



2021 POPULATION AND HOUSING CENSUS

DATA QUALITY MONITORS' MANUAL

GHANA
YOU COUNT, GET COUNTED

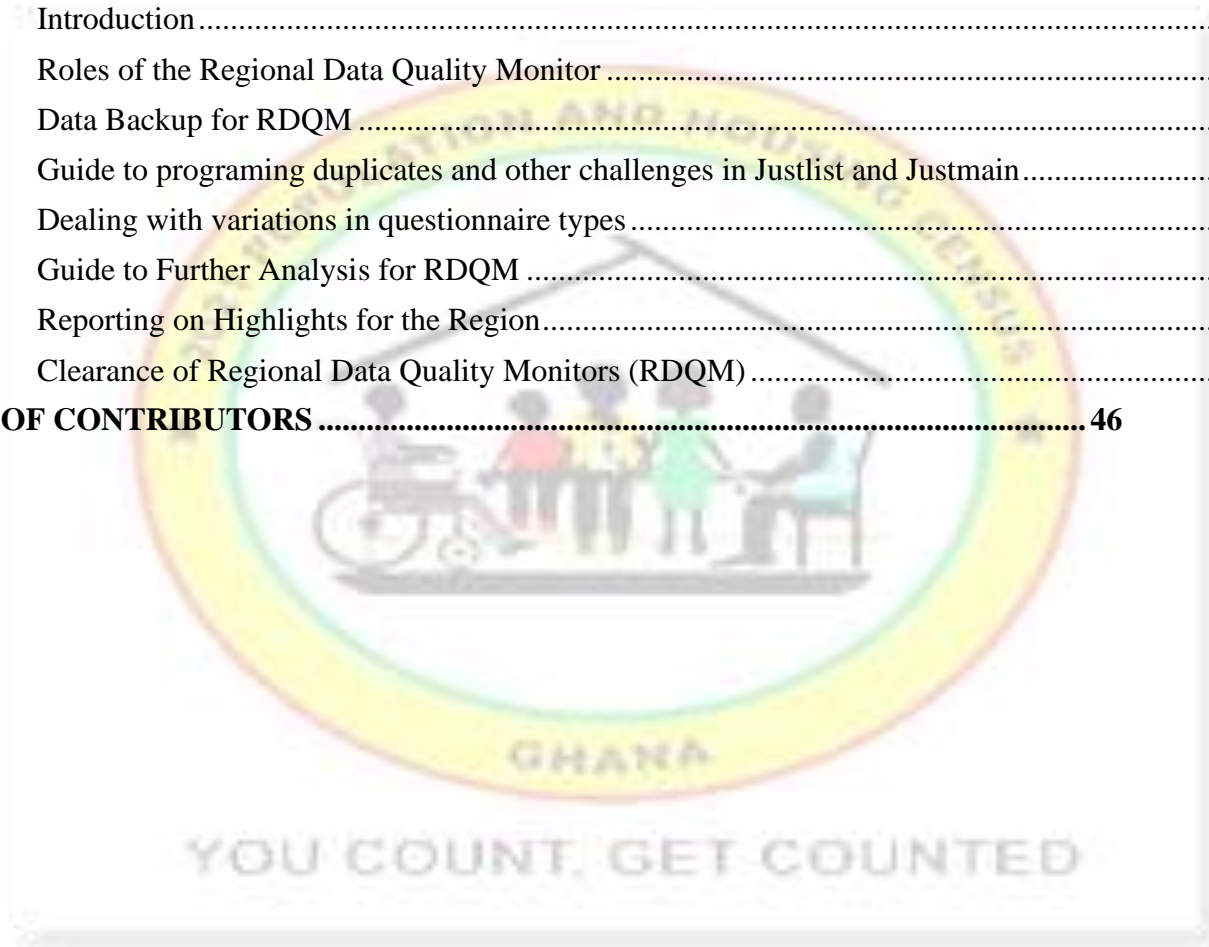
GHANA STATISTICAL SERVICE

APRIL 2021

TABLE OF CONTENTS

TABLE OF CONTENTS	i
LIST OF FIGURES	iii
ACRONYMS AND ABBREVIATIONS	iv
CHAPTER ONE: INTRODUCTION	1
1.1 Introduction.....	1
1.2 Data Quality Monitoring Process.....	1
1.3 Responsibilities of National Data Quality Monitors (NDQM).....	3
1.4 Responsibilities of Regional Data Quality Monitors (RDQM)	4
1.5 Responsibilities of the District Data Quality Monitor (DDQM)	5
1.6 Rules for a Data Quality Monitor	7
1.7 Offences and Penalties for a Data Quality Monitor.....	7
1.8 Request for Data by Unauthorized Persons outside the 2021 PHC Management	8
CHAPTER TWO: DATA QUALITY MONITORING & QUERYING IN CSPRO/STATA	9
2.1 Creating folder Icons to Start Work.....	9
2.1.1 <i>Instructions to access folders for the 2021PHC Data Quality Monitoring exercise</i>	9
2.1.2 <i>Creating shortcuts to working folders</i>	9
2.2 Accessing Data using CSPro at all levels	10
2.2.1 <i>Login</i>	10
2.2.2 <i>Processing Data</i>	10
2.2.3 <i>Functions of each Action on Menu</i>	11
CHAPTER THREE: ACTIVITIES OF DQM DURING THE LISTING STAGE	14
3.1 Activities to perform during Listing	14
3.1.1 <i>Step 1 – Accessing and Processing Data in CSPro</i>	14
3.1.2 <i>Step 2 – Analyzing Listing Data Using STATA and Excel</i>	16
3.1.3 <i>Further Analysis on Listing Data in STATA</i>	24
3.2 Step 3 - Workload Management Strategy for 2021 PHC.....	26
3.2.1 <i>Background</i>	26
3.2.2 <i>Listing Strategy</i>	26
3.2.3 <i>E-Strategy</i>	26
3.2.4 <i>Operationalization of the Strategy</i>	27
3.2.5 <i>Sensitizing Field Officers</i>	27
3.2.6 <i>Guidelines for Reassignment and Support Creation</i>	27
CHAPTER FOUR: DURING ENUMERATION	29
4.1 Procedure after Data Access – Main Enumeration	29
4.1.1 <i>Step 1 – Accessing and Data validation in CSPRO</i>	29
4.1.2 <i>Step 2 – Analyzing Main Data using STATA and Excel</i>	29

4.1.3	Step3 - Checklist after generating the monitor report during Main Enumeration.....	30
4.1.4	Step4 - Guide to further analysis of population data	32
4.1.5	Steps 5 – Standard daily reporting format	34
CHAPTER FIVE: CLEARANCE OF FIELD TEAMS.....		36
5.1	Clearance to leave the field after data collection	36
5.2	Clearance of Supervisors and Enumerators in an SA (Team) by DDQMT	36
5.3	Clearance of Supervisors and Enumerators in an SA (Team) at Headquarters	36
5.4	Clearance of the District Data Quality Monitor	37
CHAPTER SIX: REGIONAL DATA QUALITY MONITORING		39
6.1	Introduction	39
6.2	Roles of the Regional Data Quality Monitor	39
6.3	Data Backup for RDQM	40
6.4	Guide to programing duplicates and other challenges in Justlist and Justmain.....	40
6.5	Dealing with variations in questionnaire types	40
6.6	Guide to Further Analysis for RDQM	42
6.7	Reporting on Highlights for the Region.....	43
6.8	Clearance of Regional Data Quality Monitors (RDQM)	45
LIST OF CONTRIBUTORS		46



LIST OF FIGURES

Figure 1.1:	Data Quality Monitoring Process.....	2
Figure 1.2:	Data Quality Management Team Communication Channel.....	3
Figure 2.1:	DM- Home Action Menu.....	10
Figure 3.1:	Justlist file.....	14
Figure 3.2:	User Unnumbered Messages	15
Figure 3.3:	Sample Error Messages	16
Figure 3.4:	Sample Monitor Report File	18
Figure 3.5:	EA_stats sheet.....	19
Figure 3.6:	Struct-gaps_file sheet.....	22
Figure 3.7:	dups_file sheet	22
Figure 3.8:	loc_with_nopop sheet	23
Figure 3.9:	Sample bad GPS sheet	24
Figure 4.1:	Sample Main Error Messages	29
Figure 4.2:	Sample Monitor Report Main.....	30
Figure 4.3:	Main Dashboard.....	32
Figure 4.4:	Sample Daily Report – Listing	35
Figure 4.5:	Sample Daily Report – Main	35
Figure 5.1:	Sample Standard Clearance Reporting Format.....	38
Figure 6.1:	Sample of variation in questionnaire types from Listing and Main	41
Figure 6.2:	Sample Further Analysis Report - Population Pyramid	43
Figure 6.3:	Sample of Highlights - Listing	44
Figure 6.4:	Sample of Highlights - Main	44



ACRONYMS AND ABBREVIATIONS

PHC	Population and Housing Census
BF	Building Footprint
CAPI	Computer-Assisted Personal Interviewing
DCO	District Census Officer
DDQMT	District Data Quality Management Team
DFS	District Field Officer
DIT	District IT
DDQM	District Data Quality Monitor
EA	Enumeration Area
GIS	Geographic Information System
GPS	Global Positioning System
ISCO	International Standard Classification of Occupations
ISIC	International Standard Industrial Classification of all Economic Activities
NDQM	National Data Quality Monitor
RDQM	Regional Data Quality Monitor
RFS	Regional Field Supervisor
RIT	Regional IT
RDQMT	Regional Data Quality Management Team
SA	Supervisory Area

CHAPTER ONE: INTRODUCTION

1.1 Introduction

Censuses are herculean tasks that are costly and therefore require detailed planning to achieve great success. The success of a Census largely depends on the quality of training and commitment of staff, as well as mechanisms for effective Data Quality Monitoring and editing. To ensure the reliability of data, every data should be subjected to editing, which is the procedure for detecting errors in and between data records, during and after data collection and capture, and adjusting items where necessary (UN Handbook on Population and Housing Census Editing, 2010). This procedure requires a strong teamwork between subject matter specialists and data processing experts.

The strategy for the Ghana 2021 PHC Data Processing Activities would employ multiple data editing approaches for optimum results. Consistency checks will be incorporated in the data capture program (Computer Assisted Personal Interview - CAPI) to aid real-time error detection and data cleaning by the enumerators. In addition, the National, Regional and District Data Quality Monitors will run a full batch program (CSPRO and STATA) to query the synced data for structure and content errors. The errors detected (error file) will be communicated to supervisors (via Phone call, SMS and WhatsApp) through the District Data Quality Management team for possible data editing and re-syncing.

The final part of the editing would involve developing and running a full batch editing application to correct any outstanding errors in the data file. The output file from the iterative process would be checked until the data is finalized for tabulation, report writing and dissemination.

1.2 Data Quality Monitoring Process

The pressure to release results from collected data soon after the end of fieldwork, more especially, with the use of CAPI requires a speedily and effective approach to data processing. In view of this, the concept of Field Data Quality Monitors would be used in the 2021 PHC to check the synchronized data to minimize final batch edit imputations. A three-tier monitoring system will be adopted; National Data Quality Monitors (NDQMs), Regional Data Quality Monitors (RDQMs) and District Data Quality Monitors (DDQMs). Each Data Quality Monitor would be given access to the data based on their level of engagement, thus, NDQM will access

national data; RDQM, regional data; and DDQM, district data. These levels will each conduct consistency checks on data collected.

All Monitors would run batch query programs to generate error file. Each District Data Quality Monitor is expected to analyze the summary statistics and error file and ensure that all issues with the data are communicated to the respective Supervisors who will, in turn, communicate same to the Enumerator for appropriate corrections to be made. The Enumerator is expected to correct the errors and re-sync the data to the Supervisor and the Cloud. This iterative process of data download and error feedback to the Enumerators would continue until the end of the field data collection. It is expected that most of the errors would have been corrected at this stage of data processing.

Figure 1.1: Data Quality Monitoring Processes

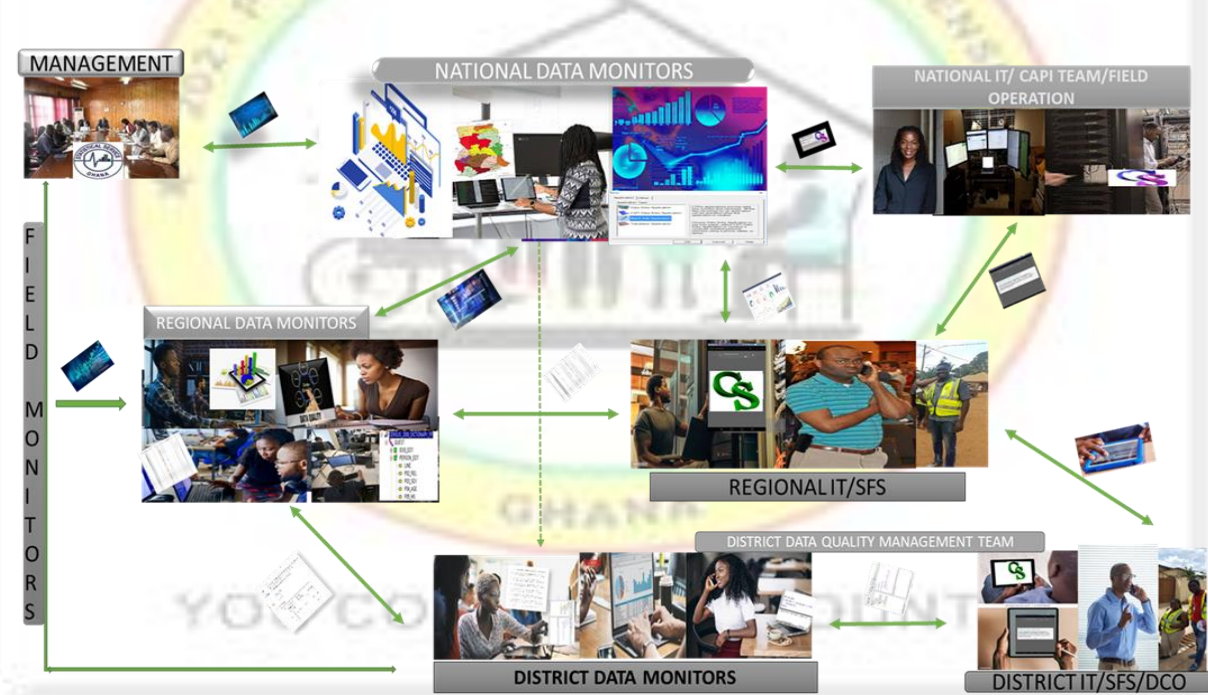
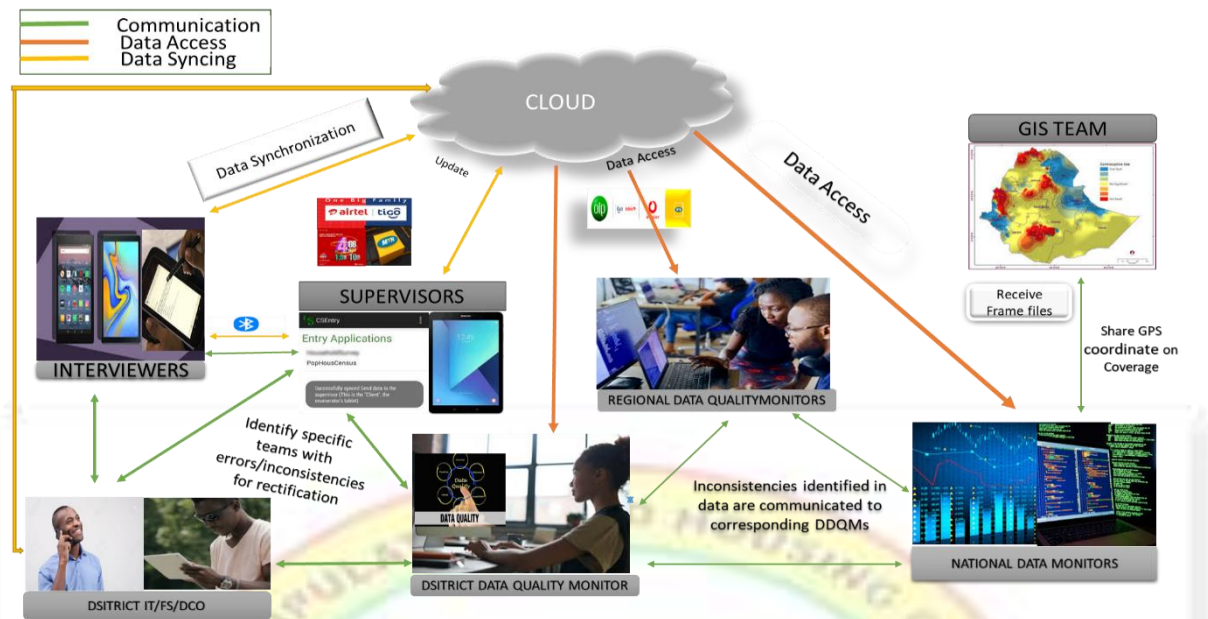


Figure 1.1: Data Quality Management Team Communication Channel



1.3 Responsibilities of National Data Quality Monitors (NDQM)

1. Prepare the CSPro batch programs, the STATA do files and the report templates
2. Train the RDQMs/Trainers
3. Download National Data
4. Check for structure and content errors
5. Liaise between all the DDQMs and the PHC Management Team
6. Send feedback to RDQMs/DDQMs
7. Receive feedback from RDQMs
8. Send regular reports to Management
9. Clear SAs/Teams for payment
10. Clear RDMT/DDQMT for final payment
11. Edit and finalize the PHC Census data
12. Generate tables
13. Convert data to REDATAM
14. Develop a National Sampling Frame
15. Archive PHC Data

1.4 Responsibilities of Regional Data Quality Monitors (RDQM)

Before Training

1. Assist in training District Data Quality Monitors (DDQMs)
2. Get to know DDQMs in their region and establish rapport for effective coordination
3. Organize an inception meeting for the DDQMs in their region to plan for the data monitoring process
4. Ensure all Logistics to be used by DDQMs are delivered to their respective districts
5. Assist to deploy DDQMs to the field
6. Create a communication platform (WhatsApp group) for DDQMs in the region
7. Report on distribution of logistics to National Data Quality Monitors

During Training

8. Ensure that DDQMs facilitate in the training of Field Officers
9. Visit DDQMs within their jurisdiction whenever the need be and ensure that all challenges that may emerge during training and fieldwork are resolved
10. Run error report on training data
11. Share error report with DDQMs for their respective districts
12. Submit the final list of DDQMs in their regions for payment of allowances
13. Ensure that daily evaluation form is completed by each DDQMs
14. Monitor the implementation of the data monitoring procedures and ensuring adherence to timelines

During Fieldwork

15. Access Data for their respective region
16. Run batch program to check for training data
17. Generate errors messages for the Region
18. Check for structure and content errors in the data
19. Assist in resolving duplicates, validate cases and records for districts in their regions
20. Communicate identified errors and inconsistencies to their respective DQDMs
21. Receive/Send feedback to DDQMs
22. Ensure errors identified at each district are resolved
23. Liaise with other members of the Regional Data Quality Management Team (RQMT)

- to resolve all data issues in the region
- 24. Generate and analyze reports and tables at the regional level
- 25. Give regular updates on progress of work to NDQMs
- 26. Seek the welfare of all DDQMs in the region
- 27. Ensure complete coverage of all EAs in the districts within the region

After Fieldwork

- 28. Clear DDQMs following the outlined protocols
- 29. Retrieve all Logistics assigned to DDQMs
- 30. Compile and forward list of cleared DDQMs
- 31. Assist to edit the final PHC data
- 32. Assist to generate tables using the agreed tabulation plan
- 33. Any other work that may be assigned by NDQM/Census Management Team

1.5 Responsibilities of the District Data Quality Monitor (DDQM)

- 1. Follow the outlined procedures for Data Quality Monitoring for the district
- 2. Access District Level data
- 3. Run batch application to ensure that all structures are uniquely identified.
- 4. Sort data to ensure that all the synced data from the different regions, districts, Sub-districts, Enumeration Areas (EAs) are properly organized in ascending order for further processing
- 5. Generate error reports in CSPro
- 6. Export data from CSPro to STATA for further processing and reporting
- 7. Check for structure and content errors in the data
- 8. Prepare and share workload management report with other members of the DDQMT
- 9. Ensure total number of localities expected for each Type3 EA have been accounted for
- 10. Ensure all listed population have been enumerated
- 11. Share error messages with other members of the DDQMT
- 12. Communicate identified errors and inconsistencies to the field officers for correction
- 13. Continuously run checks on data to ensure that errors identified have been corrected
- 14. All structures with household and non-household population have been correctly enumerated
- 15. Liaise between the Field Teams and the RDQMs
- 16. Give regular updates on progress of work to RDQMs

17. Generate and analyze reports and tables at the district level
18. Ensure complete Coverage in all EAs in the district
19. Any other related task that may be assigned

As a member of DDQMT, your roles cover 4 cardinal areas namely; before training, during training, during Fieldwork and Clearance.

Before Training

20. Assist to secure conducive venue for training of field officers at the district
21. Assess the cleanliness of the environment, class size, furniture arrangements, availability of functional electrical sockets and other related issues
22. Register and verify certificates trainees
23. Assist to check arrival and adequacy of field and training logistics
24. Compile and paste list of participants of each class before training commences
25. Assist in publicity of the 2021 PHC

During Training

26. Facilitate training of Field Officers
27. Ensure daily registration of Field Officers
28. Prepare Teams for Field Practice
29. Assist in Final Team Formation
30. Print Field officers' ID cards
31. Assist in Tablet Re-provisioning and Logistics Distribution
32. Facilitate Part Payment of Allowances
33. Deploy Teams for Fieldwork

During Fieldwork

34. Perform Monitoring, Supervision and Evaluation for listing
35. Implement workload management strategy
36. Assist in implementing Census Night activities
37. Perform Monitoring, Supervision and Validation for main enumeration

Clearance

38. Clear field teams after completion of field work
39. Compile and forward list of cleared teams/enumerators for payment
40. Produce clearance reports in the required formats

A District Data Quality Monitor's work will be deemed complete when all teams (enumerators in an SA) have been cleared. It is after this that a District Data Quality Monitor can then be cleared by the RDQM. Note that this will be done simultaneously with other members of the DDQMT.

1.6 Rules for a Data Quality Monitor

1. *A Data Quality Monitor:*
 - a) Shall be **under oath** of secrecy and, therefore, is not expected to share any part or whole data with unauthorized persons at any level;
 - b) Is **not permitted** to provide any statistics to anybody outside the DDQMT and authorized PHC monitoring team;
 - c) Shall operate at the area of jurisdiction: district monitor at the district level, regional monitor at the region and national monitors at the national level;
 - d) Shall **NOT** leave his/her area of operation without prior permission from his/her immediate supervisor – district from regional and Regional from National
2. *Your engagement as a DM is full time. Abandoning your post before completion of fieldwork will amount to gross irresponsibility and misconduct and will make you liable to be exempted from future participation in GSS activities and again, lead to loss of all financial entitlements*

1.7 Offences and Penalties for a Data Quality Monitor

Given Sections (50) and (54) of the Statistical Service Act, 2019 (Act 1003); *a person who*

- a. *being an employee of the service or a public corporation or a partner institution working in collaboration with the service, discloses, publishes or communicate to any person, information acquired in the course of employment to a third party without lawful authority contrary to section 31;*
- b. *publishes or communicates information disclosed in contravention of this Act to any person;*

- c. *fails to comply with or contravenes any terms or conditions of the oath or affirmation taken under this Act;*
- d. *destroys, defaces or mutilates any form or other documents that contain particulars obtained in the pursuance of this Act without lawful authority;*
or
- e. *Contravene any provision of the Act,*
commits an offence and is liable on summary conviction to a fine of not more than two hundred penalty units or a term of imprisonment of not more than twelve months or both.

1.8 Request for Data by Unauthorized Persons outside the 2021 PHC Management

Since you are under oath not to provide any information to any unauthorized persons in performing your duties as a Data Quality Monitor, your response to anyone outside the four (4) member District Data Quality Management Team (DDQMT) and the 2021PHC Management Team requesting for information on the ongoing Census may be as follows:

“The data you are inquiring about can only be obtained from the Government Statistician (GS) who is required by law to disseminate official statistics. The field officers sync data to Headquarters where the data is downloaded and processed by a team of experts. Only inconsistencies and error messages which have to be corrected are processed here and communicated to the field officers. I am not permitted to share that with anybody except authorized persons who are part of the Census administration. We implore you to deal directly with the GS, or 2021PHC Management Team at the Head office who are in the best position to provide you with the information you seek. Thank You”.

YOU COUNT, GET COUNTED

CHAPTER TWO: DATA QUALITY MONITORING & QUERYING IN CSPRO/STATA

DQMs will run pre-programmed codes to check the status of work and inconsistencies in the data for their districts and regions and also expected to add further checks that will be communicated. Below are steps that Data Quality Monitors would undertake to enable them access the 2021PHC data.

2.1 Creating folder Icons to Start Work

2.1.1 Instructions to access folders for the 2021PHC Data Quality Monitoring exercise

- a) Create a folder “C:\2021PHC” to store your files
- b) An email containing a zipped file (2021PHC) will be sent to each Data Quality Monitor
- c) Download the zipped file attached in the email.
- d) Extract the folders from the downloaded zipped file.
- e) Copy the five extracted folders; “Data”, “DMS”, “Listing”, “Main” and “Other” into your Drive C (“C:\2021PHC”) folder and ensure that you have the following paths for your subfolders
 - C:\2021PHC\Data
 - C:\2021PHC\DMS
 - C:\2021PHC\Listing
 - C:\2021PHC>Main
 - C:\2021PHC\Other

In the C:\2021PHC\Other folder, create the following sub-folders;

- **Main_lst** – contains backup of outputs from CSPro main error report files
- **List_lst** - contains backup of outputs from CSPro listing error report files
- **Monitor report**- contains backups of all monitor report files with dates
- **Standard report** - contains daily reports; clearance report and highlights (Highlights is for **RDQMs only**)
- **Backup** - contains all data backups (**For RQDMs only**)

2.1.2 Creating shortcuts to working folders

- a) Open “C:\2021PHC\DMS” folder
- b) Find the **DMHome** pff icon in the “C:\2021PHC\DMS\DMHome” folder.
- c) Create a shortcut of this icon to the desktop.



2.2 Accessing Data using CSPro at all levels

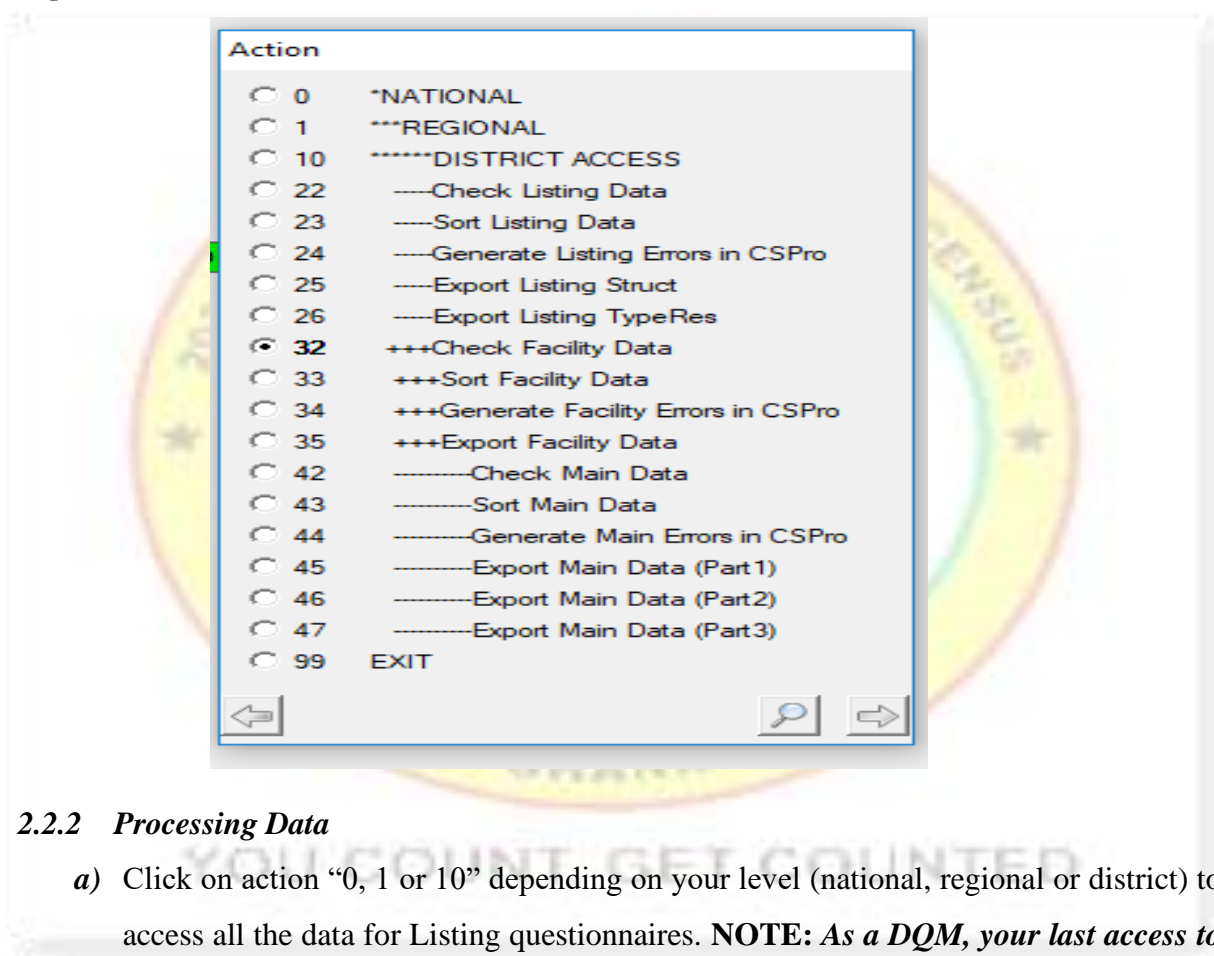
Double click on *DMHome.pff* file's shortcut on the desktop to launch the CSPro application.

Note that an ID would be required for Login before the actions can be performed.

2.2.1 Login

- a) Login with your credentials i.e.; 7-digit code
- b) An action menu pops up to perform the required actions as shown in the figure :

Figure 2.1: DM- Home Action Menu



2.2.2 Processing Data

- a) Click on action “0, 1 or 10” depending on your level (national, regional or district) to access all the data for Listing questionnaires. **NOTE: As a DQM, your last access to the data should be around midnight to ensure you work with an up-to-date file which contains all synced data for that particular day.**
- b) Run actions “22” through to “26” for Listing reports
- c) Run actions “32” through to “35” to generate a report for facilities in EAs/Localities.
- d) Run actions “42” through to “47” for Main data reports.
- e) Run action “99” to exit *DMHome* application

NOTE: Actions “42” through to “45” should be performed after CENSUS NIGHT when the main enumeration begins.

2.2.3 *Functions of each Action on Menu*

0-National: This action allows NDQMs to access national data.

1-Regional: This general process allows RDQMs to access regional data.

10-District Access: This general process allows DDQMs to access district data.

22-Check Listing Data: This action runs a batch application to ensure that all the listed structures are uniquely identified. It also ensures that the file is prepared for the sorting action. Data Quality Monitors are expected to check the “LST” output file and take note of all programmed error messages such as duplicate cases, wrong dates, training data cases and type of residence records being deleted.

23-Sort Listing Data: This procedure ensures that all the synced listing data from the different regions, districts, Sub-districts, Enumeration Areas (EAs) are properly organized in ascending order for further processing. The sort “LST” file reports duplicate cases (uid) as well as total number of questionnaires and records processed. The keys for sorting listing data are; Enumeration area code (EA code) and structure number.

24-Generate Listing Errors in CSPro: This action requires the sorted file (from action 23) as input to generate the error messages. The program compares the geographical codes from the field to the lookup file from GIS and reports error messages in case there is a mismatch. In addition, the user-defined error messages are also reported by EA code, SA, and Interviewer ID on a case-by-case basis. E.g., in the listing data a case is a structure in a particular EA. Hence the error messages are reported for each structure, where applicable. These error messages are to be shared with DDQMT, other field monitors, supervisors and enumerators as feedback. The enumerators are expected to use this feedback to correct the errors and re-sync the data to HQ.

25-26-Export Listing Data: This procedure exports the listing data from CSPro to STATA for further processing and reporting.

32-Checking Facilities Data: This procedure runs a batch application to ensure that all the facilities in the EAs and type 3 localities are uniquely identified. Also, it prepares the data file for the sorting action. Data Monitors are expected to check the output file from these actions and take note of any messages.

33-Sort Facilities Data: This procedure ensures that all the synced data on facilities from the different regions, districts, Sub-districts, Supervision Areas (SAs) and Enumeration Area (EAs) are properly organized in ascending order for further processing.

34-Generate Facilities errors in CSPro: This action requires the sorted file (from action 33) as input to generate the error messages. The program compares the geographical codes from the field to the lookup file from GIS and reports error messages in case there is a mismatch. In addition, the user-defined error messages are also reported by EA code, SA, and Interviewer ID on a case-by-case basis. E.g., in the Facilities data a case is a facility in a particular EA or locality for type three EAs. Hence the error messages are reported for each facility, where applicable. These error messages are to be shared with DDQMT, other field monitors, supervisors and enumerators as feedback to improve upon the work. The enumerators are to use this feedback to correct the errors and re-sync the data.

35-Export Facilities: This procedure exports the facilities data from CSPro to STATA for further processing and reporting.

42-Check Main Data: This action is valid during the actual enumeration, i.e., after census night. The procedure runs a batch application to ensure that the enumerated type of residence is uniquely identified. It also ensures that the data file is prepared for the sort action. Data Quality Monitors are expected to check the output file and take note of all programmed error messages such as duplicate cases, wrong dates, training data cases and wrong persons records.

43-Sort Main Data: This procedure ensures that all the synced data from the main questionnaires from the different regions, districts, sub-districts, Enumeration Areas (EAs) and structures and serial number of the type of residences are properly organized in ascending order for further processing.

44-Generate Main errors in CSPro: This action runs a batch application using the sorted file (from action 43) as input to generate error messages. The program compares the geographical codes from the field to the lookup file from GIS and reports error messages in case there is a mismatch. Also, the user-defined error messages are also reported by SA, EA and Interviewer ID on a case-by-case basis. E.g., in the main data a case is a type of residence/questionnaire in a particular structure in an EA. Hence the

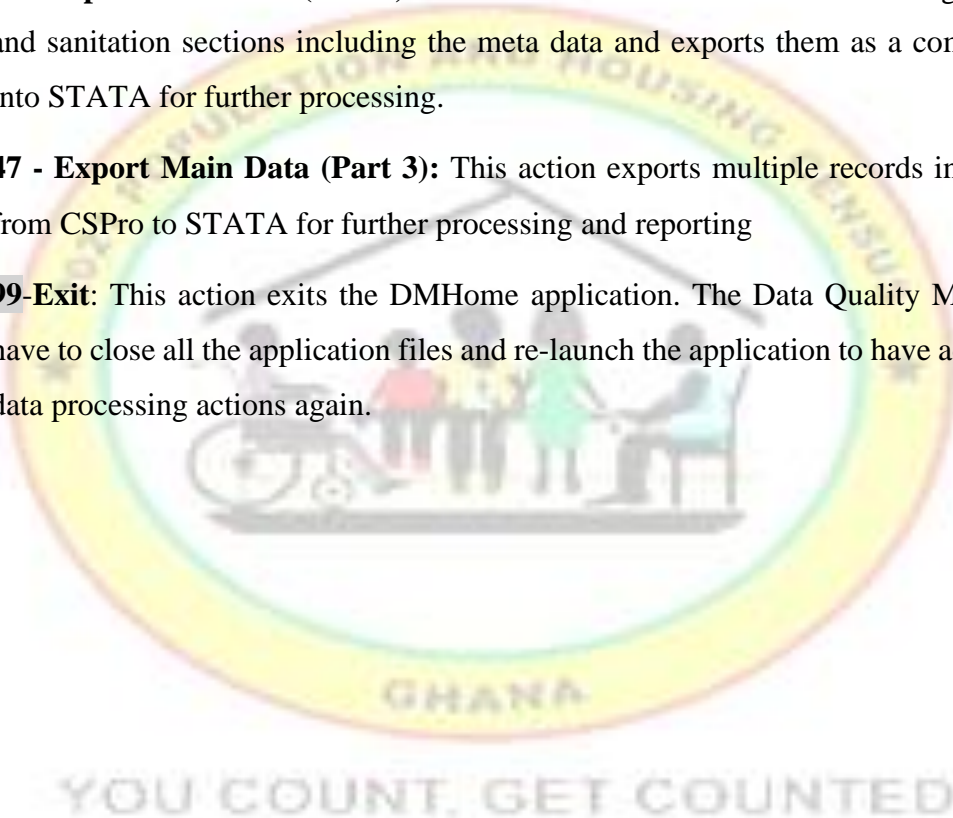
error messages are reported for each variable in the questionnaire, where applicable. The error messages are further clarified by the question number and member ID. The Data Quality Monitor is supposed to share these error messages with DDQMT, other field monitors, supervisors and enumerators as feedback to improve the work and facilitate coaching. The enumerators are expected to use this feedback to correct the errors and re-sync the data to HQ.

45--Export Main Data (Part 1): This action exports each section in main data from CSPro to STATA independently for further processing and reporting.

46 - Export Main Data (Part 2): This action selects variables from housing conditions and sanitation sections including the meta data and exports them as a composite file into STATA for further processing.

47 - Export Main Data (Part 3): This action exports multiple records in main data from CSPro to STATA for further processing and reporting

99-Exit: This action exits the DMHome application. The Data Quality Monitor will have to close all the application files and re-launch the application to have access to the data processing actions again.



CHAPTER THREE: ACTIVITIES OF DQM DURING THE LISTING STAGE

Final files to be used for the 2021 PHC exercise will be sent by NDMs through the official mail to all DQMs. Note that you are to work with files sent from HQ and not from other colleagues in the region. Once the zipped file has been downloaded from the initial email and its content extracted and put in the respective folders, the DQM is ready to begin work. The DQM is also expected to ensure that he/she has both STATA and CSPro programs installed.

3.1 Activities to perform during Listing

The listing exercise begins the census process, this is when enumerators collect data on structures and type of population present in a particular district. As soon as the listing exercise for the 2021 PHC begins, the DQM is expected to run thorough checks on the data collected and synced to ensure that the data being collected is reliable and of good quality.

3.1.1 Step 1 – Accessing and Processing Data in CSPro

Run actions 1 or 10 to access data and run reports Action 22 generates an Output file called “JUSTLIST.LST” in C:\2021PHC\Listing which will contain error messages such as duplicate cases, wrong dates, training data cases and type of residence records being deleted if there are any. The DQM is supposed to check for training data or any invalid dates and duplicates using this file, note the device IDs from which these errors are coming and prompt IT for immediate action to be taken. Below is an e.g., of the Justlist file in a case where there is a duplicate or training data.

Figure 3.1: Justlist file

Line	Freq	Pct.	Message text
11	35	-	TypeRes record deleted for Line %d
24	1	-	Duplicate structure 51 deleted from EA=1204100108
29	1	-	Duplicate structure 73 deleted from EA=1204100110
35	1	-	Duplicate structure 42 deleted from EA=1207200188
40	1	-	Duplicate structure 31 deleted from EA=1207200180
45	1	-	Duplicate structure 6 deleted from EA=1207200230
50	1	-	Duplicate structure 1 deleted from EA=1208100050
55	11	-	Duplicate SA 36 for Sadia Alhassan deleted from Dist=1203
84	15	-	Duplicate structure=%d deleted from EA=1214200052
98	7	-	Duplicate structure=%d deleted from EA=1210302210
105	104	-	EAcodes 1202100161 corrected to 1202100165
110	70	-	EAcodes 1202100165 corrected to 1202100161
134	1	-	Duplicate structure=%d deleted from 1206100020
142	2	-	Duplicate structure=%d deleted from EA=1210302074
161	14	-	Duplicate structure=%d deleted from EA=1210302077
212	45	-	Duplicate structure=%d deleted from EA=1210302005
224	5	-	Duplicate structure=%d deleted from EA=1211200128
241	11	-	Duplicate structure=%d deleted from EA=1210302069
261	14	-	Duplicate structure=%d deleted from EA=1210302072
280	5	-	Duplicate structure=%d deleted from EA=%d
334	218	-	Training data with Date=%d deleted: Northern Region, SA=%d, InterviewerID=%d & C
349	3	-	Other invalid data with Date=%d deleted: Northern Region, SA=%d, InterviewerID=%d
356	78	-	EAREdone: Old data-before 3rd August for EA=%d deleted

- Action 24 generates an output file called “LISTERRORS.LST”.
- The *CSPRO ListError File* gives the detailed error messages associated with each structure listed. Data Quality Monitors are to study the errors and identify the specific structure numbers and EAs with these errors.
- Check that the date and time in the ListErrors file corresponds to the same period the action was run.
- Then, DDQMs are to share this information with DDQMT and together, discuss these errors and communicate same to field supervisors and enumerators for rectification.
- In the situation where an enumerator/supervisor wants to find the particular case where an error message is coming from, the *CSPRO ListError File* is the one to use.
- The DQM should follow these steps in order to identify a specific listing error message:

Look for the errors with the highest frequency at the **User Unnumbered Messages** section. Below is a sample;

Figure 3.2: User Unnumbered Messages

User unnumbered messages:

Line	Freq	Pct.	Message text	Denom
38	76	0.0	SA does not match this EA in Lookup file	361727
43	0	0.0	EAType does not match this EA in Lookup file	361727
57	6	0.0	ID02(1): Structure=%d is missing for InterviewerID=%d in SA=%d	361727
65	1030	0.3	ID02(2): Structure=%d is missing for InterviewerID=%d in SA=%d	361727
80	10	0.0	LS01(1): Structure number is not same as ID structure=%d for InterviewerID=%d in SA=%d	361727
99	0	0.0	LS02(1): Structure Address is missing, InterviewerID=%d in SA=%d	361727
106	16	0.0	LS03(1): Status of GPS reading should be 1 or 2, InterviewerID=%d in SA=%d	361727
131	0	0.0	LS03(2): GPS Coordinates not taken, InterviewerID=%d in SA=%d	361727
155	5	0.0	LS04(1): Filter for GPDA is missing, InterviewerID=%d in SA=%d	361727
180	0	0.0	LS04(2): Ghana Post Address not recorded, InterviewerID=%d in SA=%d	361727
187	56	0.0	LS05(1): Level of completion is not valid for Struct=%d & InterviewerID=%d in SA=%d	361727
194	40	0.0	LS06(1): Type of structure and LS05 skip not followed, InterviewerID=%d in SA=%d	361727
198	0	0.0	LS06(2): Type of structure=%d is not valid for LS05=%d, InterviewerID=%d in SA=%d	361727
205	40	0.0	LS06o(1): Type of structure-OS %s is not valid for LS05=%d and LS06=%d, InterviewerID	361727
209	0	0.0	LS06o(2): Type of structure-OS cannot be blank for LS06=13, InterviewerID=%d in SA=%d	361727
219	40	0.0	LS07(1): Use of structure %s is not valid if LS05=%d, InterviewerID=%d in SA=%d	361727
223	0	0.0	LS07(2): Use of structure cannot be blank if LS05=%d, InterviewerID=%d in SA=%d	361727
352	2	0.0	LS07o(1): Type of structure-OS=%s, LS07=%s and LS05=%d & InterviewerID=%d in SA=%d	361727
356	0	0.0	LS07o(2): Type of structure-OS cannot be blank, use of structure has 'Other', Intervie	361727
363	0	0.0	LS07A(1): Use of Education/Health structure is not expected for LS05=%d and LS07=%s, I	361727
376	0	0.0	LS07A(2): Use of Education/Health structure is not expected for LS05=%d and LS07=%s, I	361727
390	87	0.0	LS07A(3): Use of Education/Health structure cannot be blank if LS05=%d and LS07=%s, In	361727
403	0	0.0	LS07A(4): Use of Education structure cannot exceed option 28 if LS05=%d and LS07=%s, I	361727
470	0	0.0	LS07Ao(1): Use of Education/Health structure-OS=%s, LS07A=%s and LS05=%d & Interviewer	361727
474	2	0.0	LS07Ao(2): Use of Education/Health structure-OS cannot be blank, LS07A=%s, Interviewer	361727
480	0	0.0	LS08(1): Functional toilet use %d but LS05=%d, InterviewerID=%d in SA=%d	361727
484	0	0.0	LS08(2): Functional toilet use %d but LS05=%d, InterviewerID=%d in SA=%d	361727
491	40	0.0	LS09(1): Type of functional toilet %s but LS08=%d, InterviewerID=%d in SA=%d	361727

- Copy the specific error message from the Message Text
- Ctrl+F to open the search pop up window or click on the Find icon on the Tool



bar

- c. Paste error message in search space
- d. Click on any space above the search pop up window or use the Previous and Above options to help navigate where the search should be
- e. Click find on the search pop up window
- f. Save this file after each run in " C:\2021PHC\Other" with the required format e.g., *ListErrors-20210413.lst*.

Figure 3.3: Sample Error Messages

```




11
12 CSPRO Process Summary
13 +-----+
14 | 437174 Records Read ( 100% of input file) |
15 |      0 Ignored (      0 unknown,      0 erased) |
16 |   7756 Messages (   7756 U,      0 W,      0 E) |
17 +-----+
18 | Level | Input Case | Bad Struct | Level Post |
19 +-----+
20 |      1 |      55821 |          0 |          55821 |
21 +-----+
22
23
24
25
26 Process Messages
27
28 *** Case [12011000030001] has 1 messages (0 E / 0 W / 1U)
29   U  -813 LISTDETAIL(5): Vacant Housing Unit expected in L12B, LS07=01
30     , LS10=4, Interviewer=120110000301 in SA=1
31
32 *** Case [12011000050028] has 2 messages (0 E / 0 W / 2U)
33   U  -1074 LEND: Questionnaire status can only be 1 or 2, InterviewerID=120110000503 in SA=1
34   U  -1078 LEND: Incomplete Questionnaire, InterviewerID=120110000503 in SA=1
35
36 *** Case [12011000050029] has 2 messages (0 E / 0 W / 2U)
37   U  -1074 LEND: Questionnaire status can only be 1 or 2, InterviewerID=120110000503 in SA=1
38   U  -1078 LEND: Incomplete Questionnaire, InterviewerID=120110000503 in SA=1

```

- a. The DQM is to study and understand each error message and share them with other members of the DDQMT as well as field supervisors and enumerators for correction and resyncing. After which the DQM is expected to go through actions 10 to 24 to ensure that the rectified errors have reflected in the cloud.
- b. Next, run actions 25 and 26 to export the listing data from CPro into STATA for further processing.

3.1.2 Step 2 – Analyzing Listing Data Using STATA and Excel

- Launching the do-file in STATA
 - a. Launch STATA by double clicking on STATA icon on desktop to launch it

- b. Click on the do-file icon or Ctrl + 9 to open a do-file editor. 
 - c. Click on the “Open” icon  in the do-file editor.
 - d. A dialogue box pops up.
 - e. Locate “Listing” folder in C:\2021PHC
 - f. To start work as the DQM, open the “**BACKGROUND.DO**” file and follow the instructions to setup your district codes and to work with other files. This file is what the DQM will use for any other analysis/check that will be done on the data. Once it is given, it will not be updated from HQ.
 - g. Click the “execute” icon  to run the do-file.
 - h. The excel workbook is generated in C:\2021PHC\Listing with the name “**monitor_report_(district code)**”
 - i. The report must be backed up for reference (with dates); E.g. – **monitor_report_701_20210407** in the folder “C:\2021PHC\Other”.
- Once the monitor report excel sheet is generated, the DQM is expected to study it for inconsistencies. Below is a sample of the monitor report file;

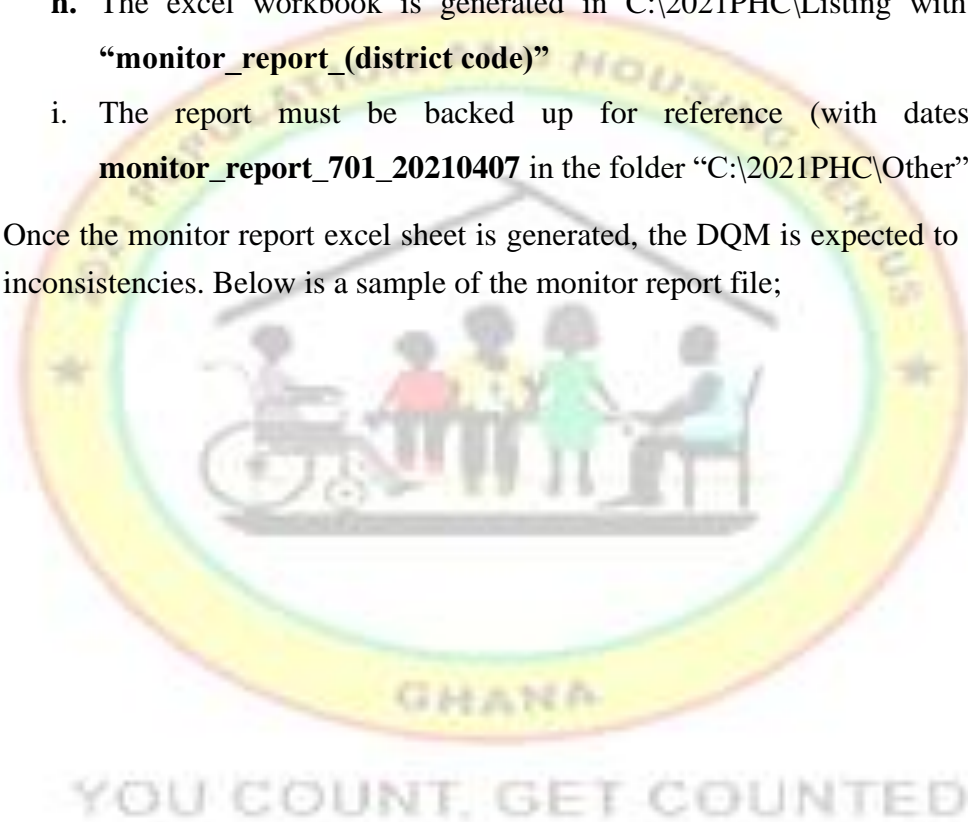


Figure 3.4: Sample Monitor Report File

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
	SNo	EA CODE	ENUMID	Enurname	Surname	L1_struct	L2_struct_wit hout_gps	L3a_Good_GP S_reading	L3b_GP-5692 0m	L2c_Bad_GPS 20mplus	L4_struct_wit hout_GPS	L5_1_Fully_co mpleted	L5_2_Comp_l roofed_unco mpleted	L5_3_Comp_l roofed_parrly_r oofed	L5_4_Roofing _level	L5_5_Lintel win_jmrov_	L5_6_Lintel without_inpr oof	L5_7_Windo w_level	L5_8_Concret e_pillar_level	L5_9_Concret e	L5_T_Total	L6_1_Single_b ld_detach	L6_2_Single storey
2	1	523100008	52310000801	TETTEH-YI	APEADU	375	0	375	0	0	329	258	73	7	8	2	6	8	10	3	375	229	1
3	1	523100009	52310000901	ASOMANII	APEADU	417	0	406	11	0	382	345	36	4	0	0	10	20	2	0	417	249	1
4	1	523100010	52310001001	ADOM	APEADU	234	0	234	0	0	171	205	14	2	0	0	0	12	1	0	234	140	0
5	1	523100011	52310001101	AMPONSA	APEADU	181	0	181	0	0	152	158	16	0	0	0	5	2	0	0	181	129	0
6	1	523100012	52310001201	OYIYE-ME	APEADU	174	0	173	1	0	129	132	11	0	3	3	8	17	0	0	174	111	1
7	1	523100013	52310001301	OKYERE	APEADU	295	0	293	2	0	242	263	8	0	0	2	5	10	7	0	295	194	0
8	1	523100014	52310001401	OBENG	APEADU	267	0	265	2	0	232	227	18	0	0	0	13	9	0	0	267	199	2
9	2	523100001	52310000101	ASARE	IBRAHIM	286	0	286	0	0	257	227	21	4	1	0	3	17	13	0	286	169	7
10	2	523100003	52310000301	BAWUSE	IBRAHIM	466	0	466	0	0	446	407	24	2	5	4	4	14	5	0	465	292	7
11	2	523100004	52310000401	OWIREDU	IBRAHIM	205	0	205	0	0	194	194	1	1	0	0	0	8	0	1	205	133	5
12	2	523100005	52310000501	DOCREY	IBRAHIM	244	0	244	0	0	234	136	62	11	0	0	17	10	8	0	244	144	10
13	2	523100019	52310001901	AGAMAH	OSEI MAN	127	0	127	0	0	102	125	1	0	0	0	1	0	0	0	127	69	2
14	3	523100020	52310002001	YEBOAH	OSEI MAN	98	0	98	0	0	69	89	3	0	0	0	1	4	0	1	98	59	1
15	3	523100021	52310002101	BOATENG	OSEI MAN	121	0	121	0	0	74	115	1	0	2	0	0	0	2	1	121	84	3
16	3	523100022	52310002201	AMEDEKA	OSEI MAN	355	0	353	2	0	340	299	26	2	3	0	16	3	5	1	355	238	0
17	3	523100023	52310002301	TWUM YE	OSEI MAN	310	0	308	2	0	276	263	12	9	0	0	12	12	2	0	310	145	2
18	3	523100024	52310002401	OKYERE	OSEI MAN	361	0	361	0	0	338	300	18	7	1	1	11	9	13	1	361	196	3
19	3	523100024	52310002402	OFOSSU AS	OSEI MAN	307	0	307	0	0	304	222	32	3	1	0	9	28	11	1	307	187	5
20	3	523100024	52310002403	BOATENG	OSEI MAN	60	0	60	0	0	48	42	10	1	0	0	2	4	1	0	60	43	1
21	3	523100024	52310002404	MARIAM	OSEI MAN	9	0	9	0	0	5	8	0	0	0	0	0	0	1	0	9	5	0
22	3	523100024	52310002405	YEBOAH	OSEI MAN	91	0	91	0	0	78	72	11	0	0	0	2	2	4	0	91	59	1
23	4	523100030	52310003001	DARKOA	ADJEI-AFR	122	0	122	0	0	114	112	4	0	0	0	2	2	2	0	122	77	0
24	4	523100031	52310003101	OSEI FRIM	ADJEI-AFR	348	0	348	0	0	305	296	32	5	0	0	11	1	3	0	348	255	0
25	4	523100032	52310003201	ENTI	ADJEI-AFR	310	0	310	0	0	261	260	18	7	0	0	5	13	7	0	310	202	1
26	4	523100033	52310003301	SEMAA	ADJEI-AFR	456	0	456	0	0	431	375	44	3	0	0	8	10	15	1	456	321	1

- The monitor report contains the following sheets;
 - a. Enum_stats
 - b. EA_stats
 - c. Struct-gap_file
 - d. listing_Dups_file
 - e. loc_with_nopop
 - f. badGPS_file
- Use these sheets to perform the DQM activities and others that may be recommended by the NDQMs/RDQMs during the exercise.
- Content of the EA_stats sheet;

Figure 3.5: EA_stats sheet

The screenshot shows an Excel spreadsheet with the following columns (A to U):

- regdist
- subdist
- Enumname
- Supname
- SAno
- EACODE
- EATYPE
- nlocexp
- nloc
- gps
- nos_duplicates
- BF
- L1_tot_struct
- BFdiff
- perBcov
- L3_struct_without_gps
- L3a_Good_GPS_reading
- L3b_GPS6to20m
- L3c_Bad_GPS20mplus
- L4_struct_without_gps
- L5_1

Key data points from the spreadsheet:

- Row 8: Enumname: AMPONSA, L1_tot_struct: 181, BF: 0, BFdiff: 181
- Row 9: Enumname: OYIYE-MEI, L1_tot_struct: 213, BF: 0, BFdiff: 213
- Row 10: Enumname: OKYERE, L1_tot_struct: 295, BF: 0, BFdiff: 295
- Row 11: Enumname: OBENG, L1_tot_struct: 267, BF: 0, BFdiff: 267
- Row 12: Enumname: ASARE, L1_tot_struct: 286, BF: 0, BFdiff: 286
- Row 13: Enumname: BAWUSE, L1_tot_struct: 480, BF: 0, BFdiff: 480
- Row 14: Enumname: OWIREDU, L1_tot_struct: 203, BF: 0, BFdiff: 203
- Row 15: Enumname: DOCREY, L1_tot_struct: 259, BF: 0, BFdiff: 259
- Row 16: Enumname: AGAMAH, L1_tot_struct: 127, BF: -11, BFdiff: 118
- Row 17: Enumname: YEBOAH, L1_tot_struct: 98, BF: 5, BFdiff: 93
- Row 18: Enumname: ROATFNG, L1_tot_struct: 121, BF: -18, BFdiff: 103

- Check from **L1_tot_struct** column that all **SAs/EAs** in the district assigned are represented with data in the report, if not, report missing SAs and EAs to DDQMT. **L1_tot_struct** gives a fair idea of the number of structures listed and synced.
- Check **nlocexp** (number of localities expected from GIS) and **nloc_in_data** (number of localities found in data). Take note, especially of Type3 EAs where the number varies from 1 to 60+ localities. We expect to see all these localities with information in the data.
- Check **BFdiff**, the difference between **L1_tot_struct** - total number of structures listed and **BF** - expected number of structures. **Positive value**

implies the enumerator/team might have covered all structures expected in the EA and added more; negative value implies the enumerator/team might have outstanding structures to cover given the structure count (BF) from GIS.

- d. Check **PerBFcov** for the rate of coverage of structures for each EA to assess work progress.
- e. Check **L3_struct_without_gps** (number of structures without GPS coordinates). Only zeros (0) should be in this column. Any other value in this column means that the Enumerator is not taking the GPS for some structures and might have a problem. You should find out the reasons why GPS coordinates have not been taken. **This should be referred to the DIT and DFS immediately** for the problem to be resolved. Check also the content of **L3b_GPS6to20m, L3c_Bad_GPS20mplus-** GPS readings with accuracy above 6 meters, zero in this column is better.
- f. Critically examine the content of the following columns:
 - i. **L05 – Levels of Completion**
 - ii. **L06 – Types of Structure**
 - iii. **L12b_1_HH(Households) = Number of households**
 - iv. **L12b_2_Vacant = Number of vacant structures**
 - v. **Prop_vacant -Proportion of vacant structures**
 - vi. **L12b_98 = Number of Secondary housing unit for same household**
 - vii. **Prop_code98 -Proportion of Secondary housing unit for same household**
 - viii. **L12b_3_GQs - stable population**
 - ix. **L12b_4_GQs - unstable population**
 - x. **List_HHsize – Household Size from listing**
- g. Check **L16c_Totpop_HH** (total household population listed) for outliers
- h. Check **L16t_Totpop_GQ** (total group quarter population) for some basic schools with resident population.
- i. Analyze the above indicators with the DDQMT. Remember that, some of the cases would have to be confirmed/validated whether they are genuine

by visiting the specific structures/institutions for verification.

NOTE: In the case of group listing, the Enum_Stats file and EA_Stats file will not be the same. The Enum_Stats sheet will present statistics on work done by each enumerator in the EA.

- **Content of Struct-gap_file;** The **Struct-gap_file** contains the number of gaps in structures for each EA. It presents the minimum and maximum structure numbers of the data reflected in the HQ database. It also gives the total count of structures listed. **The minimum structure number MUST start from 1 for each EA.** The difference between the maximum structure number and the actual structures listed gives us the total number of structures missing (structure gaps) in that EA's data file.



Figure 3.6: Struct-gaps_file sheet

regdist	subdist	Enumname	Surname	SAho	excode	minimum_str_no	maximum_str_no	strucs_listed	Bappaiff
523	523100	TETTEH-YI	APEADU	1	523100008	1	375	375	0
523	523100	ASOMANI	APEADU	1	523100009	1	417	417	0
523	523100	ADOM	APEADU	1	523100010	1	234	234	0
523	523100	AMPONSA	APEADU	1	523100011	1	181	181	0
523	523100	OYIYE-MEI	APEADU	1	523100012	1	218	213	5
523	523100	OKYERE	APEADU	1	523100013	1	295	295	0
523	523100	OBENG	APEADU	1	523100014	1	267	267	0
523	523100	ASARE	IBRAHIM	2	523100001	1	286	286	0
523	523100	BAWUSE	IBRAHIM	2	523100003	1	465	465	0
523	523100	OWIREDU	IBRAHIM	2	523100004	1	203	203	0
523	523100	DOCREY	IBRAHIM	2	523100005	1	244	244	0
523	523100	AGAMAH	OSEI MAN	3	523100019	1	127	127	0
523	523100	YEBOAH	OSEI MAN	3	523100020	1	98	98	0
523	523100	BOATENG	OSEI MAN	3	523100021	1	121	121	0
523	523100	AMEDEKA	OSEI MAN	3	523100022	1	335	333	2
523	523100	TWUM YEI	OSEI MAN	3	523100023	1	310	310	0
523	523100	OKYERE	OSEI MAN	3	523100024	1	439	396	43

- **Content of dups_file;** The **dups_file** presents cases that are duplicating, the enumerators with these duplicates and the specific structure numbers. The file also shows the date and time when the data was collected.

Figure 3.7: dups_file sheet

regdist	subdist	Enumname	Surname	SAho	excode	struct	dup_strs	start_date	start_time	end_date	end_time
523	523100	APUWENI	ARHIMAH	13	523100038	301	1	11012021	1449	11012021	1453
523	523100	YEBOAH A	ARHIMAH	13	523100038	301	1	11012021	1456	11012021	1459
523	523100	YEBOAH A	ARHIMAH	13	523100038	339	1	11012021	1707	11012021	1716
523	523100	APUWENI	ARHIMAH	13	523100038	339	1	11012021	1734	11012021	1747
523	523100	APUWENI	ARHIMAH	13	523100038	349	1	11012021	945	11012021	952
523	523100	YEBOAH A	ARHIMAH	13	523100038	349	1	11012021	1820	11012021	1827
523	523100	ASAMOAH	ARHIMAH	13	523100039	73	1	6012021	1045	6012021	1048
523	523100	AMAADI	ARHIMAH	13	523100039	73	1	11012021	1538	11012021	1548
523	523100				9999999999						

- **Content of loc_with_nopop;** The **loc_with_nopop** report shows localities in Type 3 EAs with no population. In such a situation, the enumerator might have selected

occupied or homeless at type of residence but there was no population in the locality or introduced new localities in nloc and must be corrected to clear such cases. DDQMT should critically analyze the structure details and confirm that the localities exist. This will help GSS report on all localities existing in Ghana during the 2021 Population and Housing Census.

Figure 3.8: loc_with_nopop sheet

	subdist	SAno	encode	id08	id09	id09a	structures	h12b_1	h12b_2	h12b_3	h12b_4	h12b_T	h16a	h16b	h16c
3	523100	6	5.23E+08	3		3 (MP) 83 (MP) 83	6	0	5	0	0	5	0	0	0
4	523100	6	5.23E+08	3		4 MILE 83 MILE 83	14	0	13	0	0	13	0	0	0
5	523100	15	5.23E+08	3		4 NYAVOR AKURA NYAVOR AKURA	3	0	2	0	0	2	0	0	0
6	523100	9	5.23E+08	3		1001 ASEDJA PROPER NO. 1 YILO	1	0	1	0	0	1	0	0	0
7	523100	10	5.23E+08	3		4 YILO YILO	13	0	9	0	0	9	0	0	0
8	523100	11	5.23E+08	3		1008 ABREREWANE NO. 2 ABREREWANE 1	1	0	1	0	0	1	0	0	0
9	523100		1E+10	3		1 There is No locality with zero population, ignore this message line		0	0	0	0		0	0	0

- **Content of badGPS_file;** The **badGPS_file** shows structures with GPS accuracy readings greater than 20 meters. Enumerators must be encouraged to retake them and be retrained if possible, on how to position themselves to get accurate GPS readings.

Figure 3.9: Sample bad GPS sheet

	regidst	subidst	Enumname	Surname	SAno	eacode	struct	no_job_taken	GP-Serialing_nobacepable	start_date	start_time	end_date	end_time
3	523	523100	OWUSU	OWUSU	1	523100008	3			18012021	1051	18012021	1106
4	523	523100				9999999999							

3.1.3 Further Analysis on Listing Data in STATA

After studying the sheets in the monitor report, the DQM is expected to perform further analysis to further query the data for inconsistencies.

Check Inconsistencies

- a. Check for Type 3 EAs where nloc is far more than nloc_exp in STATA.
- b. `br if eacode == "eacode number" & nloc > nloc_exp`
- c. Check id10 and id10a to see the different locality codes and names present in the data to identify invalid/unexpected localities. Note that some of the new localities may be part of the existing localities in the frame already and therefore has to be corrected. Discuss this with the DDQMT
- d. Crosstab to check for inconsistencies between;
 - i. ls05 and ls06. For example; when you tab ls05 vs ls06, do not expect to find value other than “missing” in ls06 if the options in ls05 are from 6 to 9.
 - ii. ls05 and ls08
 - iii. ls06 and ls08
 - iv. ls10 and ls11
- e. Consistency check for the option ‘other’ in some questions (ls07; ls06) to check the content of what is specified. Example;
- f. `tab ls07o`
- g. `browse eacode struct ls07o if ls07o=="KITCHEN"`. This provides a list of EAs and structure numbers where enumerators keyed in ‘KITCHEN’ as an option ‘other’ even though “KITCHEN” is an option in ls07. (Thus, checking for

options in ls07 which have been specified in “other specify”)

- h. Check the highest number of uses in LS07 selected for a particular structure and advise (flag for follow up if more than 10 uses)
- Do these further analyses as a report for the district as the DDQMT may be required to present these findings to any 2021 PHC Independent monitor who may visit the district/region.

Other Checks required for reporting.

The DQM is expected to perform these checks, compile a report and share that with the DQM he/she works under.

- a. Plot a scatter plot of Household Size per EA and discuss with DDQMT.
- b. Highlight EAcodes and Household Size and copy to a new sheet.
- c. On the new sheet, go to Insert on the toolbar and select “Scatter” from the Recommended Chart.
- d. Add the trendline from the Chart elements to detect anomalies
- e. Generate the sex ratio for each EA and check for abnormalities.
- f. What are the uses of the structures in your district?
 - i. Which ones are dominant?
 - ii. Check use of the structures with ls05
 - iii. Check use of the structures with ls06
- g. What are the different types of toilet facilities in your district?
 - i. Which ones are dominant?
 - ii. Check ls05 and type of toilet facility
 - iii. Check ls06 and type of toilet facility
- h. What type of population are in the structures listed?
 - i. Check that L12b is consistent with the type of population
 - j. Check the number of vacant units
 - k. Discuss the type of population in your district
 - l. What are the types of questionnaires expected at the enumeration phase? Discuss.

3.2 Step 3 - Workload Management Strategy for 2021 PHC

3.2.1 Background

Workload Management for the 2021 PHC comprises the distribution of work among field officers and monitoring their utilization over time. The strategy is to effectively (re)distribute work among Enumerators, not only to reduce burnout for stressed field officers, but to provide real-time insight into the workload of each EA and how an Enumerator can reasonably complete it within the allotted time frame. Thus, workload management is implemented to ensure equitable distribution of work among field personnel in order to guarantee complete coverage within the allotted period; i.e., ten (10) days for listing and fourteen (14) days for enumeration. The strategy is an operational and monitoring tool to be implemented by members of the DDQMT; regularly accessing and tracking the work done by each Enumerator on a daily basis to enable them manage the workload in the district moving forward. The Workload management is in two forms;

- the Pre-Listing strategy
- the Enumeration strategy (E-strategy).

The **Pre-Listing strategy** comprises indicators provided by GIS unit of the Service to determine on the average, the workload for the district. Using such indicators as the **Building count, accessibility score among others**, the average number of listing forms to be administered a day is predicted. The Pre-Listing strategy is to inform members of the DDQMT on the required number of Enumerators for the listing exercise in a particular district; and also Census Management on the number of Enumerators to train and select for the district.

The E-strategy comprises indicators from the Listing form and the GIS team used to predict the average questionnaires an Enumerator is likely to administer within the specified period (14-day) and also inform members of the DDQMT on the average number of Enumerators to work in each EA during main enumeration based on the number of the different types of questionnaires expected.

3.2.2 Listing Strategy

The Listing strategy will be used to manage and monitor the Listing exercise given the stated period (10 days). The predicted model is used to determine the average number of listing forms to be administered in an EA based on such indicators as Building count, accessibility score, area covered by the EA and number of expected localities for the EA.

3.2.3 E-Strategy

The enumeration strategy “E-strategy” will be used to manage the number of PHC1A questionnaires to be completed by each Enumerator during the enumeration phase within 14 days. The predicted model is used to determine how factors such as EA size and compactness (indicated by the Polsby Popper score of difficulty), number of households and, total number of structures from listing of structures, and area covered by the EA influence the average number of questionnaires to be completed in a day.

3.2.4 Operationalization of the Strategy

The models are preprogrammed as part of the Data Quality Monitor (DQM) programs and reports are generated for discussion and further implementation by DDQMT. The strategy will be monitored by tracking work progress using the daily report generated by the DQM (EA_stat report for Listing and main dashboard report for main enumeration).

3.2.5 Sensitizing Field Officers

During the training of field officers, there will be presentation on the activities of the DDQMT, where workload management will be discussed. Supervisors of the various teams will also meet after training and at the end of the Listing to discuss the implementation modalities of the workload. Notably, the discussion will focus on how to reassign Enumerators with lower workload within and across Supervisory Areas to assist Enumerators with heavy workload, taking into account the expected number of days to complete the Listing and enumeration exercise. It will also focus on the creation of assignment for EAs where more than one (1) enumerator is likely to work.

3.2.6 Guidelines for Reassignment and Support Creation

For EAs that require more than one Enumerator;

- Use standby enumerators to assist work in such EAs from the start of enumeration and Listing using the proposed intervention from the workload management strategy.
- Develop a strategy on who does chalking and who start with the listing of structures for the Listing exercise
- Correctly split assignment among Enumerators who will begin listing and enumeration Before an Enumerator will finally be reassigned to support work in another EA;
- Check that all population types identified during Listing in their EA have been correctly enumerated with all errors and inconsistencies identified resolved (to prevent Enumerators from repeating errors)- They must be Cleared by the District Data Quality Monitor.
- Consider the proximity of the EA that needs support relative to the EA of the Enumerator to be reassigned - Enumerators from nearby SAs who have completed their work will be reassigned to SAs that are close by to avoid incurring cost on travel and transport and accommodation
- Ensure work in the SA from which the Enumerator originally belongs is completed. Thus, Enumerator should support EAs within their SAs first; if the number of Enumerators to be reassigned exceeds the number who will need support, then Enumerators should be reassigned to EAs in other SAs that need support

- Ensure that assignment is split correctly for Support Enumerators. All Enumerators to be reassigned should use their Secondary Enumerator ID (supporting ID) before an Enumerator will receive support;
- Check the number of questionnaires they have administered, and adjust for expected number of questionnaires to be administered per day.

Enumerators should be explicitly informed by the DDQMT on the specific EAs that they will be reassigned to assist.



CHAPTER FOUR: DURING ENUMERATION

4.1 Procedure after Data Access – Main Enumeration

When main enumeration commences after the census night, the DQM is expected to run all procedures for the listing exercise and proceed with the activities underlined, i.e., from the DMHome application run actions 10 to 35 and proceed to run actions 42 to 47 and 99.

4.1.1 Step 1 – Accessing and Data validation in CSPRO

- **Action 42** generates the Justmain file which has information (user defined error messages) on duplicate cases, wrong dates, training data cases and type of residence records. The Action 42 also ensures that the data file is prepared for the sort action.
- **Action 43** sort the main questionnaires from the different regions, districts, sub-districts, Enumeration Areas (EAs) and structures and serial number of the type of residences in ascending order for further processing.
- Action 44 generates a file called **2021PHCME** which is also supposed to be studied and errors communicated to other members of the DDQMT as well as enumerators and supervisors for rectification and resync. Note that the same procedure was performed during the listing stage. Below is a sample of main error messages;

Figure 4.1: Sample Main Error Messages

Process Messages

```
*** Case [02051000010005011] has 4 messages (0 E / 0 W / 4U)
U   -264 This Questionnaire should be finalised, Enumerator=20510000102 in
      SA=1
U   -268 Incomplete Questionnaire, Enumerator=20510000102 in SA=1
U   -126 A07(3): Housing record expected for OCCUPIED/VACANT UNIT should
      have, Enumerator=20510000102 in SA=1
U   -136 No Person record(s) for RESTYPE=1, Enumerator=20510000102 in SA=1

*** Case [02051000010056001] has 3 messages (0 E / 0 W / 3U)
U   -264 This Questionnaire should be finalised, Enumerator=20510000101 in
      SA=1
U   -268 Incomplete Questionnaire, Enumerator=20510000101 in SA=1
U   -136 No Person record(s) for RESTYPE=11, Enumerator=20510000101 in SA=1

*** Case [02051000019052012] has 4 messages (0 E / 0 W / 4U)
U   -1089 P19B(2)-ICT ownership, is not valid for Member=1,
      Enumerator=20510000102 in SA=1
U   -1435 PDELETE: Confirm information for Member=1, Enumerator=20510000102
```

- After studying and communicating errors to the field officials concerned, run actions 45, 46, 47 to export main data from CSPro to STATA and action 99 to exit the DMHome application.

4.1.2 Step 2 – Analyzing Main Data using STATA and Excel

- Next, launch the designated dofile for main enumeration and run it to generate an updated excel workbook, “**monitor_report (district code)**”. The workbook will contain all the sheets from the listing stage as well as 5 additional sheets;
 - a. **qtype_duplicates**

- b. NoPop_on_Nights
 - c. Popstats_by_Enum
 - d. EA_Popstats
 - e. main_dashboard
- Backup “monitor_report (district code)” and 2021PHCME in the correct “other” sub folder using the appropriate format.

Figure 4.2: Sample Monitor Report Main

Srno	EA CODE	ENUMID	Enumname	Supname	L1_tot_struct	L2_struct_without_gps	L3a_Good_GPS_reading	L3b_GPSnot20m	L3c_Bad_GPS20mplus	L4_struct_without_GPSnot	L5_1_Fully_completed	L5_2_Comp_Lroofed_Lim	L5_3_Parly_roofed	L5_4_Roofing_level	L5_5_Lintel_with_improv	L5_6_Lintel_without_Lim	L5_7_Window_Level	L5_8_Concrete_pillar_Lex	L5_9_Concrete
1	52310000E	5.23E+10	TETTEH-YIK	APEADU	380	0	375	4	1	333	263	73	7	8	2	6	8	10	3
1	52310000E	5.23E+10	ASOMANIN	APEADU	418	0	406	12	0	383	346	36	4	0	0	10	20	2	0
1	52310001C	5.23E+10	ADOM	APEADU	234	0	234	0	0	171	205	14	2	0	0	0	12	1	0
1	523100011	5.23E+10	AMPONSAI	APEADU	181	0	181	0	0	152	158	16	0	0	0	5	2	0	0
1	523100012	5.23E+10	OYIYE-MEN	APEADU	174	0	173	1	0	129	132	11	0	3	3	8	17	0	0
1	523100013	5.23E+10	OKYERE	APEADU	295	0	293	2	0	242	263	8	0	0	2	5	10	7	0
1	523100014	5.23E+10	OBENG	APEADU	267	0	265	2	0	232	227	18	0	0	0	0	13	9	0
2	523100001	5.23E+10	ASARE	IBRAHIM	286	0	286	0	0	257	227	21	4	1	0	3	17	13	0
2	523100005	5.23E+10	BAWUSE	IBRAHIM	466	0	466	0	0	446	407	24	2	5	4	4	14	5	0
2	523100004	5.23E+10	OWIREDU	IBRAHIM	205	0	205	0	0	194	194	1	1	0	0	0	8	0	1
2	523100005	5.23E+10	DOCREY	IBRAHIM	244	0	244	0	0	234	136	62	11	0	0	17	10	8	0
3	523100015	5.23E+10	AGAMAH	OSEI MANU	127	0	127	0	0	102	125	1	0	0	0	1	0	0	0
3	52310002C	5.23E+10	YEBOAH	OSEI MANU	98	0	98	0	0	69	89	3	0	0	0	1	4	0	1
3	523100021	5.23E+10	BOATENG	OSEI MANU	121	0	121	0	0	74	115	1	0	2	0	0	0	2	1
3	523100022	5.23E+10	AMEDEKA	OSEI MANU	356	0	353	3	0	341	300	26	2	3	0	16	3	5	1
3	523100023	5.23E+10	TWUM YEB	OSEI MANU	310	0	308	2	0	276	263	12	9	0	0	12	12	2	0
3	523100024	5.23E+10	OKYERE	OSEI MANU	361	0	361	0	0	338	300	18	7	1	1	11	9	13	1
3	523100024	5.23E+10	OFOSU ASA	OSEI MANU	307	0	307	0	0	304	222	32	3	1	0	9	28	11	1
3	523100024	5.23E+10	BOATENG	OSEI MANU	60	0	60	0	0	48	42	10	1	0	0	2	4	1	0
3	523100024	5.23E+10	MARIAM	OSEI MANU	9	0	9	0	0	5	8	0	0	0	0	0	0	1	0
3	523100024	5.23E+10	YEBOAH	OSEI MANU	91	0	91	0	0	78	72	11	0	0	0	2	2	4	0
4	52310003C	5.23E+10	DARKOA	ADJEI-AFRA	122	0	122	0	0	114	112	4	0	0	0	2	2	2	0
4	523100031	5.23E+10	OSEI FRIMF	ADJEI-AFRA	348	0	348	0	0	305	296	32	5	0	0	11	1	3	0
4	523100032	5.23E+10	ENTI	ADJEI-AFRA	310	0	310	0	0	261	260	18	7	0	0	5	13	7	0
4	523100033	5.23E+10	SEMAA	ADJEI-AFRA	456	0	456	0	0	431	375	44	3	0	0	8	10	15	1
5	52310003E	5.23E+10	KOFI DARKI	AMPONSAI	122	0	122	0	0	88	101	5	2	0	0	2	6	6	0
5	523100037	5.23E+10	BOATENG C	AMPONSAI	225	0	224	1	0	159	172	15	9	1	0	11	15	2	0

4.1.3 Step3 - Checklist after generating the monitor report during Main Enumeration

1. Check the sheet **qtype_duplicates** which shows duplicates of each questionnaire type, the enumerators concerned and the date and time the cases were finalized (check the variable mend for completeness). Where the cases or questionnaires have been discarded and still shows up as duplicates, send the details in the agreed format to the RDQM to be programmed in CSPro justmain file and sent to you in the next updates.
2. Check the sheet with the name “**NoPop_on_Night**” to find the number of EAs with structures where usual members were neither absent nor present on census night. Discuss this information with the DDQMT for investigation and subsequent correction.
3. Next, check the sheet with name “**EA_Popstats**”. This sheet gives a summary of data from the enumeration phase for each EA. The work done by each enumerator is found in Popstats_by_Enum file.
4. Check the sheet with the name “**main_dashboard**”. On this sheet, you can find the **total number of questionnaires** to be administered from the listing exercise and what has been completed for each EA. The variables used to monitor the daily progress by EA are;

- “L12b_1_HHH” is total number of household questionnaires expected to be administered to occupied housing units and homeless households (LH12b=1;2) after listing
- “L12b_2_Vacant98” is total number of vacant housing units and Secondary housing unit for same household expected (LH12b=3;98)
- “L12b_3_GQs” is total number of Group Quarter stable questionnaires expected,
- “L12b_4_GQu” is total number of Group Quarter Unstable questionnaires expected.

Compare these variables “qHH”, “qvacant”, “qGQB”, “qCDE”, “qHGTotal” to those from point 4 above to monitor the progress for each EA where the variable:

- “qHH” is number of household questionnaires completed,
 - “qvacant” is number of vacant questionnaires completed,
 - “qGQB” is number of group quarter stable questionnaires completed,
 - “qCDE” is number of group quarter unstable questionnaires completed and
 - “qHGTotal” is the Total number of questionnaires completed.
5. Closely monitor “L12b_2_Vacant98” and “qvacant” daily to track the number of vacant structures because enumerators may tend to enumerate household as vacant to reduce their workload.
 6. Also compare the figures for, “list_hhsize” to “main_hhsize”. Graph the two to compare.
 7. Make a backup of the monitors report daily.
 8. Build the daily dashboard for enumeration from the report. At the end of the last data access, give the total questionnaire by type for each EA for the days of enumeration.

Figure 4.3: Main Dashboard

4.1.4 Step 4 - Guide to further analysis of population data

Check data consistency for the underlined variables. Some would have shown up in the error reports already:
 use "sec6_popvind.dta", clear

- **Relationship**

- browse SAno eacode struct hhnum pid rel p02** if rel==1 & p02<15 to check the ages of heads of households (persons who are household heads with ages below 15years)
- browse SAno eacode struct hhnum pid rel p02** if rel==2 & p02<15 to check the ages of spouses of households (persons below 15years who are the spouse)
- browse SAno eacode struct hhnum pid rel p02** if inlist (rel,4,5) & p02<15 to check the ages of parents and parent-in-law to the head of household.
- browse SAno eacode struct hhnum pid rel p10** if rel==2 & p10==9: to check spouses who are never married.

- **Literacy and education**

Browse SAno eacode struct hhnum pid p12b if p02<15 & in range(p12b,10,15)

For specific educational level;

- browse SAno eacode struct hhnum pid p02 p12b** if p02<15 & p12b==10
- browse SAno eacode struct hhnum pid p02 p12b** if p02<15 & p12b==11
- browse SAno eacode struct hhnum pid p02 p12b** if p02<15 & p12b==12
- browse SAno eacode struct hhnum pid p02 p12b** if p02<15 & p12b==13
- browse SAno eacode struct hhnum pid p02 p12b** if p02<15 & p12b==14
- browse SAno eacode struct hhnum pid p02 p12b** if p02<15 & p12b==15

- **Economic activity**

- browse SA_{no} eacode struct hhnum p12a p13e if p12a==1 & p13e==2:** check for persons who have never been to school but are not working because they are in full time education
- browse SA_{no} eacode struct hhnum p12a p13e if p12a==3 & p13e==2:** check for persons who are not working because they are in full-time education but have only attended school in the past
- browse SA_{no} eacode struct hhnum pid p02 p13e if p02>15 & p13e==7:** check for persons who are not working because they are ‘too young’
- browse SA_{no} eacode struct p02 p13e if p02<15 & p13e==6:** check for persons who are not working because they are ‘too old’
- browse SA_{no} eacode struct hhnum pid p13a p13c if p13a==3 & p13c==1 & p13c==2:** check for those who worked but did not receive pay/profit (under economic activity) but chose Wage/profit or paid apprentice for the type of work they engage in
- browse SA_{no} eacode struct hhnum pid p13a p13c if p13a==4 & inrange(p13c,1,5) & p13c!=. :** check for those engaged in own agricultural activity but worked other than for Own production/service work
- Check for a mismatch between purpose of production/service and occupation.
- browse SA_{no} eacode struct hhnum pid p14a:** Check for bad description of occupation (ISCO). Look out for one-word occupation descriptions such as teacher, farmer, etc. More words should be used in describing the occupation in detail; cocoa farmer, primary school teacher
- browse SA_{no} eacode struct hhnum pid p15a:** Check for bad description of industry (ISIC)
- browse SA_{no} eacode struct hhnum pid p13e p18* if p13e==4 & p18*==1:** *check for people without difficulties but are not working due to disability condition (replace * with a-f for the various categories of disability)*

- **Fertility**

- browse SA_{no} eacode struct hhnum pid p25t if p25t>3 & p25t!=.(missing),** this code will extract women who have more than 3 children as their last child(ren) born alive. This information should be shared for verification.
- Add on any that may come up.

- **Housing conditions**

- use "sec1_HS.dta", clear
 - browse SA_{no} eacode struct hhnum a07 h* if a07==1 & h*==. :** to ensure that the housing questions are answered if the type of residence is an occupied housing unit
 - browse SA_{no} eacode struct hhnum a07 h01 h02 h03 if a07==3 & h01==. | h02==. | h03==.:** to ensure that the housing questions are answered if the type of residence is a vacant housing unit

- c. **browse SAno eacode struct hhnum a07 h* if a07==98 & h*==. :** to ensure that the housing questions are answered if the type of residence is the same dwelling unit

Note: * implies that use the question number under h that applies

- **Sanitation**

- use "sec1_HS.dta", clear

browse SAno eacode struct if s03o==11, observe the answers to make sure that the answer specified is not one of the options in question s03. If an answer specified is part of the options in s03, prompt enumerators for rectification.

Compare type of structure from listing to type of dwelling

- merge 1:1 eacode struct using "C:\2021PHC\listing\struct_list.dta"
- drop if _merge==2
- a. **browse SAno eacode struct s03 ls09**
- b. **browse SAno eacode struct s05 ls09a**
- c. **browse SAno eacode struct s03 if ls08a==3:** check structures with no toilet facilities during listing
- d. **browse SAno eacode struct s03 if ls08==3 & s03!=12 & s03!=10 & s06==2**
- e. Add any other checks that may be needed.

In all these extracts of errors showing the **SA, EA, structure, household number and member**, send the information promptly to the teams/enumerators committing such errors. Follow daily to check that these anomalies have been affected in the data. Add these extracts to the "monitor_list_report_(district code)" and keep daily.

NOTE: You are to send the feedback from your further analysis report showing all the results of your checks in excel to your Regional Data Quality Monitor, at least thrice a week using the report format and naming convention: 7digit code_furtheranalysis_2021mmdd.

4.1.5 Steps 5 – Standard daily reporting format

DQMs are to report the progress of each day's work using the standard reporting format. The report should be saved using the naming format below:

2021PHC_DQM DAILY REPORT_DISTRICT CODE_YYYYMMDD_NAME

Figure 4.4: Sample Daily Report – Listing

DATE (yyyymmdd)	District	Time	Total EAs Expected	Total EAs in Data	Total Structures Listed	Partial	Dupl	Input Cases	Level Post	Records	Messages	Most Frequent Error Msg	Pct	Total count in Roster	Total Holds	Total Homeless	Total Vacant	Total Secondary HH	GQ Stable	GQ Unstable	Reports shared with DDQMT? Y/N	Action Taken	Any Comments
20210104	0523	22:49	88	4	65	24	0	65	65	234	170	LS03(1)	30.8	33.0	28.0	4.0	0.0	0.0	1.0	Yes		Start of TC field wa	
20210105	0523	18:03	88	72	2529	114	0	2529	2529	9730	426	ID02(2)	5.2	2064.0	1696.0	359.0	5.0	4.0	Yes		introduce EA 074 in SA 10 wa		
20210106	0523	18:58	88	84	7437	211	0	7437	7437	28543	1091	ID02(2)	3.4	6105.0	5043.0	1047.0	8.0	7.0	Yes		The supervisor for		
20210107	0523	17:12	88	87	10295	268	0	10295	10295	39310	1304	ID02(2)	2.5	8259.0	6883.0	1349.0	17.0	10.0	Yes				
20210108	0523	18:57	88	88	13387	285	0	13387	13387	50965	1410	ID02(2)	2.0	10605.0	8823.0	1749.0	19.0	14.0	Yes				
20210109	0523	21:28	88	88	17406	196	0	17406	17406	66092	1746	ID02(2)	1.4	13749.0	11437.0	2004.0	19.0	31.0	Yes				
20210110	0523	22:48	88	88	18626	95	0	18626	18626	70753	1585	LS08A(2)	1.4	14830.0	12364.0	2162.0	18.0	32.0	Yes		Only one EA (EA 81		
20210111	0523	15:13	88	88	18951	93	0	18951	18951	71931	1549	LS08A(2)	1.4	15035.0	12533.0	2196.0	19.0	33.0	Yes				
20210112	0523	19:18	88	88	20230	25	19	20230	20230	76551	754	LS08A(2)	1.4	15869.0	13471.0	2275.0	19.0	16.0	Yes		Two Eas (EAs 38 ar		
20210113	0523	17:48	88	88	21203	12	18	21203	21203	80175	693	LS08A(2)	1.4	16577.0	14091.0	2378.0	19.0	16.0	Yes				
20210114	0523	16:07	88	88	21281	10	0	21281	21281	80514	660	LS08A(2)	1.4	16668.0	14177.0	2384.0	19.0	16.0	Yes				
20210115	0523	19:45	88	88	21838	2	1	21838	21838	82691	656	LS08A(2)	1.5	17180.0	14617.0	2446.0	24.0	18.0	Yes				
20210116	0523	17:21	88	88	21846	0	0	21846	21846	82720	644	LS08A(2)	1.5	17182.0	14620.0	2446.0	24.0	18.0	Yes				
20210117	0523	18:42	88	88	21846	0	0	21846	21846	82719	644	LS08A(2)	1.5	17181.0	14621.0	2446.0	22.0	18.0	Yes				
20210118	0523	15:55	88	88	21853	0	0	21853	21853	82762	659	LS08A(2)	1.5	17241.0	14655.0	5.0	2446.0	69.0	47.0	19.0			
20210119	0523	5:08	88	88	21854	1	1	21854	21854	82758	653	LS08A(2)	1.5	17236.0	14654.0	5.0	2442.0	68.0	48.0	19.0			
20210120	0523	23:00	88	88	21860	1	0	21860	21860	82837	643	LS08A(2)	1.5	17295.0	14681.0	5.0	2476.0	65.0	48.0	20.0			
20210121	0523	10:01	88	88	21860	2	0	21860	21860	82860	638	LS08A(2)	1.5	17317.0	14701.0	5.0	2469.0	64.0	56.0	22.0			
20210122	0523	18:59	88	88	21865	3	0	21865	21865	82880	640	LS08A(2)	1.5	17322.0	14711.0	5.0	2464.0	64.0	56.0	22.0			
20210123	0523	8:57	88	88	21867	3	0	21867	21867	82917	644	LS08A(2)	1.5	17353.0	14731.0	5.0	2476.0	65.0	55.0	21.0			
20210124	0523	14:21	88	88	21867	4	0	21867	21867	82933	642	LS08A(2)	1.5	17367.0	14748.0	5.0	2474.0	64.0	55.0	21.0			
20210125	0523	20:35	88	88	21871	3	1	21871	21871	83022	640	LS08A(2)	1.5	17448.0	14796.0	6.0	2503.0	67.0	55.0	21.0			
20210126	0523	20:21	88	88	21871	3	1	21871	21871	83022	640	LS08A(2)	1.5	17448.0	14797.0	6.0	2502.0	67.0	55.0	21.0			
20210127	0523	12:10	88	88	21873	3	0	21873	21873	83033	634	LS08A(2)	1.5	17447.0	14790.0	5.0	2513.0	66.0	53.0	20.0			
20210128	0523	12:23	88	88	21873	3	0	21873	21873	83033	634	LS08A(2)	1.5	17447.0	14790.0	5.0	2513.0	66.0	53.0	20.0			
20210129	0523	8:09	88	88	21876	4	0	21876	21876	83078	631	LS08A(2)	1.5	17482.0	14778.0	5.0	2563.0	63.0	53.0	20.0			

Figure 4.5: Sample Daily Report – Main

DATE (yyyymmdd)	District	Time	Total EAs Expected	Total EAs in Data	Household Duplicates	Input Cases	Level Post	Records	Messages	Most Frequent Error Msg	Pct	Total count in Roster	Total Holds	Total Homeless	Total Vacant	Total Secondary HH	GQ Stable	GQ Unstable	Reports shared with DDQMT? Y/N	Action Taken	Any Comments	
20210118	0523																					
20210119	0523	5:10			119	700	700	4463	1040	Incomplete Questionnaire	20.3	220	201	0	13	3	1	2	yes		Visited Eas with no Main data	
20210120	0523	13:21			234	2113	2113	13480	2422	No Person record(s)	13.5	463	391	0	38	27	5	2	yes			
20210121	0523	7:04			255	2931	2931	18938	3021	No Person record(s)	12.1	740	590	1	59	79	5	6	yes			
20210122	0523	19:00			424	5143	5143	33086	4960	A07(5)	12.0	2091	1889	1	73	113	9	6	yes			
20210123	0523																					
20210124	0523																					
20210125	0523																					
20210126	0523																					
20210127	0523																					
20210128	0523																					
20210129	0523																					
20210130	0523																					
20210131	0523																					
20210201	0523																					
20210202	0523																					
20210203	0523																					
20210204	0523																					
20210205	0523																					
20210206	0523																					
20210207	0523																					

CHAPTER FIVE: CLEARANCE OF FIELD TEAMS

5.1 Clearance to leave the field after data collection

Your work in the district is not complete as a DDQM until data from all EAs have been synced to the server and materials distributed to teams have been received at the District Census Office. The DDQMT has the responsibility of clearing the field teams for completion of work and can therefore leave the field. Note that there are 2 levels of clearance before a team's work can be certified to be completed and ready for final payment.

- District level by DDQMT
- Headquarter level based on synced data, backup of tablet, and clearance by DDQMT

5.2 Clearance of Supervisors and Enumerators in an SA (Team) by DDQMT

Before clearance, the DDQMT should use these tools

- **DM report (listing and enumeration dashboard)**
- **Listing and main CSPro error reports generated**
- **Content of Supervisors' tablet**

to check for the following;

- a. **All structures** in the EAs have been listed and have reflected in the database.
- b. All structures with **household and non-household population** have been correctly enumerated.
- c. All data for the EAs assigned to a team have **NO**:
 - i. Duplicates
 - ii. Gaps
 - iii. Partial saves
- d. Total household and non-household populations identified during listing have been accounted for during the main enumeration
- e. Total number of EAs assigned to the SA/Team have been listed, enumerated and accounted for
- f. Total number of localities expected for each Type3 EA have been accounted for
- g. All **inconsistencies and errors** identified have been resolved for every enumerator in the team
- h. Final clearance for payment will be done by HQ after meeting all requirements set out above

5.3 Clearance of Supervisors and Enumerators in an SA (Team) at Headquarters

Clearance at HQ will check that;

- a. All data for the team has been synced
- b. The content of all the tablets have been backed up
- c. DDQMT have cleared teams and all materials collected

5.4 Clearance of the District Data Quality Monitor

The work as a DDQM is not complete until data from all EAs have been synced to the server, with no errors, gaps and duplicates. Furthermore, **ALL** teams in the district should have been cleared and the clearance report submitted to the RDQM in the specified reporting format. The RDQM proceeds to clear the DDQM after checking and confirming that:

- a. All inconsistencies and errors identified in the district have been corrected and the data re-synced to HQ
- b. All EAs and localities expected in the district have reflected with data in the database
- c. All SAs/Teams have been cleared with no duplicates, gaps and partially saved cases in data
- d. All logistics i.e., laptops and accessories have been returned in same condition as were deployed to the District IT
- e. District specific clearance report, final error reports and summary report have been submitted to the RDQM in the specified standard reporting format
- f. In the clearance report ensure that each SA has the following filled in
 - i. EAs expected and EAs covered
 - ii. No of localities expected and number of localities in data
 - iii. GAPS
 - iv. Duplicates
 - v. Building footprints and Total structures listed
 - vi. Total questionnaires listed (PHC 1A, Vacant, Group quarters)
 - vii. Number of EAs DDQMT validated the BFs
 - viii. Remarks
 - ix. Total questionnaires enumerated (PHC 1A, Vacant, Group quarters)
 - x. No population on census night (no_pop_Cnight)
 - xi. External monitors who visited
- g. List of all cleared teams in the district that have been received

CHAPTER SIX: REGIONAL DATA QUALITY MONITORING

6.1 Introduction

Regional Data Quality Monitors (RDQMs) will be stationed at the head office and work on all district data in their respective regions. RDQMs will monitor the daily activities of each district monitor by checking their daily reports and dashboard and assist them in correcting duplicates that need to be programmed. Apart from this, RDQMs will calculate and analyze a set of indicators for comparison for each district and region as a whole. **The output from the analysis must be saved daily for reporting at HQ in the format RR_monitor_report_2021mmdd.** Note that any unusual trend noticed in the course of this assignment must be **clearly communicated** to the DDQM in-charge, for investigation and possible correction. It must also be shared with other RDQMs and HQ for interrogation, logic coding and correction.

During enumerators/supervisors training, RDQMs will be with the Regional Data Quality Management Team (RDQMT). Their task includes visiting all districts and familiarizing themselves with his/her District Data Quality Monitors. After training, RDQMs will return to HQ to continue work and monitoring the DDQMs.

6.2 Roles of the Regional Data Quality Monitor

As RDQM, your work will include

- a. Check that all DDQMs have been deployed to their districts
- b. Assist in facilitating training of enumerators
- c. Ensure that all DDQMs in your region have all requirements needed before work begins
- d. Should contact all the DDQMs in your jurisdiction through phone calls and WhatsApp
- e. Communicate with other members of the RDQMT daily
- f. Check that all DDQMs are following outlined procedures to ensure complete coverage with minimal errors.
- g. Review the monitor report to ensure that the DDQMs are monitoring the data as expected.
- h. Ensure that the number of EAs expected for each district/region has been listed and enumerated
- i. Assist DMs correct all duplicates that need to be programmed in *justlist* and *justmain* programs in CSPro.
- j. Ensure that all DMs have the updated programs in their procedures
- k. Check the frequency of all variables in each module
- l. Do a further analysis of the region and district data on all modules and generate graphs and tables for discussion with NDQMs and Management

6.3 Data Backup for RDQM

After the last data access for the day, RDQMs are to take a backup of the day's data.

- a. In the C:\2021PHC\Other folder, create a folder called Backups.
- b. In the backups folder, create a subfolder each for listing and main as follows;
 - (i) C:\2021PHC\Others\Backups\data_listing
 - (ii) C:\2021PHC\Others\Backups\data_main
- c. Go to C:\2021PHC\DMS\HQData and locate the following files;
 - (i) "GHA_2021_PHCLIST" – This is the data file for listing
 - (ii) "GHA_2021_PHC1" – This is the data file for main
- d. Copy and paste the files located above into their respective folders as;
 - (i) "GHA_2021_PHCLIST" into C:\2021PHC\Others\Backups\data_listing
 - (iii) "GHA_2021_PHC1" into C:\2021PHC\Others\Backups\data_main
- e. Name the pasted files with the date of the day's work as shown below:
 - (i) "GHA_2021_PHCLIST_YYYYMMDD" – For the listing data file
 - (ii) "GHA_2021_PHC1_YYYYMMDD" – For the main data file

6.4 Guide to programing duplicates and other challenges in Justlist and Justmain

- a. Guide the DM to extract the cases that have to be corrected in the DB in the correct format using the dofile
- b. Be sure to reconcile the correct cases that are in error
- c. Copy into an excel sheet
- d. use the correct CSPro program format (e.g., date, eacode, structure number, etc.) and debug
- e. send the file to NDQM for compilation and feedback
- f. Resend the updated files to the DMs and ensure the problems are resolved.

6.5 Dealing with variations in questionnaire types

What to do when enumerated questionnaires are more than listed questionnaires (Total questionnaires is over 100% in dashboard report).

In a situation where the total questionnaires after enumeration are more than that of listing as shown in the main dashboard from the monitor report, browse in STATA the following code with the particular SA and EA it has occurred;

•br SAno eacode struct hhnum a07 if eacode == "the ea number with extra questionnaire(s)";
e.g., br SAno eacode struct hhnum a07 if eacode ==1304100234 if expected number of main questionnaires exceed listing questionnaires.

Action to take:

- a. Extract all structures and restype from the listing file (**list_report_all.dta**) and main file (**sec6_popv.dta**) to see which structure numbers the extra questionnaire(s) is/are coming from. This is done by matching the two files in an excel sheet. Place the structure listings with associated restypes from both files side by side to compare. Communicate to DIT to get back to the Enumerators/ supervisors concerned and let them step through those particular cases more than once and sync the data again. Run through the process again to check that the cases have reflected
- b. Also check the **justmain.LST** file to ensure the cases are not being deleted because of wrong dates (that is before start of enumeration/census night)

Figure 6.1: Sample of variation in questionnaire types from Listing and Main

	Listing	Main	diffTot	Main_ahhsize	popstat	head																
78	523	523100	ASIEDUA ANIWI	17	2	8	1	1	100	0	0	0	0	0	327	828	100.3038	-1	3.2198701	3.244019	6/8	209
84	523	523100	DJABANOI OHENE-AS	18	3	0	0	0	0	0	0	0	0	0	83	85	102.4096	-2	5.56	3.986842	303	76
89	523	523100	NTOW ABI DANSOA	19	2	-27	1	1	100	0	1	2	200	-1	818	819	100.1222	-1	2.614035	2.47563	1473	595

What to do when listed questionnaires are more than enumerated questionnaires (Total questionnaires is less than 100% in dashboard report).

This situation means the team has not completed their work and therefore there are more cases to be enumerated. When there are many cases, it is better to assist the team to know what cases are left to be administered. As the DQM, extract the structure and restype listings in both files and compare in excel.

- **br SAno eacode struct lh12b if eacode == “the ea number with extra questionnaire(s)”;**
e.g., If eacode 1304100234 has more Listed questionnaires relative to main questionnaires; then,
br SAno eacode struct lh12b if eacode ==1304100234

Action to take: Critically examine the cases which are causing the problem and discuss with DDQMT. Also check the **Justmain.LST** file that the enumerated cases are not being deleted due to wrong dates before census night. Escalate the problem to NDQM if the issues persist.

6.6 Guide to Further Analysis for RDQM

Further analysis is to inform Field Monitors and Census Management on the observations for the region in order to gain a better understanding and discover useful information for monitoring and decision-making. The purpose of further analysis is to extract useful information from data to take decisions based on analysis from the data. The further analysis would include the following analysis;

a. Compare statistics from listing and main data

- i. Generate the type of questionnaires from listing to be used in main enumeration for each district
- ii. Compare the type of questionnaires from enumeration and listing at the end of each day to report on progress
- iii. Outline any inconsistencies and outstanding questionnaires
- iv. Compare the number of vacant housing units from enumeration to listing

b. Population characteristics for districts and region

- i. Generate the sex ratios for your districts and region
- ii. Compare with that of the 2010 PHC, if available
- iii. Check the average household size
- iv. Show the age-sex structure of your districts and region. Comment on the results (e.g., are there any unique features? How does the pyramid look like?)
- v. Indicate the proportions of the different age groups (0-14-year-old, 15-64-year old, 65+ year old) in your districts and region
- vi. Calculate the dependency ratio for your districts and region

2. Education

Generate the following indicators

- i. Percentage of individuals who have never been to school in your districts and region
- ii. Proportion of all 3-18year old who have never been to school in your districts and region
- iii. Proportion of individuals that can read and write but have never been to school
e.g.,2- tab P11a P12a

3. Economic Activity

- i. Check the distribution of economic activity status in the last 7 days (p13a)
- ii. Check the economic activity status of the population in your districts and region
- iii. Check which type of occupation is predominant in your districts and region
- iv. Check which type of industry is predominant in your districts and region
- v. Check that the descriptions are explanatory enough (no single words)

4. Fertility

- i. Check the average number of births per woman in your districts and region
- ii. Generate the proportion of women 35years+ without children?
- iii. Graph the proportion of children ever born alive by ages of women for your region
- iv. Graph the relative age-specific fertility rate for your region
- v. Generate the Fertility Rate in your region

5. Mortality

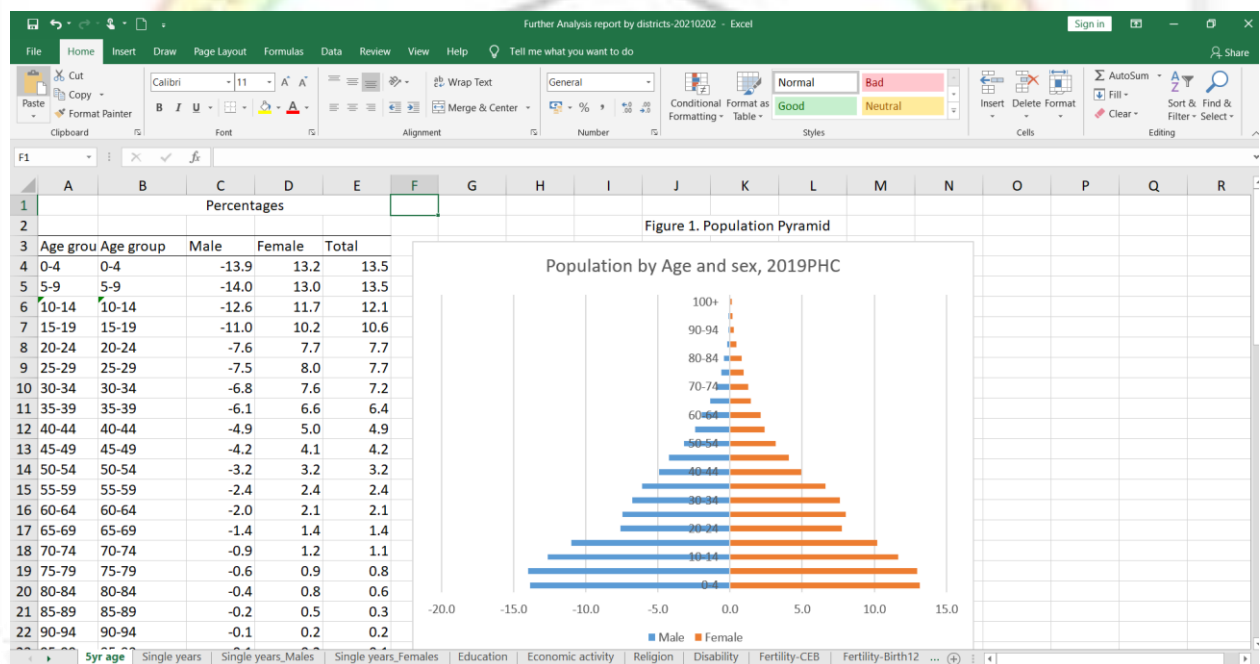
- i. Check for number/average deaths per household for your region
- ii. Check for maternal deaths for your region

6. Disability

- i. Check the disability status in your region

Other analysis may be proposed based on the further interrogation or request from Census Management.

Figure 6.2: Sample Further Analysis Report - Population Pyramid



6.7 Reporting on Highlights for the Region

RDQMs will be required to fill the contents of the highlights document – a report format to be used to track the progress of work in each region by district for each day. Highlights will be used to report the progress of work in the region to Census Management. There will be a sheet each for listing and main. The sheets for listing contain statistics for variables that are critical for data monitoring, like wise for the main sheets. It comprises two sheets on summary and frequencies of some selected variables. The sheet for each day's work is to be named using the standard DQM file naming format as; **2021PHC Highlights_District Name_YYYYMMDD**

Figure 6.3: Sample of Highlights - Listing

2021 TC3 - Fanteakwa District Highlights - Listing				
		% coverage BF in Listing after day 9	EAS	Percent
	1	<50 percent of BF	5	5.7
		50-99%	19	21.6
		100%+	64	72.7
		Structures expected	17,899	
		Total Structures listed	20,702	115.7
	2	Duplicates	Three EA with duplicate	
	3	Structure numbering Gaps in data		
		EAS Without gaps	80	90.9
		EAS with gaps	8	9.1
		1-10 structure gaps	7	87.5
		11-50 structure gaps	1	12.5
		50+	0	0

Figure 6.4: Sample of Highlights - Main

2021 TC3 - Fanteakwa District Highlights- Main Enumeration					
		Indicator	Expected	Enumerated	Percent
	1	% coverage of Eas in data after day 11	88	88	100.0
		% coverage of SAs in data after day 11	19	19	100.0
	2	Questionnaire			
		Number of Households -PHC1A	14754	9,876	66.9
		Number of Homeless-PHC1A	5	6	120.0
		Number of Vacant housing units-PHC1A	2563	1,879	73.3
		Number of Same Hh -code 98-PHC1A	63	48	76.2
		Group Quarter (stable)-PHC1B	39	9	23.1
		Group Quarter (unstable)-PHC1CDE	20	9	45.0
		Total Questionnaire	17444	11,827	67.8
		Duplicates			
		EAS without household duplicates	88	88	100.0

6.8 Clearance of Regional Data Quality Monitors (RDQM)

The RDQMs' work is not complete until all districts under their jurisdiction have data from all EAs synced to the server, with no errors, gaps and duplicates. All DDQMs within the region should have cleared their teams and all reports submitted to the NDQM. The NDQM should follow the following and any other check to clear RDQMS:

1. The RDQMs should have resolved all issues relating to duplicates, gaps, training data etc.
2. Ensure that there are no errors and inconsistencies in the region specific data
3. Region specific further analysis has been done and submitted to NDQM
4. Tabulation of key variables in region specific data have been completed
5. Clearance report from the RDQM has been received by the NDQM
6. The RDQMs' clearance report should contain for each district, the following
 - i. Number of EAs
 - ii. Number of EAs outstanding
 - iii. Number of Teams/SAs
 - iv. Number of Teams/SAs Cleared
 - v. Total structures listed
 - vi. BF percentage covered
 - vii. Total questionnaire types (PHC 1A, Vacant, Group quarters)
 - viii. Male, female and total population listed
 - ix. Number of households enumerated
 - x. Total household and non-household population
7. Logistics have been submitted to the NDQM

LIST OF CONTRIBUTORS

Jacqueline Dede Anum

Samilia Enyamah Mintah

Josephine Baako-Amponsah

Kwaku Awuah Akwaboah

Aurelia Hotor

Freda Annette Attah

Simon Onilumor Titchutab

Nana Antwiwaa Nuamah

Anointing Lartey

Ransford Kobina Barths

Michael Agyakwa

Angela Otchi



YOU COUNT. GET COUNTED

Copyright © 2021 Ghana Statistical Service