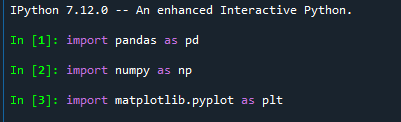
Day 7 Assignment

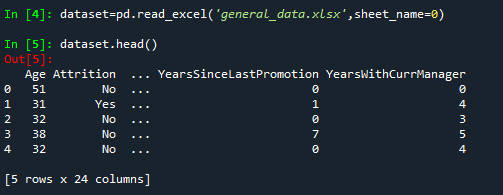
Step-1:Data Collecting

We will analyse the data and find the reasons for attrition in company

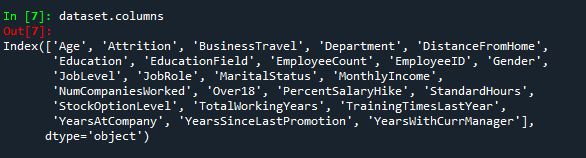
We will import numpy,pandas and matplotlib

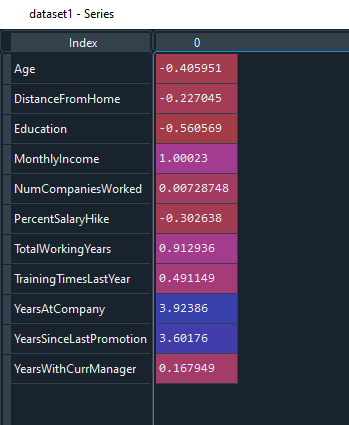
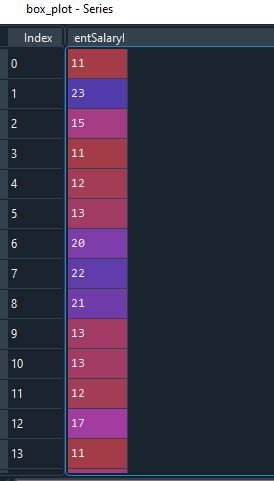


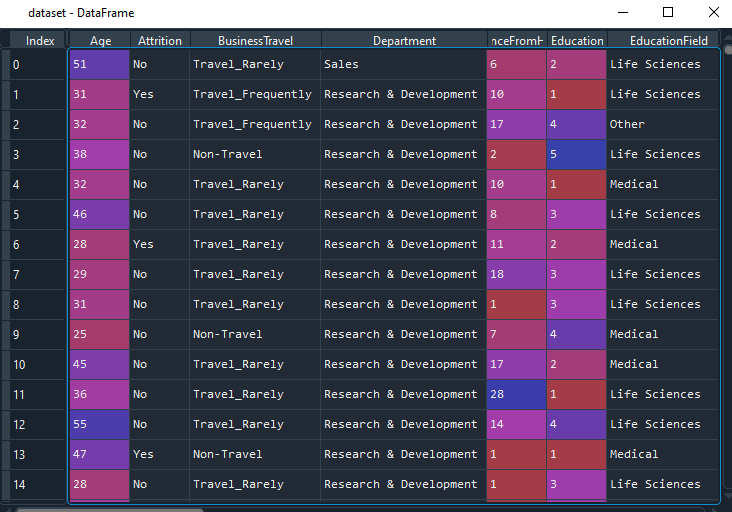
Now,declare a variable and store the file ‘general\_data.xlsx’



Dataset.columns will show the columns names

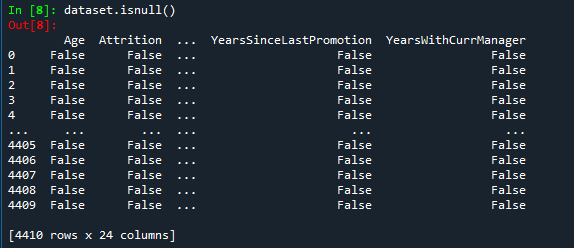




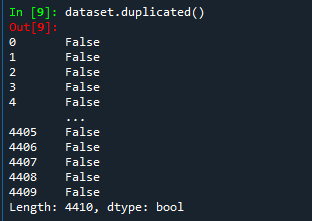
Step-2:Data Treatment

1.Checking for Null value



No Null Values Found

2.Checking For Duplicate Values

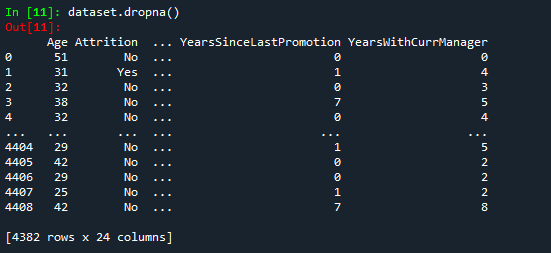


No Duplicate Value Found

If Duplicate values are found use dataset.drop\_duplicates()

To drop duplicate values.

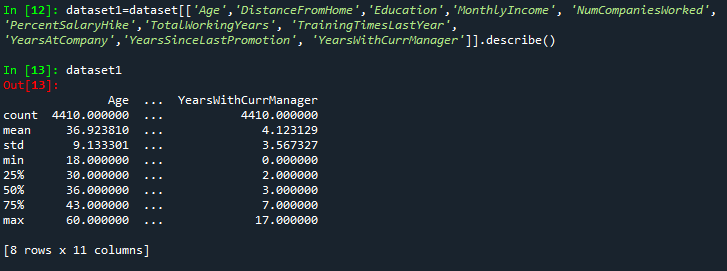
3.Checking for NA values



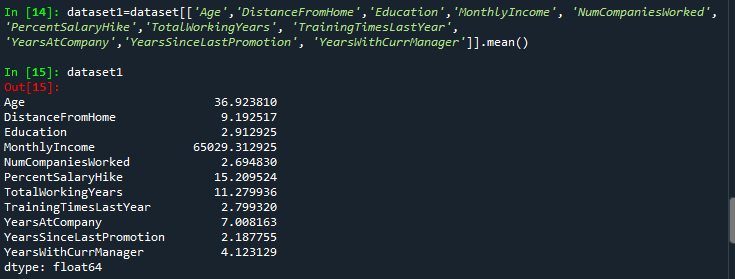
NA values are Dropped values

Step-3:Univariate Analysis

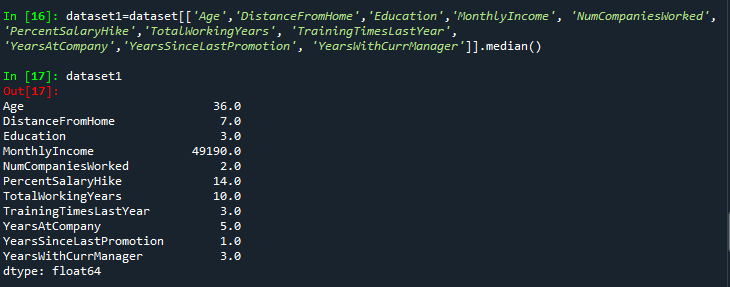
Getting the values of Descriptive Statistical Parameters



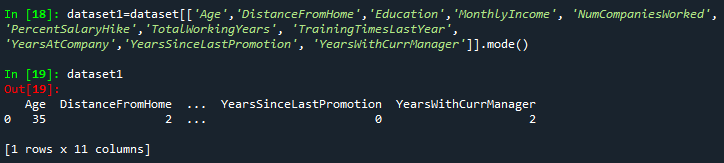
1.Mean



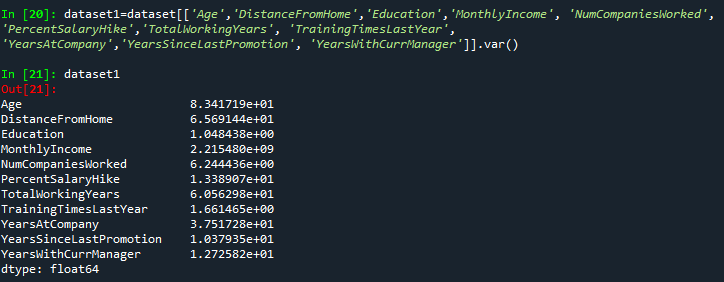
2.Median



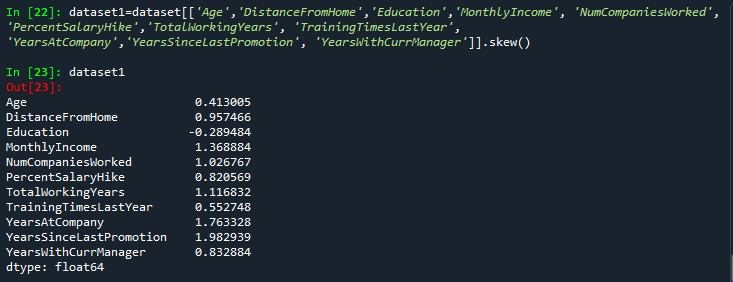
3.Mode



4.Variance

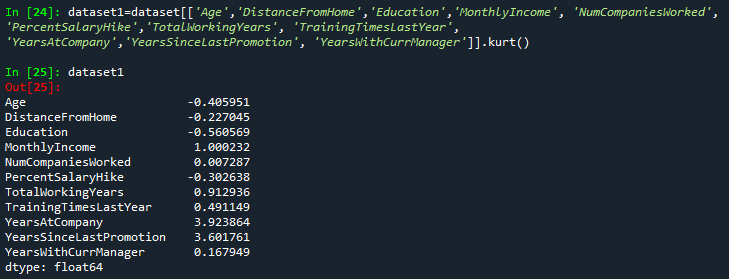


5.Skewness



Result:Education of Employeess is negatively skewed

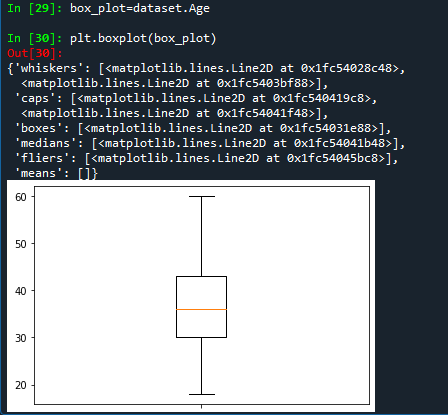
6.Kurtosis



Result:The outlier of Age,DistanceFromHome,Eductaion and PercentSalaryHike is less extreme.

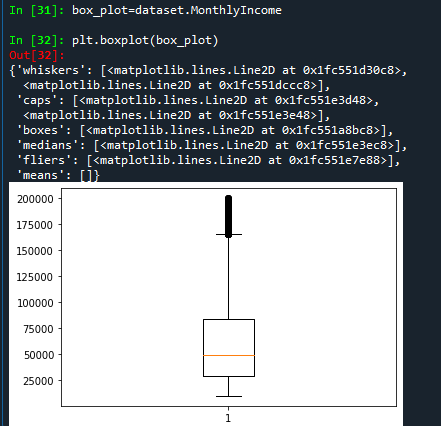
Oultiers:

Age:



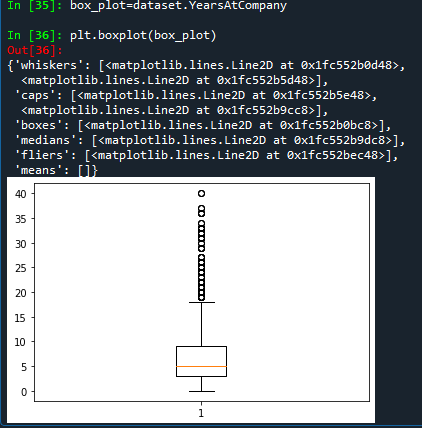
Age is Normally Distributed without any outliers

MonthlyIncome:



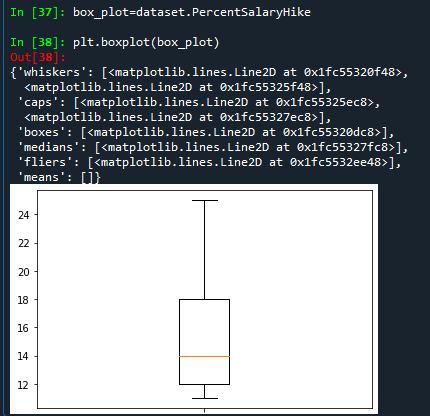
MonthlyIncome is right skewed with several outliers.

YearsAtCompany:



YearsAtCompany is also right skewed with several outliers observed

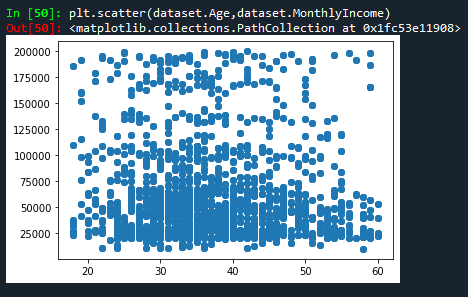
PercentSalaryHike:



PercentSalaryHike is normally distributed without any outliers.

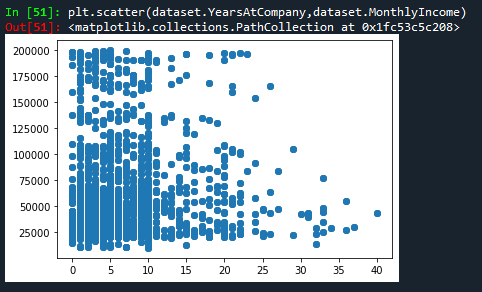
Scatter Plot:

1.Scatter plot between Age and MonthlyIncome



This Scatter Plot shows that the employees whose age is between28-38 having monthly income is between 125000-160000 are leaving company for better income package.

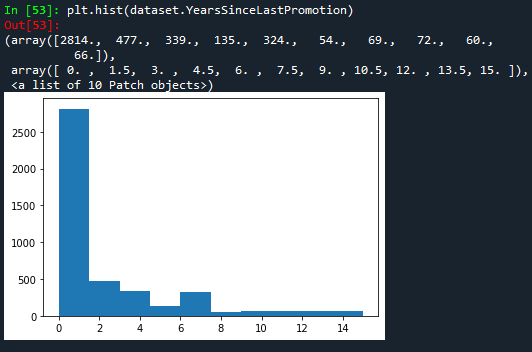
2.Scatter plot between YearsAtCompany and MonthlyIncome



This scatter plot ahows that employees who has worked for more than 25 years and still having MonthlyIncome between 25000-75000 are leaving the company

Histogram:

1.Histogram for YearsSinceLastPromotion



This Histogram shows positive skewness as peak is at left side

INFERENCE FROM ANALYSIS:

1.Analysis shows that Age and MeanDistanceFromHome are leptokurtic and all other variables are platykurtic.

2.meanMonthlyIncome is near 65k.

CONCLUSION:

1.The Employees whose age is between 28-38 are monthlyincome is between 125000-160000 are leaving company for more better package.

2.Also the Employees who are working for more than 25 years at company and still having monthly income 25000-75000 are leaving the company.

3.The Employees who are having monthly salary 175000-200000 are also leaving company for higher package.