

```
In [28]: import pandas as pd
import numpy as np
import seaborn as sns
```

```
In [2]: df=pd.read_csv('iris.csv')
```

```
In [3]: df.head()
```

```
Out[3]:
```

	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 [4]: df.tail()
```

```
Out[4]:
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
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

```
In [5]: df.describe()
```

Out[5]:

	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 [7]: `df.shape`

Out[7]: (150, 6)

In [8]: `df.isnull().any().sum()`

Out[8]: 0

In [9]: `df.dtypes`

Out[9]:

Id	int64
SepalLengthCm	float64
SepalWidthCm	float64
PetalLengthCm	float64
PetalWidthCm	float64
Species	object

dtype: object

```
In [17]: def categorize(series):
          if series.dtype == 'object':
              print("The datatype of ", series.name, " is nominal")
          else:
              print("The datatype of ", series.name, " is numeric")
```

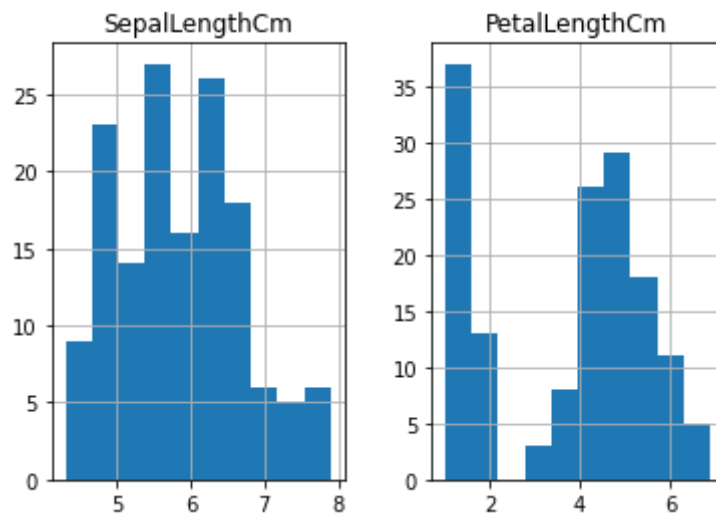
```
In [20]: df.apply(categorize,0)
```

```
The datatype of Id is numeric  
The datatype of SepalLengthCm is numeric  
The datatype of SepalWidthCm is numeric  
The datatype of PetalLengthCm is numeric  
The datatype of PetalWidthCm is numeric  
The datatype of Species is nominal
```

```
Out[20]: Id          None  
SepalLengthCm      None  
SepalWidthCm       None  
PetalLengthCm      None  
PetalWidthCm       None  
Species            None  
dtype: object
```

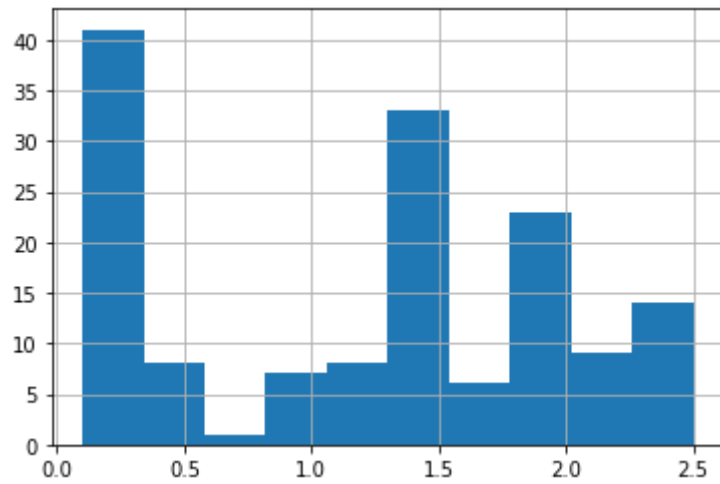
```
In [25]: df.hist(['SepalLengthCm', 'PetalLengthCm'])
```

```
Out[25]: array([[<AxesSubplot:title={'center': 'SepalLengthCm'}>,  
                <AxesSubplot:title={'center': 'PetalLengthCm'}>]], dtype=object)
```



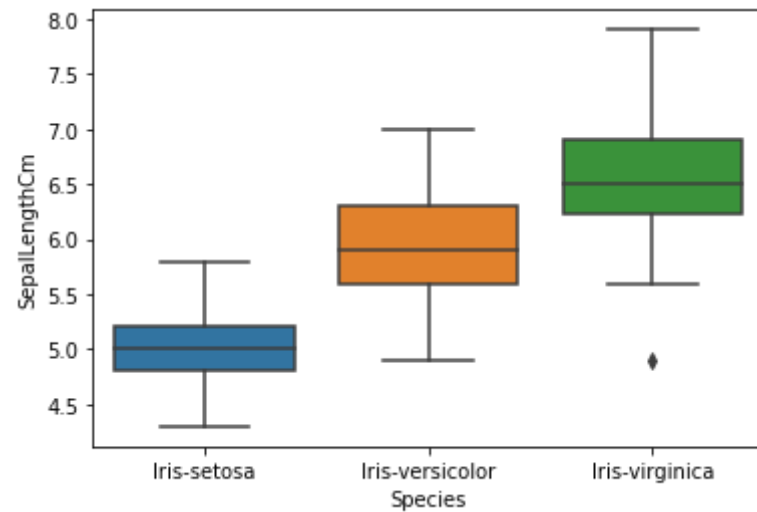
```
In [26]: df['PetalWidthCm'].hist()
```

```
Out[26]: <AxesSubplot:>
```



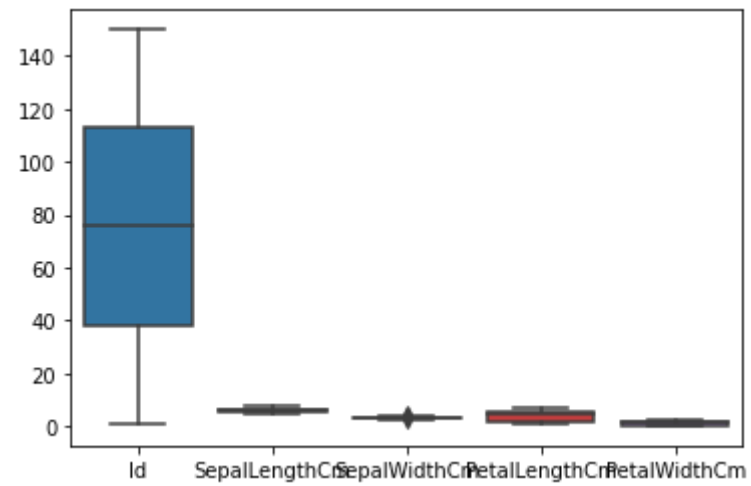
```
In [29]: sns.boxplot(x='Species', y='SepalLengthCm', data=df)
```

```
Out[29]: <AxesSubplot:xlabel='Species', ylabel='SepalLengthCm'>
```



```
In [32]: sns.boxplot( data=df)
```

```
Out[32]: <AxesSubplot:>
```



In []: