# Udacity Nanodegree – Data Analyst – Project 1

## **What is our independent variable? What is our dependent variable?**

## The independent variable is whether the words shown are congruent with the ink colors or if they are incongruent with ink colors.

## The dependent variable is the amount of time it takes to name, out loud, the color of the inks in which the words are displayed.

1. **What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.**

μC: Mean of Congruent Response Times

μI: Mean of Incongruent Response Times

Null Hypothesis:

Ho: μC - μI = 0

There will be no difference between congruent response times and incongruent response times.

Alternative Hypothesis:

H1: μC - μI ≠ 0

There will be a difference between congruent response times and incongruent response times.

For this hypothesis test, we’re going to use a dependent sample t-test.

We’re choosing a dependent t-test because the congruent and incongruent results come from the same sample. Also, we will be using a t-test as we don’t have population parameters to do a z-test. Since the Alternative hypothesis is not one sided, we’re going to use a two-tailed test.

1. **Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.**

See attached Spreadsheet!

1. **Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.**

See attached Spreadsheet for visualizations!

From the first scatter plot we can see that it appears the incongruent response times are longer then the congruent response times. From the 2nd bar graph we can verify our assumption from the first scatter plot graph that the incongruent data does in fact have a larger mean and median.

1. **Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Conclude in terms of the experiment task. Did the results match up with your expectations?**

For this test, we’ll use a two sided 99% confidence test. Our t-critical value will be 2.8073, and -2.8073. If our t-statistic is less then -2.8073 or greater then 2.8073 then we can reject the null in favor of the alternative hypothesis.

From our analysis, we get our t-statistic = -8.02, which is less then -2.8073 so we can reject the null in favor of the alternative. This meets our expectations that the incongruent test will have longer response times.

1. **Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions!**

The human brain is great at using associations to make itself operate more efficiently. That’s why habits are so important! The brain has linked the word with the color and therefore using the subconscious will be able to associate the word and color without really “consciously” thinking about it. Having the incongruent color assigned to the word we must use our conscious brain to detect this which apparently is slower in response time.

I would say this analogous to how fast you can spell your full name verse just spelling every other letter in your name. Even though there will only be half of the letters it would probably take longer to spell every other letter in your name.