```
In [1]: import numpy as np
        import pandas as pd
```

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

```
In [5]: df
```

Out[5]:

	sepal length in cm	sepal width in cm	petal length in cm	petal width in cm	class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

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

```
Out[6]: sepal length in cm    0
        sepal width in cm     0
        petal length in cm    0
        petal width in cm     0
        class                 0
        dtype: int64
```

```
In [7]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   sepal length in cm  150 non-null    float64
 1   sepal width in cm   150 non-null    float64
 2   petal length in cm  150 non-null    float64
 3   petal width in cm   150 non-null    float64
 4   class               150 non-null    object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

Features available in Dataset

```
In [8]: np.unique(df["class"])
```

```
Out[8]: array(['Iris-setosa', 'Iris-versicolor', 'Iris-virginica'], dtype=object)
```

```
In [9]: df.describe()
```

Out[9]:

	sepal length in cm	sepal width in cm	petal length in cm	petal width in cm
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

Histogram for each feature in the dataset

```python
In [10]: import seaborn as sns
         import matplotlib
         import matplotlib.pyplot as plt
         %matplotlib inline
```

```
In [11]: fig, axes = plt.subplots(2, 2, figsize=(16, 8))


         axes[0,0].set_title("Distribution of sepal length in cm")
         axes[0,0].hist(df["sepal length in cm"]);

         axes[0,1].set_title("Distribution of sepal width in cm")
         axes[0,1].hist(df["sepal width in cm"]);

         axes[1,0].set_title("Distribution of petal length in cm ")
         axes[1,0].hist(df["petal length in cm"]);

         axes[1,1].set_title("Distribution of petal width in cm")
         axes[1,1].hist(df["petal width in cm"]);
```
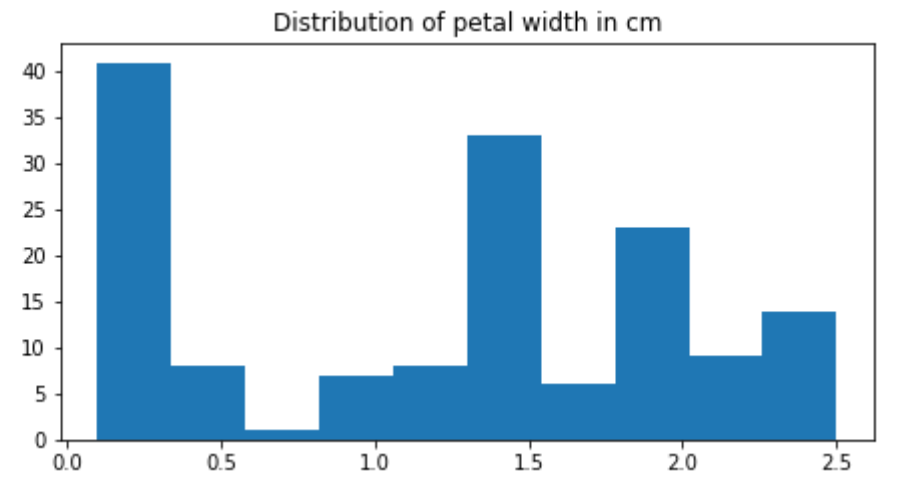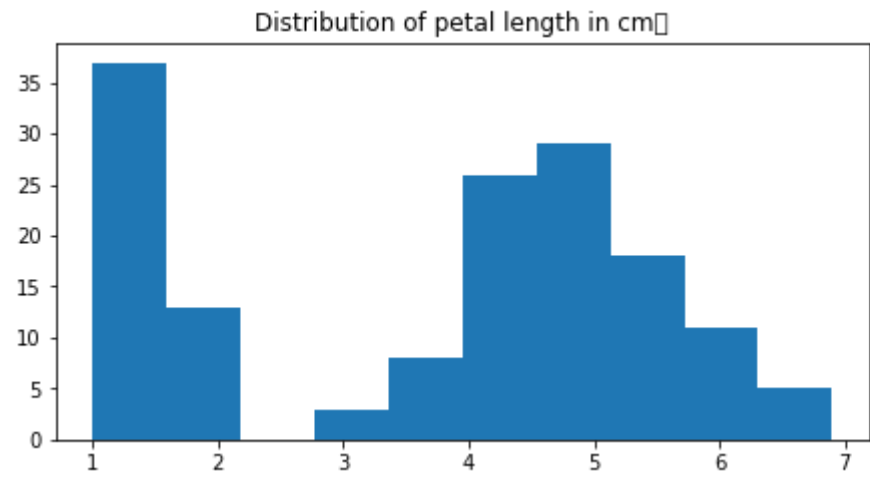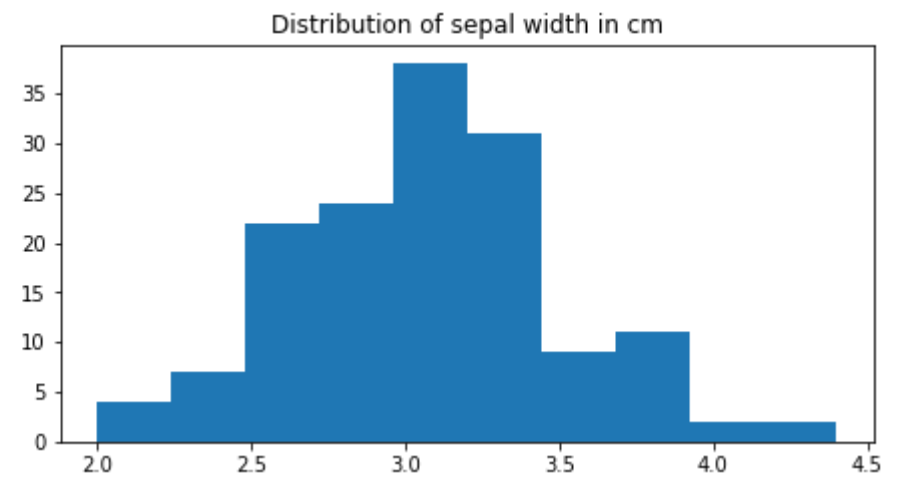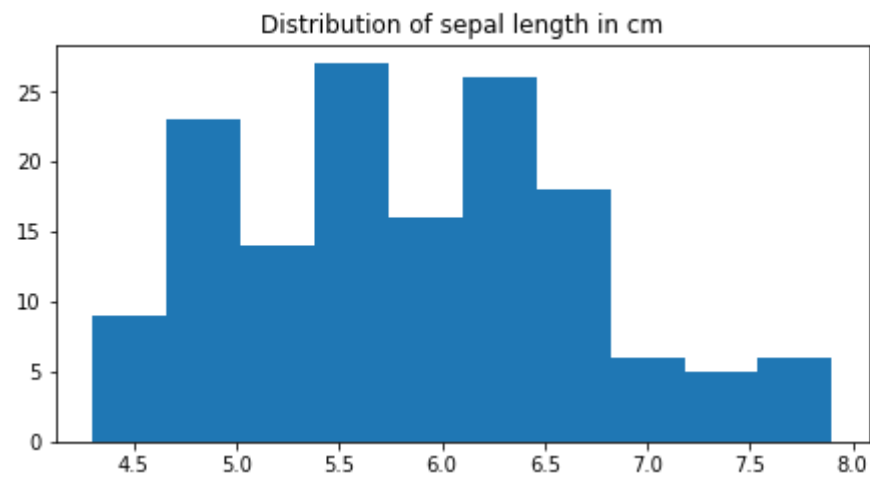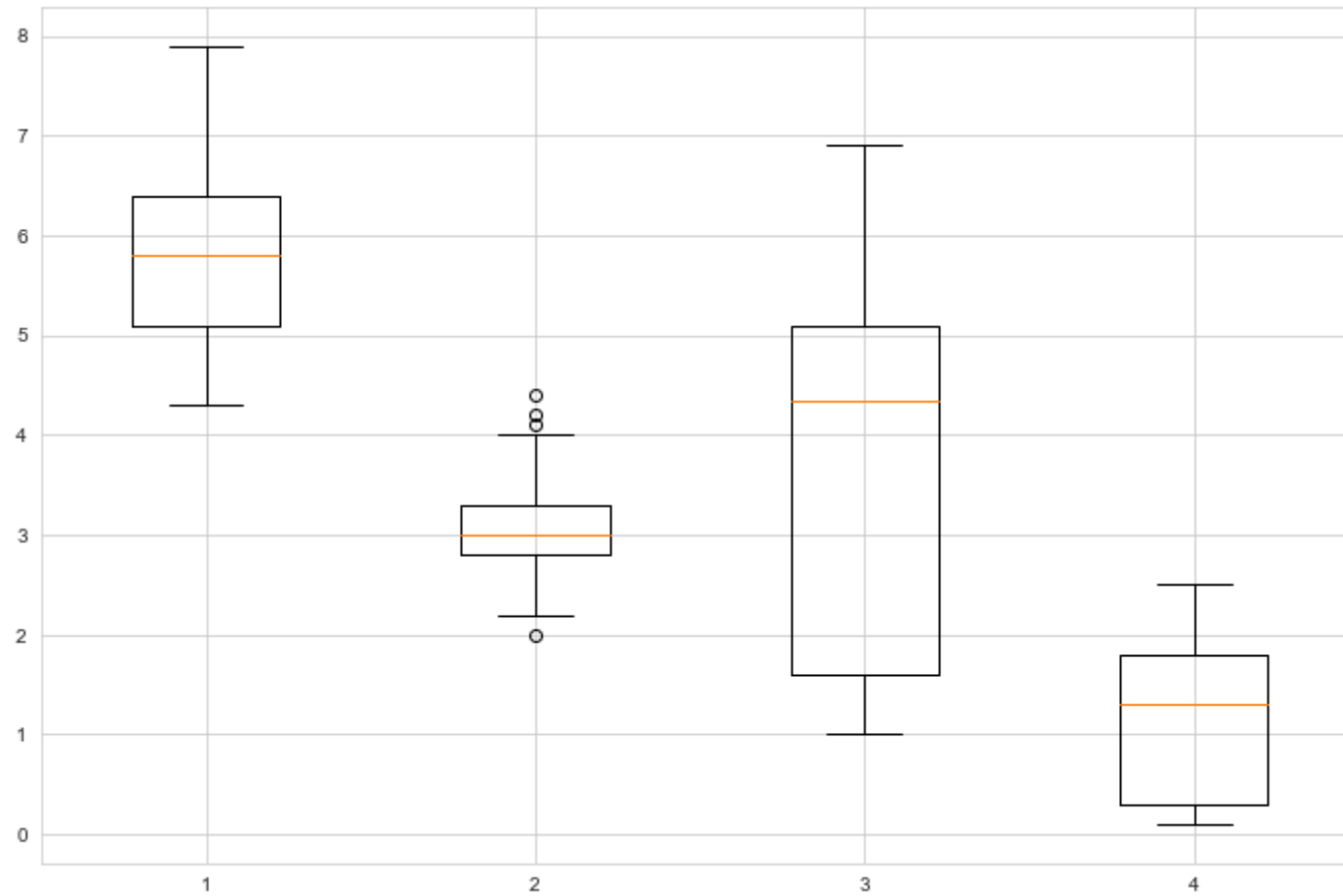
C:\Users\Joshua Deshmukh\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.9_qbz5n2kfra8p0\LocalCache\local-packages\Python39\site-packages\IPython\core\pylabtools.py:151: UserWarning: Glyph 9 (   ) missing from current font.
  fig.canvas.print_figure(bytes_io, **kw)

Boxplot for each feature in each dataset

In [12]:
```python
data_to_plot = [df["sepal length in cm"],df["sepal width in cm"],df["petal length in cm"],df["petal width in cm"]]

sns.set_style("whitegrid")

fig = plt.figure(1, figsize=(12,8))

ax = fig.add_subplot(111)

bp = ax.boxplot(data_to_plot);
```

In [ ]: