

In [1]:

```
import pandas as pd
import numpy as np
from scipy import stats
```

In [2]:

```
cutlet=pd.read_csv('Cutlets.csv')  
cutlet
```

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
5	7.3871	6.8110
6	6.8755	7.2212
7	7.0621	6.6606
8	6.6840	7.2402
9	6.8236	7.0503
10	7.3930	6.8810
11	7.5169	7.4059
12	6.9246	6.7652
13	6.9256	6.0380
14	6.5797	7.1581
15	6.8394	7.0240
16	6.5970	6.6672
17	7.2705	7.4314
18	7.2828	7.3070
19	7.3495	6.7478
20	6.9438	6.8889
21	7.1560	7.4220
22	6.5341	6.5217
23	7.2854	7.1688
24	6.9952	6.7594
25	6.8568	6.9399
26	7.2163	7.0133
27	6.6801	6.9182
28	6.9431	6.3346
29	7.0852	7.5459
30	6.7794	7.0992
31	7.2783	7.1180
32	7.1561	6.6965
33	7.3943	6.5780

	Unit A	Unit B
34	6.9405	7.3875

In [3]:

```
np.mean(cutlet['Unit A'])
```

Out[3]:

7.01909142857143

In [4]:

```
np.mean(cutlet['Unit B'])
```

Out[4]:

6.964297142857142

Formulation of Hypothesis

$H_0 : \mu_1 = \mu_2$ (There is no significance difference between two diameters)

$H_a : \mu_1 \neq \mu_2$ (There is significance difference between two diameters)

Test

We are going for 2 sample t-test because we have 2 samples and qualitative data

In [5]:

```
stats.ttest_ind(cutlet['Unit A'],cutlet['Unit B'])
```

Out[5]:

Ttest_indResult(statistic=0.7228688704678061, pvalue=0.4722394724599501)

Significance value is 0.05

Here p-value is 0.47223 which is greater than significance value

Take away

We don't reject H_0 that is there is no significance difference between two diameters

In []:

In []: