IMPS

Version 3.1

Data Dictionary and QUICKTAB for PUMS Applications

International Systems Team
Bureau of the Census
U. S. Department of Commerce
Washington, D.C. 20233-3102

October 17, 1994

ACKNOWLEDGMENTS

This Guide and the software documented herein were developed by the International Statistical Programs Center of the United States Bureau of the Census through funding from the Office of Population of the United States Agency for International Development.

Chapter 1. Introduction 1

Chapter 2. Installation 2

Chapter 3. Getting Started 3
3.1 Starting QUICKTAB 3
3.1.1 Starting Frequencies 3
3.1.2 Starting Cross-Tabulations 4
3.2 Starting the Data Dictionary 5
3.2.1 Developing a Dictionary 5
3.2.2 Listing a Data Dictionary 6
3.3 Solving Problems 6
3.4 Frequent Errors 7

Chapter 4. General Considerations 9
4.1 Running Applications 9
4.2 QUICKTAB without Menus 10
4.3 Memory Usage 10
4.4 The IMPS.SYS file 11
4.5 Using IMPS on a Network 11

Chapter 5. Introduction to Data Dictionary 12
5.1 What is a Data Dictionary? 12
5.2 Identification Section 12
5.3 Sections of a Questionnaire 12
5.4 The Questions 12
5.5 Types (Values) of Responses 13
5.6 Data Items 13
5.7 Records 13
5.8 Questionnaires 13
5.9 Data Files 14
5.10 Record Types 14
5.11 Defining Record Characteristics 14
5.11.1 Record Type Codes 14
5.11.2 Required Records (Req) 15
5.11.3 Maximum Records (Max) 15

Chapter 6. 16
6.1 Types of Data Items 16
6.2 Implicit Decimal Points 16
6.3 Length of Data Items 16
6.4 Choosing a Data Type 17
6.5 Naming Data Items 17

6.6 Data Item Values 18
6.6.1 Values and Value Names with QUICKTAB 18
6.6.2 Special Value Names 19
6.6.3 How NA and NR Are Used 19
6.7 Subitems: When and How to Use Them 20
6.7.1 Items Such as Time and Date 20
6.7.2 Split Items for Broad Categories 20
6.7.3 Defining an Item as Both Numeric and Alphanumeric 20
6.7.4 Special Ranges for QUICKTAB 21
6.7.5 Long Cross-tabulations 21
6.8 Repeating Data Items (Occ) 22

Chapter 7. Using the Data Dictionary 23
7.1 Layout or List Mode 23
7.2 Modifying the Data Dictionary 23

Chapter 8. Introduction to QUICKTAB 27
8.1 What is QUICKTAB? 27
8.2 Common Uses of QUICKTAB 27
8.3 Capabilities of QUICKTAB 27
8.3.1 Frequency Distributions 27
8.3.2 Cross-tabulations 28
8.3.3 Weighted Tabulations 28

Chapter 9. Using QUICKTAB 29
9.1 Implications of Data Dictionary Value Name Definitions 29
9.2 Selection Criteria 30
9.3 Counts Versus Percentages 30
9.4 Title of Report and Name of Tables File 30
9.5 Choosing Dimensions for Cross-tabulations 30
9.6 Data Items as Weights 30
9.7 Printing Tables from QUICKTAB 31
9.8 Rounding of Results 31

Chapter 10. A QUICKTAB Example 32
10.1 QUICKTAB Frequency 32
10.2 Weighted Frequency Counts 37
10.3. Record Selection (Defining Conditions for Tabulation) 37
10.4 Save and Reuse Selection Criteria 38
10.5 QUICKTAB Cross-tabulations 38

Appendix A. Data Dictionary System Limits 41

Appendix B. QUICKTAB System limits 42

Appendix C. Data Dictionary Run-time Error Messages 43

Appendix D. IMPS File Extensions 48

Appendix E. Utilities 49

E.1 Convert Box Characters 49
E.2 Merging Data Dictionaries 49
E.3 Recovering a Data Dictionary 50

Chapter 1. Introduction

The Integrated Microcomputer Processing System (IMPS) performs the major tasks in census and survey data processing. This subset of IMPS contains the Data Dictionary and QUICKTAB components. They allow persons with little or no computer experience to quickly tabulate data files.

IMPS requires the following hardware and software:

Microcomputer : IBM PS/2, PC, or compatible

Memory : 640 Kbytes (512 Kbytes available)

Disk Storage : This subset of IMPS requires approximately 2.5 Mbytes of disk storage, broken down as follows:

0.3 MB - IMPS system files (menus)
0.8 MB - DATADICT
0.9 MB - QUICKTAB
0.5 MB - PUMS Example application

Using the system requires some additional space for intermediate files.

Printer : Should be capable of 132 character wide display

DOS : DOS 3.3 or above

Chapter 2. Installation

The CD-ROM contains the Data Dictionary and QUICKTAB software. You can execute both systems directly from the CD-ROM without any special installation. However, if you wish to modify the PUMS data dictionary you must copy it to your hard disk.

Chapter 3. Getting Started

The two IMPS modules on the CD-ROM are Data Dictionary (DD) and QUICKTAB (QT). The Data Dictionary is used to describe the data file, variables and values. QUICKTAB is used to do frequency distributions and cross tabulations using the values defined in the Data Dictionary.

NOTE: If any screens are difficult or impossible to read because of the colors or gray scales, see Section 4.4.1 'Turning Color Off'.

In the following explanations assume that your CD-ROM drive letter is "Q".

3.1 Starting QUICKTAB

Create a "working" directory on your hard disk, for example **C:\WORK**.
From this directory, enter **Q:\TOOLS\QT** to execute the QUICKTAB module.
If requested, enter the network User-ID (any 3 characters, for example your initials).

The QUICKTAB main menu should appear which gives, among other choices, "Frequencies" or "Cross Tabulations". See chapters 8, 9, and 10 for more details about QUICKTAB.

3.1.1 Starting Frequencies

Press the ENTER key again to select 'Frequencies'.

Next to 'Data Dictionary file:' type '**Q:\TOOLS\DOC\90PUMSX2**'
and press ENTER.
Next to 'Data file:' type '**Q:\TOOLS\DOC\PUMSAXXX.TXT**'
and press ENTER.
Next to 'Saved settings file:' LEAVE BLANK
and press ENTER.

The following screen should appear:

Records Tabulate View Print DOS Load Save End F1=Help

Press ENTER and two 'record names' (HOUSING-RECORD and PERSON-RECORD) should appear. Press ENTER again, use the arrows to move and 'ENTER' to select an item or two. Press ESC key twice, once to return to the record names and again to return to the menu shown above.

Use the arrow key to move the highlight bar over the 'Tabulate' option and press ENTER. Confirm that the data file is **Q:\TOOLS\DOC\PUMSAXXX.TXT** by pressing ENTER. Confirm the report title by pressing ENTER.

Press ENTER to confirm that the name of the output file should be FREQ (with an automatic extension of .TBL). After execution a table should be displayed on the screen. You

are in the "View" program which allows users to look at ASCII files. The F1 key provides "help" information about movement in the file and other functions.

Press the ESC key to exit "view" and again to exit QUICKTAB Frequencies.

A box should appear with the question 'Save settings?'. To accept the default 'No', press the ENTER key. You should now be back at the QUICKTAB main menu.

3.1.2 Starting Cross-Tabulations

Use the arrow key to move the highlight bar over the 'Cross-Tabulations' option and press the ENTER key.

Next to 'Data Dictionary file:' type/confirm '**Q:\TOOLS\DOC\90PUMSX2**'
and press ENTER.
Next to 'Data file:' type/confirm '**Q:\TOOLS\DOC\PUMSAXXX.TXT**'
and press ENTER.
Next to 'Saved settings file:' LEAVE BLANK
and press ENTER.

The following screen should appear:

Records Tabulate View Print DOS Load Save End F1=Help

═══════════════════════════════════════

Press ENTER and two 'record names' (HOUSING-RECORD and PERSON-RECORD) should appear. Press ENTER to get to the cross-tabulation template and again to get a list if variables. Use the arrows to move and 'ENTER' to select an item for the 'Row' and the same procedure to select an item for the 'Column'. Press ESC key twice, once to return to the record names and again to return to the menu shown above.

Select the 'Tabulate' option and press ENTER. Confirm that the data file is PUMSAXXX.TXT by pressing ENTER. Confirm the report file title by pressing ENTER. Confirm that the name of the output file should be CROSSTAB (with an automatic extension of .TBL). A table should be displayed on the screen.

Press ESC key twice to exit QUICKTAB cross-tabulations.

A box should appear with the question 'Save settings?'. To accept the default 'No, press the ENTER key. You should now be back at the QUICKTAB main menu.

Press ESC key to return to the DOS prompt C:\WORK.

3.2 Starting the Data Dictionary

NOTE: This section can be skipped until you want to add or change a set of values for a data item (variable) in the Data Dictionary. QUICKTAB uses the sets of values assigned to a data item for distributions and cross tabulations.

In your "working" directory on hard disk make a copy of the Data Dictionary, 90PUMSX2.DD. (This file cannot be modified on the CD-ROM.)

From the directory **C:\WORK**, enter **COPY Q:\TOOLS\DOC*.DD**
Enter **Q:\TOOLS\DD** which executes the Data Dictionary module.
If requested, enter the network User-ID (any 3 characters, for example your initials).

The DATADICT main menu should appear with the options to "Develop" or "List" a Dictionary.

3.2.1 Developing a Dictionary

Select this option to modify the contents of the data dictionary.

Next to 'Data Dictionary file' type '**90PUMSX2**' (or press 'F2' and select)
and press ENTER.

The following screen should appear:

Common Records Layout Save End

NOTE: If the message 'Not enough memory' appears, you do not have conventional memory available to use IMPS. The Data Dictionary requires the most memory of all IMPS components, 512K bytes. See section 4.3 'Memory Usage' for suggestions on making more memory available.

Continue by pressing the ENTER key. The following screen should appear:

Item name	Type	Len	Substart	Start
SERIALNO	N	7		2

Press the ESC key.
Press the key 'E' to END the session.

A box should appear with the question 'Save Dictionary file?'. The default is 'Yes'. Since you have not changed the Data Dictionary, use the arrow key to move the highlight bar over 'No' and press the ENTER key. You should now be back at the Data Dictionary main menu.

3.2.2 Listing a Data Dictionary

Select this option to create a listing of the contents of the Data Dictionary. (Note: The "DD" cannot be printed, this option creates a "LST" which is an ASCII print file.)

Next to 'Data Dictionary file' type/confirm '**90PUMSX2**' (or press 'F2' and select) and press ENTER.

A formatted listing (90PUMSX2.LST) will be displayed on the screen. This file can be printed for documentation purposes.

Enter ESC to return to the main Data Dictionary menu and ESC again to return to the DOS prompt.

NOTE: See Chapters 5, 6 and 7 for more details on the Data Dictionary

3.3 Solving Problems

Listed below are some symptoms, causes, and solutions to problems which can occur during the use of IMPS software.

Symptom: You enter the Q:\TOOLS\QT command and receive the message 'Bad command or file name'.

Cause - The "QT" EXEcutable could not be found on the given drive or in the given directory.

Action - Check that "q" is the CD-ROM drive, and the spelling of the path to "DD" is correct.

Symptom: One of the following messages appears:

R6005 not enough memory on exec
R6008 not enough space for arguments
R6009 not enough space for environment

Cause - There is not enough conventional memory to load and execute a program in the IMPS component.

Action - See section 4.3 below for a discussion of memory problems.

Symptom: When the END option is chosen in an IMPS component, the DOS prompt appears instead of the IMPS (DATADICT or QUICKTAB) menu from which it was started.

Cause - There is not enough conventional memory to reload the appropriate IMPS menu.

Action - See section 4.3 below for a discussion of memory problems.

3.4 Frequent Errors

Problem: Message "$$XCTL$$" is in use, try again later"

Cause: User is trying to execute QUICKTAB from CD-ROM directory.

Solution: User's current directory is probably Q:\TOOLS\QUICKTAB on CD-ROM. The current directory must be on a writable device. Go to your "working" directory (C:\WORK) and restart QUICKTAB.

Problem: Message "$$XCTL$$" is Read Only"

Cause: User made files Read-Only (+R) after executing QUICKTAB.

Solution: In user's current directory enter
>**attrib -r $$*.***
Then retry.

Problem: Frequency or Cross Tab tables have no counts. All cells have dashes (-) in them.

Cause: Data file is incorrect.

Solution: The user has selected a "data file" which is not a ".TXT" file. Go back to "Tabulate" and enter a .TXT file as the data file. (can also use "\.TXT" then F2 key).

Chapter 4. General Considerations

4.1 Running Applications

To use QUICKTAB you must have a Data Dictionary (DD) which describes the data file you are tabulating. (The \TOOLS\DOC directory contains a Data Dictionary which describes the PUMS data file.) If no Data Dictionary is available, it must be created (see chapters 5 and 6).

Programs must be executed from a user's directory on your hard disk, rather than a directory on the CD-ROM, so that intermediate and print files can be written.

Enter the following:

C:
CD \WORK [All intermediate and output files will be placed in this directory]

To use DD or QUICKTAB, simply make selections from the menus. To make a selection, either move the highlight bar to the choice and press ENTER, or type the first letter (number) of the choice. When you execute any IMPS procedure you will be prompted for the necessary file names.

In most of the menus, the following function keys can be used:

F1 provides help for the current menu screen.

F2 can be used wherever a file name is entered. It gives an alphabetical listing of files names in the directory path specified. If no directory path is specified, the listing comes from the current directory. If the file name to be entered has a fixed extension, for example .DD, only files with this extension will be listed. If the file name to be entered can have any extension, then the extension specified in the current file name is used to limit the listing. If no extension is present, all files will be listed. For example, if you are trying to enter the name of a data file in the current directory with an extension of .DAT, but cannot remember the file name, you can type .DAT and then press 'F2'. IMPS will list in alphabetical order all the data files in the current directory with the extension .DAT.

F3 can be used to begin processing once all the required file names have been entered.

4.2 QUICKTAB without Menus

To use QUICKTAB outside of the menu system, the user can call one of two programs, FREQ or CROSSTAB. The batch commands below could be used

Q:\TOOLS\QT\FREQ dictionary-file data-file save-file
or
Q:\TOOLS\QT\CROSSTAB dictionary-file data-file save-file

4.3 Memory Usage

The following list shows the amount of conventional memory required by each of the components of IMPS:

Data Dictionary	512K
QUICKTAB	425K

To use either component of IMPS the user should have at least 512K of available conventional memory. In any version of DOS you can find the amount of available memory (bytes free) using the CHKDSK command. In DOS 4.0 and above, it is easier and quicker to use the MEM command.

There are several ways to increase the amount of conventional memory available:

1) Remove any unneeded memory resident programs (for example, SIDEKICK, VIRUSCAN, PRINT, or FASTOPEN). These programs stay in memory after they are used, leaving less memory available for IMPS. If you have DOS 4.0 or above, you can inspect the contents of memory for resident programs by using the command MEM/PROGRAM. If you have DOS 5.0, you can use MEM/C.

2) If you are using IMPS on a network, you will find that some of the network software is memory resident. Check your network software documentation to determine how you might make more conventional memory available by placing network software in extended memory.

3) If you are using DOS 5.0, put DOS=HIGH in your CONFIG.SYS file. This will place part of DOS in high memory, which most new computers contain.

4) If you are using DOS 6.0 and a have 80386 or 80486 processor, you may be able to free conventional memory by using the MEMMAKER program which comes with DOS.

4.4 The IMPS.SYS file

The file IMPS.SYS contains paths to the various components of IMPS and various settings required by IMPS.

```
DATADICT=DD
QUICKTAB=QT
```

4.5 Using IMPS on a Network

The IMPS Data Dictionary and QUICKTAB can be used on a network. You may wish to have data stored on a file server which can be analyzed by multiple users. The Data Dictionary files can also be stored on a file server. However, if a user attempts to modify an IMPS Data Dictionary stored on a network, no other user on the network can access that Data Dictionary while it is being modified. Therefore, you may wish to modify data dictionaries on a local hard disk.

If you are attached to a network or have a network card in your computer, IMPS programs ask you for a network User-ID. You can enter any three characters in response. If you would like to eliminate this request, enter the following DOS command before using IMPS:

```
SET IMPSNET=nnn
```

where "nnn" are any three characters.

Chapter 5. Introduction to Data Dictionary

5.1 What is a Data Dictionary?

The Data Dictionary is used to give a description or a picture of how data are (or will be) stored in the computer. It allows you to provide a meaningful name for data items and to define characteristics such as whether the data item is made up of numbers or letters, how many characters or digits there are in a data item, and whether a data item has an assumed decimal point. The Data Dictionary also allows you to define the overall structure of a data file.

5.2 Identification Section

The identification section identifies the questionnaire, usually with numeric codes. The combination of identification codes (such as province, district, place, household) on a questionnaire uniquely identifies the questionnaire. These are the codes you would need to locate a specific questionnaire.

5.3 Sections of a Questionnaire

The body of a questionnaire is often divided into sections, each of which asks a related set of questions. For example, a census questionnaire may have a section asking for information about the household or housing unit, and another section asking for information about each person in the household. Information about each person may be further grouped by topics such as education, economic status, and fertility.

The example questionnaire is divided into 2 sections:

Characteristics of the House
Characteristics of Persons

5.4 The Questions

The basic element of the questionnaire is the question. Each section of the questionnaire contains a set of questions being asked for this census or survey. For example,

What is ...'s relationship to the head of household?
What is the mortgage on this property?

We will refer to the response as the data item.

5.5 Types (Values) of Responses

Some responses are quantitative, such as age of person; and some are qualitative, such as relationship to head of household. Responses can be numeric or alphanumeric. Most descriptive responses, such as 'head of household' are equated to numeric codes which are placed on the questionnaire. However, some descriptive responses remain as alphabetic text.

Thus, numeric responses can be discrete values or quantitative values. An example of a discrete value is gender, 1 (male) or 2 (female). An example of a quantitative value is yearly income. A discrete value may be used to designate a quantity category. For example, when asking value of property, one may be asked to select from a choice of ranges of values rather than specify the exact value. Therefore, the possible responses to the property value question could be a code between 1 and 25.

An alphanumeric value consists of alphabetic and numeric characters, blanks, and special characters. For example, 'M' or 'F' for gender is an alphanumeric value.

5.6 Data Items

A data item contains the response to a question. It is the basic element of a questionnaire. AGE, INCOME, and MARITAL-STATUS are examples of data items.

5.7 Records

A record usually corresponds to a section of a questionnaire. It is a group of data items. For example, the data items of the housing questions form housing records; the data items for the person questions form the population records.

5.8 Questionnaires

Records with the same questionnaire identification codes make up a questionnaire. If the questionnaire identification codes identify a household, then all the records belonging to the household make up a questionnaire. In a typical housing and population census, a questionnaire would contain one household record and an indefinite number of person records, depending on the number of persons in the household.

In some cases, one record makes up a questionnaire. For example, a student roster might consist of a record for each student. The student identification number could serve as the questionnaire identification.

5.9 Data Files

A data file is a collection of questionnaires. Data files processed by IMPS must be sequential ASCII files, sometimes called DOS files. There is no limit to the size of a data file that IMPS can handle.

5.10 Record Types

In some instances, every record in a file has the same structure. In other cases, records have several different structures. In a housing and population census questionnaire, the data file usually has two different record structures. In general, the user has some description of the data file which the Data Dictionary is to describe.

If the data file has more than one record type, each structure must be associated with a record type, name, and code. For a typical housing and population census, the record type, names, and codes could be HOUSING (code H) and PERSON (code P). IMPS allows the record type code to be any alphanumeric value up to 12 characters. Although record type codes are usually numeric, IMPS allows them to be alphanumeric as well.

By default, the IMPS Data Dictionary sets the record type identifier as a numeric character occupying the first (or first two) positions of each record in the file, but the location of the record type identifier can be modified. IMPS allows up to 98 record types within a questionnaire.

5.11 Defining Record Characteristics

When defining records in a Data Dictionary, you must include record names, information on whether the record is required (Req), and the maximum number of records of each type within a questionnaire (Max).

5.11.1 Record Type Codes

If there is only one record type defined in the Data Dictionary, there will be no automatic field defined for record type code. If there is more than one record type defined in the Data Dictionary, a numeric field starting in position one of the record will automatically be defined to hold the value of a record type code. If there are two through nine record types, the field will be one character in length. If there are ten or more record types, the field will be two characters in length. The values of the record types will be consecutively numbered starting at '1' in the same order in which the record types are defined in the Data Dictionary. If you wish to change the value or position of the record type, you may do so in the 'Layout' mode of the Data Dictionary.

5.11.2 Required Records (Req)

'Req' is 'Y' (yes) or 'N' (no) to indicate whether at least one record of that type must appear for each questionnaire or not. This option has no meaning for QUICKTAB. Choose whichever option makes more sense to you or use the default value.

5.11.3 Maximum Records (Max)

In developing a Data Dictionary, you may define a 'Max' value for each record type. In QUICKTAB, this option has no meaning. Choose any value which is reasonable or use the default values.

Chapter 6. Defining Data

6.1 Types of Data Items

The data file can consist of data items which are either numeric or alphanumeric. This classification is important to QUICKTAB. For example, alphanumeric data cannot be used unless specific values are assigned for the data item. In the case of numeric data, IMPS only handles whole numbers. Decimal points are handled implicitly. That is, you can specify where an assumed decimal place is, but the data cannot actually contain a decimal point.

A data item may be broken into subitems. That is, one or more subitems may make up a data item. For example, the item DATE might be a 6-digit item consisting of day, month, and year. This item might be divided into three 2-digit subitems: DAY, MONTH, and YEAR. A detailed discussion of subitems appears later in this chapter.

Data items cannot have negative values. However, if it becomes necessary to handle negative data, it is possible to define the sign as a separate alphanumeric field. A QUICKTAB cross-tabulation can use the sign field as one of the fields.

6.2 Implicit Decimal Points

A data item may be defined in the Data Dictionary as having a decimal portion. Define the length of the item to include all the integer and decimal digits but not a decimal point. Define the decimal length with the number of decimal digits. For example, if the largest expected value for the 'Weight' is 99.9, then the Data Dictionary item can be defined as:

Item name	Type	Len	Decs	Occ	Substart	Start
WEIGHT	N	3	1	1		11

In QUICKTAB, the decimal point will be implied. When values are defined for data items defined as having decimal places, these values are shown on the Data Dictionary listing without decimal points. Defining a data item as having decimal places serves only for documentation purposes.

In QUICKTAB, values or value ranges in row stubs or column headings are shown without decimal points, regardless of how they were defined in the Data Dictionary.

6.3 Length of Data Items

In selecting an appropriate length for a data item definition in the Data Dictionary, one must know the largest value for the data item. A data item of any length can be defined in the Data Dictionary, but if it is longer than 15 digits/characters, it cannot be referenced in QUICKTAB.

In IMPS, values for numeric data items are assumed to be right-justified, zero filled. Values for alphanumeric data items are left-justified, blank filled. For example, if a 5-position

item is defined as numeric, and it has the value 45, it will be placed on the file or referenced, assuming it is on the file) as

0	0	0	4	5

If the item is defined as alphanumeric and has the value 45, it will be placed on the file, or assumed to be in the format:

4	5			

6.4 Choosing a Data Type

Two data types are available in IMPS: numeric and alphanumeric. Any data item which would normally contain a numeric code or value should be defined as numeric in the Data Dictionary. This includes data items that are in some cases not applicable or not reported (blank).

It is possible to define a data item as both numeric and alphanumeric in the Data Dictionary. This can be done by redefining a numeric data item with an alphanumeric subitem. The data item and the subitem have different names but define the same positions in the record.

6.5 Naming Data Items

Data item names provide a means of referencing data items in QUICKTAB. A data item name can be 1 to 16 characters long and consists of letters (A-Z), digits, and imbedded hyphens (-). The first character must be a letter. The last character must be a letter or a digit. Upper and/or lower case letters may be used. However, a lower case letter is assumed to be identical to an upper case letter. Thus, the name 'Age' is identical to the name 'AGE', and the two may be used interchangeably.

Data item names have an abbreviated form, called the short name. The short name consists of either the first eight characters of the data item name, or, if a hyphen appears within the first eight characters, those characters before the hyphen. Names for data items and subitems must be unique in their abbreviated, or short name form. The Data Dictionary system will not allow the use of IMPS commands, keywords, and special names as data items names.

Some examples of data item names and short names are:

Item Name	Short Name
SEX	SEX
Relationship	Relation
MOTHER-ALIVE	MOTHER
age	age
P04-AGE	P04

6.6 Data Item Values and Value Names

Each data item and subitem can be associated with a list of values. For example, the data item P03-SEX might be associated with the values 1 (for male), 2 (for female), and some value, possibly blank, for 'not reported'. The data item P04-AGE might be associated with the values 0 to 98 (0:98 in IMPS notation) and NR (for 'not reported'). Each value or value range may be assigned a name. 'Male' and 'Female' are reasonable value names for the data item P03-SEX. There is no obvious value name for the range of ages, so no name should be given. NR, a special name meaning, 'No Response', can be given for the value of all spaces. Values and value names affect QUICKTAB tabulations.

It helps to have a listing of the names of such values, and these appear in the Data Dictionary listing which is produced by selecting the 'Listing' option on the Data Dictionary menu.

6.6.1 Values and Value Names with QUICKTAB

QUICKTAB is a menu-driven module of IMPS which produces frequency distributions and cross-tabulations rapidly. QUICKTAB allows you to select any item or subitem that was defined in the Data Dictionary. For example, you can produce a cross-tabulation of P04-AGE and P03-SEX. There are some restrictions in QUICKTAB. Alphanumeric items which do not have values defined cannot be selected or referenced in QUICKTAB. Items and subitems with occurrences cannot be referenced in QUICKTAB. The 'Occurs' feature of the Data Dictionary is discussed later in this document.

Within the Data Dictionary definitions, you can define data item value ranges the way you would like to see them in the QUICKTAB tables. Moreover, through the use of subitems, you can define the same item several times, with a different set of value ranges for each definition. For example, suppose you produced a CROSSTAB table by single year, age, and sex, and you would like to produce other tables with 5-year age groups. If you look at a listing of the Data Dictionary EXAMPLE, you will see that AGE has been redefined twice as subitems. AGE5-YEAR defines 5-year age groups while AGE10-YEAR defines 10-year age groups.

QUICKTAB produces certain statistics (minimum value, maximum value, mean and standard deviation) for quantitative data items. In order to define a data item so that these statistics are produced by QUICKTAB, do the following: Define the item or subitem as numeric. Either omit values and value names altogether, or include only values. Including value names other than NR and NA will prevent the generation of these special statistics. In short, data items which represent quantitative values should be defined without value names. Age is an example of such an item. But numeric items which describe categories, such as sex or marital status, should be defined with value names.

6.6.2 Special Value Names

A special name is a reserved word used to name a special value. The following special names can be used as value names or values in the Data Dictionary:

NA - meaning 'not applicable'
NR - meaning 'not reported' or 'no response'

Assigning values and value names to a data item is relatively simple. In addition to the values and value names that are apparent from looking at the questionnaire or data file description, NA and NR should be considered. If a response is not applicable in all cases and therefore allowed to be blank, the value name NA and the value of spaces can be defined for the item. For example, fertility questions would not be asked of persons under 12 years of age. For these persons, the fertility responses should be blank, so NA should be defined as a value name.

The item name NR also may be associated with a value for a data item. It is appropriate to include NR as a value name if you know that the data file contains special nonresponse codes. In a population census, if the respondent is answering questions about absent members of the household, he/she may not always know their exact age or education level. Therefore, it would be appropriate to associate NR with these data items.

A data item like CHILDREN-BORN could have both NA and NR associated with it, as well as a range of numeric values like 0:15. It is not applicable to all persons, and it will sometimes not be reported. A data item like SEX would probably not have NA or NR associated with it, only MALE (value of 1) and FEMALE (value of 2) values would be on the data file. This question is applicable to all persons and is always reported.

6.6.3 How NA and NR Are Used

NA is usually defined as spaces because, when a question is not applicable, it is usually left blank. NR responses are also often left blank. If both NA and NR are associated with the value blank, then QUICKTAB will not be able to distinguish between them. The first value listed, NR or NA will 'catch' ALL the blank values in the field.

However, it is possible to associate a value other than spaces with NR and NA. Quite often, 9's are associated with NR.

6.7 Subitems: When and How to Use Them

Subitems redefine data items. They allow you to refer to the same data field in several different ways in QUICKTAB. Indicate that the data name is to be a subitem by giving the subitem's start position within the item in the 'substart' column. If the subitem starts in the same position as the item, use the same start position as the item. Make sure that the subitem length is less than or equal to the item length. Typical situations in which subitems are used are described below:

6.7.1 Items Such as Time and Date

You may wish to reference a date or time item as one entity and also in parts for easier editing or calculating. For example, in order to calculate age from date of birth, it is helpful to have day, month, and year available for separate reference. For this, you will need the date divided as day, month, and year subitems. The item BIRTHDATE might be subdivided like this:

Item name	Type	Len	Decs	Occ	Substart	Start
BIRTHDATE	N	6	0	1		19
BDAY	N	2	0	1	19	19
BMONTH	N	2	0	1	21	21
BYEAR	N	2	0	1	23	23

6.7.2 Split Items for Broad Categories

One reason for using subitems is to make data references available in larger categories. Censuses and surveys often have items of 3 or 4 digits representing categories like industry, occupation, or ethnicity. Within QUICKTAB, it could be useful to refer to different levels of detail. For occupation codes, the full value refers to a very detailed occupation, such as bus driver. The first digit alone refers to the 'major' division, such 'public service'. The first 2 digits refer to a more detailed 'minor' division, such as 'public transportation'. It may be useful to list occupations by these classes. The following code might appear in a Data Dictionary for an economic survey:

Item name	Type	Len	Decs	Occ	Substart	Start
P22-OCCUPATION	N	4	0	1		45
P22A-OCC-GRPS1	N	1	0	1	45	45
P22B-OCC-GRPS2	N	2	0	1	45	45
P22C-OCC-GRPS3	N	3	0	1	45	45

6.7.3 Defining an Item as Both Numeric and Alphanumeric

Sometimes you may have an item on a person record which, for persons from regular housing units, contains MARITAL-STATUS (codes 1 to 5). For persons in collective quarters, it contains PERM-RESIDENT (codes Y or N, for yes/no). Questionnaires for regular quarters will be keyed separately from those of collective quarters, and two different data entry applications will be needed. But it would be convenient to use only one Data Dictionary. The following definitions might appear in the Data Dictionary:

Item name	Type	Len	Decs	Occ	Substart	Start
MARITAL-STATUS	N	1	0	1		41
PERM-STATUS	A	1	0	1	41	41

6.7.4 Special Ranges for QUICKTAB

Another reason for using subitems is to provide different value ranges for tabulations done by QUICKTAB. You may wish to look at age frequency distributions in 10-year age groups and 5-year age groups. The values for AGE could be 0:98. The values for TEN-YEAR-AGE could be 0:9, 10:19, 20:29, etc. The values for FIVE-YEAR-AGE could be 0:4, 5:9, 10:14, etc.

Item name	Type	Len	Decs	Occ	Substart	Start
AGE	N	2	0	1		15
TEN-YEAR-AGE	N	2	0	1	15	15
FIVE-YEAR-AGE	N	2	0	1	15	15

6.7.5 Long Cross-tabulations

Similarly, you may wish to examine a cross-tabulation of two items but find that there are too many values for an item to allow creation of the table. You can redefine one of the items as several Data Dictionary subitems: the first subitem will define values for the lower part of the range, the next subitem will define values for the next part of the range, and so on. You can then cross-tabulate the first item by each subitem.

For example, the integer part of FARMSIZE has values 00 through 99, and the integer part of area cultivated CULT has values 00 through 99. Assume that the cross-tabulation should reflect each value, and not value intervals. If FARMSIZE becomes the rows of a QUICKTAB cross-tabulation, CULT has too many values (100) to become the columns. You can divide CULT by using subitems each with fewer values.

Item name	Type	Len	Decs	Occ	Substart	Start
FARMSIZE	N	3	1	1		11
IFARM	N	2	0	1	11	11
CULT	N	3	1	1		14
ICULTA-00-14	N	2	0	1	14	14
ICULTB-15-29	N	2	0	1	14	14
ICULTC-30-44	N	2	0	1	14	14
ICULTD-45-59	N	2	0	1	14	14
ICULTE-60-74	N	2	0	1	14	14
ICULTF-75-89	N	2	0	1	14	14
ICULTG-90-99	N	2	0	1	14	14

The values associated with ICULTA-00-14 would be 0:14; those associated with ICULTB-15-29 would be 15:29, etc.

6.8 Repeating Data Items (Occ)

The 'Occ' (occurs) option is a way to define a group of repeating data items of the same length. You cannot currently reference subscripted data items in the current version of QUICKTAB. Therefore, if you need to do frequencies or cross-tabulations with a repeated item, each of the items should be given a separate name either as an item or subitem (see below).

Item name	Type	Len	Decs	Occ	Substart	Start
CGROUP	A	18	0		11	
CCODE1	N	2	0	1	11	11
CCODE2	N	2	0	1	13	13
CCODE3	N	2	0	1	15	15
CCODE4	N	2	0	1	17	17
CCODE5	N	2	0	1	19	19
CCODE6	N	2	0	1	21	21
CCODE7	N	2	0	1	23	23
CCODE8	N	2	0	1	25	25
CCODE9	N	2	0	1	27	27

Chapter 7. Using the Data Dictionary

7.1 Layout or List Mode

Layout mode offers a graphical representation of the records which comprise a questionnaire in the data file. Whether to work in layout mode or list mode is a question of user preference. However, you must use the list mode to:

1) define valid values for a data item
2) assign value names
3) copy value lists from one item to another

You must use the layout mode to:

1) modify record type values
2) move record type values
3) change data item starting positions

7.2 Modifying the Data Dictionary

A Data Dictionary is usually designed from a record layout. For a number of reasons, you may want to change the Data Dictionary after it was created. When using a QUICKTAB application, you may wish to redefine certain fields to look at certain frequency distributions.

For example, suppose you are interested in three types of education for persons (YEARSCH on PERSON-RECORD): High school diploma only, College up to a Bachelor's degree, and Higher level advanced degrees. The user can group the codes for YEARSCH appropriately.

1) Make sure that you are in your working directory (C:\WORK) and have a copy of 90PUMSX2.DD in that directory.

2) Enter 'Q:\TOOLS\DD' (or execute an appropriate "BAT")

3) Enter network initials (3 characters, if required)

4) Select Develop Dictionary from main Data Dictionary menu

5) Enter : **90PUMSX2** (or press 'F2' and select it)

The following screen should appear:

Common Records Layout Save End

6) Select 'Records'

A 'window' should appear with the following headings:

Record name	Req	Max

7) Select 'PERSON-RECORD' from list of available records

A 'window' should appear with the following headings:

Item name	Type	Len	Decs	Occ	Substart	Start
RELAT1	N	2	0	1		9
...						

8) Use the 'down-arrow' and/or 'Page-Down' to go below YEARSCH to ANCSTRY1 and press 'F3'. This inserts a 'blank' item after YEARSCH with a 'Start' of 53.

Item name	Type	Len	Decs	Occ	Substart	Start
...						
YEARSCH	N	2	0	1		51
	N	1	0	1		53
ANCSTRY1	N	3	0	1		54
...						

9) Enter ED-GROUPS under Item name, 2 under Len, and 51 under Substart.:

Item name	Type	Len	Decs	Occ	Substart	Start
...						
┌YEARSCH	N	2	0	1		51
└**ED-GROUPS**	N	**2**	0	1	**51**	51
ANCSTRY1	N	3	0	1		53
...						

10) Move the cursor to 'ED-GROUPS' and press 'ENTER'.

A 'window' should appear with the following headings:

Value name	From	To

11) Key the following information

Value name	From	To
High School	10	
College	11	14
Post Graduate	15	17
Not HS Graduate	1	9 (Note: leading zeros are added after 'ENTER')

12) 'ESC' back to data item names 'window'.

Item name Type Len Decs Occ Substart Start

13) 'ESC' back to record names 'window'.

Record name Req Max

14) 'ESC' back to Data Dictionary menu (top line of screen)

Common **Records** Layout Save End

15) Either Save and End or 'ESC' and answer 'Yes' when the system asks if you want to Save the Data Dictionary.

16) 'ESC' back to the Main Menu.

You are now ready to enter QUICKTAB frequencies and to select ED-GROUPS from the PERSON-RECORD. After tabulation, the user should see a set of five counts for ED-GROUPS. The four specified above and 'Undefined'. Where did the Undefined counts come from? Our three groups of values covered the values from 1 to 17, but they did not include Not Applicable (NA) which is coded as zero. Since '00' is not in the list of values, it is considered 'Undefined'. To rectify the problem, go back to the Data Dictionary and add to the ED-GROUPS the data item and the value name 'na' with its value '00', note that the 'na' is converted to 'NA' when entered. Now QUICKTAB when rerun, will classify those undefined values as 'NA'.

NOTES:

- A data item can be redefined many times. (See AGE-SINGLE, AGE5-YEAR, and (AGE10-YEAR on the PERSON-RECORD.)
- Value names do NOT need to be given (See the AGE-SINGLE as above.)

- Unnamed values and Named values can be mixed (See POVERTY on PERSON-RECORD.) However, unnamed values must come before named values.

- The values in the 'From' - 'To' list do not need to be in any specific order (See the values for the HISPANIC data item on PERSON-RECORD.)

- Multiple ranges and single values can be mixed for a value name. (See values for HISPANIC as above.)

- If single values are listed they go in the 'From' column (See CITIZEN on PERSON-RECORD.)

- A blank value (spaces) can be defined for numeric data items by pressing 'ENTER' in the 'From' column. In this case there should be nothing in the 'To' column.

- If the 'From' - 'To' values are overlapping for a data item, the first 'true' value will receive the QUICKTAB count. For example, if both NA and NR are given 'blank' values the first one in the list will receive all the QUICKTAB counts for blank values and the other will receive none (zero). This can be helpful at times. Suppose the user wants counts for California (code 6) and Florida (code 12) separate from the other states (codes 1-5,7-11,13-56). The user can define values as follows:

The user is encouraged to redefine a few more data items and to test them with QUICKTAB. Use the 'List Data Dictionary' option on the Data Dictionary menu to make a listing of '90PUMSX2' and then view some of the value names and definitions. The 'view' file can be printed (90PUMSX2.LST) using 'Print' or even 'View' (See F1 Help in View).

California	06	
Florida	12	
Other States	01	56

If the data on the record = 6, QUICKTAB will add to the 'California' count then exit. If the data <> 6 but = 12, QUICKTAB will add to the 'Florida' count then exit. If the data <> 6, <> 12, but is between 1 and 56, QUICKTAB will add to the 'Other States' count and exit. If none of these are true, QUICKTAB will add to the 'Undefined' count. To summarize, even though '6' is between '1' and '56' it is NOT counted there because it 'hits' the California test first.

Chapter 8. Introduction to QUICKTAB

8.1 What is QUICKTAB?

QUICKTAB is the module of the Integrated Microcomputer Processing System (IMPS), used to quickly analyze the contents of a file. Like other components of IMPS, it requires that the file to be analyzed be described through the IMPS Data Dictionary. Both subject-matter specialists and computer programmers can easily use QUICKTAB. There is no programming language to learn. The system is entirely menu-driven and can be learned in a matter of minutes.

8.2 Common Uses of QUICKTAB

QUICKTAB can be used simply as a table-producing system to analyze the contents of a survey or census file. The tables can be used in publications, if their format is adequate, or they can be used as analysis tools.

8.3 Capabilities of QUICKTAB

QUICKTAB produces two types of tables: frequency distributions and cross-tabulations. Cross-tabulations may be 2-way or 3-way.

8.3.1 Frequency Distributions

Frequency distributions are tables which show the frequency of occurrence of values for a data item. The characteristics of the data item (its name, location in the file, length in characters, etc.) and the expected values for the item are defined during the creation of the Data Dictionary.

Frequency distributions show counts and percentages for each value defined in the Data Dictionary, and for the total of undefined values. A row is generated for each of the defined values, with one more row for undefined values. Frequency distributions can be done on a subset of a file. The user designates the selection criteria, such as women age 12 and older. The frequency distributions can be weighted.

In frequency distributions of quantitative data, such as age, descriptive statistics are also shown. Specifically, these are minimum, maximum, mean, and standard deviation. In order for QUICKTAB to produce such statistics, there can be no value names associated with the first set of the values defined for the item in the Data Dictionary.

8.3.2 Cross-tabulations

Cross-tabulations display relationships between 2 or 3 data items. For instance, QUICKTAB can produce cross-tabulations of age by sex by marital status. The user can elect to display cross-tabulations in terms of actual counts or as percentages of the total. As with frequency distributions, the counts in cross-tabulations may be unweighted or weighted.

QUICKTAB allows for the selection of a subset of the file for tabulation. The tabulations can show only the defined values of data items, or they can show both defined and undefined values.

8.3.3 Weighted Tabulations

Both frequency distributions and cross-tabulations allow the use of a weight value. The weight value must be an item on the same record with the items being tabulated. The weight item may have implied decimal places. All arithmetic is integer; the tabulation results are divided by 10, 100, 1000, etc., depending on the number of implied decimal places in the weight.

Chapter 9. Using QUICKTAB

9.1 Implications of Data Dictionary Value Name Definitions

Within the Data Dictionary definitions, you can define data item value ranges the way you would like to see them in the QUICKTAB tables. Moreover, you can define the same item several times, with a different set of value ranges for each definition. Redefinition is done by defining an item as a subitem.

For frequency distributions, a row is automatically generated for each of the defined values and an additional row for all undefined values. Undefined values are values not included in the definition for a data item in the Data Dictionary.

In creating cross-tabulations, the user may elect to include a column and row for undefined values. Cross-tabulations compare data items within one record type, but may not include data items of different record types. The system warns against crossing an item with itself, but does not prohibit it. Crossing an item by a subitem, and a subitem by another subitem, is also permitted.

Some other implications of value names and value range names are:

(a) If the names are given for all value ranges, the names and ranges appear in the tables at the intervals defined.

(b) If names are given for value ranges, and some names define split ranges, then the items for all parts of the split range will be totaled for the table.

(c) If names are not given for value ranges but there are multiple value ranges, the ranges themselves will appear as the row or column or layer heading.

(d) If names are not given for numeric value ranges, and there is only one value range, and the item is defined as 1 or 2 digits in length, then a row, column, or layer is defined for each value in the range.

(e) If names are not given for numeric value ranges, and there is only one value range. and the item is defined as 3 or more digits in length, then a row or column is defined for a group of values so there will be ten or less groups. For example, if the item is 3 digits long and the range is 000 to 999, then the 10 groups will be 000-099, 100-199, ... 900-999. If the item is 4 digits long and the range is 4731 to 6342, then the 3 groups will be 4731-4999, 5000-5999 and 6000-6342.

(f) If no values are defined for a numeric item, QUICKTAB behaves as though one value range from 0 to the maximum possible value of the item was defined.

9.2 Selection Criteria

QUICKTAB will let you store your selection criteria when running Frequencies and Cross-tabulations so you can recall and modify them later.

9.3 Counts Versus Percentages

Frequency distributions show counts and percentages for each value defined in the Data Dictionary, and for all undefined values. For cross-tabulations, you can show the table cells
as counts or as percentages of the records that were included in the cross tabulation.

In each table, the record counts and percentages are all based on a single record type, the one containing the data item which was selected.

9.4 Title of Report and Name of Tables File

QUICKTAB produces a report of frequency distributions and of cross-tabulations. You may select the default heading for these reports or you may customize your heading. In addition, users may use the default print file names (FREQ.TBL and CROSSTAB.TBL), or designate their own DOS filename.

9.5 Choosing Dimensions for Cross-tabulations

QUICKTAB allows you to specify which data item will be displayed in the rows, which in the columns, and in the case of a 3-way cross-tabulation, for which data item the tables repeat (layer item). At most, 20 columns are allowed. If the data item selected to appear in the columns has more than 20 value names defined, a message is displayed so you may reconfigure the cross-tabulation display to make that data item either appear in the rows or to make it the layer item. Therefore, the column item is usually the only one that might present a problem in terms of the number of values defined.

9.6 Data Items as Weights

For weighted tables, the weight item must be defined in the Data Dictionary as numeric and its contents must be entirely numeric. That is, it cannot include any blank characters, periods, negative signs, or other nonnumeric characters. If a weight item is not entirely numeric, QUICKTAB does not include the record in the tabulation. In other words, instead of increasing the table cell by the value of weight, it increases it by 0.

9.7 Printing Tables from QUICKTAB

Appearing on both the Frequencies and Cross-tabulations menus is an entry called PRINT. This allows an easy method of printing QUICKTAB tables. When you choose PRINT, QUICKTAB asks if your printer is capable of printing box characters which appear in the tables.

These are the solid horizontal and vertical lines in the tables. If you reply 'No', QUICKTAB converts these graphic characters to characters printable on any printer. (If you are not sure if your printer prints these characters, answer 'Yes' and print a single small table, then examine the result.) After you reply 'Yes' or 'No', if there is no printer on-line, QUICKTAB allows you to turn on the printer or cancel the request to print.

Some tables may be wider than 80 characters. If this is the case, then make sure that you are either using a wide carriage printer with wide paper, or put your printer into compressed mode. (The procedure for setting compressed mode varies depending on the type of printer.)

QUICKTAB places as many tables as possible on a page without breaking the table. A table is broken only if the number of rows make it longer than the page length.

9.8 Rounding of Results

When tabulations are weighted, QUICKTAB increases table cells by the value of weight instead of by the value 1 for each occurrence of a response. Weights with implied decimal places are treated as integers when QUICKTAB is passing the data file. For example, if a frequency distribution for SEX is being produced, and a male person record is read which has the weight of 233 with 2 implied decimal places (meaning 2.33), the value 233 is added to the count of males. After the entire file has been read, QUICKTAB divides the male count by 100 to adjust for the 2 implied decimal places. Because of this division, rounding is done. It is done in the standard way. For example, if a weighted count is 7550, division by 100 would result in 76. If a weighted count is 7549, division by 100 would result in 75.

Percentages are also rounded in the standard manner. They are calculated as the number of records having a certain response divided by the total number of records included in the table. They are shown at 1 decimal place.

Chapter 10. A QUICKTAB Example

This section demonstrates the use of QUICKTAB by describing an example which uses the Data Dictionary and data files found on the CD-ROM, these files would normally be found in the directory Q:\TOOLS\DOC (where 'q' is the CD-ROM drive).

10.1 QUICKTAB Frequency

To initiate this QUICKTAB example ENTER the following command from your working directory (C:\WORK):

C:\WORK>**Q:\TOOLS\QT**

When the QUICKTAB introductory screen appears, press any key or enter your network initials if required. (If your computer is connected to a Local Area Network, a User-ID will be requested. Enter one to three characters: letters and/or numbers.) The QUICKTAB main menu will appear. Select Frequencies from this menu.

Next to 'Data dictionary File:' and press ENTER.	type '**Q:\TOOLS\DOC\90PUMSX2**'	
Next to 'Data File:' and press ENTER.	type '**Q:\TOOLS\DOC\PUMSAXXX.DAT**'	
Next to 'Saved settings file:'	LEAVE BLANK	and press ENTER.

The screen presented has the 'Frequencies menu' at the top with the following choices:

Records Tabulate View Print DOS Load Save End

The cursor should be 'highlighting' the 'Records' option.

Records	-	select data items for frequencies
Tabulate	-	produce frequencies: select data file, give title, give name for output file (.TBL always appended)
View	-	look at a file (F1 gives Help Screen)
Print	-	Print a file
DOS	-	Execute a DOS command
Load	-	Recall settings from a previous task (see Save)
Save	-	Save settings for later use (see Load)
End	-	Exit QUICKTAB

As an example, start by choosing 'Records' from the menu, press 'Enter' when 'Records' is highlighted. A window showing the different records defined for the PUMS data appears. The example has two record types, a HOUSING-RECORD and a PERSON-RECORD. Select HOUSING-
RECORD and press 'Enter'. Another window appears that has a list of all the HOUSING-RECORD items. Move the highlight bar down (using the arrow keys) to UNITS1 and press

'Enter'. A small block appears to the left of the entry indicating that it has been selected. Now select TENURE (in 2nd column) the same way.

Next press 'Esc' to return to the record type menu. The window should say that 2 of the 117 items have been chosen from HOUSING-RECORD. Choose the PERSON-RECORD record type. When the list of items appears, press 'F5' to choose all the items. Move the highlight bar to the right to POVERTY (in 2nd column) and hold the 'Enter' key down to deselect POVERTY and all the data items which follow (the last is AINCOME8). Note that the 'Enter' key is a toggle operation, if the item is selected, then 'Enter' deselects it. On the other hand, if the item is not selected, then 'Enter' selects it. Press 'Esc' to return to the record type menu. The window should appear that says 19 from 136 PERSON-RECORD items have been selected. Press 'Esc' to return to the Frequencies menu at the top of the screen.

To generate the frequency distribution report, select the 'Tabulate' option. PUMSAXXX.TXT should appear as the data file, press 'ENTER' to accept. Next, QUICKTAB asks for a report title, which will appear at the top of each page of the frequency report. A default title of the name of the data file is provided. You can use this title or type your own title and press 'Enter'. Next, QUICKTAB asks if the name of the output file should be FREQ (with an automatic extension of .TBL). If you prefer another name for the output table, you can type it here. If a file with that name already exists, QUICKTAB will ask you if you want to overwrite it.

After you enter the name of the output file, the program is executed and the resulting tables are shown on the screen. A row is generated for each of the defined values and an additional row for all undefined values. Undefined values are values other than those defined in the Data Dictionary. The only undefined values for the example indicate negative values for income (see NOTES at end). There may be up to six columns in the report. In each table, the record counts and percentages are all based on a single record type, the one containing the data item which was selected. Use the 'Page Up', 'Page Down', and arrow keys to scroll through the tables. Press F1 for further help in viewing the tables and, in general, you might note that F1 is a 'help' key in the system.

Under 'Frequency:' Explanation of headings.

- Total the number of records which had the indicated value
- Percent the percentage of records which had the indicated value. The base of the percentage is all the selected records.
- % Def. the percentage of records which have defined values. The base of the percentage is all selected records excluding those which have values other than those defined in the Data Dictionary.
- % Valid the percentage of records which have an applicable response. The base of the percentage is all selected records excluding those which have values other than those defined in the Data Dictionary and those which have a NA (not applicable) response.

Under 'Cumulative'

- Total the cumulative number of records
- Percent the cumulative percentage calculated with total number of records selected.

The Percent Valid column is included only if NA (not applicable) is defined for the item in the Data Dictionary.

Most, but not all, data items have value 'names' (for example, the value names for 'SEX' are 'Male' and 'Female') defined for them. Numeric data items without value 'names' (See 'AGE10-YEAR' on the PERSON-RECORD for example) have the following descriptive statistics provided at the end of the frequency distribution table:

- Min the minimum value for the item in the file
- Max the maximum value for the item in the file
- N the number of values in the file
- Mean the average of the values for the item in the file
- SD the standard deviation of the value for the item in the file (uses N-1 as the denominator)

NOTE: All these statistics exclude NA value if present.

Notice that the counts given in this report are just UNweighted frequencies and will depend on the particular example file that is selected. After examining the tables, press 'Esc' to return to the Frequencies menu.

Page 1 13/01/94 15:32:00 IMPS Version 3.1
Data file: PUMSAXXX.TXT

Record: HOUSING-RECORD

Item: UNITS1

Values	Frequency				Cumulative	
	Total	Percent	% Def.	% Valid	Total	Percent
Mobil home/tralr....	-	-	-	-	0	.0
Detached.....	-	-	-	-	0	.0
Attached.....	11	11.0	11.0	11.5	11	11.0
2 Apartments.....	1	1.0	1.0	1.0	12	12.0
3-4 Apartments.....	1	1.0	1.0	1.0	13	13.0
5-9 Apartments.....	-	-	-	-	13	13.0
10-19 Apartments....	1	1.0	1.0	1.0	14	14.0
20-49 Apartments....	3	3.0	3.0	3.1	17	17.0
50+ Apartments.....	78	78.0	78.0	81.3	95	95.0
Other.....	1	1.0	1.0	1.0	96	96.0
NA.....	4	4.0	4.0		100	100.0
Undefined.....	-	-			100	100.0

Item: TENURE

Values	Frequency				Cumulative	
	Total	Percent	% Def.	% Valid	Total	Percent
Owned w/Mortgage....	8	8.0	8.0	9.0	8	8.0
Owned Clear.....	2	2.0	2.0	2.2	10	10.0
Rented for Cash.....	78	78.0	78.0	87.6	88	88.0
Rented Cash-free....	1	1.0	1.0	1.1	89	89.0
NA.....	11	11.0	11.0		100	100.0
Undefined.....	-	-			100	100.0

Page 2 13/01/94 15:32:00 IMPS Version 3.1
Data file: PUMSAXXX.TXT

Record: PERSON-RECORD

Item: AGE5-YEAR

Values	Frequency			Cumulative	
	Total	Percent	% Valid	Total	Percent
0- 4	3	2.2	2.2	3	2.2
5- 9	1	.7	.7	4	2.9
10-14	1	.7	.7	5	3.6
15-19	1	.7	.7	6	4.4
20-24	11	8.0	8.0	17	12.4
25-29	25	18.2	18.2	42	30.7
30-34	18	13.1	13.1	60	43.8
35-39	19	13.9	13.9	79	57.7
40-44	17	12.4	12.4	96	70.1
45-49	7	5.1	5.1	103	75.2
50-54	8	5.8	5.8	111	81.0
55-59	3	2.2	2.2	114	83.2
60-64	3	2.2	2.2	117	85.4
65-69	7	5.1	5.1	124	90.5
70-74	2	1.5	1.5	126	92.0
75-79	4	2.9	2.9	130	94.9
80-84	5	3.6	3.6	135	98.5
85-89	2	1.5	1.5	137	100.0
Topcode	-	-	-	137	100.0
Undefined	-	-		137	100.0

Min	Max	N	Mean	SD
0	86	137	40.4	18.3

Item: AGE10-YEAR

Values	Frequency			Cumulative	
	Total	Percent	% Valid	Total	Percent
0- 9	4	2.9	2.9	4	2.9
10-19	2	1.5	1.5	6	4.4
20-29	36	26.3	26.3	42	30.7
30-39	37	27.0	27.0	79	57.7
40-49	24	17.5	17.5	103	75.2
50-59	11	8.0	8.0	114	83.2
60-69	10	7.3	7.3	124	90.5
70-79	6	4.4	4.4	130	94.9
80-89	7	5.1	5.1	137	100.0
Topcode	-	-	-	137	100.0
Undefined	-	-		137	100.0

Min	Max	N	Mean	SD
0	86	137	40.4	18.3

10.2 Weighted Frequency Counts

The previous selection did not use the 'Weight' option which may be used to inflate the counts to be representative of the actual population. Select 'Records' on the Frequencies menu. Note above the fifth column is the word 'Weight'. Move the cursor under 'Weight' and press 'Enter'. From the list given, select 'HOUSWGT', which is the weight variable for HOUSING-RECORD. The data item, HOUSWGT, should appear below 'Weight' in the record menu. Repeat the process for the P-RECORD choosing PWGT1. Press 'Esc' to return to the Frequencies menu.

Choose 'Tabulate' to rerun the previous frequencies using weights. Confirm the same data file with 'Enter'. Change or confirm the title for the report. Change or confirm the name of the output frequencies report file. When the report is displayed you will notice that all the frequencies have increased.

These are the corresponding 'weighted' frequencies to the counts produced in the previous report. Press 'Esc' to return to the Frequencies menu.

10.3. Record Selection (Defining Conditions for Tabulation)

Suppose that you want the frequency report to include only people 18 years or older who are not attending school. These two conditions together represent a 'universe' Only data that satisfy them will be included in the frequency report.

From the Frequencies menu select 'Records' again. Note the word 'Universe' in column 4 of the menu. The user may specify a 'universe' for each record type. To enter this universe, first select the PERSON-RECORD line then either move the highlight bar to the 'Universe (Y/N)' column or simply press 'Y'. A window appears with a highlight bar in the 'Item' column. Press 'Enter' and select the data item AGE-SINGLE. Choose '>=' from the series of available operators on this item then enter '18'. You should see the following:

AGE-SINGLE >= 18

The highlight bar moves down and is now positioned at the 'And/Or' column. Since your universe consists of two conditions, these conditions must be logically linked with an AND or an OR operation. Press 'Enter' and pick 'AND'. In the same way as before, choose 'SCHOOL' and '='. You are then presented with four options, choose "Not since Feb'90" or press 'Esc' and enter '1', the 'no' code. You have now selected the universe of people who are 18 years or older AND who are not in school. QUICKTAB can handle more complex universes with conditions with inequalities and ranges of values. However, a universe always applies to a specific record type and to a single set of frequency distributions. When using universes, keep in mind that AND operations take precedence over OR operations. QUICKTAB does not have parenthesis available.

After entering your universe for the PERSON-RECORD record type, press 'Esc' to move back to the record select menu/window. Choose the HOUSING-RECORD record type. Press 'Y' to define a universe for the HOUSING-RECORD record. Choose the item ROOMS, the operation '<>' (not equal to) and the ranges FROM 0 TO 4. The ranges are chosen by pressing

'Esc' when the list of value names is displayed. The user is allowed to enter their own range of values. This will cause the frequency report for the selected HOUSING-RECORD items to include only houses that do not have between 0 and 4 rooms. In other words, only houses with 5 or more rooms will be included. Press 'Esc' to return to the Record menu/window and again for the Frequencies menu. To execute the weighted frequencies for these two universes, follow the 'Tabulate' process described above.

After examining the tables, press 'Esc' to return to the Frequencies menu and then select End. Select 'No' when asked if you wish to save the settings. (See next section.)

10.4 Save and Reuse Selection Criteria

You can save your data item and universe selections to disk by using the 'Save' option from the Frequencies menu. Saved selections can be used again later by QUICKTAB. To retrieve saved settings, which are assigned a .QTF file extension, choose the 'Load' option from the Frequencies menu. Press 'F2' key for the selection of .QTF files available or enter the name directly if it is known. This procedure does not reproduce the frequencies associated with these settings, it merely sets the data item selections, the weights, and universes to the values they contained when this .QTF file was saved.

10.5 QUICKTAB Cross-tabulations

Select a Record, then which data items from that record you wish to cross tabulate. The PUMS has two record types, a HOUSING-RECORD and a PERSON-RECORD.

You must first choose which data items you wish to examine and then run the cross-tabulation program. For example, choose 'Records' from the menu line at the top of the screen. Leave the highlight bar over the record type HOUSING-RECORD and press 'Enter'. A window appears that has boxes for three items, a column to select results to appear as Counts or Percentages, and a column to select whether or not to include a row and column in the table to count 'undefined' values.

Press 'Enter' again, and a window appears that has a list of all the HOUSING-RECORD items defined in the Data Dictionary. Move the highlight bar over UNITS1 and press 'Enter'. The name appears as the first item for the cross-tabulation. Press 'Enter' again to return to the list and select ROOMS as the second item for the table. Using the arrow keys, if necessary, move the highlight bar to the third 'slot' on the first line of the window, under the 'Layer item' column and press 'Enter'. Choose KITCHEN and press 'Enter'. You have selected a 3-way tabulation UNITS1 (row) by ROOMS (column) by KITCHEN (layer).

Press 'Esc' to return to the record type selection menu which now lists HOUSING-RECORD as having one table defined. Select PERSON-RECORD and define a table as RELAT1 by SEX, do not select a layer. Set the 'undefined' column to 'N'. Now define a table as RELAT1 by MARITAL also setting the 'undefined' column to 'N'. Define another table as RELAT1 by MARITAL this time leaving the 'undefined' column as 'Y'. Define yet another table as RELAT1 by MARITAL. This time press 'P' to select Percentages in the 'Count/Percent' column.

Press 'Esc' to return to the record type selection menu, which should now have one table defined for HOUSING-RECORD and four tables for PERSON-RECORD. If desired, you can define a universe or select the weight HOUSWGT for house records or PWGT1 for person records as described above for frequency distributions. Press 'Esc' again to return to the Cross-tabulations menu.

Choose 'Tabulate' from the menu. Select a data file as above ('F2' and/or a path). QUICKTAB asks for a report title, which will appear at the top of each page of the cross-tabulation report. A default title of the name of the data file is provided. You can use this title or type your own title and press 'Enter'. Next, QUICKTAB asks if the name of the output file should be CROSSTAB (with an automatic extension of .TBL). If you prefer another name for the output table, you can type it here. If a file with that name already exists, QUICKTAB will ask you if you want to overwrite it or not. The program is executed and the resulting file with the list of tables is shown on the screen.

The first table is UNITS1 by ROOMS layered by KITCHEN. The table cells contain counts of HOUSING-RECORD records. The first layer is the first set of rows and represents the total layer. The following four layers contain counts for households with certain values for KITCHEN; that is, they have a complete kitchen (Yes), do not have a complete kitchen (No), have this field 'NA', or have this field 'Undefined'. The first row in each layer is the total of households in that layer. The subsequent rows are the various values for UNITS1. The last row is the 'Undefined' counts (which should be zero (-). The first column is also a total. The next columns are counts of values of ROOMS in the Data Dictionary. The last column is the 'undefined' count.

The second table is RELAT1 by SEX. The table contain counts of PERSON-RECORD records. The first row is the total. The remaining rows are counts for the different relationships defined in the Data Dictionary. The columns are the total and values for sex as defined in the Data Dictionary. As requested, there is no row or column for 'undefined' values.

The third, fourth, and fifth tables are all for RELAT1 by MARITAL except that different options have been selected for each. The third table does not have an 'Undefined' row (at the bottom) or column (on the right) while the fourth table does. The fifth table is like the fourth except that the cells contain the percentages of records instead of the counts. After examining the tables, press 'Esc' to return to the Cross-tabulations menu and then select End.

NOTES:

- Some of the income values (person and household) may be negative, for example, INCOME2, RHHINC, etc. Negative values CANNOT be tabulated directly by QUICKTAB. To determine counts for negative values, the user must create a cross-tabulation of a value field by a 'sign' field. The column under the 'negative' heading gives the distribution of the negative incomes. The 'positive' column IS NOT ACCURATE. Use the full data item for distribution of the nonnegative values in which the negative values will appear in the 'Undefined' row. For example, select INCOME2 for a frequency distribution. The result is for nonnegative incomes. To obtain a distribution of negative incomes, select INCOME2 by INCOME2SGN as a cross-tabulation. The 'negative' column in the result gives the distribution of the negative incomes. The 'total' and 'positive' columns are NOT always accurate.

- QUICKTAB has some 'Help' screens which are displayed by pressing 'F1'. It has many features which are not discussed here. The best approach is to take a few minutes and 'play/practice' with it. Read the 'Help' screens and try some of the options.

- If the 'Box Characters' do not print correctly on your printer, the font needs to be changed. QUICKTAB uses the IBM PC character set. For HP laser printers the symbolic set 'PC-8' can be used.

Appendix A. Data Dictionary System Limits

Term	Valid Values
IMPS long name length	1 to 16 characters
IMPS short name length	1 to 8 characters
Length of record type code	1 to 12 characters/digits
Record types in Data Dictionary	0 to 98
Items in Data Dictionary	0 to 1000
Items per record type	1 to 1000
Common items in Data Dictionary	0 to 100
Values in Data Dictionary	0 to 2800
Number of lines on the values menu	0 to 1000
Number of values copied at one time	1 to 2000
Number of items moved at one time	1 to 2000
Record length	1 to 9999 characters
Records in any one record type (Max)	1 to 5000
Length of alphanumeric item	1 to 9999 characters
Length of numeric item	1 to 15 digits
Decimal places of numeric item	0 to 3
Item occurrences	1 to 9999

Appendix B. QUICKTAB System limits

FREQUENCY Tables

Max number of rows	=	500
Max number of item values	=	4,090
Max number of items selected		= 1,200
Max lines of universe logic	=	100

CROSSTAB Tables

Max number of rows	=	500
Max number of columns		= 20
Max number of cells	=	8,000
Max number of item values	=	4,090
Max number of tables	=	100
Max lines of universe logic	=	100

Appendix C. Data Dictionary Run-time Error Messages

Can't find Dictionary.

Cause - Requested Data Dictionary not in working directory.

Action - Check spelling and path name of file name.

Can't read Dictionary.

Cause - Data Dictionary cannot be read. DOS error or Data Dictionary corrupted beyond normal recovery.

Action - Check to see if the correct file is being used. If the Data Dictionary is unreadable, use the Recover Data Dictionary Utility to restore the Data Dictionary from a previously backed up Data Dictionary instructions file (.DI).

DISK ERROR -- Dict not saved.

Cause - Disk error. Not a Data Dictionary problem. The operating system has indicated to the program that the specified file name could not be saved. Problem could be a full disk.

Action - Check the integrity of the disk.

DISK ERROR -- Can't back up Dict.

Cause - Disk error. Not a Data Dictionary problem. The operating system has indicated to the program that a backup (.DD!) for the specified file name could not be saved.

Action - Check the integrity of the disk.

DISK FULL -- Dict not saved.

Cause - The operating system has indicated to the program that there is insufficient space on the disk to save the Data Dictionary.

Action - Use another disk or remove some files from the disk being used.

DISK FULL -- Can't back up Dict.

Cause - The operation system has indicated to the program that there is insufficient space on the disk to save the internal system Data Dictionary backup.

Action - Use another disk or remove some files from the disk being used.

Maximum record types is 98.

Cause - Attempt to create more than 98 record types.

Action - Reduce number of record types.

Maximum items is 1000.

Cause - Attempt to define more than 1000 items.

Action - Reduce the number of items.

No records defined.

Cause - No records where defined in current Data Dictionary.

Action - Define at least one record.

Invalid IMPS name.

Cause - Maximum of 16 character name can be defined. Only letters, digits and '-' are allowed. The first character must be an alphabetic letter and the last character must not be a '-'.

Action - Use another name.

Duplicate record type value.

Cause - The record type value has already been assigned to another record type.

Action - Use another value or change the value of the other record type.

Duplicate name (record type).

Cause - The record type name has already been assigned to another record type.

Action - Use another name or change the name of the other record type.

Duplicate name (COMMON item).

Cause - The common item name has already been assigned to another item.

Action - Use another name.

Duplicate name (item in '(record name)'.

Cause - The record type name has already been assigned to another record type.

Action - Use another name.

Items and subitems can't both occur.

Cause - Cannot have occurrences on both an item and redefining subitems.

Action - Remove occurrence of item or subitem.

Decimals must be <= length.

Cause - Decimal length must be less than, or equal to, item length. The item length includes the decimal length.

Action - Reduce decimal length or increase item length.

Duplicate value name.

Cause - Value names within an item must be unique.

Action - Use another name.

'From' value must be < 'to' value.

Cause - The value range specified the lower limit as greater than or equal to the upper limit.

Action - Check the range limit and correct.

Max length for numeric item is 15.

Cause - The length specified for a numeric data item was more the 15 digits.

Action - Make the item length between 1 and 15.

Subitem extends past end of item.

Cause - Subitem length cannot go beyond last character of corresponding item.

Action - Change the length of the subitem so that is falls within the item or make the item length longer.

Can't mix BLANK with a value.

Cause - BLANK cannot be part of a value range.It must be a separate value.

Action - Remove BLANK from range or have BLANK specified as a single value.

Maximum records must be <= 5000.

Cause - Maximum number of records of any one record type must be <= 5000.

Action - Change the value to be between 1 and 5000, inclusive.

Maximum record length is 9999.

Cause - The maximum record length must be less than 10,000 characters in length.

Action - Reduce the record length by removing items from record type. Also you may attempt to move excess data items from current record type to another record type, if possible.

No room for more values.

Cause - The Data Dictionary already contains the maximum number of values.

Action - Use discreet values or ranges more sparingly.

Values to append have length xxxx.

Cause - Cannot copy values to an item with a smaller item length.

Action - Enter the values individually.

Can't place item(s) here -- overlap.

Cause - Layout mode placement error.

Action - Move intervening field.

Can't move; intervening COMMON field.

Cause - Layout mode placement error.

Action - Move intervening field.

This name is reserved for IMPS use.

Cause - IMPS reserved names cannot be used.

Action - Use another name.

Appendix D. IMPS File Extensions

Data Dictionary

Extension	File Description
.DD	Data Dictionary file.
.DD!	Backup file of .DD created each time the Data Dictionary is edited.
.DI	Backup Data Dictionary instructions file of .DD created each time the Data Dictionary is edited.
.LST	Data Dictionary listing file.

QUICKTAB

Extension	File Description
.QTF	Settings for a set of frequency distributions
.QTX	Settings for a set of cross-tabulations
.TBL	Print file containing tables.

Appendix E. Utilities

In order to use any of the utilities, choose 'Utilities' from the appropriate menu.

E.1 Convert Box Characters

This utility converts 'line draw' characters (used with laser printers) to 'line printer' characters (used with nonlaser printers).

E.2 Merging Data Dictionaries

The merge utility permits the joining of two Data Dictionary files. These Data Dictionary files can contain definition of records developed by different members of a Data Dictionary design team. After each team member completes his/her task, the Data Dictionary parts can be combined to a master Data Dictionary. This utility cannot be used to add new items to an existing record type.

In order to use this utility, follow these rules:

- If COMMON data items are defined, they must have identical names, locations, and lengths.
- If only one Data Dictionary has COMMON data items defined, it must be the first Data Dictionary.
- If both Data Dictionaries have record types, the record types must be in the same starting position and be of the same length.
- If only one Data Dictionary has record type defined, it must be the first. The second Data Dictionary will be assigned the next available number.
- If neither Data Dictionaries have record type defined, the first Data Dictionary will be assigned '1' and the second will be assigned '2'.
- Record type names, codes, and item names must be unique.

If any of the above rules are violated, the Data Dictionaries will not be merged. Of course, system limits such as maximum record types must also be obeyed for the Data Dictionary merge to be successful.

To run the merge utility, select the 'Merge Dictionaries' option from the Utilities Menu of Data Dictionary.

The merge utility can also be used to update record types in a Data Dictionary. Individual record types can be developed or modified in a separate Data Dictionary file and later added to the master. The process involves the deletion of the record type in the master Data Dictionary and adding the updated record type.

E.3 Recovering a Data Dictionary

Data Dictionary files may become corrupted due to hardware, system, or user error. Occasionally, the user may also encounter an 'Integrity violation' message while developing a Data Dictionary. As a safeguard, a Data Dictionary instructions file is generated by the system. This instructions file contains a backup of the saved Data Dictionary when the developer was last exited normally. The file has the extension '.DI'. Should the Data Dictionary file (.DD) become corrupted, you can regenerate the .DD file from the .DI file by running the Data Dictionary recovery utility.

The layout of the .DI file is exactly the same as the Data Dictionary listing except that headers and blank lines are commented (period in position 1) and formfeeds a not generated. The Data Dictionary instructions file uses fixed positional parameters.

Users should not attempt to modify the .DI file. The system assumes a very strict format and does very little checking of instructions.

To run the recovery utility, select the 'Recover Dictionary' option from the Utilities Menu of Data Dictionary.