

Date: January 1991

Classification: Population

Serial No.140

DATA PROCESSING SOFTWARE FOR POPULATION CENSUSES AND SURVEYS (1)

In future, the Population Programme of the South Pacific Commission will be issuing a quarterly *Information Circular* to inform member countries of current developments in computer software and hardware for population census and survey processing.

This issue focuses on data-processing software for data entry, editing and tabulation of population data. Future issues will provide information on data security, software utilities, demographic software, computer networks, and hardware problems and solutions.

A. DATA ENTRY AND EDITING

ENTRYPOINT 90

ENTRYPOINT 90 is a menu-driven program that allows use of a microcomputer for data entry. The program is both easy to use and powerful in its error-checking capabilities. Data entry operators can enter data quickly through the workstation module by following prompts and messages displayed on the screen. As well as providing user-defined screen layouts, the program allows data to be entered logically through sequenced screens and automatic skips. All data fields can be checked for validity, consistency and structure.

The developer module provides easy-to-use, flexible full-screen formats in which to create and update data entry and editing applications. The program uses a powerful logic language which allows users to customise their applications for either simple or complex data entry routines. The system has been designed to allow logic commands to be entered through menu-driven procedure panels. After data have been entered into batches and passed all edit checks, they can be exported to ASCII files ready for use with tabulation software.

The workstation module contains a number of useful customising features, including password control; changing menu displays to limit operator accessibility; changing background, field colours and highlighting; and keyboard mapping. This software also comes with an array of useful statistics on data entry operator performance, such as elapsed time, number of keystrokes, keystroke rate, number of errors, and number of records added or deleted.

A new version of ENTRYPOINT 90 (version 1.3) is now available. This update provides some new features, such as an image editor for easier screen design, hotkeys, keyboard mapping, and a number of other improvements. The new version is available in both English and French.

ENTRYPOINT 90 is available from:

DATALEX
100 Pine Street
Suite 1600
San Francisco
CA 94111,
United States of America

FAX: 415-362-5733
TEL: 415-962-8888

or

Commercial Computer Centre
182 George St.
Parramatta 2150
Australia

FAX: (61 2) 689-2434
TEL: (61 2) 635-4544.

The cost for a developer and workstation module is approximately US\$ 1795. Additional workstation modules are available at US\$ 725 per computer.

PC-EDIT

PC-EDIT is an easy-to-use, intelligent data entry and editing package. It is designed for demographers and statisticians with little or no programming experience, requiring only a basic understanding of file layouts and the principles of data entry. The program can be used for processing almost any type of questionnaire, generating either flat or hierarchical files. It is well-suited for processing censuses and vital statistics registration or surveys, as it can handle up to 64 different record types per file. Its data files are in ASCII format and can be used as input to other editing and tabulation software packages. PC-EDIT can edit data files created with other data entry packages.

PC-EDIT has a user-friendly interface, a highly structured menu system, context-sensitive help windows, and a full-screen editor presentation. Function keys are used to access different options from the main menu. In all, PC-EDIT is very easy to learn and use in a relatively short period of time. Its demonstration programs and ready-to-print user documentation are available in both English and French.

A new version of PC-EDIT (version 3.01) was released in August 1989.

To order PC-EDIT, write to:

Project Co-ordinator
Software and Support for Population
Data Processing Project
United Nations
Department of Technical Cooperation
for Development
Room DC2-1570
New York, N.Y. 10017
United States of America

Note: PC-EDIT is distributed free of charge to UNFPA-supported projects and institutions.

B. TABULATION

PC-COXTALLY

PC-COXTALLY is a simple system for producing easy, quick cross-tabulations, in unlimited numbers, of census and survey data, either on all the records or on a specific selection for each table. This is a menu-driven system that produces frequency tables up to seven dimensions, counting records or summing values for one or two variables. There are facilities for calculating percentages

and other simple functions. Special features include the ability to process hierarchical levels (e.g. geographical and administrative subdivisions) and to produce tables summed from the lowest hierarchical level. This software is designed for users who want quick results without fussing over table layout.

A complete new version of PC-COXTALLY, which includes new features and a direct interface with PC-EDIT, is being developed by the United Nations. Tutorials are available in both English and French.

Requests for PC-COXTALLY should be addressed to:

Project Co-ordinator
Software and Support for Population
Data Processing Project
United Nations
Department of Technical Cooperation
for Development
Room DC2-1570
New York, N.Y. 10017
United States of America

Note: PC-COXTALLY is distributed free of charge to UNFPA-supported projects and institutions.

REDATAM

REDATAM stands for REtrieval of census DATA for small Areas by Microcomputer. It is a user-friendly interactive software system, designed for convenient retrieval of small-area information required for decision-making, modelling and related purposes from large census files and similarly structured data. REDATAM can store compressed data of an entire population and housing census that may involve thousands of cases, or the data of a region or city.

REDATAM requires a microcomputer operating under MS-DOS, with at least 20 Mb of disk storage (depending upon the quantity of data to be processed). It has been used successfully with WORM (write-once-ready-many times) laser disks, but for most Pacific countries a hard disk will be sufficient. Record data can be tabulated, usually in seconds, for any variables and for any geographical areas, down to enumeration districts.

Planners and statisticians can work with REDATAM without the assistance of programmers once the database of interest has been created. This development is relatively recent in the data-processing world and provides, perhaps for the first time, an integrated approach to

data-processing which supports the needs of both planners and statisticians.

REDATAM was developed by CELADE (Latin American Demographic Centre) with assistance from UNFPA (United Nations Population Fund) and IDRC (International Research Development Centre, Canada). It is presently being used in Central and South American countries for the 1990 round of censuses.

REDATAM is distributed free of cost from:

CELADE
Centro Latinoamericano de Demografia
Edificio Naciones Unidas, Casilla 91
Santiago,
Chile.

C. INTEGRATED SOFTWARE

IMPS

The U.S. Bureau of the Census developed the Integrated Microcomputer Processing System (IMPS) in the early 1980s. IMPS is a modular microcomputer based software system for data entry, editing, tabulation, and analysis of census and survey data. The IMPS system comes with an easy-to-use menu-driven interface. The system is designed around a common data dictionary (DATADICT) which allows users to name fields on the data record. The data dictionary is accessed by each of the three main modules: CENTRY (data entry), CONCOR (editing), and CENTS (tabulation).

CENTRY stands for CENSus data enTRY and is a recent development in the IMPS system. It is a menu-based system which allows users to design screen and field layouts simply. Its features include ability to process data in batches, direct retrieval of data item definitions from data dictionary, development of either single screen or multiple screen layouts, range-checking, automatic field duplication and skips, field entry sequencing, ASCII file output, protection against power failure, and operator statistics. A new version has been released which allows users to perform consistency edits during data entry as well as range checks.

CONCOR stands for CONSistency and CORrection. It was originally developed for mainframe computers, but now has been made available to microcomputers through a MS-DOS COBOL compiler. CONCOR is an integrated system of programs which can identify and change inva-

lid and inconsistent data being prepared for tabulation and analysis. Its main application has been in the editing of population and housing data from national censuses, but it has also been used for agricultural censuses, and labour force and education surveys. CONCOR is a very flexible system and can perform even the most complex consistency checks, but it does require some programming experience to develop an application.

CENTS is a CENSus Tabulation System which comprises a number of programs to tabulate, summarise, and display statistical tables. CENTS was developed to meet the need for tabulation software that would produce highly complex, good quality tables for population censuses. The user is provided with a great deal of flexibility in designing tables. For example, the placement of text, such as headings, stubs, captions and foot notes, is under the total control of the user. Another useful feature is its ability to display the table layout without needing to process the data. This allows for easier layout control. CENTS is fast and flexible, but, as for CONCOR, users require some programming experience.

As well as the above modules, IMPS includes CENTRACK, an interactive management and control system for census managers, and QUICK-TAB, a module which produces frequency distributions and two-way cross-tabulations quickly and easily.

The IMPS software package and manuals can be obtained free from:

Chief, Computer Software Branch
ISPC, Bureau of the Census
Washington, D.C. 20233
United States of America

Fax: (301) 763-7589

Note: In order to run the CONCOR and CENTS modules, a COBOL compiler will need to be purchased through a computer software dealer, at a cost of approximately US\$ 800. REALIA COBOL can be obtained from:

REALIA Inc.
10 South Riverside Plaza
Chicago
ILL 60606
United States of America

FAX: 312-346-4638
TEL: 312-346-0642.

D. INTEGRATED SOFTWARE

ISSA

The Integrated System for Survey Analysis (ISSA) is a microcomputer software package developed for all data processing tasks required for a census or survey. It has been used to process large-scale surveys in over 30 countries and is already used by some Pacific Island countries. It was developed for the Demographic and Health Surveys Program by the Institute for Resource Development/Westinghouse, USA.

ISSA uses a data dictionary to define a data file which can consist of up to four hierarchical levels. Data entry screens, data value ranges, variable and value labels are all included in the dictionary. The dictionary is developed through a menu-driven interface which also includes a form editor for designing data entry screens. Combined with applications, the dictionary provides the user with the option of interactive data entry or batch activities, such as data verification, tabulation and secondary data file construction. For complex routines, there is a logic language provided which can access up to five data files simultaneously.

The data entry component provides the user with a tool that can ensure the creation of a structurally correct data file with no range, skip or consistency errors. The tabulation component can access either data entered through the data entry component or any ASCII file(s).

ISSA can produce up to three-way cross tabulations, means and medians. Secondary analysis of tabulations can also be carried out.

ISSA can be used by either computer novices or experienced programmers. For those activities which are simple, such as creating data entry forms and simple tabulations, ISSA provides a powerful programming language for experienced programmers to create their own programs.

ISSA can be obtained for approx US\$325 from:

Institute for Resource Development/
Westinghouse
8850 Stanford Boulevard, Suite 4000
Columbia, MD 21045
United States of America

Sources

Busby L & Boxill B. (1990) *REDATAM: a database management system for Latin America and the Caribbean*.

Cushing J. (1987). The Integrated System for Survey Analysis. *POPLAC newsletter*. Chicago.

Datalex (1990) *ENTRYPOINT 90: Training and support services*.

Economic and Social Commission for Asia and the Pacific. (1988). *Statistical software packages for processing censuses and surveys*. New York.

Institute for Resource Development. (1989). *The Integrated System for Survey Analysis*.

United Nations (1990). *United Nations Population Software Notes*, Issue 3, Spring 1990.

U.S. Bureau of the Census. (1990) *DATALINE: US Census Bureau International Bulletin*. May 1989.

©Copyright South Pacific Commission, 1990.

The South Pacific Commission authorises the reproduction of this material, whole or in part, in any form, provided appropriate acknowledgement is given.

Original text: English