

DATA REPORT

GCRO THIRD QUALITY OF LIFE SURVEY

PREPARED FOR

THE GAUTENG CITY REGION OBSERVATORY

PREPARED BY

GEOSPACE INTERNATIONAL

2 APRIL 2014

Table of Contents

1. INTRODUCTION	3
2. DATA COLLECTION BACKGROUND	3
3. UNIVERSAL CODES USED	4
4. UNIQUE RECORD IDENTIFIER	5
5. DATA CORRECTIONS	5
5.1 Call backs regarding variables consisting of numerical values	5
5.2 Call backs due to the misinterpretation of a question or fieldwork selection	5
5.2.1 Trip time start and end	5
5.2.2 Transport mode	6
5.3 Call backs due to the digital data collection instrument errors.....	6
5.3.1 Walking time	6
5.3.2 Service delivery protest	6
6. RECODES	6
6.1 Main Place and Sub Place recodes	9
6.2 Employment status	10
7. COUNTRY CODES	10
8. INTERPRETATION OF THE VOTING RESULTS	11
9. GEOGRAPHIC COORDINATES	11
10. WEIGHTING	11
11. APPENDIX.....	12
11.1 Appendix 1: Country Codes	12

1. Introduction

This report serves as a summary of changes, amendments, recodes and corrections made during the data validation process for the third Quality of Life survey implemented by the Gauteng City Region Observatory (GCRO).

The report should be reviewed in conjunction with the completed SPSS dataset.

2. Data collection background

A digital data collection instrument was implemented using an open source system called Formhub and administered on a tablet device. Details about the Formhub data collection system can be found at www.formhub.org.

The questionnaire was provided by the GCRO. Once a questionnaire was administered in the field, it was uploaded in field using internet connectivity to a cloud server from where it could be accessed and downloaded online. Approximately 120 of these devices were used in the field.

This is the first time a digital questionnaire was used for the GCRO Quality of Life survey. Moreover, a different field data collection and sampling methodology was used for the third GCRO Quality of Life Survey.

In order to maximise the population spread of the sample, it was decided to use the Statistics South Africa (StatsSA) National Population Census 2011 geography, specifically the Small Area Level (SAL) polygons which were derived from the Population Census Enumerator Area polygons (EA).

The 2011 Ward boundary geography and codes used are derived from the latest Demarcation Board information. Small Area level codes and geography are derived from the latest Statistics South Africa information (Census 2011), as at August 2013.

The reason for this is simple. It enabled us to devolve the sample selection to a much lower level while also ensuring optimum population coverage within each Ward. Therefore, not only was the population spread maximised across Wards, but also within Wards. This makes for a much more representative sample regarding population and geographic spread across the entire survey geography, from Municipal to Ward level.

The main analysis and dissemination platform for the survey remained at Municipal and Ward level. The survey geography therefore constituted of Provincial; Municipality/Metro; Ward; SAL level.

There are 17840 SALs spread across Gauteng without any gaps or overlaps. The main aim was to complete a required amount of interviews per randomly sampled SAL within a Ward to reach the required number of interviews per Ward, while the spread of these SALs within each Ward made for a much more effective population spread than one would get, for example, with a number grid.

A random sample, using population size, was drawn using the following as base parameters:

- A minimum number of interviews were set for Wards falling in the District Municipalities (30 successful interviews)
- For the Metropolitan Municipalities this number was doubled to a minimum 60 interviews

- Using the adult (18+) population per ward (from StatsSA, Census 2011), a required number of interviews by Ward was calculated.
- Where the calculation provided a required number less than the specified minimum (30 or 60 interviews respectively), the required number of successful interviews was increased so the minimum requirement per Ward was met.
- In cases where the calculation provided a figure more than the minimum requirement, the figure was left as is.
- Gauteng consists of 10 Municipalities and District Municipalities subdivided into 508 Wards. The aim was to spread the SAL sample across and within Wards to cover the various types of population. Using the SALs provided the means of doing so.
- The second stage of the sampling was to apply the same parameters but this time at SAL level, drawing a random sample of SALs within which a random number of required interviews were calculated. As mentioned, this method ensured an optimum and extensive coverage of population type across Wards within a Municipality as well as within Wards.
- The end result was out of the 508 Wards, 26387 successful interviews had to be completed
- These interviews were distributed across 16400 SALs out of a total of 17840 SALs.
- Since the sample was completely random, some SALs sampled contained no private residences, such as vacant SALs, or SALs consisting of hospitals, prisons and even graveyards. These SALs were substituted with other SALs as close to the sampled SAL as possible.

3. Universal codes used

-2 : Standard as **“No response due to skip scenario”**. This indicates where a variable field was not populated due to an answer or selected option in a previous question which did not warrant an answer for the skip question.

-1 : This code is used where ever a variable field was not populated due to the following reasons

- **“No response or don’t know” (-1)** – where the respondent did not provide a response, was unsure about or did not know a specific response required
- **“Fieldwork error” (-1)**
 - Where a fieldworker made a typing error which caused an outlier and then could not be corrected during an exercise where the relevant respondents were called back to provide correct answers
 - Where the misinterpretation of questions or a miscommunication between the fieldworker and the respondent caused an invalid response which could not be corrected during a call back exercise
 - Where a question or skip question was not activated due to a development error in the digital data collection instrument and for which a valid response could not be determined during a call back exercise
 - This label pertains to the following questions:
 - D.1 How many dwellings are on this stand?
 - D.3 How many households are in the selected dwelling unit?
 - E.1 How many eligible respondents in the selected household?
 - 4.3 Think about the last time you made this trip. What time did you start?
 - 4.4 What time did you arrive at your final destination?
 - 8.9.2.1 Why are you not looking for work?
 - 10.3 Have you taken part in any service delivery protest in the past 12 months?
 - 12.2 What is your current age?

- **“Weren't asked due to fieldwork error” (-1)** – This label pertains to two questions:
 - “4.12 How far, in walking time, is the closest point to reach public transport from your home? For example, a taxi stop, bus stop or train station.” Due to a questionnaire development error, this question did not activate on the digital platform. This error was corrected later in the fieldwork phase. As a result, many responses were not captured which led to a large call back exercise. The responses for those respondents which could not be contacted during the call back exercise were coded as -1 and labelled as above.
 - “5.3.1 Why are you not paying it back?” This is a skip question which did not activate correctly. This issue was picked up early during the fieldwork exercise and did not affect many responses

4. Unique record identifier

A unique code was automatically generated by the data collection software for each questionnaire captured. This code is used as a primary key in the SPSS dataset. The field label is UUID.

5. Data corrections

This section pertains to variables needing data correction from call backs due to fieldwork or digital data collection instrument error.

5.1 Call backs regarding variables consisting of numerical values

These call backs were made mainly to verify outliers for numerical variables. The following are the variables for which call backs were made:

- 1.1 How many rooms does your household occupy? Excluding bathroom, toilet or kitchen. Six corrections were made
- 1.2 If one room only, how many other households does this household share this room with? Six corrections were made
- 1.10 How many OTHER households, if any, share this toilet with your household? Three corrections were made
- 4.16.1 Approximately how much does your household spend in total on transport each month? Provide cost in Rand. A total of 17 corrections were made
- 8.2 How old is this business? Six corrections were made
- 12.2 What is your current age? Contact was attempted with 26 respondents and 23 corrections were made

5.2 Call backs due to the misinterpretation of a question or fieldwork selection

5.2.1 Trip time start and end

It came to light that the phrasing of certain questions was misinterpreted by either the fieldworker or the respondent in some cases. A set of questions related to transport had to determine the start and end time for the most frequent trip made by a respondent in the last month, one way, from your home to your destination. The questions were phrased as follow:

- 4.3 Think about the last time you made this trip. What time did you start?
- 4.4 What time did you arrive at your final destination?

In some cases, respondents interpreted this question as a round trip, in other words when they left and then arrived back home. Fieldworkers also had finger trouble in some cases or became confused when choosing AM or PM options. This led to call backs which had to be done at the end of fieldwork. A total of 632 respondents were identified where time verification was needed and corrections were made to 410 records.

5.2.2 Transport mode

Transport mode, which again pertained to the most frequent trip in the last month, where the following questions were asked:

- 4.6 Think about the type of transport you used when you made the trip. What were all the different types of transport you used to make the trip?
- 4.7 Thinking about the LONGEST distance travelled as part of your trip, what type of transport did you use for this part of the trip?

The two questions are of course linked, meaning one of the options selected in 4.6 should appear as a chosen transport type in 4.7, as such the transport type taking the longest also had to be mentioned as a mode of transport used. This was not always the case. Sometimes the respondent was of the opinion that the type of transport taking the longest would be an alternative option which they took only sometimes, but since it took too long or was more expensive than other modes, was only used when needed. In other cases respondent saw the longest mode as that type of transport they might take back home to their family home once a month or the like. Again, many call backs had to be done in this regard to not only determine possible reasons for the discrepancies but also the correct mode. A total of 2123 respondents were successfully contacted and transport mode verified and corrected.

5.3 Call backs due to the digital data collection instrument errors

5.3.1 Walking time

“4.12 How far, in walking time, is the closest point to reach public transport from your home? For example, a taxi stop, bus stop or train station.”

As mentioned in Section 3, a questionnaire development error caused this question to not activate initially and call backs had to be made to the relevant respondents to determine a valid response. A total of 4310 corrections were made.

5.3.2 Service delivery protest

“10.3 Have you taken part in any service delivery protest in the past 12 months?”

There were some records for which this question did not activate. These cases were identified and call backs were made where relevant. A total of 858 corrections were made.

6. Recodes

The following variables were recoded. All recodes can also be viewed in the accompanying coding sheet. Note: Main Place and Sub Place recodes will be dealt with in a separate section. The employment status recode will also be dealt with separately.

- **H.3 Which type of dwelling does this household occupy?**
 - FORMAL = 1,3,4,5,6,7,10,12
 - INFORMAL = 8,9

- OTHER = 2,11,13,14
- **2.2 When did you move into Gauteng?**
 - UP TO 1984;
 - 1985 - 1994;
 - 1995 - 2004;
 - 2005;
 - 2006;
 - 2007;
 - 2008;
 - 2009;
 - 2010;
 - 2011;
 - 2012;
 - 2013
- **2.4.1 From which municipality did you move from into Gauteng**
 - 999 = Don't know
 - Note all Municipal, Main Place and Sub Place codes were derived from the official Statistics South Africa National Population Census geography and information.
- **2.7 When did you move into this particular dwelling?**
 - UP TO 1984;
 - 1985 - 1994;
 - 1995 - 2004;
 - 2005;
 - 2006;
 - 2007;
 - 2008;
 - 2009;
 - 2010;
 - 2011;
 - 2012;
 - 2013
- **4.16.1 Approximately how much does your household spend in total on transport each month? Provide cost in Rand**
 - 1-200;
 - 201-400;
 - 401-600;
 - 601-800;
 - 801-1000;
 - 1001-1500;
 - >1500
- **7.10 How satisfied are you with your life AS A WHOLE these days?**
 - SATISFIED WITH MY LIFE = 1,2
 - NEUTRAL = 3
 - DISSATISFIED = 4,5
- **8.2 How old is this business?**
 - UP TO 1 YEAR;

- 2 YEARS;
- 3-4 YEARS;
- 5-6 YEARS;
- 7-10 YEARS;
- 11-15 YEARS;
- 16+ YEARS
- **Employment sector**
 - The following recode is related to employment sector. In question 8.15 “Which sector are you employed in?” the respondent is asked to which employment sector they belong. If they choose the option “Other”, then the follow-up question, 8.15.1 “You answered "Other" to the previous question, please specify” needs to be answered. 8.15.1 is therefore an open ended question.
 - The answers provided in 8.15.1 were then recoded to the options provided in 8.15, and the resultant codes for both 8.15 and 8.15.1 were combined into one single recoded field.
- **8.27 How many hours do you work per week in your main job?**
 - 0-19 HRS;
 - 20-39 HRS;
 - 40 HRS;
 - 41-45 HRS;
 - 46-50 HRS;
 - 51-60 HRS;
 - 61+ HRS
- **10.1 In the past year, have you participated in the activities of any of the following clubs?**
 - OPTIONS 1-19 = Participated in CSO;
 - OPTION 20 = Have not participated in CSO
- **12.1 What is the highest level of school education you have completed?**
 - NONE = 1
 - PRIMARY ONLY = 2 – 9
 - SECONDARY INCOMPLETE = 10 – 13
 - MATRIC = 14
 - MORE = 15 – 18
 - 19 – UNSPECIFIED
- **12.2 What is your current age?**
 - 16-19;
 - 20-24;
 - 25-29;
 - 30-34;
 - 35-39;
 - 40-44;
 - 45-49;
 - 50-54;
 - 55-59;
 - 60-64;
 - 65+
- **12.4 How many people, including you, live in this household?**

- 1;
- 2;
- 3;
- 4;
- 5;
- 6;
- 7+

6.1 Main Place and Sub Place recodes

This recode pertains to the following question: 4.2.2 “To which place do you travel inside of Gauteng?” The relevant Municipality, Main Place and Sub Place had to be specified.

The official Statistics South Africa Municipal, Main and Sub Place codes from Census 2011 were used. The use of a digital data collection instrument made it possible to cascade place name options displayed from one level to the next. Thus, if a certain municipality is chosen, then the next digital page would display only the main places located in that municipality, according to the StatsSA Population Census geography. Similarly, if a certain main place is selected, the next digital page would only display the relevant sub places located within that main place. The codes used are the same as those used by StatsSA.

In some cases, where the respondent was unsure to which main place he or she travelled, the fieldworker would select the non-urban (NU) option for that specific municipality. Consequently, only sub places located within the specific main place boundary is displayed. If none of those options was the sub place to which the respondent travels, the fieldworker selected the NU option again. In other cases, fieldworkers simply selected the place name option containing the “NU” distinction incorrectly. In all such instances, call backs were made in an attempt to re-contact the respondent and clarify the sub place name.

If the main place or sub place name could be clarified, it was corrected in the recoded field. If not, the sub place or both main and sub place was coded as -1.

There were other cases regarding sub place options where the original response was recoded to a more appropriate response that makes geographic sense. See the following table.

Original sub place		Recoded sub place
799059090	Centurion SP1	Recode to "Centurion Central" 799059041
799059089	Centurion SP2	Recode to "Centurion Central" 799059041
799059002	Centurion SP3	Recode to "Centurion Central" 799059041
799059012	Centurion SP4	Recode to "Centurion Central" 799059041
799035104	Pretoria SP	Recode to "Pretoria Central" 799035058
764002031	Randfontein SP1	Recode to Randfontein SP 764002017

760009006	Vereeniging SP1	Recode to "Vereeniging Central" 760009032
760009026	Vereeniging SP2	Recode to "Vereeniging Central" 760009032

In the SPSS dataset, the original main and sub place fields are included as well as the recoded fields. Field headings are 4_2_2_Mainplace; 4_2_2_Subplace; and 4_2_2_MainplaceV2; 4_2_2_SubplaceV2 respectively.

6.2 Employment status

The SPSS variable, "What is your employment status (recoded)?" has been determined using various responses from previous questions. The following status options are used:

- Employed
- Unemployed
- Other

The variables used to determine this status pertains to work done for which the respondent were paid in the past seven days (from when the interview took place) as well as reasons provided if no work was done in the seven days preceding the interview.

- The status of "**employed**" was assigned to the respondent when the answers to both of the following questions were "**yes**"
 - 8.8 In the past 7 days, did you do any type of work, business or activity for which you got paid? (even if just for one hour) or expected to be paid? This could include car washing or piece work or selling things and so on.
 - 8.9.1 What is the main reason that you did not work in the past 7 days? Between jobs - have been appointed to a new job but have not started yet
- The status of "**unemployed**" was assigned to the respondent if the answer to the following question was "**yes**"
 - 8.9.2 What is the main reason that you did not work in the past 7 days? Unemployed and looking for work
- The status of "**other**" was assigned to the respondent if the answer to the following question was "**no**"
 - 8.9.2 What is the main reason that you did not work in the past 7 days? Unemployed and looking for work

The employment status recode variable has been placed after the "8.9.2.1 Why are you not looking for work?" variable in the SPSS dataset.

7. Country Codes

The country codes used are the Geopolitical Entities and Codes (formerly the US Federal Information Processing Standards (FIPS) developed by the United States federal government for use in computer systems

It has a unique two letter code for each country and this list can be found in the appendix.

8. Interpretation of the voting results

The same information has been included in the Fieldwork Report and pertains to the voting question asked: “6.1 Are you registered to vote for the 2014 general election?”

During the time of the survey, the first round of the drive for voting registration for the 2014 National Elections took place. Most respondents interpreted the question as whether they have, or are going to, register for the 2014 elections during the upcoming or current voting registration drive. In other cases some respondents were unaware that they only had to register once and therefore thought they had to re-register in order to be eligible to vote.

The end result was voting registration percentages which were completely out of sync with the current provincial estimate determined by the IEC. Call backs made in January 2014 confirmed this misinterpretation, since again the question was misinterpreted by respondents due to the second registration drive that took place in January.

The second round of fieldwork done in March 2014 reflected normalized voting registration figures, since there was no registration drive taking place during that time. This should therefore be taken into account when assessing the voting registration results for this survey.

9. Geographic coordinates

The tablet device used to administer the questionnaire had a built in GPS device which was used to capture the latitude and longitude coordinate of each stand where a visit was made or successful interview was done. The GPS trigger was part of the questionnaire, therefore, when a questionnaire was opened by the fieldworker, a GPS coordinate was required by the data collection instrument before the fieldworker could continue with the administering of the questionnaire. In this way it was ensured that each visit/interview would have a GPS coordinate.

The GPS coordinate was therefore part of the digital file for each interview and as such was geocoded on a Geographic Information System (GIS) from where it was possible to run spatial queries to determine the accuracy and coverage of visits made and interviews done.

Since the Ward geography is not in sync with the Municipal or Census Geography, it means that in many cases the Small Area Level (SAL) boundaries did not match the Ward boundaries, causing Ward boundaries to cut across SAL polygons, resulting in various parts of a SAL covered by two or more Wards. However, each SAL code was assigned to a specific Ward, and as such, interviews were done in that part of the SAL which fell in the assigned Ward to assure that interviews done in those overlapping SAL areas were done within the correct Ward.

10. Weighting

Weights (field GCROweight) calculated by race and sex based on the 2011 Census data, were applied at Ward level. Please refer to the weighting document for more info.

11. Appendix

11.1 Appendix 1: Country Codes

Country	Code
Afghanistan	af
Albania	al
Algeria	ag
American Samoa	aq
Andorra	an
Angola	ao
Anguilla	av
Antarctica	ay
Antigua & Barbuda	ac
Argentina	ar
Armenia	am
Aruba	aa
Australia	as
Austria	au
Azerbaijan	aj
Bahrain	ba
Baker I.	fq
Bangladesh	bg
Barbados	bb
Belarus	bo
Belgium	be
Belize	bh
Benin	bn
Bermuda	bd
Bhutan	bt
Bolivia	bl
Bosnia & Herzegovina	bk
Botswana	bc
Bouvet I.	bv
Brazil	br
British Indian Ocean Territory	io
British Virgin Is.	vi
Brunei	bx
Bulgaria	bu
Burkina Faso	uv

Burundi	by
Cambodia	cb
Cameroon	cm
Canada	ca
Cape Verde	cv
Cayman Is.	cj
Central African Republic	ct
Chad	cd
Chile	ci
China	ch
Christmas I.	kt
Cocos Is.	ck
Colombia	co
Comoros	cn
Congo	cf
Congo, DRC	cg
Cook Is.	cw
Costa Rica	cs
Côte d'Ivoire	iv
Croatia	hr
Cuba	cu
Cyprus	cy
Czech Republic	ez
Denmark	da
Djibouti	dj
Dominica	do
Dominican Republic	dr
East Timor	tm
Ecuador	ec
Egypt	eg
El Salvador	es
Equatorial Guinea	ek
Eritrea	er
Estonia	en
Ethiopia	et
Falkland Is.	fk
Faroe Is.	fo
Fiji	fj
Finland	fi
France	fr
French Guiana	fg

French Polynesia	fp
French Southern & Antarctic Lands	fs
Gabon	gb
Gaza Strip	gz
Georgia	gg
Germany	gm
Ghana	gh
Gibraltar	gi
Glorioso Is.	go
Greece	gr
Greenland	gl
Grenada	gj
Guadeloupe	gp
Guam	gq
Guatemala	gt
Guernsey	gk
Guinea	gv
Guinea-Bissau	pu
Guyana	gy
Haiti	ha
Heard I. & McDonald Is.	hm
Honduras	ho
Howland I.	hq
Hungary	hu
Iceland	ic
India	in
Indonesia	id
Iran	ir
Iraq	iz
Ireland	ei
Isle of Man	im
Israel	is
Italy	it
Jamaica	jm
Jan Mayen	jn
Japan	ja
Jarvis I.	dq
Jersey	je
Johnston Atoll	jq
Jordan	jo
Juan De Nova I.	ju

Kazakhstan	kz
Kenya	ke
Kiribati	kr
Kuwait	ku
Kyrgyzstan	kg
Laos	la
Latvia	lg
Lebanon	le
Lesotho	lt
Liberia	li
Libya	ly
Liechtenstein	ls
Lithuania	lh
Luxembourg	lu
Macedonia	mk
Madagascar	ma
Malawi	mi
Malaysia	my
Maldives	mv
Mali	ml
Malta	mt
Marshall Is.	rm
Martinique	mb
Mauritania	mr
Mauritius	mp
Mayotte	mf
Mexico	mx
Micronesia	fm
Midway Is.	mq
Moldova	md
Monaco	mn
Mongolia	mg
Montenegro	mw
Montserrat	mh
Morocco	mo
Mozambique	mz
Myanmar	bm
Namibia	wa
Nauru	nr
Nepal	np
Netherlands	nl

Netherlands Antilles	nt
New Caledonia	nc
New Zealand	nz
Nicaragua	nu
Niger	ng
Nigeria	ni
Niue	ne
Norfolk I.	nf
North Korea	kn
Northern Mariana Is.	cq
Norway	no
Oman	mu
Pakistan	pk
Palau	ps
Panama	pm
Papua New Guinea	pp
Paracel Is.	pf
Paraguay	pa
Peru	pe
Philippines	rp
Pitcairn Is.	pc
Poland	pl
Portugal	po
Puerto Rico	rq
Qatar	qa
Reunion	re
Romania	ro
Russia	rs
Rwanda	rw
Samoa	ws
San Marino	sm
Sao Tome & Principe	tp
Saudi Arabia	sa
Senegal	sg
Serbia	sr
Seychelles	se
Sierra Leone	sl
Singapore	sn
Slovakia	lo
Slovenia	si
Solomon Is.	bp

Somalia	so
South Africa	sf
South Georgia & the South Sandwich Is.	sx
South Korea	ks
Spain	sp
Spratly Is.	pg
Sri Lanka	ce
St. Helena	sh
St. Kitts & Nevis	sc
St. Lucia	st
St. Pierre & Miquelon	sb
St. Vincent & the Grenadines	vc
Sudan	su
Suriname	ns
Svalbard	sv
Swaziland	wz
Sweden	sw
Switzerland	sz
Syria	sy
Taiwan	tw
Tajikistan	ti
Tanzania	tz
Thailand	th
The Bahamas	bf
The Gambia	ga
Togo	to
Tokelau	tl
Tonga	tn
Trinidad & Tobago	td
Tunisia	ts
Turkey	tu
Turkmenistan	tx
Turks & Caicos Is.	tk
Tuvalu	tv
Uganda	ug
Ukraine	up
United Arab Emirates	ae
United Kingdom	uk
United States	us
Uruguay	uy
Uzbekistan	uz

Vanuatu	nh
Vatican City	vt
Venezuela	ve
Vietnam	vm
Virgin Is.	vq
Wake I.	wq
Wallis & Futuna	wf
West Bank	we
Western Sahara	wi
Yemen	ym
Zambia	za
Zimbabwe	zi