In [1]:
```
### PROBLEM STATEMENT:COMPARING THE LIFETIMES OF 4 BRAND OF TYRE(APOLLO,CEAT,BR
IDGESTONE,FALKEN) ON VARIABLE CALLED AS MILEAGE.
## FOR EACH BRAND OF AUTOMOBILE WE HAVE TAKEN 15 SAMPLE OF OBSERVATIONS.

import numpy as np
import pandas as pd
from scipy.stats import levene, shapiro, f_oneway
from statsmodels.stats.multicomp import pairwise_tukeyhsd, MultiComparison
import matplotlib.pyplot as plt
from statsmodels.formula.api import ols
from statsmodels.stats.anova import anova_lm
```

In [2]:
```
# Anova test - one way Anova Example

# Mean Scores of 4 tyre.
# Anova helps in testing if the mean scores of all tyre are the same or not
# Null Hypothesis: Means are equal across all groups
# Alternate Hypothesis: Atleast one of the means are significantly different

#Read the excel file into an object
data = pd.read_csv('tyre.csv')
```

In [3]:
```
data.head()
```

Out[3]:

	Brands	Mileage
0	Apollo	32.998
1	Apollo	36.435
2	Apollo	32.777
3	Apollo	37.637
4	Apollo	36.304

In [9]:
```
# Performing shapiro and levenes test to confirm assumptions of Normality & Equ
al Variances

# Shapiro Test
# Null Hypothesis - Data is normally distributed
# Alternate Hypothesis - Data is not normally distributed

shapiro(data[data['Brands'] == 'Apollo'].Mileage)






# P-value>0.05 hence data is  normal
# (However, ANOVA is robust to violation of normality rule, provided variances
are equal)
# Repeating  above shapiro test for other 2 archers as well
```

Out[9]:
```
(0.9691987037658691, 0.8460143804550171)
```

In [10]:
```
shapiro(data[data['Brands'] == 'Bridgestone'].Mileage)

## P-value>0.05 hence data is  normal
```

Out[10]:
```
(0.9550896286964417, 0.6078232526779175)
```

In [11]:
```
shapiro(data[data['Brands'] == 'CEAT'].Mileage)

## P-value>0.05 hence data is  normal
```

Out[11]:
```
(0.9522789716720581, 0.5610349178314209)
```

In [12]:
```
shapiro(data[data['Brands'] == 'Falken'].Mileage)

## P-value>0.05 hence data is normal
```

Out[12]:
```
(0.9691987037658691, 0.8460143804550171)
```

In [15]:
```
# Levene's Test

# H0: All variances are equal
# Ha: Atleast one of the variances not equal

levene(data[data['Brands'] == 'Apollo'].Mileage,
                      data[data['Brands'] == 'Bridgestone'].Mileage,
                      data[data['Brands'] == 'CEAT'].Mileage,
                      data[data['Brands'] == 'Falken'].Mileage)

## P-value>0.05 hence all variance are equal
```

Out[15]:
```
LeveneResult(statistic=0.69462497936714074, pvalue=0.55920638642592246)
```

In [16]:
```
f, p = f_oneway(data[data['Brands'] == 'Apollo'].Mileage,
                      data[data['Brands'] == 'Bridgestone'].Mileage,
                      data[data['Brands'] == 'CEAT'].Mileage,
                      data[data['Brands'] == 'Falken'].Mileage)

print ('One-way ANOVA')
print ('=============')

print ('F value:', f)
print ('P value:', p, '\n')

# Since p-value<0.05 , we say that atleast one of the means is significantly di
fferent
```

```
One-way ANOVA
=============
F value: 17.9415133424
P value: 2.78098919789e-08
```

In [17]:
```
mc = MultiComparison(data['Mileage'], data['Brands'])
result = mc.tukeyhsd()

print(result)
print(mc.groupsunique)


## Falken is best of all types of tyre
```

```
   Multiple Comparison of Means - Tukey HSD,FWER=0.05
=======================================================
   group1      group2   meandiff  lower   upper  reject
-------------------------------------------------------
   Apollo   Bridgestone  -3.019  -5.1289 -0.9091  True
   Apollo       CEAT    -0.0379  -2.1478  2.072  False
   Apollo      Falken    2.8255   0.7156  4.9354  True
Bridgestone     CEAT     2.9811   0.8712  5.091   True
Bridgestone    Falken    5.8445   3.7346  7.9544  True
    CEAT       Falken    2.8635   0.7536  4.9734  True
-------------------------------------------------------
['Apollo' 'Bridgestone' 'CEAT' 'Falken']
```

In [ ]:
```
## CONCLUSION: Falken is best of all types of tyre
```