

Between the devil and the deep blue sea: navigating data collection in a time of uncertainty

Inputs presented at the Gauteng City Region Observatory (GCRO) methods workshop on collecting social data: Adapting to the COVID-19 pandemic

23 October 2020

Solly Molayi and Niël Roux
Statistics South Africa



stats sa

Department:
Statistics South Africa
REPUBLIC OF SOUTH AFRICA

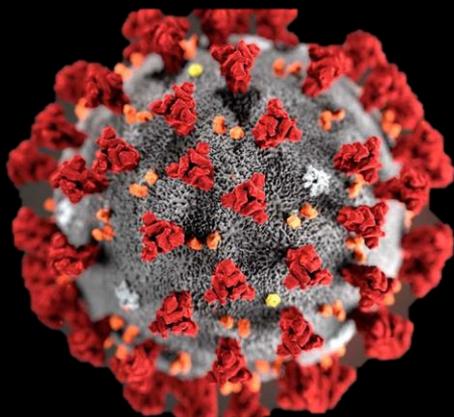


**Data collection
suspended on 20
March 2020**

**QLFS Q2 data collection
resumed mid-May 2020
using CATI**

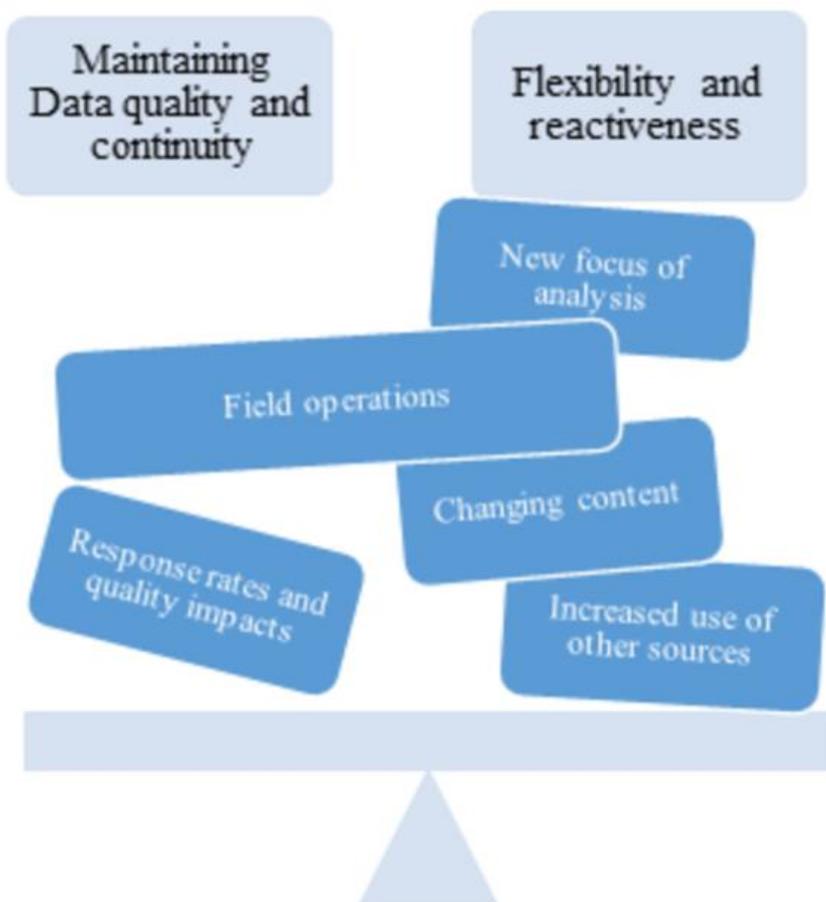
**GHS, DTS, GPSJS data
collection resumed on
September 1st using CATI**

**Uncertain
future**



The need to continue data collection and to provide survey data (to continue with business as usual) considered important enough by Stats SA to explore alternative means of data collection while...

..maintaining the balance.



Most substantial impact on ability to gather and disseminate data:

- Sampling
- Data collection
- Questionnaire content
- Data analysis
- Data continuity
- Quality

Steps taken to mediate the impact of COVID-19

Short term

- Continue with survey programmes using CATI

Medium & Long term

- Establish Quality Assurance Mixed-Mode Committee to make proposals for 2021 survey programme and beyond

Census 2021

- Test alternative listing & data collection methodologies: CATI, CAWI
- 2 test conducted
- Pilot due in February 2021

Short term: 2020 CATI surveys

- **Inadequate sample:** Master Sample based on Dwelling Units and respondent telephone numbers had to be harvested from previous surveys – suspension of household rotation, and lower response rates
- **Data collection:** Face-to-face interviews using CAPI replaced by remote data collection using CATI (telephones) – higher non-response rates
- **Data collection instruments** shortened to adapt to data collection mode
- **Inadequate time** to conduct thorough development, testing, and fieldwork training.
- **Analysis and comparability** are negatively affected by lower response rates.
- **Conclusion:** Not clear if, and how, surveys will be conducted in 2021. What is clear is that CATI, without an adjusted sample, is not an optimal solution.

Medium to long term: 2021 surveys and beyond

Medium term

- Implementation of QLFS, GHS, DTS, GPSJS in 2021
- **Weighting risk** of maintaining data quality and continuity using CAPI with a real risk of further outbreaks and its health and reputational impacts
- **Proposal** to continue with face-to-face interviews following health protocols and PPE, or limiting physical interaction to initial publicity and unpacking

Long term

- Require review of sampling frame
- Test and review methodologies post census

Census 2021:

- Initially planned for face-to-face CAPI
- Plans abruptly changed by COVID-19
- Multiple tests conducted since then focusing on CATI and CAWI
- Pilot planned for February 2021

Since we don't know the future we are faced with impossible choices

- Continue data collection using face-to-face data collection and CAPI but using PPE and following health protocols
- Benefit of using the whole master sample and updated samples
- Risk being shut down again by further outbreaks, either nationally, or in spots where infection rates are higher.
- Revert to using remote data collection using CATI (or CAWI)
- Benefit: Able to complete data collection faster and with less threat to the health and wellbeing of survey officers and respondents
- Certain risk of poorer outcomes due to smaller and warped samples, higher non-response rates, etc.

For surveys it is a case of
hoping for the best or
mitigating the worst.

FUTURE
LOADING...

0%

100%

