**ASSIGNMENT 9**

**Source Code:**

import pandas as pd

importnumpy as np

importseaborn as sns

importmatplotlib.pyplot as plt

#read the file

temp = pd.read\_csv('/home/IRIS.csv')

#describe all the data

print("The type of Data")

print(type(temp))

print("=========================================================================")

print("Data Statistical Description\n")

print(temp.describe())

print("=========================================================================")

print("Data Type Description\n")

print(temp.dtypes)

print("=========================================================================")

print("Grouping the Data by Species\n")

print(temp.groupby('species').size())

print("=========================================================================")

print("The first 5 rows of the Data Frame\n")

print(temp.head())

#DATA VISUALISATION

sns.set()

#histogram plots olfinidividual features

plt.hist(temp['sepal\_length'])

plt.xlabel('sepal-length')

plt.show()

#histrogram plots of individual features

plt.hist(temp['sepal\_width'])

plt.xlabel('sepal-width')

plt.show()

#histrogram plots of individual features

plt.hist(temp['petal\_length'])

plt.xlabel('petal-length')

plt.show()

#histrogram plots of individual features

plt.hist(temp['petal\_width'])

plt.xlabel('petal-width')

plt.show()

#grouped boxplots according to features

sns.boxplot(x=temp['sepal\_length'],y=temp['species'])

plt.show()

sns.boxplot(x=temp['sepal\_width'],y=temp['species'])

plt.show()

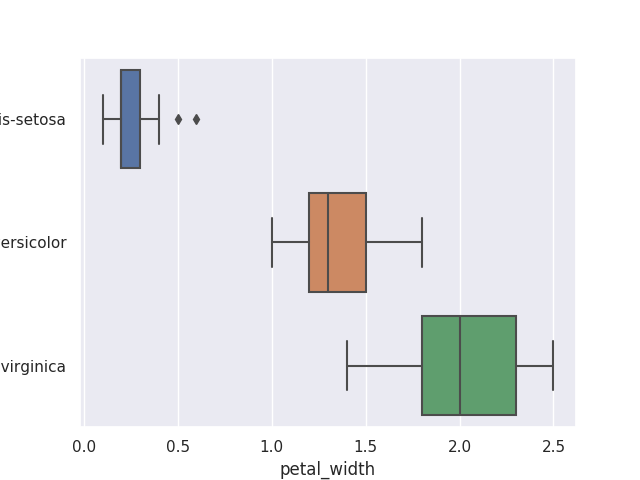
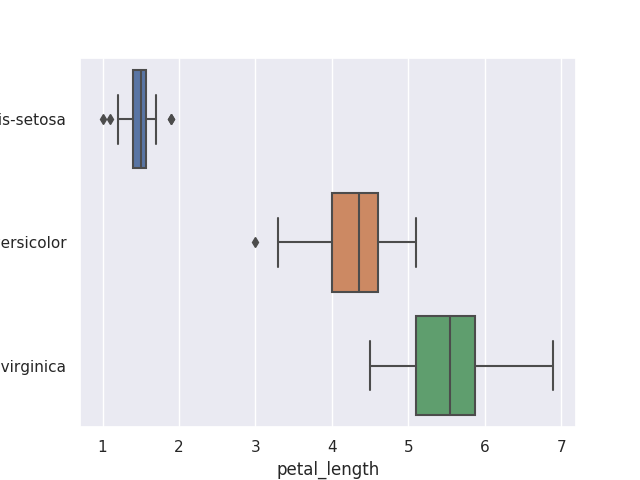
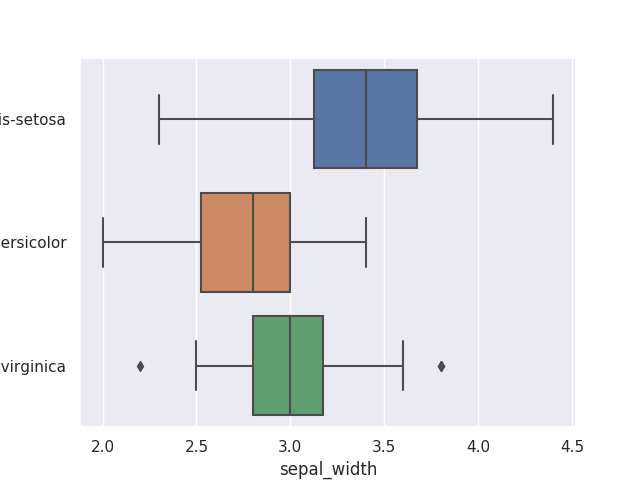
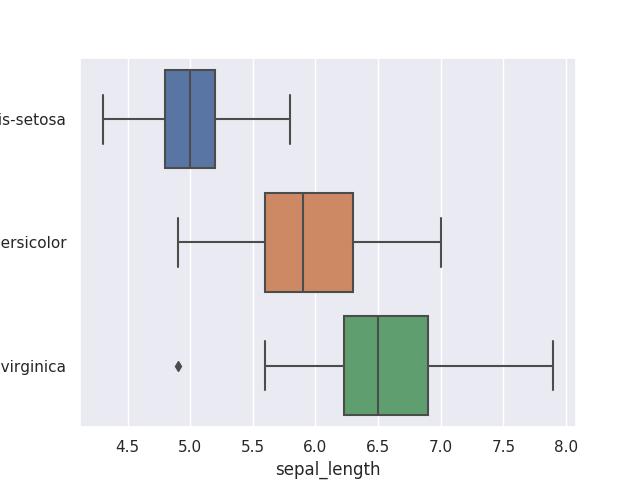
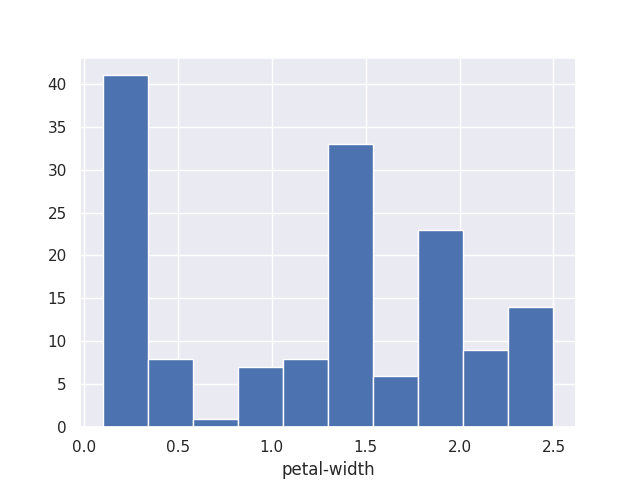
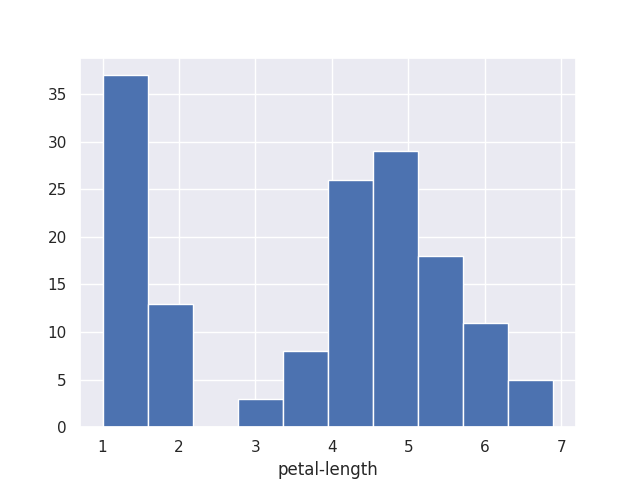
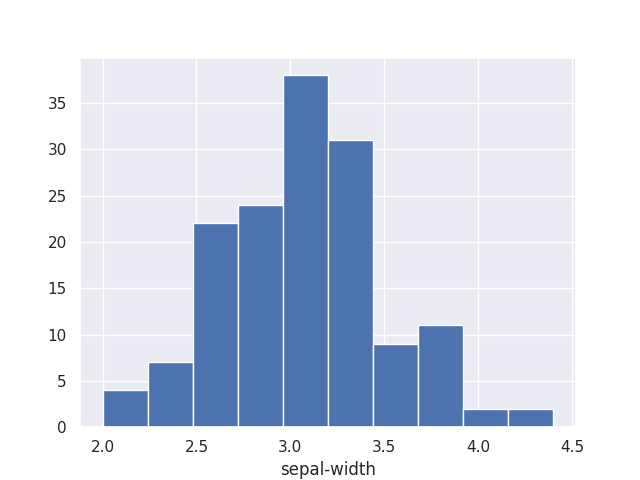
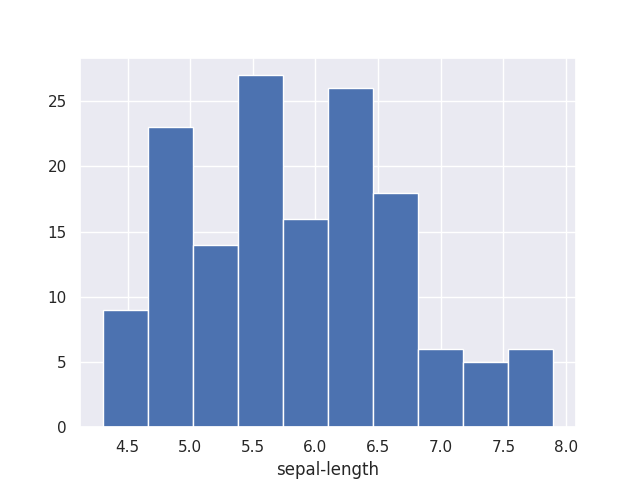
sns.boxplot(x=temp['petal\_length'],y=temp['species'])

plt.show()

sns.boxplot(x=temp['petal\_width'],y=temp['species'])

plt.show()

**Output:**

****

The type of Data

<class 'pandas.core.frame.DataFrame'>

=========================================================================

Data Statistical Description

sepal\_lengthsepal\_widthpetal\_lengthpetal\_width

count 150.000000 150.000000 150.000000 150.000000

mean 5.843333 3.054000 3.758667 1.198667

std 0.828066 0.433594 1.764420 0.763161

min 4.300000 2.000000 1.000000 0.100000

25% 5.100000 2.800000 1.600000 0.300000

50% 5.800000 3.000000 4.350000 1.300000

75% 6.400000 3.300000 5.100000 1.800000

max 7.900000 4.400000 6.900000 2.500000

=========================================================================

Data Type Description

sepal\_length float64

sepal\_width float64

petal\_length float64

petal\_width float64

species object

dtype: object

=========================================================================

Grouping the Data by Species

species

Iris-setosa 50

Iris-versicolor 50

Iris-virginica 50

dtype: int64

=========================================================================

The first 5 rows of the Data Frame

sepal\_lengthsepal\_width ... petal\_width species

0 5.1 3.5 ... 0.2 Iris-setosa

1 4.9 3.0 ... 0.2 Iris-setosa

2 4.7 3.2 ... 0.2 Iris-setosa

3 4.6 3.1 ... 0.2 Iris-setosa

4 5.0 3.6 ... 0.2 Iris-setosa

[5 rows x 5 columns]