote sensing services (focused on earth observation, satellite imagery and aerial photography, updated imagery and aerial photography of dwelling unit counts and land use information, information on areas development,	**Face-to-Face** **1.** Dr Ariane Neethling (Modelling and Benchmarking of population estimates to EA frame based on Statistics SA's mid-year population estimates) **2.** Geo Terra Image (Specialist geospatial mapping, Remote sensing services (focused on earth observation, satellite imagery and aerial photography, updated imagery and aerial photography of dwelling unit counts and land use information, information on

	dwelling unit information through aerial photography) **3.** AfricaScope (Use spatial data for additional updates that reflect new building developments and urbanisation not identified by Statistics SA, Produce the latest population estimates at an EA level)	dwelling unit information through aerial photography)	areas development, dwelling unit information through aerial photography) **Online** Panel consisting of 1.544 million individuals distributed as per Annexure A
Sampling approach	National representative of South Africa (data reflects SA's diverse country) – communities 8,000+, excluding rural	National representative of South Africa (data reflects SA's diverse country) excluding non-urban rural areas (farms and traditional)	National representative of South Africa (data reflects SA's diverse country) excluding urban and non-urban rural areas (farms and traditional)
Sampling methodology	Stratified, multistage sample design	Stratified, multistage sample design	Stratified, multistage sample design
Stratification Variables	Primary stratification variables: Community size by Province, Secondary stratification variables - population group	Primary stratification variables: Metro-Urban/Metro-Rural/Non-Metro Urban Secondary stratification variables – Province by population group Tertiary stratification variables: Community size 8,000+ by Non-Metro Urban	Primary stratification variables: Urban Metro and Urban Non-Metro Secondary stratification variables – Province by population group Tertiary stratification variables: Community size 8,000+ by Urban Non-Metro
Sampling level	EA level which is the lowest SA Geographical level	EA level which is the lowest SA Geographical level	**Face-to-Face** EA level which is the lowest SA Geographical level **Online** Sampling done on District Municipality level based on the F2F sampling structure and design
Area Classification	Community sizes: 1~250 000 Plus 2~100 000-249 000 3~40 000-99 999 4~8 000-39 999	Three area types 1~Urban Metro **(8 Metropolitan areas)** 2~Rural Metro **(Farms in Metropolitan areas)** 3~Urban Non-Metro **(Urban areas at the outer skirts of Metropolitan areas)**	Three area types 1~Urban Metro **(8 Metropolitan areas)** 2~Urban Non-Metro **(Urban areas at the outer skirts of Metropolitan areas)**

Rural	Excluded	Metro-Rural was included as of 2017 Geographical splits as classified by Statistics SA are as follow: -Metro (69.96%), -Metro-Rural (2.02%) -Urban Non-Metro (28.02%) The traditional geographical classification which covers deep rural areas, are excluded, whereby the geographical classification, urban and farms will be included	Geographical splits as classified by Statistics SA are as follow: -Metro (70%), -Urban Non-Metro (30%) The traditional geographical classification which covers deep rural areas, are excluded, whereby the geographical classification, urban and farms will be included

2. OMC research methodology

2.1 Research design

A cross sectional survey design was employed as of 2020/2022 using both personal face-to-face and online interviews. For the face-to-face section, a probability sample is drawn, and respondents are interviewed in their homes on Android powered tablets. The respondent is incentivised in the form of a lucky draw where a monetary sum of up to R5,000 is on offer. Data collection is continuous between the months of data collection. Interviews are conducted across all days of the week.

Average completes per week

Day of the week	F2F Completes	
Mon	9 145	13%
Tue	11 481	17%
Wed	12 067	17%
Thurs	11 881	17%
Fri	10 543	15%
Sat	9 000	13%
Sun	5 320	8%
Total	69 437	
Online interviews	1 582	
Total	71 019	

2.2 The OMC questionnaire

The flow of the questionnaire is as follows:

- Household selection and kish grid.
- Demographic section.
- Travel survey.

The face-to-face interview is critical for establishing the demographics of the main respondent as well as the demographics of the households in general. The face-to-face interview is completed via personal interviews with the randomly selected participants using Android powered tablets. Where appropriate, interviewers will hand over the device showing the question to the respondent to clarify certain concepts, aid recall and to introduce sensitive questions. The benefit of using tablets is that they are small enough not to be intrusive, but big enough to present a readable questionnaire to the interviewer and respondent.

The face-to-face interview is pre-programmed, so the interviewer does not need to focus on routing or look out for logical inconsistencies, as this is dealt with automatically by the software. This allows the interviewer to focus on building rapport with the respondent which will result in better quality data. The OMC questionnaire takes about 15-20 minutes to complete, depending on the number of destinations chosen by the respondents and is fully quantitative.

The interview is predominantly conducted in English. Due to interviewers residing in the areas they conduct interviews in they do at stages need to explain a word or phase in their own vernacular language to make the question more understandable to the respondent.

The questionnaire includes:

- Details regarding trips taken in the past 7 days in terms of:
    - Destination/Purpose of trip
    - Mode of transport
    - Frequency of trip
    - Time of trip (start/return trip)
    - Duration of trip (start/return trip)
    - Start and end address of each trip
    - Sequence of destinations

2.3 Questionnaire changes

Historic

The travel research conducted in 2014 was used as the pilot phase of which learnings were taken and implemented in 2015. The biggest challenge experienced in 2014 was the understanding of the meaning of a trip. Secondly was the sequential layout by the respondents of trips taken. Based on this the 2015 questionnaire was changed significantly to implement the learnings from 2014. Discussion and training topics during field briefings 2015 and 2016 was also handled differently with more emphasis placed on findings during the 2014 data collection period as well as any learnings from the preceding year.

The questionnaire changes implemented from 2017 and 2022 were based on year-on-year findings from the previous year. Questions were added to understand non-travel behaviours as well as the refusal of address information. Changes were not only implemented due to finding a better way of asking a question but also to enrich the data. Taxi rank footprints were included and later extended. Similar improvements were implemented to the Gautrain mode of transportation. In the destination specific sections, duration of a specific trip was added.

The SEM as classified by The Broadcaster Research Council (BRC) is an addition variable added to the OMC data.

For 2019B no questionnaire changes were made, however a pilot on the income question was done. The refusal answer option was taken out and was only displayed to the interviewer to use when needed. It was done to test whether having a refusal option made it easier for respondents to automatically select the option. The pilot shown a decreased in the refusal rating and a decision was made to change the income question from 20A.

A further change to the income questions were done in 2022 where respondents who refused to answer the income questions (personal and household income) are routed to a larger range of income selections. This also yielded a decrease in the refusal rates.

In 2023, a few logical checks were implemented to enhance the quality of trips:

- Specific destination numbers (e.g., Shop 1, Shop 2, etc.) were added to the questionnaire, resulting in a significant reduction in GAPS.
- Built-in logic was introduced to ensure that all destinations are mentioned in a start and end trip, leading to a considerable decrease in GAPS.

2.4 Data releases

Due to the change in the number of waves for OMC, from two waves to one wave, the data was released in September 2024. This dataset is then combined with previous releases to create a rolling sample release that includes 2017,18,19,20,22, 23 and 24.

2.5 Universe

The universe is defined as people aged 15 years and older in South Africa, living in households, hostels or similar accommodation, in communities with 8000 or more inhabitants. Institutions (for example hospitals and prisons), industrial areas, recreational areas and farms are excluded. In addition, 14 District Municipalities were also excluded as of 2017B from the universe because of annual small sample allocations to these District Municipalities. These District Municipalities are as follow:

1. Zululand
2. John Taolo Gaetsewe
3. O.R.Tambo
4. Sisonke
5. Amathole
6. Umzinyathi
7. Mopani
8. Namakwa
9. Joe Gqabi
10. Central Karoo
11. Xhariep

12. Greater Sekhukhune
13. Alfred Nzo
14. Umkhanyakude

2.6 Sample frame

A sampling frame is a complete list of sampling units that entirely covers the target population. The existence of a sampling frame allows a probability selection of sampling units. For a multi-stage survey, a sampling frame should exist for each stage of selection. The sampling frame used in the OMC travel survey is as up to date as possible (incorporating the Statistics SA' mid-year population estimates). The sampling frame covers the whole survey area (all adults 15 years and older in communities 8000+).

The sample frame used for OMC is the Enumerated Area (EA) frame. This sampling frame is jointly owned by Dr Ariane Neethling and GTI. For the 1996 and 2001 population census, South Africa was subdivided into Enumerator areas (EAs).

An EA is a small geographic unit of a manageable size (in terms of population and land area) assigned to a single person to enumerate during the census count. In their creation, it is vital to ensure that they adhere to certain standards. These standards include, among others, the following:

- EA's cover all areas of the country, i.e., no omissions and no overlapping of EA's
- EAs are accessible and of such a size that field workers can carry out their tasks within stipulated period of census taking
- EAs are as homogeneous in land use as possible

The determination of size is influenced mainly by the EA's geographical location and population density. In thinly populated areas, the EAs tend to be large in terms of area but have a low dwelling unit count of approximately 100, whereas the densely populated areas tend to have small area sizes with a dwelling count of approximately 151.

The EA sampling frame consists, for each EA, its demographic information (main-place, sub-place, municipality, district, province, EA-type, EA-Geographic type, community size, estimated population counts of the number of households, number of people as well as numbers per population group, gender and per five-year age interval). The current EA sampling frame, based on StatSA's lists of main place, sub-place and EA numbers for 2022, consists of 14 039 main places over the whole of South Africa. These main places are sub-divided into 22 108 sub-places, which are further divided into 103 576 EAs.

2.7 The current universe and realised sample

The defined universe (South African population, aged 15+ years in communities 8000+) for this survey represents 28,681,110 adult South Africans which is 47% of the total South African population of 60,414,495. These numbers are based on the 2022 StatsSA Mid-year estimates.

Universe versus sample

The tables below show the weighted and realised sample characteristics of OMC2017-2024A, aged 15+ years in communities 8000+.

The weighted sample frame by population group:

	BLACK	COLOURED	ASIAN	WHITE	TOTAL
Population	20 795 487	3 273 448	1 206 675	3 405 500	28 681 110
Gender	**MALE**	**FEMALE**			**TOTAL**
	14 140 064	14 541 046			28 681 110
Age	**AGE 15 to 24**	**AGE 25 to 34**	**AGE 35 to 49**	**AGE 50 to 64**	**AGE 65+**
	5 915 154	7 191 804	8 506 201	4 666 381	2 401 570
Province	**BLACK**	**COLOURED**	**ASIAN**	**WHITE**	**TOTAL**
EC	1 206 337	246 613	12 302	273 254	1 738 505
FS	1 470 934	52 196	0	118 346	1 641 476
GP	10 099 140	323 842	372 823	1 384 873	12 180 679
KZN	2 654 908	51 831	743 161	334 198	3 784 098
LP	470 096	11 800	3 810	52 991	538 697
MP	1 714 664	210	2 662	127 877	1 845 413
NW	1 357 880	2 687	5 580	178 070	1 544 216
NC	284 042	147 096	0	25 102	456 240
WC	1 537 485	2 437 173	66 337	910 790	4 951 786
TOTAL	**20 795 487**	**3 273 448**	**1 206 675**	**3 405 500**	**28 681 110**

The realised sample by population group:

	BLACK	COLOURED	ASIAN	WHITE	TOTAL
Population	37 893	11 166	6 242	15 718	71 019
Gender	**Male**	**Female**			
	32 386	38 633			71 019
Age	**AGE 15 to 24**	**AGE 25 to 34**	**AGE 35 to 49**	**AGE 50 to 64**	**AGE 65+**
	13 015	20 321	23 240	10 315	71 019
Province	**BLACK**	**COLOURED**	**ASIAN**	**WHITE**	**TOTAL**
EC	3 038	1 517	191	1 431	6 177
FS	2 972	343	4	845	4 164
GP	17 710	2 231	1 952	6 743	28 636
KZN	5 367	627	3 501	1 720	11 215
LP	1 108	53	96	273	1 530
MP	2 503	96	98	786	3 483
NW	1 911	190	134	867	3 102
NC	515	552	24	297	1 388
WC	2 769	5 557	242	2 756	11 324
TOTAL	37 893	11 166	6 242	15 718	71 019

2.8 Updates to the current sampling frame

Updates

Stats SA and GTI data are used to update the EA sampling frame to produce the latest population estimates at an EA level. The EA sampling frame is updated at least once a year when the new midyear population estimates are released by Stats SA.

The latest information gathered by GTI, from fieldwork reports, and other sources, is also used to update/adjust the sampling frame to reflect the latest population, housing and other

developments. Stats SA's mid-year estimates include updates on migration, HIV, death rates, birth rates etc.

Benchmarking techniques are applied in the updating process to ensure that the sum of the values in the EA sampling frame are equal to the estimates obtained for the variables in the midyear population data. These population estimates are the only official statistics in South Africa and form the base of the EA sampling frame.

GTI/EA frame updates

Dr Neethling update the EA frame twice a year. The first update is done after the June/July mid-year population release from Stats SA at a Provincial levels as well as Nov/Jan after Stats SA release the District Municipality data.

The updates are done as follow:

- GTI track changes or differences seen on an EA level on a continues basis.
- GTI typically look at dwelling type changes, area changes as well as population differences on an EA level

These changes are logged within the EA frame and Dr Neethling use this feedback to update the EA frame twice a year.

Dr Neethling indicated that it is very seldom that Provincial and District Municipality totals of the EA Frame do not tie up to Stats SA's totals. It is more on a lower level where a lot of changes appear.

Stats SA Mid-Year population estimates

Stats SA updates Provincial level data annually based on a cohort-component methodology. A cohort component technique uses the components of demographic change to project population growth. The technique projects the population by age groups, in addition to other demographic attributes such as sex, ethnicity etc. Migration, life expectancy, infant mortality, fertility and HIV prevalence annual updates are also incorporated.

No reference was made to any primary research used, such as the Quarterly labour forces of Household surveys conducted by Stats SA for the population estimates.

2.9 Sample design

A Multistage Stratified Probability Sampling methodology using probability proportionate to size (PPS) is used in the design of the TGI sample. Province (Gauteng, Free State, Limpopo, Mpumalanga, North-West, KwaZulu-Natal, Eastern Cape, Western Cape and Northern Cape), Geo-type (Urban-Metro and Urban Non-Metro-Urban) and population groups (White, Black, Coloured and Indian) are taken as the stratification variables to ensure the best possible precision for province, geo-type and population groups.

Sample design

- A stratified, multistage probability sample design
- Probability Proportion to size (PPS) with number of households as measure of size (MOS)

Stratification and allocation

- 3 Explicit Stratification variables: Primary strata: Geo-type by Province and Population group (In 2022 the Neighbourhood Lifestyle Index (NLI™©) *** was included as a stratification variable)

- Implicit Stratification variables: District Municipality, local municipality, EA Type, main-place, sub-place and EA within each Explicit Strata

Selection of units and weighting

- 1,793 EAs are selected per wave
- 6 991 HH points were selected for 2024. These points are selected systematically from a random starting point
- 90% of the selected EA need to complete 4 OMC interviews, whereas the remaining 10% selected EA's completed 3 OMC interviews
- Interviewers use a Kish grid to select households and main respondent
- GIS maps and GPS co-ordinates supplied to interviewers
- Substitution rules x+1, x-1 employed
- Weighting variables: District Municipality, Race, Age group and Gender
- Design weighting to correct the unequal inclusion probability of respondents
- RIM weighting to adjust the design weights to the demographic population estimates

***The Neighbourhood Lifestyle Index (NLI™©) are designed by Dr Neethling and GTI. The NLI™© is not a sample survey-based product but is statistically segmentation and modelled from population dwelling unit information and various other sources of information such as Credit Bureau & Stats SA. This unique product is an income based segmented classification, classifying neighbourhoods according to their income and various lifestyle characteristics represented at Enumeration Area level.

The NLI was added to the stratification list of variables. When population characteristics are diverse, such as income, one wants to ensure that every characteristic is properly represented in the sample. This helps with the generalizability and validity of the study, as well as avoiding research biases like under coverage bias.

2.10 Disproportionate sampling and power allocation

Due to the significant differences in the population sizes between the strata, it is meaningful to deviate from proportional allocation. The power allocation rule is used to determine the number of EAs to be drawn per strata. The power allocation will thus determine the optimum sample size per province, community size and population group, in each of the primary stratification split, without sacrificing the precision.

The aim of using the power allocation rule is to somewhat decrease the allocation of EAs to the larger strata and to increase somewhat the allocation to the smaller strata. Hence, with this technique one can ensure, as far as the overall sample size allows, that the sample sizes are large enough in each stratum representing the different provinces and population groups within the sample frame.

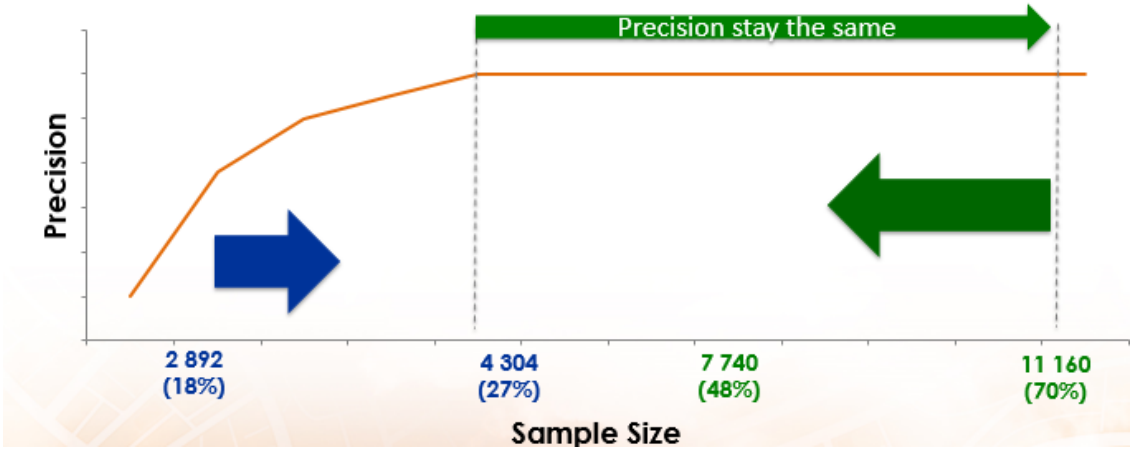
Since a sample is drawn independently in each stratum, the design in each of the nine primary strata - Provinces – will be considered separately. Different disproportional allocation, by using different power allocations were used to ensure the best possible coverage of the elements.

Why do Ask Arika employ a disproportional allocation on a population group level

Ask Afrika employ a disproportionate sample. A disproportionate sample is drawn to ensure a more optimal allocation when the strata (e.g., provinces, race groups) differ in size. One need to ensure optimal sample sizes to ensure better precision for the smaller strata. Thus, Ask Afrika boost smaller groups to get a better precision. Proportional allocation does not necessarily give optimal allocation.

Proportional allocation - Sample proportions are a true reflection of population proportions
Disproportional allocation – Sample proportions are not a true reflection of population proportions; some strata are over sampled, and some are under sampled
Use proportional allocation when:
* The strata are equal *(e.g., Provincial sizes, population group totals etc.)*
* If you are interested in overall conclusions, thus no interlocking quota data conclusions e.g., between Provinces, Province by population groups etc.
Use disproportional allocation when:
* The strata vary greatly in size *(e.g., Provincial sizes, population group totals etc.)*
* The variation within strata differs (some strata are more heterogeneous compared to other, thus the variances differ)
* If you want to report and make conclusions on a lower level than just on an overall level

Example: Population groups, Black and White



2.11 Split samples for gender

Given the nature of South African society some level of control over gender groups needs to be maintained in order not to have a bias towards one gender group that might be more likely to be at home at the time of the interview. The number of interviews to be conducted in each EA was randomly allocated to each gender group. There are an equal number of males and females in the sample.

2.12 Sampling points selection

The EAs in each of the above explicit strata are ordered according to EA type, Main-Place, Sub-Place and EA number upon which the predetermined numbers of EAs are drawn using PPS (i.e., probability proportional to size) systematic sampling with the number of persons 15 years and older per EA as measure of size. In each of the drawn EAs, 3/4 house points are selected with equal

probability. As all dwelling types is known in each EA, a systematic approach can be followed to determine the 3/4 sampling points per EA.

2.13 Selection of dwelling, household and respondent

The success of a probability sampling procedure depends on the success of the final stage of the sample design, namely the random selection of the main survey respondent. In each of the drawn EA's, 3/4 households need to be selected from which one adult 15 years or older will be interviewed.
The GPS coordinates of the four selected interviews per EA are provide for navigation purposes. These coordinates are also Geofenced in order to ensure interviewers are targeting the right sampling point. The Geofencing ensures that the questionnaire open up on the Tablet only once the interviewer is within the boundary of the Geofence.

The Kish grid is employed to select the dwelling, household and respondent if more than one of these with the predetermined gender is present at the visiting point. The Kish Grid is a well-established means of random selection and is widely used in the market research sphere. Only one person 15 years and older per household will be required and allowed to complete the questionnaire.

Dwelling selection
When more than one house/dwelling is located on the selected visiting point, the interviewer records the number of houses on the stand (from left to right). The interviewer then refers to the Kish grid to identify the dwelling that should be targeted for the interview. A dwelling unit is defined as a home, residence or sleeping place that is occupied by at least one person.

Household selection
When more than one household is located in the selected dwelling unit within the predetermined gender the interviewer records the number of households staying in the house. The interviewer then refers to the Kish grid to identify the household that should be targeted for the interview. A household is defined as people living in the same dwelling and usually eating together for at least 4 nights per week or more.

Respondent selection
When there is more than one individual within the predetermined gender aged fifteen and older in the household, the interviewer records the number of people by looking at the ages and list it from youngest to oldest. The interviewer then refers to the Kish grid to identify the individual that should be targeted for the interview.

The Kish Grid uses a pre-assigned table of random numbers to find the person to be interviewed.

The kish grid instructions are as follow:

1	2	3	4	5
The interviewer counts the number of dwellings on the stand. A dwelling	The kish grid is used to determine which dwelling to interview.	The kish grid is used to determine which household to interview.	The interviewer lists the ages of these people from youngest to oldest.	The interviewer reintroduces him/herself to the respondent (if necessary) and

is defined as a structure that is used as a home, residence or sleeping place by at least one person	After the dwelling is selected the interviewer then counts the number of households in the dwelling - within the predetermined gender. A household is defined as people living in the same dwelling and usually eating together for at least 4 nights per week or more	After the household was selected the interviewer then counts the number of people living in this household who are older than 15 - within the predetermined gender	The kish grid is used to determine which respondent to interview.	commences with the survey

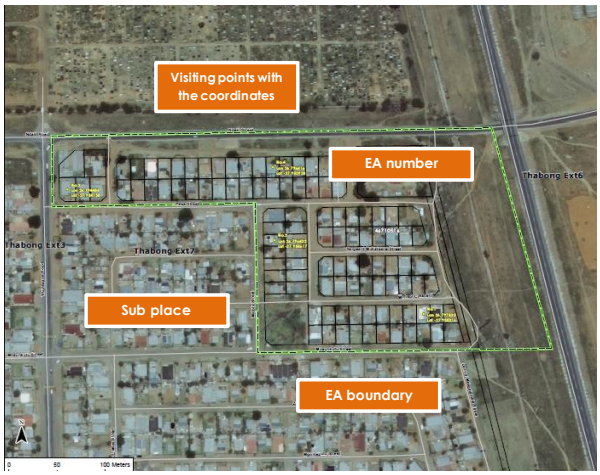
2.14 EA map production

EA maps are used for the selected EAs. These maps clearly indicate the EA boundaries as well as street names. The GPS coordinates of the EAs on the maps ensure that the fieldworkers go to the correct area. In areas with unstructured boundaries/street names, more emphasis was on the provided GPS co-ordinates which identify the selected 3/4 points within the selected EAs.

The EA maps indicate the following information:

- A density map of the selected EA which indicate the population distribution.
- Black and white maps providing navigational information into the EA interviewing points (systematic allocated) per EA according to EA density
- Preselected systematic interviewing points are allocated on each EA map. This show the exact location of interviewing (GPS coordinates will also be indicated especially in more informal settlements).

Image 1: EA Map information



2.15 Mix population groups within an EA

A 'perfect' sample is a sample that mirrors all the characteristics of the population, and it is important that enough people of all the population groups are included in the sample. Since the population size of the different population groups deviate largely in size the only way to 'control' the number of each population group in the sample, is to include population as a stratification variable and to apply power allocation.

Many EAs consist of more than one population group. The population group with the largest population number in that EA is taken as the dominant population group.

From experience over the years, Ask Afrika knows that if one does not control population group, especially the Indian and White population groups, the chance is high that one ends with very small numbers of these two population groups in a sample. The method that Ask Afrika uses, is that for EAs with dominant White and Indian (and one can add Coloured) population groups, only persons from that specified dominant population group can be interviewed in that EA. Since persons from other population groups are thus 'excluded' in these EAs, the fieldworkers must keep a record of the number of households that are 'skipped' because of a "wrong" population. To calculate the design weight scientifically in these EAs, these 'skip' numbers are included in the formula of the household design weight. This is a statistically sound and practically proven method with very satisfactory results.
Population group is a stratification variable and is automatically built into the design weight. Population group is also part of the RIM weighting and will be representative of the population figures of all individuals 15 years and older in communities 8000+. To note, all respondents with the predominant population group within the EA's will be given an equal or know opportunity to be included in the sample

2.16 Substitution and response rates

Substitution

In many surveys, non-response is handled by substituting households that are available or willing to respond. The substitution procedure is biased in the same way as non-response cases are. There is nothing to be gained in statistical accuracy by substitution over what would be achieved by over-sampling to account for non-response in the first place. The substitution method does, however, achieve the targeted sample size, thus rendering it unnecessary to over-sample for non-response. In this way sampling reliability is the same for the two methods of dealing with non-response, because the ultimate sample size in terms of completed interviews is the same (Turner, 2003).

For cost and pragmatic reasons, we believe the house to the left or right method (x+1, x-1) is suitable for the OMC and does not introduce any additional bias to the survey.

Reasons for substituting:

- The interviewer gets to the specified address and there is no household on that stand.
- The interviewer contacts the address three times to set up an appointment and each time there is no one at home.
- The interviewer attempts at least three times to contact the selected respondent, but to no avail.

- The interview is refused by the respondent.
- The interviewer cannot gain access to security estate after all attempts to obtain the interview failed.

The higher the response rate the better generally.

OMC2017-2024A Average response rates

Province	Original %	Substitute %
Eastern Cape	94%	6%
Free State	93%	7%
Gauteng	80%	20%
KwaZulu-Natal	87%	13%
Limpopo	87%	13%
Mpumalanga	90%	10%
North-West	90%	10%
Northern Cape	98%	2%
Western Cape	93%	7%
Grand Total	**87%**	**13%**

It is important to note that when substitution occurs, that a replacements point is drawn within the same area. The exact same profile of the substituted point is replaced.

2.17 Difficult to reach areas

Ask Afrika experienced refusal of access to several security estates, boomed areas and complexes in the past. As indicated in the beginning of this document, high income areas with a Land use Classification of sampling points being Retirement Village, Townhouses, Estates and Security Villages were excluded from the face-to-face sample design and online interviews was conducted. High income areas with a land use classification being a Free Hold Formal/Informal dwelling was still conducted face-to-face.

Although data collection in difficult to reach high income areas was migrated to the online platform, there was still dwelling in gated communities and complexes/estates/townhouses dwelling types in the middle to low-income areas. Similar procedures to the past were still implemented. Contact with body corporates or facility care taker was made and the reason for our interest was explained. This yet again led to successful interviews being completed. In certain cases where the body corporates required an Ask Afrika manager to speak to them, was handled on a case-to-case basis.

Stringent steps to following the sampling plan have been in place.

- Careful planning once the EA's are drawn and the maps are delivered to ensure that each area is accessible and fieldwork timing is planned
- In-Office/Out-of-Office team phoning and arranging entry into difficult to reach areas

2.18 Weighting

Universe used for the OMC 2024 data

For weighting purposes, Ask Afrika consistently utilizes the latest official population data provided by Statistics SA. The most recent data set available is from the 2022 Census. However, there has been significant concern among industry professionals and statistical experts about the dependability and quality of this data. Consequently, Ask Afrika has opted not to use the 2022 Census data but rather the Midyear Population Estimates from 2022, also published by Stats SA. This decision stems from three major concerns regarding the accuracy and reliability of the Census data.

1. **Potential over-correction of population estimates**: Concerns are expressed that the reported population figure may have been over-corrected by an estimated one million people, which raises questions about the accuracy of the adjustments made for census data. Such inaccuracies in census data can lead to flawed weighting in our survey analyses, potentially skewing research findings. This misalignment may result in delivering misleading insights to clients, who then base their strategic decisions on incorrect data, affecting their market positioning and investment strategies.
2. **Distortions in population growth figures**: Adjustments for undercounts may have misshaped the actual population growth patterns at the district and municipal levels, which could result in inaccurate market research findings. This misrepresentation can cause clients to misdirect investments and concentrate efforts on non-strategic areas, resulting in inefficient spending and lost opportunities
3. **Insufficient sample size for post-enumeration survey**: The size of the sample used for the post-enumeration survey (PES) is criticized as too small, potentially compromising the reliability of the data used to adjust the census results. A limited sample could undermine the reliability of census adjustments, potentially affecting the accuracy of the data we use for demographic weighting in our market analyses. Consequently, this can lead to erroneous forecasting and targeting, impacting the effectiveness of market strategies advised to our clients.

OMC 2024 Weighting

Weights are assigned to make weighted sample records represent the target population as accurately as possible. A weight indicates the number of population elements "represented" by a single sample element and is used to generalise the sample information to the population.

The design weights for the realised sample data were calculated and, after compensating for non-response, the weights were benchmarked to resemble the 15+ TGI/OMC population total according to the 2022 (latest released) midyear estimates of Stats SA.

To obtain a representative sample of the population, a stratified multistage sample was designed with the province, community size groups, population group and NLI groups (income-based groups) as the explicit stratification variables. The largest part of the sample was executed by using a face-to-face (CAPI) survey, with a small portion (6.7%) being done through an online survey.

The majority (93.2%) of respondents were done face-to-face and their design weights were calculated, as in the past for this group of respondents – see section 1. Hereafter the online

respondents' weights were calculated (see section 2), whereafter the final benchmarked weights were calculated for all respondents combined (see section 3).

1. Weights for Face-to-face respondents

For the CAPI survey, in the first stage primary sampling units (PSU) – i.e. the enumerated areas (EA) – were selected with probability proportional to size (PPS), with the number of households in the EA as the measure of size. The weight of an EA is given by

$$W_{PSU} = \left(n_{EA}\frac{PSU_{HH}}{POP_{HH}}.\right)^{-1},$$

where n_{EA} is the allocated number of EAs in the stratum, PSU_{HH} the number of households in the selected EA and Pop_{HH} the number of households in the selected stratum.

In the second stage, households are selected systematically in each PSU in the sample. The household weight per PSU is given by

$$W_{HH} = W_{PSU}\left(\frac{n_{HH}}{PSU_{HH}}\right)^{-1},$$

where n_{HH} is the number of selected households per PSU and PSU_{HH} is the number of households in the selected PSU. In the final stage, a person aged 15 years or older is selected from the drawn household. The respondent weight is given by

$$W_{PP} = W_{HH} * Av_{15+},$$

where Av_{15+} is the average number of persons aged 15 years and older per selected household in the EA. Since large differences were found regarding the number of persons aged 15+ years per household, which will cause too large deviation in the weights, the average number of persons aged 15 years and older in the selected households of an EA was used in the weighting process. This method is commonly used to smooth the weights.

2. Weights for Online respondents

For those strata with online panel respondents, the design weight of a respondent within the stratum is calculated by the formula that simplifies to

$$W_{PP} = \frac{POP_{Stratum}}{n_{Stratum}},$$

where $POP_{Stratum}$ is the estimated population number of persons aged 15+ years in the stratum and $n_{Stratum}$ the number of respondents in the stratum.

It should be noted that the sample can only be as good as the panel database. If the latter is skewed in some way, it could affect the sample and results.

In strata where both face-to-face and online respondents were found, the population totals applied in the above formulas were calculated according to the proportion of respondents in the two groups.

3. Final weights

To obtained integrated weights for the face-to-face and online respondents, the design weights of the face-to-face and online respondents were combined and simultaneously benchmarked to the estimated population totals of the TGI/OMC inclusive population of persons aged 15 years and older.

The design weights of the respondents were adjusted to compensate for under/over-representation of certain parts of the population, by using calibration techniques. The SAS macro, CALMAR, developed by INSEE in France, was used. The variables, District Municipality. Age, Gender and Population Group, based on the 2022 midyear estimates of StatsSA were considered.

2.19 Weighting Efficiency

Weight efficiency is a measure in survey sampling that assesses how effectively a sampling strategy utilizes assigned weights to represent the target population. Balancing the need for a representative sample with the goal of achieving high weight efficiency involves navigating several key considerations, including sample design, disproportional allocation, clustering, and non-response.

Disproportional sampling is effective for improving representation of smaller or underrepresented subgroups, enabling more accurate estimates and focused analysis. Furthermore, the use of clusters in the sample design is in most surveys a given. However, disproportional allocation and clusters often increase variance and adds complexity to the analysis, requiring careful weighting to maintain overall representativeness. Conversely, high weight efficiency, indicated by a design effect (DEFF) close to 1, or high weight efficiency reflects minimal additional variability introduced by weighting, leading to more precise and cost-effective results. Yet, it may sometimes lead to the underrepresentation of critical subgroups, affecting the reliability of estimates for these groups.

Kish's design effect provides a practical measure of weight efficiency, often used due to its simplicity in cases where weights are relatively stable. This formula captures the increase in variance due to unequal weights, but its application is limited by assumptions of equal strata means and unit variances. For more nuanced analysis, especially in complex surveys, decomposing the design effect at the stratum level—using Liu et al.'s approach—offers a detailed view of weight efficiency by accounting for variations in stratum sizes and sample allocation. This decomposition provides a more accurate measure of how weight efficiency varies across different strata, enhancing the understanding of its impact on overall survey precision.

Ultimately, managing the trade-off between achieving a representative sample and maintaining high weight efficiency is essential for aligning with the study's objectives. While lower weight efficiencies might indicate increased sampling error, they reflect the inherent complexities in balancing precision and representativeness. Thoughtful consideration of these factors, including the use of Kish's design effect and stratum-level decomposition, ensures that sampling strategies are well-suited to the study's analytical goals, yielding reliable and actionable insights.

It is advisable for researchers to also calculate the precision (margin of error) for both overall estimates and subgroup results. Even with a low overall weight efficiency, estimates may still yield precise estimates. Conversely, high weight efficiency does not guarantee acceptable precision for either overall or subgroup estimates. Therefore, it is recommended that the margin of error (precision) be evaluated alongside weight efficiency to provide a more comprehensive understanding of estimate accuracy.

Weighting Efficiency Calculation: Revised Methodology

Historically, the weighting efficiency for our survey data was calculated on a *stratum-by-stratum* basis, where efficiencies were first computed within each stratum and subsequently averaged to produce an overall weighting efficiency measure. This approach is aligned with stratified analysis practices, reflecting the design structure by accounting for variance within each stratum before aggregation. The formula used in that context is often referred to as the *Stratified Kish Calculation*.

New Approach: Overall-Level Weighting Efficiency Calculation Using the Kish Formula

Based on current industry best practices and recommendations from specialised survey methodologists, we have transitioned to calculating weighting efficiency at the *overall sample level*, rather than per stratum. This approach employs the classic *Kish formula* for weighting efficiency, which is rooted in the concept of the design effect due to weighting and does not partition the calculation by strata or clusters.

The overall weighting efficiency, *WE*, is calculated as:

$$WE = n \times \sum_{i=1} nwi2 (\sum_{i=1} nwi)2$$

where:

- *n* is the total sample size,
- *wi* is the final weight assigned to the *ith* respondent.

This calculation reflects the relative inflation of variance due to unequal weighting across the entire sample. By aggregating over the full dataset, the Kish formula captures the net effect of weighting variability, ignoring strata or clustering design complexities.

Key differences and interpretative implications

Aspect	**Stratified Kish Calculation**	**Overall-Level Kish Calculation (Current Method)**
Basis of Calculation	Efficiency calculated separately within each stratum and then averaged	Efficiency calculated once on the entire sample ignoring strata
Consideration of Survey Design	Incorporates stratification and clustering implicitly	Ignores stratification and clustering effects
Effect on Weighting Efficiency Values	Typically yields higher weighting efficiency due to within-stratum homogeneity	Produces lower weighting efficiency, reflecting true variance inflation from weighting
Interpretation	Reflects efficiency as if strata are distinct and weighted separately; may underestimate variance inflation	Reflects overall efficiency conservatively; better captures impact of weighting variability

Aspect	Stratified Kish Calculation	Overall-Level Kish Calculation (Current Method)
Use Case	Useful when strata are analytically distinct and design effects considered critical	Recommended for industry best practice and comparability; straightforward and conservative

This shift in methodology means that the new weighting efficiency estimates are generally lower than previously reported. While this might initially appear as a decrease in weighting "quality," it is rather a more robust and transparent assessment of weight-induced variance inflation.

Consequently, reporting and interpretation should clearly state the calculation method used, emphasising that the current overall-level Kish calculation:

- Provides a **more conservative measure** of weighting efficiency suitable for industry standards.
- Does **not discount the effects of stratification or clustering**, which tend to reduce variance in stratified designs.
- Represents the **net impact of weighting variability on the entire sample**, reflecting the effective sample size reduction more accurately.

Users of the data should understand this methodological change to avoid misinterpreting the lower weighting efficiency as deteriorated data quality, when in fact it reflects an improved, more realistic calculation protocol.

OMC 2017-2024

Weight Efficiency	WEff2_MY22
The formula includes Design & Calibration weights	52.73

3. Briefing and Training

3.1 Introduction

The field teams are trained at the start of February each year. The training starts with the existing supervisors / suppliers that have worked on OMC for the last couple of waves – this was deliberately done to get the teams with the least amount of training required into field first. The new teams were trained after the experienced teams to allow ample time for first time training on OMC.

During the OMC waves, twenty-two (22) field supervisors/outsource companies as well as two hundred and thirty-five (235) interviewers on average worked on the project. The field supervisors/suppliers were responsible to evaluate the interviewers' performance and oversaw the quality of each interviewer's work.

3.2 Fieldworker briefing and training

Field training and briefings took place one to two weeks prior the data collection in-field date.

Training schedule

Within these sessions, interviewers and supervisors were briefed on the objectives of the study as well as being trained on the following data collection aspects:

- Basic research principles.
- The questionnaire and understanding of interviewer instructions and skips.
- How to read and interpret the maps.
- How to use the Kish grid to select a specific respondent.
- The process to be followed in the case of non-response.
- Substitution rules to be followed.
- How Geofencing works and the settings of the tablets required
- The key questions that interviewers generally struggle with

The following steps were taken to ensure that interviewers were sufficiently skilled and trained to conduct the OMC surveys:

- Interviewers were briefed by Ask Afrika Project Managers, Regional Managers, and Field Supervisors
- Interviewers conducted pilot interviews with Respondents under supervision.
- Supervisors evaluated each interviewer's performance and delivered further training, where it was required.
- Qualified interviewers conducted independent fieldwork.

3.3 Data Collection

Data collection for OMC fieldwork commenced in February each year. Weekly schedules were set up to which the Regional Managers, Supervisors and interviewers had to adhere. We also implemented an approach of first focusing on the more challenging EAs to ensure challenging areas are given the required time to be successfully completed.

3.4 Data quality control processes

The data capture process included the following main activities:

1. Questionnaire changes and formatting
2. Printing of training materials
3. Distribution
4. Fieldwork
5. Quality assessment

3.5 Quality assessment

1. Geo-fencing

Going to the right interviewing point is the first line of quality control Ask Afrika employ when in field. To ensure this, all points are geo-fenced with a 150m radius.

What is Geo-fencing

- It is a feature in Ask Afrika's CAPI software program that uses the global positioning system (GPS) to define geographical boundaries.

Benefits of Geo-fencing
- Ensure the interview is taking place at the correct location
- No interview will be able to activate if the interviewer is not within 150 meters of the interviewing point
- Assurance that the whole interview is taking place within the selected area

The data processing team verified all interviewing point coordinates against the coordinates uploaded during the activation of the interview.

2. Covert recordings

At the start of the OMC2016B data collection period, Ask Afrika implemented a new quality control measure, namely the covert recordings. This means that predetermined questions in the interview is voice recorded.

Benefits of voice recordings:
- Quality control interview captured data against the actual respondent voice responses
- Determine if interview is taking place
- Understanding of questions

Questions are changed throughout the data collection period.

The supervisors and team leaders were responsible for the first line of quality assessments. Once the supervisors initially approved an interview, the Ask Afrika data processing team listened to 15% covert recordings per interviewer to ensure the validity of the interviews.

The supervisors thoroughly checked the questionnaire for:
- The respondent selection procedure and whether it was correctly completed
- Missing values, missing trip information and
- The flow and routing instructions provided in the questionnaire

During the OMC2024 data collection period the status of the covert recording quality control was as follow:

Row Labels	Approved	Redo	Grand Total
Eastern Cape	68	0	68
Free State	72	0	72
Gauteng	377	1	378
KZN	199	0	199
Limpopo	50	0	50
Mpumalanga	93	1	94
North-West	61	0	61
Northern Cape	62	0	62
Western Cape	52	0	52
Total	1034	2	1036

The process followed are as follow:

- Strategic questions (which requires interviewer/respondent interaction) are assigned as "Recordings" questions on Dooblo
- During fieldwork period, covert audio recordings are activated on the tablet devices for the specified questions
- These recordings are uploaded together with the data once the interviews are completed with their interviews
- 15% sample is drawn of every interview's work that is indicated as "Initially Approved"
    - **Initially Approved** interview status means that Supervisors have reviewed the data/recordings and are happy with the fieldwork
- The DP team reviews the recordings and provide feedback on the quality and authenticity of the recordings.
- The DP team's feedback is collected, and each survey is given a QA rating (%):
    - **Approve:** Clear audio evidence that the interview is conducted per brief, and that the interview is authentic.
    - **Investigate**: Any instances where interviewer quality of interview authenticity is questioned is sent back to supervisors for internal investigation.
    - **Fail:** Unmistakable evidence that the interview quality or authenticity cannot be verified is rejected(Cancelled) and calls for a re-do.
- Should an interview be coded as "failed", a further 50% of the interviewer's work is added to the sample for recording review.

3. Methodology compliance

A few methodology compliance activities are also tracked and monitored on a weekly basis while in-field. These checks include:

- Monthly sample compliance – do the interviewers keep to the monthly sampling plan
- Dominant racial group – do the interviewers keep to the dominant racial group per EA
- Weekly day of week/time interviewing – is there a spread of interviews throughout the day and week/weekend
- High income locked EA's – are the interviewers' compliant
- Interviewing length – flag all interviews with an interviewer length < 60minutes
- 100% coordinate check with house addresses (visiting points)
- Duplicated home address checks
- Duplicated contact detail checks
- Kish grid compliance

4. Address verification

All OMC destination address fields are verified on Google Earth. A total of 43 221 addresses were verified for OMC2024A. This verification process is to ensure:

- That the destination address exists
- That the correct information is provided
- That the full address fields were provided
- To improve the address quality

The OMC address quality over the past 6 waves can be seen in the table below:

	Main place/Sub place	Main place/Sub place/Street Name	Main place/Sub place/Street Name/Street #
15A	14%	35%	50%
15B	20%	37%	43%
16A	7%	45%	48%
16B	9%	31%	60%
17A	10%	32%	58%
17B	12%	30%	58%
15/16/17	12%	35%	53%
18A	16%	22%	61%
18B	14%	19%	65%
16/17/18	11%	30%	58%
19A	16%	18%	66%
19B	14%	16%	70%
17/18/19	14%	23%	63%
20/22	16%	10%	75%
23A	6%	1%	93%
24A	11%	0%	88%

Address information in Informal Settlements

It is important to note the challenge in deriving a high-quality address for some sampling points, especially sampling points falling in informal settlements. Most informal settlements do not have street names and numbers, thus the only geography information available are Province, Main-Place and Sub-Place. In these instances, the addresses of all interviews in an EA after reverse geocoding (converting a coordinate to an address) will all have the same address and are classified being cluster interviews. Cluster interviews are not allowed as interviews need to be spread throughout the EA. Cluster interviews are cancelled and these interviews need to be replaced. All cluster interviews are checked manually and in instances where street information is not registered, and the coordinates are greater than 10m from each other and visually these interviews are spread in an EA, these interviews are kept.

The photo below is an example of an Informal Settlement not having registered address information. After the reverse geocoding took place, these interviews were logged as a cluster as all interviews were linked to a single address. After these interviews were investigated it was noted that a cluster was created due to the minimal address information in this EA. These interviews were approved.

5. Data checks

The final leg of the quality control phases is the data checks which is the most intense phase of all the quality control measures.

4. Data integration

4.1 Integration/merging of different mixed methodology data into a single source data file

Integration of different mixed methodology data sets is not very complicated. It is a simple merge of the different data sets together into one dataset. However, this is only true if the same questionnaire and variable structures were followed. The purpose of creating a single source data file is not to increase the number of variables, but to increase the respondent base and to create a single source data file for analysis purposes.

It is also good practice to ensure that an overlap of respondents is captured across the different mixed methodologies. The reason for this is to examine the differences in results, if any. In most cases, differences are not because of the mixed methodologies, but are a true reflection of the behaviour, perceptions and needs of the demographic profile. This needs to be noted and reported.

A statistical technique used to determine differences between mixed methodologies, is the use of CHAID analysis. When different methodology approaches are used as independent variables together with relevant questions on a dependant variable and one of the methodologies is highlighted being a driver, then it needs to be noted that the difference is due to the methodology.

The biggest difference picked up between the online and face-to-face methodologies in OMC was the increased number of trips provided by online respondents compared to the face-to-face respondent as well as the increase in address quality provided by the online respondents.

5. Data analysis and trip totals

5.1 Data tracking and trip totals

Data were monitored on a weekly basis for the tracking of progress of field surveyors and the overview of the approved data.

Weekly reports were created that included the monitoring of the following:

- Number of completed interviews
- Number of respondents indicated that they do not travel and the reasons for not travelling
- Number of respondents indicated that they travelled
    - Overall
    - Per province
    - Per trip type
- Number of trips per mode of transport
- Destination Types selected by respondents
- Mode of transport used to travel in relation to the Destination selected
- Mode of transport used per Province
- Minibus taxi versus taxi rank
- Address level information to maximise number of full addresses

Monitoring of the data on a weekly basis enable the team to track the progress of the project and identify and pre-empt any problems that could arise in the project relating to the data and field workers.

The table below reflects the final OMC travelled population, destinations visited in past 7 days, trip totals and average number of trips taken after data cleaning and address validations:

OMC2017/2018/2019/2020/2022/2023/2024										
	Total	**EC**	**FS**	**GP**	**KZN**	**LP**	**MP**	**NW**	**NC**	**WC**
Total 2015A	7 672	708	480	3 000	1 186	136	384	320	140	1 318
Total 2015B	7 525	692	455	2 812	1 204	136	396	327	152	1 351
Total 2016A	7 707	670	475	3 067	1 182	139	360	328	139	1 347
Total 2016B	7 493	631	457	3 002	1 178	130	345	308	138	1 304
Total 2017A	7 652	688	456	3 108	1 150	162	368	302	148	1 270
Total 2017B	7 500	675	448	3 036	1 156	154	374	321	169	1 167
Total 2018A	6 921	700	495	2 439	1 255	175	315	335	160	1 047
Total 2018B	8 193	773	397	3 308	1 329	159	389	342	144	1 352
Total 2019A	7 544	749	502	2 901	1 261	152	377	335	151	1 116
Total 2019B	7 651	724	464	2 929	1 232	167	339	356	148	1 292
Total 2020/2022	10 476	734	631	4302	1539	257	574	460	191	1788
Total 17/18/19/20/22	55 937	5 043	3 393	22 023	8 922	1 226	2 736	2 451	1 111	9 032
Total 2023A	7 583	654	414	3 292	1 274	151	351	284	172	991
Total 17/18/19/20A22A/23	63 520	5 697	3 807	25 315	10 196	1 377	3 087	2 735	1 283	10 023
Total 2024A	**7 499**	**480**	**357**	**3 321**	**1 019**	**153**	**396**	**367**	**105**	**1 301**
Total 17/18/19/20A22A/23/24	**71 019**	**6 177**	**4 164**	**28 636**	**11 215**	**1 530**	**3 483**	**3 102**	**1 388**	**11 324**
% - Not travelled 2015A	7.00%	10.00%	10.00%	8.00%	6.00%	5.00%	8.00%	15.00%	6.00%	2.00%

% - Not travelled 2015B	1.00%	1.00%	0.00%	2.00%	0.00%	3.00%	0.00%	1.00%	1.00%	1.00%
% - Not travelled 2016A	1.00%	2.00%	2.00%	3.00%	0.00%	1.00%	0.00%	1.00%	0.00%	0.00%
% - Not travelled 2016B	1.00%	3.00%	2.00%	2.00%	0.00%	1.00%	0.00%	3.00%	1.00%	1.00%
% - Not travelled 2017A	0.00%	1.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.00%	2.00%	0.00%
% - Not travelled 2017B	0.20%	0.40%	0.20%	0.30%	0.20%	0.00%	0.00%	0.30%	0.00%	0.20%
% - Not travelled 2018A	0.20%	0.30%	0.00%	0.50%	0.00%	0.60%	0.00%	0.60%	0.00%	0.00%
% - Not travelled 2018B	0.10%	0.40%	0.00%	0.20%	0.00%	0.90%	0.00%	0.00%	0.60%	0.00%
% - Not travelled 2019A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
% - Not travelled 2019B	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
% - Not travelled 2020/2022	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
% - Not travelled 17/18/19/20/22	1.00%	2.00%	1.00%	2.00%	1.00%	1.00%	1.00%	2.00%	1.00%	0.00%
% - Not travelled 2023A	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
% - Not travelled 17/18/19/20A22A/23	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
% - Not travelled 2024A	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**
% - Not travelled 17/18/19/20A22A/23/24	**0.12%**	**0.23%**	**0.02%**	**0.16%**	**0.02%**	**0.20%**	**0.00%**	**0.13%**	**0.29%**	**0.06%**
Trip Destinations 2015A	23 278	2 366	1 251	8 567	3 654	473	1 175	811	503	4 478
Trip Destinations 2015B	29 550	2 784	1 620	10 841	4 647	522	1 629	1 197	607	5 703
Trip Destinations 2016A	29 826	2 516	1 588	11 433	4 604	470	1 427	1 173	563	6 052
Trip Destinations 2016B	28 839	2 242	1 814	11 652	4 569	501	1 465	1 056	520	5 020
Trip Destinations 2017A	33 421	2 833	14 284	582	751	5065	1 614	4 808	2156	1 328
Trip Destinations 2017B	33 126	2 984	2 356	13 485	5 114	786	1 819	1 503	748	4 331
Trip Destinations 2018A	28 743	3 017	2 106	10 050	5 417	747	1 424	1 443	638	3 901
Trip Destinations 2018B	34 754	3 374	14 081	611	795	5718	1 687	5 093	1844	1 551
Trip Destinations 2019A	33 525	3 266	2 248	13 587	5 553	755	1 760	1 577	571	4 208
Trip Destinations 2019B	32 815	3 137	2 169	13 179	5 043	760	1 629	1 703	521	4 674
Trip Destinations 2020/2022	43 288	3317	2777	18181	6520	1056	2191	1883	822	6541
Trip Destinations 17/18/19/20/22	239 672	21 928	40 021	69 675	29 193	14 887	12 124	18 010	7 300	26 534
Trip Destinations 2023A	29 444	2206	1540	13970	4647	581	1175	978	617	3730
Trip Destinations 17/18/19/20A22A/23	269 116	24 134	41 561	83 645	33 840	15 468	13 299	18 988	7 917	30 264
Trip Destinations 2024A	**30 201**	**1 681**	**1 374**	**14 746**	**3 866**	**628**	**1 409**	**1 324**	**373**	**4 800**
Trip Destinations 17/18/19/20A22A/23/24	**299 317**	**25 815**	**42 935**	**98 391**	**37 706**	**16 096**	**14 708**	**20 312**	**8 290**	**35 064**

The totals per District Municipality at the end of 2024 are as follow:

Province	Row Labels	Grand Total (2024)	Grand Total (2017-2024)
Eastern Cape	Buffalo City	142	1632
Eastern Cape	Chris Hani	24	233
Eastern Cape	Nelson Mandela Bay	264	3867
Eastern Cape	Sarah Baartman	50	445
Free State	Fezile Dabi	42	467
Free State	Lejweleputswa	120	635
Free State	Mangaung	153	2775
Free State	Thabo Mofutsanyane	42	287
Gauteng	City of Johannesburg	1333	10919
Gauteng	City of Tshwane	812	6721
Gauteng	Ekurhuleni	808	7231

Gauteng	Sedibeng	211	1929
Gauteng	West Rand	157	1836
KwaZulu-Natal	Amajuba	55	309
KwaZulu-Natal	eThekwini	751	8523
KwaZulu-Natal	iLembe	31	218
KwaZulu-Natal	King Cetshwayo	15	259
KwaZulu-Natal	Ugu	12	263
KwaZulu-Natal	UMgungundlovu	125	1358
KwaZulu-Natal	Uthukela	30	285
Limpopo	Capricorn	67	847
Limpopo	Vhembe	25	205
Limpopo	Waterberg	61	478
Mpumalanga	Ehlanzeni	53	427
Mpumalanga	Gert Sibande	172	1259
Mpumalanga	Nkangala	171	1797
North-West	Bojanala	155	1182
North-West	Dr Kenneth Kaunda	149	1513
North-West	Dr Ruth Segomotsi Mompati	25	156
North-West	Ngaka Modiri Molema	38	251
Northern Cape	Frances Baard	67	802
Northern Cape	Pixley ka Seme	28	272
Northern Cape	ZF Mgcawu	10	314
Western Cape	Cape Winelands	145	1047
Western Cape	City of Cape Town	956	8856
Western Cape	Garden Route	103	963
Western Cape	Overberg	37	202
Western Cape	West Coast	60	256
Grand Total		**7 499**	**71 019**

6. An overview of OMC since 2020

6.1 OMC2020

The data collection for OMC2020 was placed on hold in March 2020 after the announcement of COVID-19 and the implementation of lockdown. A decision was made to resume data collection after the relief of lockdown. With the severity of the Coronavirus epidemic, lockdown was further extension throughout the year. Lockdown had a significant impact on people's normal travel behaviour and with the Coronavirus epidemic still influencing lives till 2021 OMC decided to only resume with data collection in 2022.

6.2 Data collection challenges

Data collection of OMC 2022 was placed on hold during the floods in Kwa-Zulu Natal. A few of the sampled areas became unreachable and had to be substituted. In Gauteng due to unrest and security challenges a few interviews had to be substituted.

6.3 OMC footprints

Municipalities

SA's geography consist out of 52 District Municipalities of which 38 District Municipalities fall within the OMC footprint. 14 District Municipalities are excluded.

Also, SA's geography consist of 233 Local Municipalities of which 144 Local Municipalities (62%) fall within the OMC footprint. 26 Local Municipalities are excluded due to the community size being < 8,000 and 64 Municipalities are excluded which covers the 14 District Municipalities which are excluded.

11 out of the 144 Municipalities in the OMC footprint was not selected during the sampling stage.

EA's

SA's geography consist out of 103 576 EA's of which 55 265 EA's (53%) fall within the OMC footprint.

The following table illustrate the EA's per District Municipality falling within the OMC footprint:

District Municipalities	SA Total EA's	OMC Total	% in OMC
Alfred Nzo	2107	Excluded	0.00%
Amajuba	832	410	49.28%
Amathole	3314	Excluded	0.00%
Bojanala	3145	1168	37.14%
Buffalo City	1819	1441	79.22%
Sarah Baartman	1056	513	48.58%
Cape Winelands	1264	867	68.59%
Capricorn	2718	519	19.09%
Central Karoo	246	Excluded	0.00%
Chris Hani	3118	543	17.42%
City of Cape Town	6417	6417	100.00%
City of Johannesburg	6679	6679	100.00%
City of Tshwane	5233	5228	99.90%
Dr Kenneth Kaunda	1398	1178	84.26%
Dr Ruth Segomotsi Mompati	1012	227	22.43%
Garden Route District Municipality	1395	962	68.96%
Ehlanzeni	3237	869	26.85%
Ekurhuleni	5341	5341	100.00%
eThekwini	5621	5107	90.86%
Fezile Dabi	1086	758	69.80%
Frances Baard	718	564	78.55%
Gert Sibande	1960	1153	58.83%
Sekhukhune	2284	Excluded	0.00%
iLembe	1100	322	29.27%
Joe Gqabi	1362	Excluded	0.00%

John Taolo Gaetsewe	546	Excluded	0.00%
Lejweleputswa	1382	1066	77.13%
Mangaung	1551	1546	99.68%
Mopani	2294	Excluded	0.00%
Namakwa	592	Excluded	0.00%
Nelson Mandela Bay	2198	2198	100.00%
Ngaka Modiri Molema	1684	462	27.43%
Nkangala	2573	1629	63.31%
O.R.Tambo	3508	Excluded	0.00%
Overberg	650	351	54.00%
Pixley ka Seme	626	194	30.99%
Sedibeng	1806	1638	90.70%
Harry Gwala	1075	Excluded	0.00%
ZF Mgcawu	578	256	44.29%
Thabo Mofutsanyane	1808	929	51.38%
Ugu	1384	357	25.79%
Umgungundlovu	1725	928	53.80%
Umkhanyakude	998	Excluded	0.00%
Umzinyathi	872	Excluded	0.00%
Uthukela	1136	353	31.07%
King Cetshwayo	1457	302	20.73%
Vhembe	2574	214	8.31%
Waterberg	1666	591	35.47%
West Coast	846	479	56.62%
West Rand	1791	1506	84.09%
Xhariep	464	Excluded	0.00%
Zululand	1330	Excluded	0.00%
Grand Total	**103576**	**55265**	**53.42%**

6.4 Piloting of the OMC online questionnaire in 2021

The OMC face-to-face questionnaire was converted to an online friendly questionnaire in 2022. The questionnaire was programmed and tested extensively by Ask Afrika and Kuper Research and was signed off by all parties.

Informal pilots were conducted with Ask Afrika and OMC's own networks (family, friends, and colleagues) in 2020. After the pilot, questionnaire and programming changes were done based on the pilot feedback.

Informal pilots were conducted in 2020 and formal external pilots took place in 2021, before data collection resumed in 2022.

The purpose of the formal pilots was to:

- Evaluate the online methodology
- Evaluate the online questionnaire
- Conduct comparative analysis: 2019C vs the 2021 formal pilots to guage drastic changes by incorporating an online methodology

Main finding from the 2021 pilot were as follow:

- More trips were provided on average, in the online methodology compared to the face-to-face methodology.
- Address quality, which is one of the most important factors of the OMC research increased. 11% more respondents were able to provide proper address information up to a street number level.
- A 5% drop in work trips were noted, which is contributed to people working from home as a result of the COVID-19 pandemic. As well as a 3% drop in place of worship as well as going to shops.
- A 2% increase in going to family and friends as well as a 3% increase in going to hospitals which are also coherent to the current COVID-19 situation.
- A 6% drop in car as mode of transportation were evident, with a 5% increase in using Uber and 2% increase for Gautrain.

These differences seen between OMC2019C and the pilot in 2021, are not as a result of the methodology, but based on the impact of COVID-19.

ANNEXURE A

WEIGHTING EFFICIENCY - DETAIL

The sample design used in many surveys is often based on a complex sample design, including stratification, multistage sampling, probability proportional to size and unequal selection probabilities. A disproportionate stratified sample is frequently required to obtain a representative sample that enables the survey to provide estimates with acceptable precision for subgroups. Weights are assigned to compensate for unequal probabilities of selection, non-response, and post-stratification adjustments. These weights ensure that the sample accurately represents the

population. However, if the weights vary too much or are improperly applied, they can introduce noise, increase the standard error and the margin of error and ultimately decrease the precision.

Therefore, a trade-off between the increase of the standard error and a good representative sample, also for subgroups, is often required.

Several factors can negatively impact weight efficiency, including:

- Sample design and disproportional allocation
- The use of clusters and multistage clustering
- Large deviation in cluster sizes
- Non-response
- Post-stratification (Benchmark/Calibration), especially if the realized sample is skewed across these variables
- Mixed approach (F2F & Online)

Although these factors can increase weight variability and reduce both precision and weight efficiency, they are often essential for achieving a representative sample, compensating for nonresponse and over/under coverage of some subgroups. Additionally, in many surveys, researchers have little choice but to use single or multistage clusters, with limited control over variations in cluster sizes.

Background and Formulae of the design effect and weight efficiency.

Weight efficiency is a widely used measure that evaluates how effectively a sampling strategy utilizes assigned weights to represent the target population. However, researchers do not always fully understand the concept of weight efficiency. This section provides an overview of the practical weight efficiency formula, along with its underlying assumptions and limitations.

To comprehend weight efficiency, it is essential to introduce the concept of design effect and how weight efficiency is derived from it.

The **design effect** (DEFF) quantifies the impact of various factors such as stratification, clustering, unequal selection probabilities, and adjustments for non-response and non-coverage. It represents the ratio of the variance of a weighted estimator in a complex sample design to the variance of an unweighted estimator in a simple random sample of the same size. This can be expressed as:

$$DEFF = \frac{Var\left(\hat{Y}_w\right)}{Var\left(\hat{Y}_{srs}\right)}$$

where $\hat{Y}_w$ is the weighted estimator, and $\hat{Y}_{srs}$ is the estimator under simple random sampling without weights, ignoring the true design and post-stratification t is important to note that the design effect pertains to a specific estimate $\hat{Y}$, and will be different for different estimates in the survey. $\hat{Y}$ could be any estimate of interest, range from a mean or proportion to a total or even a regression coefficient.

One way to express the weight efficiency, Wgt_EFF) is through the design effect,

$$Wgt\_EFF = \frac{1}{DEFF}.$$

A value of weight efficiency close to 1 indicates high efficiency, meaning there is minimal loss of precision due to weighting and deviations from simple random sampling. Conversely, values below 1 signify reduced efficiency.

The following are alternative forms of these formulae, which are more readily applied to determine the effects of weights in practice.

A common approximation for the design weight is Kish's formula (Kish, 1965), which is used when the weights are relatively stable:

$$DEFF = 1 + CV^2(w)$$

where $CV(w)$ is the coefficient of variation of the weights, defined as

$$CV(w) = \frac{\text{Standard Deviation of Weights}}{\text{Mean of Weights}}.$$

Kish's design effect can also be rewritten as

$$DEFF = \frac{n\sum_{i=1}^{n} w_i^2}{(\sum_{i=1}^{n} w_i)^2}.$$

Using this equation, the commonly applied formula for weight efficiency, becomes:

$$Wgt\_EFF = \frac{1}{DEFF} = \frac{(\sum_{i=1}^{n} w_i)^2}{n\sum_{i=1}^{n} w_i^2}.$$

It is important to note that the Kish's formula only accounts for the increase in variance due to unequal weights and is derived under the extremely restrictive assumptions that the strata means are equal, and unit variances are all equal (Kalton et al., 2005).

In practice, Kish's DEFF is often used even when the sample design is more complex, due to its ease of computation. However, DEFF and thus Wgt_EFF, are not always relevant in surveys where variances differ across strata, where subgroups are intentionally sampled at different rates, and/or where different subgroups have substantially different response rates. Having unequal weights in those situations is desirable and can more nearly meet analytic goals than having equal weights will.

Moreover, Kish's design effect does not fully account for the possibility that the estimators may be more efficient than suggested by the coefficient of variation of the weights, $CV^2(w)$. This is particularly true for probability proportional to size sampling.

Decomposing Design Effects at stratum level

Since the widely used design effect formula, popularized by Kish (1965), is only applicable to the overall survey and does not leverage the available information at the stratum level, Liu et al. (2000 & 2022) and Zipf , G. and Valliant, R. (2024) proposed a decomposition of the unequal weighting effect at the stratum level, denoted as h, using the following formula:

$$DEFF_D = \sum_h \left(\frac{N_h}{N}\right)^2 \left(\frac{n}{n_h}\right) DEFF_h.$$

where,

$$DEFF_h = \frac{n_h \sum_{ij} w_{hij}^2}{\left(\sum_{ij} w_{hij}\right)^2}$$

is the stratum specific unequal weighting effect, thus the design effect calculated for stratum h, N_h and n_h represent the stratum population size and sample sizes, respectively, and N and n are the overall population size and sample size.

The weight efficiency based on this formula, is given by

$$Wgt\_EFF_D = \frac{1}{DEFF_D}.$$

This formula not only decomposes the unequal weighting effect by stratum-level design effects but also incorporates the ratios of stratum size to the total population, $\frac{N_h}{N}$, as well as the stratum sample size relative to the overall sample size, $\frac{n}{n_h}$.

It is important to note that none of the formulas presented above account for the cluster effect inherent in certain sampling designs..

For the calculation of the variance estimator of an estimator in a complex sample, the variance must first be calculated for each stratum and then aggregated to obtain the overall variance estimator. Based on sampling theory, DEFF_D and Wgt_EFF_D are more closely aligned with the calculation of variances for estimates. However, further consideration is required to assess how the values for Wgt_EFF and Wgt_EFF_D differ.

Online panel composition

Online panel partner - Cint

Cint, founded in 1998 with the belief that consumer insights are at the heart of decision-making in an ever-changing world.

Cint hosts the world's largest consumer network for digital market research, consisting of more than 155 million research participants in over 130 countries – accessing more than 4,500 individual market research panels and sample sources. The Cint platform offers a unique and transparent marketplace environment called the Insight Exchange for buyers and sellers of sample to connect and transact by matching a seller's market research panel and respondent profiling data against

the buyer's sampling needs in real-time. This makes Cint the most effective and efficient way to access sample to conduct digital market research at scale and at speed. Founded in Stockholm, Sweden in 1998, Cint is trusted by over 3,000 agencies, researchers, and brands in more than 72 countries worldwide. Since Cint's inception, innovation has been a cornerstone of the company's growth. In 2005, it launched the industry's first and only global online sample marketplace platform called CPX, now known as the Insight Exchange, which is the nucleus of the Cint platform today.

Recruitment in SA

Cint do not exclude individuals in joining the online panel. Anyone who agrees to be contacted for market research can join the panel. You will need to be over the legal age (15 years and older) and have internet access. The recruitment methods are far reaching and designed to ensure to reach a diverse group of South Africans.

The panel is random and actively recruited according to industry guidelines on conducting research over the internet.

New panellists are constantly added to keep the panel updated. Panel recruiting is an ongoing process and recruiting methods are periodically adjusted to keep the panel balanced and representative of the country's population, as much as possible.

A combination of both offline and online recruitment methods is used when recruiting panellists (Email Lists, Website recruitment, Social Media, Point of Sale, Affiliate Networks, Ad Campaigns, Lead Generation Vendors). Recruiting methods are designed to maximise the demographic diversity of the panel.

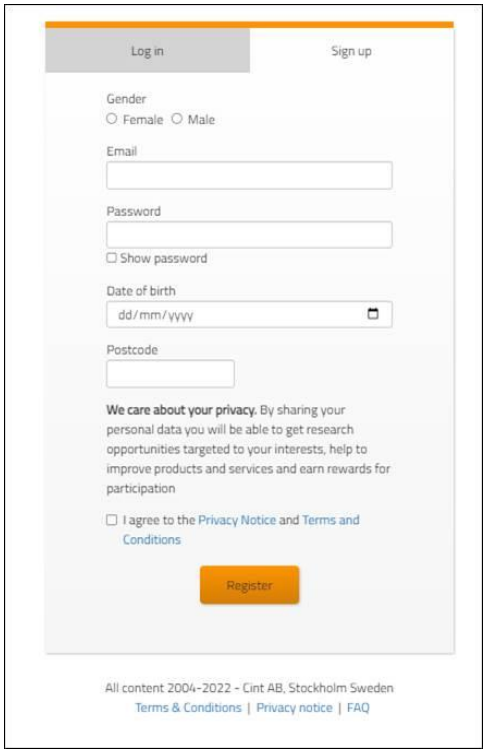
A double-opt in process is used when recruiting panellists - verification of opt-in status and email address is done using the double opt-in paradigm. Email invitations delivered to users who have consciously chosen to be part of the online panel cannot be considered as spam. Invitations to complete surveys are sent via email and in some instances notifications via WhatsApp is applicable.

The biggest limitation with online panels is that it's restricted to respondents with internet access. That said in today's South Africa internet penetration is very high if you consider smart phone users. Smart Phones have put internet in the hands of the affluent and the lowest income earners.

Recruitment process

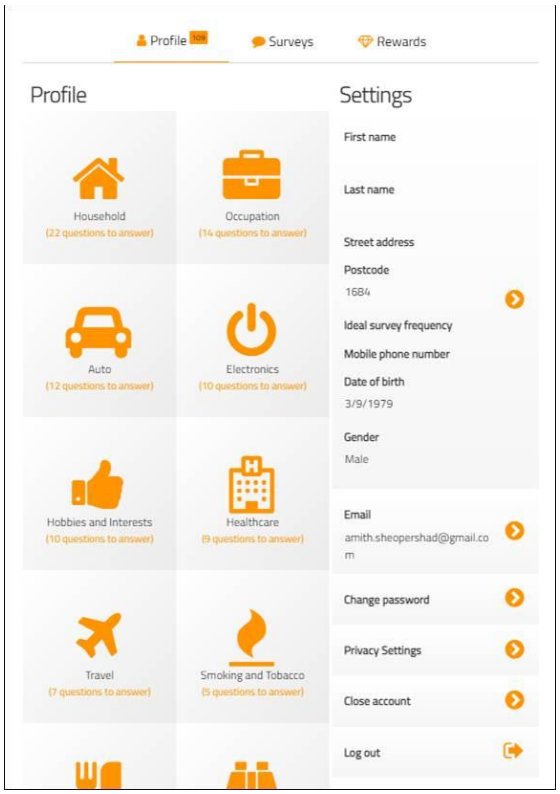
Panelists are recruited through the various channels highlighted earlier and directed to a landing page to complete the registration.

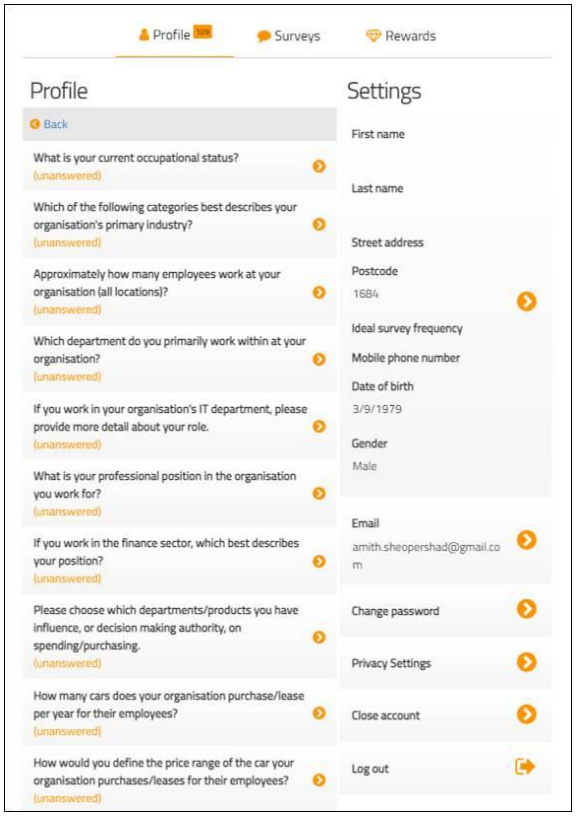
The panelist signs up using their email, setting a password, DOB + Postal code, and agree to Data Privacy requirements on PI information.



Once the registration is complete, there is a further email sent to the panelist to confirm email address.

Upon login to the panellist portal, the user can complete profiling and questions and take surveys.





Cint sends panelist survey opportunities via notification on the panel owner's website or the OpinionApp (Cint's App offering to panel owners).

Quality Charter

Respondent Quality

We evaluate respondents across multiple metrics to determine that individuals are real, unique, engaged and representative. All Cint panellists are opted in and we use a blend of personal, behavioural and demographic data points to ensure quality. A mix of advanced solutions ensure that, regardless of the number of accounts, users attempt a specific survey only once.

Operational Quality

Are we executing well and delivering on our commitments to clients and partners? We never underestimate our own, first-hand role in overall quality, so we ascribe to the golden rule when it comes to our internal operations. We measure our performance on both subjective and objective measures, such as timeliness and internal/external satisfaction.

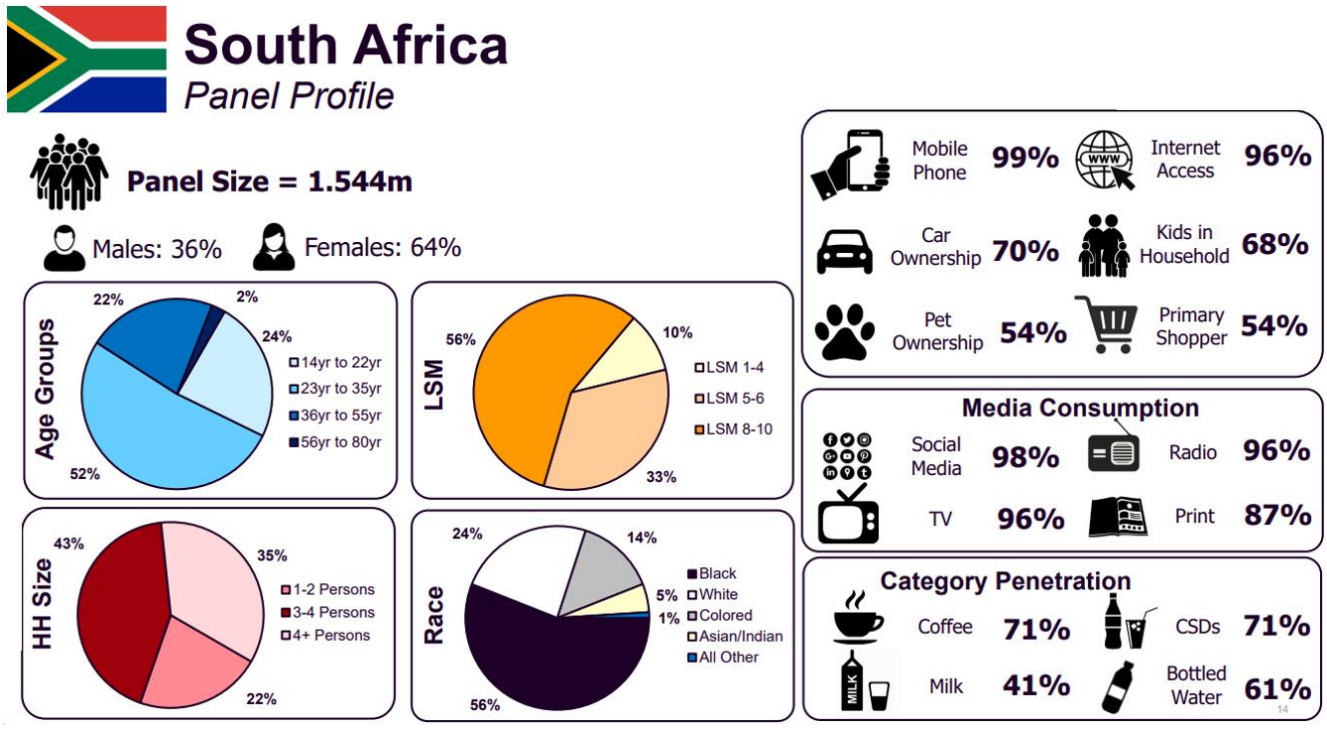
Product Quality

As a company that provides the technology backbone for the insights industry, we know our products must provide industry-leading and future-proof solutions. We seek to be ranked consistently as best in-class both through external benchmarks in the marketplace, and ratings by our clients and suppliers.

Compliance with industry bodies

The Cint platform and its products comply with ESOMAR, MRS, ARF, MRIA, AMA, AMSRO and Insights Association standards. Cint also complies with ISO 20252 and is certified in Australia.

Panel profile



–

	Demographics	TGI population	Cint
Gender	Male	49%	36%
	Female	51%	64%
Population group	Black	72%	56%
	Coloured	11%	14%
	Indian/ Asian	4%	5%
	White	12%	24%
Province	Eastern Cape	6%	6%
	Free State	6%	4%
	Gauteng	42%	40%
	KwaZulu-Natal	13%	15%
	Limpopo Province	2%	6%
	Mpumalanga	7%	6%
	North West	5%	5%
	Northern Cape	2%	1%
	Western Cape	17%	17%
Age Breaks	15-24	21%	15-22: 24%
	25-34	26%	23–35: 52%
	35-44	21%	36-55: 22%
	45-64	23%	56-80: 2%
	65+	8%	