

```
In [ ]: #omkar shinde
        #Data Visualization 3
```

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

```
In [2]: df = pd.read_csv("C:/Users/avcoe/Downloads/Iris_prac6.csv")
```

```
In [3]: df.columns = ["col1", "col2", "col3", "col4", "col5"]
```

```
In [4]: df.head()
```

Out[4]:

	col1	col2	col3	col4	col5
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

```
In [ ]: # the features and their types (e.g., numeric, nominal) available in the data
```

```
In [5]: column = len(list(df))
column
```

Out[5]: 5

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
#   Column  Non-Null Count  Dtype
---  -
0   col1    150 non-null     float64
1   col2    150 non-null     float64
2   col3    150 non-null     float64
3   col4    150 non-null     float64
4   col5    150 non-null     object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

```
In [7]: np.unique(df['col5'])
```

Out[7]: array(['setosa', 'versicolor', 'virginica'], dtype=object)

```
In [ ]: #Histogram
```

```
In [8]: import seaborn as sns
import matplotlib
import matplotlib.pyplot as plt
%matplotlib inline
```

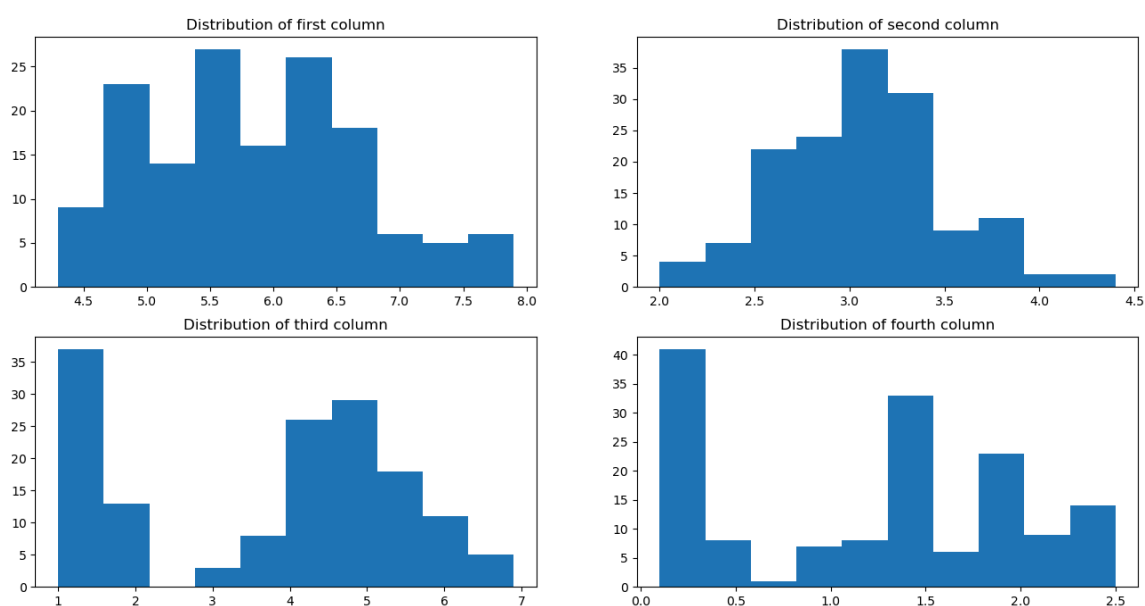
```
In [9]: fig, axes = plt.subplots(2,2,figsize=(16,8))

axes[0,0].set_title("Distribution of first column")
axes[0,0].hist(df['col1']);

axes[0,1].set_title("Distribution of second column")
axes[0,1].hist(df['col2']);

axes[1,0].set_title("Distribution of third column")
axes[1,0].hist(df['col3']);

axes[1,1].set_title("Distribution of fourth column")
axes[1,1].hist(df['col4']);
```



```
In [10]: # Create a box plot for each feature

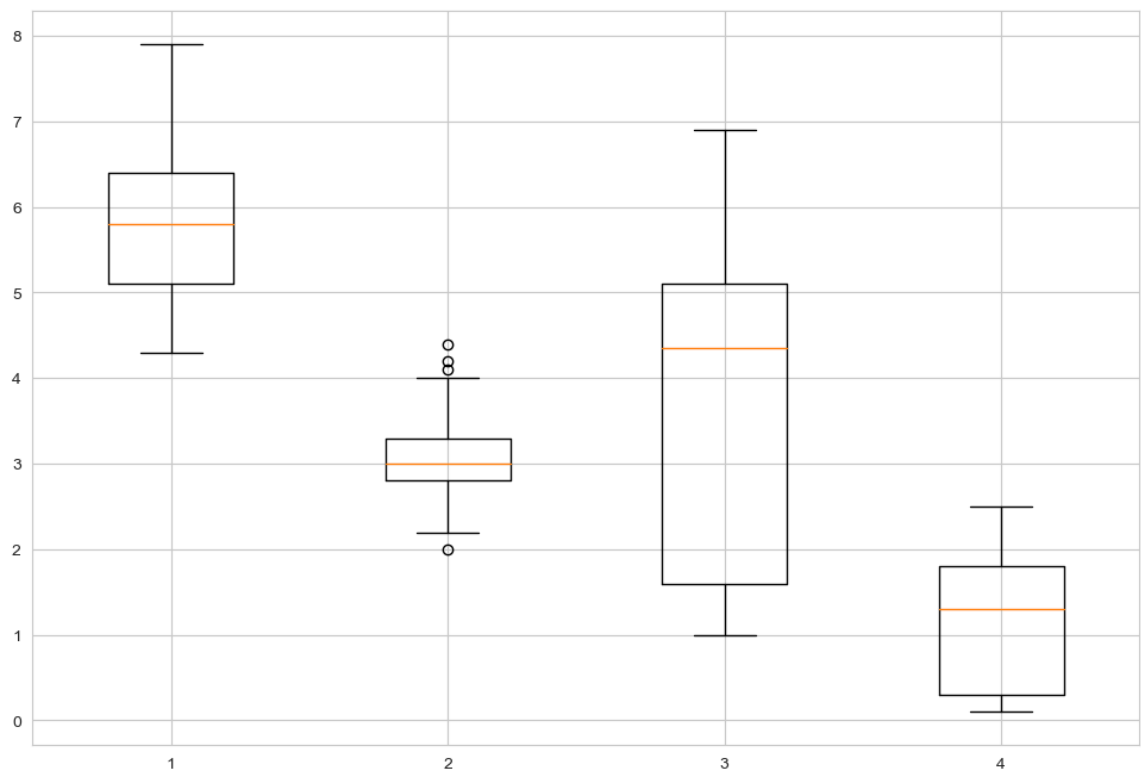
data_to_plot = [df["col1"],df["col2"],df["col3"],df["col4"]]

sns.set_style("whitegrid")

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

# creat an axes instance
ax = fig.add_subplot(111)

# create boxplot
bp = ax.boxplot(data_to_plot)
```



```
In [ ]:
```