The United Republic of Tanzania

STATISTICAL METHODS, STANDARDS AND GUIDELINES

National Bureau of Statistics

Ministry of Finance

Dar es Salaam

February, 2012

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ACRONYMS

AfDB	Africa Development Bank
ANA	Annual National Accounts
ANC	Ante Natal Care
ASDP	Agriculture Sector Development Programme
ASIP	Annual Survey of Industrial Production
BOP	Balance of Payments
CPI	Consumer Price Index
CPC	Central Product Classification
CRE	Central Register of Establishment
CSPro	Census and Survey Processing System
COICOP	Consumption of Individual Classification by Purpose
DBS	Dried Blood Spots
DP	Development Partner
EA	Enumeration Area
ESTID	Establishment's Identification
GDP	Gross Domestic Product
GDDS	General Data Dissemination System
GFS	Government Finance Statistics
GIS	Geographical Information
GPS	Global Positioning System
GVA	Gross Value Added
HBS	Household Budget Survey
HBSQF	Household Budget Survey Questionnaire Form
HCPI	Harmonised Consumer Price Index
HIV/AIDS	Human Immunodeficiency Virus/Acquired Immuno Deficiency Syndrome
IBS	Integrated Business Survey
ICP	International Comparison Program
ICR	Intelligent Character Recognition
IEC	Information, Education and Communication
ILO	International Labour Organization
ILFS	Integrated Labour Force Survey
ISIC	International Standard for Industrial Classification

JNIA	Julius Nyerere International Airport
LGA	Local Government Authority
LFS	Labour Force Survey
MAB	Ministerial Advisory Board
MAFS	Ministry of Agriculture and Food Security
MDAs	Ministries, Departments and Agencies
MDG	Millennium Development Goals
MKUKUTA	Mkakati wa Kukuza Uchumi na Kupunguza Umasikini Tanzania
MNRT	Ministry of Natural Resources and Tourism
MTEF	Medium Term Expenditure Framework
NBS	National Bureau of Statistics
NMS	National Master Sample
NPS	National Panel Survey
NPISHs	Non-Profit Institutions Serving Households
NACP	National AIDS Control Programme
NSO	National Statistical Office
NSGRP	National Strategy for Growth and Reduction of Poverty
NSS	National Statistical System
NSSF	National Social Security Fund
OCGS	Office of Chief Government Statistician
PAYE	Pay As You Earn
PIMI	Production Index of Manufacturing Industry
PORALG	President's Office – Regional Administration and Local Government
PSU	Primary Sampling Unit
QNA	Quarterly National Accounts
RSM	Regional Statistical Manager
RMSE	Roots Mean Square Error
SADC	Southern Africa Development Committee
SDDS	Special Data Dissemination System
SMD	Surveys and Mapping Division of Tanzania
SNA	System of National Accounts
SPSS	Statistical Package for Social Statistics
STIs	Sexually Transmitted Infections
STPR	Short Term Price Relative

TASAF	Tanzania Social Action Fund
TASCO	Tanzania Standard Classifications of Occupations
TDHS	Tanzania Demographic and Health Survey
TDS	Tanzania Disability Survey
TISD	Tanzania Integrated Statistical Databse
TFNC	Tanzania Food and Nutrition Centre
THMIS	Tanzania HIV/AIDS and Malaria Indicator Survey
TNA	Training Need Analysis
TNADA	Tanzania National Data Archive
TOE	Training of Enumerators
TOT	Training of Trainers
TSED	Tanzania Socio-Economic Database
TSMP	Tanzania Statistical Master Plan
TUS	Time Use Survey
UN	United Nations
UNFPA	United Nations Fund for Population Activities
UNICEF	United Nations Children Fund
UTM	Universal Transverse Mercator
VBA	Visual Basic for Application
VETA	Vocational Education Training Authority
ZCT	Zanzibar Commission for Tourism

PREFACE

This publication puts together all Methods, Standards and Guidelines that are used in the collection and compilation of official statistics in Tanzania. The objective of the publication is to harmonize and coordinate production of official statistics in the entire National Statistical System (NSS) under the umbrella of Tanzania Statistical Master Plan (TSMP).

Under the Statistics Act No. 1 of 2002, one of the main functions of the National Bureau of Statistics (NBS) is to establish statistical standards and ensure their use by all producers and users of official statistics so as to facilitate the integration and comparison of statistics produced both nationally and internationally.

Experience has shown that producers of statistics in Tanzania have been producing statistics that are sometimes conflicting. This situation has, over the years, brought difficulties for planners and decision makers to decide on which dataset to use.

In order to address this situation, the NBS has produced this document which provides Statistical Methods, Standards and Guidelines to be used by all producers and users of official statistics in Tanzania. The NBS welcomes feedback, comments and suggestions that will lead to improvement of future editions of this publication.

Dr. Albina A. Chuwa
Director General.

BACKGROUND

The Government of Tanzania has been implementing structural and institutional reforms which aim at improving socio-economic development. Ministries, Departments, and Agencies (MDAs), Local Government Authorities (LGAs), and other stakeholders are implementing interventions as outlined in national development frameworks in order to improve livelihood of the people. Tanzania's National Statistical System (NSS) comprises data producers, users, suppliers, providers, and statistical training institutions and centres in Tanzania Mainland and Tanzania Zanzibar.

The NSS has many different actors; MDAs, LGAs, and other stakeholders that collect and process data from primary providers such as villages, wards, health facilities, agricultural extension officers or schools. Much of this information is used as a basis for policy and programme formulation, planning, monitoring and evaluation analysis. The central institution within N|SS in Tanzania Mainland is the National Bureau of Statistics (NBS) while in Tanzania Zanzibar it is the Office of Chief Government Statistician (OCGS).

In order to achieve their objectives of producing statistical information and services; MDAs, LGAs and other stakeholders are required to meet agreed standards while conducting Censuses and Surveys. This is crucial for Statistics comparability and harmonization.

In line with NSS, the NBS aims to come up with methods, standards and guidelines for Tanzania Statistical Programmes to serve the interests of the MDAs, LGAs and other stakeholders. These methods, standards and guidelines outline the statistical professional principles and practices that Tanzania MDAs, LGAs and other stakeholders are required to stick to the level of quality and efforts expected in all statistical activities and services. Each standard has supplementary guidelines that present recommended best practices to fulfill the goals of the standards.

These standards and guidelines provide a means to ensure harmonization, comparability and consistency among and within statistical activities and services conducted in Tanzania. Implementation of standards and guidelines ensures that users of Tanzania statistical information, products and services are provided with details on the principles and methods employed in the planning, collection, processing, analysis, dissemination and preservation of Tanzania statistical information.

The framework for the methods, standards and guidelines presented here includes:

- Methods for statistical production
- Main stages of statistical production
- Standards and guidelines for statistical production
- Standard classifications in statistical production
- Statistical products and services

Within this framework, there are 21 standards and their related guidelines for Tanzania Statistical Surveys which focus on ensuring high quality statistical censuses and surveys that result in information products satisfying MDA`s, LGA`s and other stakeholders for ensuring and maximizing the harmonization, comparability and consistency of quality and integrity of statistical information and products disseminated in Tanzania.

When conducting a Census or Survey, the MDA`s, LGA`s and other stakeholders should engage knowledgeable and experienced survey personnel to effectively achieve the goals of the statistical standards. Persons involved should have knowledge and experience in survey methodology and dissemination as well as technological aspects of censuses and surveys.

The main goal of this document is to produce quality data that are dependent on the methods and statistical concepts used as well as quality control procedures applied in all stages of data production.

PART I: METHODS FOR STATISTICAL PRODUCTION

A: MAIN STAGES OF STATISTICAL PRODUCTION

1.0 Introduction

Statistical production depends on various methods and stages that have an impact on the quality of data, information and indicators produced and used. Therefore, it is imperative that, this process addresses user needs and concerns in the early stages. In some cases, users demand disaggregated data and statistics at the ultimate stages whereas the initial stages did not take into account these needs. Thus, it is likewise essential that various stakeholders including subject matter specialists be involved in the statistical production process to address the user needs. As a principle, quality measures have to be instituted in all stages of data production.

1.1 Methods of Statistical Production

Administrative Records and Routine Data Systems

- These are primarily established to manage operational processes - vital registration, agriculture, education, health, water, tax collection, social security, etc.
- The data generated can also be used for strategic purposes,
- They are the main sources of official statistics,
- Have built-in mechanisms in service delivery to collect data for measuring early results,
- They produce input, process and output indicators as well as outcome indicators to some extent.

Censuses and Surveys

(i) Census

- Collects data from every member of a given population,
- It aims at providing total (100%) coverage of the population,
- Conducted after every ten years in most countries,
- Costly and time consuming,
- They are a basis for intercensal sample surveys,
- Fewer questions are asked and normally produce indicators that measure medium and long term results (outcome and impact) of policies and programmes.

(ii) Survey

- Involves identifying and collecting data from a randomly selected portion (a sample) of a given population,
- Conducted annually for different subjects during the intercensal period,
- Many questions asked that produce indicators for measuring early as well as long term results (outcome and impact) of service delivery.
- Supplements census and administrative data.

(iii) Surveillance and Longitudinal Studies

- These are ongoing, systematic collection, analysis, interpretation, and dissemination of data from a specific area or population,
- Collects data for vital events (births, deaths and migration), health, education and other demographic, social and economic variables.

Examples of data produced from surveillance method include HIV testing at Ante-Natal Care clinics at Kisesa in Magu District; Demographic surveillance sites at Ifakara, Rufiji and Hai districts.

Examples of Longitudinal studies include National and Kagera Panel Surveys.

Experimental and Case studies

An experimental study involves taking measurements of the system under study, manipulating the system and then taking additional measurements using the same procedure to determine if the manipulation has modified the values of the measurements.

A case study is based on an in-depth investigation of a person, a small group, a single situation, or a specific "case,". It involves extensive research, including documented evidence of a particular issue or situation; symptoms, reactions, effects of certain stimuli, and the conclusion reached following the study. A case study may show a correlation between two factors, whether or not a causal relationship can also be proven. Case studies may be descriptive or explanatory.

1.2 Overview of Stages in Statistical Production

In producing statistics, a number of stages have to be considered. These are elaborated below.

1.2.1 Users Demand for Statistical Data

- Internal and external users approach the national statistical offices and statistical units in MDAs requiring data and indicators for planning and decision making purposes.
- The statistical experts have to discuss with data users and other stakeholders to identify data needs to be addressed.
- The statistical experts have to translate the data needs into objectives of the statistical production.
- They also need to determine what method to use to generate the required statistical data.

1.2.2 Establishing Technical Committees

- Determine the composition of technical committee based on type of data required.
- Involve expertise from different socio-economic fields and disciplines.
- Carry out critical analysis of the subject matter in question during the technical committee meetings.
- Planning including budget and roadmap for the statistical production.

1.2.3 Formulation of Statistical Problem

- Data needs are normally presented in non-statistical language,
- Need to come up with statistical formulations to produce the desired data and indicators,
- Determine appropriate study design and the type of data needed to estimate population parameters, statistical association and evaluation.

1.2.4 Information needs

- The technical committee has to determine what information has to be collected that will meet objectives and user needs,

- An overall statement of information needed for socio-economic planning and decision making purposes such as health, education, employment and welfare of the population.

1.2.5 Tabulation and Analysis Plan

Before collecting data there is need to develop tabulation / analysis plan as explained below.

- It is a planned way of summarizing and presenting the collected data,
- It includes frequency tables, cross tabulation and graphs,
- It involves computation of indicators and measures of association (correlations) and determining cause-effect relationships (regressions),
- Determine analysis variables (sex, age, locality, education, income levels, etc) in cross tabulations and regressions to unearth disparities.

1.2.6 Formulation of Statistical Questions

- Converting information needs tables and indicators to questions.
- Mock interviews among experts to test the questions.
- Pre-testing – checking how respondents understand the questions, what responses to expect, sensitivity and neutrality of questions, how to improve them, etc.
- General or specific questions for different respondent groups.

1.2.7 Data collection instruments / questionnaire design

- Combining all questions into a form, questionnaire or checklist.
- Logical flow of questions.
- Separate or single instrument for different respondent categories.
- Compile instruction manuals.
- Develop publicity and advocacy materials.

1.2.8 Sample design

- Resources can determine whether to collect data from the whole or part of total population. In addition, level of accuracy can determine the sample design.
- Identifying and selecting respondents to represent others including stratification of sub-groups.
- Adequate number of different respondent categories (adequate sample size).
- Sampling weights for estimation of population parameters.

1.2.9 Recruitment and training

- Determine number and type of personnel (e.g. supervisors and enumerators) who will be involved in data collection.
- Criteria for recruiting and selecting data collection personnel.
- Training to build the capacity of the personnel for the data collection.
- During training, impart general and specific skills to data collection personnel.

1.2.10 Pilot testing

- Testing all data production and logistical procedures before main fieldwork.
- Determine areas of strengths and weaknesses of the data collection system.
- Improvement of data production and logistical procedures.

1.2.11 Main fieldwork

- Conduct advocacy and publicity campaigns before and during data collection phase to improve the response rates.
- Dispatching data collection equipment, instruments and personnel to and from the field.
- Collect data from the earmarked respondents using appropriate instruments.
- Strengthen field supervision mechanisms and teamwork to improve data quality.
- Conduct post-enumeration survey (evaluation) immediately after main fieldwork to determine coverage, content and quality aspects of the data collected.

1.2.12 Data processing

- Institute field procedures for checking quality of data by supervisors and manual editors.
- Transfer data from data collection instruments into computer files (data entry).
- Institute office procedures for checking quality of data before, during and after data entry.
- Build capacity of data entry operators in terms of speed and accuracy.

1.2.13 Tabulation / Analysis

- Implementing the tabulation / analysis plan.
- Summarizing the collected data into tables and statistics / indicators.
- Disaggregation of data - presenting data such that socio-economic differentials are clearly seen.
- Analyzing within and among socio-economic categories – column or row totals.

- Sex and geographical location as major analysis variables – analyzing important population characteristics by sex and location such that gender and urban/rural differences are clearly reflected.

- Make statistical inference from sample data to total population.

1.2.14 Interpretation and report writing

- This involves extracting main messages from the tabulated / analyzed data.
- Composition of different experts among the authors. An expert eye / lens is very crucial at this stage to pick the critical issues.
- Write separate chapters or reports on related findings.

1.2.15 Dissemination and Statistical Literacy

- Informing users and stakeholders on the results using various means such as reports, media and website.
- General and specific packages for various users.
- Promoting the policy agenda of the produced data.
- Statistical literacy to users to understand the data.

1.2.16 Documentation and Archiving

- Preparing basic information datasheets describing the data.
- Archiving the raw data and reports.
- Institute procedures for accessing the raw data including removing identification (anonymization) of respondents.

B: STANDARDS AND GUIDELINES FOR STATISTICAL PRODUCTION

2.0 Introduction

There are various methods of data production as outlined earlier in this document. This part dwells in detail about these methods by providing relevant standards and guidelines.

2.1 SURVEY METHODOLOGY

2.1.1 Survey Planning

Standard 2.1.1: When starting a new survey or revision of an existing survey; MDAs, LGAs and other stakeholders must develop a written proposal (concept note) that sets forth a justification, including: goals and objectives, potential users, related and previous surveys, key survey estimates, the precision required of the estimates, the tabulation and analytic results that will inform decisions and other uses, steps taken to prevent unnecessary duplication with other sources of information, confidentiality of individual data, when and how frequently users need the data and public use of the data.

The guidelines for this standard are:

Guideline 2.1.1.1: Surveys (and related activities such as focus groups, pretesting, pilot studies, field tests, etc.) are collections of information subject to the requirements of an existing Statistics Act and Tanzania Statistical Master Plan (TSMP). An initial step in planning a new survey or a revision of an ongoing survey should be to contact the sponsoring Development Partner (DP), MDA, LGA and a stakeholder's most senior designated official to ensure the survey work is done in compliance with the law and regulations. NBS approval will be required before the MDAs, LGAs and other stakeholders embark on a data collection exercise from households and establishments.

Guideline 2.1.1.2: Planning is an important prerequisite when designing a new survey or implementing an amendment of an ongoing survey. Key planning activities include the following:

a) A justification for the survey
   i The rationale for the survey
   ii Relationship to previous surveys

iii Survey goals and objectives

iv Hypotheses to be tested; and

v Definitions of key variables.

vi Consultations with potential stakeholders (to identify their requirements and expectations).

b) A review of related studies, surveys, and reports of Tanzania and non-Tanzania sources to ensure that part or all of the survey would not unnecessarily duplicate available data from an existing source, or could not be more appropriately obtained by adding questions to existing Tanzania statistical surveys. The goal here is to spend Tanzania funds effectively and minimize burden to data producers.

c) A review of the confidentiality and privacy policy of an existing Statistics Act on surveys that will collect individually-identifiable data from any survey respondent.

d) A complete and review of all survey data items, the justification for each item, and the means of measurements (e.g., through questionnaires, tests, or administrative records).

e) A plan for pre-testing or cognitive interviewing, if applicable

f) A plan for quality assurance during each phase of the survey process to permit monitoring and assessing performance during implementation.

i The plan should include possibility to modify the survey procedures if design parameters appear unlikely to meet expectations (for example, if low response rates are likely).

ii Should contain general specifications for an internal project management system that identifies critical activities and

iii Key milestones of the survey that will be monitored, and the timeframes among them.

g) A plan for evaluating survey procedures and results

h) An analysis plan that identifies analysis issues, objectives, key variables and proposed statistical tests

i) An estimate of resources and target timeframe needed for completion of the survey cycle.

j) A dissemination plan that identifies target audiences, proposed major information products, and the timing of their release.

k) A data management plan for the preservation of survey data, documentation, and information products as well as the authorized disposition of survey records.

Guideline 2.1.1.3: Include standard elements of project management in the plan, including target completion dates, the resources needed to complete each activity, and risk planning.

Guideline 2.1.1.4: To maintain a consistent data series over time, use consistent data collection procedures for ongoing data collections on core statistics. Continuous improvement efforts sometimes result in a trade-off between the desire for consistency and a need to improve a data collection. If changes are needed in key variables or survey procedures for a data series, consider the justification or rationale for the changes in terms of their usefulness for policymakers, conducting analyses, and addressing information needs. Develop adjustment methods, such as crosswalks and bridge studies that will be used to preserve trend analyses and inform users about the effects of changes.

2.1.2 Survey Designing

Standard 2.1.2: MDAs, LGAs and other stakeholders must develop a survey design, including defining the study frame, target population, sampling plan, identify the data collection instruments and methods, developing a practical timetable, estimating survey cost, and selecting samples using accepted statistical methods (e.g., probabilistic methods that can provide estimates of sampling error). Any use of non-probability sampling methods (e.g., judgmental, Quota and Snowball etc. samples) must be justified statistically and be able to measure estimation error. The size and design of the sample must reflect the level of detail needed in tabulations and other data products, and the precision required of key estimates. Documentation of each of these activities and resulting decisions must be maintained in the project files for use in documentation.

The guidelines for this standard are:

Guideline 2.1.2.1: Include the following in the survey design:

a) Frame for selection

b) Proposed target population;

c) Stratification levels/domain of study and analysis
d) Response rate from previous survey or expected response rate;
e) Survey frequency
f) Timing of data collection;
g) Data collection modes (such as paper and pencil, mail survey, telephone survey, etc);
h) Sample design;
i) Precision requirements;
j) Effective sample size determination based on power analyses for key variables; and
k) Overall sample size.

Guideline 2.1.2.2: Ensure the sample design will yield the data required to meet the objectives of the survey. Include the following in the sample design:

a) identification of the sampling frame (address, name, location);
b) identify the sampling unit used (at each stage if a multistage design);
c) identify sampling strata;
d) power analyses to determine sample sizes;
e) effective sample sizes for key variables by reporting domains (Urban/Rural where appropriate);
f) criteria for stratifying or clustering, sample size by stratum, and the known probabilities of selection;
g) response rate goals (see Standard 2.1.3); estimation and weighting plan; variance estimation techniques appropriate to the survey design; and
h) Expected precision of estimates for key variables.

Guideline 2.1.2.3: When a non-probabilistic sampling method is employed, include the following in the survey design documentation:

a) a discussion of what options were considered and why the final design was chosen,
b) an estimate of the potential bias in the estimates.

Guideline 2.1.2.4: Include a statement of confidentiality along with instructions required to complete the survey.

Guideline 2.1.2.5: Include the following in the data collection plans:

a) frequency and timing of data collection;

b) methods of collection for achieving acceptable response rates;
c) training of enumerators and persons, coding and editing the data;
d) cost estimates, including the costs of pretests,
e) non-response follow-up, and
f) evaluation studies.

2.1.3 Response Rates

Standard 2.1.3: MDAs, LGAs and other stakeholders must design the survey to achieve the highest rates of response to ensure that survey results are representative of the target population so that they can be used with confidence to inform decisions. Non-response bias analyses must be conducted when unit or item response rates or other factors suggest the potential for bias to occur.

The guidelines for this standard are:

Guideline 2.1.3.1: Calculate sample survey unit response rates without substitutions.

Guideline 2.1.3.2: Design data collections that will be used for sample frames for other surveys (e.g., the Population and Housing Census enumeration areas (EAs), and the Central Register of Establishments) to meet a target unit response rate of at least 80 percent, or provide a justification for a lower anticipated rate.

2.1.4 Focus Group Discussions (for instrument development)

Standard 2.1.4: MDAs, LGAs and other stakeholders must ensure that the survey collects the required information for their intended producers and users. The purpose of this standard is to get key issues regarding the planned survey before developing the survey questionnaire.

The guidelines for this standard are:

Guideline 2.1.4.1: Identify key stakeholders in the subject matter area who will participate in focus group discussion.

Guideline 2.1.4.2: Prepare semi-structured (focused) discussion with members of the target population to expose what they know about the study that the questionnaire will cover, how they think about the study and what terms they use in talking about the study topics/variables.

Guideline 2.1.4.3: Recruit volunteers (10-20 from data collectors, producers and users side) who are at least familiar or are expected to be data producers or users of the study; to participate in a systematic discussion guided by a moderator about the survey topic(s) (questions for discussion should be prepared prior to convene volunteers). Lessons learned from the discussion will be the basis for questionnaire design.

2.1.5 Designing Survey Instrument(s) (Questionnaire(s))

Standard 2.1.5: Based on the experiences and lessons drawn from the literature review and focus group discussion, but mainly reflecting on the objectives of the proposed study, MDAs, LGAs and other stakeholders should design a questionnaire that will capture the intended information to be collected. The instrument shall probe and systematically record comprehensive information that answers the study questions.

The guidelines for this standard are:

Guideline 2.1.5.1: Identify subject matter specialists (mostly statisticians, researchers, sociologists, economists, etc) who will draft the questionnaire.

Guideline 2.1.5.2: Review literatures and instruments from previous similar studies for comparability purposes.

Guideline 2.1.5.3: Check the identification and demographic variables of the existing instruments if they meet the requirements of the intended study and update accordingly.

Guideline 2.1.5.4: Design the questionnaire using the available questions bank in Tanzania (e.g. NADA) or outside Tanzania.

Guideline 2.1.5.5: Prepare instruction manuals for data collectors and supervisors.

2.1.6 Pre-testing of Survey Instruments

Standard 2.1.6: MDAs, LGAs and other stakeholders must ensure that the draft questionnaire is pre-tested to randomly chosen participants by interviewers to probe the understanding of the study questions to respondents, time to complete one questionnaire and an attempt to learn how they formulate their answers. Then the recording of the outcome of the interview is done for questionnaire improvement. By conducting a pretest of the survey components, measurement error will be controlled.

The guidelines for this standard are:

Guideline 2.1.6.1: Randomly choose participants to participate in pre-test interviews.

Guideline 2.1.6.2: Key researchers and survey desk officers should participate fully in the cognitive interview and if possible recording the interview for quality checking.

Guideline 2.1.6.3: Arrange technical meeting to discuss the experiences learned from the cognitive interviews and use the results to improve the questionnaire.

Guideline 2.1.6.4: Record starting and finishing times for questionnaire interviews to determine the average time spent per questionnaire. The technical committee can then allocate number of questionnaires to be completed per interviewer per day.

2.1.7: Training of Trainers (TOT), Supervisors and Training of Enumerators (TOE)

Standard 2.1.7: MDAs, LGAs and other stakeholders should recruit field staff to participate in TOT and TOE (supervisors and enumerators) on the basis of their competence and experience in the planned data production exercise such as a census or a survey.

The guidelines for this standard are:

Guideline 2.1.7.1: Identify key staff (trainers) who will train supervisors and enumerators

Guideline 2.1.7.2: If supervisors and enumerators are not enough or available within an organization, consider to hire and recruit them. In addition, recruit reserve supervisors and enumerators for replacement in case of dropouts during main survey.

Guideline 2.1.7.3: Prepare conducive environment for training in terms of geography, conference facilities and accommodation for participants.

Guideline 2.1.7.4: A maximum group of 5 enumerators should be supervised by one supervisor.

Guideline 2.1.7.5: Prepare mock exam to test the understanding of trainees

Guideline 2.1.7.6: Prepare form for trainers' evaluation. This will help to show relationships of trainers-trainees and understanding of areas for improvement.

2.1.8 Pilot Testing

Standard 2.1.8: MDAs, LGAs and other stakeholders must administer at least a sample covering zonal areas using all field procedures similar to the main census or survey instruments and hold debriefing between interviewers and project team.

The guidelines for this standard are:

Guideline 2.1.8.1: Test all field instruments, logistical procedures and human resource capacity. Items to be tested include frame development, sample selection, questionnaire design, data collection, edit specifications, data processing, estimation, file creation and tabulations. A complete test of all components is desirable for all surveys that cover the whole country.

Guideline 2.1.8.2: Tabulate pilot results to see patterns of missing data values, recording and behavior coding to detect patterns of questions that are difficult to answer.

Guideline 2.1.8.3: Use the lessons learned from the pilot to improve the instruments and the approach to the study.

2.2 DATA COLLECTION

2.2.1 Sampling Frames

Standard 2.2.1: MDAs, LGAs and other stakeholders must ensure that the frames for the planned sample survey or census are suitable for the study design and are assessed against the target population for quality checking.

The guidelines for this standard are:

Guideline 2.2.1.1: Describe target populations and associated survey or sampling frames. Include the following items in this description:

a) Describe the approach in which the frame was created and the methodology of frame updating

b) Describe exclusions that have been applied to frame and target populations;

c) Describe frame problems (missing units on the frame (under-coverage), and duplicates on the frame (over-coverage)

d) Describe what was done to improve the coverage of the frame

e) Describe how data quality and item non-response on the frame may have affected the coverage of the frame; and

f) Explain limitations of the frame including the timeliness and accuracy of the frame (e.g., misclassification, eligibility, etc.).

Guideline 2.2.1.2: Conduct regular evaluations of coverage rates and coverage of the target population in survey frames that are used for ongoing surveys every 3 years.

2.2.2 Awareness to Prospective Survey Respondents

Standard 2.2.2: MDAs, LGAs and other stakeholders must ensure that all prospective survey respondents are aware of the study and they understand the purpose of the survey.

The guidelines for this standard are:

Guideline 2.2.2.1: Provide pre-notification letter / survey brochures to respondents

a) Informs potential respondents that they have been selected to participate in a survey;
b) Informs potential respondents about the name and nature of the survey; and
c) Assure them on the confidentiality of information to be collected

Guideline 2.2.2.2: Intensify Information, Education and Communication campaign (IEC) through media (such as televisions, radios, magazines, etc)

Guideline 2.2.2.3: Involve leaders at LGAs (Ward, Village Executive Officers).

2.2.3 Methods of Data Collection

Standard 2.2.3: MDAs, LGAs and other stakeholders must design and administer their data collection instruments and methods in a manner that achieves the best balance between maximizing data quality and controlling measurement error while minimizing respondent burden and cost.

The guidelines for this standard are:

Guideline 2.2.3.1: Design the data collection instruments in a manner that minimizes respondent burden, while maximizing data quality. The following strategies may be used to achieve these goals:

a) Questions should be written clearly
b) Observe logical flow of questions and design proper skip patterns;
c) Don't overload the questionnaire;
d) The questionnaire should include only items/variables that have been pre/pilot tested

Guideline 2.2.3.2: Encourage respondents to participate in order to maximize response rates and improve data quality. The following data collection strategies can also be used to achieve high response rates:

a) Ensure that the data collection reference period is of adequate and reasonable length (at most 12 months);

b) Allow three interview attempts (call-backs) before declaring unit non-response;

c) Use competent interviewers and other staff who can learn techniques for obtaining cooperation and building rapport with respondents. Techniques for building rapport include respect for respondents' rights and culture, observing appointments, follow-up skills, knowledge of the goals and objectives of the data collection and uses of the data.

d) Although incentives are not recommended and used in surveys, MDAs, LGAs and other stakeholders may consider use of respondent incentives if they believe incentives would be necessary to use for a specific survey in order to achieve data of sufficient quality for their intended use(s). Some incentives that can be offered to respondents may include:

i Small portable radios
ii T-shirts and caps
iii Hoes
iv Mosquito nets
v Key holders

vi School bags/safari bags

vii Football

viii Exercise books

ix Pens and pencils

x Survey badges

Guideline 2.2.3.3: The way data collection is designed and administered contributes to data quality. The following are important to consider:

a) Collect data at the most appropriate time of year, when relevant;

b) Establish the data collection protocol to be followed by the field staff;

c) Provide training for field staff on survey protocols,

d) Establish mechanisms to minimize interviewer falsification, such as protocols for monitoring interviewers and re-interviewing respondents;

e) Establish procedures for field edits of data collected. Enumerators and supervisors should ensure that questionnaires are duly filled before moving to another respondent or cluster.

Guideline 2.2.3.4: Develop supervision for data collection activities, with strategies to correct identified problems. The following are important to consider:

a) Design control report forms and supervision checklists.

b) Implement quality by following the process of data collection manuals

c) Use internal reporting systems that provide timely reporting of response rates and the reasons for non-response throughout the data collection.

2.3 DATA PROCESSING

2.3.1 Data Editing

Standard 2.3.1: MDAs, LGAs and other stakeholders must edit data appropriately, based on available information, to correct detectable errors.

The guidelines for this standard are:

Guideline 2.3.1.1: Check and edit data to correct errors during and after data collection. Data editing is an iterative and interactive process that includes procedures for detecting and correcting errors in the data. When electronic data collection methods are used, data are usually edited both during and after data collection. Obtain inputs from subject matter specialists in the development of edit rules and edit parameters (edit specifications). As appropriate, check data for the following and edit if errors are detected:

a) Responses that fall outside a pre-specified range (e.g., a person with 4 years old and married with 2 children);
b) Contradictory responses and incorrect flow through prescribed skip patterns;
c) Missing data that can be directly filled from other portions of the same record (including the sample frame e.g missing location identification);
d) The omission of records; and
e) The duplication of records.

Guideline 2.3.1.2: Code the data set to indicate any actions taken during editing, and/or retain the unedited data along with the edited data (e.g. adding a column in the data set to identify the imputed/edited values).

2.3.2 Data Coding

Standard 2.3.2: MDAs, LGAs and other stakeholders must add codes to collected data to identify aspects of data quality from the collection (e.g., missing data) in order to allow users to appropriately analyze the data. Codes added to convert information collected as text into a form that permits immediate analysis must use standardized codes, when available, to enhance comparability.

The guidelines for this standard are:

Guideline 2.3.2.1: Insert codes into the data set that clearly identify missing data and cases where entry is not expected (e.g., skipped over by skip pattern). Do not use blanks and zeros as codes to identify missing data, as they tend to be confused with actual data.

Guideline 2.3.2.2: When converting text data to codes to facilitate easier analysis, use standardized codes, if they exist. Use the Tanzania coding standards listed below, if

applicable. Provide cross-referencing tables to the Tanzania standard codes for any legacy coding that does not meet the Tanzania standards. Develop other types of codes using existing Tanzania MDA, LGA and a stakeholder practice or standard codes from industry or international organizations, when they exist. Current Tanzania standard codes include the following:

a) Region, District, Ward, EA/Village Codes which are maintained by NBS. The codes can be posted on the NBS website www.nbs.go.tz .

b) International Standards for Industry Classification (ISIC Codes) - Use the ISIC to classify establishments. The ISIC is UN comparability in statistics about business activities across the globe. The codes can be downloaded from a website link.

c) Classification of Individual Consumption by Purpose (COICOP).
d) System of National Accounts (SNA), 1993 and 2008
e) Hotels and Tourism – Three plus stars, etc
f) HS codes for external trade
g) Central Product Classification (CPC) codes for industry
h) Geo Information System (GIS)
i) Government Finance Statistics (GFS)
j) Tanzania Standard Classification of Occupation (TASCO)

2.3.3 Data Entry

Standard 2.3.3: MDAs, LGAs and other stakeholders must use acceptable and easy compatible software to allow data transfers to different statistical applications. Some data entry software may include CSPrO, MS Excel and MS Access.

The guidelines for this standard are:

Guideline 2.3.3.1: Data may be entered twice (double entry) to check for consistency

2.3.4 Data Cleaning – Range, Consistency Checks and Validation

Standard 2.3.4: MDAs, LGAs and other stakeholders must make sure that all data entered into the system are consistent before further analysis. All demographic variables should

reflect the data items. For example for demographic enquiries, a male should not have pregnancies in his lifetime.

The guidelines for this standard are:

Guideline 2.3.4.1: Establish rules for range, consistency and validation checks to be applied to the data during and after data entry.

Guideline 2.3.4.2: Prepare list and printout of errors found in data entered and submit to quality control personnel for further action.

Guideline 2.3.4.3: Make appropriate corrections without altering the collected data from the field. Treat the remaining erroneous data as partial non-response.

2.3.5 Data Protection

Standard 2.3.5: MDAs, LGAs and other stakeholders must observe individual data confidentiality throughout the production process to ensure that survey data are handled to avoid disclosure.

The guidelines for this standard are:

Guideline 2.3.5.1: For surveys that include confidential data, establish procedures and mechanisms to ensure the information's protection during the production, use, storage, transmittal, and disposition of the survey data in any format (e.g., completed survey forms, electronic files, and printouts).

Guideline 2.3.5.2: Ensure that

a) Individually-identifiable survey data are protected;
b) Data systems and electronic products are protected from unauthorized intervention; and
c) Data files, network segments, servers, and desktop PCs are electronically secure from malicious software and intrusion using best available information resource security practices that are periodically monitored and updated.

Guideline 2.3.5.3: Controlled access to data sets so that only specific, authorized individuals working on a particular data set can have read only, or write only, or both read and write access to that data set. Data set access rights are to be periodically reviewed by the IT manager responsible for that dataset in order to guard against unauthorized release or alteration.

2.3.6 Quality Evaluations

Standard 2.3.6: MDAs, LGAs and other stakeholders must evaluate the quality of the data and make the evaluation public (through technical notes and documentation included in reports of results or through a separate report) to allow users to interpret results of analyses, and to help designers of future surveys to focus on improvement efforts.

The guideline for this standard are:

Guideline 2.3.6.1: Include an evaluation component in the survey plan that evaluates survey procedures and results. Review past surveys similar to the one being planned to determine likely sources of error, appropriate evaluation methods, and problems that are likely to be encountered. Address the following areas:

a) Potential sources of errors, *both sampling and non-sampling* may include:
   - i Coverage error (including frame errors);
   - ii Non response error;
   - iii Measurement error, including sources from the instrument, interviewers, respondents, changes associated with time of the object or phenomenon being measured, type of questions –biased or leading ones and collection process; and
   - iv Data processing error (e.g., keying, coding, editing, and imputation error);

b) How sampling error will be measured, including variance estimation and studies to isolate error components; and

c) Make evaluation studies public to inform data users.

2.4 ESTIMATES AND PROJECTIONS

2.4.1 Developing Estimates and Projections

Standard 2.4.1: MDAs, LGAs and other stakeholders must use accepted theory and methods when deriving direct survey-based estimates, as well as model-based estimates and projections that use survey data. Error estimates must be calculated and disseminated to support assessment of the appropriateness of the uses of the estimates or projections. MDAs, LGAs and other stakeholders must plan and implement evaluations to assess the quality of the estimates and projections.

The guidelines for this standard are:

Guideline 2.4.1.1: Develop direct survey estimates by employing sampling weights appropriate for the sample design to calculate population estimates However, an MDA, LGA and a stakeholder may employ an alternative method (e.g., ratio estimators) to calculate population estimates if they have evaluated it and determined that it leads to acceptable results.

Guideline 2.4.1.2: Calculate variance estimates by a method appropriate to a survey's sample design taking into account probabilities of selection, stratification, clustering, and the effects of non-response, post-stratification and ranking. The estimates must reflect any design effect resulting from a complex design.

Guideline 2.4.1.3: Document methods used to generate estimates and projections to help ensure objectivity, utility, transparency and reproducibility of the estimates and projections (For details on documentation, see 2.7.2). Also, archive data so the estimates/projections can be reproduced.

For population projections using e.g. exponential method or natural growth method compare advantages and disadvantages of each method.

2.5 DATA ANALYSIS

2.5.1 Analysis and Report Planning

Standard 2.5.1: MDAs, LGAs and other stakeholders must develop a plan for the analysis of survey data prior to the start of a specific analysis to ensure that statistical tests are used appropriately and that adequate resources are available to complete the analysis.

The guidelines for this standard are:

Guideline 2.5.1.1: Include the following in the analysis plan:

a) An introduction that describes the purpose, the research question, relevant literature, data sources (including a brief description of the survey data and any limitations of the data), key variables to be used in the analysis, type of analysis, and significance level to be used;
b) Tables and figures that support the analysis; and
c) A framework for technical notes including,

i History of the survey program;
ii Data collection methods and procedures;
iii Sample design;
iv Response rates and the treatment of missing data;
v Weighting methods;
vi Computation of standard errors;
vii Instructions for constructed or derived variables;
viii Limitations of the data; and
ix Sources of error in the data.

2.5.2: Inference and Comparisons

Standard 2.5.2: MDAs, LGAs and other stakeholders must base statements of comparisons and other statistical conclusions derived from survey data on acceptable statistical practice.

The guidelines for this standard are:

Guideline 2.5.2.1: Specify the criterion for judging statistical significance for tests of hypotheses (Type I error) before conducting the testing.

Guideline 2.5.2.2: If part of an historical series is revised, data for both the old and the new series should be published for a suitable overlap period for the use of analysts.

2.6 REVIEW AND REBASING PROCEDURES

2.6.1 Review of Information Products

Standard 2.6.1: MDAs, LGAs and other stakeholders are responsible for the quality of information that they disseminate and must institute appropriate content/subject matter, statistical, and methodological review procedures to comply with NBS and MDAs, LGAs and other stakeholders.

The guidelines for this standard are:

Guideline 2.6.1.1: Conduct a subject-matter review of all information products that present a description or interpretation of results from the survey, such as analytic reports or "briefs." Select reviewers with appropriate expertise in the subject matter, operation, or statistical program discussed in the document. Among the areas that reviewers should consider are the following:

a) Subject-matter literature is referenced in the document appropriately;
b) Information is factually correct; and
c) Information is presented clearly and logically, conclusions follow from analysis, and no inconsistent findings are ignored.

Guideline 2.6.1.2: Conduct a statistical and methodological review of all information products. Select reviewers with appropriate expertise in the methodology described in the document. Among the tasks that reviewers should consider are the following:

a) Review assumptions and limitations for accuracy and appropriateness;
b) Ensure that appropriate statistical methods are used and reported;
c) Review calculations and formulae for accuracy and statistical soundness;
d) Review data and presentations of data (e.g., tables) for disclosure risk;

e) Review contents, conclusions, and technical (statistical and operational areas) recommendations to ensure that they are supported by the methodology used; and

f) Ensure that data sources and technical documentation, including data limitations, are included or referenced.

Guideline 2.6.1.3: Review all information products that will be disseminated electronically for compliance with existing policies and legislations governing statistical activities in the country for easy accessibility. Ensure that any product that is disseminated via special software is tested for accessibility and interpretability prior to dissemination.

Guideline 2.6.1.4: Rebase and update information products with new and current data produced from survey and routine data.

Guideline 2.6.1.5: Review and harmonize information products using common standards and levels of classification such as ISIC, SNA, etc. Avoid aggregating information products using different standards and classifications before harmonizing them. If it is necessary to customize to local conditions or branch out of main classifications, maintain the link to international standard.

2.7 DISSEMINATION

2.7.1 Information Release

Standard 2.7.1: MDAs, LGAs and other stakeholders must release information intended for the general public according to a dissemination policy that provides for equivalent, timely access to all users and provide information to the public about the MDAs, LGAs and other stakeholders' dissemination policies and procedures including those related to any planned or unpredicted data revisions.

The guidelines for this standard are:

Guideline 2.7.1.1: Dissemination procedures for major information products may include the following:

a) Develop release calendar and method for the release of information products and services;
b) Inform targeted stakeholders; and
c) Ensure equal and timely access to all users.

Guideline 2.7.1.2: Protect information against any unauthorized pre-release, and release information only according to established release procedures.

Guideline 2.7.1.3: If revisions to estimates are planned, establish a schedule for expected revisions, make it available to users, and identify initial releases as preliminary.

Guideline 2.7.1.4: Establish a policy for handling unscheduled corrections due to previously unrecognized errors. The policy may include threshold criteria (e.g., the correction will change a national level total value by more than one percent or a regional value by more than five percent) identifying conditions under which data will be corrected and re-disseminated.

Guideline 2.7.1.5: When information products are disseminated, provide users with access to the following information:

a) Definitions of key variables;
b) Source information, such as a survey form number and description of methodology used to produce the information or links to the methodology;
c) Quality-related documentation such as conceptual limitations;
d) Variance estimation documentation;
e) Time period covered by the information and units of measure;
f) Point of contact to whom further questions can be directed;
g) Software or links to software needed to read/access the information and installation/operating instructions, if applicable;
h) Date the product was last updated; and
i) Standard dissemination policies and procedures.

2.7.2 Documentation and Archiving

Standard 2.7.2: MDAs, LGAs and other stakeholders must produce survey documentation that includes those materials necessary to understand how to properly analyze data from each survey, as well as the information necessary to replicate and evaluate each survey's results. Survey documentation must be readily accessible to users, unless it is necessary to restrict access to protect confidentiality.

The guidelines for this standard are:

Guideline 2.7.2.1: Survey system documentation (metadata) includes all information necessary to analyze the data properly. It includes the following:

i. Survey instruments;

ii. Description of variables used to uniquely identify records in the data file;
iii. Description of the sample design, and sampling unit identifiers to be used for analysis;
iv. Definitions of all variables, including all modifications;
v. Data file layout;
vi. Descriptions of constructed variables on the data file that are computed from responses to other variables on the file;
vii. Description of sample weights, including adjustments for non-response and benchmarking and how to apply them;
viii. Description of how to calculate variance estimates appropriate for the survey design;
ix. Description of all editing and imputation methods applied to the data (including evaluations of the methods) and how to remove imputed values from the data;

Guideline 7.2.2: To ensure that a survey can be replicated and evaluated, the MDA, LGA and a stakeholder's internal archived portion of the survey system documentation, at a minimum, must include the following:

i. Survey planning and design decisions, including the NBS Information Collection Request package;
ii. Field test design and results;
iii. Selected sample;
iv. Sampling frame;
v. Justifications for the items on the survey instrument, including why the final items were selected;
vi. All instructions to respondents and/or interviewers either about how to properly respond to a survey item or how to properly present a survey item;
vii. Description of the data collection methodology;
viii. Sampling plan and justifications, including any deviations from the plan;
ix. Data processing plan specifications and justifications;
x. Final weighting plan specifications, including calculations for how the final weights were derived, and justifications;

xi. Final imputation plan specifications and justifications;
xii. Data editing plan specifications and justifications;
xiii. Evaluation reports;
xiv. Descriptions of models used for indirect estimates and projections;
xv. Analysis plans;
xvi. Time schedule for revised data; and
xvii. Documentation made publicly available in conjunction with the release of data.

Guideline 2.7.2.3: For recurring surveys, produce a periodic evaluation report, such as a methodology report, that itemizes all sources of identified error. Where possible, provide estimates or bounds on the magnitudes of these errors; discuss the total error model for the survey; and assess the survey in terms of this model.

Guideline 2.7.2.4: Retain all survey documentation according to appropriate Tanzania records disposition and archival policy and law.

PART II: STANDARD CLASSIFICATIONS IN STATISTICAL PRODUCTION

3.0 THE INTERNATIONAL STANDARD INDUSTRIAL CLASSIFICATION OF ALL ECONOMIC ACTIVITIES (ISIC)

ISIC is the international reference classification of a coherent and consistent structure of economic activities based on a set of internationally agreed concepts, definitions, principles and classification rules.

ISIC structure

The hierarchy in ISIC consists of:

- Section (one letter code)
- Division (two digits code)
- Group (three digits code)
- Class (four digits code)

Objective and use of ISIC

ISIC is intended to be a standard classification of productive activities. Its main purpose is to provide a set of activity categories that can be utilized for the collection and presentation of statistics according to such activities. Therefore, ISIC aims to present this set of activity categories in such a way that entities can be classified according to the economic activity they carry out.

It provides a comprehensive framework within which economic data can be collected and reported in a format that is designed for purposes of economic analysis, decision-taking and policy-making. The classification structure represents a standard format to organize detailed information about the state of an economy according to economic principles and perceptions

The ISIC has widely being used both nationally and internationally in classifying data according to kind of economic activity in the fields of production, employment, gross domestic product and other statistical areas. ISIC is a basic tool for studying economic phenomena, fostering international comparability of data, providing guidance for the development of national classifications and for promoting the development of sound national statistical systems.

Classification rule

The classification is used to classify statistical units, such as establishments or enterprises, according to the economic activity in which they mainly engage. At each level of ISIC, each statistical unit is assigned to one and only one ISIC code,

The set of statistical units that are classified into the same ISIC category is then often referred to as an industry, such as "the furniture industry", which would refer to all units classified in ISIC division 31 (Manufacture of furniture)

This standardized categorization or subdivision of the complete set of producing units in an economy makes ISIC an important tool for socio-economic statistics that need to be arranged in accordance with the productive system of the economy.

The individual categories of ISIC have been aggregated into the following 21 sections:

A 01–03	Agriculture, forestry and fishing
B 05–09	Mining and quarrying
C 10–33	Manufacturing
D 35	Electricity, gas, steam and air conditioning supply
E 36–39	Water supply; sewerage, waste management and remediation activities
F 41–43	Construction
G 45–47	Wholesale and retail trade; repair of motor vehicles and motorcycles
H 49–53	Transportation and storage
I 55–56	Accommodation and food service activities
J 58–63	Information and communication
K 64–66	Financial and insurance activities
L 68	Real estate activities
M 69–75	Professional, scientific and technical activities
N 77–82	Administrative and support service activities
O 84	Public administration and defence; compulsory social security
P 85	Educations
Q 86–88	Human health and social work activities
R 90–93	Arts, entertainment and recreation
S 94–96	Other service activities

T 97–98 Activities of households as employers; undifferentiated goods and services producing activities of households for own use

U 99 Activities of extraterritorial organizations and bodies

3.1 THE CLASSIFICATION OF INDIVIDUAL CONSUMPTION BY PURPOSE (COICOP)

COICOP is one of the functional classifications in the UN 1993 System of National Accounts (SNA93). It is used to classify individual consumption expenditures of three institutional sectors, namely households, Non-Profit Institutions Serving Households (NPISHs) and Government.

COICOP is described as a "functional" classification because it categorizes consumption expenditures according to their primary "functions" – in the sense of "purposes" or "objectives"; for example, housing, medical, transport, recreation and education.

Detailed structure

01-12 - Individual consumption expenditure of households

- 01 - Food and non-alcoholic beverages
- 02 - Alcoholic beverages, tobacco and narcotics
- 03 - Clothing and footwear
- 04 - Housing, water, electricity, gas and other fuels
- 05 - Furnishings, household equipment and routine household maintenance
- 06 - Health
- 07 - Transport
- 08 - Communication
- 09 - Recreation and culture
- 10 - Education
- 11 - Restaurants and hotels
- 12 - Miscellaneous goods and services
- 13 - Individual consumption expenditure of Non-Profit Institutions Serving Households (NPISHs)
- 14 - Individual consumption expenditure of general government

3.2 TANZANIA STANDARD CLASSIFICATION OF OCCUPATIONS (TASCO)

Tanzania Standard Classification of Occupations (TASCO) provides a systematic classification and codification structure for the civilian working population of the United Republic of Tanzania. TASCO has been fashioned after the International Standard Classification of Occupations, ISCO-1988, compiled by the International Labour Office, ILO Geneva which is the revised edition of ISCO-68. This adaptation has been done to ensure international and regional comparability of reporting and analysis of statistical data relating to occupations, manpower, population census, etc.

TASCO has been closely aligned with ISCO-88 to ensure international and regional comparability of statistical data. To maintain continuity with the country's past statistical data, based as it is on the ISCO-68, the main text of TASCO for each Major Group gives corresponding ISCO-68 Codes for titles included in the TASCO. The expanded Index, which includes all Base, Alternate and Related Titles of this Classification, will help coders, enumerators, and users of TASCO in the location, classification and codification of all occupations in the country.

Objective of TASCO

TASCO has been designed for the proper classification of the collected data of all civilian occupations in the country's labour force and to ensure its convertibility into the international standard.

TASCO, as also ISCO-88, differs from the ISCO-68 in the following aspects:

i. A new level of aggregation, called Sub-Major Groups, has been created between those represented by Major and Minor Groups in ISCO-68.
ii. The basis of occupational classifications, in the ISCO-68, was according to type of work performed. TASCO and ISCO-88 occupational classifications are based on "Skill requirement" criterion, reflected in the skill level and skill specialisation of various occupations.
iii. As a result of introduction of a new level of aggregation, viz sub-major groups, the occupational code structure consists of six (6) digits, instead of five (5) digits of ISCO-68; thus:-

- Major Groups are of one (1) digit, the extreme left numeral, at 'thousand' point of the four digits before the decimal point;
- Sub-Major Groups consist of two (2) digits, the left two numerals, at 'thousand' and 'hundred' points, of the four-digit numbers before the decimal point;
- Minor Groups consist of three (3) digits, the left three numerals, at 'thousand' 'hundred' and 'ten' points, of the four-digit numbers before the decimal point;
- Unit Groups consist of four (4) digits, all the four numerals left of the decimal point; and
- Occupational Categories consist of six (6) digits, all the four (4) numerals left of the decimal point plus the two (2) numerals right of the decimal point.

Example

2143.40 Instrument Engineer, Electrical;

Major Group is indicated by '2', one digit at the thousand' point;

Sub-Major Group is indicated by '21', two digits at 'thousand' and 'hundred' points; Minor Group is indicated by '214', at 'thousand', 'hundred' and 'ten' points; Unit Group is indicated by '2143', all the four numerals left of the decimal point; and

Occupational category is indicated by 2143.40, i.e. all the four numerals, viz '2143' left of the decimal point plus two numerals, viz.40 on the right side of the decimal point.

Basic Approach and Classification Criteria

i. TASCO provides a structure for the classification of all civilian occupations in the country's labour force. Occupational categories are, thus, the smallest segment of work which is specifically identified, classified and codified in the TASCO. At this detailed level, an occupation is a set of jobs which involve the performance of a similar or common, but not necessarily identical, set of tasks all over the country.

   An occupation, for the purpose of this classification, covers various jobs similar in their main tasks, and held by individual workers, all over the country, and in various establishments and industries, but workers may perform one or more of the different possible combinations or breakdowns of the set of tasks described in the occupational

definitions of ISCO-68, shown against TASCO occupational titles, under its Unit Groups. Workers, in a number of establishments and industries, and whose principal tasks are similar, are considered to be sharing the same type of job; and the workers, in a particular establishment, performing identical principal tasks, are considered to be manning the same type of positions. A position, in any given establishment, is a set of tasks performed by one individual. Therefore, in a given establishment, there will be as many positions as are workers employed therein. In practice, positions are normally distinguished from one another by minor differences in duties, level of responsibility for supervision, and other particularities, e.g. wages and division of labour.

ii. ISCO-68 had classified occupations according to the type of work performed which was reflected in different ways, such as occupations occurring in the same field/discipline (e.g. physical sciences, life sciences); those concerned with the same subject-matter (e.g. Chemists, Physicists, Biologists, Agronomists); and those of workers executing similar processes (e.g. metal platers and coaters), operating similar equipment (e.g. motor vehicles), performing similar services (e.g. waiters, bartenders) and fabricating similar articles (e.g. watch, clock and precision instrument makers).

iii. In TASCO, the classification criteria is based on the requirements imposed upon workers by the tasks and duties of their occupations reflected in:

(a) the knowledge and proficiency needed for performing physical and mental activities of the prescribed tasks and duties;
(b) the knowledge of working principles, methods and techniques and characteristics of materials and equipment used; and
(c) the ability to react adequately to various work situations and demands.

iv. The above technical requirements depend on the skill requirements of an occupation. Skill may be defined as the ability of an individual worker to perform a set of tasks or to fulfil the technical requirements of an occupation. Thus, for the classification of jobs into occupation and of occupations into Unit Groups, skill requirements provide the similarity criterion for sets of tasks in the group structure of TASCO. The focus is on the skill required to carry out the tasks and duties of an occupation and not on whether a worker,

having a particular occupation, is more or less skilled than another worker in the same occupation.

iv. The `skill requirement' criterion, used in TASCO, has two dimensions, viz (a) skill level and (b) skill specialisation.

(a) The skill level of an occupation is a function of the complexity and range of the tasks involved. The greater the complexity and wider the range of the set of tasks, the higher the skill level of an occupation. An occupation that involves more complex tasks than another will require a higher skill level. An occupation which requires the performance of a wide range of tasks has a higher skill level than an occupation which requires the performance of a sub-set of those same tasks. But, complexity of task is more important than mere range of tasks in determining the skill level, since complexity of tasks indicates the degree of proficiency, the level of knowledge, and the quality and character of the response required to accomplish such tasks.

(b) The skill specialisation of an occupation is a function of (i) the field of knowledge required (viz. the subject-matter essential to the performance of tasks), (ii) tools, equipment and work aids used (e.g. plant, machinery, hand-tools); (iii) materials worked on or with (such as extracted, processed, refined or manufactured), and

(c) Goods produced or services provided in relation to the tasks performed.

From the above, it can be deduced that:-

the greater the range and complexity of the set of tasks involved, the greater the duration and amount of (i) formal education; (ii) on-the-job training, and (iii) previous experience required for the satisfactory performance of the set of given tasks; but in some case, innate strength, abilities, appropriate work-experience, coupled with thorough on-the-job training, may equip a worker with the needed skills, or a part of them, for the successful performance of the set of tasks. In Tanzania, quite often, skills are acquired through informal training and work-experience.

3.3 THE SYSTEM OF NATIONAL ACCOUNTS (SNA)

The System of National Accounts (SNA) is the internationally agreed standard set of recommendations on how to compile measures of economic activity. The SNA describes a

coherent, consistent and integrated set of macroeconomic accounts in the context of a set of internationally agreed concepts, definitions, classifications and accounting rules.

The SNA provides an overview of economic processes, recording how production is distributed among consumers, businesses, government and foreign nations. It shows how income originating in production, modified by taxes and transfers, flows to these groups and how they allocate these flows to consumption, saving and investment. Consequently, the national accounts are one of the building blocks of macroeconomic statistics forming a basis for economic analysis and policy formulation.

Objective and Uses of the System of National Accounts (SNA)

The main objective of the SNA is to provide a comprehensive conceptual and accounting framework which can be used to create a macroeconomic database suitable for analyzing and evaluating the performance of an economy. The existence of such a database is a prerequisite for informed, rational policy-making and decision-taking.

Accounts and their corresponding economic activities

The sequence of accounts

Current accounts

These accounts record the production of goods and services, the generation of incomes by production, the subsequent distribution and redistribution of incomes among institutional units, and the use of incomes for purposes of consumption or saving.

Production account

The production account records the activity of producing goods and services as defined within the System. Its balancing item, gross value added, is defined as the value of output less the value of intermediate consumption and is a measure of the contribution to GDP made by an individual producer, industry or sector. Gross value added is the source from which the primary incomes of the System are generated and is therefore carried forward into the primary distribution of income account. Value added may also be measured net by deducting consumption of fixed capital.

Distribution and use of income accounts

These consist of a set of articulated accounts showing how incomes are:

- Generated by production
- Distributed to institutional units with claims on the value added created by production
- Redistributed among institutional units, mainly by government units through social security contributions and benefits and taxes
- Eventually used by households, government units or non-profit institutions serving households (NPISHs) for purposes of final consumption or saving.

Accumulation accounts

These are flow accounts that record the acquisition and disposal of financial and non-financial assets and liabilities by institutional units through transactions or as a result of other events:

- The capital account records acquisitions and disposals of non-financial assets as a result of transactions with other units or internal bookkeeping transactions linked to production (changes in inventories and consumption of fixed capital).
- The financial account records acquisitions and disposals of financial assets and liabilities, also through transactions.
- A third account, the other changes in assets account, consists of two sub-accounts. The first, the other changes in volume of assets account, records changes in the amounts of the assets and liabilities held by institutional units or sectors as a result of factors other than transactions; for example, destruction of fixed assets by natural disasters. The second, the revaluation account, records those changes in the values of assets and liabilities that result from changes in their prices.

Balance sheets

The balance sheets show the values of the stocks of assets and liabilities held by institutional units or sectors at the beginning and end of an accounting period. As already noted, the values of the assets and liabilities held at any moment in time vary automatically whenever any transactions, price changes or other changes affecting the volume of assets or liabilities held take place. These are all recorded in one or another of the accumulation accounts so that the difference between the values in the opening and closing balance sheets is entirely accounted for within the System, provided, of course, that the assets and liabilities recorded in

the balance sheets are valued consistently with the transactions and other changes - that is, at current prices.

Activities and transactions

The accounts of the System are designed to provide analytically useful information about the behaviour of institutional units and the activities in which they engage, such as production, consumption and the accumulation of assets. They usually do this by recording the values of the goods, services or assets involved in the transactions between institutional units that are associated with these activities rather than by trying to record or measure the physical processes directly. For example, the accounts do not record the physical consumption of goods and services by households - the eating of food or the burning of fuel within a given time period. Instead, they record the expenditures that households make on final consumption goods and services or, more generally, the values of the goods and services they acquire through transactions with other units, whether purchased or not.

Data on transactions provide the basic source material from which the values of the various elements in the accounts are built up or derived. The use of transactions data has important advantages. First, the prices at which goods and services are exchanged in transactions between buyers and sellers on markets provide the information needed for valuing, directly or indirectly, all the items in the accounts. Secondly, a transaction that takes place between two different institutional units has to be recorded for both parties to the transaction and therefore generally appears twice in a system of macroeconomic accounts. This enables important linkages to be established in the System.

3.4 GOVERNMENT FINANCE STATISTICS (GFS)

Government Finance Statistics (GFS) is the International Monetary Fund standard for publishing financial statistics within the Government system. It is designed to provide statistics that enable policymakers and analysts to study developments in the financial operations, fiscal position, and liquidity situation of the general government sector or the public sector in a consistent and systematic manner.

Uses of the GFS system

The GFS system is designed to provide statistics that enable policymakers and analysts to study developments in the financial operations, financial position, and liquidity situation of

the general government sector or the public sector in a consistent and systematic manner. The GFS analytic framework can be used to analyze the operations of a specific level of government and transactions between levels of government as well as the entire general government or public sector.

Compilation Framework and Data Sources

The GFS 2001 system extends the cash-based analytical framework (GFS 1986) by emphasizing the recording of data on an accrual basis to give a comprehensive picture of government activities by recording non-cash transactions, internal transactions, other economic flows, and any payment arrears. It provides an integrated framework which can be used to analyze the operations of a specific level of government and transactions between levels of government as well as the entire general government sector.

Tanzania's General Government sector is composed of Ministries, Departments and Agencies - Central Government, Regional and Local Government Authorities and extra-budgetary units.

Since financial year 2000/2001 GFS 1986 system has been used for budget preparation and accounting for central government (Tanzania Mainland) and later was extended to other government units (including local government). In 2009/10 the central government prepared its budget in GFSM 2001 system. To date, GFS system is used in all Ministries, Regions and independent governments departments (including Zanzibar government). Also to a large extent most of the extra-budgetary units and Local Government Authorities use this system.

The data categories, sources and responsible institutions are as indicated in the following table:

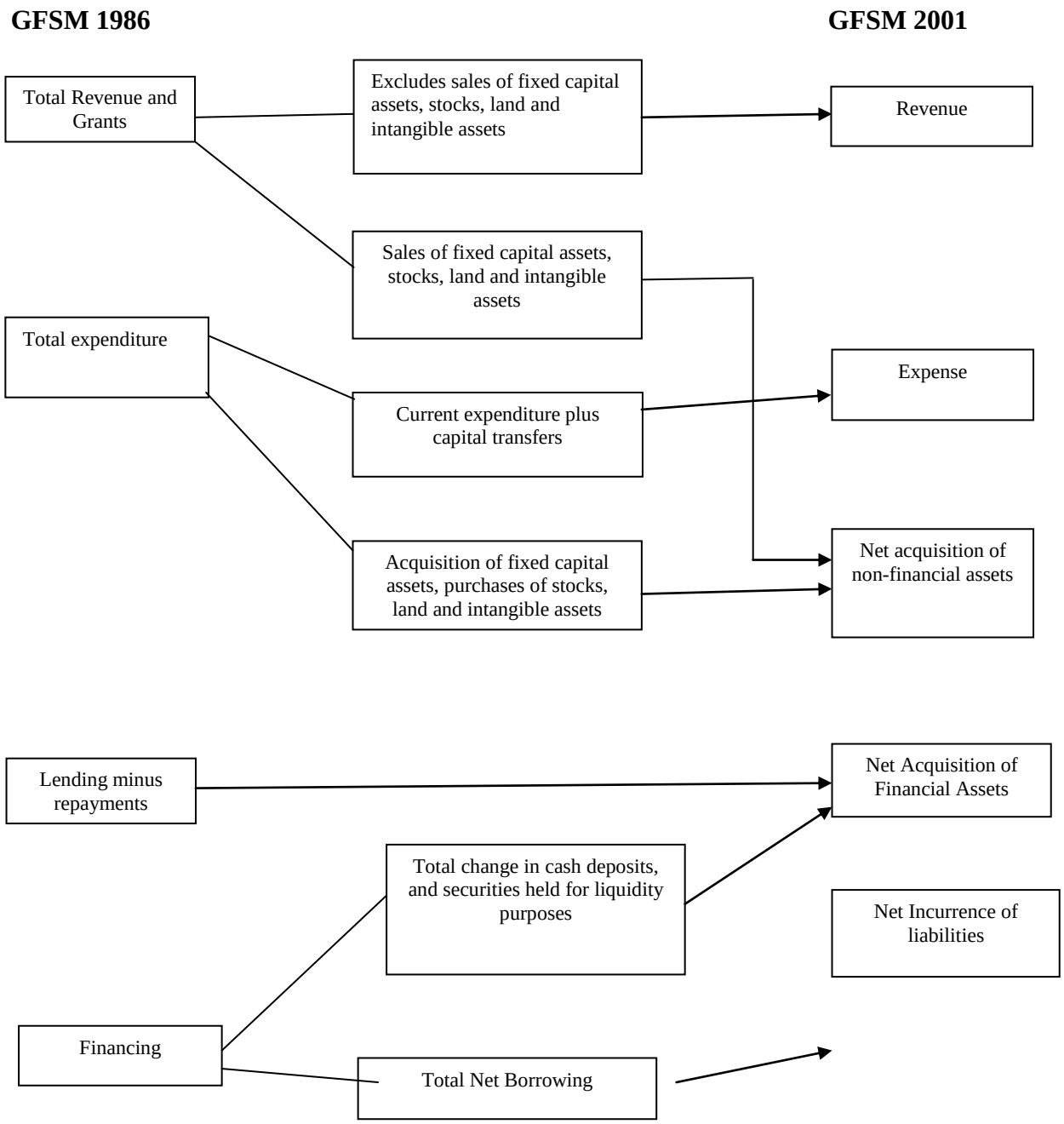
Data Category	Data Source	Responsible
Revenue		
Tax Revenue	Monthly Flash Report, Annual statement of Revenue	CPAD, ACGEN, PMO-RALG
Social Contribution	Payroll Report	CB
Grants	Monthly Reports for Grant	CEF, ACGEN
Other Revenue	Monthly Flash Report, Annual statement of Revenue	CPAD, ACGEN, PMO-RALG
Expenses		
Compensation of Employees (Wages and Salaries)	Payroll Report, Expenditure Flash Report	CB, ACGEN, PMO-RALG
Consumption of Fixed Capital (CFC)	Statement of use of Assets	DGAM, ACGEN
Interest Payments	Amortization Report, Monthly	ACGEN

	Expenditure Report	
Subsidies	TR Statement (PE & OC)	TR, CB, PMO-RALG
Grants	Statement of transfers (e.g to other levels of government, etc)	CB, PMO-RALG
Social Benefits	Statement of transfers (e.g to households)	CB
Other Expenses	Statements of Expenditure	CB, ACGEN, PMO-RALG
Non Financial Assets (NFA)		
Fixed assets		DGAM, ACGEN, PMO-RALG
Inventories		DGAM, ACGEN, PMO-RALG
Valuables		BOT
Non-Produced Assets		
Use of goods and services	Expenditure Flash Report, Consolidated Annual Financial Statements	ACGEN, PMO-RALG BOT, ACGEN
Financial Assets (FA)		
Domestic		
Currency and deposits		
Securities other than share		
Loans		
Shares and other equities		
Insurance Technical Reserves		
Financial derivatives		
Other account receivables		
Foreign		
Currency and deposits		
Securities other than share		
Loans		
Shares and other equities (Public Corporation only)		
Insurance Technical Reserves		
Financial derivatives		
Other account payables		

Classification of GFSM 1986 data to the detailed GFSM 2001 Tables

The reclassification of GFSM 1986 data to the classification categories of the GFSM 2001 framework is seen as a first step in the transition to the new GFS methodology. Broadly, the relationships between the GFSM 1986 and 2001 classification systems are as follows:

Broad overview of relationships between GFSM 1986 and GFSM 2001 Classification Systems



NBS and other MDAs engage in data production. This part describes some surveys and non surveys in data production.

4.0 SOCIAL STATISTICS

4.1 TANZANIA HIV/AIDS & MALARIA INDICATOR SURVEY (THIMS)

Objectives and Uses of Product

The main objectives of the THMIS are to:

i) Provide information on the prevalence of HIV infection among Tanzanian adults, and the prevalence of malaria infection and anaemia among children under age five years.

ii) Know the extent to which these indicators vary by characteristics such as age, sex, region, education, marital status, and poverty status.

Uses of Product

The findings of the survey are used to:

i) Evaluate ongoing health related programmes and to develop new strategies.

ii) Provides updated estimates of selected basic demographic and health indicators covered in previous surveys.

iii) Provide information to assist policy makers and programme implementers to monitor and evaluate existing programmes and to design new strategies for combating the HIV/AIDS epidemic in Tanzania.

Methodology

Sampling Frame

The THMIS utilises a two-stage sample design. The first stage involves selecting sample points (clusters) that are enumeration areas (EA's) of Population and Housing Census.

A household listing operation is undertaken in all the selected areas prior to the fieldwork.

The second stage involves systematic sampling of households from the updated lists of households.

Questionnaire

Two types of questionnaires are administered: Household Questionnaire and Individual Questionnaire based on the standard HIV/AIDS Indicator Survey and Malaria Indicator Survey.

The Household Questionnaire is used to list all the usual members and visitors of selected households. Some basic information are collected on the characteristics of each person listed, including his or her age, sex, education, and relationship to the head of the household, also to identify women and men who are eligible for the individual interview and HIV testing.

The Individual Questionnaire is used to collect information from all women and men age 15-49.

Training of Field Staff

Field staff training includes class presentations, mock interviews, field practice and tests. Field practice in malaria and anaemia testing and HIV dried blood spot collection are carried out towards the end of the training period.

Biomarkers

THMIS incorporates three 'biomarkers' for anaemia testing, malaria testing and HIV testing.

Field Work

The field work exercise is carried out in form of teams, where each team consists of a team leader, enumerators, and a driver. Field work takes a period between three and four months.

Data Processing

The processing of data consists of office editing, coding of open-ended questions, data entry, and editing of computer-identified errors. Normally, data processing is conducted concurrently with the data collection exercise.

Dried blood samples are transported for test to Muhimbili University College of Health Sciences (MUCHS) for laboratory tests.

Quality Control

Quality control teams periodically visit teams in the field to check their work and re-interview some households (spot interviews).

Analysis and Report

After all completed questionnaires have been checked, entered and finalized and analysis completed, at least two reports have to be prepared: Preliminary report and final report. Primary report has to be brief and consists primarily of major HIV and malaria indicators. The final report is more detailed.

Dissemination and Utilisation of Results

A National dissemination seminar is conducted after completion of report writing. Zonal and regional level disseminations are also done when funds are available.

4.2 TANZANIA DEMOGRAPHIC AND HEALTH SURVEY (TDHS)

Objective and use of product

The principal objective of Tanzania Demographic and Health Survey (TDHS) is to collect data on household characteristics, fertility levels and preferences, awareness and use of family planning methods, childhood mortality, maternal and child health, breastfeeding practices, antenatal care, childhood immunisation and diseases, nutritional status of young children and women, malaria prevention and treatment, women's status, female circumcision, sexual activity, knowledge and behaviour regarding HIV/AIDS, domestic violence, fistula, maternal mortality, ownership and use of mosquito bed nets and other STIs.

Methodology

Sample Design

The sample for the TDHS is designed to provide estimates for the entire country, urban and rural areas of the Tanzania Mainland and for Tanzania Zanzibar. A representative probability sample is selected in two stages. In the first stage, clusters are selected from a list of enumeration areas from the Population and Housing Census. In the second stage, a complete household listing exercise carried out within all the selected clusters, then households are systematically selected for participation in the survey.

Questionnaires

Three questionnaires are used for TDHS: the Household Questionnaire, the Women's Questionnaire, and the Men's Questionnaire. The Household Questionnaire is used to list all the usual members and visitors in the selected households some basic information are collected on the characteristics of each person listed, including age, sex, education, and relationship to the head of the household. The main purpose of the Household Questionnaire is to identify women and men who are eligible for the individual interview. The Women's Questionnaire was used to collect information from all women age 15-49. The Men's Questionnaire is administered to all men age 15-49 living in every third household in the TDHS sample. The Men's Questionnaire collects much of the same information found in the Women's Questionnaire, but did not contain a detailed reproductive history or questions on maternal and child health.

Biomarker Testing

The TDHS consists of four biomarker testing; Height and Weight Measurement, Anaemia Testing, Vitamin "A" testing and Iodine Testing.

Pre-test

All elements of the survey are to be pretested prior to the main survey.

Training of Field Staff

The training is conducted following the DHS training procedures, including class presentations, mock interviews, tests and field practice. Field practice in anthropometric measurements and biomarkers are carried out toward the end of the training period.

Fieldwork

Field work is done in form of teams, of which each consists of female interviewers, male interviewers, a supervisor, a field editor, and a driver. The field editor and supervisor are responsible for reviewing all questionnaires for quality and consistency check before the team's departure from the cluster.

Data Processing

Data processing includes office editor, coding of open-ended questions, data entry, and editing of computer identified-errors, data entry and editing accomplished by CSPro software. The Dried

Blood Spots (DBS), urine and salt samples transported to TFNC to be tested. The samples are to be logged into the CSPro Test Tracking System database.

4.3 TANZANIA DISABILITY SURVEY

Objective of product

The objective of the Tanzania Disability Survey (TDS) is to determine the prevalence of disability in the country and to determine living conditions among people with activity limitations.

Use of Product

The findings of the survey are used as a contribution to the improvement of the living conditions among people with activity limitations in Tanzania; the findings also initiate a discussion on the concepts and understanding of "disability" and monitor the impact of government policies, programmes and donor support on the well being of the population with activity limitations.

Methodology

Sampling Design

The survey covers the whole of the United Republic of Tanzania and produces estimates at regional level. Information is collected from all selected households and individuals with and without disability (measured as activity limitations).

The primary sampling unit for the survey is the census enumeration area (EA) and the ultimate sampling unit are the individual household members. Disability survey utilizes a three-stage systematic stratified random sampling design, involving clusters (EAs), households and individual household members.

Sampling Frame

The sampling frame of clusters to be used is the list of all enumeration areas (EAs) generated during the 2002 Population and Housing Census. EAs in each region are listed following the census coding system and a target sample is selected using probability proportional to size. The EA maps and other administrative information are used to identify the boundaries and features of the selected EAs. For households, the sampling frame should be the list of households (heads) constructed for each selected EA. To ensure a random scattered sample, the listing of households should be done in a serpentine manner from one end of the EA to another end.

Sample Size Determination

A total number of clusters must be selected. The targeted sample per cluster should be determined. The selection of EAs should follow the Probability Proportional to Size (PPS) sampling while the selection of households and individuals follow a simple random sampling procedure. The random spread of households is necessary for achieving a non-clustered sample.

Data Collection

Disability survey uses three types of questionnaires. First, Household Questionnaire, which attempt to collect basic socio-economic conditions of usual household members and visitors. The questionnaire also includes screening questions which used to determine persons with disabilities.

The second questionnaire is Adult Questionnaire which is used to collect information from all person aged 15 years and above who were identified in the household questionnaire as having some form of disability.

The Children Questionnaire is the third instrument which is used to collect information from all children identified as having disabilities but collect more or less the same information in the adult questionnaire.

Training of Field Staff

The training involves both theory and practise. Organisations of People with Disability are represented by participants who have a responsibility of clarifying terms used in the questionnaires. Besides participating in training representatives from Organisations of People with Disabilities is also part of supervisory teams.

Field work

The field work includes map reading and listing exercise, the supervisors and enumerators must be trained on map reading and listing. Supervisors should ensure that all households in the EA listed according to the given instructions and EA map. Supervisors are also responsible for selecting households to be interviewed.

Data Collection

Data collection is conducted in teams, each team consist of supervisor, enumerators and a driver. Supervisors are responsible for the overall administrative work in the field including checking the quality of the questionnaires before departing from the cluster.

Data Processing

Data processing starts concurrently with the fieldwork. The data processing personnel includes supervisors and a questionnaire administrator, who ensures that the expected numbers of questionnaires from all clusters are received; office editors and data entrants. The CsPro computer package is used for data entry.

4.4 POPULATION AND HOUSING CENSUS

Objective and use

The main objective of the census is to provide the Government and other stakeholders with essential statistics on the population, in terms of demographic, social and economic characteristics, housing conditions and household amenities.

The population and housing census' data is an important input for the preparation of economic and development policies, monitoring the improvement in the quality of life of the population, as well as developing a system of sustainable development in general.

Methodology

Census Budget

The cost of Census exercise is covered by the Government and development partners**.**

Census Cartographic Work

Strategies for development of census maps include: Delineation of Enumeration Area (EA), boundaries, preparation of census field map and acquisition of spatial data and development of Geographical Information System (GIS). Census mapping methodology includes the application of digital spatial data in the delineation of EAs to improve the quality of the work in terms of accuracy.

Pilot Census

A pilot census is conducted to assess the whole process of census and Post Enumeration Survey (PES) operation. The pilot census determines the workload of enumeration, logistic support, enumeration procedures, data processing and acceptability by the public in general. Administrative control and management issues also tested through the pilot census.

Staff Recruitment and Training of Field Staff

Recruitment and training of enumerators during enumeration is given a special attention as large number of staff is required. Traditionally, primary school teachers have been involved as enumerators/supervisors. Supervisors and enumerators are trained at regional level.

Census Enumeration

Census enumeration is an important operation which determines the success or failure of a census. Quality of census depends on how the exercise is conducted. Enumeration involves house to house visit by trained numerators using structured questionnaire with the objective of enumerating all persons in the country.

Data Processing

Operational control procedures are equally important in guaranteeing that all data are processed and that no data are dropped or duplicated. Data processing exercise is done through scanning technology and computers for manual data entry for questionnaires that for some reasons will be rejected by scanner.

Data Analysis and Dissemination

Data analysis and dissemination is undertaken after completion of data processing and ensuring that data is clean. The analysis is undertaken by teams of experts from Government Ministries, Research Institutions and Higher Learning and Training Institutions. A number of experts from within and outside the country provide technical assistance.

Report Writing

Preliminary report is prepared by professionals from NBS and OCGS staff and the main report is prepared by professionals from NBS, OCGS and other ministries.

4.5 EDUCATION AND HEALTH STATISTICS

Objective and Use of Product:

The health and education statistics are inputs to the Economic Survey Report which is produced annually by the Ministry of Finance.

Methodology

Method of Data Collection

The education data are collected using questionnaires. These questionnaires are sent to the Regional Statistical Offices for distribution to the Higher Learning and Training Institutions in each region. These questionnaires are filled-in by a responsible person in the respective institution under the supervision of the Regional Statistical Managers (RSM) and headquarter supervisor for quality assurance.

Health

Health Statistics are collected from the Ministry of Health and Social Welfare by staff from the Department of Social and Demographic Statistics.

The health data and information that is collected and compiled includes:

- The number of health professionals,
- Ownership of health facilities,
- The total health facilities according to ownership (Private, Government, Government agencies and Religious Health Facilities),
- The total number of health professional students who passed their exams, and
- The number of patients attended by different health facilities annually.

Education

The education data and information collected and compiled includes:

- The yearly total enrolment according to courses,
- The total number of teachers in a given year,
- The total number of students, and
- The number of lecturers in all higher learning institutions.

These statistics are usually updated every year to facilitate the production of the Economic Survey Report produced annually. As such, these data are collected from the second week of January and submitted to the National Accounts Department in March for further submission to the Ministry of Finance. They are expected to reach the Ministry of Finance not later than mid-April every year.

4.6 HOUSEHOLD BUDGET SURVEY (HBS)

Objective and Use of Product

A household budget survey is intended to collect, compile, analyze and facilitate dissemination of numerical information on household income, consumption and expenditure. Basically, the aims and objectives of the household budget survey are to produce information that will track the effectiveness of various interventions and policies towards the fight against poverty in Tanzania since mid-1990s.

Methodology

Sample design

The HBS uses a two stage sampling design with the village as Primary Sampling Unit (PSU) in rural areas and an enumeration area as a PSU in urban areas. Households are the secondary sampling units. The sample of households is divided into 12 sub-samples to be interviewed each month of a survey to capture seasonal variations. The sample is normally designed to provide estimates for the whole country disaggregated into rural and urban (Dar es Salaam and other urban) as done in 1991 and 2007 HBSs. If resources permit, the sample may also be designed to provide regional estimates as done during the 2001 HBS.

Survey Instruments

There are mainly two types of survey questionnaires namely HBS Questionnaire Form 1 (HBSQF-1) and HBS Questionnaire Form 2 (HBSQF-2). The former collects information about socio-demographic and economic information of the each household member as well as other household level information pertaining to housing conditions, sources of water and energy for lighting and cooking, sanitation and waste management, distances to various socio-economic services, assets ownership, household business, sources of household income, and others.

The Form 2 (HBSQF-2) also known as a diary is used to collect information on daily income and consumption and non-consumption expenditures of all members of the households. It is administered for the duration of one calendar month in each household. This information is recorded in the form by enumerator from the individual diary known as HBSQF-3 and the business diary known as HBSQF-4. This individual diary is given to every member of the household who is above 5 years old and able to read and write. Also HBSQF-4 is given to the households with business just for recording all transactions; that is purchases for re-selling.

On the other hand, there is an instruction manual explaining all survey details and questionnaire contents. Attempt is made to give explanation on each question that is in the questionnaires. Also codes for various items or transactions that are anticipated to be found in households are provided in this document. Each enumerator and supervisor is provided with this manual for reference each time they undertake the filling-in or reviewing the already filled-in questionnaires.

Training

There are two phases of training; the first one being the training of trainers (TOT) where all anticipated trainers gather and discuss the details of the training, the methods of training and the training schedule. These prospective trainers of enumerators and supervisors normally meet for up to 15 days and undertake class and practical training.

Training of enumerators and supervisors normally lasts for 3 weeks with many practical sessions to ensure competence and accuracy during data collection. After the training the data collection starts immediately on the first day of the month in order to have coverage of full calendar month (or 28 days) of data collection for each household. Due to a large number of enumerators and supervisors, the training is conducted in zones to ensure that all participants are properly trained.

Data Collection

Each enumerator is assigned a cluster for the whole year where he or she establishes and maintains contact with sub-sample of households for each calendar month. The data collection starts with administering HBSQF-1 and later on, explains and demonstrates to the household how to fill-in HBSQF-3 and HBSQF-4. During the survey month, the enumerator visits the households to make sure that the diaries are filled-in as well as transferring the information to HBSQF-2. The enumerator is also required to assist households without a member who can read and write filling the diaries.

Together with all these steps, there are field trips from the main office whereby the purpose is to randomly check some households' filled-in questionnaires and re-interview to check the consistence and quality of information filled-in. Such checks are done by head office staff who also check the on-going filling-in of questionnaires in some households.

Processing

Editing of filled-in questionnaires is a two stages exercise. The primary editing is done at specific regional statistical office. This ensures that all problems that require going back to the field are solved there before the questionnaires are sent to the main office where it may take time to rectify the problems. The second stage is at the main office before the questionnaires are sent for data entry in the computer laboratory.

At the data processing (computer) room, after data of every month have been entered, consistence check programs are run to detect any inconsistencies and mistakes either committed in the field and were not seen at both stages of editing or committed at the data entry stage, and fix them.

Analysis

This is an important stage of processing the survey results. This is done by NBS and other organizations or individuals with competence in specific areas such as poverty analysis, consumer price index and national accounts. Both national and international consultants may be recruited to support local staff in the analysis.

4.7 CONSUMER PRICE INDEX (CPI)

The Consumer Price Index (CPI) Compilation Model is an Excel add-in based Visual Basic for Application (VBA) program developed for providing technical assistance to fund member countries for consumer price statistics compilation. The package is a prototype model that could also be used for teaching, training, and research purposes. It can likewise serve as a framework from which price collection formats and product classification systems can be integrated to fit country-specific practices. Its main function is to compile the CPI for multiple areas at multiple levels of product items to obtain aggregate national level price statistics.

It's Methodology

1. The application uses the Modified Laspeyre's Approach to compute CPI statistics based on monthly price quotations (or monthly average price quotations) and annual expenditure information using a combination of manual and computer tabulation procedures.

$$I_{t\to 0} = \sum_{i=1}^{n} w_{0,i} * \left(\frac{p_{t,i}}{p_{t-1,i}}\right) * \left(\frac{p_{t-1,i}}{p_{0,i}}\right)$$

which can be rewritten as:

$$I_{t\to 0} = \sum_{i=1}^{n} w_{0,i} * STPR_{t\to t-1,i} * LTPR_{t-1\to 0,i}$$

where $W_{0,i}$ is base period weight for item i, $STPR_{t\to t-1,i}$ is the short-term price relative of item i for current period (= $p_{t,i}/p_{t-1,i}$) and $LTPR_{t-1\to 0,i}$ is the long-term price relative of item i for previous period (= $p_{t-1,i}/p_{0,i}$).

2. CPI is calculated using the equivalent of a recursive procedure, in which previous period's base-weighted long-term price relatives, p_{t-1}, q_0 are updated by the current period's price relatives.

3. The geometric mean method is used in computing the price level at aggregate level in view of its multiple advantages.

4. The program adopts the Matching Price Observation method in imputing the areas' average prices, in which the price averages are calculated on the basis of "matched observations". Whenever a particular price observation is missing from either the previous month or the current month, the corresponding price observations will be dropped from the other period. This is to ensure consistent sample of price quotations in each period.

5. The program calculates missing variety prices based on Short Term Price Relative STPR (actual or imputed) and previous period price, and stores them in the database with a flag. These calculated price data can be retrieved into spreadsheets for the next period imputation process.

4.8 INTERNATIONAL COMPARISON PROGRAM (ICP)

The International Comparison Programs is a global project, managed centrally by the World Bank, with an organised hierarchy of regional management structure. The responsibility of managing its African component ICP-Africa lies with the African Development Bank (AfDB). The main objective of the ICP is to compare the economic aggregate and the volumes of gross domestic product (GDP) between the countries.

The targeted population is the set of all goods and services that are consumed by household from outlets during the benchmark year, covering the whole country (including rural and urban areas), except expenditure of housing rent on residential (building), education and health services which are subsidized in most of the countries.

The ICP price survey is integrated with the existing system of price collection for the CPI except to some items which are not available in CPI but are needed in ICP according to the agreement of country members.

The ICP price collections are carried out in urban and rural areas (weekly market) of the sampled centres. Stratification allocation and purpose are the methods used to select the region for ICP price collection survey. In Tanzania, there are seven zones including Zanzibar which constitute Unguja and Pemba.

The country members capture and validate ICP price collection using the same method as used in CPI but further analysis is done by African Development Bank (AFDB).

4.9 HARMONISED CONSUMER PRICE INDEX (HCPI)

The Harmonised Consumer Price Index (HCPI) Compilation Model is an Excel add-in based Visual Basic for Application (VBA) program developed for providing technical assistance to Fund member countries of SADC for consumer price statistics compilation. The package is a prototype model that could also be used for teaching, training, and research purposes. It can likewise serve as a framework from which price collection formats and product classification systems can be integrated to fit country-specific practices.

Its main function is to compile the HCPI for multiple areas at multiple levels of product items to obtain aggregate national level price statistics.

It's Methodology.

1. The Modified Laspeyres Formula is used to compute HCPI statistics based on monthly price quotations (or monthly average price quotations) and annual expenditure information using a combination of manual and computer tabulation procedures.

$$I^{t/0} = \sum_i p_i^t q^b_i \ / \sum_i p_i^0 q^b_i$$

which can be written

$$I^{t/0} = \sum_i w_i \, . \, I_i^t \, / \, I_i^0$$

where

w_i is the weight used for product i,

$I^{t/0}$ is the price index for product i between the price reference period 0 and period t;

p_i^t is the price of product i in period t;

p_i^0 is the price of the same product i in period 0;

q^b_i is the base-period quantity of product i expressed as a proportion of the expenditure on product i to total expenditures covered in the HCPI.

2. HCPI is calculated using the equivalent of a recursive procedure, in which previous period's base-weighted long-term price relatives, p_{t-1}, q_0 are updated by the current period's price relatives.

3. The geometric mean method is used in computing the price level at aggregate level in view of its multiple advantages.

4. The program adopts the Matching Price Observation method in imputing the areas' average prices, in which the price averages are calculated on the basis of "matched observations". Whenever a particular price observation is missing from either the previous month or the current month, the corresponding price observations will be dropped from the other period. This is to ensure consistent sample of price quotations in each period.

5. The program calculates missing variety prices based on Short Term Price Relative STPR (actual or imputed) and previous period price, and stores them in the database with a flag. These calculated price data can be retrieved into spreadsheets for the next period imputation process.

4.10. INTEGRATED LABOUR FORCE SURVEY.

Objective of Labour Force

The Integrated Labour Force Survey intends to obtain comprehensive data on the current status of National Labour Market and to provide up to date data needed by the government and other stakeholders on human economic activities particularly those related to the informal sector and its magnitude, unemployment, underemployment child labour and time use.

Use of Product

The findings of the survey are used in planning, policy implementation, monitoring and evaluation of government programmes aimed to determine the magnitude of the Labour force in the country and to collect information on employment status so as to introduce necessary changes in the country's employment policies where needed.

Methodology

Design of the Sample

The Integrated Labour Force Survey used the existing National Master Sample (NMS). The NMS is a generalized set of area units that can be used as Primary Sampling Units (PSUs) for conducting various household surveys. It is a fixed sample of rural and urban clusters, which, among other things, makes it possible for the performance of a continuous Survey Programme as well as ad hoc sample surveys.

Frame of the Sample

The sampling frame for the current NMS is based on the preliminary results of the 2002 Population and Housing Census. The Primary Sampling Unit (PSU) is the village for the rural and EA for urban areas respectively. A probability proportional to size without replacement (ppswor) – systematic sampling procedure is used for the selection of PSU. About two months before the commencement of the field work a household listing exercise is done. All households within each cluster are listed. The household listings give the sampling frame of households for each cluster.

Sample Size Determination

The selections of EAs follow the Probability Proportion to Size (PPS) sampling while the selection of households and individuals follow a simple random sampling procedure.

Estimation Procedure

The sampling procedure for both the urban and rural samples suggests good estimates at national, and cluster levels. Regional estimates can also be worked out. For urban sample it is possible to get estimates for the three domains of study, i.e. Dar Es Salaam city, nine municipalities and other towns. Estimation of individual towns and households' sizes can also be obtained by some imputation methods.

Data Collections

Labour force survey uses five types of questionnaires. First, questionnaire (LFS1) is administered to the head of the household or his/her representative intending to collect household particulars. The second questionnaire (LFS2) aims to collect the information of labour force details for individuals. The third questionnaire (CLS1) is administered to parents or guardians of all child aged 5 to 17 years. The fourth questionnaire (CLS2) aims to collect information of children age 5 to 17 years. The fifth questionnaire on the time use (TUS) is designed to collect the information on the routine activities of the respondents and administered on seven consecutive days to each member aged 5 years and above of the selected households.

Field work

The field work includes listing exercise, the supervisors and enumerators must be trained on map reading and listing. The regular field visits for ensuring close supervisors is made by national and regional supervisor. Supervisors are also responsible for ensuring the quality control of the data at all stages of data production.

Data Collection

Data collection is conducted in teams, each team consist of supervisor, enumerators and a driver. Supervisors are responsible for the overall administrative work in the field including checking the quality of the questionnaires before departing from the cluster.

Data Processing

Data processing starts soon after receiving questionnaires from the filed. The data processing personnel includes supervisors and a questionnaire administrator, who are responsible for checking

the number of clusters (EAs) in a region and number of each household in the cluster. Followed by manual editing, coding questionnaires, data entry using CSPro as a package to capture data and data cleaning and validation is done by experienced data processing personnel.

4.11 EMPLOYMENT AND EARNINGS SURVEY.

The Employment and Earnings Survey, is an annual survey conducted by the National Bureau of Statistics. The enumeration covers three main categories of employing establishments in both private and public sectors. The categories involved are: All establishments of public sector; all registered private establishments employing at least 50 persons; and a sample of all registered private establishments whose employment capacity is between 5 to 49 persons in Tanzania Mainland. The survey does not include domestic servants in Private households, non-salaried working proprietors and non-salaried family workers.

Objective of the Survey

The main objective of employment and earnings is to obtain a comprehensive data on the annual status of employment and earnings as well as data on socio-economic characteristics of the labour market.

Use of Product

The findings of employment and earnings survey are used for estimating the labour market indicators that could be used for planning, policy formulation and examining the contribution to Gross Domestic Product (GDP) of different categories of employment.

Methodology

(a) The Selection of Establishments

The Employment and Earnings Survey used the existing Central Register of Establishments (CRE) frame. The selection of establishments from the CRE frame falls under the following groups: -

i). All establishments of public sectors found in the current CRE frame are taken;

ii). All establishments of private sector with at least 50 employees found in the current CRE frame are taken;

iii). The list of surveyed establishments of private sector employing persons in the range of 5 to 49 is based on a sample.

(b) The Sample Design

i). A sample of 10 percent of establishments is selected in the employment size group of 5 to 9 employees;

ii). A sample of 33 percent of establishments is selected in the employment size group of 10 to 49 employees.

(c) Sample Selection

A random sampling method is used to select the number of establishments to be enumerated according to the sample size in each employment group.

4.12 NATIONAL PANEL SURVEY

1 Background

The Technical Committee of the MKUKUTA Monitoring System has requested the National Bureau of Statistics (NBS) in late 2006 to establish a technical team to clarify the way forward in conducting a potential series of regular annual National Panel Surveys. The team is formed, and is inclusive of Government, research organizations, and development partners. The team is chaired by NBS and links to the Survey and Routine Technical Working Group of MKUKUTA Monitoring.

2 Purpose of Tanzania's Panel Survey

The National Panel Survey (NPS) is intended to achieve multiple objectives. These objectives are:

- To provide information required for MKUKUTA monitoring, monitoring of other development objectives (MDG, PAF) and monitoring of specific programs;
- To provide high quality nationally representative information on income dynamics at the household level, to provide annual consumption estimates to monitor poverty in years between HBSs and to provide reliable agricultural statistics to feed into the National Accounts;
- To provide a flexible survey instrument able to accommodate ad hoc data requests and to assess impact of (new) policy interventions

3 What is a Panel Survey?

Panel surveys collect data about individuals, households and communities over time, in order to assess change. Because panel surveys usually collect data from the same sources each year they are well suited to assess change, to provide information about the causes of change and poverty dynamics and to assess impact of interventions.

4 Why is it important for monitoring MKUKUTA and other development initiatives?

For MKUKUTA (but also PAF), information from Panel Surveys will help to provide indications of the degree of chronic and transitory poverty, and allows to measure some MKUKUTA outcomes annually (see annexes 1 and 2). This could complement the 5-year HBS cycle, e.g. poverty incidence, inequality or access to clean water.

It may also be a useful instrument/platform to reduce the numerous surveys planned for specific projects, by consolidating them into a single operation. Therefore for MKUKUTA monitoring it potentially links to output reporting via the Annual Implementation Report.

The Panel Survey will also regularly provide information on the state of the agricultural sector and rural livelihoods, and supply information for the monitoring and evaluation of the ASDP and other sectoral interventions.

5 Timeframe

The timeframe reviewed here for the NPS consists of the first 3 years of the panel. This includes 4 months of planning (described below) and 3 annual rounds (12 months each), each including a preparation phase for the subsequent data collection.

When implemented successfully, the panel may be maintained for a much longer period. Some panels have been maintained for as long as 20 years. For outer years, budget information as presented for year 3 is indicative.

6 Institutional Arrangements

The National Bureau of Statistics takes overall responsibility for the National Panel Survey. In doing so they will report progress and seek technical advice and financial approvals through the Survey and Routine Data Technical Working Group, and its on-ward link to the MKUKUTA Technical Committee.

The NBS may choose to form partnerships with other institutions or may contract out elements of the preparation, implementation, data processing or analysis of the National Panel Survey.

Public procurement procedures will be followed in the implementation of this work.

7 Survey Sample

The overall sample design will provide for annual poverty estimates for three strata in mainland Tanzania (rural, Dar es Salaam, other urban areas) plus Zanzibar, and will enable the study of poverty dynamics. The sample will supply annual production estimates for main crop and livestock at the national level and, if feasible and deemed of policy relevance, for major agro-ecological zones. Given the proposed sample size, agricultural data can also be analyzed by ex-post farm typologies e.g. by farm-size, degree of market integration, etc. Information on average cost of production for selected crops at the same level of geographical disaggregation will also be supplied, though less frequently.

The sampling will emphasize randomization to ensure full representative ness at the selected domains of inference. The NPS sample will be drawn from the National Master Sample (NMS). The draft sample design is described in Table 1 below and was based on inputs from sampling/survey experts. It entails interviewing 3,456 households (2,240 rural and 1216 urban) in mainland Tanzania and 576 in Zanzibar (288 rural and 288 urban) for a total annual sample of 4,032 households. The total rural sample will be 2,528 households.

The precise and final sample design will be finalized in the coming months with the support of a sampling expert. At that time, the exact PSUs and sequencing of fieldwork will be identified.

Table 1: Sample Design

	Households	Primary Sampling Units (PSUs)
Mainland	3,456	216
Rural	2,240	140
Dar es Salaam	608	38
Other urban	608	38
Zanzibar	576	36

Rural	288	18
Urban	288	18
Total	4,032	252
Rural	2,528	138
Dar es Salaam	608	38
Other urban	896	56

8 Survey Instruments

To suit its multiple objectives the NPS will be comprised of a set of survey instruments. Specifically, there will be four questionnaires:

- household questionnaire
- agriculture questionnaire (for the subset of NPS households engaged in agricultural activities)
- Community questionnaire
- Price questionnaire

The household questionnaire will include a set of *core modules* (sections) which will be implemented annually. Table 2 presents an overview of the proposed core modules. They deal with the dynamics of poverty (per capita consumption, cash and non-cash income, savings, assets, food security, health and education), vulnerability and social protection and program participation). The core modules will be revised after year 1 to account for the changing socioeconomic environment, but it is expected that they are by-and-large the same in every annual survey round to ensure comparability. The final decisions about what will be core modules will need to be taken by the NPS Technical Committee. The core modules will largely exclude in-depth information on topics covered in on-going survey plans (such as maternal and child health issues covered at length in the DHS).

In addition to the core modules, which are fairly comprehensive, the design of the NPS allows for the introduction of thematic modules on a rotating basis. *Rotating modules* can be added to the NPS to capture information on outcomes that either might not change quickly over time (e.g. perceptions such as currently captured in the Views of the People Survey) or that reflect recent policies and programs of interest. There may also be questions that should be included because there is a special interest such as a module to test the abilities of primary school students. These non-core modules

(sometimes they could be single questions) could be included upon request in which case there is scope for the NPS to become (partially) self-financing as it seems reasonable that those requesting to add non-core modules contribute to the cost of running the NPS. To date various such requests were made including to capture gender and gender violence (UNFPA), nutrition (TFNC/UNICEF), community investments (TASAF) or agriculture (Gates Foundation / ASDP).

Table 2: Overview of Household Questionnaire

Core Modules		
Section		**Level of obs.**
A	Survey information	Household
B	Household member roster (incl. antropometrics)	Individual
C	Education	Individual
D	Health	Individual
F	Labor and Time Use (Adults and Children 10+)	Individual
G	Housing	Household
H	Water, Sanitation and Energy Use	Household
I	Consumption of Food (recall)	Household
J	Household non-food consumption	Household
K	Agriculture (core module)	Household
L	Assets	Household
M	Household Enterprises	Enterprise
N	Transfers and remittances	Household
O	Credit	Loans
P	Recent Shocks	Household
Q	Mortality of household members	Deceased
Rotating (optional) Modules		
Non-core topics *Examples:* Perceptions about government services, sanitation, crime		

In addition, the instruments will include an extensive agricultural module which will be administered to all households in the main sample identified during the main interview as being involved in crop and/or livestock production. Thus, the main questionnaire will contain a selected number of questions on land and livestock ownership and main crops to enable the identification of the relevant households for the follow-up interview which will be administer soon after the core survey by a interviewer with good knowledge of the sector.

The extensive agricultural module with allow, among other things, for the annual estimation of land area, both owned and cultivated, as well as production figures for main crops and livestock. Detailed cost of production for selected crops will also be collected but less frequently. In intermittent years, the collection of additional information for the characterization of the sector, e.g. access to extension services and irrigation facilities, will also be pursued.

The choice of specific agricultural-related topics for inclusion in certain years will be done by NBS in close collaboration with the relevant line ministries. Over a three-year cycle, the instrument will thus provide an array of sectoral indicators at the national and sub-national level, each at the frequency deemed most appropriate also in light of specific reporting requirements.

In addition to a household questionnaire the NPS comprises other questionnaires including a community questionnaire (to assess community characteristics including, economic infrastructure, and the presence of programs and interventions). The inclusion of a price questionnaire is also possible, and will depend on the availability of reliable and timely price information, currently being collected by the Ministry of Industry and Trade.

Draft questionnaires have been prepared and the technical team will revise these taking into account the need to align with other national surveys (HBS, DHS, ILFS) to enhance comparability. The Standards and Methods Department of the NBS will ensure quality of the questionnaire and will also assist with ensuring consistency with other surveys.

9 Project Team

The set up of the NPS project team is intended to ensure that the NPS will provide timely and high-quality data on an annual basis. The team consists of two groups: the Management Team and office support staff, and the mobile Field Team.

The Management Team consists of a Project Manager who oversees all aspects of the NPS. The Project Manager is supported by a full-time International Resident Advisor who will provide, among other things, technical and analytical support for the first two years of the panel. There is a Field Manager who oversees field work by the interview teams and a Data Manager who oversees all aspects of data processing and production. Administrative support is provided by an office assistant and a (part-time) financial officer.

The Field Teams are designed to be small, mobile units comprised of well-trained enumerators who will be employed on an annual basis and possibly retained for several years. By having small field staff in continuous employment and extensive supervision, the project seeks to ensure the quality of the work and the commitment of its staff. There will be six Field teams comprised of: 1 Field Supervisor, 2 Household questionnaire interviewers, 1 Agriculture questionnaire interviewer, 1 Data Entry Operator, and 1 Driver. Data entry will be carried out in the field and concurrently with the fieldwork to checking or errors and corrections in almost real time, as well as speed up the preparation of the final dataset for analysis.

10 Release Calendar and Data access

Within 3 months of completion of the field work of every survey round, a cleaned and anonymized data set will be made publicly available through the NBS web site.

5.0 ECONOMIC STATISTICS.

5.1 ANNUAL SURVEY OF INDUSTRIAL PRODUCTION (ASIP).

Objective and use of product

The survey aims at producing basic industrial statistics in accordance with the UN International Recommendations for Industrial Statistics so as to ensure national and international comparability. The survey is designed to:

- Collect statistics on the operating characteristics and structure of the industrial sector in the country;
- Assess the contribution of industrial sector to the overall economy;
- Obtain data for computing national accounts estimates;
- Obtain data for the construction of Input-Output tables
- Assess the current status of the industrial sector in the country;
- Identify conditions that affect firm-level productivity and competitiveness;
- Secure inputs for use by the Government and other organizations in developing and reviewing economic policies, programmes and strategies that support sector-productivity and growth;
- Provide facts to support dialogue with the Government and other development partners so as to enhance public-private sector partnership; and
- Update the existing data for Monitoring and Evaluation at the level of industrial sub-sectors.

Methodology of data production

Reference period

Data are requested for and in many cases supplied on a calendar year basis. Those units of enquiry that could not supply data according to calendar year due to allocation problems are advised to supply the data according to the financial years that cover larger part of the reference calendar year.

Unit of enquiry

In principal, the statistical unit is an establishment defined as an economic unit, which engages under a single ownership or control in one, or predominantly one kind of economic activity at a single physical location i.e. an individual firm, mine, factory or workshop. However, for certain firms which cannot furnish separate data at an establishment level, the enterprise is used as the unit of enquiry. It should be noted here that, due to record keeping practices of some firms, it has not been possible in every case to strictly follow the definition of an establishment as stated above. Therefore, in a few cases, the restrictions especially on location should be relaxed.

Scope and coverage

Scope

The survey covered all industrial establishments with ten or more workers.

According to the latest UN International recommendations, the industrial sector is comprised of establishments engaged in "Mining and quarrying" (ISIC Revision 4, Section B), "Manufacturing" (Section C), "Electric power generation, transmission and distribution" (Section D) and "Water collection, treatment and supply" (Section E).

Manufacturing, according to international recommendations, is defined as the physical or chemical transformation of materials or components into new products, whether the work is performed by power-driven machines or by hand, whether it is done in a factory or in the worker's home, and whether the products are sold at wholesale or retail prices.

Coverage

Due to budget constraint, the survey covers all industrial establishments with ten or more workers **(**persons engaged**).** The other cluster with 1-9 workers is covered during a census of industrial production supposed to be carried out after every ten years depending on the availability of financial resources.

The questionnaire

The questionnaire is designed according to "The International Recommendations for Industrial Statistics" of the United Nations and takes into consideration the requirements of the stakeholders. The information collected satisfies the needs of National Accounts and are considered to be useful to both the government, researchers and the business community.

The following, is the information collected through the questionnaire:-

General Information, Employment, Labour costs, Input/ Purchases, Output, Inventory of working capital and Expenditure on fixed assets.

Data collection

Data collection is undertaken by a trained team of enumerators under the supervision of supervisors from NBS headquarters. Training of field-staff takes place after the Training of Trainers (ToT). The NBS is responsible for the training of supervisors. The supervisors are then responsible for the training of enumerators in the regions.

Fieldwork is conducted by total enumeration to the establishments. The field-staff visit all the establishments and assist the filling in of the questionnaires. Respondents are asked to provide the data from their accounting records for calendar year of the reference year.

Data processing

Completed questionnaires are sent to the NBS headquarters in Dar es Salaam where they are edited and coded by NBS staff. The data entry system used is CSPro after which the data are transferred to Microsoft Access for generation of tables. Estimates are made by means of substitutions using the secondary data of previous ASIP survey. The cleaned dataset are then used to generate the preliminary tables as per tabulation plan.

Organisation of the Report

The report presents the industrial statistics according to the latest UN recommendations (2008). The tables are grouped separately at the end of the report and refer to establishments with 10 or more workers.

The report also presents the main results relating to both quantitative and qualitative information that were also collected through the survey. Tables referring to such information are integrated within the section giving the major findings. The tables therefore refer to the respondents only as indicated therein.

Reference to applicable standards and classifications

Industrial classification

Establishments are classified into industries on the basis of major activity in conformity with the International Standard Industrial Classification (ISIC) Revision 4. Each industry is basically

defined in terms of its principal products or services, these being similar in nature or commonly associated in production

5.2 THE PRODUCTION INDEX OF MANUFACTURING INDUSTRY (PIMI)

Objective and use of Product

The survey aims at producing basic quarterly statistics on major commodities according to UN International Recommendations for Industrial Statistics. The survey is intended to:

- Collect quarterly commodity statistics on the operating characteristics and structure of industrial sector in the country;
- Assess the contribution of industrial sector to the overall economy;
- Obtain data for computing National Accounts estimates;
- Obtain data for the construction of Input-Output tables
- Assess the current status of the industrial sector in the country;
- Identify conditions that affect firm-level productivity and competitiveness;
- Secure inputs for use by the Government and other organs in developing/reviewing policies, programmes and strategies that support sector-productivity growth;
- Provide the private sector with facts to support dialogue with Government and other partners so as to enhance public-private sector partnership; and
- Update existing Monitoring and Evaluation data at the level of industrial establishments.

Methodology of Data Production

Coverage

The Production Index of Manufacturing Industry (PIMI) measures developments in the production in manufacturing establishments that on average have a labour force of 50 persons or more. However, the index includes some establishments with a labour force of less than 50 persons to cover those activities that otherwise would not be represented. In total, slightly more than 300 establishments are surveyed on quarterly basis, contributing slightly more than 85% to total manufacturing gross output according to the latest industrial census results. The index covers Tanzania Mainland only.

Data Collection and Processing

Data for the Quarterly Survey of Industrial Production is collected through questionnaires which are sent at the end of the reference period to the Regional Statistical Managers and later distributed to the respective establishments. Questionnaires are collected in the first month after the end of the reference period. Follow-ups are made by physical visits and telephone calls. Where necessary, a team of supervisors is sent to the regions to speed-up the process.

The survey collects monthly data on quantities produced with standard units for all major commodities produced by the establishment, together representing at least 85% of output. Both the establishments and the commodities are classified according to International Standard Classification of All Economic Activities (ISIC) Revision 3. In addition, the questionnaire asks for the establishment's number of employees, total labour cost and the turnover during the quarter. The accuracy of the responses is verified both at the regional offices and at NBS headquarters.

Data Processing and Analysis

Data processing takes place centrally in Dar es Salaam. The basket weights refer to annual output valued at ex-factory selling prices for each establishment separately. The lowest level indices are calculated as the production of the current quarter, expressed in units like kilograms, square meters, litres, etc., divided by production in the base year (divided by four to bring it to a quarterly basis). These lowest level indices are aggregated according to the Laspeyre’s formula to obtain group and overall indices.

There are various methods to deal with missing data on quantities produced. If the figure for one-month is missing from the quarterly questionnaire, it is usually estimated on the basis of the data reported for the two other months of the same quarter. If no data is reported, the production changes for the corresponding period in the previous year are applied on the last quoted production. If no other method is applicable, the missing production data is carried forward.

Periodicity

The PIMI is compiled and disseminated on a quarterly basis.

Timeliness

In principle, the PIMI is released six weeks after the reference period

Reference to applicable standards and classifications

Industrial classification

Establishments are classified into industries on the basis of major activity in conformity with the International Standard Industrial Classification (ISIC). Each industry is basically defined in terms of its principal products or services, these being similar in nature or commonly associated in production. The index is compiled at 4-digit ISIC levels, but published in 3-digit ISIC detail.

5.3 INDUSTRIAL CENSUS

Objective and use of Product

It is done after every ten years depending on the availability of financial resources.. The main objective of the Industrial Census is to obtain comprehensive data on the contribution of the Industrial sector to Gross Domestic Product (GDP) and to up-date the previous industrial census data. Broadly, the survey provides baseline data on the economic characteristics of the industrial sector for use in planning, policy formulation, monitoring and evaluation of government programmes aimed at improving the sector and the economy.

Specific Objectives are:

- To obtain an updated list of operating industrial establishments;
- To measure the size of employment in the industrial sector;
- To measure workers remuneration;
- To measure cash income;
- To measure the amount of stocks held;
- To measure value added;
- To measure expenditure on fixed assets;
- To collect information on the principal activity, size and ownership;
- To get benchmark data for other subsequent industrial surveys to be used for updating data series; and
- To enhance technical capacity and workplace skills development for departmental staff.

Methodology of Data Production

Scope and Coverage

The **Scope** is the entire industrial sector (Mining, Manufacturing, Electricity, Gas and Water). The **Coverage** is also the entire industrial sector with total coverage of the 10+ cluster (establishments engaging 10 persons or more) and a sample of the 1-9 cluster (establishments engaging 1-9 persons).

Organization

After the decision to carry out a census, it becomes essential to socialize it as a project. As such, a number of committees have to be formed namely:

- The **Central Census Committee** whose role is to steer the operation. It consists of knowledgeable members who have authority to take decisions. They normally come from government institutions, data producers and data users.

- The **Technical Committee** whose role is to deal with all technical aspects. It is responsible for decisions on recruitment and training of census staff, questionnaire design, sampling procedure, tabulation plan, field work, data processing, analysis and publication of reports.

- **Inter- Institutional Team** comprising subject-matter specialists who are able to coordinate census logistics with crucial role of determining the demand structure of consumers.

Sampling

Sampling is done for those establishments engaging 1 – 9 persons as indicated above.

Census Instruments

Two kinds of questionnaires are used for data collection, long and short questionnaires. The long or the detailed questionnaire is used for establishments engaging ten or more persons and the short questionnaire is used for establishments engaging one to nine persons.

Data collection

Data collection is done through interview. Enumerator is supposed to visit all establishments involved in Industrial Census dealing with Mining and quarrying, manufacturing, Generation and distribution of electricity, gas and water. When the completed questionnaires are received from the

regions, each questionnaire is registered by recording the establishment's Identification number (ESTID), ISIC Code, Region Code, district code and size code. Then the questionnaires are filled region wise ready for data entry.

Manual editing, coding, data entry, validation and tabulation

Editing and coding is done immediately after the arrival of questionnaires at the NBS headquarters. Data entry is done using personal computers to the designed data entry system. Data validation is done to detect outliers' figures and miss-keying in of data by data entry operators. Tabulation is done using statistical packages.

Standards and clarifications

All establishments are classified by ISIC code by considering the main activity of the establishment.

5.4 BUSINESS SURVEY

Business Survey covers all businesses operating on fixed premises and collects basic structural information about each business namely; name and address; location with corresponding Geographical Positioning System (GPS) coordinates, type of premises, main activity and ownership, employment and year of commencement.

Objective

The overall objective of the Business Survey is to provide a sampling frame for business surveys as well as to generate information needed for the construction of a directory of businesses that can be used as a guide to investors, researchers, the business community and the general public.

Specific objectives include the improvement of National Accounts aggregates particularly in the estimation of the country's Gross Domestic Product (GDP) and to supplement the Central Register of Establishments (CRE) particularly on the part of informal industry whose small businesses have always been difficult to capture through the normal registration system of the country due to their extreme volatility and high birth and death rates.

Methodology

1. **Advance Planning**

The first step involves setting up of an efficient administrative system that will directly be responsible for the overall management of the survey.

The second step is to form a dedicated technical team that can manage the technical aspects of the project.

The third step is to establish a stakeholders' committee.

2. Sampling

The best assumption made is to cover the most concentrated areas with business activities using the Central Register of Establishments as a guide.

3. Questionnaire Design

A total of five questionnaires are used namely:

- Questionnaire number one is for the listing of the geographical area names within wards.
- Questionnaire number two is the main questionnaire of the survey and is used to collect business data.
- Questionnaire number three is complementary to questionnaire number two in the sense that when it happens a respondent cannot provide all the information instantly as requested in questionnaire number two then, questionnaire number three is left behind for completion by the respective respondent.
- Questionnaire number four is designed to collect information of branches of head quarter businesses or enterprises.
- Questionnaire number five is designed to list stallholders in recognized markets.

5. Field Organization

Field enumeration is organized in such a way that all levels of Regional Administration and Local Government are involved in one way or another. In addition, the NBS survey team including enumerators and field supervisors are officially introduced to the District Authorities by the respective Regional Administrations before the enumeration. District authorities through their Ward Executive Officers and Street (Mtaa) Chairpersons are responsible for providing escort to the enumerators during enumeration at a very minimal cost.

6. Data Processing

The questionnaires are first pre-edited in the field by supervisors to make sure that they are complete and accurate enough before delivering them to NBS head office. In the case of missing information, the questionnaires are returned to the field for correction.

Questionnaires received at NBS head office are re-edited and entered into the data capture system. Special queries are developed by the IT technician into the system for validation checks of the data variables.

5.5 INTERGRATED BUSINESS SURVEY (IBS)

Objective and use of product

Integrated Business Survey measures the performance of key indicators of the distributive trade and construction sectors of the economy. The information collected is used in the compilation of National Accounts aggregates such as Gross Domestic Product (GDP).

Specific Objectives

- To obtain information on the number and geographic distribution of construction, distributive trade and service establishments;
- To estimate the number of persons engaged in these sectors;
- **O**btain data on the type and flow of goods and services in order to allow policy formulation to strengthen the sectors;
- To get estimates of the contribution of the construction, distributive trade and services activities to Gross Domestic Product (GDP); and
- To obtain data that could be helpful in formulating socio-economic plans and programmes.

Methodology of Data Pproduction

1. Advance Planning

The **first step** involves setting up of an efficient administrative system that would directly be responsible for the overall management of the survey.

The **second step** is to form a dedicated technical team that can manage the technical aspects of the project.

The **third step** is to establish a technical committee.

2. Questionnaire Design

In carrying out the survey, a set of questionnaires are developed taking into account stakeholders' requirements and UN recommendations. For the distributive trade module, two questionnaires, DTS 01 and DTS 02 (large and small) are used as instruments for data collection. The small questionnaire covers establishments with 5 to 9 persons engaged and the large questionnaire covers establishments with 10 or more persons engaged.

Likewise, for the construction module, two questionnaires (small and large) are used for data collection. The small questionnaire (CIS 1 FORM A) covers establishments of contractors in classes 5 to 7 (with an average investment cost of less than 500 million shillings***)*** and the large questionnaire (CIS 2 FORM B) covers establishments of contractors in classes 1 to 4 (with an average investment cost of more than 500 million shillings)

3. Sample design

The classification of activities is based on the International Standard Industrial Classification (ISIC) Revision 4. In practice, complete coverage is impossible due to limited financial resources, the survey is done on a sample basis. Most of the small establishments doing business in those kinds of activities operate in a manner that makes it very difficult to locate during field work.

For the construction activity, all establishments in classes 1 – 4 are fully covered while those establishments in classes 5 – 7 are sampled.

All establishments engaging 10 or more persons are fully covered, while those establishments engaging 5 to 9 persons are sampled. Distributive trade~~,~~ stratification is by the size of establishment i.e. number of persons engaged. The sampling strategy ensures representation by class, size and classification of the economic activities at national level. The stratification in construction is based on the activity and class of the establishment.

4. Field Work

Data is collected from the establishments through interviews. Managers of the respective establishments are visited by enumerators and asked to supply the required information.

5. Data Processing and Management

Data processing centralized at NBS Head Office, starts one month after the commencement of fieldwork and continues concurrently with the field enumeration exercise. The data capture system is developed using CSPro version 4.0, MS Access and MS Excel. The data processing team comprise two permanent IT department staff who become data processing supervisors, three data editors, two data verifiers and nine data entry operators who are temporarily employed. Data entry is accomplished in two months' time.

Data editors are supposed to edit the questionnaires received from the regions by checking the correctness and filling in the respective codes. Edited questionnaires are entered into the computer by data entry operators. On completion of data entry, the data are cleaned by the IT staff assisted by data verifiers for validation checks of the data variables.

The main investigation variables used for checking consistence of data entered in the database are employment, labour cost, expenditures and receipts.

Data files are created into two databases, the first one is for the distributive trade and the second one is for the construction industry.

6. Report Writing

The Integrated Business Survey produces three reports namely; the Methodology or the Technical Report; and two main survey reports namely; the Distributive Trade Report and the Construction Report.

7. National Stakeholders Workshops and Dissemination

After the completion of the report writing, the reports are disseminated through stakeholders' workshops for comments and recommendations before the preparation of final reports.

5.6 TANZANIA TOURISM SECTOR SURVEY

Tanzania Tourism Sector Survey is managed and implemented by five participating institutions namely, the Bank of Tanzania (BOT), the National Bureau of Statistics (NBS), the Ministry of

Natural Resources and Tourism (MNRT), Immigration Department and Zanzibar Commission for Tourism (ZCT). These institutions carried out the International Visitors' Exit Surveys since 2001.

The following are the primary objectives of the survey:

- To provide an up-to-date price information for estimating tourist expenditure in Tanzania in order to improve compilation of National Accounts and Balance of Payments (BOP) Statistics; and

- To collect information for tourism promotion and macro-economic policy formulation.

Methodology

Scope of the Survey

The survey targeted international visitors to Tanzania. A person is considered as an international visitor if he/she travels to a country other than that of his/her usual residence, and outside his/her usual environment for a period not exceeding twelve months and whose main purpose of visit is other than an activity remunerated from within the country visited.

Sample size

The survey sample size is normally designed to capture about one percent of tourists arrivals annually. In order to allow comparability across the years, the sample size has remained constant to the one used during the 2004 International Visitors' Exit Survey.

Survey period

The survey is usually conducted during the tourist peak season, in order to be able to capture most of tourists' characteristics. In that regard, a two-week survey is normally done during the fourth and fifth weeks in the peak season.

Coverage

Since the survey aims at obtaining data on *inter alia,* expenditure and length of stay, it is necessary that it is carried out at departure points. The survey is normally conducted at four major departure points, namely Julius Nyerere International Airport (JNIA), Kilimanjaro International Airport (KIA), Zanzibar Airport (ZAA) and Namanga (NAM) border.

The Questionnaire

The questionnaire used in the survey comprises of four main parts namely: visitor profiles, travel behaviour, expenditure patterns and visitor comments. (Appendix I)

Questions 1 and 2 aimed at establishing visitor profiles (nationality, country of residence and age group).

Questions 3 to 6 targeted at obtaining travel behaviour namely purpose of visit, type of tour (package/non-package), items in the package and nights spent.

Questions 7 to 11 were structured to capture expenditure of visitors and cost of package tour associated with Tanzania. In addition, the questions requested for details on the amount of money spent in Tanzania.

Question 12 was aimed at obtaining comments from the visitors on their perception of Tanzania's tourism industry.

5.7 DISTRIBUTIVE TRADE SURVEY IN TANZANIA

The envisaged survey objectives is to ensure effective compilation, analysis and dissemination of acceptable distributive trade statistics (including hotels, restaurants and tourist service providers) for use by government, business community and general public. Other specific survey objectives are as follows:

- Provide up to date statistical information for the improvement of the compilation of National Accounts (GDP).
- Set benchmarks for distributive trade statistics in Tanzania
- Collect information for planning and macro-economic policy formulation such as employment, marketing and financial services.
- Provide indicators for MKUKUTA monitoring.

Methodology

The survey collects distributive trade statistics using questionnaires designed to obtain basic information from distributive trade outlets. The survey is done on the sample basis covering municipalities, cities, districts headquarters and towns, so only selected wholesale and retail outlets throughout the country are covered. The classification of commodities are based on the International Standard Industrial Classification (ISIC) Revision 3.

The assumption is that, in practice, complete coverage of retail, wholesale is not possible taking into account financial constraints. Also many of the establishments doing business in those kinds of activities operate in a manner that makes it very difficult to find them, even during the time of field work.

The sources of data are wholesale and retail trade outlets selected throughout the country. Questionnaires are administered by a team of the trained enumerators with close supervision from NBS offices throughout the Country. The enumerators do the data collection exercise by using the structured questionnaires.

Neither the mailing system nor the uses of telephones are considered practical for data collection in Tanzania. Therefore the collection medium chosen are personal delivery and recovery of documents by a small group of well-trained enumerators who visit all selected Regions.

5.8 AGRICULTURE SAMPLE CENSUS

The Agriculture Sample Census is designed to meet the data needs of a wide range of users down to district level including policy makers at local, regional and national levels, rural development agencies, funding institutions, researchers, NGOs, farmer organizations, etc. As a result the dataset is both more numerous in its sample and detailed in its scope compared to surveys. To date, Tanzania Agricultural Census is the most detailed in Africa. The census *is* normally carried out in order to:

- Identify structural changes if any, in the size of farm household holdings, crop and livestock production, farm input and implement use. It also seeks to determine if there are any improvements in rural infrastructure and the level of agriculture household living conditions;

- Provide benchmark data on productivity, production and agricultural practices in relation to policies and interventions promoted by the Ministry of Agriculture and Food Security and others stake holders.

- Establish baseline data for the measurement of the impact of high level objectives of the Agriculture Sector Development Programme (ASDP), National Strategy for Growth and Reduction of Poverty (NSGRP) and other rural development programmes and projects.

- Obtain a benchmark data that will be used to address specific issues such as: food security, rural poverty, gender, agro-processing, marketing, service delivery, etc.

Methodology

1. Census Organisation

The Census is conducted by the National Bureau of Statistics (NBS) in collaboration with the sector Ministries of Agriculture, and the Office of the Chief Government Statistician in Zanzibar (OCGS). At the National level the Census is headed by the Director General of the National Bureau of Statistics with assistance from the Director of Economic Statistics. The Planning Group oversee the operational aspects of the Census and this consists of staff from the Department of Agriculture Statistics of NBS and representatives of the Department of Policy and Planning of the Ministry of Agriculture and Food Security (MAFS). At the regional level, implementation of census activities is overseen by the Regional Statistical Office of NBS and the Regional Agriculture Supervisor from the Ministry of Agriculture and Food Security. At the District level, the Census activities are managed by two Supervisors from the President's Office, Regional Administration and Local Government (PORALG). The supervisors manage the enumerators who also come from PORALG.

2. Tabulation Plan Preparation

The tabulation plan is developed following user group workshops and thus reflects the information needs of the end users. It takes into consideration the tabulations from previous censuses and surveys to allow trend analysis and comparisons.

3. Sample Design

The Mainland sample consists of 3,221 villages. These villages are drawn from the National Master Sample (NMS) developed by the National Bureau of Statistics (NBS) to serve as a national

framework for the conduct of household based surveys in the country. The National Master Sample was developed from the 2002 Population and Housing Census. The total *Tanzania* Mainland sample is 48,315 agricultural households. In *Tanzania* Zanzibar, a total of 317 EAs are selected and 4,755 agriculture households are covered. Nationwide, all regions and districts are sampled except three urban districts (two from Mainland and one from Zanzibar).

4. Questionnaire Design and Other Census Instruments

The questionnaires are designed following user meetings to ensure that the questions asked are in line with users data needs. Several features are incorporated into the design of the questionnaires to increase the accuracy of the data:

- Where feasible all variables are extensively coded to reduce post enumeration coding error.

- The definitions for each section are printed on the opposite page so that the enumerator could easily refer to the instructions whilst interviewing the farmer.

- The responses to all questions are placed in boxes printed on the questionnaire, with one box per character. This feature makes it possible to use scanning and Intelligent Character Recognition (ICR) technologies for data entry.

- Skip patterns are used to reduce unnecessary and incorrect coding of sections which do not apply to the respondent.

- Each section is clearly numbered, which facilitated the use of skip patterns and provide a reference for data type coding for the programming of CSPro, SPSS and the dissemination applications.

Three other instruments are used:

- Village Listing Forms are used for listing households in the village and from this list a systematic sample of 15 agricultural households are selected.

- A Training Manual which is used by the trainers for the cascade/pyramid training of supervisors and enumerators

- Enumerator Instruction Manual which is used as reference material.

5. Field Pre-testing of the Census Instruments

The Questionnaire should *be* pre-tested in areas where different characteristics of farming and livestock keepers are found. This is conducted in order to test the wording, flow and relevance of the questions and to finalise crop lists, questionnaire coding and manuals. In addition to this, several data collection methodologies are to be finalised, namely, livestock numbers in pastoralist communities, cut flower production, mixed cropping, use of percentages in the questionnaire and finalising skip patterns and documenting consistency checks.

6. Training of Trainers, Supervisors and Enumerators

During training, cascade/pyramid training techniques are employed to maintain statistical standards. The top level of training is provided to national and regional supervisors. The trainers are members of the Planning Group from the National Bureau of Statistics and the sector Ministries of Agriculture. In each region, three training sessions are conducted for the district supervisors and enumerators. In addition to training them in field level Census methodology and definitions, emphasis is placed on training the enumerators and supervisors in consistency checking. Tests are given to the supervisors and enumerators and the best 50 percent of the trainees are selected for the enumeration of the smallholder questionnaire and the community level questionnaire. The household listing exercise is carried out by all trained enumerators.

7. Information, Education and Communication (IEC) Campaign

Radios, televisions, newspapers, leaflets, t-shirts and caps are used to publicize the Agriculture Sample Census. This helps in sensitizing the public for the field level activities. The t-shirts and caps are given to the field staff and the village chairpersons. The village chairpersons help to locate the selected households.

8. Data Collection

The data collection methods used during the census is by interview and no physical measurements, e.g., crop cutting and field area measurement is taken. Field work is monitored by a hierarchical system of supervisors at the top of which is the Mobile Response Team followed by the Regional Supervisors and District Supervisors. The Mobile Response Team consists of Principal Supervisors who provide overall direction to the field operations and responded to queries arising outside the

scope of the training exercise. The mobile response team consists of the Manager of Agriculture Statistics Department, the Desk Officer for the Census and senior officers. Decisions made on definitions and procedures are then communicated back to all enumerators via the Regional and District Supervisors.

9. Field Supervision and Consistency Checks

Enumerators are trained to probe the respondents until they are satisfied with the response given before they record them in the questionnaire. The first check of the questionnaires is done by enumerators in the field during enumeration. The second check is done by the district supervisors followed by Regional and National Supervisors. Supervisory visits at all levels of supervision focus on consistency checking of the questionnaires. Inconsistencies encountered are corrected, and where necessary call backs are made by the enumerator to obtain the correct information. Further quality control checks are made through a major post enumeration checking exercise where all questionnaires are checked for consistencies by supervisors in the district offices.

10. Data Processing and Analysis

Data processing consists the following processes:

- Data entry
- Data structure formatting
- Batch validation
- Tabulation

11. Data Entry

Scanning and ICR data capture technology for the small holder questionnaire are used. This not only increases the speed of data entry, it also increases the accuracy due to the reduction of keystroke errors. Interactive validation routines are incorporated into the ICR software to trap errors during the verification process.

Prior to scanning, all questionnaires undergo a manual cleaning exercise. This involves checking that the questionnaires have a full set of pages, correct identification and good handwriting.

CSPro is used for data entry of all Large Scale Farm and Community based questionnaires due to the relatively small number of questionnaires. It is also used to enter small holder questionnaires that are rejected by the ICR extraction application.

12. Data Structure Formatting

A program is developed in visual basic to automatically alter the structure of the output from the scanning or extraction process in order to harmonise it with the manually entered data. The programmes automatically checks and change the number of digits for each variable, the record type code, the number of questionnaires in the village, the consistency of the Village ID Code and saves the data of one village in a file named after the village code.

13. Batch Validation

A batch validation program is developed in CSPro in order to identify inconsistencies within a questionnaire. This is in addition to the interactive validation during the ICR extraction process. The procedure varies from simple range checking within each variable to more complexes checking between variables. After the long process of data cleaning, the tabulations are prepared based on a pre-designed tabulation plan.

14. Tabulations

Statistical Package for Social Sciences (SPSS) is used to produce the Census tabulations and Microsoft Excel is used to organize the tables and compute additional indicators. Excel is also used to produce charts while ArcView and Freehand are used for the maps.

15. Analysis and Report Preparation

The analysis in the reports focuses on regional comparisons, time series and national production estimates. Microsoft Excel is used to produce charts; ArcView and Freehand are used for maps, whereas Microsoft Word is used to compile the report.

16. Data Quality

A great deal of emphasis is placed on data quality throughout the whole exercise from planning, questionnaire design, training, supervision, data entry, validation and cleaning/editing. As a result of this, it is believed that the census is highly accurate and representative of what is experienced at field level during the Census year. With very few exceptions, the variables in the questionnaire are within the norms for Tanzania and they follow expected time series trends when compared to historical data.

Large Scale Farms

The survey covers all Large Scale Farms in Tanzania Mainland. There is a total of 968 Large **Scale Farms in Tanzania Mainland.**

Methodology

The designed questionnaires that are used in the Agriculture Sample Census have been reviewed based on the census results *and* are posted to the Regional Statistical Managers who collect the data from all the 968 large scale farms. The reference period is the agriculture year. An agriculture year in Tanzania commences on 1st October and ends on 30th of September of the following year. A list of large scale farms is updated annually by the Regional Statistical Managers. The questionnaires are edited by the Regional Statistical Managers before they are posted to *NBS headquarters,* Dar es Salaam for data entry and analysis.

Data entry

The data is being entered manually using CSPro. Data cleaning is done by the subject matter specialist and programmer.

Tabulation

Table format is being given to the programmer who produces the tables for analysis.

Analysis

Analysis takes place in the Department *of Agriculture*.

Routine Data

Quarterly agricultural production and producer unit prices are *normally* collected by the Extension Officers from the Villages.

Methodology

The data is collected at the village level by the extension officers who estimate production of various crops grown in the respective villages. They *then* send the data to the Districts who compile them to give District estimates of production and producer unit price of each crop. The Regional Statistical Managers collect the returns and send the data to the *NBS* head office.

Data Entry

The Regional Statistical Managers enter the data in the e-excel spread sheet and send them to the NBS head office.

Data Cleaning

The Regional Statistical Managers check the data for consistency before entering them in the spread sheet.

Analysis

The Department of Agriculture do*es* not produce any report on the data collected but compile*s* them to give Regional Production and unit price of each crop and submit to the National Accounts Department for the calculation of GDP.

5.9 GROSS DOMESTIC PRODUCT (GDP)

5.9.1 ANNUAL GROSS DOMESTIC PRODUCT

Gross Domestic Product (GDP) is the sum of values added of all domestic producers in the economy. It represents the money value of all goods and services produced within a country out of economic activity during *a* specified period usually a year, before the provision of the consumption of fixed capital.

The basic formula for calculating the GDP is:

$$Y= C+ I + E + G$$

Where

Y = GDP

C = Consumer Spending

I = Investment made by Industry

E = Excess of Export over Imports (X – M)

G = Government Spending

Objectives of the Gross Domestic Products are

1. To measure the performance of each economic activity in the country.
2. To measure the contribution of each economic activity in the economy.
3. To compare economic performance among different economic territories.

Methodology of data production

Annual Gross Domestic Product of Tanzania is mainly compiled using production and expenditure approaches.

(i) Production Approach

In this approach, GDP estimates at market prices are derived by summing up the gross value added at basic prices of each industry and adding taxes less subsidies on products. This approach is used for the compilation of all sectors except the government and *the* Central Bank.

(ii) Expenditure Approach

Expenditure approach considers the income expenditure for the purpose of consumption or capital formation. It is used in estimating GDP for the Government and the Central Bank. Not all products purchased by domestic buyer come from domestic production; some come from outside the country. Therefore, the GDP at market prices by this approach is derived by adding up the purchases that are made for final consumption, capital formation, and imports less export (net export).

Reference to applicable Standards and Classifications

Annual Gross Domestic Product uses the International Standards of Industrial Classification (ISIC *Revision.* 3) which classifies economic activities into 15 categories namely; (A) Agriculture, Hunting and Forestry, (B) Fishing, (C) Mining and Quarrying; (D) Manufacturing; (E) Electricity Gas and Water Supply; (F) Construction; (G) Wholesale and Retail Trade, Repair of motor vehicles, motorcycles, personal and household goods, (H) Hotels and Restaurants; (I) Transport, Storage and Communications; (J) Financial Intermediation, (K) Real Estate, Renting and Business Services (L) Public Administration and Defense; Compulsory Social Security; (M) Education, (N) Health and Social Work; (O) Other Community, Social and Personal Services.

5.9.2 QUARTERLY GROSS DOMESTIC PRODUCTS

Quarterly National Accounts (QNA) provides up-to-date information for monitoring economic cycles and short term changes in the economy. It is also *a* requirement for Tanzania to graduate from GDDS to SDDS.

Objectives of Quarterly Gross Domestic Product

(i) The main purpose of QGDP is to provide a picture of current economic development *which* is more timely than that provided by *Annual National Accounts (ANA)*, and *is* more comprehensive and coherent than that provided by individual short-term indicators.

(ii) Therefore, QGDP should be timely, coherent, accurate, comprehensive, and reasonably detailed

Methodology

The compilation procedures adopt the same principles, definitions, and structure as the Annual Gross Domestic product (GDP). Currently in Tanzania, only the production approach is applied in the quarterly estimations. Compilation of quarterly GDP can either be done directly or indirectly depending on the data available. Indirect approach uses short term indicators within a framework of a statistical model to generate quarterly estimates.

Compilation of quarterly GDP can either be done directly or indirectly depending on the data available. Indirect approach uses short term indicators within a framework of a statistical model to generate quarterly estimates. Most of the data available for quarterly compilations are on quantities of output, thus, the quarterly value added estimates by activity are mainly output-indicator based. QNA for Tanzania Mainland mainly relies on the indirect approach where quarterly *Gross Value Added (GVA)* estimates are based on output indicators obtained from administrative records. In cases where more than one indicator is used, a weighted index is generated using base year weights.

Reference to applicable Standards and Classifications

Quarterly Gross Domestic Product uses the International Standard Industrial Classification (ISIC Revision 3) which classifies economic activities into 15 categories namely; (A) Agriculture, Hunting and Forestry, (B) Fishing, (C) Mining and Quarrying; (D) Manufacturing; (E) Electricity Gas and Water Supply; (F) Construction; (G) Wholesale and Retail Trade, Repair of motor vehicles, motorcycles, personal and household goods, (H) Hotels and Restaurants; (I) Transport, Storage and Communications; (J) Financial Intermediation, (K) Real Estate, Renting and Business Services (L) Public Administration and Defense; Compulsory Social Security; (M) Education, (N) Health and Social Work; (O) Other Community, Social and Personal Services

5.9.3 REGIONAL GROSS DOMESTIC PRODUCT/REGIONAL NATIONAL ACCOUNTS

Refers to coherent, consistent and integrated sets of macroeconomic accounts and tables designed for a variety of analytical and policy purposes but compiled at regional level. In Tanzania Mainland, there are 25 regions and each region constitutes an economic territory used for the compilation of Regional National Accounts.

Objectives of regional gross domestic product

a. To measure the performance of each economic activity in the region.
b. To measure *the contribution* of each economic activity in the economy.
c. To compare economic performance among different regions in the country.

Methodology

Regional National Accounts are compiled in accordance with International Standards (United National System of National Accounts *(SNA) 1993*. The following two approaches have been recommended for the 2001 revision of National Accounts.

Production Approach

In this approach, GDP estimates at market prices are derived by summing up the gross value added at basic prices of each industry and adding taxes less subsidies on products. This approach is used for the compilation of all sectors except the government and the Central Bank.

Expenditure Approach

Expenditure approach; considers the income expenditure for the purpose of consumption or capital formation. It **is** used in estimating GDP for the Government and The Central Bank. Not all products purchased by domestic buyer come from domestic production; some come from outside the country. Therefore, the GDP at market prices by this approach is derived by adding up the purchases that are made for final consumption, capital formation, and imports less export (net export)**.**

Reference to applicable Standards and Classifications

Regional National Accounts are compiled in accordance with International Standards (United National System of National Accounts *(SNA) 1993*.

5.9.4 NATIONAL ACCOUNTS PUBLICATION

This "National Accounts of Tanzania Mainland 1999-2009" publication gives consolidated estimates of Gross Domestic Product (GDP) and other allied aggregates of Tanzania Mainland.

Methodology

The methodologies used are based on the 1993 SNA. The following two approaches are recommended for the 2001 revision of National Accounts.

Production Approach

In this approach, GDP estimates at market prices are derived by summing up the gross value added at basic prices of each industry and adding taxes less subsidies on products. This approach is used for the compilation of all sectors except the government and the Central Bank.

Expenditure Approach

Expenditure approach; considers the income expenditure for the purpose of consumption or capital formation. It is used in estimating GDP for the Government and the Central Bank. Not all products purchased by domestic buyer come from domestic production; some come from outside the country. Therefore, the GDP at market prices by this approach is derived by adding up the purchases that are made for final consumption, capital formation, and imports less export (net export)**.**

Reference to applicable Standards and Classifications

Tanzania Mainland uses the International Standards of Industrial Classification (ISIC Revision 3) which classifies economic activities into 15 categories namely; (A) Agriculture, Hunting and Forestry, (B) Fishing, (C) Mining and Quarrying; (D) Manufacturing; (E) Electricity Gas and Water Supply; (F) Construction; (G) Wholesale and Retail Trade, Repair of motor vehicles, motorcycles, personal and household goods, (H) Hotels and Restaurants; (I) Transport, Storage and Communications; (J) Financial Intermediation, (K) Real Estate, Renting and Business Services (L) Public Administration and Defense; Compulsory Social Security; (M) Education, (N) Health and Social Work; (O) Other Community, Social and Personal Services.

6.0 STATISTICAL SERVICES

6.1 CENTRAL REGISTER OF ESTABLISHMENTS (CRE)

Introduction

The Central Register of Establishments (CRE) is a directory or a list of all operating establishments in Tanzania Mainland. An establishment is defined as an enterprise or part of an enterprise that is situated in a single location and in which only a single (non-ancillary) productive activity is carried out or in which the principal productive activity accounts for most of the value added. The register includes business entities operated by working proprietors and unpaid family workers or co-operators.

It was established at the Bureau of Statistics (BOS) in 1978 when a Technical Committee was formed under the chairmanship of the Head of Industrial Statistics Section to harmonise the various registers which had existed then. The various registers were merged into one standard register – the CRE.

The register was created after a need had been felt by BOS to have a permanent and standard register using the same codes classifying region, district, economic activities (using ISIC codes) and sector for all the Departments of the BOS. Since then the register has been updated and maintained by the National Bureau of Statistics.

Objectives and Use of Product

The main objective of CRE is to be used as sampling frame for different establishments based surveys. This implies that CRE shall cover all economic activities in the country.

Another use is to show statistical information on distribution of all economic units, employment by sector and geographical areas. This is very useful to planners and decision makers of the economy.

Also, several public institutions concerned with enforcement of employment laws and regulations find the register useful e.g. VETA, NSSF, PPRC, etc.

Outside NBS the following institutions are the major users of the register: Ministry of Labour for factory and industry inspection. Income Tax department uses for income tax administration and verification. NSSF uses CRE for the administration of worker's social security.

Methodology of Data Production

The main source of information is the National Bureau of Statistics through its surveys like the Survey of Employment and Earnings, Large Scale Farms, Annual Survey of Industries, Hotel Statistics, etc.

Also, it has become evident that data on existing and operating establishments is collected through different sources like visiting the site, using Tax returns, Ministry of Labour, Ttelephone Ddirectories, Post Ooffice Directory, Pay As You Earn (PAYE), News Papers, Municipalities and District Councils, Registrar of Companies, Transport Licencing Board, Board of Internal Trade, Regional and District Trade Officers.

Updating the CRE

CRE is being updated continuously throughout the year but a report is produced every June and December each year. The updating activity involves adding new establishments when they start operating and delete the ones which are no longer in existence and nothing changes in the establishments' characteristics. A Technical and Statistical Report on CRE is released annually, starting with data from 1992.

Reference of Applicable Standards and Classifications

The register consists of a list of a single line records. Each line as a complete record describing fully an establishment by region, district serial number in the district, the name and postal address, location, industrial code, ownership and number of employees of the establishment just to mention a few.

The location describes the physical area where the establishment is situated. The four digit industrial codes describe the main economic activity of the establishment. The codes are according to the United Nations International Standard Industrial Classification (ISIC) of all economic activities. The sector code classify the establishments according to types of ownership i.e. Private, Parastatals, Governments, etc.

6.2 REGIONAL AND DISTRICT SOCIAL ECONOMIC PROFILES

Introduction

Regional and District Authorities are responsible for preparing Social Economic Profiles within their respective areas with technical support from the National Bureau of Statistics. The process of preparing such profiles includes compilation of social economic activities in the Region and Districts.

The profiles cover the location of region and district, area and all social economic development activities undertaken in the respective region or district. In social activities the profiles cover sectors like education, health, water and sanitation In economic activities there are sectors like agriculture, industry, trade, transport and tourism.

Main Objective and Use

The main objective of Regional and District Social Economic Profiles is to provide regions and districts with useful information for planning, policy formulation and appropriate decision making. The information enables regions and ~~or~~ districts to identify areas of improvement and also to identify investment opportunities, available in the regions and districts.

Methodology

Regional and District socio-economic profiles use~~d~~ secondary data from all sectors or departments of the district councils. Templates which cover all these sectors are prepared and distributed to all heads of departments to be filled in. The filled in templates are then collected for compilation, processing, analysis and production of the report for the respective region or district.

6.3 TANZANIA SOCIO-ECONOMIC DATABASE (TSED)

The Tanzania Socio-Economic Database (TSED) is a comprehensive and up-to-date socio-economic database system. It is a powerful tool for organizing, storing and presenting data in a uniform way, allowing data to be easily and quickly shared across government departments, UN agencies and other development organizations. The database is compliant with International Statistical Standards and operates both as a desktop application (on CDs) as well as on the website (www.tsed.org).

Objectives and Uses of Products

The main objective of TSED is to democratize access, use and dissemination of accurate data on a wide range of socio-economic indicators in a user-friendly manner, and thus:

a) Facilitates the systemization, storage and analysis of performance indicators in different thematic areas that are defined by users;

b) Allows for user-friendly analysis of data through tables, graphs and maps for inclusion in reports, presentation and advocacy materials

c) The system allows the creation of global, regional, national and local maps, with allowances up to 10 geographic levels;

d) Allows the grouping of indicators in different frameworks, sectors and sub-sectors, by themes (poverty disparities etc), by institutions, by sources, by international commitments and goals;

e) Provides updated time series data as well as multiple estimates from various sources, disaggregated data to village level, by sex and urban/rural strata whenever these are available; and

f) Allows the creation, modification and merging of indicator databases without the need for specialised programmes or technical support. The system contains the wizard that provides step by step orientation for these tasks.

Methodology of Data Production

TSED is an indicator Database that stores data from various statistical reports published by NBS and other MDAs. Indicators available in TSED developed in collaboration with all MDAs within Tanzania Mainland and the data uploaded into TSED are from recognised official sources.

6.4 TANZANIA INTEGRATED STATISTICAL DATABASE (TISD)

The NBS library is supported by the Tanzania Integrated Statistical Database (TISD) established in 2003. The TISD comprises a large computer network within the NBS where data is stored. The main objective of TISD is to have a “one stop shop” for datasets (in soft and hard copies) and other documents published within and outside the country, thus provide policy makers, administrators,

academicians, Civil Society Organisations, development partners and the general public with more reliable statistical data in a timely manner

6.5 TANZANIA NATIONAL DATA ARCHIVE (TNADA)

Tanzania National Data Archive (TNADA) is a web-based cataloging system that serves as a portal for researchers to browse, search, compare, apply for access and download relevant census or survey information.

It provides a powerful instrument that facilitates the process of releasing study metadata and micro data to the user community. It allows for:

- Increased quality and diversity of research;
- Improved reliability and relevance of data;
- Reduced duplication of data collection activities;
- Improved visibility of the institution as their data becomes more frequently used and is more readily accessible;
- Increased donor and public confidence in the institution;
- Improved publishing and dissemination efficiency of the National Statistical Office (NSO); and
- Access to survey information such as reports, tables, and micro data.

Twelve Studies have already been uploaded and more are still in the process

6.6 NBS WEBSITE

This is the tool developed to assist the National Bureau of Statistics to disseminate its products and services to statistical users and producers in the Government and private sectors.

The presence of NBS website has made it possible for staff and stakeholders to benefit from the shared global information resources and knowledge. The availability of NBS Website (www.nbs.go.tz) has also increased the possibility to collaborate and share information with participating countries and development partners.

Objective

The main objective of the NBS website is to improve access and visibility of the products produced by the NBS such as survey reports ;which include ; Households Budget Surveys, Economic Surveys, Population and Housing Census, Agricultural Surveys, and Demographic Surveys.

The focus target groups are;- Government officials including; planners, policy analysts and parliamentarians, researchers, NGOs, media and the public in general.

Methodology

The document format used to display and deliver different types of documents on a website is PDF. The end use needs to have Acrobat Reader for PDF to open the document and the internet connection would be necessary for downloading.

6.7 SAMPLING

Sampling refers to selecting a subset of elements from a population or a full set of element. The usual goal in sampling is to produce a representative sample. A sample is similar to the population on all characteristics, except that it includes fewer elements because it is a sample rather than the complete population. A perfect representative sample would be a "mirror image "of the population from which it was selected, except it would include fewer elements. The selection of a sample can either be through probability or non – probability mechanism.

Objective

To design a representative sample that is cost effective within the agreed timeframe.

Methodology

NBS maintains two types of frames, the Population and Housing Census Frame is used to design households based surveys while the CRE is used for designing establishment based surveys. Normally, the households based surveys design uses multistage cluster sampling and elements are selected systematically or a complete coverage is done within the selected cluster. In establishments survey, cut off sampling is common to some threshold establishments revenue while the remaining establishments may be selected using probability proportional to size sampling or systematic sampling.

Before enumeration updating of selected clusters is done to reflect the current situation.

Reference to applicable standards

All sampling procedures and methods follow international standards using L. Kish (1966) and W. Cochran.

6.8 ENVIRONMENT STATISTICS

Environment is a cross-cutting issue that involves all sectors of the economy.

Environment Statistics describe the state and trends of the environment in the media of the natural environment, air, climate, water, land/soil, the biota within the media and human settlements. Environment Statistics are integrative in nature, measuring human activities and natural events that affect the environment as well as the impact of these activities and events, also social responses to environmental impacts and the quality and availability of natural assets.

Main Objective

The main objective of producing environment statistics is to enable planners and decision makers to find sustainable solutions to the growing pressure on environment for basic needs. This also helps to increased environment awareness by the general public.

Methodology

Environment Statistics cover all sectors which in one way or another are related to environment. The information is collected through Routine Data System which comprises all Ministries, Departments and Agencies (MDAs) by sending official letters with required templates. The collected information is compiled, edited, entered to the computer followed by analysis and production of report.

The frequency of producing the Environment Statistics Document in Tanzania Mainland is two years,

6.9 RECRUITMENT AND TRAINING OF STAFF ON PERMANENT AND TEMPORARY TERMS

6.9.1 RECRUITMENT

Recruitment Policy of NBS is guided by the Public Service Act of 2002 and its Regulations of 2003 as amended from time to time.

The objective

The objective of the Policy is to ensure that recruitment is based on merit and open competition in order to get competent and right staff to do the right job at the right time.

Methodology

(i) Advertising

Vacancies in NBS are advertised in newspapers or public notice boards calling upon interested and qualified applicants to apply for the post.

Applicants for the post are required to submit letters of application together with copies of certificates, testimonials, curriculum vitae and other relevant documents to the Director General.

(ii) Short listing.

After receiving applications, NBS prepares a full list of all the applicants indicating those who best qualified to appear before Recruitment Committee for interview on a specific date, time and place.

(iii) Interview and Selection

Recruitment Committee prepares questions and marking schemes to be used for interview to ensure objectivity, consistency and to measure the right qualities of the applicants. Applicants who appear for interview are required to submit their original certificates, testimonials and other relevant documents for further scrutiny.

After the interview, the Recruitment Committee selects suitable candidates for the post and submit their names in order of merit to the Appointing Authority (Director General), who will effect appointments in accordance with the required number. In

selecting candidates for employment, Director General takes into account any technical advice given by Recruitment Committee.

Then, Director General issues the letter of appointment to the appointed candidates.
The probationary period for staff employed on permanent and pensionable terms is required to be not more than 12 months from the date of appointment.

6.9.2 STAFF TRAINING AND DEVELOPMENT

Training Objectives:

The fundamental objectives of training shall be:

- Need for trained and skilled manpower to reach organizational objectives;
- Develop the competencies of employees and improve their performance;
- Help staff to grow within the NBS in order to meet future need of human resources;
- To solve a particular performance gap as identified through performance appraisal.; and
- Motivating staff;

Methodology

- Training Programmes based on Training Needs Assessment and Performance Appraisal findings will be prepared, maintained and coordinated.

- There shall be a regular Training Needs Analysis (TNA) exercise, which will be used to determine the training programme and plans.

- NBS will endeavour to train its staff locally and on job training. However, NBS may sponsor a limited number of employees to participate in training programmes outside Tanzania to share experiences with other relevant people and where such training can not be offered locally.

- The Training Programme will be presented before the Workers Council Meeting for discussion/sharing and then back to the Management for its implementation.

- Training will take place both in training institutions and through on the job training.

- Each department shall provide on-the-job training programmes to improve job knowledge and skills.

- Induction Training will be conducted to newly employed staff to orient them to the organization.

- Staff may be allowed to attend any short course after serving the NBS for at least 6 months. However depending on the nature and importance of the course to NBS, staff may be allowed to attend a course even after serving a lesser period as per Director General's approval.

- For a long course, the employee must have been confirmed and served the NBS for at least two (2) years.

- No employee shall be allowed to proceed on training before the expiry of at least two (2) years after finishing a long course.

- The NBS Staff undertaking various courses shall be required to submit report, research paper (Dissertation/Thesis) to the Director General after completion of their studies.

- Where an employee secures a private or self sponsored training he/she will be required to get an approval from the Director General. Such approval shall be given only if such training is relevant to the organization.

- The NBS will refund to the privately self sponsored employee the full cost of the course provided he/she qualifies for and obtains accreditation in terms of performance standard specified by the institution running the training course. However, this refund does not cover employees who have secured scholarships.

- All applications for training by NBS staff shall be made after getting the approval of the Director General.

6.10 NBS LIBRARY

Is a room set apart where printed and none printed materials special for dissemination and retrieval of information are kept. These materials include; books, journals, periodicals, newspapers, CDs, and DVDs, for studying, reference and reading.

Library Objectives are:

- To provide current and semi-current information to statistical users, e.g.; Researchers, Academicians, Universities, Government institutions, NGOs, and other stakeholders.

- To support research activities, conducted by different stakeholders.

- To provide statistical information to domestic and external clients.

Reference to applicable Standards and Classifications

NBS library has several reference materials (Publications) for use by its users.

Classification of NBS library materials

Classification of NBS Library materials is categorized into 10 groups as follows:-

Groups; 0: General

1. Population, Labour and Housing
2. Agriculture, Forest and Fisheries
3. Mining, Manufacturing, Electricity, Gas, Water supply and Construction
4. Transport, Communications, Commerce and Expenditure
5. External trade
6. Wage, Price, Family Income and Expenditure
7. Enterprise and Establishments
8. Banking, public Finance and National Accounts
9. Education, Health and Social Statistics
10. Others
11. Thesis

6.11 FINANCE

6.11.1 FINANCIAL RESOURCES

Introduction

Financial resources are one of the key elements for successful implementation of any institutional goal. In this respect, the NBS strives for availability of adequate financial resources from the government and other sources including commissioned work, loans, and grants from development partners. Government subvention is mainly for paying staff salaries, production of core statistics and other administrative costs .

Budget process

The NBS's Budgets are prepared in accordance with Medium Term Expenditure Framework (MTEF) and budget guidelines as issued by the Treasury. The Heads of Departments have overall responsibilities for ensuring that departmental budgets are produced and submitted to the Financial Manager in accordance with agreed timeframe.

The budget committee is formed to facilitate and consolidate budget inputs from all departments and submit to NBS management and later on scrutinized by the Directorate of Finance. The drafted budget is thereafter shared to all staff through Workers Council before submitting to the MAB for approval and later on , to the responsible ministry for funding.

Commissioned Work

Apart from Government subvention, the NBS generates its own revenue from commissioned work which forms an important source of internal revenue for NBS.

6.12 GEOGRAPHICAL INFORMATION SYSTEM- (GIS)

Introduction

A Geographical Information System (GIS) is a computer based system which is used to digitally produce and/or reproduce and analyze the features present on earth surface and the events that take place on it. A GIS is an information system designed to work with data referenced by spatial /geographical coordinates. In other words, GIS is both a database system with specific capabilities

for spatially referenced data as well as a set of operations for working with the data. It may also be considered as a higher order map.

Components of GIS

GIS constitutes of five key components:

- Hardware - Consists of the computer system on which the GIS software will run
- Software - GIS software provides the functions and tools needed to store, analyze, and display geographic information
- Data - Geographic data and related tabular data can be collected in-house or purchased from a commercial data provider. The digital map forms the basic data input for GIS
- People - GIS users range from technical specialists who design and maintain the system to those who use it to help them perform their everyday work
- Method - The models and operating practices unique to each organization

1.0 Mapping Concepts, Features and Properties

A map represents geographic features or other spatial phenomena by graphically conveying information about locations and attributes. Location information describes the position of particular geographic features.

The basic objective of mapping is to provide:

- Descriptions of geographic phenomenon
- Spatial and non-spatial information
- Map features like Point, Line, and Polygon.

1.1 Map Features

Location information is usually represented by points for features such as wells and telephone pole locations, lines for features such as streams, pipelines and contour lines and areas for features such as lakes, counties and census tracts.

Point feature

A point feature represents a single location. It defines a map object too small to show as a line or area feature. A special symbol of label usually depicts a point location.

Line feature

A line feature is a set of connected, and ordered coordinates representing the linear shape of a map object that may be too narrow to display as an area such as a road or feature with no width such as a contour line.

An area feature is a closed figure whose boundary encloses a homogeneous area, such as a state country soil type or lake.

1.2 Map Characteristics

In addition to feature locations and their attributes, the other technical characteristics that define maps and their use include:

- Map Scale
- Map Accuracy
- Map Extent and
- Data- Base Extent

1.3 Scale

To show a portion of the Earth's surface on a map, the scale must be sufficiently adjusted to cover the objective. Map scale or the extent of reduction is expressed as a ratio. The unit on the left indicates distance on the map and the number on the right indicates distance on the ground. Map scale indicates how much the given area has been reduced. For the same size map, features on a small-scale map (1:1, 000, 000) will be smaller than those on a large-scale map (1:1,200). A map with less detail is said to be of a smaller scale than one with more detail. Cartographers often divide scales into three different categories:

Small-scale maps have scales smaller than 1: 1,000,000 and are used for maps of wide areas where not much detail is required.

Medium-scale maps have scales between 1: 75,000 and 1: 1,000,000.

Large-scale maps have scales larger than 1: 75,000. They are used in applications where detailed map features are required.

1.3.1 Scale in Digital Maps

With digital maps, the traditional concept of scale in terms of distance does not apply because digital maps do not remain fixed in size. They can be displayed or plotted at any possible magnification.

In digital mapping, the term scale is used to indicate the scale of the materials from which the map was made. For example, if a digital map is said to have a scale of 1:100,000, it was made from a 1:100,000-scale paper map.

1.4 Types of Information in a Digital Map

Three general types of information can be included in digital maps:

- Geographic information, which provides the position and shapes of specific geographic features.
- Attribute information, which provides additional non-graphic information about each feature.
- Display information, which describes how the features will appear on the screen.

1.5 Map Resolution

Map resolution refers to how accurately the location and shape of map features can be depicted for a given map scale. Scale affects resolution. In a larger-scale map, the resolutions of features more closely match with real-world features because the extent of reduction from ground to map is less. As much map scale decrease, the map resolution diminishes because features must be smoothed and simplified, or not shown at all.

1.6 Map Accuracy

Many factors besides resolution, influence how accurately features can be depicted, including the quality of **source data**, **the map scale drafting skill** and the **width of lines** drawn on the ground.

2. Data Automation

Computer Aided Mapping has its limitations. Goal of GIS is not only to prepare a good map but also perform map analysis. Maps are the main source of data for GIS

2.1 Source Data

Source maps should meet the minimum mapping requirements as set forth by the Surveys and Mapping Division of Tanzania (SMD). If possible, only source maps which meet the SMD and those that are in good condition (preferably originals or archival copies) should be used. If a source map meeting these requirements cannot be obtained, it is recommended that source data be drafted onto base maps which meet these standards, or onto a transparent overlay which has been punch registered to an appropriate base map

For regional land based analysis, the scales at or near 1:10,000 shall be the primary base scale for input into the GIS database. Agencies may use smaller scales if deemed appropriate for the project and its intended analysis.

For local or municipal land based analysis, scales of 1:2,500 or less shall be the primary base scale for input into the GIS database.

Large scale source maps should be in satisfactory condition and have a minimum of 4 tics coordinate in order to achieve an accurate registration (Georefencing).

2.2 Data Capture

2.2.1 Scanning

Scanning of source data must be done at the high resolution of 300mp, and 24 bits; and should be saved in Tiff format

2.2.2 Registration (Georeferencing)

A minimum of 4 coordinate tics at the extremes of the data (i.e. map corners) should be used to register a map. The registration Root Mean Square Error (RMSE) value should be less than or equal to 002. All RMSE values must be recorded in appropriate documentation forms and shapefiles.

Where new coordinate tics are to be established, tic locations should be entered through keyboard entry of exact geographic coordinates (i.e. UTM or Longitude-Latitude coordinates). Addition of new tics should be evenly spaced. A tic match tolerance should be set in the tolerance file for coverage

2.2.3 Digitizing

A map manuscript should initially be digitized on screen

Features should be captured using a minimum number of coordinates needed to accurately represent the cartographic feature within a scale of 1:10,000 Points should be entered by geographic coordinates, if available. When manually digitizing points from a source map, one is required to enter the exact center of the point or map symbol.

2.3 Cartographic Accuracy

Positional accuracy should be as follows:

- 90% of the planimetric features on the digital map must be within 0.010 inch of the centerline of that feature on the source map when plotted at the original scale

- 100% of all features must be within 0.020 inch

The 0.010 inch interval is equivalent to a standard 0.010 plotter pen width. When a proof hardcopy plot of a digital map is overlaid on the original base map, discrepancies will be seen as an open space between the plotted feature and the original source map.

3. Attribute Coding

3.1 Item Definitions

The NBS GIS Database Design should be consulted for item definitions in a new thematic layer. If a new layer is not identified in the Database Design document, or if there are questions relating to item definitions, please contact Geo-Information staff.

Item definitions must be consistent from coverage to coverage within a layer.

Item names and codes should be unique from Layer to Layer

3.2 Attribute Integrity

Attribute values must be 100'% correct when verified against source map attribute values, or compared with a list of source attribute values.

3.3 Map Coordinates

3.3.1 Universal Transverse Mercator

Universal Transverse Mercator (UTM), Zone 35 to 37 souths (depending on the location) is the projection used in Country's library The map coordinate system of the Census Library is Universal Transverse Mercator (UTM). Units are meters. The datum is Arc 1960.

3.4 Edge matching

Features must be accurately edge matched to adjoining features.

3. 5. Spatial Topology

Data submitted to the Tanzania country wide GIS shall be topologically clean and free of error.

A topologically clean coverage will contain:

- No slivers
- No overshoots
- No undershoots
- No open polygons
- No unlabeled polygons
- No more than one label per polygon
- No unresolved line segment intersections

4.0. Testing for Accuracy

Transform an empty coverage to test positional accuracy of tics before digitizing features.

Positional accuracy of features must be tested with verification plots after automation. Attribute values should also be plotted with features at a legible scale and verified against the source map, or printed and checked against the digital database.

A final plot of the digital map, or a legible verification plot, as well as a printout of its attributes should be kept on file as reference.

The accuracy of attribute values must also be tested by comparing them to source data.

An accuracy rating for each layer must be provided in the layer's documentation, as well as a description of the rating method used. The type of accuracy test performed, a description of the method used, and the result of the test must be indicated in the Layer's documentation.

The specification of accuracy and the extent to which that accuracy is assessed is subjective. Several factors will determine how accurate the data must be and the rigor of the tests performed. How the information will be used, the consequences of inaccuracies and project constraints, such as costs and time limitations, can all play a role in evaluating the degree of testing.

The cost of accuracy assessment must be weighed against the benefits of the accuracy information. The assessment method used to test a layer's accuracy should be based on its intended use. Where consequences of error are less critical, verification plots and an assignment of an accuracy level (rating) may suffice without any further testing. For other layers, it may be important to report an accuracy measurement which indicates the likely error for an individual point and also includes a level of confidence.

4.1 Verification Plots

Positional accuracy of features must be tested by overlaying a plot of the digital data in inches with the source map from which it was digitized.

A proof hardcopy of the digital map must be plotted in inches prior to transforming the coverage into its final projection. When a proof plot is overlaid on the original base map, discrepancies will be seen as an open space between the plotted feature and the original source map. If more than 10 percent of the total features are found to be more than 0.010 inches from the centerline of the source map features, then they must be re-digitized. All features found to be more than 0.020 inch away from source map features must be corrected. A pen size of .25 mm or .010 inch must be used. A light table is recommended to facilitate verification.

4. 2 Normal Distribution

Where consequences of error are critical, a more rigorous test, such as a normal distribution model, should be performed to indicate probability of error.

The normal distribution mathematical model has been shown to be a good predictor of error distribution for digital spatial data. To model the normal distribution, the average positional errors

must be found for a set of test points. Test points should be easily verifiable (i.e. road intersections, corners of large buildings, etc.) and must be verified against an independent source of higher accuracy, such as field verifications. The variation of error values for those points must be determined using the standard deviation. The standard deviation is then incorporated into the normal distribution model.

To quote an accuracy measurement of a data set, a confidence level must be chosen. The higher the level of confidence is, the lower the accuracy will be. For example, at a confidence level of 80%, the predicted error of a data set might be 1.75 meters. But at 95%/0, the predicted error may be 2.15 meters. This last statement means that 95 percent of the time features will be within 2.15 meters of their true position and 5 percent of the time, features will be more than 2.15 meters away from their true position.

4. 3 Tic Coordinate Registration

If a transformation RMSE value is higher than .005, additional measures should be performed to lower the RMSE before proceeding.

Several steps can be performed to lower this value and/or increase positional accuracy:

i) Re-mount the source map and flatten it against the digitizing board. Secure it adequately with tape at the corners and at several other areas around the edge. Overlay the map with a piece of clear mylar or translucent paper. Re-register the map carefully.

ii) Additional tics, such as 2.5 minute interior tics, can be digitized but should be evenly spaced. Using 7 or 8 tics will reveal more information about which tics(s) are bad. A "bad" tic coordinate can be thrown out.

iii) If new coverage tics are needed, the coordinates should be entered by keyboard only.

iv) Tics should only be drafted onto source maps by those experienced with this technique.

5.0. Documentation

5.1 Lineage Report

The accuracy, processing history, and lineage of a layer must be recorded at the time of data capture, and must be available to other users in the form of a Lineage Report for evaluation of the data. This report will be attached to the database.

Existing data which were captured prior to establishment of these procedures for lineage reporting and data quality reporting must be identified as such. In such cases, lineage and quality estimates should only be made when there is some degree of certainty and confidence in this information. Estimated lineage items should be identified as "estimated". Unknown lineage and quality parameters must be listed as "unknown". Any database merging data from more than one source should identify those sources by providing a lineage reference on each one used.

5. 2. Layer Summary

Each layer of data collected shall have a Layer Summary file or attribute table attached to the database.

Each user who collects one or more databases will develop and maintain a layer summary form for each layer collected to be attached to the database. This form will describe or provide a reference for the classification, attributes, and attribute coding scheme for each database.

6.0. Census Map Library

The main objective of census mapping is to subdivide the country into small Standard Enumeration Areas (EAs), each of which will be assigned a unique identification code on the basis of the country's administrative divisions.

6.1. The demarcation of EAs should be such that they satisfy the following requirements:

a) Boundaries and starting points are clearly identifiable.

b) Boundaries of EAs should not cut across existing Region or District Boundaries. All social Boundaries should be respected.

c) EAs should not overlap with one another.

d) Each EA should be of a reasonable size (in number of household and land area).
Each census library has the geographic extent of a county and is partitioned, or tiled, by the county boundary. The map coordinate system of the Census Library is Universal Transverse Mercator (UTM). All the data are stored in zone 35 to 37. Units are meters. The datum is Arc1960.

6.2. Naming Conventions

Layer under development should be named to identify their data content and be easily recognized by the user.

Initial and processed coverage names should conform to standardized naming conventions recommended by ESRI.

Intermediate coverages that are generated during processing should be named so they are clearly identified as the most current coverage.

The first letter of a file must be an alpha term. It is recommended that coverage names be limited to six to eight characters, and not contain any extensions.

6.3 Data Output Specifications

6.3.1. Development of sets of EAs and SAs

- The sets should be in PDF format
- Sets should be sized to fit A3 paper
- Sets should be presented in a layout as provided in the templates

6.3.2. Printing of EAs and SAs

- Paper size should be A3 for all EAs and SAs and should be colour printed in two copies each.
- The paper to be used should be one suitable for map printing; preferably waterproof paper type.

6.3.3. Map Output

A plotted map must have a textual clause, or disclaimer, which states the original intent of the data, or warns about the map's usage. It must also include the plot date, the agency which composed the map, the data sources, and the latest revision of the database.

Other map elements which may be included optionally are:

Scale bar

Symbol key

North arrow

National emblem and/or NBS logo

Corner grid tics

Tic coordinates

Note: A LEGEND directory groups together with coverage such as scale bars, north arrows and the Organization logo are for general use.

National Bureau of Statistics

Director General	P.O. Box 796 Dar es Salaam	Telephone +255 22 2122724
General Office	P.O. Box 796 Dar es Salaam	Telephone +255 22 2122722/3
Fax:	+255 22 2130852	
E-mail:	dg@nbs.go.tz,	
	[Website www.nbs.go.tz]	

Vision

To be a preferable source of official statistics in Tanzania

Mission

To facilitate informed decision-making process, through provision of relevant, timely and reliable user-driven statistical information, coordinating statistical activities and promoting the adherence to statistical methodologies and standards