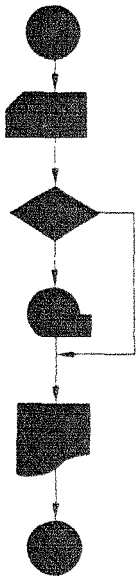
THE NORTHEASTERN FOREST-INVENTORY DATA-PROCESSING SYSTEM. IX. OPERATION OF SUBSYSTEM OUTPUT.



by
Robert C. Peters and
Robert W. Wilson Jr.

U. S. FOREST SERVICE RESEARCH PAPER NE-77
1967

NORTHEASTERN FOREST EXPERIMENT STATION, UPPER DARBY, PA.
FOREST SERVICE, U.S. DEPARTMENT OF AGRICULTURE
RICHARD D. LANE, DIRECTOR

About the Authors

ROBERT C. PETERS obtained his Bachelor's degree from the University of California in 1960 and his Master's at Yale University in 1961. He joined the Forest Service in 1961 as a research forester, and was assigned to the Station's biometrics unit from 1961 until 1965, when the unit was discontinued. Mr. Peters played a key role in the development of the data-processing system reported here.

ROBERT W. WILSON, JR. took his Bachelor's degree at The Pennsylvania State University in 1947 and his Master's and Ph.D. degrees at Yale University in 1948 and 1965, respectively. He joined the U. S. Forest Service in 1948 and has worked in various research capacities for the Northeastern Forest Experiment Station. From 1961 to 1965 he was in charge of the Station's biometrics unit at New Haven, Conn. He is assigned at present to the Forest Insect and Disease Laboratory at West Haven, Conn.

THE NORTHEASTERN FOREST-INVENTORY DATA-PROCESSING SYSTEM. IX. OPERATION OF SUBSYSTEM OUTPUT.

■

Contents

PREFACE

THIS paper is the ninth in a series of ten papers prepared to describe the forest-inventory data-processing system of the Northeastern Forest Experiment Station. This system was devised for using modern, large-scale, high-speed computers in processing forest-inventory data. The series will comprise the following papers:

I. Introduction.
II. Description of subsystem EDIT
III. Operation of subsystem EDIT.
IV. Information for programmers — subsystem EDIT.
V. Description of subsystem TABLE.
VI. Operation of subsystem TABLE.
VII. Information for programmers — subsystem TABLE.
VIII. Description of subsystem OUTPUT.
IX. Operation of subsystem OUTPUT.
X. Information for programmers — subsystem OUTPUT.

IX-A. INTRODUCTION

ONE of the major projects of the U.S. Forest Service is a nationwide forest survey, which is designed to obtain useful and timely information about the timber resources of the United States. In the course of the surveys, which are made mainly on a state-by-state basis, great masses of detailed data are collected about timber volumes, growth, timber cut, and other characteristics of the timber resource.

In recent years the volume of information obtained from forest-survey field plots has increased greatly. The task of compiling and analyzing this mass of data with mechanical computing machines was both cumbersome and time-consuming.

A solution to this problem was seen in the development of the high-speed electronic computers. The Northeastern Forest Experiment Station, which was responsible for conducting the forest survey of the heavily forested Northeastern States, investigated the possibilities and devised the Northeastern Forest-Inventory Data-Processing System.

This paper tells about the operation of a part of the system. Program OUTPUT is designed for use in conjunction with program TABLE (see part V of this series). Its sole function is to produce and print fully labelled tables of statistics for sampled populations from tabular summaries produced in program TABLE. Detailed instructions for solving estimation problems with the standard version of the program are given in the following chapters.

A general description of the program outputs, inputs, logic and procedures is given in part VIII of this series. The material cov-

ered there should be thoroughly digested before attempting to use the program. Presumably, the user will already be fully familiar with the use of program TABLE.

The program is written in the standard IBM FORTRAN IV language, and is operative at the Yale University Computer Center on an IBM 7094/7040 Direct Coupled System under the IBSYS DCS operating system with IBJOB processor.[1] It will operate with little or no modification on other comparable systems. Part X in the series contains a selection of programming information that will be useful if the standard version of the program must be modified for any reason.

Copies of these publications and information on the FORTRAN IV program decks may be obtained from the Northeastern Forest Experiment Station, 6816 Market Street, Upper Darby, Pennsylvania 19082.

In what follows, the concept of a "job" as the unit of processing is completely arbitrary. It depends on the nature of the estimating problem and the user's preferences. Any number of jobs may be included in a processing run. The only requirement is that within each job the same processing and output options, as expressed through the job control cards, sec. 410, apply to each population processed in the job; and that all populations for which the tables of output statistics are to be summed to population group totals must be contained in a single job.

Similar remarks apply to the concept of the "population" and its relation to the input data. Ordinarily, the population will be defined in its normal statistical sense, but in certain circumstances, such as estimation for parts of a "normal" population under appropriate assumptions, the population may be defined as the part for which estimates are required. Similarly, it may be defined as a group of "normal" populations. The only requirement is that the processing specifications be consistent with the definitions. It should also be noted that the same body of input data can be used repeatedly, even within the same job, in these special cases.

[1] Mention of a particular product should not be construed as an endorsement by the Forest Service or the U. S. Department of Agriculture.

IX-B. CONTROL CARD FORMATS

The description and specification of each processing job in a processing run is presented to the computer through a special deck of data cards referred to as the control deck. Each card in this deck contains specific pieces of information arranged in a definite format.

In this chapter each type of control card is described. The description gives the format of the cards, the information they must contain, and — where appropriate — the purpose and use of the required information. Consequently, this chapter may be used both as a detailed list of instructions for coding the description of a job, and as an outline to follow in the initial stages of job specification in order that the specifications be complete.

Run Title Card (Sec. 400)

This card must precede the first job control deck in the file. It simply gives a descriptive title to be printed at the top of each of the job summary pages.

Run Title Card — Item 401

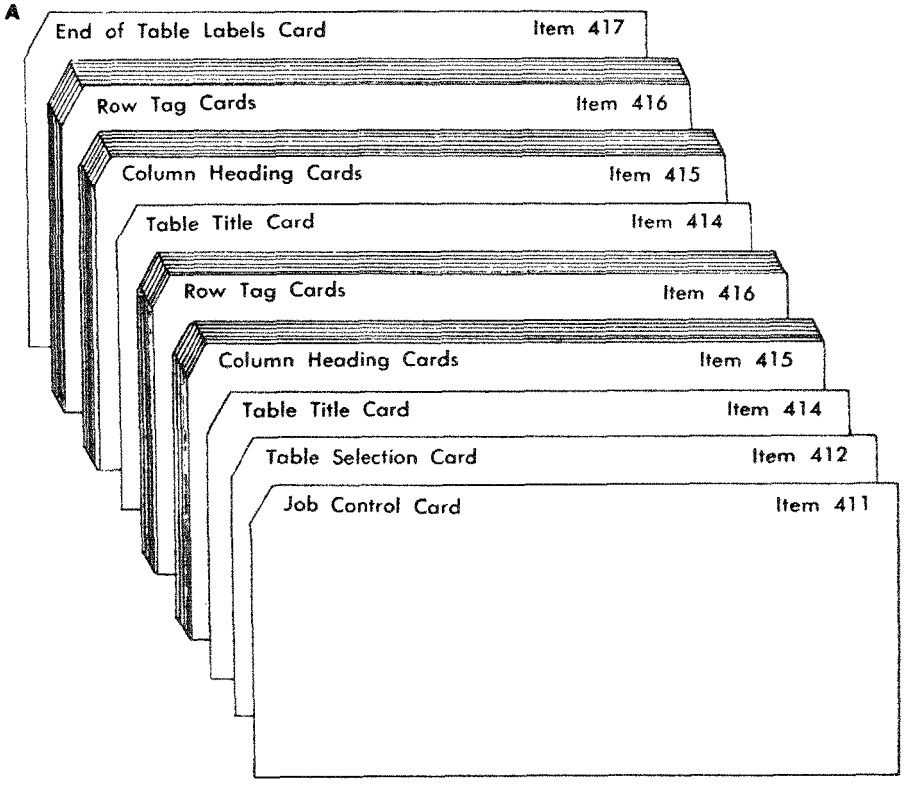
Columns	*Contain* —	*Explanation*
1-72	AAA . . . A	= 72 alphameric characters, giving a descriptive title for the entire run.

Job Control Cards (Sec. 410)

The job control cards are the header cards for each job in a processing run. They contain the general specifications for the job as a whole (including table labels). Each group of job control cards in the control deck is followed by other groups of cards (described below) that give detailed descriptions of job segments.

The first card in the group, item 411, gives the number of populations to be processed in the job, the processing option to be used, and the output options. The second card, item 412, extends the output options by showing which of the sample summary tables given as input are to be processed and printed. These two cards must always be in the control deck.

The remaining cards of this group are used to provide labels



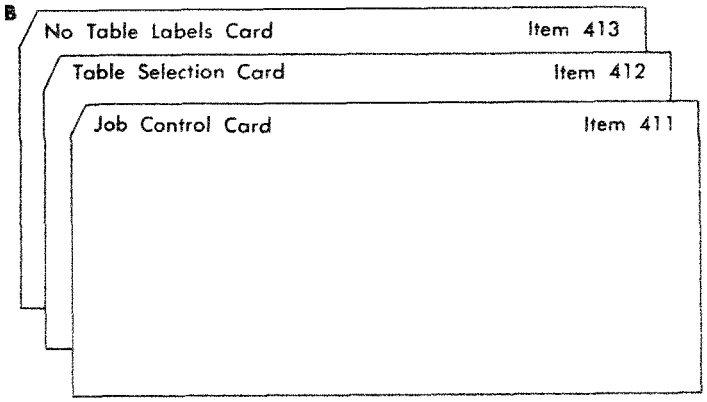


Figure 1. — Order of job control cards in control deck: **A**, the standard setup for two output tables; **B**, optional setup which may be used for any job in a run after the first.

for each table to be printed as output. These labels must always be supplied for the first job in a processing run, but they may be omitted in any subsequent job. The last set of labels read will always be used when labels are omitted for a job. This condition requires that the no table labels card (item 413) be the third and final card in the group of job control cards (fig. 1-B).

The table labels are supplied in complete sets, table by table, and in the order in which the tables occur in the input data file (fig. 1-A). A set of labels for a table consists of a table title card (item 414) that contains the table name and a descriptive title; a column heading card (item 415) for each column in the table (including the column of row totals at the extreme right of the table); and a row tag card (item 416) for each row of the table (including the row of column totals at the bottom of the table).

The entire group of table labels must always be followed by the end of table labels card (item 417).

Job Control Card — Item 411

Columns	*Contain —*	*Explanation*
1-11	JOBbCONTROL	= Card label.
12	b	
13-15	XXX	= 3 numeric characters, giving the number of populations to be processed in this job. The number must be right-justified in the field.
16-18	000	= Do not sum the population tables over all populations in the job.
	001	= Sum the tables and print tables of these sums.
19-21	001	= The data input is to be processed as a 100 percent inventory.
	002	= The data input is to be processed as a simple, random sample, using the direct-estimating procedure.
	003	= The data input is to be processed as a stratified, random sample with known weights, using the direct-estimating procedure.
	004	= The data input is to be processed as a stratified double-random sample with estimated stratum weights, using the direct-estimating procedure.

-	005	= The data input is to be processed as a stratified double-random sample with estimated stratum weights, using the separate weighted ratios estimating procedure.
-	006	= The data input is to be processed as a stratified double-random sample with estimated stratum weights, using the combined ratio estimating procedure.
22-24	000	= Do not produce tables of standard error estimates from the tables of variance estimates. This option must be taken if option 001 is taken in columns 19-21, because neither the variance nor the standard error estimates is appropriate to that option.
-	001	= Produce and print tables of standard error estimates, expressed as percentages of the tables of totals.
-	002	= Produce and print tables of standard error estimates, expressed in the same units as the tables of totals.

Table Selection Card — Item 412

Columns	*Contain —*	*Explanation*
1-12	TABLEbSELECT	= Card label.
13	b	
14	0	= Do not print the first table in the data input.
-	1	= Produce and print the first table in the data input. This option may be exercised sequentially for up to 67 tables in the data input, using the following columns:
15-80	X	= Repetitions of the column 14 format, giving the print option for the remaining tables in the data input.

No Table Labels Card — Item 413

Columns	*Contain —*	*Explanation*
1-9	NObLABELS	= A control word, signifying that table labels are not in the control deck for this job. This card cannot be used for the first job of a run since the labels must be present for that job. However, subsequent jobs may use the same labels, in which case this card may be used.

Table Title Card — Item 414

Columns	*Contain —*	*Explanation*
1-6	AAAAAA	= 6 alphameric characters, giving a unique name by which an input-output table may

			be referenced. The name must be left-justified in the field, and must appear exactly as given in columns 14-19 of the output table definition card (item 321 in part VI).
7-78	AAA . . . A	=	72 alphameric characters, giving a descriptive title for an output table. This title will be printed on each page of output for the table.

Column Heading Card — Item 415

Columns	*Contain —*		*Explanation*
1-20	AAA . . . A	=	20 alphameric characters, giving a label for a column in an output table. The label should be centered in the field. There must be one card for each column of the table named in columns 1-6 of the appropriate table title card (item 414), and the last card of the group must be the label for the column of row totals.

Row Tag Card — Item 416

Columns	*Contain —*		*Explanation*
1-20	AAA . . . A	=	20 alphameric characters, giving a label for a row on an output table. The label should be left-justified in the field. There must be one card for each row of the table named in columns 1-6 of the appropriate table title card (item 414), and the last card of the group must be the label for the row of column totals.

End of Table Labels Card — Item 417

Columns	*Contain —*		*Explanation*
1-13	ENDbOFb LABELS	=	A control word, signifying the end of the deck of table labels. This card must always be present unless a no table labels card (item 413) has been used.

Population Description Cards (Sec. 420)

The control cards described in this section contain all of the information that is relevant in describing a population as a whole. There must be one group of these cards for each population to be processed in a job. The arrangement of the individual cards in the group for one population is shown in figure 2.

The population title card (item 421) must always be the first

card in the group. It simply contains a descriptive title for the population that will be printed at the top of every page of output for that population.

The data input identification card (item 422) must always be the next card in the group. It contains the identification of the first sample to be processed for the population, exactly as it is in the input data file (see part VI-B, sec. 330 and part VII-E). The identification is used to search the input file for the required sample summary tables. If there is more than one sample for a population, it is assumed that the input data for the additional samples follow immediately after the identified sample in the input file.

The expansion factor card (item 423) is the third card in the group, and it too must always be included. In addition to the expansion factor by which the population mean is to be multiplied, it contains the number of samples to be processed for the population, and the sum of the sample weights (see sec. 430).

Finally, if and only if processing option 6 is being used, the group of population description cards must contain a set of cards (item 424) each of which contains the independent estimates of

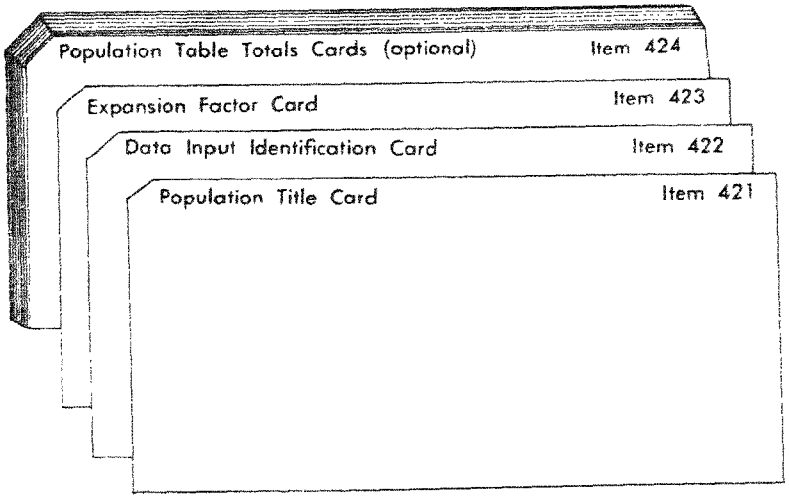


Figure 2. — Order of the population description cards in the control deck. A set of cards like this must be available for each population in a job using processing option 6. With other processing options, the last set of cards (item 424) are not used.

the population grand mean and grand variance for one of the output tables. These cards must be ordered as are the equivalent sample summary tables in the input data file.

Population Title Card — Item 421

Columns	*Contain* —		*Explanation*
1-72	AAA . . . A	=	72 alphameric characters, giving a description of a population.

Data Input Identification Card — Item 422

Columns	*Contain* —		
1-11	INPUTbIDENT	=	Card label.
12-13	bb		
14-21	XXX . . . X	=	8 numeric characters, giving the first identification word in the identification record of the first sample in the block of input data for the population described in the corresponding population title card (item 421). The word must appear exactly as it does in the data input (see part **VII-E**). The remaining 4 identification words in the record must be reproduced in the following columns of this card. This identification record is used to search the input tape for the block of data to be processed.
22-53	XXX . . . X	=	Repetitions of the columns 14-21 format, giving the remaining identification words in the first identification record in the block of input data for a population.

Expansion Factor Card — Item 423

Columns	*Contain* —		*Explanation*
1-12	EXPANb-FACTOR	=	Card label.
13	b		
14-28	bX.XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the value of the expansion factor (normally the size of the population that has been sampled) by which every cell of every output table for the population will be multiplied. This value is referred to by the symbol "wt" in the summary of estimating procedures (part X-D).
29	b		
30-32	XXX	=	3 numeric characters, giving the number of sampling strata in the population. The

			number must be right-justified in the field. If the sampling option given in columns 19-21 of the job control card (item 411) is 002, then this field must contain 001.
33	b		
34-48	bX XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the sum of the sample weights. If the sampling option given in columns 19-21 of the job control card (item 411) is 001 or 002, then this value must be 0.00000000E 00. This value is referred to by the symbol "n" in the summary of estimating procedures.

Population Table Totals Card (optional) — Item 424

Columns	*Contain —*		*Explanation*
1-12	TABLEbTOTALS	=	Card label.
13	b		
14-28	bX.XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the mean value for a population of the grand total cell of an output table. This value is referred to by the symbol "t" in the summary of estimating procedures (part X-D).
29-	b		
30-44	bX.XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the variance of the mean value in columns 14-28. This value is referred to by the symbol "vt . . ." in the summary of estimating procedures.

Sample Description Cards (Sec. 430)

The control cards described in this section appear in the control deck only if the information they contain — sample weights, and independent estimates of the sample means and variances of the grand total cell of each output table — is required by the processing option being used. If these cards are required, they must appear as a set following each set of population description cards (sec. 420). The arrangement of the individual cards in the group is shown in figure 3.

The sample weight card (item 431) is not used when processing option 1 or 2 is specified in columns 19-21 of the job control card

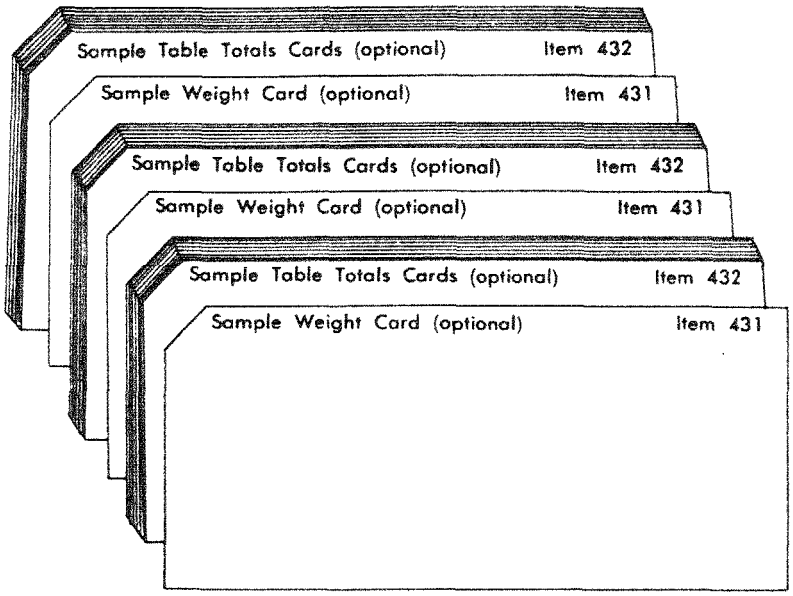


Figure 3. — Order of the sample description cards in the control deck. The setup for a population with three samples, using processing option 5, is illustrated.

(item 411). Otherwise, there must be one card of this type for each sample in a given population. Each card contains the weight by which sample values will be multiplied before summing to population values. It also contains a second weight that may be used when estimates for a segment of the population are being made. The order of these cards in the set must be the same as that of the sample summaries in the data input file.

The sample table totals card (item 432) is used only when processing option 5 is specified in columns 19-21 of the job control card (item 411). It contains the independent estimates of sample (or stratum) mean and variance for one output table. There must be as many of these cards in the set for a sample as there are output tables, and they must be ordered in the set according to the order listed on the table selection card (item 412). The set for each sample immediately follows the corresponding sample weight card (item 431) for that sample.

Sample Weight Card (optional) — Item 431

Columns	*Contain —*		*Explanation*
1-12	STRATUMbWTS.	=	Card label.
13	b		
14-28	bX.XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the value of the weight to be applied to the data for a sampling stratum. This value is referred to by the symbol "n_j" in the summary of estimating procedures (part X-D).
29	b		
30-44	bX.XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the value of the adjustment factor to be applied to the variances for a sampling stratum when compiling estimates for a fraction of the population. This value is referred to by the symbol "n/n_j" in the summary of estimating procedures. If the value is O.OOOOOOOOEbOO, it indicates that estimates for the population as a whole are being compiled, so the adjustment factor will not be applied.

Sample Table Totals Card (optional) — Item 432

Columns	*Contain —*		*Explanation*
1-12	TABLEbTOTALS	=	Card label.
13	b		
14-28	bX.XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the mean value for a sampling stratum of the grand total cell of an output table. This value is referred to by the symbol "$\bar{t}_j$" in the summary of estimating procedures (part X-D).
29	b		
30-44	bX.XXXXXXXXEb-XX	=	A 15-character numeric field (E specification), giving the variance of the mean value in columns 14-28. This value is referred to by the symbol "$v\bar{t}_j$" in the summary of estimating procedures.

Population Group Title Card (Sec. 440)

This card is used in the control deck only if sums over all populations in the job (see item 411, columns 16-18) are required as

output. It gives a descriptive title for the group of populations represented by these sums. The title is printed at the top of every page of the "sums" output. If used, the card is placed in the control deck following all other control cards for a given job.

Population Group Title Card (optional) — Item 441

Columns	*Contain —*		*Explanation*
1-72	AAA . . . A	=	72 alphameric characters, giving a descriptive title for the tables of sums over populations in the group.

End of Run Card (Sec. 450)

This card must always be the last card in the control deck to show that there are no more jobs to be processed.

End of Run Card — Item 451

Columns	*Contain —*		*Explanation*
1-10	ENDbOFbRUN	=	A control word, signifying the end of the control deck.

IX-C. OPERATING INSTRUCTIONS

The use of program OUTPUT to obtain population statistics by processing sample summary data is covered in the information given below. To facilitate checking the setup of processing runs, some of the information given is a resume of material covered elsewhere.

Program Restrictions

The standard version of the program carries limitations on the overall size and on certain dimensions of processing problems that can be handled in a single-processing run. They are:

1. The total number of summary tables for a given sample in the data input file cannot exceed 40.
2. The total number of storage locations available to produce output tables of statistics for a given population is 15,000. The limitations on the numbers of cells in all output tables for the population is more stringent and depends on the processing option being used (see part X-B).

3. The number of rows in an individual output table cannot exceed 50.
4. The number of columns in an individual output table cannot exceed 50.
5. The number of samples processed for a given population cannot exceed 999.

These restrictions result primarily from the way in which the available storage capacity of the computer has been allocated to various uses in the standard version of the program. However, the program has been constructed so that the more important of these allocations can readily be changed if a problem of substantially different relative dimensions is encountered. The modification of dimensioned space is described in part X.

Control Deck Setup

The control deck consists of all the punched cards through which processing specifications, necessary constants, and other data (exclusive of the data to be processed) are entered into the computer. These cards, and the logical groups into which they fall, have been described in the previous chapter. The assembly of the groups of control cards to form the control deck, as well as the placement of the control deck in the monitor input deck, are shown in figure 4.

It should be noted that the monitor input deck consists of the program deck, followed by the control deck, with system control cards interspersed. The latter cannot be described in detail here because they will vary from one computer installation to another. For more information about them, see the systems representative at the computer center where the processing will be done.

Input Data Setup

The normal data input is a magnetic tape file written in the binary mode and containing all the tables of sample statistics to be processed in a given run (see part VII-E of this series). The tables will be grouped in known order (see the order of the output table definition cards in the control deck for program TABLE)

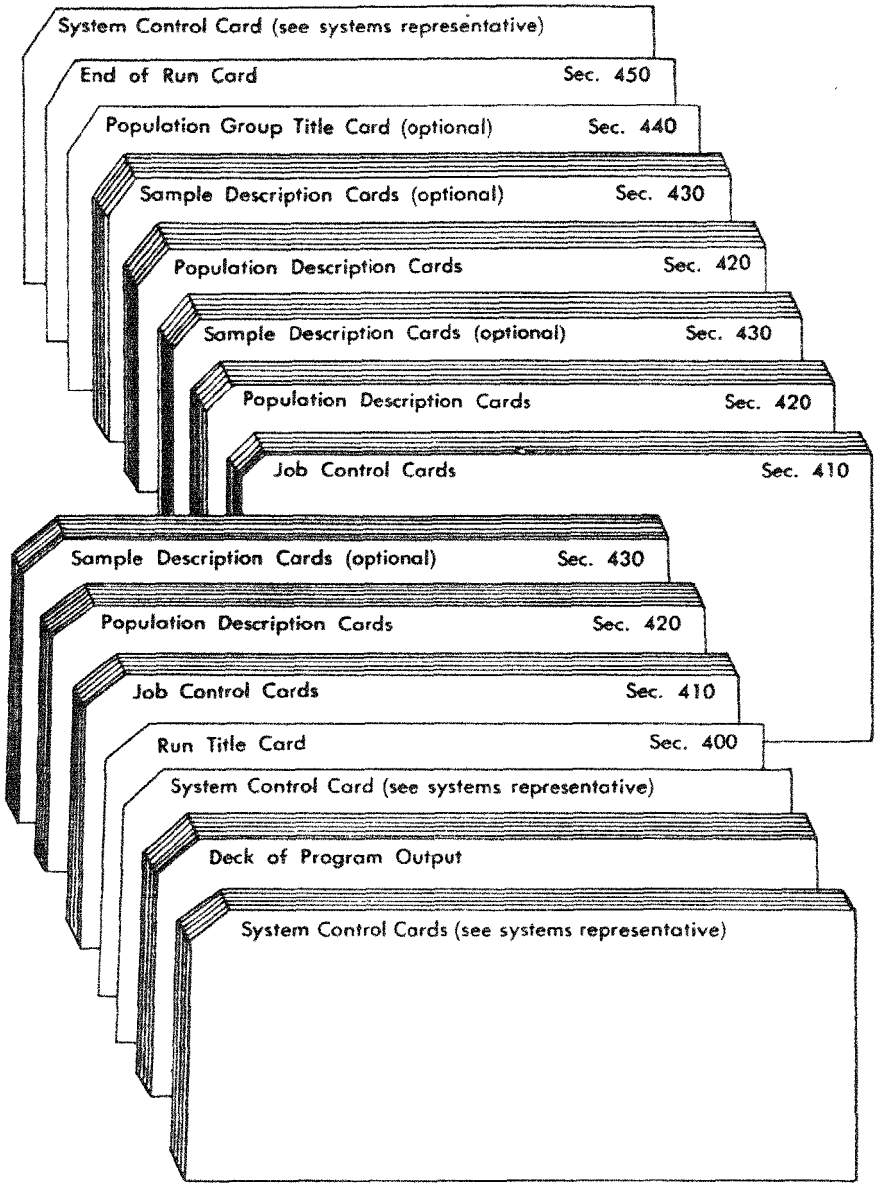


Figure 4. — The control deck setup, illustrating the kinds of cards that are necessary and the order in which they must be arranged. The example is for a run that contains two jobs, the first requiring statistics for one population, and the second requiring statistics for two populations and the sum of the two groups.

within each sample, and the samples will be grouped in known order (see the order of the data sets in the data input file to program TABLE) within the population. Because the input data file is searched for the proper data each time a population is to be processed, the population data may be in any order and the file may also contain extraneous data which will be skipped in processing. However, all data to be processed must be contained in a single file. If it is in multiple files, then it must be processed in multiple passes.

In addition to the above requirements for the input data, they must have been produced in program TABLE under an output option that is consistent with the processing option to be used in program OUTPUT. The necessary relationships are:

OUTPUT processing option	TABLE output option	Sample summary table contain
1	1 or 2	Sample sums or means.
2	3	Sample means and their variances.
3	3	Sample means and their variances.
4	3	Sample means and their variances.
5	4	Sample means, their variances and covariances of individual means with grand means within the table.
6	4	Sample means, their variances, and covariances of individual means with grand mean within each table.

Tape Assignments

In the standard version of program OUTPUT the FORTRAN logical tape assignments are as follows:

Unit	*Use*
3	Scratch tape.
5	Monitor input for program deck and control deck.
6	Monitor print for job summary.
12	Output printed tables of population statistics.
19	Input of sample summaries written in the binary mode.

These tape assignments can be changed to fit local conditions by loading appropriate file routines with the program. See your systems representative or the section entitled FORTRAN files in the IBM IBJOB processor manual, file number 7090-27.

Use of Sense Switches and Sense Lights

No sense switches are used in the program. All sense switches will be set at normal monitor settings. No sense lights are used.

Use of Program Halts

There are no halts in program OUTPUT.

Use of Overlay Feature

The program is not constructed to use the overlay feature. All subprograms must be loaded at once.

Messages Printed During Execution

The messages listed below are those printed by the program during execution. Each message, with its consequences, is described; and the appropriate action, if any, is indicated.

Other messages may also appear in the printed summary of the run. They will be produced by the operating system under which this program is being executed. For the meaning and consequences of any message not found in the list below, see your computer systems representative.

TABLE SELECT CARD INCORRECT

Message 1. Message prints if columns 1-12 of the table selection card (item 412) are not punched TABLEbSELECT. Correct and start processing from beginning of job.

EXPANSION FACTOR CARD INCORRECT

Message 2. Message prints if columns 1-12 of the expansion factor card (item 423) are not punched EXPANbFACTOR. Correct and start processing from beginning of job.

TABLE TOTAL CARD INCORRECT

Message 3. Message prints if columns 1-12 of the survey unit table totals card (item 424) are not punched TABLEbTOTALS. Correct and start processing from beginning of job.

STRATUM WEIGHT CARD INCORRECT

Message 4. Message prints if columns 1-12 of the sampling stratum weight card is not punched STRATUMbWTS (item 431). Correct and start processing from beginning of the job.

INPUT IDENTIFICATION CARD INCORRECT

Message 5. Message prints if columns 1-11 of the data input identification card are not punched INPUTbIDENT. Correct and start processing from beginning of the job.

JOB CONTROL CARD INCORRECT

Message 6. Message prints if columns 1-11 of the job control card (item 411) are not punched JOBbCONTROL. Correct and start processing from beginning of the job.

THE TABLE NAMED AAAAAA HAS ITS TITLES OUT OF ORDER

Message 7. Message prints if the table name punched in columns 1-6 of the table title card (item 414) does not correspond with the table name on the data input tape (see part VII-E). AAAAAA is the table name which appears on item 414. Correct and start processing from the beginning.

MACHINE CAPACITY EXCEEDED — REDUCE SIZE OR NUMBER OF TABLES

Message 8. Message prints if the total number of cells necessary to produce the desired output table is greater than 15,000. Correct by reducing number of tables or altering dimension (see part X-B). Start processing from beginning of job.