Int BJH
QC LG

8940

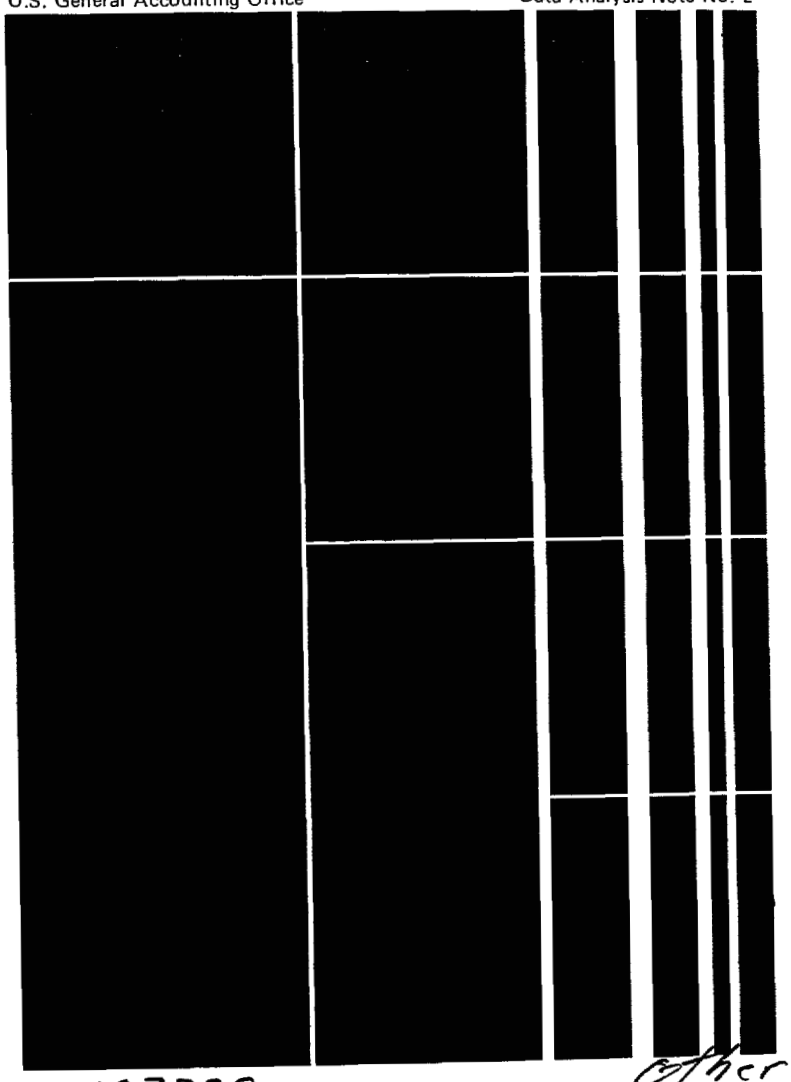
PROGRAMMING GUIDE
To Version 6.0 Of
Statistical Package
For The Social Sciences (SPSS)

Ident

Systems Analysis Group
Federal Personnel And Compensation Division
U.S. General Accounting Office

GAO 00150

Data Analysis Note No. 2



Other

FOREWORD

As part of its assistance function in the Federal P
sonnel and Compensation Division of the U.
Accounting Office, the Systems Analysis G
prepared this guide summarizing Version 6
Statistical Package for the Social Sciences (SPSS
purpose is to facilitate the use of SPSS for those who a
currently using version 6.0 and to encourage those wh
are using older versions of SPSS to convert to
updated version. This guide contains general instru
tions, formats of SPSS statements, and instructions
executing SPSS programs.

NOTATIONS

1. Braces { } indicate that a choice is to be made from the enclosed items.
2. Brackets [ ] indicate the enclosed items are optional.
3. Underlining indicates the default parameters.
4. Special characters such as slashes / and parentheses ( ) must be used as shown.
5. UPPERCASE words must be used in the statement while lowercase words indicate user-supplied information.
6. The following abbreviations are used:

   nf = maximum number of functions.

   cp = minimum required cumulative percentage of eigen values.

   sig = maximum tolerated significance level.

   mdrp = missing data replacement where $0 < mdrp < 1$ (e.g., mdrp = .4 omits cases with greater than 40% missing data.

REFERENCE

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. SPSS: Statistical Package for the Social Sciences, 2nd ed. New York: McGraw Hill, 1975.

GENERAL RULES

1. Control field is columns 1-15. Contents of this field must begin in column 1 and are inflexible regarding spacing and spelling of words. The specification field is columns 16-80.

2. File and variable names have a maximum length of 8 characters; however, some systems only allow 6 characters. The first character must be alphabetic. Names may be composed of letters and numbers but may not include blanks and must be contained on one card.

3. Values for variables may not contain blanks.

4. Alphanumeric values must be enclosed in apostrophes.

5. The following keywords may not be used as variable names:

ABS	EQ	LT	SIN
ALL	EXP	MOD10	SQRT
AND	GE	NE	SUBFILE
ATAN	GT	NOT	THRU
BY	LE	OR	TO
CASWGT	LG10	RND	TRUNC
COS	LN	SEQNUM	WITH

6. Mathematical symbols are (in order of execution):
   Exponentiation**
   Multiplication * and division /
   Addition + and subtraction -

7. Names, values, keywords, and labels must be separated by commas or blanks. The delimiters ( ),=, and / are used when specified.

8. Data definition cards VARIABLE LIST and INPUT FORMAT must be used together but may be replaced by a DATA LIST card.

9. Data definition card N OF CASES is required unless the SUBFILE LIST card is used.

10. A maximum of 500 variables per file is allowed (except for the MINI and MAXI versions of SPSS).

11. A maximum of 100 subfiles per file is allowed. The first four characters of subfile names must be unique.

12. Labels and variable names may not include ( ) or /.

13. The TO convention may be used with VARIABLE LIST, DATA LIST, MISSING VALUES, VALUE LABELS, and PRINT FORMATS.

14. The ALL convention may be used with MISSING VALUES, VALUE LABELS, PRINT FORMATS, OPTIONS, STATISTICS, and RUN SUBFILES.

15. Values of SUBFIL, SEQNUM, and CASWGT cannot be altered using COMPUTE, IF, and RECODE statements.

FORMAT ELEMENTS

For FIXED data:

nFw.d
nAw
nX
Tc
/

Where:

n = number of adjacent elements with the same format
F = indicates a numeric value
w = column width including sign and decimal
d = number of digits to the right of the decimal
A = indicates an alphanumeric value
X = skip columns
T = transfer to a specified column
c = column number transferred to
/ = go to the next record or data card

For BINARY data:

nV
nS
Pn
/

Where:

n = number of adjacent elements with the same format
V = indicates input to SPSS
S = skip spaces
P = proceed to the specified variable
/ = go to the next record or data card

PRECEDENCE TABLE

(from page 541 of the 2nd edition of SPSS)

0	EDIT	13	VAR LABELS VALUE LABELS DOCUMENT
1	NUMBERED		
2	RUN NAME PRINT BACK PAGESIZE	14	*SAMPLE
		15	*COMPUTE *IF *SELECT IF *COUNT *RECODE
3	GET FILE GET ARCHIVE MERGE FILES		
4	FILE NAME	16	*WEIGHT
5	ADD VARIABLES ADD CASES ADD DATA LIST DATA LIST VARIABLE LIST	17	ASSIGN MISSING MISSING VALUES PRINT FORMATS
		18	RUN SUBFILES RAW OUTPUT UNIT
6	ADD SUBFILES DELETE SUBFILES SUBFILE LIST	19	LIST CASES
		20	procedure card
7	INPUT MEDIUM INPUT FORMAT	21	OPTIONS STATISTICS
8	N OF CASES		
		22	READ INPUT DATA READ MATRIX
9	SAMPLE WEIGHT	23	DELETE VARS KEEP VARS REORDER VARS
10	DO REPEAT		
11	COMPUTE IF SELECT IF COUNT RECODE	24	SORT CASES
		25	SAVE FILE SAVE ARCHIVE
12	END REPEAT	26	FINISH

NONPROCEDURE CONTROL CARDS

Starred cards, such as *COMPUTE, have the same format as the unstarred versions. Starred commands are temporary and last only for one procedure card.

ADD CASES	number of cases
ADD DATA LIST	FIXED (number of cards per case)/ card number variable list starting column [- ending column (variable type) variable list. . ./card number. . .]
ADD DATA LIST	BINARY (number of cards per case)/ card number variable list variable position [variable list variable position. . ./card number. . .]
ADD SUBFILES	subfile name$_1$ (n_1), subfile name$_2$ (n_2), . . .subfile name$_b$ (n_b)
ADD VARIABLES	list of new variable names
ALLOCATE	TRANSPACE = number of bytes
ASSIGN MISSING	variable list (value)
COMMENT	any text
COMPUTE *COMPUTE	computed variable=arithmetic expression
COUNT *COUNT	variable name = criterion variable list (value list)
DATA LIST	FIXED (number of cards per case)/card number variable list starting column [- ending column (variable type) variable list. . ./card number. . .]

DATA LIST	BINARY (number of cards per case)/ card number variable list variable position [variable list variable position. . ./card number. . .]
DELETE SUBFILES	list of subfile names
DELETE VARS	variable list
DO REPEAT	stand-in variable name$_1$ = variable list/ stand-in variable name$_2$= variable list/. . .
END REPEAT	[any comment]
DOCUMENT	any text, but only one such card is allowed per run
EDIT	
FILE NAME	file name [file label up to 64 characters]
FINISH	
GET ARCHIVE	FILE=system or archive file name$_1$ VARIABLES = {variable list or ALL}/ FILE = system or archive file name$_2$ VARIABLES= . . .
GET FILE	file name
IF *IF	(logical expression) computed variable = arithmetic expression
INPUT FORMAT	{FIXED or FREEFIELD or BINARY} (format list if FIXED or BINARY)

INPUT MEDIUM	{CARD or DISK or TAPE or OTHER}
KEEP VARS	variable list
LIST ARCHINFO	{VARLIST or COMPLETE}
LIST CASES	CASES = number of cases/ VARIABLES = variable list
LIST FILEINFO	{ COMPLETE or LABELS VARLIST SORTVARS VARINFO SUBDIRECTORY DOCUMENTS }
MERGE FILES	FILE=file name$_1$ VARIABLES={ variable list of ALL }/ FILE=file name$_2$ VARIABLES=. . .
MISSING VALUES	variable list (up to 3 missing values)/ variable list. . .
N OF CASES	number of cases or UNKNOWN
NUMBERED	YES or NO -columns 73-80 are for identification
OPTIONS	{options number list or ALL}
PAGESIZE	{number of lines or NOEJECT}
PRINT BACK	{NO or CONTROL or FORMAT}
PRINT FORMATS	variable list (value 0-5 or A)/. . .

RAW OUTPUT UNIT file number 9 or 15 through 20

READ INPUT DATA

READ MATRIX

RECODE
*RECODE
variable list
{(value list = new value)
or (CONVERT)}/...

REORDER VARS variable list

RUN NAME run label of up to 64 characters

RUN SUBFILES {(subfile name or list) or EACH or ALL}

SAMPLE
*SAMPLE
sampling factor

SAVE ARCHIVE output archive file name, [label]/
FILE = input file$_1$
VARIABLES = {variable list or ALL}/
FILE = input file$_2$
VARIABLES=...

SAVE FILE [file name, file label]

SELECT IF
*SELECT IF
(logical expression)

SORT CASES variable list {(A) or (D)}

STATISTICS {statistic number list or ALL}

SUBFILE LIST subfile name$_1$ (n_1) subfile name$_2$ (n_2)...

TASK NAME	label of up to 64 characters
VALUE LABELS	variable list ($value_1$) $label_1$ up to 20 characters ($value_2$) $label_2$. . ./ variable list ($value_1$) $label_1$ ($value_2$) $label_2$. . ./. . .
VAR LABELS	variable $name_1$, variable label up to 40 characters/ variable $name_2$, variable label . . ./. . .
VARIABLE LIST	variable list
WEIGHT *WEIGHT	variable name
WRITE FILEINFO	VARIABLES={variable list or ALL}/ {ALL or FILE NAME VARLIST SUBFLIST NCASES MISVALS PRNTFMTS VARLABS VALLABS DOCUMENTS}/ CHAR= BCD or EBCD / NUMBERED

AGGREGATE

```
GROUPVARS=variable list/
VARIABLES=variable list/
AGGSTATS=(any one or a combination of:
          VALIDN
          SUM
          MEAN
          SD
          MAX
          MIN
          SKEW
          KURT
          PCTGT (value)
          PCTLT (value)
          PCTBTN (low value, high value))/
[RMISS=value/]
VARIABLES=.../
AGGSTATS=...
```

OPTIONS:

1. Includes missing data.
2. Listwise deletion of missing data. If there is a missing value for any variable on the list, that entire case is omitted.
   Default: Variable-by-variable exclusion. Missing data is excluded from all calculations for a given variable.
3. Output the values of the grouping variables on the aggregated file.
   Default: Output the program-generated sequential aggregation group identification number and number of cases.
4. Generate a compositional aggregated file.
   Default: Generates a file with one case per aggregation unit.

STATISTICS:

1. Reports program-generated group identification number, total cases, values of the grouping variables for each aggregation group.
2. Prints all of (1) plus the exact contents of each case in the aggregated file, including the values of each of the aggregated variables.
3. Prints the statistics in (2) for the first 10 cases only. This overrides statistic (2).

ANOVA

dependent variable list BY independent variable list (min, max) independent variable list (min, max) ... [WITH covariate list]/ dependent variable list ...

OPTIONS:

1. Includes missing values
   Default: Listwise deletion. (See REGRESSION for description)
2. Suppresses printing of value labels.
3. Ignore 2-way and higher interactions among factors.
4. Ignore 3-way and higher interactions among factors.
5. Ignore 4-way and higher interactions among factors.
6. Ignore 5-way and higher interactions among factors.
   Default of options 3 to 6: examine all interaction effects.
7. Process covariates concurrently with the main effects for the nonmetric factors.

8. Process covariates after the main effects for the nonmetric factors.
   Default of options 7 and 8: introduce covariates before factors.
9 All effects are assessed simultaneously. Ignores options 7 and 8.
10. Uses stepdown procedure within blocks of nonmetric factors and covariates.
    Default for options 9 and 10: adjust each factor main effect for other factors and each covariate effect for other covariates.

STATISTICS: 1. Outputs a table of category means to be expressed as deviations from the grand mean.
2. Outputs standardized partial regression coefficients for the covariates.

BREAKDOWN
(integer mode)

VARIABLES = variable list (low value, high value)
variable list (low value, high value) . . ./
{CROSSBREAK or TABLES} = variable list BY variable list BY. . ./
variable list BY variable list BY. . ./. . .

BREAKDOWN
(general mode)

TABLES = variable list BY variable list BY . . ./
variable list BY variable list BY . . ./. . .

OPTIONS:

Default for options 1 and 2 is omission of a case from all computations for a table if it has missing data for any variable in the table.

1. Includes missing values.
2. Excludes missing values for only dependent variables.
3. Suppresses printing of all labels.
4. Prints output in a modified tree diagram. Available only in the general mode.

The following options are available for the CROSSBREAK facility which may be used only in the integer mode.

5. Deletes cell frequencies.
6. Deletes sums.
7. Deletes standard deviations.
8. Deletes value labels but prints variable labels.

STATISTICS:

1. Gives a one-way analysis of variance including an Eta squared statistic, sum of squares, degrees of freedom, mean square.
2. Causes a test of linearity to be calculated on one-way breakdowns; includes F ratio, Pearson's r,r^2, sum of squares, degrees of freedom, mean square.

The following statistics are available for the CROSSBREAK facility and measure relationships between row and column variables:

3. Chi square
4. Phi (for 2 x 2), Cramer's V (for others)
5. Contingency coefficient
6. Lambda, symmetric and asymmetric
7. Uncertainty coefficient, symmetric and asymmetric
8. Kendall's tau b
9. Kendall's tau c
10. Gamma, conditional not partial or zero order
11. Somer's D, symmetric and asymmetric
12. Eta

CANCORR

VARIABLES = variable list/
RELATE = (parameters) variable set_1 WITH variable set_2/
[CANVAR [=mdrp]]/
RELATE = . . .

OPTIONS:

1. Includes missing values.
2. Pairwise deletion of missing data. (See REGRESSION for description and default).
3. Computes a covariance matrix.
4. Indicates matrix input.
5. Computes matrix about the origin.
6. Continues computations if either variable set is linearly independent.
7. Continues computations even if a matrix is not positive-definite.
8. Outputs a simple correlation matrix.

9. Indicates the input correlation matrix is indexed by the VARIABLE LIST card. May only be used in conjunction with option 4.

10. Weighted estimate of missing data in computing canonical scores. Only applies if option 2 is also used.

STATISTICS:

1. Prints means, standard deviations, and number of valid cases of all variables in the VARIABLES = list.
2. Prints correlation matrix of all variables on the VARIABLES = list.
3. Prints correlation matrix even when noncomputable correlations are encountered.
4. Outputs canonical variate scores.

CONDESCRIPTIVE variable list

OPTIONS:

1. Includes missing values. Default: excludes missing values.
2. Suppresses printing of variable labels.
3. Causes Z-scores to be written to a raw-output-data file.
4. Causes a reference dictionary to be printed at the end. It informs the user of locations of statistics for variables; it lists variables alphabetically and in the order they appear in the system file.

STATISTICS:

1. Mean
2. Standard error
5. Standard deviation
6. Variance
7. Kurtosis
8. Skewness
9. Range
10. Minimum
11. Maximum

Default: all statistics are given.

CROSSBREAK See BREAKDOWN

CROSSTABS
(integer mode)

VARIABLES = variable list (low value, high value) variable list (low value, high value) . . ./
TABLES = variable list BY variable list BY . . ./. . .

CROSSTABS
(general mode)

TABLES = variable list BY variable list BY . . ./
variable list BY variable list BY . . ./. . .

OPTIONS:

1. Includes missing values.
2. Suppresses printing of variable and value labels.
3. Deletes row percentages
4. Deletes column percentages.
5. Deletes total percentages.

Options 6, 7, and 8 are only available in the integer mode:

6. Deletes value labels but prints variable labels.
7. Includes missing values in the tables but not included in the statistics. Missing values are not included in row and column percentages.

8. Causes values of the row variable to be printed in descending rather than ascending order.
9. Causes an index to be printed listing all tables produced and beginning page numbers.

STATISTICS:

1. Chi-square*
2. Phi (for 2 x 2 tables) Cramer's V (for larger tables)
3. Contingency coefficient
4. Lambda, symmetric and asymmetric.
5. Uncertainty coefficient, symmetric and asymmetric
6. Kendall's tau b
7. Kendall's tau c
8. Gamma** (partial and zero-order gamma for tables with 3 to n variables)
9. Somer's D, symmetric and asymmetric.
10. Eta, for numeric data only

*for 2 x 2 tables, Fisher's test is used for less than 21 cases while Yate's test is used for larger numbers of cases.

**available in integer mode only.

DISCRIMINANT

GROUPS = {variable name (min, max) or SUBFILES}/
VARIABLES=variable list/
ANALYSIS=variable list
[(level)] . . . variable list
[(level)]/
[METHOD = DIRECT or WILKS or MAHAL or MAXMINF or MINRESID or RAO /
TOLERANCE = value/
MAXSTEPS = value/
FIN = value/FOUT = value/
PIN = value/POUT = value/
VIN = value/
FUNCTIONS = nf, cp, sig/
PRIORS = {EQUAL or SIZE or probabilities list}]

OPTIONS:

1. Includes missing values.
2. Includes cases with missing values during classification.
3. Suppresses step-by-step output.
4. Suppresses stepwise summary table.
5. Prints classification results table.
6. Prints discriminant scores and classification information.
7. Prints a single plot of cases.
8. Prints a separate plot for each group.
9. Suppresses classification phase features for cases initially unclassified.

10. Prints a territorial map. Ignored if there are more than 2 discriminant functions.
11. Prints unstandardized discriminant function coefficients.
12. Prints classification functions.
13. Rotates the discriminant functions; standardized coefficients are printed. Unstandardized rotated coefficients are printed if option 11 is requested.
14. Uses individual group covariance matrices for classification.
15. Outputs matrix materials. Prints means and covariance matrices for each group if option 14 is requested.
16. Indicates matrix input.

    Options 17, 18, and 19 produce raw output in 80 column records with subfile name and case sequence numbers first and sequence numbers in columns 73-80. Formatting information is included as part of the output.
17. Outputs discriminant scores.
18. Outputs membership probabilities for all groups.
19. Outputs actual group and classified group numbers.

STATISTICS:
1. Prints means for each group and total.
2. Prints standard deviations for each group and total.

3. Outputs pooled within-groups covariance matrix.
4. Outputs pooled within-groups correlation matrix.
5. Outputs matrix of pairwise F ratios.
6. Outputs univariate F ratios.
7. Computes test for equality of group covariance matrices.
8. Computes group covariance matrices.
9. Outputs total covariance matrix.

FACTOR

VARIABLES = variable list/
[TYPE = { PA1 / PA2 / RAO / ALPHA / IMAGE / BYPASS }/
DIAGONAL=value list/
NFACTORS=value/
MINEIGEN=value/
ITERATE=value/
STOPFACT=value/
ROTATE={ VARIMAX / QUARTIMAX / EQUIMAX / OBLIQUE / NOROTATE }/
DELTA=value
FACSCORE [=mdrp]]/
VARIABLES=...

OPTIONS:

1. Includes missing data.
2. Pairwise deletion of missing data. A case is omitted from calculations if either of the variables being considered has a missing value.

Default: Listwise deletion of missing data. A case is omitted from all calculations if it contains missing data.

3. Input is a correlation matrix.
4. Input is a factor matrix and communalities.
5. Outputs a simple correlation matrix.
6. Outputs a factor matrix and communalities.
7. Outputs a factor-score coefficient matrix.
8. Outputs means and standard deviations.
9. Indicates the order of variables on an input matrix is defined on the VARIABLE LIST card. May be used only if option 3 is also used.
10. Produces weighted factor scores when missing data are encountered.
11. Sequences the output factor scores.

STATISTICS:

1. Means and standard deviations.
2. Correlation matrix.
3. Inverse and determinant of correlation matrix.
4. Communalities, eigenvalues, and proportion of total and common variance.
5. Initial-factor matrix.
6. Rotated-factor matrix and transformation matrix.
7. Factor-score coefficient matrix.
8. Plot of rotated factors.

FREQUENCIES (general mode)

GENERAL = {variable list or ALL}

FREQUENCIES (integer mode)

INTEGER = {variable list or ALL} (low value, high value) variable list. . .

OPTIONS:

1. Includes missing values.
2. Suppresses printing value labels.
3. Causes all output to be printed in an 8½ x 11 inch space to the left.
4. Causes all printed output to be written on a permanent file instead of on a line printer. Option 3 is automatically included.
5. Causes tables for all variables to be printed in condensed format.
6. Causes tables filling more than one page to be printed in condensed format.
7. Deletes the frequency tables and prints only statistics.
8. Causes histograms to be printed following the table and preceding the statistics.
9. Causes a reference dictionary to be printed at the end. It informs the user of locations of statistics for variables; it lists variables alphabetically and in the order they appear in the system file.

STATISTICS:

1. Mean
2. Standard error
3. Median
4. Mode

5. Standard deviation
6. Variance
7. Kurtosis
8. Skewness
9. Range
10. Minimum
11. Maximum

GUTTMAN SCALE scale name = variable name (decision point)
variable name (decision point)../
scale name = variable name...

OPTIONS:
1. Includes missing data in all scales and computations.
2. Suppresses printing of variable labels.
3. Suppresses automatic ordering of variables from most to least difficult.

STATISTICS:
1. Inter-item and part-whole correlation coefficients.
2. Coefficient of reproducibility.
3. Minimum marginal reproducibility.
4. Percent improvement achieved by Guttman scale.
5. Coefficient of scalability.

NONPAR CORR variable list [WITH variable list]/...

OPTIONS:
1. Includes missing values.
2. Causes listwise deletion. A case is omitted from all calculations if it contains a missing value for any variable on the list.

Default: Pairwise deletion. A case is omitted from computation of a given coefficient if the value of either variable being compared is missing.

3. Causes computation of a 2-tailed rather than 1-tailed test of significance for each coefficient.
4. Causes a matrix of coefficients to be written for all specified lists. The keyword WITH cannot be used with this option.
5. Yields only Kendall correlations.
6. Yields both Kendall and Spearman correlations.
   Default: Yields only Spearman correlations.

STATISTICS: No statistics card available.

ONEWAY

dependent variable list BY independent variable list (min, max)/
POLYNOMIAL=number/
CONTRAST=coefficient list/
RANGES=test name (significance level)/

OPTIONS:

1. Includes missing values.
2. Causes listwise deletion of missing values.
3. Suppresses printing of variable labels.
4. Outputs number of cases, mean, standard deviation for each group on a raw-output-data file.

6. Uses the first eight characters of the value label for the independent variable as the group labels.
7. Inputs category frequencies, means, and standard deviations instead of raw data.
8. Inputs category frequencies, means, pooled variance, and degrees of freedom for the pooled variance.

STATISTICS:
1. Gives number of cases per category, mean, standard deviation, standard error, minimum, maximum, and a 95% confidence interval for the mean.
2. Outputs fixed- and random-effects measures.
3. Outputs statistics on homogeneity of variables.

PARTIAL CORR

correlation list BY control list (order values)/
correlation list BY control list (order values)/. . .

OPTIONS:
1. Includes missing data.
2. Pairwise deletion of missing data. (See REGRESSION for description).
   Default: Listwise deletion of missing data.
3. Two-tailed tests of statistical significance.
4. Matrix input.
5. Punches or writes matrices for future access.

6. Specifies the order of variables on input matrices by the VARIABLE LIST card. May be used only if option 4 is used.
7. Omits degrees of freedom and significance.
8. Prints only nonredundant partials, in a serial string format.

STATISTICS:
1. Prints simple or zero-order correlations used in computing partial correlations along with degrees of freedom and statistical significance.
2. Prints means and standard deviations of variables on the PARTIAL CORR card plus number of valid cases.
3. Prints zero-order coefficients if and only if any of these coefficients are noncomputable.

PEARSON CORR variable list [WITH variable list]/...

OPTIONS:
1. Includes missing values.
2. Causes listwise deletion. A case is omitted from all calculations if it contains a missing value for any variable on the list.
   Default: Pairwise deletion. A case is omitted from computation of a given coefficient if the value of either

variable being compared is missing.

3. Causes computation of a 2-tailed rather than 1-tailed test of significance for each coefficient.
4. Causes a matrix of coefficients to be written for all specified lists. The keyword WITH cannot be used with this option.
5. Deletes the number of cases and significance figures.
6. Prints only nonredundant coefficients.

STATISTICS:

1. Means and standard deviations, unaffected by option 2.
2. Cross-product deviations and covariance for each pair of variables for which a correlation coefficient was requested.

REGRESSION

VARIABLES= variable list/
REGRESSION= dependent variable [(inclusion level)]
WITH independent variable list (inclusion level)
[RESID={number or mdrp}] /
REGRESSION=. . ./
VARIABLES=. . .

OPTIONS:

1. Includes missing values.
2. Causes pairwise deletion of cases. A missing value for a particular variable causes that case to be eliminated from calculations involving that variable.
   Default: Listwise deletion. Cases with missing values are deleted from all calculations.

3. Suppresses printing of variable labels.
4. Indicates the matrix of correlation coefficients will be input by the user.
5. Indicates means and standard deviations are to be read in preceding the input correlation matrix. This is only valid when option 4 is also used.
6. Suppresses step-by-step output and prints only the summary table.
7. The opposite of option 6.
8. Causes the correlation matrix to be output on a unit of the user's choice.
9. Indicates the user inputs one large correlation matrix and that subsets of the variables on the matrix will be used in the regressions. Must be accompanied by option 4.
10. Causes sequencing information to be entered in columns 1 to 20 of each record on the raw-output-data file.
11. Prints standardized residuals on the raw-output-data file.
12. Outputs standardized predicted dependent variable values.
    Note: Options 11 and 12 should not be used simultaneously if only plots are desired.
13. Creates standardized predictors which are a weighted product of the existing data. In effect only when data replacement and option 2 are requested.

14. Suppresses printing of axes on the plots of standardized predictor versus standardized residual (requested by statistic 6).
15. Outputs means and standard deviations to the raw-output-data file in separate sets corresponding to the variable list.

STATISTICS:
1. Correlation matrices.
2. Means, standard deviation, number of valid cases.
3. Forces printing of correlation matrix even if some coefficients cannot be calculated. Warns of bad data items that should be omitted.
4. Outputs a plot of standardized residuals against the sequence of cases for residuals.
5. Durbin-Watson statistic for residuals.
6. Requests a plot of standardized residuals against standardized predictor values with residuals on the vertical axis.
7. Prints correlation matrix and number of cases.

SCATTERGRAM variable list [(low value or LOWEST, high value or HIGHEST)]
variable list . . . [WITH] variable list . . .

OPTIONS:

1. Includes missing values.
2. Causes listwise deletion of cases. Deletes cases if any of the variables on the list contains a missing value.
   Default: Pairwise deletion. Deletes cases from a graph if one variable in the graph is missing.
3. Suppresses printing of variable labels.
4. Suppresses plot grid lines.
5. Prints diagonal grid lines.
6. Computes a 2-tailed test of significance if statistic 3 is chosen.
7. Uses integer plot labels. A specified range on the control card will override this option.
8. If not enough core storage is available for all cases, plots will be produced on as many cases as possible.

STATISTICS:

1. Pearson's r
2. r^2
3. Significance of r
4. Standard error of the estimate
5. Intercept
6. Slope

T-TEST

GROUPS = group specification/
VARIABLES = variable list

OPTIONS:

1. Includes missing values.
2. Causes listwise deletion of cases. Completely omits cases if any variable on the variables list contains a missing value.

Default: Pairwise deletion. For groups, if the value of the variable is missing, the case is left out of computations. For pairs, if either variable contains a value that is missing, the case is left out of computations.

3. Suppresses printing of variable lables.

STATISTICS: There are no optional statistics for T-TEST.

WRITE CASES {(format list) or BINARY} variable list

OPTIONS:
1. Suppresses output of cases with missing data for any variables listed.

STATISTICS: No statistics card available.

EXECUTING SPSS PROGRAMS

On the NIH system, the mimimum JCL needed are:

```
// Jobname JOB (account, box, class), name
//    EXEC RUNSPSH6
// GO.SYSIN DD *
     SPSS deck
/*
```

Additional JCL required are:

If GET FILE or GET ARCHIVE are used

```
// GO.FT03F001 DD UNIT= {FILE or 2420}
//    DISP= {SHR or (OLD, KEEP)},
//    VOL= (SER= file or tape number),
//    DSN=AAAAIII. filename
```

If SAVE FILE or SAVE ARCHIVE are used

```
// GO.FT04F001 DD UNIT= {FILE or 2420},
```

```
//    DISP= (NEW, KEEP),
//    VOL= (SER= file or tape number),
//    DSN=AAAAIII. filename,
{ //    DCB=BLKSIZE= 4000,
  //    SPACE= (4000, (10,5), RLSE)
                    or
  //    DCB= (RECFM=  FB, LRECL= 80,
           BLKSIZE= 3120) }
```

If INPUT MEDIUM is used

```
// GO.FT08F001 DD UNIT=FILE,
//    DISP= SHR,
//    VOL= SER= file number,
//    DSN=AAAAIII. filename,
//    DCB=(RECFM= FB, LRECL= 80,
         BLKSIZE= 3120)
```

If WRITE CASES is used or correlation matrices are output

```
// GO.FT09F001 DD UNIT= {FILE or 2420},
//    DISP= (NEW, KEEP),
//    VOL= (SER=file or tape number),
//    DSN=AAAAIII. filename,
//    DCB= (RECFM=  FB, LRECL= 80,
         BLKSIZE= 3120)
```

When writing on DISK the following parameter must also be used

```
//    SPACE= (3120, (10,5), RLSE)
```

Where: 2420 represents a 9-track 1600 BPI tape
AAAA represents a valid account number
III represents a registered set of initials.

On the INFONET system, SPSS programs are run with the following statement:

```
!SPSS   IN$:fn1 OUT$:fn2 FTN8:fn3 FTN9:fn4
```

Where: IN$ indicates the SPSS program name.
OUT$ denotes a file name to which the SPSS output is written. This file must be saved if the user wishes to access it in the future.
FTN8 denotes the raw-input data file.
FTN9 denotes a file name to which SPSS output resulting from WRITE CASES or special output from some procedures is to be written.
fn1, fn2, fn3, fn4 denote file names.

The shortest statement needed to run SPSS is:

```
!SPSS  IN$:  file name
```