

Explanation

```

from sklearn.cluster import KMeans
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
Loading... otlib import pyplot as plt
from sklearn.datasets import load_iris
%matplotlib inline
✓ 2.0s

```

```

Downloading kiwisolver-1.4.5-cp312-cp312-win_amd64.whl (56 kB)
56.0/56.0 kB 1.4 MB/s eta 0:00:00
Downloading pillow-10.2.0-cp312-cp312-win_amd64.whl (2.6 MB)
2.6/2.6 MB 811.3 kB/s eta 0:00:00
Downloading pyparsing-3.1.2-py3-none-any.whl (103 kB)
103.2/103.2 kB 2.0 MB/s eta 0:00:00
Installing collected packages: pyparsing, pillow, kiwisolver, fonttools, cycler, contourpy, matplotlib
Successfully installed contourpy-1.2.0 cycler-0.12.1 fonttools-4.49.0 kiwisolver-1.4.5 matplotlib-3.8.3 pillow-10.2.0 pyparsing-3.1.2
PS C:\Users\acer>

```

So first I checked these first lines and I encounter an error where I don't have 'matplotlib' so, I open the terminal and used the 'pip install matplotlib' to install it.

```

iris = load_iris()
✓ 0.0s

```

```

df = pd.DataFrame(iris.data, columns=iris.feature_names)
df.head()
✓ 0.0s

```

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2

Loading 'iris' data frame from the sklearn.datasets.

```
df['flower'] = iris.target
df.head()
```

✓ 0.0s

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	flower
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

Creating another column for the 'flower' label

```
import matplotlib.pyplot as plt
from sklearn.datasets import load_iris
iris = load_iris()

df = pd.DataFrame(iris.data, columns=iris.feature_names)

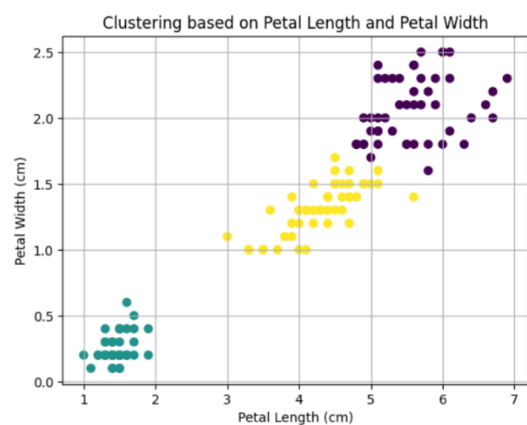
df['flower'] = iris.target

scaler = MinMaxScaler()
scaled_features = scaler.fit_transform(df[['petal length (cm)', 'petal width (cm)']])

kmeans = KMeans(n_clusters=3, random_state=42)
df['cluster'] = kmeans.fit_predict(scaled_features)

plt.scatter(df['petal length (cm)'], df['petal width (cm)'], c=df['cluster'], cmap='viridis')
plt.xlabel('Petal Length (cm)')
plt.ylabel('Petal Width (cm)')
plt.title('Clustering based on Petal Length and Petal Width')
plt.grid(True)
plt.show()
```

✓ 0.1s



Forming clusters of flowers using petal width and length features.

```

from sklearn.cluster import KMeans

X = df.drop('flower', axis=1)

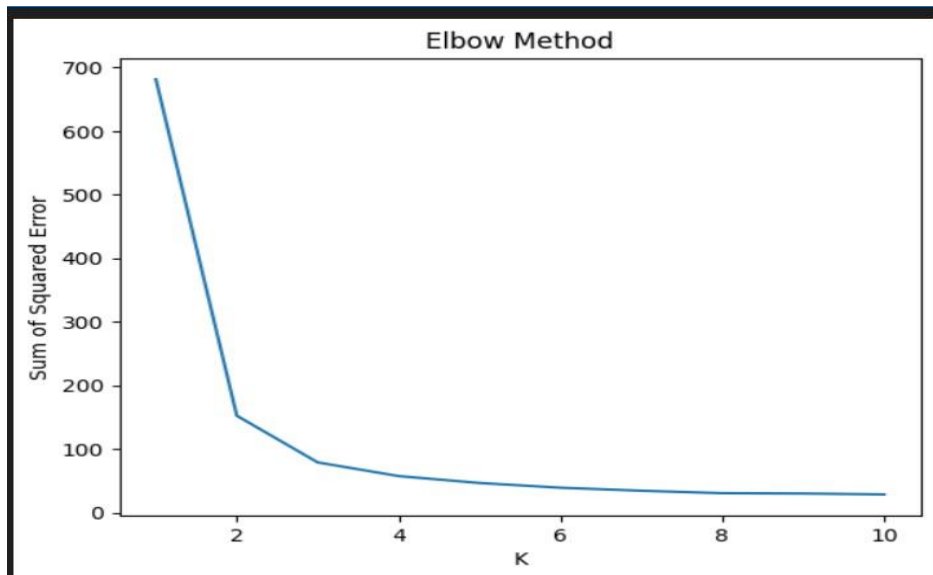
sse = []

for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, init='k-means++', random_state=42)
    kmeans.fit(X)
    sse.append(kmeans.inertia_)

plt.plot(range(1, 11), sse)
plt.title('Elbow Method')
plt.xlabel('K')
plt.ylabel('Sum of Squared Error')
plt.show()

```

✓ 0.1s



Elbow plotting for determining the value of K.