

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import skew
```

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

```
In [3]: 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 [4]: df.shape
```

```
Out[4]: (150, 6)
```

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

```
Out[5]:
```

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

```
In [6]: df.head()
```

```
Out[6]:
```

	<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 [7]: df.tail()
```

```
Out[7]:
```

	<b>Id</b>	<b>SepalLengthCm</b>	<b>SepalWidthCm</b>	<b>PetalLengthCm</b>	<b>PetalWidthCm</b>	<b>Species</b>
<b>145</b>	146	6.7	3.0	5.2	2.3	Iris-virginica
<b>146</b>	147	6.3	2.5	5.0	1.9	Iris-virginica
<b>147</b>	148	6.5	3.0	5.2	2.0	Iris-virginica
<b>148</b>	149	6.2	3.4	5.4	2.3	Iris-virginica
<b>149</b>	150	5.9	3.0	5.1	1.8	Iris-virginica

```
In [8]: df.isna().sum()
```

```
Out[8]: Id          0
SepalLengthCm      0
SepalWidthCm       0
PetalLengthCm      0
PetalWidthCm       0
Species            0
dtype: int64
```

```
In [9]: print("\n\nThe features in the dataset are as follows : ")
print("1. Sepal length : ",df['SepalLengthCm'].dtype)
print("2. Sepal width : ",df['SepalWidthCm'].dtype)
print("3. Petal length : ",df['PetalLengthCm'].dtype)
print("4. Petal width : ",df['PetalWidthCm'].dtype)
print("5. Species : ",df['Species'].dtype)
```

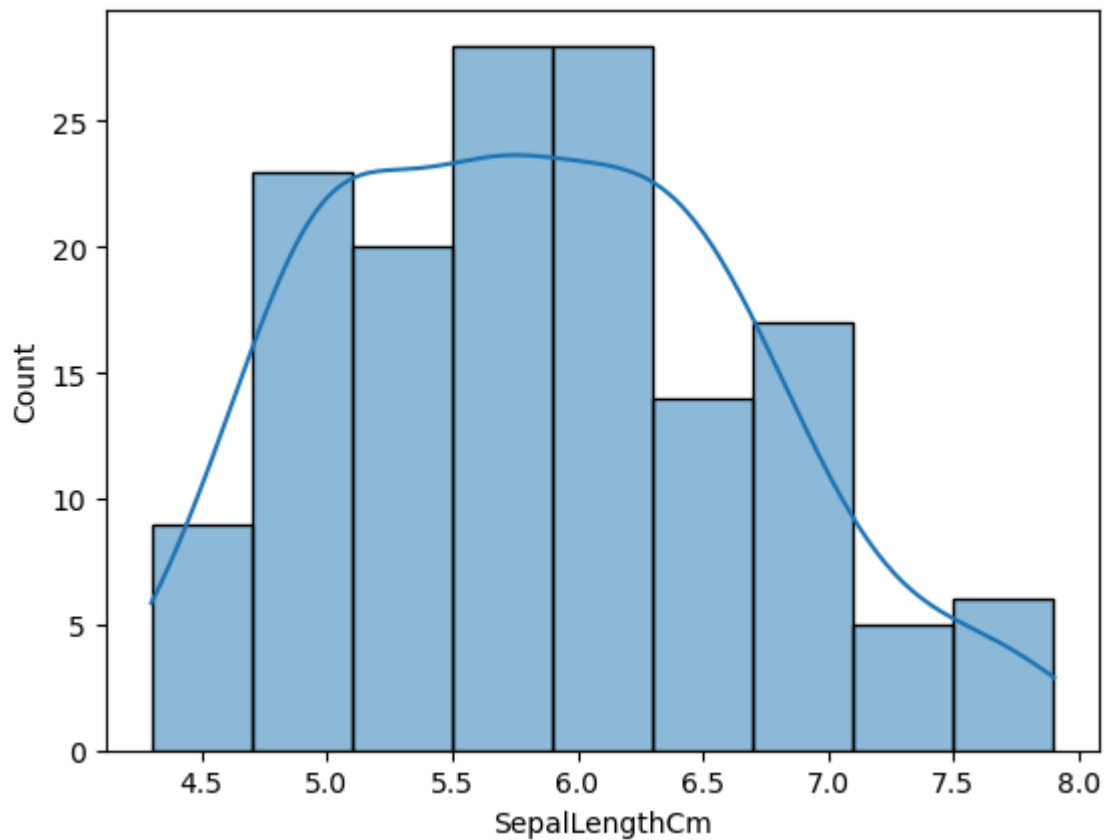
The features in the dataset are as follows :

1. Sepal length : float64
2. Sepal width : float64
3. Petal length : float64
4. Petal width : float64
5. Species : object

```
In [10]: sns.histplot(data=df,x='SepalLengthCm',kde=True)
```

```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

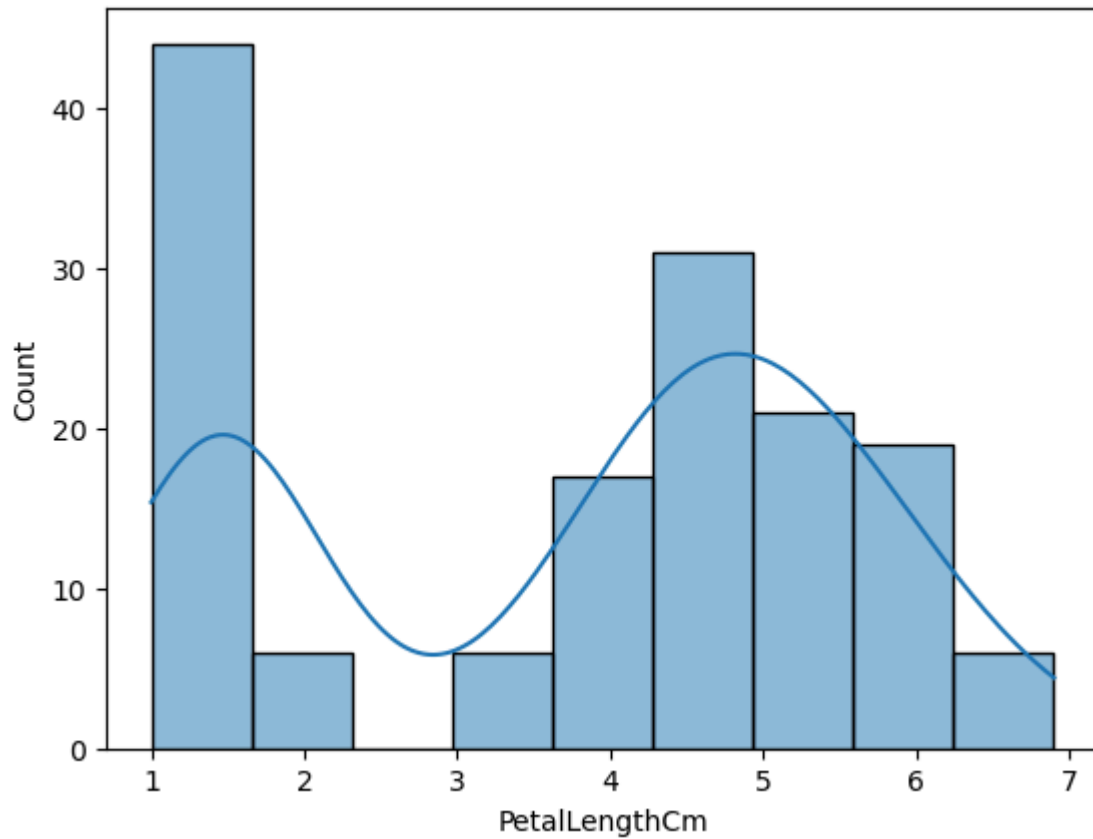
```
Out[10]: <Axes: xlabel='SepalLengthCm', ylabel='Count'>
```



```
In [11]: sns.histplot(data=df, x='PetalLengthCm', kde=True)
```

```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```

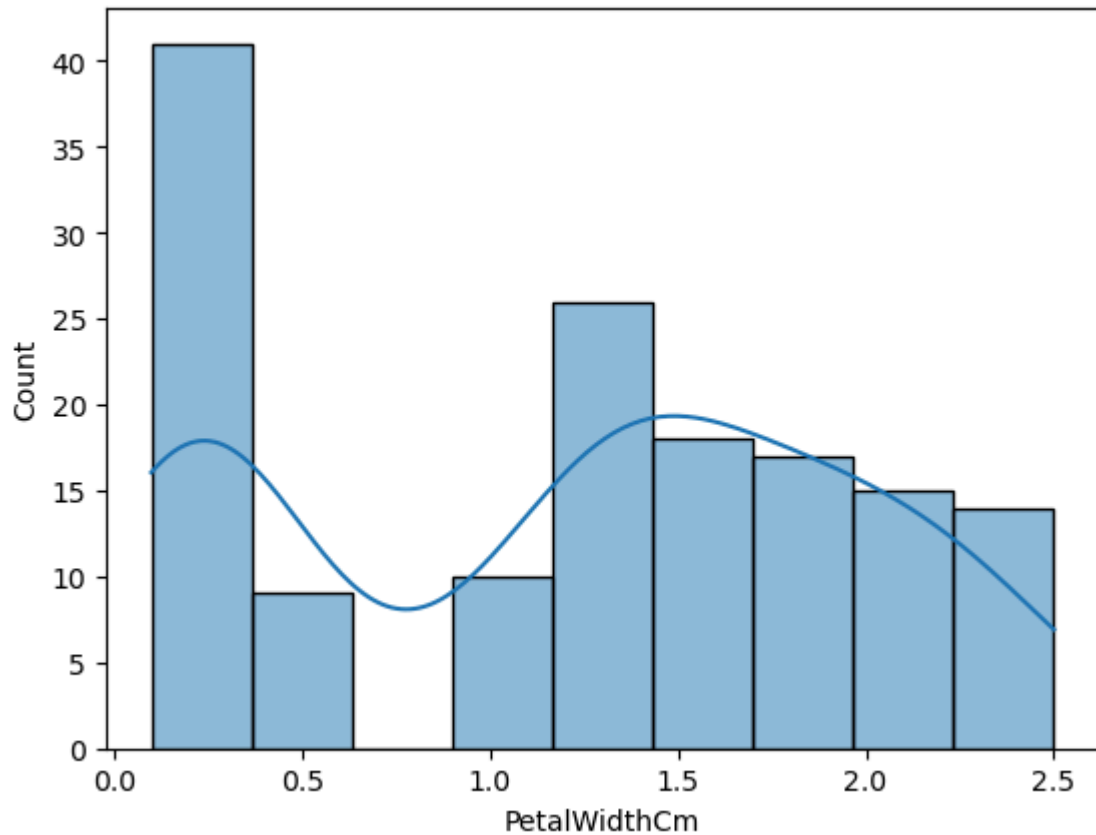
```
Out[11]: <Axes: xlabel='PetalLengthCm', ylabel='Count'>
```



```
In [12]: sns.histplot(data=df,x='PetalWidthCm',kde=True)
```

```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:  
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f  
uture version. Convert inf values to NaN before operating instead.  
  with pd.option_context('mode.use_inf_as_na', True):
```

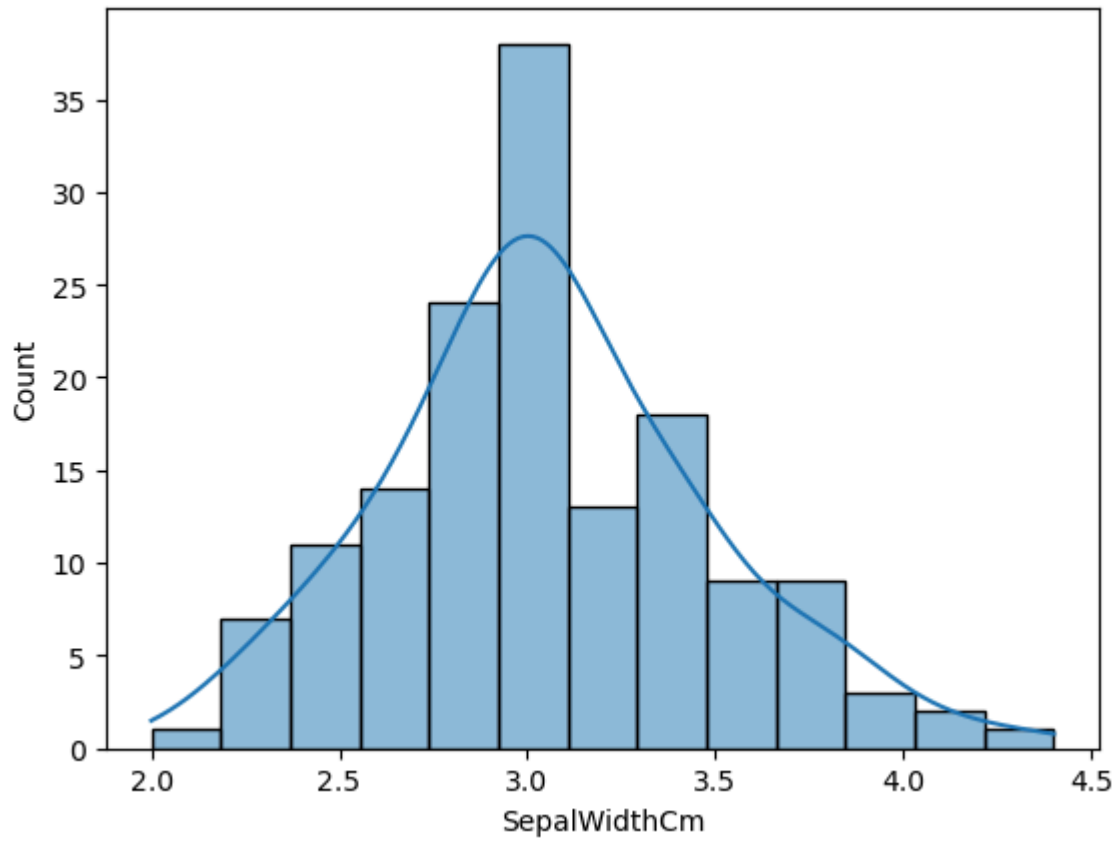
```
Out[12]: <Axes: xlabel='PetalWidthCm', ylabel='Count'>
```



```
In [13]: sns.histplot(data=df,x='SepalWidthCm',kde=True)
```

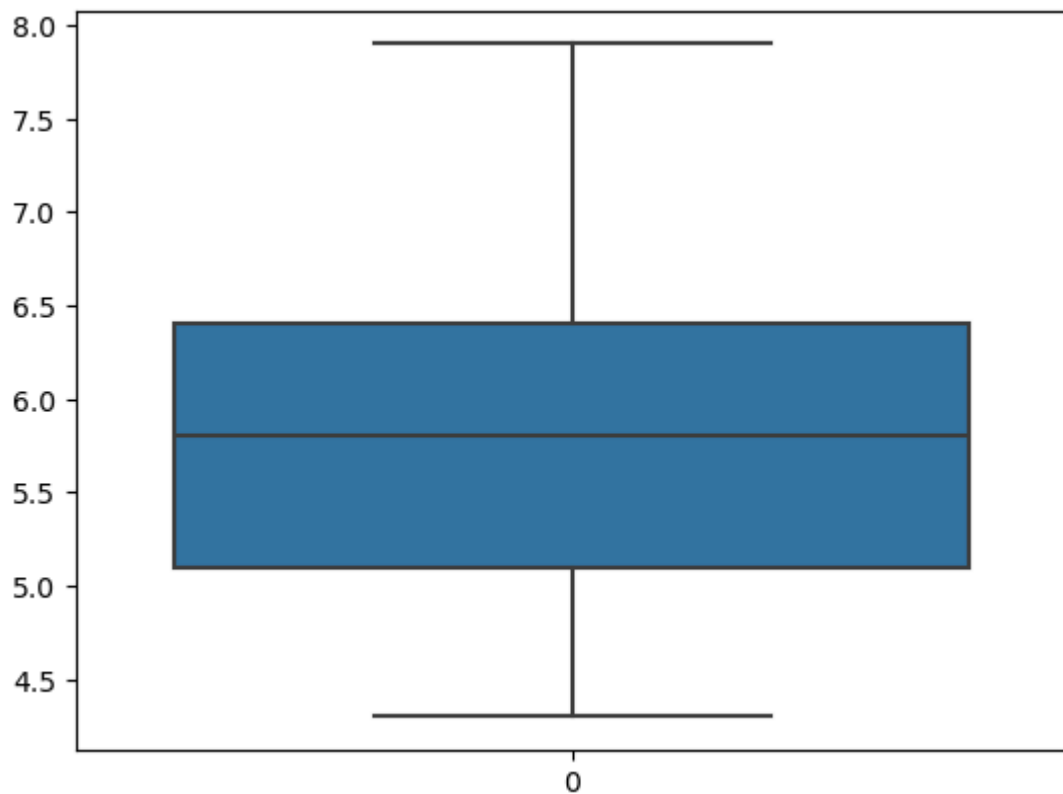
```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:  
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f  
uture version. Convert inf values to NaN before operating instead.  
  with pd.option_context('mode.use_inf_as_na', True):
```

```
Out[13]: <Axes: xlabel='SepalWidthCm', ylabel='Count'>
```



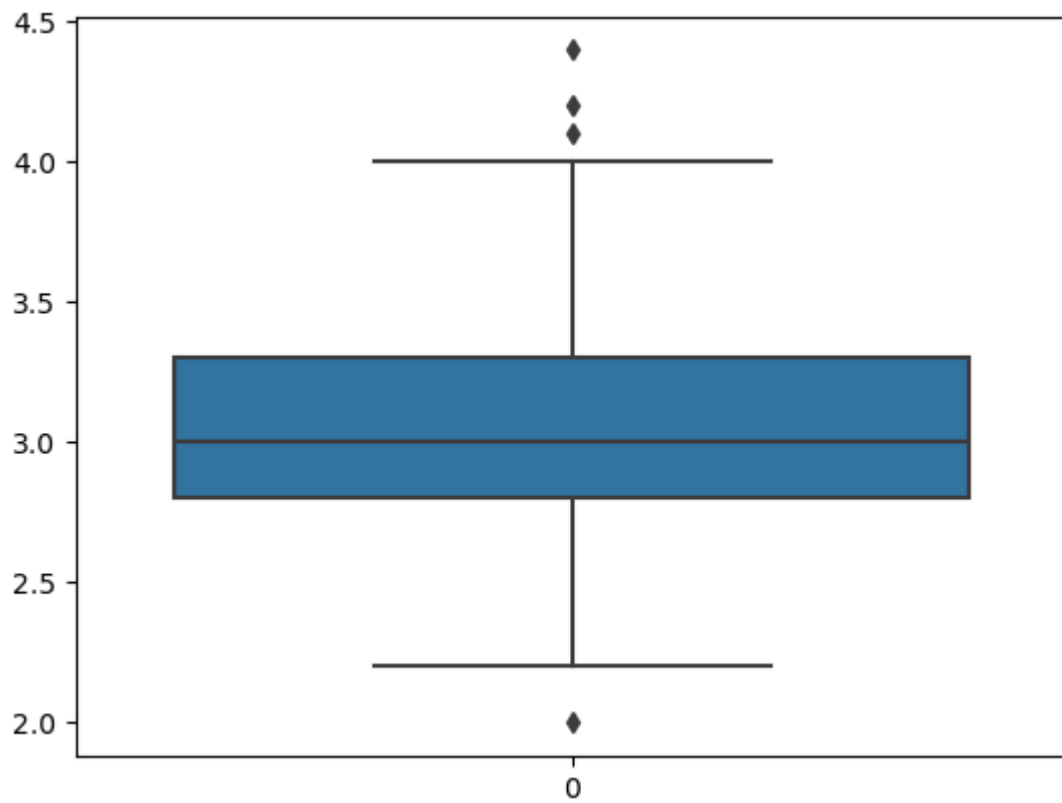
```
In [14]: sns.boxplot(df['SepalLengthCm'])
```

```
Out[14]: <Axes: >
```



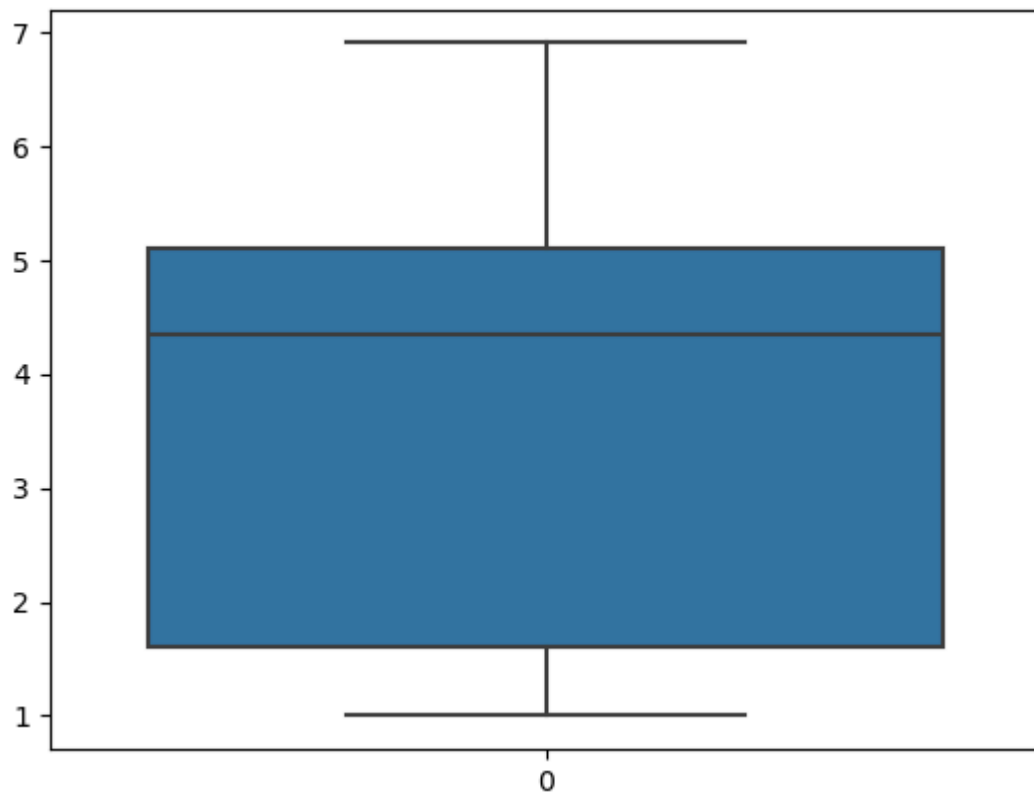
```
In [15]: sns.boxplot(df['SepalWidthCm'])
```

Out[15]: <Axes: >



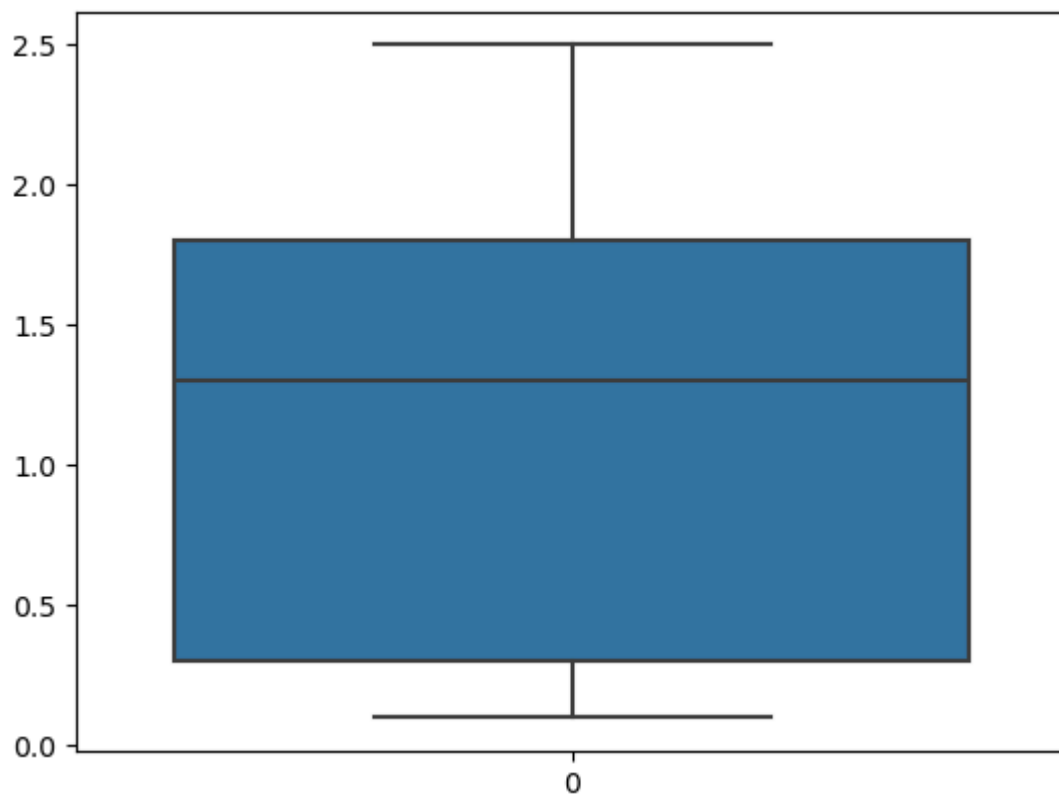
```
In [16]: sns.boxplot(df['PetalLengthCm'])
```

Out[16]: <Axes: >



```
In [17]: sns.boxplot(df['PetalWidthCm'])
```

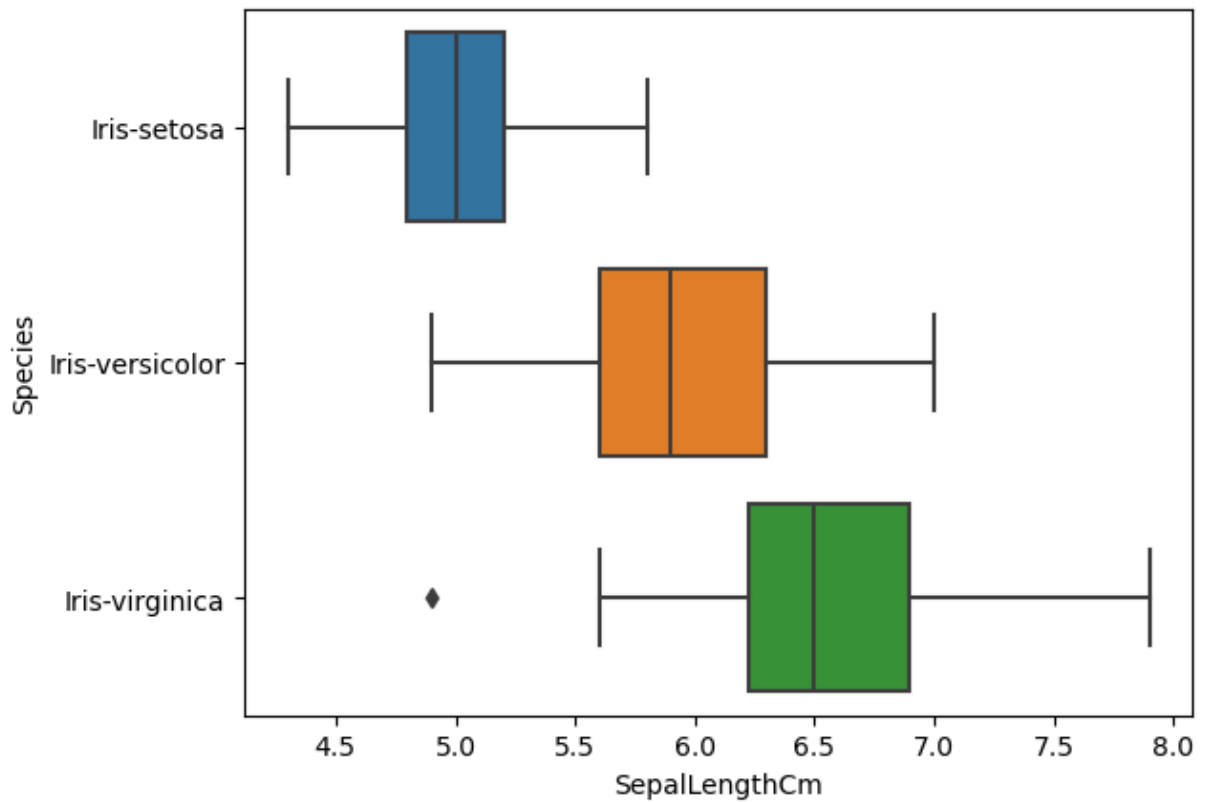
```
Out[17]: <Axes: >
```



```
In [18]: sns.boxplot(x='SepalLengthCm',y='Species',data=df)
```

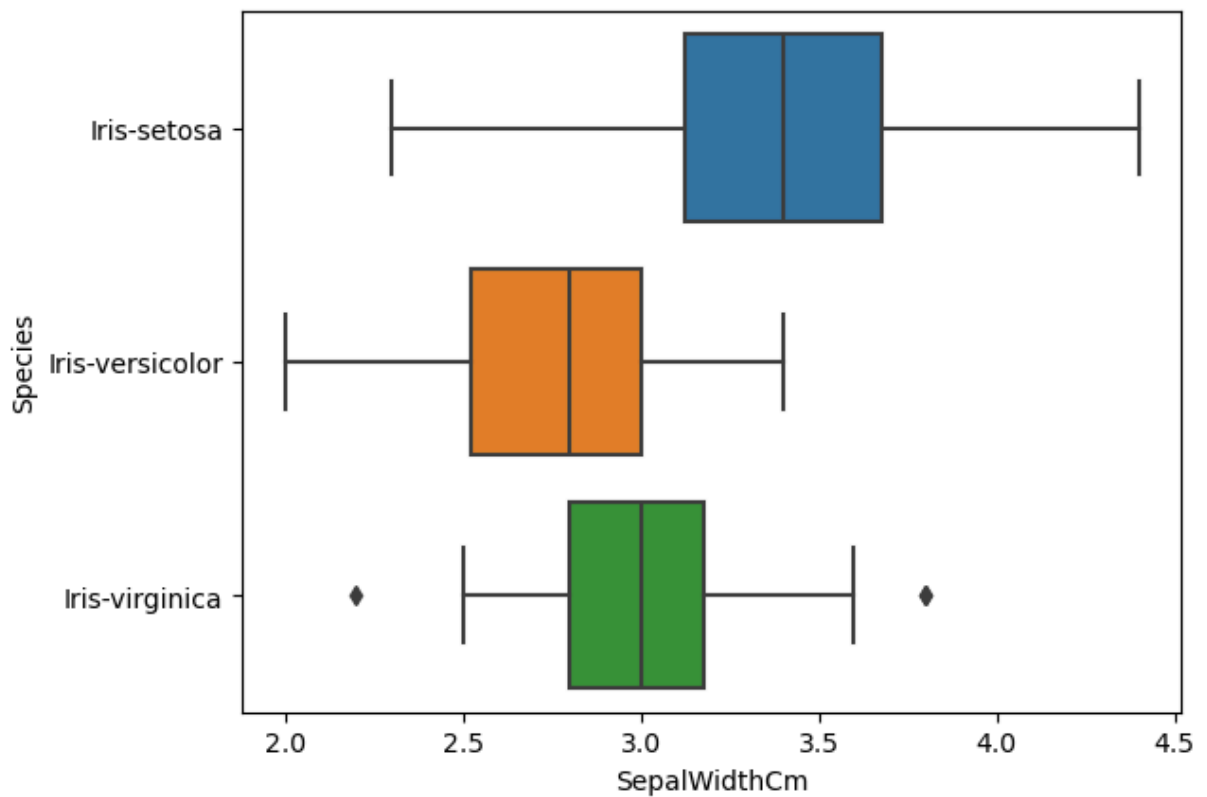


Out[18]: <Axes: xlabel='SepalLengthCm', ylabel='Species'>



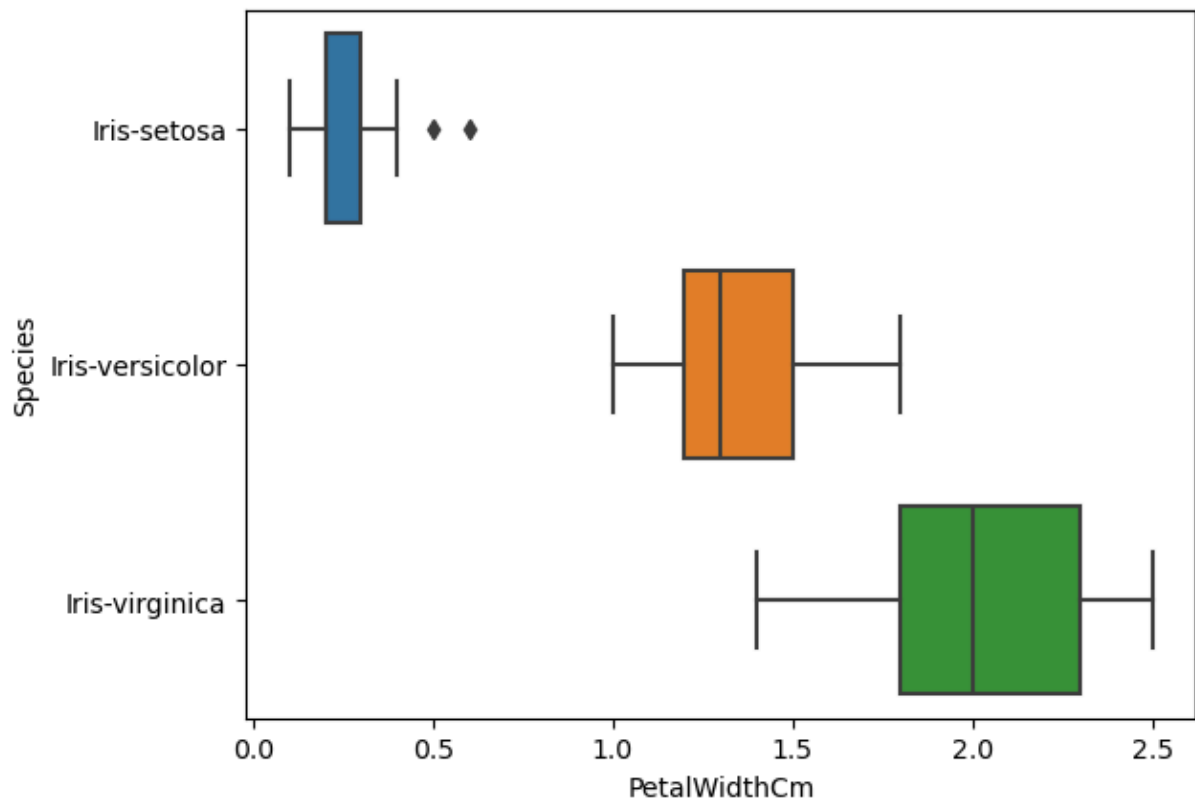
In [19]: `sns.boxplot(x='SepalWidthCm',y='Species',data=df)`

Out[19]: <Axes: xlabel='SepalWidthCm', ylabel='Species'>



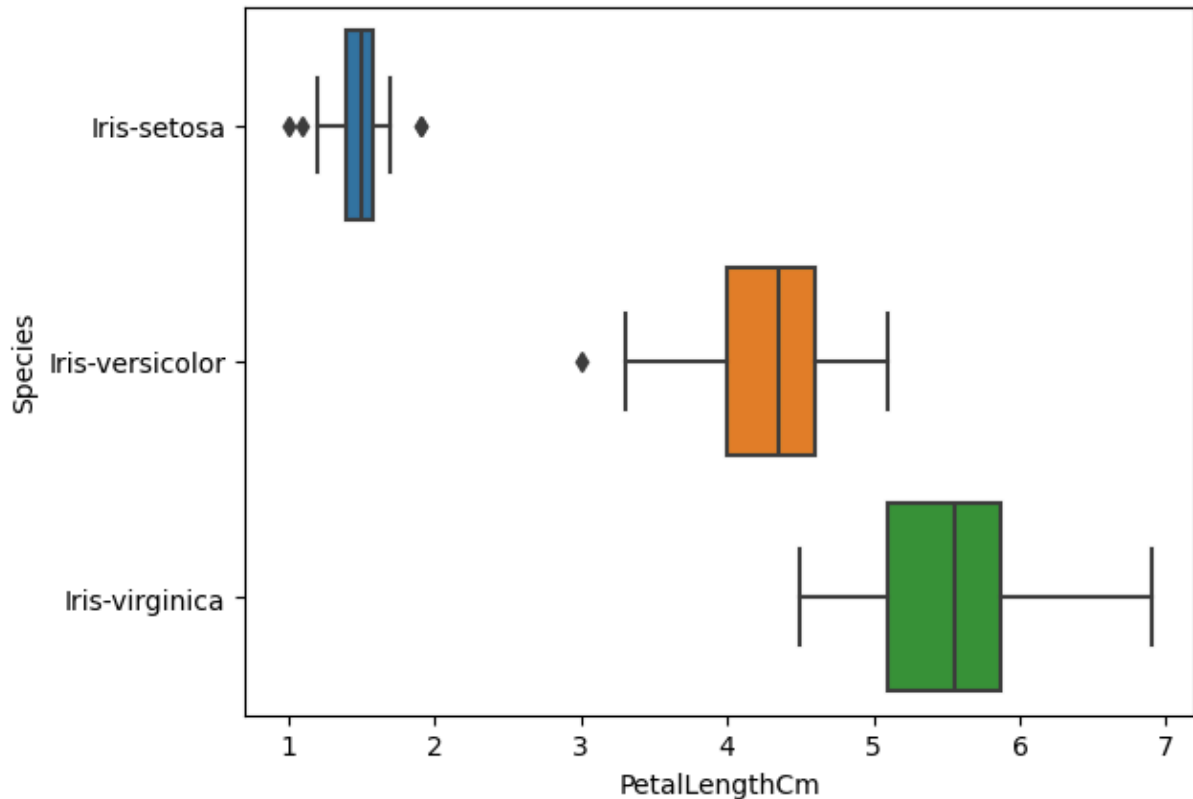
```
In [20]: sns.boxplot(x='PetalWidthCm',y='Species',data=df)
```

```
Out[20]: <Axes: xlabel='PetalWidthCm', ylabel='Species'>
```



```
In [21]: sns.boxplot(x='PetalLengthCm',y='Species',data=df)
```

```
Out[21]: <Axes: xlabel='PetalLengthCm', ylabel='Species'>
```

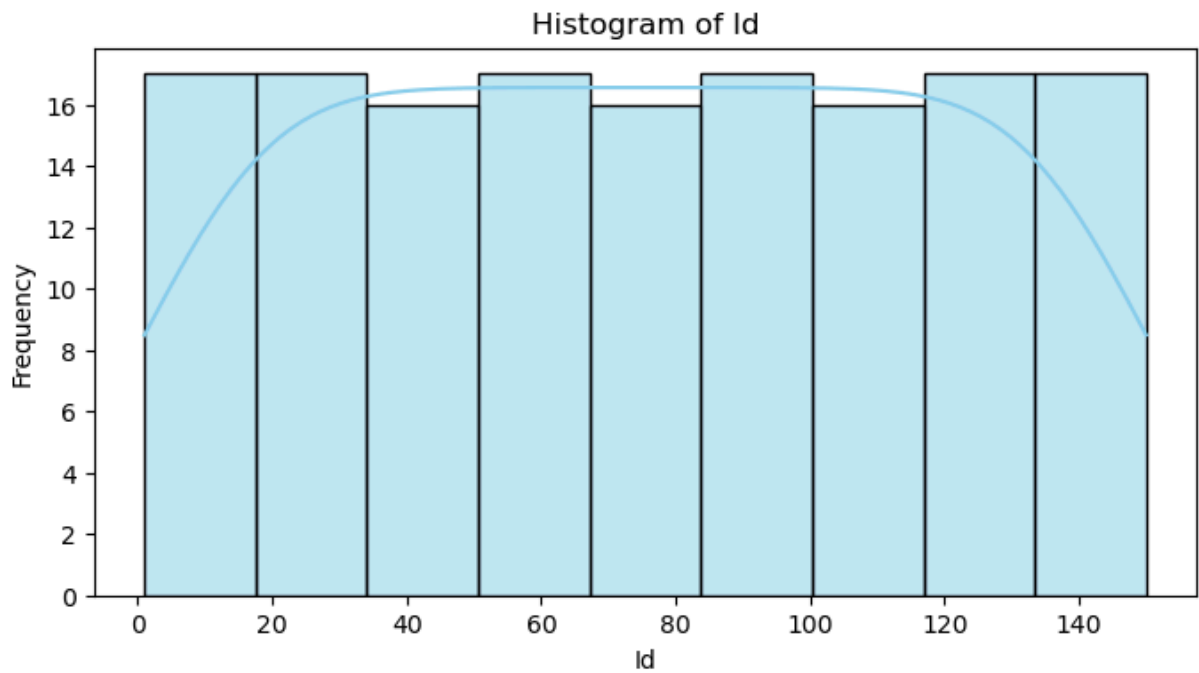


```
In [22]: numeric_columns = df.select_dtypes(include=np.number).columns.tolist()
for column in numeric_columns:
    plt.figure(figsize=(8, 4))
    sns.histplot(df[column], kde=True, color='skyblue')
    plt.title(f'Histogram of {column}')
    plt.xlabel(column)
    plt.ylabel('Frequency')
    plt.show()

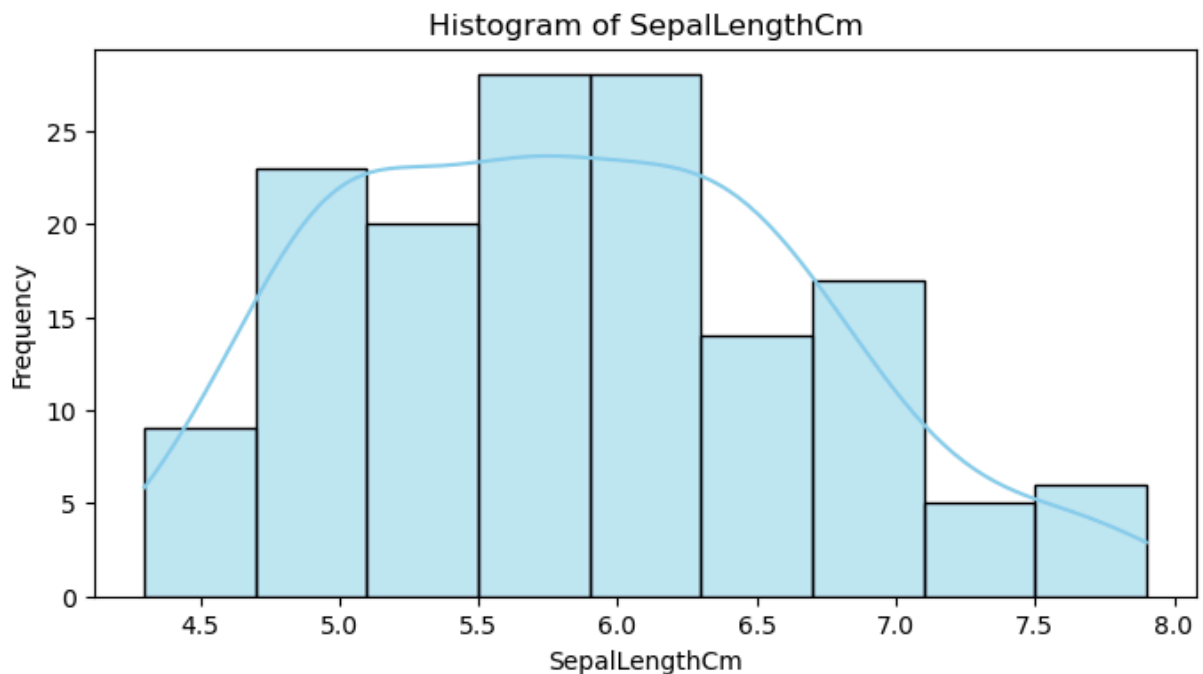
    skewness = skew(df[column])
    print(f"Skewness of {column}: {skewness}")

    if skewness > 0:
        print(f"{column} is positively skewed.")
    elif skewness < 0:
        print(f"{column} is negatively skewed.")
    else:
        print(f"{column} is normally distributed.")
```

/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/\_oldcore.py:1119:  
FutureWarning: use\_inf\_as\_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.  
with pd.option\_context('mode.use\_inf\_as\_na', True):

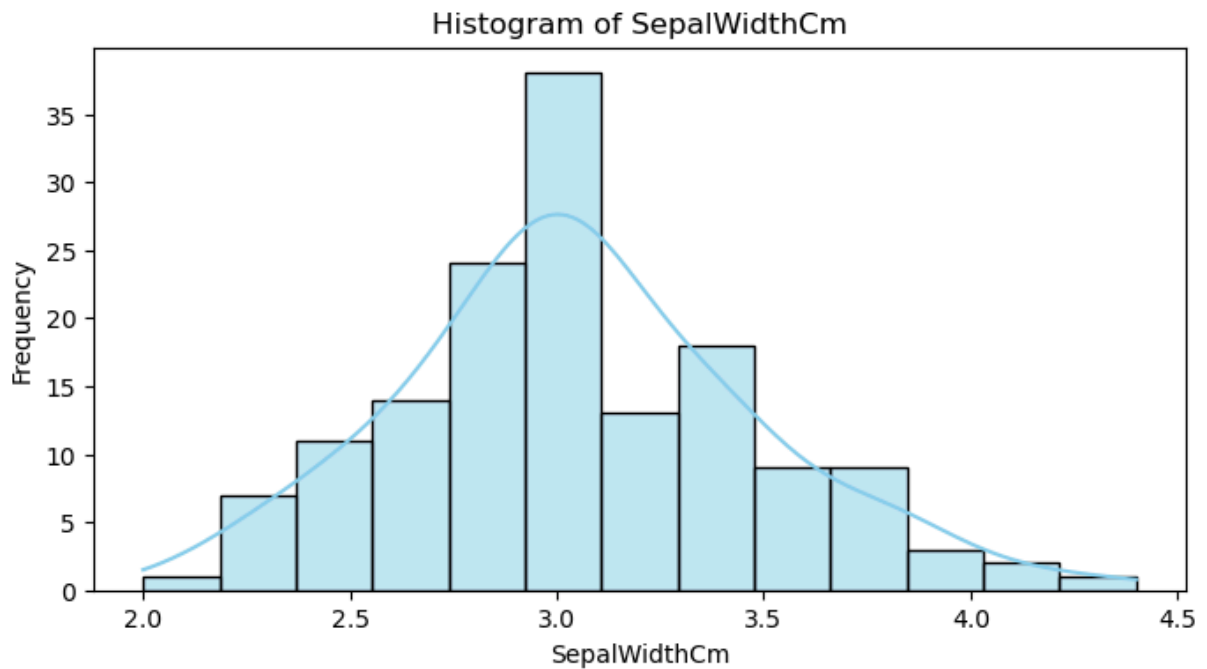


```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:  
FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.  
with pd.option_context('mode.use_inf_as_na', True):  
Skewness of Id: 0.0  
Id is normally distributed.
```



```
Skewness of SepalLengthCm: 0.3117530585022963  
SepalLengthCm is positively skewed.
```

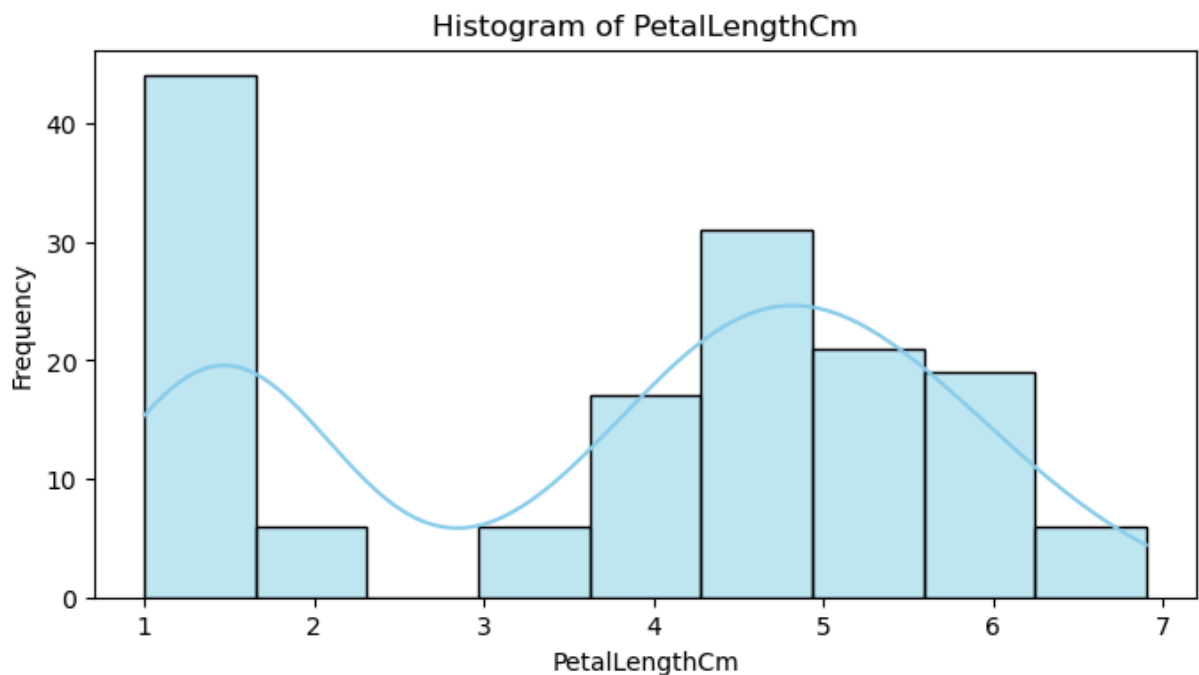
```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:  
FutureWarning: use_inf_as_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.  
with pd.option_context('mode.use_inf_as_na', True):
```



Skewness of SepalWidthCm: 0.330702812773315

SepalWidthCm is positively skewed.

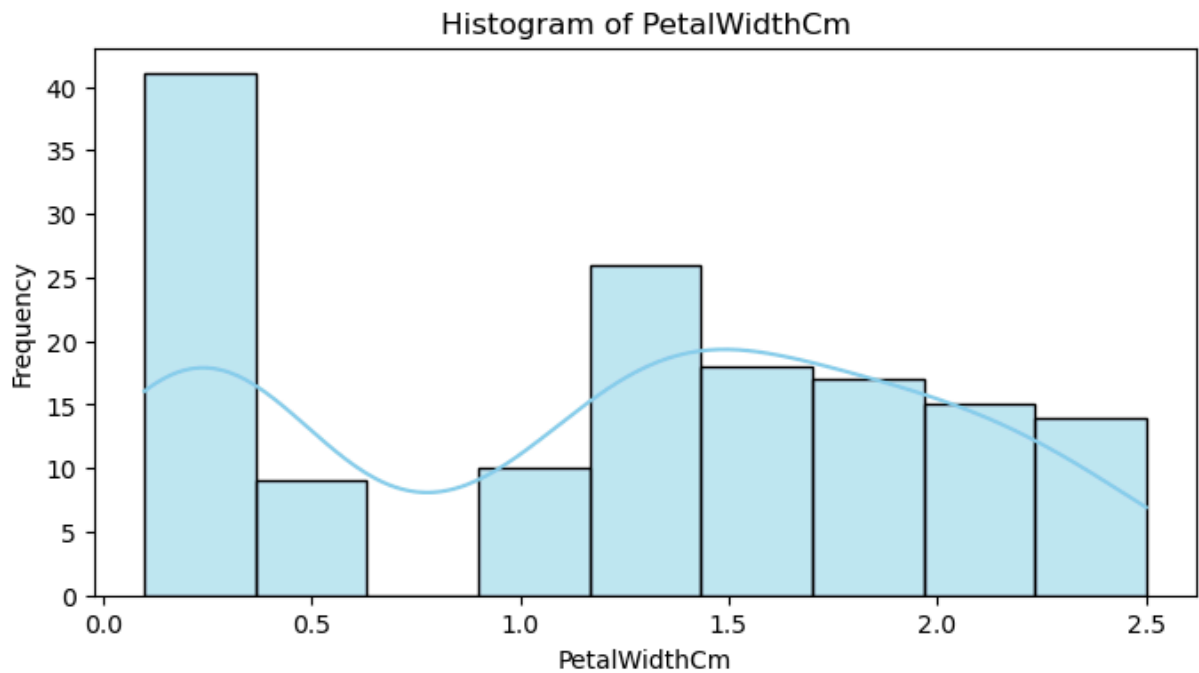
```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:  
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f  
uture version. Convert inf values to NaN before operating instead.  
with pd.option_context('mode.use_inf_as_na', True):
```



Skewness of PetalLengthCm: -0.2717119501716388

PetalLengthCm is negatively skewed.

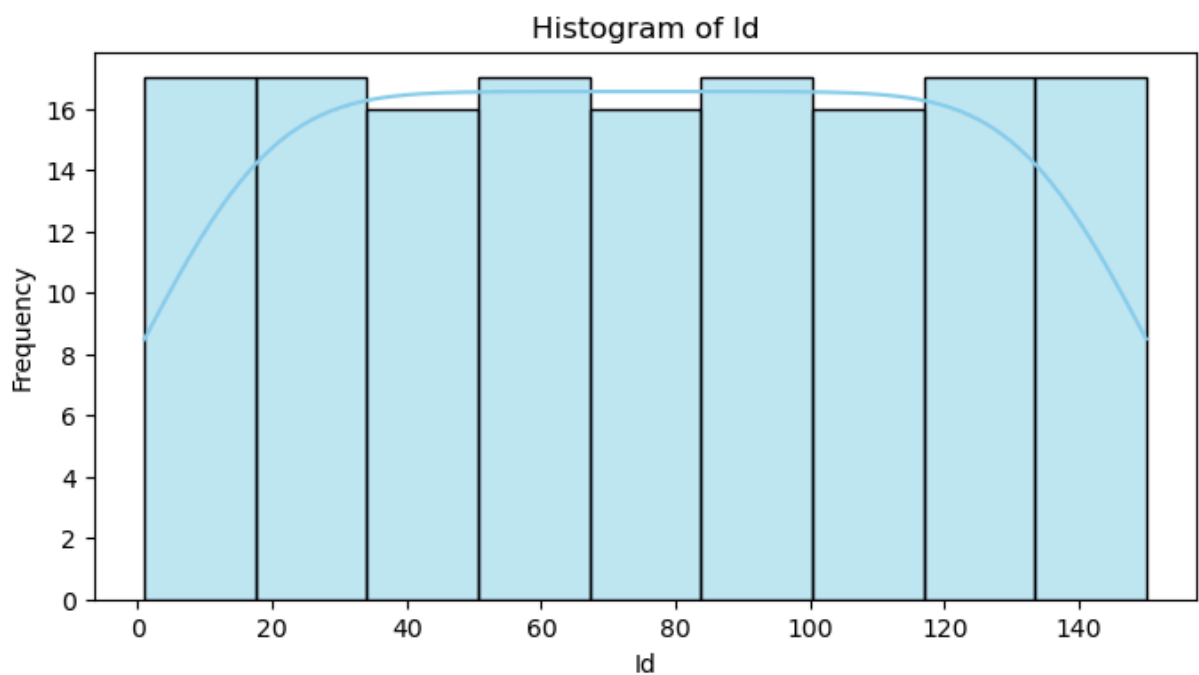
```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:  
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f  
uture version. Convert inf values to NaN before operating instead.  
with pd.option_context('mode.use_inf_as_na', True):
```



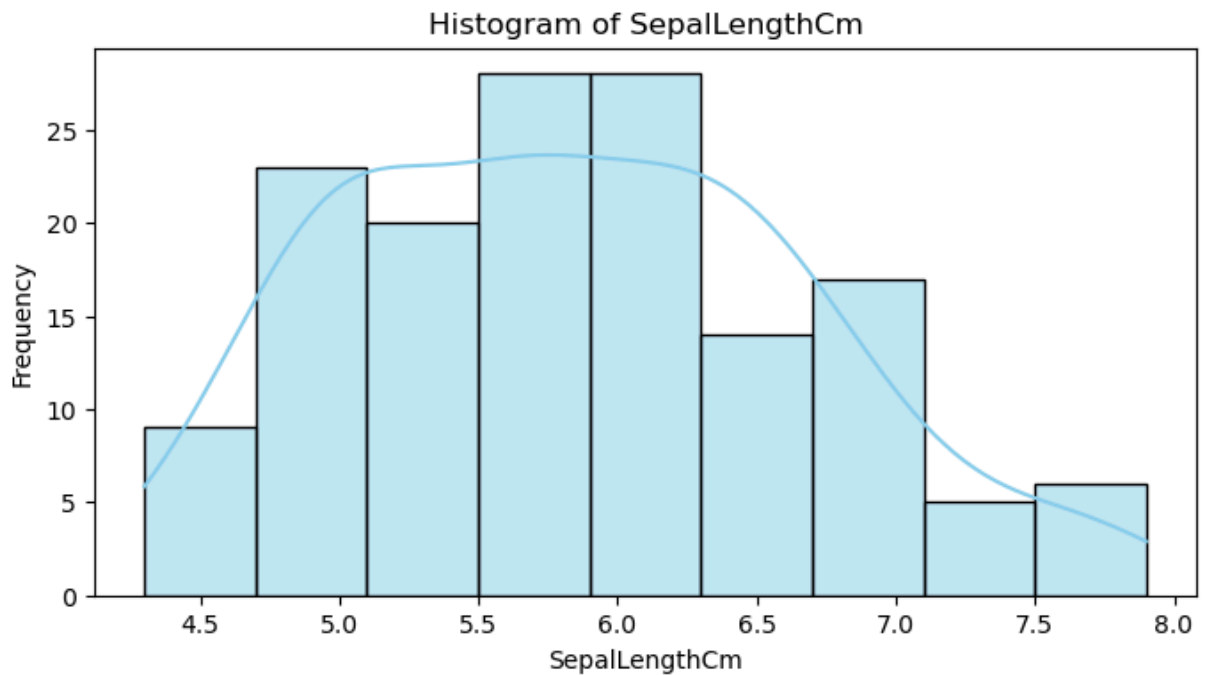
Skewness of PetalWidthCm:  $-0.10394366626751729$   
PetalWidthCm is negatively skewed.

```
In [25]: for column in df.columns[:-1]: # Exclude the target variable
plt.figure(figsize=(8, 4))
sns.histplot(df[column], kde=True, color='skyblue')
plt.title(f'Histogram of {column}')
plt.xlabel(column)
plt.ylabel('Frequency')
plt.show()
```

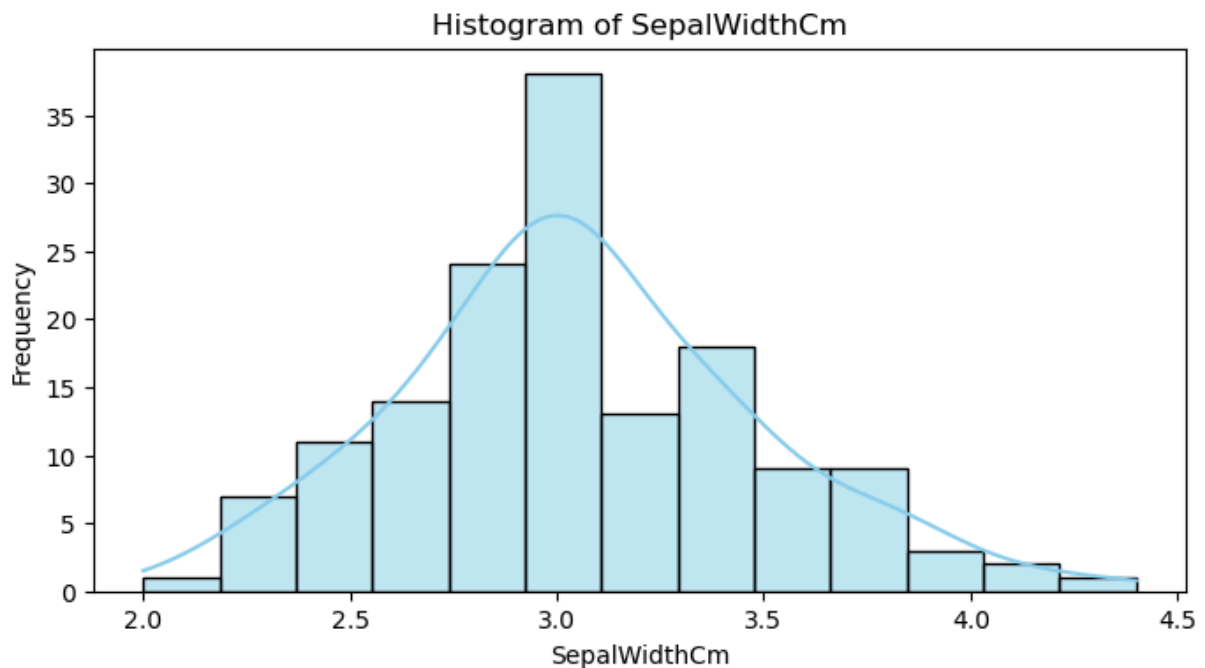
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/\_oldcore.py:1119:  
FutureWarning: use\_inf\_as\_na option is deprecated and will be removed in a future version. Convert inf values to NaN before operating instead.  
with pd.option\_context('mode.use\_inf\_as\_na', True):



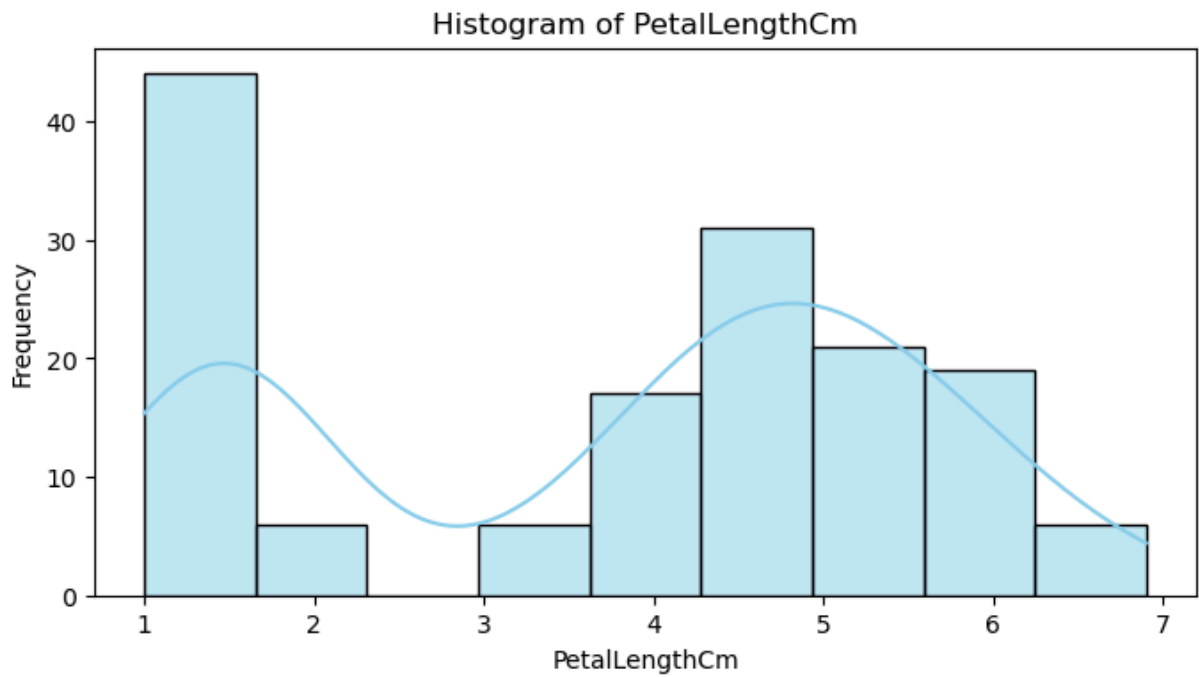
```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f
uture version. Convert inf values to NaN before operating instead.
with pd.option_context('mode.use_inf_as_na', True):
```



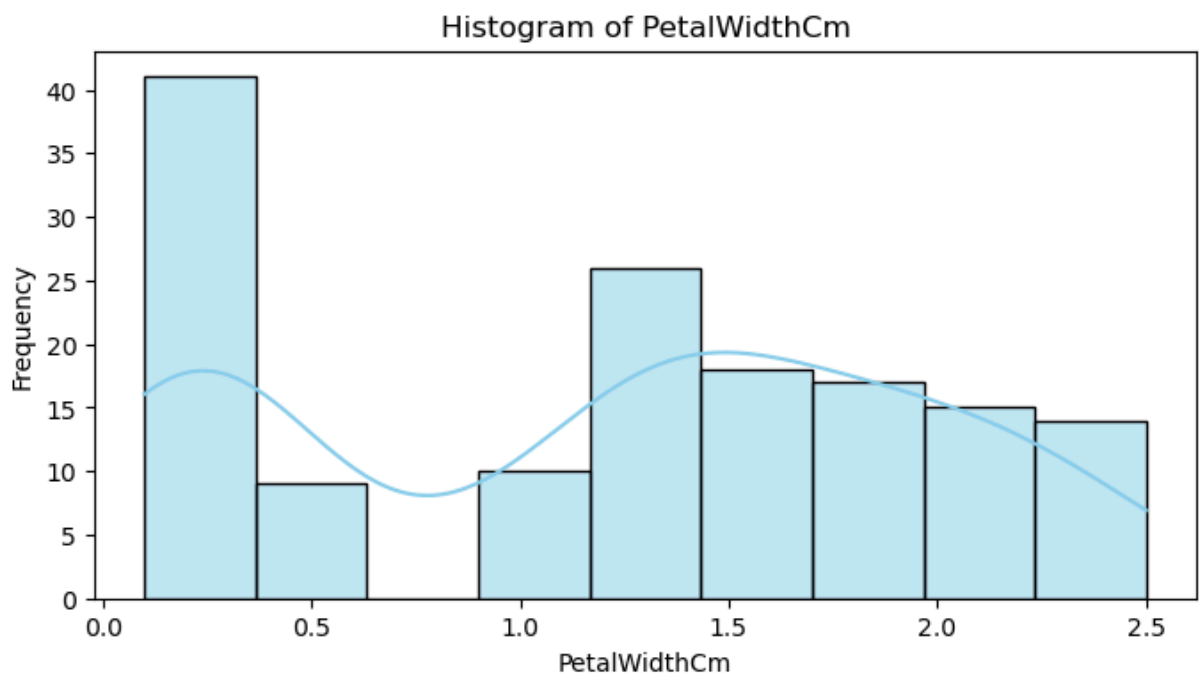
```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f
uture version. Convert inf values to NaN before operating instead.
with pd.option_context('mode.use_inf_as_na', True):
```



```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f
uture version. Convert inf values to NaN before operating instead.
with pd.option_context('mode.use_inf_as_na', True):
```



```
/Users/anvi/anaconda3/lib/python3.11/site-packages/seaborn/_oldcore.py:1119:  
FutureWarning: use_inf_as_na option is deprecated and will be removed in a f  
uture version. Convert inf values to NaN before operating instead.  
with pd.option_context('mode.use_inf_as_na', True):
```



In [ ]: