**CORRELATION ANALYSIS**

In order to find the correlation of the variables Age, DistanceFromHome, MonthlyIncome, TotalWorkingYears, YearsAtCompany, YearsWithCurrManager from that of Attrition, we executed the Correlation Analysis as follows.

import pandas as pd

from scipy.stats import pearsonr

dataset = pd.read\_csv("general\_data.csv")

dataset.drop\_duplicates()

dataCleaned= dataset.dropna()

**Attrition and DistanceFromHome**

H0 : There’s no significant correlation between Attrition & DistanceFromHome

Ha : There’s significant correlation between Attrition & DistanceFromHome

from scipy.stats import pearsonr

stats, p=pearsonr(dataCleaned.Attrition, dataCleaned.DistanceFromHome)

print(stats,p)

-0.009448638515156248 0.5317715668019558

As r = -0.009, there’s low negative correlation between Attrition and DistanceFromHome

As the P value of 0.518 is > 0.05, we are accepting H0 and hence there’s no significant correlation between Attrition & DistanceFromHome

**Attrition and Age**

H0 : There’s no significant correlation between Attrition & Age

Ha : There’s significant correlation between Attrition & Age

stats, p=pearsonr(dataCleaned.Attrition, dataCleaned.Age)

print(stats,p)

-0.1583986795409671 5.1265982193780794e-26

As r = -0.158, there’s low negative correlation between Attrition and Age

As the P value of 5.1265982193780794e-26 < 0.05, we are accepting Ha and hence there’s significant correlation between Attrition & Age

**Attrition and MonthlyIncome**

H0 : There’s no significant correlation between Attrition & MonthlyIncome

Ha : There’s significant correlation between Attrition & MonthlyIncome

stats, p=pearsonr(dataCleaned.Attrition, dataCleaned.MonthlyIncome)

print(stats,p)

-0.030160293808460664 0.045890862744719166

As r = -0.031, there’s low negative correlation between Attrition and MonthlyIncome

As the P value of 0.045 is < 0.05, we are accepting Ha and hence there’s significant correlation between Attrition & MonthlyIncome

**Attrition and TotalWorkingYears**

H0 : There’s no significant correlation between Attrition & TotalWorkingYears

Ha : There’s significant correlation between Attrition & TotalWorkingYears

stats, p=pearsonr(dataCleaned.Attrition, dataCleaned.TotalWorkingYears)

print(stats,p)

-0.1696699168472392 1.1645434967091854e-29

As r = -0.169, there’s low negative correlation between Attrition and TotalWorkingYears

As the P value of 1. 1645434967091854e-29 is < 0.05, we are accepting Ha and hence there’s significant correlation between Attrition & TotalWorkingYears

**Attrition and YearsAtCompany**

H0 : There’s no significant correlation between Attrition & YearsAtCompany

Ha : There’s significant correlation between Attrition & YearsAtCompany

stats, p=pearsonr(dataCleaned.Attrition, dataCleaned.YearsAtCompany)

print(stats,p)

-0.1330026184252154 9.476118084836507e-19

As r = -0.133, there’s low negative correlation between Attrition and YearsAtCompany

As the P value of 9.476118084836507e-19 is < 0.05, we are accepting Ha and hence there’s significant correlation between Attrition & YearsAtCompany

**Attrition and Years With Current Manager**

H0 : There’s no significant correlation between Attrition & YearsWithCurrManager

Ha : There’s significant correlation between Attrition & YearsWithCurrManager

stats, p=pearsonr(dataCleaned.Attrition, dataCleaned.YearsWithCurrManager)

print(stats,p)

-0.15469153690287274 7.105369646771178e-25

As r = -0.154, there’s low negative correlation between Attrition and YearsWithCurrManager

As the P value of 7.105369646771178e-25 is < 0.05, we are accepting Ha and hence there’s significant correlation between Attrition & YearsWithCurrManager