

# Table of Contents

# Pages	Topic
4	Overview of Stats & SPSS Procedures
DATA MANAGEMENT	
2	Labeling & Missing Values
3	Compute
2	Recode
10	Data Dictionary
DESCRIPTIVE STATISTICS	
1	Descriptives
4	Frequencies
4	Examine
2	Means
GRAPHING	
2	Bar Graph
2	Line Graph
2	Error Bar Graph
1	Scatterplot
INFERENCEAL STATISTICS	
2	One-Sample T-Test
4	Independent Groups T-Test
2	Repeated Measures T-Test
3	One-Way ANOVA
4	Repeated Measures ANOVA
3	Two-Way ANOVA
4	Mixed Model ANOVA
2	Correlation
2	Regression
4	Multiple Regression
2	Chi Square
MISCELLANEOUS	
2	Selecting Cases
2	Reliability

# Overview of Stats and SPSS Procedures

Updated 4/17/02

Below is a brief description of each of the documents that you can view online, download, or print for help with your statistical analyses. This is not intended to be a statistics text, it is meant to help you implement procedures on SPSS and find the pieces of the output that you will need for reporting and interpretation.

If you have comments or suggestions, please drop me a note (rusciojp@etown.edu). For example, I would be happy to include additional procedures if you come across situations not covered here.

## DATA MANAGEMENT

### Labeling & Missing Values

These commands allow you to label your variables and, for categorical variables, their values. These labels are then used in most output so that you can easily interpret it without remembering cryptic notation and codes that you used. The missing values command allows you to specify when data are missing so that SPSS knows not to include those cases in an analysis.

### Compute

This command allows you to create a new variable using any existing variables and virtually any type of computation. This is useful, for example, when you want to average (or sum) several questionnaire responses to form a scale score.

### Recode

This command allows you to transform or reclassify scores on a variable. Under certain circumstances, for example, you might want to dichotomize a continuous variable into “absent” vs. “present” form. Or, you might need to reverse-score some negatively-worded items on a questionnaire before combining them with others.

### Data Dictionary

This command shows you information on each variable in your data set, including variable and value labels, missing values, and so forth.

## DESCRIPTIVE STATISTICS

### Descriptives

This procedure provides the basics: number of cases, minimum value, maximum value, mean, and standard deviation. It will provide this information in a compact table for any number of variables that you specify.

### Frequencies

This is perhaps the most versatile descriptive statistics procedure. You can get any stats you want by naming them (or using the term “all” to get everything) on the /stats subcommand. You will get a table listing all the values for a variable and the number and percentage of cases at each value. You can obtain a histogram showing you the distribution of scores, with a superimposed normal curve to help you judge normality if you like.

### Examine

This procedure will, by default, give you a lot of information, including a comprehensive table of descriptive statistics, a stem & leaf plot, and a boxplot for each variable.

### Means

This procedure will provide you with the means and standard deviations of variables, broken down across the levels of one or more categorical variables if you choose. This is particularly handy when you are analyzing an ANOVA design for which SPSS does not give you means or standard deviations. Also, you can easily hand calculate the standard error (standard deviation divided by square root of sample size) of each mean for constructing a graph.

## GRAPHING

### Bar Graph

This procedure will plot a bar graph, useful for either a one-way or two-way design.

### Line Graph

This procedure will plot a line graph, also useful for either a one-way or two-way design.

### Error Bar Graph

This procedure will plot an error bar graph, which includes the mean value as a point surrounded by vertical bars representing the standard error. This is helpful in judging whether means differ from one another at a glance.

### Scatterplot

This procedure will generate a scatterplot between two variables, including the best-fit regression line to aid in your interpretation. You can readily see the direction of a relationship, and you can judge its strength by noting how closely the points cluster around the line. You can also spot factors that influence the correlation coefficient, such as restriction of range, outliers, nonlinearity, or heteroscedasticity.

## INFERENCEAL STATISTICS

### One Sample T-Test

This procedure will test whether the sample mean on one variable differs reliably from a population mean that you specify.

### Independent Groups T-Test

This procedure will test whether the means of two groups of participants differ reliably from one another. This is used for a between-subjects design with two groups.

### Repeated Measures T-Test

This procedure will test whether the means of two conditions differ reliably from one another. This is used for a within-subjects design with two conditions.

### One-Way ANOVA

This procedure will test whether the means of more than two groups of participants differ reliably from one another. It will also conduct post-hoc tests (e.g., Tukey's HSD). This is used for a between-subjects design with more than two groups.

### Repeated Measures ANOVA

This procedure will test whether the means of more than two conditions differ reliably from one another. It will not conduct post-hoc tests, but it will give you output with which you can easily do this by hand. This is used for a within-subjects design with more than two conditions.

### Two-Way ANOVA

This procedure will test for main effects and an interaction. It will conduct post-hoc tests on any factor with more than two levels. This is used for a factorial design two between-subjects factors.

### Mixed Model ANOVA

This procedure will perform ANOVAs for mixed designs. A mixed design is one that contains at least one between-subjects factor and at least one within-subjects factor. The procedure tests all main effects and interactions in the design.

### Correlation

This procedure will calculate correlation coefficients between variables. The output includes a test of whether the correlation is reliably different from 0. Note that you should usually look at a scatterplot to help interpret a correlation coefficient.

### Regression

This procedure will determine the best-fit regression line to predict one variable from another. The output includes the correlation coefficient, the coefficient of determination, and the coefficients for the regression equation itself.

### Multiple Regression

This procedure allows you to use multiple predictor variables (binary/dichotomous or continuous) to predict a continuous dependent variable. This is often used to do the equivalent of a factorial ANOVA when one or more of the "factors" is not categorical, but continuous.

### Chi Square

This procedure will test whether two categorical variables are associated with one another beyond chance levels. The output includes a table of observed frequencies, the chi square statistic, plus row, column, and total percentages to aid in interpreting the results.

## MISCELLANEOUS

### Selecting Cases

This command will allow you to conduct analyses on a subsample. For example, if you want to determine how well two variables are correlated for women, and then for men, you can temporarily select only women and then only men for the correlational analyses.

### Reliability

This procedure will determine the internal consistency (Cronbach's  $\alpha$ ) of a measure. It will also provide item analysis information, such as the correlation of each item with the remainder of the items ("corrected item-total correlations") and the value of  $\alpha$  if an item is deleted.