# MANAGING DATA QUALITY IN CAPI SURVEYS

## HOUSEHOLD SURVEYS IN A CHANGING DATA LANDSCAPE:

### CHALLENGES, OPPORTUNITIES AND AN AGENDA FOR THE FUTURE
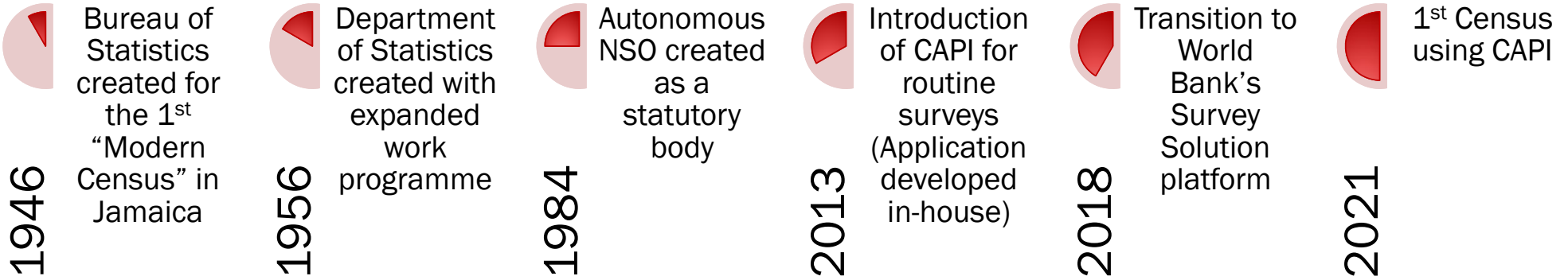
FEBRUARY 28, 2020

Presented by:
Leesha Delatie-Budair
Deputy Director General
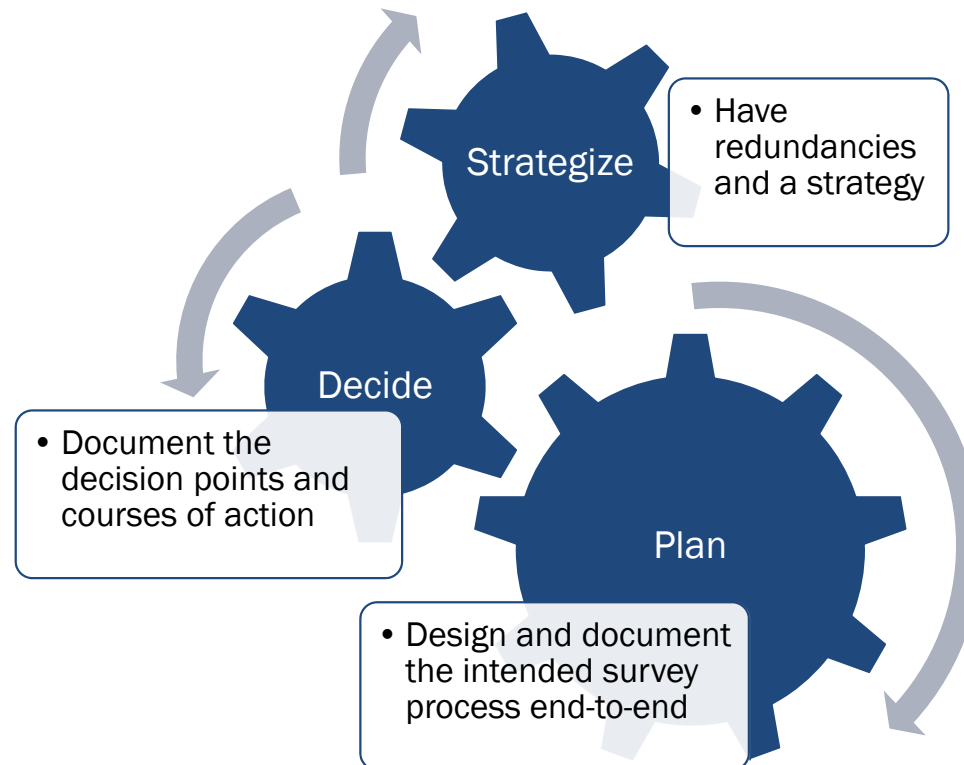Statistical Institute of Jamaica

# BACKGROUND

Statistical Institute of Jamaica

## PAPI

**1946** — Bureau of Statistics created for the 1st "Modern Census" in Jamaica

**1956** — Department of Statistics created with expanded work programme

**1984** — Autonomous NSO created as a statutory body

## CAPI

**2013** — Introduction of CAPI for routine surveys (Application developed in-house)

**2018** — Transition to World Bank's Survey Solution platform

**2021** — 1st Census using CAPI

# BEFORE DATA COLLECTION

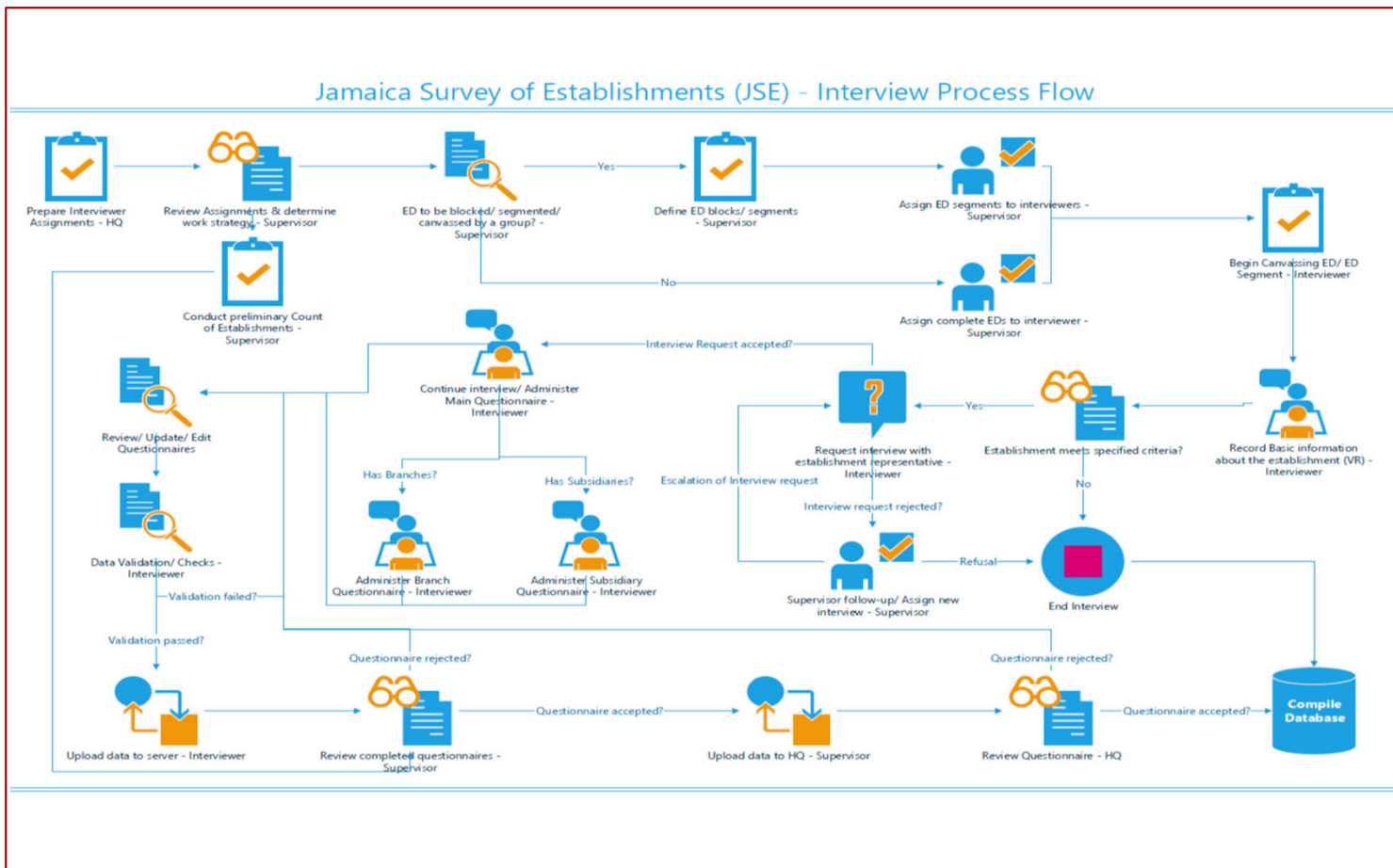Statistical Institute of Jamaica

# DESIGN THE QUESTIONNAIRE WITH THE RESULTS IN MIND

- Use simple language

- Ask exactly what you want?

- Include probing questions where possible?

- Be mindful of the limitations of:
    - The interviewers who will administer the questionnaire
    - The respondent
    - The environment for field work
    - The device being used

# Have an enumeration plan!

# DECISIONS! DECISIONS! DECISIONS!

| What are the data quality checks to be applied? | At what level should quality checks be applied? | Where should the quality checks be applied? |
|---|---|---|
| • Range checks, consistency checks, outliers etc. | • Unit, group, PSU etc. | • Within CAPI, data processing, data analysis |

| Who is responsible for each quality check? | Document the most plausible scenarios, and the mitigating strategies? | How do you define a completed interview? |
|---|---|---|
| • Interviewer, supervisor, headquarter user, analyst etc. | • If X then Y | • % of questions, or key questions? |

# STRATEGIZE

Define escalation procedures and strategies for field issues

Have a strategy for HQ users

- When do they start?
- How many are needed and how will they be supervised?

Have a communication strategy!

- How does information flow to and from the field?
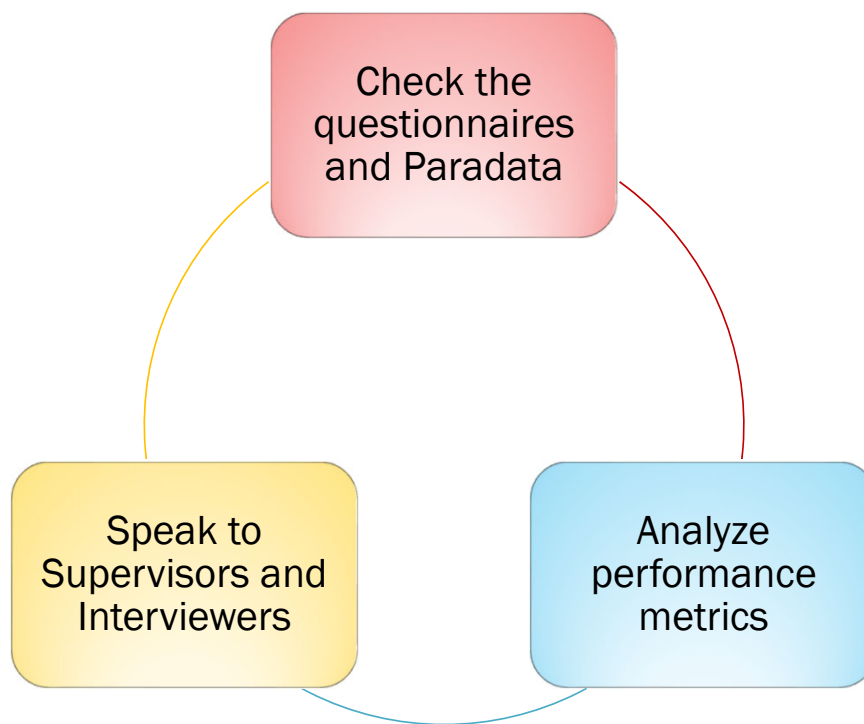- What should be communicated to the public, how, when and by whom?

Statistical Institute of Jamaica

# DURING DATA COLLECTION
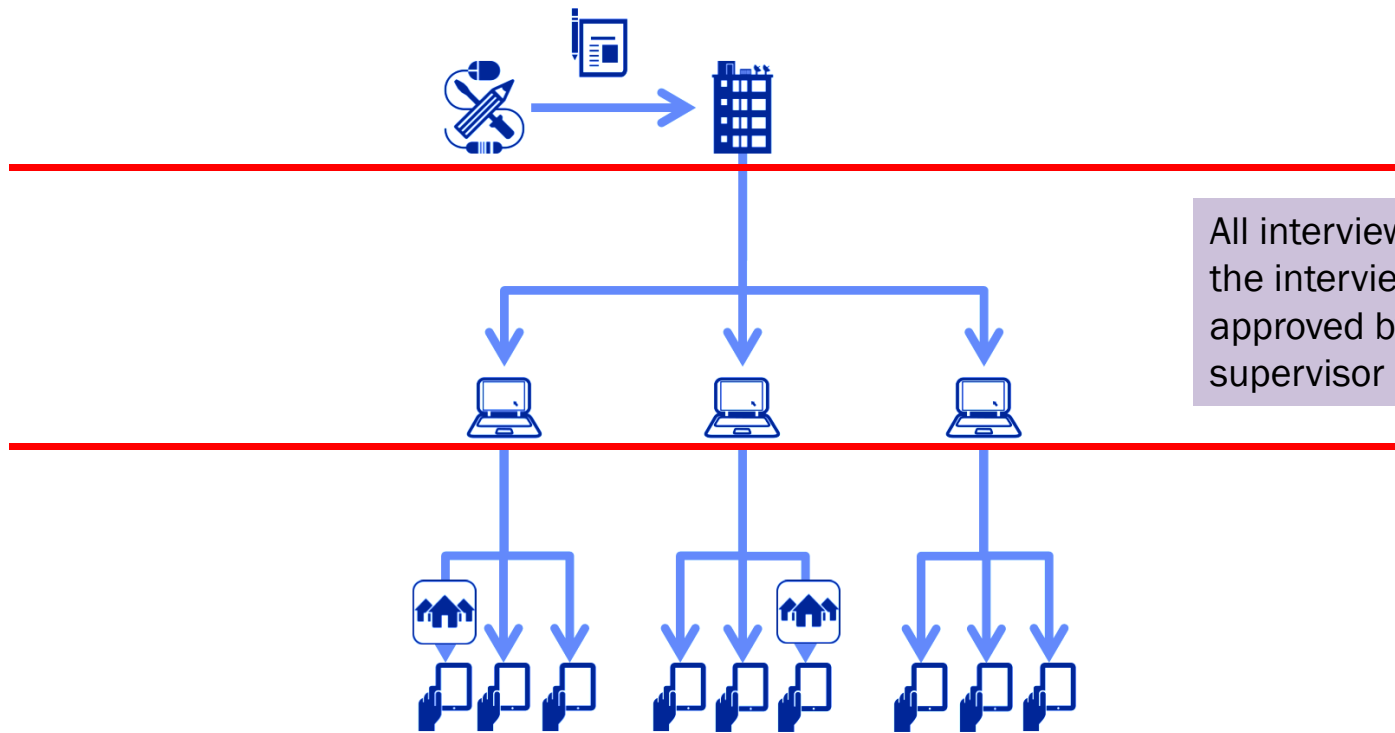
# SURVEY SOLUTIONS PROCESS FLOW

All interview data submitted by the interviewers must be approved by 2 persons: supervisor and HQ

# NECESSARY ACTIONS DURING COLLECTION

- Supervise the Interviewers!
  - Monitor Completion Rates
  - Monitor Rejection Rates
  - Monitor data quality of questionnaires
- The Survey Solutions Approach
  - Quick access to common problems: not answered or erroneous questions.
  - Interview-specific bookmarks for follow up after corrections.
  - Comments of interviewers and supervisors are recorded and preserved for analysis.



**Survey Solutions: Supervisor View**

# NECESSARY ACTIONS DURING COLLECTION

- Supervise the Supervisors!
  - Monitor Completion Rates
    - Are teams lagging?
    - Are interviewers moving too fast
  - Monitor Rejection Rates
  - Monitor data quality of approved questionnaires
- The Survey Solutions Approach
  - HQ approves Questionnaires approved by Supervisors
  - Standard Reports
    - Map Reports
    - Teams and Statuses
    - Quantity
    - Speed
    - Other standard reports



Interviews

| | INTERVIEW KEY | Hide | Identifying Questions | RESPONSIBLE | LAST UPDATE | HAS ERRORS | STATUS |
|---|---|---|---|---|---|---|---|
| ☐ | 74-55-39-63 | | | WilliamC | 5/3/2017 | — | ApprovedBySupervisor |
| ☐ | 03-21-61-17 | | HEAD OF HOUSEHOLD:Albert, ADDRESS:36 Orange Blossom Way | GraceR | 4/28/2017 | — | ApprovedBySupervisor |
| ☐ | 69-20-31-02 | | Identification number of household:754 | WilliamC | 4/28/2017 | Yes | ApprovedBySupervisor |
| ☐ | 37-79-13-21 | | HEAD OF HOUSEHOLD:Frank, ADDRESS:243 Columbia Rd | LilyC | 4/28/2017 | — | ApprovedBySupervisor |
| ☐ | 05-72-56-68 | | HEAD OF HOUSEHOLD:Vernon , ADDRESS:98 Yale Street | LilyC | 4/28/2017 | — | ApprovedBySupervisor |
| ☐ | 79-06-59-40 | | HEAD OF HOUSEHOLD:Edith, ADDRESS:66 Orange Blossom Way | OwneG | 4/28/2017 | — | ApprovedBySupervisor |
| ☐ | 60-78-25-10 | | | | | | |

**Survey Solutions: HQ User View**

# NECESSARY ACTIONS DURING COLLECTION

- Analyze Key variables
  - Location
  - Date and Time
  - Time between responses (Paradata is a Gold Mine!)

- The Survey Solutions Approach
  - Produced automatically.
  - Shows who changed the data, when and to what value.
  - Allows low-level quality control and check of interviewers' behavior:



| interview_id | order | event | responsible | role | timestamp | offset | parameters |
|---|---|---|---|---|---|---|---|
| f98fad3a65ab491d9949d4e98a7d0386 | 1 | SupervisorAssigned | interviewer1 | Interviewer | 2020-02-10T17:30:16 | -05:00:00 | |
| f98fad3a65ab491d9949d4e98a7d0386 | 2 | InterviewerAssigned | interviewer1 | Interviewer | 2020-02-10T17:30:16 | -05:00:00 | interviewer1 |
| f98fad3a65ab491d9949d4e98a7d0386 | 3 | KeyAssigned | | | 2020-02-10T17:30:16 | -05:00:00 | 93-83-75-16 |
| f98fad3a65ab491d9949d4e98a7d0386 | 4 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:30:31 | -05:00:00 | parish\|\|St. Thomas\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 5 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:30:31 | -05:00:00 | oType\|\|Wholesale\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 6 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:30:31 | -05:00:00 | oType_code\|\|1007\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 7 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:30:31 | -05:00:00 | outlet_name\|\|L. SALMON\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 8 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:30:31 | -05:00:00 | outlet_code\|\|03073\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 9 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:30:31 | -05:00:00 | town\|\|SPRINGFIELD\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 10 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:30:31 | -05:00:00 | address\|\|SPRINGFIELD\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 11 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:32:19 | -05:00:00 | estab_open\|\|3\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 12 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:32:30 | -05:00:00 | sub_estab_check\|\|1\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 13 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:33:44 | -05:00:00 | estab_sup_check\|\|1\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 14 | AnswerSet | interviewer1 | Interviewer | 2020-02-10T17:34:06 | -05:00:00 | estab_sub_address\|\|Springfield Road\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 15 | AnswerRemoved | | | 2020-02-10T17:34:42 | -05:00:00 | estab_sub_address\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 16 | Paused | | | 2020-02-10T17:34:47 | -05:00:00 | |
| f98fad3a65ab491d9949d4e98a7d0386 | 17 | Resumed | interviewer1 | Interviewer | 2020-02-11T08:31:20 | -05:00:00 | |
| f98fad3a65ab491d9949d4e98a7d0386 | 18 | Paused | | | 2020-02-11T08:31:23 | -05:00:00 | |
| f98fad3a65ab491d9949d4e98a7d0386 | 19 | AnswerRemoved | | | 2020-02-11T09:56:50 | -05:00:00 | estab_sup_check\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 20 | AnswerSet | interviewer1 | Interviewer | 2020-02-11T09:57:05 | -05:00:00 | sub_estab_check\|\|2\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 21 | Completed | | | 2020-02-11T09:57:43 | -05:00:00 | |
| f98fad3a65ab491d9949d4e98a7d0386 | 22 | Restarted | | | 2020-02-11T09:58:15 | -05:00:00 | |
| f98fad3a65ab491d9949d4e98a7d0386 | 23 | AnswerSet | interviewer1 | Interviewer | 2020-02-11T09:58:38 | -05:00:00 | visit_number\|\|1\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | 24 | AnswerSet | interviewer1 | Interviewer | 2020-02-11T09:59:19 | -05:00:00 | visitTimestamp\|\|2020-02-11 09:59:19\|\|1 |
| f98fad3a65ab491d9949d4e98a7d0386 | 25 | AnswerSet | interviewer1 | Interviewer | 2020-02-11T09:59:29 | -05:00:00 | contact_inf\|\|99\|\|1 |
| f98fad3a65ab491d9949d4e98a7d0386 | 26 | AnswerSet | interviewer1 | Interviewer | 2020-02-11T09:59:49 | -05:00:00 | contact_inf_oth\|\|Establishment out of business\|\|1 |
| f98fad3a65ab491d9949d4e98a7d0386 | 27 | VariableDisabled | | | 2020-02-11T10:00:04 | -05:00:00 | iCode_12\|\| |
| f98fad3a65ab491d9949d4e98a7d0386 | | | | | | | |
| f98fad3a65ab491d9949d4e98a7d0386 | | | | | | | |
| f98fad3a65ab491d9949d4e98a7d0386 | | | | | | | |
| f98fad3a65ab491d9949d4e98a7d0386 | | | | | | | |
| f98fad3a65ab491d9949d4e98a7d0386 | | | | | | | |
| f98fad3a65ab491d9949d4e98a7d0386 | | | | | | | |

**Survey Solutions: ParaData**

# SURVEY SOLUTIONS: LISTEN TO THE FIELD TEAMS

- Listen to the Field Teams, take corrective action and provide feedback

- Get regular updates from the field staff regarding:
  - questionnaire logic and skips;
  - false or missing error messages and their interpretation;
  - limits and formats;
  - navigation/ergonomics;
  - software bugs/updates.

- Introduce corrections and updates to improve.

# AFTER DATA COLLECTION

# UNDERSTAND WHAT HAPPENED?

## Debrief

- Field Teams
- Survey Managers
- Analysts

## Assess

- Rejection percentages
- Location, Date and Time of Interviews
- Key variables

## Document

- What worked?
- What Didn't?
- Lessons Learned

**STATISTICAL INSTITUTE OF JAMAICA**

(STATIN)

**7 Cecelio Avenue**

**Kingston 10**

**Jamaica W.I.**

**(876) 630-1600**

INFO@STATINJA.GOV.JM | STATINJA.GOV.JM

@STATINJA

Thank You for Your Attention