DOCUMENT RESUME

ED 352 024 IR 015 862

AUTHOR MacFarland, Thomas W.
TITLE Examples of Data Analysis with SPSS-X.
INSTITUTION Nova Univ., Fort Lauderdale, FL. Center for Computer and Information Sciences.
PUB DATE Dec 89
NOTE 44p.; For related materials, see IR 015 861-866. The floppy disk provided with the original document is not included in this document.
PUB TYPE Guides - Classroom Use - Instructional Materials (For Learner) (051)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Analysis of Variance; Chi Square; *Computer Assisted Instruction; *Computer Software; Correlation; Higher Education; *Statistical Analysis; Statistics; Study Guides
IDENTIFIERS *Statistical Package for the Social Sciences; *VAX Computers

ABSTRACT

Intended for classroom use only, these unpublished notes contain computer lessons on descriptive statistics using SPSS-X Release 3.0 for VAX/UNIX. Statistical measures covered include Chi-square analysis; Spearman's rank correlation coefficient; Student's t-test with two independent samples; Student's t-test with a paired sample; One-way analysis of variance, with Tukey mean comparison; Two-way analysis of variance; and Pearson's product moment coefficient of correlation. (ALF)

Nova University

Center for Computer and Information Sciences

EXAMPLES OF DATA ANALYSIS

WITH SPSS-X

THOMAS W. MACFARLAND, Ed.D.

FOR CLASSROOM USE ONLY

EXAMPLES OF DATA ANALYSIS WITH SPSS-X

Thomas W. MacFarland, Ed.D.

Unpublished notes of the author

December, 1989

Test: Descriptive Statistics

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 48

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario: Determine mean, standard deviation, and other descriptive statistics for two data sets. To make this more useful, include the concept of missing data.

X	Y
4	1
2	6
1	2
0	2
1	3
	0
	3

Files:
1. mean-sd.dat
2. mean-sd.r01
3. mean-sd.o01

Command: At the Unix prompt (%), key:

```
%spssx <mean-sd.r01> mean-sd.o01
```

Outcome: Descriptive statistics for X include:

Mean = 1.600

SD = 1.517

Further analysis can be found in the output file.

Descriptive statistics for Y include:

Mean = 2.429

SD = 1.902

Addendum:

```
***********
mean-sd.dat
***********

4 1
2 6
1 2
0 2
1 3
9 0
9 3

***********
mean-sd.r01
***********

SET WIDTH = 80
DATA LIST FILE = 'mean-sd.dat' FIXED
     / X           1
       Y           3

variable lables
       X           "Criterion X"/
       Y           "Criterion Y".

MISSING VALUES X Y (9)

FREQUENCIES VARIABLES=X Y
    /STATISTICS All

***********
mean-sd.o01
***********

116 Nov 89   SPSS-X Release 3.0 for VAX/UNIX
 15:09:41    Nova University                VAX 8550        ULTRIX 2.3

 For ULTRIX 2.3          Nova University                    License Number 19439
 This software is functional through November 30, 1989.
```

Try the new SPSS-X Release 3.0 features:

* Interactive SPSS-X command execution
* Online Help
* Nonlinear Regression
* Time Series and Forecasting (TRENDS)
* Macro Facility

* Improvements in:
* REPORT
* TABLES
* Simplified Syntax
* Matrix I/O

See SPSS-X Users Guide, Third Edition for more information on these features.

```
   1  0  SET WIDTH = 80
   2  DATA LIST FILE = 'mean-sd.dat' FIXED
   3       / X           1
   4         Y           3
   5
```

THE COMMAND ABOVE READS 1 RECORDS FROM mean-sd.dat

VARIABLE	REC	START	END	FORMAT	WIDTH	DEC
X	1	1	1	F	1	0
Y	1	3	3	F	1	0

END OF DATALIST TABLE

```
   6  variable lables
   7         X          "Criterion X"/
   8         Y          "Criterion Y".
   9
  10  MISSING VALUES X Y (9)
  11
  12  FREQUENCIES VARIABLES=X Y
  13       /STATISTICS All
```

There are 198480 bytes of memory available.
The largest contiguous area has 198480 bytes.

***** MEMORY ALLOWS A TOTAL OF 7087 VALUES, ACCUMULATED ACROSS ALL VARIABLES.
THERE ALSO MAY BE UP TO 886 VALUE LABELS FOR EACH VARIABLE.

X Criterion X

VALUE LABEL	VALUE	FREQUENCY	PERCENT	VALID PERCENT	CUM PERCENT

0	1	14.3	20.0	20.0
1	2	28.6	40.0	60.0
2	1	14.3	20.0	80.0
4	1	14.3	20.0	100.0
9	2	28.6	MISSING	
TOTAL	7	100.0	100.0	

MEAN	1.600	STD ERR	.678	MEDIAN	1.000
MODE	1.000	STD DEV	1.517	VARIANCE	2.300
KURTOSIS	1.456	S E KURT	2.000	SKEWNESS	1.118
S E SKEW	.913	RANGE	4.000	MINIMUM	.000
MAXIMUM	4.000	SUM	8.000		

VALID CASES 5 MISSING CASES 2

- -

Y Criterion Y

VALUE LABEL	VALUE	FREQUENCY	PERCENT	VALID PERCENT	CUM PERCENT
	0	1	14.3	14.3	14.3
	1	1	14.3	14.3	28.6
	2	2	28.6	28.6	57.1
	3	2	28.6	28.6	85.7
	6	1	14.3	14.3	100.0
	TOTAL	7	100.0	100.0	

MEAN	2.429	STD ERR	.719	MEDIAN	2.000
MODE	2.000	STD DEV	1.902	VARIANCE	3.619
KURTOSIS	1.795	S E KURT	1.587	SKEWNESS	.967
S E SKEW	.794	RANGE	6.000	MINIMUM	.000
MAXIMUM	6.000	SUM	17.000		

VALID CASES 7 MISSING CASES 0

PRECEDING TASK REQUIRED 0.15 SECONDS CPU TIME; 0.31 SECONDS ELAPSED

14

14 COMMAND LINES READ.
0 ERRORS DETECTED.

```
0 WARNINGS ISSUED.
0 SECONDS CPU TIME.
1 SECONDS ELAPSED TIME.
  END OF JOB.
```

Test: Chi-square Analysis (using all data)

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 671

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario: A study on stock movement, based on January indices as an indicator of yearly movement:

1. January = Up
   Feb.-Dec. = Up ..... n = 25 (Group 1 1)

2. January = Up
   Feb.-Dec. = Down ..... n = 10 (Group 1 2)

3. January = Down
   Feb.-Dec. = Up ..... n = 09 (Group 2 1)

4. January = Down
   Feb.-Dec. = Down ..... n = 09 (Group 2 2)

Ho: There is no difference between January stock prices and stock prices throughout the other 11 months.

Files:
1. z.dat
2. z.r01
3. z.o01

Command: At the Unix prompt (%), key:

```
%spssx <z.r01> z.o01
```

Outcome: Computed chi-square = 2.373

Criterion chi-square = 3.841 (alpha = .05, df = 1)

Computed chi-square (2.373) < Criterion chi-square (3.841)

Accept the Null Hypothesis ..... there is no difference between January stock prices and stock prices throughout

the other 11 months.

Addendum:

```
*****
z.dat
*****

01 1 1
02 1 1
03 1 1
04 1 1
05 1 1
06 1 1
07 1 1
08 1 1
09 1 1
10 1 1
11 1 1
12 1 1
13 1 1
14 1 1
15 1 1
16 1 1
17 1 1
18 1 1
19 1 1
20 1 1
21 1 1
22 1 1
23 1 1
24 1 1
25 1 1
26 1 2
27 1 2
28 1 2
29 1 2
30 1 2
31 1 2
32 1 2
33 1 2
34 1 2
35 1 2
36 2 1
37 2 1
38 2 1
39 2 1
40 2 1
41 2 1
42 2 1
42 2 1
```

```
43 2 1
44 2 2
45 2 2
46 2 2
47 2 2
48 2 2
49 2 2
50 2 2
51 2 2
52 2 2

*****
z.r01
*****

SET WIDTH = 80
DATA LIST FILE = 'z.dat' FIXED
     / Year        01
       January     04
       Other_11    06

variable lables
       Year        "Years, from 1927-1981"/
       January     "Stock Movement during January"/
       Other_11    "Stock Movement during Feb.-Dec.".

value labels
       January    1 'Up' 2 'Down'/
       Other_11   1 'Up' 2 'Down'.

CROSSTABS TABLES = January by Other_11
     /STATISTICS=CHISQ

*****
z.o01
*****

114 Nov 89    SPSS-X Release 3.0 for VAX/UNIX
 23:04:06     Nova University                VAX 8550          ULTRIX 2.3


 For ULTRIX 2.3          Nova University                    License Number 19439
 This software is functional through November 30, 1989.

 Try the new SPSS-X Release 3.0 features:

 * Interactive SPSS-X command execution          * Improvements in:
 * Online Help                                   *   REPORT
 * Nonlinear Regression                          *   TABLES
 * Time Series and Forecasting (TRENDS)          *   Simplified Syntax
 * Macro Facility                                *   Matrix I/O
```

See SPSS-X Users Guide, Third Edition for more information on these features.

```
0  SET WIDTH = 80
DATA LIST FILE = 'z.dat' FIXED
    / Year        01
      January     04
      Other_11    06

```

```
THE COMMAND ABOVE READS    1 RECORDS FROM z.dat
```

VARIABLE	REC	START	END	FORMAT	WIDTH	DEC
YEAR	1	1	1	F	1	0
JANUARY	1	4	4	F	1	0
OTHER_11	1	6	6	F	1	0

END OF DATALIST TABLE

```
variable lables
       Year        "Years, from 1927-1981"/
       January     "Stock Movement during January"/
       Other_11    "Stock Movement during Feb.-Dec.".

value labels
       January    1 'Up' 2 'Down'/
       Other_11   1 'Up' 2 'Down'.

CROSSTABS TABLES = January by Other_11
     /STATISTICS=CHISQ
```

```
There are 198368 bytes of memory available.
The largest contiguous area has 198368 bytes.

 ***** GIVEN WORKSPACE ALLOWS FOR  5509 CELLS WITH    2 DIMENSIONS FOR CROSSTA

114 Nov 89 SPSS-X Release 3.0 for VAX/UNIX                              Page
 23:04:07  Nova University                VAX 8550          ULTRIX 2.3


- - - - - - - - - -   C R O S S T A B U L A T I O N   O F   - - - - - - - - - -
    JANUARY   Stock Movement during January
BY  OTHER_11  Stock Movement during Feb.-Dec.
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -    PAGE  1 OF  1

                   OTHER_11
          COUNT   I
                  IUp        Down       ROW
                  I                     TOTAL
```

JANUARY		1	2	
Up	1	25	10	35
				66.0
Down	2	9	9	18
				34.0
	COLUMN TOTAL	34	19	53
		64.2	35.8	100.0

CHI-SQUARE	D.F.	SIGNIFICANCE	MIN E.F.	CELLS WITH E.F.< 5
1.53307	1	0.2157	6.453	NONE
2.37340	1	0.1234	(BEFORE YATES CORRECTION)	

NUMBER OF MISSING OBSERVATIONS = 0

PRECEDING TASK REQUIRED 0.22 SECONDS CPU TIME; 0.26 SECONDS ELAPSED

18

18 COMMAND LINES READ.
0 ERRORS DETECTED.
0 WARNINGS ISSUED.
0 SECONDS CPU TIME.
1 SECONDS ELAPSED TIME.
END OF JOB.

Test: Chi-square Analysis (using WEIGHT, to avoid excessive keying)

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 667

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario: A study on the incidence of obesity in children and obesity of parents (either one or both parents):

1. Parent = Obese
   Child = Obese ..... n = 34 (Group 1 1)

2. Parent = Obese
   Child = Nonobese ... n = 29 (Group 1 2)

3. Parent = Nonobese
   Child = Obese ..... n = 16 (Group 2 1)

4. Parent = Nonobese
   Child = Nonobese ... n = 21 (Group 2 2)

Ho: There is no difference in the incidence of obesity among children and the incidence of obesity with their respective parents.

Files:
1. zz.dat
2. zz.r01
3. zz.o01

Command: At the Unix prompt (%), key:

```
%spssx <zz.r01> z.o01
```

Outcome: Computed chi-square = 1.072

Criterion chi-square = 3.841 (alpha = .05, df = 1)

Computed chi-square (1.072) < Criterion chi-square (3.841)

Accept the Null Hypothesis ..... there is no difference

between the incidence of obesity among children and the incidence of obesity with their respective parents.

Addendum:

```
******
zz.dat
******

34 1 1
29 1 2
16 2 1
21 2 2

******
zz.r01
******

SET WIDTH = 80
DATA LIST FILE = 'zz.dat' FIXED
     / Freq      01-02
       Parent       04
       Child        06

WEIGHT by Freq

variable labels
       Parent      "Parent, either one or both"/
       Child       "Child".

value labels
       Parent      1 'Obese'
                   2 'Nonobese'.
     /
       Child       1 'Obese'
                   2 'Nonobese'.

CROSSTABS TABLES = Parent by Child
 /STATISTICS=CHISQ

******
zz.o01
******

115 Nov 89    SPSS-X Release 3.0 for VAX/UNIX
 13:15:18     Nova University                VAX 8550          ULTRIX 2.3

 For ULTRIX 2.3          Nova University                      License Number 19439
 This software is functional through November 30, 1989.
```

Try the new SPSS-X Release 3.0 features:

* Interactive SPSS-X command execution
* Online Help
* Nonlinear Regression
* Time Series and Forecasting (TRENDS)
* Macro Facility

* Improvements in:
* REPORT
* TABLES
* Simplified Syntax
* Matrix I/O

See SPSS-X Users Guide, Third Edition for more information on these features.

```
   1  0  SET WIDTH = 80
   2  DATA LIST FILE = 'zz.dat' FIXED
   3       / Freq     01-02
   4         Parent      04
   5         Child       06
   6
```

THE COMMAND ABOVE READS 1 RECORDS FROM zz.dat

VARIABLE	REC	START	END	FORMAT	WIDTH	DEC
FREQ	1	1	2	F	2	0
PARENT	1	4	4	F	1	0
CHILD	1	6	6	F	1	0

END OF DATALIST TABLE

```
   7  WEIGHT by Freq
   8
   9  variable labels
  10         Parent     "Parent, either one or both"/
  11         Child      "Child".
  12
  13  value labels
  14         Parent     1 'Obese'
  15                    2 'Nonobese'.
  16       /
  17         Child      1 'Obese'
  18                    2 'Nonobese'.
  19
  20  CROSSTABS TABLES = Parent by Child
  21   /STATISTICS=CHISQ
```

There are 198368 bytes of memory available.
The largest contiguous area has 198368 bytes.

 ***** GIVEN WORKSPACE ALLOWS FOR 5509 CELLS WITH   2 DIMENSIONS FOR CROSSTA

13:15:19 Nova University VAX 8550 ULTRIX 2.3

- - - - - - - - - - - C R O S S T A B U L A T I O N O F - - - - - - - - - -
PARENT Parent, either one or both
BY CHILD Child
- PAGE 1 OF 1

COUNT	CHILD Obese	Nonobese	ROW TOTAL
PARENT	1	2	
1 Obese	34	29	63
			63.0
2 Nonobese	16	21	37
			37.0
COLUMN TOTAL	50	50	100
	50.0	50.0	100.0

CHI-SQUARE	D.F.	SIGNIFICANCE	MIN E.F.	CELLS WITH E.F.< 5
0.68640	1	0.4074	18.500	NONE
1.07250	1	0.3004	(BEFORE YATES CORRECTION)	

NUMBER OF MISSING OBSERVATIONS = 0

115 Nov 89 SPSS-X Release 3.0 for VAX/UNIX Page
13:15:19 Nova University VAX 8550 ULTRIX 2.3

PRECEDING TASK REQUIRED 0.13 SECONDS CPU TIME; 0.18 SECONDS ELAPSED

22

22 COMMAND LINES READ.
0 ERRORS DETECTED.
0 WARNINGS ISSUED.
0 SECONDS CPU TIME.
1 SECONDS ELAPSED TIME.
END OF JOB.

Test: Spearman's Rank Correlation Coefficient

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 619

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario:

Type of Wine	Judge 1 Rating	Judge 2 Rating
1	6	5
2	5	6
3	1	2
4	3	1
5	2	4
6	4	3

Files:

1. yy.dat
2. yy.r01
3. yy.o01

Command: At the Unix prompt (%), key:

```
%spssx <yy.r01> yy.o01
```

Outcome: Computed r = .657

Criterion r = .829 (alpha = .05, n = 6)

Computed r (.657) < Criterion r (.829)

By casual observation, it would appear that there is a moderate association between the ratings of Judge 1 and Judge 2 (r > .60).

Yet, r is not significant. Recall, however, that correlation is very sensitive to n, especially when n is low, such as n = 6 in this example.

Addendum:

******

```
yy.dat
******

1 6 5
2 5 6
3 1 2
4 3 1
5 2 4
6 4 3

******
yy.r01
******

SET WIDTH = 80
DATA LIST FILE = 'yy.dat' FIXED
     / Brand          01
       Judge_1        03
       Judge_2        05

variable lables
     Brand       "Type of Wine"/
     Judge_1     "Judge #1"/
     Judge_2     "Judge #2".

NONPAR CORR VARIABLES=Judge_1 with Judge_2

******
yy.o01
******

14 Nov 89    SPSS-X Release 3.0 for VAX/UNIX
 13:08:53    Nova University                    VAX 8550        ULTRIX 2.3

 For ULTRIX 2.3          Nova University                    License Number 19439
 This software is functional through November 30, 1989.

 Try the new SPSS-X Release 3.0 features:

 * Interactive SPSS-X command execution          * Improvements in:
 * Online Help                                    *   REPORT
 * Nonlinear Regression                           *   TABLES
 * Time Series and Forecasting (TRENDS)           *   Simplified Syntax
 * Macro Facility                                 *   Matrix I/O

 See SPSS-X Users Guide, Third Edition for more information on these features.

    1  0  SET WIDTH = 80
    2     DATA LIST FILE = 'yy.dat' FIXED
    3          / Brand          01
    4            Judge_1        03
```

```
   5          Judge_2        05
   6

THE COMMAND ABOVE READS    1 RECORDS FROM yy.dat

          VARIABLE    REC  START     END          FORMAT   WIDTH   DEC

          BRAND         1      1       1          F            1     0
          JUDGE_1       1      3       3          F            1     0
          JUDGE_2       1      5       5          F            1     0

END OF DATALIST TABLE


   7  variable lables
   8        Brand       "Type of Wine"/
   9        Judge_1     "Judge #1"/
  10        Judge_2     "Judge #2".
  11
  12  NONPAR CORR VARIABLES=Judge_1 with Judge_2

There are 198384 bytes of memory available.
The largest contiguous area has 198384 bytes.

 ***** WORKSPACE ALLOWS FOR    9916 CASES FOR NONPARAMETRIC CORRELATION PROBLEM

114 Nov 89 SPSS-X Release 3.0 for VAX/UNIX                              Page
 13:08:54  Nova University                VAX 8550         ULTRIX 2.3

- - - S P E A R M A N   C O R R E L A T I O N   C O E F F I C I E N T S - - -


              JUDGE_2


JUDGE_1          .6571
              N(    6)
              SIG .078


" . " IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.
114 Nov 89 SPSS-X Release 3.0 for VAX/UNIX                              Page
 13:08:54  Nova University                VAX 8550         ULTRIX 2.3


PRECEDING TASK REQUIRED       0.07 SECONDS CPU TIME;       0.10 SECONDS ELAPSED

   13

     13 COMMAND LINES READ.
      0 ERRORS DETECTED.
```

```
0 WARNINGS ISSUED.
0 SECONDS CPU TIME.
1 SECONDS ELAPSED TIME.
  END OF JOB.
```

Test: Student's t-test with two independent samples

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 423

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario:

Group 1	Group 2
1.2	4.2
3.1	2.7
1.7	3.6
2.8	3.9
3.0	

Files:
1. x.dat
2. x.r01
3. x.o01

Command: At the Unix prompt (%), key:

```
%spssx <x.r01> x.o01
```

Outcome: Computed t = -2.39

Criterion t = -1.895 (alpha = .05, df = 7)

Computed t (-2.39) > Criterion t (-1.895)

Reject Null Hypothesis ... differences are significant

Group 2 is "more" than Group 1

Addendum:

*****
x.dat

```
*****

1 1.2
1 3.1
1 1.7
1 2.8
1 3.0
2 4.2
2 2.7
2 3.6
2 3.9

*****
x.r01
*****

SET WIDTH = 80
DATA LIST FILE = 'x.dat' FIXED
     / Group_ID     01
       Datum      03-05

variable labels
       Group_ID   "Group ID for the Study"/
       Datum      "Measured Score for Each Subject".

value labels
       Group_ID   1 'Group 1'
                  2 'Group 2'.

T-TEST GROUPS=Group_ID
      /VARIABLES=Datum


*****
x.o01
*****

113 Nov 89    SPSS-X Release 3.0 for VAX/UNIX
 15:52:55     Nova University                    VAX 8550          ULTRIX 2.3


 For ULTRIX 2.3          Nova University                    License Number 19439
 This software is functional through November 30, 1989.

 Try the new SPSS-X Release 3.0 features:

 * Interactive SPSS-X command execution          * Improvements in:
 * Online Help                                    *   REPORT
 * Nonlinear Regression                           *   TABLES
 * Time Series and Forecasting (TRENDS)           *   Simplified Syntax
 * Macro Facility                                 *   Matrix I/O
```

See SPSS-X Users Guide, Third Edition for more information on these features.

```
  1  0  SET WIDTH = 80
  2  DATA LIST FILE = 'x.dat' FIXED
  3       / Group_ID     01
  4         Datum      03-05
  5
```

THE COMMAND ABOVE READS 1 RECORDS FROM x.dat

VARIABLE	REC	START	END	FORMAT	WIDTH	DEC
GROUP_ID	1	1	1	F	1	0
DATUM	1	3	5	F	3	0

END OF DATALIST TABLE

```
  6  variable labels
  7         Group_ID    "Group ID for the Study"/
  8         Datum       "Measured Score for Each Subject".
  9
 10  value labels
 11         Group_ID    1 'Group 1'
 12                     2 'Group 2'.
 13
 14  T-TEST GROUPS=Group_ID
 15        /VARIABLES=Datum
```

There are 198464 bytes of memory available.
The largest contiguous area has 197952 bytes.

THE T-TEST PROBLEM REQUIRES 56 BYTES OF WORKSPACE

- - - - - - - - - - - - - - - - - - T - T E S T - - - - - - - - - - - - - - - - - -

GROUP 1 - GROUP_ID EQ 1
GROUP 2 - GROUP_ID EQ 2

VARIABLE		NUMBER OF CASES	MEAN	STANDARD DEVIATION	STANDARD ERROR
DATUM	Measured Score for Each Subject				
	GROUP 1	5	2.3600	0.856	0.383
	GROUP 2	4	3.6000	0.648	0.324

		*	POOLED VARIANCE ESTIMATE			*	SEPARATE VARIANCE ESTIMATE		
F VALUE	2-TAIL PROB.	*	T VALUE	DEGREES OF FREEDOM	2-TAIL PROB.	*	T VALUE	DEGREES OF FREEDOM	2-TAIL PROB.
1.75	0.675	*	-2.39	7	0.048	*	-2.47	7.00	0.043

113 Nov 89 SPSS-X Release 3.0 for VAX/UNIX Page
15:52:56 Nova University VAX 8550 ULTRIX 2.3

PRECEDING TASK REQUIRED 0.12 SECONDS CPU TIME; 0.17 SECONDS ELAPSED

16

16 COMMAND LINES READ.
0 ERRORS DETECTED.
0 WARNINGS ISSUED.
0 SECONDS CPU TIME.
1 SECONDS ELAPSED TIME.
END OF JOB.

Test: Student's t-test with a paired sample

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 432

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario:

	Starting Salary	
Pair	Male	Female
01	14300	13800
02	16500	16600
03	15400	14800
04	13500	13500
05	18500	17600
06	12800	13000
07	14500	14200
08	16200	15100
09	13400	13200
10	14200	13500

Files:
1. xx.dat
2. xx.r01
3. xx.o01

Command: At the Unix prompt (%), key:

```
%spssx <xx.r01> xx.o01
```

Outcome: Computed t = 2.91

Criterion t = 1.833 (alpha = .05, df = 9)

Computed t (2.91) > Criterion t (1.833)

Reject Null Hypothesis ... differences are significant

Males had a higher starting salary than females

Addendum:

```
******
xx.dat
******

01 14300 13800
02 16500 16600
03 15400 14800
04 13500 13500
05 18500 17600
06 12800 13000
07 14500 14200
08 16200 15100
09 13400 13200
10 14200 13500

******
xx.r01
******

SET WIDTH = 80
DATA LIST FILE = 'xx.dat' FIXED
     / Pair        01-02
       Male_$      04-08
       Female_$    10-14

variable labels
       Pair        "Paired Group ID Number"/
       Male_$      "Male Starting Salary"/
       Female_$    "Female Starting Salary".

T-TEST PAIRS=Male_$,Female_$

******
xx.o01
******

113 Nov 89    SPSS-X Release 3.0 for VAX/UNIX
 16:39:03     Nova University                VAX 8550          ULTRIX 2.3


 For ULTRIX 2.3          Nova University                    License Number 19439
 This software is functional through November 30, 1989.

 Try the new SPSS-X Release 3.0 features:

 * Interactive SPSS-X command execution          * Improvements in:
 * Online Help                                    *   REPORT
 * Nonlinear Regression                           *   TABLES
 * Time Series and Forecasting (TRENDS)           *   Simplified Syntax
 * Macro Facility                                 *   Matrix I/O
```

See SPSS-X Users Guide, Third Edition for more information on these features.

```
  1  0  SET WIDTH = 80
  2  DATA LIST FILE = 'xx.dat' FIXED
  3       / Pair        01-02
  4         Male_$      04-08
  5         Female_$    10-14
  6
```

THE COMMAND ABOVE READS 1 RECORDS FROM xx.dat

VARIABLE	REC	START	END	FORMAT	WIDTH	DEC
PAIR	1	1	2	F	2	0
MALE_$	1	4	8	F	5	0
FEMALE_$	1	10	14	F	5	0

END OF DATALIST TABLE

```
  7  variable labels
  8         Pair        "Paired Group ID Number"/
  9         Male_$      "Male Starting Salary"/
 10         Female_$    "Female Starting Salary".
 11
 12  T-TEST PAIRS=Male_$,Female_$
```

There are 198376 bytes of memory available.
The largest contiguous area has 198024 bytes.

THE T-TEST PROBLEM REQUIRES 64 BYTES OF WORKSPACE

- - - - - - - - - - - - - - - - - - T - T E S T - - - - - - - - - - - - - - - - -

VARIABLE	NUMBER OF CASES	MEAN	STANDARD DEVIATION	STANDARD ERROR
MALE_$	Male Starting Salary			
	10	14930.0000	1734.647	548.544
	10	14530.0000	1526.834	482.827
FEMALE_$	Female Starting Salary			

(DIFFERENCE) MEAN	STANDARD DEVIATION	STANDARD ERROR	* CORR.	2-TAIL PROB.	T VALUE	DEGREES OF FREEDOM	2-TAIL PROB.
400.0000	434.613	137.437	* 0.972	0.000	2.91	9	0.017

---

113 Nov 89 SPSS-X Release 3.0 for VAX/UNIX Page
16:39:04 Nova University VAX 8550 ULTRIX 2.3

PRECEDING TASK REQUIRED 0.08 SECONDS CPU TIME; 0.23 SECONDS ELAPSED

13

13 COMMAND LINES READ.
0 ERRORS DETECTED.
0 WARNINGS ISSUED.
0 SECONDS CPU TIME.
1 SECONDS ELAPSED TIME.
END OF JOB.

Test: Oneway Analysis of Variance, with Tukey Mean Comparison

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 503

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario: A study on the observation that the ability to tolerate pain (i.e., pain threshold) may be related to hair color (the measured score).

Light Blond	Dark Blond	Light Brunette	Dark Brunette
62	63	42	32
60	57	50	39
71	52	41	51
55	41	37	30
48	43		35

Note. Notice how group n is not equal.

Ho: There is no difference in the ability to tolerate pain between people with various hair colors (e.g., light blond, dark blond, light brunette, dark brunette).

Files:
1. anova1.dat
2. anova1.r01
3. anova1.o01

Command: At the Unix prompt (%), key:

```
%spssx <anova1.r01> anova1.o01
```

Outcome: Computed F = 6.79

Criterion F = 3.49 (alpha = .05, df = 3,15)

Computed F (6.79) > Criterion F (2.49)

Reject the Null Hypothesis ..... there are differences between groups.

Consult the mean comparison (e.g., Tukey, LSD) part of the outcome file to better understand differences between groups:

(*) DENOTES PAIRS OF GROUPS SIGNIFICANTLY DIFFERENT AT THE 0.050 LEVEL

Mean	Group	Grp 4	Grp 3	Grp 2	Grp 1
37.4000	Grp 4				
42.5000	Grp 3				
51.2000	Grp 2				
59.2000	Grp 1	*	*		

Using traditional methodology, you could also visually present mean comparisons by using underscores:

Group 4	Group 3	Group 2	Group 1
Dark Brunette	Light Brunette	Dark Blond	Light Blond
37.40	42.50	51.20	59.20

Addendum:

****
anova1.dat
****

```
1 62
1 60
1 71
1 55
1 48
2 63
2 57
2 52
2 41
```

```
2 43
3 42
3 50
3 41
3 37
4 32
4 39
4 51
4 30
4 35

**********
anova1.r01
**********

SET WIDTH = 80
DATA LIST FILE = 'anova1.dat' FIXED
     / Group         01
       Score      03-04

variable lables
     Group       "Group, Based on Hair Color"/
     Score       "Pain Threshold Score".

value labels
     Group 1 'Light Blond'
           2 'Dark Blond'
           3 'Light Brunette'
           4 'Dark Brunette'.

ONEWAY Score by Group (1,4)
 /RANGES=TUKEY

**********
anova1.o01
**********

115 Nov 89   SPSS-X Release 3.0 for VAX/UNIX
 22:25:44    Nova University                VAX 8550          ULTRIX 2.3


 For ULTRIX 2.3          Nova University                    License Number 19439
 This software is functional through November 30, 1989.

 Try the new SPSS-X Release 3.0 features:

 * Interactive SPSS-X command execution          * Improvements in:
 * Online Help                                   *   REPORT
 * Nonlinear Regression                          *   TABLES
 * Time Series and Forecasting (TRENDS)          *   Simplified Syntax
 * Macro Facility                                *   Matrix I/O
```

See SPSS-X Users Guide, Third Edition for more information on these features.

```
  1  0  SET WIDTH = 80
  2  DATA LIST FILE = 'anova1.dat' FIXED
  3       / Group          01
  4         Score       03-04
  5
```

THE COMMAND ABOVE READS 1 RECORDS FROM anova1.dat

VARIABLE	REC	START	END	FORMAT	WIDTH	DEC
GROUP	1	1	1	F	1	0
SCORE	1	3	4	F	2	0

END OF DATALIST TABLE

```
  6  variable lables
  7       Group       "Group, Based on Hair Color"/
  8       Score       "Pain Threshold Score".
  9
 10  value labels
 11       Group 1 'Light Blond'
 12             2 'Dark Blond'
 13             3 'Light Brunette'
 14             4 'Dark Brunette'.
 15
 16  ONEWAY Score by Group (1,4)
 17   /RANGES=TUKEY
```

448 BYTES OF MEMORY REQUIRED FOR ONEWAY PROCEDURE.

There are 198472 bytes of memory available.
The largest contiguous area has 198472 bytes.

- - - - - - - - - - - - - - - O N E W A Y - - - - - - - - - - - - - - - -

Variable SCORE Pain Threshold Score
By Variable GROUP Group, Based on Hair Color

ANALYSIS OF VARIANCE

SOURCE	D.F.	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB.
BETWEEN GROUPS	3	1360.7263	453.5754	6.7914	.0041

```
WITHIN GROUPS              15     1001.8000       66.7867

TOTAL                      18     2362.5263
```



```
- - - - - - - - - - - - - - - - O N E W A Y - - - - - - - - - - - - - - - - -

     Variable  SCORE      Pain Threshold Score
  By Variable  GROUP      Group, Based on Hair Color


MULTIPLE RANGE TEST


TUKEY-HSD PROCEDURE
RANGES FOR THE 0.050 LEVEL -

          4.07   4.07   4.07

THE RANGES ABOVE ARE TABLE RANGES.
THE VALUE ACTUALLY COMPARED WITH MEAN(J)-MEAN(I) IS..
        5.7787 * RANGE * DSQRT(1/N(I) + 1/N(J))


   (*) DENOTES PAIRS OF GROUPS SIGNIFICANTLY DIFFERENT AT THE 0.050 LEVEL

                          G G G G
                          r r r r
                          p p p p

     Mean       Group     4 3 2 1

    37.4000     Grp 4
    42.5000     Grp 3
    51.2000     Grp 2
    59.2000     Grp 1     * *
```



```
PRECEDING TASK REQUIRED       0.18 SECONDS CPU TIME;       0.29 SECONDS ELAPSED

  18

    18 COMMAND LINES READ.
     0 ERRORS DETECTED.
     0 WARNINGS ISSUED.
     0 SECONDS CPU TIME.
```

```
1 SECONDS ELAPSED TIME.
  END OF JOB.
```

Test: Twoway Analysis of Variance

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 533

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario: A study on the effect of type of golf club and type of golf ball on distance of hit.

		Brand			
		A	B	C	D
Club	Driver	226.4	238.3	240.5	219.8
		232.6	231.7	246.9	228.7
		234.0	227.7	240.3	232.9
		220.7	237.2	244.7	237.6
	Five-Iron	163.8	184.4	179.0	157.8
		179.4	180.6	168.0	161.8
		168.6	179.5	165.2	162.1
		173.4	186.2	156.5	160.3

Ho: There is no difference between golf clubs (driver and five-iron) and golf balls (Brands A, B, C, and D) and the distance of hit.

Files:
1. anova2.dat
2. anova2.r01
3. anova2.o01

Command: At the Unix prompt (%), key:

```
%spssx <anova2.r01> anova2.o01
```

Outcome:

Club ..... F = 936.752 (Significance .00)

Ball ..... F = 7.791 (Significance .001)

Club by Ball
Interaction ... F = 7.452 (Significance .001)

Reject the Null Hypothesis ..... there are differences between clubs, balls, and interaction of clubs and balls.

Bonferroni methodology would be needed to further analyze where differences and similarities exist.

Addendum:

```
**********
anova2.dat
**********

1 1 226.4
1 1 232.6
1 1 234.0
1 1 220.7
1 2 238.3
1 2 231.7
1 2 227.7
1 2 237.2
1 3 240.5
1 3 246.9
1 3 240.3
1 3 244.7
1 4 219.8
1 4 228.7
1 4 232.9
1 4 237.6
2 1 163.8
2 1 179.4
2 1 168.6
2 1 173.4
2 2 184.4
2 2 180.6
2 2 179.5
2 2 186.2
2 3 179.0
2 3 168.0
2 3 165.2
```

```
2 3 156.5
2 4 157.8
2 4 161.8
2 4 162.1
2 4 160.3

**********
anova2.r01
**********

SET WIDTH = 80
DATA LIST FILE = 'anova2.dat' FIXED
     / Club          01
       Ball          03
       Distance  05-09

VARIABLE LABELS
       Club        "Type of Golf Club"/
       Ball        "Type of Golf Ball"/
       Distance    "Distance of Hit".

VALUE LABELS
       Club        1 'Driver'
                   2 'Five-Iron'

     / Ball        1 'Brand A'
                   2 'Brand B'
                   3 'Brand C'
                   4 'Brand D'.

ANOVA Distance by Club (01,02) Ball (01,04)

**********
anova2.o01
**********

16 Nov 89   SPSS-X Release 3.0 for VAX/UNIX
 13:13:07   Nova University              VAX 8550        ULTRIX 2.3


 For ULTRIX 2.3          Nova University                License Number 19439
 This software is functional through November 30, 1989.

 Try the new SPSS-X Release 3.0 features:

 * Interactive SPSS-X command execution        * Improvements in:
 * Online Help                                  *   REPORT
 * Nonlinear Regression                         *   TABLES
 * Time Series and Forecasting (TRENDS)         *   Simplified Syntax
 * Macro Facility                               *   Matrix I/O

 See SPSS-X Users Guide, Third Edition for more information on these features.
```

```
   1  0  SET WIDTH = 80
   2  DATA LIST FILE = 'anova2.dat' FIXED
   3       / Club          01
   4         Ball          03
   5         Distance  05-09
   6
```

THE COMMAND ABOVE READS 1 RECORDS FROM anova2.dat

VARIABLE	REC	START	END	FORMAT	WIDTH	DEC
CLUB	1	1	1	F	1	0
BALL	1	3	3	F	1	0
DISTANCE	1	5	9	F	5	0

END OF DATALIST TABLE

```
   7  VARIABLE LABELS
   8         Club          "Type of Golf Club"/
   9         Ball          "Type of Golf Ball"/
  10         Distance      "Distance of Hit".
  11
  12  VALUE LABELS
  13         Club          1 'Driver'
  14                       2 'Five-Iron'
  15
  16       / Ball          1 'Brand A'
  17                       2 'Brand B'
  18                       3 'Brand C'
  19                       4 'Brand D'.
  20
  21  ANOVA Distance by Club (01,02) Ball (01,04)
```

'ANOVA' PROBLEM REQUIRES 992 BYTES OF MEMORY.

* * * A N A L Y S I S O F V A R I A N C E * * *

DISTANCE Distance of Hit
by CLUB Type of Golf Club
BALL Type of Golf Ball

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	32893.847	4	8223.462	240.031	.000
CLUB	32093.111	1	32093.111	936.752	.00

BALL	800.736	3	266.912	7.791	.001
2-Way Interactions	765.961	3	255.320	7.452	.001
CLUB BALL	765.961	3	255.320	7.452	.001
Explained	33659.809	7	4808.544	140.354	.000
Residual	822.240	24	34.260		
Total	34482.049	31	1112.324		

32 cases were processed.
0 cases (.0 pct) were missing.

PRECEDING TASK REQUIRED 0.34 SECONDS CPU TIME; 0.37 SECONDS ELAPSED

22

22 COMMAND LINES READ.
0 ERRORS DETECTED.
0 WARNINGS ISSUED.
1 SECONDS CPU TIME.
1 SECONDS ELAPSED TIME.
END OF JOB.

Test: Pearson's Product Moment Coefficient of Correlation

Source: McClave, James T., and Frank H. Dietrich, II. STATISTICS, 4th edition. San Francisco, California: Dellen Publishing Company, 1988. ISBN 0-02-379260-4 Page 712

Software: SPSS-X Release 3.0 for VAX/UNIX

Scenario:

Year	Number of Salespeople	Dollar Value of Sales
1978	15	135000
1979	18	163000
1980	24	233000
1981	22	241000
1982	25	263000
1983	29	293000
1984	30	341000
1985	32	326000
1986	35	363000
1987	38	415000

Files:

1. y.dat
2. y.r01
3. y.o01

Command: At the Unix prompt (%), key:

```
%spssx <y.r01> y.o01
```

Outcome: Computed r = .987

Criterion r = .549 (alpha = .05, n = 10)

Computed r (.987) > Criterion r (.549)

Reject the Null Hypothesis ..... the association between x (Number of Salespeople) and y (Dollar Value of Sales) is significant.

That is to say, as the number of salespeople increases the dollar value of sales also increases.

Addendum:

```
*****
y.dat
*****

1978 15 135000
1979 18 163000
1980 24 233000
1981 22 241000
1982 25 263000
1983 29 293000
1984 30 341000
1985 32 326000
1986 35 363000
1987 38 415000

*****
y.r01
*****

SET WIDTH = 80
DATA LIST FILE = 'y.dat' FIXED
     / Year        01-04
       Number      06-07
       Sales       09-14

variable lables
       Year        "Year"/
       Number      "Number of Salespeople"/
       Sales       "Total Sales".

CORRELATION Number with Sales

*****
y.o01
*****

114 Nov 89    SPSS-X Release 3.0 for VAX/UNIX
 12:47:49     Nova University                VAX 8550        ULTRIX 2.3


 For ULTRIX 2.3          Nova University                    License Number 19439
 This software is functional through November 30, 1989.

 Try the new SPSS-X Release 3.0 features:

 * Interactive SPSS-X command execution          * Improvements in:
 * Online Help                                   *   REPORT
 * Nonlinear Regression                          *   TABLES
 * Time Series and Forecasting (TRENDS)          *   Simplified Syntax
 * Macro Facility                                *   Matrix I/O
```

See SPSS-X Users Guide, Third Edition for more information on these features.

```
   1  0  SET WIDTH = 80
   2  DATA LIST FILE = 'y.dat' FIXED
   3       / Year        01-04
   4         Number      06-07
   5         Sales       09-14
   6

THE COMMAND ABOVE READS    1 RECORDS FROM y.dat

          VARIABLE   REC  START     END          FORMAT  WIDTH   DEC

          YEAR         1      1       4          F           4     0
          NUMBER       1      6       7          F           2     0
          SALES        1      9      14          F           6     0

END OF DATALIST TABLE


   7  variable lables
   8         Year        "Year"/
   9         Number      "Number of Salespeople"/
  10         Sales       "Total Sales".
  11
  12  CORRELATION Number with Sales

*****PEARSON CORR PROBLEM REQUIRES       80 BYTES WORKSPACE *****

114 Nov 89 SPSS-X Release 3.0 for VAX/UNIX                          Page
 12:47:51  Nova University                VAX 8550        ULTRIX 2.3

- - - -  P E A R S O N   C O R R E L A T I O N   C O E F F I C I E N T S  - - -

              SALES

NUMBER         .9870
              (   10)
              P= .000


 (COEFFICIENT / (CASES) / 1-TAILED SIG)

 " . " IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED
114 Nov 89 SPSS-X Release 3.0 for VAX/UNIX                          Page
 12:47:51  Nova University                VAX 8550        ULTRIX 2.3
```

```
RECEDING TASK REQUIRED        0.08 SECONDS CPU TIME;        0.75 SECONDS ELAPSED

 13

   13 COMMAND LINES READ.
    0 ERRORS DETECTED.
    0 WARNINGS ISSUED.
    0 SECONDS CPU TIME.
    2 SECONDS ELAPSED TIME.
      END OF JOB.
```