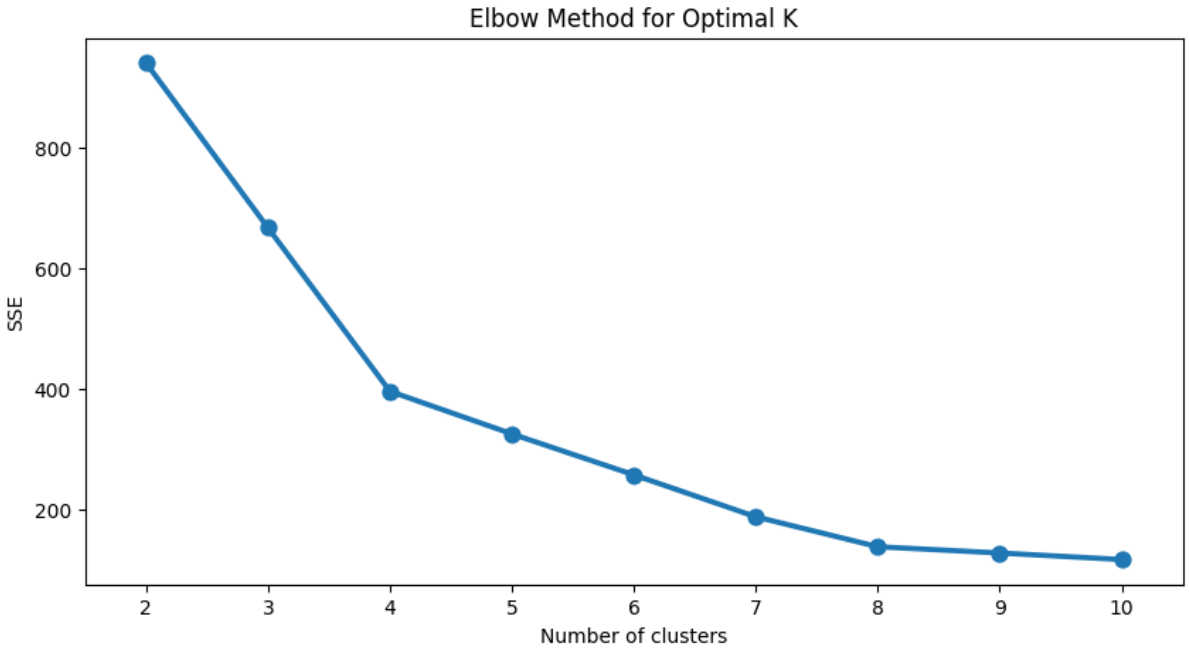
Task 3: Analysis of Clustering Results



Number of Clusters:

The Elbow Method graphic suggests that 4 clusters is the ideal quantity. This is because there is a noticeable "elbow" or bend in the plot where there are four clusters. After this, there are diminishing returns in clustering accuracy, as evidenced by the Sum of Squared Errors (SSE) decreasing less significantly.

Other Relevant Clustering Metrics:

Although they are not displayed directly in the plot, the following additional clustering metrics can be used to assess the calibre of the clustering results:

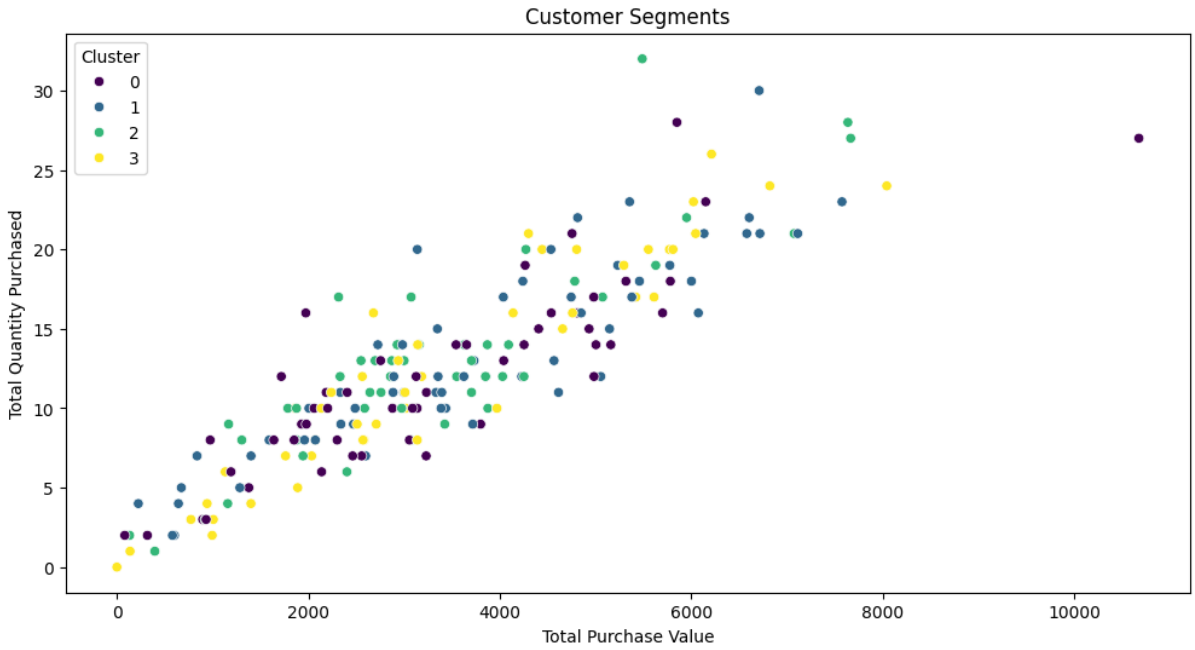
1. Silhouette Score: A data point's silhouette score indicates how similar it is to its own cluster in relation to other clusters. Better-defined clusters are indicated by a higher Silhouette Score.

2. Calinski-Harabasz Index: Calculates the proportion of variance inside a cluster to variance between clusters. Better-separated clusters are indicated by a higher value.

3. Davies-Bouldin Index: The average similarity between each cluster and its most comparable cluster is measured by the Davies-Bouldin Index. Better-separated clusters are indicated by a lower value.

Overall Interpretation:

Although the number of clusters can be accurately estimated using the Elbow Method, it is advised to validate and improve the clustering results using a mix of criteria Davis-Bouldin Index.



Number of Clusters:

We can see from the scatter plot that the data points are arranged into four separate clusters, each of which is represented by a different colour.

DB Index Value:

The degree of separation between the clusters is shown by the given DB Index value of 0.7434534867259787.

- In this instance, a DB Index of **0.7434534867259787** indicates a moderate degree of separation between the clusters.
- In general, a lower DB Index number indicates better-separated clusters.

Other Relevant Clustering Metrics (Inferred from the Scatter Plot):

- **Visual Inspection:** The scatter plot indicates that there may be some overlap between the clusters, suggesting that the separation may not be flawless.
- **Potential for Improvement:** To perhaps enhance cluster separation and obtain more distinct customer segments, additional research and possibly other clustering methods may be investigated, contingent on the particular business objectives.

Overall Interpretation:

The four clusters that were created indicate different client segments according to the overall amount and value of their purchases. To learn more about their tastes, shopping patterns, and business value, these categories can be further examined.