1. **Importing necessary libraries**

**import** **pandas** **as** **pd**

**import** **matplotlib.pyplot** **as** **plt**

**import** **numpy** **as** **np**

**import** **seaborn** **as** **sns**

**from** **pandas** **import** DataFrame

1. **Loading Dataset**

**import** **os**

os.getcwd()

data1=pd.read\_excel("admission.xlsx")

1. **Data Exploration/Data Visualization**

data1.count()

data1.isnull ().sum ()

data1.describe()

corr= data1.corr()

data1.corr()

**Heatmap**

%matplotlib inline

**import** **seaborn** **as** **sns**

sns.heatmap(corr, vmin=-**1.0**, vmax=**1.0**, square=True, cmap="BuPu")

admit\_chance = admit\_chance.drop(['Chance of Admit ','Serial No.'])

admit\_chance

**pairplot**

**from** **pandas.plotting** **import** scatter\_matrix

a = scatter\_matrix(data1, figsize=(**8**, **8**))

**scatter blot to show correlation between different variables**

data1.plot.scatter("CGPA", "GRE Score")

data1.plot.scatter("CGPA", "SOP")

1. **Data Splitting**

x = data1[['TOEFL Score','CGPA','University Rating','GRE Score','Research']]

y = data1['Chance of Admit ']

**from** **sklearn.model\_selection** **import** train\_test\_split

X\_train, X\_test, y\_train, y\_test = train\_test\_split(x,y,test\_size=**0.25**)

x\_train.shape

x\_train.head()

1. **Running the model**

**from** **sklearn.linear\_model** **import** LinearRegression

lm = LinearRegression()

lm.fit(X\_train,y\_train)

LinearRegression()

y\_pred = lm.predict(X\_test)

lm.score(x,y)

**from** **sklearn.ensemble** **import** RandomForestRegressor

rf= RandomForestRegressor()

rf.fit (x\_train, y\_train)

y\_prdict2= rf.predict(x\_test)

score= metrics.r2\_score(y\_test, y\_prdict2)

**from** **sklearn.metrics** **import** mean\_absolute\_error

mae = mean\_absolute\_error(y\_pred, y\_test)

mae