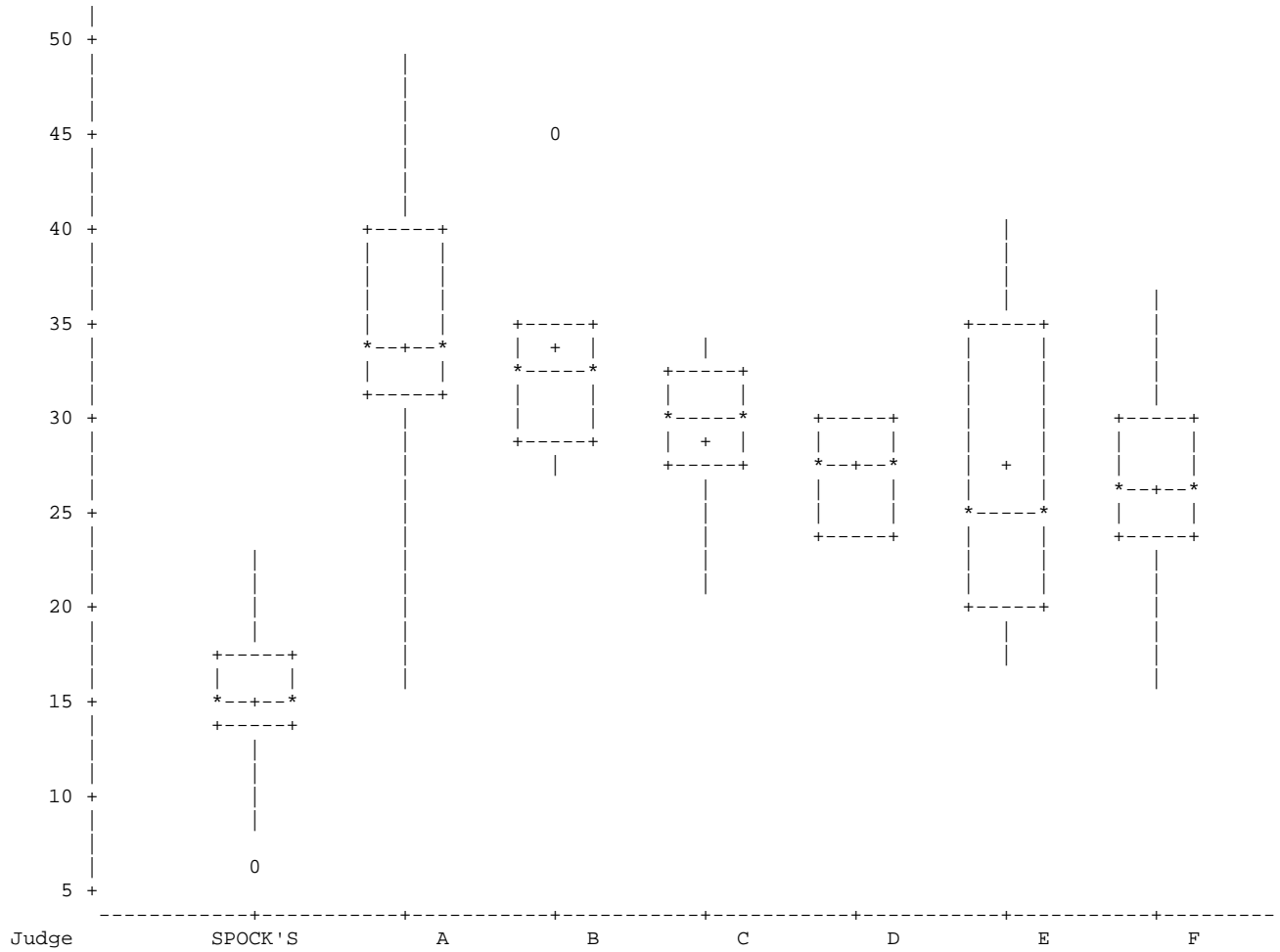


```
1      /*
2      Fit an analysis of variance to determine if the jury for
3      spock trial was unusually low in the number of women.
4      */
5
6      /*
7      SAS applications of note:
8          By statement with NOTSORTED
9          use of if statement to create new variable
10     DATA and SET statement to modify data set
11     */
12
13
14     dm'log;clear;output;clear';
15     options nodate nocenter nonumber ps=512 ls=132;
16     ODS HTML style=minimal rs=none
body='C:\geaghan\Current\EXST3201\Fall2005\SAS\SpockTrial01.html' ;
NOTE: Writing HTML Body file:
C:\geaghan\Current\EXST3201\Fall2005\SAS\SpockTrial01.html
17
18     Title1 'Chapter 5 : Spock Conspiracy Trial';
19     filename input
'C:\Geaghan\Current\EXST3201\Datasets\ASCII\CASE0502.csv';
20
21     data Jury; infile input missover DSD dlm="," firstobs=2;
22         input Percent Judge $;
23         datalines;
NOTE: The infile INPUT is:
File Name=C:\Geaghan\Current\EXST3201\Datasets\ASCII\CASE0502.csv,
RECFM=V,LRECL=256
NOTE: 46 records were read from the infile INPUT.
The minimum record length was 6.
The maximum record length was 28.
NOTE: The data set WORK.JURY has 46 observations and 2 variables.
NOTE: DATA statement used (Total process time):
real time          0.03 seconds
cpu time           0.03 seconds
25     options ps=52 ls=111;
26     Title2 'Proc univariate by judge';
27     proc univariate data=jury plot normal; by judge notsorted;
28         var percent;
29     run;
NOTE: The PROCEDURE UNIVARIATE printed pages 1-15.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
real time          0.40 seconds
cpu time           0.28 seconds
30
31     options ps=512 ls=132;
32     proc sort data=Jury; by judge percent; run;
NOTE: There were 46 observations read from the data set WORK.JURY.
NOTE: The data set WORK.JURY has 46 observations and 2 variables.
NOTE: PROCEDURE SORT used (Total process time):
real time          0.02 seconds
cpu time           0.03 seconds
33
34     Title2 'Raw data listing';
35     proc print data=Jury; run;
NOTE: There were 46 observations read from the data set WORK.JURY.
NOTE: The PROCEDURE PRINT printed page 16.
NOTE: PROCEDURE PRINT used (Total process time):
real time          0.07 seconds
cpu time           0.03 seconds
```

Chapter 5 : Spock Conspiracy Trial
Proc univariate by judge

The UNIVARIATE Procedure
Variable: Percent

Schematic Plots



Chapter 5 : Spock Conspiracy Trial
Raw data listing

| Obs | Percent | Judge | Obs | Percent | Judge | Obs | Percent | Judge |
|-----|---------|-------|-----|---------|-------|-----|---------|---------|
| 1 | 16.8000 | A | 16 | 30.5000 | C | 32 | 26.4000 | F |
| 2 | 30.8000 | A | 17 | 31.9000 | C | 33 | 26.7000 | F |
| 3 | 33.6000 | A | 18 | 32.5000 | C | 34 | 29.5000 | F |
| 4 | 40.5000 | A | 19 | 33.8000 | C | 35 | 29.8000 | F |
| 5 | 48.9000 | A | 20 | 33.8000 | C | 36 | 31.9000 | F |
| 6 | 27.0000 | B | 21 | 24.3000 | D | 37 | 36.2000 | F |
| 7 | 28.9000 | B | 22 | 29.7000 | D | 38 | 6.4000 | SPOCK'S |
| 8 | 32.0000 | B | 23 | 17.7000 | E | 39 | 8.7000 | SPOCK'S |
| 9 | 32.7000 | B | 24 | 19.7000 | E | 40 | 13.3000 | SPOCK'S |
| 10 | 35.5000 | B | 25 | 21.5000 | E | 41 | 13.6000 | SPOCK'S |
| 11 | 45.6000 | B | 26 | 27.9000 | E | 42 | 15.0000 | SPOCK'S |
| 12 | 21.0000 | C | 27 | 34.8000 | E | 43 | 15.2000 | SPOCK'S |
| 13 | 23.4000 | C | 28 | 40.2000 | E | 44 | 17.7000 | SPOCK'S |
| 14 | 27.5000 | C | 29 | 16.5000 | F | 45 | 18.6000 | SPOCK'S |
| 15 | 27.5000 | C | 30 | 20.7000 | F | 46 | 23.1000 | SPOCK'S |
| | | | 31 | 23.5000 | F | | | |

```

37      Title2 'Proc univariate for all data combined';
38      proc univariate data=jury plot normal;
39          var percent;
40      run;
NOTE: The PROCEDURE UNIVARIATE printed page 17.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time          0.07 seconds
      cpu time           0.04 seconds

```

Chapter 5 : Spock Conspiracy Trial
Proc univariate for all data combined

The UNIVARIATE Procedure
Variable: Percent

Moments

| | | | |
|-----------------|------------|------------------|------------|
| N | 46 | Sum Weights | 46 |
| Mean | 26.5826087 | Sum Observations | 1222.8 |
| Std Deviation | 9.17911401 | Variance | 84.2561339 |
| Skewness | 0.03841528 | Kurtosis | -0.03463 |
| Uncorrected SS | 36296.7399 | Corrected SS | 3791.52603 |
| Coeff Variation | 34.5305238 | Std Error Mean | 1.35338652 |

Basic Statistical Measures

| Location | | Variability | |
|----------|----------|---------------------|----------|
| Mean | 26.58261 | Std Deviation | 9.17911 |
| Median | 27.50000 | Variance | 84.25613 |
| Mode | 17.70000 | Range | 42.50000 |
| | | Interquartile Range | 12.80000 |

NOTE: The mode displayed is the smallest of 4 modes with a count of 2.

Tests for Location: Mu0=0

| Test | -Statistic- | -----p Value----- | |
|-------------|-------------|-------------------|--------|
| Student's t | t 19.64155 | Pr > t | <.0001 |
| Sign | M 23 | Pr >= M | <.0001 |
| Signed Rank | S 540.5 | Pr >= S | <.0001 |

Tests for Normality

| Test | --Statistic-- | -----p Value----- | |
|--------------------|---------------|-------------------|---------|
| Shapiro-Wilk | W 0.987632 | Pr < W | 0.9013 |
| Kolmogorov-Smirnov | D 0.07902 | Pr > D | >0.1500 |
| Cramer-von Mises | W-Sq 0.042645 | Pr > W-Sq | >0.2500 |
| Anderson-Darling | A-Sq 0.254957 | Pr > A-Sq | >0.2500 |

Quantiles (Definition 5)

| Quantile | Estimate |
|------------|----------|
| 100% Max | 48.9 |
| 99% | 48.9 |
| 95% | 40.5 |
| 90% | 36.2 |
| 75% Q3 | 32.5 |
| 50% Median | 27.5 |
| 25% Q1 | 19.7 |
| 10% | 15.0 |
| 5% | 13.3 |
| 1% | 6.4 |
| 0% Min | 6.4 |

Chapter 5 : Spock Conspiracy Trial
 Analysis among judges
 Analysis of variance with PROC GLM

The GLM Procedure

| Class Level Information | | | | | | |
|-----------------------------|-----------|----------------|--------------|---------|--------|-------------|
| Class | Levels | Values | | | | |
| Judge | 7 | A | B | C | D | E F SPOCK'S |
| Number of Observations Read | | | | | | 46 |
| Number of Observations Used | | | | | | 46 |
| Dependent Variable: Percent | | | | | | |
| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F | |
| Model | 6 | 1927.080772 | 321.180129 | 6.72 | <.0001 | |
| Error | 39 | 1864.445255 | 47.806289 | | | |
| Corrected Total | 45 | 3791.526027 | | | | |
| R-Square | Coeff Var | Root MSE | Percent Mean | | | |
| 0.508260 | 26.01027 | 6.914209 | 26.58261 | | | |
| Source | DF | Type I SS | Mean Square | F Value | Pr > F | |
| Judge | 6 | 1927.080772 | 321.180129 | 6.72 | <.0001 | |
| Source | DF | Type III SS | Mean Square | F Value | Pr > F | |
| Judge | 6 | 1927.080772 | 321.180129 | 6.72 | <.0001 | |

Chapter 5 : Spock Conspiracy Trial
 Analysis among judges
 Analysis of variance with PROC GLM
 Listing of PROC GLM residuals

| Obs | Percent | Judge | e | yhat |
|-----|---------|-------|----------|---------|
| 1 | 16.8000 | A | -17.3200 | 34.1200 |
| 2 | 30.8000 | A | -3.3200 | 34.1200 |
| 3 | 33.6000 | A | -0.5200 | 34.1200 |
| 4 | 40.5000 | A | 6.3800 | 34.1200 |
| 5 | 48.9000 | A | 14.7800 | 34.1200 |
| 6 | 27.0000 | B | -6.6167 | 33.6167 |
| 7 | 28.9000 | B | -4.7167 | 33.6167 |
| 8 | 32.0000 | B | -1.6167 | 33.6167 |
| 9 | 32.7000 | B | -0.9167 | 33.6167 |
| 10 | 35.5000 | B | 1.8833 | 33.6167 |
| 11 | 45.6000 | B | 11.9833 | 33.6167 |
| 12 | 21.0000 | C | -8.1000 | 29.1000 |
| 13 | 23.4000 | C | -5.7000 | 29.1000 |
| 14 | 27.5000 | C | -1.6000 | 29.1000 |
| 15 | 27.5000 | C | -1.6000 | 29.1000 |
| 16 | 30.5000 | C | 1.4000 | 29.1000 |
| 17 | 31.9000 | C | 2.8000 | 29.1000 |
| 18 | 32.5000 | C | 3.4000 | 29.1000 |
| 19 | 33.8000 | C | 4.7000 | 29.1000 |
| 20 | 33.8000 | C | 4.7000 | 29.1000 |
| 21 | 24.3000 | D | -2.7000 | 27.0000 |
| 22 | 29.7000 | D | 2.7000 | 27.0000 |
| 23 | 17.7000 | E | -9.2667 | 26.9667 |
| 24 | 19.7000 | E | -7.2667 | 26.9667 |

| | | | | |
|----|---------|---------|----------|---------|
| 25 | 21.5000 | E | -5.4667 | 26.9667 |
| 26 | 27.9000 | E | 0.9333 | 26.9667 |
| 27 | 34.8000 | E | 7.8333 | 26.9667 |
| 28 | 40.2000 | E | 13.2333 | 26.9667 |
| 29 | 16.5000 | F | -10.3000 | 26.8000 |
| 30 | 20.7000 | F | -6.1000 | 26.8000 |
| 31 | 23.5000 | F | -3.3000 | 26.8000 |
| 32 | 26.4000 | F | -0.4000 | 26.8000 |
| 33 | 26.7000 | F | -0.1000 | 26.8000 |
| 34 | 29.5000 | F | 2.7000 | 26.8000 |
| 35 | 29.8000 | F | 3.0000 | 26.8000 |
| 36 | 31.9000 | F | 5.1000 | 26.8000 |
| 37 | 36.2000 | F | 9.4000 | 26.8000 |
| 38 | 6.4000 | SPOCK'S | -8.2222 | 14.6222 |
| 39 | 8.7000 | SPOCK'S | -5.9222 | 14.6222 |
| 40 | 13.3000 | SPOCK'S | -1.3222 | 14.6222 |
| 41 | 13.6000 | SPOCK'S | -1.0222 | 14.6222 |
| 42 | 15.0000 | SPOCK'S | 0.3778 | 14.6222 |
| 43 | 15.2000 | SPOCK'S | 0.5778 | 14.6222 |
| 44 | 17.7000 | SPOCK'S | 3.0778 | 14.6222 |
| 45 | 18.6000 | SPOCK'S | 3.9778 | 14.6222 |
| 46 | 23.1000 | SPOCK'S | 8.4778 | 14.6222 |

```

54      Title3 'Analysis of variance with PROC MIXED';
55      proc mixed data=Jury;
56          class judge;
57          model percent = judge / outp=resids;
58      run;

```

NOTE: The data set WORK.RESIDS has 46 observations and 9 variables.

NOTE: The PROCEDURE MIXED printed page 21.

NOTE: PROCEDURE MIXED used (Total process time):

```

real time      0.20 seconds
cpu time       0.15 seconds

```

```

59
60      Title4 'Proc univariate check of the residuals from ANOVA';
61      proc univariate data=resids plot normal;
62          var resid;
63      run;

```

NOTE: The PROCEDURE UNIVARIATE printed page 22.

NOTE: PROCEDURE UNIVARIATE used (Total process time):

```

real time      0.09 seconds
cpu time       0.04 seconds

```

```

64
65      Title4 'Scatterplots of PROC MIXED residuals';
66      options ps=52 ls=99;
67      proc plot data=resids;
68          plot resid * judge / vref=0;
69          plot resid * pred / vref=0;
70      run;
71      options ps=256 ls=99;

```

Chapter 5 : Spock Conspiracy Trial
 Analysis among judges
 Analysis of variance with PROC MIXED

The Mixed Procedure

Model Information

| | |
|---------------------------|-------------|
| Data Set | WORK.JURY |
| Dependent Variable | Percent |
| Covariance Structure | Diagonal |
| Estimation Method | REML |
| Residual Variance Method | Profile |
| Fixed Effects SE Method | Model-Based |
| Degrees of Freedom Method | Residual |

Class Level Information

| | | |
|-------|--------|---------------------|
| Class | Levels | Values |
| Judge | 7 | A B C D E F SPOCK'S |

Dimensions

| | |
|-----------------------|----|
| Covariance Parameters | 1 |
| Columns in X | 8 |
| Columns in Z | 0 |
| Subjects | 1 |
| Max Obs Per Subject | 46 |

Number of Observations

| | |
|---------------------------------|----|
| Number of Observations Read | 46 |
| Number of Observations Used | 46 |
| Number of Observations Not Used | 0 |

Covariance Parameter Estimates

| | |
|----------|----------|
| Cov Parm | Estimate |
| Residual | 47.8063 |

FitStatistics

| | |
|--------------------------|-------|
| -2 Res Log Likelihood | 274.0 |
| AIC (smaller is better) | 276.0 |
| AICC (smaller is better) | 276.1 |
| BIC (smaller is better) | 277.6 |

Type 3 Tests of Fixed Effects

| | | | | |
|--------|-----|-----|---------|--------|
| | Num | Den | | |
| Effect | DF | DF | F Value | Pr > F |
| Judge | 6 | 39 | 6.72 | <.0001 |

Chapter 5 : Spock Conspiracy Trial
 Analysis among judges
 Analysis of variance with PROC MIXED
 Proc univariate check of the residuals from ANOVA

The UNIVARIATE Procedure
 Variable: Resid (Residual)

| Moments | | | |
|-----------------|------------|------------------|------------|
| N | 46 | Sum Weights | 46 |
| Mean | 0 | Sum Observations | 0 |
| Std Deviation | 6.43677845 | Variance | 41.4321168 |
| Skewness | 0.01530119 | Kurtosis | 0.44839315 |
| Uncorrected SS | 1864.44525 | Corrected SS | 1864.44525 |
| Coeff Variation | . | Std Error Mean | 0.94905121 |

| Basic Statistical Measures | | | |
|----------------------------|----------|---------------------|----------|
| Location | | Variability | |
| Mean | 0.00000 | Std Deviation | 6.43678 |
| Median | -0.25000 | Variance | 41.43212 |
| Mode | -1.60000 | Range | 32.10000 |
| | | Interquartile Range | 8.11667 |

NOTE: The mode displayed is the smallest of 2 modes with a count of 2.

| Tests for Location: Mu0=0 | | | | |
|---------------------------|-------------|----|-------------------|--------|
| Test | -Statistic- | | -----p Value----- | |
| Student's t | t | 0 | Pr > t | 1.0000 |
| Sign | M | -1 | Pr >= M | 0.8830 |
| Signed Rank | S | -7 | Pr >= S | 0.9400 |

| Tests for Normality | | | | |
|---------------------|------|----------------|-----------|-------------------|
| Test | | --Statistic--- | | -----p Value----- |
| Shapiro-Wilk | W | 0.988172 | Pr < W | 0.9167 |
| Kolmogorov-Smirnov | D | 0.074758 | Pr > D | >0.1500 |
| Cramer-von Mises | W-Sq | 0.029352 | Pr > W-Sq | >0.2500 |
| Anderson-Darling | A-Sq | 0.207188 | Pr > A-Sq | >0.2500 |

| Quantiles (Definition 5) | |
|--------------------------|-----------|
| Quantile | Estimate |
| 100% Max | 14.78000 |
| 99% | 14.78000 |
| 95% | 11.98333 |
| 90% | 8.47778 |
| 75% Q3 | 3.40000 |
| 50% Median | -0.25000 |
| 25% Q1 | -4.71667 |
| 10% | -8.10000 |
| 5% | -9.26667 |
| 1% | -17.32000 |
| 0% Min | -17.32000 |

Chapter 5 : Spock Conspiracy Trial
 Analysis among judges
 Analysis of variance with PROC MIXED
 Scatterplots of PROC MIXED residuals

Plot of Resid*Pred. Legend: A = 1 obs, B = 2 obs, etc.

