**Download the Iris flower dataset or any other dataset into a DataFrame. (e.g.,** [**https://archive.ics.uci.edu/ml/datasets/Iris**](https://archive.ics.uci.edu/ml/datasets/Iris) **). Scan the dataset and give the inference as:**

1. **List down the features and their types (e.g., numeric, nominal) available in the dataset.**
2. **Create a histogram for each feature in the dataset to illustrate the feature distributions.**
3. **Create a boxplot for each feature in the dataset.**
4. **Compare distributions and identify outliers**

**-----------------------------------------------------------------------------------------------------------------------------------------**



**import** pandas **as** pd

**import** numpy **as** np

**import** matplotlib.pyplot **as** plt

**import** seaborn **as** sns



df **=** pd**.**read\_csv("iris dataset.csv")



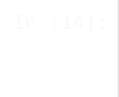
df**.**info()

<class 'pandas.core.frame.DataFrame'> RangeIndex: 150 entries, 0 to 149 Data columns (total 5 columns):

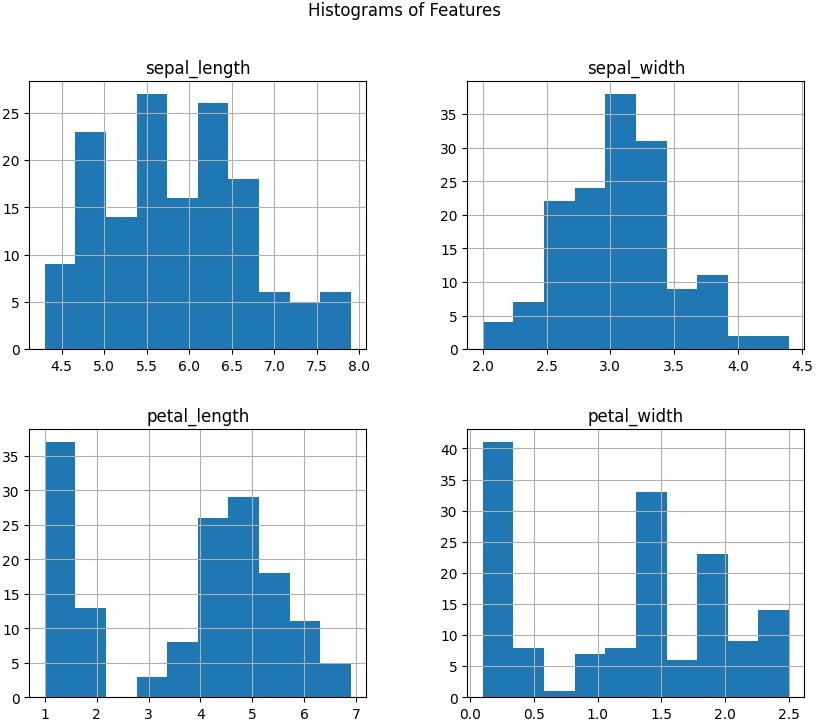
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| # |  | Column | Non-Null Count |  | Dtype |
| 0 |  | sepal\_length | 150 non-null |  | float64 |
| 1 |  | sepal\_width | 150 non-null |  | float64 |
| 2 |  | petal\_length | 150 non-null |  | float64 |
| 3 |  | petal\_width | 150 non-null |  | float64 |
| 4 |  | species | 150 non-null |  | object |

dtypes: float64(4), object(1) memory usage: 6.0+ KB

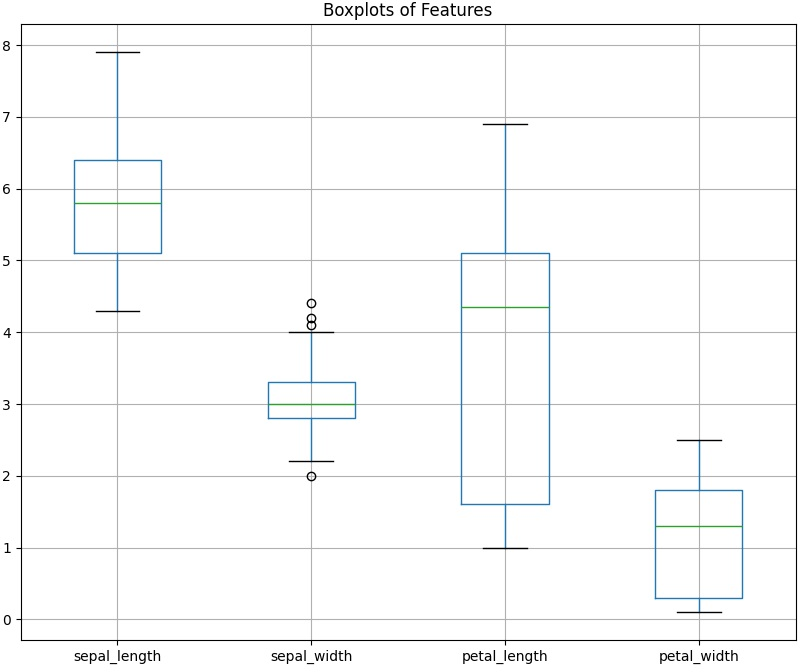
# **Create a histogram for each feature in the dataset to illustrate the feature distributions**

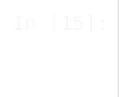


df**.**hist(figsize**=**(10, 8)) plt**.**suptitle('Histograms of Features') plt**.**show()



1. **Create a boxplot for each feature in the dataset.**





df**.**boxplot(figsize**=**(10, 8)) plt**.**title('Boxplots of Features') plt**.**show()

**If we observe closely for the box 2, interquartile distance is roughly around 0.75 hence the values lying beyond this range of (third quartile**

**+ interquartile distance) i.e. roughly around 4.05 will be considered as outliers. Similarly outliers with other boxplots can be found.**