**Develop a program to create histograms for all numerical features and analyze the distribution of each feature. Generate box plots for all numerical features and identify any outliers. Use California Housing dataset.**

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.datasets import fetch\_california\_housing

# Load California Housing Dataset

data = fetch\_california\_housing()

df = pd.DataFrame(data.data, columns=data.feature\_names)

df['MedHouseVal'] = data.target # Adding target variable

# Generate histograms for all numerical features

def plot\_histograms(df):

df.hist(figsize=(12, 10), bins=30, edgecolor='black')

plt.suptitle("Histograms of Numerical Features", fontsize=16)

plt.show()

# Generate box plots for outlier detection

def plot\_boxplots(df):

plt.figure(figsize=(12, 8))

for i, column in enumerate(df.columns, 1):

plt.subplot(3, 4, i)

sns.boxplot(y=df[column])

plt.title(column)

plt.tight\_layout()

plt.show()

# Identify outliers using IQR method

def detect\_outliers(df):

outliers = {}

for column in df.columns:

Q1 = df[column].quantile(0.25)

Q3 = df[column].quantile(0.75)

IQR = Q3 - Q1

lower\_bound = Q1 - 1.5 \* IQR

upper\_bound = Q3 + 1.5 \* IQR

outliers[column] = df[(df[column] < lower\_bound) | (df[column] > upper\_bound)][column].count()

return outliers

# Run analysis

plot\_histograms(df)

plot\_boxplots(df)

outlier\_counts = detect\_outliers(df)

print("Outlier counts per feature:")

print(outlier\_counts)