Chapter 7

COMPUTERS

Computers provide a way to process research data quickly and economically. In this section you will find information on different uses of computers, how to prepare data to be processed by computers, and a guide to the different choices one faces in using the computer. Sample printouts are given to illustrate the capabilities of available computer program packages. Sometimes researchers avoid using computers because of the difficulty presented by technical terms and equipment. The explanations that follow attempt to take the "mystery" out of computers by defining computer-related terms and describing the equipment available.

The computer can be a useful tool for community-based research for its ability to process data and generate reports. However, you can either save or waste time and money, depending on the wisdom of your decisions concerning its use. This chapter outlines some of the factors to consider in choosing methods of processing data and specific applications of computer technology.

Information collected systematically by such means as survey interviews, needs assessments, questionnaires, and record keeping systems is called **data**. The term **data processing** refers to the operations performed on data, usually by means of automatic equipment. Today, electronic equipment for the processing of data can vary from an inexpensive calculator to a high speed computer system. The project collecting a hundred or less data items (to be analyzed using descriptive statistical techniques) might well use a programmable calculator for satisfactory and cost effective results. Such calculators are available with ready-to-run statistical program for under $200 (1981). Calculators can satisfy a multitude of office and accounting needs, thus maximizing the use of program funds. In making the decision to use a computer, you need to consider the things computers can and can't do.

Computers Can

* Process large quantities of information very quickly
* Perform statistical operations with great accuracy
* Process large quantities of data inexpensively, relative to human labor
* Retrieve stored information in a selective way
* Print reports according to a specified format
* Print multiple copies of reports at low cost

Computers Can Not

* Think
* Make decisions, other than those based on the numerical calculations programmed
* Draw conclusions
* Process data without accurate, very specific instructions
* Process information with human error

It is a common practice today to "blame the computer" when things go wrong with paperwork, telephones, traffic lights, and other automated things in our lives. The computer, however, is only a machine that performs exactly according to instructions. The errors in computerized tasks are, with very rare exceptions, caused by human mistakes in instructions to them. An old saying among computer users that describes this is, "garbage in , garbage out." Before undertaking any work with a computer, the user should be prepared to accept sole responsibility for errors, and possibly some great frustration in the learning process. Once the

basic concepts of computer use and the specific details of a data processing application are learned, the computer can serve as a very efficient and cost effective means of data analysis.

DATA PREPARATION

The interpretation of information from an idea to coded data can take many forms. One of the common fears held about using a computer concerns the taking away of the "human element" in the process. Keeping the human element in research depends largely upon the quality and thoroughness of the choices made concerning the representation of data and the interpretation of results. Planning for the coding of data (representing data according to a set format) and planning for the later analysis (examination of the relationship between the data items) should be completed before the data collection begins. Otherwise, the researcher may, more than likely, find needed information missing or incomplete when the time for processing is at hand. Presented below is a discussion of different types of data and some examples of data preparation for automated processing by calculator or computer.

In designing a research topic, a set of questions is developed. This set of questions is then answered for each unit included in the study. That basic unit, or case, can be either a single respondent in a group survey, or a larger unit such as one of several communities being compared. Each of these questions answered or measurements taken within each unit in the study is called a **variable**. In other words, the responses or measurements can vary from unit to unit, but for each case within the unit there will be recorded one value for each of the variables. Two examples of variables that might be included in a survey are "age" and "sex." The value of "age" for a given respondent would be recorded in years; whereas, the value of the variable "sex" would be either male or female. Once values have been recorded for the variables, the cases can be compared to determine relationships according to different procedures. Several of these procedures were discussed in the preceding chapter on statistics. The use of the calculator or a computer can speed these procedures, but the process of determining the questions to be asked (that is, defining the variables), and recording the values (according to some pre-defined, consistent plan) must occur whether or not a computer is used. There is **no computer magic**, and the human decision-making that takes place before any data are collected will largely determine the quality of the outcome.

Although the computer can read alphabetic characters and symbols, as well as numbers, the data are usually represented as numbers for the purposes of calculations. This does not limit the types of data that can be processed by means of a computer to characteristics that can be measured quantitatively. To summarize some of the concepts introduced in Chapter 6, STATISTICS, data which are measurable or do follow a consistent

progression are described as **quantitative**. Examples of quantitative variables are "age," "highest grade level completed," and "income." Data which do not follow a measurable, orderly progression, but are more abstract or unordered are described as **qualitative**. Examples of qualitative variables are "sex," "area of residence," and "marital status." Assigning values to the possible responses to qualitative variables requires careful thought and planning. In other words, the different qualities that may be answers to a particular question must be thought of ahead of time and included during the design of the questionnaire, interview form, or other instrument, if all possible responses are to be indicated. Categorizing and assigning numbers to the values is a step in preparing the setup for data processing.

For the purposes of creating a step-by-step coding example, a few questions will be chosen from an imaginary questionnaire titled, "Urban Adjustment Survey." The variables are "Respondent Number," "Sex," "Age," "Marital Status," "Tribe," "Blood Quantum," and "Frequency Speak Native Language." The data might be collected according to the format below.

Figure 7.1   Example of a Questionnaire Page

URBAN ADJUSTMENT SURVEY

<u>OO3</u> Respondent   <u>25</u> Age   <u>1</u> Sex   <u>172</u> Tribe <u>Sioux</u>
1-3   4-5   6   7-9

                              1) Female
<u>3</u> Blood Quantum   2) Male
10
    1) Less than ¼
    2) ¼ to ½   <u>2</u> Marital Status
    3) Over ½ to ¾   11
    4) Over ¾ to Full Blood   1) Never Married
                                    2) Married
                                    3) Widowed
<u>3</u> Frequency speak Native language 4) Divorced
12   5) Seperated
    1) Speak only English   6) Common law
    2) Speak mostly English
    3) Speak English and Native language equally
    4) Speak mostly Native language
    5) Speak only Native language

In this example, the codes for the variable "Tribe" would be too lengthy to list on the questionnaire and a separate key would be used for coding purposes. A sample coding key for tribes is presented in Appendix B. The small numbers under the space for responses represent the card columns where the information will be punched on computer cards, one digit per column. If the punched card method is used, then 80 columns are available per card (96 columns for some newer machines, but most facilities still use the 80 columns).Deciding the card column assignments is known as defining the **fields of data**. The system for coding is decided upon by the researcher prior to the collecting of data. Specific examples of this process are shown in the previous chapter on SURVEY RESEARCH.

The following may be useful to you to remember in defining the data fields:

* A respondent identification number simplifies changing or adding information to a case

* Other identifying information, such as date completed or interviewer, may be useful in clearing up any later difficulties with the data

* Enough digits or spaces are important to allow for the maximum responses (e. g. two digits would allow a recording of 99 maximum, three digits would allow 999)

* If more than one 80-column card is required for each case, the cards should be numbered (often the last one or two digits of the 80 columns are reserved for a card number)

* If more responses are possible than those identified, another category called "other" can be created for a variable (to keep the value of the information, a space for comments can be created, and these comments can be quoted or summarized in the research report)

A key is sometimes compiled which defines the fields of data and corresponding card columns. Such a key may be useful in coding the data, but may not be necessary if the card columns are indicated on the form used for data gathering. In Figure 7.2, an example of a key defines the fields of data presented earlier in Figure 7.1.

Figure 7.2   Example of a Key Defining Fields of Data

Variable	Card Column
Respondent	1-3
Age	4-5
Sex	6
Tribe	7-9
Blood Quantum	10
Marital Status	11
Frequency speak Native language	12

Continuing the above example, a sample coding form is shown in Figure 7.3, filled out according to the data recorded above (for one respondent) in the sample questionnaire. Tablets with the card columns numbered at the top are readily available. These coding forms help minimize keypunching mistakes and are usually the only data format acceptable to keypunching services.
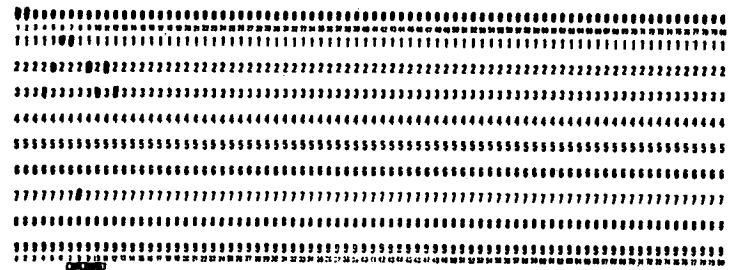
Figure 7.3   Coded Data

A card punched with the sample data would appear as follows.

Figure 7.4   Punched Data

Although other mediums are available for recording and storing data, cards are most often used by the first time or small scale computer user. The advantages and disadvantages of the various input and output mediums are discussed in the following section.

INPUT AND OUTPUT

The computer, an information processing machine, simply accepts information or input and gives results or output. This process occurs according to a set of instructions given by the user. These instructions are called programs. Devices for reading input, storing information (mass storage), and generating output are called peripheral equipment. This equipment is connected to the central processing unit, or CPU, which performs arithmetic operations, directs temporary storage of information in memory, and controls the functioning of the peripheral units. The diagram in Figure 7.5 shows a basic computer system.

Figure 7.5
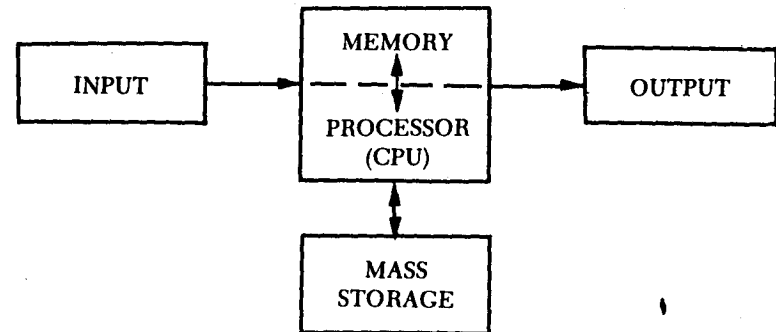
To use a computer, you do not need to know the mechanical details about the machine; however, decisions do need to be made about the peripherals that will be used to input data and output results. Devices that input data include punched card readers, mark sense readers (which detect pencil marks on pre-printed forms), and terminal keyboards. Examples of output devices that display or write the results are line printers, card punches, plotters, terminal printers and CRT screens. The input devices most commonly chosen by those using the computer for the first time are punched cards and timesharing terminals (keyboard and CRT terminal display). Output is commonly printed on line printers. Some of the commonly used mediums include:

Punched cards: Holes punched in a pattern represent numbers, letters, and symbols on a paper card. Data cards are punched by the user or a commercial service, on a keypunch machine for input purposes and the cards that are the result of computer output are prepared automatically on a card punch. Machines called verifiers are sometimes used to check the correctness of the punched data. The type of card used most frequently has 80 columns; however, IBM introduced in 1970 a card that contains 96 columns of data.

Magnetic tape: Data are recorded on magnetic tape as magnetized marks. The recording and reading of data is performed by the same machine, a tape drive, which is operated by an employee of the computer center. Magnetic tape is usually ½ inch wide (or, more recently, cassettes are used). It is a very popular medium for storing data for long periods of time, where cards might be subject to warping or other age problems. Tapes are so inexpensive that the user often purchases and stores the tapes himself.

Magnetic disk: The term used for a stack or pack of round disks, upon which data can be recorded as magnetized spots. Access to the data is usually faster than tape, but use charges generally are greater than for tape. This medium is used to store data and programs which are being used frequently, where the user does not want to carry around and constantly re-read large card decks.

Printers: Typically 14 5/8" wide line printer output with 132 columns of characters across the page and 55 to 60 rows vertically per page. Terminal printers usually have 8½" wide paper with 72 to 80 characters across the page.

In Figure 7.6, the advantages and disadvantages of the most commonly used input, output, and storage mediums are detailed.

Figure 7.6 INPUT, OUTPUT, AND STARAGE MEDIUMS

MEDIUM / DEVICE	ADVANTAGES	DISADVANTAGES
Punched Cards	Data preparation is easy for beginners Keypunch machines are readily available Data can be visually checked for accuracy Cost is low Corrections are easy to make	Storage is difficult, since cards may warp Card readers vary in speed, and may be slow Card readers sometimes tear up cards Record length is fixed (80 or 96 characters)
Magnetic Tape	Data is transferred rapidly Tape may be erased and reused Storage is easy and warpage problems eliminated Cassette tapes available Moderate cost for long-term storage Records can be almost any length	More experience needed to use Initial input medium needed, such as teletype or punched cards All data are examined sequentially, even if separate files are created
Magnetic Disk	Fast access due to direct access to selected data, without examining each record sequentially Records can be almost any length Moderate cost for short-term storage Disk may be erased and reused	Relatively expensive for long-term storage Another initial input medium needed, then transfer to disk
Printed Paper	Fast, because an entire line is printed at one time Easy for the user to read Produces output ready for reports Can be duplicated with an office copier	Storage takes room, deterioration of paper with time For graphs, greater accuracy can be obtained with a plotter

MEDIUM / DEVICE	ADVANTAGES	DISADVANTAGES
Plotted Paper	Plotter produces graphs and drawings with great accuracy Different ink colors readily available	Special machinery needed, may not be available at a particular facility Cost much higher than line printer Plot programs necessary
Microfilm & Microfiche	Stores extremely large amounts of data is a very small space Output doesn't deteriorate Data recording occurs at higher speed than printing Colors available Low cost per unit of output	Very considerable delay in getting output from a job Special reader needed to view output from a job Special reader needed to view output, not readily available
CRT Terminal	Keyboard allows rapid entry of data Display, similar to a television set, is rapid and provides visual access for scanning of large amount of data Use for rapid editing Interactive, allows for rapid results and immediate feedback	No hard copy (permanently recorded data) Extra charges for connect time to computer and disk storage
Teletypwriter Terminal	Hard copy produced for either input or output. Interactive	Relatively slow---slower than printer for output--slower than CRT for interactive work, also noisy and more likely to break down

Once the decision is made as to the desired input and output format, instructions are given to the computer in the form of a computer program, or set of instructions for the processing of data. Fifteen years ago, those persons using a computer usually had to learn a computer language in order to write a program or processing the data and performing calculations. Today many already written or "packaged," programs are available for processing data relating to social and economic studies. Some of these packages include good operating instructions for the beginning user. In addition to performing calculations, these package programs usually label the results for clear reporting. With these programs, in addition to simply providing coded data, the user must create a set of commands to the program for reading the data, requesting the specific calculations to be performed, and labeling and writing the output. This is much easier than writing a program from scratch and generally the diagnostics for detecting errors in the set-ups are very good.

PROGRAM PACKAGES

Program packages are highly recommended for the user who needs to process data, yet doesn't want to invest large amounts of time required to learn computer programming. One such package that is readily available at colleges, universities, and some computer service businesses is the *Statistical Package for the Social Sciences*. This set of "canned" programs will perform a wide variety of calculations and will label the output clearly, according to a set of given instructions. The package is well known for its excellent documentation and user's guide (see under the author, Nie, in Additional Sources). Some examples of the capabilities and kind of report output that is available with SPSS follow.

Example 1: This sample printout, Figure 7.7, is from the program *FREQUENCIES* of the *Statistical Package for the Social Sciences* (SPSS). The program has labelling capabilities and is particularly useful for reporting purposes. In this example, a progress report is being printed for the project, "American Indian Center---Urban Adjustment Survey." The variable summarized is "Tribe," with the numbers listed under "Code" having been defined as representing the tribe labels listed to the left. The absolute frequency is the number of times each tribe is listed in the sample. The relative frequency represents a percentage of the total, for each tribe listed. The adjusted frequency refers to the percentage calculated with missing values excluded. For example, if we had asked that the value "0" (zero) mean "not certain," this response would be calculated excluding the cases having that response. This provides the ability to selectively disregard certain answers in calculating the results. The cumulative frequency totals the frequencies, or accumulates the values vertically down the column.

Figure 7.7   EXAMPLE OF A FREQUENCES TABULATION

AMERICAN INDIAN CENTER------URBAN ADJUSTMENT SURVEY

TR   TRIBE

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)
NOT CERTAIN	0.	1	1.5
ALABAMA	4.	1	1.5
APACHE	8.	5	7.7
CHEYENNE	31.	1	1.5
CHIPPEWA	36.	4	6.2
CHOCTAW	38.	1	1.5
HOPI	64.	2	3.1
KICKAPOO	76.	1	1.5
MONO CA	106.	1	1.5
NAVAJO	115.	3	4.6
PAIUTE CA	133.	4	6.2
PIMA	144.	2	3.1
PONCA	146.	1	1.5
QUECHAN	152.	3	4.6
SEMINOLE	162.	2	3.1
SIOUX	172.	20	30.8
UTE	192.	5	7.7
WINNEBAGO	204.	3	4.6
YAQUI	211.	1	1.5
MISSION	219.	1	1.5
CREEK	220.	2	3.1
OTHER CALIFORNIA INDIAN	221.	1	1.5
TOTAL		65	100.0

Example 2:   This example of a *FREQUENCIES* output from *SPSS* illustrates the value of the cumulative frequency calculation. For example, if community planning efforts required information concerning the educational level of community members, this table would indicate that 49.2% of the sample had completed the 11th grade or less.

Figure 7.8   EXAMPLE OF CUMULATIVE FREQUENCY

AMERICAN INDIAN CENTER-----URBAN ADJUSTMENT SURVEY

GRADCOM   HIGHEST SCHOOL GRADE COMPLETED

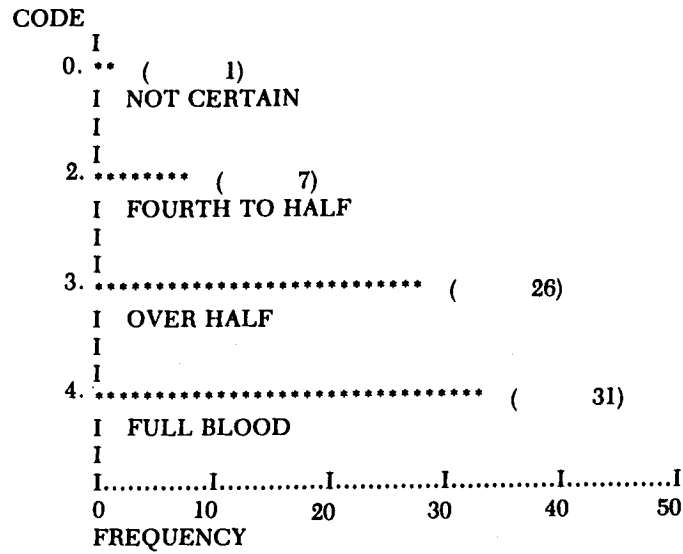| CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
	0.	7	10.8	10.8	10.8
	6.	3	4.6	4.6	15.4
	7.	1	1.5	1.5	16.9
	8.	2	3.1	3.1	20.0
	9.	4	6.2	6.2	26.2
	10.	3	4.6	4.6	30.8
	11.	12	18.5	18.5	49.2
	12.	31	47.7	47.7	96.9
	13.	1	1.5	1.5	98.5
	14.	1	1.5	1.5	100.0
TOTAL		65	100.0	100.0	

VALID CASES   65   MISSING CASES   0

Example 3:   This histogram of the variable "Indian Blood Quantum" illustrates on option of the program *FREQUENCIES from SPSS*. This type of visual display of data is useful in research reports to present information in an interesting, as well as descriptive, format.

Figure 7.9   EXAMPLE OF A HISTOGRAM

AMERICAN INDIAN CENTER------URBAN ADJUSTMENT SURVEY

BLDQT   BLOOD QUANTUM

```
        CODE
           I
        0. ••  (      1)
           I   NOT CERTAIN
           I
           I
        2. ........  (      7)
           I   FOURTH TO HALF
           I
           I
        3. ...........................  (      26)
           I   OVER HALF
           I
           I
        4. .............................  (      31)
           I   FULL BLOOD
           I
           I............I............I............I............I............I
           0      10      20      30      40      50
           FREQUENCY
```

VALID CASES   65   MISSING CASES   0

Example 4:   This example of output from program *CROSSTABS* illustrates the type of breakdowns of data that can be obtained and displayed. This chart shows a breakdown of the variables "Frequency Speak Native Language" by "Blood Quantum," illustrating the number and percentage of responses for "Blood Quantum" for each frequency interval represented in the sample.

Figure 7.10   EXAMPLE OF A CROSSTABULATION

FREQ COUNT ROW PCT COL PCT TOT PCT	ONLY ENGLISH 1.	MOSTLY ENGLISH 2.	ENGL & N ATIVE EQ 3.	MOSTLY NATIVE 4.	ONLY NAT IVE LANG 5.	ROW TOTAL
BLOOD 0. NOT CERTAIN	1 100.0 20.0 1.5	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	1 1.5
2. FOURTH TO A HALF	4 57.1 80.0 6.2	3 42.9 7.1 4.6	0 0.0 0.0 0.0	0 0.0 0.0 0.0	0 0.0 0.0 0.0	7 10.8
3. HALF TO 3 FOURTH	0 0.0 0.0 0.0	19 73.1 45.2 29.2	7 26.9 50.0 10.8	0 0.0 0.0 0.0	0 0.0 0.0 0.0	26 40.0
4. FULL BLOOD	0 0.0 0.0 0.0	20 64.5 47.6 30.8	7 22.6 50.0 10.8	3 9.7 100.0 4.6	1 3.2 100.0 1.5	31 47.7
COLUMN TOTAL	5 7.7	42 64.6	14 21.5	3 4.6	1 1.5	65 100.0

A minimal setup for SPSS will require the user to define a variable list (a list of short names for the variables used), labels for the variables (a list of labels to indicate a longer label to be printed out as identification for variables), values for the responses (a list of words describing what the numerical responses stand for), an input format (set of instructions indicating which card columns have which variable values), a definition of the input medium (for example, whether the data would be read from cards, tape, or disk file), an indication of the number of cases to be processed, and an indication of the SPSS program(s) to be used for processing. In Appendix B, an example SPSS setup for the sample *Needs Assessment* questionnaire presented in Chapter 2 is listed in detail.

A WORD ON PROGRAMMING LANGUAGES

For the majority of community projects, program packages can fill data processing needs. For others that require specific applications, the development of an original program may be necessary. For example, some accounting systems set up with a small home or office computer might require an original program. To learn a programming language you will need much study and practice, with the emphasis on practice. Several self-teaching books are listed in the bibliography for this chapter. Here is a brief description of the more common programming languages:

Basic   (Beginner's All-purpose Symbolic Instruction Code)

A simple and straight-forward language, with relatively few rules, which is particularly appropriate for beginners. BASIC is a versatile language which is commonly used for time-sharing systems, medium to the largest computers as well as mini-and microcomputers. It has unusually poor control of printing format.

Cobol   (Common Business-Oriented Language)

This language was originally developed for commercial purposes, and is particularly popular for business and accounting applications. COBOL reads quite like English, yet has more rules than many other languages. It is more awkward than most for calculations, but excellent control of printing format.

Fortran (Formula Translator)

Although this language was originally developed for mathematical capabilities, FORTRAN has become a common language for publishing

social science applications. The arithmetic statements in this language are similar to algebraic statements.

RPG (Report Program Generator)

This language has a specialized function, that of preparing reports. The instructions are straightforward and simple, yet specific or formalized. Although mathematical computations are possible, the advantages of RPG lie with the input and output, or file management and report writing features.

Other programming languages in common use include PL/1 (Programming Language 1) and APL (A Programming Language). These languages are used more frequently for mathematical applications.

Some knowledge of programming is always very useful to the user setting up an already existing package program, like SPSS. The decision to learn a computer program depends on the interests of the user, the availability of assistance, the availability of computer resources, and (above all) the availability of a large amount of time. The description of several program packages in Figure 7.11, may be useful in making the decision to use an available package. Another alternative to consider, if you have a specialized computer application, is that of hiring a computer programmer to write the program for you. This may save a great deal of time and funds in generating a custom program, but will still probably cost much more than use of a packaged program.

Summary

The main advantage of using a computer for data processing relates to the speed and accuracy at which large amounts of data can be handled. The instructions and methods for examining the data remain the same whether a pencil, a small calculator, or a large computer is used. Constructing a computer setup and preparing data take time. Some of the advantages of investing this time are the report writing capabilities and the speed at which intermediate runs and output can be gained. Particularly with community-based research, the intermediate reports on data collection progress can be valuable in providing information back to researchers and to the community before the study is completed. Cross-tabulations can be valuable in discovering possible correlations at an early stage in the project, enabling the researcher and/or advisory committee to focus on a particular topic if the crosstabs indicate a likely relationship. When progress is reported gradually, advisory committees can truly provide advice on redirection or confirm the course of the research.

Figure 7.11   PROGRAM PACKAGES

Name	Type	Description
BMD (Bio-Med)	Statistical	A package of statistical programs designed for the use of social science and biomedical data, intended for the user familiar with advanced statistics. Includes programs for descriptive statistics, linear and nonlinear least squares, multiple group stepwise discriminant analysis, factor analysis, principal component analysis, time series, univariate and multivariate analysis of variance, and variance component estimation.
DATA-TEXT	Statistical	A general purpose system for the statistical analysis of social science data, with advantages in processing an unlimited number of variables in correlation and factor analysis routines. Relatively expensive to run, yet more cost effective than other packages with large amounts of data.
ESP (Economic Software Package)	Statistical	Package for the statistical analysis of time series by ordinary least squares and two-stage least squares methods, designed for computational and data processing requirements encountered in econometeric research.
FAMULUS	Bibliographic	Designed to process reference collections maintained by researchers. FAMULUS contains programs for editing, sorting, merging, galley printing, indexing, and searching according to keywords. FAMULUS also allows for the storing and printing of abstracts.
MINITAB	Statistical	General purpose, easy for beginners in statistical and line plotter package. Capabilities include univariate parametric statistics (standard deviation, t-test, one-way analysis of variance, multiple regression, etc) and nonparametric statistics. Also offers random data generation capabilities, flexible histogram, and two and three dimensional plotting routines. Error scanning & diagnosis is extensive and explicit for beginners.

Name	Type	Description
OSIRIS	Statistical	Package designed for the analysis of social science data, with capabilities including data generation, correction and modification of data files, transformation of data, and univariate and multivariate analysis programs. Other features include unique statistical programs which perform multivariate analysis using qualitative data, provisions for handling missing data, weighting, case selection, and file documentation.
PSTAT	Statistical	Package for the analysis of social science data, including capabilities for stepwise regression, factor analysis, crosstabulation, and multivariate analysis of variance. System is well adapted for time-sharing systems.
SAS (Statistical Analysis System)	Statistical	Capabilities include transforming or deleting variables or cases, flexible commands for reading, printing, sorting, merging, sorting data, and report writing. Package also has procedures for converting SPSS, OSIRIS, BMDP, or DATA-TEXT files into files for use with SAS. Adapted for time-sharing systems.
SOUPAC	Statistical	The package of statistical data processing programs includes populations statistics, analysis, and nonparametric statistics. May be expensive to run.
SPSS (Statistical Package for the Social Sciences)	Statistical	Developed for the analysis of social science data, the package includes descriptive statistics, aggregate statistics, crosstabulation, correlation, regression, discriminant analysis, and factor analysis. The package use is designed for those without previous programming experience and the explanations of statistical procedures in the accompanying manual are excellent.

WHERE TO GO FOR SERVICES

One of the most available and least expensive sources for computer services is a college or university. Although some limit service to research projects funded through the college, many do accomodate outside users. Costs may be more reasonable because the college does recover some funds for computer use through student tuition, but one of the main reasons why this source is recommended is the availability of consultant services. Most college facilities provide a reasonable amount of consultant time, free of charge. The billed cost for computer time generally also includes the use of keypunch machines and CRTs and, sometimes, verifier machines for checking keypunch accuracy. For persons in remote areas, time-sharing computing by telephone connected equipment, can often be purchased from college or computer service facilities. These facilities usually have provisions for mailing computer output to the user.

Private companies or individuals can provide greater personal assistance, but usually at greater cost. Many nationally used record-keeping systems have been adapted to computer processing by consultants who are willing to process data on a regular basis. Purchasing such services can be more economical overall than designing a data processing system and training a staff. Also, computerized accounting systems are readily available. Usually, the general collecting and summarizing of data for program needs can often be purchased; whereas, special applications sometimes needed in research must be designed.

A difficulty commonly encountered in obtaining computer services lies in the terminology connected with computer concepts and equipment. Although the terms may seem numerous, many consultants are sensitive to the needs of the novice users and will explain the meanings when asked. Figure 7.12 is a short glossary of the more common terms encountered in obtaining computer services. As computers become more incorporated into everyday life, the development of good self-teaching books, program packages, and consultants increases.

Figure 7.12   GUIDE TO COMPUTER RELATED TERMS

Analysis:	An examination of anything into parts or elements, separately or in relation to the whole.
Bug:	Slang for an error in a set of instructions given to the computer.
Card:	Common term for punched card, a paper card which holds 80 to 96 columns of alphabetic or numerical information, according to a pattern created by the presence or absence of punched holes.

Case:	The basic comparative unit included in a study, such as a respondent or a community.
Coding:	Writing data or instructions in a form which can be interpreted by the computer program.
Computer:	Machine which accepts information (numbers or letters or the alphabet), processes the information according to a set of instructions, and produces results.
CRT:	A Cathode Ray Tube is a television screen display and may be utilized as a computer-linked terminal.
Data:	A group of facts or statistics, sometimes referred to as raw data. The term "data" refers to the plural, i. e. "the data are...".
Data Processing:	Operations performed on data, usually by means of automatic equipment.
Debugging:	Finding and correcting the errors in a set of computer instructions.
Disk:	A pack or stack of magnetic disks, upon which data can be recorded as a series of magnetic spots.
File:	A group of related records, so defined by the user.
Flowchart:	A diagram showing the logical sequence of events in a set of computer instructions.
Hard Copy:	A copy of the output results printed on paper.
Hardware:	The physical machines that read, process, store, and write data.
High Level Language:	A set of symbols, often resembling English, which can be used to input instructions to a compiler. Examples of some of these languages are BASIC, FORTRAN, COBOL, PL/1, and APL. The compiler is a program which then translates the instructions into computer machine language.
Input Device:	A machine that reads previously recorded data. Examples of input devices are punched card readers, mark sense document reader, keyboard, and light pen.
Mark Sense:	A process of recording data which is read according to the presence or absence of pencil marks in specified locations on a form.
Medium:	The physical material upon which data are recorded. Examples of mediums are punched cards, magnetic tape & magnetic disk.

Microcomputer: A small computer, often costing only a few hundred dollars, and used primarily for home or small business use.

Output Device: A machine that displays or writes computer output. Examples include the line printer, card punch, terminals, and digital plotter.

Peripherals: Equipment connected to the computer (such as printer, card readers, tape drives), to perform input-output operations.

Plotter: A peripheral device which can draw graphic representations of data according to computer program output.

Program: A set of instructions in a particular computer language, which gives commands to perform a series of operations. Examples of these operations include the reading of data, performing calculations, and printing results.

Read: To sense data from a source medium, such as cards or tape.

Record: A set of related data, such as a case, considered as a unit.

Software: Instructions given by the user to the computer.

Tape: Often the short term for magnetic tape, a recording medium which stores data as magnetized marks.

Teletypewriter: A peripheral interactive terminal which produces hard copy.

Time-sharing system: A system of computer use by more than one user at the same time, usually by means of several terminals.

Value: The response indicated for each variable.

Variable: A measurement or quality obtained for each unit in a study.

Write: To record data on a medium.

Cost for computer services depends largely on the experience of the user. A great deal of processsing can be obtained for under $100 when the user is experienced. Otherwise, a great deal of computer time can be wasted in debugging, or removing the errors from a program or package setup. Purchasing the services of a consultant can in many instances reduce the costs of wasted computer time, in addition to helping the user avoid frustration. Most computing facilities offer keypunch services by experienced operators. Machine verification, or checking for errors through the use of a verifying machine, is usually included in the charges. The decision of whether to do data recording in-house or to hire the services depends on the availability of staff, training, and the need for accuracy.

In a remote area, a time-sharing setup with data entry at your organization center may be a more efficient system. Choice may not be simply based on dollars and cents, for the training that staff can receive from doing data processing in-house can be a valuable resource for the future.

Personal computers are very useful for **word processing** (or the creation, modification, and storage of text). For organizations that have a great deal of report or proposal writing, the word processor enables quick changes and error-free printing. This can greatly assist the organization when many comments from community members or staff members are added to the text. The ease and time saved in editing tend to encourage more comments, and, therefore more community participation in the writing process. Also, many people find word processing systems useful for improving writing skills, due to the ease of rewriting. At this time, statistical program packages for personal computers are not very well developed.

Personal computers (from such manufacturers as Apple, IBM, Osborne, Radio Shack) are economical for individual organizations to purchase; however, the technical expertise needed to program and to maintain them is expensive and difficult to obtain in non-urban locations. For those persons contemplating a home computing system, a glance at Libes' *Small Computer Systems Handbook* may help in making the decision. Personal computers are generally owned by persons who are willing to invest a great deal of time in learning and experimenting with their use.

Computers can be an economical way to process data and produce reports. Alternatives, such as programmable calculators and processing services might be considered carefully before undertaking computer processing, for working with a computer requires time, funds and patience in exchange for the end product. A certain amount of frustration is experienced by everyone who uses a computer, when things do not go as planned. Perhaps this is caused by the difficulty in remembering that the computer is only a machine; it doesn't think. The functioning of the computer is only as good as the quality of the instructions given to it.

Positive aspects of investing time in learning to process data by computer include the accuracy gained, and the technical knowledge gained for future projects. As a tool for community-based research, the computer can be of great value in producing intermediate reports for community discussion and involvement in the research process.

ADDITIONAL SOURCES

Ashley, Ruth *ANS COBOL: A Self-Teaching Guide* (New York: John Wiley & Sons, 1979).

This is an introduction to COBOL statements, arithmetic coding, accessing card files & tape files, mass storage (disk files), printed reports for beginners.

Blatt, John M., *Introduction to FORTRAN IV Programming* (Pacific Palisades, CA: Goodyear Publishing Company, 1971).

This discusses FORTRAN statements, arrays, subprograms, debugging, machine and assembly language. Emphasis is on the WATFOR, WATFIV compilers. (Beginner to intermediate level.)

Friedman, Jehosua, Philip Greenberg, and Alan M. Hoffberg *FORTRAN IV: A Self-Teaching Guide* (New York: John Wiley & Sons, 1975).

This is beginners guide to computer concepts, input & output and FORTRAN statements. This includes exercises for self-teaching and is written clearly.

Harris, Martin, *Introduction to Data Processing: A Self-Teaching Guide* (New York: John Wiley & Sons, 1979).

The text discusses the concept of automated data processing, an overview of computer systems, storage systems, flowcharting, the impact of computers, systems analysis, writing a computer program in BASIC, and characteristics of selected computer systems. (Beginner level).

Libes, Sol, *Small Computer Systems Handbook* (Rochelle Park, NJ: Hayden Book Company, Inc., 1978).

This includes discussions on the advantages of the small or home computer, digital logic, central processors, microprocessors, interfacing, mass storage systems, computer software, and applications. This is an optimistic guide aimed at those who already decided to develop a small system, yet the discussion should be valuable in assisting persons to decide if available resources are adequate. (Intermediate to advanced level).

London, Keith, *Introduction to Computers* (Boston: Faber and Faber, 1968).

This discusses the concept of computers, the central processor, peripheral units, programming, basic computer techniques, the role of computers in commerce and industry, and the use of computers for scientific work. (Beginner to intermediate level).

Nie, Norman, C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale Bent, *Statistical Package for the Social Sciences*, Second Edition (New York: McGraw-Hill Book Company, 1975).

Although an introduction to the SPSS program package, this text give excellent descriptions of coding data for input, and several statistical procedures such as descriptive (frequencies, histograms, mean, standard deviation), Chi-square, T-Test, Bivariate Correlations, Multiple Regression, Analysis of Variance, Discriminant Analysis, and Factor Analysis. This is the comprehensive text for the package, and is written at the beginner to intermediate level.

Spencer, Donald, *Computers in Action* (Rochelle Park, NJ: Hayden Book Company, Inc., 1974).

This is a helpful guide in making the decision of whether or not to use computers. This discusses the history of computer technology; how computers work (digital computer organization); getting information in and out of the computer; introduction to BASIC computer programming language. This is well written and realistic. (Beginners level).

Stein, Marvin L. and William Munro, *A FORTRAN Introduction to Programming and Computers* (New York: Academic Press, 1966).

This presents elements of FORTRAN, input and output, data statements and modes for arithmetic, introduction to modern digital computing systems. (Intermediate level).

Stern, Nancy B., *Flowcharting: A Tool for Understanding Computer Logic* (New York: John Wiley & Sons, 1975).

This is a self-teaching guide. This describes the logic of computer processing, the relationship between a job description, a flowchart, a program, and elements of flowcharting. (Beginner to intermediate level).

Worth, Thomas, *BASIC for Everyone* (Englewood Cliffs, NJ: Prentice Hall, Inc., 1976).

This includes information on BASIC statements, files processing, flowcharting, subroutines, libraries, and paper tape usage. This includes problems for self-instruction. (Beginners level).