import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline

# Importing DataSet and take a look at Data

Boston = pd.read\_csv("boston.csv")

Boston.head()

Boston.info()

Boston.describe()

Boston.plot.scatter('RM', 'MEDV');

X = Boston[Boston.columns[:-1]]

Y = Boston['MEDV']

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

from sklearn.linear\_model import LinearRegression

from sklearn.preprocessing import StandardScaler

# Split DataSet

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size=0.3)

sc\_X = StandardScaler()

X\_train\_ = sc\_X.fit\_transform(X\_train)

X\_test\_ = sc\_X.transform(X\_test)

print(f'Train Dataset Size - X: {X\_train.shape}, Y: {Y\_train.shape}')

print(f'Test Dataset Size - X: {X\_test.shape}, Y: {Y\_test.shape}')

# Model Building

lm = LinearRegression()

lm.fit(X\_train\_, Y\_train)

predictions = lm.predict(X\_test\_)

# Model Visualization

plt.scatter(Y\_test, predictions);

plt.xlabel('Y Test');

plt.ylabel('Predicted Y');

plt.title('Test vs Prediction');

sns.regplot(x = X\_test['RM'], y = predictions);

plt.scatter(X\_test['RM'], Y\_test, marker = '+');

plt.xlabel('Average number of rooms per dwelling');

plt.ylabel('Median value of owner-occupied homes');

plt.title('Regression Line Tracing');

from sklearn import metrics

print('Mean Absolute Error:', metrics.mean\_absolute\_error(Y\_test, predictions))

print('Mean Square Error:', metrics.mean\_squared\_error(Y\_test, predictions))

print('Root Mean Square Error:', np.sqrt(metrics.mean\_squared\_error(Y\_test, predictions)))