Q.1)

Independent Variable: The conditions in the Stroop Task: Congruent condition and Incongruent condition are Independent Variables

Dependent Variable: The time measured for each observation (naming the ink colour of the given word) is dependent variable

Q.2)

Hypothesis:

Null Hypothesis: There is no difference in time taken by the individual in naming the colour in congruent condition and Incongruent condition

Alternate Hypothesis: There is difference in time taken by the individual in naming the colour in congruent condition and Incongruent condition

H_0: $\mu_{Con} = \mu_{Incon}$

(Mean of Congruent condition is equal to Mean of Incongruent condition)

H_A: $\mu_{Con} \neq \mu_{Incon}$

(Mean of Congruent condition is not equal to Mean of Incongruent condition)

Statistical Test:

The dependent t-test (also called the paired t-test or paired-samples t-test) compares the means of two related groups to determine whether there is a statistically significant difference between these means. As same participants are tested more than once, it is dependent t-test.

(Source: https://statistics.laerd.com/statistical-guides/dependent-t-test-statistical-guide.php)

Assumptions:(Source:http://www.psychology.emory.edu/clinical/bliwise/Tutorials/TOM/meanstests/assump.htm)
1. Sampling is random
2. The dependent variable is normally distributed in both congruent and Incongruent conditions.
3. Samples or sets of data used to produce the difference scores are linked in the population through repeated measurement, natural association, or matching

Assumptions: (Source: http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/hypothesis-tests/tests-of-means/why-use-2-sample-t/)

- Determine whether the means of two independent groups differ.
- Calculate a range of values that is likely to include the difference between the population means.
- The observations from the first sample must not have any bearing on the observations from the second sample.

Q.3)

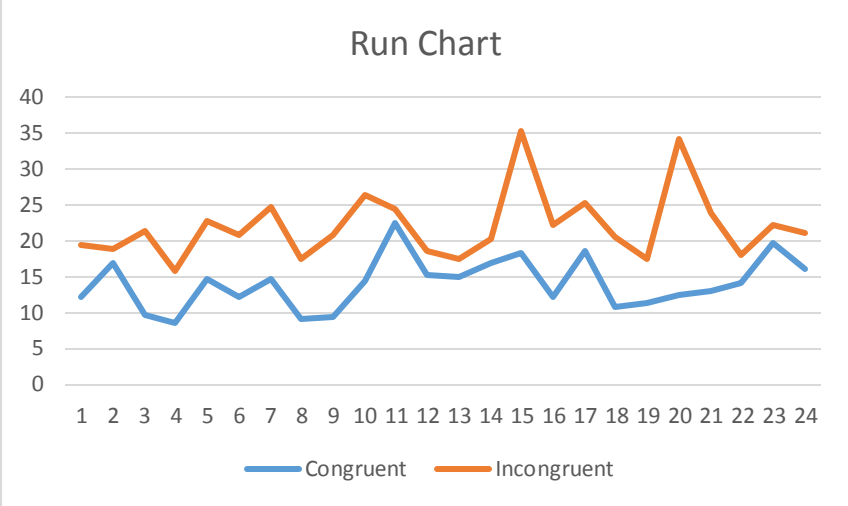
Descriptive Statistics: (Calculated in R Studio)

```
> setwd("F:/UB/Sem3/TA/HW 6")
> P1<-read.csv("stroopdata.csv")

> summary(P1$Congruent)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   8.63   11.90   14.36   14.05   16.20   22.33
> summary(P1$Incongruent)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  15.69   18.72   21.02   22.02   24.05   35.26

> sd(P1$Congruent)
[1] 3.559358
> sd(P1$Incongruent)
[1] 4.797057
```

It can be seen that Mean, Median, 1st Quartile, 3rd Quartile, Min and Max are higher for Incon gruent condition dataset.

Q.4)

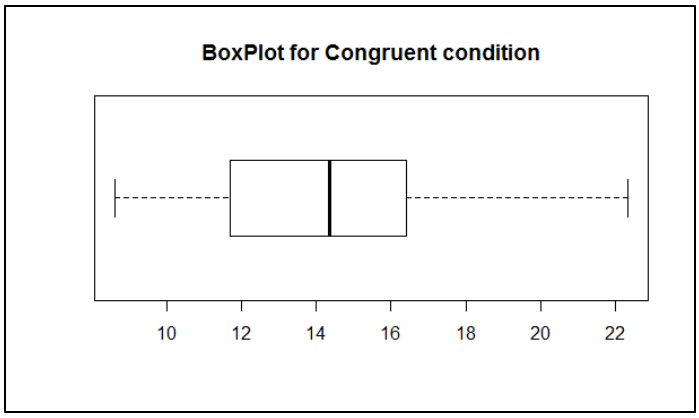
Run Chart (Source: Excel)



From this char we can say that time taken in Incongruent condition (orange line) is higher at e ach and every point of observations

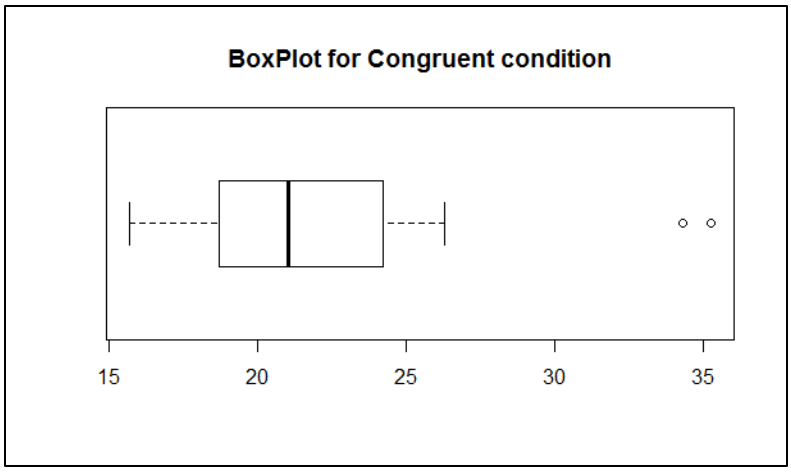
Boxplot (source: R Studio)

```
> boxplot(P1$Congruent, horizontal=TRUE, main="BoxPlot for Congruent condi
tion")
```



From Boxplot we can see that Median is '14', IQR is around 4 and there are no outliers.

```
> boxplot(P1$Incongruent, horizontal=TRUE, main="BoxPlot for Congruent con
dition")
```



This boxplot shows that Median is around '21' , IQR around '5' with 2 outliers.

Q.5)

Congruen	Incongrue	D	D-Dmean	(D-Dmear	Dmean	-7.96479
12.079	19.278	-7.199	0.765792	0.586437	S^2	23.67
16.791	18.741	-1.95	6.014792	36.17772	Std dev	4.86
9.564	21.214	-11.65	-3.68521	13.58076		
8.63	15.687	-7.057	0.907792	0.824086		
14.669	22.803	-8.134	-0.16921	0.028631		
12.238	20.878	-8.64	-0.67521	0.455906		
14.692	24.572	-9.88	-1.91521	3.668023		
8.987	17.394	-8.407	-0.44221	0.195548		
9.401	20.762	-11.361	-3.39621	11.53423		
14.48	26.282	-11.802	-3.83721	14.72417		
22.328	24.524	-2.196	5.768792	33.27896		
15.298	18.644	-3.346	4.618792	21.33324		
15.073	17.51	-2.437	5.527792	30.55648		
16.929	20.33	-3.401	4.563792	20.82819		
18.2	35.255	-17.055	-9.09021	82.63189		
12.13	22.158	-10.028	-2.06321	4.256829		
18.495	25.139	-6.644	1.320792	1.744491		
10.639	20.429	-9.79	-1.82521	3.331385		
11.344	17.425	-6.081	1.883792	3.548671		
12.369	34.288	-21.919	-13.9542	194.7199		
12.944	23.894	-10.95	-2.98521	8.911469		
14.233	17.96	-3.727	4.237792	17.95888		
19.71	22.058	-2.348	5.616792	31.54835		
16.004	21.157	-5.153	2.811792	7.906172		
14.05113	22.01592	-7.96479				

This is paired t-test, So difference between two samples is found out (D). This is subtracted and squared to get $(D-Dmean)^2$. This is divided by n-1=24 to get S^2

Std dev is found by its taking sqrt which is equal to 4.86

t-statistic = $(x1bar - x2bar)/(\frac{s(x1-x2)}{\sqrt{n}})$

x1bar = 14.05, x2bar=22.01

t-statistic =(14.05-22.01)/(4.86/sqrt(25))=-8.03, p-value=0.0001

(Source : https://graphpad.com/quickcalcs/pValue2/)

T-critical at alpha=0.05 2-tailed test =± 2.069 (df=23)

As t-statistic(-8.03) is less than T-critical (-2.069), and also p-value (0.001)<0.05, so Null hypothesis is rejected at 95% C.I

We accept the alternate hypothesis that there is a statistically significant difference in m=naming the colour of the ink in both the conditions. This is also evident from the Run chart above and so the results had match up.

Q.6)

The words themselves have a strong influence over your ability to say the color. The interference between the different information (what the words say and the color of the words) your brain receives causes a problem. There are two theories that may explain the Stroop effect:

1. Speed of Processing Theory: the interference occurs because words are read faster than colours are named.
2. Selective Attention Theory: the interference occurs because naming colors requires more attention than reading words.

(Source: https://faculty.washington.edu/chudler/words.html)

Interactive Animal Stroop Effect Experiment also has the same results.

(Source: https://faculty.washington.edu/chudler/java/readya.html)