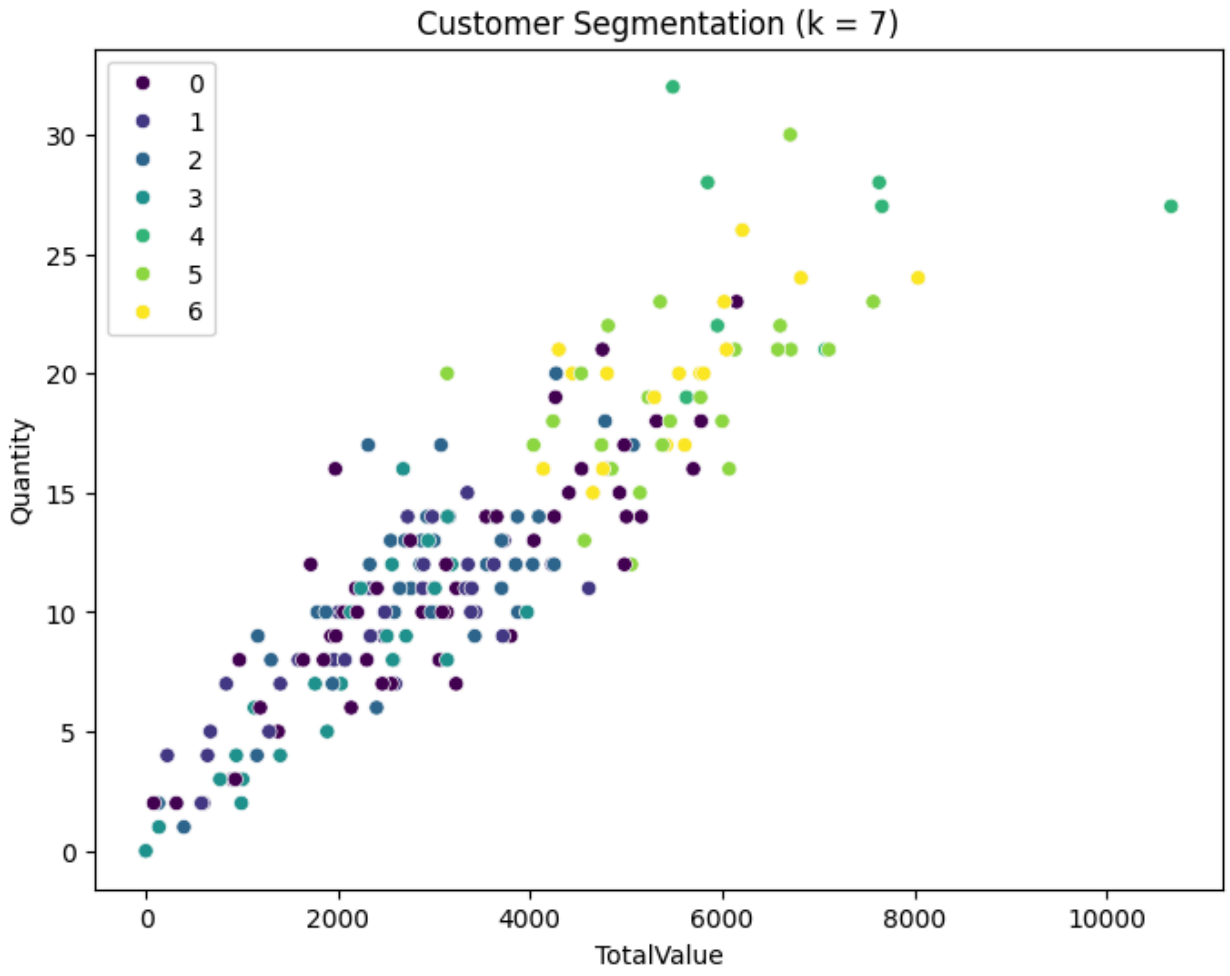
Task 3: Customer Segmentation / Clustering

Customer Segmentation Report

KMeans Clustering:

Optimal number of clusters: 7
Davies-Bouldin Index: 0.67
Silhouette Score: 0.58



Agglomerative Clustering:

Optimal Number of Clusters: 8
Davies-Bouldin Index: 0.59
Silhouette Score: 0.57

Key Observations:

* KMeans: The KMeans algorithm identified 7 optimal clusters based on the Silhouette Score. A Davies-Bouldin Index of 0.67 suggests a relatively good separation between the clusters.

* Agglomerative Clustering: Agglomerative clustering recommends 8 as the best number of clusters. The Silhouette Score and Davies-Bouldin Index show a level of performance similar to KMeans for this data set.

* Visualization: The scatter plots depict the customer segments by TotalValue and Quantity. Note the clear clusters generated by each algorithm. Additional analysis can be conducted to identify the nature of each cluster.

Further Analysis Recommendations:

1. Cluster Profiling: Examine the attributes of each customer segment (e.g., demographics, buying habits, geographic location).

2. Feature Importance: Determine the most important features driving the cluster separation.
3. Actionable Business Insights: Generate actionable business insights to support targeted marketing campaigns, product suggestions, or customer retention initiatives based on the segments identified.
4. Other Clustering Methods: Investigate alternative clustering methods such as DBSCAN to see how they perform.
5. Parameter Tuning: Try varying parameter values (e.g., distance measures, linkage types) for Agglomerative Clustering and KMeans to see if there are performance gains.