

Practical 0

Title : Data Visualization III

```
[ ]: import pandas as pd  
import seaborn as sns  
import matplotlib.pyplot as plt
```

```
[ ]: # Download the Iris dataset and load it into a DataFrame  
url = 'https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data'  
column_names = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width',  
    ↴'class']  
iris_df = pd.read_csv(url, names=column_names)
```

1. List down the features and their types (e.g., numeric, nominal) available in the dataset.

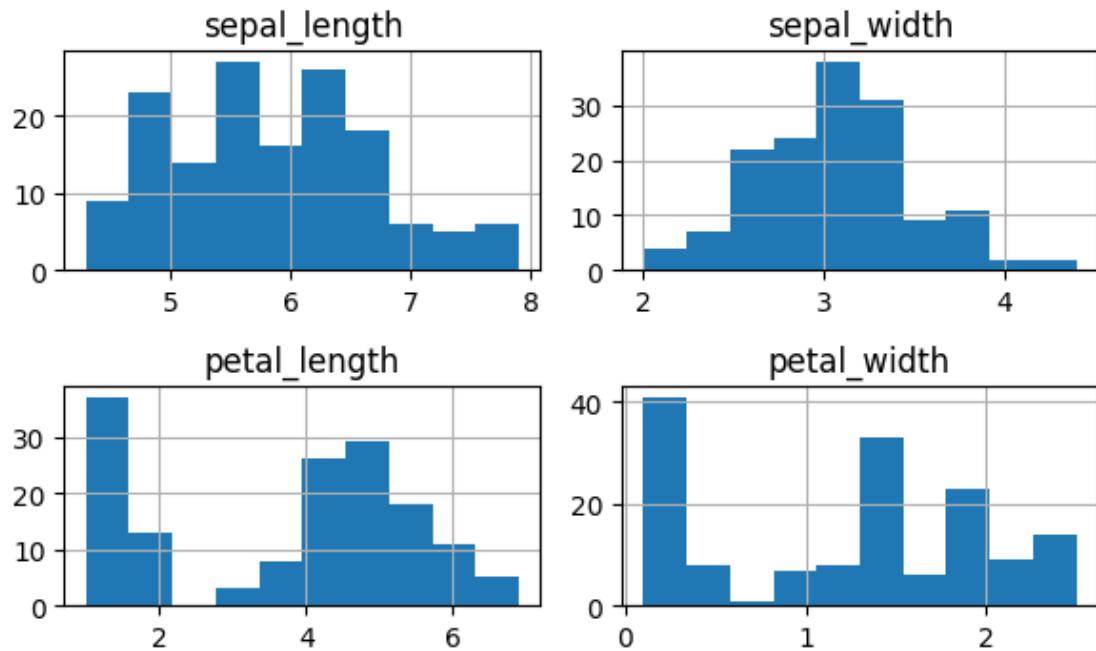
```
[ ]: feature_types = iris_df.dtypes  
print("Features and their types:")  
print(feature_types)
```

```
Features and their types:  
sepal_length    float64  
sepal_width     float64  
petal_length    float64  
petal_width     float64  
class           object  
dtype: object
```

2. Create a histogram for each feature

```
[ ]: iris_df.hist(figsize=(6, 4))  
plt.suptitle("Histograms of Iris Dataset Features")  
plt.tight_layout()  
plt.show()
```

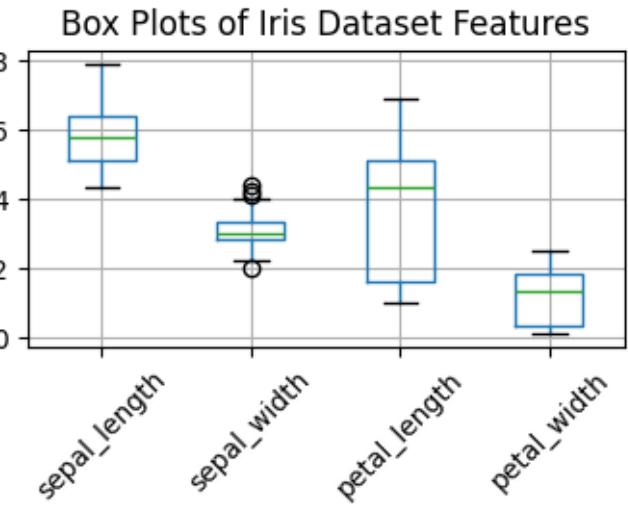
Histograms of Iris Dataset Features



3. Create a box plot for each feature

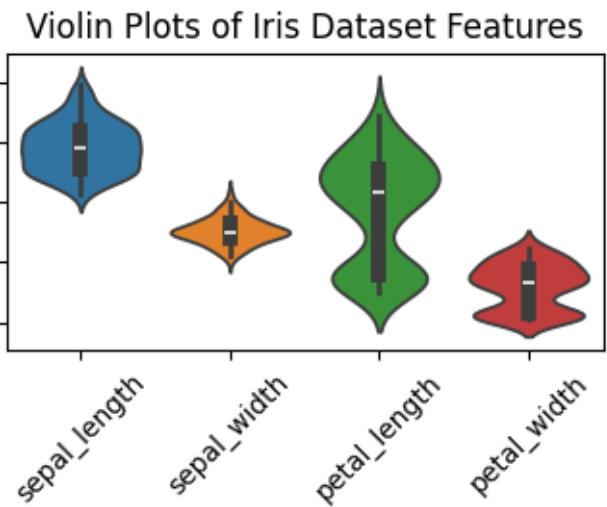
```
[ ]: plt.figure(figsize=(6, 4))

iris_df.boxplot()
plt.title("Box Plots of Iris Dataset Features")
plt.xticks(rotation=45)
plt.show()
```

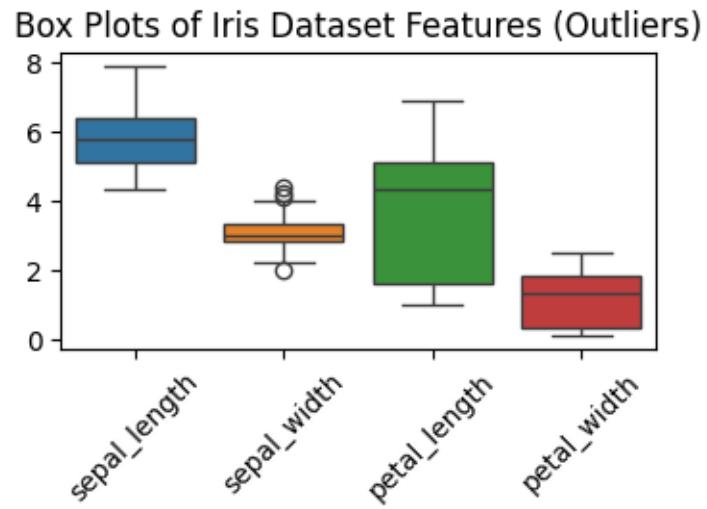


4. Compare distributions and identify outliers

```
[ ]: # Compare distributions using violin plots
plt.figure(figsize=(6, 2))
sns.violinplot(data=iris_df)
plt.title("Violin Plots of Iris Dataset Features")
plt.xticks(rotation=45)
plt.show()
```



```
[ ]: # Identify outliers using box plots
plt.figure(figsize=(4, 2))
sns.boxplot(data=iris_df)
plt.title("Box Plots of Iris Dataset Features (Outliers)")
plt.xticks(rotation=45)
plt.show()
```



```
[ ]:
```