

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans
from sklearn.cluster import AgglomerativeClustering
from sklearn.metrics import silhouette_score
from scipy.cluster.hierarchy import dendrogram, linkage
from sklearn.preprocessing import LabelEncoder

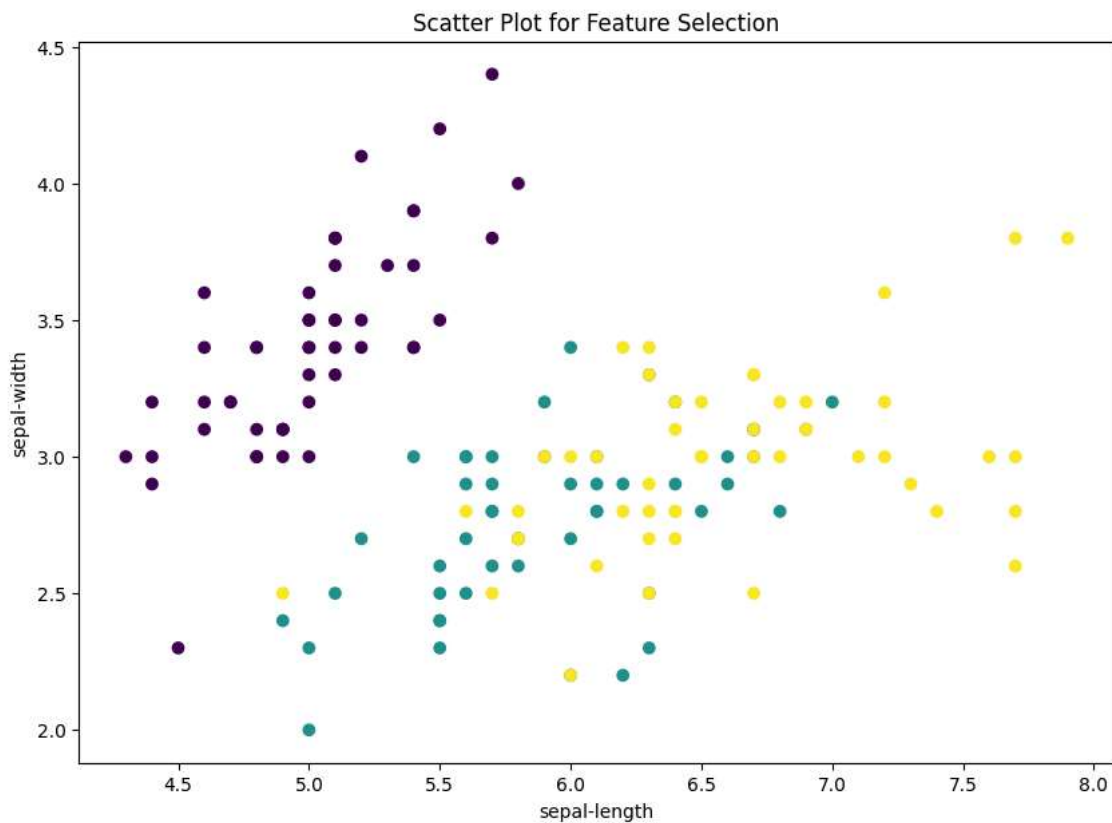
# Load Iris dataset
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"
names = ['sepal-length', 'sepal-width', 'petal-length', 'petal-width', 'class']
dataset = pd.read_csv(url, names=names)

# Encode class labels for plotting
le = LabelEncoder()
y_encoded = le.fit_transform(dataset['class'])

# Remove class labels for clustering
X = dataset.iloc[:, :-1].values

# Scatter plot for feature selection
plt.figure(figsize=(10, 7))
plt.scatter(X[:, 0], X[:, 1], c=y_encoded, cmap='viridis')
plt.xlabel(names[0])
plt.ylabel(names[1])
plt.title('Scatter Plot for Feature Selection')
plt.show()

```

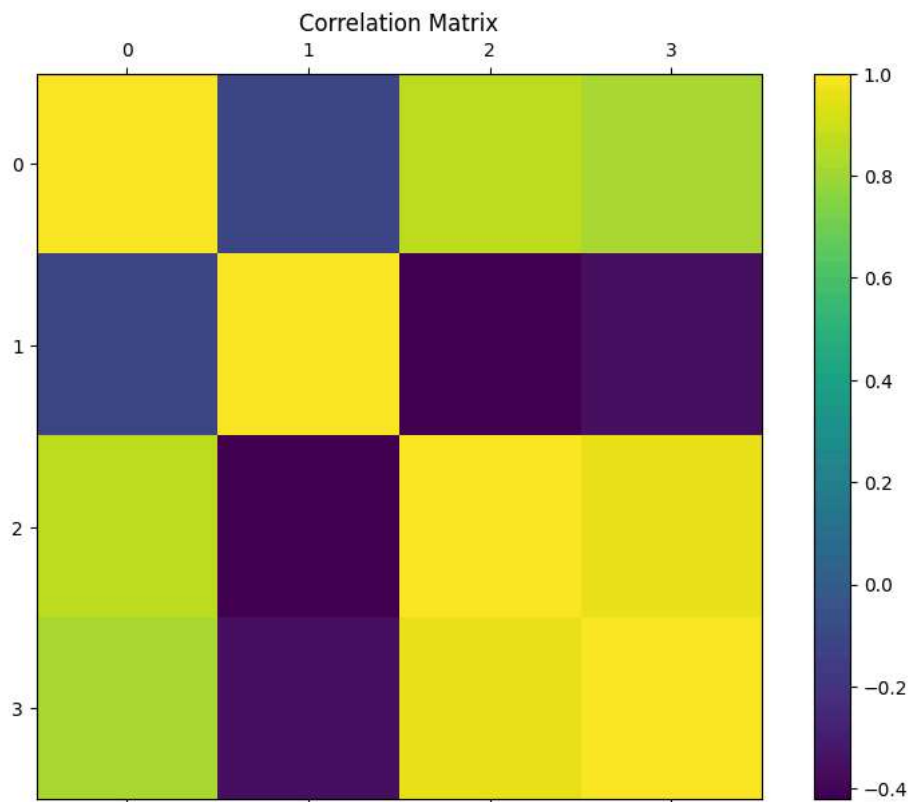


```

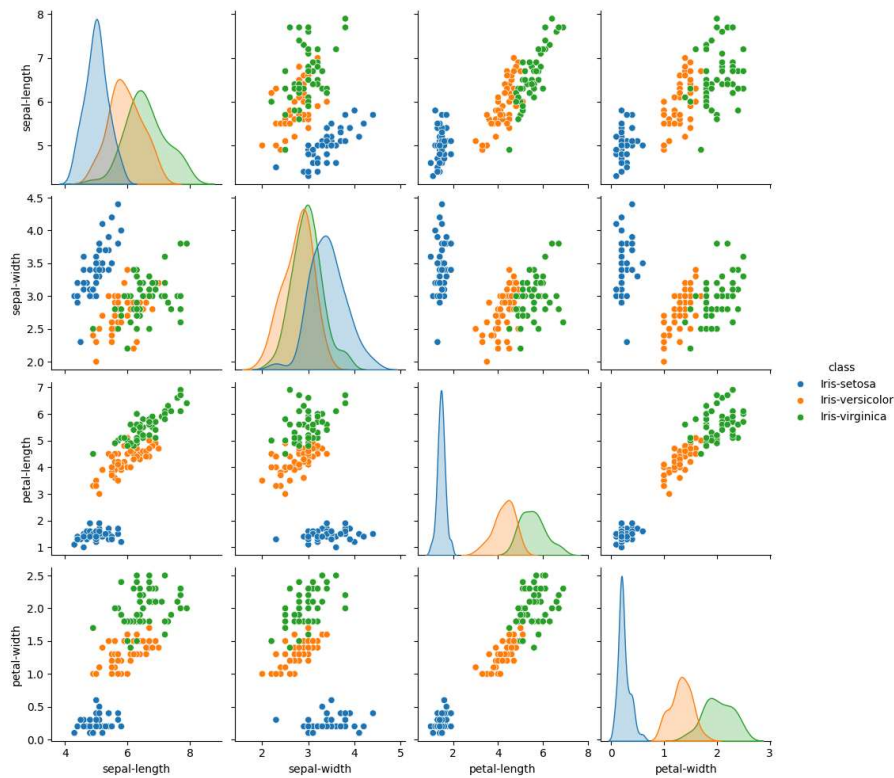
# Correlation matrix
plt.figure(figsize=(10, 7))
plt.matshow(dataset.corr(), fignum=1)
plt.colorbar()
plt.title('Correlation Matrix')
plt.show()

```

```
<ipython-input-7-1c62a1e8c08e>:3: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, plt.matshow(dataset.corr(), fignum=1)
```



```
# Pair plot among all attributes
sns.pairplot(dataset, hue='class')
plt.show()
```



Based on the scatter plot and correlation matrix, we can visually inspect the data and select features.

```
selected_features = [0, 1]
```

K-means Clustering

```
kmeans = KMeans(n_clusters=3, random_state=0).fit(X[:, selected_features])
```

```
labels_pred = kmeans.labels_
```

```
print("Silhouette Score for K-means: %f" % silhouette_score(X[:, selected_features], labels_pred))
```

Silhouette Score for K-means: 0.443469

/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: FutureWarning: The default value of `n_init` will change from 10 to 100 in version 1.2. For now, use `n_init=10` to silence this warning.

Hierarchical Clustering

```
linkage_methods = ['single', 'complete', 'average', 'ward']
```

```
for method in linkage_methods:
```

```
    Z = linkage(X[:, selected_features], method=method)
```

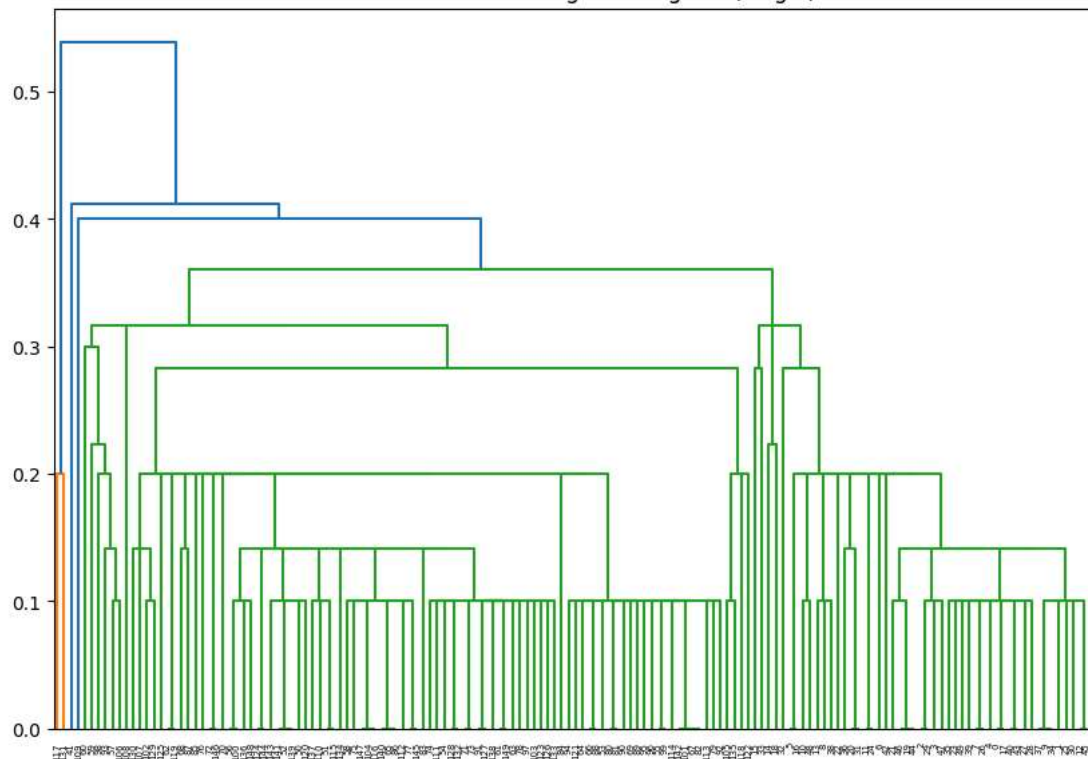
```
    plt.figure(figsize=(10, 7))
```

```
    dendrogram(Z, leaf_rotation=90)
```

```
    plt.title(f'Hierarchical Clustering Dendrogram ({method})')
```

```
    plt.show()
```

Hierarchical Clustering Dendrogram (single)



Hierarchical Clustering Dendrogram (complete)

