import libraries

In [2]:
```
import pandas as pd
```

dataset from kaggle

In [3]:
```
iris_dataset = pd.read_csv(r"C:\Users\khush\Desktop\Iris.csv")
```

In [5]:
```
iris_dataset.head()
```

Out[5]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

getting values

In [6]:
```
iris_dataset.describe()
```

Out[6]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.843333	3.054000	3.758667	1.198667
std	43.445368	0.828066	0.433594	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000

operations of entire dataset

In [7]:
```
iris_dataset.mean()
```

```
C:\Users\khush\AppData\Local\Temp\ipykernel_17568\906983207.py:1: FutureWarni
ng: The default value of numeric_only in DataFrame.mean is deprecated. In a f
uture version, it will default to False. In addition, specifying 'numeric_onl
y=None' is deprecated. Select only valid columns or specify the value of nume
ric_only to silence this warning.
  iris_dataset.mean()
```

Out[7]:
```
Id               75.500000
SepalLengthCm     5.843333
SepalWidthCm      3.054000
PetalLengthCm     3.758667
PetalWidthCm      1.198667
dtype: float64
```

In [8]:
```
iris_dataset.median()
```

```
C:\Users\khush\AppData\Local\Temp\ipykernel_17568\543178892.py:1: FutureWarni
ng: The default value of numeric_only in DataFrame.median is deprecated. In a
future version, it will default to False. In addition, specifying 'numeric_on
ly=None' is deprecated. Select only valid columns or specify the value of num
eric_only to silence this warning.
  iris_dataset.median()
```

Out[8]:
```
Id               75.50
SepalLengthCm     5.80
SepalWidthCm      3.00
PetalLengthCm     4.35
PetalWidthCm      1.30
dtype: float64
```

In [9]:
```
iris_dataset.Species.mode()
```

Out[9]:
```
0        Iris-setosa
1    Iris-versicolor
2     Iris-virginica
Name: Species, dtype: object
```

In [10]:
```
iris_dataset.groupby(['Species']).count()
```

Out[10]:

Species	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
Iris-setosa	50	50	50	50	50
Iris-versicolor	50	50	50	50	50
Iris-virginica	50	50	50	50	50

In [11]:
```
iris_dataset.SepalLengthCm.std()
```

Out[11]:
```
0.8280661279778629
```

In [12]:
```
iris_dataset.SepalWidthCm.std()
```

Out[12]:
```
0.4335943113621737
```

In [13]:
```
iris_dataset.PetalLengthCm.std()
```

Out[13]:
```
1.7644204199522617
```

In [14]:
```
iris_dataset.PetalWidthCm.std()
```

Out[14]:
```
0.7631607417008414
```

all statistical values according to species

In [15]:
```
setosa_stats = iris_dataset[iris_dataset['Species'] == 'Iris-setosa'].describe
```

In [16]:
```
print("Iris-setosa statistics:")
print(setosa_stats)
```

```
Iris-setosa statistics:
             Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count  50.00000       50.00000     50.000000      50.000000      50.00000
mean   25.50000        5.00600      3.418000       1.464000       0.24400
std    14.57738        0.35249      0.381024       0.173511       0.10721
min     1.00000        4.30000      2.300000       1.000000       0.10000
25%    13.25000        4.80000      3.125000       1.400000       0.20000
50%    25.50000        5.00000      3.400000       1.500000       0.20000
75%    37.75000        5.20000      3.675000       1.575000       0.30000
max    50.00000        5.80000      4.400000       1.900000       0.60000
```

In [17]:
```
versicolor_stats = iris_dataset[iris_dataset['Species'] == 'Iris-versicolor'].
```

In [18]:
```
print("\nIris-versicolor statistics:")
print(versicolor_stats)
```

```
Iris-versicolor statistics:
             Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count   50.00000      50.000000     50.000000      50.000000     50.000000
mean    75.50000       5.936000      2.770000       4.260000      1.326000
std     14.57738       0.516171      0.313798       0.469911      0.197753
min     51.00000       4.900000      2.000000       3.000000      1.000000
25%     63.25000       5.600000      2.525000       4.000000      1.200000
50%     75.50000       5.900000      2.800000       4.350000      1.300000
75%     87.75000       6.300000      3.000000       4.600000      1.500000
max    100.00000       7.000000      3.400000       5.100000      1.800000
```

In [19]:
```
virginica_stats = iris_dataset[iris_dataset['Species'] == 'Iris-virginica'].de
```

In [20]:
```
print("\nIris-virginica statistics:")
print(virginica_stats)
```

```
Iris-virginica statistics:
               Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count   50.00000       50.00000     50.000000      50.000000      50.00000
mean   125.50000        6.58800      2.974000       5.552000       2.02600
std     14.57738        0.63588      0.322497       0.551895       0.27465
min    101.00000        4.90000      2.200000       4.500000       1.40000
25%    113.25000        6.22500      2.800000       5.100000       1.80000
50%    125.50000        6.50000      3.000000       5.550000       2.00000
75%    137.75000        6.90000      3.175000       5.875000       2.30000
max    150.00000        7.90000      3.800000       6.900000       2.50000
```

interquartile range

In [22]:
```
setosa_data = iris_dataset[iris_dataset['Species'] == 'Iris-setosa']
```

In [24]:
```
setosa_q1 = setosa_data['SepalLengthCm'].quantile(0.25)
```

In [25]:
```
print("First Quartile (Q1) for Iris-setosa (sepal_length):", setosa_q1)
```

```
First Quartile (Q1) for Iris-setosa (sepal_length): 4.8
```

In [27]:
```
setosa_data = iris_dataset[iris_dataset['Species'] == 'Iris-setosa']
```

In [28]:
```
setosa_q3 = setosa_data['SepalLengthCm'].quantile(0.75)
```

In [29]:
```
print("Third Quartile (Q3) for Iris-setosa (sepal_length):", setosa_q3)
```

```
Third Quartile (Q3) for Iris-setosa (sepal_length): 5.2
```

In [30]:
```
setosa_iqr = setosa_q3 - setosa_q1
```

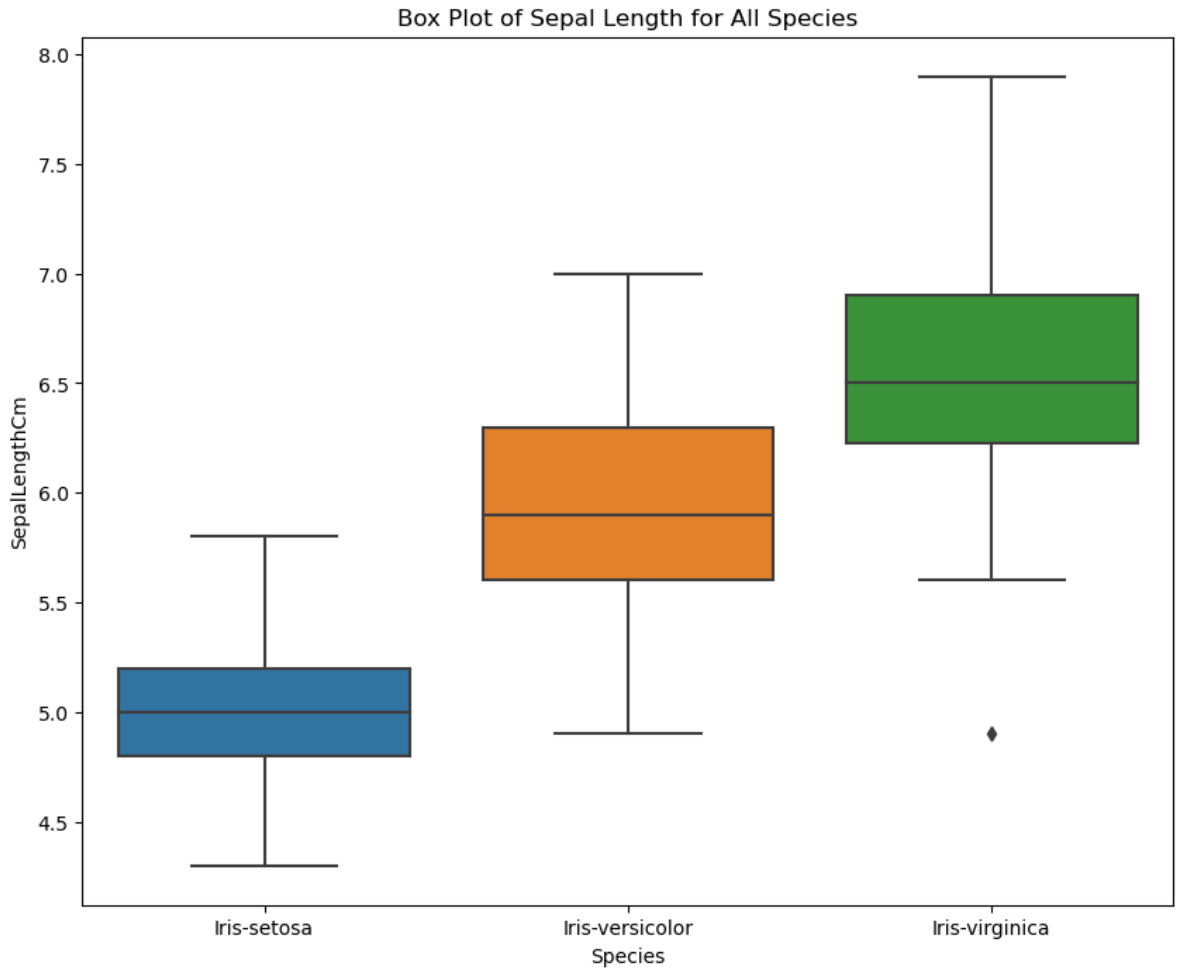
In [31]:
```
print("Interquartile Range (IQR) for Iris-setosa (sepal_length):", setosa_iqr)
```

```
Interquartile Range (IQR) for Iris-setosa (sepal_length): 0.40000000000000036
```

box plot

In [32]:
```
import seaborn as sns
import matplotlib.pyplot as plt
```

In [34]:
```
plt.figure(figsize=(10, 8))
sns.boxplot(x='Species', y='SepalLengthCm', data=iris_dataset)
plt.title('Box Plot of Sepal Length for All Species')
plt.show()
```



In [ ]: