**CYCLE 3-PART 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

1.Shape of the data set

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

3.Size of dataset

4.No:of samples available for each variety

5.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','type']

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

print("Shape of the data set")

print("shape:",iris.shape)

print("First five and last five rows of data set")

print("First 5 rows")

print(iris.head())

print("Last 5 rows")

print(iris.tail())

print("Size of the Dataset")

print("Size:",iris.size)

print("No of samples available for each type")

print(iris["type"].value\_counts())

print("Description of the data set")

print(iris.describe())

**OUTPUT**

**Shape of the data set**

**shape**: (150, 5)

**First five and last five rows of data set**

**First 5 rows**

sepal\_length sepal\_width petal\_length petal\_width type

0 sepal.length sepal.width petal.length petal.width variety

1 5.1 3.5 1.4 .2 Setosa

2 4.9 3 1.4 .2 Setosa

3 4.7 3.2 1.3 .2 Setosa

4 4.6 3.1 1.5 .2 Setosa

**Last 5 rows**

sepal\_length sepal\_width petal\_length petal\_width type

146 6.7 3 5.2 2.3 Virginica

147 6.3 2.5 5 1.9 Virginica

148 6.5 3 5.2 2 Virginica

149 6.2 3.4 5.4 2.3 Virginica

150 5.9 3 5.1 1.8 Virginica

**Size of the Dataset**

Size: 755

**No of samples available for each type**

Setosa 50

Versicolor 50

Virginica 50

variety 1

Name: type, dtype: int64

**Description of the data set**

sepal\_length sepal\_width petal\_length petal\_width type

count 151 151 151 151 151

unique 36 24 44 23 4

top 5 3 1.5 .2 Setosa

freq 10 26 13 29 50

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

**PROGRAM**

matplotlib inline

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

sns.pairplot(iris)

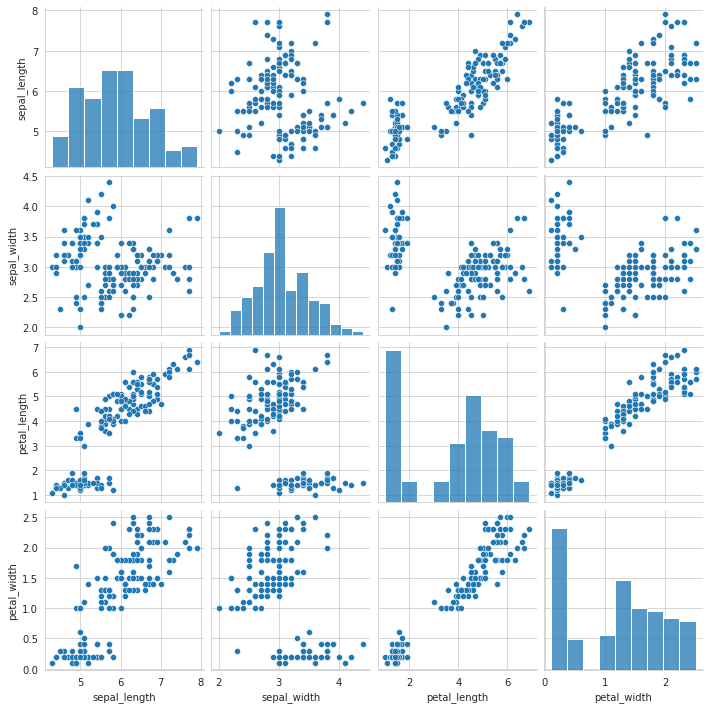
sns.pairplot(iris,kind="scatter")

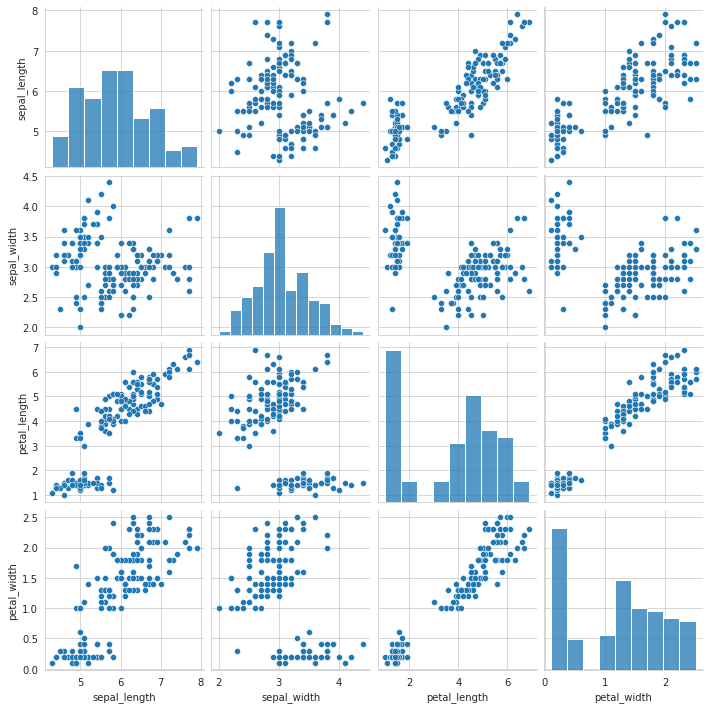
sns.pairplot(iris,kind="kde")

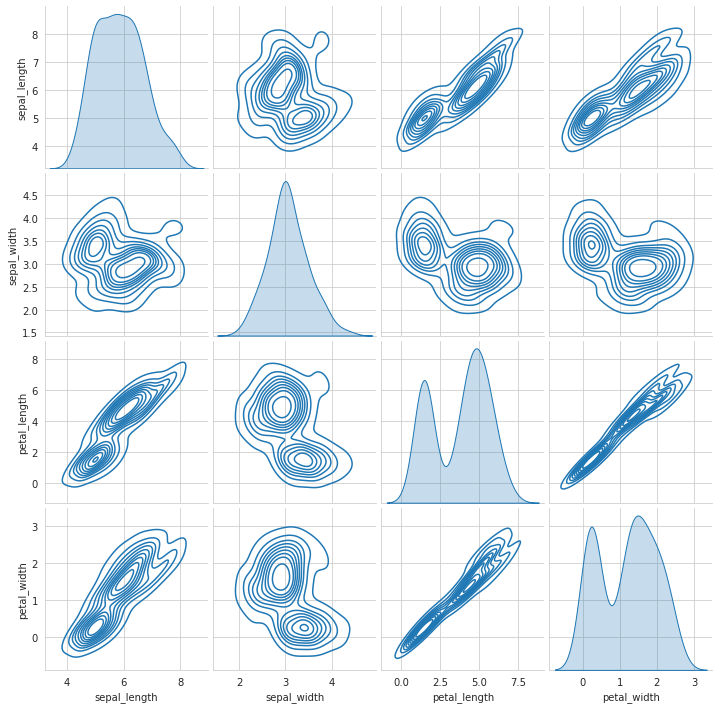
sns.pairplot(iris,kind="hist")

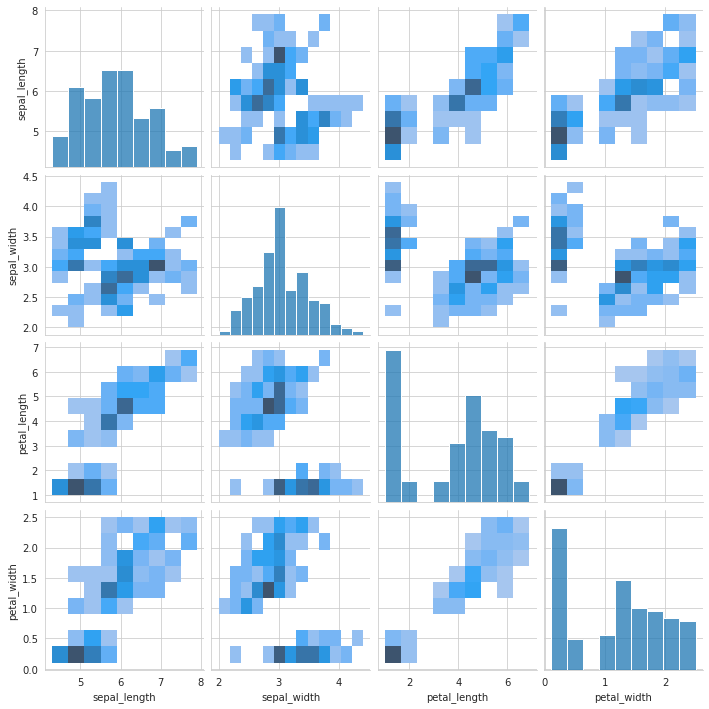
sns.pairplot(iris,kind="reg")

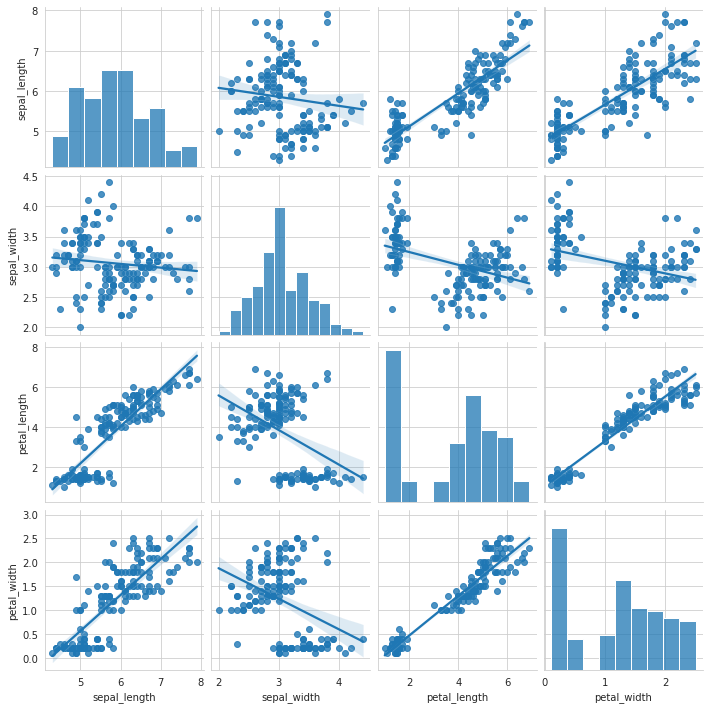
**OUTPUT**

****

****

****

****

****

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

**1)displot()**

**2) histplot()**

**3) relplot()**

**PROGRAM**

sns.displot(iris)

sns.histplot(iris)

sns.relplot(data=iris,kind="line")

plt.style.use("dark\_background")

sns.displot(iris.sepal\_length,bins=20,color="g")

plt.title("Distribution\_sepal\_length",fontsize=20,color="red")

plt.show()

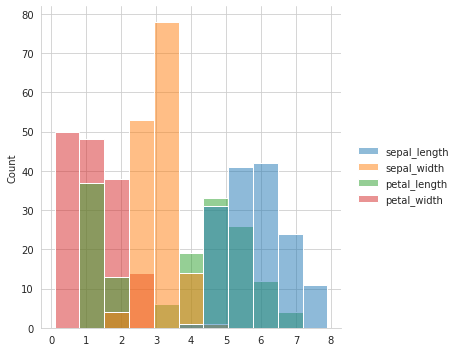
plt.style.use("dark\_background")

sns.displot(iris.sepal\_length,bins=20,color="g")

plt.title("Distribution\_sepal\_width",fontsize=20,color="red")

plt.show()

**OUTPUT**

****

