

A Doc to answer questions that aren't included in the Google Slides.

- Implement a method to determine the optimal number of clusters using methods like the Elbow Method or Silhouette Analysis. Discuss different techniques for finding the optimal number of clusters and explain your chosen method.

The optimal number of clusters using the Elbow Method is 5. Different techniques, such as the Elbow Method and the Silhouette Analysis, have different ways of finding the optimal number of clusters for a clustering algorithm. The Silhouette Analysis calculates a silhouette score of the data points to determine which cluster they fit into best. Another technique is the dendrogram, which determines which longest vertical lines do not cross horizontal lines to indicate the optimal number of clusters. The method I chose, the Elbow Method, determines the best cluster amount by identifying the elbow of the curve in the plot that bends.

```
[12]: sse = []
      k_range = range(1,11)
      for k in k_range:
          km = KMeans(n_clusters=k, random_state=42)
          km.fit(X_scaled)
          sse.append(km.inertia_)

      plt.figure(figsize=(8, 4))
      plt.plot(k_range, sse, 'bo-')
      plt.xlabel('Number of clusters (k)')
      plt.ylabel('Sum of Squared Errors (SSE)')
      plt.title('Elbow Method for Optimal k')
      plt.grid(True)
      plt.show()
```

