# Data Analyst Nanodegree – Project 1

Report by Thomas Dräbing

## 1. What is our independent variable? What is our dependent variable?

The dependent variable is the time needed by the participants to state the colors of the ink. The independent variable is binary and represents which list, either the congruent or incongruent, it is, since the time a participant needs depends on which list he or she is working on.

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

*Now it's your chance to try out the Stroop task for yourself. Go to **this link**, which has a Java-based applet for performing the Stroop task. Record the times that you received on the task (you do not need to submit your times to the site.) Now, download **this dataset** which contains results from a number of participants in the task. Each row of the dataset contains the performance for one participant, with the first number their results on the congruent task and the second number their performance on the incongruent task.*

The question that is asked in this test is whether the participants are faster or slower in finishing one of the two lists. Thus the null hypothesis is:

$$\hat{t}_{con} = \hat{t}_{incon}$$

Where t_hat represents the mean time, con for congruent and incon for incongruent.

Thus the two sided alternative hypothesis would be:

$$\hat{t}_{con} \neq \hat{t}_{incon}$$

When we can assume a normal distribution, a two-sided paired Student's t-test would be a good choice to test whether the sample means are significantly different. The test has to be paired, since each participant provides both samples. The test is two-sided, when we do not want to prove whether one sample mean is significantly larger or smaller than the other, but just different.

We might also assume that finishing the incongruent list takes longer than finishing the congruent list and might want to prove that. The null hypothesis would be:

$$\hat{t}_{con} >= \hat{t}_{incon}$$

And the alternative hypothesis:

$$\hat{t}_{con} < \hat{t}_{incon}$$

Then a one-sided paired Student's t-test would be used.

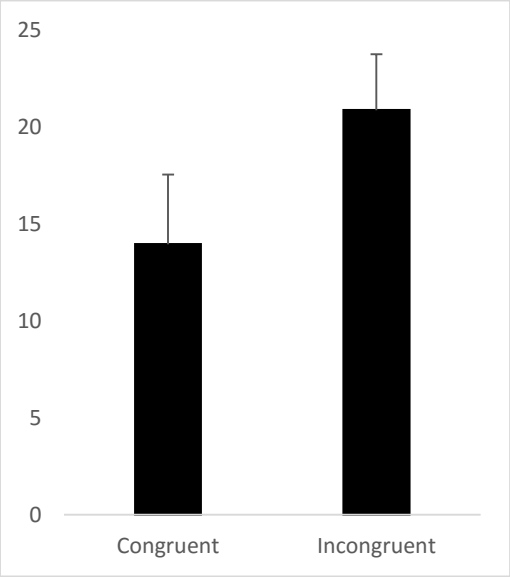
The Shapiro-Wilk test was used to test for normality (http://contchart.com/goodness-of-fit.aspx). The test resulted in a p-value of 0.69 for the congruent data and 0.003 for the incongruent data. Thus the null-hypothesis of normality cannot be rejected for the incongruent data.

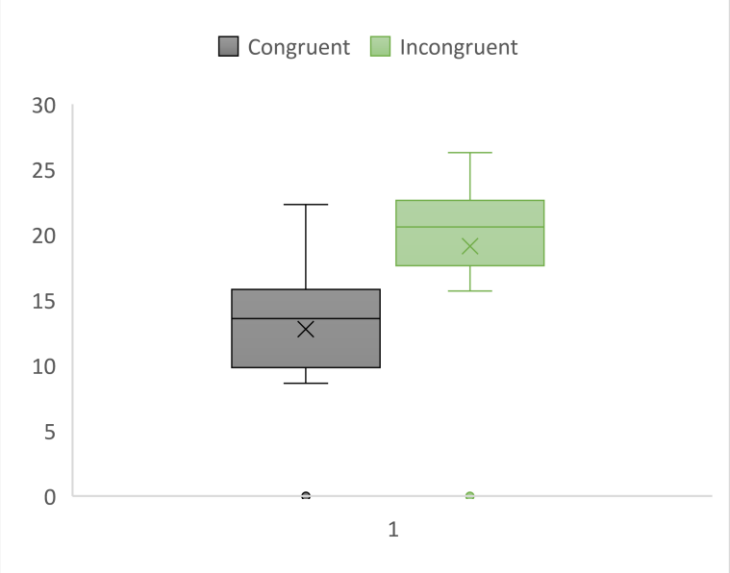
This could be due to outliers, which could be removed. If this does not work a nonparametric test like Mann-Whitney-U, which tests whether two distributions belong to the same population, could be used instead of the t-test.

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

| | **Congruent** | **Incongruent** |
|---|---|---|
| **Mean** | 14.051 | 22.016 |
| **Median** | 14.357 | 21.018 |
| **SD** | 3.559 | 4.797 |
| **Variance** | 12.669 | 23.012 |

## 4. 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.




The first plot is a bar chart of the mean times needed by the participants to finish the respective lists. The error bars represent the standard deviation. The second plot is a box plot, displaying the median, the quartiles and outliers. The plots show that the participants apparently need longer to finish the incongruent list, but there also seem to be two outliers for this list, which could influence this observation.

Removing the two sample pairs containing outliers results in a p-value of the Shapiro-Wilk-Test of 0.618 for the congruent sample and of 0.767 for the incongruent sample. Thus by removing the outliers a t-test can be used as a statistical test.

5. 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? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

The two-sided paired t-test results in a p-value of 4.382E-09. The Mann-Whitney-U-test (https://ccb-compute2.cs.uni-saarland.de/wtest/?id=www/www-ccb/html/software/wtest/) of the data still containing the outliers resulted in a p-value of 7.38E-09. At a confidence level of 99.9 %, which would result in a critical statistic value of 0.001, the null hypothesis can be rejected. Thus according to the t-test the means of the two samples are significantly different and according to the Mann-Whitney-U-test the two samples stem from two different populations. Together with the observation that the mean time to finish the incongruent list is higher that means that the participants needed significantly longer to finish the incongruent list, which would match my expectations.

6. 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 participants concentrated on reading the words and thus did not just look at the colors, which might confuse them, leading to longer response times. A child, before it learns to read, would probably perform much better. The same could be done with shapes, which have a label containing their shape or another shape respectively inside them.

The English Wikipedia page to 'Stroop effect' (https://en.wikipedia.org/wiki/Stroop_effect) further lists emotional tests, where facial expressions are mixed with emotional words, or numerical, where numbers are written in different font sizes either matching their numerical value or not.

It further sates that there are several theories to how this effect comes to pass, like the word processing being much faster than color processing or reading taking more attention, like I guessed above.