In [79]:
```
import pandas as pd
import numpy as np

url="E:\Iris.csv"
df = pd.read_csv(url)
df.head()
```

Out[79]:

	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

In [77]:
```
#List down the features and their types
#(e.g., numeric, nominal) available in the dataset.
```

In [80]:
```
column = len(list(df))
column
#Clearly, dataset has 6 column indicating 6 features about the data
```

Out[80]:
```
6
```

In [81]:
```
df.info()
#Hence the dataset contains 5 numerical columns and 1 object column
```

```
<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
```

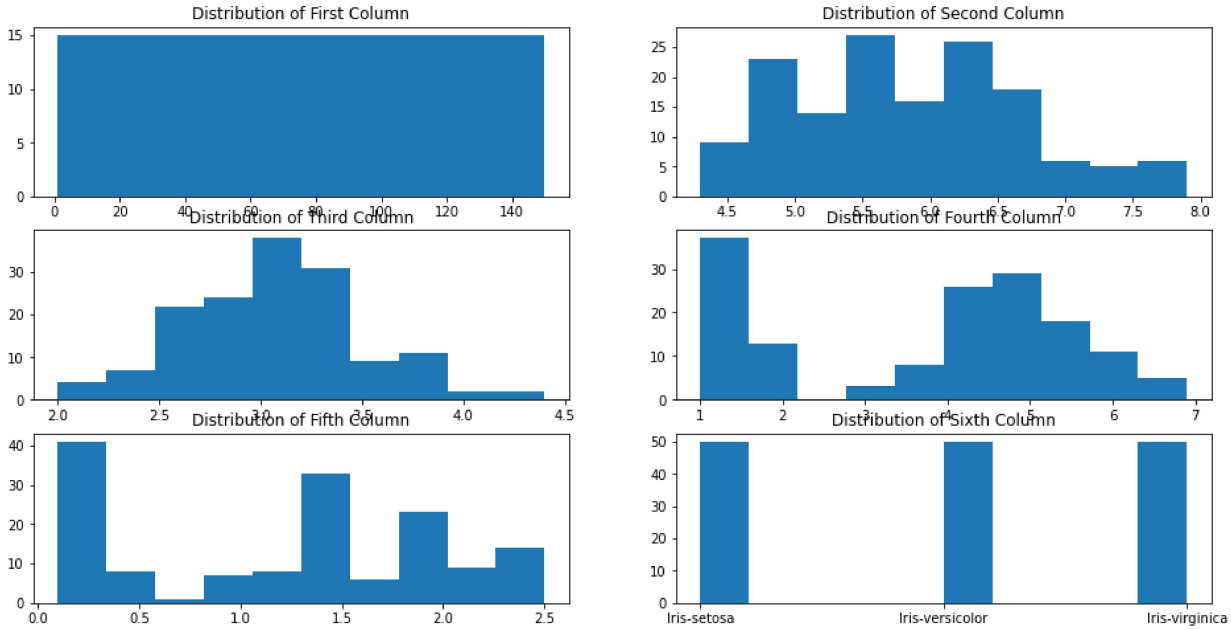
In [ ]:
```
#Data Visualization-Create a histogram for each feature in the dataset to
#illustrate the feature distributions. Plot each histogram.
```

In [86]:
```
import seaborn as sns
import matplotlib
import matplotlib.pyplot as plt
%matplotlib inline
df.head(2)
```

Out[86]:

	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

In [90]:
```
fig, axes = plt.subplots(3, 2, figsize=(16, 8))


axes[0,0].set_title("Distribution of First Column")
axes[0,0].hist(df["Id"]);

axes[0,1].set_title("Distribution of Second Column")
axes[0,1].hist(df["SepalLengthCm"]);

axes[1,0].set_title("Distribution of Third Column")
axes[1,0].hist(df["SepalWidthCm"]);

axes[1,1].set_title("Distribution of Fourth Column")
axes[1,1].hist(df["PetalLengthCm"]);

axes[2,0].set_title("Distribution of Fifth Column")
axes[2,0].hist(df["PetalWidthCm"]);

axes[2,1].set_title("Distribution of Sixth Column")
axes[2,1].hist(df["Species"]);
```



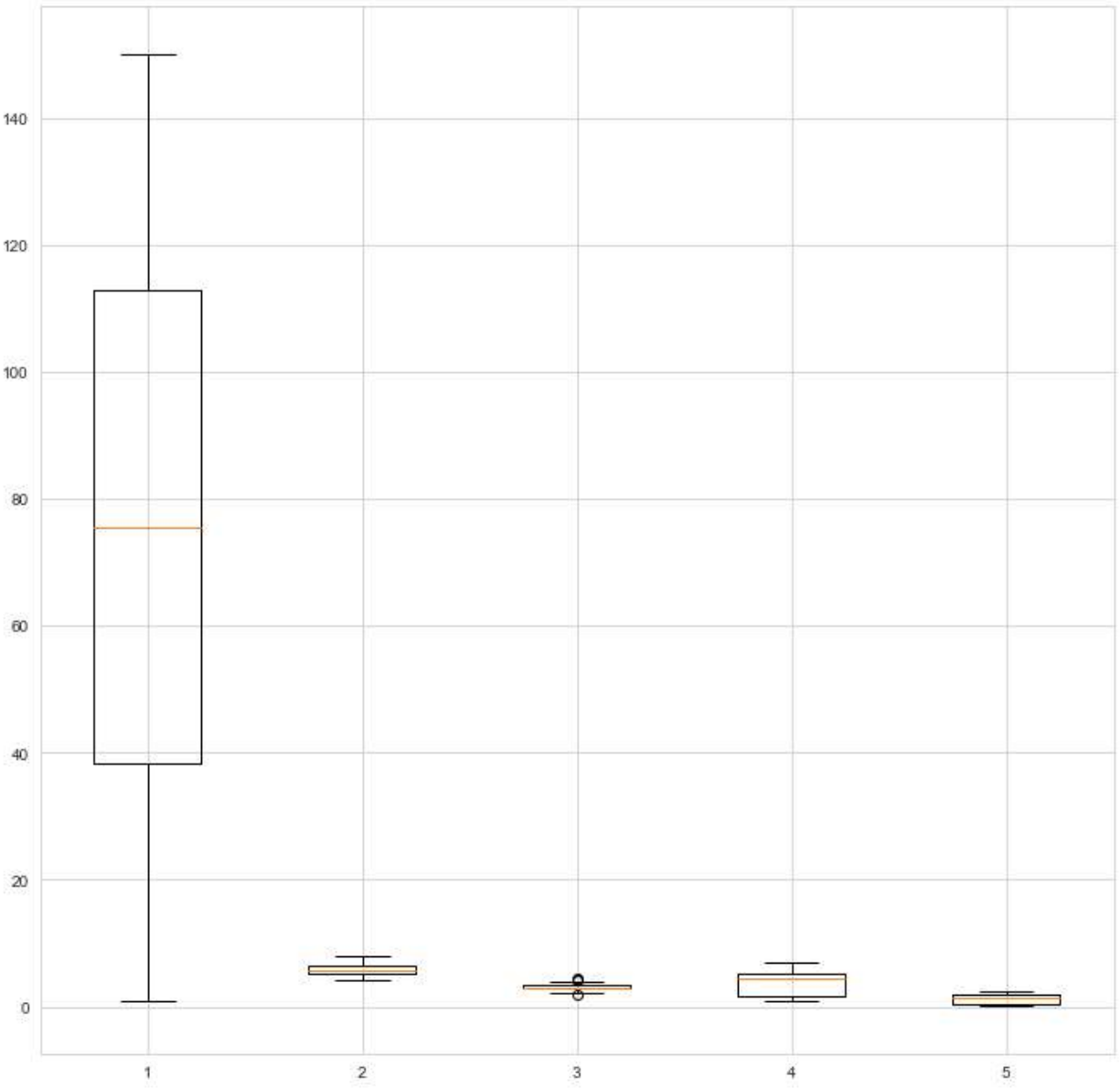
In [91]:
```
#Create a boxplot for each feature in the dataset.
#All of the boxplots should be combined into a single plot.
#Compare distributions and identify outliers.
df.head(2)
```

Out[91]:

	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

In [104…

```
data_to_plot = [df["Id"],df["SepalLengthCm"],df["SepalWidthCm"]
                ,df["PetalLengthCm"],df["PetalWidthCm"]]

sns.set_style("whitegrid")

# Creating a figure instance
fig = plt.figure(1, figsize=(12,12))

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

# Creating the boxplot
bp = ax.boxplot(data_to_plot);
```



In [ ]: