

Table of Contents

Introduction and Telephone Numbers.....	3
JCL Introduction.....	4
EDIT Format.....	5
LRECL Format.....	6
Other Considerations.....	7
JCL for SPSS.....	10
Input.....	10
Output.....	12

INTRODUCTION

At present the Statistical Software Section supports the following programs on the IBM System 370.

- o SAS, SAS/STAT, SAS/GRAPH, SAS/ETS, SAS/OR, SAS/FSP, SAS/AF, SAS/IML, SAS/DB2, SAS/ASSIST, SAS/CONNECT, SAS/CBT001
- o BMDP
- o SPSS, SPSS/TABLES
- o IMSL
- o GLIM
- o LISREL

In addition, the SSS supports SAS, SPSS, BMDP, and IMSL on the IBM PC and IMSL on the Convex C240. Additional information about these packages can be obtained from the SSS. Additional documentation for many of these packages is available from the DCRT Technical Information Office.

Telephone numbers that may be useful are:

Statistical Software Section/LSM: (301) 496-6037

The SSS provides support for the statistical software packages listed above. The LSM is located in Building 12A, Room 3045.

Programmer Assistance and Liason (PAL) Unit: (301) 496-5525

The PAL Unit provides technical assistance, information, and consultation to Computer Center users. All questions regarding WYLBUR should be directed to the PAL Unit. It is located in Building 12A, Room 1017.

DCRT Technical Information Office (TIO): (301) 496-5431

The Technical Information Office provides users with publications and documentation related the services and facilities of the NIH Computer Utility. Users can use WYLBUR's ENTER PUBWARE to order publications. The TIO is located in Building 12A, Room 1017.

Training Unit: (301) 496-2339

The Training Unit provides courses for Computer Center users. Users can use WYLBUR's ENTER TRAINING to see information on Computer Center courses and scheduling. The Training Unit is located in Building 12A, Room 1025.

User Area: (301) 496-6888

The User Area contains computer terminals that are available for all Computer Center users. The User Area is located in Building 12A, Room 1018.

Information Media Library: (301) 496-6021

The Information Media Library maintains a large supply of magnetic tapes for use with the IBM System 370 computer. You may request up to two tapes per day by contacting them by phone. The Information Media Library is located in Building 12, Room 1100.

JOB CONTROL LANGUAGE INTRODUCTION

Jobs run in batch mode at NIH must begin with Job Control Language (JCL) to tell the IBM System 370 who the user is, what kind of program he wants to run, and what data sets he wants to use. JCL must be typed in upper case, begin in column 1, and extend no farther than column 71.

The first line of JCL is the job statement. It has the following format:

```
//iii JOB (aaaa,box,class),lastname
```

where the words shown in lower case are to be replaced with the user's information:

iii	registered initials
aaaa	account number
box	box number
class	class of the job
lastname	user's last name

The job class is determined by the resources needed for the job. Small jobs not requiring tapes (but possibly using the Managed Storage System (MSS)) get fast turnaround time from class A or class E. Class B or C is needed to mount a tape. Other parameters, such as maximum CPU time or maximum lines, may appear after the job class. For more information, see the Computer Center User's Guide (CCUG), available from the Technical Information Office.

The following abbreviations are used in this handout:

ddname	single word of user's choice (1-8 alphanumeric or national (\$,#,@) characters beginning with an alphabetic or national character), used as a file reference in a data definition (DD) statement
dsname	operating system data set name
serial	serial number of a tape
seqnum	sequence number of a data set on tape (defaults to 1)
lab	type of label on tape (SL, NL, AL--defaults to SL)
len	maximum line length (integer)
b	blocksize (calculation discussed below)
s1	primary space allocation (calculation discussed below)
s2	secondary space allocation (calculation discussed below)

In accordance with the Computer Center's move toward an 'All-Cataloged Environment,' all JCL discussed will assume that all input and output files are cataloged. For a discussion of how to catalog files, see the WYLBUR Fundamentals Manual and the CCUG, both available from the Technical Information Office. For information about how to move existing files to the all-cataloged environment, see INTERFACE, Number 166, also available from the Technical Information Office. Utilities that may be useful in converting to the all-cataloged environment are CATDS, UNCATDS, and DSSCR.

One consequence of the all-cataloged environment is that during program development, it is possible to accidentally create multiple copies of permanent files of which only the first copy is cataloged. To avoid this, include an EXEC DSSCR statement in the JCL to delete previous copies of the permanent file. To see a list of files on different disks with the same name, use WYLBUR's SHOW DSNS DUPLICATE command.

Most of the examples discussed in this handbook assume that if a tape is being used, it is an NIH standard tape. An NIH standard tape is a standard-labeled 3480 cartridge tape which has 18 tracks and stores information at 38,000 BPI. If you are interested in using another type of tape, refer to the CCUG. To have a tape assigned, call the Information Media Library at (301) 496-6021.

At NIH, WYLBUR data sets are stored in one of two formats on disk and the MSS: EDIT format or LRECL format.

EDIT FORMAT

EDIT format data sets are compressed to save space, while LRECL format data sets are not. EDIT format is the default storage format. If a data set is saved using the WYLBUR SAVE command, it is stored in EDIT format.

```
? SAVE AS dsname
```

EXAMPLE:

```
? COLLECT CLEAR
  1. ? 4 5
  2. ? 0 1
  3. ? 2 6
? ***
? SAVE AS RAWDATA1
'RAWDATA1' SAVED AND CATALOGED ON FILE
```

To find out if a data set on a disk is stored in EDIT format, use the WYLBUR SHOW DSNAME command.

```
? SHOW DSNAME dsname
```

EXAMPLE:

```
? SHOW DSNAME RAWDATA1  
FILE
```

```
RAWDATA1 ON FILE51
```

```
CREATED 11/26/91, LAST USED 11/26/91  
1 TRACK (1 USED), NO. OF AREAS=1, SECONDARY SPACE=1 BLOCK  
DSORG=PS, RECFM=U, LRECL=11476, BLKSIZE=11476  
KEYLEN=0, RKP=0, OPTCD=C(20)  
EXPIRATION DATE=11/26/91, NO PASSWORD, NOT RACF PROTECTED
```

If the data set is stored in EDIT format, this command will report RECFM=U, LRECL=11476, and BLKSIZE=11476.

To find out if a data set on the MSS is stored in EDIT format, use the NIH procedure ADSMAP, described in the CCUG.

EXAMPLE:

```
// (JOB statement)  
// EXEC ADSMAP  
//SYSIN DD *  
FIND DSNAME
```

If the data set is stored in EDIT format, this command will report RECFM=U, LRECL=11476, and BLKSIZE=11476.

LRECL FORMAT

To save a disk data set in LRECL format, add the LRECL option to WYLBUR's SAVE command.

```
? SAVE AS dsname LRECL=len
```

where 'len' is greater than or equal to the length of the longest data line.

EXAMPLE:

```
? COLLECT CLEAR  
1. ? 4 5  
2. ? 0 1  
3. ? 2 6  
?  
? ***  
? SAVE AS RAWDATA1 LRECL=6  
'RAWDATA1' SAVED AND CATALOGED ON FILE
```

To find out if a data set on disk is stored in LRECL format, use the WYLBUR SHOW DSNAME command.

EXAMPLE:

```
? SHOW DSNAME RAWDATA1  
FILE
```

```
RAWDATA1 ON FILE51
```

```
CREATED 11/26/91, LAST USED 11/26/91  
1 TRACK (1 USED), NO. OF AREAS=1, SECONDARY SPACE=1 BLOCK  
DSORG=PS, RECFM=FB, LRECL=6, BLKSIZE=11472  
KEYLEN=0, RKP=0, OPTCD=C(20)  
EXPIRATION DATE=11/26/91, NO PASSWORD, NOT RACF PROTECTED
```

OTHER CONSIDERATIONS

SPSS does not read raw data files saved in EDIT format. Thus, if your data is saved in EDIT format, you will need to apply one of the following:

- o Use the WYLBUR USE FROM and RESAVE commands on the data set

```
? USE FROM dsname  
? RESAVE LRECL=len
```

where 'len' is greater than or equal to the length of the longest data line.

- o Include the EDSIN utility in the JCL

There are several places in JCL where space allocations and blocksizes must be determined.

The blocksize must be determined when using EXEC EDSIN to read in EDIT format files from disk or the MSS when the records extend past column 80. To determine the blocksize (BLKSIZE=b) to use for a given logical record length (LRECL=len), use the following WYLBUR command:

```
T 11476/len*len
```

where 'len' is greater than or equal to the length of the longest record.

The logical record length, blocksize, and space allocations must be determined when outputting raw data from SPSS. In general, the secondary space allocation ('s2') can be 1. Set LRECL equal to an integer greater than or equal to the length of the longest record. To calculate blocksize and primary space allocation, use WYLBUR's ENTER DISKCALC facility. The example below shows how to use ENTER DISKCALC when LRECL=133 and there are 500 records to be output.

? enter diskcalc

DISKCALC Command Procedure:

Select a function from the following menu (or strike ENTER):
(For fixed length, NON-keyed data sets)
(Function 2 is applicable to all data set types)

- 1 What is the Computer Center recommendation for BLKSIZE for a given LRECL?
- 2 Compute how many physical records (blocks) may be stored on a track or cylinder with a given BLKSIZE.
- 3 Show records and blocks per track and cylinder for a selection of possible BLKSIZEs for a given LRECL.
- 4 Compute total space needed to contain a specified number of records of a given LRECL for a selection of possible BLKSIZEs.
- 5 Change default disk model (currently 3380K, 2,635 cylinders per volume, 15 tracks/cylinder, 47,476 bytes per track).

END Return to WYLBUR.

Select a function (or press ENTER to see the function menu): 4

Enter the LRECL (logical record length) of the record (or press Enter):

LRECL: 133

Enter the number of logical records (or press ENTER):

Number of records: 500

Space required for 500 records with LRECL=133 on a 3380K volume which contains approximately 2,635 cylinders per volume.

BLKSIZE	Number of Tracks	Number of Cylinders
3,059	2	
3,458	2	
3,857	2	
4,256	2	
4,788	2	
5,453	2	
6,251	2	
7,448	2	
9,044	2	
11,438	2	<=NIH Recommendation
15,428	2	
23,408	2	

We will use the NIH recommendation. In this example then, LRECL=133, BLKSIZE=11438, s1=2, and s2=1.

When creating an SPSS system file, SPSS automatically defines the record format, the logical record length, and the blocksize. To calculate the primary space parameter, use the following steps:

1. Let $X = 8 * (\text{number of variables} + 7) * (\text{number of observations})$.
2. Let $Y1 = X/5120$. Let $Y2$ be $Y1$ rounded up to the next integer.
3. Let $Z = Y2/8$. The primary space allocation 's1' is Z rounded up to the next integer.

For example, suppose there are 12 variables and 5000 observations. Then

$$\begin{aligned} X &= 8 * 19 * 5000 = 760,000 \\ Y1 &= 760,000 / 5120 = 148.44 \\ Y2 &= 149 \\ Z &= 149 / 8 = 18.625 \\ s1 &= 19 \end{aligned}$$

and the space parameter in the JCL would be `SPACE=(TRK,(19,1))`.

JOB CONTROL LANGUAGE FOR SPSS

NIH currently uses Release 4 of SPSS. The system is described in the SPSS Reference Guide, 4th Edition, and augmented by the SPSS INFO command. SPSS is used in batch mode from WYLBUR.

Any reference to the FILE HANDLE command in the SPSS Reference Guide should be interpreted as a reference to the DD statement in the JCL concerning that file. The FILE HANDLE command is not needed on the NIH IBM System 370.

SPSS assumes that columns 73 through 80 of each command line are reserved for line numbers. The WYLBUR line numbers will appear to the left of each SPSS command. If you want SPSS to use all 80 columns to read commands, specify UNNUMBERED as the first SPSS command and be sure to use RUN UNN to submit the job. If the data is stored with the program and either the data extends past column 72 or if the DATA LIST keyword FREE is used, you must specify RUN UNN to WYLBUR and UNNUMBERED as the first SPSS command.

SPSS has the following syntax rules at NIH:

- o Commands must begin in column one
- o Any command may continue onto several lines, but continued lines of a command must begin past column one
- o Command terminators are not needed

The following abbreviations and terms are used in SPSS JCL.

input data file	raw data file which is stored outside the SPSS command file or program
System file	file created by a SPSS job
foreign file	non-SPSS System file created by other software that can be used by SPSS

SPSS can read data stored in a SPSS System file faster than from a foreign file or input data file. SPSS System files can be stored on disk, MSS or tape using the SPSS 'SAVE OUTFILE=ddname' command and appropriate JCL.

INPUT

1. Input data in the job stream

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//SYSIN DD *
```

2. Input data on disk in LRECL format or in a SPSS System file

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//ddname DD DSN=aaaaaaa.dsname,DISP=SHR
//SYSIN DD *
```

3. Input data on disk in EDIT format with data lines not extending past column 80

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC EDSIN,NAME='aaaaaaa.dsname'
// EXEC SPSS
//ddname DD DSN=&INPUT,DISP=(OLD,DELETE)
//SYSIN DD *
```

4. Input data on disk in EDIT format with data lines extending past column 80

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC EDSIN,NAME='aaaaaaa.dsname',LRECL=len,BLKSIZE=b
// EXEC SPSS
//ddname DD DSN=&INPUT,DISP=(OLD,DELETE)
//SYSIN DD *
```

5. Input data on the Managed Storage System (MSS) in LRECL format or in a SPSS System file

```
// (JOB statement)
/*ROUTE XEQ MSS
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//ddname DD DSN=aaaaaaa.dsname,DISP=SHR
//SYSIN DD *
```

6. Input data on the MSS in EDIT format with data lines not extending past column 80

```
// (JOB statement)
/*ROUTE XEQ MSS
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC EDSIN,NAME='aaaaaaa.dsname'
// EXEC SPSS
//ddname DD DSN=&INPUT,DISP=(OLD,DELETE)
//SYSIN DD *
```

7. Input data on the MSS in EDIT format with data lines extending past column 80

```
// (JOB statement)
/*ROUTE XEQ MSS
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC EDSIN,NAME='aaaaiii.dsname',LRECL=len,BLKSIZE=b
// EXEC SPSS
//ddname DD DSN=&INPUT,DISP=(OLD,DELETE)
//SYSIN DD *
```

8. Input data on NIH standard tape

```
// (JOB statement with class B or C)
/*ROUTE XEQ TAPE
/*MESSAGE serial,R
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//ddname DD DSN=aaaaiii.dsname,DISP=SHR
//SYSIN DD *
```

OUTPUT

9. Save a calculated file created by SPSS procedures or a case-ordered file created by WRITE or PRINT with a FILE subcommand

- A. Raw output data set to be saved on disk

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC DSSCR,NAME='aaaaiii.dsname'
// EXEC SPSS
//ddname DD UNIT=FILE,DISP=(NEW,CATLG),
// DSN=aaaaiii.dsname,SPACE=(TRK,(s1,s2)),
// DCB=(RECFM=FB,LRECL=len,BLKSIZE=b)
//SYSIN DD *
```

where

'len' is 80, except when using WRITE or PRINT, in which case substitute the length of the longest record to be written. PRINT separates the variables by blanks and also leaves column 1 blank.

For a discussion of how to calculate 'b,' 's1,' and 's2,' see the JCL Introduction.

- B. Raw output data set to be printed

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//ddname DD SYSOUT=A,DCB=(RECFM=UA,BLKSIZE=133)
//SYSIN DD *
```

10. Save an SPSS system file or SPSS portable system file (the output of the SAVE, XSAVE or EXPORT commands).

A. SPSS system file, or portable file, to be saved on disk

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC DSSCR,NAME='aaaaiii.dsname'
// EXEC SPSS
//ddname DD UNIT=FILE,DISP=(NEW,CATLG),
// DSN=aaaaiii.dsname,SPACE=(TRK,(s1,s2))
//SYSIN DD *
```

B. SPSS system file on disk; new SPSS system file, or portable file, to be saved on disk

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC DSSCR,NAME='aaaaiii.dsname2'
// EXEC SPSS
//ddname1 DD DSN=aaaaiii.dsname1,DISP=SHR
//ddname2 DD UNIT=FILE,DISP=(NEW,CATLG),
// DSN=aaaaiii.dsname2,SPACE=(TRK,(s1,s2))
//SYSIN DD *
```

C. Input data on a standard label 6250 BPI tape; new SPSS system file or portable file, to be saved on the MSS

```
// (JOB statement with class B or C)
/*ROUTE XEQ 9TRACKHI
/*ROUTE XEQ MSS
/*MESSAGE serial,RS
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC DSSCR,NAME='aaaaiii.dsname2'
// EXEC SPSS
//ddname1 DD DSN=dsname,DISP=SHR,
// VOL=(PRIVATE,SER=serial),UNIT=9TRACKHI,
// LABEL=(seqnum,lab)
//ddname2 DD UNIT=MSS,DISP=(NEW,CATLG),
// DSN=aaaaiii.dsname2,SPACE=(TRK,(s1,s2))
//SYSIN DD *
```

D. Input data on disk in LRECL format; SPSS portable file to be saved on an NIH standard tape

```
// (JOB statement with class B or C)
/*ROUTE XEQ TAPE
/*MESSAGE serial,W
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//ddname1 DD DSN=aaaaiii.dsname,DISP=SHR,
//ddname2 DD UNIT=TAPE,DISP=(NEW,KEEP),
// VOL=(PRIVATE,SER=serial),DSN=aaaaiii.dsname,
// LABEL=(seqnum,lab),DCB=(RECFM=FB,LRECL=80,BLKSIZE=3200)
//SYSIN DD *
```

GET SAS and GET BMDP Commands

11. The GET SAS command allows the use of Version 5 SAS data sets stored on disk or the MSS in an SPSS program. The Version 5 SAS data sets must not be in transport format.

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//ddname1 DD DSN=aaaaiii.dsname1,DISP=SHR
//ddname2 DD DSN=aaaaiii.dsname2,DISP=SHR
//SYSIN DD *

UNNUMBERED
GET SAS DATA=ddname1.membername SASLIB=ddname2
  /KEEP=variables
  /RENAME=(oldname=newname)
  /MAP
```

where

'ddname1' represents the SAS library stored on aaaaiii.dsname1, and 'membername' is the specific SAS member to be read into the SPSS program. SASLIB specifies the DD name for the library containing SAS formats, which may contain value labels for some or all of the variables.

12. The GET BMDP command allows the use of BMDP save files stored on disk or the MSS.

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//ddname DD DSN=aaaaiii.dsname1,DISP=SHR
//SYSIN DD *

UNNUMBERED
GET BMDP FILE=ddname
  /KEEP=variables
  /RENAME=(oldname=newname)
  /MAP
```

where

'ddname' refers to the BMDP file to be read in. BMDP save files are identified within each data set by code, content, and label fields. Code and label are specified by the user within BMDP; content is supplied by the BMDP program to identify the type of file. To print these fields and other information about the save files in a data set, use the SCAN subcommand. The SCAN subcommand must appear immediately following the FILE subcommand. SCAN=YES prints the information and then reads the file, while SCAN=ONLY prints the information and stops. The CODE, CONTENT, and LABEL subcommands may be used to specify a particular save file within a data set, and they must follow the FILE and SCAN subcommands. If they are not used, SPSS reads the first save file with content DATA.

CREATING TEMPORARY SPSS SYSTEM FILES

SPSS run in batch mode has the capability of handling up to ten different temporary SPSS system files in one session. This feature is especially useful when several files need to be manipulated or altered before executing a MATCH or UPDATE command. No additional user-written JCL is necessary to accomplish this. Simply specify one of ten reserved ddnames (TEMP1-TEMP10) on any SAVE commands used. The system will create a scratch file containing the modified SPSS system file, which can be accessed for the remainder of your session. This avoids the cost and confusion of creating permanent file space for SPSS system files you may never need to access in subsequent jobs.

13. Two SPSS system files, currently stored on disk, need to be sorted by the same variable before the MATCH command will execute properly.

```
// (JOB statement)
// JCLLIB ORDER=ZABCRUN.PROCLIB
// EXEC SPSS
//FY89 DD DSN=aaaaiii.TRIAL89,DISP=SHR
//FY90 DD DSN=aaaaiii.TRIAL90,DISP=SHR
//SYSIN DD *
```

UNNUMBERED

```
GET FILE=FY89
SORT CASES BY ID
SAVE OUTFILE=TEMP1
```

```
GET FILE=FY90
SORT CASES BY ID
```

```
MATCH FILES FILE=* / FILE=TEMP1 /
      BY ID / MAP
```

LIST