```python
import torch
import numpy as np
from sklearn.cluster import KMeans
from sklearn.metrics import silhouette_score
from sklearn.preprocessing import StandardScaler

tensor_features = torch.load('/kaggle/input/task2dataset-3/stacked_tensor_2.pt')

flattened_data = tensor_features.view(tensor_features.size(0), -1).numpy()

scaler = StandardScaler()
normalized_data = scaler.fit_transform(flattened_data)

wcss = []
silhouette_scores = []

num_clusters_range = range(2, 25)

for num_clusters in num_clusters_range:
    kmeans = KMeans(n_clusters=num_clusters, random_state=42)
    clusters = kmeans.fit_predict(normalized_data)
    wcss.append(kmeans.inertia_)
    silhouette_avg = silhouette_score(normalized_data, clusters)
    silhouette_scores.append(silhouette_avg)
    print(f"For n_clusters = {num_clusters}, the average silhouette_score is : {silhouette_avg}")
```

```
/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 2, the average silhouette_score is :
0.22860950231552124

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 3, the average silhouette_score is :
0.1278744786977768

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(
```

```
For n_clusters = 4, the average silhouette_score is :
0.09982550889253616

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 5, the average silhouette_score is :
0.10801851749420166

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 6, the average silhouette_score is :
0.10363930463790894

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 7, the average silhouette_score is :
0.1017104834318161

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 8, the average silhouette_score is :
0.08396127074956894

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 9, the average silhouette_score is :
0.0863175094127655

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
```

```
to suppress the warning
  warnings.warn(

For n_clusters = 10, the average silhouette_score is :
0.07486061006784439

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 11, the average silhouette_score is :
0.08222465962171555

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 12, the average silhouette_score is :
0.08291265368461609

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 13, the average silhouette_score is :
0.08030898869037628

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 14, the average silhouette_score is :
0.07978122681379318

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 15, the average silhouette_score is :
0.07906251400709152
```

```
/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 16, the average silhouette_score is :
0.08901362121105194

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 17, the average silhouette_score is :
0.0742369145154953

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 18, the average silhouette_score is :
0.07914140820503235

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 19, the average silhouette_score is :
0.08023254573345184

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

For n_clusters = 20, the average silhouette_score is :
0.07500628381967545

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(
```
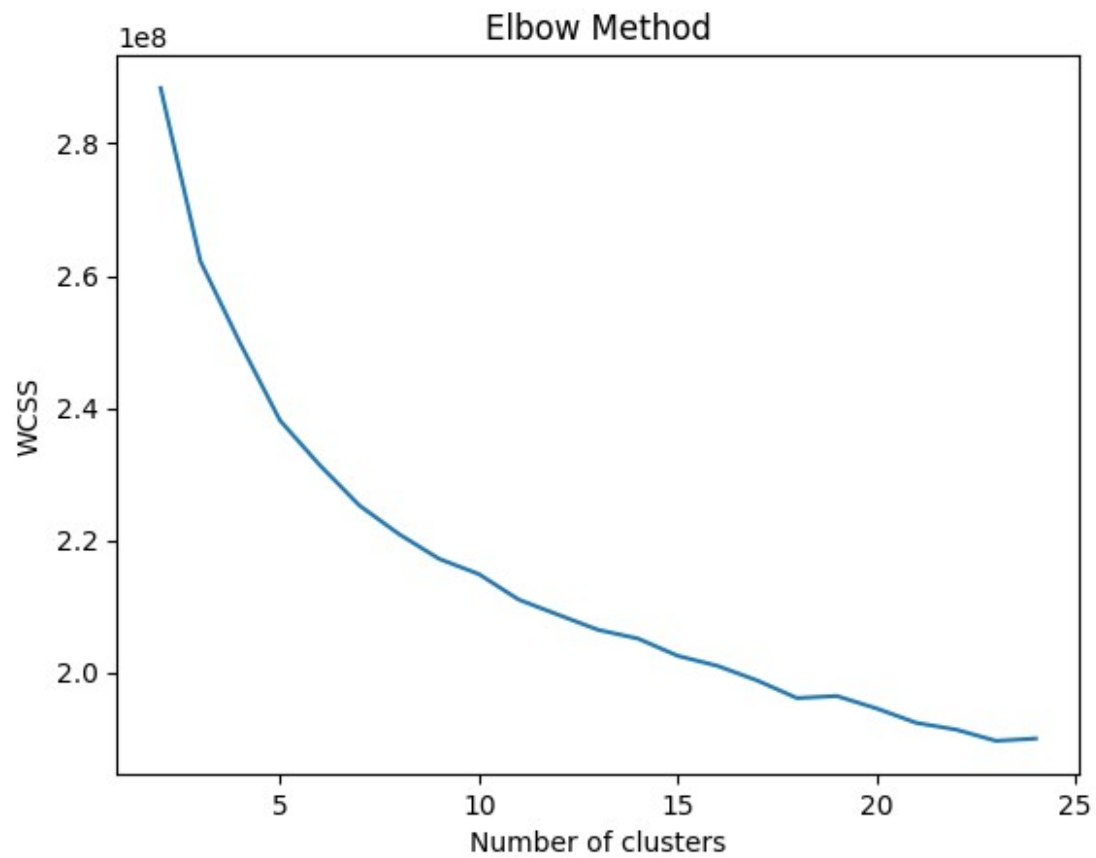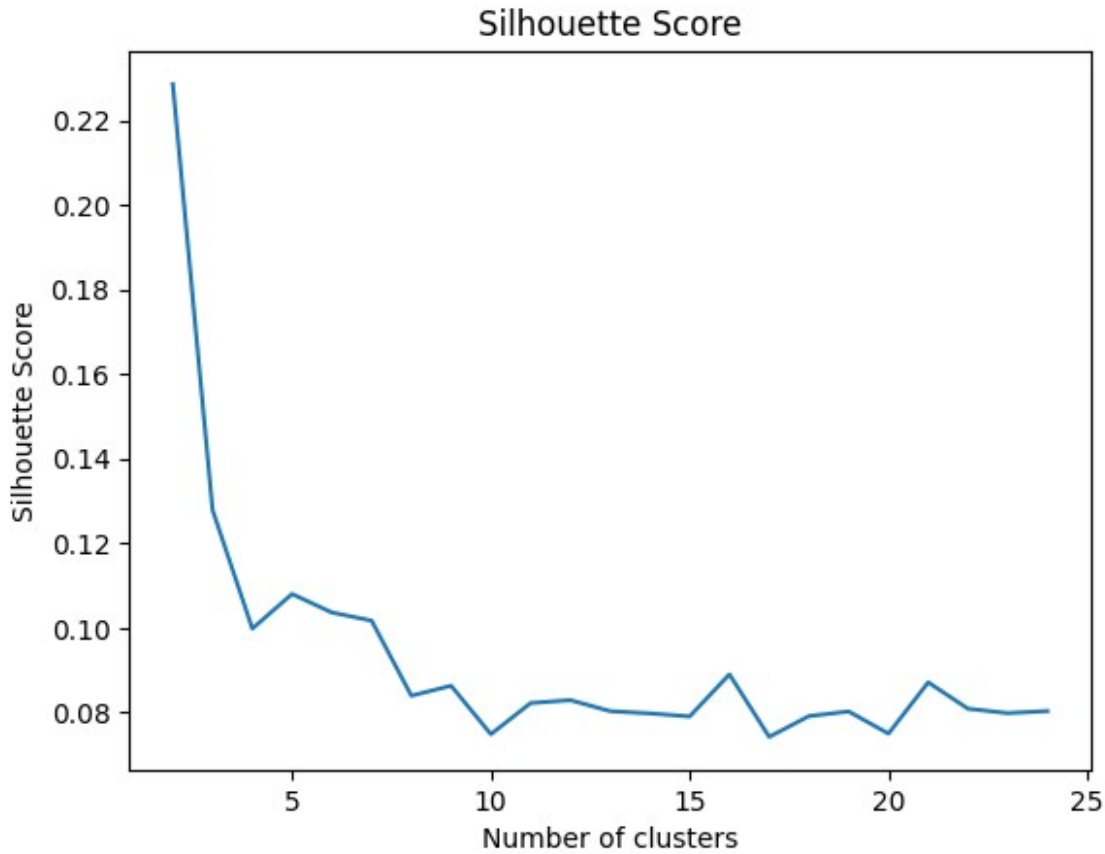
```
For n_clusters = 21, the average silhouette_score is :
0.08712560683488846
```

```
For n_clusters = 22, the average silhouette_score is :
0.08088894188404083
```

```
For n_clusters = 23, the average silhouette_score is :
0.07984220236539841
```

```
For n_clusters = 24, the average silhouette_score is :
0.08033424615859985
```

```python
import torch
import numpy as np
from sklearn.cluster import KMeans
from sklearn.metrics import silhouette_score
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt

plt.plot(num_clusters_range, wcss)
plt.title('Elbow Method')
plt.xlabel('Number of clusters')
plt.ylabel('WCSS')
plt.show()
```

Elbow Method

```
plt.plot(num_clusters_range, silhouette_scores)
plt.title('Silhouette Score')
plt.xlabel('Number of clusters')
plt.ylabel('Silhouette Score')
plt.show()
```

## Silhouette Score



```
num_clusters = 5

kmeans = KMeans(n_clusters=num_clusters, random_state=42)

clusters = kmeans.fit_predict(normalized_data)

/opt/conda/lib/python3.10/site-packages/sklearn/cluster/
_kmeans.py:870: FutureWarning: The default value of `n_init` will
change from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly
to suppress the warning
  warnings.warn(

silhouette_avg = silhouette_score(normalized_data, clusters)
print(f"Silhouette Score: {silhouette_avg}")

Silhouette Score: 0.10801851749420166

from sklearn.decomposition import PCA
import matplotlib.pyplot as plt
pca = PCA(n_components=2)
reduced_data = pca.fit_transform(normalized_data)

plt.figure(figsize=(10, 8))
for cluster_label in range(num_clusters):
```
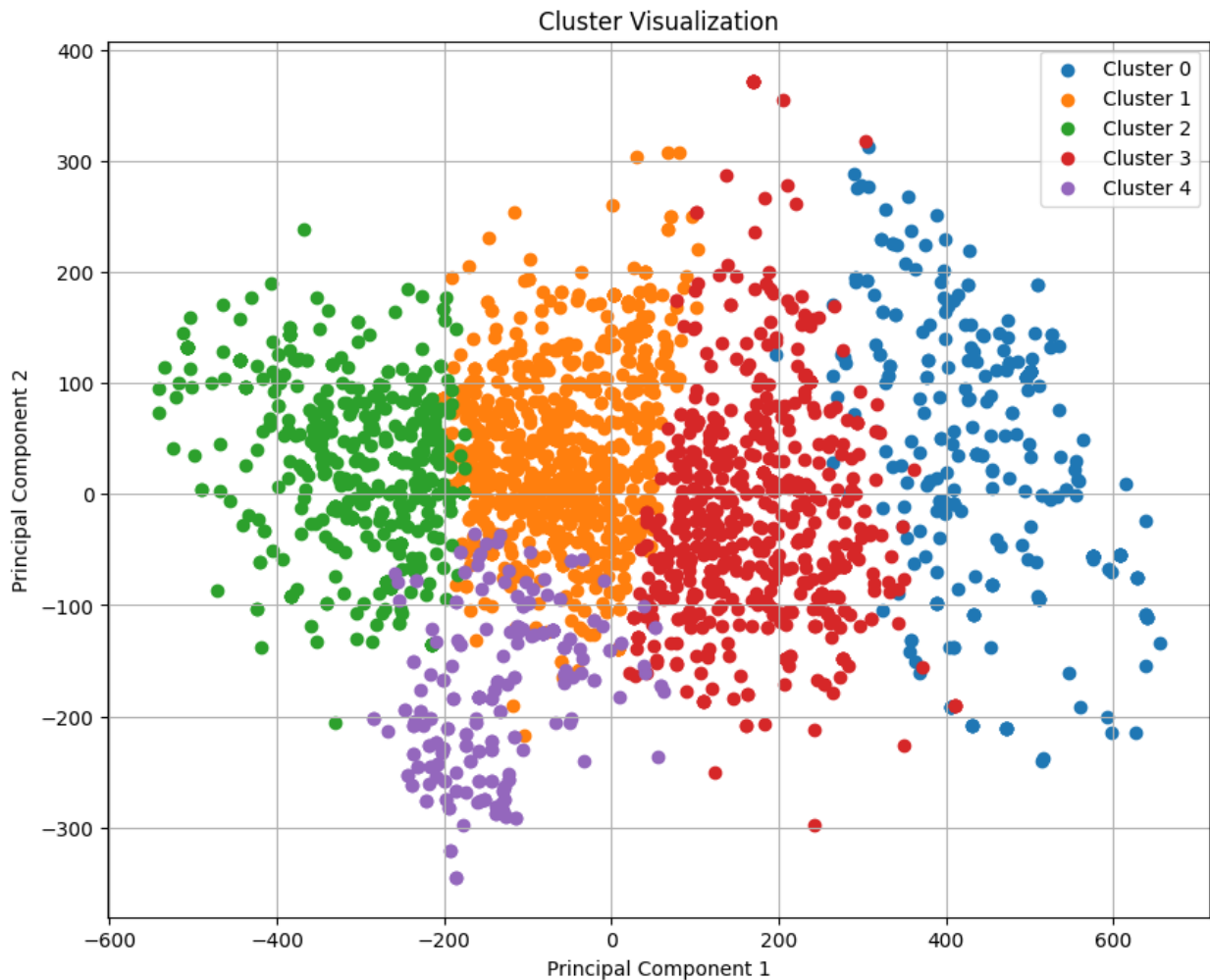
```
    cluster_points = reduced_data[clusters == cluster_label]
    plt.scatter(cluster_points[:, 0], cluster_points[:, 1],
label=f'Cluster {cluster_label}')
plt.title('Cluster Visualization')
plt.xlabel('Principal Component 1')
plt.ylabel('Principal Component 2')
plt.legend()
plt.grid(True)
plt.show()
```



Cluster Visualization

```
import matplotlib.pyplot as plt
from sklearn.manifold import TSNE
from sklearn.metrics import silhouette_samples
import seaborn as sns
import pandas as pd
```