

Data Visualization - III

In []:

In []:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In []:

```
df1=pd.read_csv('/content/iris.csv')
df1
```

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
...
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In []:

```
df=pd.DataFrame(df1)
df.head()
```

Out[6]:

	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

In []:

df.describe()

Out[7]:

	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

In []:

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Id               150 non-null   int64
1   SepalLengthCm    150 non-null   float64
2   SepalWidthCm     150 non-null   float64
3   PetalLengthCm    150 non-null   float64
4   PetalWidthCm     150 non-null   float64
5   Species          150 non-null   object
dtypes: float64(4), int64(1), object(1)
memory usage: 7.2+ KB
```

In []:

df.columns

Out[9]:

```
Index(['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm',
      'Species'],
      dtype='object')
```

In []:

df['SepalLengthCm'].max()

Out[10]:

7.9

In []:

```
df['SepalLengthCm'].min()
```

Out[11]:

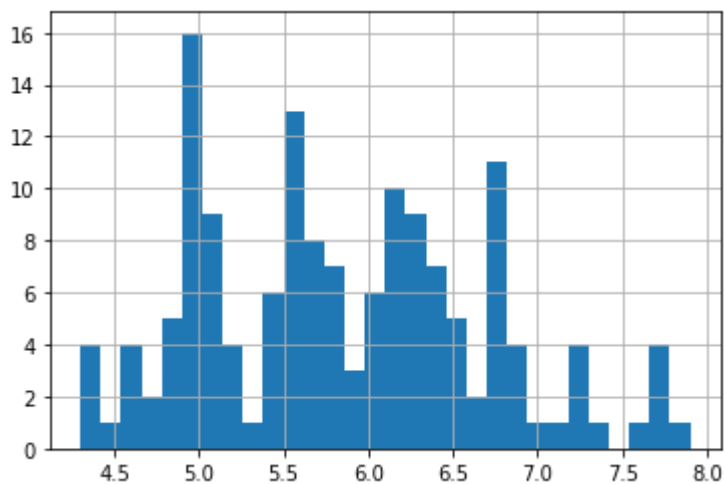
4.3

In []:

```
df['SepalLengthCm'].hist(bins=30)
```

Out[12]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb58296150>



In []:

```
df['PetalLengthCm'].max()
```

Out[13]:

6.9

In []:

```
df['PetalLengthCm'].min()
```

Out[14]:

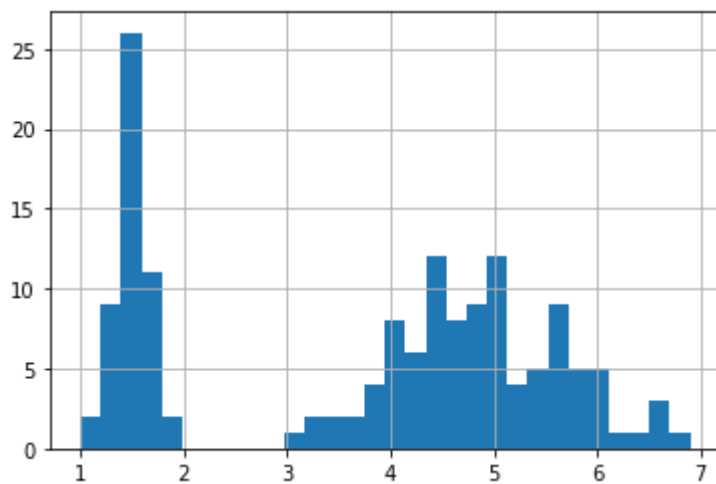
1.0

In []:

```
df['PetalLengthCm'].hist(bins=30)
```

Out[15]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb581536d0>



In []:

```
df['PetalWidthCm'].max()
```

Out[16]:

2.5

In []:

```
df['PetalWidthCm'].min()
```

Out[17]:

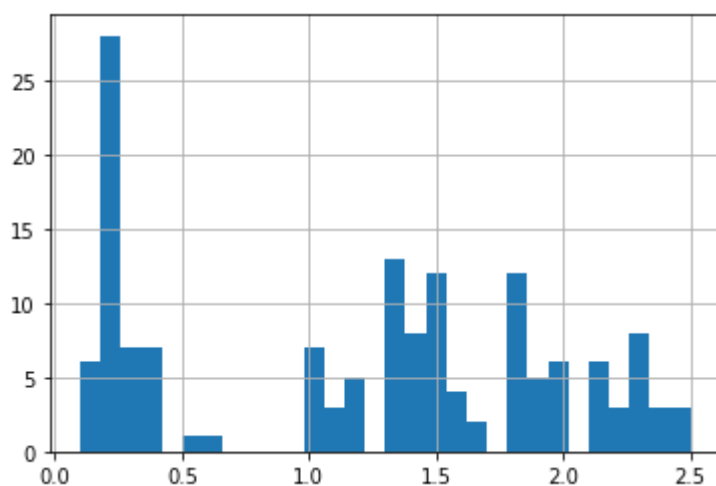
0.1

In []:

```
df['PetalWidthCm'].hist(bins=30)
```

Out[18]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb57beff50>



In []:

```
df['SepalWidthCm'].max()
```

Out[19]:

4.4

In []:

```
df['SepalWidthCm'].min()
```

Out[20]:

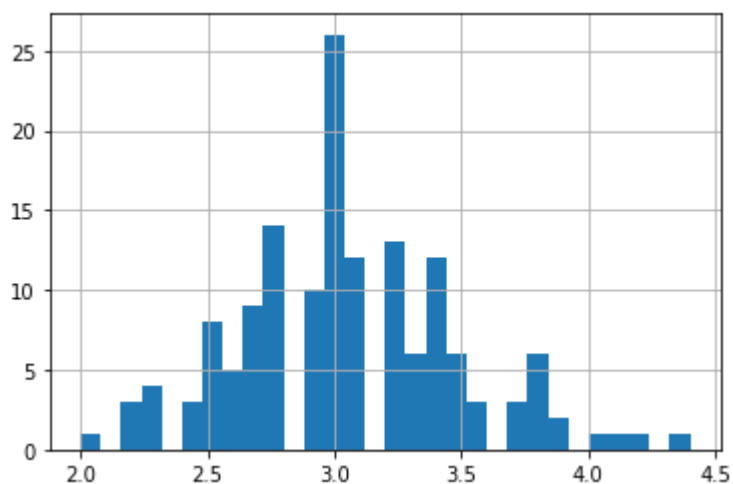
2.0

In []:

```
df['SepalWidthCm'].hist(bins=30)
```

Out[21]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb57b6da50>



In []:

```
df['Species'].value_counts()
```

Out[22]:

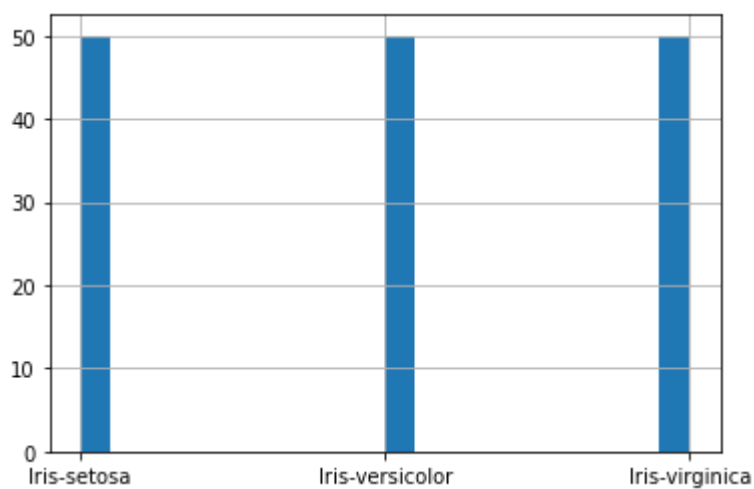
```
Iris-setosa      50
Iris-versicolor  50
Iris-virginica   50
Name: Species, dtype: int64
```

In []:

```
df['Species'].hist(bins=20)
```

Out[23]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb57a7e250>

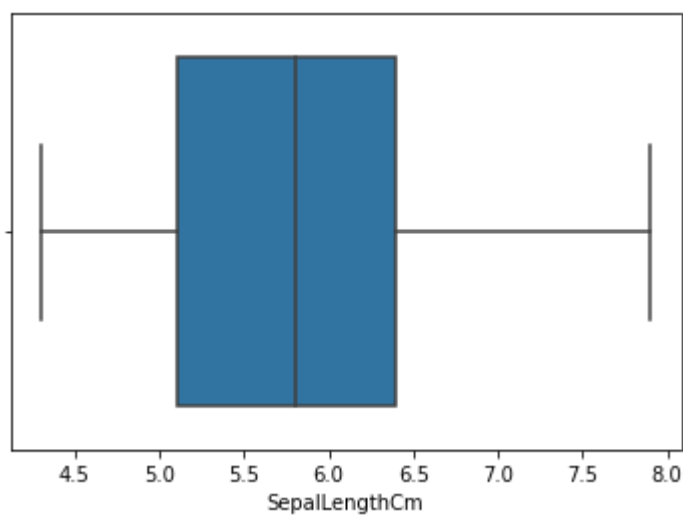


In []:

```
sns.boxplot(x='SepalLengthCm', data=df)
```

Out[24]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb57a50bd0>

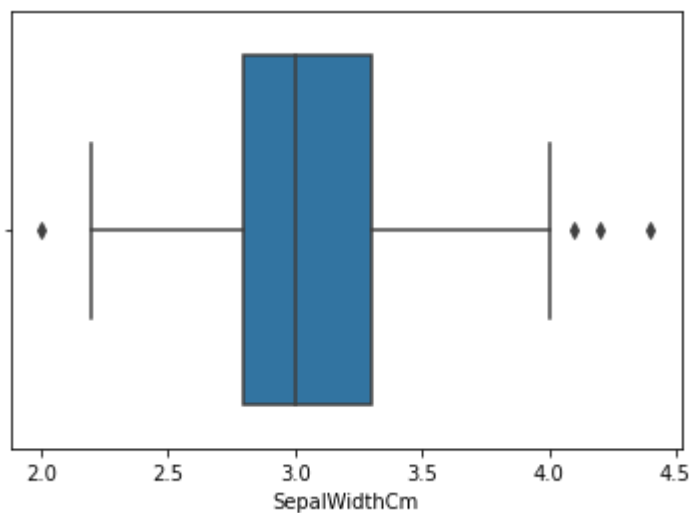


In []:

```
sns.boxplot(x='SepalWidthCm',data=df)
```

Out[25]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb57976390>

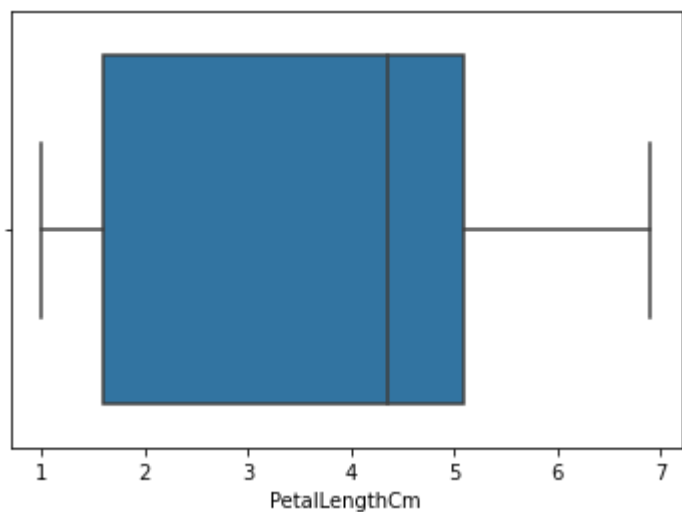


In []:

```
sns.boxplot(x='PetalLengthCm',data=df)
```

Out[26]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb578f5d50>



In []:

```
sns.boxplot(x='PetalWidthCm',data=df)
```

Out[27]:

<matplotlib.axes._subplots.AxesSubplot at 0x7fdb57868e50>

