from sklearn import datasets

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load the Iris dataset from sklearn

iris = datasets.load\_iris()

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

df['species'] = iris.target

# Convert species numbers to names

df['species'] = df['species'].map({0: 'setosa', 1: 'versicolor', 2: 'virginica'})

# 1. List down the features and their types

print("1. Features and Their Data Types:")

print(df.dtypes, "\n")

# 2. Create histograms for each feature

print("2. Generating Histograms for Each Feature...")

df.hist(figsize=(10, 6), bins=20, edgecolor="black")

plt.suptitle("Histograms of Iris Dataset Features")

plt.show()

# 3. Create boxplots for each feature

print("3. Generating Boxplots for Each Feature...")

plt.figure(figsize=(10, 6))

sns.boxplot(data=df.drop(columns=["species"]))

plt.title("Boxplots of Iris Dataset Features")

plt.show()

# 4. Identify outliers using the IQR method

print("4. Identifying Outliers in Each Feature Using Boxplots...")

outliers = {}

for column in df.drop(columns=["species"]).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].tolist()

# Print detected outliers

for feature, values in outliers.items():

if values:

print(f"Outliers detected in {feature}: {values}")

else:

print(f"No significant outliers detected in {feature}")