CASP V3.0

ANTHROPOMETRIC SOFTWARE PACKAGE TUTORIAL GUIDE AND HANDBOOK

By Michael Dennis Jordan

The complete instruction set for using the anthropometric analysis software based on the

CDC Standard Deviation-Derived Growth Reference Curves
Derived from NCHS/CDC Reference Population

NCHS Growth Curves for Children, Birth-18 years, United States
Series 11-number 165, DHEW Publication number (PHS) 78-1650

Software by the Div. of Nutrition, CCDPHP, Centers for Disease Control
Supported, in part, by the Office of Nutrition
Agency for International Development (AID)
under RSSA # BST-1064-R-HC-2174-03

THE CENTERS FOR DISEASE CONTROL
Center for Chronic Disease Prevention and Health Promotion
Division of Nutrition, Statistics Branch
1600 Clifton Road
Atlanta, Georgia 30333

US Dept of Health and Human Services US Public Health Service

THE CDC ANTHROPOMETRIC SOFTWARE PACKAGE
Version 3.0
TUTORIAL GUIDE AND HANDBOOK

By Michael Dennis Jordan

The complete instruction set for using the anthropometric analysis software based on the

CDC Standard Deviation-Derived Growth Reference Curves Derived from NCHS/CDC Reference Population

NCHS Growth Curves for Children, Birth-18 years, United States Series 11-number 165, DHEW Publication number (PHS) 78-1650

Software by Michael D. Jordan. Anthropometric subroutine set supported by Norman W. Staehling, Div. of Nutrition, Center for Health Promotion and Education

Supported, in part, by the Office of Nutrition, Agency for International Development (AID) under RSSA # BST-1064-R-HC-2174-03

This software is in the public domain and available without cost.
Technical assistance may be obtained by writing to the author.

THE CENTERS FOR DISEASE CONTROL
Center for Health Promotion and Education
Division of Nutrition, Statistics Branch
1600 Clifton Road
Atlanta, Georgia 30333

Second Printing
April 1987

ACKNOWLEDGEMENTS

Foremost to
NORMAN W. STAEHLING, MS, Statistician
who maintains the FORTRAN version of the anthropometric subroutine
upon which the microcomputer version is based.

Furthermore, the author extends gratitude to the following individuals
and institutions for their contributions and guidance.

The Centers for Disease Control

Frederick Trowbridge, MD
Ronald Fichtner, PhD
Barbara Sullivan
Faye Wong, MPH, RD
Richard Peck, MPH
Steve Ilardi, BS

The Johns Hopkins University School of Hygiene and Public Health

Shiriki Kumanyika, PhD
Sandra Huffman, ScD

The International Food Policy Research Group

Bruce Cogill, PhD

TABLE OF CONTENTS

WAIT!

Before you use any of the software in the package, protect your materials:

COPY YOUR DISKETTES!

The Program Diskette contains the most essential parts of the Anthropometric Software Package. The Supplemental Files Diskette has sample files that will help you learn to use the package properly.

The files on the two diskettes are not copy protected. In fact, we urge you to copy them before you do anything else. Some of the sample files are used in the tutorial portions of the documentation. You will have to read from and write to some of these files. If they are lost or you accidentally ruin them in the learning process, you will not be able to finish the tutorial.

Use only copies of the diskettes, not the originals.

If you have a computer with a non-removable hard disk, you should create a directory (consult your DOS manual) and copy the entire package into your new directory.

INVENTORY OF CASP PACKAGE

1 User Registration Form (MUST be filled out and sent to CDC!)
1 User's Tutorial and Guidebook
1 5.25 inch DS/DD Program Diskette containing:

ENTRY.EXE
TAB.EXE
BATCH.EXE
DATA

1 5.25 inch DS/DD Supplemental Files Diskette containing:

DEMOBAT.TXT
DEMOTAB.TAB
SAMPLE.ETB
SWAP.EXE
LCHANGE.EXE

FOR VERSION 2 USERS

In Mid-1985, The Centers for Disease Control, Center for Health Promotion and Education, distributed the second release of the Anthropometric Statistical Package (CASP). The patients files you've made with Version 2 are not compatible with Version 3.0.

SWAP, a supplemental program found on the supplemental files diskette of this package, will convert your old Version 2 files into Version 3.0 files.

The conversion program makes files with three different extensions: ETB, TXT and FLD.

ETB files are SWAP-converted files that can be used immediately in the Version 3.0 ENTRY and TAB programs. They are direct conversions of the Version 2 data including the anthropometric indices. Note: since SWAP performs no recomputations there is no way for the program to detect when it is copying an Edit Flagged record. (See EDIT FLAGS in the user handbook). There is no Edit Flag field in the Version 2 files to copy therefore the Edit Flag field in the Version 3.0 ETB files will all be blank.

TXT are conversions of Version 2 files into an ASCII file without the anthropometric indices. Some programmers call them batch files or rectangular files. The ASCII file is a bonus. There are many applications for this TXT file because it can be easily read and edited by other programs.

TXT files contain all the data from the individual records in Version 2 files except the anthropometric indices. The BATCH program in this package can reprocess the TXT files and apply the computed indices. The processed files are also output as ASCII files. BATCH also produces ETB files.

The advantage of reprocessing the TXT into a new ETB is that you will obtain anthropometric indices that are calculated by the algorithms identical to those in the Centers' central computer.

FLD is a field location **template and partner** to TXT file. You'll need it should you decide to reprocess the TXT file with BATCH. FLD contains all the field locations for I.D., age, sex, height, weight, and all extra classification variables, if any. Read BATCH for further details about FLD files.

Lotus 1-2-3* compatibility **promised in Version 2** was never fully successful. In Version 3.0 there is a program called LCHANGE on the supplemental files diskette that will convert ETB and TAB files to 1-2-3 PRN files. Read Appendix D for further details.

* Lotus 1-2-3 and DBase used throughout this manual are copyrighted trademark names used by Lotus Corporation and Ashton-Tate, Inc., respectively.

INTRODUCTION

WHAT CAN I DO WITH THE ANTHROPOMETRIC SOFTWARE PACKAGE?

The CDC Anthropometric Software Package consists of three programs called ENTRY, TAB and BATCH, and a reference data set. The programs perform several related tasks that allow the user to go from raw, disorganized data to precisely calculated and statistically analyzed listings, cross-tabulations, and presentation graphics.

The principal uses of the software package are record maintenance, anthropometric calculation and statistical analysis. Although there are already many statistical analysis packages available, this one, developed by the Division of Nutrition, Center for Health Promotion and Education at the Centers for Disease Control, is customized for public health professionals.

The programs were made to operate on IBM-compatible micro-computers including the portables. For field-oriented work there are enough features in the software to operate a fully independent field site capable of cross-tabulating 50,000 patient records.

The output files from ENTRY and BATCH are saved in a format that allows them to be transmitted to mainframe and micro computers and used by other statistical analysis programs.

A micro-computer operating under IBM DOS 2.1 or higher with a minimum of 256K and a dot matrix printer are the basic equipment. If you can get one, a color monitor is recommended over a monochrome.

The next few pages describe in detail the capabilities of the three programs in the package ENTRY, TAB and BATCH.

THE ENTRY PROGRAM

The ENTRY program is principally used for performing anthropometric calculations and creating patient files. When given age, sex, height and weight, the program returns the percentiles, z-scores, and percent of medians for height/age, weight/age, and height/weight. These calculations can be saved in a file named by the user with the patient's identification if desired. The program can manage 5000 records per file. Also the user may create 24 additional fields to collect other data.

Among other features are:

1. a provision for weighting data
2. a probable error flag
3. the capability to search and separate all error-flagged records or records with common characteristics
4. sub-setting and merging of files
5. listing by fields

THE TAB PROGRAM

The TAB program is used for making cross-tabulations of patient records saved with ENTRY or converted from other sources with the BATCH program. The user chooses three variables from a patient file and divides them into intervals. The program presents the results on screen or paper. The results can be saved and presented again for briefings or reprints.

Among the major features are:

1. the capability to crosstabulate weighted data
2. histogram (bar graphs) for color monitor owners
3. save/recall features for tab results and intervals used for results
4. uniformly formatted printed tabulations with titles and footnotes
5. the capability to calculate cell-by-cell means on any numeric variable

THE BATCH PROGRAM

The BATCH program is used to perform calculations and conversions upon raw files of data called "batch files." Batch files can come from mainframe computers or other PC software. If the batch file contains enough information to derive the patient's sex, age, height and weight, you only have to provide column location of this information and the files will be processed one after another without user interaction. The finished files may be cross-tabbed with TAB and edited with ENTRY.

Among the major features are:

1. on/off filter for edit-flagged records
2. available printout of all flagged records
3. full graphic display of batch records and trial processing
4. field locations can be saved as a template for easy re-use
5. the capability to produce processed ASCII output files

HOW TO USE THIS HANDBOOK AND SURVIVE THE LEARNING PROCESS

Much to the dismay of programmers and technical writers, the typical computer user doesn't like to read. Most over-anxious would rather jump into a program and read the instructions after the damage is done.

Unfortunately, using this software package--successfully, at least---won't be that simple. You must use this handbook if you intend to be a proficient user.

This handbook is more of a tutorial than a reference manual. The user performs the functions in the programs as he reads about them. The handbook uses step-by-step examples. The user is expected to participate as he reads. The user is expected to apply the experience gained from using the handbook examples to real-world cases.

You will need a copy of the program diskette and the supplemental diskette. Do not use your original diskettes.

If you have a computer with a non-removable hard disk, you should have already created a directory and you should use your hard disk copy.

Ideally you should try all the exercises in the handbook. However the applicability of all the material in the handbook may differ from user to user.

FOR USERS WHO WILL ENTER DATA MANUALLY: Read Chapters 1 thru 3. You may read Chapter 4 and practice the examples for familiarity.

FOR ADVANCED USERS, ESPECIALLY THOSE WITH MAINFRAME ACCESS OR BATCH FILES: Read Chapter 4 first and move on to Chapters 1, 2 and 3.

Do not start in the middle of a chapter. To be fully proficient, you must read and practice all the examples in each chapter.

CHAPTER I USING THE ENTRY PROGRAM

To start the data entry and computation program, insert the program diskette in drive A and type ENTRY. If you have copied the programs to hard disk, boot up in the usual manner and type ENTRY.

When the program starts, the first thing it will do is load up the growth curve reference set from the disk file DATA. After this is done and the menu appears, you may remove the program diskette if you are using it on a floppy disk drive.

The menu offers six choices. The last choice [6] Escape to DOS turns off the program and returns the computer to its start-up state. (DOS stands for Disk Operating System).

1.1 MAKING AN ANTHROPOMETRIC CALCULATION

Anthropometric calculations are performed in function [1] Data Entry and Interactive Computation.

Press [1].

The program may be used as an anthropometric calculator or as a data entry program. It's called Interactive because it pauses for responses from the user.

First, get acquainted with the program by studying the screen. Notice how many fields there are on the screen.

Press the return/enter and backspace keys a few times and watch how the triangular pointer moves from field to field. Make sure you use both the enter and backspace keys enough times to understand what they do before you continue.

The program needs to know HEIGHT, WEIGHT, SEX, and AGE in order to compute the percentiles, z-scores and percent of median for HEIGHT/AGE, WEIGHT/AGE, and WEIGHT/HEIGHT. The program has several ways of getting the four pieces of information it needs.

NAME: name is an optional field. Some users require it, some don't. It has no affect on the computations.

SEX: the sex for boys can be either M, m, or 1
the sex for girls can be either F, f, or 2

When you start making your patient files you must be consistent. Use only one of the symbols for each sex.

Move the pointer to the sex field. Type 'M' and press return.
IMPORTANT: Even the most careful user will make a typo error sooner or later. With this program you can move the cursor to the error with the cursor arrow keys, NOT BACKSPACE, and just type over the error. You cannot move the cursor to the first character in the string. If you make a mistake on the very first character you type in any field you must:

1. Go ahead and enter the mistake so you can get the field pointer back on the screen.

2. Press the backspace key so the pointer will move back to the field with the mistake. Don't confuse the backspace key with the move-left cursor arrow key. They ARE NOT the same thing to this program.

3. Retype the data for that field--carefully.

AGE: the program will accept ages in months, years or combinations of both or by computing the time difference between VISIT DATE and BIRTHDATE.

The proper formats for a birth and visit date are:

mm/dd/yy as in 07/26/86
yymmdd as in 860726

Both formats are valid for July 26, 1986. The computer will recognize only these two formats for dates. When day is unknown, 00 or 99 may be used. The program assumes the 15th.

IMPORTANT: There is a difference between 'age unknown' and age '0 months.' Do not use '0' to indicate an unknown age. The age, visit date and birthdate fields must all be left blank to indicate an unknown age. The only exception is '999.99' months which will be accepted as age unknown.

When age is missing measurements are assumed to be recumbent below 85 cms and standing there after.

Move the pointer to the birthdate field and type exactly as shown:
06/15/83

Move the pointer to the visit date field and type exactly as shown:
06/30/86

WEIGHT AND HEIGHT: the program will accept any number in these fields. Zero or a null field is used for unknowns. The program defaults to the metric measurements of kilograms and centimeters first. If you wish to use pounds/ounces and feet/inches you may but it would serve no purpose to mix centimeters and inches or kilograms and ounces.

For weight, move the pointer to the appropriate fields and type
36 (pounds)
2 (ounces)

For height, move the pointer to the appropriate fields and type

3 (feet)
2.5 (inches)

Don't forget to press the return key after the last entry.

Double check your screen. You should have entered:

SEX: M
BIRTH DATE: 06/15/83
VISIT DATE: 06/30/86
WEIGHT: 36 lbs, 2 ounces
HEIGHT: 3 ft, 2.5 inches

To process this record, press [*]. The instruction for this is also on the lowest line on the screen.

Notice that the anthropometric indices were performed and quickly returned. What happened was that the program plotted the patient on its stored memory of growth curves and displayed the computed indices.

If the program put an [*]asterisk on the screen instead of processing the data, it's because you didn't press [ret] after the last entry!

Do you think in metric or U.S. standard?

Study the results on the screen. Notice that pounds/ounces and feet/inches aren't displayed anymore. Users accustomed to U.S. standards are probably wondering why the program no longer showed the height and weight measurements after it computed the anthropometric indices.

The package was compiled with the display default permanently set to metric. Metrics are much more precise and easier to calculate than U.S. standard. The growth curve reference data is in metric. With the default set as it is, all the data associated with anthropometric measures must be entered in metric.

If this does not fit your purpose you may change the default to U.S. standard very easily by going back to the main menu (Press [Esc] escape). Whenever you have the main menu on the screen you have the opportunity to change from metric to U.S. or back to metric by just pressing 'U' or 'M' on your keyboard.

If you choose U.S. then all your anthropometric outputs will be in pounds/ounces and feet/inches. You will stay in U.S. standard until you stop the program or press 'M' while at the main menu.

Measurements must be input in the same mode you selected for output. In other words, don't select to display the output in U.S. mode if you are inputting your data in metric. This will cause confusion if you choose to edit or re-enter data.

1.2 EDITING YOUR ENTRIES

What if you look at the results on the screen and discover that your patient data is incorrect? You could clear the screen by pressing [RET] and start over, or you could fix only the part of the record that is incorrect.

At the bottom of the screen, three options are presented. [RET], as stated above, will clear the screen, [Esc] will take you back to the menu, and any other key will bring back the pointer and allow you to re-enter any field.

Assume that age is incorrect and that the error was caused by a faulty visit date. Let's assume that the patient was really visited on 05/30/86 instead of 06/30/86. To correct that error:

1. Press any key (except return or backspace) to get the pointer back on the screen.
2. Press [RET] or the backspace key to move the pointer to the visit date field.
3. Type and enter the new date (05/30/86).
4. Press [*] to process the data for the anthropometric indices.

You will notice a change in the computed age. Also, there have been some changes in the indices for weight/age and height/age. You could, if you desired, use the entry program as a anthropometric calculator by just changing the data in any field an indefinite number of times.

To invalidate fields (remove any data) blank out the field with spaces.

1.3 EDIT FLAGS

Sooner or later, if you enter enough data, you'll see a message called an Edit Flag. The words "EDIT FLAG" will appear flashing near the display of computed anthropometric indices.

Assume that you accidentally entered 32.5 feet for height instead of 3 feet and 2.5 inches. There wouldn't be any curve data for such an extreme height so there would be and Edit Flag warning and unknowns as indices for height/age and weight/height.

The Edit Flag warning flashes every time the patient data cannot be plotted on any of its curves. Common causes of Edit Flags are:

1. No height or weight found in the record
2. No sex indicated
3. Illegal dates in the birth or visit fields
4. Computed ages over 18 years old

The computed indices will be percentiles of 999.99, z-scores of 9.99 and medians of 999.99. These are tokens for unknowns.

Sometimes actual patient data will cause an edit flag to occur. There may be no errors in the data but still the Edit Flag appears and there are real anthropometric indices on the screen.

You may still get an edit flag message if:

1. Patient's weight and height are at some extreme that doesn't fit any of the reference curves.
2. The z-scores for weight/age and height/age are greater than 6 or less than -6.
3. The z-score weight/height is less than -4 or greater than 6.
4. Boys are taller than 145 cms; girls are taller than 137 cms.
5. Boys are older than 138 months; girls are older than 120 months

1.4 MAKING PATIENT RECORD FILES

In the previous section of this booklet, you learned to use the entry program as an anthropometric calculator without saving the results. You learned how to correct an individual field and what causes an Edit Flag.

In this section you will learn how to make a file to hold patient data and how to have even more fields than just name, sex, height, weight, and age.

First you must go back to the main menu. Press [Esc] to get the main menu, if you haven't already got it.

Press selection [3].

A small window appears. In the window you are prompted to create a name for your file. Acceptable names are any combination of up to eight alphabetic or numeric characters (no spaces). You must include the disk drive upon which you intend to create the file and you must not make up extensions for your file names. An example of a legal name is:

C:MYFILES. There is 'C' as the disk drive, followed by a colon, and the file name. That is the only correct format.

PATIENTRECS is illegal because the name is too long.
C:NEWFILE.MDJ is illegal because there is an extension

At this, time mount disk drive A and enter the name A:PRACTICE.

1.5 CREATING EXTRA CLASSIFICATION VARIABLES

A large window appears after you have entered the name of the file. In this window you may create the names of up to 24 new classification variables. If you do not want extra variables then you may press [*].

To familiarize you with the rules for extra variables, we'll make a few extra variables. Let's assume that we need to collect data on race or nationality, residence and mother's age when the child was born. That would be three extra classification variables. In the program they are called user-selected variables.

All you have to do is type a short name for the field (no longer than 7 characters). Notice we're using the pointer again. Pressing return or backspace will move the pointer to 24 stops in the window. This is to allow you to go back and change a name. Type and enter;

ETHNIC
REGION
#MAT AGE (for maternal age)

Notice that MAT AGE has a # sign before it. The # sign means that any information going into this field will be numeric. You could put numbers in any of the other fields, too, but the computer would treat them as words, instead of numeric values.

A $ sign is the opposite of a # sign. A $ sign means that the field is an alphanumeric character field. We didn't use $ in front of ETHNIC or REGION because when a field is created, the program assumes that it's an alphanumeric character field. You don't have to type the $ sign unless you are changing a numeric field to a character field.

If you make a mistake, remember from your practice with data entry how to correct it; enter the mistake and move the pointer back to the field with the backspace key.

When you have created all three labels press [*] to register them. The disk drive will activate for a moment and then the menu will return.

You have now created a file called PRACTICE. If you look at the disk's directory you will see PRACTICE.ETB. That is the file you have created and there are three extra variable names stored with the file name.

Right now the file is empty so we are going to put a few records in it.

Press selection [1].

By now you are familiar with this screen. One of the choices shown in the line menu at the bottom of the screen is [/] ACTIVATE A FILE.

Press [/] (that's the slash on the same key with the question mark).

A prompt appears in the top left corner of the screen. The program wants to know what file you are opening. You must answer this prompt with the DISK DRIVE DESIGNATION and the FILE NAME. Since you want the practice file you must type and enter:

A:PRACTICE (Later, when you are making your own files, you may replace the 'A' with whatever disk drive you wish to use).

Notice that your extra classification fields have appeared near the bottom.

Before we enter data in your new file, let's assume that you've decided you really need more than just these three fields. It's not too late to add more.

1. Press [Esc] to get back to the menu.
2. Choose selection [3] as if you were creating a new file.
3. Type and enter A:PRACTICE.

The program detects that there is already a file called PRACTICE so it will present the current list of extra variables.

At this point you may add more variables, over-strike and change the name of a current variable, or delete a variable by using the space bar.

For this exercise, let's add POLIO and MEASLES (for vaccination recording) as extra fields. Remember, press [ret] after the last entry and press [*].

Rules for creating variables are simple. You don't need to memorize them right away, but you will need to know them when you are through practicing and you start making files of your own.

1. Extra classification fields are five columns long. (This is permanently fixed.) You may have 24 extra classification fields.

2. A field label can be seven characters long. Anything beyond seven will be truncated. The # and $ signs are not counted as part of the length.

3. A # sign at the beginning of a field label means that the field will contain numeric information.

4. A $ sign at the beginning of a field cancels a # sign. You would use a $ sign to switch a numeric field back to an alphanumeric character field.

1.6 ENTERING RECORDS INTO PATIENT FILES

Go back to the menu if it isn't already there.
Decide now whether you want metric or U.S. (Press U for U.S. or M for metric.)
Press selection [1].

The data entry screen appears.
To get records into a file you must have an opened file. One of the choices at the bottom of the screen is [/] Activate a file.

Press [/].

In the upper left corner, the prompt FILE appears.

Mount a new formatted diskette in A drive then type and enter A:PRACTICE (B or C drive will serve just as well if you prefer).

The extra variables now appear near the bottom of the screen.

Making patient records is a three-step process: TYPE, PROCESS, SAVE.

TYPE in a name, sex, birthdate, visit date, weight, and height. Follow the same pattern you used in interactive computation. Make up information for the extra variables, too.

PROCESS by pressing [*]. Check for Edit Flag (accompanied by a beep). Read anthro indices.

SAVE by pressing the return key to write the information to the practice file. Each record takes up 221 bytes of disk space.

Perform these three steps for four example patients of your own. Creating this file will get you accustomed to entering records.

You can see that, while this method of moving pointer type data entry takes a little getting used to, it can be mastered with practice. Using the [*] for processing may seem a bit odd, but it was selected because there's a [*] key in easy reach of the right little finger just above and to the right of the space bar on most IBM-type keyboards (except the IBM AT).

1.7 SETTING TABS

You probably don't need to stop at every field on the screen. If you are filling out birth and visit dates, you surely don't need to stop at age in month and years. Likewise, if you're in the metric mode, you don't need to stop at pounds/ounces and feet/inches.

The pointer can be programmed to stop only at the fields you want to fill out. Let's assume you want to stop only at NAME/I.D., SEX, BIRTH, VISIT, LBS, OZS, FT and IN. It would speed up data entry to have the pointer stop only at those fields. To set tabs:

1. Move the pointer to the first field you wish to tab. (By the way, fields can be tabbed in any order!)

2. Press TAB (the double arrows next to the Q key). When a field is tabbed, there is a beep. You cannot tab the user-created extra classification variables. The logic assumed here is that you wouldn't have created the fields if you didn't plan to fill them out.

3. When you have tabbed the chosen fields, press the caret character [ ] (the symbol above the number 6). [ ] activates the tabs. This key also unsets all the tabs. If you unset the tabs you'll have to reset all of them.

Now when you move the pointer from field to field it will only stop at the fields you've tabbed.

Try entering four more records now that you have tabs. You'll see how much more efficient using tabs can be.

1.8 SEARCHING AND VIEWING RECORDS

After you've entered several records or have had someone else do it for you, it's not easy to grasp the prevailing characteristics of these patients as a group. This function is probably the most useful feature for just getting acquainted with your records.

Get back to the main menu and press [2].

The first thing you must do is tell the computer what file you wish to search. As always you do this by typing and entering a file name. DON'T FORGET TO INDICATE DISK DRIVE BEFORE THE FILE NAME.

We will review your training file PRACTICE.

The data entry screen returns but this time it has a different function. It's used to search for specific kinds of records. You may command the program to:

Find all records that were edit-flagged (press [@]).
Find all records that were not edit-flagged (press [#]).
Get all records flagged or not (press [*]).
Inclusive or exclusive search for very field-specific records (input sought after info in field it is expected to be found and press [*]).

Press [*]

You are immediately told you have eight records (unless you made more or less than instructed).

You have options for presenting the results of the search:

[1] and [2] allow you to make a listing of the records on the screen or the printer. This is the way to dump the records but in an orderly formatted manner with only the fields you choose to see.

Press [2] if you have a printer on line or [1] if you don't.

If you are going to use your printer, set the paper so that the print head starts on the first line of the page.

You are presented with a list of the fields in your records. You select the field you wish to see by the corresponding number, typing and entering each number one at a time.

Select NAME/ID
AGE
SEX
ETHNIC
REGION
W/H percentile
H/A percentile

Press return alone or type and enter '0' to begin listing.

When the listing is on the screen, the screen displays twenty records at a time. When the listing is on the printer, the printer writes fifty records per page. Each page will have a header and page number.

The search results menu returns after the printer has finished.

Selection number [3] on the menu allows you to look at a patient's total record. Press [3].

The data entry screen is back again but this time it displays the first record and the anthropometric indices. To examine other records in the file, you may press return to see the next record, backspace to see the previous record, or type and enter the record number you wish to see.

A counter in the upper right corner shows you which record you are viewing. The record numbers apply to the index of records that fit the search, not to the record's physical location on the disk. In this case, they are the same only because all the records fit the search criteria. Your printed listing of selected fields serves as an index. The consecutive numbers in the leftmost column correspond to the record selected for viewing on the screen.

1.9 EDITING RECORDS

You can perform editing functions on a record while you have it up on the data entry screen. Along the bottom of the screen you will see a small menu for selection of functions.

Press [E] for edit.

You will see that the pointer has returned. You can move the pointer to any field and change the data.

Move the pointer to the SEX field and change the data to the opposite sex.

Keep your eye on the anthropometric indices and press [*].

Notice that the anthropometric indices changed slightly. Not only did they change on the screen but the change in sex and indices were also written to the file. You have just edited a record.

As long as the pointer is still on the screen, you are in the edit mode and able to change any item in a patient's record. Pressing [*] will process the record for any changes in the anthropometric indices and write it to the file. If you wish, you may change the entry in the sex field again and watch the anthropometric indices change again when you press [*].

Just like in data entry, the edit routine will detect probable errors and flash EDIT FLAG. If you get an edit flag, remember it will be written on the record that way until you change it.

1.10 DELETING RECORDS FROM FILE

[D] is for delete.

It works very simply. If you put a record on the screen and push [D], you erase the record. The record doesn't really disappear from the disk. A special byte of information is written to the record and you won't have access to the record anymore. It's still out there taking up space. If you had a file of 1000 records and you deleted the second one, it would take too long to pull all the records back one address to cover the gap.

There are other procedures you can use to 'tighten up' your files after a lot of deleting. Merging or sub-setting files automatically cause records coded as 'deletable' not to be copied. If you do a lot of deleting, be sure to read the section 1.12 Transferring and Subsetting Files.

1.11 MAKING FIELD-SPECIFIC INCLUSIVE AND EXCLUSIVE SEARCHES

Besides being able to separate edit-flagged from non-edit-flagged records, you can search and separate records based on criteria as specific as 'all children from village XYZ who were vaccinated against measles and had w/h percentiles over 70 and were under 60 months old'.

To perform such searches you must escape to the results display selection menu. Press [Esc] if you are still in the edit/review/delete mode. We could press [5] for New search but we are going to use a larger file so we have something substantial on which to experiment.

Press [Esc] until you get back to the main menu.

Press [2] for Edit and Search

We will use a practice file called SAMPLE. It was supplied with your supplemental files diskette. Fetch it from the diskette. By now you should know that you always specify the disk drive along with the file name.

We are going to find all males, between 12 and 36 months who are between the 40th and 80th percentile for height and age, and whose birthweights were between 1800 and 6700 grams at birth.

1. Move the pointer to the sex field and input "M"
2. Move the pointer to the months field and input "12"

After you pressed the return key, you noticed that the pointer stayed at the months field. Because months is a numeric field the program expects a lower and upper range. 12 was the lower range. 36 will be the upper range and you will enter it in the months field.

3. Input 36 as the upper range in the months field.
4. Move the pointer to the h/a percentile area in the indices display bars. Input 40 as the lower range, press return, and enter 80.99 as the upper range. Press return to get the pointer back.

5. Move the pointer to the BIRTHWT field in the extra classification fields area. Notice that the pointer changed to a # sign. This is a reminder that the information in this field is mathematic and will require you to input a lower and upper numeric range.

Type and enter 1800. That is the lower range.

The # sign will stay fixed at the BIRTHWT field. Type and enter 6700. That is the upper range.

6. Press [*].

The prompt [I] INCL [E] EXCL appears near the top right corner of the screen. This prompt asks if your search will be INCLUSIVE, meaning children within or equal to the search specifications or EXCLUSIVE, meaning children outside of or not equal to the search specifications.

For this search to find the patients we've describe the search will have to be inclusive. Type and enter the letter [I].

The disk drive is activated as the program searches each record to see if it fits the search criteria. In the upper right corner of the screen you will see a pair of numbers. The number ascending on the left tells how many records were searched. The number ascending on the right is the number of records found that fit the search criteria.

When the program has searched all the records you will be told how many records fit the search criteria and a menu to select how you want to view the results of the search.

If you choose to look at the records individually or list them out on the printer or the screen, you will see that all the records found are all males between 12 and 36 months old, with h/a percentiles between 70 and 80.99, and weight between 1800 and 6700 grams at birth.

If the search had been [E] Exclusive, the computer would have found children who did not fit within the search described above. Exclusive searching allows the users search two sides of a curve simultaneously.

1.12 TRANSFERRING AND SUB-SETTING FILES

You may put 5000* patients into a single file. You cannot fit 5000 records on a diskette (unless using a 1.2 mb high-density drive) but you could keep that many records in a single file if you had a hard disk.

Rarely are users interested in all their records at once. The typical user, for instance, may only want to concentrate on the abnormal patients or patients fitting a narrow classification.

Using option [4] Transfer Subset to a New File, you can make a new file of the records you've found during a search. We recently identified a number of 1, 2 and 3-year-old males between the 70th and 80th percentile for height-for-age who weighed between 1800 and 6700 grams at birth. We could copy them to a new file very easily.

On the Search Results Menu press [4].

You are prompted to make up a name. Let's call this group of patients A:GROUP1. Indicate a disk drive and type A:GROUP1.

After you press the return key, the program will write a copy of the records that fit this search to the file GROUP1.

This will be a fully independent patient file that can be searched, edited and added to just like the file it came from.

Note: You could use this function to copy an entire file for back-up. To do this you would fetch all records instead of making a field-specific search and then use the transfer function to copy all of the records into another file.

In the section on DELETING RECORDS, it was mentioned that none of the deleted records were ever really erased--just specially coded so you couldn't access them.

***The program was designed to manage 5000 records at a time. Attempts to save more than 5000 records per file may bring unpredictable results.**

They are still taking up space on the file. If you've deleted more than a few records, you can use this function to 'tighten-up' your file. The records coded as deleted will not be transferred to the new file.

1.13 SAMPLE WEIGHTS

In more complex tabulations some users find it necessary to 'weight' their data. In some instances the user determines that a specific group of respondents to a survey are under-or over-represented.

By using sample weights in the ENTRY program you cause a special field to be written on each record. In that field you may place any number that will fit in five columns or less (e.g. .0025, 1.276, 45.1). When the file is cross-tabulated or tallied in some way by another program, each patient will represent one times whatever value is in the sample weight field.

To demonstrate, we will weight the records in the practice file you made earlier.

Make sure you are at the main menu, then press [4].

When you receive the prompt for a file name, enter the disk drive designator and the name of the file.

The first thing you are told is that your practice file has no provision for weighting data. This is nothing to be concerned about. This means there is no sample weight field in the records at present, but you can have one.

Press [2] to create the sample weight field.

Immediately you have your field. There is another prompt in the window. From here you could [1] go back to the menu or [2] continue which means now that you've established a sample weight field you could apply the weights.

There are two ways to weight your data. You could go back to the menu, use the edit function and individually weight every record one at a time. Now that the field exists as part of your record, it would be very easy to edit in the numbers just as you would edit any other field. This would be a very slow process.

The second and most efficient way is to "describe" to the program which records require weighting and what that weight should be.

Press [2] to continue.

You are given three choices:

1 apply one weight to all records.
2 apply weight to selected records
3 change old weight to new weight

With selection [1] you can give all the records the same weight. To demonstrate this function we will apply the weight of 1 to all records.

Press [1].

You are asked to enter what weight is to be applied to all records.

Type and enter '1'.

The disk drive activates, and you may actually hear the head writing the number 1 to the sample weight field on every record. We will look at the results as soon as the computer finishes writing on every record. The menu will return after the process is complete.

When the menu is back on the screen use the edit function [2], give the name of the practice file, call in all your records (by pressing [*]), and select [3] to view every record individually.

Look in the user-selected classification variables. The last field is '@SW'. @SW is a very specific label. The computer is programmed to recognize only that label as a sample weight field. Notice that every record has a '1' in that field.

Go back to the menu.

Select [4]. We will demonstrate the third function in sample weighting. When you get the prompt for a file name, reply with the practice file you just finished weighting.

Select [3].

With this function you must search for the kind of records you plan to weight. To specify the records, you use the same methods used to perform a field-specific search. For a simple example, assume you want to apply the weight of .75 to all males.

1. Move the pointer to the SEX field. Type and enter 'M'.

2. Press [*].

Just as in field-specific searching previously demonstrated in this chapter, the search for males or any other characteristic could be inclusive or exclusive. You have to specify inclusive by typing and entering [I].

After finding all the males in the file, you are prompted to give the weight to be applied.

3. Type and enter .75

The program proceeds and writes '.75' in the sample weight field of the records of every male in the file.

You can confirm this by using the edit function to view each record individually.

The second feature in the sample weight function is changing one sample weight to another.

First, go back to the main menu and select [4] again. Reply to the prompt with the practice file name and select feature number [2].

The first prompt asks for the old weight--that is the number it's suppose to find and change. For this example type and input '1'.

The next prompt asks for the new weight. In other words, the computer wants to know what it should change '1' into. Reply with '1.333'

After you have answered that prompt, the program will tell the computer to search for and change all '1's to '1.333'.

Sample weighting has no influence on the computed anthropometric indices. The only time the sample weight influences any process in the package is in cross-tabulation. The TAB program will detect weighted data. During tabulation, each record that fits the cross-tab will not be counted as one, but instead it will count for the value in the sample weight field.

There are a few considerations you'll want to remember when using sample weights.

1. Sample weights are an extra variable counted as one of the 24 allowed user-selected classification variables. You cannot have a sample weight field if you've already created 24 extra variables.

2. You can create a sample weight at the same time you create your other extra fields by just using the label '@SW'. Remember that the 'S' and 'W' must be capital letters.

3. Once you've created a sample weight field you are obligated to give every record a weighted value--even if that value is '1'. Without a value in the sample weight field, the record counts for zero during cross-tabulation. It is recommended that all weighted records be assigned a value of '1' before weighting any specific groups.

1.14 MERGING FILES

Merging files in the entry program is the process of taking several like files and tying them together. If several people are working on the same project from different work stations, it might be useful to merge the separate files before cross-tabulating them.

You may never need to merge a file but the function is there if you do.

To begin merging files, press [5] on the main menu.

A window will appear. In this window you must list the disk drives and file names of the files you intend to merge. After the last file name has been entered, press [*] to go to the next step.

The merged files will be kept under a different name. The next prompt asks what you wish the name of the merged files to be. YOU CANNOT USE ONE OF THE NAMES OF THE FILES YOU ARE MERGING FOR THE NEW NAME.

As the files are merged, you see two numbers appear beside the file names in the window. The left number is how many records are coming from the file being merged. The right number is total number of records written to the new file.

If any of the files you've listed are incorrectly named or non-existent, the process will stop and prompt you to make a correction.

If your merger causes more than 5000 records to be written to the new file, an overflow file called TEMP.ETB will be automatically created. TEMP.ETB will contain the excess. It is important that you change TEMP.ETB's name since it will be over-written next time this situation occurs.

If you would like to test this function, make a few files with a few records in them and merge them.

CHAPTER II USING THE TAB PROGRAM

What can this Tabulation program do?

Cross-tabulations are the documents that support and illustrate the health professional's conclusions. BATCH and ENTRY are basically the tools for collecting the research data. TAB is the program that presents the findings of that research.

The Anthropometric Statistical Package will make a three-dimensional cross-tabulation using any three variables in an ETB file created with BATCH or ENTRY.

Because of the considerable amount of time involved in making a cross-tab, Version III has been enhanced so that the user may:

1. save the results of a cross-tab and present it or print it as many times as desired.

2. create a library of frequently used tabulation intervals.

3. create a cross-tab, replace any variable, change any interval and rerun as new cross-tab.

4. batch-load up to ten ETB files for one cross-tab.

Completely new in Version III is the capability to present the results of a cross-tab as a bar graph or histogram as it is often called. The histogram is a feature recommended only for users with color monitors or monochrome monitors compatible with color option cards.

Every effort has been made to make Version III of the Tab program the most flexible and user-friendly program in the package. In order to get the most out of the program you must carefully follow the instructions provided in this chapter.

2.1 PLANNING A CROSS-TABULATION

To start the program, boot up the computer, insert the Program Diskette (if you're not using a hard disk copy) and enter the word **TAB.**

The principle product of the TAB program is a three-dimensional cross-tabulation. It's called three-dimensional because the tabulation has width in the form of columns across the page, height in the form of rows going down the page, and depth in the number of pages to the tabulation.

You cannot pull a cross-tab out of nowhere. With 16 permanent variables and up to 24 extra classification variables, there are hundreds of possible combinations for a cross-tabulation.

It may be helpful to rough draft your cross-tabulation on plain paper before you start choosing variables and defining intervals.

Each dimension represents one of three variables you choose from a patient file. Each variable is then broken into intervals. The second element in a cross-tabulation are the intervals of the three variables.

The maximum number of intervals depends on what dimension the variable occupies. The allowed number of intervals for each dimension are:

1. COLUMN (maximum 8 intervals)
2. ROW (maximum 6 intervals)
3. PAGE (maximum 14 intervals)

There are a few steps to making a cross-tabulation. You must first:

STEP 1. Prepare a "template" of the variables and their intervals.

STEP 2. Choose what file(s) will be cross-tabulated against the template.

STEP 3. Perform the tabulation and save the results.

STEP 4. Display, print or graph the results.

Provided with the supplemental files diskette is an ETB file called SAMPLE. This will be our practice file.

2.2 THE INTERVAL TEMPLATE

You must start with STEP 1, choosing variables and making intervals. Variables don't come out of thin air. Any variables in an ETB patient file can be used in cross-tabulation but the program has no way of detecting what those variables could be if it has never "seen" the file.

While the menu is present (it's present whenever the word SELECT is blinking), press [5] Create or Edit Tab Intervals.

You are prompted to give the name of an ETB file so the program can "look" at the available variables. Reply with A:SAMPLE (or B:, or C: depending where you put the package's supplemental files diskette).

Notice the windows descending across the screen. You are told how many variables are available to you and that there is one window each for COLUMN, ROW, and PAGE.

First, let's look at the variables. You can do this by pressing the spacebar or the backspace key. DO NOT PRESS ANYTHING ELSE BUT THESE TWO KEYS. You will see all the variables in the file appear one at a time. Keep using the spacebar and backspace keys until you have seen all the variables.

We are going to choose the first variable. Press the spacebar or the backspace key until you reach the variable **AGE.**

Press the return key. You have chosen AGE as your Column variable. The pointer has moved to the Row window.

Just as you did in the Column window, find the variable **W/HZile** and press the return key.

Just as you did in the Row window, find the variable **STATE** but do not press return.

You have chosen the three variables. To register these variables, press [*].

The pointer is now in the window for the Column dimension. Hold your finger down on the return key and watch the pointer move through the windows. The pointer can move to all the windows but will make only eight stops in this window--meaning you can have eight intervals of AGE.

Type and enter EXACTLY AS SHOWN

0-23(Under 2)
24-35(2 yrs)
36-47(3 yrs)
48-59(4 yrs)
60-71(5 yrs)

Press the enter key until the pointer is in the Page dimension window.

Type and enter

0-5.00(5th -)
5.01-94(NORMAL)
95-100(95th +)

Move the pointer to Page dimension.

Type and enter:

A(NY)
B(PA)
C(NJ)
D(VT)

For reasons of privacy the true identity of states A, B, C and D have been deleted. NY, PA, NJ, and VT are not the actual states. They are merely examples. You may use other initials for states.

Press [*]

Another window appears near the bottom of the screen. You are given the opportunity to choose one of the numeric variables in the file for calculation as an accumulated mean.

You may use the spacebar or the return key to see all the variables in the patient file that are numeric (including the user selected extra variables).

When you have reached BIRTHWT (birth weight) stop and press [*].

What we have done is tell the program to read the birth weights of every patient fitting the cross-tabulation and, in background, calculate the average birth weight for every cell.

The next prompt asks whether or not we want to include the sample weight factor in counting the patients. This window appears only if there is a sample weight field in the patient file from which you selected to choose the variables.

Your reply to this question will be [Y]

The menu returns. We have made the cross-tabulation template. This was a simple example but it could have been a very involved procedure. You could have had eight columns, six rows and 14 pages. It would not be pleasant to make one of these templates every time you needed to cross-tabulate a file. Therefore you should save it.

Press [1] Save Tabulation Intervals.

You are prompted to provide a disk drive and name. Since the file we read to get the variables was called SAMPLE, we shall also name the interval template SAMPLE. The SAMPLE template will automatically be assigned the extension INT. DO NOT USE EXTENSIONS OF YOUR OWN.

Before we continue, we are going to go back and look at the intervals you entered. We are going to study the rules of their structure because the only way you will ever get the cross-tabulation you want is to know how the program will interpret what you have entered.

Let's dissect the interval 0-23(Under 2)

It has three parts: the bottom range [0]
the top range [23]
the user label [(Under 2)]

the bottom range and top range are always separated by a hyphen. Since age is a numeric field it requires a bottom and top range.

The program interprets 0-23 as 0 thru and including 23.99. The program will apply .99 to the top range of all whole number numeric intervals.

The program will apply .09 to the upper range of all numeric intervals that are precise to the nearest tenth (e.g. 25.3) would be interpreted by the program as 25.39.

The program will not add .99 or .09 if the top range of an interval that is explicitly precise to two decimals (e.g. 12.55 will be interpreted as 12.55 because of the two decimal precision). 23 would be interpreted as 23.99 but 23.00 would remain 23.00 because the user was explicit about the decimal precision. The bottom range in this interval is not changed.

The program recognizes the mathematical operators <, >, =>, and <=. There is important information in Appendix A about the effects of using these operators. You don't need to read it now, but YOU WILL NEED THE INFORMATION IN APPENDIX A FOR FUTURE REFERENCE. COPY IT!

When making intervals for non-numeric variables such as SEX and the character-type extra variables, you must use intervals that precisely match the information you are looking for.

In our SAMPLE file we are looking for STATE and the intervals are A, B, C, and D. We didn't type the actual state names because none of the state names are in the fields for STATE in any of the records. A, B, C or D are the letters in the STATE field so that is what the intervals should be. The same would apply to SEX. If you have 1 and 2 in the fields for SEX then that is what you will type as the intervals. You may also apply a descriptive label if desired such as 1(Male) and 2(Female) so that the report will read more clearly to someone else.

How can you be sure you'll get what you asked for?

Typing those intervals and getting them right the first time is going to take experience. There is a way to check and see precisely how the computer interpreted your intervals.

Press [5] Edit and Create Intervals.

The variables and the intervals come back on the screen.

At this time TURN ON YOUR PRINTER. Press [?] and you will see a printout of the ranges the computer is using to tabulate the patient records.

It is not a particularly neat printout. It's a raw unformatted technical readout of exactly what values the program is using. You must read it as is.

2.3 EDITING INTERVAL TEMPLATES

Any interval template can be edited. If it is already in the computer's memory you can [Esc] Escape back to the menu, press [5] and re-edit the template. If you don't know whether or not there is a template in working memory, look at the menu. If there is a ninth selection then there is a template in the computer's working memory.

You can recall and edit a stored template, too.

1. First you'd have to press [9] (Clear memory) and recall the interval template by name with selection [3] on the menu.

2. Once the interval template is in memory, press [5] to display it. Use the spacebar to change the variables, press return to select them and [*] to move to intervals. You can re-type the intervals or delete them with spaces.

3. After you have edited the template, you can save them under the former name or make a new name.

2.4 PERFORMING THE CROSS-TABULATION

If you have made a useful template you are ready to perform the cross-tabulation. The template should be in working memory. If not, recall the template SAMPLE with selection [3].

Press [6] Perform a cross-tabulation.

The window for listing files will appear on the right side of the screen. Here you may list up to ten files. With each file having a capacity of 5000 records, you have the potential for cross-tabulating 50,000 patients.

We have one file. Therefore we would enter A:SAMPLE on the first line. After you press enter the pointer will move down to the next line. Since we have no more files, press [*].

The screen changes to a display similar to a spreadsheet. It is a spreadsheet of sorts. But with this spreadsheet you can monitor the results of a cross-tabulation in progress.

Press the spacebar.

Notice that numbers are rapidly appearing or changing on the screen. This is page one of the cross-tabulation. Press the press bar again and you will see the progress on page two. Press the spacebar everytime you wish to monitor a different page of the cross-tabulation.

The cross-tabulation is complete when **PAGE** appears in the lower left corner of the screen.

There are four percentage calculation modes. They are PAGE, ROW, COLUMN and GLOBAL. The mean birthweights we collected can be seen through the MEAN mode. A list of the pages in the tabulation can be seen in the LIST mode. In all you have six selections: [P], [R], [C], [G], [M] and [L]. Pressing the first letter of each selection name calls the selected mode.

Press [R] (to select the ROW mode) then press and enter [1] (for page 1).

The numbers appearing on the screen show the number of patients in each cell. The percentage numbers show the distribution of patients by row. Each row stands alone as 100 percent.

Press [C] then press and enter [1].

The percentage of distribution is now shown by column.

Press [P] then press and enter [1].

The percentage of distribution is now shown for the entire page.

Press [G]

Note that the word GLOBAL is in inverse video.

Press [2] and enter.

Notice that we are on page 2 and the percentage of the final total is not 100 percent. It is less than 100 percent because the calculations were made based on a total which included every patient from all pages. This a GLOBAL TOTAL of page 2.

Press [M].

MEAN BIRTHWT is in inverse video. Select any page number (1, 2, 3 or 4) and press enter.

You will see the average birth weight for each cell.

You could go on indefinitely comparing the Global and Page totals and the mean birth weights of these patients. BUT YOU SHOULD SAVE YOUR RESULTS FOR SAFEKEEPING AS SOON AS POSSIBLE. You can escape the cross-tabulation spreedsheet by pressing the [Esc] key. [Esc] brings back the menu. Pressing [4] on the menu will bring back the spreedsheet.

2.5 SAVING CROSS-TABULATION INTERVALS

Press [2] save tabulation results.

A window appears and you are prompted to provide a name for the results. You could call the results any eight-character name you choose but for the sake of continuity you may consider giving the results the same name as the patient file. The extension TAB will be assigned to the results file so there will be no chance of confusion with SAMPLE.ETB, the patient file and SAMPLE.INT, the interval template.

2.6 PRINTING THE CROSS-TABULATION

The cross-tabulation you've performed still sits in memory until you press [9] Clear memory on the main main menu. You should always save the interval template and the results of the tabulation.

Press [7] Print/Graph cross-tabulation

If there are no results in working memory, you will be prompted to retrieve some from your library of tabulation results. Tabulation results have the extension TAB but you must not include the extension when answering the prompt.

You will receive a print menu.

The report may be printed to show relative percentages several ways.

1. [C] The Column Mode. Each column total is 100%. The percentages for the cell figures are relative to the column total.

2. [R] The Row Mode. Each row total is 100%. The percentages for the cell figures are relative to the row total.

3. [G] The Global Mode. N = all the patients in the report. The page total percentage is always a portion of the whole population.

4. [P] The Page Mode. N = the total number of patients on a given page. The page total is always 100% and all figures on that page reflect each figures percentage of the page total.

5. [M] The Mean Mode. The number of patients in each cell is accompanied by the mean of whatever numeric variable was chose for mean calculation.

6. [H] Histogram. Histograms are available but will be discussed in greater detail in the next chapter.

You may select any of these modes (except Histogram).

The user may select a two-line title. The title is optional. If no title is desired, press [ret] key. To delete a previous title (titles are saved with tabulation results), input one space. You cannot change just a part of the title. Titles must be completely retyped for any changes.

Footnotes may also be added to the bottom of the printed report. You are allowed three lines for footnotes. If footnotes are not desired, press return three times. To delete previous footnotes (footnotes are saved with tab results), input a space for each line before pressing [ret] key. You can change one line of the footnotes without affecting the other lines.

If the tabulation results you are going to print come from the working memory and have not been stored, the program will automatically store your results in a file called TEMP.TAB. You can change the name of TEMP later (See your DOS manual). If you do not change the name of TEMP.TAB it will be overwritten the next time you attempt to print a cross-tabulation that was not saved.

If you change or add a title or footnote, the program will update your results file.

The printing begins automatically. Make sure your printer is on. To cancel the printout press [Esc].

CHAPTER III MAKING AND USING HISTOGRAMS

Histograms--or bar charts---are another way of viewing cross-tabulation results. Your cross-tabulation printouts present cold precise calculations. Histograms fall in the realm of presentation graphics. Presentation graphics are often used to present data symbolically--in this case as a bar chart. If you have no interest in histograms, you may skip this chapter. None of the information is necessary in using other parts of the software package.

NOTE: Graphic capabilities for this version were given a low priority. It takes a lot of memory to perform graphics. Not knowing how useful the feature would be, two memory-saving measures were taken:

1. There is no output to ink-pen or dot matrix printers and plotters.
2. There is no adaptation for monochrome monitors.

The histogram feature is for desk top presentation use only. If the function becomes popular among users, direct plotter compatibility may be developed in future versions.

3.1 WHERE TO BEGIN

Before you can make a histogram, you have to have something to graph. You need data. The data comes from the results of a performed cross-tabulation. The results may be either stored among your [.TAB] files on disk or still in memory from a recently performed cross-tabulation. To speed things along, we have provided a set of tab results with your supplemental files diskette. Pay close attention to what you are doing in the demonstration and follow the instructions. You will need to draw on your experience with this practice when graphing one of your files.

3.2 FETCHING THE DATA

Get back to the main menu. Clear memory with selection 9 (if 9 is present). Press selection [7].

You'll be prompted to identify the tabulation result file you wish to print or graph. Your answer will be A: (or whatever disk drive) DEMOTAB.

After the computer has fetched the file, you will get a report overview. The overview is a very general summary of what's in the results. It reads that these results came from a tabulation performed on a file of U.S patients. It shows the three variables in the report and how many patients fit the cross-tabulation; approximately 2.9 million. The data were weighted. No one actually interviewed nearly 3 million patients. Also it shows that a mean W/H percentile was collected.

ST-200
N86
04678
COMMUNITY HEALTH CELL
LIBRARY AND DOCUMENTATION

The next prompt asks whether or not you want a printout or a histogram. Before we jump into making histograms, we need some cross-tabulation printouts for reference.

Answer the prompt with [P] and enter.

The printout menu appears. You will need the [C] Column Total printout and the [M] mean H/A z-score printout. They are four pages long. Turn on your printer and make them both at this time.

When you have the printouts, study them briefly. You'll see that the report is about the distribution of W/H z-scores in children up to 5 years old in four different states. The Mean printout shows the average W/H z-scores for each of the cells.

Keep the printout close by. You will need them for reference.

3.3 MAKING THE HISTOGRAM

When the program has returned you to the menu, press [7] again.

This time, select [H] from the menu.

The screen changes immediately to the bar graphics display screen. The only thing on the screen is a percentage scale on the left edge and a small menu in the lower right hand corner.

The tabulation results data is in working memory in a three-dimensional array. But this bar graph is only two-dimensional. It has only an x and y axis. You'll have to choose the two variables for each axis yourself.

Press [1] Set axis.

The variables (or three sides--row, column, page) in the array are presented. The first axis to be chosen will be the X axis. The X axis is for the labels that go along the bottom of the screen.

For this demonstration, choose [3] STATE.

The next prompt asks for the variable you want to be the bars (the Y axis). For this demonstration, select [1] W/H z-score.

You may see the word STANDBY flash briefly. This is not something you should be concerned with. The program is letting you know there will be a delay while it shuffles around some figures internally.

The only variable left is AGE which was not chosen for either axis. You'll notice in the printed reports that age was broken up into five groups. We may chose to look at one of the age groups or combine the age groups just as the printed reports do on the bottom line of each page.

For this demonstration, we will [2] combine the ages.

Now that you've positioned all the variables you must select the intervals you wish to see.

Press [2] LABELS.

Since we chose STATE as the X axis, we are presented with a list of available states. There are four, but let's look at only the first three; Georgia, Texas and Alabama*.

You must choose the three states one at a time by corresponding number. Press [ret] after each choice. To indicate that you have finished choosing, press [ret] without inputting a number.

Notice that as you chose each state, its label appeared along the X axis.

After choosing all the three states, you are presented the list of intervals in the W/H z-score variable. We are going to choose six of seven groups: 1,2,3,5,6 & 7.

Enter each number one at a time. Press [ret] alone when all six groupings are chosen.

A color code appeared as you selected each label. The numbers accompanying them are showing how many thousand of children fit into that label (e.g. The total number of children equal to or below -2 W/H z-score is 10,000 for the 3 states selected).

You may have also noticed a working title at the top of the screen. It shows that you are comparing the W/H percentiles in three different states combining all the age groups. The number 2982 followed by the letter K indicates thousand (2.9 million children).

WHENEVER THE TOTAL NUMBER OF CHILDREN IN A REPORT EXCEED 1 MILLION, ALL NUMBERS ARE REPRESENTED TO THE NEAREST THOUSAND.

Press [4] EXECUTE.

EXCUTE means 'draw something.' Now it should be very clear what you have been building up to. Each grouping of bars stands alone as a 100% population--just like your Row Total printout. The percentages represented on the screen are the same percentages you will find on the bottom lines of the first three pages in the report.

*The states are hypothetical.

The bars aren't very tall. None exceeds 40%. A 100% scale is more than you need. So let's change the scale first.

Press [5] SCALE.

You have a list of five pre-set scales or you can make your own with selection 6. Let's choose our own.

Press [6].

The prompt %age? wants you to provide a number. You may give any positive number less than 100. Type '40' and enter.

The scale changes immediately with 40% now as the top. The screen is blank.

Press [4] EXECUTE

The bars are back and with a finer definition in the difference in percentages. Compare the figure on the Row Total printout with the bars on the screen.

3.4 PRECISION BAR HEIGHTS

The precision of the bar heights may sometimes be a little off. Remember that the printed report's percentage are rounded to the nearest tenth. The bar percentages are rounded to the nearest 1/40th of the scale.

The general rule is that the smaller the scale the finer the precision. It is difficult to get satisfactory precision in the bars that rise only to 8 percent when they must share the scale with bars that rise to 80 percent.

It is always best to concentrate on the bars that are closely related in height. Let's re-label the histogram so we can use a smaller scale.

Press [2] LABELS.

This time we will choose all four states for the X-axis labels and only the equal to and less than -1 W/H z-score for bars. We are only using one bar this time.

Press [5] SCALE.

Our scale will be 20% which is a pre-set percentage.

Press [4] 20%

After the scale has changed press [4] EXECUTE.

You could even go to a smaller scale if you want, but by now you should understand the effect choosing different scales has on accuracy.

3.5 N-SHIFT

N-SHIFT is for number shift. With N-SHIFT you change the N number. Right now we are in a combined ages mode. All the ages of the children were taken as a whole. But we do not have to look at all the age groups combined.

We can concentrate on one of the age groups.

Press [3].

The program asks if you want to combine the ages or select an interval. We are going to select an interval.

Press [1].

The list of intervals is presented. For this example we shall choose the one-year-olds.

Press [2], the corresponding number for the one-year-olds.

We have focused the histogram on the 958,000 one-year-olds in the cross-tabulation results. You won't see the new N number unless you press [2] and put out the labels again.

We ought to clean up the screen since we are starting fresh. You may press [L] or [H] for low or high intensity. Low intensity draws grey lines on the screen. Some monitors can't draw that color so there is high intensity which draws light blue lines across the screen. The intensity setting may be used at anytime to clear the scale.

You may select the labels again for the two axes.

Select the first three states, choose all the W/H z scores except ZERO and then execute. Histograms only seem difficult to make the first few times. If you stop and think about it, the menu is almost a step-through guide.

3.6 SUMMARY OF STEPS

The steps to making a histogram are:

1. SET AXIS. Decide what two dimensions are going to the X and Y axes. Choose the X-axis first. The program will ask you to choose the bars--which is the Y-axis. Then the program will ask what you want to do about the unchosen dimension---combine it or choose one of the intervals.

2. LABELS. You are just choosing the intervals. Remember that you're limited for space on the screen. The more X-axis intervals you select, the fewer bars you can have.

3. EXECUTE. Draw the bars. You can always execute as long as you have selected axes and set out the labels.

4. SCALE. If after drawing the bars you decide that you don't need a 100% scale, you simply choose another.

5. N-SHIFT. You are choosing to combine or interval the third dimension. The N number at the top of the screen changes accordingly.

3.7 GRAPHING THE MEANS

Since we have the capability to collect the average of any numeric quantity on a cell-by-cell basis, we, appropriately enough, have the capability to plot a histogram of those means.

Press [6] MEAN PLOT to plot the W/H z-score groups.

The only difference in drawing bars for means and drawing bars for percentages is the scale. We aren't using a percent scale. We are using a range. Since the mean we collected was on W/H z-score, our top range would be the reasonably expected highest z-score in the results.

Let's choose the W/H z-score of 4 as the upper range.

Type '4' and enter.

Now we have to choose a bottom range. So that scale will be symmetric, we'll choose -4 as the low W/H z-score.

Type and enter '-4'

Notice the scale. It ranges from -4 to 4. In the middle is the zero line. That line is brighter than the others.

We would select AXIS and LABELS the same as we did when we chose percentages.

Choose the three previous states for the X axis. Choose the six previous W/H z-scores as the bars. Combine all the ages. Then [4] EXECUTE.

Just like when you chose different scales for percentages, you can choose different ranges for Mean Plots. This time set the mean plot scale to 3 at the top and -3 at the bottom. The screen will clear and you can execute right away.

Histograms are pleasing illustrations of numeric data but they can be difficult to make, especially when you're trying to represent figures on your printed cross-tabulation. For what it's worth, the histogram is never wrong. Many times what is really wrong is your interpetation of what you're seeing. You must practice and experiment with the histogram function. Repeat this chapter again but use different axes.

CHAPTER IV USING THE BATCH PROGRAM

What is Batch processing?

Batch processing is generally described as the method of processing several pieces of formatted data from a disk or tape reptitiously rather than interactively with a human operator entering one record at a time.

This is the way our bank statements and telephone bills are processed--in large quantities in repetition--hence the term 'batch'.

Batch will serve the user who has a database of unprocessed records and only needs a way to process then. Batch may seem a bit more difficult to follow for the inexperienced user. The typical Batch user probably has access to a mainframe computer, uses other file management software and perhaps knows a great deat about personal computers.

Batch file data on a diskette is easy to spot. The file "DEMOBAT.TXT" on the Supplmental Files diskette provided with this package is a batch file. You can actually see the data in the file by putting the documentation diskette in drive A and entering.

TYPE A:DEMOBAT.TXT

What you'll see are long strings of numbers. Some batch files have numbers and letters. The records scroll down the screen in a pattern because there is an embedded code for carriage return and line feed at certain intervals. This embedded interval marks the break point between records. Because you are able to make this file appear on the screen so easily, programmers usually call this a text file, or an ASCII (pronounced "ASKee") file--ASCII stands for American Standard Code for Information Interchange.

If you've purchased and used DBASE or LOTUS, you've probably made several ASCII files. A systems analyst or programmer could assist you in getting ASCII files from a mainframe computer. The point is, that if you've got an ASCII file that has at least height, weight, sex and enough information to derive the age of a patient, the BATCH program can read it, process it for anthropometric indices, attach the results and write it to another file.

4.1 PREPARING RAW DATA FOR BATCH PROCESSING

Batch files come from a mainframe computer or from a PC program. Where you get the data isn't important. Getting the batch file on a diskette so you can get it into the PC is important.

The user who moves data from another system or "host" computer to a PC is "downloading." One who downloads must already have the means of making this data transfer. No program or procedure in this package can perform a system-to-system link.

Data already on a diskette must meet certain simple criteria:

1. It must be a columnar ASCII file.
2. It must have the standard two-byte end-of-record marker CHR$(13)+CHR$(10) so the program knows where one record ends and the next begins.
3. A standard EOF mark (CHR$(27) or CHR$(26)) must be at the end of the file.

These are not complex requirements. Most off-the-shelf PC database programs meet these requirements and virtually all mainframe data is stored in ASCII with the appropriate file markers.

4.2 PROCESSING THE RAW DATA FILE

From this point the raw data will be called the SOURCE DATA. The product we are making will be called the DESTINATION FILE. You will need to boot up your computer provide a formatted diskette for your destination file.

After the computer has booted up and DOS is present, type
BATCH and [enter].

Shortly after typing BATCH, you will briefly see a red box introducing the program and telling you there will be a slight delay for loading the data table that calculates the indices.

You will see a menu appear after the data table is loaded. When this has occurred you may remove your program diskette and put it away for the rest of the session.

YOUR MENU CHOICES ARE:

1. Examine the Batch File
2. Retrieve Field Location Template
3. Create Field Location Template
4. Calculate and Convert Source Files
5. Escape to DOS

As your understanding of Batch grows you will use all these choices. For the sake of this exercise, we will concentrate on four tasks:

Task I.
Examine a source data file to determine if it is suitable for batch processing.

Task II.
Create a field location template and use it to see if we can trial process a typical record.

Task III.
Convert a batch file into a specialized "ETB" file compatible with the Anthropometric Statistical Package

Task IV.
Convert a batch file into another standard batch file for use on mainframes or with other statistics programs.

To succeed you must follow only the instructions given. Don't anticipate procedures.

4.3 EXAMINING SOURCE DATA FILE

Step 1.

Press [1]

At this point the screen changes. A four-line scale appears near the top half of the screen and the lower half is dominated by a bright blue field and a red window.

In the window there is the prompt:
DISK DRIVE AND NAME OF SOURCE FILE?
EXCEPTED FORMAT B:YOURFILE.EXT

Your reply will be: **A:DEMOBAT.TXT [return]**

NOTE: A and B are disk drives. Throughout this exercise you will use whatever letter represents the disk drive you are using.

Appearing in the same window, the next prompt is:
RECORD LENGTH?

Your reply will be **[return]**

Pressing [ret] means that you don't know what the record length is and you want the program to determine it by searching the first record for an end-of-record mark. There could be many occasions when you aren't well-acquainted with the source data and the record length is not certain. Never 'guess' or 'estimate' a record length. You either know it or you don't.

You'll notice that the computer has displayed a string of numbers along the lines of the scale. The scale is tick-marked in 5 and 10-byte increments. This is a graphic representation of the first record in the DEMOBAT.TXT source file. The computer reports that the record is 246 bytes long (the term 'bytes' is preferred over 'columns' when speaking of disk and tape files).

The blinking colon in the blue box is the actual location of the end-of-record marker. For our purpose it is not counted as part of the record's usable length. The program, however, does recognize it and processes the record's length as 248. This offset need not concern you.

Just because the computer reported that the first record is 246 bytes long doesn't mean that every record in the source file is of equal length. For the program to work all the records must be the same length. You **must** check a few records at random to make sure that they are all the same length.

Step 2.

There is a prompt on the screen that asks if you would like to check another record. You may enter another number: Let's say 45

Doing this causes the program to pick the 45th record -- or what is hoped to be the 45th based on the premise that every 246 bytes (plus 2-byte maker) is the start of a new record.

The computer reports back that this random record is indeed 246 bytes just like the first record. Had any of the records before 45th record been shorter or longer than 246, the synchronization would be off and the computer would have "landed" in the middle of a record. Some number other than 246 would have been reported in that event.

ANYTIME THE PROGRAM CHECKS A RANDOM RECORD THE LENGTHS MUST ALWAYS BE THE SAME. IF NOT, YOU CANNOT PROCESS THE FILE.

This information, though fringing on the technical, is useful. When you start to examine your own data files for suitability you will need to repeat this task to assure yourself that your file is uniform. Uniformity is fundamental to batch file processing.

4.4 IDENTIFYING FIELD LOCATIONS

In order to perform the functions in this task, you must have already examined batch file and fetched a record to the screen.

In this critical step you tell the computer the column locations for various pieces of information within the record. You'll prove that you column locations are correct by test processing a record.

Step 1. Press [3].

The first fields you must identify are the ones necessary to compute the anthropometric indices. A large window appears and displays the fields the program recognizes when processing a record. It will not be necessary to fill out every field. At a minimum the computer must know the patient's:

SEX: recognizes M and F or 1 and 2 as male and female respectively.
AGE: computes from birth and visit date or age in months.
HEIGHT: recognizes HT cm or feet and inches.
WEIGHT: recognizes WT kgs or pounds and ounces.

Notice the triangle-shaped pointer. This pointer can be moved from field to field by pressing [return] and [backspace]. This way the user can skip around and correct mistakes easily.

Using the figures below, fill out the field locations for this record. If you make a mistake, press [return], the backspace key (NOT left arrow), and type over the error.

I.D. 1-11
SEX 59
VISIT DT (yr) 45-46 VISIT DT (mo) 41-42 VISIT DT (da) 43-44
BIRTH DT (yr) 55-56 BIRTH DT (mo) 51-52 BIRTH DT (da) 53-54

HT cms 87-90,1
WT kgs 84-86,1

The computer monitors your inputs. You'll hear a beep if you input a negative range like 56-44. If you type in an incorrect field, go back to that field and void it out with a single slash [/]. The slash causes the computer to ignore the field.

You may have noticed as you entered your field locations that data in the record display was highlighted. On monochrome monitors the highlighting is in inverse video. On color monitors:

FIELD	COLOR
I.D	White
Sex	Blue
Visit	Magenta
Birth	Red
WT cm	Green
HT kgs	Yellow

In fields followed by ",1" (as height and weight are) the computer will assume one decimal place. ",2" would indicate two decimal places and so forth. The program will also accept data that already has decimal points in the record. Don't use the comma extension on fields that already have decimals in the data.

Step 2. Press [*] asterisk.

The window reappears except there aren't any field labels. This is another field location routine except this time you will indicate User-Selected Fields. User-Selected Fields are extra classification fields beyond those needed for anthropometric computation. You may have up to 24 of these extra fields. In our example we are using only seven.

Step 3. Create extra variables from the table below.

```
HOMESTD  (for homestead)
#HHSIZE  (household size. # mean this a numeric field)
INCOME   (main source of income)
POLIO
DIPTH
MEASLES
VACERT   (vaccination certificate)
#SIBS     (brothers and sisters)
@SW
```

Press [*] to quit.

Notice that two of the fields have a '#' in front of them. This means that this field is for numeric data and that mathematical calculations can be made on whatever numbers are in that field.

The '#' sign will not be saved with the label but the program will still treat it as a numeric data field.

The other unusual label is '@SW'. That is the exclusive label for sample weight. It tells the program that anything in this field is to be the patient's true count in a cross tabulation. (see: SAMPLE WEIGHTS, Chapter 1). IF EVER YOU HAVE SAMPLE WEIGHTS IN YOUR BATCH FILE YOU MUST USE '@SW' AS THE FIELD LABEL.

Now that you have identified or "labeled" the fields, you must put in their field locations.

Step 4. Follow the information below. Type in the field location data beside the indicated label.

```
HOMESTD   7-9
#HHSIZE   67-68
INCOME    222-223
POLIO     78
DIPTH     79
MEASLES   80
VACERT    76
#SIBS      224-225
@SW       171-175
```

Press [*] to quit.

Step 5. Press the [ret] key.

Notice that another window appears quickly followed by the computed anthropometric indices. What has happened is that the computer performed a test computation on the on-screen record based upon the information you supplied in the field locations.

The information has not been saved anywhere but it serves as a way to tell you how the computer interpreted the data in the record. This is the way it will interpret all your records. If anything was wrong, this would be the time to fix it. When you process your own source data you will need to use this task as a guide.

Step 6. Press the [ret] key.

Another prompt window appears.
This time you are told: **THESE FIELD LOCATIONS WILL BE SAVED**
ASSIGN DISK DRIVE AND A NAME

Any name up to eight characters is acceptable but for the sake of simplicity and easy coordination, you may use the same name as the first eight characters of your source data file.

For this exercise your reply will be: **A:DEMOBAT** There is no default name for a template but it is still recommended that you name it after your source file.

You have just made a small file of the field location data. This little file is called a "template." It contains the field locations, the decimal placement and the symbols for male and female. Making a file like this will save you from having to re-enter field locations the next time a DEMOBAT.TXT batch file is processed.

Even though the batch file is also called DEMOBAT, the computer knows the difference because it attaches the extension ".FLD" to the template you just made. Therefore:

IT IS IMPORTANT THAT YOU NEVER MAKE EXTENSIONS OF YOUR OWN !

The program will not be able to find your templates if you start applying extensions to your template names. Only the batch file that was downloaded or created by some other PC software is allowed to have its own extension. The program assigns all other extensions.

4.5 CONVERT AND COMPUTE SOURCE FILE

In the previous two tasks you examined your source data file for suitability; made a template of the field locations, including those of seven user selected variables; and submitted a record to a test computation. Now that you are sure one record will process correctly you should be able to process the whole batch file.

Step 1. Press [4].

The screen has cleared and changed again. There is a coral colored border at the top of the screen and a yellow prompt window presenting you with three choices.

Choices B and C are for making batch files into other reformatted batch files for use in other systems. Choice A is for making ETB patient files compatible with ENTRY and TAB. We are going to make [A] first.

Step 2. Press [A] and [enter].

Another window appears near the top of the screen. It reads:

SOURCE FILE [A:DEMOBAT.TXT]

The file name in the brackets is a default name. If you press [ret] that is the file it will process. Since you examined DEMOBAT.TXT, it is assumed that is the file you wish to convert. If not, enter the name of a new file.

WARNING: It is not wise to convert batch files to ETB files on the same disk drive. Depending on the size of the batch file, there is the possibility of running out of room on the diskette.

If you have only one disk drive you will have to use short batch files. Each ETB record requires 221 bytes of disk space. You'll have to experiment to see how many records fit on a diskette that already has a batch file.

Step 3. The prompt asks for the **LENGTH OF SOURCE RECORDS? [246]**
The answer is 246 unless you've decided to process a different file. If you press return to skip this prompt, the program will default to 246.

The prompt asks **HOW MANY RECORDS ARE IN SOURCE FILE?**

Its not often that anyone knows or remembers the exact number of records in a batch file. If you don't know you would just press [return]. The default answer is 5000. The computer would process until it ran out of batch records or wrote 5000 records, whichever occured first. If you wish to cut off your batch file at a certain point you would enter that cut-off number. The program will not make more than 5000 ETB records per file.

Your reply to the prompt is just [ret]

STEP 4. The prompt asks **BEGIN CONVERSION WITH WHAT RECORD?**

You are able to start at any point in the source file. This enables you to process source files with thousand of records and divide it into several ETB-compatible files. The default answer is 1.

Reply to the prompt with a null [return] answer.

STEP 5. The prompt asks **PRINT FLAGGED RECORDS (Y/N)?**
Here you choose whether or not you want a hard copy of records that receive Edit Flags. The answer must be N or Y in capital letters. You may answer this prompt either way you choose. The default is "N".

STEP 6. The prompt asks **WRITE FLAGGED RECORDS TO FILE (Y/N)?**
A positive reply means that the Edit Flagged records, though possibly the result of a bad data entry will be included with the file. You may answer this prompt either way. The default is "Y".

The process has begun. The batch files are being read, computed, converted and saved on another file that can be appended, searched and edited with ENTRY or cross-tabulated and printed with TAB.

The resulting indices are displayed near the bottom of the screen. The flagged records are being displayed in metric measurements near the top. Deficient records were purposely included in this file. Included in this display is a count of the flagged records and what percentage of the batch file contained flagged records.

If you would like to see the flagged data in U.S. Standard rather than metric, press [U]. The change may not be immediate but don't press [U] more than once.

To switch back to metric, press [M].
To turn off the display of indices, press [D]. **The program works faster with the display of indices turned off.**

To escape the conversion process, press [Esc]. No damage will result to the file. The program will finish converting whatever record it happens to be processing and then close the file. Any batch records sent to this file in the future will be added at the end.

The return of the menu indicates the successful completion of the conversion process.

Review the work accomplished so far. Practice them again with a few variations until you are comfortable with converting batch files into data files useful to this package.

4.6 CONVERTING BATCH FILES TO OUTPUT FILES

This task covers the B and C conversion options mentioned in Task III. Use of either of these options causes the batch file data to be computed and the results to be output as another batch file.

Advanced users and programmers using DBASE, SPSS*, SASGRAPH*, etc., would be the only ones most likely to need this capability. Just as when the original batch file was downloaded to the PC diskettes, these new batch files can be "uploaded" to the mainframe.

Users who don't have access to a mainframe or some other sophisticated statistical process may wish to practice this task for experience but there will be no useful application for the files it produces using ENTRY and TAB.

Step 1. Escape back to DOS and restart the program. Connect and turn on your printer. Make sure paper is loaded and ready.

* SPSS and SASGRAPH are copyrighted trademark names.

Though this task is very similar to TASK III, we are going to approach the procedure by a quicker route.

Step 2. When you have the menu on the screen Press [4].

You are given a choice of A, B or C. You already tried A. B or C produces essentially the same product: a plain ASCII file. As indicated beside the two choices, one includes decimals and the other does not.

For this practice, choose C. (You must use upper case C)

When prompted for the name of the source file, reply with [A:DEMOBAT.TXT]. The NO DEFAULT message reminds you that you must provide a file name. The program cannot default on the very first prompt after start-up.

The next prompt asks you for the field location template. The default response is A:DEMOBAT. You may press the return key and the program will load the template.

The next prompt asks for the name you wish to give the new batch file you are making. The default is A:DEMOBAT. The program will attach the extension ".SRC" to it. Press the return key to use the default name.

The next prompt asks for the length of the batch records. The default is 246. The program found that number in the template file DEMOBAT.FLD. Press the return key to use the default.

The next prompt asks how many records are in the batch file. You shouldn't even answer this prompt unless you purposely wish to cut off the processing at a certain point.

The next prompt asks at what record you wish to start processing. The default is the first record. You wouldn't answer this prompt unless you want the program to start somewhere other than the beginning.

The next two prompts ask whether you want to print out your flagged batch records and if you want to save the flagged records. The defaults are No and Yes respectively.

The next prompt asks if you want a printout of the field locations. The default answer is [Y].

The printer will print out the field locations for the file it is making. Field location information is important to the programmer receiving the processed output file.

This program writes records that place all the results of the anthropometric computation, computed age, and three-position sequence number at the beginning of the record followed by a copy of the batch record it processed. In other words, you get a copy of the original batch file with new computations attached to the beginning of every record.

When the records are processed, you may escape to DOS and examine what the computer has made by entering TYPE A:DEMOBAT.SRC. TYPE is a DOS command. Using PRINT instead or TYPE will send the output to the printer. Consult your DOS manual if you have any difficulties.

Because we used option [C], there are decimals in the computed part of the record. If you had used selection [B] you would have obtained the same information in the same sequence except with no decimals.

The correct choice depends on your needs. Be warned, however, that both B and C write the extension .SRC to the file they create. You cannot have a B and C-created file of the same name unless you create them on separate disk drives.

APPENDIX A

Mathematical Operators Recognized by TAB

Less than (<): The program would interpret <23 as all values of 23 minus .01-- in this case all values 22.99 and less. The value .01 is always subtracted from the value given; therefore <-8 means all values starting at -8.01 and less.

Less than/equal (<=): The program would interpret <=23 as all values 23 and less.

Greater than (>): The program would interpret >23 as all values 23.00 plus .01-- in this case all values 23.01 and more. The value .01 is always added in this case no matter what the value; therefore >55.68 means all values starting at 55.69 and up.

Equal/greater than (=>): The program would interpret =>23 as all values 23 and greater starting at 23.00.

Hyphen (-): The hyphen is used to indicate the entire span from one value to another value. The program adds .09 to upper range values that are already precise to one decimal places. The programs adds .99 to upper range values that have no decimal places.

When positioned properly the program detects the difference between a hyphen and a negative sign. Therefore it is correct to use a label like -5.52--4.01 to mean numbers between negative 5.52 and negative 4.01. The program "knows" which is the hyphen. DO NOT USE BRACKETS OR PARENTHESIS.

Parentheses ( ): The parentheses is for descriptive labeling. The interval 2.00-2.99 might be given (+2 Zscr). "+2 Zscr" will be printed on the finished report instead of 2.00-2.99. LABELS MAY BE SEVEN CHARACTERS AND SPACES WIDE.

CROSS-TABULATING FOR EDIT FLAGS

It isn't wise to cross-tabulate with edit flags in the ETB files. However you can cross-tabulate for Edit Flags. Select the variable "EF" as a page, row or column. The interval would be 'X'.

CROSS-TABULATING FOR DATES

It is possible to cross-tabulate by visit date (V'DATE) or birthdate (B'DATE). If, for example, you wanted to make a longitudinal study to look at patient progress over the course of several months you would select V'Date as either a page, row or column variable, The intervals would be yymmdd-yymmdd (e.g. 850701-850831(JUL-AUG). This example searches from July 1, 1985 to August 31, 1985. The yymmdd format must be used because that is the way dates are written on the record. Also: the early date must come before the later date.

APPENDIX B

Files made by Version 3

If you were a former user of Version 2, you've noticed that Version 3 makes and calls upon a staggering variety of files. This is a summary of those file and their purpose.

DATA: This is a sequential file of growth curves and a permanent part of the package. It cannot be written to by the Version 3 programs. Loss of this file would render your Batch and Entry programs unless.

ETB: These are patient records. They are made with ENTRY and BATCH. Each record is 221 bytes long regardless of how little or how much information is put in them. They are ASCII and can be used "as is" in any outside program that will tolerate decimals in the data.

FLD: These are field location files made and used by the Batch program. They contain the column locations of key data in a batch file that the Batch program already processed. They can be recalled, edited, and reused in several variations to process other batch files.

INT: These are cross-tabulation interval files made and used by the Tab program. They contain the variables and the intervals and the variable selected for calculation of a mean. INT files are known as templates and they can be recalled, edited and stored in several variations by the Tab program.

SRC: These are ASCII output files of processed batch records. They have the Version 3-computed anthropometrics, age, sex, and a sequence number in front followed by the original batch record. They can be used in any mainframe or PC statistical, graphics, or tabulation program that uses standard rectangular files.

TAB: These are stored cross-tabulation results. They contain every tabulated figure in each cell and various control data used by the Tab program.

TEMP.TAB: This file is produced while using the TAB program when a user attempts to print the results of a cross-tabulation without saving the results. The file can be renamed (keeping the TAB extension) and added to the user's library of save results.

TXT: These are 'deprocessed' ASCII batch records made exclusively by the SWAP Version 2 to Version 3 conversion program. They are 88 columns wide and can be used in other PC and mainframe application. They are made especially for the Batch program which reprocess these records into ETB and SRC files.

APPENDIX C

Accuracy of Anthropometric Indices

The anthropometric indices are calculated by a subroutine that is identical to a FORTRAN language subroutine in the central computer at the Centers for Disease Control. It will more than likely differ from other anthropometric programs used at universities and other government agencies. This software package does not supercede or refute the calculation routines of any other program.

The only time the accuracy of the anthropometric indices may come into question within this package is when users use U.S. standard feet/inches and pounds/ounces rather than metric.

Although the option to use U.S. standard exists in the Entry and Batch programs, the program computes and stores information in metric. The anthropometric indices, computed age, height and weight are rounded up to the nearest two decimal places (values are rounded down in the case of negative numbers).

When U.S. standard measures are used the program will make its anthropometric calculations with full precision metric equilavents and then round off the metric equivalent to the nearest hundredth before writing the record to the patient file.

When using the Edit function of Entry, you may notice a sudden shift in the computed anthropometric indices when the record is restored to the disk. The shift may be as great as +- .03 because the heights and weights entered in U.S. Standard cannot be recalled and precisely represented regardless of what display mode you are using. If the slight decrepencies cannot be tolerated then it is suggested that you re-enter the U.S. Standard measurements anytime you edit a record for any reason.

APPENDIX D

Lotus Compatibility

CONVERTING LOTUS TO CASP

Lotus print files (PRN) are already compatible with Batch, the file processing program explained in Chapter IV of this handbook. As stated in Chapter IV, the main requirement for processing a text file is that all the records be uniform in length. In most ways, Lotus print files, meet that requirement. No special conversion program is needed.

The main obstacle to processing a Lotus print file is that most users apply titles, sub-headings and assorted banners and carriage returns ahead of the data. Since the Batch program interpets a horizontal row of cells as a single record, the various headings and carriage returns are useless to the process and, in fact, may even prevent examination of the file.

If you are using Lotus 1-2-3 as a database for patient records:

1. each record must contain sex, age, height and weight horizontally.
2. the database worksheet must be stored as a print file.
3. all extraneous documentation (headings, legends, etc) must be deleted.

Using the knowledge you acquired in Chapter IV, you should be able to examine one of the records from the Lotus print file, identify the column locations of the individual fields, and process the entire print file for output as ETB files or computed text files.

CONVERTING CASP TO LOTUS

There is a conversion program on the supplemental files diskette of the CASP package called LCHANGE. LCHANGE is designed to convert ETB patient files and TAB cross-tabulation results files to Lotus print files.

The purpose for the conversion program is to give users the option of using the graphic output capabilities in Lotus. The program's instructions appear on the screen.