

SECRETARIAT OF THE PACIFIC COMMUNITY
STATISTICS FOR DEVELOPMENT PROGRAMME

UNFPA-SPC REGIONAL WORKSHOP REVIEWING THE 2010 ROUND OF POPULATION
AND HOUSING CENSUSES IN THE PACIFIC
(Noumea, New Caledonia, 21–25 May 2012)

Experience from the Pacific – Lessons learned for 2020

**Session 10 : CENSUS DOCUMENTATION AND MICRODATA DISSEMINATION IN
2010 ROUND**

(Document presented by Scott Pontifex, Database/Data Mining Specialist - Development Statistics,
Secretariat of the Pacific Community)

1. INTRODUCTION

This brief document describes the wider importance of documentation and metadata in the overall Census lifecycle, and examines benefits and barriers to making Pacific - Islands census micro-data more accessible to data users across the region and further afield.

2. BACKGROUND WHAT IS MICRODATA?

Microdata are defined as unit records or low level statistical data that have not been sufficiently aggregated to be made publicly available; they contain personal and confidential information about individuals and households contained in NSO census and household survey databases. However there is an increasing demand for access to Microdata by data users as it is an important input to evidenced-based decision making and policy development.

At the First Regional Meeting for Heads of Planning and Statistics in March 2005, planners and statisticians acknowledged that greater emphasis should be placed on data utilization, and making the dissemination of census and survey information an integral part of statistical planning. Furthermore the 2009 Statistical Benchmarking study identified that many Pacific Island NSOs in the region *lack centralized archives of data and metadata for storing of final datasets such as a modern data warehouse and noted that metadata was poor or did not exist.*

3. TYPES OF TECHNICAL ASSISTANCE AND TRAINING PROVIDED

Training comprised of a 2 week sub-regional training in the use of the International Household Survey Network (ISHN) Microdata Toolkit held in Noumea for two staff from most NSOs around the region: SPC has been involved in conducting two in-country training workshops in the Micro Management Toolkit in the Pacific Islands region (Fiji Bureau of Statistics and Vanuatu National Statistics Office). These workshops aimed to prepare documented datasets as part of an initial pilot project in the region. Other examples of training included UNSIAP organised training in the Toolkit as part of their fifth Group Training Course in Modules on Fundamental Official Statistics, 5 October 2009 – 19 March 2010.

The training provides an introduction of the ADP data documentation project and brief introduction to the micro-data toolkit and its various features, including components of survey documentation, dataset documentation and variable documentation. The metadata editor toolkit is used to prepare documentation for the census/survey and then archived onto CD using the CD-ROM builder application. A demonstration of the web application (NADA) was made to the participants. This application enables the documentation to be published on a web server for dissemination purposes (e.g. PRISM, NSO web-sites and intranet servers). In the Vanuatu training, the main operational focus was to provide hands-on application of the training documenting microdata from the NSO and other household surveys undertaken in-country (1999 Census, 2009 Pilot Census, 2006 HIES, 2006 Agricultural Census and the 2007 MICS survey). Documenting datasets is a time-consuming but worthwhile activity as it provides participants the opportunity to understand the need for documentation for recent surveys making any problems encountered easier to resolve.

Key objectives and activities of these training were:

- Introduce participants to the importance of good survey documentation, applying recognized international standards (DDI);
- Train participants in the use of the IHSN Microdata Management Toolkit (see: <http://ihsn.org/home/index.php?q=tools/toolkit>).
- For each participant to document, in the MMTToolkit, a recent Household survey undertaken in their respective country.
- Discuss the need for an official data access and protection policy for microdata sets;
- Archive documented census/survey datasets on NSO file server and CD-Rom;
- Document microdata from recent household surveys and censuses

In terms of tangible workshop outcomes and outputs, participants are expected to:

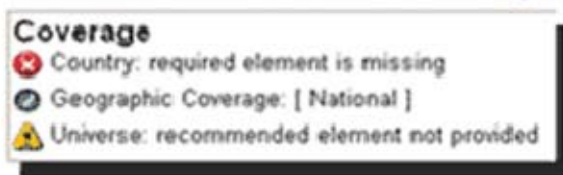
- know how to use the MMTToolkit and are familiar with the DDI;
- bring back a copy of the MMTToolkit to their colleagues for further use;
- bring back a fully documented survey, on DVD or CDROM;
- know where to get technical assistance on future survey documentation, should that be required.

- 1 With the **Metadata Editor**, import data from various standard format (SPSS, STATA, ASCII, others), and provide comprehensive metadata in user-friendly screens.



Data and metadata become one entity, saved in a single file. The free **Nesstar Explorer** program allows users to view metadata and re-export data to various common formats.

- 2 Generate various "Metadata Diagnostic" reports...



- 3 ... and automatically produce detailed survey documentation in PDF format.

- 4 With the **CD-ROM Builder**, produce a user-friendly html-output for sharing and preserving your data and metadata.



4. HIGHLIGHTS/ACHIEVEMENTS

A full inventory of past surveys and censuses, including gathering related data and metadata has been undertaken.

The development of an on-line catalog of Census and Surveys (NADA). The on-line catalog application has been installed and based in-country in Vanuatu and Samoa and a regional version is also hosted by the NSO website (PRISM). See SPC PRISM website:

<http://www.spc.int/prism/reports/data-catalog>

Documenting censuses (and survey) using the Microdata Management Toolkit, has taken place in some countries insuring a proper archiving of datasets.

The project provides training and technical support as well as the necessary software and hardware. Training and support has so far usually been provided at the sub-regional national level, to allow participation of a larger number of national agencies and staff. More concentrated in-country training has so far been undertaken in two Pacific Island Countries, and a *further series of sub-regional workshops in Fiji, Noumea and Pohnpei is being planned in collaboration with Paris21*, involving other countries yet to document their recently completed censuses and other household surveys.

The screenshot displays the Prism Pacific Data Catalog website. The header features the Prism logo and the text "Explore the Statistics of the Pacific Island Countries and Territories". On the right, it identifies the site as part of the "National Minimum Development Indicators Statistics for Development Programme" under the "Secretariat of the Pacific Community". A navigation bar includes links for "News", "Events & Activities", "Regional Data", "Reports", and "Feedback". The main content area is titled "Pacific Data Catalog" and contains a search interface with a "Filter by Country" dropdown, two search input fields (one for "study description" and one for "variable description"), and checkboxes for "Variable description includes" (Name, Label, Question, Classification). Below the search fields, it shows "Sort results by: Country | Year | Title" and "Found 11 studies out of 11". Two search results are visible: "Cook Islands - 2005-06 Household Expenditure Survey" and "Federated States of Micronesia FSM Household Income and Expenditure Survey 2005". Each result includes a brief description and links for "Browse metadata" and "Access policy". On the right side, there is a "Main Menu" sidebar with a list of site sections and a "MetaData Catalog (IHSN)" section with a brief announcement and the IHSN logo and URL.

5. DEVELOPMENT OF DRAFT DATA ACCESS AND DISCLOSURE POLICY FOR NATIONAL STATISTICS OFFICES

Pacific Island NSOs have the task to gather, manage and disseminate huge amounts of statistical information. Legislation and policies for Microdata dissemination are different from country to country and are often lacking or vague. For instance the Statistical Law or Act may either explicitly prohibit the dissemination of Microdata, or not mention the issue at all. In some cases even if the dissemination of Microdata is allowed the specific rules and regulations may be lacking. One of the main outputs from this project are discussions with Government Statisticians concerning the need for a data access and protection policy to be in place to ensure that the release of Microdata complies with Government regulations and legislation (Statistics Act), in particular protecting confidentiality and privacy of personal information. In Vanuatu a draft policy was prepared.

6. DATA DISSEMINATION

Data dissemination increases the use and potential impact of data, by:

- Making it possible for analytical work to be replicated.
- Creating the potential to use old data to test new ideas.
- Reducing the costs of data collection and the burden on respondents, by avoiding the need for researchers to undertake their own surveys.
- Demonstrating transparency and credibility in data production, which are at the heart of good governance.
- Improving the relevance and quality of data by incorporating user's feedback in future data collection.

7. TRAINING AND CAPACITY BUILDING IN DOCUMENTATION

Data documentation serves several important functions. It helps:

- data producers (NSOs) build institutional memory,
- researchers to help find the data they are interested in by making it easier to locate the datasets and variables that meet their research requirements;
- users understand what the data are measuring and how they data have been created, and assess their quality; and it
- gives background on the survey design and the methods used when collecting and processing the data, thereby reducing the risk that data will be misunderstood or misused.

In Vanuatu for example with Census of Population and Household occurring once every ten years a bonus of being able to document both the 1999 Census and 2008 Pilot Census was recognition that documentation needs to be part of the overall census or survey cycle and be done on an ongoing basis not ad-hoc or as a one off exercise, to ensure new NSO staff involved in census and household survey activities have the opportunity to familiarize themselves with past procedures and methods. Another outcome from the training workshops was the suggestion that documentation of Census and Survey requires input from all stakeholders involved at different stages of the data lifecycle which includes technical assistance providers, therefore becoming a requirement for external TA providers to complete sufficient documentation of the work being undertaken.

8. DATA PRESERVATION AND DEVELOPMENT OF A SURVEY CATALOG AND NATIONAL DATA ARCHIVE (NADA)

Statistical datasets are fragile and their long term availability to users is not always maintained. Micro-datasets can be damaged or lost because of human error, because of technical problems, or because of disasters such as fire, cyclones or floods.

New technologies can also render old data unreadable, because of either hardware or software advances. Datasets must be properly archived to ensure they will be available for future use. Participants from the training in Fiji and Vanuatu indicated a strong interest to incorporating their documentation of data sources in a central catalogue of Microdata and Metadata using the Microdata Toolkit. The use of NADA (a National Data Archive, user searchable- metadata repository) was seen as a tangible way to store all completed Microdata documentation to be published on a web server for dissemination purposes.

9. DISCUSSION POINTS

The meeting is invited to comment on:

- i. The usefulness of this initiative, and seeks a growing involvement and interest by NSOs to participate particularly in the area of training and capacity building in documentation and metadata and improved access to microdata.
- ii. The importance of data management, data preservation and need for accurate documentation as an ongoing activity by Pacific Island NSOs.
- iii. The development of a census/survey catalog and National Data Archive (NADA) to provide information on Pacific Island Census and Household Surveys, on conditions and procedures for accessing the Microdata, and to help disseminate documentation and Microdata.