Hypothesis Testing Exercise

1) 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.

- Minitab File : **Cutlets.mtw**

ANS:-

```
import pandas as pd
import numpy as np
import scipy
from scipy import stats
culter=pd.read_csv("C:\\Users\\kiran\\Downloads\\Cutlets.csv")
culter.head()
```

	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

```
stats.ttest_ind( culter["Unit A"], culter["Unit B"]) [1]
```

```
0.472239472459950

Hence p value is greatern then 0.05.
We will go with the h0
There is no any significant difference in diameter of the culter betwee
n two units.
```

2) A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports of 4 laboratories. TAT is defined as sample collected to report dispatch.

Analyze the data and determine whether there is any difference in average TAT among the different laboratories at 5% significance level.

- Minitab File: **LabTAT.mtw**

ANS:-

```
import pandas as pd
import numpy as np
import scipy
from scipy import stats
LabTaT=pd.read_csv("C:\\Users\\kiran\\Downloads\\LabTAT.csv")
LabTaT.head()
```

	Laboratory 1	Laboratory 2	Laboratory 3	Laboratory 4
0	185.35	165.53	176.70	166.13
1	170.49	185.91	198.45	160.79
2	192.77	194.92	201.23	185.18
3	177.33	183.00	199.61	176.42
4	193.41	169.57	204.63	152.

```
stats.f_oneway( LabTaT["Laboratory 1"],LabTaT["Laboratory 2"],LabTaT["Laboratory3"],LabTaT["Laboratory4"])
```

F_onewayResult(statistic=118.70421654401437, pvalue=2.1156708949992414e-57)

Hence the p value is smaller then 0.05

We will go with h1

There is difference in average TAT of reports of the laboratories.

3) Sales of products in four different regions is tabulated for males and females. Find if male-female buyer rations are similar across regions.

ANS:- BuyerRatio=pd.read_csv("C:\\Users\\kiran\\Downloads\\BuyerRatio.csv")

BuyerRatio

Observed Values	East	West	North	South
0 Males	50	142	131	70
1 Females	435	1523	1356	750

table=BuyerRatio.iloc[:,1:6]

table

East	West	North	South
50	142	131	70
435	1523	1356	750

stats.chi2_contingency(table) [1]

0.660309490709188

Hence we know that p value is greater then 0.05

We will reject the h0

Means ,all proportional are not equal.

4) TeleCall uses 4 centers around the globe to process customer order forms. They audit a certain % of the customer order forms. Any error in order form renders it defective and has to be reworked before processing. The manager wants to check whether the defective % varies by centre. Please analyze the data at 5% significance level and help the manager draw appropriate inferences

Minitab File: **CustomerOrderForm.mtw**

ANS:- telecall=pd.read_csv("C:\\Users\\kiran\\Downloads\\Costomer+OrderForm.csv")

Telecall

```
telecall['Phillippines'].value_counts()
telecall['Indonesia'].value_counts()
telecall['Malta'].value_counts()
telecall['India'].value_counts()

df= { 'Phillippines':[271, 29],
    'Indonesia' : [267, 33],
    'Malta':[269,31] ,
    'India':[280,20]  }

tab=pd.DataFrame(df)
tab

stats.chi2_contingency(tab)
```