

Q1. Data Visualization and Statistical Measures:

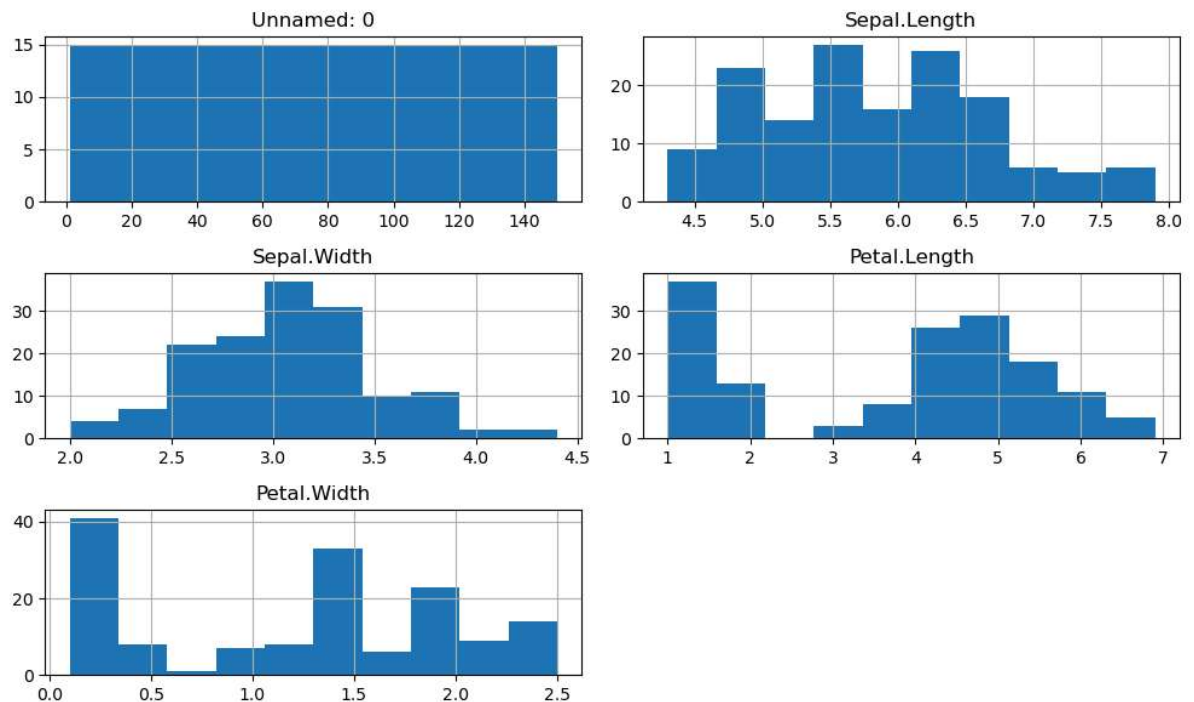
For this question, you are required to analyse the iris dataset (iris.csv) using Python. Perform all possible data visualization techniques (histograms, scatter plots, box plots, etc.) on all numerical columns of the dataset. Additionally, calculate all possible statistical measures (mean, median, mode, standard deviation, etc.) for each numerical column.

```
In [2]: 1 import pandas as pd
        2 import matplotlib.pyplot as plt
        3 import seaborn as sns
        4 from scipy import stats
        5 data = pd.read_csv(r"C:\Users\sneha\Downloads\iris1.csv")
```

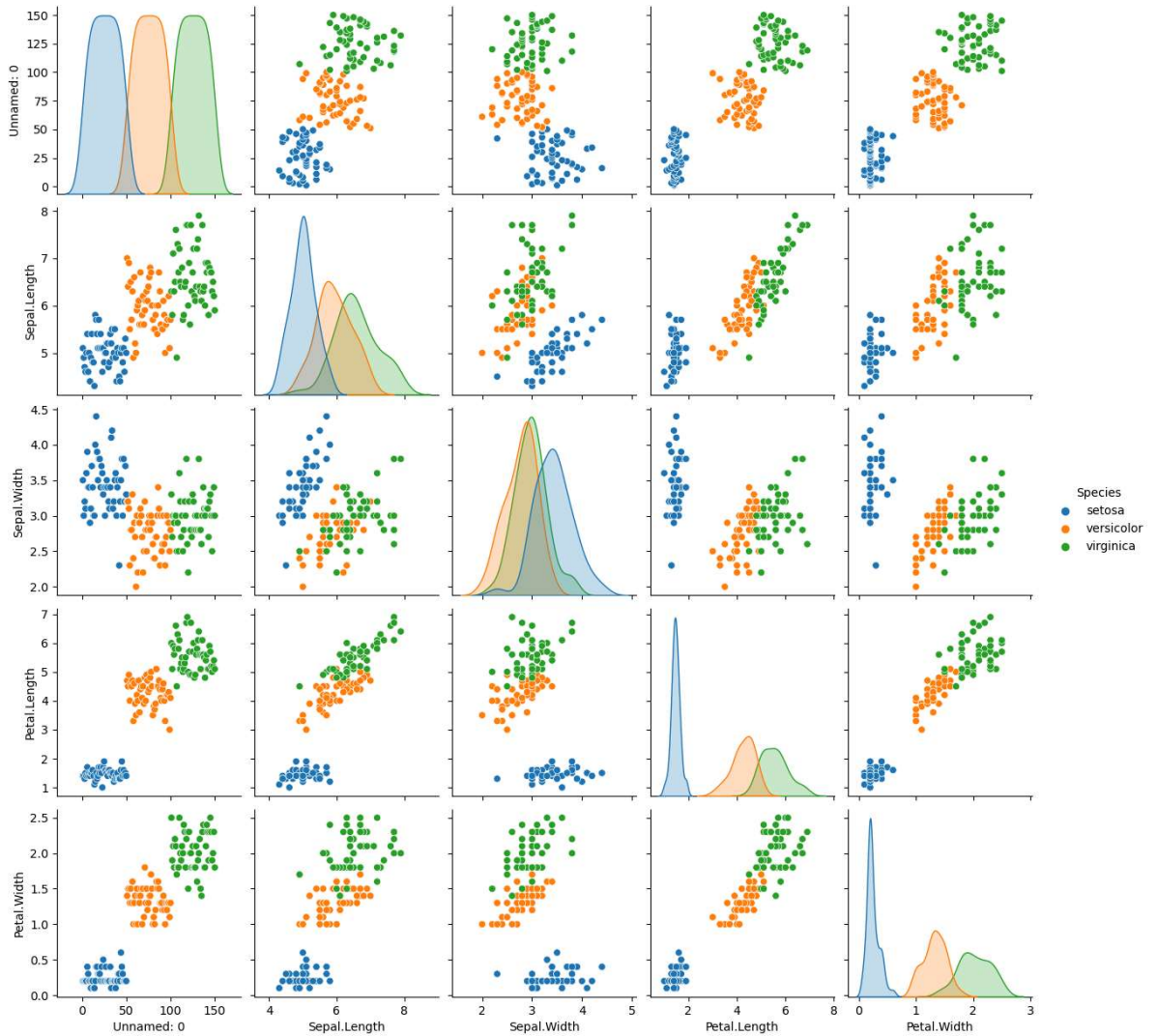
```
In [3]: 1 print(data.head())
```

	Unnamed: 0	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
0	1	5.1	3.5	1.4	0.2	setosa
1	2	4.9	3.0	1.4	0.2	setosa
2	3	4.7	3.2	1.3	0.2	setosa
3	4	4.6	3.1	1.5	0.2	setosa
4	5	5.0	3.6	1.4	0.2	setosa

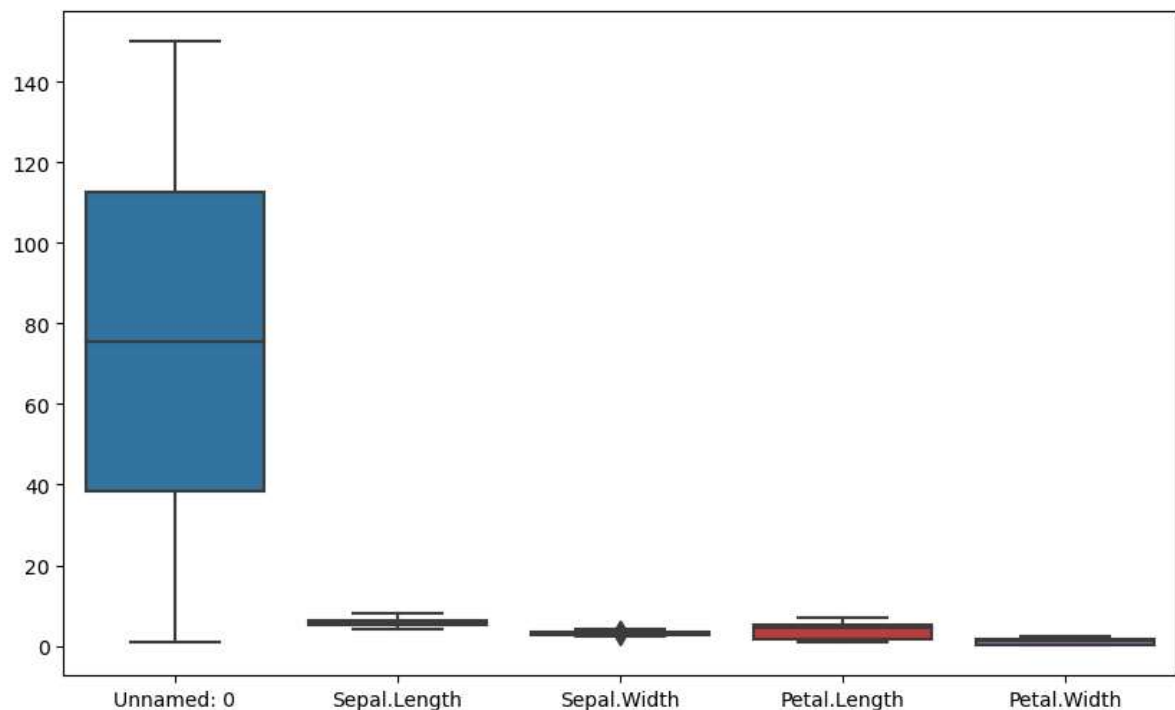
```
In [4]: 1 data.hist(figsize=(10, 6))
        2 plt.tight_layout()
        3 plt.show()
```



```
In [5]: 1 sns.pairplot(data, hue='Species', height=2.5)
        2 plt.show()
```



```
In [6]: 1 plt.figure(figsize=(10, 6))
        2 sns.boxplot(data=data.drop('Species', axis=1))
        3 plt.show()
```



```
In [7]: 1 data
```

Out[7]:

	Unnamed: 0	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
0	1	5.1	3.5	1.4	0.2	setosa
1	2	4.9	3.0	1.4	0.2	setosa
2	3	4.7	3.2	1.3	0.2	setosa
3	4	4.6	3.1	1.5	0.2	setosa
4	5	5.0	3.6	1.4	0.2	setosa
...
145	146	6.7	3.0	5.2	2.3	virginica
146	147	6.3	2.5	5.0	1.9	virginica
147	148	6.5	3.0	5.2	2.0	virginica
148	149	6.2	3.4	5.4	2.3	virginica
149	150	5.9	3.0	5.1	1.8	virginica

150 rows × 6 columns

```
In [8]: 1 new1=data['Unnamed: 0']  
2 new2=data['Sepal.Length']  
3 new3=data['Sepal.Width']  
4 new4=data['Petal.Length']  
5 new5=data['Petal.Width']
```

```
In [9]: 1 print("mean of unnamed:",new1.mean())  
2 print("mean of sepal length :",new2.mean())  
3 print("mean of sepal width:",new3.mean())  
4 print("mean of petal length :",new4.mean())  
5 print("mean of petal width:",new5.mean())
```

```
mean of unnamed: 75.5  
mean of sepal length : 5.8433333333333334  
mean of sepal width: 3.0573333333333337  
mean of petal length : 3.7580000000000005  
mean of petal width: 1.1993333333333336
```

```
In [10]: 1 print("median of unnamed:",new1.median())  
2 print("median of sepal length :",new2.median())  
3 print("median of sepal width:",new3.median())  
4 print("median of petal length:",new4.median())  
5 print("median of petal width:",new5.median())
```

```
median of unnamed: 75.5  
median of sepal length : 5.8  
median of sepal width: 3.0  
median of petal length: 4.35  
median of petal width: 1.3
```

```
In [11]: 1 print("mode of unnamed:",new1.mode().iloc[0])  
2 print("mode of sepal length :",new2.mode().iloc[0])  
3 print("mode of sepal width:",new3.mode().iloc[0])  
4 print("mode of petal length:",new4.mode().iloc[0])  
5 print("mode of petal width:",new5.mode().iloc[0])
```

```
mode of unnamed: 1  
mode of sepal length : 5.0  
mode of sepal width: 3.0  
mode of petal length: 1.4  
mode of petal width: 0.2
```

```
In [12]: 1 print("std_dev of unnamed:",new1.std())  
2 print("std_dev of sepal length :",new2.std())  
3 print("std_dev of sepal width:",new3.std())  
4 print("std_dev of petal length:",new4.std())  
5 print("std_dev of petal width:",new5.std())
```

```
std_dev of unnamed: 43.445367992456916  
std_dev of sepal length : 0.8280661279778629  
std_dev of sepal width: 0.435866284936698  
std_dev of petal length: 1.7652982332594667  
std_dev of petal width: 0.7622376689603465
```

In []:

1	
---	--