**Assignment No 10**

**Code:**

# Assignment number 10:

# Load the Iris dataset

# Display basic information

# Display statistical information

# Display null values

# Fill the null values

# Display and iterpret Boxplot & Histogram

# Identify outliers

#---------------------------------------------------------------------------------------

# Importing libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

#---------------------------------------------------------------------------------------

# Reading dataset

df = pd.read\_csv('iris.csv')

df = df.drop('Id', axis=1)

df.columns = ('SL', 'SW', 'PL', 'PW', 'Species')

#---------------------------------------------------------------------------------------

# Display basic information

print('Information of Dataset:\n', df.info)

print('Shape of Dataset (row x column): ', df.shape)

print('Columns Name: ', df.columns)

print('Total elements in dataset:', df.size)

print('Datatype of attributes (columns):', df.dtypes)

print('First 5 rows:\n', df.head().T)

print('Last 5 rows:\n',df.tail().T)

print('Any 5 rows:\n',df.sample(5).T)

#---------------------------------------------------------------------------------------

# Display Statistical information

print('Statistical information of Numerical Columns: \n',df.describe())

#---------------------------------------------------------------------------------------

# Display and fill the Null values

print('Total Number of Null Values in Dataset:', df.isna().sum())

#---------------------------------------------------------------------------------------

fig, axis = plt.subplots(2,2)

sns.boxplot(ax = axis[0,0], data = df, y='SL')

sns.boxplot(ax = axis[0,1], data = df, y='SW')

sns.boxplot(ax = axis[1,0], data = df, y='PL')

sns.boxplot(ax = axis[1,1], data = df, y='PW')

plt.show()

fig, axis = plt.subplots(2,2)

sns.boxplot(ax = axis[0,0], data = df, y='SL', hue='Species')

sns.boxplot(ax = axis[0,1], data = df, y='SW', hue='Species')

sns.boxplot(ax = axis[1,0], data = df, y='PL', hue='Species')

sns.boxplot(ax = axis[1,1], data = df, y='PW', hue='Species')

plt.show()

fig, axis = plt.subplots(2,2)

sns.histplot(ax = axis[0,0], data = df, x='SL', multiple = 'dodge', shrink = 0.8, kde = True)

sns.histplot(ax = axis[0,1], data = df, x='SW', multiple = 'dodge', shrink = 0.8, kde = True)

sns.histplot(ax = axis[1,0], data = df, x='PL', multiple = 'dodge', shrink = 0.8, kde = True)

sns.histplot(ax = axis[1,1], data = df, x='PW', multiple = 'dodge', shrink = 0.8, kde = True)

plt.show()

fig, axis = plt.subplots(2,2)

sns.histplot(ax=axis[0,0], data=df, x='SL', hue='Species', element='poly', shrink=0.8, kde= True)

sns.histplot(ax=axis[0,1], data = df, x='SW', hue = 'Species', element = 'poly', shrink = 0.8, kde = True)

sns.histplot(ax=axis[1,0], data = df, x='PL', hue = 'Species', element = 'poly', shrink = 0.8, kde = True)

sns.histplot(ax=axis[1,1], data = df, x='PW', hue = 'Species', element = 'poly', shrink = 0.8, kde = True)

plt.show()

**Output:**

nformation of Dataset:

 <bound method DataFrame.info of       SL   SW   PL   PW    Species

0    5.1  3.5  1.4  0.2     Setosa

1    4.9  3.0  1.4  0.2     Setosa

2    4.7  3.2  1.3  0.2     Setosa

3    4.6  3.1  1.5  0.2     Setosa

4    5.0  3.6  1.4  0.2     Setosa

..   ...  ...  ...  ...        ...

145  6.7  3.0  5.2  2.3  Virginica

146  6.3  2.5  5.0  1.9  Virginica

147  6.5  3.0  5.2  2.0  Virginica

148  6.2  3.4  5.4  2.3  Virginica

149  5.9  3.0  5.1  1.8  Virginica

[150 rows x 5 columns]>

Shape of Dataset (row x column):  (150, 5)

Columns Name:  Index(['SL', 'SW', 'PL', 'PW', 'Species'], dtype='object')

Total elements in dataset: 750

Datatype of attributes (columns): SL         float64

SW         float64

PL         float64

PW         float64

Species     object

dtype: object

First 5 rows:

               0       1       2       3       4

SL          5.1     4.9     4.7     4.6     5.0

SW          3.5     3.0     3.2     3.1     3.6

PL          1.4     1.4     1.3     1.5     1.4

PW          0.2     0.2     0.2     0.2     0.2

Species  Setosa  Setosa  Setosa  Setosa  Setosa

Last 5 rows:

                145        146        147        148        149

SL             6.7        6.3        6.5        6.2        5.9

SW             3.0        2.5        3.0        3.4        3.0

PL             5.2        5.0        5.2        5.4        5.1

PW             2.3        1.9        2.0        2.3        1.8

Species  Virginica  Virginica  Virginica  Virginica  Virginica

Any 5 rows:

                 54          87         137         63         149

SL              6.5         6.3        6.4         6.1        5.9

SW              2.8         2.3        3.1         2.9        3.0

PL              4.6         4.4        5.5         4.7        5.1

PW              1.5         1.3        1.8         1.4        1.8

Species  Versicolor  Versicolor  Virginica  Versicolor  Virginica

Statistical information of Numerical Columns:

                SL          SW          PL          PW

count  148.000000  150.000000  150.000000  148.000000

mean     5.856757    3.054000    3.758667    1.197973

std      0.825459    0.433594    1.764420    0.760278

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

Total Number of Null Values in Dataset: SL         2

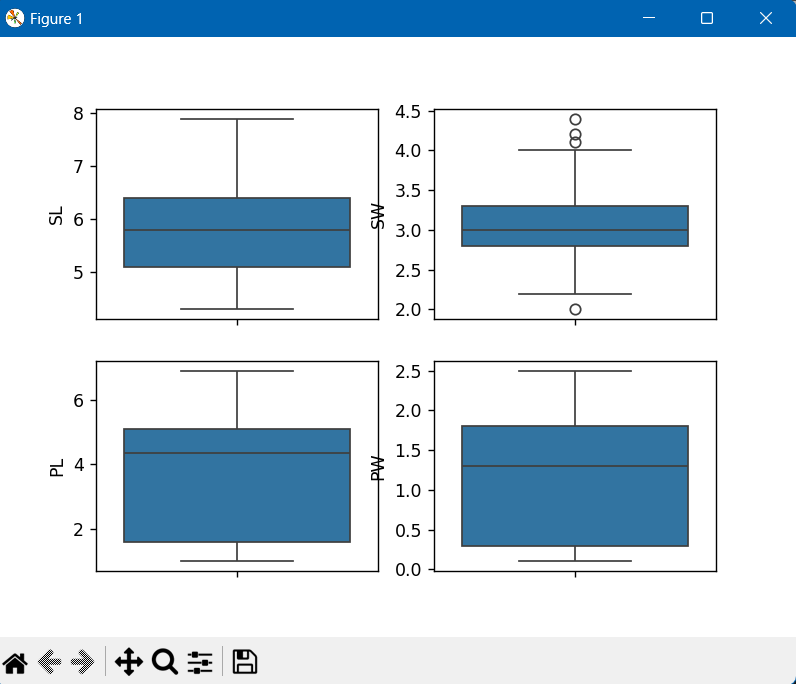
SW         0

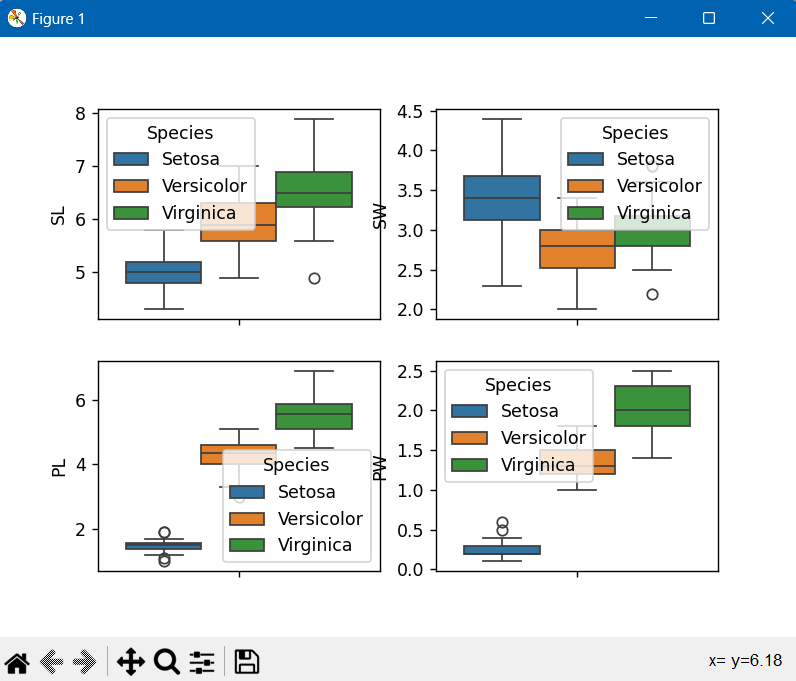
PL         0

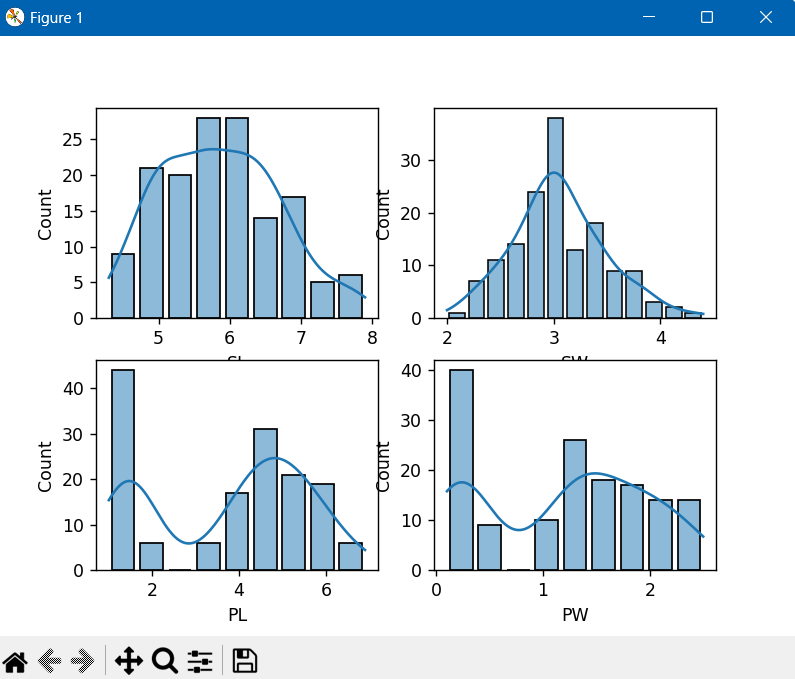
PW         2

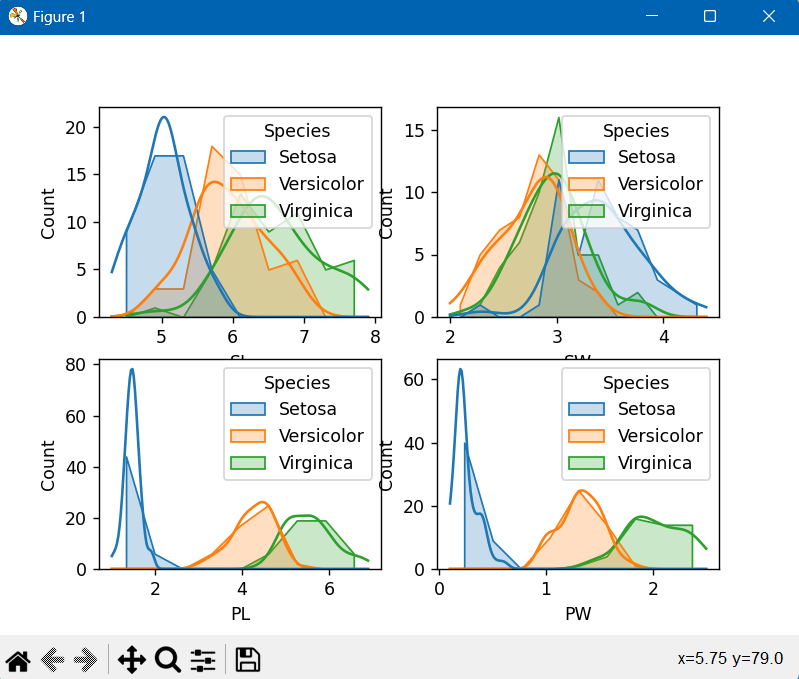
Species    0

dtype: int64

****

****

****

****