



The Republic of Rwanda



# **NISR Survey Operating Standard Procedure Manual**

## **SOSPM**

National Institute of Statistics of Rwanda

**NISR Survey Operating Standard  
Procedure Manual  
SOSPM**

**April, 2018**



# NISR Survey Operating Standard Procedure Manual

---

## Table of content

### Contents

I. Purpose of the Manual.....	5
I. 1. Primary Objectives .....	5
I. 2. Secondary Objectives.....	5
II. Introduction .....	7
III. Rationale and Objectives of the survey .....	9
III.1 Research Rationale .....	9
III.2.The Objectives of the survey .....	9
III.3. Consultation of stakeholders from NSS .....	10
III.4. Identification of core indicators of the survey from stakeholders .....	11
IV. Survey planning.....	13
IV.1. Planning .....	13
IV.1.1. Required logistics .....	13
IV.1.2 Calendar of the logistics.....	14
IV.1.3. Budget of the survey.....	14
V. Questionnaire Design and review .....	17
VI.1. Population .....	20
VI.2. Sampling frame .....	21
VI.3.The sampling.....	21
VI.3.1.Simple random sampling .....	27
VI.3.2. Systematic sampling .....	27
VI.3.3. Stratified sampling .....	28
VI.3.4. Cluster Sampling .....	28
VI.3.5. Multi-Phase Sampling .....	29
VI.4. Sample Size Determination and Allocation.....	31
VI.4.1 Estimation .....	31
VI.5.Weightting .....	32
VI.5.1 Design Weight.....	33
VII. Data Collection Procedures.....	35
VII.1. Affection of teams .....	35
VII.2. The contact of the local authorities before any activity .....	36
VII.3. Arranging Transportation, Accommodations and Security .....	36
VII.4. Delimitation of the boundaries of the enumeration area .....	37
VII.5.The listing of the households.....	37
VII.6. Writing Instruction Manuals of Enumerators and Supervisors.....	38

# NISR Survey Operating Standard Procedure Manual

---

VII.7. Agreeing terms with any intermediate collection bodies such as Local Administrative Authorities .....	38
VIII. Recruitment process.....	39
VIII.1. Recruitment of the enumerators .....	39
VIII.3. Training organization and structure .....	41
VIII.4. Methods and content of training.....	41
IX. Field work organisation and operations.....	43
IX.1. Piloting the survey instruments and tools .....	43
X. Deployment plan of human resources on field .....	43
XI. Field Supervision plan and structure .....	49
XII. Listing procedure and organization.....	51
XII.1. Contacting the village authority .....	51
XII.2. Locating a selected Enumeration areas (EA)/ Cluster .....	51
XII.3. Preparing Map Spotting.....	52
XII.4. Listing of Units.....	52
XIII. Field Data collection.....	55
XIV. Data processing.....	57
XV. Survey Plan and roadmap/timetable and tasks distribution .....	61
XVI. Data analysis framework .....	63
XVI.1. Tabulation Plan .....	63
XVI.2. Constitution of the analysis plan .....	63
XVII. Report writing.....	65
XVII.1 The structure of a report .....	65
XVIII. Data Dissemination and Archiving .....	67
XVIII.1. Archiving.....	67
XVIII.2. Dissemination .....	70
XVIII.2.1. Methods of dissemination .....	71
XVIII.2.2 Dissemination plan .....	71
XVIII.2.3. NISR Overview on Dissemination.....	72

# **NISR Survey Operating Standard Procedure Manual**

---

## **I. Purpose of the Manual**

### **I. 1. Primary Objectives**

The principal purposes of this MANUAL are to:

- a. Secure an optimum degree of Statewide uniformity in surveying.
- b. Establish and maintain survey standards.
- c. Improve the overall efficiency of the Division's survey function.
- d. Provide a single reference source for Division-wide surveying policies, procedures, and information.

### **I. 2. Secondary Objectives**

The Secondary aims are to provide:

- a. A comprehensive reference for new employee orientation.
- b. Source material for training.
- C. Help employees in other institutions gain a better understanding of surveying and its proper relationship to other Administration activities.



# **NISR Survey Operating Standard Procedure Manual**

---

## **II. Introduction**

The survey operating standard procedure manual is written for researchers who need to do surveys for their studies, primarily for enumerators involved in the day to day operations of survey. It is also written for surveyors themselves and their respective employees. The document is designed to serve multiple purposes, source of information, a reference, a guide and training in the effective and successful operation of research survey<sup>1</sup>. In general, surveying is performed to determine the relative location or positioning of points on or near the earth's surface.

More specifically, surveying is the science of making measurements, relative to known or assumed datum's and standards, and applying the principles of mathematics to such measurements to determine existing or future horizontal and vertical position, from area, magnitude, boundaries, and extent of land parcels and topographical features<sup>2</sup>. Objectivity refers to whether information is accurate, reliable, unbiased, and is presented in an accurate, clear, and unbiased manner.

It involves both the content of the information and the presentation of the information. This includes complete, accurate, and easily understood documentation of the sources of the information, with a description of the sources of any errors that may affect the quality of the data, when appropriate. Objectivity is achieved by using reliable information sources and appropriate techniques to prepare information products.

---

<sup>1</sup> "Survey Manual Introduction." 8 Mar. 2007, <http://www.state.nj.us/transportation/eng/documents/survey/Chapter1.shtm>.

<sup>2</sup> "Surveying, Science or Art. - Tennessee Land Development Services." 13 Aug. 2011, <http://www.tnlds.com/2011/08/13/surveying-science-or-art/>.



## **III. Rationale and Objectives of the survey**

### **III.1 Research Rationale**

In the beginning of any research survey, the research rationale is required in order to show the research problem, define the key terms, objectives and the importance of the subject. It will review and summarize the available literature, note any gaps in this literature or in the study of the topic. Its main role is to establish what you propose to do, how you will do it, and it needs to be done. It also needs to be specific and ideally be able to explain the importance of the research, which is conducted by providing valid arguments. It has to provide a brief explanation of why your research topic is worthy of study and may make a significant contribution to the body of already existing research.

### **III.2. The Objectives of the survey**

The objectives of a research project summarize what is to be achieved by the study. These objectives should be closely related to the research problem. It states what researchers expect to achieve by the study in general terms. It is possible (and advisable) to break down a general objective into smaller, logically connected parts<sup>3</sup>. These are normally referred to as specific objectives. Specific objectives should systematically address the various research questions. They should specify what to be done in the study, where and for what purpose. The objectives will guide all subsequent survey steps that should be planned to ensure that the final results meet the original objectives. This step of the survey determines what is to be included in the survey and what is to be excluded; what the client needs to know versus what would be nice to know.

### **Why should research objectives be developed?**

The formulation of objectives will help us to:

- Focus the study (narrowing it down to essentials);
- Avoid the collection of data which are not strictly necessary for understanding and solving the problem you have identified;

---

<sup>3</sup> "Project Planning: Establishing Survey Objectives | Survey Connection." <https://survey.umn.edu/best-practices/project-planning-establishing-survey-objectives>.

# **NISR Survey Operating Standard Procedure Manual**

---

- Organize the study in clearly defined parts or phases.

Properly formulated, specific objectives will facilitate the development of your research methodology and will help to orient the collection, analysis, interpretation and utilization of data<sup>4</sup>.

## **How objectives should be stated?**

When the project is evaluated, the results will be compared to the objectives that have to be spelled out clearly, in order to facilitate the evaluation. The objectives of the study have:

- To cover the different aspects of the problem and its contributing factors in a coherent way and in a logical sequence;
- To be clearly phrased in operational terms, specifying exactly what you are going to do, where, and for what purpose;
- To be realistic considering local conditions;

## **III.3. Consultation of stakeholders from NSS**

In order to succeed in the research study, one of the most significant issues is to collaborate with all stakeholders related with the study to be undertaken. All stakeholders are supposed to claim their institutional problems, complaints and propose solutions and suggestions with the institution in charge.

The institution should also provide feedback to these opinions and suggestions regularly. The management of the institutions should support suggestions and projects and do rewarding workshops<sup>5</sup>.

Project managers should initially identify all the people or groups that have the potential to affect the work in progress, either positively or negatively. Keeping a list of specific interests helps the project manager anticipate stakeholder feedback. Project managers rate stakeholder importance so that input provided can be prioritized.

---

<sup>4</sup>"Develop Clear Research Objectives for Your Survey Research Project." 5 Mar. 2012, <https://www.snapsurveys.com/blog/develop-clear-research-objectives-for-your-survey-research-project/>.

<sup>5</sup>"Stakeholder power analysis - Power Tools." [http://www.policy-powertools.org/Tools/Understanding/docs/stakeholder\\_power\\_tool\\_english.pdf](http://www.policy-powertools.org/Tools/Understanding/docs/stakeholder_power_tool_english.pdf).

Additionally, project managers should consider establishing a contract with stakeholders so that everyone knows what roles they can be expected to play during the project life cycle, such as reviewing documents, contributing resources, and approving reports. Often, an interdisciplinary survey team is given responsibility for the planning, design, implementation, and evaluation of the survey and of its planned products. It is composed of members having different technical skills required to plan and implement the survey.

Survey teams are typically composed of a survey manager, an expert in the field of study covered by the survey (a subject matter expert), a survey statistician, a computer systems analyst, and a data collection and operations expert.

### **III.4. Identification of core indicators of the survey from stakeholders**

A technical committee has to be established in order to harmonize the utilization of indicators regarding sectors and if necessary try to find out if the produced indicators have no duplication and if necessary define sources/surveys and censuses within NISR and NSS it is coming from.

Statistical indicators are known as any quantitative data that provide evidence about the quality of measured topics that need to be monitored and help to determine whether objectives for a specific program or project have been achieved<sup>6</sup>. Before any research, indicators should be identified by stakeholders and then submitted to the institution in charge in order to be included in any designed methodology of the research.

These indicators need to be easily understood and meaningful to those who seek to use the information they provide. They should be specific, measurable, achievable and attributable, relevant and realistic, and time-bound. They should capture the essence of the desired result by clearly relating to the achievement of an objective, on the same way it should be framed in a way that enables comparison over time.

Where appropriate, the provided indicators should be similar in similar contexts to enable comparison.

---

<sup>6</sup> "The DHS Program - Survey Indicators." <https://dhsprogram.com/data/Survey-Indicators.cfm>.

# **NISR Survey Operating Standard Procedure Manual**

---

This should be done early, during the formation of a detailed adaptation plan, in order to ensure that:

- Indicators are linked to specific actions and/or objectives and should be considered during the preparation of adaptation plans, and included as [ part of the plan.
- The identification of Indicators by stakeholders to be included in a research will involve the development of an assessment plan to measure the performance of adaptation efforts in terms of relevance, effectiveness, efficiency, results and sustainability of the results.

The monitoring and evaluation of indicators establish the levels of performance that are likely to be achieved in a practical manner and that reflect the expectations of stakeholders .It allows progress to be tracked in a cost-effective manner at the desired frequency for fixed period.

## **IV. Survey planning**

### **IV.1. Planning**

#### **IV.1.1. Required logistics**

Before starting any data collection, survey materials have to be identified, listed, collected and stored in order to facilitate the field Interviewers. The officer in charge of logistics is responsible for ensuring that all the field materials are available and in good condition<sup>7</sup>. The distribution plan has to be established in advance in order avoid any gap that could occur during the distribution.

The following materials will be essential to conduct the fieldwork:

- Recommendation letter from the institution Authority
- ID badges
- Copies of the survey description letter to give to respondents
- Copies of the Questionnaire
- Tablets if CAPI is used
- Field Interviewer and Supervisor Manual
- Pens or pencils
- Funds to cover field expenses
- Team expenses
- Fuel and vehicle expenses
- Guides
- Advances for per diem allowances for the team
- Maps of enumeration areas
- List of enumeration areas
- Reporting forms
- Listing forms if applicable
- Umbrella

---

<sup>7</sup>"Methods of Survey Sampling - What sampling method should you use?."  
<https://explorable.com/methods-of-survey-sampling>.

## **IV.1.2 Calendar of the logistics**

Before starting any statistical survey, a detailed calendar of activities should be designed in advance in order to show and define all activities that will be done. The mentioned calendar will show clearly the time allocated to each included activity from the beginning to the end.

The survey schedules show the timing and sequence of tasks within a survey, as well as the survey duration and consist of tasks, dependencies among tasks, durations, constraints, milestones and other time-oriented project information.

The schedules specify the relative beginning and ending times of activities and their occurrence times. The schedules may be presented on a calendar framework or on an elapsed time scale.

## **IV.1.3. Budget of the survey**

The survey budget will include all activities that will be undertaken from the beginning to the end. The preparation of the survey budgets starts with the listing of the needed resources. The budget can be referred to the number of day and equipment that any group of personnel needs in order to achieve the planned survey activities.

Indeed the cost of survey material has to be checked properly based on previous surveys and a regular monitoring of budget execution has to be put in place. Underestimating the value of some items, or failing to anticipate fluctuations in the price of materials between drafting the budget and purchasing the materials, can seriously endanger the success of the project.

Hence, the following activities could influence the estimation of the budget:

- Data collection materials (computers, handhelds, telephones)
- The training costs of field staff, data collection ,processing ,analysis, archiving, and dissemination of microdata;
- The size of the sample, the period of data collection, and the length of the interview.;
- The staff costs such as outlays on transport and travel allowances, and equipment used could influence the budget cost.,

## **NISR Survey Operating Standard Procedure Manual**

---

- The costs of the listing of households within selected primary units or the listing of persons within selected households, interviewing and field supervision.
- The cost of data collection including the cost of travel both between and within PSUs, the length of the questionnaire, whether or not interpreters are used, and the number of units to be interviewed.

Furthermore, the budget must take into account the potentially significant costs associated with monitoring survey activities and results, for example, listing and subsampling procedures carried out in the field, the response rates for key domains of interest against pre-specified levels, etc. Effective monitoring of such activities enables survey implementers to take corrective measures, if necessary, during data collection, instead of discovering deficiencies after data collection, when it might be too prohibitively expensive to compensate for them.

On other hand, when using CAPI system of data collection, take into consideration the cost of electronic materials to be used on field, internet, appropriate software and the whole channel of data transmission. As part of data processing, data entry, edit and imputation work may involve a mixture of fixed and variable costs, depending on the degree of automation used in this process.

The other principal costs of data processing are arguably fixed, and include the costs of computing equipment and software; and the development of weights, and variance estimators and other data analysis work. For instance, weights would be computed regardless of the number of PSUs or households sampled; and after a weighting procedure has been developed and programmed, the incremental cost of computing a weight for an additional household would be negligible.

The budget is generally divided into staff costs, equipment (e.g., computers, freezers, centrifuges) and running expenses (e.g., office expenses, computer and laboratory consumables, travel costs). Each item must be justified and of reasonable cost. This will often involve discussions with the personnel office about staff costs and consulting price lists to get the best prices for

# **NISR Survey Operating Standard Procedure Manual**

---

supplies and equipment. In certain countries, allowances should be made for inflation and currency fluctuations when calculating the final budget.

## V. Questionnaire Design and review

A questionnaire (or form) is a group or sequence of questions designed to obtain information on a subject from a respondent. Questionnaires have a major impact on data quality since the questionnaire is the process whereby data are collected.

They also influence the image that the statistical agency projects to the public. The design of questionnaires has a significant impact on response rates and depend on how easily statistical office staff are able to convert the responses into electronic form.

The questions posed must conform to the survey's objectives and has to include all needed indicators and able to provide information that is useful for data analysis. All information requirements must be met, but each question should have an explicit rationale for being on the questionnaire and referred to the indicators that need to be measured. It must be known why each question is being asked and how the information is to be used. The question wording must be clear and sequenced logically for the respondent. The questions must be designed so that they are easily understood and can be accurately answered by respondents. Finally, the questionnaire should be tested before implementation using cognitive testing, focus groups<sup>8</sup>.

The questionnaire includes open-ended and closed-ended questions. An open-ended question asks the respondent to formulate his own answer, whereas a closed-ended question has the respondent pick an answer from a given number of options<sup>9</sup>. During the design of the questionnaire, previous similar surveys have to be consulted if necessary insert changes by including questions related to indicators that are necessary to meet the targeted objectives of the survey. These indicators must include the NSS, National (NST & SSPs), Regional (EAC & Agenda 2063) and Global (SDGs) indicators<sup>10</sup>. The use of the international standard statistical classifications

---

<sup>8</sup>"Defining The Purpose of a Survey, and The Importance of Goals." 5 Jan. 2016, <https://www.surveygizmo.com/resources/blog/purpose-of-survey/>.

<sup>9</sup>"Comparing Closed-Ended and Open-Ended Questions - FluidSurveys." 7 Aug. 2013, <http://fluidsurveys.com/university/comparing-closed-ended-and-open-ended-questions>.

<sup>10</sup> "SDG Indicators — SDG Indicators - UNSD - the United Nations." <https://unstats.un.org/sdgs/indicators/indicators-list/>.

# **NISR Survey Operating Standard Procedure Manual**

---

in the design of questions will be necessary in order to harmonize the contents of the questionnaire with the international standards and to allow the international comparison of data.

The design of the questionnaire will follow the following steps:

- consult with data users and respondents
- To list all the objectives and what information is required in order that they are achieved.
- To list all the questions that could inserted into the questionnaire
- To Refine the question phrasing
- To develop the response format including pre-coded list of answers or open ended.
- To put the questionnaires into appropriate sequences in order to brings logic and flow to the interview.
- To finalize the layout of the questionnaire by formatting with clear instructions to the interviewer, including a powerful introduction, routings and probes.
- To pretest and revise the questionnaire in order to find out if changes are needed so that final revisions can be made.

The design of a questionnaire will help to collect exploratory information (i.e. qualitative information for the purposes of better understanding or the generation of hypotheses on a subject) or quantitative information (to test specific hypotheses that have previously been generated).

## VI. Sample Design

Sampling is a means of selecting a subset of units from a target population for the purpose of collecting information. This information is used to draw inferences about the population as a whole. The subset of units that are selected is called a sample. The sample design encompasses all aspects of how to group units on the frame, determine the sample size, allocate the sample to the various classifications of frame units, and finally, select the sample. Choices in sample design are influenced by many factors, including the desired level of precision and detail of the information to be produced, the availability of appropriate sampling frames, the availability of suitable auxiliary variables for stratification and sample selection, the estimation methods that will be used and the available budget in terms of time and resources<sup>11</sup>.

### Design

When determining sample size, take into account the required levels of precision needed for the survey estimates, the type of design (e.g., clustering, stratification) and estimator to be used, the availability of auxiliary and contact information, budgetary constraints, as well as other factors, such as nonresponse, presence of out-of-scope units, attrition in longitudinal surveys, etc. For periodic surveys, take into account expected births and deaths of units within the changing survey population. It's worth noting that the precision of survey estimates is usually influenced more by the total sample size than by the sampling fraction (ratio of the sample size to the population size).

It is important to remember that most surveys produce estimates for many different variables, and optimizing the sample for one particular variable may have detrimental effects on other important variables. Handle this problem by first identifying the most important variables and then using this subset of variables to determine the sampling strategy to be adopted, which often requires a compromise between optimal strategies for the variables in the subset. See Bethel (1989). Stratification consists of dividing the population into subsets (called strata). Within each stratum, an

---

<sup>11</sup> "Sample design - Statistics Canada." <https://www.statcan.gc.ca/pub/12-539-x/2009001/sample-plan-eng.htm>.

independent sample is selected. The choice of strata is determined by the objectives of the survey, the availability of variables on the frame, the distribution of the variable of interest, and the desired precision of the estimates.

Most surveys produce estimates for various domains of interest (e.g., provinces). If feasible, take this into account in the design by stratifying appropriately (e.g., by province). Otherwise, it will be necessary to consider special methods at the estimation stage to produce estimates for these domains. To achieve statistical efficiency, create strata in such a way that each stratum contains units that are as homogeneous as possible with respect to the information collected in the survey. For longitudinal surveys, choose stratification variables that correspond to characteristics that are stable over time. Conduct studies to evaluate alternative sampling methods, stratification options and allocation schemes. The usefulness of these studies depends on the availability and vintage of data used to conduct the studies, whether from previous censuses, surveys or administrative data and their relation to the variables of importance to the survey.

Establish an expected response rate using a pre-test or data from previous occasions of the same or similar surveys. This rate can in turn be used in sample size determination. A sample can be divided into waves and additional waves of sample can be released as needed based on the achieved sample by stratum. For longitudinal surveys, expected cumulated attrition for the given number of cycles must be used.

## **VI.1. Population**

Before any sampling procedure in any study, the population has to be identified. A population, in statistics is a discrete group of people, or things that can be identified by at least one common characteristic for the purposes of data collection and analysis.

It will serve as the basis from which the sample will be drawn during any research or study. It includes all people or items with the characteristic one wish to understand. It is important to clearly identify the population being

studied or referred to, so that you can understand who or what are included in the data.

The population needs to be clearly identified at the beginning of a study. The study should be based on a clear understanding of who or what is of interest, as well as the type of information required from that population. However due to financial constraint it will not always be possible to do the study on the whole population, the reason why the sample has to be drawn from the identified population.

## **VI.2. Sampling frame**

Before any sampling, the sampling frame has to be identified from the list of all the items in the population. It is known as a complete list of individuals that need to be studied where the sample has to be drawn. The difference between a population and a sampling frame is that the population is general and the frame is specific. The source of the sampling frame has to be specified in advance and updated in order to capture enough information during the study.

## **VI.3. The sampling**

In survey methodology, sampling is the selection of a subset (a statistical sample) of individuals from within a statistical population, to estimate characteristics of the whole population. Two advantages of sampling are that the cost is lower and data collection is faster than measuring the entire population. Each observation measures one or more properties (such as weight, location, color) of observable bodies distinguished as independent objects or individuals. In survey sampling, weights can be applied to the data to adjust for the sample design, particularly stratified sampling. Results from probability theory and statistical theory are employed to guide the practice. In business and medical research, sampling is widely used for gathering information about a population.<sup>[2]</sup> Acceptance sampling is used to determine if a production lot of material meets the governing specifications.

The sampling process comprises several stages:

- Defining the population of concern

# NISR Survey Operating Standard Procedure Manual

---

- Specifying a sampling frame,
- Determining the sample size
- Implementing the sampling plan
- Sampling and data collection

When the Survey Design starts, the planning phase of the statistical agency must decide whether to conduct a census or sample survey. If the decision is a sample survey, then the agency needs to plan how to select the sample by specifying the target population. During the sampling subsets of units from a population will be selected for the purpose of collecting information for those units to draw inferences about the population as a whole. Within any of the types of frames identified which will be used, a variety of sampling methods can be employed, individually or in combination:

## Sampling Methods

The following paragraphs provide brief explanations of technical terms used in sampling that a survey manager should be aware of. They can be read in advance of completing the form or/and referred to when completing the information for this section. They are included here using simple language and without mathematical formula<sup>12</sup>.

<b>Sampling</b>	Sampling is the process of selecting a number of cases from all the cases in a particular group or universe.
<b>Probability sampling</b>	A probability sample is a sample selected by a method based on the theory of probability (random process), that is, by a method involving knowledge of the likelihood of any unit being selected.

---

<sup>12</sup> "The OECD Glossary of Statistical Terms - OECD.Stat." <https://stats.oecd.org/glossary/>.

## NISR Survey Operating Standard Procedure Manual

<b>Non-probability sampling</b>	<p>A sample of units where the selected units in the sample have an unknown probability of being selected and where some units of the target population may even have no chance at all of being in the sample.</p> <p>Forms of non-probability sampling are numerous, such as voluntary samples (only responses of volunteers are used), quota samples, expert samples.</p> <p>Data collected using non-probability sampling cannot provide valid conclusions about the whole population – their results are only valid about the members of the sample.</p>
<b>Sampling error</b>	<p>That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a sample of values is observed; as distinct from errors due to imperfect selection, bias in response or estimation, errors of observation and recording, etc.</p> <p>The totality of sampling errors in all possible samples of the same size generates the sampling distribution of the statistic which is being used to estimate the parent value.</p>
<b>Standard error</b>	<p>Measures the variability of the estimate, or precision. The larger the standard error of an estimate the less precise it is.</p>
<b>Non-sampling error</b>	<p>An error in sample estimates which cannot be attributed to sampling fluctuations.</p> <p>Non-sampling errors may arise from many different sources such as defects in the frame, faulty demarcation of sample units, defects in the selection of sample units, mistakes in the</p>

## NISR Survey Operating Standard Procedure Manual

	<p>collection of data due to personal variations or misunderstanding or bias or negligence or dishonesty on the part of the investigator or of the interviewer, mistakes at the stage of the processing of the data, etc.</p>
<b>Sample size and errors in estimation</b>	<p>Standard errors are inversely proportional to the square root of the sample size. This means that the gain in precision for every extra unit in the sample size is bigger when the sample size is small. As the sample size increases, the benefits of every extra unit in the sample become smaller quite quickly. The practical</p>

	<p>consequences of this are that reasonable precision may be affordable, but extreme precision can be very, very expensive!</p> <p>While large sample sizes tend to provide smaller more precise estimates, survey managers should be aware that large sample sizes increase the occurrence of non- sampling errors. Therefore the gains in precision must be carefully balanced against the risk of introducing errors that cannot be measured.</p>
<b>Sampling unit</b>	<p>A sampling unit is one of the units into which an aggregate is divided for the purpose of sampling, each unit being regarded as individual and indivisible when the selection is made.</p>

## Stratification

Stratification refers to the division of a population into strata. Strata are non-overlapping subsets of the whole population (often, but not always, geographically defined) within each of which a separate sample is selected.

Stratification is usually done with one of these two objectives:

To potentially improve the overall precision of the estimates by gaining control over the composition of the sample. For instance, we may want to ensure that the sample contains certain predefined proportions of households headed by men and women, or in urban and rural areas, or in different regions of the country.

To produce estimates for subgroups of the population that otherwise could be poorly represented in the sample. For instance, a non-stratified sample of Argentina will contain a lot of households from Buenos Aires but very few from a less populated province such as Tierra del Fuego. If we want estimates of sufficient precision for all provinces, we need to ensure that our sample contains enough households from each of them.

These objectives are not complementary:

If the objective is to obtain precise estimates for the population as a whole, the sample should be allocated among strata more or less in proportion to their population;

If the objective is to obtain estimators of

## NISR Survey Operating Standard Procedure Manual

	<p>comparable precision for all strata, the sample should be of about the same size in each of them.</p>
<b>Sampling frame</b>	<p>A list of all members of a population used as a basis for sampling. In multi-staged sampling, sampling frames may be constructed for different stages in the sampling process.</p>
<b>Multi-stage sampling</b>	<p>Multi-stage sampling is a sampling method by which a sample is selected in stages. The sampling units at each stage are sub-sampled from the units chosen at the previous stage. The sampling units belonging to the first stage are called primary or first stage units; and similarly for second stage units, etc. The sampling units at the last stage of the process are called the final or ultimate sampling units.</p>
<b>Sampling weights</b>	<p>Because a sample is used to estimate characteristics of the population, each value in the sample makes a contribution to the estimation of the population parameter. This contribution is its weight. Because of the complex sampling designs that are used, in most cases sampling units carry the different weights, and these weights need to be derived. The derivation of weights is based on the probability of selection of a sampling unit. Weights can be derived as soon as a sampling scheme has been designed, but these weights will need some adjustment after data have been collected to take account of non-responses.</p>

## VI.3.1. Simple random sampling

In a simple random sample (SRS) of a given size, all such subsets of the frame are given an equal probability. Each element of the frame thus has an equal probability of selection: the frame is not subdivided or partitioned. Furthermore, any given pair of elements has the same chance of selection as any other such pair (and similarly for triples, and so on). This minimizes bias and simplifies analysis of results. In particular, the variance between individual results within the sample is a good indicator of variance in the overall population, which makes it relatively easy to estimate the accuracy of results.

SRS can be vulnerable to sampling error because the randomness of the selection may result in a sample that doesn't reflect the makeup of the population. For instance, a simple random sample of ten people from a given country will on average produce five men and five women, but any given trial is likely to over represent one sex and underrepresented the other. Systematic and stratified techniques attempt to overcome this problem by "using information about the population" to choose a more "representative" sample.

## VI.3.2. Systematic sampling

Systematic sampling relies on arranging the study population according to some ordering scheme and then selecting elements at regular intervals through that ordered list.

Systematic sampling involves a random start and then proceeds with the selection of every  $k$ th element from then onwards. In this case,  $k = (\text{population size} / \text{sample size})$ . It is important that the starting point is not automatically the first in the list, but is instead randomly chosen from within the first to the  $k$ th element in the list.

As long as the starting point is randomized, systematic sampling is a type of probability sampling. It is easy to implement and the stratification induced can make it efficient, *if* the variable by which the list is ordered is correlated with the variable of interest. 'Every 10th' sampling is especially useful for efficient sampling from databases.

### **VI.3.3. Stratified sampling**

When the population embraces a number of distinct categories, the frame can be organized by these categories into separate "strata." Each stratum is then sampled as an independent sub-population, out of which individual elements can be randomly selected. There are several potential benefits to stratified sampling, it is very useful and considered during the sample design.

First, dividing the population into distinct, independent strata can enable researchers to draw inferences about specific subgroups that may be lost in a more generalized random sample.

Second, utilizing a stratified sampling method can lead to more efficient statistical estimates (provided that strata are selected based upon relevance to the criterion in question, instead of availability of the samples). Even if a stratified sampling approach does not lead to increased statistical efficiency, such a tactic will not result in less efficiency than would simple random sampling, provided that each stratum is proportional to the group's size in the population.

Third, it is sometimes the case that data are more readily available for individual, pre-existing strata within a population than for the overall population; in such cases, using a stratified sampling approach may be more convenient than aggregating data across groups (though this may potentially be at odds with the previously noted importance of utilizing criterion-relevant strata).

Finally, since each stratum is treated as an independent population, different sampling approaches can be applied to different strata, potentially enabling researchers to use the approach best suited (or most cost-effective) for each identified subgroup within the population.

### **VI.3.4. Cluster Sampling**

Cluster sampling is the process of randomly selecting complete groups (clusters) of population units from the survey frame. It is usually a less statistically efficient sampling strategy than SRS and is performed for several reasons. The first reason is that sampling clusters can greatly reduce the cost of collection, particularly if the population is spread out and

personal interviews are conducted. The second reason is that it is not always practical to sample individual units from the population. Sometimes, sampling groups of the population units is much easier (e.g., entire households). Finally, it allows the production of estimates for the clusters themselves (e.g., average revenue per household).

Cluster sampling is a two-step process:

- First, the population is grouped into clusters (this may consist of natural clustering, e.g., households, schools).
- The second step is to select a sample of clusters and interview all units within the selected clusters.

The survey frame may dictate the method of sampling. Until now, the focus has been on sampling individual units of the population from a list frame.

If the units of the population are naturally grouped together, it is often easier to create a frame of these groups and sample them than try to create a list frame of all individual units in the population. For example, the client may be interested in sampling teachers but only have available a list of schools. In the case of household or farm surveys, many countries do not have complete and up-to-date lists of the people, households or farms for any large geographic area, but they do have maps of the areas. In this case an area frame could be created, with the geographical areas divided into regions (clusters), the regions sampled and everyone within the region interviewed. Different sample designs can be used to select the clusters, such as SRS, SYS or PPS. A common design uses PPS where sampling is proportional to the size of the cluster.

### **VI.3.5. Multi-Phase Sampling**

Despite the similarities in name, multiphase sampling is quite different from multi-stage sampling. Although multi-phase sampling also involves taking two or more samples, all samples are drawn from the same frame and the units have the same structure at each phase. A multi-phase sample collects basic information from a large sample of units and then, for a subsample of these units, collects more detailed information. The most common form of multi-phase sampling is two-phase sampling (or double sampling), but three or more phases are also possible. However, as with

## **NISR Survey Operating Standard Procedure Manual**

---

multi-stage sampling, the more phases, the more complex the sample design and estimation.

Multi-phase sampling is useful when the frame lacks auxiliary information that could be used to stratify the population or to screen out part of the population.

For example, suppose information is needed about cattle farmers, but the survey frame only lists farms, with no auxiliary information. A simple survey could be conducted whose only question is: 'Is part or all of your farm devoted to cattle farming?' With only one question, this survey should have a low cost per interview (especially if done by telephone) and consequently the agency should be able to draw a large sample. Once the first sample has been drawn, a second, smaller sample can be drawn from amongst the cattle farmers and more detailed questions asked of these farmers. Using this method, the statistical agency avoids the expense of surveying units that are not in scope (i.e., who are not cattle farmers).

Multi-phase sampling can also be used to collect more detailed information from a subsample when there is insufficient budget to collect information from the whole sample, or when doing so would create excessive response burden. The Canadian Quarterly Retail Commodity Survey (QRCS) is one example. The first phase of the survey is the Monthly Wholesale Retail Trade Survey (MWRTS). Each month, MWRTS asks wholesale and retail establishments for two variables: Their monthly sales and inventories. QRCS subsamples the retail establishments and asks them to report their sales by retail commodity, for example, clothing, electronics, foodstuffs, etc. Similarly, multi-phase sampling can be used when there are very different costs of collection for different questions on a survey. Consider a health survey that asks some basic questions about diet, smoking, exercise and alcohol consumption.

In addition, suppose the survey requires that respondents be subject to some direct measurements, such as running on a treadmill and having their blood pressure and cholesterol levels measured. It is relatively inexpensive to ask a few questions, but the medical tests require the time of a trained health practitioner and the use of an equipped laboratory, so are relatively

expensive. This survey could be done as a two-phase sample, with the basic questions being asked at the first phase and only the smaller, second phase sample receiving the direct measurements.

## VI.4. Sample Size Determination and Allocation

One of the questions most frequently asked of a statistician is: how big should the sample be? Managers are anxious to obtain an answer to this fundamental question during the planning phase of the survey since it impacts directly on operational considerations such as the number of interviewers required. There is no magical solution and no perfect recipe for determining sample size. It is rather a process of compromise in which the precision requirements of the estimates are weighed against various operational constraints such as available budget, resources and time. Determining the size of the population is one of the most difficult decisions to make in your survey. A larger sample can yield more accurate results but the more responses you collect, the more expensive it gets.

The best way to calculate a sample size is to use the size of the population you are surveying. The precision of the survey estimates and the sample size are interrelated: since the sampling variance decreases as the sample size increases, the greater the precision required of the estimates, the larger the sample size needed. Therefore, the appropriate sample size depends on the desired precision of the survey estimates. This precision of an estimate,  $t$ , may be expressed in terms of the allowable standard error,  $SE(t)$ , the margin of error,  $z \times SE(t)$ , or the coefficient of variation  $SE(t) / t$ . The determination of the sample size for a survey often includes the specification of desired precision in terms of one or more of these measures.

### VI.4.1 Estimation

Introduction to Surveys states that the client is typically motivated by a wish to study the characteristics of a population. Estimation is the means by which the statistical agency obtains values for the population of interest so that it can draw conclusions about that population based on information gathered from only a sample of the population. The principle behind estimation in a probability survey is that each sample unit represents not

only itself, but also several units of the survey population. It is common to call the average number of units in the population that a unit in the sample represents the design weight of the unit. Determining this weight is an important part of the estimation process. While the design weights can be used for estimation, most surveys produce a set of estimation weights by adjusting the design weights.

The two most common reasons for making adjustments are to account for nonresponse and to make use of auxiliary data.

Once the final estimation weights have been calculated, they are applied to the sample data in order to compute estimates. Summary measures of the population such as totals, averages and proportions are typically estimated for a wide range of characteristics collected from the sample units.

These characteristics, often referred to as variables in statistical theory, may be qualitative ( for example, sex or marital status ) or quantitative ( for example, age or income). Depending on the type of data, different formulas are appropriate for the estimation of summary measures. An important part of estimation is estimating the magnitude of the sampling error in the estimate. This provides a measure of the quality of the survey's estimates for the specific sample design. Sampling error can only be estimated if probability sampling is used. The purpose of this chapter is to illustrate how to calculate weights, generate estimates of summary measures and estimates of their sampling error for probability sample surveys.

## VI.5. Weighting

The survey must produce estimates for a large number of variables, but all those estimates will be based on a set of weights attached to each responding household's data record, and determined by the product of three straightforward calculations:

- The design weight, calculated from the dwelling's probability of selection;
- An adjustment factor for nonresponse;
- A post-stratification adjustment to demographic benchmarks.

**Weighting for Unequal Probability Sample Designs:** Sample designs are called self-weighting when the design weights are the same for all units in the sample. This occurs when each unit has the same inclusion probability. For self-weighting designs, if no subsequent adjustments are made to the design weights (e.g., for nonresponse or auxiliary data), weights can be

ignored in the production of some statistics, such as proportions and averages. The production of totals simply requires that the sample total be inflated by the design weight. Although the simplicity of self-weighting designs is appealing, it is not always possible or desirable to select a self-weighting sample. For example, in the case of a stratified design for a national survey for smaller regions, N-proportional allocation may result in inadequate sample sizes and for larger region; it may result in unnecessarily large samples.

## **VI.5.1 Design Weight**

The first step is to calculate the design weight, which as its name suggests is derived directly from the sample design, namely the design weight is the inverse of the probability of selection. The sample design in this case is a stratified two-stage design with PSUs selected using PPS at the first stage and SSUs (dwellings) selected using systematic sampling at the second stage. The design weight is calculated for the selected dwelling. The household living in that dwelling, and each person in that household, will have the same design weight. In a multi-stage probability design, the selection probability of a final-stage unit is simply the product of the selection probabilities at each stage. Equivalently, the design weight can be seen as the product of the weights at each stage, since each of them is the inverse of the corresponding probability.



## **VII. Data Collection Procedures**

Data collection is the process of gathering and measuring information on targeted variables in an established systematic way, which then enables one to answer relevant questions and evaluate outcomes. It is also a process of collecting information from all the relevant sources to find answers to the research problem, and evaluate the outcomes. During the data collection the questionnaire will be administered to the selected people through interviews. The frequent research interviews structured, semi-structured and unstructured.

The purpose of the research interview is to explore the views, experiences, beliefs and/or motivations of individuals on specific matters. Interviews are, therefore, appropriate for exploring sensitive topics, where participants may not want to talk about such issues in a group environment.

Therefore, any data collection will be undertaken through the following steps:

- Affectation of teams
- The distribution of questionnaires and listing forms
- The contact of the local authorities before any activity
- Delimitation of the boundaries of the enumeration area
- The listing of the households within the household
- Arranging Transportation, Accommodations and Security

### **VII.1. Affectation of teams**

In the beginning of any data collection the teams have to be dispatched in advance and the team members appointed. The number of interviewers required depends on the sample size, on the number of days to be spent interviewing and the on the number of respondents one interviewer can interview in a day.

You can estimate this number from the length of a working day divided by the amount of time it takes to complete one interview (determined during pretest of the questionnaire).

## **VII.2. The contact of the local authorities before any activity**

In rural areas it is customary before embarking on a formal interviewing survey to notify the relevant public authorities, e.g. Districts authority, village head, etc. to ensure cooperation from respondents.

Sometimes individuals may refuse to co-operate unless they are convinced that the interviewer has permission and approval to conduct the survey from the recognized local authorities.

Contacting Local Authorities (District, Sector, Cell and Village) officials is responsibility of the supervisor before the team begins to work in an area. The Letter of Authorization has to be designed for this purpose, and a copy should be given to the appropriate authorities. In addition, Field enumerators should have copies of this letter in case they need to show it to authorities, such as local police officers, that they encounter during their field visits.

## **VII.3. Arranging Transportation, Accommodations and Security**

Transportation may be provided in advance by using government vehicles or renting cars, ensure that they are well maintained during the fieldwork period. Allow funds in the study budget for fuel, maintenance, and eventual repairs, or ensure that funds have been availed is an important step. The transport means have to be arranged in the way of helping enumerators to move easily from their accommodation to the field and vice versa. It is often possible to arrange for the team's accommodations with local communities: teams may often sleep in church buildings, army quarters or even the mayor's house.

If private accommodations are arranged, make sure that the interviewers get a daily allowance that is sufficient to cover their costs. Arrangements should be made for paying the field workers and supervisors, as well as providing them with "pocket money" for meals, accommodations and other unexpected expenses. Many places have no commercial restaurants, so meal arrangements will have to be made by enumerators themselves. Security issues are also important. Field work and accommodation may take place in area where security of enumerators is guaranteed. Local guides are often useful.

## **VII.4. Delimitation of the boundaries of the enumeration area**

The boundaries of each Enumeration Area (EA) must be clearly defined, so that there are no overlapping EA's or missing segments. The map should be the most detailed and up-to-date available with reference points (rail lines, bridges, waterways, street names, etc) delineated as well. Obtaining and Preparing Enumeration Area Maps before the field work begins, in order to know exactly the area the data collection is going to operate. This will allow to locate the assigned segment areas and plan the best travel routes within them. It will determine the precise boundaries of each assigned segments and indicate the location of each structure on a map.

The supervisor will provide to the enumerator a map of each segment that will help to recognize the precise boundaries of the segment. When preparing the fieldwork it is essential to be familiar with the area within the segment. After you are certain of the location of the sampled segment and its defined boundaries, you must determine the most efficient and thorough way to travel the roads and byways of the segment. When traveling within the segment, you must completely cover every side of every street and roadway within the entire segment, always looking for structures on one side of the street.

All boundaries of the segment must be followed (including railroad, water, power lines, and non-visible boundaries) because they may contain structures. This map will enable the field interviewers who perform data collection in this segment to be certain they are working in the proper area and it will also serve as a quality assurance tool for your supervisor.

## **VII.5. The listing of the households**

After the delimitation of the boundaries of the enumeration area (EA), the listing of households is required and should be done according some techniques that had to be trained to the enumerators in advance. Interviewers are often responsible for both listening and interviewing the member of selected households in the enumeration area.

Before listing, each interviewer should be given:

# **NISR Survey Operating Standard Procedure Manual**

---

- The instructions on how to plot a route through the EA, covering the entire area without backtracking (which risks duplication) or missing areas. It also includes instructions on how to identify an in-scope sample unit. For example, if the sample unit is a dwelling, there should be instructions on how to find and identify hidden and multi-unit dwellings, the definition of an unoccupied dwelling, etc.
- A listing form with the information to be collected for each Enumeration area. Once listing has been completed, data on the listing forms are generally captured electronically and sent back to head office so that the next stage of sampling can be performed.

The interviewer is instructed to begin at the Northwest corner and proceed in a clockwise direction around each block, listing all habitable dwellings on his or her right. The same general listing method (beginning at a predetermined starting point, proceeding clockwise, listing units on the right) is used in rural areas as well.

## **VII.6. Writing Instruction Manuals of Enumerators and Supervisors**

Any survey will need an instruction manual that will include all instruction regarding the process of data collection. The instruction had to be on two types, one for the enumerators and the other one for the supervisors. The instruction manuals have to be drafted clearly in order to explain the detail of the tasks that will be assigned to each enumerator and supervisors described above and will be used during the training and consulted in the data collection on the field.

These tasks have to be understood scrupulously in order to allow the collection of reliable data and avoid any technical issue which could occur during the data collection. Before the data collection all enumerators and supervisors have to be given a copy that had to be consulted any time during the data collection.

## **VII.7. Agreeing terms with any intermediate collection bodies such as Local Administrative Authorities**

Before the data collection different authorities have to be informed on the calendar of the implementation of the data collection of survey in order to allow the people sensitization of the activity.

## **VIII. Recruitment process**

Before any data collection, the term of reference (ToR) of recruitment of consultant, supervisors and enumerators have to define in advance in order to put in place an adequate team which will lead and implement all the field activities of the survey, the calendar has to be established in advance.

### **VIII.1. Recruitment of the enumerators**

The interviewers and supervisors play a critical role in surveys, as they are members of the research team who implement the survey design. They are often required to perform multiple tasks with a high level of accuracy. In a face-to-face survey, the interviewer may be required to physically locate the sampled household and to update the sample frame. In both telephone and face-to-face surveys, the interviewer has to contact the household, explain the purpose of the study, enumerate household members, select the respondent, motivate the respondent to participate in the required manner, and accurately record the respondent's answers as well as any other required information.

The interviewers and supervisors must be recruited based on the evaluation in order to determine whether they are capable of performing the task that is assigned to them. The recruitment of interviewers may be based on different criteria such as level of study, experience in data collection and so on. In addition an evaluation test should be organized in order to select the best elements among the group. The recruited interviewers will be able to understand and interpret easily the questionnaire and especially be fluent with the language that is going to be used during the interview. In addition reserves have to be added on the number of enumerator to replace those who could leave during the data collection.

During the recruitment:

- The research organization is requested to have its own in-house interviewing staff in order to select suitable interviewers for the particular survey.

## **NISR Survey Operating Standard Procedure Manual**

---

- The interviewing component of the study may be defined in advance and should be specified in the term of reference relative to the survey.
- Persons with other jobs, bad behavior may not be selected to work on the study.
- The research organization had to recruit more than the number of interviewers needed for data collection to allow any replacement in the case of the dismissal of candidates who prove to be unsuitable.
- When possible, select interviewers who have previously worked on similar studies and have good recommendations based on their performance.

In another hand, the supervisors must be similarly evaluated and recruited based on their level of leadership and experience to perform their task. Interviewers should have at least completed the full period of schooling within their country and be fluent in the main language of the country. Individual countries must decide what further level of education is required as well as what formal assessments will be carried out prior to selection.

It must be mentioned here that the nature of the training must be adapted to the tasks that the supervisors are expected to perform such as refusal conversions, cross-checking and verification of selected interviews and editing of interviews.

Detailed protocols for these procedures must be drawn up and clearly explained during the training process. The mechanism of informing candidates in need to apply for the job has to be put in place in order to avail equal opportunity to be selected, this has to be released through advices on different medias. Once interviewers recruited they should be ongoing evaluation during the initial data-collection period and at the conclusion of the fieldwork. The interviewers can be given a formal assessment at the end of training and some form of certification provided to each successful interviewer.

## **VIII.3. Training organization and structure**

Once the enumerators recruited, the training will be the following phase that will enable the enumerators to understand the content of the questionnaire in detail, how it has to be administered to the people. The content of the instruction manual has to be well explain to the enumerators as it will be the main instrument that had to be consulted any time ,when necessary in the household. Collecting high-quality data will only be possible if enough time is allowed to train the supervisors and enumerators thoroughly. Training is normally provided by the consultant and senior survey staff. A separate trainer a professional with experience in the field may be necessary for training in anthropometry.

It is recommended that the training be carried out in a location near the project office, with all enumerators and supervisors receiving the same training in a single classroom. Training is usually conducted over five days. At least two days should be dedicated to anthropometry training. Supervisors may also need additional training in checking questionnaires and ensuring quality control in anthropometry. Training should not exceed eight hours per day, at the end of which trainers should meet to evaluate the day. The training sessions provided for the survey team will need to be adapted to the particular country context, as well as to the abilities of the trainees.

Approximately five days should be allowed for the training. Cultural sensitivity is an important aspect of the training programmer. Training should be provided at all levels of the survey team involved in the survey, from interviewers to trainers and supervisors, as well as to the central team overseeing the process nationally. This will ensure that all involved persons are clear with regard to their role in ensuring good quality of data. The calendar should be established in advance and had to be respected scrupulously.

## **VIII.4. Methods and content of training**

Training should be long enough for the interviewers to become familiar with not only the techniques for successful interviewing, but also the

# **NISR Survey Operating Standard Procedure Manual**

---

content of the questionnaire to be used. For experienced interviewers, the training should be shorter than for less experienced ones. The training should be done in the language which is used in questionnaire, the language which is supposed to be used in the fieldworks in order to communicate well to interviewers. The recommended length of training may be from three to five days, with three days being appropriate for experienced interviewers requiring training on the questionnaire only. The longer period of training is recommended for all other interviewers.

All the training should be carried out as far as possible by the same team to ensure a standard training either for all interviewers in one session or for different groups at different times and places. To cut down costs and provide for regional training, training may be decentralized and cascaded. However, these costing benefits are then outweighed by the disadvantages of a diluted or varying training. The training methods should include as much role playing in interviews as possible (with a minimum of one per interviewer). This method provides assimilation of interviewing techniques more effectively.

For role playing to be effective, different scripts must be prepared in advance of the training so that the different branching structures of the interview, the nature of explanations that are permitted, and anticipated problems during an interview with difficult respondents can be illustrated. Training materials should be provided to all interviewers to use as reference material. Any material provided should be comprehensively reviewed during the training and, where relevant, should be translated into the languages used in the country.

The content of training should include the following:

- Administrative issues
- Planning of fieldwork
- Review of all materials provided
- Contacting procedures, consent forms and confidentiality
- Conducting an interview should encompass:
  - Interview procedures in the field
  - Supervision in field and reporting procedures

## **IX. Field work organisation and operations**

### **IX.1. Piloting the survey instruments and tools**

The pilot study is a small study to test research protocols, data collection instruments, sample recruitment strategies, and other research techniques in preparation for a larger study. The pilot study will be one of the important stages in a research project and will be conducted to identify potential problem areas and deficiencies in the research instruments and protocol prior to implementation during the full study.

It will also help members of the research team to become familiar with the procedures in the protocol, and will help them decide between two competing study methods, such as using interviews rather than a self-administered questionnaire. During the pilot survey, the calendar of the activities has to be taken into consideration by introducing all activities that had to be tested like the data collection, the survey CAPI data collection application, the data quality control mechanism and the pilot report..

Basically, pilot testing allow to find out if your survey, key informant interview guide or observation form will work in the “real world” by trying it out first on a few people. For large or complex surveys it’s a good idea to do a full pilot before starting actual data collection. When doing a pilot you will test all the survey steps from start to finish with a reasonably large sample. The size of the pilot sample depends on how big your actual sample is, and how many data collectors you have. For a typical baseline or endline survey a sample of around 30-50 people is usually enough to identify any major bugs in the system. The purpose is to make sure that everyone in your sample not only understands the questions, but understands them in the same way. This way, too, you can see if any questions make respondents feel uncomfortable. You’ll also be able to find out how long it takes to complete the survey in real time.

## **X. Deployment plan of human resources on field**

Once you have created a great survey, you need to deploy field staff. Field data collection is a complex process that often requires lots of time, money,

and people. Deploying your survey in the field effectively will not only help save these resources, but it will also make your data more reliable.

The following steps are required to help the deployment of staff in the field survey:

## **Step1: Identify the resources**

The first step of creating a field plan is to estimate the resources you have for your project. The main resources to consider are budget and time.

Budget and time will determine how many responses you should collect, how many surveyors you should recruit, and how much you should spend on resources like devices, documents, projectors, refreshments, training logistics, salaries, and more. Your Scope of Work (SOW) will help determine the budget and time available for your project. The Scope of Work (SOW) is a formal document that describes the work activities, deliverable, timelines and milestones, pricing, quality requirements, governance terms and conditions, etc. The SOW should also outline all the parties involved in the project, the budget, and the timelines.

## **Step2: Recruit your field staff**

The field staffs are crucial; they are the main investigators on field, as well as the people collecting data. That is why it is vital to find reliable and trustworthy field staff. Before recruiting field staff, it is important to consider their responsibilities.

What should they know in advance?

How will they be trained? How will they collect data?

Who will be responsible for monitoring the data collection process?

You must create a fixed hierarchy by ensuring that a clear reporting and management system is in place, project timelines are being met, and field staff are supported at every step of data collection.

The most common hierarchy uses three types of people:

- Field Managers,
- Monitors, and
- Surveyors.

# **NISR Survey Operating Standard Procedure Manual**

---

Field managers lead the entire team working in the field. They supervise and oversee tasks of field employees, run training programs, and ensure that everyone works as effectively as possible.

Monitors directly manage and support the surveyors. Surveyors collect data from the field by directly interacting with the sample population and recording survey responses.

The last step is to recruit field staff at all levels. When recruiting field staff, consider the following four factors:

1. Education qualifications
2. Ability to handle technology
3. Previous experience
4. Willingness to learn new things

### **Step 3: Create a plan**

An implementation plan is the complete plan for training and data collection. Having a plan in place ensures that everyone is on the same page, which makes data collection go more smoothly.

First, estimate the number of days required for training field staff. Ideally you should not train more than 50-60 people in one room at one time. Based on how complex and long your survey is, you could do a one-day training, two-day training, or trainings spread out over multiple days. By multiplying the duration of training (in days) and number of batches (with each batch 50-60 staff members), you can calculate the number of total days needed for training.

The training will allow to understand very well the content of the questionnaire and estimate what materials and people will be required and the number of days for the data collection.

It will also allow calculating:

1. Total number of responses required (also known as the sample size)
2. Total number of surveyors
3. Average time taken to fill each questionnaire

The number of surveyors will influence the time allocated to the data collection, this means that hiring more surveyors will decrease the time needed for data collection.

# **NISR Survey Operating Standard Procedure Manual**

---

It is a good practice to include morning and evening sessions as a part of your team's daily schedule. Morning sessions allow for a quick check-in with the team on how they are feeling, their goals for the day, and the progress of their project. Evening sessions allow for feedback and review. A good practice is to ask the team how their day went, if they faced any challenges, and if they have any feedback for the questionnaire or data collection app during the evening session.

## **Step 4: Train field staff**

After creating an implementation plan, you need to prepare for and conduct training for your staff. Important preparations include the following:

- Phones/tablets should be procured and given to the field staff.
- Make field staff aware of the fact that training will happen on tablets.
- Create a plan for setting up the training material.
- Field staff should be aware of the purpose of the training.

The best way to conduct training for mobile-based data collection is to divide the training in three parts: technical training, survey training, and data collection training. Technical training focuses on learning the mobile-based data collection technology. Before conducting the technical training, it is helpful to go through the data collection app yourself. Be sure to try the full functionalities of the app, make a questionnaire from scratch and fill out an entire questionnaire to understand how the application functions. Survey training should happen only once your field staff is comfortable with the device and the app. Then the next step is to train the staff about the survey itself. Survey training should focus on explaining the logic behind the survey questionnaire and how to ask the different types of questions on the survey.

## **Step 5: Monitor your team**

It is necessary to monitor the team on the ground while they are collecting the data. The surveyors who are piloting or collecting data need to be observed and advised on common mistakes, doubts, and best practices. You can share these learnings in the morning or evening sessions so everyone can benefit from them. The data being collected need to be monitored

## **NISR Survey Operating Standard Procedure Manual**

---

regularly in order to check for any inconsistencies or problems or if they are facing any challenges, or require any additional resources.

The feedback and experiences from enumerators have to be taken into consideration. Give regularly a feedback to surveyors by giving a constructive criticism as you observe them on the ground. Don't be mean, be specific, and be sure to explain the rationale behind your feedback and how each person can improve.



## **XI. Field Supervision plan and structure**

The field supervision plan is in responsibilities of the supervisor and it follows the structure of the survey plan. The supervisor is expected to make the arrangements of the field activities with enumerators to the assigned households. The field Supervisor is a critical link between the field enumerators and the central office. The field enumerators depend on field supervisor for guidance and support. The central office also depends on field supervisor not only to monitor production and performance, also to communicate any field issues that may have an impact on the timely completion of NISR.

As a senior member of the field data collection team, a field supervisor is entrusted with ensuring that the team's data are collected according to NISR data collection protocol and are of the highest quality. The field supervisor position is intended to provide supervision of field enumerators.

The position's major objectives are to ensure the following:

- ❖ Field data collection activities produce data of the highest possible quality.
- ❖ Field data collection activities are conducted efficiently, within established cost and time frame parameters.
- ❖ All fieldwork is performed in a professional manner.

You are responsible for managing the efforts of your staff with respect to all mapping and listing activities that may be required for NISR. From an operational point of view, the success of every project depends on two basic factors:

- (1) Following established and proven survey protocols to collect quality data and
- (2) Managing the time and resources spent to collect the data.

NISR field supervisor is responsible for managing the unit screening and interviewing operations. Supervisor is required to follow these standard supervisory procedures:

- Set goals and manage work plans. Assign and manage units.
- Mentor and meet with Field Interviewers.

# **NISR Survey Operating Standard Procedure Manual**

---

- Monitor data collection.
- Handle reluctant respondents and refusals.
- Manage “nobody home” and “selected respondent not at home” households.
- Assign non-interview final result codes. Distribute passwords and unlock codes.
- Solve problems.
- Prepare a field supervision plan/calendar

## **XII. Listing procedure and organization**

The listing requires the visit to each selected segment to record a description of each household. The member of a team will be appointed and trained to perform the listing operation; he/ she is supposed to prepare the listing and do his or her work at the same time as the other. The Units listing operation is a fundamental operation in surveys. After the Enumeration areas (EAs) are selected for the survey, a complete listing of dwelling units/households in the selected EAs is conducted prior to the selection of households. The listing operation consists of visiting each of the selected clusters, collecting geographic coordinates of the cluster, drawing a location map of the cluster as well as a sketch map of the structures in the cluster, recording on listing forms a description of every structure together with the names of the heads of the units in the structures and other characteristics.

The listing operation is an important procedure for reducing non-sampling errors in the survey, especially when the sampling frame is outdated. The listing operation provides a complete list of units in the EA. This information is necessary for an equal probability random selection of households in the second stage. With the units listing prior to the main survey, it is possible to pre-select the sample households in advance and the interviewers are asked to interview only the pre-selected households without replacement of non-responding.

### **XII.1. Contacting the village authority**

When the team visits an EA/cluster especially in rural areas, it is necessary to contact the chief of village, unit committee members or opinion leaders and inform him/her of your mission before you start with your listing activities. Briefly explain the objectives of survey. Impress upon them that the information gathered will be treated as confidential and there will be a follow up by another group for the main fieldwork.

### **XII.2. Locating a selected Enumeration areas (EA)/ Cluster**

Each team will be given a location map of the EA /cluster ie, a basic EA /cluster map and a description form indicating the boundaries to the

# **NISR Survey Operating Standard Procedure Manual**

---

EA/cluster selected. He has to read and follow these descriptions carefully to trace the boundaries. In most EAs /clusters, the boundaries follow easily recognizable natural features such as streams or rivers, and constructed features such as roads or railroads. However, lines may be invisible (especially in rural areas) of the selected EA /cluster. However, since the maps were drawn some years ago some basic features might have changed and new landmarks found.

Include all useful information to find the EA /cluster and its boundaries directly on the map and on the description form attached so that these maps can be updated. Listing is done by visiting every building/structure and writing in chalk the serial number of the building/structure on the front door or on the wall near the front door or at a conspicuous spot on the building.

Listing will ensure that every house in the EA has been covered. Remember that this serial number on the structure should be the same as on the household listing form.

## **XII.3. Preparing Map Spotting**

The supervisor will complete the top portion of the Location Identification Form (Form village/1) by filling in the identification information for the EA /cluster: the locality name, the two digit code for the region, the two digit code for the district, and the one-digit code that designates whether the place is a large city, small city, town or rural area . The remaining information will be completed after the completion of the EA /cluster. All information needed for filling in the identification box is provided to the team. Form Village/1 will also serve as a control sheet for the Supervisor to monitor the progress within each EA /cluster.

## **XII.4. Listing of Units**

Each interviewer will use the Listing Form ( Form village) to record all units found in the sub-area of the EA /cluster assigned to him. Begin by entering your name and the number assigned to you. Then enter the identification codes of the EA /cluster.

## **NISR Survey Operating Standard Procedure Manual**

---

Because the sample design is based on clusters which have been delineated on maps, the first task for the interviewing teams is to identify the selected clusters and go into the field to list the dwellings in them with a sufficiently detailed description that they can be uniquely identified and easily found again if selected for interviewing.

Once the lists are completed, the supervisor of district select the appropriate number of dwellings, using systematic sampling from the lists, with sampling intervals and random starts communicated to them by the regional managers.



## **XIII. Field Data collection**

Data collection is the process of gathering data by establishing systematic way, in which all steps have to be prepared in advance in order to collect successfully reliable data. Then before deploying field data collectors the process of affecting teams, the distribution of questionnaire or tablets has to be done in advance in order to allow a reliable organization. The teams of enumerators had to be headed by supervisors that have to organize and handle all concerns that could occur during the fieldwork. He will have in his attribution a role of collecting information from all the relevant sources and find answers to the research problem, and evaluate the outcomes.

During the data collection the questionnaire will be administered to the selected people through interviews. The purpose of the research interview is to explore the views, experiences, beliefs and/or motivations of individuals on specific matters. Interviews are, therefore, appropriate for exploring sensitive topics, where participants may not want to talk about such issues in a group environment. The team on the ground collecting data will need to be observed and advised on common mistakes, doubts, and best practices. During the data collection, the feedback and experiences from enumerators have to be taken into consideration.

The supervisor is requested to give regularly a feedback to surveyors by giving a constructive criticism observed on the ground and the way of improving. This could be shared in the morning or evening sessions so everyone can benefit from them. The data being collected need to be monitored regularly in order to check for any inconsistencies or problems or if they are facing any challenges, or require any additional resources. The enumerator is requested start interviews in data collection, to introduce the reason and the sponsor of the survey before starting any interview with the respondent.

During the data collection, the enumerators have to be guided by the contents of the instruction manual, which has to be well understood and consulted during the data collection process. Upon the team arrived in the area of survey, the supervisor has to contact the local authorities before the team begin to work in order to beneficiate their support regarding the awareness of the people that had to be interviewed .Sometimes individuals

## **NISR Survey Operating Standard Procedure Manual**

---

may refuse to co-operate unless they are convinced that the interviewer has permission and approval to conduct the survey from the recognized local authorities.

The Letter of Authorization has to be designed for this purpose, and a copy should be given to the appropriate authorities. In addition, Field enumerators should have copies of this letter in case they need to show it to authorities, such as local police officers, that they encounter during their field visits.

The enumerators will be asked to go along the delimitation boundaries of the EA, in order to list all households, work that has to be done by filling the listing forms which has to include all the head of households within each EA.

It will determine the precise boundaries of each assigned segments and indicate the location of each structure on a map which will be provided to each enumerator in order to recognize the precise boundaries of the segments.

Transportation means of field staff may be arranged in advance by availing well maintained vehicles or renting cars. This will help enumerators to move easily from their accommodation to the field and vice versa. If private accommodations are arranged, make sure that the interviewers get a daily allowance that is sufficient to cover their costs. Arrangements should be made for paying the field workers and supervisors, as well as providing them with "pocket money" for meals, accommodations and other unexpected expenses. Many places have no commercial restaurants, so meal arrangements will have to be made by enumerators themselves. Security issues are also important. Field work and accommodation may take place in area where security of enumerators is guaranteed. Local guides are often useful.

## **XIV. Data processing**

All the various steps of data processing should be planned early at the design stage of the study. This should be done in consultation with a statistician and a computer programmer. During the field activities, the enumerators collect information with the purpose of the study and save the data using maybe questionnaire or tablet. After data collection, data included in questionnaires or tablets are returned back to be entered by using different software like SPSS, Excel, etc. They are processed to obtain a datafile so that tables may be produced and the survey results analysed. The steps to be carried out at this stage of the survey include coding, data capture, editing and imputation, outlier detection and treatment, and database creation. In addition, a variety of quality control and quality assurance procedures are developed and put into effect.

The Survey System offers offline survey software for Android smartphones, tablets, Windows 6.5, earlier smartphones and personal digital assistant (PDAs). These small, light devices are ideal for conducting in-person surveys at malls, amusement parks, movie theaters, airports, and other public locations. The computer assisted personal interviewing (CAPI) software is also ideal for conducting interviews in remote locations. The Survey System's Mobile Survey Software Module offers superior features for conducting both simple surveys and complex interviews with sophisticated logic.

The closed questions are pre-coded on the questionnaire, so that the data capture operators simply need to key in the code numbers found beside the circles or boxes corresponding to the response. The open questions dealing with quantities (respondent's age in years, number of hours worked, wages, etc.) can also be captured directly from the questionnaire. On the other hand, the questionnaire contains several questions, such as type of work done, which have so many possible answers that the categories and codes cannot be included on the questionnaire. In these cases, the list of codes has been prepared and given to office personnel who write the appropriate code on the questionnaire prior to data capture.

## **NISR Survey Operating Standard Procedure Manual**

---

Therefore, a code list of economic sectors had to be provided to the office staff to use to assign the most appropriate value to the response before data capture.

After coding completed, batches of questionnaires are sent to the data capture operator who keys the coded responses into the computer to create the preliminary datafile. This is done using data capture screens prepared by the NISR team's informatics support team. Editing began when the interviewers, and then their supervisors, groomed the questionnaires to check that they were properly completed, and to see whether follow-up with the respondent was needed. Some further grooming is done in the office prior to coding and data capture. More editing occurs during data capture, since the capture screens were programmed to detect certain errors at the time of keying. These edits include validity checks for out-of-range codes.

Once the data are in electronic format, more complete editing is done. As soon as the questionnaire was finalized, the team began preparing edit rules to check for unreasonable combinations of codes for groups of related questions. Because of the grooming and manual edit steps; very few records have so many remaining edit failures at this stage that they will have to be declared unusable. Any such case is treated as a total nonresponse. As in the case of refusals and other instances of total nonresponse at the data collection stage, these cases are compensated for by adjustment to the survey weights.

For useable questionnaire records, all fields failing an edit rule and all item non-responses are analyzed for possible imputation. Depending on the type of question, the method of imputation varies. Once imputation has been completed, the datafiles may be again passed through the edit programs, to ensure that the data are fully consistent. Imputation must be run again on the file, to eliminate the few remaining records which fail the edit rules, and the file is then edited again to be sure of consistency. At the same time, diagnostic indicators such as the number of cases imputed for each field, the type of imputation used, the number of eligible donors and the

## **NISR Survey Operating Standard Procedure Manual**

---

frequency of their use, and other measures, are recorded as input to the survey evaluation process.

At the evaluation stage, these indicators will be used to calculate edit failure rates and imputation rates for the key survey variables and those variables that showed the greatest number of problems. The edit and imputation flags will also be maintained on the file to aid in diagnosing the overall quality of the final database. The last steps in processing are to create a flat file to serve as the basic survey datafile, and to compute the weights and add them to the file. The survey file then stores the file in a suitable format for the NISR' main statistical processing software (to speed up production of tables and other survey outputs).

Once processed and organized, the data may be incomplete, contain duplicates, or contain errors. The need for data cleaning will arise from problems in the way that data is entered and stored. Data cleaning is the process of preventing and correcting these errors. Common tasks include record matching, identifying inaccuracy of data, overall quality of existing data, reduplication, and column segmentation. Such data problems can also be identified through a variety of analytical techniques.

For example, with financial information, the totals for particular variables may be compared against separately published numbers believed to be reliable.

Unusual amounts above or below pre-determined thresholds may also be reviewed. There are several types of data cleaning that depend on the type of data such as phone numbers, email addresses, employers etc. Quantitative data methods for outlier detection can be used to get rid of likely incorrectly entered data. Textual data spell checkers can be used to lessen the amount of mistyped words, but it is harder to tell if the words themselves are correct.



## **XV. Survey Plan and roadmap/timetable and tasks distribution**

Survey scheduling is one of the critical management tasks as it dictates the time frames in which the project will be completed, the budgets/costs in terms of resource requirements and the sequence of tasks to be completed. Survey scheduling is defined as the process of determining when project activities will take place depending upon defined durations and precedent activities. Schedule constraints specify when an activity should start or end, based on duration, predecessors, external predecessor relationships, resource availability, target dates or other time constraints.

Survey schedules show the timing and sequence of tasks within a project, as well as the project duration and consists of tasks, dependencies among tasks, durations, constraints, milestones and other time-oriented project information. The schedules specify the relative beginning and ending times of activities and their occurrence times. The National Institute of Statistics of Rwanda (NISR) prepares and updates the requirements for surveys. When the researcher is back from the field, collected data are processed ready to be analysed and then produce the report and send it to NISR for approval before dissemination, in order to check whether the work proposed is well done.



## **XVI. Data analysis framework**

Data analysis involves summarizing the data and interpreting their meaning in a way that provides clear answers to questions that initiated the survey. Often, it consists of interpreting tables and various summary measures, such as frequency distributions, means and ranges, or more sophisticated analyses may be performed. The analyst may just want to describe the observed units, in which case all tools from elementary and intermediate statistics are available (tables, diagrams and charts, elementary measures of location and spread, basic modeling, classification models, etc.).

The analyst wants to describe the population and test hypotheses about it, in which the sample design must be properly accounted during the analysis. It is inadvisable for data analysis to commence without the production of a Data Analysis Plan because such a plan is part of good statistical practice. The lack of a plan for analysis can often produce results that are not relevant to the survey objectives, or more importantly, omit tables that are needed.

### **XVI.1. Tabulation Plan**

The Data Tabulation Plan provides model tables that set forth the major findings of a survey in a manner that will be useful to policy makers and program managers. It also helps provide guidance concerning the most important indicators that should be presented in the survey report, the level of analysis expected, and ensures timely dissemination of survey results.

### **XVI.2. Constitution of the analysis plan**

During the survey study, there is a schedule of activities depending on the time of the researcher. In this, the data analysis period should be in plans and different needs including softwares. Every study has the objectives and indicators that the analysis is based on. In data analysis activity, the analyst's activities are scheduled depending on the calendar of the research and the steps have to be followed systematically.

## **NISR Survey Operating Standard Procedure Manual**

---

In such cases, it would be important to ensure that the methodological procedures have been documented in the Data Analysis Plan and implemented during the actual analysis of the data. Other issues can also arise such as the way in which outliers are dealt with, and actions to take in dealing with missing values.

Once all of the items to be measured have been identified, the next task is to determine how much detail is required for each item and the format of the results. What measures, counts, indices, etc. are needed? Are estimates for subpopulations required? The detailed plan of the way the data are to be analyzed and presented is referred to as the analysis plan and, in addition to planned analyses, requires the creation of proposed tabulations.

An analysis plan greatly facilitates the design of the questionnaire. For example, with respect to the details of the final results, is it necessary to be able to distinguish between different age groups within seniors? Does the client need to distinguish between men and women, or between different types of transportation (bus, car, bicycle, etc.)? Should continuous or categorical data be used? For example, does the client need to know what a senior's exact income is, or is a range of incomes adequate? (If the client is interested in calculating averages, then exact income is more appropriate.) It is a good idea to prepare a preliminary set of proposed tabulations and other key desired outputs.

Determining how the results are to be presented helps define the level of detail and indeed the whole scope of the survey. Without a clear analysis plan, it may be possible at the end of the survey to generate hundreds of analysis tables, but only a few may relate directly to the survey's objectives. The proposed tabulations should specify each variable to be presented in a table and its categories. The purpose of this step is to create and retain mock-ups of those tables that will form the analysis. Specification at this level allows the statistical agency to begin drafting questions for the survey questionnaire.

## **XVII. Report writing**

A survey report is written for a clear purpose and to a particular audience. Specific information and evidence are presented, analyzed and applied to a particular problem or issue. The information is presented in a clearly structured format making use of sections and headings so that the information is easy to locate and follow. When you are asked to write a report you will usually be given a report brief which provides you with instructions and guidelines. The report brief may outline the purpose, audience and problem or issue that your report must address, together with any specific requirements for format or structure.

Survey findings need to be presented in a way that is readable and technically acceptable. Good reporting is essential to ensure the importance and usefulness of the findings are not lost. The overall aim of the written report is to communicate the survey findings. It will provide a formal record of the survey research, and can provide a foundation for future research efforts. The written report is often the only aspect of a survey research project that gets exposure.

Consequently, the overall quality of a project is often judged on how well the report is written and presented. A good report should:

- Meet the brief and address the objectives,
- Present clear findings with strong conclusions (and recommendations if appropriate)
- Be accurate and free of errors
- Use tables, graphs and text appropriately
- Be written in a way that will convey the results without overwhelming or boring the reader.

### **XVII.1 The structure of a report**

Depending on the type of report, the structure can include:

- A title page.
- Executive summary.
- Contents.

# **NISR Survey Operating Standard Procedure Manual**

---

- An introduction.
- Terms of reference.
- Procedure.
- Findings.
- Conclusions.
- Recommendations.
- References/Bibliography.
- Appendices.
- The sections, of a report usually have headings and subheadings, which are usually numbered

Once you have your structure, write down the headings and start to fill these in with the information you have gathered so far. By now you should be able to draft the terms of reference, procedure and findings, and start to work out what will go in the report's appendix.

It is always important to revise your work. Things you need to check include:

- If you have done what you were asked to do. Check the assignment question, the instructions/guidelines and the marking schedule to make sure.
- That the required sections are included, and are in the correct order.
- That your information is accurate, with no gaps.
- If your argument is logical. Does the information you present support your conclusions and recommendations?
- That all terms, symbols and abbreviations used have been explained.
- That any diagrams, tables, graphs and illustrations are numbered and labelled.
- That the formatting is correct, including your numbering, headings, are consistent throughout the report.
- Organize proofreading of the report, and your writing is as clear and effective as possible.

You might need to prepare several drafts before you are satisfied. If possible, get someone else to check your report.

## XVIII. Data Dissemination and Archiving

### XVIII.1. Archiving

When preparing any study the archiving plan has to be prepared in order to store collected data that could be consulted in the future. There is different way which researchers are required to archive their data. The archived data should be stored through questionnaire or electronically using software that have been conceived by computer programmers.

Before any action, the following are questions could be asked:

- What to archive;
- Where to archive;
- How to archive;
- Which Timeframe/Timeline;
- When to archive,
- Which means?

Then continue by identifying the Value of institutional record, memories and cultural heritage of institutions.

Predict the impact of loss of records, the cost of replacement just in case?

The legal requirements of what to archive?;

How much effort would we put into recovery just in case?

Analyze the Cost of preservation; Access and security requirements as well as attendant costs just in case

Analyze the Hardware docs, software's docs' dependencies and stability

**Archiving system:** Achieving is the process of moving inactively information to less accessible less storage medium. It is a vital records management process that keeps the crucial information for long periods of time. These three following type of data have to be accurately recorded:

1. The main survey or census planning data
2. The whole journey of acquiring process (documentation of the whole process of the research study)
3. The data set and relevant metadata

The archiving process should be beginning at the first steps of any research subject and end at results publication, that process as well as its related

## **NISR Survey Operating Standard Procedure Manual**

---

data, data set and metadata might be kept on two mean of long time storage one for online archiving system and another for offline archiving medium (including magnetic tape and /or disk RW DVD disk)

Currently, NISR is using online archiving system: National Data Archive (NADA) which is with an input of DDI and RDF xml files developed using nesstar publisher as for formulating template that might be very easily adapted to the structure of database used and facilitating the insertion activities. Unfortunately, this archiving is not complete because there are some documents that are not uploaded to our dissemination system. Those are including the planning documents, budgeting and cost estimation documents and some workflow documents.

The latter need to be achieved in a very systematic system because we experienced a loss of documents which cost us a lot time especially at the beginning of each survey and census. The archiving is not the end point of survey instead it is an ongoing inseparable activities which should always keet into consideration at any stage of research project. We use only one archiving tool as we said above but it is not sufficient, the suggestions are to use also a combination of different media where application code and all syntax could be archived and served as references later and for more improvement and innovation.

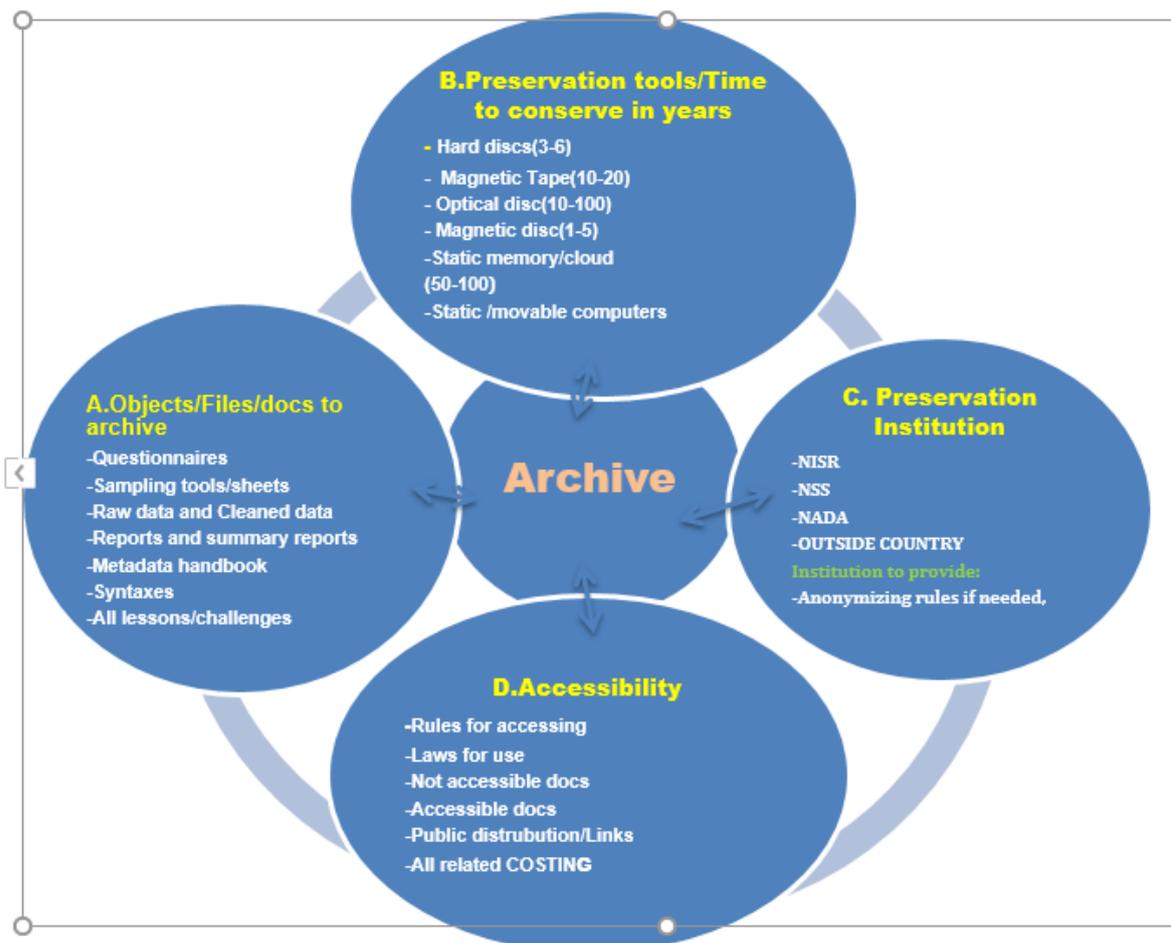
# NISR Survey Operating Standard Procedure Manual

## MAIN POINTS TO CONSIDER WHILE ARCHIVING SOFT DOCUMENTS

Action	Brief description
1.Backup System(in case of soft on fixed machines)	Must be Frequent, Must be redundant Must be geographically distributed for just in case. .
2.Disposition of Originals and Copies	Must Retain original Must also retain all conformed derivative versions(copies) Must also retain all migrated versions (changed formats or versions).
3.Planned regular Checks	Must Frequently checks, 2-3 times per year for volatile storage like NADA Must check if any copies are stored on magnetic tape, sticks, hard discs, by specifically checking the readability as well as fixity.
4.Format Support and Migration	Must store in formats which are best to preserves the integrity of the original file and the intellectual content of any migration copies in the future Must store in reduced files like zipping, or any other. Must migrate original copy in the most preferred preservation format
5.General Description and Laws	Must Critically precise legal requirements to keep materials indefinitely. Must state the laws for any loss or corrupted files which may constitute a major compromise in future understanding of the history of the institution or activities

# NISR Survey Operating Standard Procedure Manual

7.Preservation of Metadata	Must keep as much as possible details, mainly the Keep full records.
8.Able to practice recovery tests	Must verify on recovery regardless of system changes or IT Structure or problems at least once /year.



## XVII.2. Dissemination

Dissemination goes beyond publication of your research. Research may aim to sow the seeds of changes in policies, services, or beliefs about whoever has been studied.

- Does that apply to your research?
- Do you think it could inform wider changes?
- Do you want it to?

Dissemination raises important ethics questions. Not least, if research is disseminated, it has to involve people, and it may challenge or upset them. But while dissemination raises complex issues, it has also been argued that people have an ethical duty to try to make their research findings widely known and, if possible, acted upon. This consideration is particularly important for social science researchers, because of the nature of our work.

## **XVIII.2.1. Methods of dissemination**

As the importance of dissemination and impact becomes increasingly recognized, researchers are using ever more innovative methods – including websites, video, conferences, drama, and even exhibitions. As with other forms of reporting, think about the best way to reach your desired target audiences. What is their availability in terms of time? How can you accommodate their levels of interest, and levels of understanding? Would people come to a seminar or meeting? Could you reach workers through an article or news item in a professional magazine? Are there any organizations that could help you with dissemination?

Those question cloud be answered before dissemination and its necessary to identify stakeholders and user groups to whom the survey products would be disseminated, and also highlighted the need to set up and agree on policies and procedures for dissemination. However, on completion of the production of survey outputs, it is well to re-visit these decisions and consider whether changes are required. Hence this list of user groups and methods for dissemination to be stated and considers the timeliness of dissemination activities.

## **XVIII.2.2 Dissemination plan**

After the validation of the report, it has to be disseminated and the dataset availed on the NISR website ready to be used by any one. The stakeholders and media have to be aware of the content of the report, and have to be invited in dissemination workshop that had to be organized in order to be sensitized on the topic included in the report . Then the distribution of hard copies of the report has to be distributed to the stakeholders and other partners that will need to use it.

## **XVIII.2.3. NISR Overview on Dissemination**

The main task of NISR is to produce official statistics: to collect, process, analyse, publish and archive data. The data Globalisation and the development of information technology have caused changes in the need for statistics: users need an increasing amount of relevant and reliable information more quickly. As the amount of information in the world is continuously growing, it is more and more important to present statistics in a compact and understandable way.

NISR is one of the providers on the information market and has a unique position as the main producer of official statistics. The user's decision whether official statistics are useful for him/her depends on the extent to which the available statistics meet his/her needs. As far as respondents are concerned, NISR is one data collector among others. In the collection of data from private individuals, NISR competes with research companies.

As concerns legal persons, data collection is regulated by the law, but the respondents' cooperative attitude is also a prerequisite for successful cooperation. It is the task of the Marketing and Dissemination Department to identify users' needs and to develop and disseminate the statistical output accordingly. This task is fulfilled in cooperation with other departments of NISR.

For this purpose, dissemination activities have been defined, comprising dissemination (incl. sales), marketing surveys, product design, price policy, selection of distribution channels, and support measures.

The following principles are observed in the dissemination and distribution of official statistics: All users must have equal access to official statistics: this means that the dissemination dates of official statistics have to be announced in advance and no user category (incl. Eurostat, state authorities and mass media) can have access to official statistics before other users. The dissemination of official statistics must be impartial and transparent: this means that neither political comments nor comments of governmental authorities are added to official statistics.

## **NISR Survey Operating Standard Procedure Manual**

---

Official statistics are first published in the Statistical Database. In case a news release is published based on the same data, the relevant news release is published simultaneously with the first release of the data in the Statistical Database. In news releases, statistical publications, etc. the data from the Statistical Database are used. This guarantees that the user finds identical data in all channels

Official statistics are always distributed with metadata and unambiguous explanations in order to avoid misinterpretations and misuse. NISR comments on the misinterpretation, misuse and criticism of official statistics. To ensure statistical confidentiality and prevent human errors, NISR has established specific procedures (the employees sign a confidentiality agreement; the Statistical Database is updated automatically). IT-related requirements are described in the IT-system user guidelines. The production of statistics included in the official statistical programme is financed from the state budget and these statistics are available free of charge. Users have to pay for hard copies of statistical publications and for orders for information not included in the statistical programme. All materials published and information letters (mass mail) sent to clients are subject to editing and proofreading (excl. materials published on the Intranet).

In the production of official statistics, it is important to identify users' needs, to foresee the information need of the society and to act accordingly. Dissemination and communication have been always considered as important activities of the National Institute of Statistics of Rwanda (NISR). They are related to its mandate of producing, disseminating and archiving official statistics on one hand, and coordinating the National Statistical system on another hand. The guiding principle has been that if statistics are not used, then all resources spent in their production will be taken as a lost.

