[redacted]

Examples

ANOVA

This example, illustrating the use of ANOVA, is a three-way analysis of variance with one covariate. The data are 500 cases from the 1980 General Social Survey. The variables are

- PRESTIGE—the respondent's occupational prestige scale score. PRESTIGE is the dependent variable.
- EDUC—the respondent's education in years.
- RACE—the respondent's race, coded 1=WHITE, 2=BLACK, and 3=OTHER.
- SEX—the respondent's sex, coded 1=MALE and 2=FEMALE.
- REGION—the respondent's residence, coded as one of nine regions.

The task is twofold: determine the degree to which the American occupational structure differs across race, sex, and region; and measure the effect of the respondent's educational level, since education might prove to be a concomitant influence. The data are in an external file named AANOVA.DAT. The SPSS/PC commands in the command file named on the INCLUDE command are

```
DATA LIST FILE='AANOVA.DAT'
 / PRESTIGE 1-2 EDUC 3-4 RACE 5 SEX 6 REGION 7.
VARIABLE LABELS PRESTIGE "Resp's Occupational Prestige Score"
                EDUC 'Highest Year School Completed'
                REGION 'Region of Interview'.
ANOVA PRESTIGE BY REGION(1,9) SEX,RACE(1,2) WITH EDUC
      /STATISTICS 2
      /OPTIONS 10,11.
FINISH.
```

- The DATA LIST command assigns variable names and gives column locations for the variables in the analysis (see DATA LIST).
- The VARIABLE LABELS command completes the file definition (see VARIABLE LABELS).
- The ANOVA command names PRESTIGE as the dependent variable; REGION, SEX, and RACE as the factors; and EDUC as the covariate. The minimum and maximum values for REGION are 1 and 9, and the minimum and maximum values for both SEX and RACE are 1 and 2. Since variable RACE actually has values 1, 2, and 3, cases with value 3 are eliminated from the model.
- Statistic 2 requests the regression coefficient for the covariate EDUC.
- Option 10 requests the hierarchical approach for decomposing sums of squares. The covariate EDUC is assessed first to establish statistical control. Then the effect of REGION is assessed; next, the effect of SEX adjusted for REGION; and next, the effect of RACE adjusted for REGION and SEX. Finally, each of the interaction effects is assessed.
- Option 11 tells SPSS/PC to display the results within a narrow width.

The display produced by ANOVA is on the following page. The exact appearance of the printed output depends on the characters available on the printer used.

ANOVA display

```
         PRESTIGE Resp's Occupational Prestige Score
    BY   REGION   Region of Interview
         SEX
         RACE
  WITH   EDUC     Highest Year School Completed
```

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
COVARIATES	23715.522	1	23715.522	191.701	0.000
EDUC	23715.522	1	23715.522	191.70[illegible]	0.000
MAIN EFFECTS	2708.380	10	270.838	2.18[illegible]	0.018
REGION	1260.552	8	157.569	1.274	0.255
SEX	22.413	1	22.413	0.181	0.671
RACE	1425.415	1	1425.415	11.522	0.001
2-WAY INTERACTIONS	3144.833	17	184.990	1.495	0.092
REGION SEX	1349.220	8	168.653	1.363	0.211
REGION RACE	1138.839	8	142.355	1.151	0.328
SEX RACE	534.154	1	534.154	4.318	0.038
3-WAY INTERACTIONS	1663.399	6	277.233	2.241	0.039
REGION SEX RACE	1663.399	6	277.233	2.241	0.039
EXPLAINED	31232.135	34	918.592	7.425	0.0
RESIDUAL	52205.957	422	123.711		
TOTAL	83438.092	456	182.978		

COVARIATE	RAW REGRESSION COEFFICIENT
EDUC	2.331

```
500 CASES WERE PROCESSED.
 43 CASES (  8.6 PCT) WERE MISSING.
```

CLUSTER

This example clusters cities using data from the 1982 *Information Please Almanac.* These are the most populous 25 cities in the U.S. in 1980. The clustering variables are

- CHURCHES—number of churches.
- PARKS—number of parks. Some cities only report total acreage and have a missing-value code of 9999 for number of parks.
- PHONES—number of telephones.
- TVS—number of television sets.
- RADIOST—number of radio stations.
- TVST—number of television stations.
- POP80—city population in 1980.
- TAXRATE—property tax rate in city.

We use these variables to cluster cities into groups that are relatively homogeneous with respect to these variables. Cities differ on these variables simply as a function of their population. We therefore rescale the variables to number of parks etc., per person. The data are in an external file named ACLUS.DAT. The SPSS/PC commands in the command file named on the INCLUDE command are

```
DATA LIST FILE='ACLUS.DAT'
   / CITY 6-18(A) POP80 53-60
   / CHURCHES 10-13 PARKS 14-17 PHONES 18-25 TVS 26-32
     RADIOST 33-35 TVST 36-38 TAXRATE 52-57(2)
   /.
MISSING VALUE PARKS (9999).
COMPUTE CHURCHES=CHURCHES/POP80.
COMPUTE PARKS=PARKS/POP80.
COMPUTE PHONES=PHONES/POP80.
COMPUTE TVS=TVS/POP80.
COMPUTE RADIOST=RADIOST/POP80.
COMPUTE TVST=TVST/POP80.
CLUSTER CHURCHES TO TAXRATE /
        METHOD = BAVERAGE/
        ID = CITY /
        PRINT = CLUSTER(3,5) DISTANCE SCHEDULE
        PLOT = VICICLE HICICLE DENDROGRAM.
FINISH.
```

- DATA LIST names the file that contains the data and gives variable names and column locations. There are three records per case, but no data are read from the third record, so an extra slash is included to skip the unread record for each case (see DATA LIST).
- The MISSING VALUE command tells SPSS/PC to treat the value 9999 as a user-missing value for the variable PARKS (see MISSING VALUE).
- The COMPUTE commands divide each measure by the population of the city in 1980. This yields the number of churches, phones, etc., per person (see COMPUTE).
- The CLUSTER variable specification names six variables using the TO convention.
- METHOD clusters cases by the method of average linkage between groups. It uses the default squared Euclidean distances over the six specified variables.
- The ID subcommand requests that the string variable CITY be used to label CLUSTER output.
- The cluster membership table is part of the default display (Figure A).
- PRINT requests the computed distances between cases, the cluster to which each case belongs for the 3-, 4-, and 5-cluster solutions, and the cluster agglomeration schedule (see Figure B).
- PLOT presents the cluster solution both as a horizontal (Figure C) and a vertical icicle plot (Figure D) as well as a dendrogram (Figure E).

The display is shown on the following pages. The exact appearance of the printed display will depend on the characters available on the printer used.

A Cluster membership of cases

Data Information

22 unweighted cases accepted.
3 cases rejected because of missing value.

Squared Euclidean measure used.

1 Agglomeration method specified.

Cluster Membership of Cases using Average Linkage (Between Groups)

Label	Case	Number of Clusters 5	4	3
Baltimore	1	1	1	1
Chicago	2	2	2	2
Cleveland	3	1	1	1
Columbus	4	3	1	1
Dallas	5	4	3	3
Denver	6	4	3	3
Detroit	7	3	1	1
Houston	8	4	3	3
Indianapolis	9	5	4	2
Jacksonville	10	4	3	3
Los Angeles	11	4	3	3
Memphis	12	3	1	1
Nashville	13	1	1	1
New Orleans	14	2	2	2
New York	15	2	2	2
Philadelphia	16	1	1	1
Phoenix	17	4	3	3
San Diego	18	3	1	1
San Francisco	19	3	1	1
San Jose	20	4	3	3
Seattle	21	4	3	3
Washington	22	4	3	3

B Agglomeration schedule for clustering

Agglomeration Schedule using Average Linkage (Between Groups)

Stage	Clusters Combined Cluster 1	Cluster 2	Coefficient	Stage Cluster 1st Appears Cluster 1	Cluster 2	Next Stage
1	5	20	.213571	0	0	8
2	17	22	.261159	0	0	6
3	4	18	.292620	0	0	11
4	8	21	.656814	0	0	8
5	3	13	3.067433	0	0	14
6	10	17	3.173483	0	2	16
7	1	16	5.655860	0	0	14
8	5	8	8.083633	1	4	12
9	7	12	13.270877	0	0	11
10	14	15	16.843185	0	0	13
11	4	7	33.221954	3	9	15
12	5	11	38.605591	8	0	17
13	2	14	48.604095	0	10	19
14	1	3	73.268372	7	5	18
15	4	19	88.134521	11	0	18
16	6	10	97.164627	0	6	17
17	5	6	250.892181	12	16	20
18	1	4	651.010742	14	15	20
19	2	9	1026.891846	13	0	21
20	1	5	1710.032959	18	17	21
21	1	2	5559.281250	20	19	0

C Horizontal icicle plot

```
Horizontal Icicle Plot Using Average Linkage (Between Groups)

                          Number of Clusters

                                     111111111122
         C A S E           123456789012345678901
Label               Seq    +++++++++++++++++++++

Indianapolis          9    XXXXXXXXXXXXXXXXXXXXX
                           XXX
                           XXX
New York             15    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXX
                           XXXXXXXXXXXX
New Orleans          14    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXX
                           XXXXXXXXX
Chicago               2    XXXXXXXXXXXXXXXXXXXXX
                           X
                           X
Washington           22    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXXXX
Phoenix              17    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXX
Jacksonville         10    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXX
                           XXXXXX
Denver                6    XXXXXXXXXXXXXXXXXXXXX
                           XXXXX
                           XXXXX
Los Angeles          11    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXX
                           XXXXXXXXXX
Seattle              21    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXX
Houston               8    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXX
                           XXXXXXXXXXXXXX
San Jose             20    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXXXXX
Dallas                5    XXXXXXXXXXXXXXXXXXXXX
                           XX
                           XX
San Francisco        19    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXX
                           XXXXXXX
Memphis              12    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXX
                           XXXXXXXXXXXXX
Detroit               7    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXX
                           XXXXXXXXXXX
San Diego            18    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXXXX
Columbus              4    XXXXXXXXXXXXXXXXXXXXX
                           XXXX
                           XXXX
Nashville            13    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXXXX
Cleveland             3    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXX
                           XXXXXXXX
Philadelphia         16    XXXXXXXXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXX
                           XXXXXXXXXXXXXXX
Baltimore             1    XXXXXXXXXXXXXXXXXXXXX
```

D Vertical icicle plot

```
Vertical Icicle Plot using Average Linkage (Between Groups)

   (Down) Number of Clusters   (Across) Case Label and number

    I N N C W P J D L S H S D S M D S C N C P B
    n e e h a h a e o e o a a a e e a o a l h a
    d w w i s o c n s a u n l n m t n l s e i l
    i     c h e k v   t s   l   p r   u h v l t
    a Y O a i n s e A t t J a F h o D m v e a i
    n o r g n i o r n l o o s r i i i b i l d m
    a r l o g x n   g e n s   a s t e u l a e o
    p k e   t   v   e     e   n     g s l n l r
    o   a   o   i   l         c     o   e d p e
    l   n   n   l   e         i             h
    i   s       l   s         s             i
    s           e             c             a
                              o

      1 1   2 1 1   1 2   2   1 1   1   1   1
    9 5 4 2 2 7 0 6 1 1 8 0 5 9 2 7 8 4 3 3 6 1
 1 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
 2 +XXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
 3 +XXXXXXX XXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXX
 4 +X XXXXX XXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXXXX
 5 +X XXXXX XXXXXXXXXXXXXXXXX XXXXXXXXX XXXXXXX
 6 +X XXXXX XXXXXXX XXXXXXXXX XXXXXXXXX XXXXXXX
 7 +X XXXXX XXXXX X XXXXXXXXX XXXXXXXXX XXXXXXX
 8 +X XXXXX XXXXX X XXXXXXXXX X XXXXXXX XXXXXXX
 9 +X XXXXX XXXXX X XXXXXXXXX X XXXXXXX XXX XXX
10 +X XXX X XXXXX X XXXXXXXXX X XXXXXXX XXX XXX
11 +X XXX X XXXXX X X XXXXXXX X XXXXXXX XXX XXX
12 +X XXX X XXXXX X X XXXXXXX X XXX XXX XXX XXX
13 +X X X X XXXXX X X XXXXXXX X XXX XXX XXX XXX
14 +X X X X XXXXX X X XXXXXXX X X X XXX XXX XXX
15 +X X X X XXXXX X X XXX XXX X X X XXX XXX XXX
16 +X X X X XXXXX X X XXX XXX X X X XXX XXX X X
17 +X X X X XXX X X X XXX XXX X X X XXX XXX X X
18 +X X X X XXX X X X XXX XXX X X X XXX X X X X
19 +X X X X XXX X X X X X XXX X X X XXX X X X X
20 +X X X X XXX X X X X X XXX X X X X X X X X X
21 +X X X X X X X X X X X XXX X X X X X X X X X
```

E Dendrogram

```
Dendrogram using Average Linkage (Between Groups)

                            Rescaled Distance Cluster Combine

    C A S E         0         5        10        15        20        25
  Label         Seq +---------+---------+---------+---------+---------+

Dallas          5   -+
San Jose       20   -+
Houston         8   -+-+
Seattle        21   -+ |
Los Angeles    11   -+ +-----------+
Phoenix        17   -+ |           |
Washington     22   -+ |           |
Jacksonville   10   -+-+           |
Denver          6   -+             +----------------------------------+
Cleveland       3   -+             |                                  |
Nashville      13   -+---+         |                                  |
Baltimore       1   -+   |         |                                  |
Philadelphia   16   -+   +---------+                                  |
Columbus        4   -+   |                                            |
San Diego      18   -+   |                                            |
Detroit         7   -+---+                                            |
Memphis        12   -+                                                |
San Francisco  19   -+                                                |
New Orleans    14   -+                                                |
New York       15   -+-------+                                        |
Chicago         2   -+       +----------------------------------------+
Indianapolis    9   ---------+
```

CORRELATION

This example analyzes 1979 prices and earnings in 45 cities around the world, compiled by the Union Bank of Switzerland. The variables are

- FOOD—the average net cost of 39 different food and beverage items in the city, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- RENT—the average gross monthly rent in the city, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- SERVICE—the average cost of 28 different goods and services in the city, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- PUBTRANS—the average cost of a three-mile taxi ride within city limits, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- TEACHER, COOK, ENGINEER, MECHANIC, BUS—the average gross annual earnings of primary-grade teachers in public schools, cooks, electrical engineers, automobile mechanics, and municipal bus drivers; working from 5 to 10 years in their respective occupations. Each of these variables is expressed as a percentage above or below those of Zurich, where Zurich equals 100%.

This example analyzes the degree to which variation in the costs of goods and services in a city is related to variation in earnings in several occupations. CORRELATION is used to compute correlations between the average costs of various goods and services and the average gross earnings in five different occupations. The data are in an external file named ACORR.DAT. The SPSS/PC commands in the file named on the INCLUDE command are

```
DATA LIST FILE='ACORR.DAT'
 /FOOD 2-4 RENT 6-8 PUBTRANS 10-12 TEACHER 14-16 COOK 18-20
  ENGINEER 22-24 SERVICE 26-28 MECHANIC 30-32 BUS 34-36.
CORRELATION VARIABLES=FOOD RENT PUBTRANS TEACHER COOK ENGINEER/
             SERVICE PUBTRANS WITH MECHANIC BUS
             /STATISTICS=1,2
             /OPTIONS=6.
FINISH.
```

- The DATA LIST command assigns variable names and gives column locations for the variables in the analysis (see DATA LIST).
- The CORRELATION command requests two correlation matrices. The first variable list produces correlation coefficients for each variable with every other variable. However, the redundant coefficients will be suppressed with Option 6. The second variable list produces four coefficients, pairing SERVICE with MECHANIC and BUS, and PUBTRANS with MECHANIC and BUS.
- The STATISTICS subcommand requests the mean, standard deviation, and number of nonmissing cases for each variable, and the cross-product deviations and covariance for each pair of variables. The statistics for all the variable lists precede all the correlation matrices in the CORRELATION display.
- Option 6 suppresses redundant coefficients in both correlation matrices.

The display produced by CORRELATION is on the following page. The exact appearance of the printed display will depend on the characters available on the printer used.

A CORRELATION statistics

Variable	Cases	Mean	Std Dev
FOOD	43	70.9767	18.8319
RENT	43	122.2558	95.7054
PUBTRANS	43	48.8837	24.9258
TEACHER	43	38.3023	25.4797
COOK	43	64.9767	30.5626
ENGINEER	43	60.1163	26.4802
SERVICE	43	73.5116	18.9892
PUBTRANS	43	47.8372	24.0445
MECHANIC	43	50.4884	31.0762
BUS	43	42.9535	27.3652

Variables		Cases	Cross-Prod Dev	Variance-Covar
FOOD	RENT	43	18459.2558	439.5061
FOOD	PUBTRANS	43	11121.8837	264.8068
FOOD	TEACHER	43	11244.3023	267.7215
FOOD	COOK	43	9574.9767	227.9756
FOOD	ENGINEER	43	10106.1163	240.6218
RENT	PUBTRANS	43	-3527.7209	-83.9934
RENT	TEACHER	43	-2515.3256	-59.8887
RENT	COOK	43	17443.2558	415.3156
RENT	ENGINEER	43	26859.7209	639.5172
PUBTRANS	TEACHER	43	18659.5116	444.2741
PUBTRANS	COOK	43	19394.8837	461.7829
PUBTRANS	ENGINEER	43	17621.5814	419.5615
TEACHER	COOK	43	20326.3023	483.9596
TEACHER	ENGINEER	43	18627.4884	443.5116
COOK	ENGINEER	43	25524.1163	607.7171

Variables		Cases	Cross-Prod Dev	Variance-Covar
SERVICE	MECHANIC	43	11965.2558	284.8870
SERVICE	BUS	43	12806.0233	304.9053
PUBTRANS	MECHANIC	43	23496.4186	559.4385
PUBTRANS	BUS	43	21561.6744	513.3732

B CORRELATION matrices

```
Variable            Variable            Variable
Pair                Pair                Pair
--------            --------            --------

FOOD        .2439   FOOD        .5641   FOOD        .5579
WITH     N(   43)   WITH     N(   43)   WITH     N(   43)
RENT     SIG .058   PUBTRANS SIG .000   TEACHER  SIG .000

FOOD        .3961   FOOD        .4825   RENT       -.0352
WITH     N(   43)   WITH     N(   43)   WITH     N(   43)
COOK     SIG .004   ENGINEER SIG .001   PUBTRANS SIG .411

RENT       -.0246   RENT        .1420   RENT         2523
WITH     N(   43)   WITH     N(   43)   WITH     N(   43)
TEACHER  SIG .438   COOK     SIG .182   ENGINEER SIG .051

"." is printed if a coefficient cannot be computed.

Variable            Variable            Variable
Pair                Pair                Pair
--------            --------            --------

PUBTRANS    .6995   PUBTRANS    .6062   PUBTRANS    .6357
WITH     N(   43)   WITH     N(   43)   WITH     N(   43)
TEACHER  SIG .000   COOK     SIG .000   ENGINEER SIG .000

TEACHER     .6215   TEACHER     .6573   COOK        .7509
WITH     N(   43)   WITH     N(   43)   WITH     N(   43)
COOK     SIG .000   ENGINEER SIG .000   ENGINEER SIG .000

"." is printed if a coefficient cannot be computed.

Variable            Variable            Variable
Pair                Pair                Pair
--------            --------            --------

SERVICE     .4828   SERVICE     .5868   PUBTRANS    .7487
WITH     N(   43)   WITH     N(   43)   WITH     N(   43)
MECHANIC SIG .001   BUS      SIG .000   MECHANIC SIG .000

PUBTRANS    .7802
WITH     N(   43)
BUS      SIG .000

"." is printed if a coefficient cannot be computed
```

CROSSTABS

This example uses CROSSTABS to examine how respondents in different age groups answer a question on alcohol-drinking habits. The data are drawn from a 500-case sample from the 1980 General Social Survey. The variables are

- AGE—the respondent's age recoded to four categories.
- DRUNK—the response to the question, Did you ever drink too much?

The raw data are stored in an external file named AXTABS.DAT. The SPSS/PC commands in the command file named on the INCLUDE command are

```
DATA LIST FILE='AXTABS.DAT'
 /DRUNK 1 AGE 2-3.
RECODE AGE (LOW THRU 29=1) (29 THRU 40=2) (40 THRU 58=3)
           (58 THRU HI=4)  DRUNK (1=1) (2=2) (ELSE=8
MISSING VALUE DRUNK(8).
VARIABLE LABELS AGE 'Age in Four Categories'
                DRUNK 'Ever Drink Too Much'.
VALUE LABELS AGE 1 'YoungestQuarter' 4 'Oldest  Quarter'
             DRUNK 1 'Yes' 2 'No' 8 "Don't Drink/NA"
CROSSTABS TABLES=DRUNK BY AGE
           OPTIONS=3,4
           STATISTICS=1,4,6,7,8,9.
FINISH.
```

- The DATA LIST command identifies the data file, and gives the names and locations of the variables to be analyzed (see DATA LIST).
- The RECODE command redefines the variable AGE into four categories and recodes all missing values for the variable DRUNK to one missing value (see RECODE).
- The VARIABLE LABELS, VALUE LABELS, and MISSING VALUE commands complete the data definition for AGE and DRUNK (see VARIABLE LABELS, VALUE LABELS, and MISSING VALUE). The value labels for AGE are formatted to print appropriate labels for a column variable. Note that the label "Youngest-Quarter" has no blanks between the words and the label "Oldest Quarter" has two blanks separating the words.
- The CROSSTABS command sets up the table. It requests a table with DRUNK as the row variable and AGE as the column variable.
- Options 3 and 4 request row and column percents.
- The STATISTICS subcommand requests chi-square, lambda, Kendall's tau-*b*, Kendall's tau-*c*, gamma, and Somers' *d*.

The results produced by CROSSTABS are on the following page. The exact appearance of a printed display will depend on the characters available on the printer used.

CROSSTABS display

```
Crosstabulation:     DRUNK      Ever Drink Too Much
                  By AGE        Age in Four Categories

              Count  |Youngest|        |        |Oldest  |
   AGE->      Row Pct|Quarter |        |        |Quarter |  Row
              Col Pct|   1    |   2    |   3    |   4    | Total
DRUNK         -------+--------+--------+--------+--------+
                 1   |   62   |   33   |   36   |   16   |  147
   Yes               |  42.2  |  22.4  |  24.5  |  10.9  |  38.5
                     |  57.9  |  34.7  |  37.9  |  18.8  |
                     +--------+--------+--------+--------+
                 2   |   45   |   62   |   59   |   69   |  235
   No                |  19.1  |  26.4  |  25.1  |  29.4  |  61.5
                     |  42.1  |  65.3  |  62.1  |  81.2  |
                     +--------+--------+--------+--------+
             Column     107      95       95       85       382
              Total     28.0     24.9     24.9     22.3     100.0

 Chi-Square     D.F.     Significance        Min E.F.      Cells with E.F.< 5
 ----------     ----     ------------        --------      ------------------

  31.57228       3          0.0000           32.709           None
                                              With DRUNK              With AGE

         Statistic                  Symmetric         Dependent          Dependent
         ---------                  ---------         ---------          ---------

Lambda                               0.09716           0.11565            0.08727
Uncertainty Coefficient              0.04150           0.06382            0.03075
Somers' D                            0.23546           0.19222            0.30381

         Statistic                    Value          Significance
         ---------                    -----          ------------

Kendall's Tau C                      0.28768           0.0000
Gamma                                0.39632

Number of Missing Observations =      118
```

This example analyzes 1979 prices and earnings in 45 cities around the world, compiled by the Union Bank of Switzerland. The variables are

- NTCPUR—the city's net purchasing-power level, calculated as the ratio of labor expended (measured in number of working hours) to the cost of more than 100 goods and services, weighted by consumer habits. NTCPUR is expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- FOOD—the average net cost of 39 different food and beverage items in the city, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- RENT—the average gross monthly rent in the city, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- APPL—the average cost of six different household appliances, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- SERVICE—the average cost of 28 different goods and services in the city, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- WCLOTHES—the cost of medium-priced women's clothes, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- MCLOTHES—the cost of medium-priced men's clothes, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.
- CLOTHES—the average cost of medium-priced women's and men's clothes, expressed as a percentage above or below that of Zurich, where Zurich equals 100%.

In this example, we obtain univariate summary statistics about purchasing power and the costs of various goods and services in cities. The data are in an external file named ADESC.DAT. The command file specified on the INCLUDE command contains the following SPSS/PC commands:

```
DATA LIST FILE='ADESC.DAT'
 /  NTCPUR 1-3 FOOD 12-14 RENT 23-25 APPL 34-36
    SERVICE 45-47 WCLOTHES 56-58 MCLOTHES 67-69.
VARIABLE LABLES NTCPUR 'Net Purchasing Level'/
                FOOD 'Avg Food Prices'/
                RENT 'Normal Rent'/
                APPL 'Price of Appliances'/
                SERVICE 'Price for Services'/
                WCLOTHES "Medium-Priced Woman's Clothes"/
                MCLOTHES "Medium-Priced Men's Clothes"
COMPUTE CLOTHES=(WCLOTHES + MCLOTHES)/2.
VARIABLE LABELS CLOTHES 'Ave. Cost of W and M Clothes'
DESCRIPTIVES VARIABLES=NTCPUR, FOOD, RENT TO SERVICE, WCLOTHES,
             MCLOTHES, CLOTHES
             /STATISTICS=1,5,9
             /OPTIONS=5.
FINISH.
```

- The DATA LIST command defines the names and locations of the variables to be analyzed (see DATA LIST).
- The VARIABLE LABELS command completes the definition of the variables (see VARIABLE LABELS).
- The COMPUTE command creates the variable CLOTHES by adding the values for WCLOTHES and MCLOTHES and dividing by 2 (see COMPUTE).
- The VARIABLE LABELS command assigns a label to the new variable CLOTHES (see VARIABLE LABELS).
- The DESCRIPTIVES command requests statistics for the variables named.
- The associated STATISTICS subcommand requests the mean, standard deviation, and range for each variable in the analysis.
- Option 5 specifies listwise deletion of missing values. A case missing on any variable specified on the DESCRIPTIVES command is excluded from the computation of statistics for all variables.

- Because the default width of 79 is used, DESCRIPTIVES displays the statistics and variable labels for each variable on one line using only 79 columns.

The display produced by DESCRIPTIVES is shown below. The exact appearance of the printed display will depend on the characters available on the printer used.

DESCRIPTIVES display

```
Number of Valid Observations (Listwise) =        44.00

Variable       Mean     Std Dev     Range   Label

NTCPUR        58.70      28.81     100.00   Net Purchasing Level
FOOD          71.00      18.61      90.00   Avg Food Prices
RENT         121.75      94.65     413.00   Normal Rent
APPL          78.70      22.23     111.00   Price of Appliances
SERVICE       73.68      18.80      71.00   Price for Services
WCLOTHES      81.20      30.36     153.00   Medium-Priced Women's Clothes
MCLOTHES      87.86      25.91     125.00   Medium-Priced Men's Clothes
CLOTHES       84.53      26.75     139.00   Ave. Cost of W and M Clothes
```

FACTOR This example uses six abortion items from a 500-case sample of the 1980 General Social Survey. Respondents indicate whether they favor or oppose abortion in the following contexts:

- ABHLTH—if the woman's health is seriously endangered.
- ABRAPE—if the woman is pregnant as a result of rape.
- ABDEFECT—if there is a strong chance of a serious defect in the child.
- ABPOOR—if the woman has a low income and cannot afford more children.
- ABSINGLE—if the woman is not married and doesn't want the child.
- ABNOMORE—if the woman is married and wants no more children.

The data are in an external file named AFACTOR.DAT. The SPSS/PC commands in the command file named on the INCLUDE command are

```
TITLE  FACTOR ANALYSIS OF ABORTION ITEMS.
DATA LIST FREE FILE='AFACTOR.DAT'
 /ABDEFECT ABNOMORE ABHLTH ABPOOR ABRAPE ABSINGLE.
RECODE  ABDEFECT TO ABSINGLE(1=1)(2=0)(ELSE=9).
MISSING VALUE ABDEFECT TO ABSINGLE (9).
VALUE LABELS  ABDEFECT TO ABSINGLE
  0 'NO' 1 'YES' 9 'MISSING'
FACTOR  VARIABLES=ABDEFECT TO ABSINGLE/
 MISSING=MEANSUB/
 WIDTH=100/
 FORMAT=SORT BLANK(.3)/
 PLOT=ROTATION(1 2)/
 EXTRACTION=ULS/
 ROTATION=OBLIMIN.
FINISH.
```

- The TITLE command puts the title, FACTOR ANALYSIS OF ABORTION ITEMS, at the top of each page of output for this session (see TITLE).
- The DATA LIST command assigns variable names and indicates the data are in freefield format (see DATA LIST).
- The RECODE, MISSING VALUE, and VALUE LABELS commands redefine the abortion items and label the redefined responses for this session (see RECODE, MISSING VALUE, and VALUE LABELS).
- The FACTOR command invokes the FACTOR procedure. The VARIABLES subcommand names all the variables that are used in this FACTOR procedure.
- The MISSING subcommand forces mean substitution for missing data.
- The WIDTH subcommand limits the width of the display to 100 columns.
- The FORMAT subcommand displays the factor loadings in descending order of magnitude and suppresses the printing of factor loadings less than 0.3.
- The PLOT subcommand requests a plot of the variables in factor space, where 1 and 2 are the factor numbers to be plotted.
- The EXTRACTION subcommand specifies unweighted least squares as the method of extraction.
- The ROTATION subcommand specifies an oblimin rotation.

Portions of the output produced by this set of commands appear in Figures A through D. The exact appearance of the printed display will depend on the characters available on the printer used.

- *Initial Statistics.* Figure A contains the initial statistics which are produced by default. Initial statistics are the initial communalities, eigenvalues of the correlation matrix, and percentage of variance explained.
- *Extraction Statistics.* Figure B contains the extraction statistics which are produced by default. Extraction statistics are the communalities, eigenvalues, and unrotated factor loadings. Note the effect of sorting and blanking produced by the FORMAT subcommand.
- *Rotation Statistics.* Figure C contains the rotation statistics which are produced by default if the model is rotated. They are the rotated factor pattern and structure matrices (since this is an oblimin rotation), and the factor correlation matrix.

• *Factor Plot.* Figure D contains a plot of the variables in rotated factor space. Although the rotation is oblimin, the plot axes are orthogonal. Since Factor 1 overlaps with Factor 3 and Factor 2 overlaps with Factor 4, they do not appear on the plot.

A Initial statistics

INITIAL STATISTICS:

VARIABLE	COMMUNALITY	FACTOR	EIGENVALUE	PCT OF VAR	CUM PCT
ABDEFECT	.44988	1	3.38153	56.4	56.4
ABNOMORE	.66747	2	1.19287	19.9	76.2
ABHLTH	.35555	3	.50823	8.5	84.7
ABPOOR	.66600	4	.40867	6.8	91.5
ABRAPE	.39760	5	.28847	4.8	96.3
ABSINGLE	.60394	6	.22024	3.7	100.0

B Extraction statistics

FACTOR MATRIX:

	FACTOR 1	FACTOR 2
ABPOOR	.81970	-.32158
ABNOMORE	.81704	-.34273
ABSINGLE	.78051	
ABDEFECT	.65764	.46888
ABRAPE	.62257	.33014
ABHLTH	.53469	.44445

FINAL STATISTICS:

VARIABLE	COMMUNALITY	FACTOR	EIGENVALUE	PCT OF VAR	CUM PCT
ABDEFECT	.65233	1	3.05462	50.9	50.9
ABNOMORE	.78502	2	.81821	13.6	64.5
ABHLTH	.48344				
ABPOOR	.77532				
ABRAPE	.49658				
ABSINGLE	.68014				

C Rotation statistics

PATTERN MATRIX:

	FACTOR 1	FACTOR 2
ABNOMORE	.90007	
ABPOOR	.88136	
ABSINGLE	.80050	
ABDEFECT		.80734
ABHLTH		.72675
ABRAPE		.63777

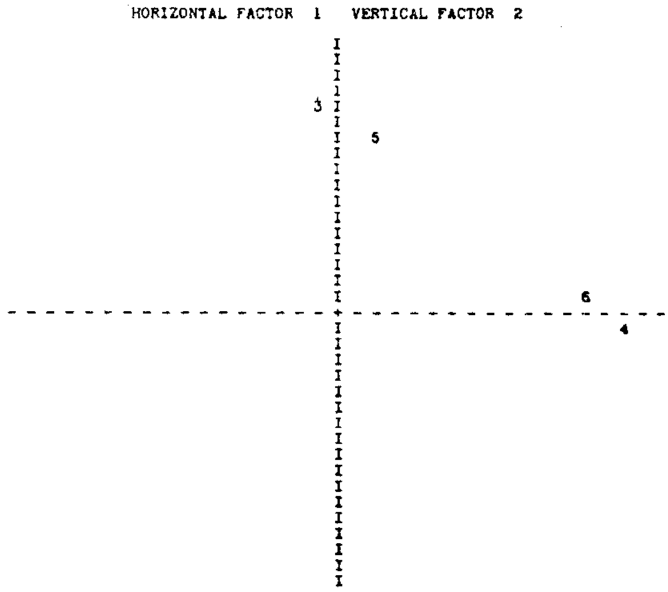
STRUCTURE MATRIX:

	FACTOR 1	FACTOR 2
ABNOMORE	.88574	.46583
ABPOOR	.88052	.48028
ABSINGLE	.82393	.48046
ABDEFECT	.44195	.80767
ABRAPE	.45976	.69852
ABHLTH	.33632	.69342

FACTOR CORRELATION MATRIX:

	FACTOR 1	FACTOR 2
FACTOR 1	1.00000	
FACTOR 2	.54567	1.00000

D Factor plot



SYMBOL	VARIABLE	COORDINATES	SYMBOL	VARIABLE	COORDINATES
1	ABDEFECT	(.00060, .80734)	2	ABNOMORE	(.90007, -.02621)
3	ABHLTH	(-.06097, .72675)	4	ABPOOR	(.88136, -.00153)
5	ABRAPE	(.11111, .63777)	6	ABSINGLE	(.80050, .04285)

FREQUENCIES

The following example demonstrates the use of FREQUENCIES to do some preliminary checks on a newly defined file. The file is based on employment data from Hubbard Consultants Inc. Variables include: date employee was hired, employee's department, salary, job category, name, age, and sex, as well as salary increases from 1980 to 1982. The data are in an external file named AFREQ.DAT. The SPSS/PC commands in the command file named on the INCLUDE command are

```
DATA LIST FILE='AFREQ.DAT'
               /MOHIRED YRHIRED 12-15 DEPT79 TO DEPT82 SEX 16-20
               /SALARY79 TO SALARY82 6-25
                AGE 54-55 RAISE80 TO RAISE82 56-70
               /JOBCAT 6 EMPNAME 25-48 (A).
MISSING VALUE   SALARY79 TO SALARY82 AGE (0)
                JOBCAT (9).
VARIABLE LABELS  SALARY79 'Salary in 1979'
                 SALARY80 'Salary in 1980'
                 SALARY81 'Salary in 1981'
                 SALARY82 'Salary in 1982'
                 JOBCAT 'Job Categories'
VALUE LABELS   SEX 1 'Male' 2 'Female'
               JOBCAT 1 'Officials & Managers' 2 'Professionals'
               3 'Technicians' 4 'Office & Clerical' 5 'Craftsmen'
               6 'Service Workers'.
FREQUENCIES   VARIABLES=SALARY79 TO SALARY82 SEX AGE JOBCAT
              /FORMAT=LIMIT(10)
              /STATISTICS=DEFAULT MEDIAN.
FINISH.
```

- The DATA LIST command assigns variable names and gives column locations for the variables in the analysis (see DATA LIST).
- MISSING VALUE, VARIABLE LABELS, and VALUE LABELS complete the file definition (see MISSING VALUE, VARIABLE LABELS, and VALUE LABELS).
- FREQUENCIES displays frequency tables for variables having 10 or fewer categories and statistics for all the variables named. The default statistics are the mean, standard deviation, minimum, and maximum.
- The output uses the default format of 79 characters.

The FREQUENCIES display produced by this job is on the facing page. The exact appearance of the printed display will depend on the characters available on the printer used.

FREQUENCIES display

SALARY79 Salary in 1979

Mean	12247.323	Median	10140.000	Std Dev	6665.182
Minimum	6337.000	Maximum	45500.000		

Valid Cases 158 Missing Cases 117

- -

SALARY80 Salary in 1980

Mean	12123.725	Median	10400.000	Std Dev	6316.356
Minimum	5720.000	Maximum	48100.000		

Valid Cases 273 Missing Cases 2

- -

SALARY81 Salary in 1981

Mean	15096.212	Median	12359.500	Std Dev	8074.387
Minimum	7605.000	Maximum	52000.000		

Valid Cases 160 Missing Cases 115

- -

SALARY82 Salary in 1982

Mean	17161.552	Median	15132.000	Std Dev	8695.734
Minimum	5830.000	Maximum	50700.000		

Valid Cases 145 Missing Cases 130

- -

SEX

Value Label	Value	Frequency	Percent	Valid Percent	Cum Percent
Male	1	83	30.2	30.2	30.2
Female	2	192	69.8	69.8	100.0
	TOTAL	275	100.0	100.0	

Mean	1.698	Median	2.000	Std Dev	.460
Minimum	1.000	Maximum	2.000		

Valid Cases 275 Missing Cases 0

- -

AGE

Mean	37.158	Median	34.000	Std Dev	11.335
Minimum	20.000	Maximum	69.000		

Valid Cases 272 Missing Cases 3

- -

JOBCAT Job Categories

Value Label	Value	Frequency	Percent	Valid Percent	Cum Percent
Officials & Managers	1	48	17.5	17.5	17.5
Professionals	2	62	22.5	22.5	40.0
Technicians	3	98	35.6	35.6	75.6
Office & Clerical	4	67	24.4	24.4	100.0
	TOTAL	275	100.0	100.0	

Mean	2.669	Median	3.000	Std Dev	1.030
Minimum	1.000	Maximum	4.000		

Valid Cases 275 Missing Cases 0

HILOGLINEAR

Consider a market research analysis of laundry detergent preferences. Consumers may prefer either BRAND M or BRAND X detergent. The study in this example considers three additional variables: water softness, previous use of BRAND M, and washing temperature. The data are in an external file named AHILOG.DAT. The SPSS/PC commands in the command file named on the INCLUDE command are

```
TITLE DETERGENT PREFERENCES RIES & SMITH(1963).
DATA LIST FREE FILE='AHILOG.DAT'
 / WATSOFT BRANDPRF PREVUSE TEMP FREQ.
VARIABLE LABELS WATSOFT  'WATER SOFTNESS'
                BRANDPRF 'BRAND PREFERENCE'
                PREVUSE  'PREVIOUS USE OF M'
                TEMP     'WATER TEMPERATURE'
                FREQ     'NUMBER IN CONDITION'.
VALUE LABELS WATSOFT 1 'SOFT' 2 'MEDIUM' 3 'HARD'
             BRANDPRF 1 'BRAND X' 2 'BRAND M' /
             PREVUSE 1 'YES' 2 'NO' /
             TEMP 1 'HIGH' 2 'LOW'.
WEIGHT BY FREQ.
HILOGLINEAR WATSOFT (1,3) BRANDPRF PREVUSE TEMP (1,2)
            PRINT = ALL /
            PLOT = DEFAULT /
            METHOD = BACKWARD /
            CRITERIA = MAXSTEPS(24)
            DESIGN/DESIGN.
FINISH.
```

- DATA LIST defines four variables and indicates that the data are in freefield format (see DATA LIST).
- The VARIABLE LABELS and VALUE LABELS commands complete the file definition (see VARIABLE LABELS and VALUE LABELS).
- The WEIGHT command weights the observations by FREQ, the variable containing the number of observations for each combination of values (see WEIGHT).
- HILOGLINEAR specifies four variables. The variable WATSOFT has three levels and the other three variables each have two.
- The PRINT subcommand requests all available displays: observed, expected, and residual values (display not shown); the result of the iterative proportional fitting algorithm and tests for order of the saturated model and for effects of each order (Figure A); measures of partial association for effects (Figure B); and parameter estimates (Figure C).
- The PLOT subcommand requests the default plots of residuals against observed and expected values and a normal probability plot (partial display in Figure F).
- The METHOD subcommand requests backward elimination. The CRITERIA subcommand requests a maximum of 24 steps. And the DESIGN subcommand successively eliminates terms from the default saturated model. (Partial display in Figure D.) The observed and expected frequencies for the final model are also displayed (Figure E).

The exact appearance of the printed display will depend on the characters available on the printer used.

A Tests for order of saturated model and for effects of each order

Tests that K-way and higher order effects are zero.

K	DF	L.R. Chisq	Prob	Pearson Chisq	Prob	Iteration
4	2	.738	.6915	.738	.6915	3
3	9	9.846	.3631	9.871	.3611	3
2	18	42.926	.0008	43.902	.0006	2
1	23	118.626	.0000	115.714	.0000	0

Tests that K-way effects are zero.

K	DF	L.R. Chisq	Prob	Pearson Chisq	Prob	Iteration
1	5	75.701	.0000	71.812	.0000	0
2	9	33.080	.0001	34.031	.0001	0
3	7	9.108	.2450	9.133	.2433	0
4	2	.738	.6915	.738	.6915	0

B Partial associations

Tests of PARTIAL associations.

Effect Name	DF	Partial Chisq	Prob	Iter
WATSOFT*BRANDPRF*PREVUSE	2	4.571	.1017	3
WATSOFT*BRANDPRF*TEMP	2	.162	.9223	3
WATSOFT*PREVUSE*TEMP	2	1.377	.5022	3
BRANDPRF*PREVUSE*TEMP	1	2.222	.1361	3
WATSOFT*BRANDPRF	2	.216	.8977	3
WATSOFT*PREVUSE	2	1.005	.6051	3
BRANDPRF*PREVUSE	1	19.892	.0000	3
WATSOFT*TEMP	2	6.095	.0475	3
BRANDPRF*TEMP	1	3.738	.0532	3
PREVUSE*TEMP	1	.740	.3898	3
WATSOFT	2	.502	.7780	2
BRANDPRF	1	.064	.7996	2
PREVUSE	1	1.922	.1656	2
TEMP	1	73.211	.0000	2

C Partial display of parameter estimates for saturated model

Estimates for Parameters.

WATSOFT*BRANDPRF*PREVUSE*TEMP

Parameter	Coeff	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	-.0086293293	.04833	-.17856	-.10335	.08609
2	-.0296475092	.04734	-.62629	-.12243	.06313

WATSOFT*BRANDPRF*PREVUSE

Parameter	Coeff.	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	.0925171313	.04833	1.91437	-.00221	.18724
2	-.0318024179	.04734	-.67182	-.12458	.06098

WATSOFT*BRANDPRF*TEMP

Parameter	Coeff.	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	-.0203612840	.04833	-.42132	-.11508	.07436
2	.0048119005	.04734	.10165	-.08797	.09759

WATSOFT*PREVUSE*TEMP

Parameter	Coeff.	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	-.0474552194	.04833	-.98194	-.14218	.04727
2	.0488797573	.04734	1.03257	-.04390	.14166

BRANDPRF*PREVUSE*TEMP

Parameter	Coeff.	Std. Err.	Z-Value	Lower 95 CI	Upper 95 CI
1	-.0504586672	.03363	-1.50056	-.11637	.01545

D Partial display of final statistics

```
Backward Elimination for DESIGN 1 with generating class

  WATSOFT*BRANDPRF*PREVUSE*TEMP

 Likelihood ratio chi square =     0.0        DF = 0  P = 1.000


Step 8

  The best model has generating class

      WATSOFT*TEMP
      BRANDPRF*TEMP
      BRANDPRF*PREVUSE

  Likelihood ratio chi square =    11.88633    DF = 14  P =  .615

If Deleted Simple Effect is              DF   L.R. Chisq Change    Prob  Iter

 WATSOFT*TEMP                             2               6.098   .0474     2
 BRANDPRF*TEMP                            1               4.361   .0368     2
 BRANDPRF*PREVUSE                         1              20.578   .0000     2

Step 9

  The best model has generating class

      WATSOFT*TEMP
      BRANDPRF*TEMP
      BRANDPRF*PREVUSE

  Likelihood ratio chi square =    11.88633    DF = 14  P =  .615

The final model has generating class

    WATSOFT*TEMP
    BRANDPRF*TEMP
    BRANDPRF*PREVUSE

The Iterative Proportional Fitting converged at iteration 0
```

E Observed and expected frequencies for selected model

Observed, Expected Frequencies and Residuals.

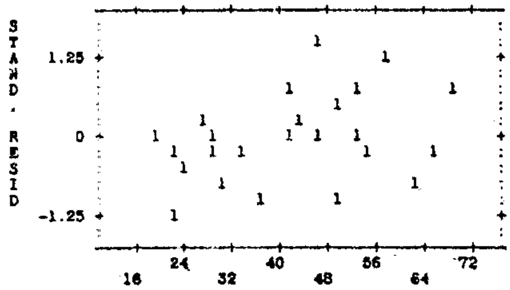
Factor	Code	OBS count	EXP count	Residual	Std Resid
WATSOFT	SOFT				
BRANDPRF	BRAND X				
PREVUSE	YES				
TEMP	HIGH	19.0	19.5	-.52	-.12
TEMP	LOW	57.0	47.8	9.15	1.32
PREVUSE	NO				
TEMP	HIGH	29.0	28.4	.61	.11
TEMP	LOW	63.0	69.6	-6.58	-.79
BRANDPRF	BRAND M				
PREVUSE	YES				
TEMP	HIGH	29.0	30.8	-1.85	-.33
TEMP	LOW	49.0	57.5	-8.52	-1.12
PREVUSE	NO				
TEMP	HIGH	27.0	25.2	1.76	.35
TEMP	LOW	53.0	47.1	5.94	.87
WATSOFT	MEDIUM				
BRANDPRF	BRAND X				
PREVUSE	YES				
TEMP	HIGH	23.0	23.7	-.65	-.13
TEMP	LOW	47.0	47.0	.01	.00
PREVUSE	NO				
TEMP	HIGH	33.0	34.4	-1.40	-.24
TEMP	LOW	66.0	68.3	-2.32	-.28
BRANDPRF	BRAND M				
PREVUSE	YES				
TEMP	HIGH	47.0	37.4	9.63	1.57
TEMP	LOW	55.0	56.5	-1.48	-.20
PREVUSE	NO				
TEMP	HIGH	23.0	30.6	-7.68	-1.37
TEMP	LOW	50.0	46.2	3.79	.56
WATSOFT	HARD				
BRANDPRF	BRAND X				
PREVUSE	YES				
TEMP	HIGH	24.0	26.1	-2.09	-.41
TEMP	LOW	37.0	42.9	-5.89	-.90
PREVUSE	NO				
TEMP	HIGH	42.0	37.9	4.06	.66
TEMP	LOW	68.0	62.4	5.63	.71
BRANDPRF	BRAND M				
PREVUSE	YES				
TEMP	HIGH	43.0	41.2	1.77	.28
TEMP	LOW	52.0	51.6	.44	.06
PREVUSE	NO				
TEMP	HIGH	30.0	33.7	-3.73	-.64
TEMP	LOW	42.0	42.2	-.18	-.03

Goodness-of-fit test statistics

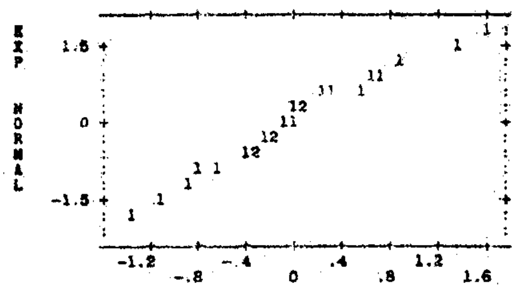
Likelihood ratio chi square =	11.88633	DF = 14	P = .615
Pearson chi square =	11.91780	DF = 14	P = .613

F HILOGLINEAR residuals plots

Observed counts VS Standardized residuals



Normal Plot



LIST

This example demonstrates the use of LIST to display values of selected cases in a newly defined file. The file is based on employment data from Hubbard Consultants Inc. Variables include: date employee was hired, employee's department, salary, job category, name, age, and sex, as well as salary increases from 1979 to 1982. The data are in an external file named ALIST.DAT. The SPSS/PC commands in the command file named on the INCLUDE command are

```
SET WIDTH=WIDE.
DATA LIST FILE='ALIST.DAT'
   / EMPLOYID 1-5 MOHIRED YRHIRED 12-15 DEPT79 TO DEPT82 SEX 16-20
   / SALARY79 TO SALARY82 6-25 HOURLY81 HOURLY82 40-53(2) PROMO81 72
     AGE 54-55 RAISE82 66-70
   / JOBCAT 6 NAME 25-48 (A).
LIST VARIABLES=MOHIRED YRHIRED DEPT82 SALARY79 TO SALARY82 NAME/
  CASES FROM 50 TO 100 BY 5  FORMAT=SINGLE,NUMBERED.
FINISH.
```

- The SET command requests a width of 132 (WIDE) so the column titles will be horizontal instead of vertical (see SET command).
- The DATA LIST command assigns variable names and gives column locations for the variables in the analysis. The variable NAME is defined as a string variable (see DATA LIST).
- LIST displays values for the variables MOHIRED, YRHIRED, DEPT82, SALARY79 to SALARY 82, and NAME. Beginning with the 50th case, every 5th case will be listed. The display is forced to have only line per case, and the sequence number for each case will be displayed.

The LIST display produced by this job is shown below. The exact appearance of the printed display will depend on the characters available on the printer used.

LIST display

Cas	MOHIRED	YRHIRED	DEPT82	SALARY79	SALARY80	SALARY81	SALARY82	NAME
50	1	79	0	14300	14300	15730	0	EVA ELDER
55	6	79	0	0	15600	0	0	EDWARD GREEN
60	12	79	0	0	8840	9503	0	LOVEY E. HUDSON
65	5	80	0	0	13520	0	0	PATRICIA SMITH
70	8	79	0	0	8255	0	0	HELEN D. SMITH
75	10	70	4	14300	18850	21450	26182	MONICA C. RIVERS
80	1	79	0	0	7442	0	0	THOMAS P. JOHNSON
85	4	80	3	0	18200	18395	19682	ANN JOHNSON
90	10	79	0	0	5720	0	0	CHRISTINA P. NORRIS
95	5	79	0	7670	9490	0	0	M. ELLIOT KRAFT
100	2	70	3	11830	12545	13799	18083	FANNIE SMITH

Number of cases read = 100 Number of cases listed = 11