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

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

Out[2]:

	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
...	...	...	...	...	...	...
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [3]:
```
column = len(list(df))
column
```

Out[3]: 6

In [4]:
```
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
```

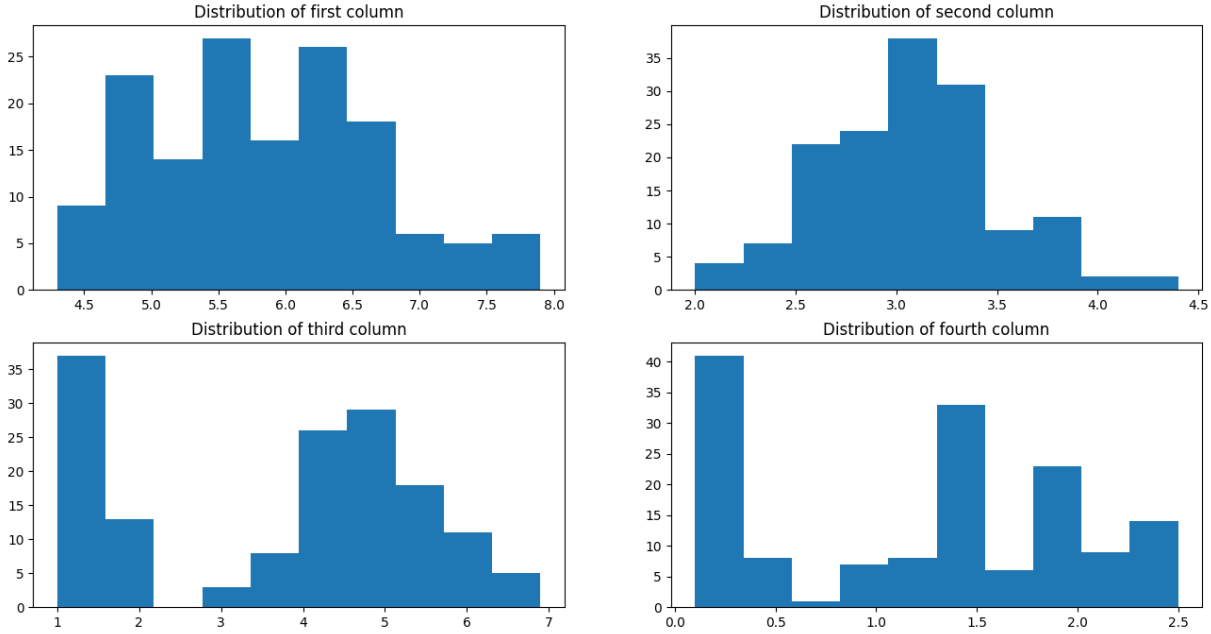
In [5]:
```
np.unique(df['Species'])
```

Out[5]:
```
array(['Iris-setosa', 'Iris-versicolor', 'Iris-virginica'], dtype=object)
```

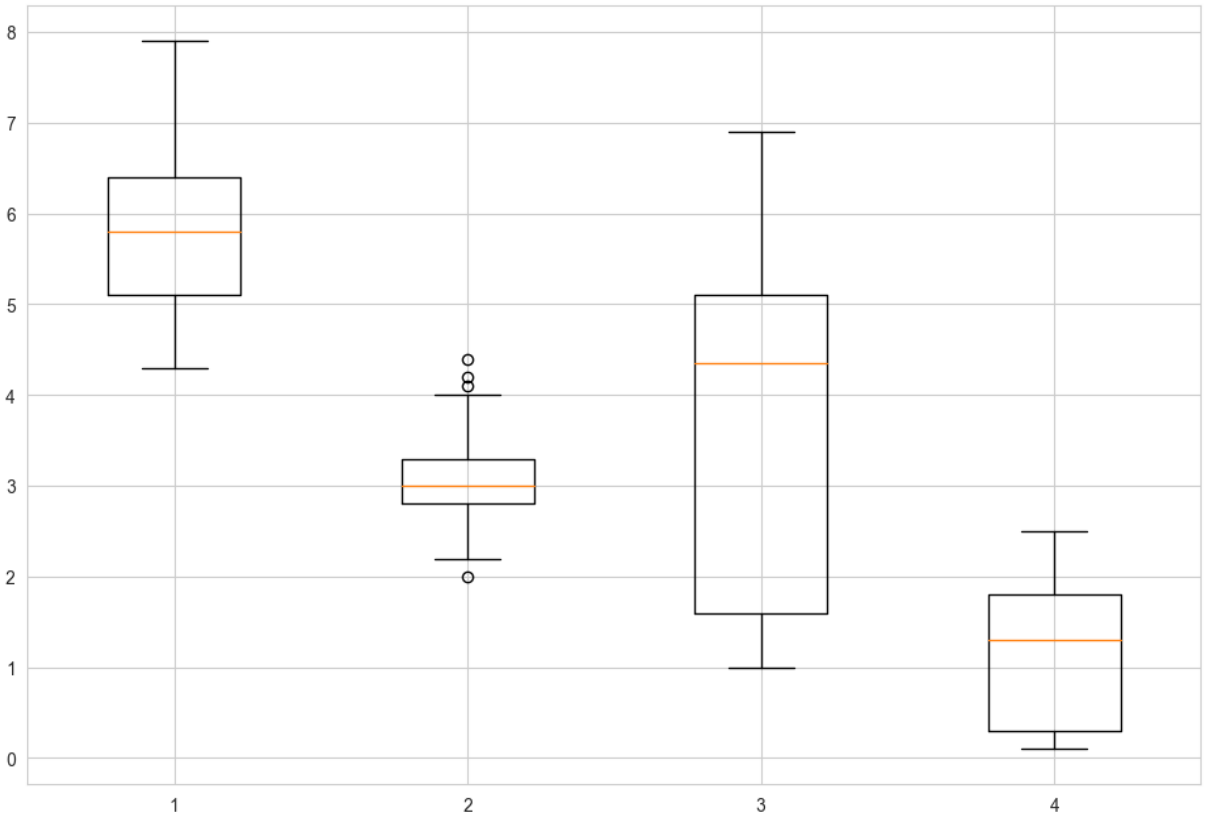
In [6]:
```
import seaborn as sns
import matplotlib
import matplotlib.pyplot as plt
%matplotlib inline
```

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

axes[0,0].set_title("Distribution of first column")
axes[0,0].hist(df['SepalLengthCm'])

axes[0,1].set_title("Distribution of second column")
axes[0,1].hist(df['SepalWidthCm'])

axes[1,0].set_title("Distribution of third column")
axes[1,0].hist(df['PetalLengthCm'])

axes[1,1].set_title("Distribution of fourth column")
axes[1,1].hist(df['PetalWidthCm'])
```

Out[8]:
```
(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>)
```



In [10]:
```
data_to_plot = [df['SepalLengthCm'],df['SepalWidthCm'],df['PetalLengthCm'],df['Peta
sns.set_style('whitegrid')

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

ax=fig.add_subplot(111)

bp=ax.boxplot(data_to_plot)
```



In [11]:
```
df.describe()
```

Out[11]:

	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