Import neccessery libraries

In [1]:
```
import pandas as pd
import numpy as np
from scipy import stats
```

Problem

A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports of 4 laboratories. TAT is defined as sample collected to report dispatch

Analyze the data and determine whether there is any difference in average TAT among the different laboratories at 5% significance level

Data description

α = 0.05 (95% Confidence)

Y = Continious X = Discrete

Is Y1, Y2, Y3 and Y4 normal ?

H0 = Y1, Y2, Y3 and Y4 are normal H1 = Y1, Y2, Y3 and Y4 are not normal

H0: The null hypothesis: It is a statement of no difference between sample means or proportions or no difference between a sample mean or proportion and a population mean or proportion. In other words, the difference equals 0.

Ha: The alternative hypothesis: It is a claim about the population that is contradictory to H0 and what we conclude when we reject H0.

Import data

In [2]:
```
import os
```

In [3]:
```
os.getcwd()
```

Out[3]:
```
'C:\\Users\\Akarsh\\assignment 3'
```

In [4]:
```
os.chdir('C:\\Users\\Akarsh\\Desktop\\assignments')
```

In [5]:
```
os.getcwd()
```

Out[5]:
```
'C:\\Users\\Akarsh\\Desktop\\assignments'
```

Normality test

In [7]:
```
df = pd.read_csv('LabTAT.csv')
```

In [8]:
```
df
```

Out[8]:

	Laboratory 1	Laboratory 2	Laboratory 3	Laboratory 4
0	185.35	165.53	176.70	166.13
1	170.49	185.91	198.45	160.79
2	192.77	194.92	201.23	185.18
3	177.33	183.00	199.61	176.42
4	193.41	169.57	204.63	152.60
...	...	...	...	...
115	178.49	170.66	193.80	172.68
116	176.08	183.98	215.25	177.64
117	202.48	174.54	203.99	170.27
118	182.40	197.18	194.52	150.87
119	182.09	215.17	221.49	162.21

120 rows × 4 columns

In [10]:
```
stats.shapiro(df['Laboratory 1'])
```

Out[10]:
```
ShapiroResult(statistic=0.9901824593544006, pvalue=0.5506953597068787)
```

In [ ]:
```
#P value for Laboratory 1 = 0.55 > α
```

In [11]:
```
stats.shapiro(df["Laboratory 2"])
```

Out[11]:
```
ShapiroResult(statistic=0.9936322569847107, pvalue=0.8637524843215942)
```

In [ ]:
```
#P value for Laboratory 2 = 0.86 > α
```

In [12]:
```
stats.shapiro(df["Laboratory 3"])
```

Out[12]:
```
ShapiroResult(statistic=0.9886345267295837, pvalue=0.4205053448677063)
```

In [14]:
```
#P value for Laboratory 3 = 0.42 > α
```

In [13]:
```
stats.shapiro(df["Laboratory 4"])
```

Out[13]:
```
ShapiroResult(statistic=0.9913753271102905, pvalue=0.6618951559066772)
```

In [15]:
```
#P value for Laboratory 4 = 0.66 > α
```

HO is accepted. Thats is Y1, Y2, Y3, and Y4 are normal

Variance

H0 = Variance of all 4 laboratories are the same

H1 = Variance of all 4 laboratories are the not same

In [18]:
```
stats.levene(df["Laboratory 1"], df["Laboratory 2"], df["Laboratory 3"], d
```

Out[18]:
```
LeveneResult(statistic=2.599642500418024, pvalue=0.05161343808309816)
```

In [17]:
```
#P Value of Variance test is == 0.051 > α
```

HO is accepted.

One way anova test

model

H0 = Mean TAT for 4 laboratories equal (There is no significance difference between TAT of the laboratories)

H1 = Mean TAT for 4 laboratories not equal (There is a significance difference between TAT of the laboratories)

In [19]:
```
stats.stats.f_oneway(df["Laboratory 1"], df["Laboratory 2"], df["Laboratory
```

Out[19]:
```
F_onewayResult(statistic=118.70421654401437, pvalue=2.1156708949992414e-57)
```

In [21]:
```
#P value of the One way Anova test is == 2.116 < α
```

H1 is accepted.

Mean TAT for 4 laboratories not equal (There is a significance difference between TAT of the laboratories)

In [ ]:
```

```