Stroop Effect

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

The dependent variable is the time it takes to name the ink colors between congruent and incongruent samples. The independent variable is the actual color and whether the word is congruent or incongruent.

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

Null Hypothesis

The average amount of time it takes to identify incongruent words is less than or equal to the average amount of time it takes to identify congruent words.

H_0: μincongruent $\leq$ μcongruent (μincongruent represents the population mean of the time spent naming the incongruent words and μcongruent represents the population mean of the time spent naming the congruent words)

Alternative Hypothesis
The average amount of time it takes to identify incongruent words is greater than the average amount of time it takes to identify congruent words.

H_a: μincongruent $>$ μcongruent (μincongruent represents the population mean of the time spent naming the incongruent words and μcongruent represents the population mean of the time spent naming the congruent words)

My Time: It took me 22.581 seconds to identify the congruent words and 37.232 seconds for the incongruent words.

Statistical Test
The sample size is 24. The t-test would be used as the sample size is less than 30. The degrees of freedom would be 23 for a right tailed test.

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

    The mean value of the difference column is 7.96 and its standard deviation is 4.86.

```r
describe(stroopdata)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Congruent	1	24	14.05	3.56	14.36	13.88	3.49	8.63	22.33	13.70	0.37	-0.62	0.73
Incongruent	2	24	22.02	4.80	21.02	21.29	3.89	15.69	35.26	19.57	1.36	1.52	0.98
difference	3	24	7.96	4.86	7.67	7.40	5.17	1.95	21.92	19.97	0.94	0.80	0.99

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 boxplot below shows the incongruent column having a much larger median compared to the congruent column. We can also see some outliers present within the Incongruent column as well.
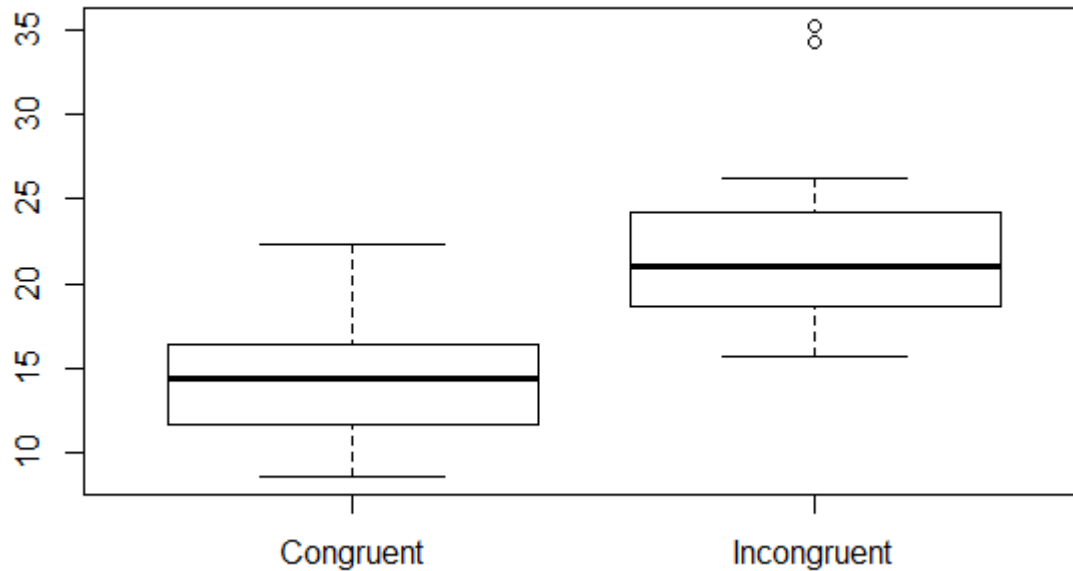
The histogram below shows the frequency of the differences among the incongruent and congruent columns. It shows the difference as having positive skewness.

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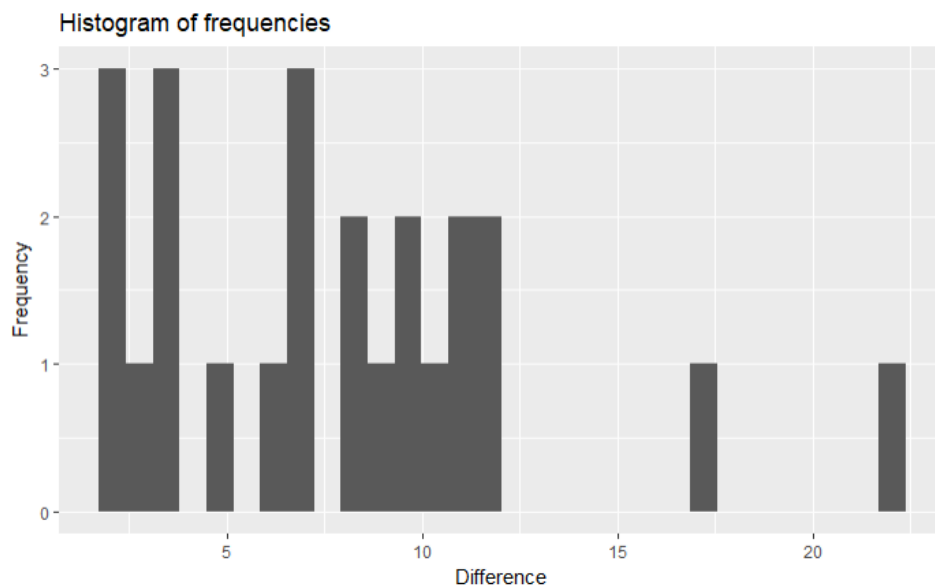
5. Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

The t-stat result from our data is 8.0207 at the 95% confidence interval for 23 degrees of freedom. The t-critical value is 1.714 at the .05 alpha level with 23 degrees of freedom. As the t-stat is larger than the t-critical, we are able to reject the null hypothesis and accept our alternative hypothesis. As a result,

there's a noticeable difference in the amount of time it takes a participant to identify congruent and incongruent words.

One tailed test:

```{r}
t.test(stroopdata$Incongruent,stroopdata$Congruent, paired = TRUE, alternative = "greater")
```

```
        Paired t-test

data:  stroopdata$Incongruent and stroopdata$Congruent
t = 8.0207, df = 23, p-value = 2.052e-08
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 6.262868      Inf
sample estimates:
mean of the differences
            7.964792
```

```{r}
abs(qt(0.05, 23))
```

```
[1] 1.713872
```

References:

http://personality-project.org/r/r.guide.html#descriptive

http://r4ds.had.co.nz/index.html

https://sapa-project.org/blog/2013/06/29/Descriptives-With-Psych/

https://stackoverflow.com/questions/11526041/critical-t-values-in-r

https://www.r-bloggers.com/one-sample-students-t-test/