In [1]:
```
#import the libraries
import pandas as pd
import scipy
import numpy as np
from scipy import stats
import warnings
warnings.filterwarnings("ignore")
```

In [2]:
```
Cut = pd.read_csv("Cutlets.csv")
Cut.head(5)
```

Out[2]:

	Unit A	Unit B
0	6.8090	6.7703
1	6.4376	7.5093
2	6.9157	6.7300
3	7.3012	6.7878
4	7.4488	7.1522

In [3]:
```
#Random sample of Unit A
sample_size= 10
Cut_Asample= np.random.choice(Cut.iloc[:,0],sample_size)
Cut_Asample
```

Out[3]:
```
array([6.5797, 7.2783, 7.2705, 6.9405, 6.8755, 7.2854, 6.8755, 7.2828,
       6.9952, 7.2705])
```

In [4]:
```
#Random sample of Unit B
sample_size= 10
Cut_Bsample= np.random.choice(Cut.iloc[:,1],sample_size)
Cut_Bsample
```

Out[4]:
```
array([6.8889, 6.7478, 7.1688, 7.0503, 6.9182, 6.9399, 6.73  , 6.8889,
       7.5459, 7.4314])
```

Independent T-Test

In [5]:
```
Ttest = stats.ttest_ind(Cut_Asample,Cut_Bsample)
Ttest
```

Out[5]:
```
Ttest_indResult(statistic=0.2940034151032175, pvalue=0.772118608262254)
```

In [6]:
```
#p-value
p_value = stats.ttest_ind(Cut_Asample,Cut_Bsample)[1]
p_value
```

Out[6]:
```
0.772118608262254
```

In [7]:
```
if p_value <0.05:                                    # alpha value is 0.05 or 5%
    print(" we are rejecting null hypothesis")
else:
    print("we are accepting null hypothesis")
```

```
we are accepting null hypothesis
```

Since the p- value is always > 0.05, there is a significant difference in diameter of Cutlet between 2 units.