**ASSIGNMENT-4**

**FUNDAMENTAL OF MACHINE LEARNING**

**NAME- CHETHAN Y**

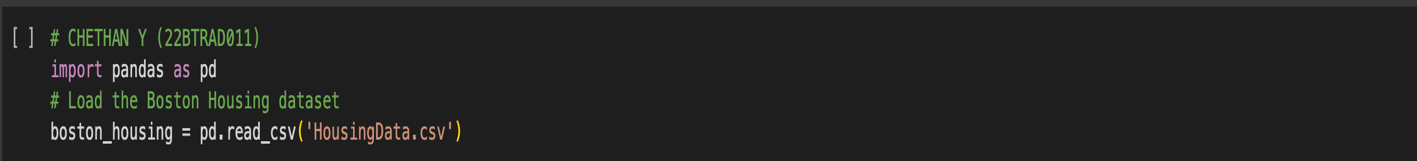
**USN-** 22BTRAD011

**QUES.** Load a dataset with outliers values (Boston Housing Dataset).

CODE:

import pandas as pd

# Load the Boston Housing dataset

boston\_housing = pd.read\_csv('HousingData.csv') 

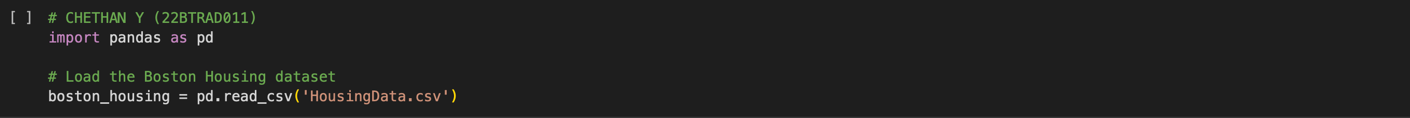
**QUES.** Implement one-hot encoding

CODE:

import pandas as pd

# Load the Boston Housing dataset

boston\_housing = pd.read\_csv('HousingData.csv')



**QUES.** Create visualizations for different aspects of a dataset using Matplotlib or Seaborn.

**CODE:**

import matplotlib.pyplot as plt

# Scatter plot of MEDV vs RM

plt.scatter(boston\_housing['RM'], boston\_housing['MEDV']) plt.xlabel('RM') plt.ylabel('MEDV')

plt.show()

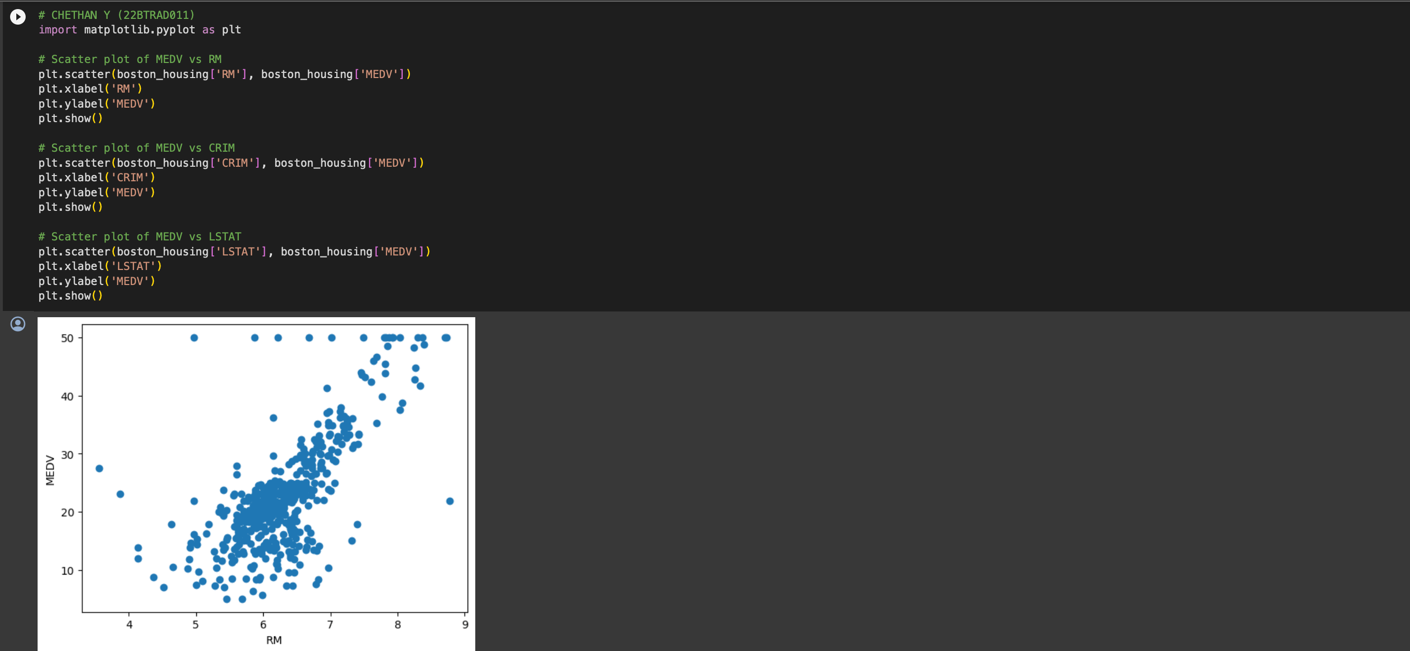
# Scatter plot of MEDV vs CRIM

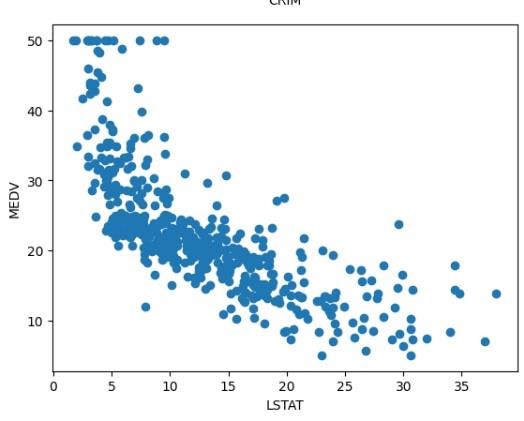
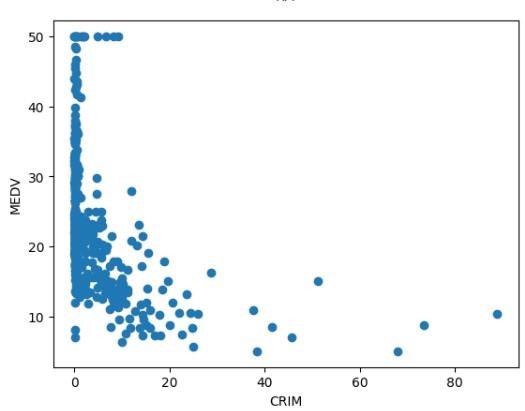
plt.scatter(boston\_housing['CRIM'], boston\_housing['MEDV']) plt.xlabel('CRIM') plt.ylabel('MEDV')

plt.show()

# Scatter plot of MEDV vs LSTAT

plt.scatter(boston\_housing['LSTAT'], boston\_housing['MEDV']) plt.xlabel('LSTAT') plt.ylabel('MEDV') plt.show()





**QUES.** Interpret the visualizations to gain insights into the dataset.

CODE:

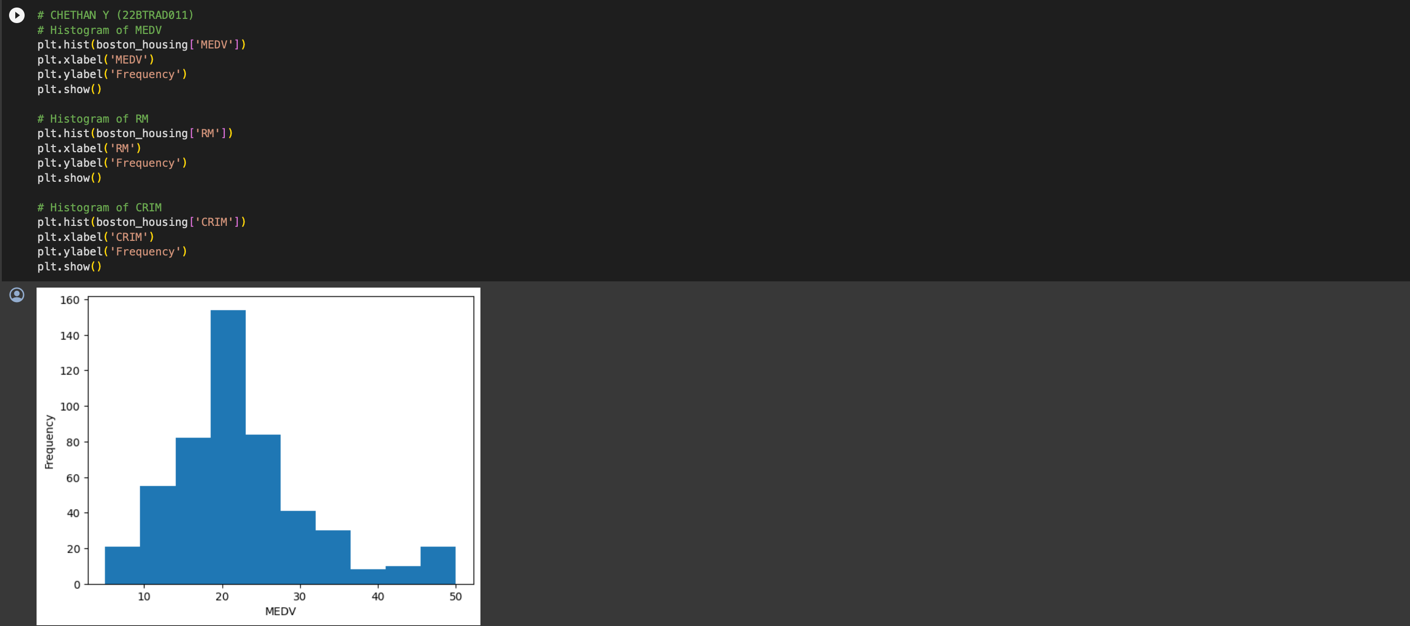
# Histogram of MEDV plt.hist(boston\_housing['MEDV']) plt.xlabel('MEDV') plt.ylabel('Frequency') plt.show()

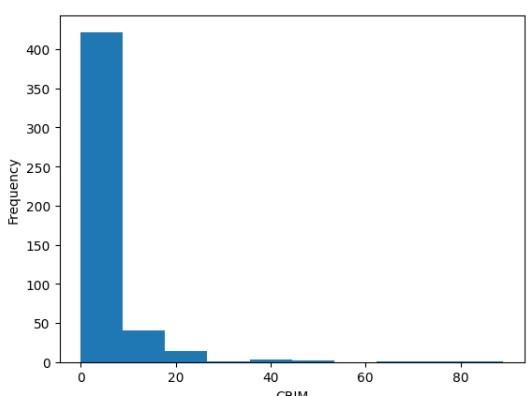
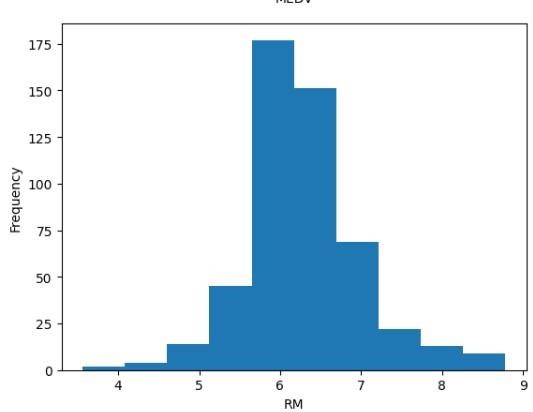
# Histogram of RM plt.hist(boston\_housing['RM'])

plt.xlabel('RM') plt.ylabel('Frequency') plt.show()

# Histogram of CRIM plt.hist(boston\_housing['CRIM'])

plt.xlabel('CRIM') plt.ylabel('Frequency') plt.show()





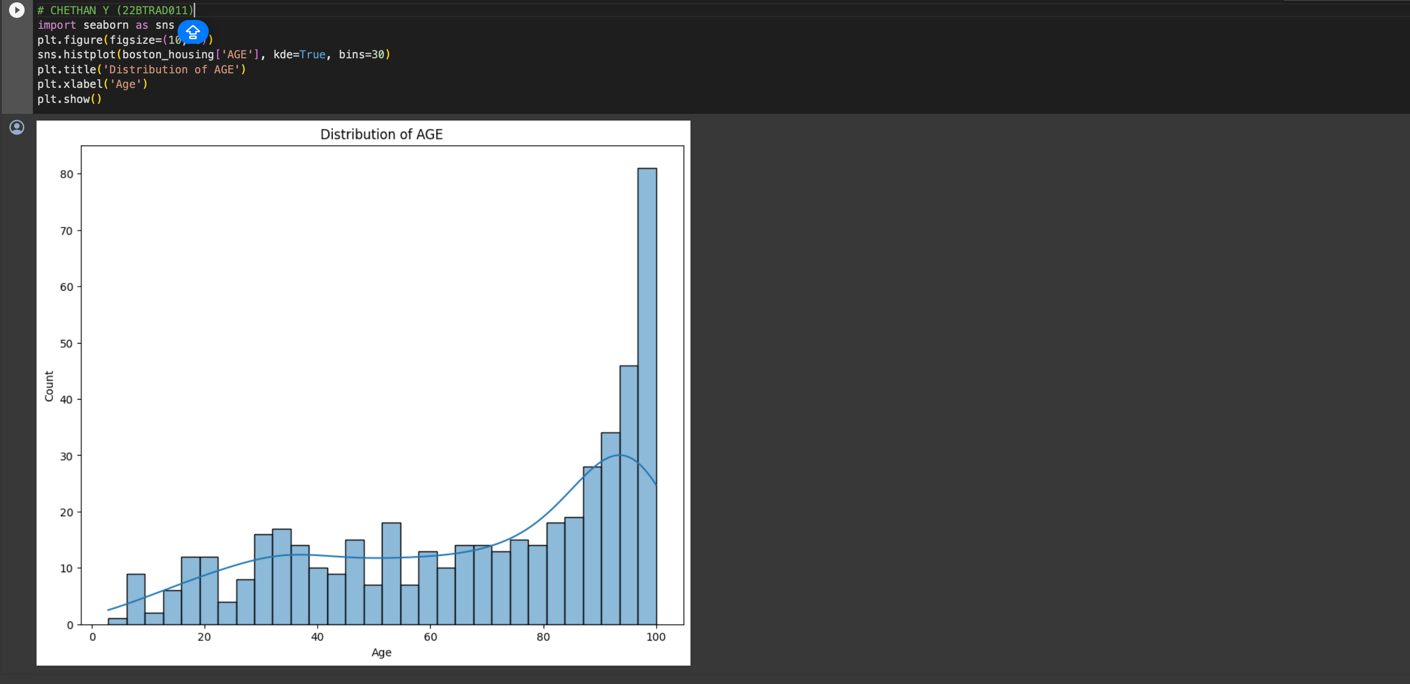
**QUES.** Perform Univariate and multivariate analysis for the dataset.

CODE:

import seaborn as sns plt.figure(figsize=(10, 8))

sns.histplot(boston\_housing['AGE'], kde=True, bins=30)

plt.title('Distribution of AGE') plt.xlabel('Age') plt.show()



GITHUB:

https://github.com/CHETHAN /MACHINE-LEARNING