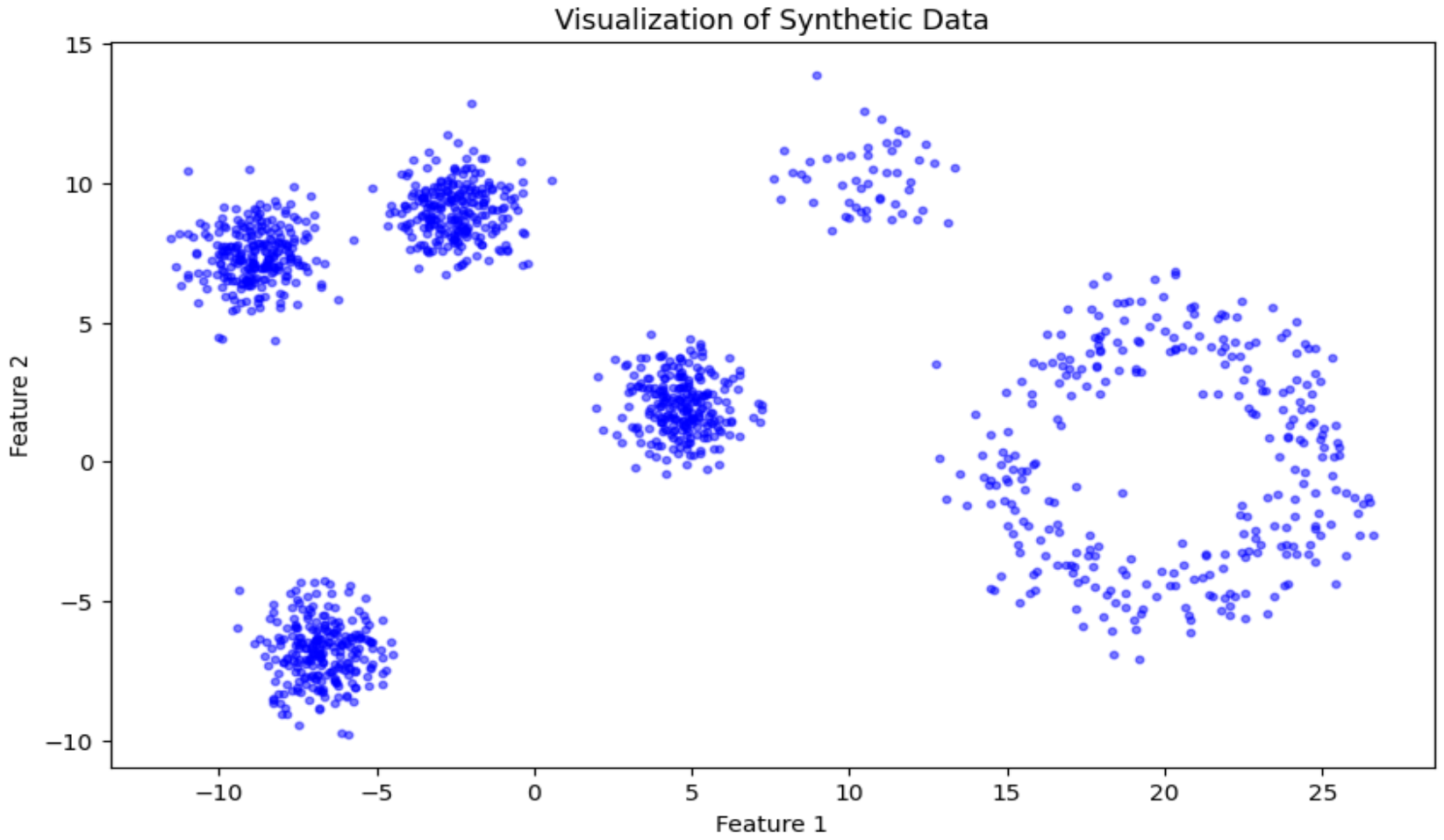
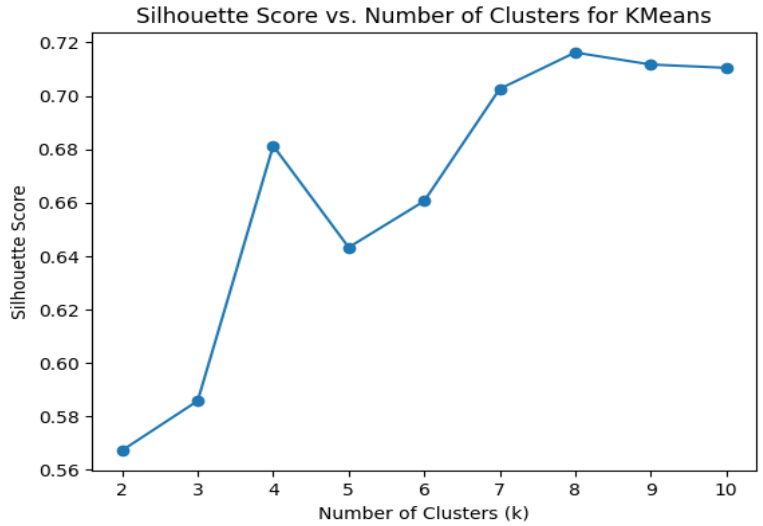
מטלה 3

- שם סטודנט: סייד אחמד אנואר
- ת.ז:206710212
- קורס: למידת מכונה
- שם מרצה: ד"ר איגור נור

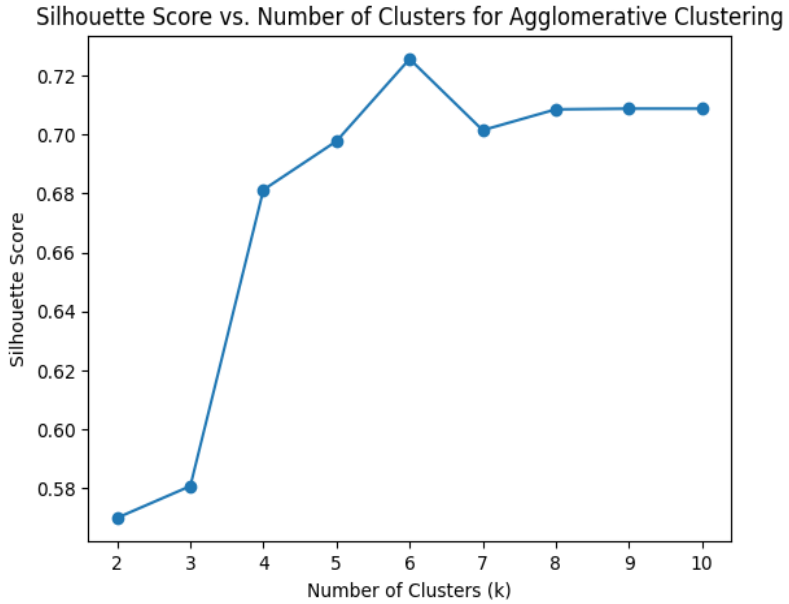
```
X = create_data(1.0, 1000, 50, 300, 200)


plt.figure(figsize=(10, 6))
plt.scatter(X[:, 0], X[:, 1], s=10, color='b', alpha=0.5)
plt.title("Visualization of Synthetic Data")
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.show()
```



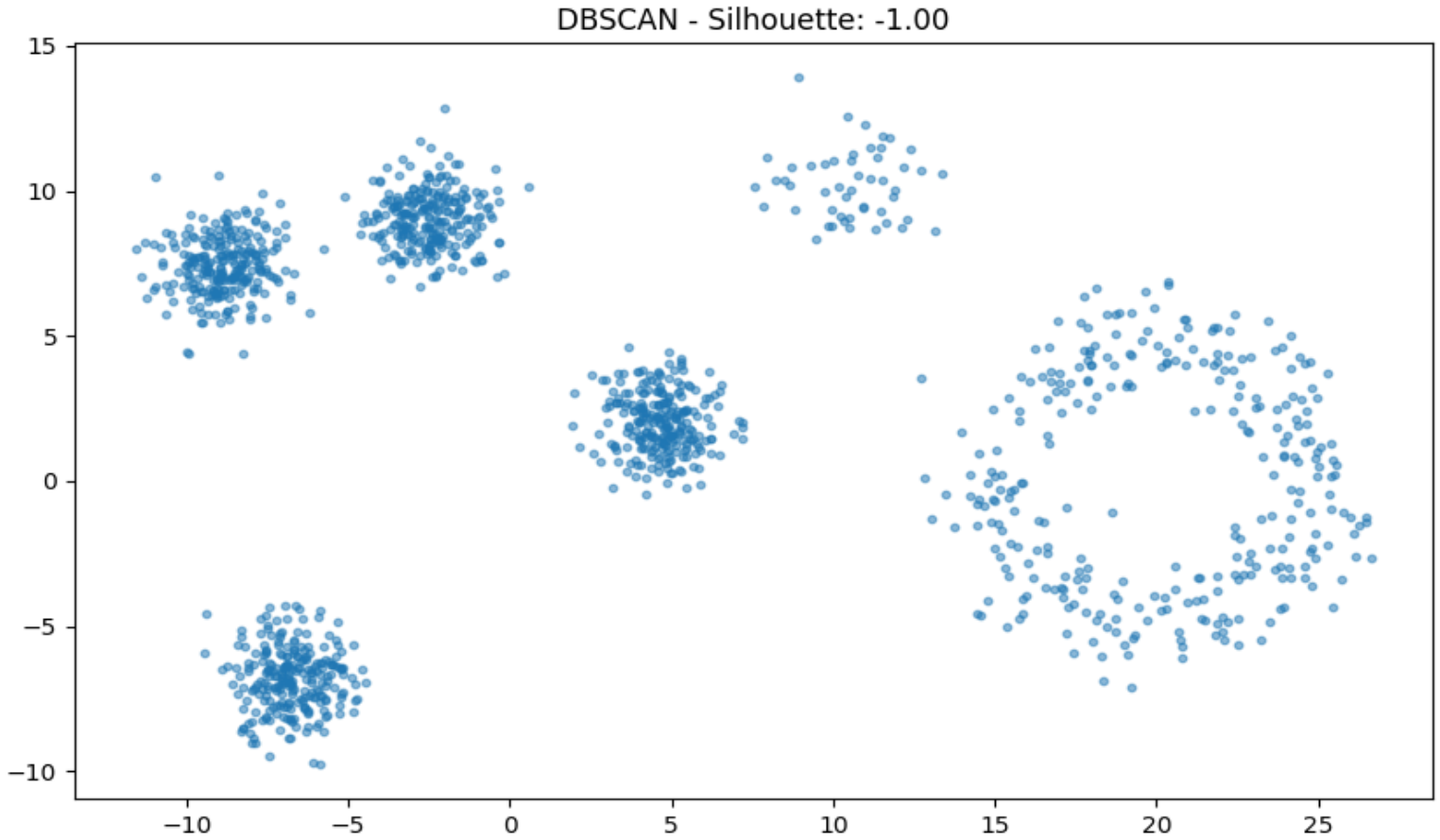
1



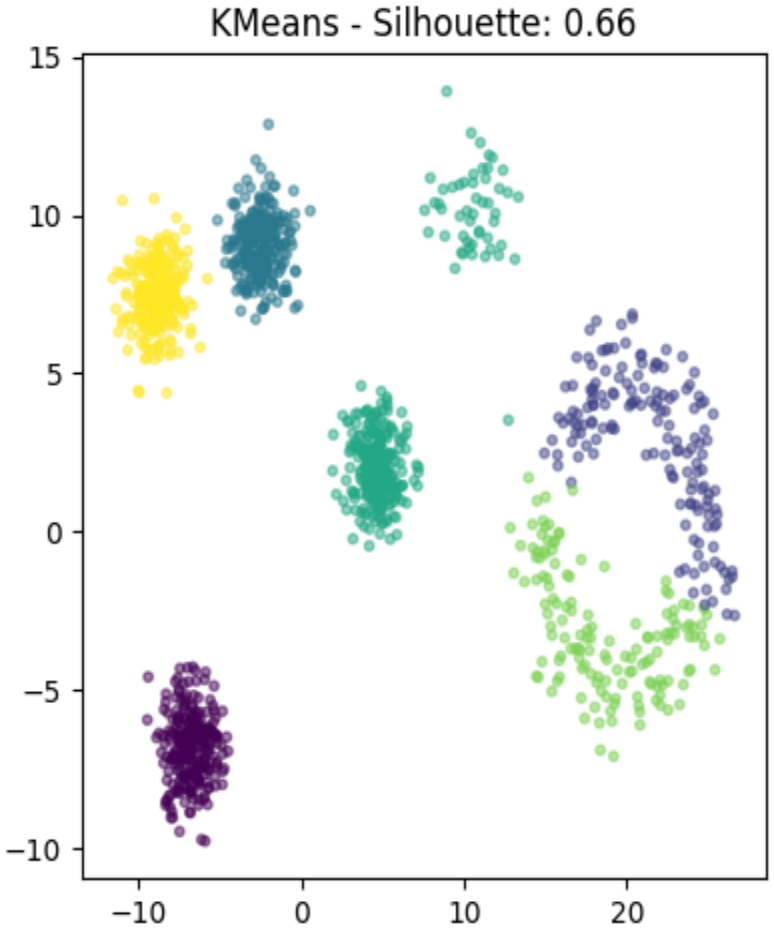

```
The optimal number of clusters is: 8
```

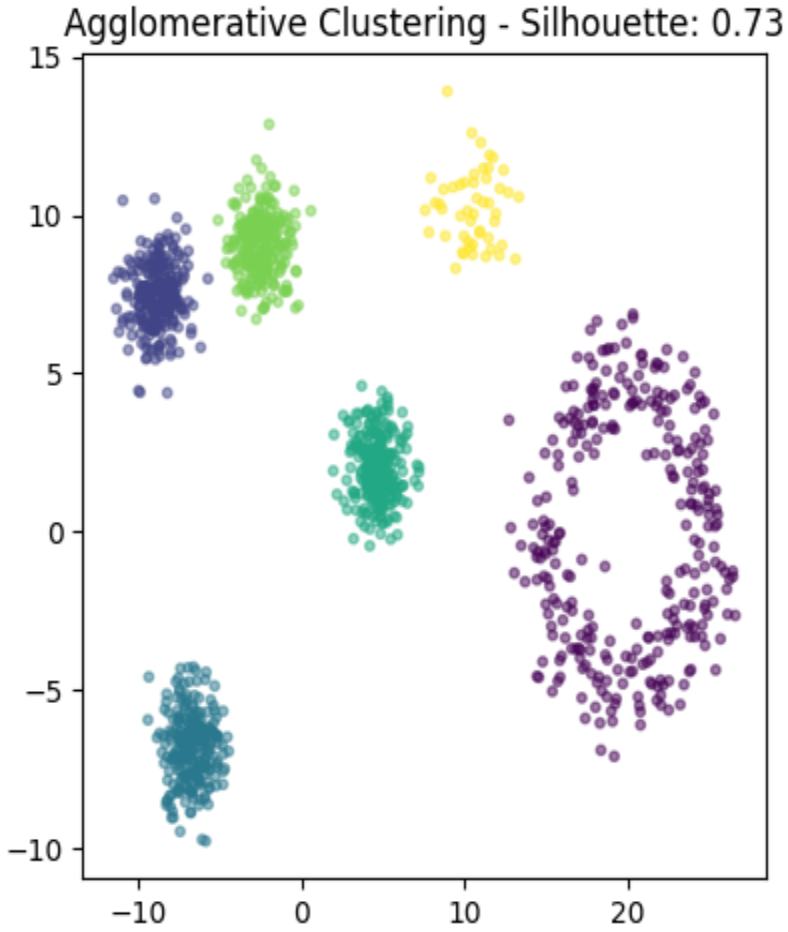


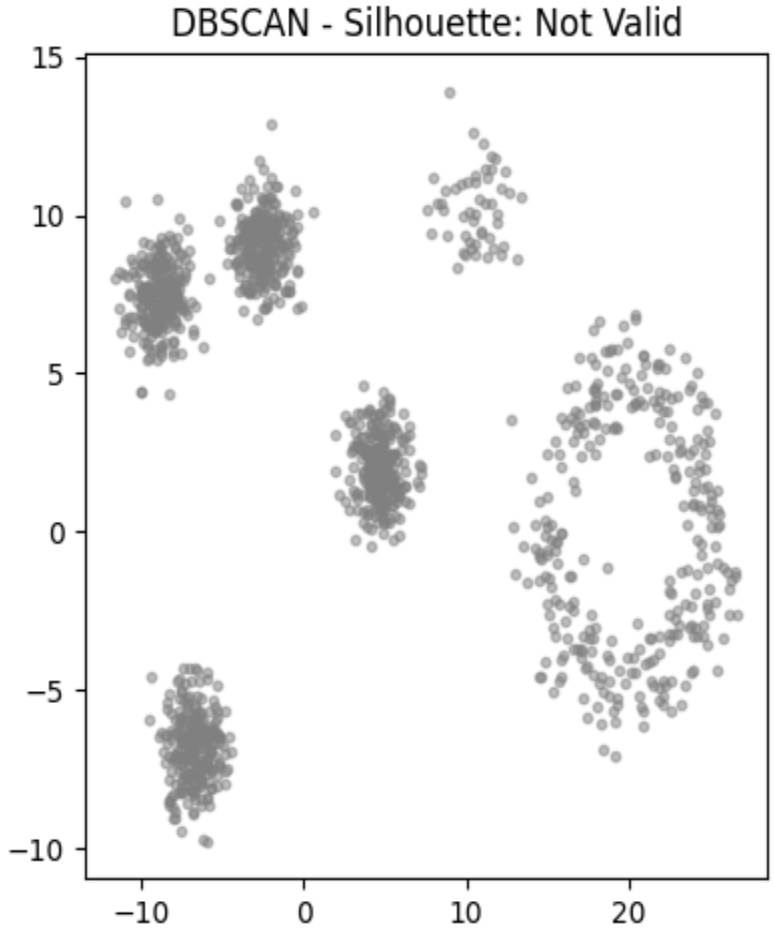

```
The optimal number of clusters is: 6
```



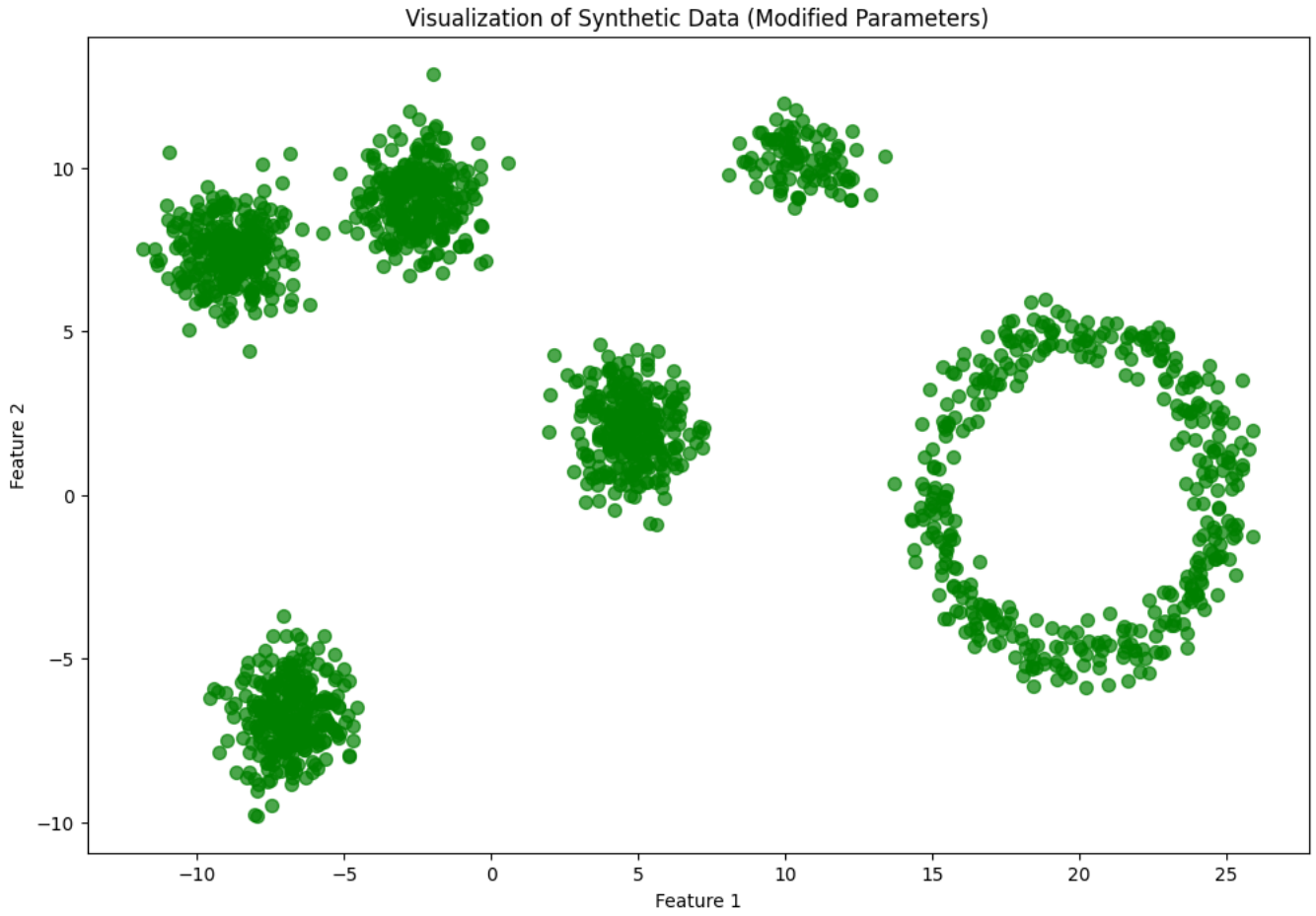
```
KMeans Silhouette Score: 0.6605985626907828
Agglomerative Clustering Silhouette Score: 0.7256900714288006
DBSCAN Silhouette Score: -1
The best method based on Silhouette Score is: Agglomerative Clustering
```



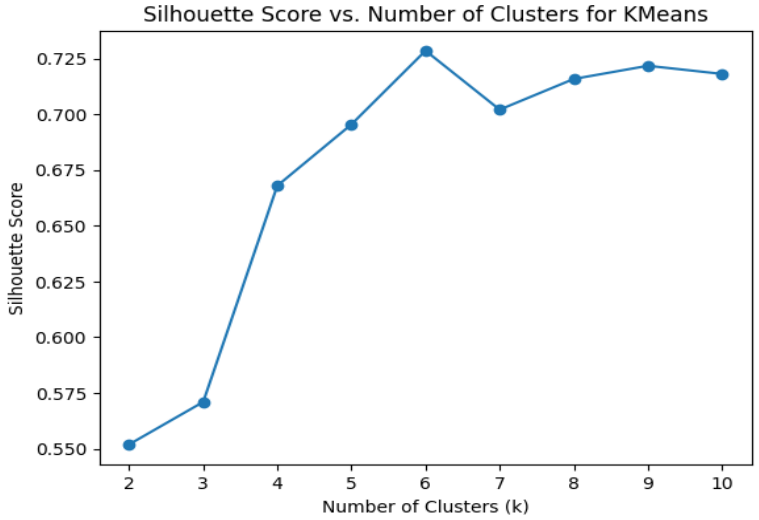




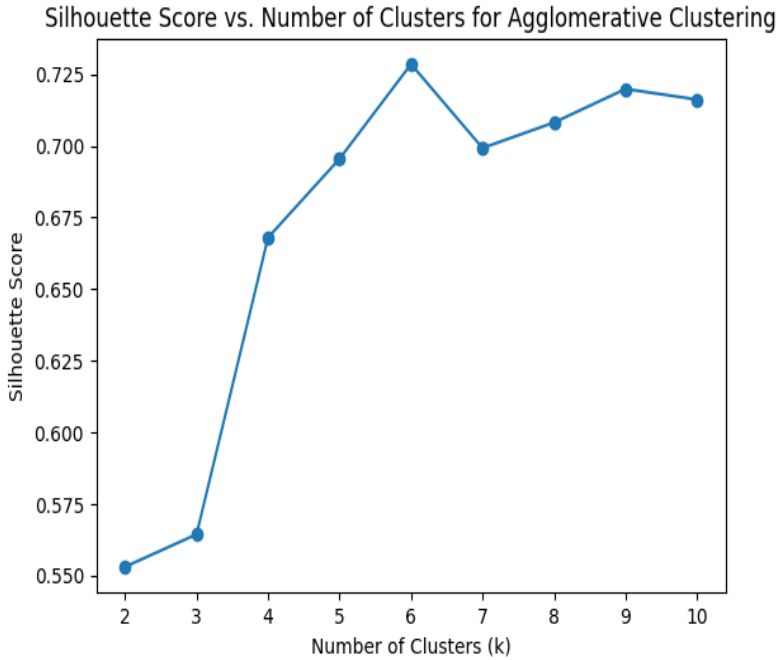
```python
X = create_data(0.5, 1200, 100, 400, 250)

plt.figure(figsize=(12, 8))
plt.scatter(X[:, 0], X[:, 1], s=50, color='g', alpha=0.7)
plt.title("Visualization of Synthetic Data (Modified Parameters)")
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.show()
```



2

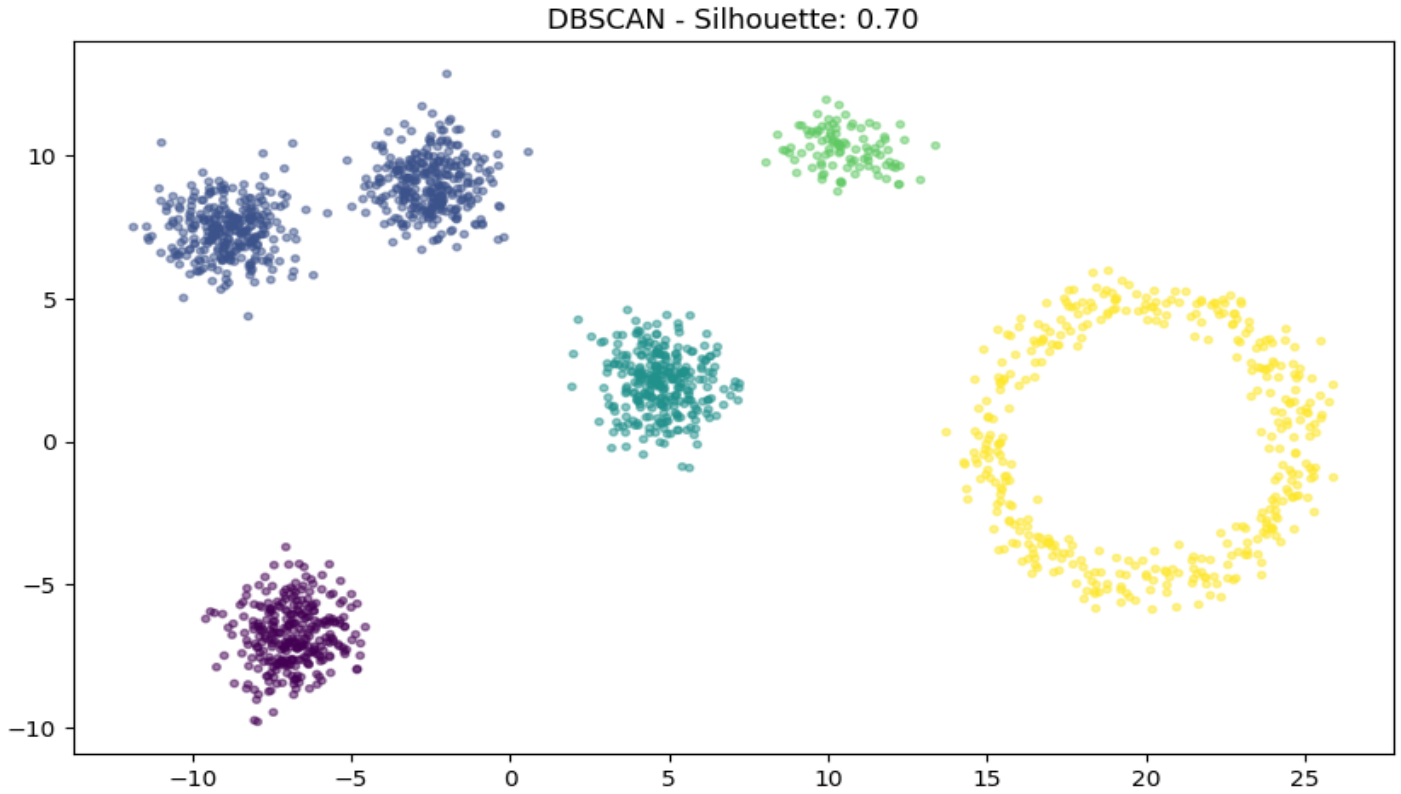



```
The optimal number of clusters is: 6
```

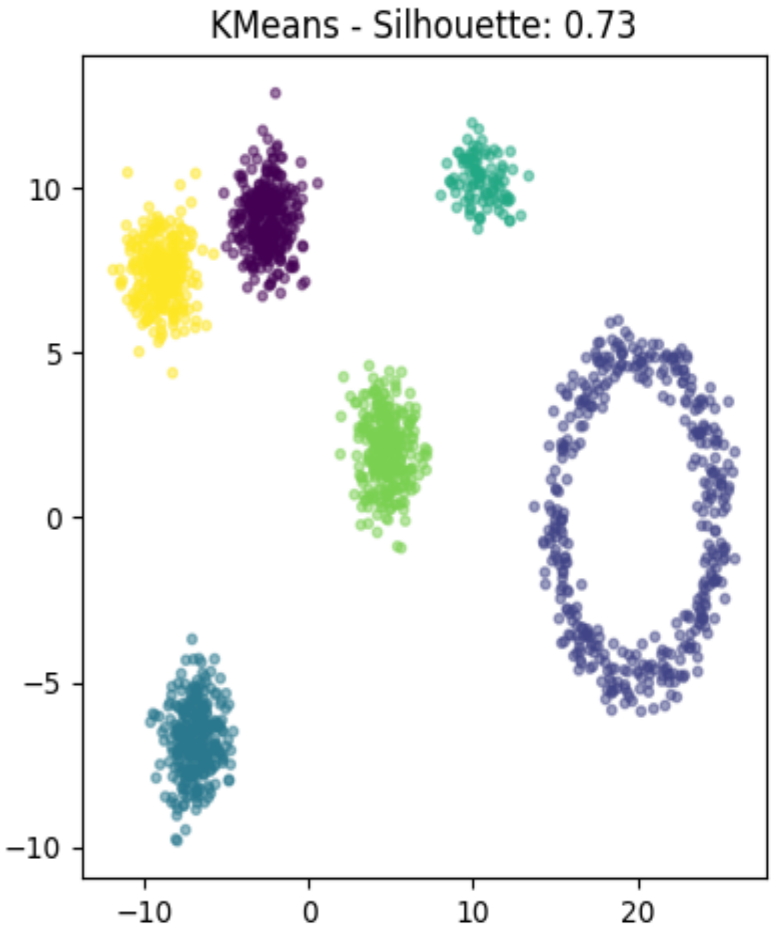


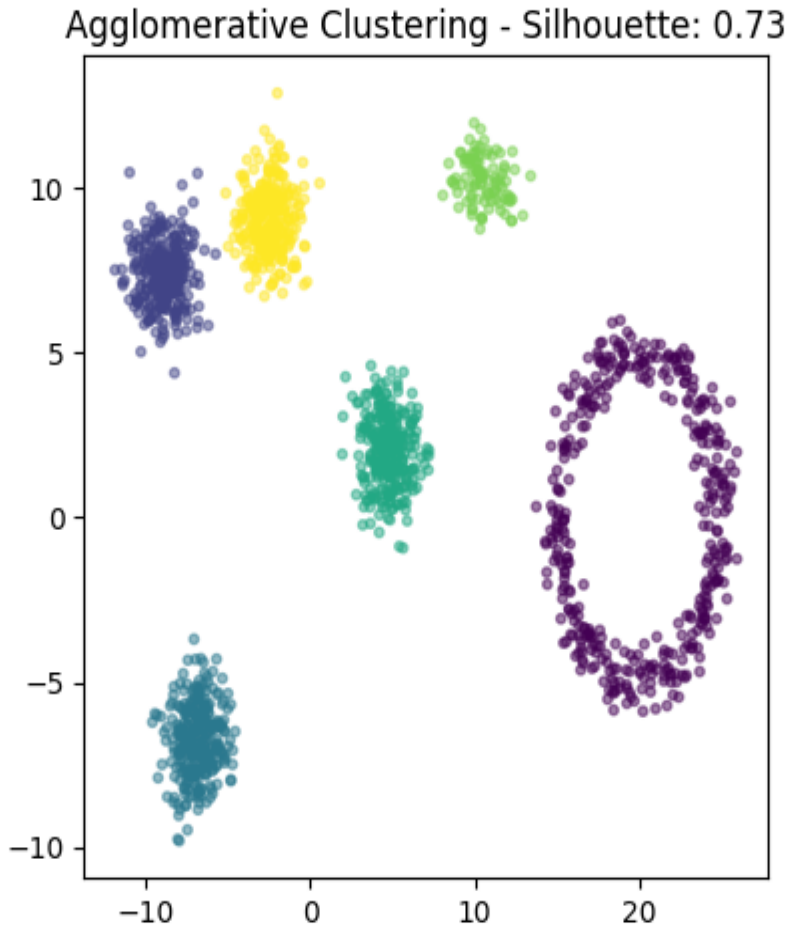

```
The optimal number of clusters is: 6
```

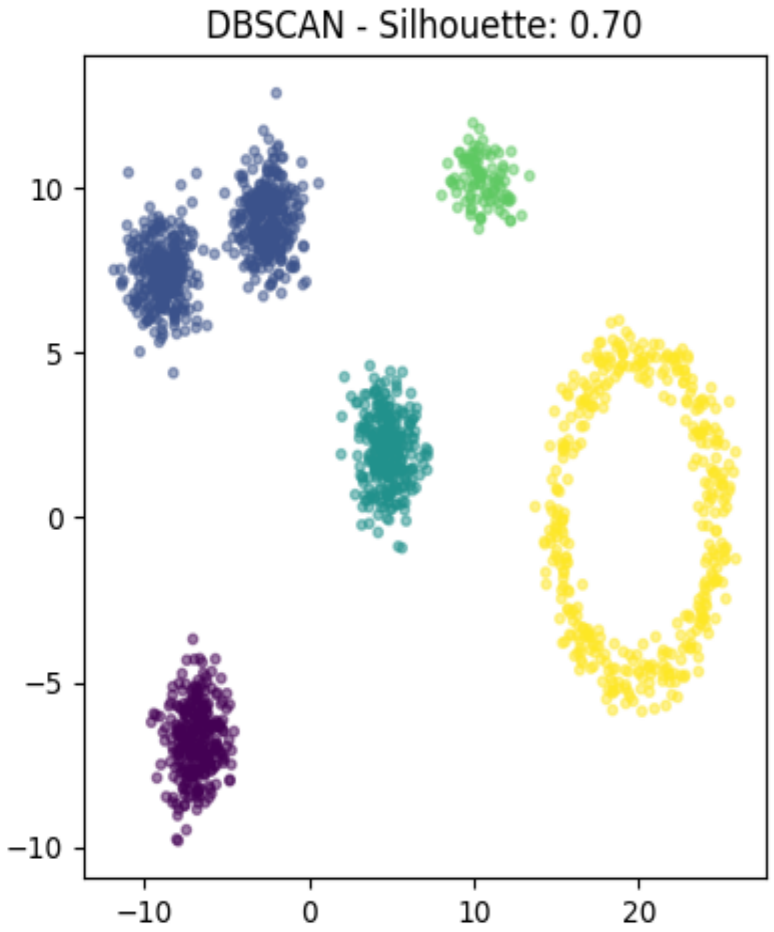
```
Best Silhouette Score: 0.6954765028529278
Best eps: 1.7, Best min_samples: 3
```



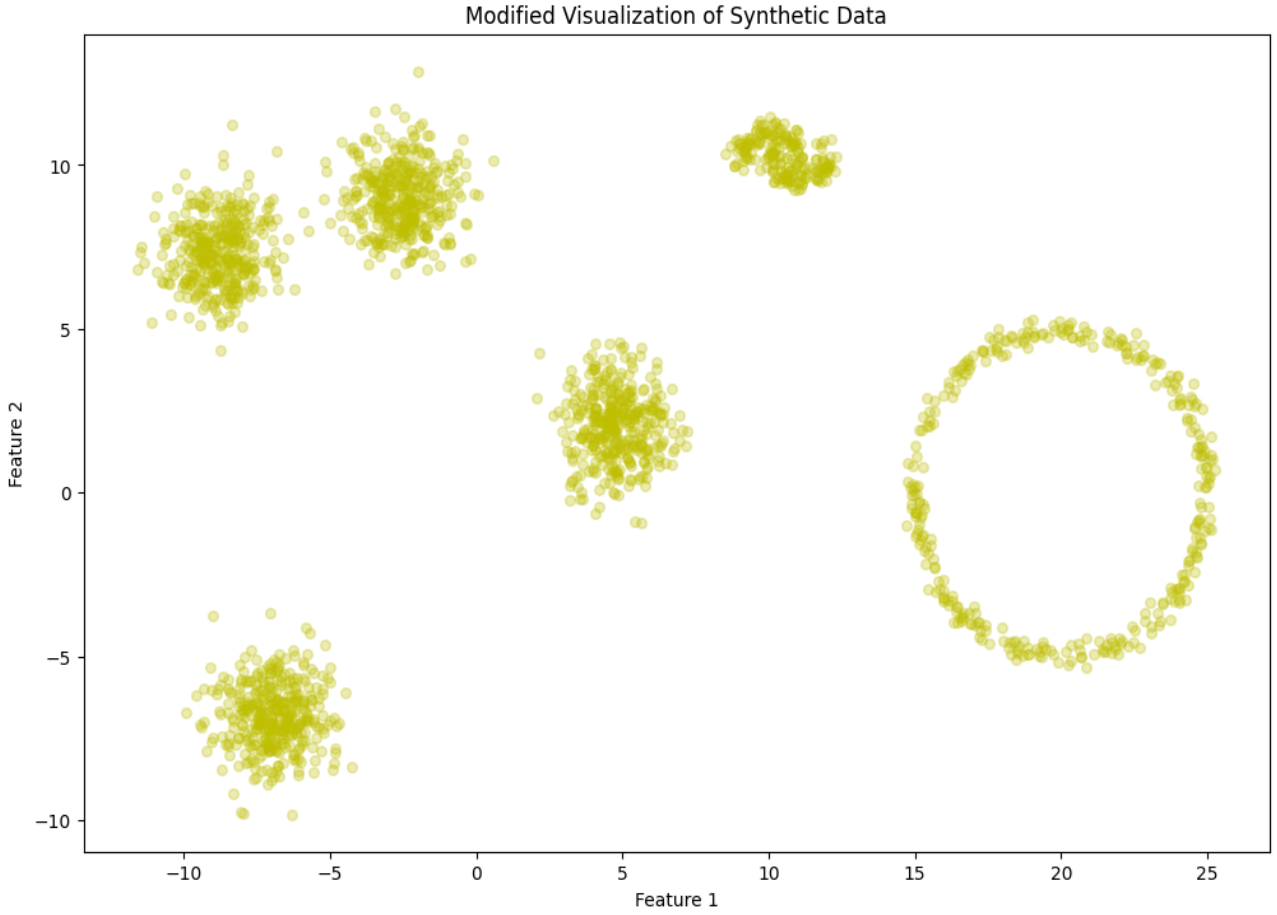
```
KMeans Silhouette Score: 0.7285246501052782
Agglomerative Clustering Silhouette Score: 0.7285246501052782
DBSCAN Silhouette Score: 0.6954765028529278
The best method based on Silhouette Score is: KMeans
```



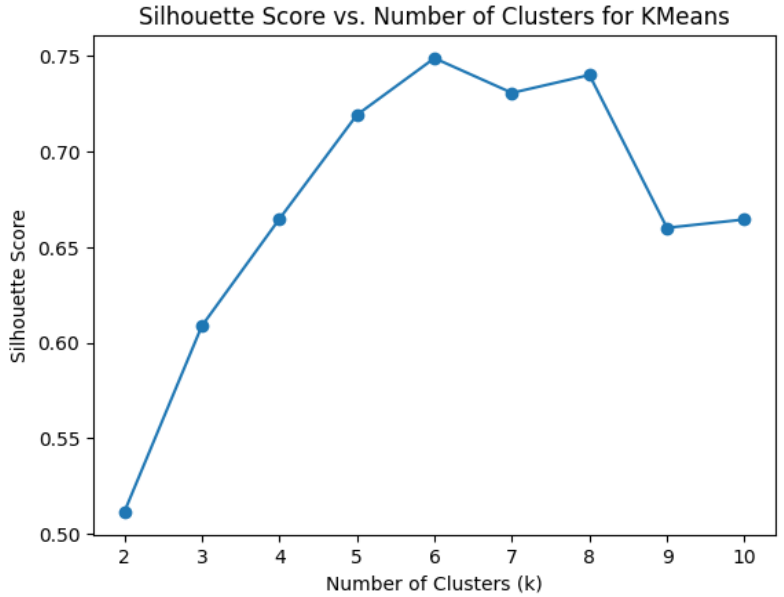




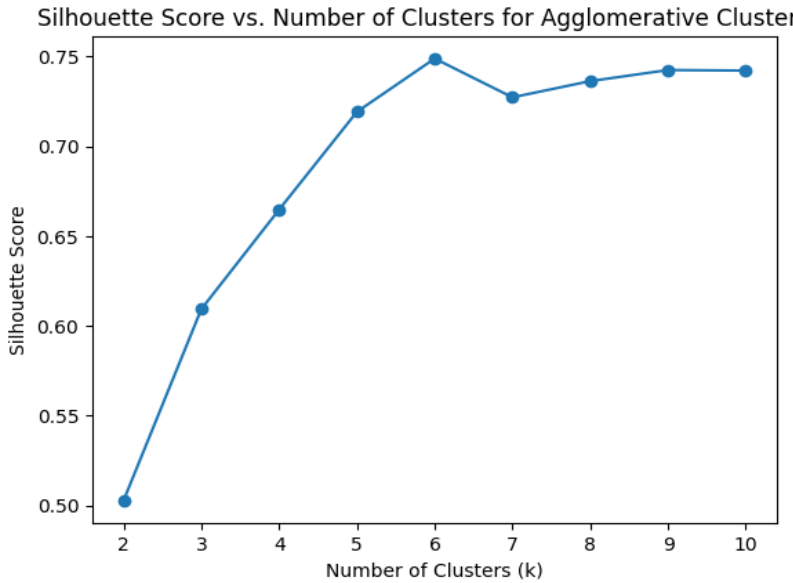
```python
X = create_data(0.2, 1500, 200, 350, 300)

plt.figure(figsize=(12, 8))
plt.scatter(X[:, 0], X[:, 1], s=30, color='y', alpha=0.3)
plt.title("Modified Visualization of Synthetic Data")
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.show()
```



3

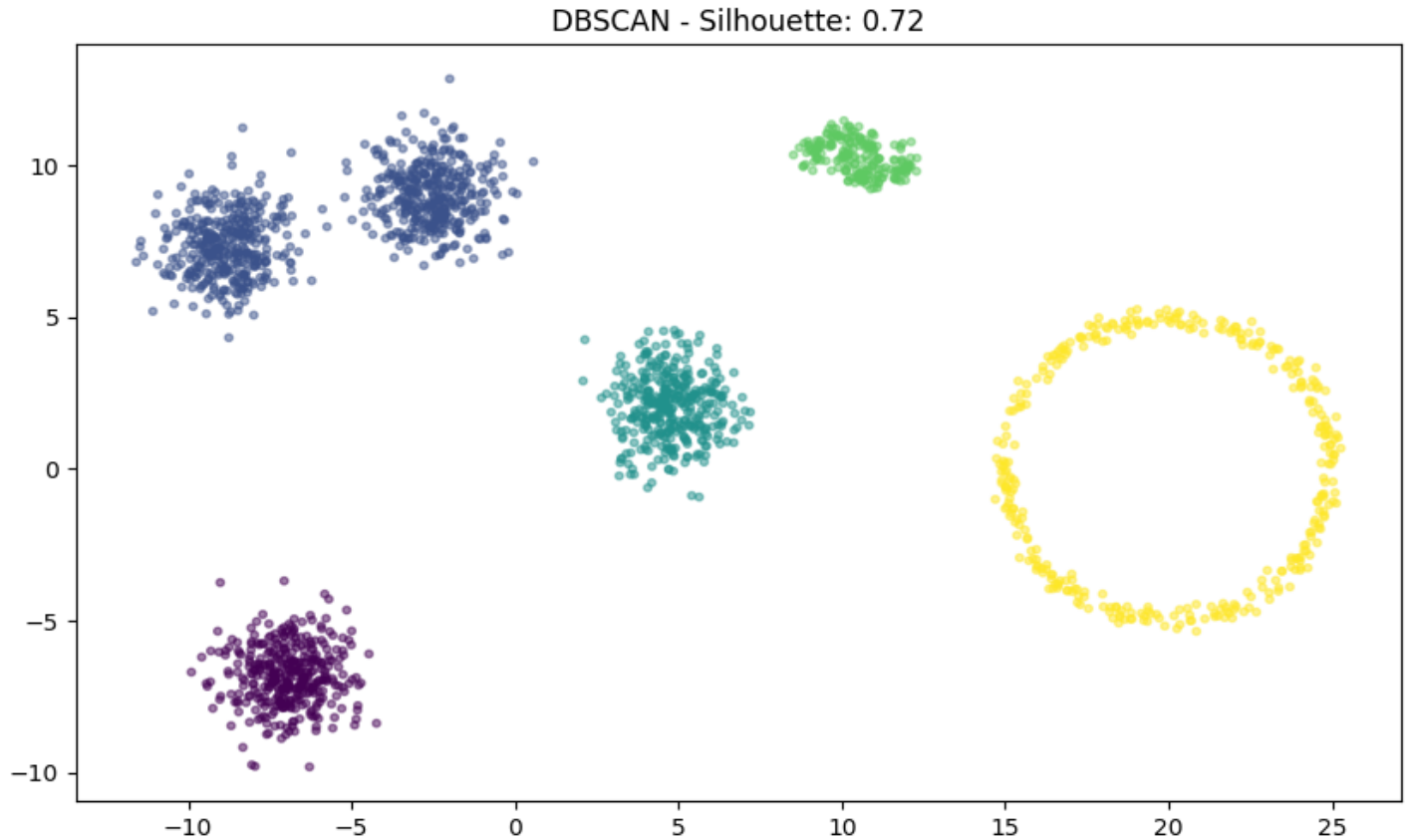



```
The optimal number of clusters is: 6
```

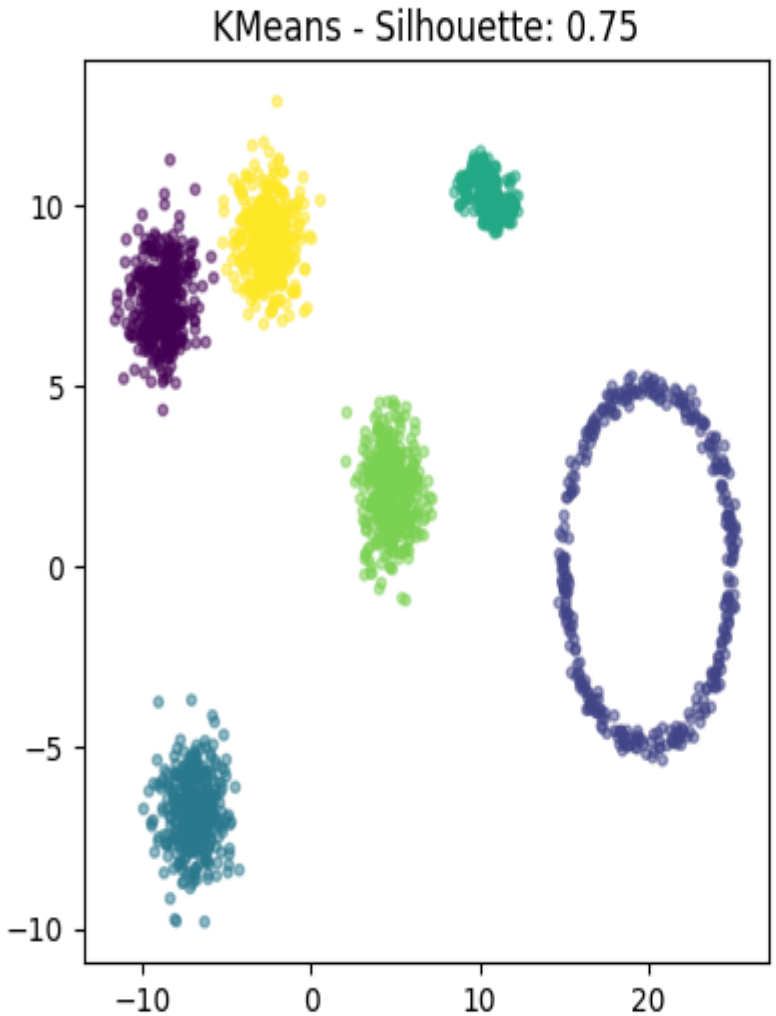


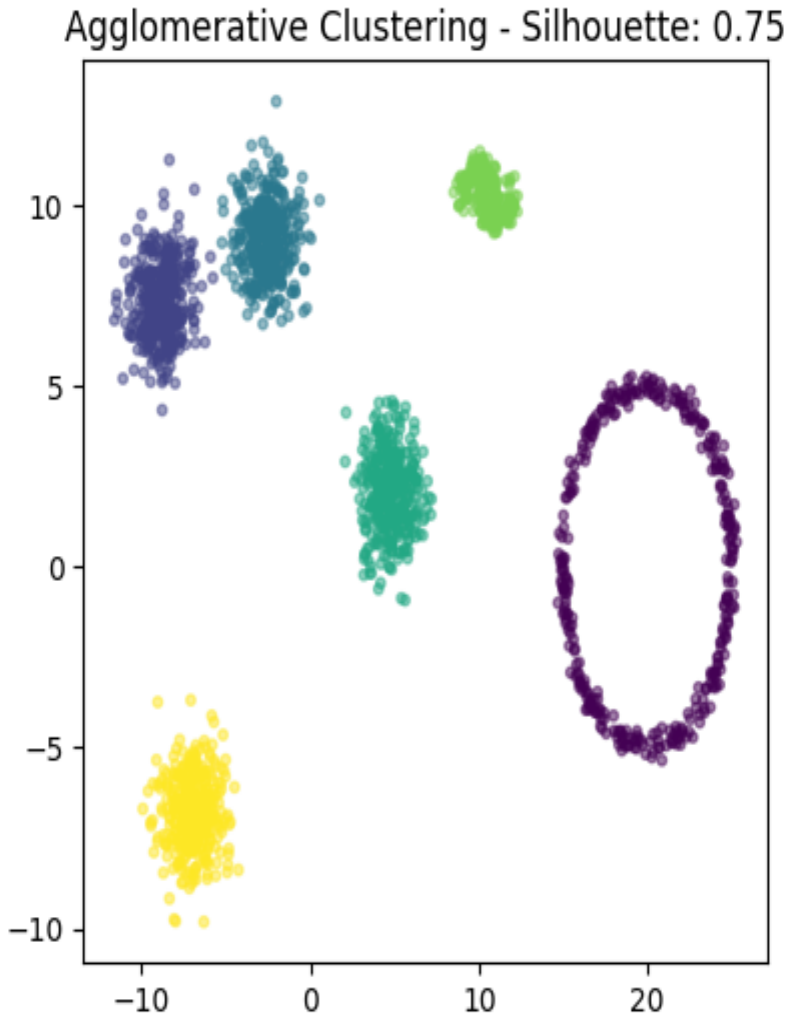

```
The optimal number of clusters is: 6
```

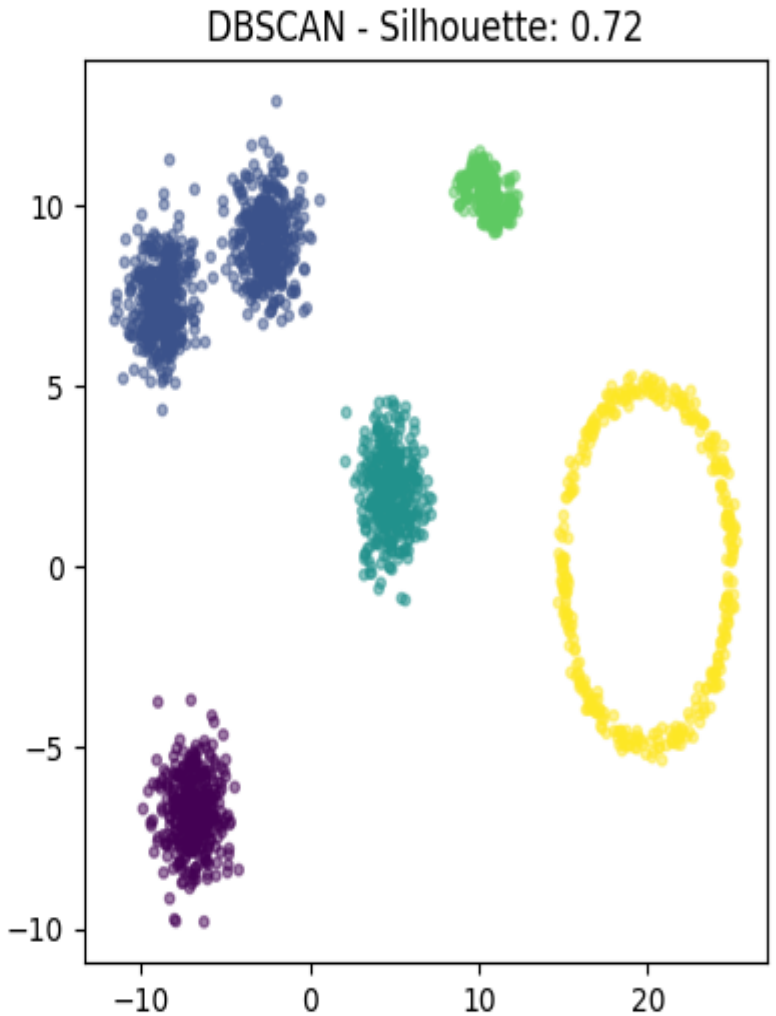
```
Best Silhouette Score: 0.719149549273575
Best eps: 1.7, Best min_samples: 3
```



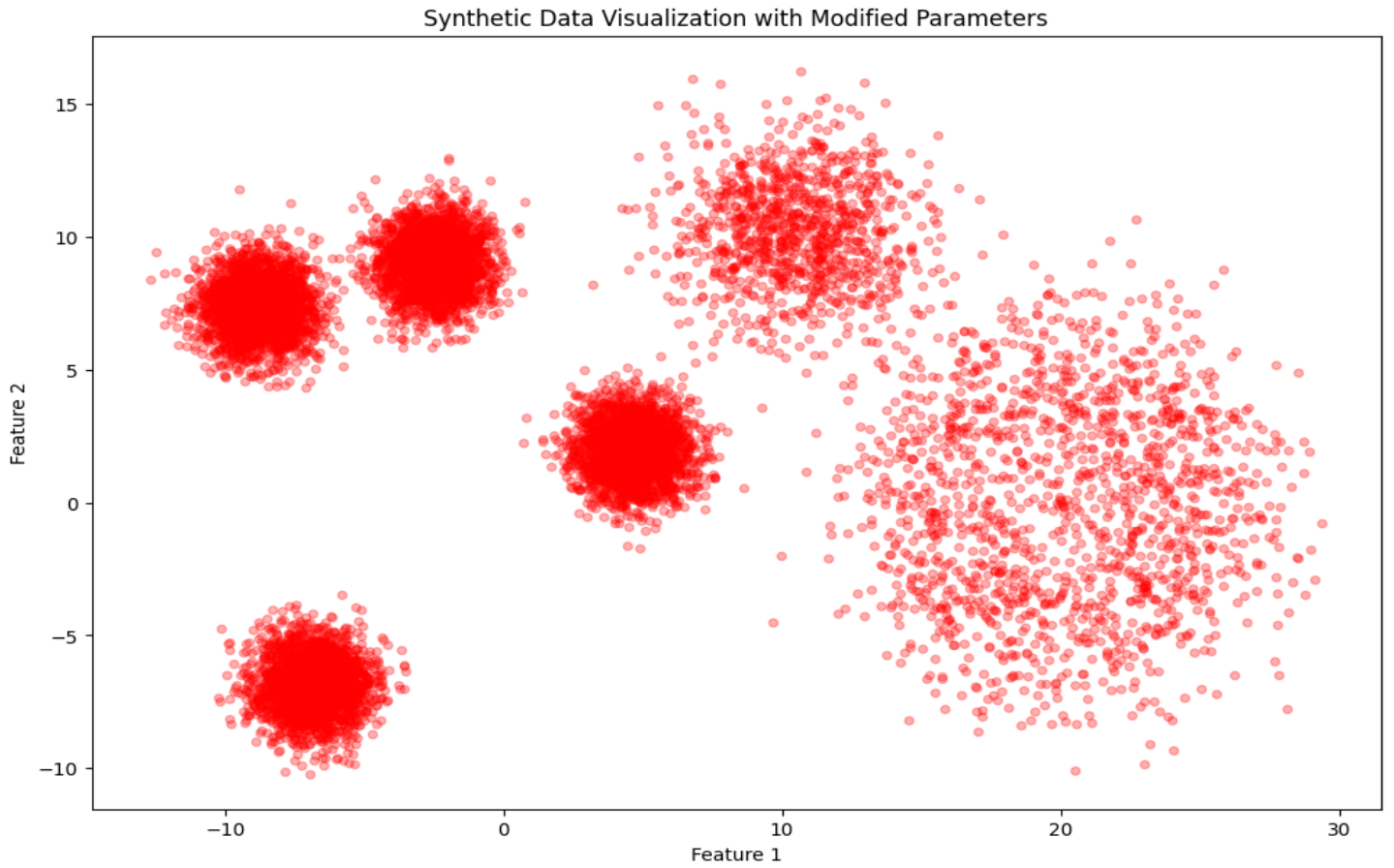
```
KMeans Silhouette Score: 0.7488249923862215
Agglomerative Clustering Silhouette Score: 0.7488249923862215
DBSCAN Silhouette Score: 0.719149549273575
The best method based on Silhouette Score is: KMeans
```



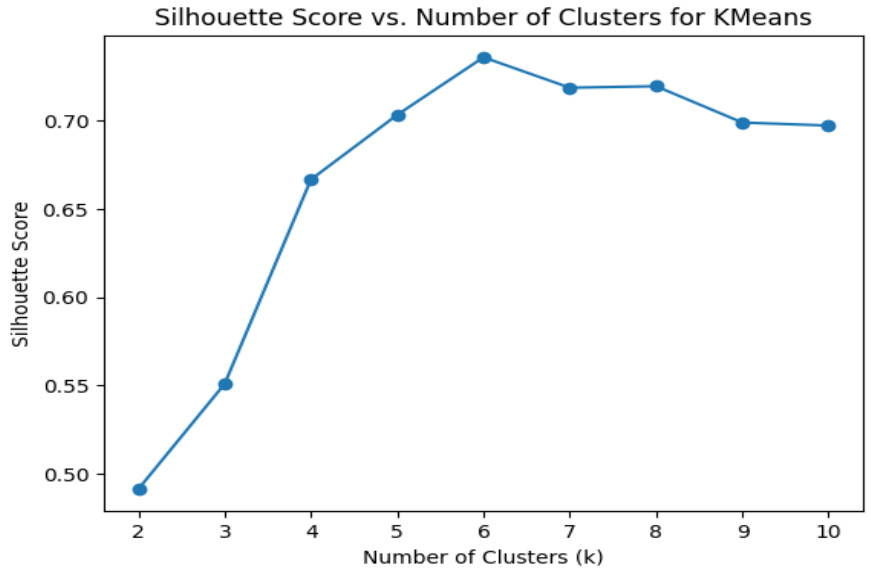




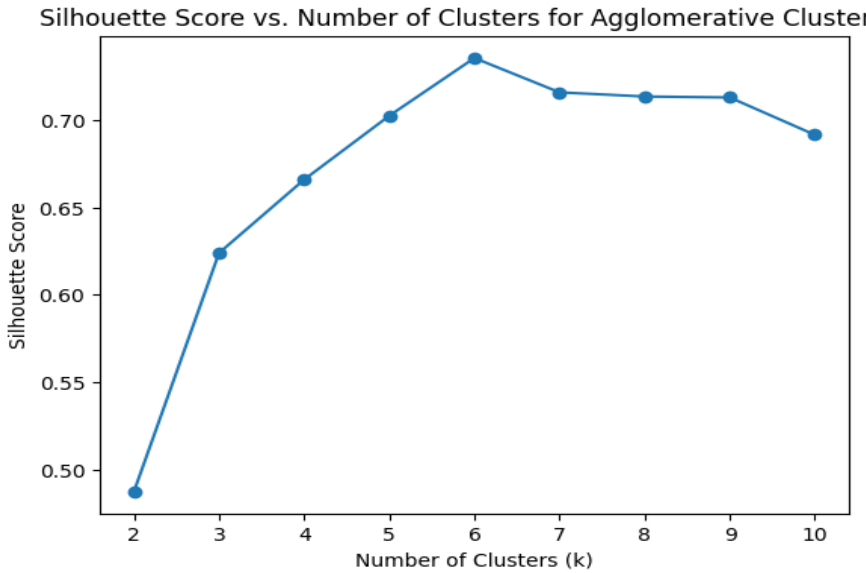
```python
X = create_data(2.0, 10000, 1000, 1500, 800)

plt.figure(figsize=(12, 8))
plt.scatter(X[:, 0], X[:, 1], s=20, color='r', alpha=0.3)
plt.title("Synthetic Data Visualization with Modified Parameters")
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.show()
```



4

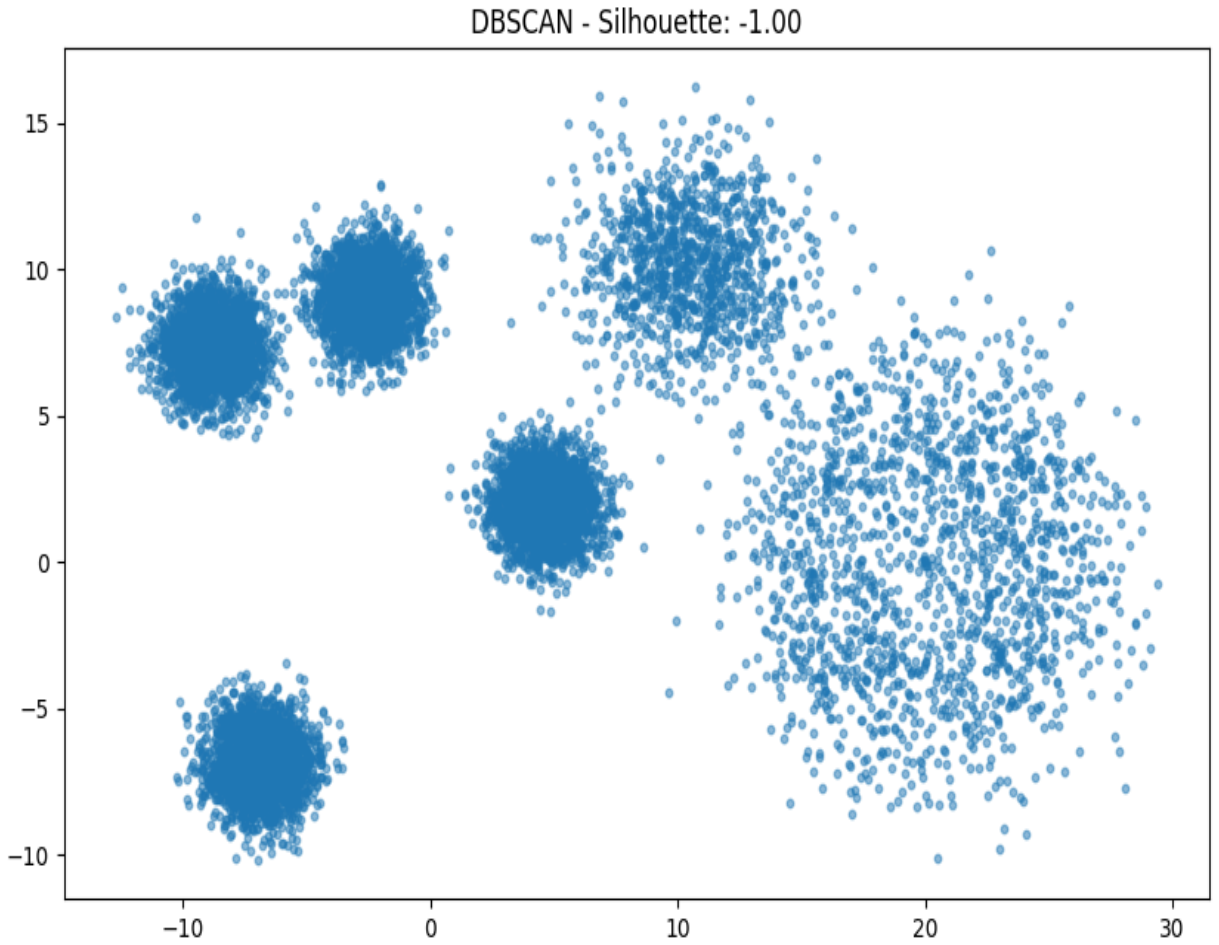



```
The optimal number of clusters is: 6
```

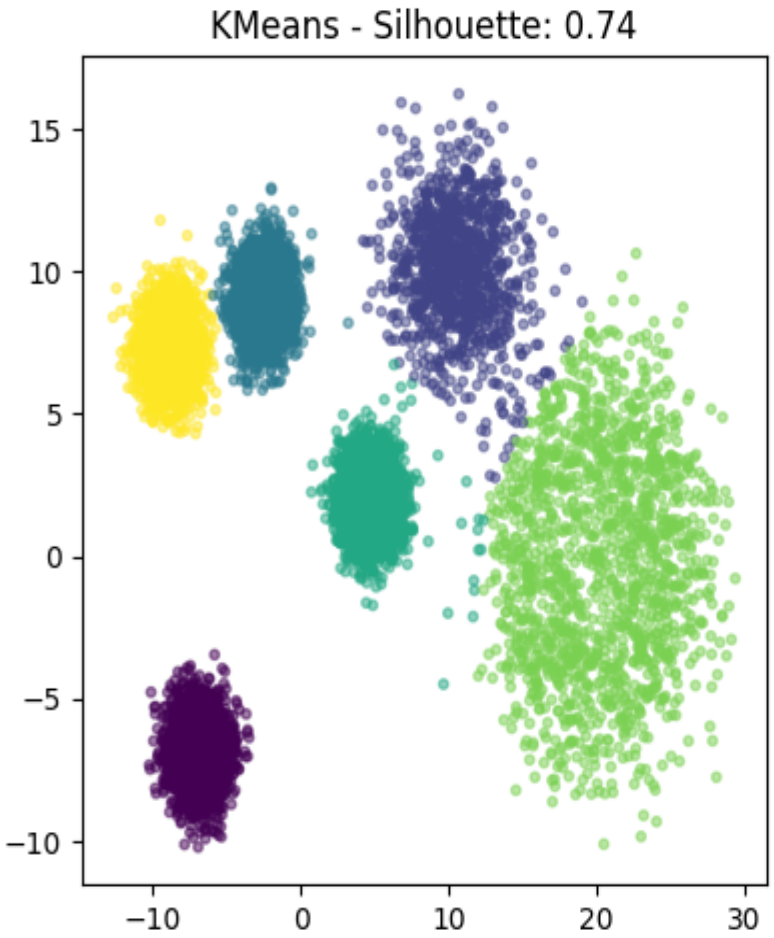


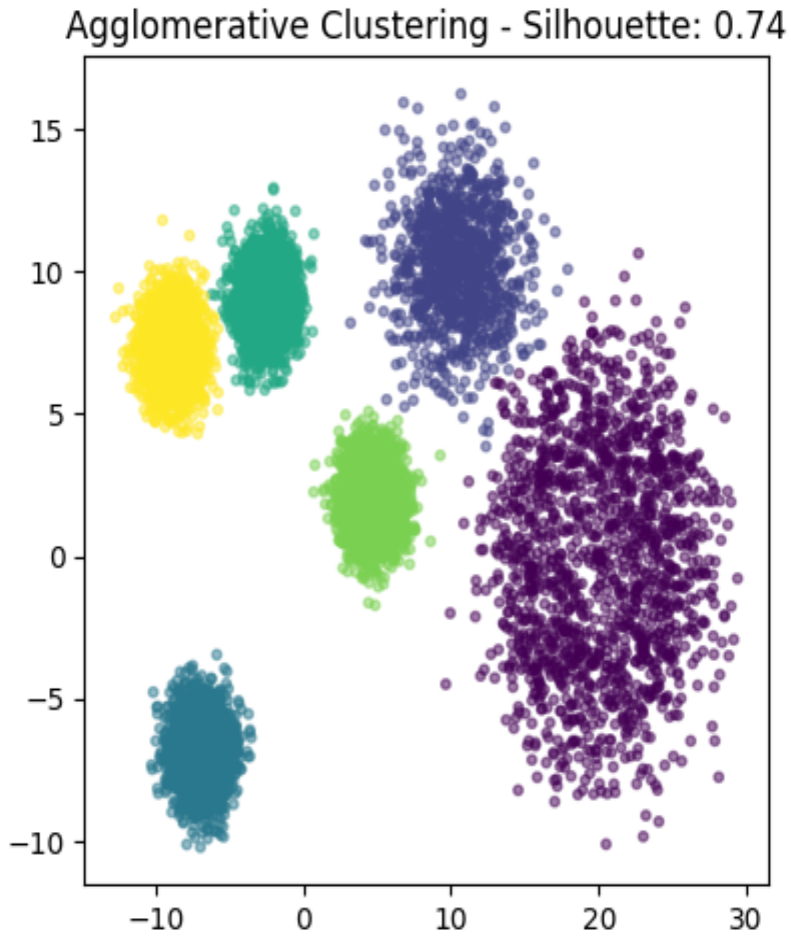

```
The optimal number of clusters is: 6
```

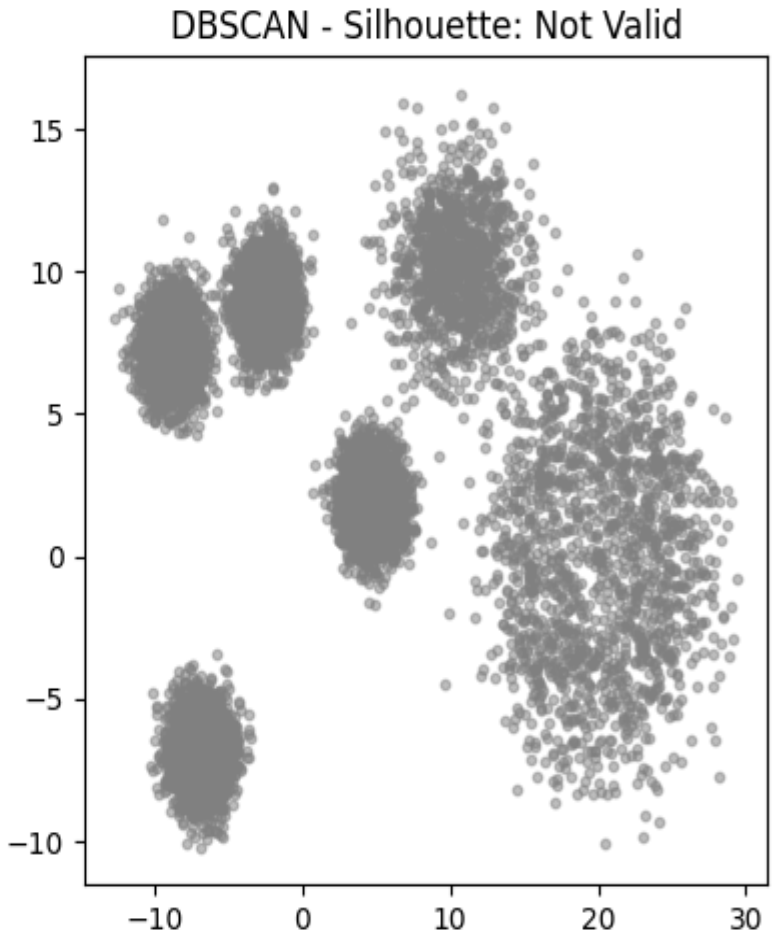
```
Best Silhouette Score: -1
Best eps: 0, Best min_samples: 0
<ipython-input-5-7d84c03185f3>:36: UserWarning: No data for colormapping provided via 'c'. Parameters 'cmap' will be ignored
  plt.scatter(X[:, 0], X[:, 1], c=best_labels, cmap='viridis', s=10, alpha=0.5)
```



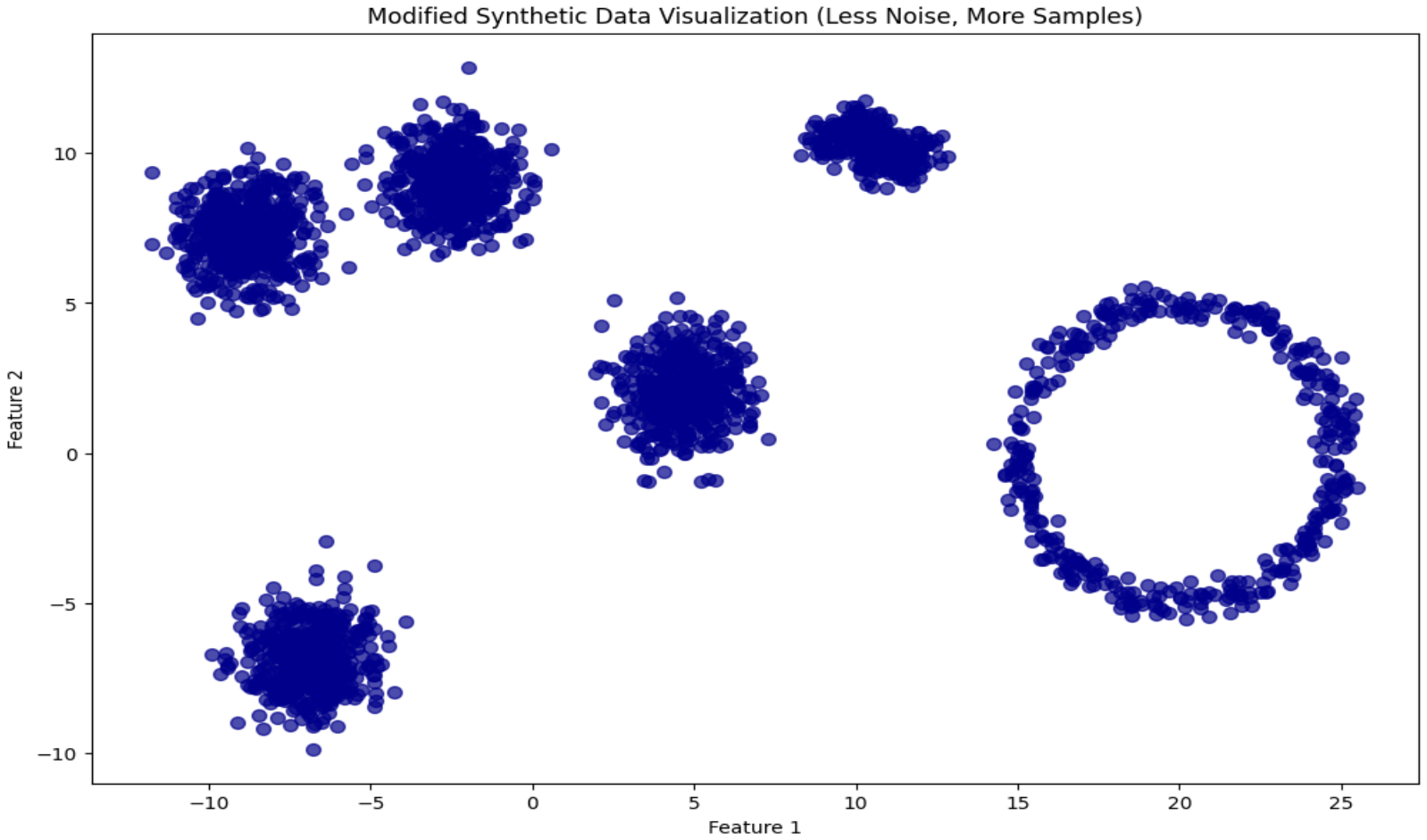
```
KMeans Silhouette Score: 0.7358972642465359
Agglomerative Clustering Silhouette Score: 0.7356286933917509
DBSCAN Silhouette Score: -1
The best method based on Silhouette Score is: KMeans
```



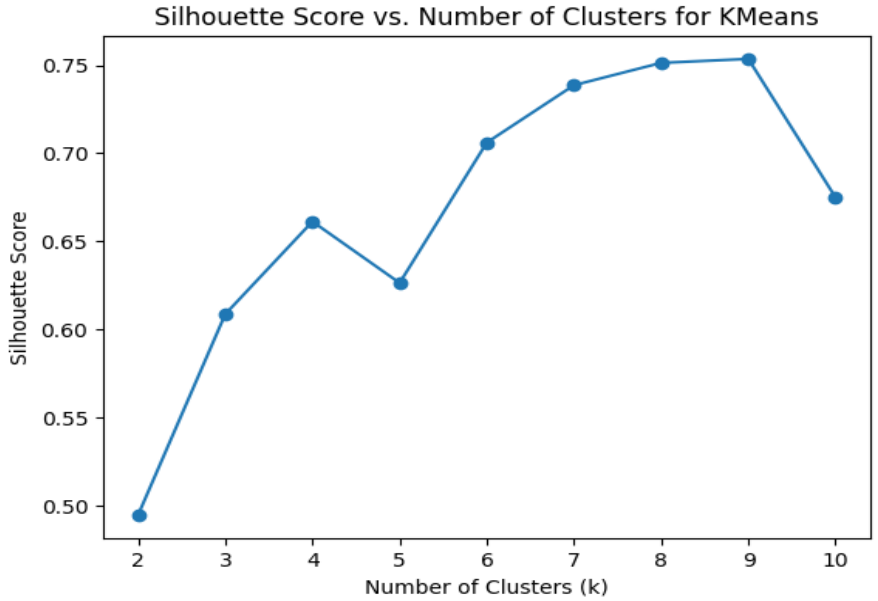




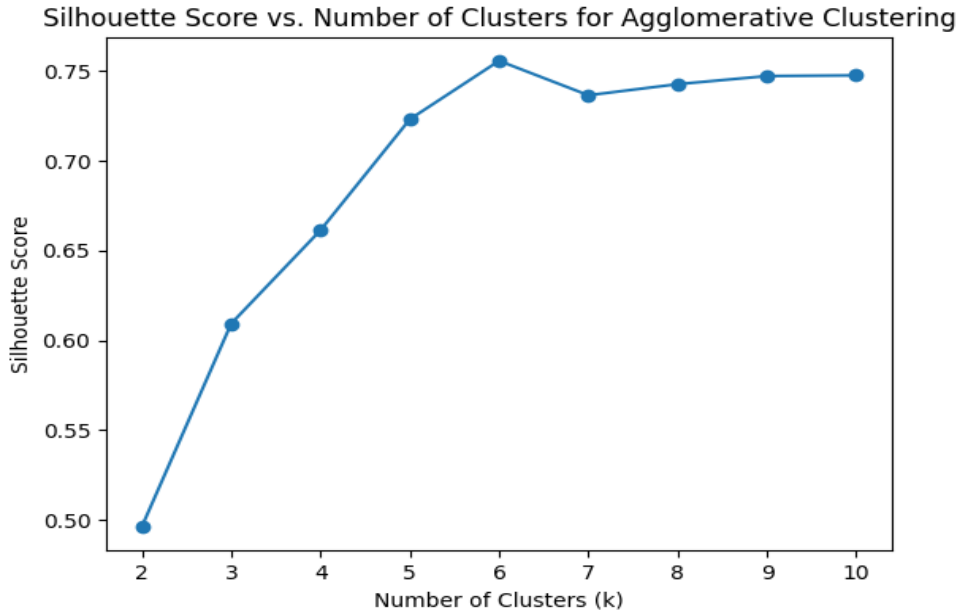
```python
X = create_data(0.3, 2000, 300, 400, 1000)

plt.figure(figsize=(12, 8))
plt.scatter(X[:, 0], X[:, 1], s=50, color='darkblue', alpha=0.7)
plt.title("Modified Synthetic Data Visualization (Less Noise, More Samples)")
plt.xlabel("Feature 1")
plt.ylabel("Feature 2")
plt.show()
```



5

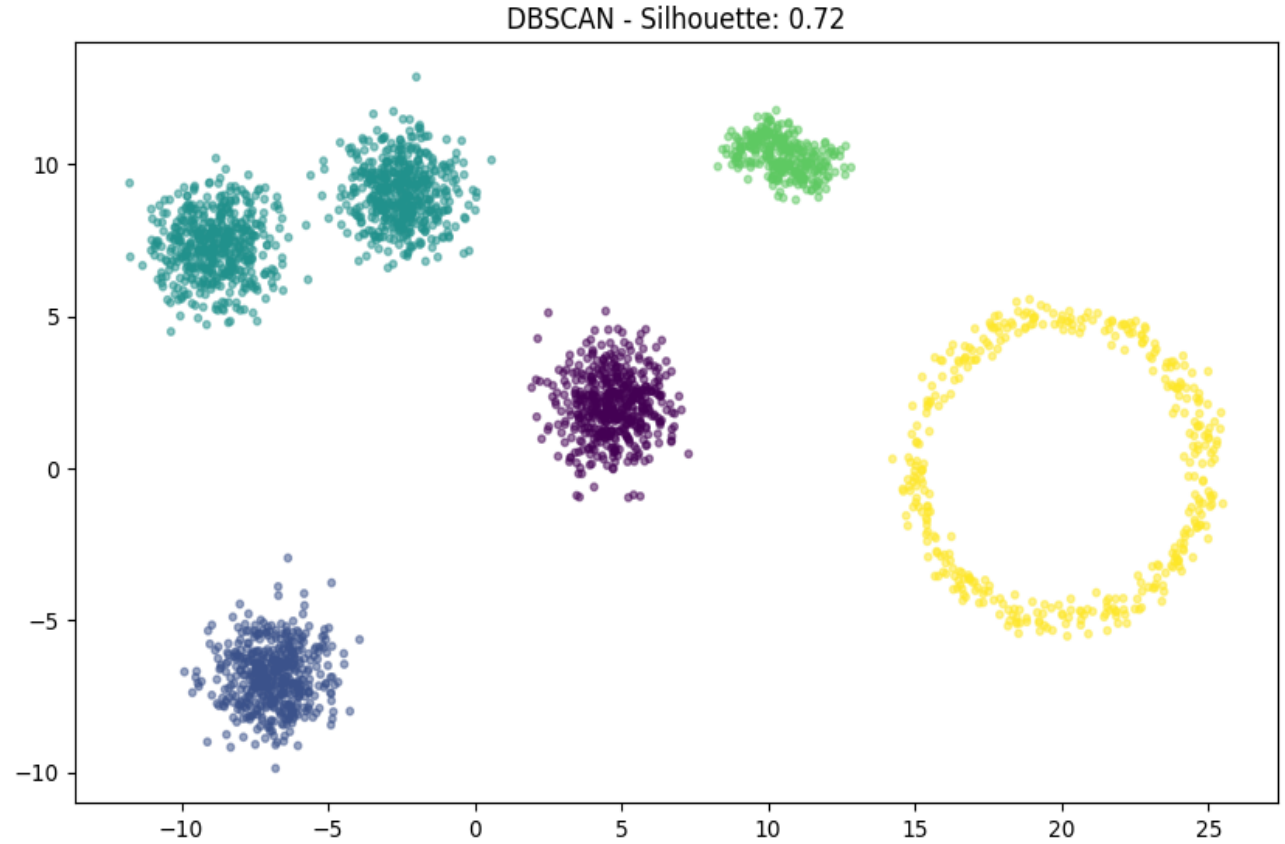



```
The optimal number of clusters is: 9
```




```
The optimal number of clusters is: 6
```

```
Best Silhouette Score: 0.7230126060091746
Best eps: 1.4, Best min_samples: 3
```



```
KMeans Silhouette Score: 0.7060778632974658
Agglomerative Clustering Silhouette Score: 0.7555585669877418
DBSCAN Silhouette Score: 0.7230126060091746
The best method based on Silhouette Score is: Agglomerative Clustering
```