Import neccessery libraries

In [2]:
```python
import pandas as pd
import numpy as np
from scipy import stats
```

Problem

A F&B manager wants to determine whether there is any significant difference in the diameter of the cutlet between two units. A randomly selected sample of cutlets was collected from both units and measured? Analyze the data and draw inferences at 5% significance level. Please state the assumptions and tests that you carried out to check validity of the assumptions

Data description

α == 0.05 (95% Confidence)

Y == Continious X == Discrete

Is Y1 and Y2 normal ?

H0 = Y1 and Y2 are normal H1 = Y1 and Y2 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 [3]:
```python
import os
```

In [4]:
```python
os.getcwd()
```

Out[4]: `'C:\\Users\\Akarsh\\assignment 3'`

In [5]:
```python
os.chdir('C:\\Users\\Akarsh\\Desktop\\assignments')
```

In [6]:
```python
os.getcwd()
```

Out[6]: `'C:\\Users\\Akarsh\\Desktop\\assignments'`

Normality test

In [8]:
```python
df = pd.read_csv('Cutlets.csv')
```

In [16]:
```python
df
```

Out[16]:

	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

	Unit A	Unit B
31	7.2783	7.1180
32	7.1561	6.6965
33	7.3943	6.5780

In [9]:
```python
stats.shapiro(df["Unit A"])
```

Out[9]:
```
ShapiroResult(statistic=0.9649458527565002, pvalue=0.3199819028377533)
```

In [12]:
```python
#P value for Unit A == 0.32 > α
```

In [10]:
```python
stats.shapiro(df["Unit B"])
```

Out[10]:
```
ShapiroResult(statistic=0.9727300405502319, pvalue=0.5224985480308533)
```

In [13]:
```python
#P value for Unit B == 0.52 > α
```

HO is accepted. Thats is both Y1 and Y2 are normal

Paired t test

model

H0 = Mean for Y1 and Y2 are equal (There is no significance difference between diameter of the Culets)

H1 = Mean for Y1 and Y2 are not equal (There is a significance difference between diameter of the Culets)

In [14]:
```python
stats.ttest_rel(df["Unit A"], df["Unit B"])
```

Out[14]:
```
Ttest_relResult(statistic=0.7536787225614314, pvalue=0.4562300768038412)
```

In [15]:
```python
#P value of the Paired T Test is == 0.45 > α
```

H0 is accepted.

Mean of both Y1 and Y2 are equal

In [ ]: