Name: Ankit Sharma

Email: kumarankitx022@gmail.com

Mob no.: +91-7677241423

ASSIGNMENT – 02

SIGNIFICANCE LEVEL/ VALUE:

The significance level, also denoted as alpha or α, is a measure of the strength of the evidence that must be present in your sample before you will reject the null hypothesis and conclude that the effect is statistically significant. The person conducting the hypothesis test determines the significance level before conducting the experiment.

The significance level is the probability of rejecting the null hypothesis when it is true. For example, a significance level of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference. Lower significance levels indicate that you require stronger evidence before you will reject the null hypothesis.

Common choices for the level of significance are $\alpha = 0.05$ and $\alpha = 0.01$.

P-VALUE:

P-value provides a convenient basis for drawing conclusions in hypothesis-testing applications. The *p*-value is a measure of how likely the sample results are, assuming the null hypothesis is true; the smaller the *p*-value, the less likely the sample results. If the *p*-value is less than α, the null hypothesis can be rejected; otherwise, the null hypothesis cannot be rejected. The *p*-value is often called the observed level of significance for the test. In other words, the evidence in your sample is strong enough to be able to reject the null hypothesis at the level.

In [1]:

```
import numpy as np
import pandas as pd
import scipy.stats as stats
import matplotlib.pyplot as plt
%matplotlib inline
```

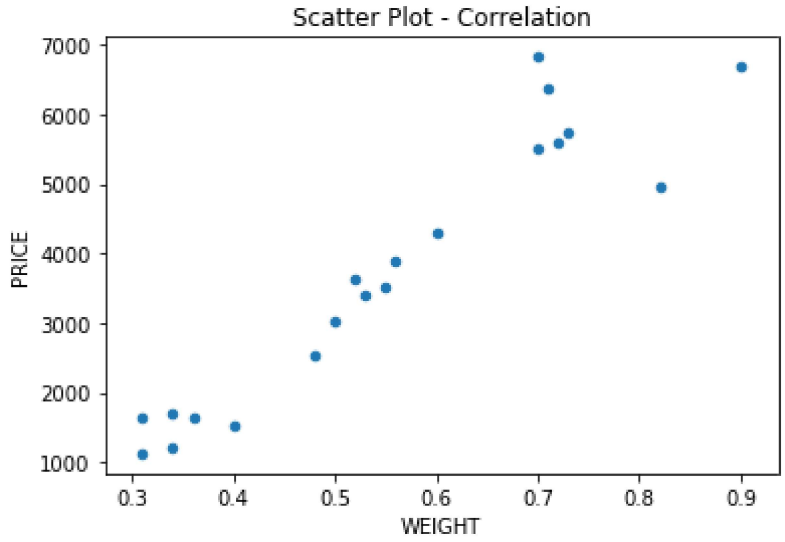
In [2]:

```
diamond = pd.read_excel('datasets/Inferential_Statistics.xlsx', sheet_name = 'Sheet1')
```

Correlation Coefficients

In [3]:

```
ScatterPlot = diamond.plot(kind='scatter',x='WEIGHT',y='PRICE',title='Scatter Plot - Correl
```



In [4]:

```
diamond['WEIGHT'].corr(diamond['PRICE']) #positive correlation
```

Out[4]:

```
0.9457713913032358
```

Correlation Coefficient Matrix

Here all the diagonal values will be 1 while correlation coefficient will be there for all the combination of the numerical variables. There are two methods of calculating Correlation Coefficient and its matrix – Pearson and Spearman.
diamond.corr(method= 'pearson')

In [5]:

```
diamond.corr(method= 'spearman')
```

Out[5]:

	IDNO.	WEIGHT	PRICE
IDNO.	1.000000	0.928115	0.921805
WEIGHT	0.928115	1.000000	0.925104
PRICE	0.921805	0.925104	1.000000

In [6]:

```
diamond.corr(method= 'pearson')
```

Out[6]:

	IDNO.	WEIGHT	PRICE
IDNO.	1.000000	0.893409	0.925019
WEIGHT	0.893409	1.000000	0.945771
PRICE	0.925019	0.945771	1.000000