In a Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participant's task is to say out loud the color of the ink in which the word is printed. The task has two conditions: a congruent words condition, and an incongruent words condition. In the congruent words condition, the words being displayed are color words whose names match the colors in which they are printed: for example RED, BLUE. In the incongruent words condition, the words displayed are color words whose names do not match the colors in which they are printed: for example PURPLE, ORANGE. In each case, we measure the time it takes to name the ink colors in equally-sized lists. Each participant will go through and record a time from each condition.

Q1: Question response correctly identifies the independent and dependent variables in the experiment.

The variable that is manipulated in this experiment: The conditions (Congruent and Incongruent, the way words are displayed)

And that is why, it is an independent variable.

Using dependent variable, the experiment tries to measure the change and in this case, it is the time
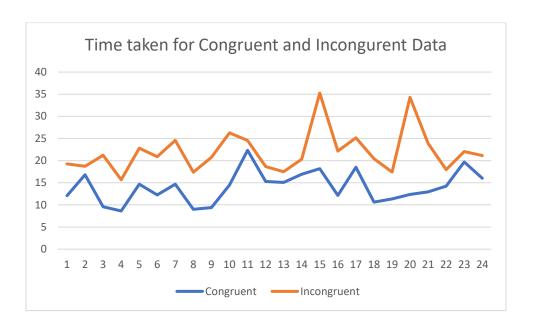
taken to name the ink colors.

Q2a: Null and alternative hypotheses are clearly stated in words and mathematically. Symbols in the mathematical statement are defined.

Ho - **Null Hypothesis**: ($\mu - \mu_i = 0$) There is **not** significant difference in the average response time in reading out the congruent words(μ) vs viewing the incongruent(i) words.

Ha - **Alternative Hypothesis**: ($\mu - \mu_i <> 0$)
The time taken to name the ink colors will vary for the two conditions.

Q2b: A statistical test is proposed which will distinguish the proposed hypotheses. Any assumptions made by the statistical test are addressed.

Time taken for Congruent and Incongurent Data

It is clear from the graph, that time taken for incongruent data is more than time taken for congruent data. We shall use T Test for dependent samples to verify this.

Since the same participants are present in both groups, we are going forward with this.

Moreover, we don't have data for the population, so we can't use Z Test.

Assumptions: There is sufficient difference between the time used to recognize colors under the two conditions so that one is not influenced by the other.

Q3: Descriptive statistics, including at least one measure of centrality and one measure of variability, have been computed for the dataset's groups.

	Congruent	Incongruent
Mean	14.051125	22.01591667
Median	14.3565	21.0175
Standard Deviation	3.559357958	4.797057122
Variance	12.66902907	23.01175704

Mean and Median are measures of centrality. Std Dev and Variance are measures of variability.

They have been calculated in Excel.

Q4: One or two visualizations have been created that show off the data, including comments on what can be observed in the plot or plots.

getwd()

stroop <- read.csv("stroopdata.csv")

library(ggplot2)

library(tidyr)

head(stroop,3)

Congruent Incongruent

1	12.079	19.278
2	16.791	18.741
3	9.564	21.214

 # To add a new column Number which has values from 1 to 24.

stroop$Number<- 1:nrow(stroop)

#Checking the data

head(stroop,3)

Congruent Incongruent Number

1	12.079	19.278	1
2	16.791	18.741	2
3	9.564	21.214	3

#Making the dataframe usable with Tidyr and gather with the below command

stroop_effect <- gather(stroop, "Congruency", "Time", 1:2)

#Checking the data now

head(stroop_effect,3)

```
  Number Congruency   Time
1      1  Congruent 12.079
2      2  Congruent 16.791
3      3  Congruent  9.564
```

tail(stroop_effect,3)

```
   Number  Congruency   Time
46     22 Incongruent 17.960
47     23 Incongruent 22.058
48     24 Incongruent 21.157
```
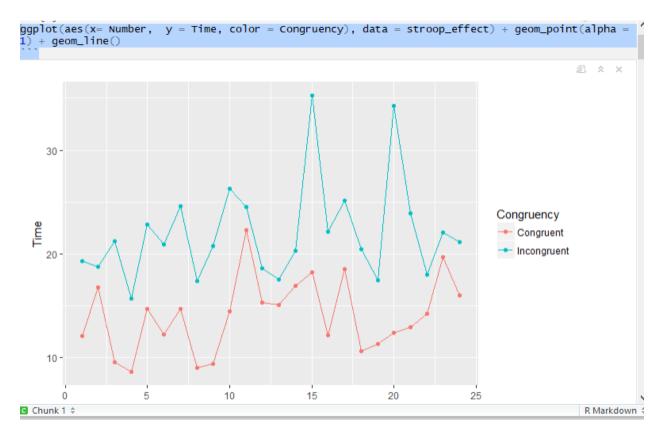
#Plotting the data

library(RColorBrewer)
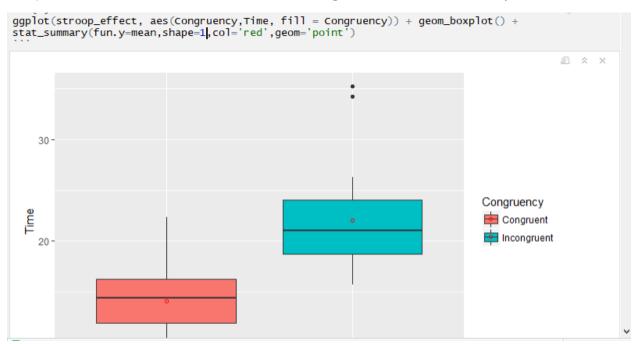
library(ggplot2)

ggplot(aes(x= Number, y = Time, color = Congruency), data = stroop_effect) + geom_point(alpha = 1) + geom_line()

```

```
ggplot(aes(x= Number, y = Time, color = Congruency), data = stroop_effect) + geom_point(alpha =
1) + geom_line()
```



This plot makes it clear that time taken to read incongruent data is more for everyone.

```
ggplot(stroop_effect, aes(Congruency,Time, fill = Congruency)) + geom_boxplot() +
stat_summary(fun.y=mean,shape=1,col='red',geom='point')
```


To confirm the time again for the two types of methods, the mean and median are also visible in this plot.

Q5: A statistical test has been correctly performed and reported, including test statistic, p-value, and test result. The test results are interpreted in terms of the experimental task performed.

Congruent	Incongruent	Difference				
12.079	19.278	-7.199	Average			
16.791	18.741	-1.95	-7.964791667			
9.564	21.214	-11.65				
8.63	15.687	-7.057				
14.669	22.803	-8.134	Std Dev			
12.238	20.878	-8.64	4.86482691			
14.692	24.572	-9.88				
8.987	17.394	-8.407	df = 23(24-1)		Degrees of freedom	
9.401	20.762	-11.361	n = 24		Total Value of samples	
14.48	26.282	-11.802				
22.328	24.524	-2.196	Dividing Std Dev by sqr root of n			
15.298	18.644	-3.346				
15.073	17.51	-2.437	0.993028635			
16.929	20.33	-3.401				
18.2	35.255	-17.055	Calculating T stats			
12.13	22.158	-10.028	-8.020706944	(Average/(Std dev/(sqrt(number of samples))		
18.495	25.139	-6.644	Calculating p value			
10.639	20.429	-9.79				
11.344	17.425	-6.081				
12.369	34.288	-21.919				
12.944	23.894	-10.95				
14.233	17.96	-3.727				
19.71	22.058	-2.348				
16.004	21.157	-5.153				

P Value Results
t=-8.0207 DF=23
The two-tailed P value is less than 0.0001
By conventional criteria, this difference is considered to be extremely statistically significant.

Chart Title

Congruent   Incongruent

Calculating T critical for two sided alpha = 0.05

It comes out to be 2.069

So, we can ignore the Null hypothesis now.

Which means that, reading congruent and incongruent data does not take the same amount of time.

Also, confidence interval

avg mean + - (t critical*(std dev/sqrt(n))

-5.910215421		so the confidence interval is (-10.019, -5.910)	
-10.01936791			

So, we can easily reject the null hypothesis and conclude that the difference between congruence and incongruence group time difference is statistically significant, namely, the Stroop effect is present.

Q6: Hypotheses regarding the reasons for the effect observed are presented. An extension or related experiment to the performed Stroop task is provided, that may produce similar effects.

Since we are familiar with those words and use them on daily basis, our cognitive powers are not that strong and hence we end up taking more time in the case of incongruent method.

An extension: We can add meaningless words to the list. (words that do not exist). In my opinion, it would take less time to read out the colors for those words.