D**ATA SCIENCE & MACHINE LEARNING**

**LAB CYCLE 3.2**

# DATA HANDLING USING ‘Pandas’ and DATA VISUALIZATION USING ‘Seaborn’

# Using the pandas function read\_csv(), read the given ‘iris’ data set.

**1. Use appropriate functions in pandas to display**

**(i) Shape of the data set**

**(ii) First 5 and last five rows of data set(head and tail)**

**(iii) Size of dataset**

**(iv) No:of samples available for each variety**

**(v) Description of the data set( use describe**

**Program**

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

import pandas as pd

col=['sepal.length','sepal.width','petal.length','petal.width','variety']

iris=pd.read\_csv("iris.csv",names=col)

print(col)

#Shape of the data set

print("\n Shape \n")

shape = iris.shape

print(shape)

#First 5 and last five rows of data set(head and tail)

print("\n Head \n",iris.head())

print("\n Tail \n",iris.tail())

# Size of dataset

print("\nSize\n",iris.size)

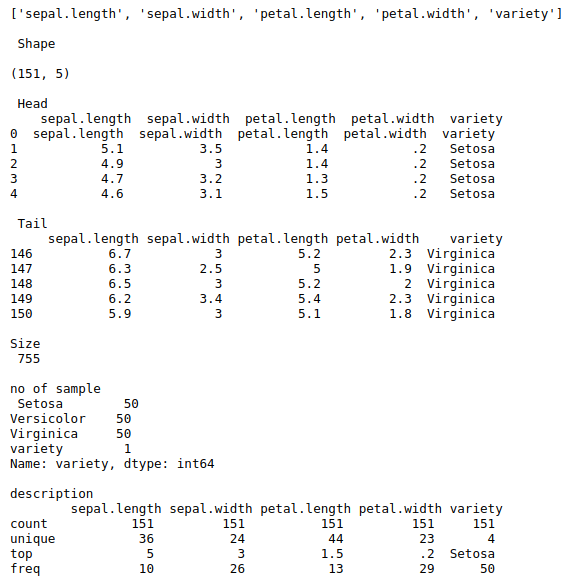
# No:of samples available for each variety

print("\nno of sample\n", iris["variety"].value\_counts())

# Description of the data set( use describe

print("\ndescription\n", iris.describe())

**Output**



**2. Use pairplot() function to display pairwise relationships between attributes. Try different kind of plots {‘scatter’, ‘kde’, ‘hist’, ‘reg’} and different kind of markers**

**Program**

import numpy as np

import pandas as pd

import seaborn as sns

col=['sepal.length','sepal.width','petal.length','petal.width','variety']

iris = sns.load\_dataset("iris")

my\_data\_frame = pd.DataFrame(iris)

g = sns.pairplot(my\_data\_frame, hue = "species")

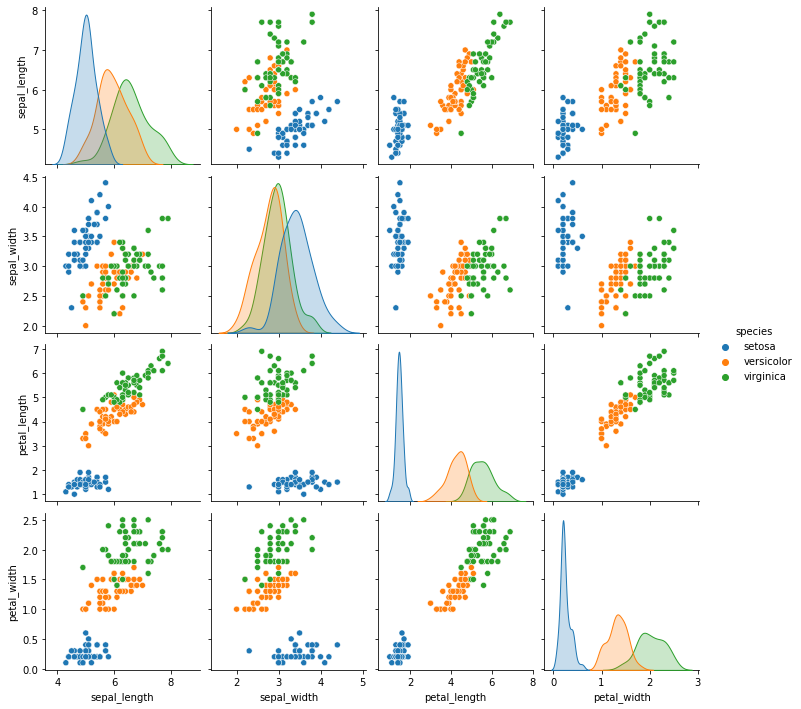
g3 = sns.pairplot(my\_data\_frame, kind="scatter", hue = "species", markers=["o", "s", "D"])

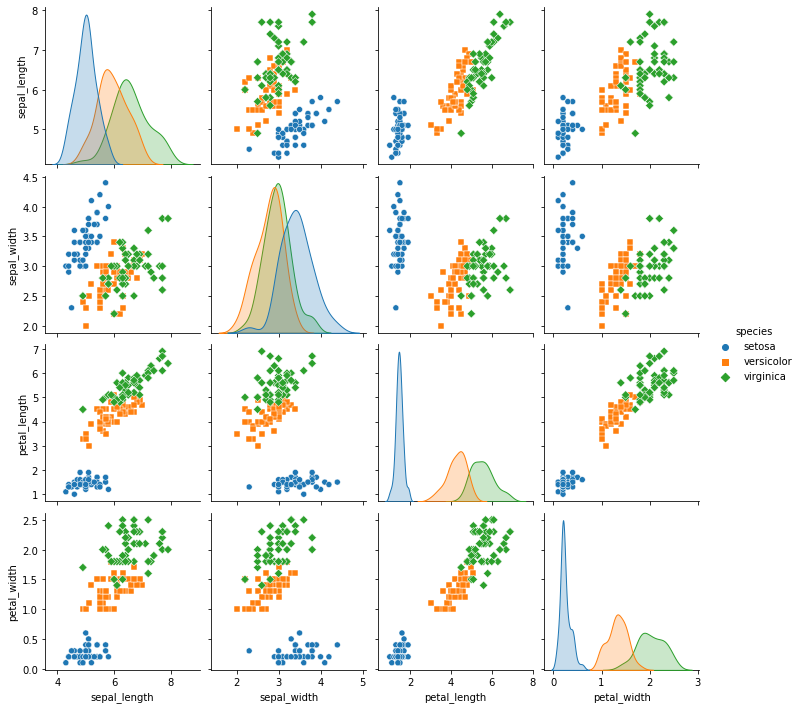
g2 = sns.pairplot(my\_data\_frame, kind="kde", hue = "species")

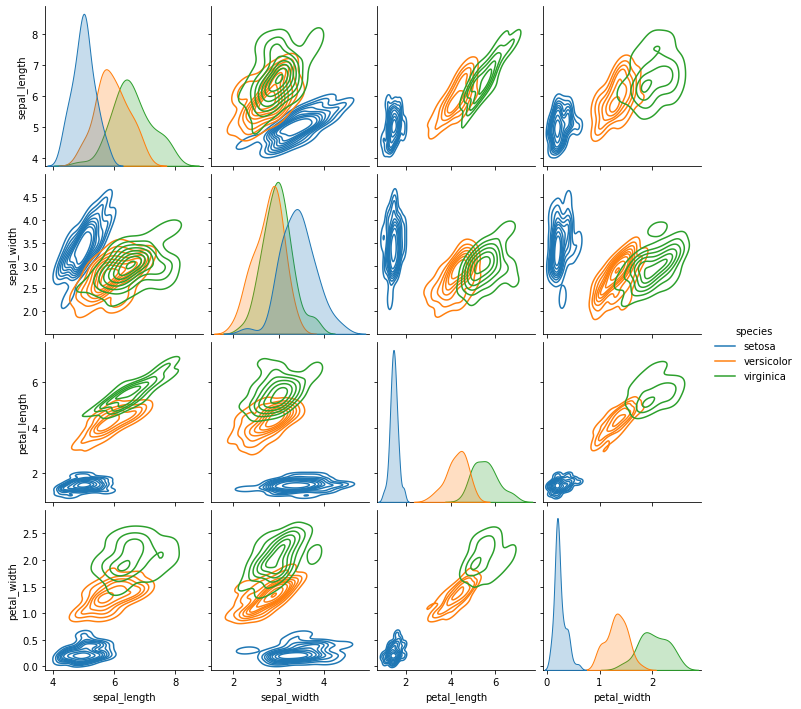
g4 = sns.pairplot(my\_data\_frame, kind="hist", hue = "species")

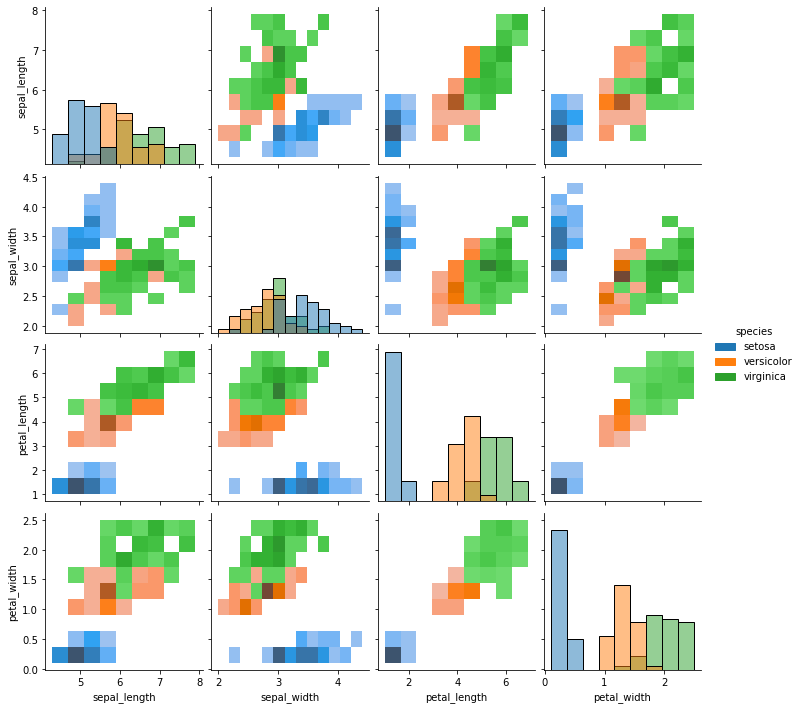
g5 = sns.pairplot(my\_data\_frame, kind="reg", hue = "species", markers=["X", "H", "<"])

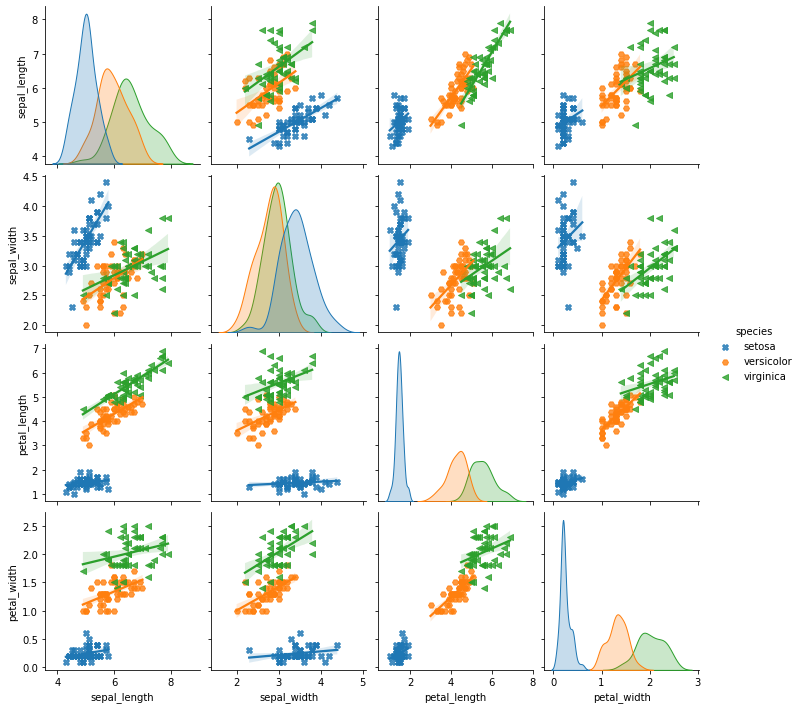
**Output**







****

****

**3. using the iris data set,get familiarize with functions:**

**1)displot()**

**2) histplot()**

**3) relplot()**

**Program & Output**

import numpy as np

import pandas as pd

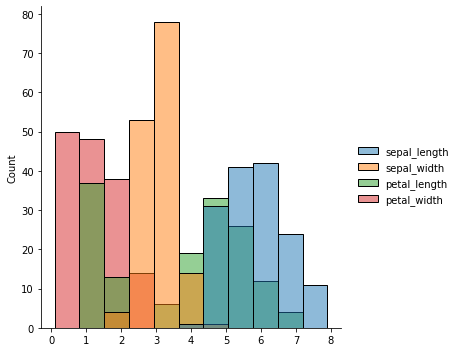
import matplotlib.pyplot as plt

import seaborn as sns

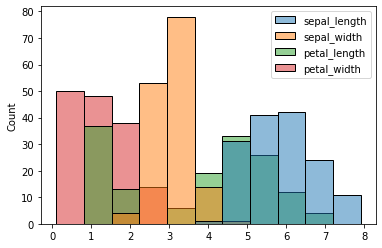
iris = sns.load\_dataset("iris")

my\_data\_frame = pd.DataFrame(iris)

displot = sns.displot(data = iris)

****

histplot = sns.histplot(data = iris)



relplot = sns.relplot(data = iris)

