



## Application Program

### **1401 Data Analysis and Reduction System**

**(1401-CA-04X), Version 2**

#### **Application Description**

The program is a general purpose statistical reporting, data analysis, and information retrieval system of wide use in areas such as medical research, capital investment, stock analysis, market research, inventory, and others. It is extremely useful in analyzing and preparing data for the study of complex systems.

This manual contains a general description of the program options, machine configuration, and general system chart.

Copies of this and other IBM publications can be obtained through IBM branch offices. Address comments concerning the contents of this publication to IBM, Technical Publications Department, 112 East Post Road, White Plains, N.Y. 10601

# CONTENTS

Introduction . . . . .	1
General Description of Application . . . . .	2
Input Programs . . . . .	2
Data Manipulation Programs . . . . .	2
Data Extract Program . . . . .	2
Report Programs . . . . .	2
System Programs . . . . .	3
System Generation and Modification Programs . . . . .	3
System Description . . . . .	3
Input . . . . .	3
Processing . . . . .	4
Output . . . . .	4
Program Modifications . . . . .	4
Tape Labels . . . . .	4
Operating Programs . . . . .	4
Storage Assignments . . . . .	5
System Configuration . . . . .	5
Programming Systems . . . . .	5
General System Chart of Application . . . . .	6



## INTRODUCTION

The IBM 1401 Data Analysis and Reduction System is a powerful series of programs to generate tabular and statistical reports from existing card or magnetic tape files. Three types of programs, modified by control cards to suit individual requirements, operate under the control of a monitor program.

1. The data preparation programs accept punched cards or magnetic tape in a variety of formats, and add, delete, and/or change the format and content of each record or user-selected records from the data file. Records may be selectively extracted for information retrieval purposes. Communications-based messages of fixed- or variable-length variable format may be used as input to create fixed-length, fixed-format output records.
2. The report programs produce, as specified in control cards, matrix-type, frequency tabulation, and statistical reports. The matrix reports, consisting of horizontal columns and vertical rows, present crossfooted distributions expressed in percentages, counts, or summaries as required. The frequency tabulation reports contain the frequency of occurrence of the values in a specified field of every data record. The quantitative statistical reports contain parameters such as the mean, variance, and standard deviation, along with a matrix-type presentation of frequency, count, and percentage for each defined group. Tabulations up to ten fields, maximum length of 18 positions each, may be specified as an independent function, or with simultaneous listing and/or punched output.
3. The utility programs provide the ability to list records, control program operating sequence, prepare and modify the systems tape, and perform various other functions to facilitate using the system.

These programs can be used in almost any industry to provide a concise analysis of subjects such as:

- Inventory activity
- Traffic characteristics
- Sales analysis
- Communications systems
- Engineering studies
- Medical reports
- Payroll analysis
- Personnel studies

The system, with its ability to accept records from cards or magnetic tape in a variety of formats, can quickly provide management with meaningful reports without extensive programming and data reduction effort.

## GENERAL DESCRIPTION OF APPLICATION

The programs in the system may be divided into six general groups:

1. Input Programs
2. Data Manipulation Programs
3. Data Extract Program
4. Report Programs
5. System Programs
6. System Generation and Modification Programs

A brief description of the functions of each of these programs follows:

### Input Programs

Data Input Control — accepts fixed- or variable-length, fixed-format card or tape records and converts them to a format suitable for the data preparation or report programs. Input fields are defined by their position in the record. Serial numbers may be assigned to each record if desired.

Variable Input Control — accepts variable-length, variable-format tape or card records and converts them to a format suitable for the data preparation or report programs. Input fields are defined by field divider codes, consisting of length, occurrence, and character representations.

### Data Manipulation Programs

Field Edit — provides right or left adjusting of fields in formatted tape to prepare for sorting.

Field Redefinition — combines, separates, adds, or deletes fields in formatted tape records.

### Data Extract Program

Extract — selects records and creates a new file based on relational conditions, testing all records or only those records occurring at user-specified intervals. This program is a valuable aid in statistical sampling and information retrieval.

### Report Programs

Matrix Report — produces reports expressing counts, percentages, or summation of values.

Data Record List — prints a formatted listing of input records by fields. All or selected fields may have decimal or comma punctuation, sign control, and suppression of nonsignificant zeros. Up to ten fields may be selected for tabulation. Records may also be punched and listed.

Data Record Tabulate — produces a report on the frequency of occurrence of the values in a specified field of every data record. Each value that occurs within the field is listed, along with the frequency of occurrence, that is, the number of times the same value appears in successive records.

Quantitative Statistical Parameters Report — computes and prints a comprehensive tabular report of statistical parameters such as mean, variance, and standard deviation.

## System Programs

Monitor — controls system programs.

I/O Program — common input/output routines for all system programs using magnetic tape.

Blank Identification — provides for immediate reuse of any operational program, or produces a formatted print of the system program tape, or prints the program identification table from the system program via the monitor.

## System Generation and Modification Programs

Tape Load — prepares a program system tape from cards.

General Utility — provides a storage printout with optional printouts of the input, output, and system tapes.

Modify System — duplicates and/or modifies the system tape by adding, deleting, or replacing system programs.

## SYSTEM DESCRIPTION

### Input

Input data may consist of:

1. Cards — with one or more cards used to create one record.
2. Blocked or unblocked fixed-length tape records. The maximum single record size is 999 characters. Maximum block size is 1500 characters.
3. Unblocked variable-length tape records. The maximum usable portion is the first 999 characters. These records may not contain record marks, tape marks, or group marks as data characters.
4. Variable-length records with variable-length fields, such as communications-based messages, reports, or data may be entered from cards or magnetic tape, and subsequently reformatted to fixed-length, fixed-field records for processing.

## Processing

A monitor controls the system, permitting the running of operational programs with minimum operator intervention. The same input data may be used to produce a variety of reports. The output of one program can be used as the input to another, to provide a multilevel analysis. Original input data may be saved on request.

Operator intervention is minimized by internally scheduling the output tape unit of the current program to become the input tape unit for the next program. Programs may be run in any logical order by arranging the control cards accordingly.

Control cards are checked for format and reasonableness. The detection of an error that may cause a program malfunction or loss of data will set a system error switch on. Subsequent control cards in the card reader are checked, and error messages printed, but the programs are not run. This prevents dependent programs from attempting to process improper data.

## Output

Reports are produced on the 1403 Printer.

## Program Modifications

User-written modifications (additions, deletions, or changes) may be made to the system as required. A system program is included to facilitate locating errors in new additions to the operational program file.

The system can duplicate itself on a second reel of magnetic tape.

## Tape Labels

An 80-character header label is assumed on all input tape reels. If a single reel of input is to be processed, a control card is used to indicate the presence or absence of a header label.

## Operating Programs

The operating programs are stored on magnetic tape and loaded as required by a system monitor program. Control cards are analyzed by the system monitor, which selects and loads the appropriate programs. At the end of the program run, control is returned to the system monitor.

Operating programs on the system tape are written in load mode, one program or program segment per block. Programs requiring more than 4000 storage positions are written as two or more blocks.



## Storage Assignments

For this system, 1401 core storage is logically divided into three areas:

Area 1. Positions 001-599. The monitor occupies positions 340 to 599. Positions 001 through 332 are reserved for the card reader, punch, printer, and index registers.

Area 2. Positions 600-4599. This area is reserved for the current operational program. Only one program from the system tape may occupy this area at any one time.

Area 3. Positions 4600-7999. This area is reserved for system constants, tables, a 1500-position common tape input/output area, the common tape input/output routine, an 80-character header and trailer label area, and the calling sequence for the general utility program.

## SYSTEM CONFIGURATION

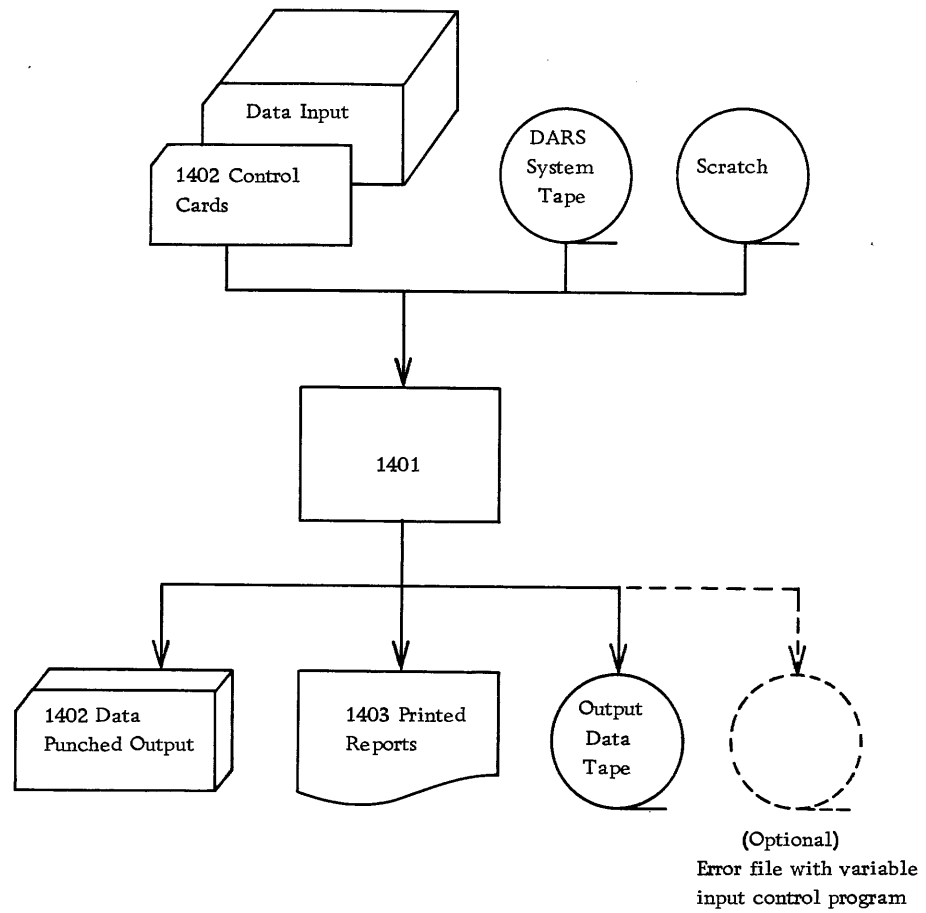
The Data Analysis and Reduction System operates on the IBM 1401 Data Processing System with the following minimum configuration:

- 1401 Processing Unit, 8K
- Three 729 or 7330 Magnetic Tape Units (Four if Error File option is used with Variable Input Control Program)
- 1402 Card Read Punch
- 1403 Printer, Model 2
- Advanced Programming Feature
- High-Low-Equal Compare Feature
- Multiply-Divide Feature
- Sense Switches (required for Variable Input Control Program)

## PROGRAMMING SYSTEMS

- The source program language for DARS is the 1401 Tape Autocoder language.
- IBM 1401 Tape-to-Card Utility Program (1401-UT-028) is required to punch the card image deck to create the DARS operating system tape. The utility program may also be used to punch the Autocoder source and object cards.
- IBM 1401 Tape-to-Printer Utility Program (1401-UT-026) is used to print the 1401 Tape Autocoder listings on the distributed DARS magnetic tape.

# GENERAL SYSTEM CHART OF APPLICATION





International Business Machines Corporation  
Data Processing Division  
112 East Post Road, White Plains, New York 10601