**Data Loading and Treatment:**

**General data**

import pandas as pd

dataset=pd.read\_excel("general\_data.xlsx",sheet\_name=0)

dataset.dropna()

**Attrition\_Yes**

import pandas as pd

dataset1=pd.read\_excel("general\_data.xlsx",sheet\_name=1)

dataset1.dropna()

**Attrition\_No**

import pandas as pd

dataset2=pd.read\_excel("general\_data.xlsx",sheet\_name=2)

dataset2.dropna()

**Wilcoxon Sign Test🡪 To compare 2 paired samples**

H0🡪 There is no significant difference between the YearsSinceLastPromotion and YearsWithCurrManager of the Employee

Ha🡪 There is significant difference between the YearsSinceLastPromotion and YearsWithCurrManager of the Employee

import pandas as pd

dataset=pd.read\_excel("general\_data.xlsx",sheet\_name=0)

dataset.dropna()

from scipy.stats import wilcoxon

stats,p=wilcoxon(dataset.YearsSinceLastPromotion,dataset.YearsWithCurrManager)

print(stats,p)

699082.5 7.578916343856332e-279

Ha 🡪Accepted

**Freidman Test 🡪 To compare more than 2 paired samples**

H0🡪 There is no significant difference between the NumCompaniesWorked, YearsSinceLastPromotion and YearsWithCurrManager of the Employee

Ha🡪 There is significant difference between the NumCompaniesWorked, YearsSinceLastPromotion and YearsWithCurrManager of the Employee

import pandas as pd

dataset=pd.read\_excel("general\_data.xlsx",sheet\_name=0)

dataset.dropna()

from scipy.stats import friedmanchisquare

stats,p=friedmanchisquare(dataset.NumCompaniesWorked,dataset.YearsSinceLastPromotion,dataset.YearsWithCurrManager)

print(stats,p)

1214.1914754953116 2.196196065710543e-264

Ha 🡪Accepted

**Mann Whitney Test🡪 To compare 2 independent samples**

H0: There is no significant differences in the Distance From Home between attrition (Y) and attirition (N)

Ha: There is significant differences in the Distance From Home between attrition (Y) and attirition (N)

from scipy.stats import mannwhitneyu

a1=dataset1.DistanceFromHome

a2=dataset2.DistanceFromHome

stat, p=mannwhitneyu(a1,a2)

print(stat, p)

1312110.0 0.4629185205822659

H0🡪 Accepted

**KRUSHKAL WALLIS TEST:**

H0 - There is no significant difference the in the monthlyIncome for managers, HR and Sales Executive.

H1 - There is a significant difference the in the monthlyIncome for managers, HR and Sales Executive.

import pandas as pd

dataset3=pd.read\_excel("general\_data.xlsx",sheet\_name=3)

dataset3.dropna()

from scipy.stats import kruskal

stats,p =kruskal(dataset3.Manager,dataset3.HR,dataset3.SalesExec)

print(stats,p)

1.9723833757421623 0.3729944689982891

H0🡪 Accepted

**Chi Square Test🡪 To check the dependency of the variable. Variables should be categorical**

H0🡪 There is no dependency between Gender and MaritalStatus

Ha🡪 There is dependency between Gender and MaritalStatus

import pandas as pd

dataset=pd.read\_excel("general\_data.xlsx",sheet\_name=0)

dataset1=dataset.dropna()

from scipy.stats import chi2\_contingency

chitable=pd.crosstab(dataset1.Gender,dataset.MaritalStatus)

chitable

Out[3]:

MaritalStatus Divorced Married Single

Gender

Female 364 841 551

Male 606 1166 854

stats,p,dof,expeted=chi2\_contingency(chitable)

print(stats,p)

5.849261068501047 0.053684522841547

H0🡪Accepted