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.

File : **Cutlets.csv**

BUSSINESS PROBLEM:

Is there any significance difference between the diameter of the two units?

Normality test:

H0: Data is normal

Ha: Data is not normal

95%Confidence level (1-alpha)

5% confidence level(alpha)

Unit. A: P value>alpha=0.32>0.05=P high H0 fly=accept H0

Unit. B: P value>alpha=0.5225>0.05=P high H0 fly= accept H0

Accept H0=>data is normally distributed

Variance test:

H0: variance of Unit. A =variance of Unit. B

Ha: variance of unit. A ≠variance of unit. B

P value=0.3136>0.05=> P high H0 fly => accept H0

2 Sample t test:

H0: diameter of the cutlet from Unit. A = diameter of the cutlet from Unit. B

Ha: diameter of the cutlet from Unit. A ≠ diameter of the cutlet from Unit. B

P value=0.4723>0.05=> P high H0 fly=>accept H0

This means that there is no significant difference in the diameters of the cutlets between both the units.

**Hypothesis Testing Exercise**

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.

File: **LabTAT.csv**

BUSSINESS PROBLEM:

Is there any difference in the average Turn Around Time (TAT) of reports of the laboratories.

NORMALITY TEST:

H0: Data is normal

Ha: Data is not normal

95%Confidence level (1-alpha)

5% confidence level(alpha)

Laboratory.1: P value>alpha=0.5508>0.05=P high H0 fly=accept H0

Laboratory.2: P value>alpha=0.8637>0.05=P high H0 fly=accept H0

Laboratory.3: P value>alpha=0.4205>0.05=P high H0 fly=accept H0

Laboratory.4: P value>alpha=0.6619>0.05=P high H0 fly=accept H0

Conclusion: Accept H0=>data is normal.

VARIANCE TEST:

H0: variance of laboratory 1= variance of laboratory 2

Ha: variance of laboratory 1 ≠ variance of laboratory 2

P value=0.1675

H0: variance of laboratory 2= variance of laboratory 3

Ha: variance of laboratory 2 ≠ variance of laboratory 3

P value=0.2742

H0: variance of laboratory 3= variance of laboratory 4

Ha: variance of laboratory 3 ≠ variance of laboratory 4

P value=0.3168

H0: variance of laboratory 4= variance of laboratory 1

Ha: variance of laboratory 4 ≠ variance of laboratory 1

P value=0.1408

P high =>H0 fly=>accept H0

Conclusion: variances of all laboratories are assumed to equal.

Anova test:

H0: average Turnaround time of all the laboratories are equal.

Ha: average turnaround time of at least one laboratory is different.

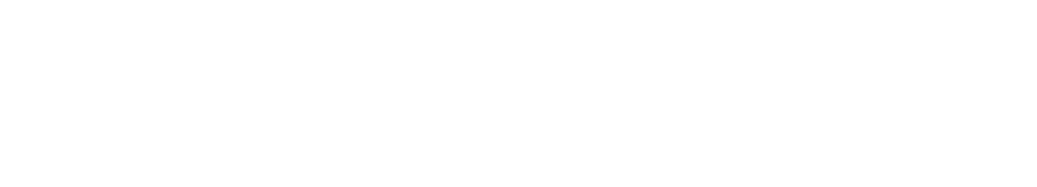
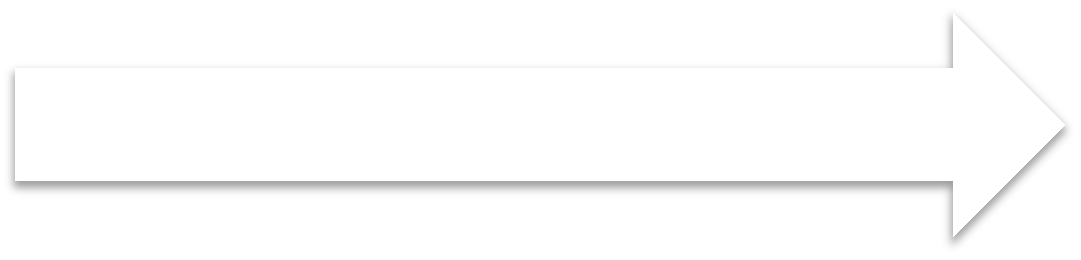
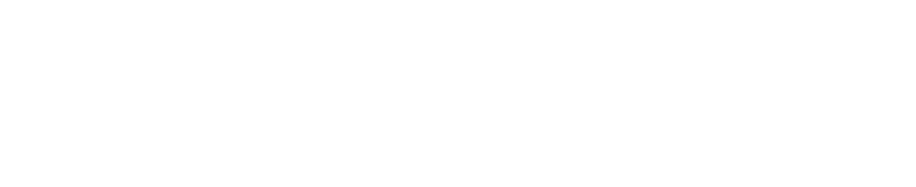
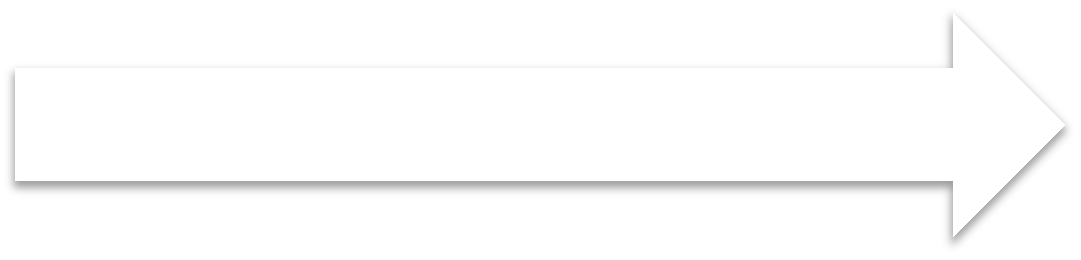
P value = 2e-16

P low =>H0 go=>accept Ha

Conclusion average turnaround time of at least one laboratory is different.

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **East** | **West** | **North** | **South** |
| Males | 50 | 142 | 131 | 70 |
| Females | 550 | 351 | 480 | 350 |



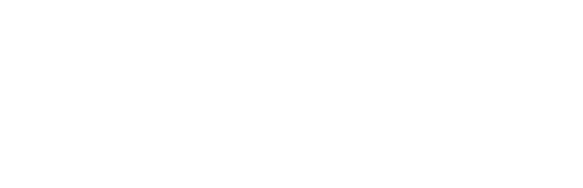
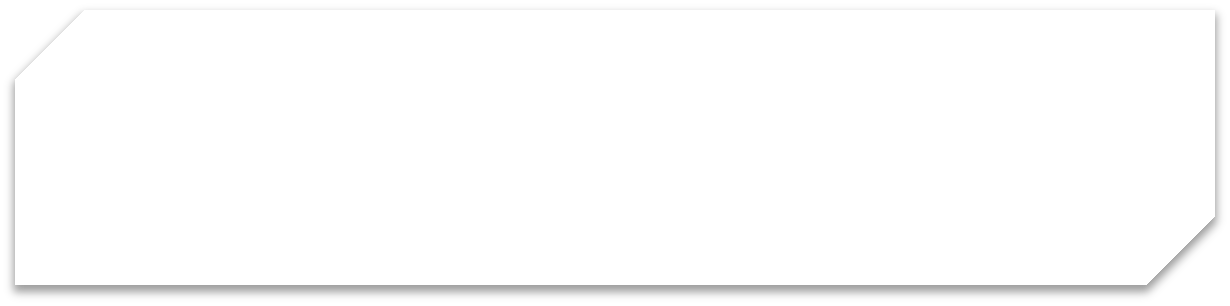
H0

* All proportions are equal

Ha

* Not all Proportions are equal

1. Check p-value
2. If p-Value < alpha, we reject Null Hypothesis



Buyer Ratio.csv

* H0: All proportions are equal
* Not all Proportions are equal

Unstacked data converted to stacked data using stack function.

Chi-square test:

p-value = 0.2931

p high =>H0 fly=>accept H0

conclusion: male-female buyer rations are similar across regions

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 must be reworked before processing. The manager wants to check whether the defective % varies by center. Please analyze the data at *5%* significance level and help the manager draw appropriate inferences

File: **Customer OrderForm.csv**

H0: Defects among all the centers are equal.

Ha: Defects among all the centers are not equal.

As there is no vector column, we cannot stack data, so we need to create vector column for the further processing.

Stack the data

create contingency table

**chi-square test:**

p-value = 0.2771>0.05=>p high =>H0 fly=>accept H0

conclusion: Defects among all the 4 centers are equal.

Fantaloons Sales managers commented that *%* of males versus females walking into the store differ based on day of the week. Analyze the data and determine whether there is evidence at *5 %* significance level to support this hypothesis.

File: **Fantaloons.csv**

H0: proportion A = proportion B

Ha: proportion A not= proportion B

2 proportion test: (two sided)

p-value = 8.543e-05<0.05=>p low =>Ho go=>accept Ha

# Unequal proportions

Conclusion: males entering the store = females entering the store

2 proportion test: (greater)

H0 -> proportion A <= proportion B

Ha -> proportion A > proportion B

p-value = 4.272e-05<0.05=> p low Ho go=>accept Ha