Main Events

New Technologies in Census Data Processing

This article discusses some new techniques for the management and processing of census and survey data. The methodologies outlined combine the use of a high volume document image scanner and forms management software that can improve the speed of data capture by 75%, while ensuring a very high level of data accuracy based on the verification techniques implemented. This technique, by combining the roles of the coder/editor and the data entry clerk, significantly reduces the tasks associated with data capture. The software facilitates the logical digital storage and quick retrieval of census and/or survey forms and the security of these forms.

The census of 1991 in the Caribbean saw the introduction of scanning technology to improve the speed of processing and to ensure the availability of census results on a timely basis. In 1991, Caribbean countries utilised Optical Mark Reading (OMR) scanning technology for data capture. While this technology was an improvement over the manual data capture, allowing the generation of a clean database within a year and a half, there were some drawbacks. These included:

* the high sensitivity of the scanner to the quality of the paper used,
* the susceptibility of the scanner to dust and other environmental factors
* the quality of the shading,
* the fact that it is a high volume processing scanner which may lead to future under utilisation, and
* the forms had to be exclusively designed and printed by the company from which the scanner was purchased.

The Technological Upgrading of the Statistics Trade Information System –TradSys

During 2001 the Statistics Sub-programme obtained funding from the research and advisory budget of the Secretariat for the upgrade of its Trade Information System - TradSys. This system was originally created in 1995 to process detailed commodity level trade data received from Member Countries and comprised modules, written in dBase IV, which accepted the raw data and perform conversions, correlation and standardization as required to produce a regional trade database. The programme was designed to process data using the 1993 version of the CET; therefore, with the implementation of the 1996 version of the CET by member states, it became necessary to update the system. To make full use of the technology available at this time it was decided that the upgraded system would be based on an SQL platform using Microsoft Access as the user interface.

The upgraded TradSys Version 2.0 allows for greater flexibility in generating reports in response to data requests from users, an important aspect of the work done by the Statistics Sub-programme. By utilising an Access front end, additional queries and modules can be written by staff of the Statistics Sub-programme using Access or Visual Basic for Applications, thus allowing for ongoing enhancement of the system.

A further phase to the project will see the creation of queryable databases on the web, allowing users to actually input their query parameters and retrieving the results immediately. ■

Contents

New Technologies in Census Data Processing and Dissemination—cont'd

(Continued from page 1)
***The New Face of Data Capture*:**

In the 2000 round of censuses, the Saint Lucia Statistical Office pioneered the use of Intelligent Character Recognition (ICR) scanning technology to process the data. The diagram below outlines the process involved in data capture using the new modality of ICR technology as applied to the census in the Philippines.

Some of the advantages inherent in the application of the new technology are:
* the use of non-proprietary scanners which are faster and less sensitive to paper quality and other environmental conditions,
* the images scanned are stored logically in folders created during the process based on variables exported from the census questionnaire. PDF (Portable Document Format) files of scanned questionnaires are exported to folders on the ICR software server,
* the technology is easily adaptable to other censuses and surveys through internal development of forms as opposed to the exclusivity of the OMR process, and
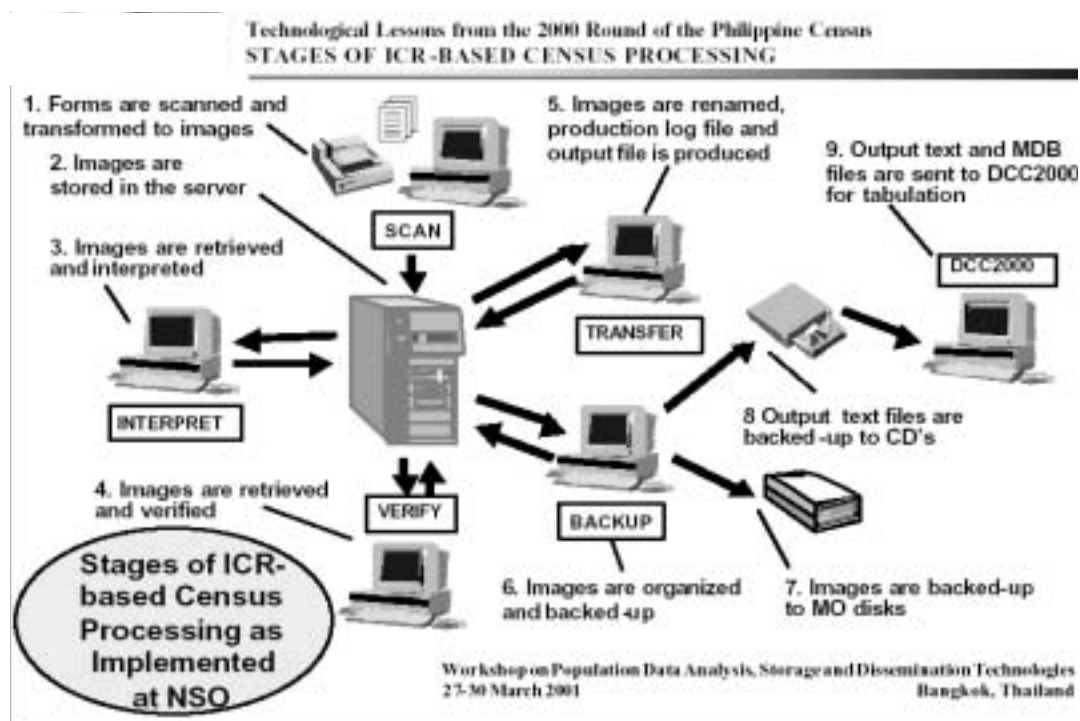* forms stored electronically can be retrieved speedily for examination.

By applying this technology in the Census of 2001, Saint Lucia reduced the processing time considerably, scanning the entire census in about four months, compared to one year for the 1991 census using the OMR technology. This is even more remarkable when one considers that the number of variables in the current census is twice as much as existed in 1991. Preliminary data for the Saint Lucia Statistics Office were therefore produced in record time.

Image scanning is also cost effective as can be justified by the amount of data entry, editing and coding cost which could be saved, as the software allows code tables to be built into the data entry screens presented to the verifiers. The software therefore enables the fusion of the function of the data entry clerk and editor/coder, resulting in a redefinition of these separate roles into a single one of "verifier". ICR software can also be used to:
* design forms for websites in HTML (Hypertext Markup Language), XML (Extensible Markup Language) or PDF format and to capture data from these forms,
* to capture data from faxed survey forms,
* to capture data emailed into the Statistics Department,
* to generate data from forms designed outside the software by novices,
* to develop interactive PDF forms for data capture via internet and email and
* to pre-fill forms sent out to establishments based on data captured from those establishments in the past, which allows data already captured to be verified by the respondent, along with the addition of new data

ICR software generally requires a LAN (local area network), Pentium computers, a high capacity image/database server, document scanners, and high capacity storage devices such as CD writers and magneto optical drives. ■

Technological Lessons from the 2000 Round of the Philippine Census
STAGES OF ICR-BASED CENSUS PROCESSING

1. Forms are scanned and transformed to images
2. Images are stored in the server
3. Images are retrieved and interpreted
SCAN
INTERPRET
4. Images are retrieved and verified
VERIFY
Stages of ICR-based Census Processing as Implemented at NSO
5. Images are renamed, production log file and output file is produced
TRANSFER
6. Images are organized and backed-up
BACKUP
7. Images are backed-up to MO disks
8 Output text files are backed-up to CD's
9. Output text and MDB files are sent to DCC2000 for tabulation
DCC2000
Workshop on Population Data Analysis, Storage and Dissemination Technologies
27-30 March 2001
Bangkok, Thailand

Saint Lucia Releases it's Preliminary Census Report

The Census Office of the Saint Lucia Statistics Department released the preliminary report on the Census of May 2001 at the end of March 2002. The report indicates that the preliminary count of Saint Lucia's enumerated population as at midnight on Census Day, the 22nd May 2001, amounted to 156,635 persons. This total consisted of 77,664 males and 78,971 females. Of the total enumerated population, there were 151,143 individuals residing in private households, 586 persons living in institutions and 4,906 guests in hotels or guesthouses. This report is available on the web site of the Saint Lucia Statistical Department at http://www.stats.gov.lc/cenpub.pdf. The census database will also be available for query through the department's web site by July 2002. ∎

International Trade in Services

The CARICOM Secretariat is undertaking developmental work in the area of services statistics to strengthen statistical capabilities in Member States. The Project, CARICOM Protocol II- Trade Policy and Facilitation Project is funded by the Canadian International Development Agency (CIDA) and originates in the Single Market and Economy programme. The statistical component is led by the Statistics sub-programme and the primary aim is the establishment of a foundation for the collection, compilation and dissemination of statistics on production and international trade in services.

Activities that have been completed include a situational assessment on the status of services statistics in the region, conduct of a regional workshop in Barbados and the conduct of a pilot national workshop in Grenada. The regional workshop was hosted by the government of Barbados through the Barbados Statistical Service and the Centre for International Services of the University of the West Indies (UWI). There were several other sponsoring organizations including the Caribbean Centre for Monetary Studies (CCMS) of UWI, the Central Bank of Barbados, the Caribbean Export Development Agency, Ernst and Young, and Price Waterhouse Coopers. These organizations provided additional funding, resource personnel, equipment and other services for the successful conduct of the workshop.

A second pilot workshop will be held in St. Kitts and Nevis in early May. An Advisory Group has been set up principally to enable the production of databases on services statistics and to document common guidelines for the production of these statistics. The issue of a common classification system for data collection is a priority in the development of these statistics. The activities of this project will form the basis for the further development of services statistics in the region. ∎

Twenty-Sixth Meeting of the Standing Committee of Caribbean Statisticians (SCCS)- Nassau, The Bahamas

The Twenty-Sixth SCCS Meeting took place in Nassau, The Bahamas during October 24-26, 2001. Among the issues discussed during this meeting were a status report on the Harmonisation of Statistics in the Region and on the development of a core data set, Classification Systems, developments in Merchandise Trade statistics and Trade in Services, the 1993 Systems of National Accounts (SNA 1993) and Social and Environment Statistics. The meeting also heard presentations on the status of the CARICOM Single Market and Economy and the Caribbean Court of Justice, the use of technology in the processing of statistical information and the management of statistical offices in the region. Representatives from organizations such as ILO, PAHO, CARTAC, UNSD and UNECLAC also presented information on projects and initiatives underway in the region.

The meeting endorsed the approach to the harmonisation of statistics in the region, and the need for statistical offices to document statistical methodologies to establish a core set of statistics for the region. The meeting also welcomed the CARICOM National Accounts Workshop (November 2001) and the developmental work in the area of Services and Social and Environment Statistics, appreciation being expressed to the UNSD for supporting both the activities in Social and Environment Statistics and in National Accounts. It was recognized however that an infusion of resources was required to upgrade the statistical offices across the region to enable the production of the statistics required by users and to support and monitor the impact of the CSME and CCJ. The Twenty-Seventh SCCS Meeting will be held in Grenada around the end of October, 2002. ∎

Workshop in Advanced Access

The CARICOM Secretariat once again participated in an Advanced Access Workshop conducted by the UNSD under the project "Strengthening Capacity in the Compilation of Statistics and Indicators for Conference Follow-up in the CARICOM Region". One member of the Statistics Sub-programme benefited from this workshop, which was held in Barbados during 19-29 March 2002, and included topics such as relational database concepts and design, SQL, Visual Basic for Applications (VBA), Active X Objects (ADO) and Component Object Model (COM). ∎

National Accounts Workshop

CARICOM Secretariat also conducted a workshop in National Accounts in November 2001, aimed at advancing the

Past Events and Outcomes, cont'd

implementation of the 1993 Systems of National Accounts in CARICOM member states, as well as the wider Caribbean region. The United Nations Statistical Division (UNSD) provided the services of the Inter-regional Advisor on Macro Accounting for Policy Analysis. The workshop was jointly hosted by the Central Bank of Trinidad and Tobago and the Central Statistical Office, Trinidad and Tobago with support from UNECLAC. The focus of the workshop was practical, demonstrating what countries were able to implement since the last workshop in 1999. Arising out of this workshop was the formation of an Advisory Group to provide support in the different areas of the 1993 SNA and to monitor its implementation. A work plan for future training activities was also designed and funding is being sought to activate this plan. The Secretariat would like to remind Member States to submit a status report of activities that have taken place since the November workshop. The next workshop is scheduled to take place in November 2002 and the Secretariat would also like to invite any of the countries inclusive of those in the wider Caribbean who are interested in hosting this workshop to communicate this information to the Secretariat. ■

Future Events and Expectations

Meetings Of Importance To Statisticians

* Pilot national workshop in services statistics – St Kitts and Nevis, May 2002.
* CARICOM Advisory Group Meeting on the establishment of a Programme for Social/Gender and Environment Statistics - Dominica, May 2002.
* CARICOM meeting to establish a core data set for the region, July 2002.
* Twenty Seventh SCCS Meeting, October 2002.
* Tenth RCCC meeting (Date to be advised).
* National Accounts Workshop, November 2002. ■

Publications

CARICOM's Selected Economic Indicators, 1985, 1995-1999 is available from the Documentation Centre of the Secretariat for US$25.00 on CD, and US$35.00 in hard copy.

You can expect to see these and other publications within the next six months!

* *Intra-regional Trade Report Vol. I*
* *External Debt of CARICOM Member States*
* *Women and Men in CARICOM Member States, 1980, 1990 and 2000 Rounds of Population Censuses*
* *International Trade in Services Vol. I* ■

Question and Answer Segment

In our first Question and Answer Segment we discussed the Concept of Value Added. The value added was measured by Gross Output less Intermediates in one approach. In this issue we highlight the question:

How is output defined, measured and valued?

Output is the total value of goods and services produced by an industry-establishment, enterprise or producing unit. It is defined initially relative to a producing unit and aggregated to obtain the total output of an industry and further of the economy relative to the production account. Output comprises those goods and services that are produced within an establishment and that become available for use outside that establishment.

Those goods and services produced by the processes of production including what is termed ancillary activities, and are used up within the same establishment during the accounting period by other processes, are not included as output. These goods do not leave the establishment and are therefore not made available for use outside of it.

Output is normally recorded when it is completed. However there are cases in which particular goods and services may take a long time to be produced spanning more than one accounting period. In these cases in order to recognize the output produced within a specific accounting period, the work-in-progress that has occurred within that period of accounting is used to measure output.

Work-in-progress consists of output that is not complete and therefore is not yet in a stage that it is readily marketable. Outputs may be used in several ways. They may be sold, bartered, placed in inventories, supplied to affiliate establishment owned by the same enterprise as intermediates, retained by producer as final consumption or capital formation and may be supplied free. In the next issue we will elaborate more on these uses, measurement and valuation of output. ■