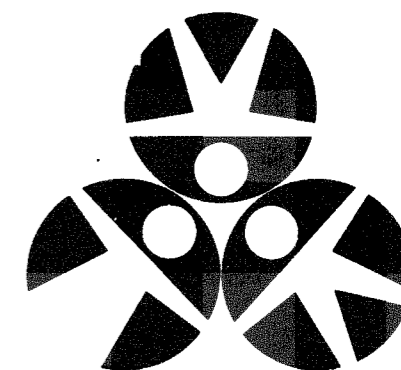


ACSPRI newsletter



Australian Consortium
for Social and
Political Research
Incorporated

Number 3 March 1980

ISSN 0158-6882

Male Australian citizens born at sea, characters 1473-1476 of block 1, would be in characters 33-36 of the 19th 80-character record (1473 = (18x80) + 33).

The number of dwellings in which coal, coke or briquettes was used for cooking is written as characters 2541-2544 in block 2, or characters 61-64 in the 32nd 80-character record from that block (2541 = (31x80) + 61).

Although this seems tedious, it may well be easier than an approach which considers the data to consist of variable length records. Also once a decision has been made about the data to be used from each CD, these would be read from the CD tapes with, for example, a FORTRAN program, and written to a tape or disk file which would then be used for subsequent analyses.

Bill Craig, ANU Survey Research Centre.

* * *

CONTENTS

A Social Science Data Archive in Australia	1
Contributions to the Newsletter	1
Guidelines for Depositing Data with ACSPRI	2
Information Sources on Australian Data Collections	3
Summer Schools and Conferences	4
Overseas Visitors	5
Agreements with Overseas Archives	6
Using CD Summary File data from the 1976 Australian Census	7

A SOCIAL SCIENCE DATA ARCHIVE IN AUSTRALIA

Since the foundation of ACSPRI in 1976 the ANU Survey Research Centre has acted as the agent for ACSPRI in the servicing of requests for data held by ICPSR and other overseas archives, and in storing, maintaining and distributing any Australian data sets deposited with ACSPRI. Late in 1979, a Committee was appointed by the Vice-Chancellor to review the Centre. As a result of this review, it was recommended that:

- (1) the Survey Research Centre should cease to exist as an independent Centre within the University;

- (2) the majority of the resources currently allocated to the Centre should be transferred to the Research School of Social Sciences to establish and maintain a 'Data Archive'. The Data Archive would continue the secretarial function for the Australian Consortium for Social and Political Research Inc. and would provide, throughout the University, advice and assistance for surveys. It would not however conduct surveys itself;

- (3) an Advisory Committee for the Data Archive should be established to oversee its activities.

Standing Committee of Council accepted these recommendations on 12 December 1980. The final arrangements for the establishment of the Data Archive within the Research School of Social Sciences are currently being made, and the details of its development program are being considered. These will be presented in the next issue of the Newsletter.

CONTRIBUTIONS TO THE NEWSLETTER

Contributions from readers are encouraged and reports on the use of data supplied through ACSPRI would be of particular interest. Contributions and enquiries about ACSPRI should be addressed to:

Roger Jones
The ANU Survey Research Centre,
Australian National University,
PO Box 4, Canberra 2600
Telephone (062) 49 4400

GUIDELINES FOR DEPOSITING DATA WITH ACSPRI

Machine-readable research data represent a valuable resource in terms of both human effort and cash funds, and ought therefore to be protected and utilised fully. Depositing a data set with an archive is insurance against loss or damage of the data. In addition, depositing data in an archive is of direct benefit to the depositor. The archive will usually check and clean the data and improve the documentation, and the researcher will have access to these improvements. Publication of the availability of the data may assist in developing contacts with researchers working in the same area. The researcher also has the satisfaction of knowing that his/her work can be built upon by others, and that students can use the data to learn methods of analysis.

ACSPRI welcomes copies of all machine-readable research data that are

- about Australia or part of it.
- collected by Australian investigators, whether pertaining to Australia or not
- likely to be of interest to Australian users.

ACSPRI will accept all data in machine-readable form (preferably card or magnetic tape) and requires in addition only accurate accompanying documentation. No requirements are imposed about coding, punching conventions or the format of the tape.

Depositors may impose an embargo period on the use of his/her data, although this must be limited, normally to a maximum of two years. During this time the data will be advertised but will only be made available with the written permission of the depositor. Alternatively, the depositor can choose to be kept informed of the use being made of the data and will be provided with details of requests and copies of any publications resulting from work with the data. The final form of

deposit is where the depositor does not wish to monitor access or be informed of instances of use.

Data deposited with ACSPRI will be provided to academic researchers in member institutions or overseas at cost, subject to the conditions specified by the depositor and after the requestor has signed a "User's Undertaking". In this, the user agrees to preserve the confidentiality of survey respondents and their replies, to acknowledge both the original depositors and ACSPRI in any work based on the data, to refrain from redistributing the data to others and to agree, when required, to obtain written permission for publication of any interpretation of the data.

Depositing a data set should involve the researcher in very little additional work, particularly if it is done while the methodology and document preparation are still fresh in memory. The basic requirement is that the accompanying documentation should enable a stranger to make sense of the data. Thus the materials required are:

- (i) a copy on cards or magnetic tape of the raw data file (not SPSS system file) and where available an SPSS set-up deck. Where magnetic tape is used, a Magnetic Tape Description Form should be completed.
- (ii) a copy of the coding frame or codebook
- (iii) a clean questionnaire
- (iv) a copy of any written documents (such as interviewers and coders instructions, details of the sampling frame) and publications
- (v) a completed survey description form

Once these materials have been received by ACSPRI, a limited amount of checking will be undertaken on the data set (e.g. the correct number of cards for each case, the range of codings, etc.) and any queries will be notified back to the depositor. The survey description

[2]

provided by the depositor will be used to publicise the availability of the data and a 'neat and tidy' codebook will be prepared for distribution with the data file. A decision as to whether further checking and correction of the data is intended will be made and notified to the depositor. Since errors discovered in these processes may only be rectified by referral back to the original collection instruments and documents, researchers are urged to retain these materials until that time.

* * *

INFORMATION SOURCES ON AUSTRALIAN DATA COLLECTIONS

Information on the data collected by government bodies and academic researchers has been relatively scarce in Australia and one of the primary objectives of ACSPRI is to collect and disseminate such information to interested researchers. Some recent developments have considerably improved access to this type of information.

1. Government Data Collections

In 1978, the Secretariat Section, Australian Bureau of Statistics, began compilation of a catalogue of 'Statistical Data Collections undertaken by Commonwealth Government Departments and Authorities'. A second survey, covering collections undertaken wholly or in part during the twelve months to 30 June 1980, has also been completed and results will be available shortly. The first survey covers 1978 collections, and contains only the title and type of collection, the collection authority, publications, and whether or not further information is available from the collection. For the second survey, additional details of the method and source of collection, number of units, frequency, and geographical coverage were also obtained, and the name and telephone number of the officer to be contacted for further information will be published.

In New South Wales, a 'Directory of Statistical Collections, 1980' compiled by the State Statistical Coordination

Unit is on sale at the Government Information Service. Collections are organised under subject areas with entries covering the topic, area, frequency, time period covered, availability, and the contact officer. An index of the Departments covered and a subject index are included.

The Queensland State Statistics Coordinating Committee has recently released 'Statistical Collections of State Government Departments and Authorities' which is based on results of a survey undertaken between June 1977 and November 1979. Key words are used to describe data collections and an index is included. Entries contain the title, frequency and size of the collection, a brief description of the content, the method of compilation, availability, medium of distribution, and a contact officer.

The State Statistical Priorities Committee, South Australia, has developed a 'Survey Register' based on a survey of State Government Departments in 1978. Entries are organised by key words and include the responsible department, title, coverage and date of the survey, and a contact officer. An updated, improved survey is planned for July 1981 in respect of the 1980-81 financial year.

In Western Australia, the State Statistical Requirements and Coordination Committee is preparing a booklet on data collections of State Government Departments and Authorities which should be available in mid 1981. This booklet will include data collections and research publications with a significant statistical base produced since January 1979.

Work on a Directory of Data Sources for Tasmania was unfortunately abandoned and details of the work carried out in Victoria have not yet been received. Nevertheless it is clear that the Commonwealth and the States have recognised the importance of improving awareness of the statistical collections they undertake and the potential value of these data as a resource for administration, business and research.

[3]

2. Academic Data Collections

The ANU Survey Research Centre's publication 'Australian Social Surveys: Journal Extracts 1974-78' is based on a systematic search of thirty Australian social science journals published between January 1974 and December 1978 for articles reporting the use of survey data. Entries are organised under subject headings and include the author's name and address, the title and journal reference for the article and, where available from the article, the geographical coverage, date of fieldwork, time covered, population sampled and sample size of the survey. A keyword-in-context index based on the titles of the journal articles, and an author index are also included.

A second project conducted by the ANU Survey Research Centre, the 'Inventory of Australian Surveys', has involved the collection of more detailed information about Australian surveys. The aims of this project were to assess the extent of survey activity by academic staff and post-graduate students, to obtain detailed methodological information about each survey located, and to locate survey data potentially suitable for secondary analysis. The information collection began as a survey of heads of social science departments in universities and colleges of advanced education requesting names and contact addresses of researchers who had conducted surveys from that department since 1970. Individual researchers were then contacted by mail and requested to give a more detailed description of their work on an inventory questionnaire. The survey resulted in 571 survey descriptions, and further entries have been received and added to bring the present total to 760 entries.

Most of these entries refer to surveys conducted between early 1970 and 1978. Entries are organised under subject headings (as in the Journal Extracts) and include the name and address of the investigator(s), the survey title, sponsors, details of the fieldwork and sampling procedures, a summary of results, list of publications, and the status of the data. Supporting indices

include names of principal investigators and a keyword-in-context index based on the survey titles.

A computer printout of the full list of entries, comprising over 700 pages, is available for reference at the Survey Research Centre and will be supplied to interested researchers or institutions at the cost of reproduction, about \$20, plus postage. An abbreviated listing of survey titles listed by investigator within subject areas can be provided for \$2 plus postage, and the keywords-in-context index for \$4 plus postage. For further details, contact Roger Jones or Jane Muford at the ANU Survey Research Centre (Tel. 062 494400).

* * *

SUMMER SCHOOLS AND CONFERENCES

ICPSR Summer Training Program / Survey Research Centre Summer Institute

These training programs conducted each year at the University of Michigan were described in the March 1980 issue of the Newsletter. This year, the first session runs from 29 June to 24 July and the second session from 27 July to 21 August. A limited amount of financial assistance is available to participants from ACSPRI member institutions. Applications from non-member institutions are also welcome. For further details contact Roger Jones at the ANU Survey Research Centre (Tel. 062 494400).

Fourteenth Essex Summer School in Social Science Data Analysis and Collection

The Fourteenth Essex Summer School will take place between 10 July and 21 August 1981. The school will consist of three two-weekly sessions in each of which a variety of courses will be offered, with the more introductory courses in the first session and the more advanced courses in the final session. Each course lasts for two weeks and consists of a two-hour class each weekday. Participants will normally only be allowed to take one course per session. The majority of courses deal with techniques of data analysis and model building. A course in Survey Design and Analysis

[4]

will run independently but parallel to the Summer School in the first two sessions (10 July - 7 August) and a special two week course covering problems and techniques of data management will be held from 27 July to 7 August. Interested persons should write to:

The Organising Secretary, 14th Essex Summer School, Department of Government, University of Essex, Colchester CO4 3SQ, England.

IFDO/IASSIST 1981

The International Federation of Data Organisations (IFDO) and International Association for Social Science Information Service and Technology (IASSIST) will hold a joint conference on 'The Impact of Computerization on Social Research: Data Bases and Technological Development' at the Universite des Sciences Sociales, Grenoble, France, on 14 - 18 September 1981. Conference themes are: new types of research, new institutions, new tools, relations between data producers and users, and social science teaching. One day of the conference will be devoted to three concurrent workshops on the topics 'The assay/evaluation of survey, ecological, satellite and cartographic data', 'The organisation and management of data services', and 'The formation of a professional data archivist and librarian'. For further details, write to M. Frederic Bon, C.E.R.A.T., Institut d'Etudes Politiques, BP34, 38 401 St Martin d'Heres Cedex, France.

The American University - 1981 Institute

The topic for this year's institute, to be held on 10 - 11 June, is 'Federal Databases: Identification, Evaluation and Access'. The institute will address such questions as: What databases are available, and in what format? How can I evaluate the quality and usefulness of federal databases? What are the means of access? What is the outlook for improvements in accessibility and in the technology of access? The fee for the institute is US\$190. For further information, contact Prof. Lowell H. Hattery, Center for Technology and

Administration, The American University, Washington D.C. 20016.

* * *

OVERSEAS VISITORS

Professor Charles F. Cannell, Research Scientist, Survey Research Center, University of Michigan, USA.

The Australian-American Educational Foundation has approved the appointment of Professor Charles F. Cannell as a Short-Term Senior Scholar in response to a proposal put forward jointly by the Sample Survey Centre, University of Sydney, and the Centre for Applied Social and Survey Research (CASSR), Flinders University of South Australia. He will be in Australia during October and November, 1981.

Professor Cannell is best known for his extensive research and writing in the area of survey interviewing techniques, and his books, The Dynamics of Interviewing: Experiments in Interviewing Techniques, and A Technique for Evaluating Interviewer Performance.

During his six weeks stay in Australia, Professor Cannell will be based in the Sample Survey Centre and CASSR where he will be discussion leader in a program of seminars and workshops on interviewing techniques. He will also visit the Australian National University and other universities for meetings with survey researchers. Details of Professor Cannell's program will be circulated to members of ACSPRI in coming months.

- - -

Professor D. Holt, Professor of Social Statistics, University of Southampton, U.K.

Professor Holt will be visiting Australia for a short period during July, 1981, as a result of initiatives being taken by the Sample Survey Centre, University of Sydney. His visit will be sponsored by the SSC, other universities, and by ACSPRI. Seminars are being planned for Sydney, Canberra, and Melbourne, where Professor Holt should

[5]

be accessible to ACSPRI members with an interest in his research.

Likely topics for Professor Holt's meetings are:

1. Models for Regression Analysis Using Survey Data
2. Imputation for Non-Response
3. Logistic Models for Survey Data

Details will be circulated to ACSPRI members when Professor Holt's itinerary has been set.

- - -

Further details of these visits can be obtained from Dr T. Beed, Sample Survey Centre, University of Sydney.

* * *

AGREEMENTS WITH OVERSEAS ARCHIVES

Data Program and Library Service (DPLS)

ACSPRI has further extended its links with overseas archives by establishing contacts with the Data and Program Library Service (DPLS), University of Wisconsin-Madison. DPLS acquires social science data files generated or obtained by researchers at the University. While its principal mandate is as a local campus service, DPLS also accepts suitable data donated by researchers and institutions elsewhere and assumes responsibility for publication of the availability of the data file through journals concerned with such data.

Following requests from Australian researchers, ACSPRI has recently acquired three data files from DPLS and obtained agreement to distribute these data files to other interested researchers in Australia. Descriptions of these three data sets and the costs of acquiring them are given below.

- - -

Charles F. Westoff et al.

American Family Growth 1957-1967

This longitudinal study examines the fertility history of white, urban, native-born American couples with two children whose second child was born during September 1956. Data were gathered during three interviews conducted in 1957, 1960, and between 1963 and 1967. Phase I looks at the social and psychological factors thought to relate to differences in fertility. Phase II focuses on why some couples stopped at two children while others had a third or fourth child. Phase III examines how well attitudes and events of the early marriage determined the record of the later years of childbearing. The initial sample contained 1165 couples, of which 905 were successfully interviewed in the second phase and 814 in the third phase. The data file and accompanying documentation is available from ACSPRI for \$90 plus the tape cost.

- - -

Charles F. Westoff and Norman B. Ryder

National Fertility Study 1975

The study is based on data collected from a national (USA) probability sample of continuously married women who were white, married less than 25 years and before the age of 25, and whose husbands were also once-married. For marriages of more than five years, 2361 women selected from the 1970 National Fertility Study sample were reinterviewed. A proportionate sample of 1042 new respondents was added to represent marriages that occurred after 1970. Detailed data are available on contraceptive efficacy, fecundability, work history, family planning intentions, sterilisation, opinions on abortion, current population problems, attitudes on individual efficacy, and detailed background information on self and spouse. The data and accompanying documentation are available from ACSPRI for \$90 plus the tape cost.

- - -

Charles F. Westoff and Norman B. Ryder

National Fertility Study 1970

The study is based on data collected from a national probability sample of ever-married women born since July 1, 1925. Completed interviews were obtained from 6752 women, a response rate of 80%. The information collected was similar to that in the 1975 study. The data and accompanying documentation is available from ACSPRI for \$75 plus the tape cost.

* * *

USING CD SUMMARY FILE DATA FROM THE 1976 AUSTRALIAN CENSUS

The following is a suggestion to those who would be involved in the use of CD summary file data at the data processing level, i.e. people who would specify to a computer the way in which CD data is to be read.

The Australian Bureau of Statistics publication CD Summary Files Final Data Technical Details describes the summary data as consisting of variable length records (p3). However, each record of a given type is of the same length, and more importantly, the position of each data item (total) within a data block is fixed. This enables a user of these data to consider them as consisting of fixed length records, thus avoiding the difficulties inherent in processing records of variable length.

Data for each CD is written as two consecutive blocks of 3200 ASCII characters each. The first block of the pair contains records of type 0, 1, 2, and 3, and the second contains records of types 4-9. Each record consists of four characters which give its length, two characters specifying the record type, two filler characters, and a number of data characters. The length of the record includes the first eight characters. Record lengths are given on page 3 of the Technical Details. The lengths of the data items in record type 0 are given on page 5, and each total in extracted from that block.

record types 1-9 is written as 4 characters.

The approach suggested here consists of ignoring the record structure specified for each block, and considering only the position of a data item within the block. In effect, each block is considered to be one record, and each CD comprises 2 records.

The data item 'state' is written as character 9 in record type 0, which means that it is the 9th character in the first of the pair of blocks which comprise the data for each CD.

For each CD, the number of male Australian citizens recorded as having been born at sea is written as characters 1473-1476 of record type 3 (Tech. Details p26). This means that this item is in characters 3033-3036 of the first block of data for each CD. The calculation involves adding the lengths of the preceding records in the block, record types 0, 1, and 2, and the starting character position of the desired total, i.e. $80 + 488 + 568 + 1473 = 3033$. Since the total consists of four characters, it ends at character 3036.

The number of dwellings in each CD which mainly use coal, coke or briquettes for cooking is written as characters 9-12 in record type 9 (Tech. Details p32), or in characters 2541-2544 of block 2 for each CD. 2541 is the sum of the lengths of records 4-8, and the position of the required total in record 9, i.e. $592 + 444 + 576 + 512 + 408 + 9$.

If your computing system constrains record lengths to be less than the length of the blocks in the CD data, it would be possible to split each block into a number of smaller records, and then select the required fields with calculations similar to those described above. Each 3200 character block could be split into 40 80-character records, for example, and the previously mentioned data items selected as follows:

'state', the 9th character in the first block of each pair, would be in character 9 of the first 80-character record