STATISTICS BOTSWANA

OPERATIONS MANUAL

DIRECTORATE OF STANDARDS, METHODS AND INFORMATION SYSTEMS

Statistics Botswana. Private Bag 0024 Botswana
Tel: (267) 367 1300. **Fax:** (267) 395 2201. **Email:** info@statsbots.org.bw **Website:** www.statsbots.org.bw

BOTSWANA 50 UNITED AND PROUD

DIRECTORATE OF STANDARDS, METHODS AND INFORMATION SYSTEMS OPERATIONS MANUAL

DIRECTORATE OF STANDARDS, METHODS AND INFORMATION SYSTEMS

Published by

Statistics Botswana
Private Bag 0024, Gaborone
Tel: 3671300/ 3671400
Fax: 3952201
E-mail: info@statsbots.org.bw

Contact Division
Standards, Methods, Quality Assurance, Training and Certification Manager
Statistics Botswana

December 2016

PREFACE

The Directorate of Standards, Methods and Information Systems (DSMIS) is a service Directorate that coordinates the overall organisational Sampling Frames Support Services (Households and Business Surveys); Standards, Methods, Quality Assurance, Training and Certification Services; Fieldwork and Regional Services Coordination; as well as Data Management and Information Systems. These services are extended to the key stakeholders as and when required. It is constituted by four (4) Divisions which are:

1. Sampling Frames Support (SFS)
2. Standards, Methods, Quality Assurance, Training and Certification (SMQATC)
3. Fieldwork and Regional Services Coordination (FRSC) and
4. Data Management and Information Systems (DMIS)

The divisions have distinct roles and responsibilities all geared towards the provision of timely, reliable and quality data for analysis in order to drive evidence based decision making and support investment initiatives. This Operations Manual is divided into four (4) chapters, each detailing the operations of a division as outlined above.

A.N Majelantle
Statistician General
December 2016

TABLE OF CONTENTS

CHAPTER 1: SAMPLING FRAMES SUPPORT DIVISION

The Division of Sampling Frames Support has two (2) units, namely: Business Registers and Cartography and Geographical Information Systems (GIS). It is responsible for the provision of sampling frames for both household and business surveys programmes, cartographic and GIS services for field enumeration area maps, statistical thematic maps, geo-coding and production of Population and Housing Census Atlas. The Business Registers is mandated to provide an up-to-date and reliable statistical business register, which forms the basis for sampling frame for business surveys.

1.1 CARTOGRAPHY AND GEOGRAPHICAL INFORMATION SYSTEMS (GIS) UNIT

1.11 Background

The Cartography Unit has been in existence as an ad hoc unit for some time. The unit was established as a permanent unit after the 1991 Population and Housing Census conducted in Botswana to service the entire mapping needs of the organisation, stakeholders and the public at large.

The unit is charged with the responsibility to oversee all matters relating to cartography, mainly the production of census enumeration area maps, as well as the production of maps with thematic indicators such as population distribution, densities, migration, price collection points, poverty, disability or illiteracy for publications and all reports produced by the organisation.

The enumeration area maps form the basis for development of the household based sampling frame. The mapping is done using both conventional mapping methods and modern geospatial technologies such as Geographical Information Systems (GIS), Global Positioning System (GPS) and Remote Sensing.

1.12 Data Collection

It is worth mentioning that before producing maps, necessary data used to produce and update these

maps are collected. The work of developing maps is done in conjunction with the Department of Surveys and Mapping as the sole provider of GI data. The Department serves as the major source of baseline data, notwithstanding other sources such as Botswana Power Corporation (BPC) for power line data, Land Boards for locality areas (cattlepost and lands), Local Authorities, private data vendors and Department of Town and Regional Planning (DTRP). Data collection is also done in the field using Global Positioning Systems devices (GPS) where geographical coordinates are collected for geo-referencing and map updating purposes. Other data are captured through remote sensing techniques.

1.13 Data Capture, Editing and Manipulation

Since data used in map production comes from diverse sources and in different formats, they are thoroughly checked, edited, and transformed into the same coordinate system to conform to the current format for ease of reference, data sharing and dissemination between agencies. The Department of Surveys and Mapping as the sole provider of spatial data in the country developed the Botswana National Geographic Referencing System in 2002 (BNGRS02) based on the Universal Transverse Mercator projection (UTM and LO coordinate systems). This system is meant to control coordinates within the country. All the maps produced in Statistics Botswana are based on the BNGRS02 coordinate system. This is also in line with the establishment of the national geographic data standards facilitated by the Botswana Bureau of Standards (BOBS) aimed at improved data integration and sharing amongst the Geographical Information community and the reduction of duplication of efforts.

1.14 Cartographic Fieldwork Operations

The cartographic work is done eighteen (18) months before commencement of the Population and Housing Census as per the recommendation by the United Nations. It entails updating of maps, systematical numbering of dwellings and building structures, as well as recording of the estimated population in each dwelling or building structure countrywide.

It has two main objectives and these are as follows:

a) To collect data and information used to produce reliable maps which will constitute the frame within which orderly enumeration will be conducted, and also produce data for estimating the workload of enumerators and supervisors;
b) To ensure that every person in the country is enumerated once at the time of the Population and Housing Census to satisfy its objectives.

Information collected during this exercise is vital during the planning stages of the main Population and Housing Census such as resource mobilization, delineation of enumeration areas, determining the location of the population and locality type, updating and establishing the sampling frame for inter-census household surveys, as well as the dissemination of the census results at the lowest geographical level for rural development planning Therefore, the success in conducting the population and housing census or a household survey is dependent on the quality of the cartographic fieldwork.

1.141 Mapping Relevant Documentation

In the process of documentation, a project memorandum (PM) defining the scope and objectives of the project, resources needed and request for the project funds is developed. Other key documents such as invitation to tender (ITT) documents, instructional manuals and guideline documents are prepared for field mapping, map preparation and production together with forms to facilitate data collection in the field and maintain records in the office.

There are documents necessary to be availed before the commencement of the mapping exercise and these include: mappers instructions manual; map production specifications manual; dwelling listing forms; mapping field return recording forms; map editing guidelines manual, and various maps - topographical maps (digital and analogue), aerial photographs and satellite images, and physical plan maps.

1.142 Data collection Instruments

a) Pre-listing Form: this form is used to collect dwelling information such as the dwelling number assigned, status and number of persons found during the cartographic listing exercise. This information is central to enumeration area delineation **(See example on Appendix I)**.

b) Site Condition Form: provides information on number of persons from the pre-listing form, total residential places occupied, total residential places vacant and total residential places; it also provides type of locality and main language spoken. The information collected is used for the general planning of the census operations **(See example on Appendix I)**.

1.143 The Geographical Coding Scheme

In the development of the coding scheme, unique codes are assigned to different geographical entities such as districts, villages, cities/towns, localities, lands area in order that they are identified uniquely. For the convenience of Population and Housing Census operations, all cities and towns are considered as census districts and are allocated **code 0**. For example, Gaborone is identified by code 01, Francistown 02, Lobatse 03 and others follow the same pattern.

1.144 Coding of Rural Localities

Each locality is assigned a three (3) digit code, the first digit code gives an ecological definition of the locality, and next two digits are the serial number of the locality. The entire locality codes are shown in the table below;

Locality Type	Locality Code
Cities and Towns	0
Village	1
Lands Area	2 or 7
Cattle Post	3 or 8
Mixture (lands and cattle post)	5
Freehold Farm	4 or 9
Border Post, work camps etc.	6

If it happens that the locality codes for lands areas reach the maximum code of 299, the preceding locality is assigned a code of 701. This applies to both cattle posts and freehold farms as indicated in the above table.

1.145 Coding an Enumeration Area

Using the geographical codes stated above, it is easy to assign codes to enumeration areas in different localities. For Example, a code of an EA in Mochudi in Kgatleng District would be coded 40 - 01- 0011 where 40 would be a District code, 01 would be a code indicating the first village in the District whilst 0011 would indicate the EA code.

1.146 Coding a Locality

A similar procedure is followed when coding localities. A code for a lands area in Mochudi village would be 40 – 01 - 201, where 40 would be the district code, 01 would be the village code (Mochudi) and 201 for the locality code for the first lands area of Mochudi. The following example shows codes for different localities by their ecological type within Kgatleng district and associated to Mochudi village;

40 - 01- 201 would be for the lands
40 - 01- 301 would be for the cattle post
40 - 01- 401 for freehold farm;
40 - 01- 501 for mixture of lands and cattle posts; and
40 - 01 -601 would be for border post

1.147 Codes for Census Districts

The census districts were assigned original codes with an interval of ten (10) to allow for subdivision in the event a district expanded. For example, Southern (10) was subdivided twice to have Barolong (11) and Ngwaketse West (12) and the original code was maintained by the parent district.

The following new Sub Districts have not yet been assigned geographical codes because they have physical boundary problems;
a. Tlokweng
b. Moshupa
c. Central Serowe/Palapye - The district needs to be divided into two sub-districts: being Serowe and Palapye
d. Tonota

1.148 Planning the Mapping Exercise

The cartographic fieldwork of a Population and Housing Census is a very huge and expensive project that requires a well thought planning process in terms of timing the execution of the work plan and optimal resource allocation to ensure that it is completed on time and meets the required standards. The planning process includes stock taking of all the existing cartographic products such as determining their relevance to the mapping exercise, resources availability, administrative logistics, publicity and advocacy campaign for the cartographic fieldwork, fieldwork management strategy in order to enable timely delivery of quality products.

The success of cartographic mapping depends on the availability of mapping equipment and other supporting materials such as base maps, vehicles, GPS, computers, peripheral devices, and stationery. An inventory of available items is made as a basis for determining additional equipment requirements.

1.149 Administrative Issues and Logistics

The enormity of the tasks of field mapping requires a large number of personnel resources. Since the GIS/ Cartography unit cannot raise the required number of personnel for field mapping, suitable temporary staff are recruited and trained for at least three-weeks on field mapping techniques, including map reading, EA identification, Geo-coding scheme, dwelling listing, map updating, aerial photo interpretation and the use of global positioning systems (GPS) to augment staff compliment in the Unit. Procurement of suitable fieldwork vehicles, equipment, and vehicle mounted radio telecommunication systems, computer software and hardware, stationery and materials to support fieldwork activities will form part of the logistics to complete the tasks of field mapping. Newly developed technologies in areas of cartography and geographical information systems such as Smart Client would also greatly improve the cartographic fieldwork and hence reduce costs.

1.1410 Commencement of Field Mapping

The cartographic mapping needs to commence beforehand to ensure that all EA maps are ready three (3) to four (4) months before the work of carrying out of the Population and Housing Census begins, in conformity with the United Nations recommendations.

The required number of teams depends on the projected number of enumeration areas and on the experience from the previous mapping exercise. Generally, the number and composition of teams should be manageable for effective monitoring to deliver quality cartographic products.

1.1411 Cartographic Inspection

This exercise is carried out by experienced technical officers from various divisions and units within the

organization six (6) months after the commencement of the mapping exercise to quality check the mapping exercise. Each technical officer is assigned at least five (5) enumeration areas of different ecologically classifications to check the quality of the mapping exercise particularly:
a) completeness, i.e., all the dwellings within an EA have been affixed metal plates
b) EA boundaries are clearly marked and visible on the ground
c) Locality codes and names, e.g., duplication of codes
d) Mix-up of localities across enumeration areas
e) Census numbers i.e. whether they are listed systematically, visible and readable

The cartography inspection is carried out twice before the census enumeration and the feedback from this exercise is very important in terms of planning, monitoring, management and quality control of the entire mapping exercise. In addition, it enables smooth running of the Population and Housing Census enumeration exercise, by ensuring that every person in the country is counted with minimal possibilities for over-enumeration or under-enumeration.

1.15 EA Delineation Procedures

The principal objective of enumeration area delineation is to sub-divide the country into contiguous small geographic areas (EAs), each being reasonably big enough in terms of ideal population size and area. An ideal population size is the number of dwellings and people that one enumerator can satisfactorily enumerate within the time period scheduled for data collection.

The maps that are to be used are updated using the previous Population and Housing Census data in preparation for the upcoming census through data collection and fieldwork activities. In doing so, the following factors are observed, among others:

- EAs should cover the entire land surface of a country without gaps;

- The EA boundaries should be marked explicitly on the map and should not overlap;

- The EA boundaries should not cross a road, river, railway line or stream because in most cases the boundaries should follow these features because they are discernible on the map as well as on the ground;

- Each EA should be assigned a unique code tied to the district code and its locality of affiliation;

- EAs should be accessible and be ideal in terms of population size and areal extent so that an enumerator can be able to enumerate it within a scheduled time period; and

- EAs in rural areas should range between 75 – 150 dwellings with a population size of between 330 and 660; while those in urban areas and major villages range between 220 – 440 dwellings with a population size of between 440 and 880 people.

1.16 GIS Acquisition and Implementation

A Geographical Information System (GIS) is defined as a collection of computer hardware, software, people and geographic data for capturing, storing, checking, integrating, manipulating, analysing and displaying data related to positions on the earth's surface. The system was acquired immediately after the 2001 Population and Housing Census and was implemented in 2002 for data capturing.

Typically, a GIS is used for handling maps of one kind or another. These might be represented as several different layers where each layer holds data about a particular kind of feature class, e.g. point, line or polygon. Each feature is linked to a position on the graphical image on a map and a record in an attribute table. This unique capability of GIS to exhibit spatial relationships between features in a database distinguishes it from other information systems such as AutoCAD.

The system was also used to produce thematic maps for the Census Atlas two years after the Population and Housing Census of 2001. As for the 2011 Population and Housing Census, all the cartographic work was based largely on GIS technology with digital orthophoto maps and remotely sensed data used as backdrops for topographic mapping especially in rural areas. The Global Position System (GPS) devices were also used to

collect geographical coordinates of rural localities and plot them to update the district maps. The system was also used to produce final enumeration area maps for the 2011 census.

1.17 Cartographic Technical Support

As mentioned earlier, Cartography Unit was established mainly to satisfy the mapping needs of the organisation. During the mapping exercise, permanent staff from the unit provides technical support to teams in the field in all matters relating to cartography and also participate as quality controllers for the entire field work.

In addition to preparing of enumeration files and printing of EA maps during the conduct of censuses and household surveys, cartographic unit staff participate as trainers, enumerators, technical officers and supervisors as well as providing technical support during main Population and Housing Census enumeration. It also participates during the cleaning, editing of the census data, and during the dissemination of the census result.

1.18 Production of Reports

The cartographic fieldwork data have a bearing or input into a number of reports produced in Statistics Botswana by subject matter specialists in collaboration with the Cartography and GIS Unit. These are listed below;

1.181 Cartographic Technical Officers Report

The Cartographic Technical Officers report is produced after cartographic inspection to assist or guide on areas that need attention or where there are gaps still to be filled.

1.182 Country Code List Report

The country Code List is used to guide the cartographic fieldwork and the main enumeration by listing all the localities and their geographical codes. This list, which is normally produced after the Population and Housing Pilot Census, is also used for checking during cartographic inspection exercise.

1.183 Population of Towns, Villages and Associated Localities Report

The unit plays a key role in the release of this report before the release of the main census results by providing geographical codes, locality names and population by district. It is released before the Population and Housing Census results.

1.184 Production of a Census Atlas

Cartographic unit is mandated to produce a Population and Housing Census Atlas report two years after the successful conduct of a Population and Housing Census. A Population and Housing Census Atlas is defined as a collection of maps accompanied with text in a book form. This is in line with a recommendation by SADC that urges member states to produce such a document after conducting a census.

1.19 Customer Requests

Immediately after the census, the unit receives a lot of requests for geographical datasets and enumeration area maps. These come from different stakeholders such as researchers, students, government ministries and departments, academia, private sector, politicians, elections office, non-governmental organisations and individual for various uses. There is also a high demand of thematic maps for publications within Statistics Botswana and stakeholders.

1.10 Challenges

The biggest challenge the unit faces is to address customer queries regarding boundary issues and association of localities immediately after the release of the main census results. However, the responsibility of boundary identification lies with other Ministries not Statistics Botswana

1.2 BUSINESS REGISTER UNIT

1.21 Introduction

The statistical business register is a critical baseline data for sampling of economic/business based surveys. It is also used by entrepreneurs/investors and academic researchers in order to derive evidence based decision making for investment purposes. An effective, efficient and reliable statistical business register has been a problem to most of National Statistics Offices especially in Africa, Botswana being one of them. This is mainly due to the unreliability and fragmented data sources and the software/system used to administer and update the business register the result of which created a problem of accuracy of the information. The business register contain a list of registered companies (establishments and enterprises). The variables that are included in the business register include the following components of business:

a) Incorporation/company number
b) Tax/PIN number
c) Registered name
d) Trading name
e) Postal addresses
f) Physical addresses
g) Emails
h) Telephone numbers
i) Fax number
j) Nature of principal business - BSIC (business activity)

k) Turnover
l) Number of employees
m) Nature of Company (Institutional type)
n) Ownership by nationality
o) Activity status

The current international standard under the statistical business register is the International Standard Industrial Classification Revision 4 (ISIC Rev. 4) which took effect in August 2008 after the lower version Revision 3. Botswana has since adapted this classification to form Botswana Standard Industrial Classification Revision 4 (BSIC Rev. 4). The classification is shared with stakeholders for coding where data is being sourced.

1.22 Source data (primary and secondary)

The secondary data for the business register is sourced from various administrative organisations which include Botswana Unified Revenue Services (BURS), Department of Industrial Affairs, Census of Establishments and Enterprises (CEE), Department of Consumer Affairs and the local authorities (councils). Primary data is sourced through the survey feedback, phone calls and physical visits to the companies. The statistical business register is therefore continuously updated.

1.23 Outputs

The unit is responsible for the production of business based survey sampling frames. It also produces business statistics digest and provides statistical business requests to various customers.

The basic concepts and definitions can be found in the Compendium of Concepts and Definitions publication

CHAPTER 2: STANDARDS, METHODS, QUALITY ASSURANCE, TRAINING & CERTIFICATION DIVISION

The Division of Standards, Methods, Quality Assurance, Training and Certification (SMQATC) has three (3) units which are: Standards and Methods, Quality Assurance, Training and Certification; and is responsible for the following:

a. Provision of expertise on statistical methodologies, principles procedures and concepts for all surveys and decennial censuses conducted by Statistics Botswana and other stakeholders;

b. Training and certification on official statistics for the organisation and stakeholders; and

c. Quality Assurance with respect to standards, control processes and procedures, compliance and classifications.

2.1 STATISTICAL METHODS UNIT

This operation manual details guidelines in the form of general activities involved in preparatory work for a survey with much emphasis on the sampling and methodology. The following are the table of contents of the operation manual:

Background
Sample design
Survey Design Evaluation
Statistical Methods Research

2.11 Background

The Statistical Methods Unit (SMU) of the Statistics Botswana is charged with the responsibility to provide expertise on the areas of methodologies, principles, approaches, processes, procedures, concepts

techniques and units of scope to various projects. Among the specific tasks of the SMU is to drive the design and the development of the statistical methods through supervising the sample design activities which entail reviewing censuses and survey operations. This also involves recommending appropriate methodologies and overseeing the application of the statistical principles, methods and approaches.

2.12 Sample Design

Generally a sampling design should be limited to a two stage plan. That is, a stratified two stage probability sampling design is generally employed for the selection of samples. The first stage is the selection of EAs as Primary Sampling Units (PSUs) and is selected with probability proportional to measure of size, the measure of size (MOS) being the total number of households as defined by the recent Population and Housing Census, in this case the Botswana 2011 Population and Housing Census.

The second stage is the selection of the actual households to be enumerated after listing within the selected EAs. That is to say, the households are systematically selected from a fresh list of households for each selected EA. Section is entirely dependent upon an up-to-date sampling frame. The sequential processes are detailed in the sub-sections that will follow.

2.13 Target Population

Generally the target population depends largely on the objectives of the household survey in reference and hence differ from one household to the other. However, it covers all households selected based on the selection or eligibility criteria determined by the objectives of the survey in question. For instance, the 2013 Literacy survey covered all households with persons aged between 10 and 70 years of age who are currently not attending formal school, had never attended school, have attained standard four (4) or below or its equivalent or currently attending non formal education; **WHEREAS** the 2015 Agricultural Census covered all households having at least one cow, goat/sheep, donkey/mule, horse, and/or having a piece

of land whether planted or laying fallow; **WHEREAS** for the 2014 Botswana Information and Communication Technology survey, all households with persons aged 10 years and above, and are usual members who spent last night with the household are targeted.

2.14 Sampling Frames

2.14a Household Based Surveys Sampling Frame

The sampling frame is a list of all the population elements from which a sample is drawn or selected. A sampling frame is normally current and based on the preceding Population and Housing Census (PHC). The sampling frame, thus, comprises a list of all Enumeration Area (EA) as demarcated during the latest PHC. The setup is such that all household surveys undertaken after the census use that sampling frame as baseline. For instance, the 2011 Population and Housing Census formed the basis for the 2013 Botswana Literacy survey and 2014 Botswana Information and Communication Technology survey.

It is for these basic reasons that the sampling frame should be up-to-date and adequate, needless to mention that the quality of data obtained from subsequent surveys depends largely on it.

Household surveys samples exclude institutional EAs such as security/armed forces barracks (prisons, police and army), nursing homes, other distinct institutions like schools.

The current sampling frame is defined and constituted by all Enumeration Areas (EAs) found in three geographical regions viz. (i) Cities & Towns (ii) Urban Villages , and (iii) Rural Districts as defined by the 2011 Population and Housing Census. It consisted of 5,063 EAs being the total number of Enumeration Areas (EAs) delineated during the 2011 Population and Housing Census.

2.14b Business or Economic Based Surveys Sampling Frame

Business based surveys utilise a different sampling frame. This kind of a frame constitutes a list of businesses or establishments in the country. The frame is well defined in Chapter 1 on item 1.2 of this manual.

2.15 Master Sample

A master sample is a sample of EAs selected from the main sampling frame as constructed from the Population and Housing Census such that its size will allow sub-samples to be drawn from it which meet the needs of different surveys.

Advantages of a master sample, accrues by way of the smaller number of primary sampling units (PSUs) which would, among other things, allow clearly delineated units which would be easier to update frequently compared with the entire frame. The 2001 to 2011 inter-census period of household surveys conducted in Botswana on the other hand, does not use a master sample safe for agricultural related surveys (censuses and annual agricultural surveys); where a master sample is selected based on all agricultural holders instead of the sampling frame constructed from the Population and Housing Censuses, i.e., 2011 PHC constitute the sampling frame from which sample selections of various surveys are taken.

2.16 Stratification

Stratified Sampling is a technique whereby the entire population of sampling units is divided into distinct sub-populations called strata, and within each strata a separate sample is selected. In most developing countries such as Botswana, and hence Statistics Botswana, a two stage Stratified Sampling is applied/employed as an appropriate procedure. It has an important advantage of reducing the sampling error, which is a critical aspect in household surveys to attain quality statistics. That is why it has an edge over other designs. Other benefits of stratification include the following:

- Administrative convenience and flexibility; and

- Renowned for guaranteed representation of important domains and special sub-populations.

In almost all household surveys carried out by Statistics Botswana, EAs are stratified into three (3) or four (4) broad strata (usually referred to as domains) depending on the type and objectives of the survey undertaken at that time; and these are:

- Towns and Cities (at times separated depending on the objectives of the survey),
- Urban Villages, and
- Rural Villages

The maximum benefits of using Stratified Sampling Procedure are optimally derived when it is applied correctly together with the Systematic Sampling Procedure for further selection of surveying units (households/ dwellings/ holders). It is also commonly justified in the field of Statistics as "Implicit", basically because it is applied at each stage until the required sample is attained, and ensures objective PSU categorization. PSUs form part of critical aspects in sampling designs; they affect the quality of all subsequent phases of the survey process. Therefore, the PSUs should be mutually exclusive and collectively exhaustive strata because it is at a level where benefits of Implicit Stratification are particularly strong and derived.

In any survey, it is of paramount importance to determine the number of such PSUs to be selected in the sample. Subsequently cluster size (number of households) has to be determined and controlled to a certain minimum in order to obtain reliable estimates. A reasonable cluster size in the design would yield more reliable estimates because the sample will be more spread out.

In that regard, a recommended cluster size of 15 – 20 households per EA will be ideal in order to manage the PSUs and make enumeration easy. (UN Publication of 2005, "Household Sample Surveys In Developing and Transition Countries")

Statistics Botswana uses a cluster size in the range of 15 -20 households per selected EA in almost all surveys except in rare cases where a 12 is used; for example the 2015 Botswana Multi Topic Household Survey due to

the complexity of the survey; while 2013 Botswana Literacy survey enumerated 20 households per selected EA.

As a Rule of Thumb: It is better to have more PSUs and smaller cluster sizes. In conclusion, PSUs in general should possess the following characteristics:

- Clearly identifiable boundaries which are stable over time. In Statistics Botswana, this information is provided through mapping using the modern GIS technologies. Admittedly, the number of households in a PSU may change overtime, this require regular updating whenever a survey is undertaken;
- cover the target population completely;
- have a measure of size for sampling purposes as mentioned before; and
- Should be large in number to be segmented and easy to explore.

2.17 Determination of Sample Size and Allocation

Clustering of the geographical areas in the country have a direct bearing in the sample design, especially in determining the sample size ***n*** for the national household surveys. These strata are normally referred to as domains, in the estimation procedures. Within the process of determining the sample size, several factors are taken into consideration, including the sample size increase factor.

Turner GA (2000) has developed an adjustment factor that is used with number of domains as part of the process of determining the sample size, that is, $\mathbf{D}^{0.65}$, where **D** is the number of domains that the sampling process would have adopted. This factor is now being used in almost all household surveys carried out including those conducted by Statistics.

The decision on the number of domains has to be taken by the entity based on the objectives of the survey and after involving the relevant stakeholders. For example, in the case of the domains, (rural areas, urban villages and cities/towns), the Domain will be 3, implying $\mathbf{3}^{0.65} = 2.04$. This plays a very pivotal role because the sample size ***n*** will be increased by 2.04 for that particular survey.

The domains (strata) in the sample are key phenomena to the analytical needs for a country and driving the objectives of the survey as well as achieving its intent. This procedure allows the encompassing of sub-populations in the sample and this enhances policy decisions.

This procedure is applied in Statistics Botswana whenever a sample size is determined; see the example below as was applied in the 2015 Botswana Multi Topic Household Survey (**BMTHS**) methodology design.

Turner G A (2000) suggested that allowable margin of error for domain estimates to be 40% greater than national. If margin of error is 5 percentage points and increase level n by factor D0.65 (D = number of domains) then;

D	2	**3**	4	5	6	10	15	50
Sample size increase factor	1.57	**2.04**	2.46	2.85	3.2	4.47	5.81	12.72

2.18 Sample Size Determination

The calculation of the sample size is based on the three main factors, which are;

i. Basic precision
ii. Number and type of estimation domains
iii. Needs for estimates, entailing indicators and objectives of the survey
iv. Cost and time required for data collection, processing and dissemination of survey results.

Precision requirements involve specifying the sampling errors within which the survey estimates are wanted. The survey estimates may be subject to both bias and variable errors that affect the expected value of the survey estimates, taking it away from the true value of the parameter. These sampling errors should be minimised or made negligible by making the sample size large and adhering to adequate sampling procedures, by way

of increasing the reliability and making the sample more representative. This, however, is challenged by resources needed to carry out the survey (financial, human and other materials necessary).

A standard practice is to compute the requirements for each of the key indicators needed to be measured, and in the process, the largest sample size will be derived from such calculations, and hence chosen. In general, the rarer the indicator, the larger the sample size n required for the measurement of the estimates in the survey. It is pertinent to choose an indicator that will yield the largest sample. For instance, in the 2015 BMTHS sample size determination, the rarest indicator was "breast feeding" based on the 2007 Botswana Family Health Survey. The margin of error for each of the key indicators will then be no bigger than the margin of error for the estimates requiring the largest sample size.

Other important factors to be considered are such that:

a) There has to be sufficient information on the key related indicators (variables) and the corresponding prevalence from the previous surveys undertaken. For example, the 2007 Botswana Family Health survey and 2011 Population and Housing Census information was used to determine sample size for 2015 BMTHS. This could play a role in informing envisaged designs if they drive the similar sets of objectives.

b) There has to be information available to calculate the clustering or the Survey Design effect ***Deff***. ***Deff*** is defined as the ratio of sample variances from the survey to that of a simple random sample of the same size. The ***Deff*** from the previous surveys with similar objectives that has been carried out in the country may be assumed to be equal, otherwise an acceptable range of ***Deff*** should be between 1.5 and 2.5, (William G. Cochran, 3rd Edition, 1977; Sampling Techniques). It is also important to note that failure to take account of the survey design effect in the estimates of standard errors can lead to invalid interpretation of the survey results. For instance, in the case of the 2015/16 Botswana Multi Topic Household Survey, the ***Deff*** employed was 1.55, while for 2014 ICT survey it was 2.

c) Specification of the level of confidence is critical, it is the risk level based on ideas encompassed under the Central Limit Theorem that is when a population is repeatedly sampled, the average value of the

attribute obtained by those samples is equal to the true population value. The recommended confidence level for the surveys is approximately 95-percent; and this is what Statistics Botswana applies. This means, if 95 percent confidence level is selected, 95 out of 100 samples will have the true population value within the range of the precision specified.

d) Adjustment for potential non-response rate is also one of the elements that factor in the calculation of the sample size ***n***. Using this type of factor is necessary to raise the sample size with a specific percentage of non-response, which is normally **1- 0.05 = 0.95**. It will give the final number of households in a sample after adjustment of the non-response. The justification for this step within the procedure emanates from the fact that survey response is not likely to be **100** percent, hence it is inevitable. Moreover, this will ensure that the sample size for the interviewed households will still attain the required n, even though the non-response households will contribute to a non- sampling bias to the results. See example in Table 4.51 for more details.

e) It is worth noting that the calculation of the sample size n has to take into consideration the required resources. These may be done by either increasing the allowable margin of error, reducing the number of domains or the ***Deff***. It is an objective trade-off and risk worth taking rather than having the envisaged survey failing to be implemented due to resource constraints.

2.19 Sample Selection and Allocation

The selection of PSUs or EAs involves procedures and formulae to ensure a well-balanced representative sample. Therefore, the selection is done using probability proportional (**pps**) to a measure of size, being the number of households. In general, a wide variability in the number of households across PSUs has adverse effects on the fieldwork logistical planning.

The procedure for selecting the EAs in each stratum consists of:

i. Calculating the sampling interval for the stratum:
   $k = \sum X_i / n$
   Where;
   $\sum X_i$ is the size of the stratum (total number of households in the stratum according to the recent Population and Housing Census, i.e., 2011 Population and Housing Census) and 'n' is the number of EAs to be selected in the stratum;
ii. Calculating the cumulated size of each EA;
iii. Calculating the sampling numbers R, R+k, R+2k…R+ (n-1) k, where R is the random number between 1 and k; and
iv. Comparing each sampling number with the cumulated size.

In the process of selection, the first EA selected is the first whose cumulated size is greater or equal to the random number (sampling number). In this way the sampling interval is increasingly cumulated and added to the constant random as explained in (iii) above. Subsequent EAs are selected thereof until the required sample size has been reached as per the Census Districts within the three domains (in this case cities/towns, urban villages and rural areas). It should be noted that sample selection and allocation are done simultaneously.

The allocation at district level is necessitated by the fact that, the Government being the most user of statistical information needs such information to substantiate evidence based decision making.

The procedure reduces sampling error to the extent that the strata which are set up are internally homogenous with respect to the variables of interest, hence lower standard errors. However, heterogeneity will exist across strata or is expected as indicated before, that the strata may be rural, urban and cities/towns. It goes without saying that the livelihoods, trends and patterns may vary significantly (Kish, L (1965), Survey Sampling).

2.110 Quality Control

The involvement of a Sampling Statistician in quality control operations is to guard against the violation of the sample requirements, procedures and principles during fieldwork implementation. After establishing the procedures for selecting the households/dwellings or holdings, the next stage is to partake in the listing of the enumeration units/households. The listing process is done as a separate exercise before enumeration to ensure complete frame coverage. The information from household listings is of essence during the calculation of the weights, that is, it feeds into the weights calculation formulae.

2.111 Weights applications

Weighting is a necessary process which is applied in household surveys. It is defined as the grossing up of the sample data to the population size for estimation of the population parameters. The weights are mathematically equal to the reciprocal of the probabilities of selection, and they reflect the sample design used. Therefore they are calculated separately for each survey undertaken. A non-response adjustment is applied as necessary to make the sample counts conform to the original sample size.

A further adjustment which might be considered is when information is available from an external source on the distribution of enumeration units by area. This may be an independent data that can be suitable to be used in a post-stratification scheme to improve the estimates.

2.112 Survey Design Evaluation

After every survey, standard errors form part of the statistical analysis. They are computed and presented together with other descriptive statistics such as mean, median and standard deviation to measure changes, precision and differences for comparison purposes. Kish (1995) states that standard errors are generally considerably more expensive and complicated to compute compared to other descriptive statistics for surveys undertaken, that is why they are often not presented.

Another aspect is the calculation of the confidence intervals within which the true value for the population can be reasonably expected to lie.

At the end of every survey design, all processes are documented for future reference. A Chapter on survey methodology is prepared and forms part of the report to be published.

2.113 Statistical Method Research

The Statistical Methods Unit is mandated to undertake continuous research to improve methodology and operations of statistical surveys and censuses in Statistics Botswana. Exploring various statistical methodologies and designs will in the long run improve the methodologies if done regularly. These should conform to the international standards and match the trends on the on-going surveys in the statistical globe and, as such, should be enhanced through benchmarking with other regional and international statistical agencies and attending conferences where various presentations on methodologies are made. One of the statistical methods research undertaken in Statistics Botswana is traffic count.

2.2 TRAINING AND CERTIFICATION UNIT

2.21 Introduction

The vision of Statistics Botswana states that, the organisation shall be 'a world class provider of quality official statistics and related services'. The mission also states that the organisation shall enable stakeholders formulate policies, plan and make decisions through, promotion of use of official statistics; development and maintenance of the national statistical system; provision of statistical and technical services; to mention a few.

To fully achieve and adhere to this vision and mission, Statistics Botswana has a Training and Certification Unit that ensures that all employees strive to execute a high level of competence, compatible with their level of expertise, experience and knowledge. This, in turn, will ensure that the organisation's mandate is done to the best of the ability of the workforce and is consistent with the best practices across the world. It, thus, follows that employees will require continuous training in new operational methods, techniques and skills that are suitable for their various professions within the organisation and the greater National Statistical System. It is, therefore, crucial that Statisticians, Demographers, Data Processors and Analysts and any other personnel working in the Statistics Botswana are provided with opportunities to participate in local and international workshops, training courses as well as seminars. This should be extended to the greater NSS. This will ensure that they continually enhance their professional growth which, in turn, will give the organisation a pool of grounded experts that can take it to greater heights.

Training and Certification exists as a unit responsible for:

- Planning, delivery and assessment of short-term in-house training in any one of the following statistics related operations: Applied Statistics, Field Administration, Data Collection, Data Input, Data Processing and Analysis, Use of Information and Technology Systems for Data Capture and Report–writing Skills, Portals, as well as application of Classifications;

- Training is further extended to the entire National Statistical System (NSS) to ensure credibility, compliance and application of statistical methods, standards, concepts, procedures and processes.
- Certification of quality statistics across the entire statistical value chain, within the NSS.
- Certification has also been extended to enumerators who are trained and successfully enumerate in surveys and censuses.

The Training and Certification unit comprises of two major sections (Training Section and Certification).

2.22 Training Section

The training section of the unit is charged with the mandate of facilitating both the theory and practical training programmes. The unit adheres to targeted and relevant staff training needs to realise positive outcome, impact and returns on training coupled with effective assessments.
Training usually takes place in a conducive place for learning. This also dictates that an array of training aids is available to further stimulate understanding and grasping of concepts and principles of the subject. Subject matter specialists lead training in terms of content to be taught. Training and Certification provides support in these training sessions to ensure that appropriate and up-to-date instructional methods are practised. After training, evaluation is conducted and feedback is availed in the form of a report. Training undertaken includes the following:

2.221 Training of Trainers (TOT)

In undertaking an organisational fieldwork surveys and censuses, Fieldwork and Regional Services Coordination Division in close collaboration with subject matter specialists including constant liaison with Training & Certification assemble trainers from different units, i.e. Data Management, Field Services, Sampling Methodologies and Quality Assurance to train them in order to become trainers in that particular project. The duration for Training of Trainers usually takes an average of 3 to 5 days.

In case of the Population and Housing Census (PHC) undertaking, there is a different set up, where an office is setup and is made to spearhead all census related training. In all the census stages including training, different stakeholders are engaged in the different processes.

2.222 Training of Supervisors (TOS)

Training of supervisors is usually undertaken with the aid of an Enumerators Training Manual, Supervisors Manual and Questionnaires and is conducted by members of the Technical Working Group. The length of the training differs from one survey to another and at times depends on the complexity of the questionnaire or other factors prevailing at the time.

For the PHC undertaking, TOS is usually conducted by Trainers in collaboration with Technical Officers from Statistics Botswana. Evaluation is usually done by means of a test that is set by the officers from Training and Certification Unit accompanied by a marking key. The test is written to screen those who qualify to be field work supervisors.

2.223 Training of Enumerators (TOE)

Training is conducted by trainers, officers from various directorates and supervisors, including subject matter specialists (project owners) with Training and Certification officers performing the facilitation role. Training of Enumerators is undertaken with the aid of an Enumerators Training Manual and Questionnaires and the period of training differs from one survey to the other. Evaluation and assessment is done by means of a test that is set by Training and Certification Unit accompanied by a marking key and also leads the practical observation selection role.

Training usually takes place in one week or two, depending on the content to be taught and other logistical considerations. Training of enumerators is undertaken with an aid of an enumerator training manual, cartographic material and questionnaires. The period of training differs from one survey to the other.

2.223.1 Evaluation and Assessment of Enumerators and Supervisors

After the training, the trainees are assessed and evaluated to screen for those who will be capable of undertaking the work of enumeration. This is done through three processes, of taking a theory test, undertaking a practical observation test, and finally grading and selection.

- Theory Test is done by means of a written test assessed using a marking key that is set by Training and Certification Unit.
- The other is a Practical Observation exercise, where trainees are observed for their ability to carry out practical interviews in the field. (See 2.274).
- The last one is the Grading and Selection of Enumerators and Supervisors. Selection is done by adding up marks of theory and those of Practical Observation and averaging the two to get the final mark. Marks are graded accordingly by ascending order from the highest to lowest. A cut off point is determined by Human Resources, guided by the numbers needed in the survey.

2.224 Non-Project Trainings

Ideally, any training that takes place in the organisation should be known and facilitated by Training and Certification Unit, hence the provision of training request form. This is a form that helps those who request for training to link with Training and Certification Unit in terms of what subject they want to be resourced on, by who and how the training will benefit the organisation in addressing skills gap. These trainings are in-housed.

External stakeholders such as Statistics Department of University of Botswana, schools and other stakeholders can make request for training on the National Statistical Systems (NSS), and this forms part of the non-projects training. In the NSS, training could be on the Application of Classifications, Concepts and Definitions, etc.

2.23 Certification Section

Certification is a process of confirming certain characteristics of an object, person or organisation. This confirmation is often provided but not always by some form of external review, education, assessment or audit. Accreditation is a specific organisation's process of certification.

The process of certification can be categorised into two phases. The first phase is at the micro level, where certification is done to appreciate attendance and or participation. The other category is that of certifying professionals. Professional certification is one of the most common types of certification in modern society, where a person is certified as being able to competently complete a job or task by the passing of an examination and/or the completion of a programme of study.

Other professional certification require that one obtain work experience in a related field before the certification can be awarded, i.e., data collection, data processing and analysis. Some professional certification is valid for a lifetime upon completing all certification requirements, while others expire after a certain period of time and have to be maintained with further education and/or testing. It should be noted that certification does not attest that a person has sufficient knowledge in a subject area but that he/she has passed a test.

Up to this end, certificates in Statistics Botswana are issued to those who deserve, i.e., enumerators who have successfully completed all the stages of data collection processes in a project, for example, the enumerator has gone through theory training on data collection, has been assessed using the assessment forms of training and certification, has also fully participated in the data collection. Issuing of these kinds of certificates to training that is not project inclined is being considered particularly within the NSS.

In general certification shall be awarded to all employees, temporary and permanent, who have handled or are handling duties prescribed by Statistics Botswana in a statistical environment.

2.24 Certification at Micro Level

There are a number of certificates to be awarded but some are within the jurisdiction of Human Resources Division since they are more on performance than participation, for example, Certificate of Appreciation (HR Responsibility), Certificate of Attendance (T&C responsibility), and Certificate of Participation (both HR and T&C Responsibility). This manual is confined to certificates under T&C jurisdiction only.

2.241 Certificate of Attendance

This shall be issued after completion of training of duration of a minimum of seven (7) consecutive days. However, certificates shall be silent about competencies. The certificate shall be issued within seven (7) days after end of training, and shall be mostly issued to non-projects training participants.
Participants within the NSS shall be provided with certificates by Statistics Botswana if, and only if, they have gone through the statistical standards and classification trainings.

2.242 Certificate of Participation

Certificates shall be issued to deserving candidates depending on the length of the survey/census trainings. They are issued after a training duration of at least seven (7) consecutive days.

The certificate of participation is also extended to offers who are employed on temporary basis i.e. a survey accountant whose sole responsibility is to work on financial matters of the survey for a stipulated period of time.

2.25 Certification at Macro Level

These certificates shall be issued to participants after they have been graded through competency assessments. All key players are expected to complete an assessment form and submit it for further assessment by the accrediting body, in this case, the Botswana Qualifications Authority.

In addition to the above, the organisation also has to go through the accreditation process where availability of certain facilities and operating documents including the legal framework form part of the assessment. Statistics Botswana is currently in the process of being accredited, both at individual (officers) and corporate levels.

2.251 Certificate of Appreciation

This certificate will be issued to officers who are engaged on temporary basis, either to augment processes of the project or daily exigencies of the organisation for a period of 3 months or more. This shall be done in liaison with Human Resource Department, for example a temporary survey Accountant.

2.26 Accreditation

Accreditation is the process in which certification of competency, authority, or credibility is presented. Organisations that issue credentials or certify third parties against official standards are themselves formally accredited by accreditation bodies such as Botswana Qualification Authority. The accreditation process ensures that they are competent to assess and certify third parties, behave ethically and employ suitable quality assurance. In this case, accreditation forms part of certification process. In the organisation such certification process is accorded to professionals at Diploma level and above who are normally engaged on the trainings mentioned above.

2.27 Requisition, Training and Evaluation Forms

The following forms **(see Appendix II)** are used at different stages of training processes and are;

2.271 Training Request Form

The Training Request Form is used by the person or entity that requires to be trained. It collects information about the entity/organisation, that is, name and address, number of participants, date of training, duration,

venue, source of funding - where applicable, what they need to be trained on, as well as tools/resources to be used in the training. The entities here refer to Statistics Botswana directorates conducting household surveys, any department of the University of Botswana (usually Statistics Department), Government schools and other stakeholders within the NSS.

2.272 Training Preparation Form

The form is used by a trainer or any identified resource person to scheme and prepare the content to be taught. The form is administered just before the actual training takes place and is given to training assessors, in this case T&C staff. It outlines how the training is going to be conducted based on the following three broad topics:

- Introduction,
- Presentation, and
- Summary/conclusion

2.273 Training Assessment/Observation Form

The form is used by the Training and Certification personnel (assessors/trainers) for a comprehensive assessment of trainers by the trainees as they train on a particular subject.

2.274 Survey Practical Observation Form

This form is used to assess potential enumerators/supervisors/trainers in a survey or census. The main objective of the form is to screen competencies in the following attributes:

- Introduction,
- Logical sequencing,
- Confidence,

- Clarity and thought expression, and
- General mastery of the questionnaire.

The observation on the above aspects is done during the practical training after theory has been taught. That is, after surveys/censuses trainings, be it for enumerators, data entry clerks, supervisors and training of trainers. Participants are tested on the administration of the questionnaire in the field set up as well as typing master test for data entry clerks.

2.275 Questionnaire Administration Form- Field-Follow Up

This form is used by a T&C official/assessor or any identified personnel (usually quality control officers) to observe the actual questionnaire administration on the field. The purpose of this tool is to further check impact and effect of training and further observe lapse in the acquired competency. Upon completion, information on the forms is used for writing a report on observations from the field and making recommendations for improvements.

2.3 QUALITY ASSURANCE UNIT

2.31 Introduction

This operations manual details guidelines in the form of general activities involved in the Quality Assurance Unit in Statistics Botswana. The overall purpose of the document is to be of assistance to those working in the office, particularly those who are directly responsible for the data that Statistics Botswana produces. It further outlines what really falls under the QA Unit and its functionality is clearly presented.

2.32 Background

The Quality Assurance Unit is within the Division of the Standards, Methods, Quality Assurance, Training and Certification which, in turn, falls within the Directorate of Standards, Methods and Information Systems (DSMIS).

In the previous setup (Central Statistics Office era), quality assurance was subsumed under the Surveys Unit which had its roles centred only on data collection quality control. Under Statistics Botswana, Quality Assurance has been made separate mainly because its role is not only centred on data collection but also around all the other activities or processes in the Statistical value chain.

Quality Assurance Unit covers two main areas which are Classifications and Compliance. It is the responsibility of the Unit to ensure that Statistics Botswana uses the latest versions of classifications. It does so by customising the internationally approved classifications to Botswana situation, either through adoption or adaption as may be necessary. This will at the end ensure that data produced in Statistics Botswana is comparable with data from other countries in the region and internationally. The compliance section mainly focuses on making sure that the set standards in the organisation and internationally adapted or adopted ones are adhered to.

2.33 Data Quality Assurance Framework (DQAF)

2.331 Definition of Quality Assurance

Traditionally, Quality Assurance is defined as a systematic process of ensuring that a product or service being developed meets specified requirements. The important thing is to put in place all necessary measures so that at the end of the process there is quality.

The main purpose of DQAF is to assess the quality of data; provide data quality criterion and clear procedures for designation of data as official statistics as prescribed by Statistics Act. It can also be utilised by producers of statistics to self-evaluate the quality of their data and produce quality declarations to guide users of their data. Data quality is further defined in terms of prerequisites and the dimensions of quality, namely relevance, accuracy, timeliness, accessibility, interpretability, coherence, methodological soundness and integrity. The assessment of the quality of data is done in accordance with the UN Fundamental Principles of Official Statistics.

The target outcome of QAF may be summarised as to have 'more satisfied users, using better statistics to make more informed decisions.' Evidence based decision making rely entirely on the quality of data provided. While this is so, Statistics Botswana does not have a DQAF yet but will, through the Botswana Strategy for Development of Statistics implementation, commence the development of the framework. This will address data quality issues, among others, within the organisation and the National Statistical System (NSS).

2.332 Quality Assurance Framework Overall Objective

The target outcome is to be realised through achievement of the overall QAF objective, which aims at:

- Introducing a comprehensive quality assessment program that summarises the quality of NSS data outputs for the benefit of both data producers and users; and
- Identifying quality challenges and potential improvements and then bring them to the attention of data producers for action.

The specific objectives of QAF are:

- To assess, on a rotational basis, the quality of all NSS products and of the systems that produce them;
- To identify quality problems and to highlight the major ones;
- To propose quality improvements;
- To ensure that quality problems and potential improvements are brought to the attention of senior management;
- To review the extent to which quality problems have been addressed (during the next round of assessment);
- To provide producers and users with quality summaries, including quality scores by dimension; and
- To provide quality scores enabling labelling of official statistics.

These quality assessments will be based on the dimensions of data quality. They will be done on annual or bi annual basis by the QA Unit, depending on the length the sector has been within the system. For the first few years of the assessment programme, the first four items will be given higher priority. As the programme

matures, the remaining items will be accorded equal priority or gradually added into the programme, at the same time balancing the workload of the unit.

Once the sector assessments are made, quarterly progress reports and annual reports will be prepared and presented to senior management and the Executive Committee. The results of the assessments will be presented to and discussed with data producers and key users.

2.333 Dimensions of Data Quality

ISO standard 9000:2005 defines quality as the "degree to which a set of inherent characteristics fulfils requirements". Statistics Botswana, like other Statistical Agencies, has adopted the definition of data quality in terms of "fitness for use". The quality of statistical data can be determined by the extent to which they meet user needs. Data quality is, therefore, defined as 'fitness for use'. It is further defined in terms of prerequisites of the dimensions of quality namely:

i) **Relevance -** The degree to which the statistical product meets the needs of the users in the areas of coverage, content and detail;

ii) **Accuracy -** This refers to the closeness between the estimated (sample) values from the true (population) parameter/value, that is, how correct the estimated value is.

iii) **Timeliness and Punctuality -** Timeliness is the degree to which data produced are up-to-date and available for use, as well as frequency of publications as per the set standards. Punctuality, on the other hand, refers to the time lag between the actual delivery date and the expected delivered date.

iv) **Accessibility and Clarity -** Accessibility refers to the ease with which the users are able to have the data, in a user friendly data format(s) and the supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and capacity building with respect to metadata.

v) **Comparability –** This refers to the degree to which the available data can be compared overtime with same data from other sources, regionally and globally.

vi) **Coherence -** Data derived from different sources or methods, referring to the same phenomenon, should be consistent, otherwise, the said data is incoherent.

Given the dimensions of data quality, the Quality Assurance Unit is developing 'Guidelines for Producing Quality Statistics' and these checks the above aspects at each stage of the statistical value chain. This means that at every stage of the statistical value chain, there are key indicators that are collected that are used to check if the dimensions are adhered to or not. Currently, the guidelines are undergoing editorial processes and would be implemented in due course.

2.334 Classifications

Generally a statistical classification is a way to group a set of related categories in a meaningful, systematic and standard format. The statistical classification is usually exhaustive, has mutually exclusive and well-described categories, and has either a hierarchical or a flat structure. A statistical classification usually contains codes and descriptors.

2.341 Importance of classifications

Statisticians classify data in order to facilitate the compilation, description and analysis of large sets of individual statistical observations. The process involves the systematic grouping of individual observations related, for example, to individuals, organisations, entities, objects, transactions, concepts, etc., within a much smaller set of classes or categories in accordance with actual or perceived similarities in the characteristics displayed by each individual observation. Statistical classifications are generally developed to support policymaking and because of that, to organise and present statistics.

2.342 Benefits of using Standard Classifications

i. **Data Integration:** The use of statistical classifications allows the integration of data from a variety of sources and between different periods of time.

ii. **Improved Survey Efficiency:** Using standard classifications reduces the resources and time requirements associated with many aspects of survey development and maintenance. They assist in planning for new survey developments and in reviewing existing surveys.

iii. **Meeting User Requirements:** Classifications are created in close consultation with stakeholders who are involved in developing statistical surveys.

iv. **Data Comparability:** Standard classifications ensure that data is comparable over time and between different surveys, as such they can be used to provide a link between different datasets and to compare data at regional, sub-regional and globally.

v. **Efficient Use of Resources:** The use of standard code structures simplifies data processing and means only one version of the code numbers needs to be stored. As greater use of standard classifications increases and it thus enriches the data produced by an organisation.

vi. **More Reliable Statistics:** The use of standard classifications gives reason for increased confidence in the credibility of statistics.

Within Statistics Botswana, Classifications are useful at different stages of a survey, from questionnaire design to training of supervisors and enumerators, coding of questionnaires, online editing and the analysis stage. Classifications are regularly updated to keep pace with the international community in data analysis and packaging.

2.343 List of Classifications

There are a number of classifications used by various Units within Statistics Botswana, most of which have been internationally adapted.

i. Botswana International Standard of Industrial Classification (BISIC)

It is an adapted version of the International Standard of Industrial Classification (ISIC Rev 4). The main purpose of BISIC is to provide a set of activity categories that can be utilised for the collection and reporting of statistics according to economic activities in the fields of production, employment, gross domestic product and other statistical areas. ISIC is a basic tool for studying economic phenomena, fostering international comparability of data, providing guidance for the development of national classifications and for promoting the development of sound national statistical systems. This classification is used in almost all household surveys carried out by Statistics Botswana, e.g., Botswana Aids Impact Survey, Literacy Survey, ICT survey, Agricultural Census and Surveys, Statistical Business Register, etc.

ii. Botswana Standard Classification of Occupations (BOSCO)

BOSCO has been revised in line with ISCO-08 which has been developed to facilitate international comparison of occupational statistics and to serve as a model for countries developing or revising their national occupational classifications. The purpose of the ISCO classification has been described as a tool for organizing jobs into a clearly defined set of groups according to **tasks** and duties undertaken in the jobs. It is intended for use in statistical applications and in a variety of client oriented applications. Client oriented applications include the matching of job seekers with job vacancies, the management of short or long term migration of workers between countries and the development of vocational training programmes and guidance. The classification is widely used in Statistics Botswana in almost all household surveys undertaken by the organization.

iii. International Classification of Diseases (ICD 10)

The International Classification of Diseases (ICD) is the standard diagnostic tool for epidemiology, health management and clinical purposes. This includes the analysis of the general health situation of population groups. It is used to monitor the incidence and prevalence of diseases and other health problems, providing a picture of the general health situation of countries and populations.

ICD is used by physicians, nurses, researchers, health information managers and coders, health information technology workers, policy-makers, insurers and patient organizations to classify diseases and other health problems recorded on many types of health and vital records, including death certificates and health records. In addition to enabling the storage and retrieval of diagnostic information for clinical, epidemiological and quality purposes, these records also provide the basis for the compilation of national mortality and morbidity statistics by WHO Member States. Finally, ICD10 is used for reimbursement and resource allocation decision-making by countries, including Botswana. The classification is widely used by Health Statistics Unit of Statistics Botswana housed in the Ministry of Health.

iv. Classification Of Individual Consumption by Purpose (COICOP)

The classification of individual consumption by purpose (COICOP) is a classification used to classify both individual consumption expenditure and actual individual consumption. This classification is mostly used by Price Statistics and National Accounts Unit within Statistics Botswana.

v. Central Product Classification (CPC)

Central Product Classification is intended to provide a framework for international comparison of various kinds of statistics dealing with goods, services and assets. Basically, CPC is intended to be used for different types of statistics, for example, industrial statistics and national accounts, price statistics, foreign trade statistics (including trade in services) and balance-of-payments statistics.

vi. International Standard Classification of Education (ISCED Revision 11)

ISCED is the reference classification for organising education programmes and related qualifications by education levels and fields. It is designed to serve as a framework to classify educational activities as defined in programmes and the resulting qualifications into internationally agreed categories. This is mostly used by Ministry of Education and Skills Development (MESD), Botswana Examinations Council and other related sectors. Within Statistics Botswana, Education Statistics Unit uses this classification and it is housed under MESD.

vii. Classification of Crimes

It is used for classifying criminal activities according to their severity. The classification is used by the Crime Statistics Unit within Statistics Botswana and other related stakeholders in the area of justice.

viii. Harmonised System of Codes (HSC)

The Harmonised System is an international nomenclature for the classification of products. It allows participating countries to classify traded goods on a common basis for customs purposes. At the international level, the Harmonised System (HS) for classifying goods is a six-digit code system. This is widely used by Merchandise Trade Statistics Unit of Statistics Botswana and other stakeholders dealing with trade issues.

2.335 Compliance

Within the Quality Assurance Unit, there is the compliance section which mainly deals with standards. The section's work centres on monitoring the adherence to standards by the internal and external stakeholders.

Compliance in the statistical context basically means monitoring the adherence to the set standards, nationally, regionally and globally. It is perhaps not possible to mention compliance when there are not set guidelines that have to be followed, this then brings about the question of what are these set standards that an organisation and its stakeholders have to abide by, such that at the end issues of compliance comes into

the picture. There are set standards within the organisation though uncoordinated but are internationally agreed ones of which need to be adhered to. The organisation, through the implementation of BSDS, is in the process of developing comprehensive and documented standards.

Some of the key internationally adhered to standards are the GDDS initiatives governing the production and dissemination of Consumer Price Index (CPI), National Accounts estimates, Trade Statistics and Employment figures. The periodicity and timeliness of these products are of essence to the Statistics Botswana.

2.351 Definition of Standards

A statistical standard is a set of rules used to standardise the way data are collected and statistics are produced. They provide information about data collected on a particular topic that assists in the understanding and interpretation of such data. There are however, some of the uncoordinated standards existing within.

2.352 The need for standards

Standards are used to ensure that data about the same characteristic are collected and communicated in the same way every time. This enables data from different sources to be compared on a consistent basis and enables meaningful comparisons to be made over time.

There are four key advantages from having widespread use of approved statistical standards, viz:

i. ensuring the quality of statistical outputs,
ii. creating a meaningful statistical picture of society and economy,
iii. reducing costs, and
iv. improving transparency.

The following are the internationally set statistical standards that countries have to adopt and adapt. It should be noted that these standards are structured according to the different processes of the statistical value chain and are equally applied in Statistics Botswana.

2.353 Development of Concepts, Methods, and Design

i. **Survey Planning:** an organization initiating a new survey or major revision of an existing survey must develop a written plan that sets forth a justification.

ii. **Survey Design:** A survey design must be developed, including defining the target population, designing the sampling plan, specifying the data collection instrument and methods, developing a realistic timetable and cost estimate, and selecting samples using generally accepted statistical methods.

iii. **Survey Response Rates:** Organizations must design the survey to achieve the highest practical rates of response and in line with the international standards, commensurate with the importance of survey uses, respondent burden and data collection costs, to ensure that survey results are representative of the target population so that they can be used with confidence to inform decisions.

iv. **Pretesting Survey Systems:** All components of a survey function as intended when implemented in the full-scale survey and that measurement error is controlled by conducting a pretest of the survey components or by having successfully fielded the survey components on a previous occasion.

2.354 Collection of Data

i. **Developing Sampling Frames:** Ensure that the frames for the planned sample survey or census are appropriate for the study design and are evaluated against the target population for quality.

ii. **Required Notification:** Each information collecting instrument must clearly state the reasons for collecting the planned information; the way such information is planned to be used; whether responses are voluntary

or mandatory (citing authority); the nature and extent of confidentiality to be provided (if any, citing authority); an estimate of the average respondent burden together with a request that the public direct to the agency any comments concerning the accuracy of this burden estimate; and any suggestions for reducing this burden; the OMB control number; and a statement that an agency may not conduct the interview; a person is not required to respond to an information collection request unless it displays a currently valid OMB control number.

In Statistics Botswana the questionnaire is gazzetted as a public notice and as per the Statistics Act requirement before the commencement of any survey. Other public notices are made through various media including TV, Radio broadcast, special TV or Radio programmes, as well as local newspapers.

iii. **Data Collection Methodology:** Organizations must design and administer their data collection instruments and methods in a manner that achieves the best balance between maximizing data quality and controlling measurement error while minimizing respondent burden and cost.

2.355 Processing and Editing of Data

i. **Editing:** Data must be edited appropriately, basing on available information, to mitigate or correct detectable errors. Editing manuals are prepared and editors trained before the commencement of the activity.

ii. **Nonresponse Analysis and Response Rate Calculation:** Organizations must appropriately measure, adjust for, report, and analyze unit and item nonresponse to assess their effects on data quality and to inform users. Response rates must be computed using standard formulas to measure the proportion of the eligible sample that is represented by the responding units in each study, as an indicator of potential nonresponse bias.

iii. **Coding:** Agencies must add codes to collected data to identify aspects of data quality from the collection (e.g., missing data) in order to allow users to appropriately analyze the data. Codes added to convert

information collected as text into a form that permits immediate analysis must use standardized codes, when available, to enhance comparability. Coding manuals are prepared and coders trained before the commencement of the activity.

iv. **Data Protection:** Agencies must implement safeguards throughout the production process to ensure that survey data is handled in a way that prevents unintended disclosure.

v. **Evaluation:** The quality of the data must be evaluated and the evaluation made public (through technical notes and documentation included in reports of results or through a separate report) to allow users to interpret results of analyses, and to help designers of recurring surveys to focus improvement efforts.

vi. **Developing Estimates and Projections:** Accepted theory and methods must be used when deriving direct survey-based estimates, as well as model-based estimates and projections that use survey data. Error estimates must be calculated and disseminated to support assessment of the appropriateness of the uses of the estimates or projections. Agencies must plan and implement evaluations to assess the quality of the estimates and projections.

2.356 Data Analysis

i. **Analysis and Report Writing Planning:** A plan must be developed for the analysis of survey data prior to the start of a specific analysis to ensure that statistical tests are used appropriately and that adequate resources are available to complete the analysis. Usually, Tables Specifications are prepared to guide the process.

ii. **Inference and Comparisons:** Statements of comparisons should be based on other statistical conclusions derived from survey data on acceptable statistical practice.

2.357 Review Procedures

i. **Review of Information Products:** The organization is responsible for the quality of information that they disseminate and must institute appropriate content/subject matter, statistical and methodological review procedures to comply with the Information Quality Guidelines.

2.358 Dissemination of Information Products

i. **Releasing Information:** Information which is released should be intended for the general public according to a dissemination plan that provides for equivalent, timely access to all users and provides information to the public about the agencies' dissemination policies and procedures including those related to any planned or unanticipated data revisions.

ii. **Data Protection and Disclosure Avoidance for Dissemination:** When releasing information products, the organization must ensure strict compliance with any confidentiality pledge to the respondents and all applicable legislation and regulations. Data are usually anonymized for this reason.

iii. **Survey Documentation:** Survey documentation should include those materials necessary to understand how to properly analyze data from each survey, as well as the information necessary to replicate and evaluate each survey's results.

iv. **Documentation and Release of Public-Use Micro data:** Micro data that a released to the public must include documentation clearly describing how the information is constructed and provide the metadata necessary for users to access and manipulate the data.

2.359 Statistical Compendium

With the development of the Botswana Strategy for the Development of Statistics (BSDS) within Statistics Botswana, the organisation has found the need to develop a compendium of statistical concepts and definitions as an effective tool that facilitates harmonisation, comparability and use of standard statistical concepts and definitions. The Quality Assurance Unit has published the first issue of the compendium but the expectation is that it will continuously be updated as the NSS stabilises. The main benefit of having a Statistical Compendium is that it ensures the standardisation of statistical terms for internal and external stake holders. This makes it easier for users to interpret the statistical concepts. Currently, the unit is compiling the compendium of classifications.

2.36 Quality Assessment Reports

Quality Assurance Unit has a role of conducting quality assessments on all the statistics products in order for them to be designated as official statistics as prescribed by Statistics Act of 2009. After the production of a statistical report/output, all stages of the statistical value chain are audited to ensure their compliance to the set standards. Each stage of the statistical value chain is checked against the standards which apply to it.

The assessment is done following the Data Quality Assurance Framework, the various quality aspects of the entire statistical value chain (i.e. need, design, collection, processing, analysis and dissemination), and certifies national statistics on one of the four assessment levels; where level 4 certification (quality statistics) indicates optimal conditions for statistical production, while Level 1 (poor statistics) indicates the least favourable conditions. At the end of the assessment, which is done by the Quality Assurance team, a report is produced on the entire quality of the product and it is at that final stage that the statistics product can qualify either as official statistics product or there is need for improvement.

2.37 Quality Management System

Quality assurance Unit addresses all issues pertaining to the quality of data and also the quality aspects of all other processes that are not data related but are within the organisation through quality management system. As a result, the Unit is in the process of developing a Quality Management System (QMS) through the BSDS initiative, whose main objective is to enable the organisation to achieve the goals and objectives set out in its policy and strategy. It provides consistency and satisfaction in terms of methods, materials, equipment, etc., and interacts with all activities of the organisation, beginning with the identification of customer requirements and ending with their satisfaction, at every transaction interface.

A QMS integrates the various internal processes within the organization and intends to provide a **process approach** for project execution. A Process Based QMS enables organizations to identify measure, control and improve the various core business processes that will ultimately lead to improved business performance.

A complete ISO 9001 Quality Management System must address all the **requirements of ISO 9001** and, more specifically, must meet the ISO 9001 Documentation Requirements.

The Concept of QMS

A Quality Management System in its basic concept seeks to:

- Recognise the external quality related requirements specified in Licenses to Trade, guidelines, specified customer requirements, and the chosen management system standard(s);
- Ensure that all requirements have been documented within the management system in the appropriate location in terms of defined specific system requirements;
- Confirm that employees receive applicable training in the quality system requirements;
- Outline performance processes, where applicable, to the quality system requirements;
- Produce records or evidence that system requirements have been met;
- Measure, monitor and report the extent of compliance with these performance procedures;

- Continually monitor and analyse changes to the requirements and confirm that all changes are reflected in changes to the specific requirements when necessary;
- Execute the audit and analyse the system processes and correct them when necessary;
- Include processes that will help continually improve the quality system.

In general, QMS can be thought of as a system that controls the quality levels in an organisation.

CHAPTER 3: FIELDWORK AND REGIONAL SERVICES COORDINATION (FRSC)

3.1 Introduction

This operations manual provides insight into the mandate and operations of the Division of Fieldwork and Regional Services Coordination (FRSC). The FRSC is a division of the Department of Standards Methods, and Information Systems. The manual is an internal document that provides processes and procedures of collecting statistical data from the different sources such as administrative records from institutions, business entities, prices outlets and households.

The manual is structured according to the various data sets emanating from the different surveys that are conducted to fulfil the operations that will ensure that Statistics Botswana delivers its mandate; which is, provision of official statistics for evidence based policy and program formulations, as well as monitoring and evaluation of development and intervention programmes. The definitions, processes and procedures stated in the manual shall assist the field staff in understanding the objectives of the different surveys and provide guidance to ensure that the data collected is of good quality and is delivered on time.

3.2 Background

Statistics Botswana is the principal data collecting, processing and disseminating agency responsible for coordinating, monitoring and supervising the National Statistical System. It, thus, has the statutory mandate to produce and provide the Government, private sectors, NGOs, parastatal, international organisations, civil society and the general public with statistical information for decision-making, policy formulation and planning purposes. The statutory mandate also includes the responsibility of providing advisory and technical service to all users on statistical matters.

Being a perennial source of data for policy and programme formulation, monitoring and evaluation, Statistics Botswana is mandated to conduct on regular basis, business surveys, household surveys, censuses of population and housing, censuses of agriculture as well as obtaining data from administrative records. All

of these activities are undertaken to compile statistical information that provides continuum socio economic and demographic data required for evidence-based decision making. Under the Directorate of Standards, Methods and Information Systems (DSMIS), Statistics Botswana has in its functional organisational structure, the Division of Fieldwork and Regional Services Coordination (FRSC). The FRSC is responsible for data collection operations covering prices, business surveys, administrative records, household surveys and censuses. The Division has two (2) units which are: Regional Coordination-South based in Gaborone and Regional Coordination-North based in Francistown. The division is headed by a Principal Statistician, while the units are headed by Senior Statisticians, referred to as Regional Coordinators. Below the head of the unit are the field supervisors at the Statistician level, followed in hierarchical order by enumerators.

For administrative and operational efficiency, the country has accordingly been divided into two regions aligned with the two main regional offices of Gaborone and Francistown.

3.21 Regional Sub-divisions:

For field operations on routine data collection of prices, business surveys and administrative data, the regions have been aligned with census districts as shown below.

a) The South Region:

i. Gaborone – Headquarters for the South Region
ii. Lobatse
iii. Jwaneng
iv. Ngwaketse East
v. Ngwaketse West
vi. South East
vii. Kweneng East
viii. Kweneng West

ix. Barolong
x. Kgatleng
xi. Central Mahalapye
xii. Kgalagadi South
xiii. Kgalagadi North
xiv. Ghanzi

b) The North Region:

i. Francistown – Headquarters for the North Region
ii. Selebi-Phikwe
iii. Orapa
iv. Sowa
v. Central Bobonong
vi. Central Serowe-Palapye
vii. Central Boteti
viii. Central Tutume
ix. North East
x. Ngamiland East
xi. Ngamiland West
xii. Chobe

At the transformation of Central Statistics Office to Statistics Botswana in 2012, a recommendation was made for opening new small and medium offices in order to facilitate efficient data collection. The offices were proposed to be opened at the following places: Kanye, Tsabong, Ghanzi, Palapye, Kasane and Maun, and would be coordinated at the two regional offices of Gaborone and Francistown, respectively.

3.22 Data Collection Methods:

It is imperative to keep in mind that statistical information is derived from different sources through the following methods of data collection:

i. Direct interviews (face-to-face or canvas),
ii. Telephone interview,
iii. Mail questionnaires (with telephone or email follow-ups). This is also referred as self-fill questionnaire,
iv. Transcribing data into questionnaires from administrative records,
v. CD and Email, and
vi. Electronically interfaced systems.

Traditionally, the direct interview, mail questionnaire and transcription have been widely used in Botswana, and still are used. However, with modern technology, the last two methods of data collection, CD and email and electronically interfaced systems, are also becoming prominent in Statistics Botswana, though at a slow pace.

Call-in visits and telephone appointments are made when collecting data. The sources of the data are retail outlets, business establishments, commercial farms, government and non-government institutions and households as well.

Where questionnaires are sent by mail responses from establishments remain low, and follow-up visits method is used to collect the data on non-response. The desire is for each survey to have at least 75 percent response rate as per international standards. However, depending on the capacity of Statistics Botswana, the response rate of 50 percent is still acceptable. The canvas method of data collection is used for all censuses and household based surveys.

Note: The two basic problems with self-fill questionnaire are that a respondent might not understand the questions or, even if he/she does, he/she might not feel inclined to answer. However, the advantages of

the self-fill approach are that they are cheaper, quicker and less demanding upon manpower as such a questionnaire is supposed to be simple and short.

Noteworthy is that the data sources, particularly business establishments are so engaged with their activities that they have little time to give attention to statistical enquiry. Therefore, use of telephone, email or physically visiting the establishments to intensify follow-ups becomes necessary. Hence it is imperative for resources such as telephones, emails and transport to be available to field staff, both the enumerators and field supervisors.

The data collection takes place on the three (3) broad areas which are; Prices collection; follow-ups on business surveys, data collection from administrative records such as Health facilities, transport permits, commercial farms; and household surveys and censuses.

3.3 Prices Collection

Prices are collected every month from selected representative retail outlets in selected towns (including the two cities) and villages throughout the country. During the Central Statistics era, the sample was a purposive one, dictated largely by the availability of resources and the need for monthly data collection so that Consumer Price Index (CPI) was published every month. However, with the 2012 transformation of CSO into Statistics Botswana, a decision was taken to follow sampling procedures. This initiative is implemented in the CPI rebasing project which is in progress.

Prices collection is carried out during the first two weeks of each month in all outlets in the sample. All types of shops are represented in the price collection exercise – large, medium and small. In all, about 1292 different retail outlets are visited each month. Only cash purchases are taken into account. Discounts, sales prices and free-market prices for items, which also have official prices, are used whenever they occur.

The prices collection exercise also includes quarterly data collection from selected households in the cities of Gaborone and Francistown. The quarterly prices include wages of domestic workers and rentals as well as input values of construction works. The latter is collected from a sample of construction companies.

3.31 Editing the Completed Questionnaire

When completed questionnaires arrive from the field, they are visually inspected for consistency checks. This is mainly a comparison of last month's prices with the current month. If a price change is more than the figure that is reasonably expected to be, and there are no written comments from the price collector, the price collector is asked for an explanation, including going back to the field to confirm the figure as may be necessary or a decision made as to whether to allow this price quote to stand as is or whether the previous month's price should be carried forward instead. The procedure could be summarised as follows:

i. Check for mistakes, e.g., prices written on wrong slots,

ii. Price too high or too low compared to last month price, and

iii. Avoid too many zero prices if possible.

As with all forms of data collection, price collectors encounter problems in the field. Occasionally when the price outlet owner is away, the shop assistants may not be willing to give field staff permission to take prices. This happens mostly in small shopping outlets, bars and night clubs.

The purpose of collecting prices data on monthly basis is to compute the Consumer Price Index (CPI). The index essentially measures price changes of commodities from month to month. In order for the index to produce the required results we need to price the same item every month. Questionnaires contain clearly defined item descriptions, which remain unchanged. For example, the price of a 12.5kg bag of maize meal will be monitored continuously over time. However, if an item disappears from the market a similar replacement item is selected and the appropriate adjustment is made to the base price in office.

If an outlet (a shop for example) is closed for the first time, it will not be deleted immediately. It will take three months using the previous month prices, because it might be opened again by a different or the same operator in a short time.

3.4 Business Surveys Data Collection and Follow-Ups

Business surveys are conducted on quarterly and annual bases to measure the performance of the economy. The base for the business survey is the Statistical Business Register (SBR) which is internationally recognised and this replaces the Enterprise and Establishment Register (EER). It contains listing of all registered business entities as well as government institutions such as councils for employment data.

The register is segmented into large and small establishments, where large establishments/companies comprise at least 50 employees while small establishments refer to those with less than 50 employees.

The large companies are all surveyed while the small ones are sampled for the different survey types.

The business surveys that are conducted include the following:

i. Employment Survey (ES) or Survey of Employment and Employees
ii. Survey of Industrial Production (SIP)
iii. Survey of Recent Trends (SRT)
iv. Annual Economic Survey (AES)
v. Information and Communication Technology
vi. Annual Agricultural Surveys (traditional and commercial farming)

The ES, SRT and SIP are quarterly surveys while the annual agricultural and economic surveys, as implied, are conducted every year.

Establishments are given one month to respond to the initial enquiry. After a month of dispatch of questionnaires, reminders are sent together with another questionnaire. In a period of 2 - 3 weeks after the dispatch of the reminders, physical and telephone follow-ups are undertaken.

The sampled establishments are the small organisations and are the more difficult to follow-up and obtain replies from. However, for the results to be representative and hence meaningful, it is imperative that a reasonable sample of respondents in all industries, where possible, is obtained.

3.41 Employment Surveys (ES)

Employment Surveys are surveys of Employers who are sent questionnaires through the mail asking details of their employment and wages. The aim is to obtain accurate estimates of total formal sector employment and earnings at regular intervals. The major target is paid employment but questions are asked about working proprietors (the self-employed). Informal sector data is not covered in the business surveys but it is collected through household surveys.

Employment Surveys provide information on the size and movement in the numbers of persons in paid employment and is a major economic indicator for Government and other planners. Unemployment is generally more difficult and expensive to measure than employment and thus movements in formal employment are also taken as the major indicators of likely changes in unemployment.

Wages and salaries are the major data item collected besides number of employees and are also of major policy interest. The level and movement of average earnings is a major policy item as well. The classification of average earnings by citizen/non-citizen is also of major use for policy purposes. The use of employment data includes setting minimum wages, interest rates, rentals for Botswana Housing Corporation (BHC) houses, and taxation levels. The government and private sector are major users.

Internally within Statistics Botswana the employment and wages figures are an input into the National Accounts which is the major economic data summary for the country.

3.42 Survey of Industrial Production (SIP)

The Survey of Industrial Production (SIP) collects quarterly data from enterprises and establishments about production of goods and services in quantity and value. The enterprises and establishments are defined by geographical location.

An enterprise is defined as an institutional unit or a combination of institutional units that enclose all functions required to, directly or indirectly, control all functions required to carry out production activities.

An establishment is a technical unit in which production activities are carried out in a defined geographical location, and for which activities, data are available or can be meaningfully derived and compiled. This, thus, means that an establishment must, at the least, be practising some minimal records keeping of its activities. An establishment can engage in one predominant economic activity and also include other kinds of economic activity of secondary or tertiary status.

An industry is a category of establishments that produce a similar type of goods and services.

The data collected through SIP is used to measure economic performance in terms of output by different industries. The survey also identifies and measures the potential production of the existing enterprises and establishments.

Derived also from the SIP is data on employment size as well as gross remuneration in terms of wages, salaries, overtime, rations, housing benefits, gratuities and leave pays. These data are categorised by type of workers (proprietors and employees).

3.43 Survey of Recent Trends (SRT)

The Survey of Recent Trends (SRT) data is collected monthly as reference time from private businesses and public institutions in the country. The SRT data forms the source of quarterly GDP estimates by sector which

provides information for statistical and economic analysis of industries. The data collected is on the turnover and intermediate costs of large enterprises and establishments engaged in the activities such as building construction; trade, hotels and restaurants; roads construction, rail, air transport, posts and telecommunications, other services of transport, banks, insurance, real estate, business services; education, health, recreation, and other social and personal services that are market and non-market oriented.

Industries not covered by SRT:

The SRT coverage is not extended to cover water and electricity activities because it was expected that the existing information would be adequate. However, the only information collected on water and electricity is on production in quantity and values.

Quarterly data for general government production are available and obtained from the budgets, as are taxes and subsidies on production. Similarly, data on agricultural activities, particularly livestock and crops are obtained from the records of the Ministry of Agriculture and the Agricultural Statistics Unit of Statistics Botswana. These data sets are used for the annual estimates.

3.44 Annual Economic Surveys

The annual Economic Survey (AES), previously known as Census of Production and Distribution (CPD) covers industries such as Agriculture, Mining and Quarrying, Electricity and Water, Construction, Agriculture (freehold farming only), Wholesale and Retail Trade, Transport and Communication, Finance and Business Services and Community and Personal Services. Quite a number of these are in the informal sector, and some would be small retail outlets which do not keep any records of business transactions. The AES is conducted every two years and traditionally draws its sample from the Enterprise and Establishment Register (EER), the latter which is now referred to as Statistical Business Register (SBR). The EAS data is used, among others, for rebasing the constant price estimates to more current period for quarterly national accounts (QNA).

3.45 Information and Communication Technology (ICT):

With the millennium advent of the ICT equipment and infrastructure, demand for access and usage of ICT as a means of modern communication has risen dramatically since the year 2000. As a result, there has been a commensurate demand for data and information that are required to measure access, usage and guide policy decision making in the area of ICT. The ICT data is collected and released on quarterly and yearly basis from the relevant institutions and business entities, e.g., Botswana Telecommunications Corporation (BTC), Botswana Communications Regulatory Authority (BOCRA), Botswana Postal Services, media houses and households. Data from the latter is obtained through household survey and subscription. Statistics obtained from these ICT outlets include, among others, access, usage and traffic volumes.

3.46 Annual Agricultural Surveys

Annual agricultural surveys are a component of a continuous programme of household surveys administered by Statistics Botswana. This component of the programme of household surveys is specific to the agricultural sector aimed at establishing trends in agricultural production and mode of operations. The main objective of the programme is to provide time series data on basic information about crop production and livestock population and general agricultural data.

Agricultural Censuses are conducted every ten (10) years and are aimed at providing benchmarking data on agricultural sector for policy formulation, programmes and projects monitoring and evaluation.

The processes and preparations for annual surveys are similar to those of agricultural censuses except that the scope is broader on coverage objectives.

Data is collected through two sub-components and these are:

i. the traditional sector or subsistence farming
ii. the commercial sector

a. Traditional Sector

The data from the traditional sector is collected through a questionnaire administered by enumerators. The surveys/censuses cover agricultural holdings and their principal characteristics such as:

- Demographic characteristics of the holder and holding (age, sex, marital status, education, economic activity etc.);
- System of land tenure,
- Farming practices,
- Area planted, harvested and crop production by type,
- Livestock population and production by type (Cattle, Goats, Sheep, etc.),
- Sale of livestock and crop harvests,
- Other Livestock Population (Poultry, Donkeys/Mules, Horses, Dogs, others),
- Availability of farm labour,
- Water supply, farm equipment and machinery inventory, and
- Farm enterprises.

b. Commercial Sector

The data from the commercial sector is collected through a mailed questionnaire administered by farmers themselves. Follow-ups are usually made to those farmers who did not respond to the questionnaire. The sector covers basically information similar to the one collected under the traditional sector except that additional information for some variables is collected such as:

i. Cash crops (beans/pulses, sunflower, groundnuts, wheat, cotton, Lucerne etc.).
ii. Vegetables (potatoes, tomatoes, onions, cabbage, spinach, carrots etc.).
iii. Fruits (oranges, naartjies, lemons, grapefruit, bananas, watermelons, green mealies, etc.).
iv. Dairy production, etc.

3.461 Survey preparations

The preparation cycle of the annual survey starts as soon as the current survey comes to an end. Basically the following plan is followed in carrying out the annual agricultural surveys as well as decennial censuses (every ten years).

3.462 Annual Agricultural Surveys and Decennial Censuses Design

There are two different frames used for selecting the sample during the annual agricultural surveys/decennial censuses. The first frame is the Enumeration Area Sampling Frame (EASF), which consists of a list of enumeration areas (or simply blocks in agricultural statistics terms) constructed in all the villages, land areas and the cattle post areas. The agricultural sampling frame is constructed from enumeration areas (EAs) with agricultural holdings as identified and enumerated during the previous population and housing census. This excludes enumeration areas for cities, towns, commercial farms, camps, Tribal Grazing Land Policy (TGLP) farms, and constitutes a master sample for agricultural statistics from which a sample for the agricultural census can be selected. Similarly, the agricultural census sampling frame becomes a master sample from which an annual agricultural survey sampling frame is selected.

The second frame consists of a list of commercial farms and is obtained from the Statistical Business Register database and supplemented by other lists of farms during follow-ups made on annual agricultural surveys by the Agricultural Statistics Unit (ASU) and Fieldwork and Regional Services Coordination Units of Gaborone and Francistown.

The annual survey/decennial census is divided into three phases in which questionnaires for traditional farmers are administered by enumerators while commercial farmers complete mail questionnaires. Phases I and II are collection from traditional practice while phase III is of commercial practice. Follow-ups are made due to high non-responses by commercial farmers and, in fact, the response rate from mail questionnaires is normally low and responses are, therefore, mainly obtained from intensive follow-ups which are conducted between August and December months of the year.

3.463 Questionnaires

There are four sets of questionnaires used during agricultural surveys/censuses, namely, Forms I, II, III and IV. The first three forms (questionnaires) are used for the traditional sector while the fourth one is for the commercial sector.

Form I is used for identification of Agricultural Holders where all dwellings within each selected enumeration area in a particular agricultural district are listed. Enumerators are expected to identify the boundary of the enumeration area (block) before listing. It is in this form that agricultural holders are identified and selected. Agricultural holders selected from this form are interviewed during the administration of Forms II and III. For a household to qualify as an agricultural holding, it has to have the following characteristics:

- A piece of land allocated for ploughing purposes, whether planted or fallow and/or
- At least one cow, goat, sheep, donkey, mule, horse or pig.

Form II is administered to collect data on:

- Characteristics of the agricultural holder and other members of his/her holding (age, sex, marital status, education levels, economic activity etc.);
- Land use in terms of crop and fallow land and measurements of such; and
- Farming practices.

Form III is administered to collect data on:

- Livestock population (cattle, goats, sheep, donkeys, horses, pigs etc.),
- Livestock (cattle, goats and sheep) production (births, deaths, sales, home consumption, purchases and off take),
- Livestock water supply,
- Crop production and sales,

- Farming implements and equipment, and machinery inventory,
- Farm Enterprises (economic section and farm labour).

Form IV is administered to farmers through self-enumeration method of data collection. Forms are sent to farmers to complete. The follow-ups are usually made as a result of high non-response rates. Form IV questionnaire collects the same information that is in Form III except that it includes dairy cattle, cash crops, fruits and vegetables. This information is collected from commercial farmers only.

3.464 Fieldwork

Training of enumerators is conducted once a year at the beginning of the agricultural season and it usually takes place during the first two weeks of January. There are eleven (11) Agricultural Statistics Unit (ASU) teams within the six (6) agricultural regions (Southern, Gaborone, Central, Francistown, Maun and Western). The teams cover the twenty six (26) agricultural districts. Each team is allocated a vehicle for ease of operation and consists of a driver, a supervisor and three enumerators.

The following table gives a breakdown of teams by region and districts.

Region	Number of Teams
Southern (5 districts)	2
Gaborone (5 districts)	3
Central (7 districts)	3
Francistown (3 districts)	1
Maun (3 districts)	1
Western (3 districts)	1
Total number of teams	**11 (44 field staff and 11 drivers = 55)**

Although each team has a vehicle, the fieldwork is normally interrupted and delayed by continuous vehicle breakdowns due to bad terrains.

3.465 Data Collection Phases

Data collection is done in three phases. The first two phases are confined to the traditional sector, while the third phase is confined to the commercial sector and follow ups. The first phase runs from January to June, while the second phase runs from June to September. The third phase runs from October to December.

Phase I deals with the administration of Forms I and II, while **Phase II** deals with the administration of **Form III: Phase III** deals with the administration of **Form IV** by farmers. Forms are dispatched to commercial farmers in August and follow ups are usually done from October to December of the current survey/census year.

3.466 Data Quality Control Checks

Data quality control checks are done by Technical Officers who are based at the Head Office in the Ministry of Agriculture. Scheduled trips to visit teams at intervals of three months are undertaken by the Technical Officers. However, for agricultural censuses quality control is done collectively by teams from the Ministry of Agriculture and Statistics Botswana.

3.5 Data from Administrative Records

Administrative records data is collected from government institutions such as schools, health facilities, transport offices, Immigration forms (cross border movement), parastatal organisations as well as annual agricultural surveys (commercial farming only). The data is collected on annual basis, save for schools data which is collected about first term data and later the overall data collected after all enrolments of the first have been completed.

The administrative records data covers the following:

3.51 Education

Data on enrolments, progress, number of teachers, school facilities are collected from pre-school, primary and secondary schools, as well as tertiary institutions. Education data collection has, however, been taken by the Ministry of Education and Skills Development since 2013 as a major user, partner and collaborating stakeholder.

3.52 Health Facilities

Data are collected from clinics (government owned and private) and hospitals. The data covers outpatients and in-patients visits illness and treatment, as well as midnight census (visits at night due to emergency). The data also covers information of health facilities such as number of doctors, nurses, hospital beds, maternity wings and others.

3.53 Transport

Data covers passenger and haulage permits, goods traffic by volume and passenger traffic by road and air. Goods and passenger traffic data collection has, since 2013, been left to be collected electronically by the subject matter units and divisions. The remaining data is on transport permits for passenger and haulage.

3.54 Tourism

Data on tourism includes cross border movements of people at the border posts. The immigration forms are collected from the major border posts as well as medium and small ones. The forms are sorted by entry point, date of entry and month. The sorted forms are delivered to Data Management Unit for data capture and processing.

3.6 Household Survey and Census Projects

Household surveys became a programme feature of Statistics Botswana, the former Central Statistics Office (CSO), in 1983. The institution of a programme of household surveys was a national response to the United Nations global household survey programme, the aim of which, at inception, was and still is, to develop sustainable household survey capability programme in developing countries. Prior to this development, household surveys in Statistics Botswana (the former CSO) were conducted on an ad-hoc basis, with little or no conscious effort for capability and capacity building.

The Household Surveys Unit was established within the general philosophy of the African Household Survey Capability Programme (AHSCP). The Unit started operations on a continuous basis in 1983, as opposed to before then, when household surveys were planned and conducted on an ad-hoc basis.

The planning of a household survey is an important task, since the quality of the survey results depends considerably on the preparations made before the survey is conducted. The amount of planning needed varies greatly with the type of literature material available from past similar surveys and the nature of the information sought to address the objectives of the survey.

The design and execution of household surveys involve a wide range of technical and operational activities including the following:

- Decision on the survey to be undertaken as per schedule of household surveys,
- Project proposal and budgeting,
- Establishing the survey administrative structure (personnel and reporting levels),
- The formulation of the objectives of the survey and major variables,
- Determination of specification of the output and the selection of topics as guided by major variables,
- Decision on the domains of estimation – national, sub-national, urban-rural, major settlements, or all of these,
- Designing tabulation plan,

- Designing the questionnaire(s),
- Develop enumerators manuals,
- Developing systems for data processing, capture and validation,
- Developing manual editing and coding manuals,
- Determining the frame and exclusions if need be,
- Designing the sample,
- Budget estimate and resource mobilisation,
- Preparing project memorandum,
- Drawing work-plan reflecting activities and major milestones,
- Pre-testing and finalising the design of questionnaire(s) and the preparing manuals and instructions,
- Determining numbers of field staff and transporting requirement,
- Designing quality assurance procedures for training during fieldwork and data processing,
- Gazetting of the questionnaire(s),
- Procurement of fieldwork materials and supplies,
- Procurement of transport,
- Designing publicity programme,
- Preparing procedures and processes for distribution of fieldwork materials and supplies, and transport,
- Recruitment and training of field staff,
- Data collection fieldwork,
- Supervision of data collection,
- Quality assurance,
- Tabulation and production of preliminary tables,
- The analysis and interpretation of results,
- Report writing, and
- Dissemination of the results (printing report, website, CD rom, dissemination seminar, etc.)

It is generally not possible to give rules of thumb regarding the actions to be taken at the different stages of the survey since these would depend very much on (i) the data requirements, (ii) available resources and (iii) the operational conditions of individual surveys.

The important aspects requiring attention at planning and operating stages of a sample survey are discussed in sequence in the following sections:

3.61 Objectives of the Survey

The first task when designing a household survey is to lay down its objectives and put them down in writing. The objectives should be precise, for example, it is not enough to say that the survey is intended to find out about the living conditions of old persons. One should exactly define what is meant by the living conditions and what is meant by old person. Failure to think of the survey objectives fully and precisely must inevitably undermine its ultimate value. Once the objectives are settled, the plan of the survey is directed to achieving them with the required accuracy and within the given resources and time period.

The objectives of the survey should define (i) the population the survey intends to cover, (ii) data relevant to the purpose of the survey to be collected, and (iii) Geographic coverage.

3.62 Justification

The justification of the survey should be established in close consultation with planners and the major data user (stakeholders) of the survey.

i. If it is relevant to collect the existing basic informative data related to survey;
ii. Review existing data and literature with respect to
- Data availability (gaps and omission if any);
- The extent to which existing data has been utilised and its major constraints;

iii. The extent to which the output of the survey overcomes the above constraints;
iv. Relationship with other studies.

3.63 Variables of Study

Plan the Subject-Matter Content of the survey, for instance:

i. Select the variables (topics) to be studied by the survey,
ii. Consider the concepts to be measured, observed or reported and their theoretical basis if any. The concepts should be as simple and unambiguous as possible and meaningful and understandable to the population being studied, and
iii. Consider the level of detail and depth of information required.

3.64 Period of Data Collection

i. Determine the period of data collection by sampling units (village, households, etc.). From a study conducted in one season it is not possible to make generalisation about events taking place throughout the years (if seasonal variation exists). For example, to minimise any seasonality variation the data collection for Household Income and Expenditure Survey (HIES) and Botswana Labour Force Survey (BLFS) are for 12 months period;
ii. Determine the method of interview and data collection (whether personal visits or mail questionnaire, etc.);
iii. Determine the general approach to be used in covering the population i.e. whether de-jure or de-facto approach is to be used).

3.65 Output Design

Specify the output of the survey. Outline the Tabulation Plan:

i. List the tables required (keeping publications in mind!),
ii. Indicate the table formats, and
iii. Give the main classificatory variables.

For example in most of the national household surveys the main classificatory variables considered are

i. **Relationship with Head of Household:** Mother, father, son, daughter etc.
ii. **Sex:** Male and Female
iii. **Age-groups**: Age groups at interval of 5 years each (viz. <1, 1-4, 5-9, 10-14…)
iv. **Place of Birth**
v. **Place of Residence**: (i) Cities/Towns, (ii) Urban Villages and (iii) Rural
vi. **Education**: (i) Pre-Primary, (ii) Primary (ii) Secondary, and (iii) Tertiary
vii. **Training**:
viii. **Economic Activities**: As per 2001 Census Categories
ix. **Marital Status**: (i) Never Married, (ii) Married (iii) Living Together , (iv) Divorced, and (v) Separated
x. **Geographical Area**: Administrative Districts/Sub-Districts
xi. **Parents Survival**
xii. **Language**
xiii. **Religion**

Extra classificatory variables may be added as and when needed.
Determine the Levels of Estimation i.e. give list of tables by:

i. Stratum
ii. Cities
iii. District
iv. Region
v. Country

Specify the types of tabulation and statistical analysis intended, i.e. whether

i. Totals
ii. Proportions

iii. Ratios
iv. Measures of location (means, median, mode, etc.),
v. Measures of dispersion (variance, standard error, coefficient of variation, range),
vi. Frequency distributions,
vii. Associations (correlation, regression); multivariate analysis,
viii. Trend and time-series analysis etc., and
ix. Multivariate analysis

3.66 Work-Plan

Drawing up a realistic plan of all the survey activities from the planning stage until the report preparation and dissemination, plays an important role in close supervision of survey activities by the survey manager.

3.67 Survey Activities

i. List the survey activities and determine their timings,
ii. Schedule of preparatory activities,
iii. Appointment of staff (if any),
iv. Training of field personnel,
v. Training of office personnel,
vi. Schedule of data collection,
vii. Schedule of data processing, and
viii. Schedule of report preparation and dissemination of survey results.

If some activities are delayed because of reasons beyond control, it will affect other activities too. Thus, it is recommended that work-plan should be updated from time to time and remedial measures noted for their implementation.

3.68 Manpower

The minimum manpower requirements for the survey must be established. Manpower requirements should include the number of:

i. Professional and technical staff,
ii. Field personnel (a) enumerators (b) supervisors, and (c) drivers,
iii. Data processing personnel (a) editors (b) coders, and (c) data entry operators.

3.69 Administrative Structure

Establish the administrative structure of the survey.

i. Organisational Chart

(a) Linking the survey programme to a specific Government department or agency that is the major user of the data,
(b) Linking the various units (sections) created for the design and execution of the survey; viz. (i) survey operational section, and (ii) analysis and planning section etc.

ii. Manpower allocation, specifying job description line of responsibilities and reporting.

3.610 Budgeting

The activity of budgeting lies with the subject matter specialists in consultation with fieldwork and data processing service providers to ensure inclusion of budget provision of their activities.

Preparation of a preliminary budget estimate is a priority activity that should be planned and executed at an early stage. The draft budget shall be based on the assumption about (i) the number of household (or

other elements of the population as the case may be) to be covered, and (ii) the time needed to interview a household.

The major items for budget estimate include:
1. Salaries (depends upon the number of field staff, viz.: Enumerators Supervisors) and drivers,
2. Transport cost,
3. Training,
4. Equipment and supplies,
5. Fuel, maintenance and insurance,
6. Printing of questionnaires and manuals,
7. Pretesting,
8. Data collection,
9. Data processing software,
10. Publicity cost,
11. Consulting fee (data analysis and report writing),
12. Dissemination seminar,
13. Archiving, and
14. Miscellaneous activities.

This will be an estimate since the details of some costs will be unknown and as such will be examined or revised as the survey activities progress. Once the budget is drafted, the funds required should be found.

Remark: (a) A good way to start a budget is to look at the budget of similar surveys already done in the country. (b) VAT and Contingency costs should be added.

3.611 Questionnaire Design

Questionnaire design is the responsibility of subject matter specialist in consultation with all key players. Once the objectives and the tabulation plan have been prepared, a relevant questionnaire is developed. The

information collected through a questionnaire ultimately depends on what questions are asked and how they are asked. One of the most difficult parts in survey preparation is the translation of complex ideas and concepts into usable questions for survey purposes. As is often the case in translating from one language to another, something is frequently lost in the process. Questionnaires which are mailed to respondents should be attractive and simple in order to increase the response rate.

There is a great deal of experience in questionnaire design and most of the specific issues and problems are well known. Nevertheless, questionnaire design is more of an art than a science. One has to try to optimise the requirements of three categories of audiences (i.e. respondents, the enumerators, and the data processors).

The way in which a questionnaire is designed will have a crucial influence on the accuracy of the responses obtained and the quality of the data collected and statistical information derived through analysis. As such all key players in this regard should form part of the structure designing the questionnaire, i.e., system developers, sampling experts, analysts, data collectors and processors experts.

3.612 Pre-test and Piloting

No household survey questionnaire should be finalised without being tried out on a small number of households. Pre-test involves administering a questionnaire to selected few households or trying out selected sections of the questionnaire.

Piloting is a comprehensive field test of a draft questionnaire which covers more households. The households should belong to several households that represent the population of interest. For example, for surveys which intend to cover both urban and rural, the pilot should be done in different parts of the country so that it covers people very much similar to those to be sampled.

Pilot test serves many purposes but the main ones are given below:

i. To check the adequacy of the questionnaire. This is probably the most valuable function of the pilot and pre-test survey;
ii. To test the editing and tabulation programs;
iii. To check the efficiency of the survey procedures, i.e., manuals, instructions; and
iv. To mark the time to complete a questionnaire, thus, determining data collection period.

All members of the survey team should be involved in the pilot testing process. Usually pre-test/pilot is conducted by a team which designed the questionnaire and the field supervisors. After the pre-test/piloting, the survey team should hold meetings to report their findings and revise the questionnaires and survey procedures.

3.613 Gazetting

The main purpose of gazetting a questionnaire is to commission the study under the Law, in this case **Statistics Act 2009 section 39.** There are clauses that give the Statistician General powers to conduct any survey. The Act also protects the organisation conducting the survey, its staff, enumerators and respondents taking part in the survey. The gazetting processes are spearheaded by the subject matter specialists.

3.614 Recruitment and Training of Field Staff

The quality of household survey depends to a significant extent on the quality of the field staff. Training for this group should be adequate. Usually, the decisions on whether to employ temporary staff or permanent staff depend on the periodicity of the survey. If it is a continuous survey, e.g., annual agricultural surveys, permanent field staff is usually considered. For surveys which are conducted once after a number of years, say five or 10 years, temporary staff is usually used.

The period of training will depend on numerous factors including the number of trainees, complexity of the survey, as well as the work experience of the trainees and length of the questionnaires. Usually those who

gained experience for the same survey will not require more time for training, whereas those who are new will need thorough training so that the data obtained is of good quality.

It is important to train more than required enumerators and administer some kind of a test as well as practical interviews involving real household, the results of which can then be used to select the required number of enumerators.

Remark: The quality of training will affect the quality of the survey and ultimately the quality of data collection.

3.615 Fieldwork

The actual work of going out to the sampled areas and interviewing the sampled households is typically referred to as the fieldwork. Fieldwork should start as soon as possible after training in order to minimise any forgetting of what was learned in the training.

All survey materials and equipment such as questionnaires, forms, tents, etc., should be made available on time so that interviewers are not hampered by lack of materials or equipment.

A detailed fieldwork plan must be drawn up that matches the household that has been selected with the survey team. The plan needs to be realistic and it should be drawn based on past experience. The field work should follow the work plan drawn. It is important to prepare a time table indicating a sequence and estimated duration of various operations of the survey and also to keep track of the sequence and summary returns of the data collection process.

Field staff is usually organised in teams led by supervisors and each team is assigned a portion of the sample to cover. The supervisor is responsible for ensuring that households in his/her assigned portion are interviewed. Adequate transportation should be provided for each field team.

3.615.1 Data Collection

a. Consider the Technical Aspects of data collection

i. Prepare instructions and manuals for:
(a) Sample selection,
(b) Completing the questionnaires (for enumerators),
(c) Supervision of field work (for supervisors), and
(d) Training of field personnel (including demonstration and testing materials).
ii. Prepare procedures for the supervision of data collection and quality control. The supervisor should check listings coverage, sample selection, spot checking of enumerators, re-interview a small sub-sample of respondents for cross checking, etc.

b. Consider the Operational Aspect of data collection.

i. Establish the criterion for the selection of field personnel (enumerators, supervisors, etc.);
ii. Plan for the training of field personnel (arrange for training materials, duration of training, etc.);
iii. Consider organisation:
(a) Distribute the sample by geographical locations,
(b) Allocate enumerators,
(c) Allocate supervisors (ratio of 1:3 to 1:4 is desirable in most cases), and
(d) Organise logistics-transport etc.

c. Consider the Publicity and Public Relations aspect of data collection, for instance

i. Addressing Kgotla meetings,
ii. Addressing schools, and
iii. Radio and Television programs; Newspapers; Pamphlets and Posters in public places.

3.615.2 Interviewers' Supervision

The quality of work done by enumerators is of crucial importance to any household survey and interviewer supervision is a key to good quality on the part of the interviewing staff. Assuring quality is not an easy task since some interviewers may simply not be able to do the work and others may not put forth their full effort. Important to maintaining quality work is an effective system of field work supervision. The general management of the survey team lies upon supervisors.

Some of the Responsibilities of a Supervisor include:

i. General management of the team,
ii. Create a good atmosphere among team members for better management and work relations,
iii. Identification of EA's and the selected households,
iv. Use the checklist provided for completed questionnaires on errors and omissions, and
v. Ensure that all procedures are followed when interviewing.

3.615.3 Quality Control

Controlling and assuring the quality of the survey is important. Many activities such as training and supervision are an integral part of quality control; however, there are some specific activities usually referred to as quality control which relates to controlling the quality of the interviewer's work. This is where by survey team members visits the field team members for field editing, validation, observation of interviews and addressing any field work problem.

a. Quality Assurance Measures at Data Collection Stage

The major sources of errors in data collection are usually those related to the completion of interviews. There are a number of operational controls at the disposal of a survey manager (both preventive measures and

corrective methods) that could be used in order to minimise errors that would otherwise arise during the process of survey operations.

i. Errors associated with incomplete listings or incorrect listings of households within a village EA could be prevented by rechecking the supervisors in the field and correcting the records by adding or deleting sampling units (household) to or from the list, respectively. Careful attention has to be paid to this survey operation since it could be an important source of error which could upset the random selection of the sample as well as the sampling weight.

ii. Non-interviews of all kinds also pose problems. The seriousness of this source of error basically depends on the proportion of the sample not completed and also on the extent to which the household not covered differs from those included in the sample in terms of the subject under investigation. These problems could be handled by repeated call-backs by an enumerator or supervisor. In all national household surveys, three call-backs are attempted and considered adequate provided they are made on strategic dates and at strategic times of the day when members of the household are expected to be home.

iii. Blank items (i.e., missing data) can distort survey results . Preventive measures for this problem include (i) the tabulation of missing data by enumerators, (ii) training of enumerators, and (iii) the revision of the questionnaire format. Corrective measures could include checking back the respondents by the supervisors.

iv. To improve the quality of the survey data, it is desirable that energy and resources are concentrated on minimising errors in (i) household listing, (ii) the rates of non-interview of households, (iii) the rates of non-response items in the questionnaires, (iv) respondent error, and (v) enumerators error in asking the questions and recording information at data collection stage rather than relying on the adjustment of the data at a later stage. These adjustments, in any case, would entail the use of assumptions and might not hold true.

REMARK: After completion of quality assurance exercise a report on the findings should be compiled for future use.

3.615.4 Fieldwork Problems and Possible Solutions

i. Refusals

Some people who have been selected for interview may refuse to be interviewed. The reasons that they give may be that they do not have time or that the enumerators invade their privacy. Some of them may answer the household questionnaire but refuse to answer some questions in the individual questionnaire. Attempts to interview them through revisits sometimes fail.

The conduct of the field staff may also contribute to refusals. Field staff are expected to conduct themselves in the manner that reflects a descent behaviour that is worthy of a government official. This covers such things as the way one dresses when they visit households to conduct interview and where, for instance, an enumerator involves themselves in a dispute at social places.

Refusals may be minimised through publicity of the survey through the media, use of cover letters that explains the purpose of the enumerator's visit.

ii. Terminations

A number of enumerators may terminate a contract for various reasons. This becomes a critical problem especially for surveys such as Household Income and Expenditure Survey (HIES) which takes twelve months as it requires properly trained staff. To overcome this problem, more than required enumerators and supervisors should be recruited and trained.

iii. Transport Constraints

Reliable transportation is crucial to the work of the survey teams. Shortage of vehicles and continuous vehicle breakdown are usually encountered during the field work period. If transport problems are unattended, they will result in delaying work progress. Each team should have dependable transportation so that it can move from one area to another. Emergency transportation must also be planned in case of breakdowns.

iv. Budgetary Constraints

Even though a survey is budgeted for, financial shortage may be encountered and may result in shortage of equipment as well as late payments for field staff. Therefore, it is important to draw a realistic budget and it should be based on past experience of a similar survey.

v. Enumeration Area Maps

The enumeration maps used in the survey may not be up-to-date such that some dwelling reflected in them may not exist during the survey period. For example, some demolished dwellings may be on the map but not on the ground. In some cases the floods may wipe away some dwellings and this will result in fewer households been listed during the listing exercise.

Finally, some important issues, procedures and concepts arising from conducting household surveys and censuses are discussed in the appendices (See Appendix III).

3.7 Other Data Requirement and Continuous Improvement

In addressing the dynamism of decision making Statistics Botswana continuously broadens its coverage areas of statistical information in order to fulfil its mandate. One of the areas to be considered is the provision of information on informal sector businesses from administrative records.

It is increasingly becoming advantageous to use Information and Communication Technology (ICT) as a driver rather than as hitherto known to be an enabler. In this respect Statistics Botswana continues to harness ICT to improve service delivery particularly in data collection and processing. The use of ICT to move away from paper based methods of data collection and processing and adopt the Computer Assisted Personal Interviewing (CAPI) is the trend that National Statistical Office follow to combine data collection, processing and data entry. The CAPI also removes the traditional transportation of paper questionnaires to office. It is therefore imperative for Statistics Botswana to adopt CAPI within no time.

CHAPTER 4: DATA MANAGEMENT AND INFORMATION SYSTEMS (DMIS)

4.1 Introduction

Data Management and Information Systems Division is composed of two units, namely: Data Management and Information Systems, mainly responsible for the overall provision of data management services for both primary and secondary data to facilitate the statistical function of Statistics Botswana; and IT related support services including database administration, IT infrastructure, maintenance of the website, system development and maintenance of existing systems.

4.2 DATA MANAGEMENT UNIT

Official statistics has become more important to many than before and as such there is an overwhelming demand for it. Thus, there is need to be disseminated accurately, timely and useably. These statistics are mainly used by government, private, and parastatals organisations and the general public for decision making and policy formulation. The provision of such is through carrying out various household and other related surveys.

Data Management Unit is responsible for the overall management of data within Statistics Botswana and the same services are extended to the entire NSS as a way of providing guidance as and when requested.

Data management describes the procedures and processes for organising and taking control of survey, census or administrative data. As such data management is key component of any national statistical office. Proper planning for data management processes and procedures are required and should be made known as the survey/census is undertaken.

It is against this backdrop that any data collected in Statistics Botswana, be it through surveys, censuses or administrative records are subjected to data processing procedures. Data processing is one of the key components of data management and it is an integral part of a survey/census undertaking from the initial

stages of planning to dissemination of statistical information. As such, a survey/census is not complete until data has been processed, analysed, reported and disseminated.

Data processing entails all the tasks that are required to produce meaningful information. The tasks includes development of training manual, training of relevant personnel, manual editing, coding, data capture and verification, online editing, data organisation and tabulation. In surveys/censuses undertaken in the organisation, a schedule is normally developed detailing the start and completion of data processing, with specific timelines as well as the necessary resources required. Timelines for the operation are communicated and progress reports prepared continuously until the process is completed.

For all surveys/censuses conducted in Statistics Botswana, Data Management personnel participates in the design of the questionnaire(s) to ensure that the variables of interest observed by the survey are properly recognised and identified and that skip patterns for the interviewers are explicit and correct.

4.21 Data Processing Operations

Data processing can be done manually or using a computer. When the volume of data is small, it is faster and cheaper to process the data manually; but when the volume is larger, it is ideal, faster and cost effective to use a computer, this leads to better data handling and fewer errors. The organisation normally uses both, that is, for survey and censuses, computer based processes are used while manual processes are used for secondary or administrative data.

Carrying out the steps in this process requires as much preparation and training of staff as does the actual enumeration because it is critical to the final quality of data. It is of great importance to adequately plan for data processing well in advance and build quality control systems into all stages of the data processing phases. These systems should allow for coordination and control of the flow of the questionnaires. All these call for the full involvement of data processing personnel in the survey/census right from its inception. In this regard, Data processing officers are normally included in the training of the field instruments for better understanding of the questionnaire and to bring them to the same level with their counterparts. They are also

part of the quality control team for the same reason alluded to above. For these reasons, Data Management Unit always appoints a focal point officer for each survey or census conducted and works closely with subject matter specialist statistician.

Operational control procedures are followed for smooth process of receiving questionnaires from the field and also during the various steps of data processing (manual editing and coding, data entry and online editing). Control procedures are equally important in guaranteeing that all data are processed and that no data are dropped or duplicated. There are control forms designed for monitoring purpose of all processes such as receiving data from fieldwork personnel, editors and coders, data entry personnel, verification officers, online editors as well as a reporting format on daily basis **(see Appendix IV)**.

4.211 Development of Editing and Coding Manual and Coding of Questionnaires

The coding manual is aimed at translating word classification into numerical numbers to facilitate transfer of data from the questionnaire to machine process. This is done by assigning codes to possible responses to questions. Assignment of codes depends on the type of questions in the questionnaire; the types are number/ value questions, closed response questions, open ended or semi open closed questions. The manual as a guide to coders is produced by subject matter specialists and should be shared with data management. Training is, thereafter, conducted by subject matter specialist and Data Management staff in readiness for coding.

Part of this task will be to elaborate the coding schemes to be used for items related to geographic location, subject of training, occupation (uses Botswana Standard Classification of Occupations (BOSCO) and industry (uses Botswana Standard Industrial Classification (BSIC)). There are, however, other standard codes commonly used, some of which are International Classification of Diseases (ICD10) – used by Health Statistics, International Standard Classification of Education (ISCED) – Used by Education Statistics, etc.

4.212 Manual Editing of Questionnaire

A lot of care goes into the design of the questionnaire, preparation of instructions and close supervision of field operations; despite all these, errors still creep in the field data. Such errors include omissions or blanks, impossible entries, unreasonable or inconsistent entries.

During manual editing, questionnaires are checked for completeness and consistency. Trained personnel review the forms in readiness for data entry to ensure that each questionnaire has correct and unique identification; that responses in the questionnaires are legible to the data entry clerk; and that unnecessary entries and blank pages are clearly marked so that they will not be inadvertently captured. A work unit is usually a single Enumeration Area (EA). Work is assigned by the supervisors and control forms used to record what has been issued to the editors and returned. The work of editors will be continuously verified by the supervisors. Editors whose work is consistently sub-standard is usually retrained or re-assigned to other duties, or even dismissed depending on the level of the errors he/she commits.

The guiding principles in questionnaire editing include minimising the number of changes made in the originally recorded information, eliminating inconsistent entries and supply entries for certain missing data. Such inconsistencies can be minimised by thorough checks at the field by supervisors and quality control teams.

During the coding and editing exercises, questionnaires are sorted and batched or divided into work units; the batching has an advantage as is based on some geographical level that is easily identifiable. For example, data for enumeration area X belonging to a particular district, say Kweneng East, are batched together and placed in one location. A batch control form is attached to each batch showing a district, EA Number, number of questionnaires per EA and total records by type of questionnaire (see Appendix IV).

4.213 Data Capture or Entry

Data entry aims to convert data from the questionnaires into the computer in a readable format using CSPro

software. The main method of data entry in Statistics Botswana is Keyboard data entry method, where desktop computers are linked to a local area network. For surveys and censuses, project owners or subject matter specialists provide necessary equipment such as computers and printers to be used during data capture processes. Ideally there should be a well-established and furnished data entry centre, however, due to financial constraints; Statistics Botswana is yet to establish one. Captured data reside in servers and back-up copies are written to any appropriate media, e.g., tapes, hard drives, CDs, etc.

Permanent data entry operators or clerks are currently certificate holders and are well conversant with operations of a computer. However, for surveys and censuses, temporary clerks are normally at least Botswana General Certificate of Secondary Education (BGCSE) certificate holders. In this case a keyboard test is administered during the process of selection.

Like others, Data Entry Operators/Clerks are trained in the application that will be used to capture data. Their daily work for each individual is monitored through a monitoring tool (see Appendix IV) and feedback is given on the error rates as well as on the overall performance. Operators whose work is sub-standard are retrained, re-assigned to other duties or dismissed (particularly when employed on temporary basis), according to the individual circumstances. Information on operators' performance is maintained in the computerised control system.

4.214 Online Editing

This involves editing and correction of captured and verified data through a system. The entered or captured data pass from the data capture phase to the consistency edit/checks and correction phase. This phase consists of programmes, written to fulfil the edit specifications prepared by the subject-matter specialists. This process helps to minimise inconsistencies during tabulation and analysis phases. The edit specifications are developed as per the questionnaire. The edit program is developed by System Developers and tested using the pre-test/Pilot data; and finalised before the actual online editing of census/ survey data. Online editing normally commences some few weeks after data capture to avoid delays in releasing results. No data are released to subject matter specialists before completion of online editing.

4.215 Production of Summary Tables

Summary tables are produced to further identify and check any inconsistencies that may be in the different tables. Where there are inconsistencies, the data is further corrected, either manually or through an edit programme depending on the magnitude of the inconsistencies identified. This is the last stage of data processing and once done, the data is then concatenated by system developers, exported to SPSS or any application that may be requested, weighted (weights prepared by Statistical Methods unit) and then delivered to the subject matter units for analysis and report writing.

4.216 Conclusion

Data processing is a crucial stage in the data life cycle of an organisation. If not given attention, may give irrelevant statistical information. Data from the field needs to go through data processing steps and procedures towards production of clean datasets. A good working relationship between data processing officers and subject matter specialist should be maintained for common understanding of the processes and issues towards production of clean data. All officers should be involved at the initial stage of the project as to be on common platform until the process is completed.

4.3 INFORMATION SYSTEMS UNIT

The Information System unit is responsible for the system development and maintenance of the existing systems, general IT support services, network and database administration and IT infrastructure. It has two sections, namely, System Development and IT Support Services. The manual details how the services are provided.

4.31 System Development Section

The section is responsible for system development and maintenance of existing systems in Statistics Botswana. System development processes are elaborated and presented below (see appendix 4). For surveys and

censuses undertaken in the organisation, System developers together with subject matter specialist engage each other from the onset as to get insight information on what should be done; should be part of questionnaire design team to understand the flow of questions; and should be part of training of enumerators exercise for the same reason. Any changes to the questionnaire during this period should be communicated to the developer within a reasonable time frame for incorporation into the data entry and/or online edit programme or system. Also, any new categories spotted during coding should be passed to the developers. A sound and proper data entry system and/or online edit programme requires the commitment from subject matter to provide guidance as and when required.

The subject matter specialists, in the process, prepare "edit and tabulation specifications" to guide the system development processes. Once the system(s) are ready, they are installed at the site and tested. The data entry personnel are trained and implementation and maintenance of these systems follow accordingly. Documentation of all processes is made at this juncture for future references. The systems, once completed and tested, are submitted to Data Management unit for implementation as per schedule in terms of household surveys and censuses. However, supervision of data entry personnel is the responsibility of system developers as well as maintenance of existing systems.

4.311 Creation of Backup and User Data Files

At all times during data entry and online editing systems development processes, including data entry and online editing exercises, back-ups of all versions of the programmes and data files are fully maintained. It must always be possible to return to any earlier version and recover the data or programme as they were should any problem be encountered in the process.

Upon completion of online editing, EA batches are concatenated. A final product is a clean dataset, which is submitted to the subject matter specialist in an SPSS format for analysis and report writing.

4.32 IT Support Section

The services provided by this section include general IT support, website maintenance, as well as network and database administration. The services are generally guided by the IT policy, IT guidelines and Helpdesk SLA. All these are in existence in Statistics Botswana and are being implemented.

APPENDICES

a) A5DICES

APPENDIX I: Pre Listing and Site Condition Forms

A. Prelisting Form

Republic of Botswana

Census District: …… Code
Village …….. Code
Locality …… Code
Enumeration Area: …… Code

Date Started EA

D	D	M	M	Y	Y

Date Completed EA

D	D	M	M	Y	Y

No.	Plot/ Lolwapa Number	Plot Type	Main use of plot/ Lolwapa	Non Residential	Name of contact person in Lolwapa	No. of persons living in the Lolwapa	Comments
		01 = Residential 02 = Non Residential ***(skip to A5)*** 03 = Both	01=Residential Occupied 02 = Residential Vacant 03 = Residential under Construction	01 = Business 02 = School 03 =Church 04=Govt. Offices 05= Others **(Specify)**			
(A1)	**(A2)**	**(A3)**	**(A4)**	**(A5)**	**(A6)**	**(A7)**	**(A8)**
01							
02							
03							
04							
05							
05							
07							
08							
09							
10							

TOTAL IN THIS PAGE

Site Condition Form

No.	Plot/ Lolwapa Number	Plot Type	Main use of plot/ Lolwapa	Non Residential	Name of contact person in Lolwapa	No. of persons living in the Lolwapa	Comments
		01 = Residential 02 = Non Residential *(skip to A5)* 03 = Both	01=Residential Occupied 02 = Residential Vacant 03 = Residential under Construction	01 = Business 02 = School 03 =Church 04=Govt. Offices 05= Others **(Specify)**			
(A1)	**(A2)**	**(A3)**	**(A4)**	**(A5)**	**(A6)**	**(A7)**	**(A8)**
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
35							
37							
					GRAND TOTAL		

APPENDIX II: Requisition, Training and Evaluation Forms

Statistics Botswana
Private Bag 0024, Gaborone, Botswana
Tel: (267) 367 1300. Fax: (267) 395 2201

A. Training Request Form

DATE_____

Surname, First Name	
Position	
Name of Organisation/ Institution/Unit	
Contact Details (Telephone Number)	
E-mail address	
Number of Participants	
Suggested Date for Training	
Duration of Training	
Venue [In-house/outside]	

a) Topic to be resourced on, specify, modules, sub-topics etc.	
a) _____	
b) _____	
c) _____	
d) _____	
Training Materials required: Flipcharts, Laptop, Projector, other (specify)	
2. To be used by training and Certification Officials	
Surname, First Name	
Position	
Telephone and email address	
Request Status	

Statistics Botswana
Private Bag 0024, Gaborone, Botswana
Tel: (267) 367 1300. Fax: (267) 395 2201

B. Training Preparation Form

Name of Trainer: _____ Date: _____

Organisation: _____

Topic: _____

Survey/Programme/Study: Instructional Objectives

_____

_____

Material/Training Aids:

_____

_____

*Lesson to be presented under 3 headings: Introduction, Presentation and Summary.

TIME	CONTENT (WHAT)	TRAINEE ACTIVITY (HOW)
Introduction		
Presentation		
Summary/Conclusion		

Statistics Botswana
Private Bag 0024, Gaborone, Botswana
Tel: (267) 367 1300. Fax: (267) 395 2201

C. Training Assessment/Observation Form

Name of Trainer: _____

Date: _____

Organisation: _____

Topic: _____

a) ORGANISATIONAL ATTRIBUTES /COMPETENCIES	MARKS	COMMENTS
a) Record of work and Planning 1. Highly organised scheme of work, all material available. e.g., objectives, content, methods of learning activities, etc. 2. Materials present, some explanation required, functional sequencing 3. Insufficient, incomplete, poorly presented. **b) Planning** 2 Clear and appropriate objectives, timing, introduction and development. 3 Minimal function with respect to the above criteria, good in some areas. 4 Some areas not attended to, dull, no evidence of research, clarity and thought of expression.		
b) TRAINING COMPETENCIES **1 Introduction** ***a)*** Effective /Integrates with lesson plan ***b)*** Some integration and relevance to lesson plan. ***c)*** No integration and poor sequence.		

b) **Development** 1. High interest, confidence and mastery of subject matter, appropriate and logical sequence, trainees logically involved, well thought out questions, varied. 2. Reasonable mastery of subject matter, some logical sequence followed, recall or observational questions. 3. Unsure /uncomfortable with subject matter, one teaching approach/trainer dominated the lesson. **c) Trainees centred activities** 2 Trainees are the focus of the lesson, actively involved, their interest involved. 3 No very active involvement of trainees 4 Serious lapses in trainee's activity, no involvement.						
b) PROFESSIONAL COMPETENCIES 1 Professional attitude/ethics/creativity /resourcefulness ***a)*** Interest, tolerance, respect, innovativeness, competency in control. Evidently striving to do his/her best, very professional. ***b)*** Some interest in what she/he is doing, some classroom control, tries to be professional. ***c)*** No sign of genuine interest, unprofessional.						
c) CONCLUSION 1. A comprehensive summary of the lesson using appropriate techniques, innovations and creativity. 2. Inconclusive lesson, time not adequate for summary.						
TOTAL		**Signature of assessor**				
0 – 49 **Fail**	**50 – 59** **Pass**	**60 – 69** **Credit**	**70 – 79** **Merit**	**80+** **Distinction**		

Statistics Botswana
Private Bag 0024, Gaborone, Botswana
Tel: (267) 367 1300. Fax: (267) 395 2201

D. Practical Observation Form

NAME OF ASSESSOR _____

NAME OF ENUMERATOR _____

DATE _____

TOTAL MARKS: 20

COMPETENCY	1	2	3	4	5
1.Introduction • Who you are • Whom you represent • Your purpose/why you are here • Objectives of the survey • Confidentiality aspect					
2. Logical thinking /Sequencing • Ability to identify who to administer questionnaire to • Following skip patterns • Asking appropriately as row and column wise. • Relationship to the HH. • Recalling(Remembering what was previously said)					
3.Confidence • Drawing attention of the respondent • Comfortable with subject matter • Showing high interest					
4.Clarity and thought of expression • No fumbling • Tactful • Audible					

• Articulate • Control					
5. General Mastery of the questionnaire • Mentioning sections of the questionnaire • Showing full knowledge of the questionnaire					
Comments : _____ _____ _____					

Statistics Botswana
Private Bag 0024, Gaborone, Botswana
Tel: (267) 367 1300. **Fax:** (267) 395 2201

E. Questionnaire Administration Form

NAME OF ASSESSOR: _____

NAME OF ENUMERATOR: _____

TEAM NO: _____

DISTRICT: _____

LOCALITY: _____

COMPETENCIES:

_____

WEAKNESSES:

_____

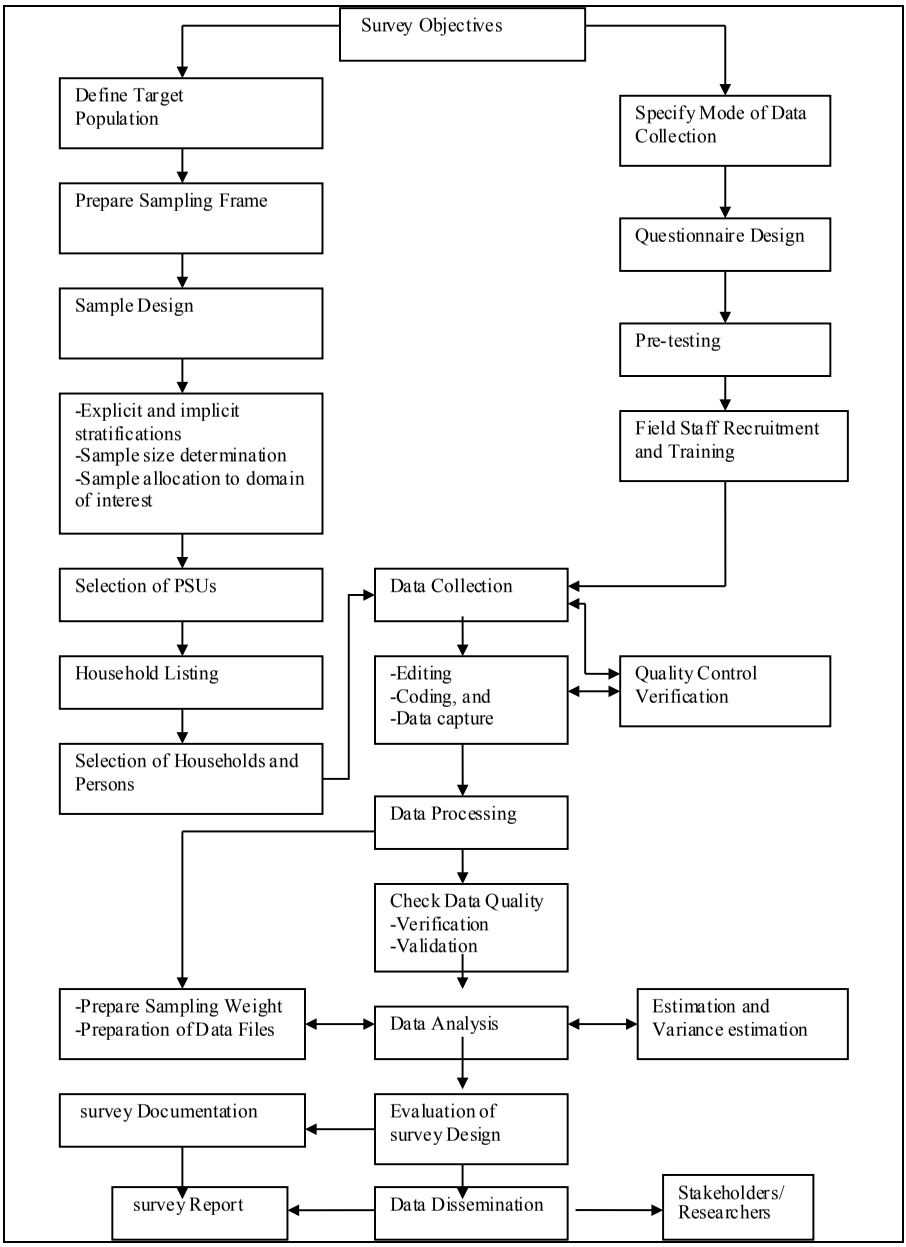
_____

WAYS TO IMPROVE:

_____

APPENDIX III: Household Surveys and Censuses Procedures and Concepts

A. Flow Chart of a Household Survey Process



B. Estimation of Number of Enumerators

The detailed calculation process for the estimation of number of enumerators required to complete the survey in specified period is given below. The example given below pertains to 2008 Botswana AIDS Impact Survey-III. This estimation helps in budgeting and procurement of survey resources (tents, transport, etc.).

BASIC INFORMATION REQUIRED FOR ESTIMATION OF NUMBER OF ENUMERATORS	
Number of days that one person will list one EA (listing of all households as the requirement for second stage sampling frame)	
Total number of selected EAs to be listed (as per sample design)	
Number of working days for enumeration in a month (22 days)	
On average travelling time to and from and within the EA per household (in minutes)	
Time to administer the Household questionnaire (in minutes)	
Time to administer the Individual questionnaire per respondent (in minutes)	
Average number of qualified individual respondent -10-64 years (as per target population)	
Time to complete the dry blood sample (DBS) (in minutes)	
Average number of qualified individual for DBS (≥18 Months)	
Selected EAs in Cities/Towns and Urban Villages (as per sample design)	
Selected EAs in Rural Areas (as per sample design)	
Number of Household to be selected per EA in Cities/Towns and Urban Villages (as per sample design)	
Number of Household to be selected per EA in Rural Areas (as per sample design)	
At this point the assumption is that an interviewer will find all the respondents at home for every single Household visited. We know that not to be the case. Households sometimes have to be revisited latter the same day or on following days or weekends. To compensate for this, let an additional time cost to be added is	
Duration of the survey (in months)	
Adjustment of resignations and absence of enumerators and supervisors	
Number of enumerators per team (ranging from 3-5)	
CALCULATION PROCESS	
ITEMS	**Value**
(A) Listing Time	

(1) Number of days that a team of 2 enumerators lists all the households in **one** EA (past surveys experience)	1 day
(2) Thus one EA will be listed (listing of all households) by one enumerator in (2x1)	2 days
(3) Number of selected EAs to be listed	459
(4) =(2)x(3) Number of days taken by one enumerator for listing of total selected EAs	918 days
(5) Number of working days in a month	22 days
(6)=(4)/(5) Number of months taken by one enumerator for listing of total selected EAs	41.7 Months
(B) Expected Time Spent per HH per Interviewer	
(7) Travelling time to and from and within the EA per household (past surveys experience)	25 Minutes
(8) Time to administer the Household questionnaire (pre-test estimate)	45 Minutes
(9) Time to administer the Individual questionnaire per respondent (pre-test estimate)	120 Minutes
(10) Estimating the average number of qualified individual respondents (10-64 Years)**1** per household	2.58 individual per household
(11)=(9)x(10) Total time to administer the Individual questionnaires per HH	310 Minutes
(12) Time to administer the Individual questionnaire per respondent-DBS (pre-test estimate)	15 Minutes
(13) Estimating the average number of qualified individual respondent (≥18 Months)-for DBS**2** (pre-test estimate)	3.08 individual per household
(14)=(12)x(13) Total time to administer the Individual questionnaires per HH – DBS	46 Minutes
(15)=(7)+(8)+(11)+(14) The total expected time spent per Household per interviewer	426 Minutes
(C) Expected Number of Households to be Interviewed per Interviewer per	

Qualified individual respondent (10-64 Years)= Household size x Proportion of 10-64 years respondents x Individual response rate = 4.1 x0.7 x 0.9 = 2.58

Qualified individual respondent (≥18 Months)-for DBS= Household size x Proportion of respondent (≥18 Months) x Individual response rate = 4.1*0.94*0.8= 3.08

(16) Usual working hours per day for the interviewers = 8 hours	480 Minutes
(17)=(16)/(15) Number of HH to be completed per interviewer per day	1.1 households
(18)=(17)x(5) Thus the expected Number of Households to be Interviewed per Interviewer per month	22.5 households
(19) To be conservative side: Number of Households to be Interviewed per Interviewer per month	22 households
(D) Expected Number of Months required by One interviewer to complete the survey field work all by him/herself	
Two basic settlements type strata are used in the PSU sample (EAs) and they have the following EAs allocation:	
(20) Number of EAs selected in Cities/Towns and Urban Villages	320 EAs
(21) Number of EAs selected in Rural Areas	139 EAs
(22) Number of Household to be selected per EA in Cities/Towns and Urban Villages	15 Households/EA
(23) Number of Household to be selected per EA in in Rural Areas	25 Households/EA
(24)=(19)/[(22) x 0.9] The estimated EAs that can be completed per month by one enumerator in Cities/Towns and Urban Villages = (adjusted for 90% household response rate)	1.6 EA
(25)=(19)/[(23) x 0.9] The estimated EAs that can be completed per month by one enumerator in Rural Areas = (adjusted for 90% household response rate)	1 EA
(26) Using the information (20) and (21) in (24) and (25), the number of months by one enumerator to complete the survey in each settlement is given as	
(27) =(20)/(24) Cities/Towns and Urban Villages	196.4 months
(28)=(21)/(25) Rural Areas	142.2 months
(29)=(27)+(28) Months for one enumerator to complete the survey fieldwork all by him/herself	338.6 months
(30) At this point the assumption is that an interviewer will find all the respondents at home for every single Household visited. We know that not to be the case. Households sometimes have to be revisited later the same day or on following days or weekends. To compensate for this, let an additional time cost to be added is	30 percent

(31) Total number of increased months= (29)+ 30% of (29)	440.1 months
[E] Expected Number of interviewers (excluding drives and supervisors)	
(32)= (6)+(31) Total time to do listing and the survey itself for **one enumerator** would therefore be	481.8 months
(33) The survey is to be completed in	2.75 months
(34) Percent resignations and absence of enumerators and supervisors	15 percent
(35)=(adjustment for 34 i.e. 1.15)x(32)/(33) Number of enumerators required to complete the survey	201 enumerators
(36) Number of enumerators per team (may vary from survey to survey)	4
(37)=(35)/(36) Number of teams	50

C. Format of Survey Budget

The provision of sufficient funds for (i) the employment, training and travel costs of survey personnel, (ii) purchase of equipment and supplies etc., is a pre-requisite for the design and execution of any survey. The salary and allowance of survey personnel would depend on the number of personnel, salary/allowance rate and duration of employment. The supply of stationary, printing materials, medicines, camping equipment is needed by all surveys. However, the expenditure on non-expendable items such as survey equipment (measuring and weighing scale), personal computers would depend on the nature of the survey. The cost components for the design and execution of household surveys and the suggested format are given as follows:

DESCRIPTION	Estimated Period	Quantity	Unit Cost	Total
1	2	3	4	5
1. Personnel Emoluments and allowances				
1.1 Salaries:				
Temporary Enumerators, team supervisors, coders and editors				
1.2 Overtime:				
Enumerators, supervisors and drivers				
1.3 Commuted Allowances				
Enumerators, supervisors, drivers				
1.4 Subsistence, Training and Meal Allowances				
Quality control trips, training allowance, meal allowance				
Sub-Total-1				

2. Charges for Hired Vehicles				
Main survey				
(i)Pool Vehicle Hire Rate-Enumeration				
(ii) Pool Vehicle Hire Rate-Quality control				
(iii)Fuel Cost				
(iv) Pool Driver hire				
(v) Mileage				
Sub-Total-2				
3. Training to Field Staff				
(I) Meals charges (person meals)				
(ii) Conference Hall				
(iii) Photocopying-Questionnaire for training				
(iv) Photocopying-100 Enumerators Manual-100 pages each				
Sub-Total-3				
4. Stationary and Camping Materials				
4.1 Materials/Supplies				
(I) Scientific Calculators				
(ii) Service Kit				
(iii) Pumps				
(iv) Paper files				
(v) Blue Pens				
(vi) Green Pens				
(vii) Pencils				

(viii) Eraser				
(ix) Clip boards				
(x) Enumerators Bag				
(xi) Photo copying Papers (A4)-Reams for pre-test and training (Questionnaires and manuals)				
4.2 Camping equipment etc.				
Sub-Total-4				
5. Procurements of PCs, and Printers				
(i) Micro computers				
(ii) Printer				
Sub-Total-5				
6. Other: Questionnaire Printing for Main survey				
Questionnaire Printing				
Sub-Total-6				
7. Consultancy: Analysis of Survey data and report writing				
Consultancy-Data analysis and report writing (may be a team of 4 experts)				
Sub-Total-7				
8. Dissemination Seminar on the survey results				
Conference facilities etc.				
Printing/ photocopying charges for 275 reports (includes papers, printing and binding charges)				

Sub-Total-8				
9. Printing of Final Survey Report				
Sub-Total-9				
10. Advertising and publicity of survey				
Sub-Total-10				
Total				
Contingencies				
VAT				
GRAND. TOTAL				

APPENDIX IV: System Development Process (including Pictorial Presentation)

The system development processes are used to guide the analysis, design, development, and maintenance of information systems of Statistics Botswana.

a) Process Model

Process capability is unpredictable because the software process is constantly changed or modified as the work progresses. Schedules, functionality and product quality are generally inconsistent. Performance depends on the capabilities of individuals and varies with their skills, knowledge, and motivations. There are few stable software processes in evidence, and performance can be predicted only by individual rather than organisational capability.

b) Tasks in the Development Process

Professional system developers and the subject matter specialists share a common goal of building information systems that effectively support Statistics Botswana business process objectives. In order to ensure that cost-effective, quality systems are developed which address an organisation's business needs, developers employ some kind of system development Process Model to direct the project's lifecycle. The activities performed include the following:

- System conceptualisation
- System requirements and benefits analysis
- Project adoption and project scoping
- System design
- Specification of software requirements
- Detailed design
- Software integration and testing
- System integration and testing

- Installation at site
- Site testing and acceptance
- Training and documentation
- Implementation
- Maintenance

c) The System Development Processes and their explanation

i) **System Conceptualisation:** System Conceptualisation refers to the consideration of all aspects of the targeted Statistics Botswana and other departments (stakeholders) functions or processes, with the goals of determining how each of those aspects relates with one another, and which aspects will be incorporated into the system. It is valid only if other departments are engaged like Police, Prisons, High court, Immigration, NACA, etc.

ii) **Systems Analysis:** This step refers to the gathering of system requirements, with the goal of determining how these requirements will be accommodated in the system. Extensive communication between the subject matter specialists (statisticians, different stakeholders) and the developer is essential.

iii) **System Design:** Once the requirements have been collected and analysed, it is necessary to identify in detail how the system will be constructed to perform necessary tasks. More specifically, the System Design phase is focused on the data requirements (what information will be processed in the system), the software construction (how will the application be constructed?), and the interface construction (what will the system look like? What standards will be followed?). This is basically aligned to database construction.

iv) **Coding:** Also known as programming, this step involves the creation of the system software. Requirements and systems specifications from the System Design step are translated into machine readable computer code.

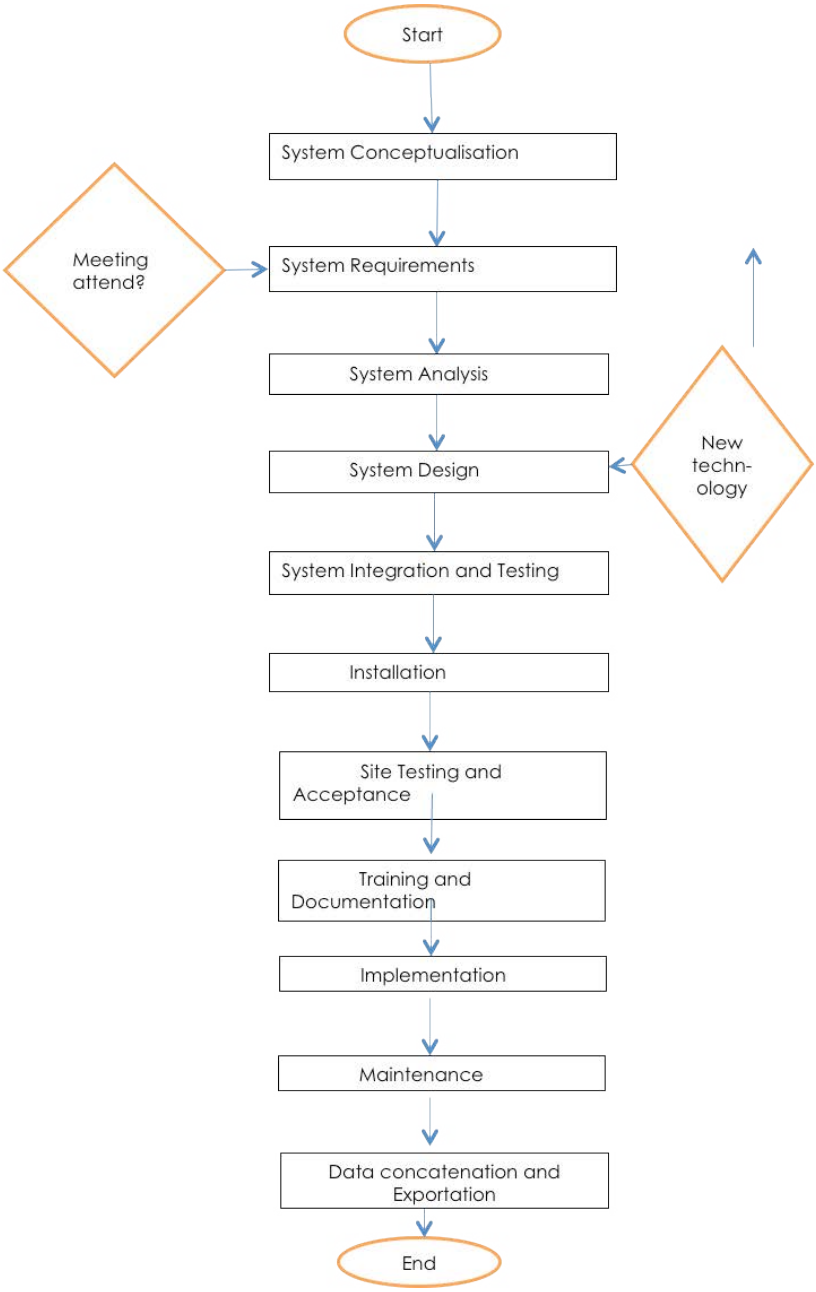
v) Testing: As the software is created and added to the developing system, testing is performed to ensure that it is working correctly and efficiently. Testing is generally focused on two areas: internal efficiency and external effectiveness. The goal of external effectiveness testing is to verify that the software is functioning according to system design, and that it is performing all necessary functions or sub-functions. The goal of internal testing is to make sure that the computer code is efficient, standardised and well documented.

vi) Training and documentation: As the testing is done and the subject matter specialists approve the system, the training which also involves user manuals starts. This is done to ensure that the system is stable, thus, data entry capturers fully understand the flow of the system. After system's stability, the final documentation is completed.

vii) Implementation: System implementation covers activities from a detailed workflow analysis to the formal go-live of the new system. During system implementation, special matter specialists may refine the initial workflow analysis that had been completed as part of the requirements analysis phase. In addition to the workflow analysis, full system testing is completed during this phase.

viii) Maintenance: Even though the system is stable and works excellently, there are always changes in functional and non-functional requirements that need to be attended to. These changes are managed in such a way that the system need not to be redone completely since it will affect schedule, budget.

ix) Data concatenation and Exportation: After data capturing and online editing are completed, the entire data are concatenated and exported to SPSS or any other format of which the analysts or subject matter specialists are comfortable with. These depend on which programming language has been used and how data was externally captured (in terms of standalone or networked machines) since it may not need this process.


Start
System Conceptualisation
Meeting attend?
System Requirements
System Analysis
System Design
New techn-ology
System Integration and Testing
Installation
Site Testing and Acceptance
Training and Documentation
Implementation
Maintenance
Data concatenation and Exportation
End

Statistics Botswana. Private Bag 0024 Botswana
Tel: (267) 367 1300. **Fax:** (267) 395 2201. **Email:** info@statsbots.org **Website:** www.cso.gov.bw

BOTSWANA 50
UNITED AND PROUD