

# Assignment 1

Due: Friday, February 15, 2008

## Dataset

The data for this assignment is provided for you in “assignment1.sav”.

The (fictional) data was taken from a study comparing three visualizations: TreeMap, Phylotrees, and ArcTrees. Forty-four Participants (23 Female, 21 Male) were asked to find the root node in one of the three visualizations as quickly as possible. The total time to find the node was recorded for each participant.

## SPSS Hints

- Both the T-Test and the one-way ANOVA are run through the “Analyze → Compare Means...” menu
- We discussed two types of T-Tests in class. They have names in SPSS:
  - One-sample T-Test:
    - This corresponds to tests with a null hypothesis that compares a sample mean to a constant value ( $H_0: \mu = \mu_0$ )
  - Independent Samples T-Test:
    - This corresponds to tests with a null hypothesis that compares two means of *independent sample populations* ( $H_0: \mu_A = \mu_B$ )
- A dependent variable is sometimes called the “Test Variable”
- An independent variable is called the “Grouping Variable” or “Factor”
- You may notice that there are two values provided in the output for a T-Test: one for equal variance (discussed in class) and one for unequal variance. These are achieved by using a different denominator to calculate  $t$ , and yes the latter value can be used to avoid our assumption about equal variance :). Note, however, that when equal variance is not assumed, a different formula is also used to calculate the degrees of freedom.

## Questions

1. What is the dependent variable?
2. What are the independent variables (there are two)?
3. What is the sample size?
4. Describe the population.
5. Perform an *Independent Samples T-Test* using SPSS on the independent variable “Gender”.
  - a. What null hypothesis are you testing?
  - b. What are the degrees of freedom for each gender?
  - c. What is the total degrees of freedom?
  - d. What is the t-value?
  - e. What is the probability of a Type I error?
  - f. Was one mean significantly larger than the other? Which one?
  - g. How would you report the results in a paper? Begin with the phrase: “Males/Females were (not) significantly faster than Females/Males...”
6. Perform a *one-way ANOVA* using SPSS on the independent variable “Visualization”. When performing this test, enable two additional options:
  - Under “Post Hoc...” select “LSD”. “LSD” stands for “Least Significant Difference” and performs pairwise t-tests as a part of the analysis.
  - Under “Options” select “Descriptive”: this produces means and standard error values for each visualization.
    - a. What null hypothesis are you testing?
    - b. What are the means for each visualization?
    - c. What is the F-score?
    - d. What is the probability of a Type I error? What conclusion would you reach about your data based on this result?
    - e. Which pairs of means are different?
    - f. How would you report the main result of the ANOVA in a paper? Begin with the phrase: “A one-way ANOVA revealed a difference in speed for the three visualizations...”
    - g. How would you report the pairwise differences? Begin with the phrase: “Post-hoc analysis revealed that participants found nodes significantly faster with <VisName1> than with <VisName2>...”

7. Perform a one-way ANOVA using SPSS on the independent variable “Gender” (NOTE: use the “GenderNum” variable instead of the “Gender” variable).
  - a. What is the F-score?
  - b. What is the probability of a Type I error?
  - c. Compare the output of this test to the T-Test performed in question 5.
    - i. Were the p-values different?
    - ii. How does  $F$  relate to  $t$ ?