

# 31444 Mehul

## Assignment - 10

```
In [ ]: import pandas as pd
import seaborn as sns
from matplotlib import pyplot as plt
import numpy as np
import warnings
warnings.filterwarnings('ignore')
```

Categorical Nominal Variable -> Species

Quantitative Continous Variable -> Sepal length , Sepal Width, Petal Length, Petal Width

```
In [ ]: data = pd.read_csv("Iris.csv")
```

```
In [ ]: data.head()
```

```
Out[ ]:
```

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

```
In [ ]: data.isnull().sum()
```

```
Out[ ]: Id          0
SepalLengthCm    0
SepalWidthCm     0
PetalLengthCm    0
PetalWidthCm     0
Species          0
dtype: int64
```

```
In [ ]: data.dtypes
```

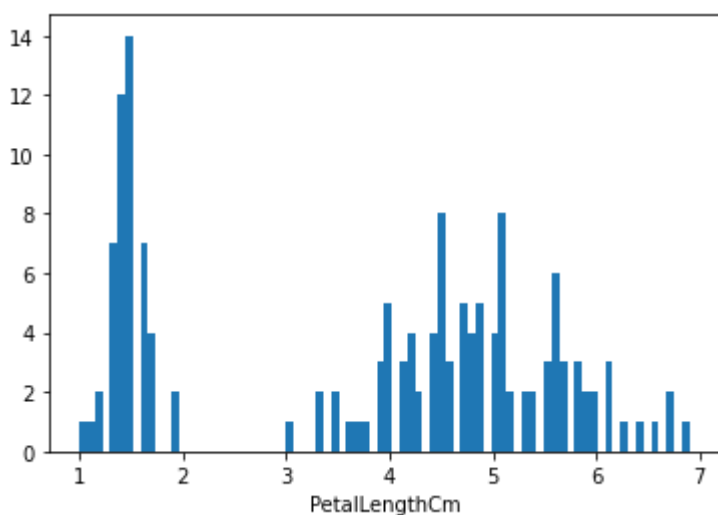
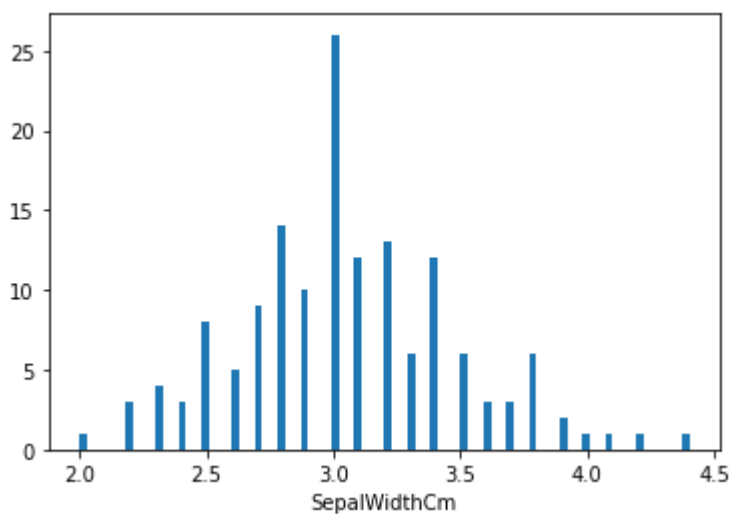
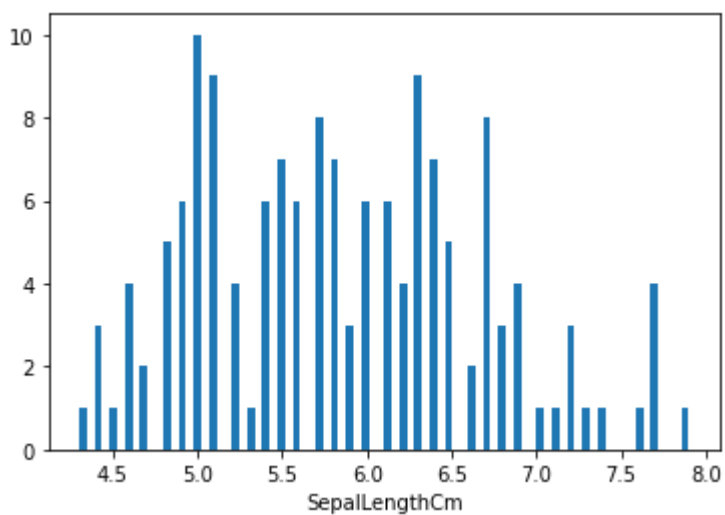
```
Out[ ]: Id          int64
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm    float64
PetalWidthCm     float64
Species          object
dtype: object
```

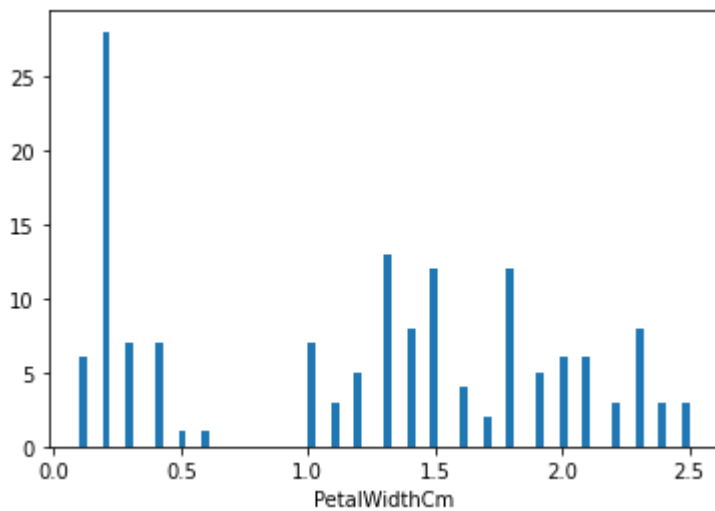
```
In [ ]: features_col = ['SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm']
```

## 2. Histogram for each feature

```
In [ ]:
```

```
for i in features_col:  
    plt.hist(data[i], bins=80)  
    plt.xlabel(i)  
    plt.show()
```

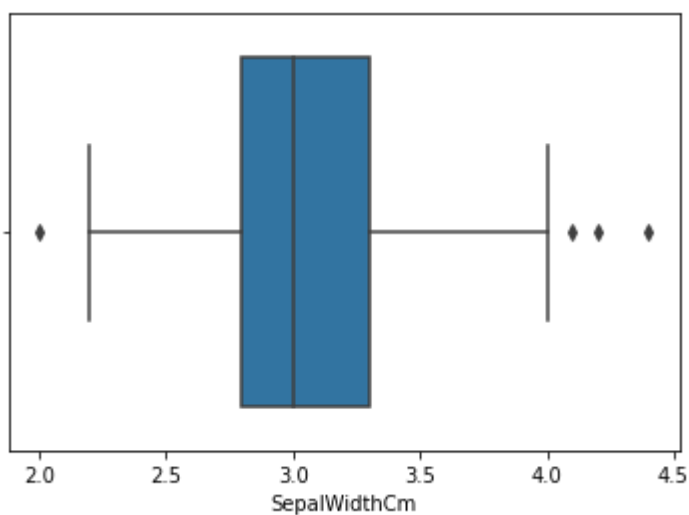
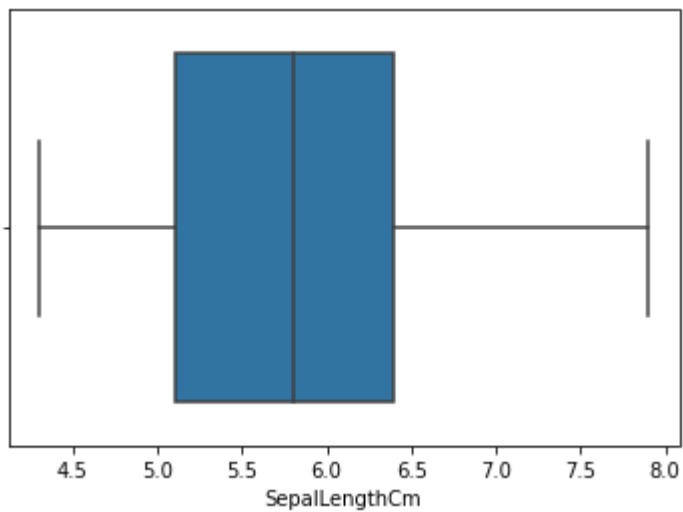


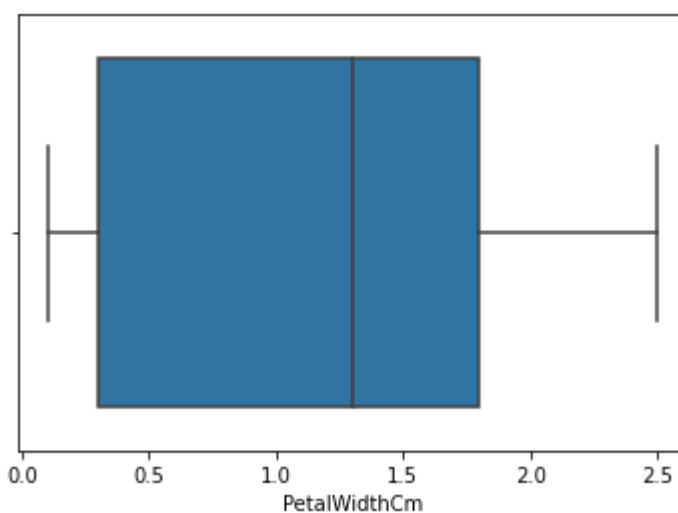
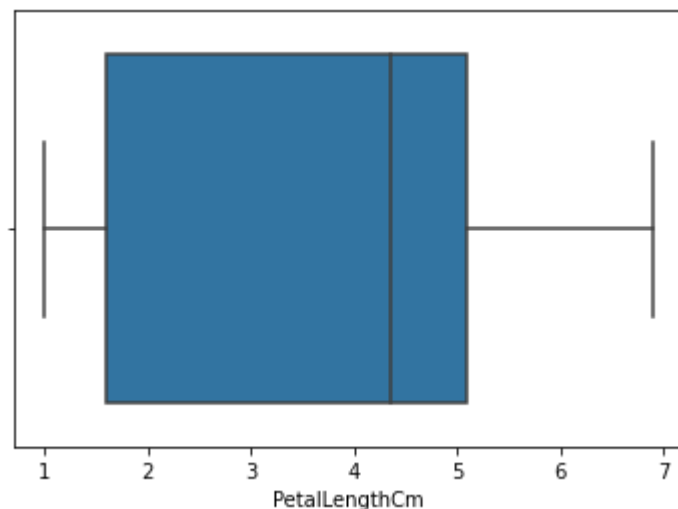


### 3. BoxPlot for each feature

In [ ]:

```
for i in features_col:  
    sns.boxplot(data[i])  
    plt.show()
```





Only sepalWidth has outliers

#### 4. Comparing distributions and identifying outliers

```
In [ ]: def showOutliers(col):
    Q3= np.quantile(data[col],0.75)
    Q1= np.quantile(data[col],0.25)
    IQR = Q3 - Q1
    lower_range = Q1 - 1.5 * IQR
    upper_range = Q3 + 1.5 * IQR
    outlier_list = [x for x in data[col] if (
        not (x>lower_range and x<upper_range)
    )]
    print(col, " => ",outlier_list)
```

```
In [ ]: showOutliers('SepalWidthCm')
showOutliers('PetalWidthCm')
showOutliers('SepalLengthCm')
showOutliers('PetalLengthCm')
```

```
SepalWidthCm  => [4.4, 4.1, 4.2, 2.0]
PetalWidthCm  => []
SepalLengthCm => []
PetalLengthCm => []
```