```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

data1 = pd.read_csv(r"C:\Users\Jayditya\Downloads\Lab-
20250210T092855Z-001\Lab\Experiments\Datasets\13Iris.csv")
data1.head()
```

```
   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
```

```
print(data1.columns)
```

```
Index(['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm',
'PetalWidthCm',
       'species'],
      dtype='object')
```

```
#anotherway
column = list(data1)
print(column)
```

```
['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm',
'PetalWidthCm', 'species']
```

```
data1.info()
```

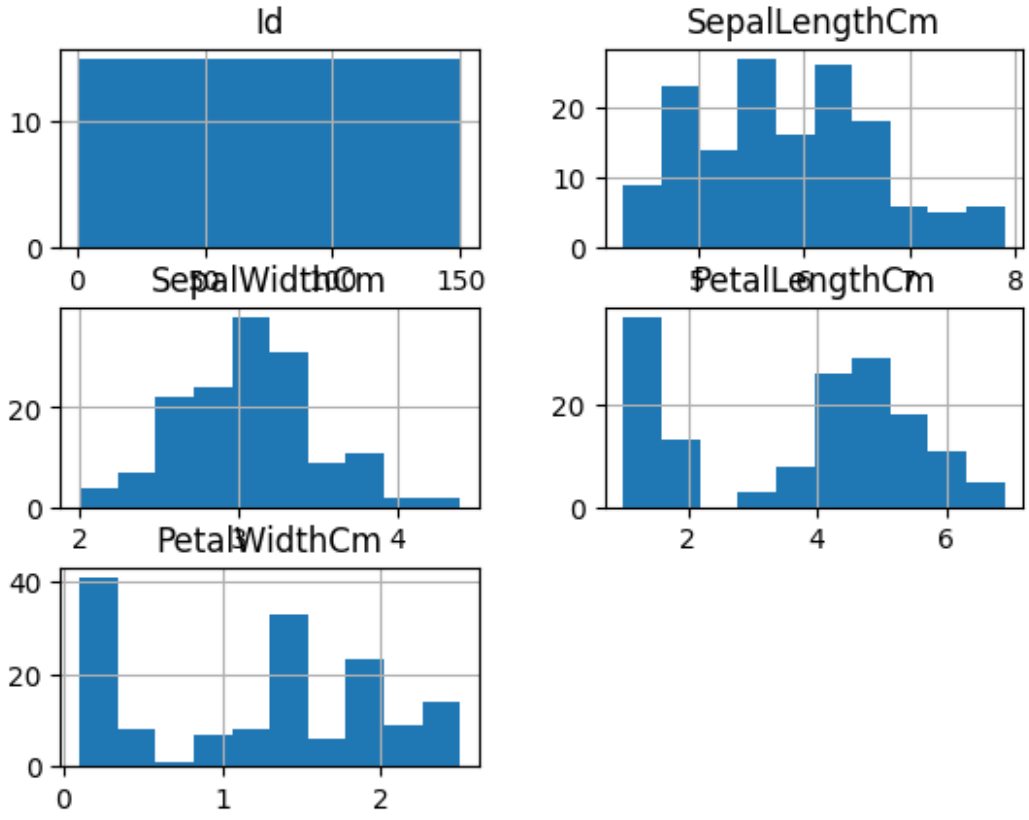
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Id             150 non-null    int64
 1   SepalLengthCm  150 non-null    float64
 2   SepalWidthCm   150 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
```

```
data1.dtypes
```

```
Id                 int64
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm    float64
PetalWidthCm     float64
species           object
dtype: object
```

```
data1.hist()
```

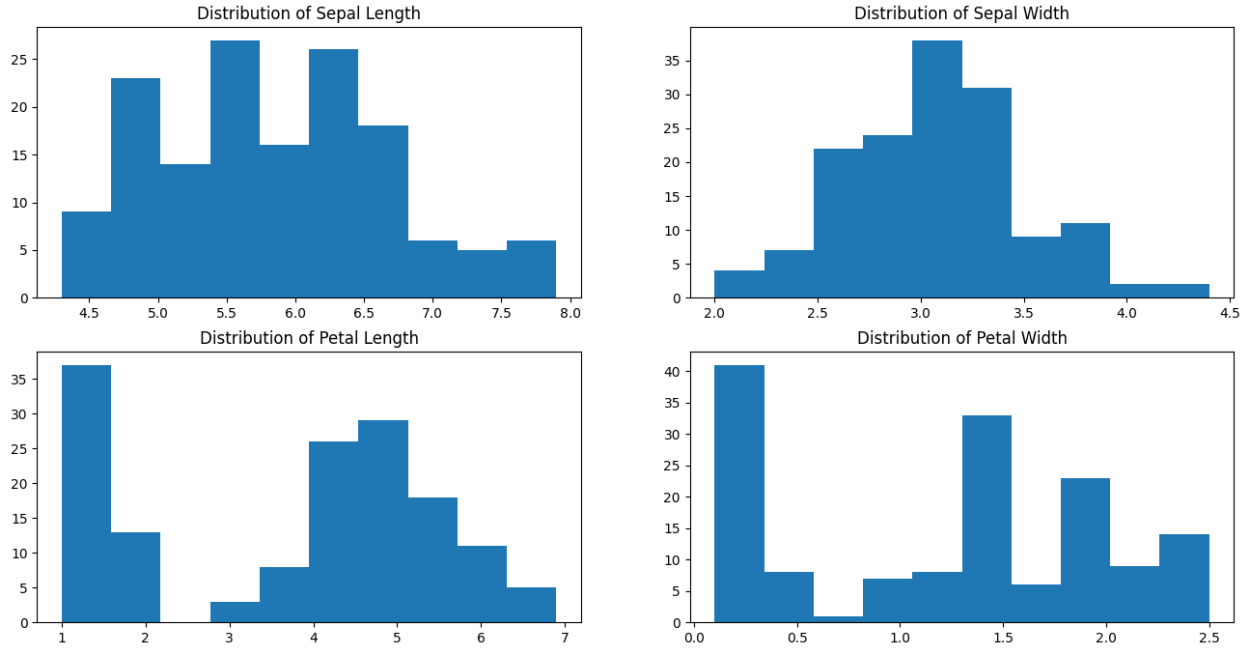
```
array([[<Axes: title={'center': 'Id'}>,
        <Axes: title={'center': 'SepalLengthCm'}>],
       [<Axes: title={'center': 'SepalWidthCm'}>,
        <Axes: title={'center': 'PetalLengthCm'}>],
       [<Axes: title={'center': 'PetalWidthCm'}>, <Axes: >]],
dtype=object)
```




```
fig, axes = plt.subplots(2, 2, figsize=(16, 8))

axes[0,0].set_title("Distribution of Sepal Length")
axes[0,0].hist(data1["SepalLengthCm"])
```

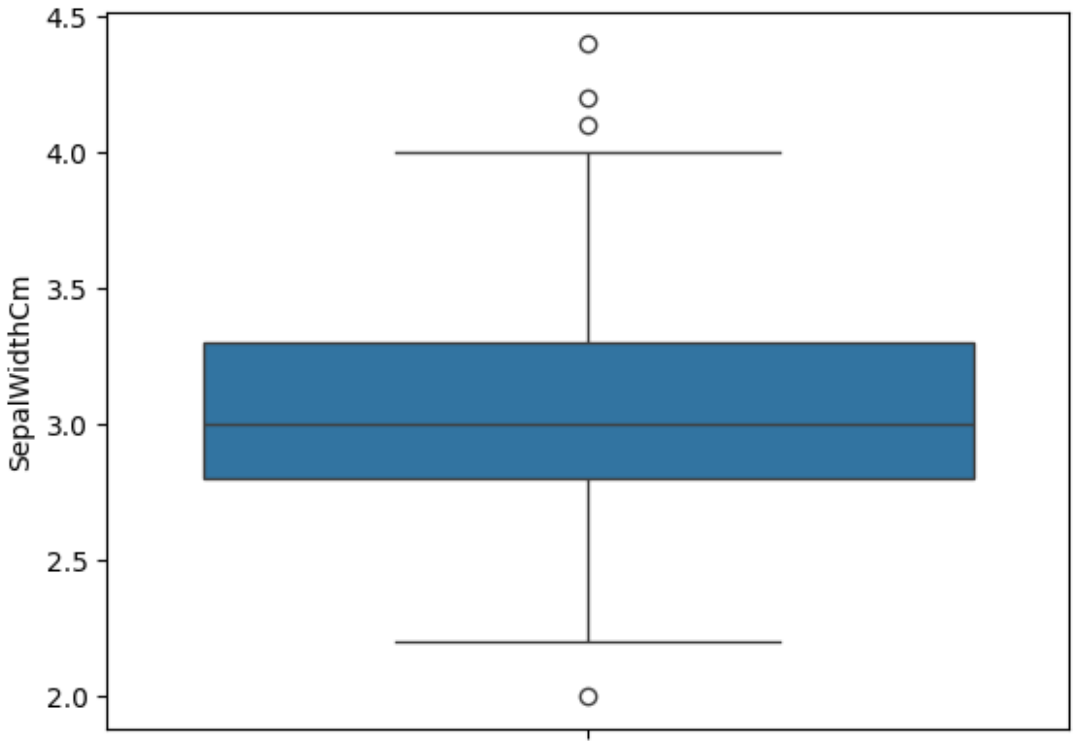
```
axes[0,1].set_title("Distribution of Sepal Width")
axes[0,1].hist(data1["SepalWidthCm"])

axes[1,0].set_title("Distribution of Petal Length")
axes[1,0].hist(data1["PetalLengthCm"])

axes[1,1].set_title("Distribution of Petal Width")
axes[1,1].hist(data1["PetalWidthCm"])
```

```
(array([41.,  8.,  1.,  7.,  8., 33.,  6., 23.,  9., 14.]),
 array([0.1 , 0.34, 0.58, 0.82, 1.06, 1.3 , 1.54, 1.78, 2.02, 2.26,
2.5 ]),
 <BarContainer object of 10 artists>)
```




```
sns.boxplot(data1['SepalWidthCm'])
```

```
<Axes: ylabel='SepalWidthCm'>
```




```
print(np.where(data1['SepalWidthCm']>4.0))
```

```
(array([15, 32, 33]),)
```

```
col = 'SepalWidthCm'

# Calculate IQR
Q1 = data1[col].quantile(0.25)
Q3 = data1[col].quantile(0.75)
IQR = Q3 - Q1

# Define bounds
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

# Get outliers
outliers = data1[(data1[col] < lower_bound) | (data1[col] >
upper_bound)]

# Print info
print(f"Outliers in column '{col}':")
print(outliers)
```

```
Outliers in column 'SepalWidthCm':
    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
```