

Table of Contents

Statistics:	2
1.	2
Ans:	2
2.	3
HYPOTHESES	3
STATISTICAL TESTS:	3
3.	4
> summary(data).....	4
> stat.desc(data).....	4
4.	5
R CODE:	5
RESULT:	5
EXCEL:.....	6
5.	7
HYPOTHESES:	7
T STATISTIC:.....	7
CRITICAL T VALUE:.....	7
DECISION:.....	7
RESULTS:	7
REFERENCES:	8

Statistics: The Science of Decisions

Project Instructions

Background Information

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.

Questions For Investigation

As a general note, be sure to keep a record of any resources that you use or refer to in the creation of your project. You will need to report your sources as part of the project submission.

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

Ans:

Dependent variable: Time

Independent Variable: Lists

Explanation: The variable is termed dependent when it depends on an independent variable. For instance, in the example presented, the time measured is dependent on the lists presented to the participants.

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

HYPOTHESES:

There should be two hypotheses H_0 (null hypotheses) and H_A (alternate hypotheses).

H_0 : There is no significant difference between the time taken for the congruent words from the incongruent words

H_A : The time taken for the different words are significantly different

STATISTICAL TESTS:

As both the word lists are dependent because the same participant will once take the congruent word test and then the in-congruent word test. So, I would perform a two sample T test to determine whether there is a significant difference in the scores for both the tests.

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

I used a statistical package in R called Hmisc to summarize the data set as follows:

```
> summary(data)
```

	Congruent	Incongruent
Min.	: 8.63	Min. :15.69
1st Qu.:	11.90	1st Qu.:18.72
Median :	14.36	Median :21.02
Mean :	14.05	Mean :22.02
3rd Qu.:	16.20	3rd Qu.:24.05
Max.	:22.33	Max. :35.26

The measure of central tendencies are:

Mean, Median

I used the statistical package called pastecs to view the variability measures of the data as follows:

```
> stat.desc(data)
```

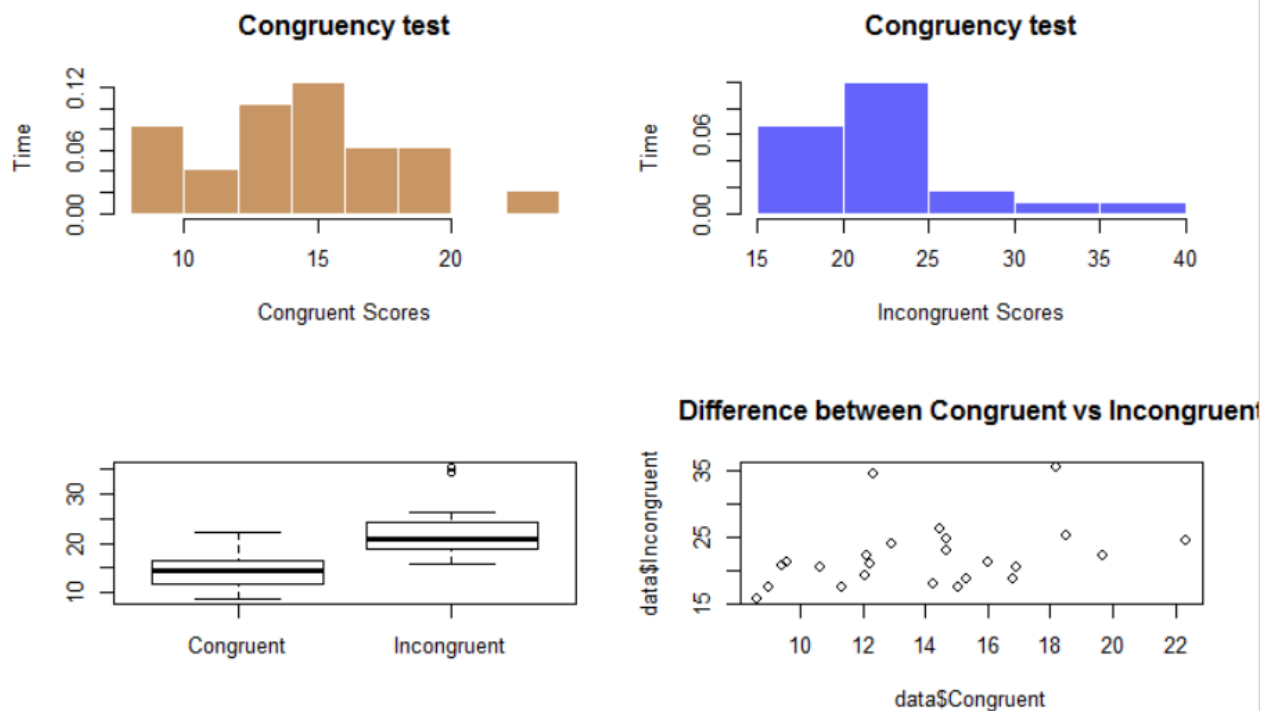
	Congruent	Incongruent
nbr.val	24.0000000	24.0000000
nbr.null	0.0000000	0.0000000
nbr.na	0.0000000	0.0000000
min	8.6300000	15.6870000
max	22.3280000	35.2550000
range	13.6980000	19.5680000
sum	337.2270000	528.3820000
median	14.3565000	21.0175000
mean	14.0511250	22.0159167
SE.mean	0.7265509	0.9791952
CI.mean.0.95	1.5029851	2.0256196
var	12.6690291	23.0117570
std.dev	3.5593580	4.7970571
coef.var	0.2533148	0.2178904

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.

R CODE:

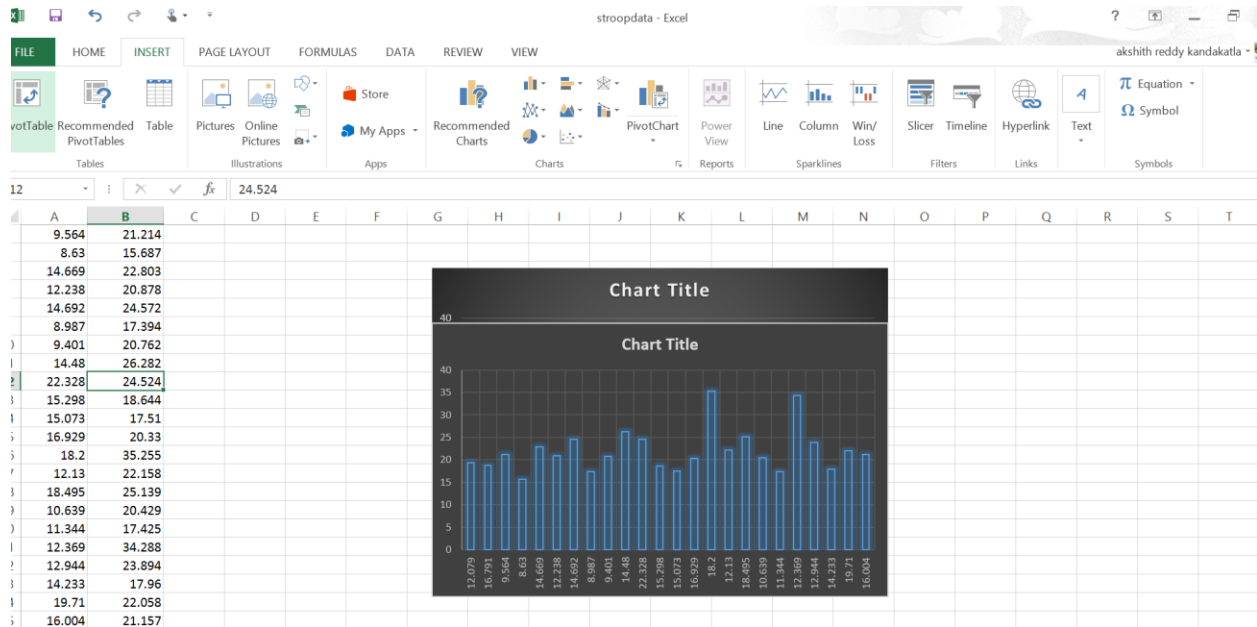
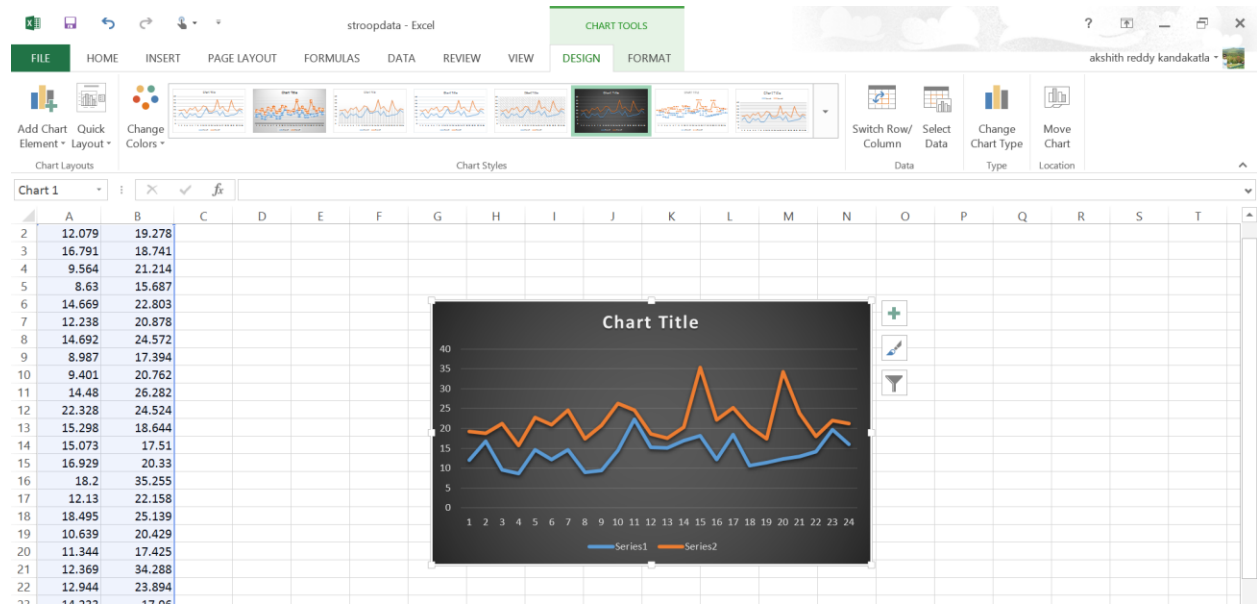
```
> hist(data$Congruent, xlab="Congruent Scores",ylab="Time", main="Congruency  
test",col=rgb(200,150,100,max=255),border=0,prob=TRUE)  
> hist(data$Incongruent, xlab="Incongruent Scores",ylab="Time", main="Congrue  
ncy test",col=rgb(100,100,250,max=255),border=0,prob=TRUE)  
> par(mfrow=c (2, 2)  
+ )  
> hist(data$Congruent, xlab="Congruent Scores",ylab="Time", main="Congruency  
test",col=rgb(200,150,100,max=255),border=0,prob=TRUE)  
> hist(data$Incongruent, xlab="Incongruent Scores",ylab="Time", main="Congrue  
ncy test",col=rgb(100,100,250,max=255),border=0,prob=TRUE)  
> boxplot(data)  
> plot(data$Congruent,data$Incongruent, main="Difference between Congruent vs  
Incongruent")
```

RESULT:



EXCEL:

However, we can also visualize the data using excel as follows:



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?

HYPOTHESES:

H_0 : There is no significant difference between the time taken for the congruent words from the incongruent words

H_A : The time taken for the different words are significantly different

T STATISTIC:

The T statistic can be computed by using the formula $(\mu_C - \mu_I) / (S / \sqrt{n})$ which was computed as follows to be : -8.020

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Congruent	Incongruent	diff btw cong & incong		Average Cong	Average Incong	avg of diff	Std dev of diff	denominator	t statistic			
2		12.079	19.278	-7.199	14.051125	22.01591667	-7.964791667	4.86482691	0.9930286348	-8.020706944			
3		16.791	18.741	-1.95									
4		9.564	21.214	-11.65									
5		8.63	15.687	-7.057									
6		14.669	22.803	-8.134									
7		12.238	20.878	-8.64									

CRITICAL T VALUE:

Assuming a 95% Confidence Interval

We have $\alpha = 0.05$ for a two sided T Test which is +/- 2.064

DECISION:

As the T statistic is way beyond the T Critical value we reject the Null Hypotheses meaning that the participants performed better on the Congruent tests

RESULTS:

The results were as expected.

REFERENCES:

All references were from the Udacity Statistics course materials.