# Stroop effect

**1. What is our independent variable? What is our dependent variable?**
Independent variable – words conditions (congruent / incongruent)
Dependent variable – time to test completion

**2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.**
Null hypothesis:
Incongruent mean – congruent mean = 0

Alternative hypothesis:
Incongruent mean – congruent mean > 0

I expect to perform one-tailed t-test.

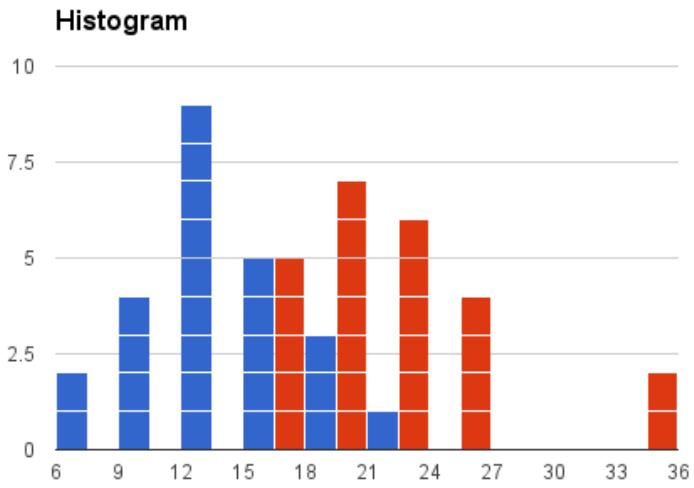
I chose a set of hypotheses and appropriate statistical test because intuitively it seems to me that people should perform far worse on incongruent condition.

**3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.**
Congruent:
Mean – 14.05
Standard deviation (sample) – 3.56

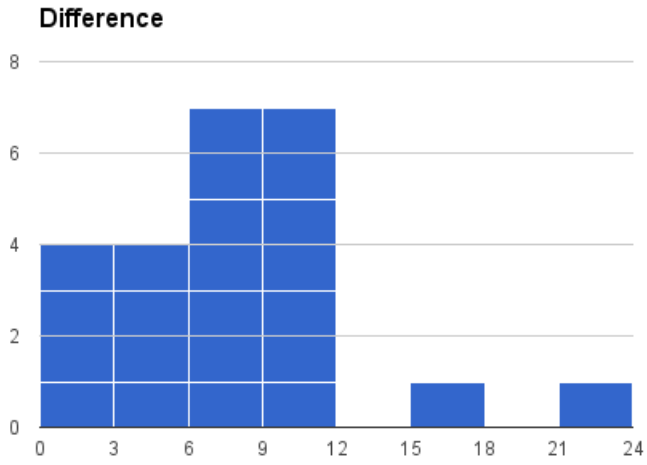
Incongruent:
Mean – 22.02
Standard deviation (sample) – 4.80

Difference:
Mean – 7.96
Standard deviation (sample) - 4.86

**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.**




Judging by this visualization on average people have better task completion time with congruent data and worse task completion time with incongruent data. This is consistent with computed means. Variability looks more or less equal except a few outliers for incongruent data which probably explain larger computed standard deviation.




Difference between incongruent and congruent data seems to be in 0-12 secs range, except few outliers

**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?**
Let's say we want 95% confidence that incongruent data does influence task completion time. For one-tailed test with alpha = 0.05 and df = 23,  t-critical value equals 1.714.

Standard error of the mean =  standard deviation / sqrt (n samples) = 0.99
t-statistic = mean / standard error = 8.04

We reject the null hypothesis because t-statistic > t-critical meaning $p<0.05$

Results did match with my expectations, though I am a little surprised with big t-statistic value/