untitled10

May 7, 2024

[3]:
```
import numpy as np
import pandas as pd
import seaborn as sns
```

[4]:
```
df = pd.read_csv('Iris.csv')
```

[5]:
```
df
```

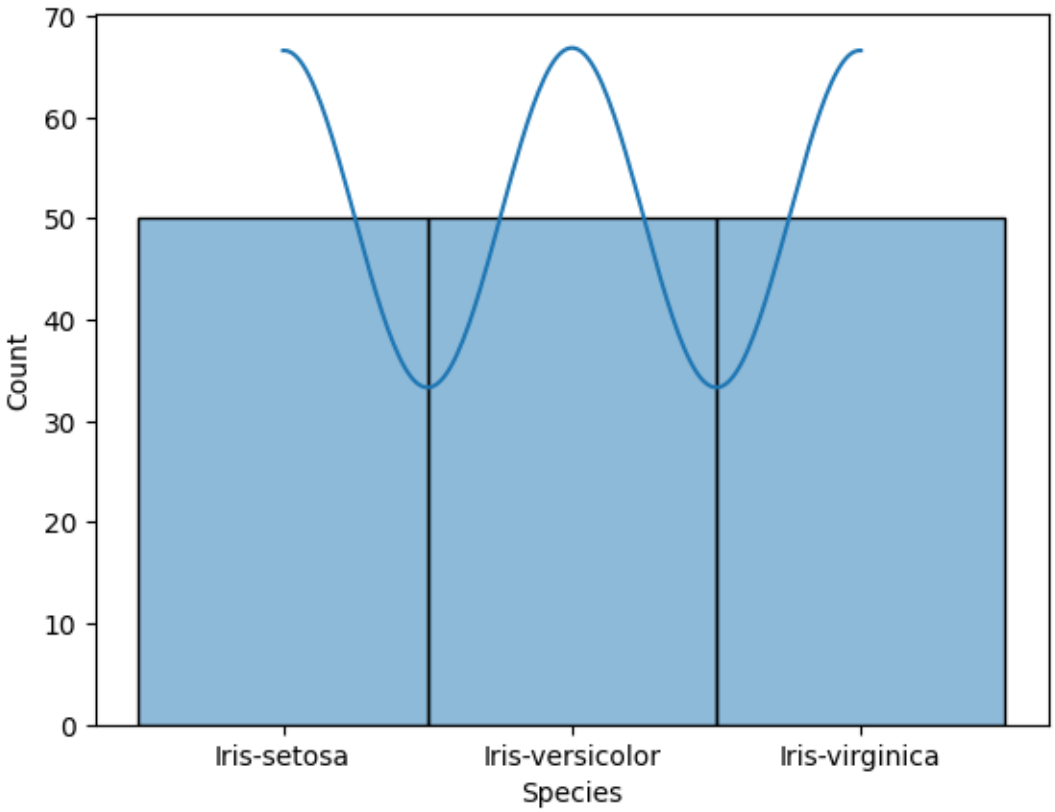
[5]:
```
       Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
0       1            5.1           3.5            1.4           0.2
1       2            4.9           3.0            1.4           0.2
2       3            4.7           3.2            1.3           0.2
3       4            4.6           3.1            1.5           0.2
4       5            5.0           3.6            1.4           0.2
..    …              …             …              …             …
145   146            6.7           3.0            5.2           2.3
146   147            6.3           2.5            5.0           1.9
147   148            6.5           3.0            5.2           2.0
148   149            6.2           3.4            5.4           2.3
149   150            5.9           3.0            5.1           1.8

            Species
0       Iris-setosa
1       Iris-setosa
2       Iris-setosa
3       Iris-setosa
4       Iris-setosa
..              …
145  Iris-virginica
146  Iris-virginica
147  Iris-virginica
148  Iris-virginica
149  Iris-virginica

[150 rows x 6 columns]
```

[6]:
```
df.isnull().sum()
```

[6]:
```
Id               0
SepalLengthCm    0
SepalWidthCm     0
PetalLengthCm    0
PetalWidthCm     0
Species          0
dtype: int64
```

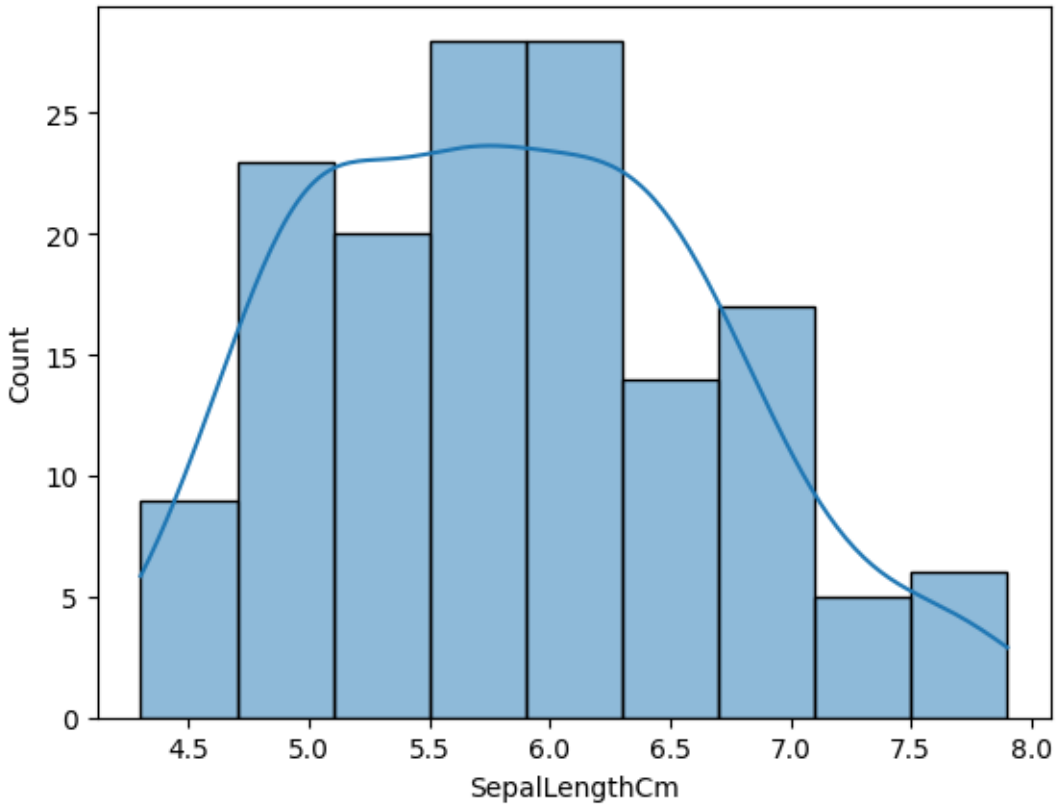
[32]:
```
sns.histplot(data=df['Species'],kde=True)
```

[32]:
```
<Axes: xlabel='Species', ylabel='Count'>
```



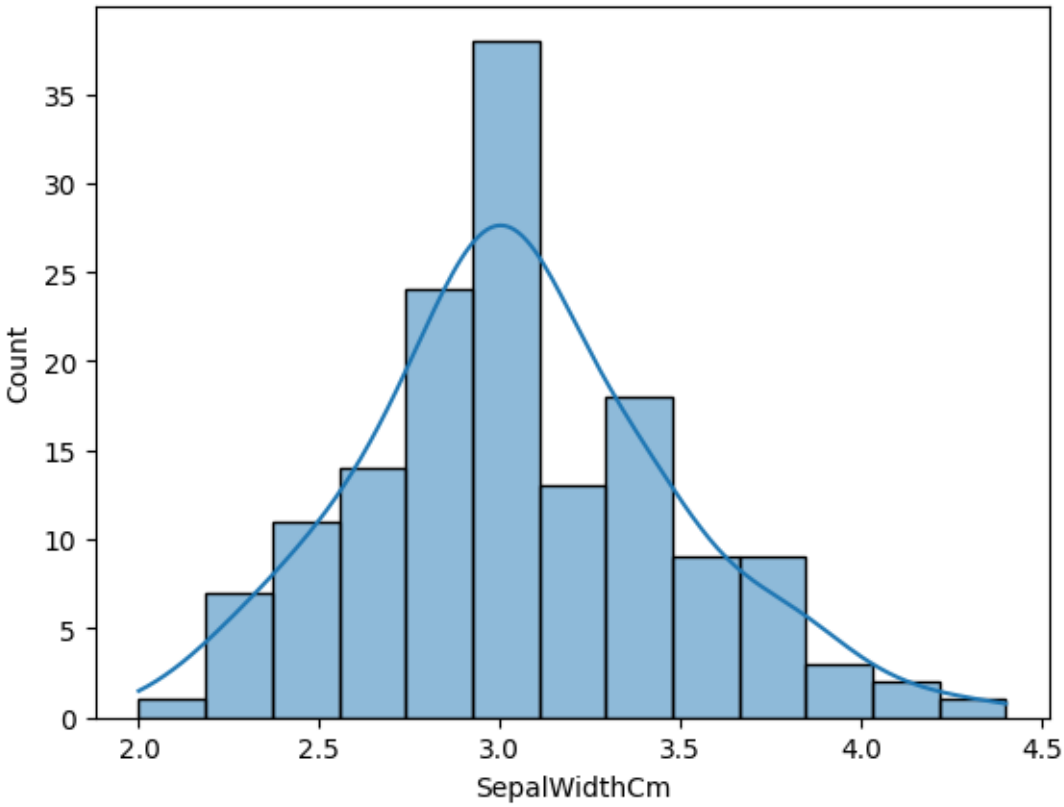
[33]:
```
sns.histplot(data=df['SepalLengthCm'],kde=True)
```

[33]:
```
<Axes: xlabel='SepalLengthCm', ylabel='Count'>
```



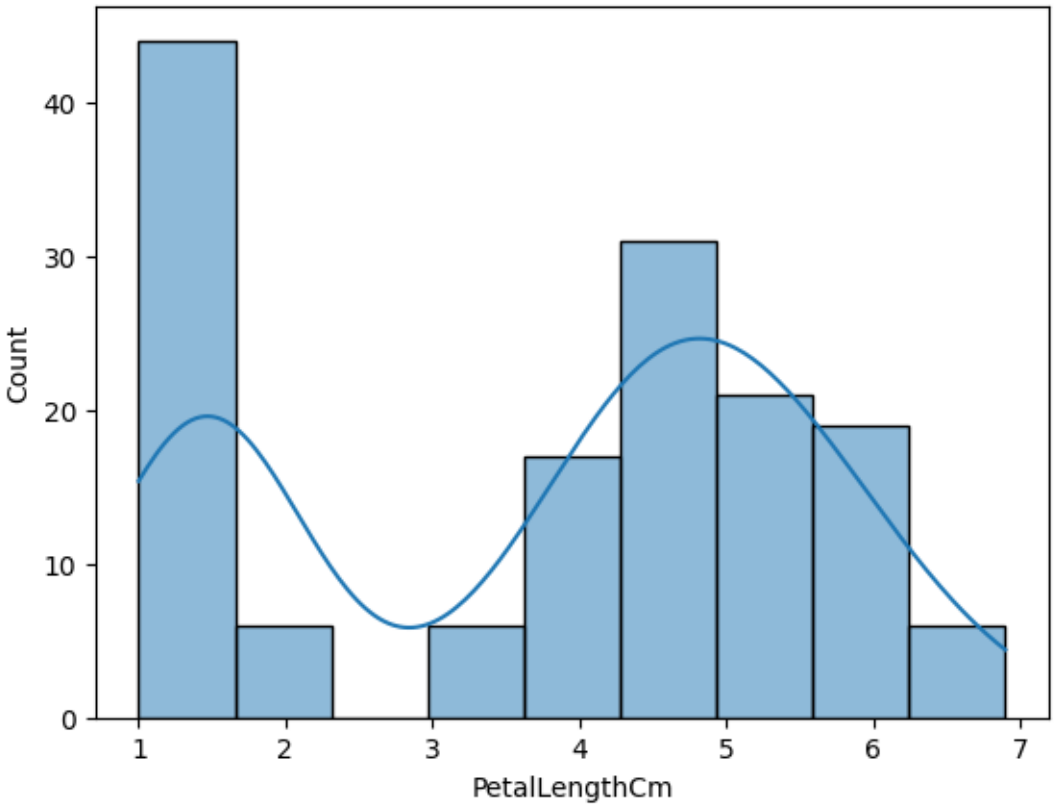
[34]:
```
sns.histplot(data=df['SepalWidthCm'],kde=True)
```

[34]:
```
<Axes: xlabel='SepalWidthCm', ylabel='Count'>
```



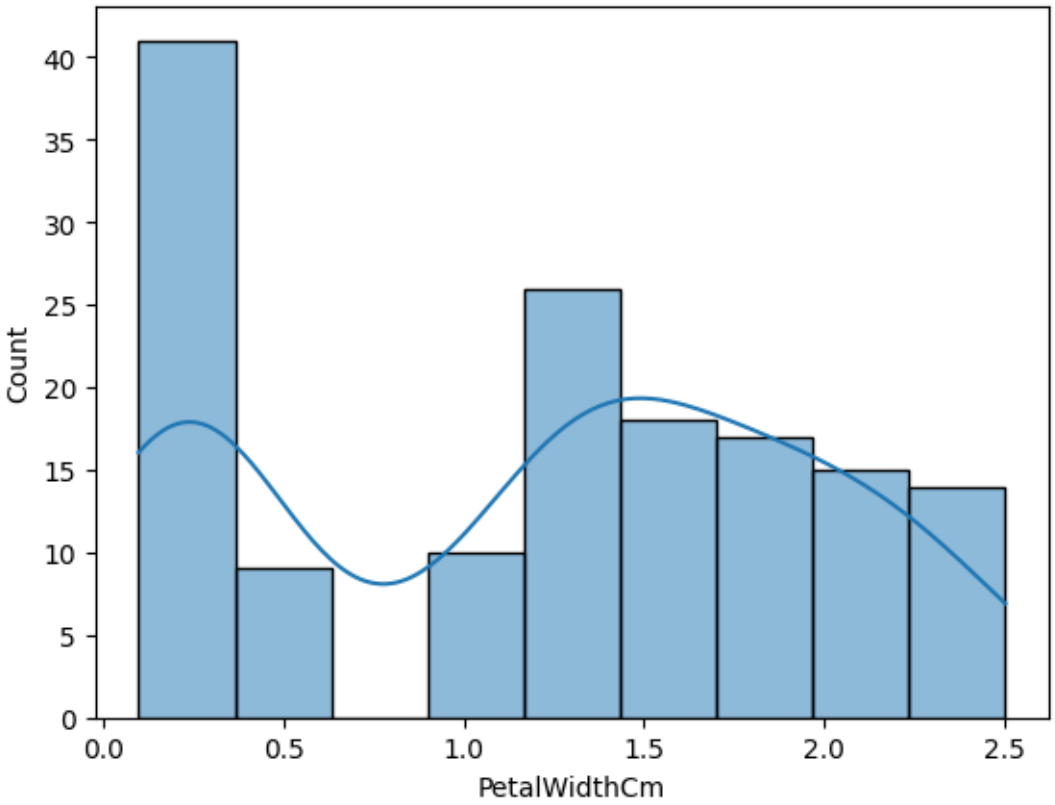
[35]:
```
sns.histplot(data=df['PetalLengthCm'],kde=True)
```

[35]:
```
<Axes: xlabel='PetalLengthCm', ylabel='Count'>
```



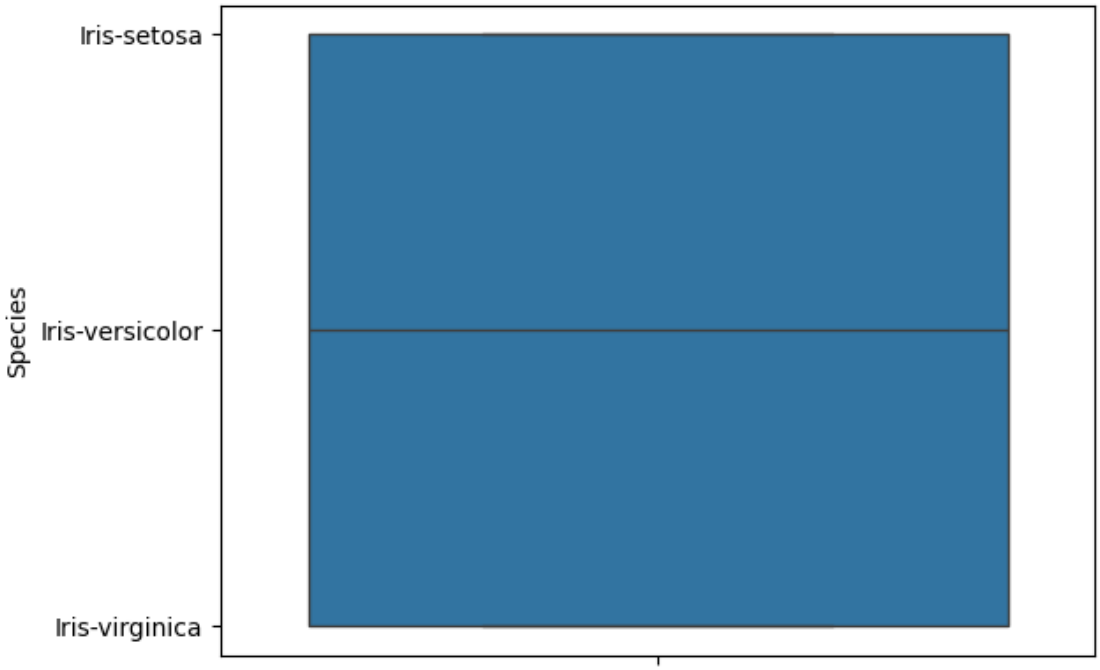
[36]:
```
sns.histplot(data=df['PetalWidthCm'],kde=True)
```

[36]: <Axes: xlabel='PetalWidthCm', ylabel='Count'>



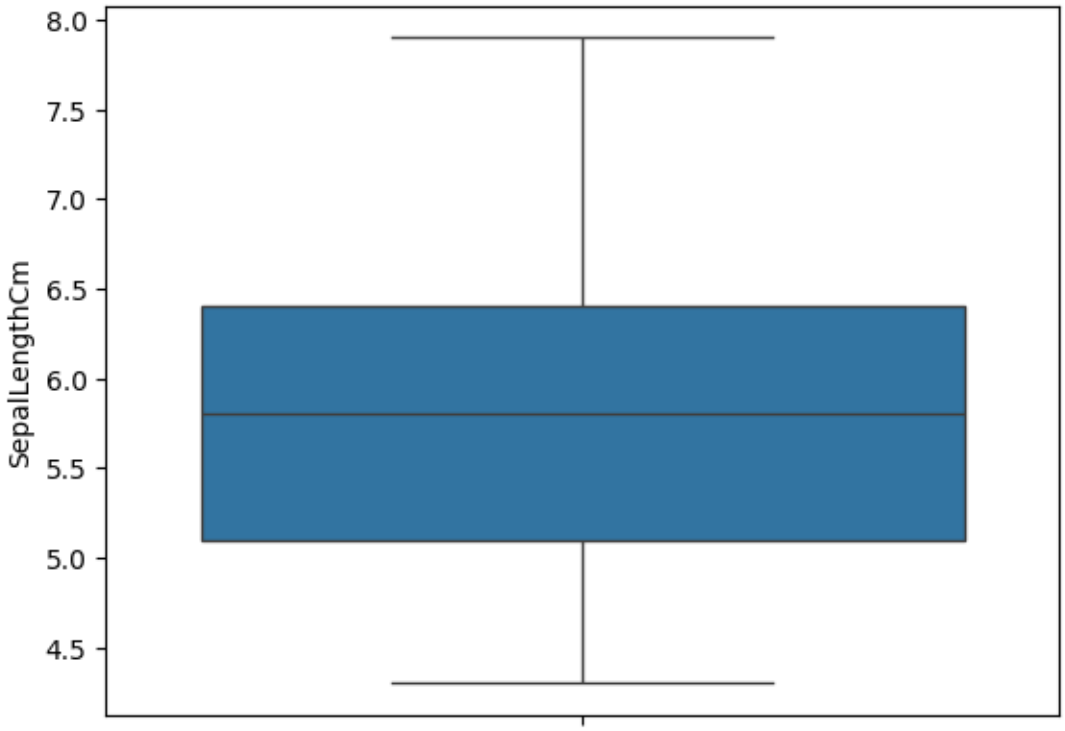
[37]:
```
sns.boxplot(data=df['Species'])
```

[37]:
```
<Axes: ylabel='Species'>
```



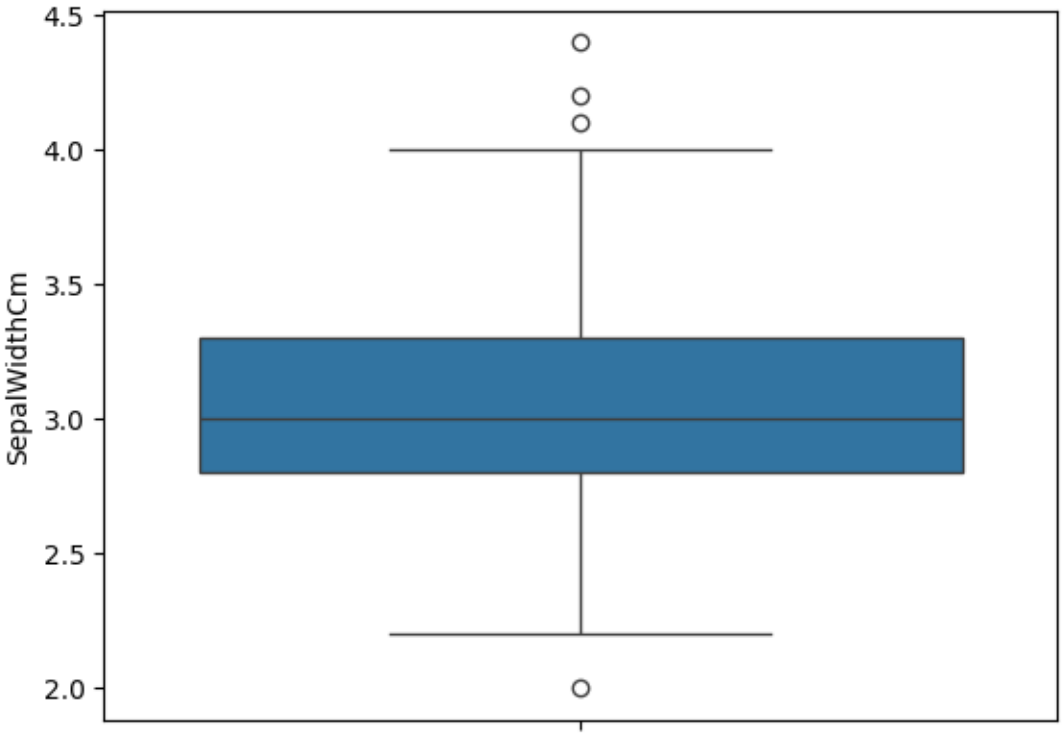
[38]:
```
sns.boxplot(data=df['SepalLengthCm'])
```

[38]:
```
<Axes: ylabel='SepalLengthCm'>
```



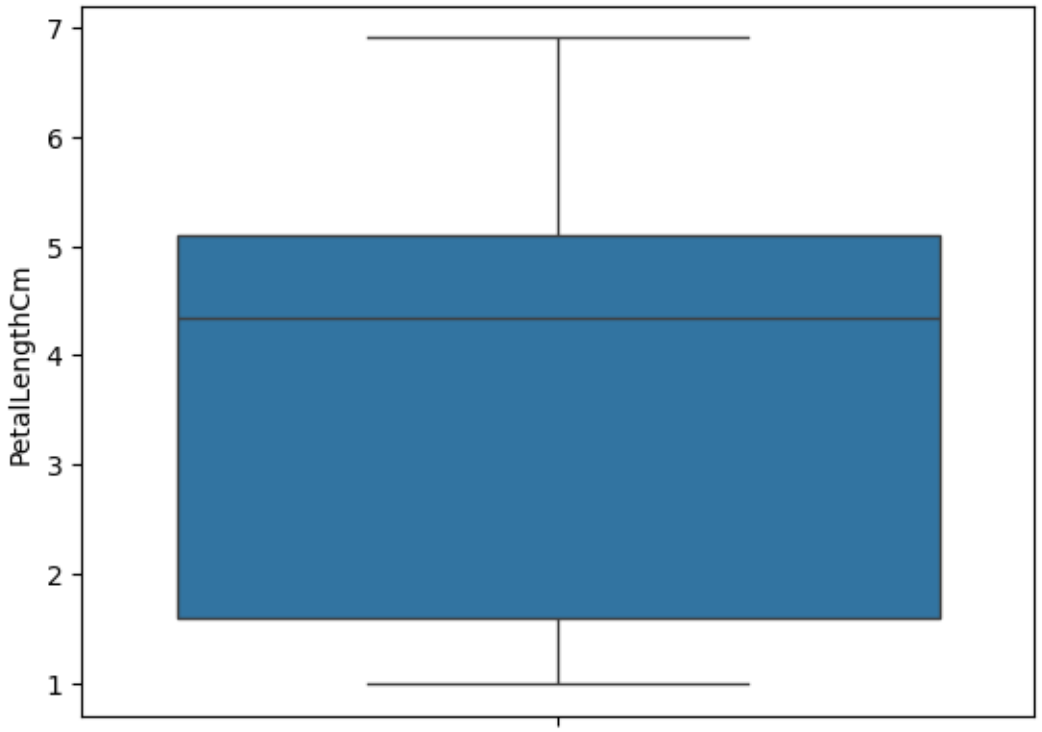
[39]:
```
sns.boxplot(data=df['SepalWidthCm'])
```

[39]: <Axes: ylabel='SepalWidthCm'>



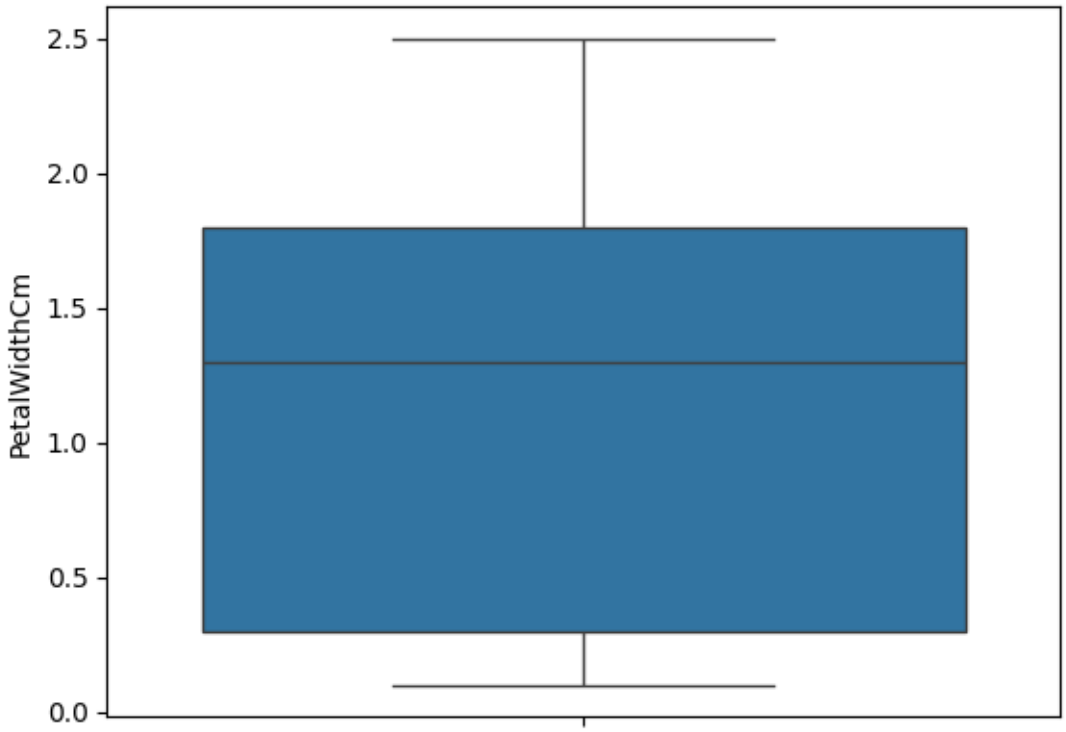
[40]:
```
sns.boxplot(data=df['PetalLengthCm'])
```

[40]: <Axes: ylabel='PetalLengthCm'>



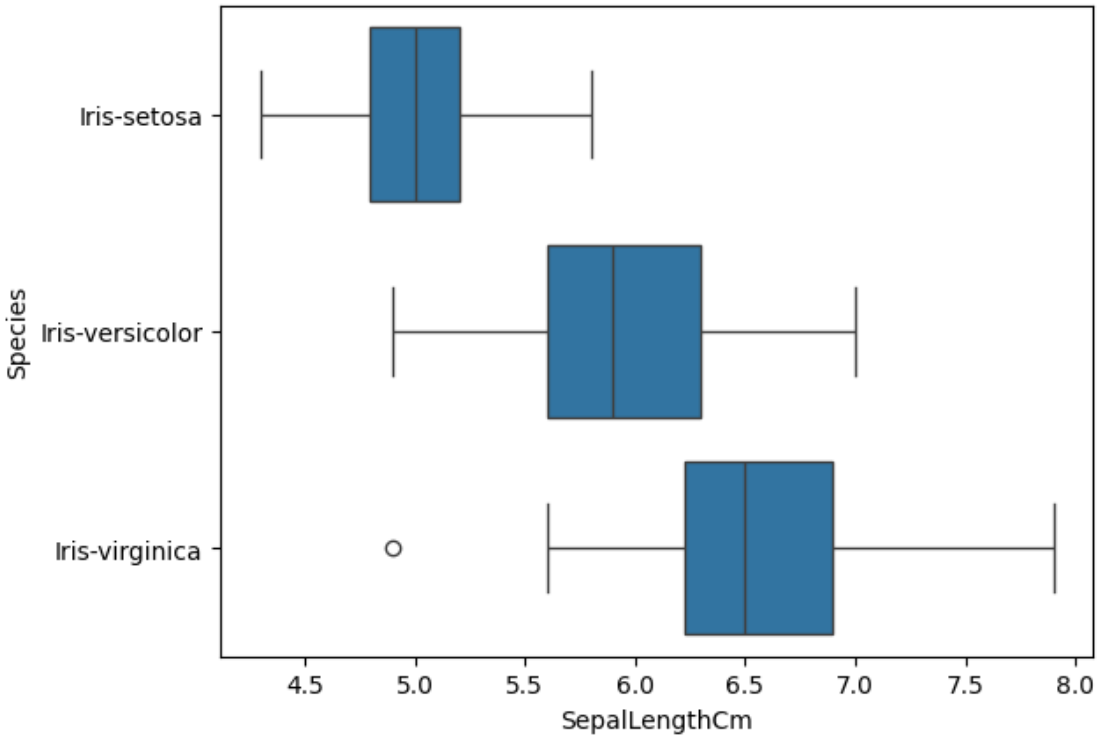
[42]:
```
sns.boxplot(data=df['PetalWidthCm'])
```

[42]: <Axes: ylabel='PetalWidthCm'>



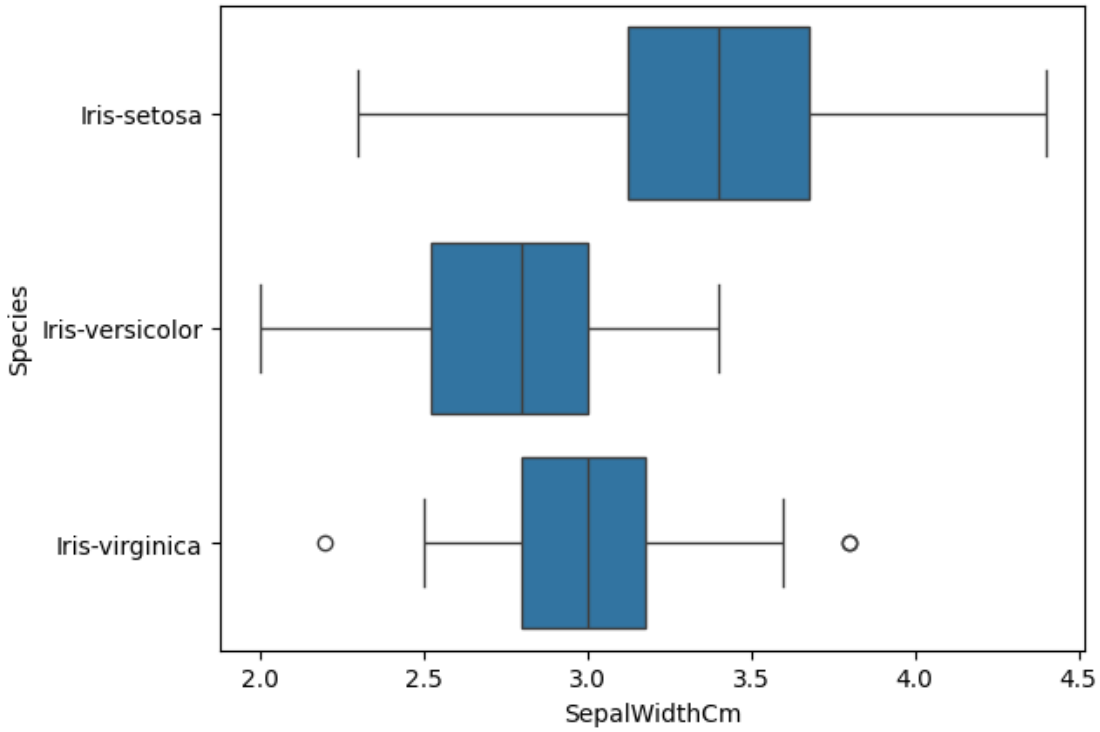
[66]:
```
sns.boxplot(x='SepalLengthCm',y='Species',data=df)
```

[66]:
```
<Axes: xlabel='SepalLengthCm', ylabel='Species'>
```



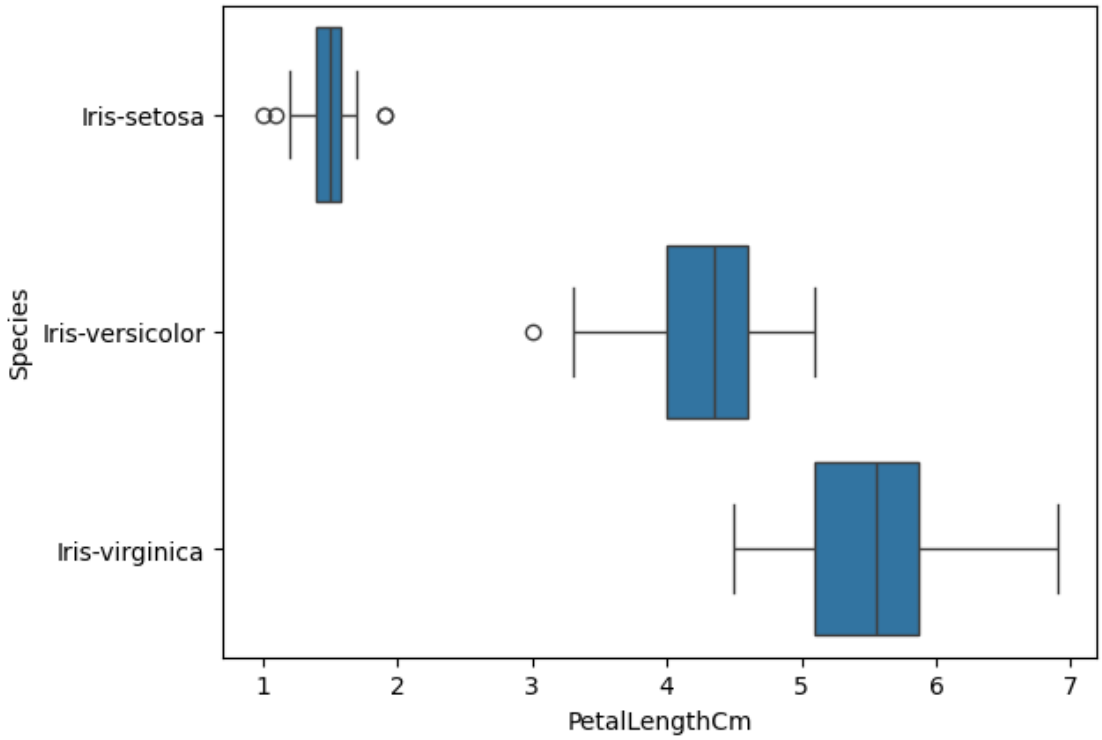
[44]:
```
sns.boxplot(x='SepalWidthCm',y='Species',data=df)
```

[44]: <Axes: xlabel='SepalWidthCm', ylabel='Species'>



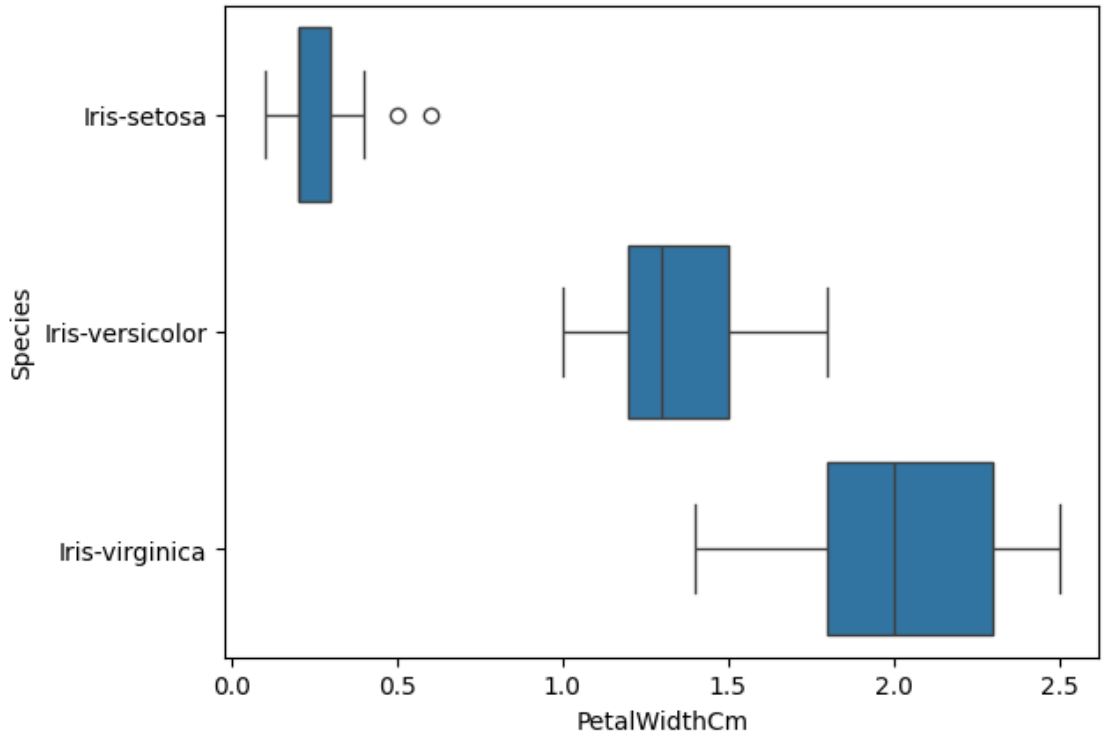
[65]:
```
sns.boxplot(x='PetalLengthCm',y='Species',data=df)
```

[65]:
```
<Axes: xlabel='PetalLengthCm', ylabel='Species'>
```



[68]:
```
sns.boxplot(x='PetalWidthCm',y='Species',data=df)
# Filter the DataFrame for 'Iris-setosa'
setosa_data = df[df['Species'] == 'Iris-setosa']['PetalWidthCm']

# Calculate IQR for 'Iris-setosa'
Q1 = setosa_data.quantile(0.25)
Q3 = setosa_data.quantile(0.75)
IQR = Q3 - Q1

# Determine lower and upper bounds for outliers
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

# Identify outliers for 'Iris-setosa'
outliers = setosa_data[(setosa_data < lower_bound) | (setosa_data >␣
 ↪upper_bound)]
print("Outliers for Iris-setosa:")
print(outliers)
```

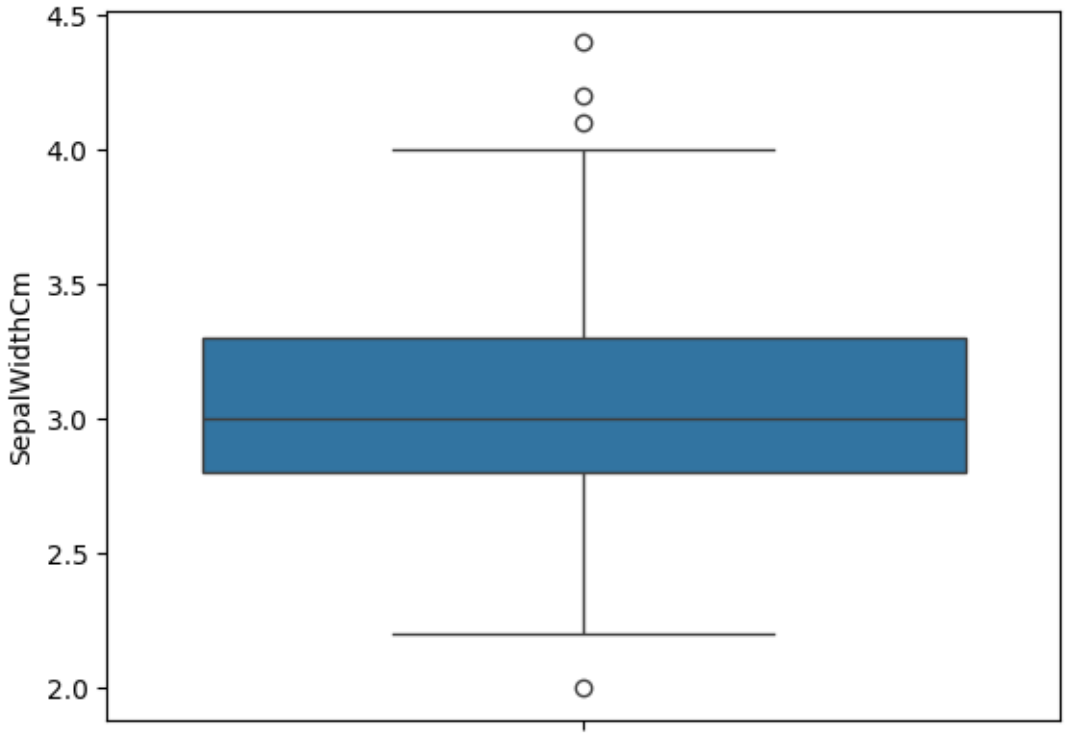
```
Outliers for Iris-setosa:
23    0.5
43    0.6
Name: PetalWidthCm, dtype: float64
```



[62]:
```
sns.boxplot(data=df['SepalWidthCm'])
Q1 = df['SepalWidthCm'].quantile(0.25)
Q3 = df['SepalWidthCm'].quantile(0.75)

IQR = Q3 - Q1

IQR

lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

lower_bound
upper_bound


outliers = df[(df['SepalWidthCm'] < lower_bound) | (df['SepalWidthCm'] >␣
 ↪upper_bound)]
print(outliers)
```

```
    Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
15  16            5.7           4.4            1.5           0.4
32  33            5.2           4.1            1.5           0.1
33  34            5.5           4.2            1.4           0.2
```

```
60  61            5.0           2.0            3.5           1.0

            Species
15      Iris-setosa
32      Iris-setosa
33      Iris-setosa
60  Iris-versicolor
```



[ ]: