- Import Essential Libraries

In [ ]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.impute import SimpleImputer

sns.set_theme(style="darkgrid", palette= 'muted')
%matplotlib inline
```

- File path

In [ ]:

```
file_path = '/content/drive/MyDrive/Datasets/Internship/Task-01/Iris.csv'
```

- Convert CSV file to Pandas Dataframe

In [ ]:

```
raw = pd.read_csv(file_path)
raw.head()
```

Out[ ]:

	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

- Basic Info about file
- Check if there is any missing value exist or not

In [ ]:

```
data = raw
data.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 [ ]:

```
data.describe()
```

Out[ ]:

	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 [ ]:

```
data.columns
```

Out[ ]:

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

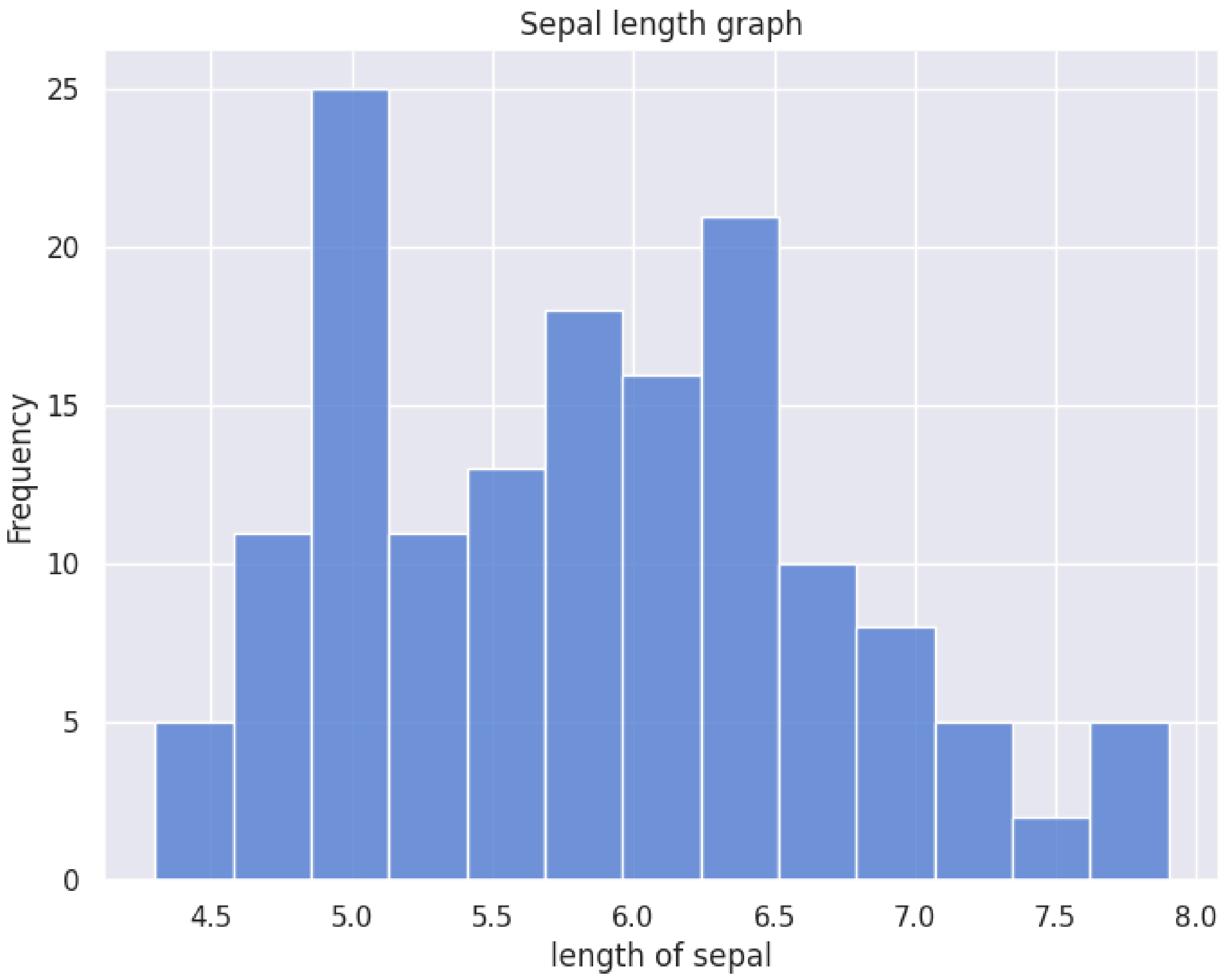
- Visualize the data using Seaborn

In [ ]:

```
fig, ax = plt.subplots(figsize=(8, 6))
sns.histplot(data=data,ax=ax, x="SepalLengthCm", bins='sqrt')
plt.title("Sepal length graph")
plt.xlabel("length of sepal")
plt.ylabel("Frequency")
```

Out[ ]:

```
Text(0, 0.5, 'Frequency')
```

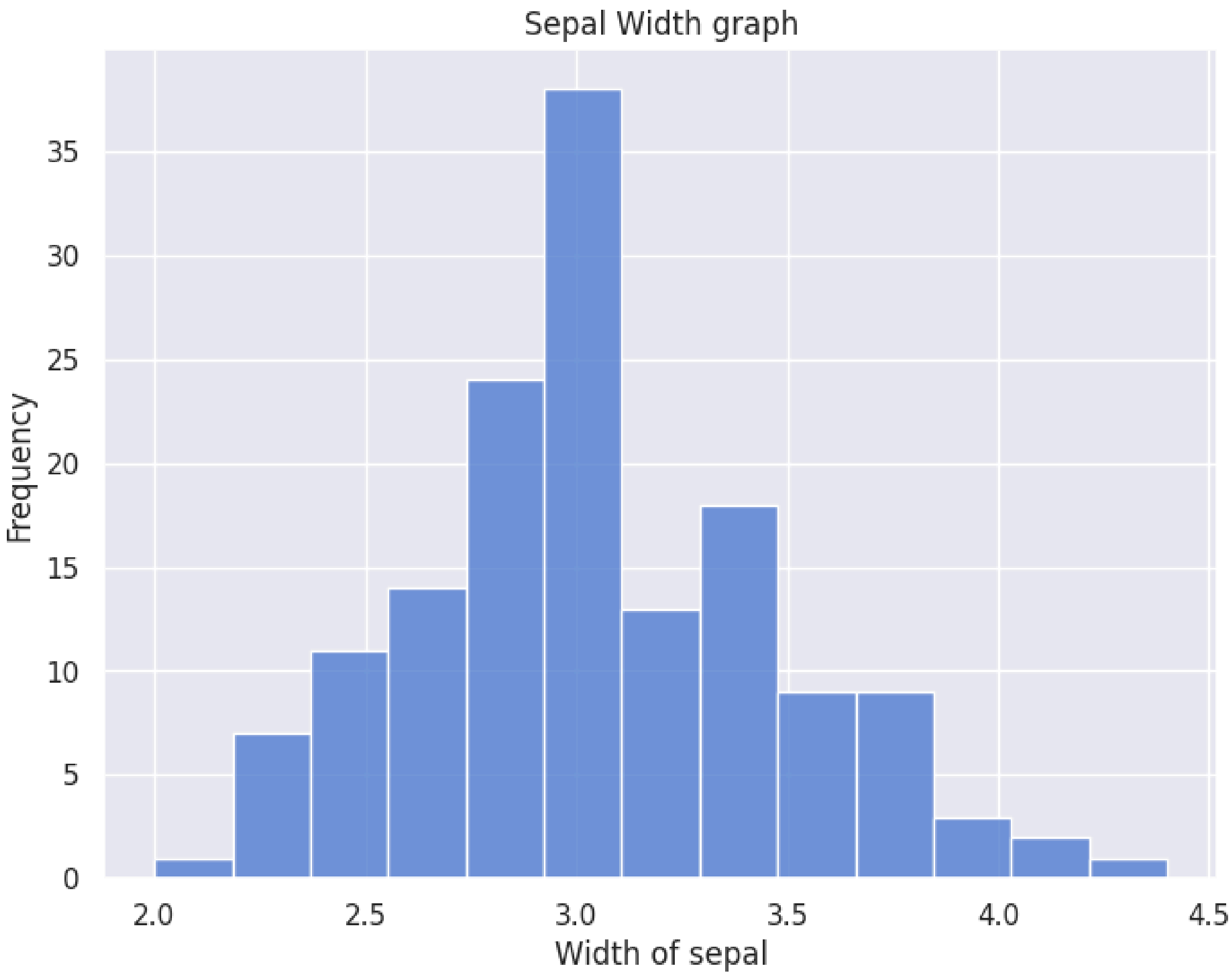


In [ ]:

```
fig, ax = plt.subplots(figsize=(8, 6))
sns.histplot(data=data,ax=ax, x="SepalWidthCm", bins='sqrt')
plt.title("Sepal Width graph")
plt.xlabel("Width of sepal")
plt.ylabel("Frequency")
```

Out[ ]:

```
Text(0, 0.5, 'Frequency')
```

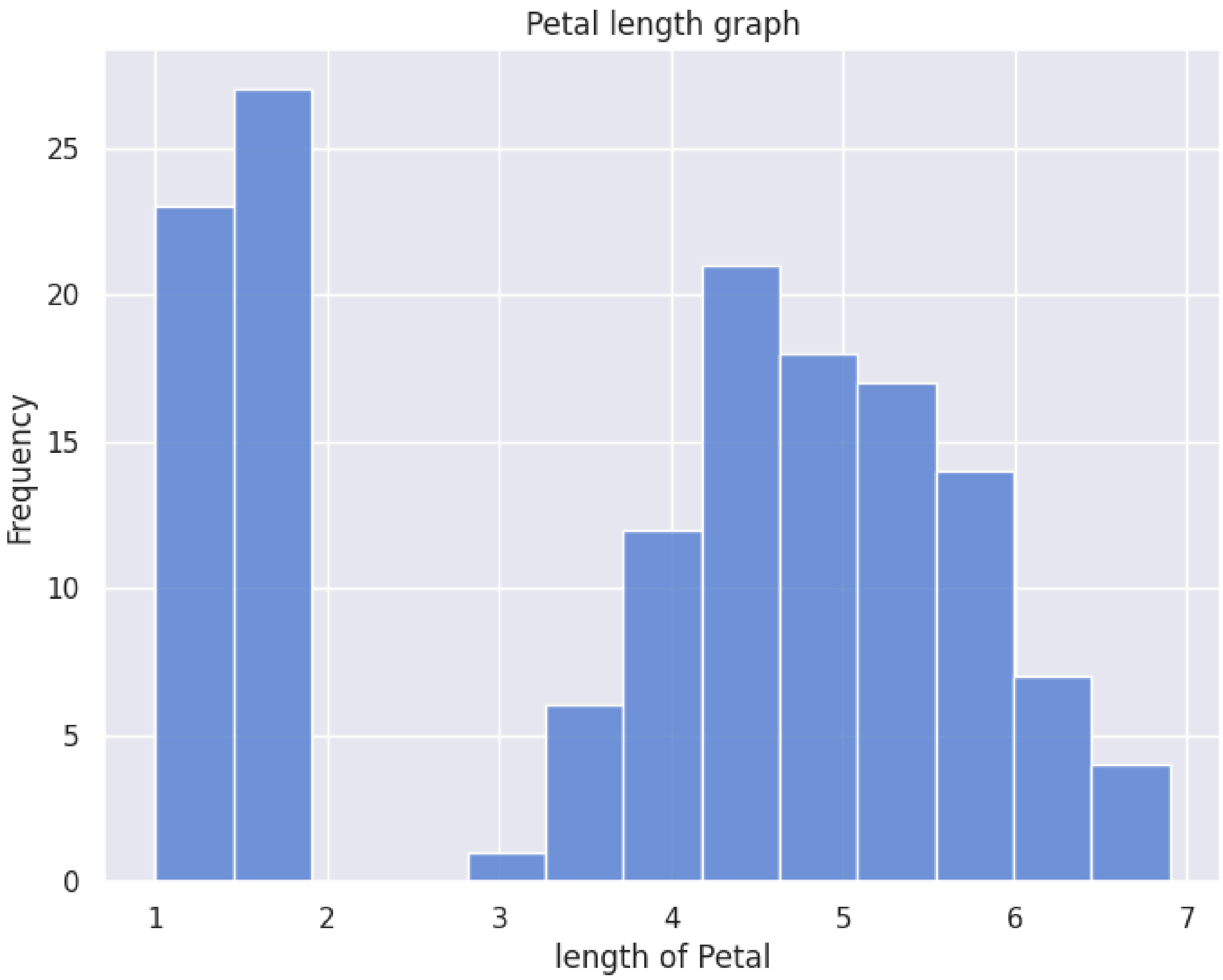


In [ ]:

```
fig, ax = plt.subplots(figsize=(8, 6))
sns.histplot(data=data,ax=ax, x="PetalLengthCm", bins='sqrt')
plt.title("Petal length graph")
plt.xlabel("length of Petal")
plt.ylabel("Frequency")
```

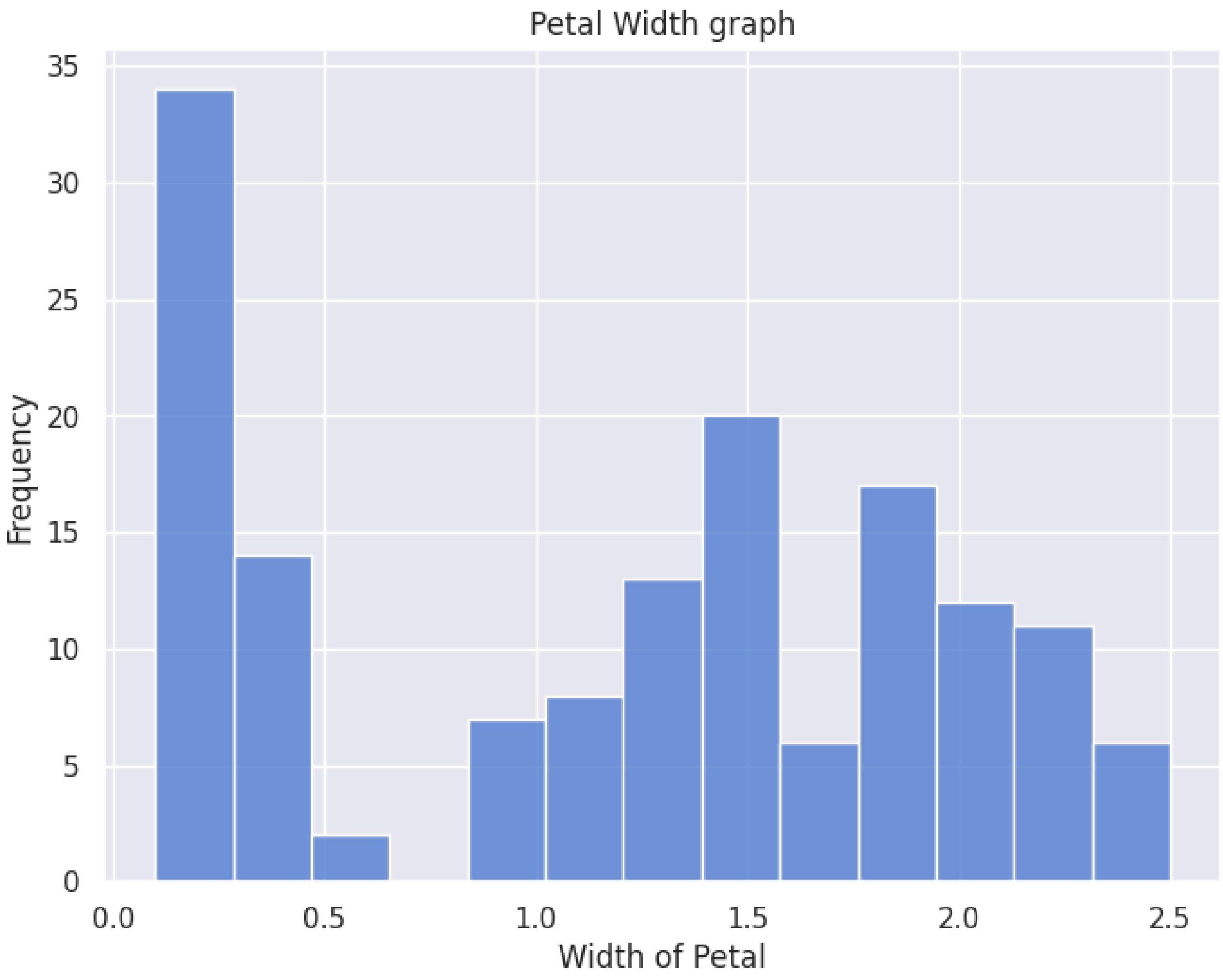
Out[ ]:

```
Text(0, 0.5, 'Frequency')
```



```
In [ ]: fig, ax = plt.subplots(figsize=(8, 6))
        sns.histplot(data=data,ax=ax, x="PetalWidthCm", bins='sqrt')
        plt.title("Petal Width graph")
        plt.xlabel("Width of Petal")
        plt.ylabel("Frequency")
```

```
Out[ ]: Text(0, 0.5, 'Frequency')
```



```
In [ ]:
```