MF

Staff Paper 82-32
May 1982

COMPARISONS AND NOTES ON STATISTICS PROGRAM PACKAGES FOR APPLE, TRS-80, AND CP/M MICROCOMPUTERS

by

Robert D. Stevens and Valerie Auserehl Kelly

Contents

IV. List of Microcomputer Statistics Packages.

A. More Comprehensive Packages

B. Less Comprehensive Packages

C. Specialized Application Packages

V. Related Papers

VI. Recommended Reading

I. Introduction

The advent of a large number of microcomputers in the last three years has led to numerous efforts to produce computer programs to capitalize on the potential of these machines. Statistics program packages are of particular interest to professionals in many fields.

This short paper presents an outline of the more detailed information about microcomputer statistics packages for the TRS-80, Apple II, and CP/M operating systems which is contained in Staff Paper 82-23. In Section II, we compare a number of microcomputer packages with respect to system requirements and range of statistical routines. We also compare the microcomputer packages to each other and to SPSS with respect to five selected statistical routines. A comparison of data management capabilities and a summary of special attributes ends the comparisons.

Section III provides a one to two page summary of information contained in Staff Paper 82-23 for each package compared in Section II and for a few additional packages which we consider promising. These latter packages were not included in Section II due to inadequate information. Sample printouts and extracts from manuals are not included here, but may-be found in Staff Paper 82-23.

Section IV contains a list of all statistics packages identified to date. The list is presented using three groupings: 1) more comprehensive packages, 2) less comprehensive packages and 3) specialized applications. The list is, in effect, the table of contents to an annotated "diskography" of statistics packages found in Staff Paper 82-23.

Computer journals and personal contacts were used to attempt an up-to-date list of statistics program packages. The focus of the search was on relatively comprehensive packages of integrated statistics programs. No attempt was made to identify the many single purpose limited programs, some of which are in the public domain and may be obtained at no cost through computer clubs.

II. Comparisons of Statistics Program Packages

The comparison of statistics packages which follows is presented with the following objectives in mind:

1) to illustrate how the microcomputer system requirements (hardware configurations) differ from one statistical package to another (Table I),
2) to illustrate the range of statistical routines contained in microcomputer packages and the variability of routines included from one package to another (Table II),
3) to illustrate the variability of statistics packages, both mainframe and microcomputer, vis a vis 5 selected statistical routines (descriptive statistics, frequencies, crosstabs, ANOVA, and regression). (Tables III-VII),
4) to illustrate the varibility of data management capabilities among selected microcomputer statistical packages (Table VIII),
5) to offer a format for others to use when attempting to evaluate packages or routines which are not covered in this paper.

Tables I and II are self-explanatory. Tables III - VII describe the contents of the mainframe package SPSS and five microcomputer packages. The table format allows the reader to easily compare the packages and

identify those which provide features most appropriate to particular applications. We have not presented any evaluation of the packages described. During the course of our study, it became evident that, given the current state of the art, a microcomputer package is best selected with a particular application in mind and a careful analysis of whether or not a given package can perform all the desired functions.

The following comparative tables on selected statistics packages are based on our access to the sources of information summarized below:

NAME OF PACKAGE	SOURCES OF INFORMATION USED IN COMPARISONS		
	PROGRAM: VERSION USED	MANUAL: VERSION USED	CURRENT USERS
SPSS	MSU MAINFRAME	SECOND EDITION	
A-STAT	VERSION 79.6, 1981	VERSION 79.6, 1981	Stevens/Kelley Ag. Economics - MSU
ECONOMETRIC LINEAR FORECASING (ELF)	VERSION 3.4, 1981	VERSION 4.0, 1982	Stevens/Kelley Ag. Economics - MSU
STATISTICAL PROCESSING SYSTEM (SPS)	VERSION 3.0, 1980	VERSION 3.0, 1980	Geoff Remes Life Sciences - MSU
MICRO-STAT		RELEASE 1.0, & PARTS of 2	Leighton Price Lifelong Education - MSU
STAT PAC	NOT AVAILABLE	1981	Len Malczynski Ag. Economics - Purdue

TABLE I

SYSTEM REQUIREMENTS

Package Name	SPSS	ASTAT	ELF	MICRO-STAT	SPS	STAT-PAC
Package Price	--	$145	$200			$285
REQUIREMENTS						
Type of Computer	MAINFRAME	APPLE II+	APPLE II+	VARIABLE	TRS-80 II[1]	TRS-80 I/III
Disk Operating System		DOS	DOS	CP/M	TRSDOS	I:NEWDOS III: TRSDOS
Language Used		APPLESOFT	APPLESOFT	BASIC-80[2]		
Memory (RAM) Required		48K	48K	48K	48K	48K
DISK DRIVE(S) (5-1/4")						
MINIMUM		1	1	1	1	2
RECOMMENDED		2	2	2[3]	2	2
Printer		Optional	Optional	Optional	Optional	Parallel or RS-232

[1] Also available for Model I.

[2] Available in 3 versions of CP/M basics

[3] One 8" disk drive would also be sufficient in lieu of 2 5-1/4" drives.

It should be noted that selection of packages compared in this section was determined by our access to user manuals. We originally attempted to include other promising packages (STAT-PRO, AIDA, SPM) for which we had obtained fairly detailed program descriptions from software distributors. This plan was abandoned, however, when it became obvious that comparisons based on these non-equivalent information bases risked being more misleading than helpful. Although we believe that the packages compared are representative of the better ones available, the reader is encouraged to refer to section III of this paper and to sections III and IV of Staff-Paper 82-23 for information collected on the numerous packages not included in this comparison. (These packages are listed in Section IV of this paper.)

In making our comparisons, we decided to give a package credit for a particular routine or characteristic only if the manual explicity told the user how to perform the task. This decision will perhaps mask the potential of some programs in the hands of seasoned statisticians and microcomputer users, however, it will protect the average user from being misled about what he or she will be able to do with a package.

TABLE II

RANGE OF STATISTICS ROUTINES

PACKAGE NAME	ASTAT	ELF	MICRO-STAT	SPS	STAT-PAC
TYPE OF ROUTINE					
DESCRIPTIVE STATISTICS	Y	Y	Y	Y	Y
FREQUENCIES	Y		Y	Y	Y
PROBABILITIES	N	Y	Y	Y	N
HYPOTHESIS TESTING (T-TEST)					
Means		Y			
Proportions	N	N	Y	Y	Y
CROSSTABS/CONTINGENCY TABLES	Y	Y	Y	Y	Y
CORRELATION ANALYSIS	Y	Y	Y	Y	Y
REGRESSION					
MULTIPLE REGRESSION[1]	Y	Y	Y	Y[2]	Y
STEPWISE	N	Y	N	N	N
PATH ANALYSIS	Y	N	N	N	N
N-WAY LEAST SQUARES	Y	N	N	N	N
ANALYSIS OF VARIANCE:					
ONE WAY	N	Y	Y	Y	Y
TWO WAY	N	Y	Y	Y	
RANDOMIZED BLOCK	N	N	Y	N	N
DISCRIMINANT ANALYSIS	N	Y	N	N	N
IMAGE ANALYSIS (GUTTMAN)	N	N	N	Y	N
FACTOR ANALYSIS	N	Y	N	Y	N
PRINCIPAL COMPONENT ANALYSIS	N	N	N	Y	N
COMBINATIONS, PERMUTATIONS, FACTORIALS	N	N	Y	N	N
TIME SERIES	N	N	Y	N	N
SCATTERGRAM	N	Y	Y	Y	Y
GENERATION OF MONTE CARLO DATA SETS	N	N	N	Y	N
TOTAL 'YES'	7	12	13	14	8
TOTAL 'NO'	14	9	8	7	13

[1] All packages with multiple regression also do simple regressions.

[2] SPS has 3 regression modules including a General Linear Models module which provides much more extensive output vis a vis test and checks than other packages.

TABLE III

DESCRIPTIVE STATISTICS[1]

PACKAGE NAME	SPSS	ASTAT	ELF	MICRO-STAT	SPS	STAT-PAC
MODULE	CONDESCRIPTIVE	DATA/DES.	STAT	DES. STAT	DES. STAT	DES. STAT
CAPABILITIES						
SELECT CERTAIN VARIABLES ONLY	Y	N[2]	Y	Y	Y	Y
OUTPUT TO PERMANENT FILE	Y	Y	N	N	N	
Z SCORE TRANSFORMATIONS TO PERM. FILE	Y	N	N	N	N	N
NUMBER VALID CASES READ	Y	Y	N	N	N	Y
NUMBER MISSING VALUES	Y	N	N	N	N	Y
NUMBER RECORDS READ	N	N	Y[3]	Y	Y[3]	N
RESPONSE PERCENT						Y
STATISTICS						
MIN/MAX	Y	Y	Y	Y	Y	Y
RANGE	Y	N	Y	N	Y	Y
SUM	Y	Y	Y	Y	N	Y
MEAN (ARITHMETIC)	Y	Y	Y	Y	Y	Y
MEDIAN	N	N	N	N	Y	Y
MODE	N	N	N	N	Y	Y
SUM OF SQUARES						
VALUES	N	N	N	Y	N	N
DEVIATIONS	N	N	N	Y	N	N
VARIANCE	Y	N	Y	Y	Y	Y
STANDARD DEVIATION	Y	Y	Y	Y	Y	Y
STANDARD ERROR	Y	N	Y	Y	N	Y
SKEWNESS	Y	N	Y	N	Y	Y
KURTOSIS	Y	N	Y	N	Y	Y
CONFIDENCE INTERVALS	N	N	N	N	N	Y
TOTAL 'YES'	14	6	11	10	11	16
TOTAL 'NO'	6	14	9	10	9	4
TOTAL 'UNCERTAIN'	1	1	1	1	1	0

1 BLANKS REPRESENT UNCERTAINTY DUE TO LACK OF INFORMATION ON THIS AND FOLLOWING TABLES.

2 LACK OF THIS OPTION MEANS YOU GET ALL STATISTICS FOR ALL VARIABLES INCLUDING SUCH VARIABLES AS ID NUMBER. TIME CONSUMING AND REQUIRES EDITING OF PRINT-OUT.

3 THE MANUAL FOR ELF SAYS "NUMBER OF OBSERVATIONS IN THE FILE;" FOR SPS MANUAL SAYS N= _ ; WHILE NOT ENTIRELY CLEAR, WE ARE ASSUMING THAT THE PROGRAMS HAVE NOT MADE CORRECTIONS FOR MISSING DATA VALUES.

TABLE IV

FREQUENCIES

PACKAGE NAME	SPSS	ASTAT	ELF	MICRO-STAT	SPS	STAT-PAC
MODULE NAME	FREQ.	FREQ.	TABLES	FREQ.	HISTO-GRAM & FREQ.DIST.	FREQ. DISTRIB.
CAPABILITIES						
SELECT CERTAIN VARIABLES ONLY	Y	Y	Y[1]	Y[2]	Y	Y
OUTPUT TO PERMANENT FILE	Y	N	N		N	N
NUMBER OF VALID CASES READ	Y	Y			N	N
NUMBER OF MISSING VALUES	Y	N	Y[3]	N	N	Y
NUMBER OF RECORDS READ	N	Y	N		N	N
RESPONSE PERCENT	N	N	N	N	N	Y
STATISTICS[3]	Y	N	N	N	Y	N
NOMINAL DISTRIBUTIONS	Y	Y	Y	Y	Y	Y
ABSOLUTE FREQUENCY	Y	Y	Y[4]	Y	Y	Y
RELATIVE FREQUENCY	Y	Y	Y[4]	Y	N	Y
ADJUSTED REL.FREQ. (FOR MISSING VALUES)	Y	N	Y		N	N
CUMULATIVE PERCENT	Y	Y	N	Y	N	N
CUMULATIVE ABSOLUTE FREQ.	N	N	N	Y	N	N
HISTOGRAM	Y	N	Y	Y	Y	N
RELATIVE (PERCENT)	N	N	N	Y	N	N
ABSOLUTE (COUNT)	Y	N	Y	N	Y	N
DATA RESTRICTIONS						
NO LIMIT ON NUMBER OF VALUE CODES	Y	N[5]		N[6]	N[7]	
COMPUTES FRACTIONAL VALUE CODES	Y					
IGNORES MISSING VALUES IF SPECIFIED	Y	Y	Y		Y	Y
GROUPED DISTRIBUTIONS	N[8]	N[8]	N[8]	Y	Y	N[9]
ABSOLUTE FREQUENCY	N	N	N	Y	Y	N
RELATIVE FREQUENCY	N	N	N	Y	N	N
ADJUSTED RELATIVE FREQ.	N	N	N		N	N
CUMULATIVE PERCENT	N	N	N	Y	N	N
CUMULATIVE ABSOLUTE FREQ.	N	N	N	Y	N	N
HISTOGRAM	N	N	N	Y	Y	N
RELATIVE (PERCENT)	N	N	N	Y	N	N
ABSOLUTE (COUNT)	N	N	N	N	Y	N
INTERVAL RESTRICTIONS						
HANDLES UNEQUAL INTERVALS	N	N	N	N	N	N
NO LIMIT ON NUMBER OF INTERVALS	N	N	N		N[7]	N
LISTS NO. VALUES OUTSIDE INTERVALS	N	N	N	Y	N	N
TOTAL 'YES'	15	8	10	16	11	7
TOTAL 'NO'	16	22	19	7	19	22
TOTAL 'UNCERTAIN'	0	1	2	8	1	2

[1] EACH VARIABLE MUST BE RUN AS A SEPARATE TABLE: SLOW AND AWKWARD.

[2] SELECT CERTAIN CASES IS ALSO AN OPTION.

[3] SPSS FREQUENCIES MODULE PROVIDES ALL STATISTICS AVAILABLE IN CONDESCRIPTIVE MODULE PLUS MEDIAN & MODE. MICRO PACKAGES DO NOT PROVIDE STATISTICS WITH THEIR FREQUENCIES MODULES; DESCRIPTIVE STATISTICS MODULES MUST BE USED. (S.P.S. DOES, HOWEVER, PROVIDE MIN. AND MAX. WITH FREQUENCIES.)

[4] RELATIVE & ABSOLUTE BOTH AVAILABLE BUT MUST BE RUN AS SEPARATE DISTRIBUTIONS; OTHER PROGRAMS COMBINE BOTH IN SINGLE OUTPUT.

[5] DEFAULT SET FOR CODE $\leq$ 20 BUT CAN BE CHANGED TO MUCH LARGER VALUE IF COMBINED NUMBER OF OBSERVATIONS AND NUMBER OF VALUES DOES NOT EXCEED AVAILABLE MEMORY.

[6] LIMIT = 50

[7] LIMIT = 36

[8] GROUPED DISTRIBUTIONS REQUIRE REDEFINITION OF VARIABLES & VALUE LABELS; OTHER PACKAGES PERFORM ENTIRE OPERATION WITHIN FREQUENCIES MODULE.

[9] CAN BE ACCOMPLISHED BY CREATING A NEW SUBFILE OF DATA USING EXTENSIVE DATA MANAGEMENT RECODING CAPABILITIES, BUT NOT WITHIN FREQUENCIES ROUTINE.

TABLE V

CONTINGENCY TABLES/CROSS TABS

PROGRAM NAME	SPSS	ASTAT	ELF	MICRO -STAT	SPS	STAT -PAC
MODULE NAME	CROSS TABS	TABLES	TABLES	CROSS TABS	CROSS TABS	CROSS TABS
CAPABILITIES						
VARIABLE LIMIT PER CROSS TAB	8-10	3[1]	3	2	2[2]	2[3]
SELECT CASES OPTION AVAILABLE	Y			Y		
TREATMENT OF MISSING VALUES						
TOTAL NUMBER MISSING VALUES LISTED	Y		N		Y	Y
SHOWN IN TABLE BUT NOT CALCULATED	Y		Y		N	N
TABLE CONTENTS						
PER CELL						
COUNT (FREQUENCY)	Y	Y	Y[4]	Y[4]	Y	Y
% OF ROW	Y	Y	Y[4]	Y[4]	Y	Y
% OF COLUMN	Y	Y	Y[4]	Y[4]	Y	Y
% OF TOTAL	Y	N	N	Y[4]	N	Y
EXPECTED VALUE	N	Y	N	Y	N	N
EXPECTED PERCENT	N	N	N	Y	N	N
PER COLUMN/ROW						
TOTALS R AND/OR C COUNT	Y	Y	Y[4]	Y[4]	Y	Y
R AND/OR C AS % OF TOTAL	Y	Y	N	Y[4]	Y	Y
MAXIMUM TABLE SIZE		15x15		20Rx5c[5]	10x10	
STATISTICS						
DEGREES OF FREEDOM	Y	Y	Y	Y	Y	Y
CHI-SQUARE (OVERALL AND/OR CELL)	Y	Y	Y	Y	Y	Y[6]
FISHER'S EXACT TEST	Y			Y	N	N
YATES CORRECTED CHI-SQUARE	Y				N	N
PHI (2 x 2 ONLY)	Y	Y	Y	N	Y	N
CONTINGENCY COEFFICIENT	Y	Y	Y	N	Y	N
LAMBDA	Y	Y	N	N	N	N
UNCERTAINTY COEFFICIENT	Y	Y	N	N	N[7]	N
TAU (B AND/OR C)	Y	Y	N	N	N[7]	N
GAMMA (ZERO ORDER AND/OR PARTIAL)	Y	Y	N	N	N[7]	N
SOMER'S D	Y	Y	N	N	N[7]	N
ETA	Y	N	N	N	N	N
KENDALL'S S	N	Y	N	N	N	N
G STATISTIC	N	N	N	N	N[7]	N
SIGNIFICANCE OF TEST	Y	Y	Y[8]	N	N[7]	N
CRAMER S V	Y	Y	Y	N	Y	N
KOLMOGOROY-SMIRNOV GOODNESS OF FIT	N	N	N	Y	N	N
TOTAL 'YES'	23	18	11	11	11	9
TOTAL 'NO'	5	5	14	14	16	18
TOTAL 'UNCERTAIN'	0	5	3	3	1	1

[1] REQUIRES USE OF A "FILTERED MARGINALS" PROCEDURE AND MULTIPLE RUNS OF THE TABLES MODULE FOR MORE THAN 2 VARIABLES.

[2] USE OF SPS QUALITATIVE ANALYSIS SERIES OF MODULES ALLOWS FOR ANALYSIS OF 3 VARIABLES CODED ON A SCALE OF 0-9.

[3] GREATER THAN TWO WAY COULD BE DONE BY CREATING SUBFILES AND DOING MULTIPLE ANALYSES.

[4] THESE OPTIONS ARE OBTAINED BY RUNNING A SEPARATE TABLE COMMAND FOR EACH--ALL OUTPUTS CANNOT BE PLACED ON THE SAME TABLE.

[5] PROGRAM ALLOWS YOU TO SPECIFY VALUES TO BE COUNTED FOR SPECIFIC ROWS AND COLUMNS THEREBY ADJUSTING CONTINUOUS VARIABLES TO FIT TABLE SIZE AS DESIRED.

[6] ALSO GIVES PROBABILITY OF CHANCE.

[7] AVAILABLE WITH QUALITATIVE ANALYSIS MODULES MENTIONED IN FOOTNOTE 2. OTHERS AVAILABLE BUT NOT LISTED ON TABLE V: PEARSON'C C, TSCHRUPROW'S T, GOODMAN KRUSKAL TAU Y.

[8] MUST PUT TABLES CHI-SQUARE AND DEGREES OF FREEDOM OUTPUT INTO PROBABILITIES MODULE.

TABLE VI

REGRESSION

PACKAGE NAME	SPSS	ASTAT	ELF	MICRO-STAT	SPS	STAT-PAC
MODULE NAME	REGRES.	PATHS OR REG.	STEP-WISE	REGRES.	REGRES.[1]	REGRES.
CAPABILITIES						
APPLICATIONS DESCRIBED IN MANUALS						
BIVARIATE REGRESSION[2]	Y	N	N	Y	Y	Y
MULTIVARIATE REGRESSION	Y	Y	N	Y	Y[3]	Y
STEPWISE	Y	N	Y	N[4]	N	N
PATH ANALYSIS	Y	Y	N	N	N	N
HIERARCHICAL	Y	N	N[5]	N[5]	N	N
NON-LINEAR	Y	N[5]	N[5]	N[5]	Y[6]	N
DUMMY REGRESSION	Y	N	N	N	N	N
ANALYSIS OF VARIANCE	Y	N	N	N	N	N
ANALYSIS OF COVARIANCE	Y	N	N	N	N	N
GENERAL LINEAR MODELS	N	N	N	N	Y	N
FORECASTING	N	N	Y	N	N	N
NATURE OF OUTPUT						
'b' COEFFICIENT	Y	Y[7]	Y	Y	Y	Y
BETA COEFFICIENTS	Y	Y	Y N	Y	Y[8]	N
CORRELATION MATRIX	Y	Y[9]	Y[9]	Y[9]	Y[9]	Y
FORCED CORRELATION MATRIX	Y	N	N	N	N	N
MEANS & STANDARD DEVIATION	Y	Y[9]	Y	Y	N[10]	Y
NUMBER OF VALID CASES	Y	Y	N[11]	N[11]	N	Y
ACTUALS, PREDICTED VALUES & ERRORS	Y	Y	Y	Y	Y	N
PLOT OF RESIDUALS	Y	N[12]	Y	Y	Y[13]	N
STATISTICS						
SIMPLE R (BETA)	Y	Y	N	Y	Y[8]	N
S.E. 'b' & PROBABILITY 'b' = 0	Y	Y	N	Y	Y	N
MULTIPLE R	Y	Y	N	Y	N	Y
R^2 + S.E. OF ESTIMATE (REGRESSION)	Y	Y	Y	Y	Y	Y
PARTIAL/TOLERANCE/F FOR VARIABLES NOT IN EQUATION	Y	N	N	N	N	N
ADJUSTED R^2	Y	Y	Y	N	Y	Y
ANOVA TABLE	Y	Y	N	Y	Y	N[14]
F TEST	Y	Y	Y	Y	Y	Y
T TEST OF VARIABLE COEFFICIENT	Y	N	Y	Y	Y	N
DURBAN WATSON STATISTIC	Y	Y	Y	Y	Y	N
VON NEUMANN RATIO	N	N	Y	N	N	N
FIRST ORDER AUTOCORRELATION TEST	N	N	Y	N	N	N
VARIANCE/COVARIENCE MATRIX OF COEFFICIENT ESTIMATES	N	N	N	Y	Y	N
'b' CONFIDENCE INTERVALS	N	N	N	N	Y	N
'Y' CONFIDENCE INTERVALS	N	N	N	N	Y	N
TESTS FOR MULTICOLLINEARITY	N	N	N	N	Y	N
TOTAL 'YES'	27	16	14	17	21	10
TOTAL 'NO'	8	19	21	18	14	25
TOTAL 'UNCERTAIN'	0	0	0	0	0	0

[1] THIS MODULE INCLUDES 3 SEPARATE SECTIONS: SIMPLE, MULTIPLE AND GENERAL LINEAR MODELS (GLM). THE LATTER HAS MORE SOPHISTICATED ANALYSIS CAPABILITIES THAN THE FIRST TWO, BUT REQUIRES PUTTING DATA IN TWO SEPARATE MATRICES.

[2] ALL PROGRAMS DO BIVARIATE ANALYSIS EVEN WHEN NOT SPECIFICALLY DESCRIBED AS A SEPARATE ROUTINE.

[3] SPECIFIES LIMIT OF 10 INDEPENDENT VARIABLES.

[4] VERSION OF MANUAL USED IN OUR REVIEW LISTED STEPWISE IN CONTENTS BUT NEVER DESCRIBED IT.

[5] DATA ENTRY MODULE PERFORMS TRANSFORMATIONS ON NON-LINEAR FUNCTIONS WHICH CAN THEN BE PUT IN REGRESSION ROUTINE.

[6] EXPLAINS AND PERFORMS TRANSFORMATIONS (LOG, POWER, EXPONENTIAL).

[7] REQUIRES SPECIAL PROCEDURE AS A-STAT NORMALLY CALCULATES BETA COEFFICIENT.

[8] IN GLM MODULE ONLY.

[9] OBTAINED BY USING CORRELATION MODULE.

[10] COULD BE OBTAINED FROM MODULE PROVIDING DESCRIPTIVE STATISTICS.

[11] PACKAGES DO NOT ALLOW FOR MISSING VALUE CODES.

[12] CAN BE DONE BY SAVING RESIDUALS TO AN OUTPUT FILE AND INTERFACING IT WITH APPLE PLOT PROGRAM.

[13] GLM MODULE ALLOWS PLOTS OF ORIGINAL AND NORMALIZED RESIDUALS AGAINST Y OR ANY X VARIABLE.

[14] PROVIDES SUMS OF SQUARES AND F RATIO AND DEGREES OF FREEDOM BUT NOT MSE.

TABLE VII

ANALYSIS OF VARIANCE

PACKAGE NAME	SPSS	ASTAT	ELF	MICRO -STAT	SPS	STAT -PAC
MODULE NAME	ANOVA	NONE[1]	ANOVA	ANOVA	ANOVA 1-WAY[2]	ANOVA
CAPABILITIES						
ACCEPTABLE DESIGN CHARACTERISTICS		NONE				
EQUAL CELL SIZES	Y		Y	Y[3]	Y	
UNEQUAL CELL SIZES	Y		Y	Y[3]	Y	
CLASSICAL ANALYSIS	Y					
HIERARCHICAL ANALYSIS	Y					
REGRESSION ANALYSIS	Y					
FIXED EFFECT	Y		Y[4]		Y	
RANDOM EFFECT			Y[4]			
APPLICATIONS DESCRIBED IN MANUAL						
ONE-WAY ANALYSIS OF VARIANCE	Y		Y	Y	Y	Y
RANDOMIZED BLOCK				Y	N	N
N-WAY (NUMBER SHOWN = N)	5		2	2	N	N
NATURE OF OUTPUT						
ANOVA TABLES (SS,DF,MSE,F)	Y		Y	Y	Y	Y
PROBABILITY F	Y		N[5]	N	N[6]	Y
NO. OF CASES READ	Y		Y	Y	N	Y
NO. OF MISSING CASES	Y		Y[7]	N	N	N
MULTIPLE CLASSIFICATION ANALYSIS TABLE	Y		N	N	N	N
FOR ONE-WAY ONLY						
APRIORI/APOSTERIORI CONTRASTS	Y				N	
NO. CASES/TREATMENT	Y		Y	Y		Y
TREATMENT MEANS	Y		Y	Y	Y	Y
TREATMENT S.D.	Y		Y	N	Y	Y
TREATMENT MIN/MAX	Y		N	N	N	N
95% CI FOR MEAN	Y		N	N	N	N
FIXED EFFECTS MEASURES	Y		N	N	N	N
RANDOM EFFECTS MEASURES	Y		N	N	N	N
COCHRANS C	Y		N	N	N	N
BARTLETT-BOX F	Y		N	N	N	N
HARTLEY'S F MAX	Y		N	N	N	N
POOLED ESTIMATE	Y		N	N	N	N
SEPARATE ESTIMATE	Y		N	N	N	N
BARTLETS CHI-SQUARE	N		N	N	Y	N
ETA2	N		N	N	Y	N
T-TEST BETWEEN GROUPS	N		N	N	N	Y
FOR N-WAY ANALYSIS						
MAIN EFFECTS	Y		Y	Y	N	N
COVARIATE EFFECTS	Y		N	N	N	N
RESIDUAL	Y		Y	Y	N	N
FACTOR-COVARIATE INTERACTION	N[8]		N	N	N	N
N-WAY INTERACTION (NO. = N)	5		2	2	N	N
COLUMN/ROW/CELL MEANS + FREQUENCIES	Y		N	Y	N	N
BETA FOR COVARIATES	Y		N	N	N	N
TOTAL 'YES'	32		14	13	9	8
TOTAL 'NO'	4		18	19	24	22
TOTAL 'UNCERTAIN'	2		6	6	5	8

[1]THE REGRESSION ROUTINE OF ASTAT PRINTS OUT AN ANALYSIS OF VARIANCE TABLE SHOWING REGRESSION AND RESIDUAL SUM OF SQUARES WITH APPROPRIATE STATISTICS. AS THIS IS NOT COMPARABLE TO PROGRAMS WITH SPECIFIC ANOVA ROUTINES, WE HAVE NOT INCLUDED IT HERE.

[2]SPS ALSO HAS A GENERAL LINEAR MODEL ANOVA ROUTINE WHICH HAS NO APRIORI RESTRICTIONS ON THE MODEL. ITS FEATURES ARE NOT DESCRIBED ON THIS TABLE.

[3]ONE-WAY CAN BE UNEQUAL; TWO-WAY MUST BE EQUAL.

[4]FOR TWO-WAY F STATISTIC CALCULATION.

[5]CAN BE OBTAINED BY USING PROBABILITY MODULE.

[6]CAN BE OBTAINED BY SUMMING NUMBER OF CASES/TREATMENT.

[7]ONE-WAY ONLY.

[8]MAY BE DONE WITH REGRESSION ROUTINE.

TABLE VIII

DATA MANAGEMENT

PACKAGE NAME	ASTAT	ELF	MICRO -STAT	SPS	STAT -PAC
CAPABILITY					
DATA ENTRY	Y[1]	Y	Y	Y	Y[2]
DATA EDITING	N[3]	Y	Y	Y	Y
ADD NEW OBSERVATIONS AND/OR VARIABLES	Y	Y	Y	Y	Y
NUMERIC KEYPAD SIMULATOR	N	Y	N	N	N
DATA TRANSFORMATIONS					
ADD/SUB/MULT/DIVIDE	Y	Y	Y	Y[4]	N
STANDARDIZE	N	N	N	Y	N
TRANSFORM VALUE TO RANK-ORDER			Y		N
LINEAR (A + B* x 1)			Y		N
EXPONENTIAL	Y	Y	Y	Y	N
LOGARITHMS COMMON	Y	Y	Y	Y[5]	N
NATURAL	Y	Y	Y	Y	N
SQUARE ROOTS	Y	Y	N	Y	N
SINES	Y	Y	N		N
COSINES	Y	Y	N		N
TANGENTS	Y	Y	N		N
LAGS AND OR TIME SERIES ANALYSIS		Y	Y		N
RECIPROCAL		N	Y	Y	N
SUM ACROSS VARIABLES		N	Y	Y	N
SUM ACROSS OBSERVATIONS		N	Y	Y	N
SELECT VARIABLE VALUE AND RECODE	Y[6]	N		Y	Y[6]
OUT OF BOUNDS DATA CHECK	N	N	N	N	Y
MERGE FILES	Y	N	Y	Y	N
SORT DATA	Y	N	Y	Y	N
SUB-GROUP FILE CREATION	Y	N	Y		Y
STRIP FILES (CREATE SUBFILES W. SELECTED VARIABLES)		N		Y	Y
LIST DATA TO PRINTED	Y	Y	Y	Y	Y
RANK-ORDER FILE DATA		N	N	Y	
REORDER OBSERVATIONS		N		Y	
AGGREGATE	Y	N			
ESTABLISH MISSING VALUE INDICATORS	Y	N	N	Y	Y
DELETE CASES FROM CALCULATION ONLY	Y	N			
TOTAL 'YES'	18	14	17	20	9
TOTAL 'NO'	4	15	9	2	18
TOTAL 'UNCERTAIN'	9	2	5	9	4

[1]LIMIT OF 45 VARIABLES PER FILE.

[2]ALLOWS FIXED OR FREE FORMAT DATA ENTRY.

[3]RECOMMENDS USE OF COMPATIBLE TEXT FILE EDITORS--SOME ARE PUBLIC DOMAIN PROGRAMS.

[4]ADDING AND MULTIPLING A VARIABLE BY A CONSTANT ARE STRAIGHT FORWARD, VARIABLES/VARIABLE TRANSFORMATIONS REQUIRE MORE COMPLEX PROCEDURES USING BOTH TRANSFORM AND FORM INDICES PROGRAMS.

[5]ALLOWS FOR CHOICE OF BASE.

[6]ASTAT ALLOWS RECODE TO A MATHEMATICAL EXPRESSION. STAT-PAC RECODES TO SPECIFIC NUMBER ONLY.

General Summary Comments on Comparisons of Packages

We would like to point out here a few observations which we were unable to incorporate in other parts of section II. In general, we found that manuals are not consistent in information provided. Some particularly common deficiencies are:

-- failure to provide formulas used for calculations.

-- failure to indicate limits on data base configurations (eg. max number variables and/or observations for specific routines)

-- failure to clearly state missing value capabilities and procedures.

-- failure to include detailed indexes for manuals.

A few particular strengths or weaknesses of specific packages which may not have been revealed in our tables are noted below.

1. ASTAT

   Data handling and transformation capabilities of ASTAT are very extensive compared to most other packages. ASTAT interfaces with Apple Plot, File Cabinet and Visicalc DIF files. Compiled version now available. Program is not menu driven.

2. ELF

   Manual is easy to read and leads user through program prompts. Package interfaces with TWG-ARIMA to perform time series analysis. Users have had numerous problems with bugs in early versions.

3. Micro-Stat

   Good data handling. Excellent documentation on formulas used in calculations (this is lacking in most packages). Manual leads user through prompts.

4. SPS

   Good documentation on formulas used. Quite a bit of instructional information on proper use of various routines. Interfaces with Mainframe SPSS easily.

5. STAT-PAC

   Data management is a strong point. Excellent cross-tabs output.

We invite all readers who have knowledge of packages described in this paper to contact us with any additional observations and/or corrections to material presented. Interpreting a manual (as we have done in most cases) and actually running a program will not necessarily result in identical descriptions of a package. Feedback from those with hands-on experience is very much desired.

III. Outline of Detailed Program Documentation Contained in Staff Paper 82-23.

A. A-Stat

1. General Information

   a. Descriptive Publicity
   b. User Manual Table of Contents
   c. Minimum System Requirements
   d. Default Limits on Data Size

2. A-Stat Disk Catalog

3. Sample Page of Manual Instructions

4. File Formats and Configuration Procedures (Sample Printouts)

   a. A-Stat data file format (ADATA.filename)
   b. A-Stat dictionary file of variables (ADICT.filename)
   c. System configuration procedures

5. A-Stat Statistics and Output Tables - Examples

   a. Descriptive Statistics (NO.OP IN=) - Minimum, Maximum, sum, mean, std. dev., and number of observations.

   b. Frequency Distributions (FREQ=) - Frequency and culmulative frequency.

   c. Correlations (CORRELATE IN=)

   d. Multiple Regression (PATHS COR=) - Beta, simple R, multiple R, R squared, and Determinant.

   e. Analysis of Variance Table (PATHS COR=) - Beta, simple R, multiple Regression with multiple R, R squared, Determinant, 'b'-weight, Std. Error, probability, Std. Error of Estimate, Adjusted R square, Deg. of Freedom, Sum of Squares, Mean square, F-test, probability.

   f. Cross-tabulations (TABLES=) - Cell contents: Frequ., Row %, col. %, Expected Value, cell Chi square, also table Chi square, Prob. of Chi-square, Phi, Contingency.

6. Data Management

   a. List of Data (LIST IN=)
   b. Transformation with Descriptive Statistics (NO.OP IN=)
   c. A Data File with New Variables Added (Dictionary + Data Files)
   d. A-STAT Descriptive Statistics File
   e. A-STAT Correlation Matrix File
   f. Purge A-STAT Files (PURGE=)

7. USER Comments

B. AIDA - Apple Interactive Data Analysis

Index of Materials, Routines, Statistics and Tables

PAGE

C. ELF

GENERAL INFORMATION, CATALOG, FILE FORMATS, STATISTICS, AND EXAMPLES OF OUTPUT TABLES

Index of Materials, Routines, Statistics and Tables

Page

D. MICROSTAT

1. User's Manual Table of Contents
2. Forward to User's Manual
3. Sample of User Instructions for ADOVA-One Way & Block (pg. 26)
4. Sample Output Presented in User's Manual
5. Sample Explanation of Computational Equations ANOVA
6. User Comments

E. S.P.S. - Statistical Processing System

GENERAL INFORMATION, MENUS, STATISTICS AND EXAMPLES OF OUTPUT TABLES

Index of Materials, Routines, Statistics, and Tables

19. Output Tables (Model II Version 3.2) - Examples

    Data File
    Descriptive Statistics
    Histograms and Frequency Distribution

20. S.P.S. User Notes (10/81 adenda)
    Version 3.2 for TRS-80 Model II

F. STAT-PRO

1. Vendor's Comparison with Mainframe Packages
2. Description of STAT MOD
3. Description of Plot MOD
4. Description of QVEST MOD
5. Description of Inventory System
6. Description of BIBLIOMOD
7. Description of MAILING SYSTEM
8. Price List
9. User Agreement
10. Sample Output via Plot MOD

IV. List of Microcomputer Statistics Packages

A. More Comprehensive Statistics Packages

Package Name	Apple II	TRS-80	CP/M
A-STAT + *	X		
ABSTAT			X
Apple Interactive Data Analysis (AIDA)*	X		
Econometric Linear Forecasting (ELF) + *	X		in process
Ed-Sci Statistics	X		
DB MASTER STAT PAK	X		
HSD STATS, ANOVA, REGRESSION°	X		
INTERSTAT/ADVANCED MATH ROUTINES	X		
MICROSTAT+ *			X
STAT PAC + *		I,III	
STAT PRO *	X		
STATISTICAL ANALYSIS (or Advanced Stat Ana.)		I,II,III	
STATISTICAL PACKAGE FOR MICROCOMPUTERS (SPM)°	in process	I,III	in process
STATISTICAL PROCESSING SYSTEMS (SPS)+*		I,II,III	
STATISTICIAN, THE°		I,II,III	in process
STATISTICS WITH DAISY°	X		

+ These packages were included in the comparisons made in Section II of this paper.

* These packages are described in detail in Section III of Staff Paper 82-23. Sample printouts and extracts from manuals are provided.

° The annotated "diskography" in Section IV of Staff Paper 82-23 includes reviews of these programs and/or substantial descriptive materials received from software houses.

B. Less Comprehensive Statistics Packages

PROGRAM NAME	APPLE	TRS-80	CP/M
Interactive Statistical Inquiry System (ISIS)	X		
Micro Data Management & Analysis System (MIDMAS)	X		
SAFOR		X	
Scientist	X		
SOFTWARE PAC #3 - Statistics	X	X	
STATISTICS PAC	X	I,III	

C. Single or Specialized Application Programs

PROGRAM NAME	APPLE	TRS-80	CP/M
ANOVA*	X	X	X
Curve Plotting	X		
Multilinear Regression*	X	X	X
Multiple Regression	X		
Regression I*	X	X	X
Regression II*	X	X	X
Sistema de Estadistica Agricola			X
STATPOWER (Designed to teach Statistics)	X		
TREND STATISTICAL PACKAGE	X		
TSP	X		
TW6-ARIMA (box Jenkins)	X		

---

*These four programs are offered by Dynacomp, Inc. separately or as a package. There is a single annotation for all of them which is listed under title of 'ANOVA' in Section IV, C of Staff Paper 82-23.

V. Related Papers

MICROCOMPUTER SOFTWARE FOR PROFESSIONAL USE

-- NOTES ON PROGRAMS WITH EMPHASIS ON

APPLE, TRS-80, AND CP/M

PART I Word Processors for Microcomputers - Comparisons and References by Robert D. Stevens, Staff Paper #82-16, pp. 160.

PART II Data Base Management and Other Programs for Professionals Using Microcomputers - Comparisons and Reference Materials, Robert D. Stevens, (pp. 80), Staff Paper #82-22.

PART III Statistics Packages for Microcomputer by Robert D. Stevens and Valerie Kelly, (pp. 70) Staff Paper #82-23.

Separate Summary Comparisons

1. Word Processor Software of the Apple II Microcomputer - Summary of Comparisons by Robert D. Stevens, Staff Paper #82-21, pp. 9.

2. Comparisons of Data Base Management Programs for the Apple, TRS-80, and CP/M Microcomputer Systems by Robert D. Stevens, Staff Paper #82-31, pp. 16.

3. Comparisons of Statistics Program Packages for the Apple and TRS-80 Microcomputers by Robert D. Stevens and Valerie Kelly, Staff Paper #82-32, pp. 10.