Questions For Investigation

Test a Perceptual Phenomenon

Q1] What is our independent variable? What is our dependent variable?

Independent Variable : Congruent words & Incongruent words conditions
Dependent Variable : Time taken to read word, i.e. the response time.

Q2] What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

In the question no information is given about the population, just 24 sample datasets are given. The ability to compare the means of the dataset for the pre and post test validates the benefit of this test selection & since t-test is best fit for this case so two tailed test is done.

Hypothesis:-

Null Hypothsis,H0 :- No change in time between two reading tasks (congruent or incongruent)

Alternate Hypothesis,Ha :- Incongruent task take more time than Congruent.

H0 : μi ≤ μc (μi - population mean of incongruent values, μc - population mean of congruent values)

Ha: μi > μc (μi - population mean of incongruent values, μc - population mean of congruent values)

Q3] Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

In [1]:
```
import pandas as pd
from scipy.stats import t as pt
from matplotlib import pyplot as plt
pj = pd.read_csv("stroopdata.csv")
pj
```

Out[1]:

	Congruent	Incongruent
0	12.079	19.278
1	16.791	18.741
2	9.564	21.214
3	8.630	15.687
4	14.669	22.803
5	12.238	20.878
6	14.692	24.572
7	8.987	17.394
8	9.401	20.762
9	14.480	26.282
10	22.328	24.524
11	15.298	18.644
12	15.073	17.510
13	16.929	20.330
14	18.200	35.255
15	12.130	22.158
16	18.495	25.139
17	10.639	20.429
18	11.344	17.425
19	12.369	34.288
20	12.944	23.894
21	14.233	17.960
22	19.710	22.058
23	16.004	21.157

In [2]:
```
# this will make a new column with name difference
#new column will contain the difference between the congruent and incongruent

pj['Difference'] = pj['Incongruent'] - pj['Congruent']
pj
```

Out[2]:

	Congruent	Incongruent	Difference
0	12.079	19.278	7.199
1	16.791	18.741	1.950
2	9.564	21.214	11.650
3	8.630	15.687	7.057
4	14.669	22.803	8.134
5	12.238	20.878	8.640
6	14.692	24.572	9.880
7	8.987	17.394	8.407
8	9.401	20.762	11.361
9	14.480	26.282	11.802
10	22.328	24.524	2.196
11	15.298	18.644	3.346
12	15.073	17.510	2.437
13	16.929	20.330	3.401
14	18.200	35.255	17.055
15	12.130	22.158	10.028
16	18.495	25.139	6.644
17	10.639	20.429	9.790
18	11.344	17.425	6.081
19	12.369	34.288	21.919
20	12.944	23.894	10.950
21	14.233	17.960	3.727
22	19.710	22.058	2.348
23	16.004	21.157	5.153

In [3]:
```
#this calculates mean and standard deviation for congruent case, incongruent case
#and also for the Difference case

c_mean = pj.Congruent.mean()
c_std = pj.Congruent.std()
i_mean = pj.Incongruent.mean()
i_std = pj.Incongruent.std()
d_mean = pj.Difference.mean()
d_std = pj.Difference.std()
print('Congruent mean : ' + str(c_mean))
print('Congruent std dev : ' + str(c_std))
print('Incongruent mean : ' + str(i_mean))
print('Incongruent std dev : ' + str(i_std))
print('Difference mean : ' + str(d_mean))
print('Difference std dev : ' + str(d_std))
```

```
Congruent mean : 14.051125000000004
Congruent std dev : 3.5593579357645195
Incongruent mean : 22.015916666666667
Incongruent std dev : 4.797057122469138
Difference mean : 7.964791666666667
Difference std dev : 4.864826910359056
```

In [4]:
```
# it gives the descriptive statistics regarding this dataset
pj.describe()
```

Out[4]:

	Congruent	Incongruent	Difference
count	24.000000	24.000000	24.000000
mean	14.051125	22.015917	7.964792
std	3.559358	4.797057	4.864827
min	8.630000	15.687000	1.950000
25%	11.895250	18.716750	3.645500
50%	14.356500	21.017500	7.666500
75%	16.200750	24.051500	10.258500
max	22.328000	35.255000	21.919000

Q4] 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.

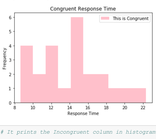
In [5]:
```
print(pj)
```

```
    Congruent  Incongruent  Difference
0      12.079       19.278       7.199
1      16.791       18.741       1.950
2       9.564       21.214      11.650
3       8.630       15.687       7.057
4      14.669       22.803       8.134
5      12.238       20.878       8.640
6      14.692       24.572       9.880
7       8.987       17.394       8.407
8       9.401       20.762      11.361
9      14.480       26.282      11.802
10     22.328       24.524       2.196
11     15.298       18.644       3.346
12     15.073       17.510       2.437
13     16.929       20.330       3.401
14     18.200       35.255      17.055
15     12.130       22.158      10.028
16     18.495       25.139       6.644
17     10.639       20.429       9.790
18     11.344       17.425       6.081
19     12.369       34.288      21.919
20     12.944       23.894      10.950
21     14.233       17.960       3.727
22     19.710       22.058       2.348
23     16.004       21.157       5.153
```

In [6]:
```
# It prints the Congruent column in histogram form

plt.hist(pj["Congruent"], color = "Pink")
plt.title("Congruent Response Time")
plt.xlabel("Response Time")
plt.ylabel("Frequency")
plt.legend(["This is Congruent"])
```

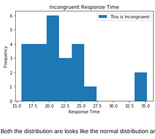
Out[6]: <matplotlib.legend.Legend at 0x215fe876c18>



In [7]:
```
# It prints the Incongruent column in histogram form

plt.hist(pj["Incongruent"])
plt.title("Incongruent Response Time")
plt.xlabel("Response Time")
plt.ylabel("Frequency")
plt.legend(["This is Incongruent"])
```

Out[7]: <matplotlib.legend.Legend at 0x215fe49e128>

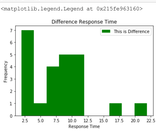


Both the distribution are looks like the normal distribution and we can see that the mean is different for both the distributions.

In [8]:
```
# It prints the Difference column in histogram form

plt.hist(pj["Difference"], color = "green")
plt.title("Difference Response Time")
plt.xlabel("Response Time")
plt.ylabel("Frequency")
plt.legend(["This is Difference"])
```

Out[8]: <matplotlib.legend.Legend at 0x215fe963160>



The difference of these 2 conditions is also like a normal distribution

Q5] 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 critical value of t at 95% confidence level:-

Sample Size,(n) = 24

Degree of freedom(df) = n - 1 = 23

In [9]:
```
#t-critical value for 95% confidance interval and 23 degree of freedom for two tailed test
#since it is a 2 tailed test so with 95% we will have 2.5% and 2.5% therefore inside ppf value is ta
ken as 0.975
#ppf is percent point function

print("The Critical value for the two tailed test is: ",round(pt.ppf(0.975,23),3))
```

The Critical value for the two tailed test is: 2.069

In [10]:
```
# Here first we calculate the Standard Error form the Standard deviation
# SE = SD/root(n), where n is the Sample Size
# SE is Standard Error (d_std_err)
# SD is Standard Deviation (d_std)

d_std_err=d_std/(24**0.5)
t = d_mean/(d_std/(24**0.5))
print("t-value = {}".format(t))
```

```
t-value = 8.020706944109957
```

The calculated t-value is 8.020706944109957 which is much larger than the t-critical value, i.e. 2.069

Since t-value is in the critical region (very far), so the null hypothesis is rejected.

Q6] What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions!

The Stroop effect is a phenomenon that occurs when you must say the color of a word but not the name of the word. For example, blue might be printed in red and you must say the color rather than the word. The brain has an image association between the shape of the word and the colour. When there is a mismatch, additional time is necessary for the prefrontal cortex to process the information and decide on its meaning.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.

Similar effect to the this effect can be reserve stroop effect.In this test the participant's task is to say out loud the word that is