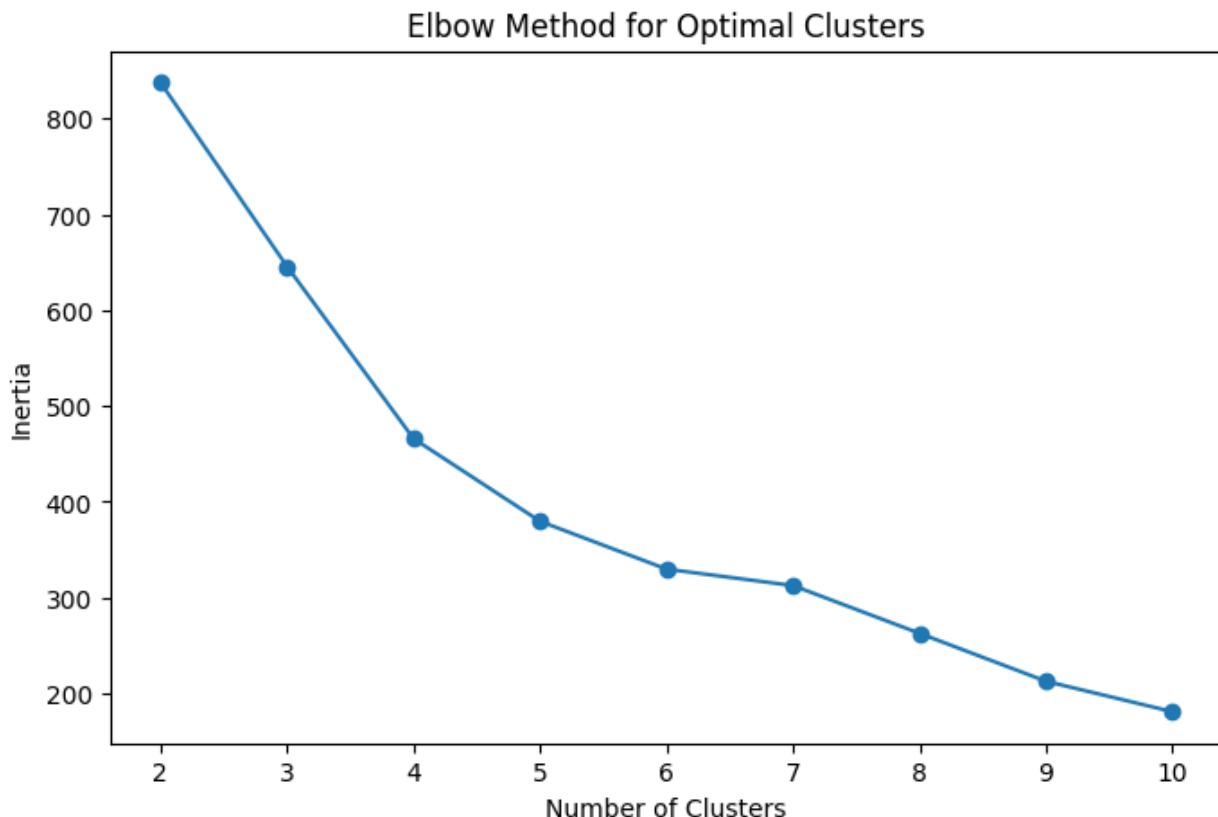


Task 3: Customer Segmentation / Clustering

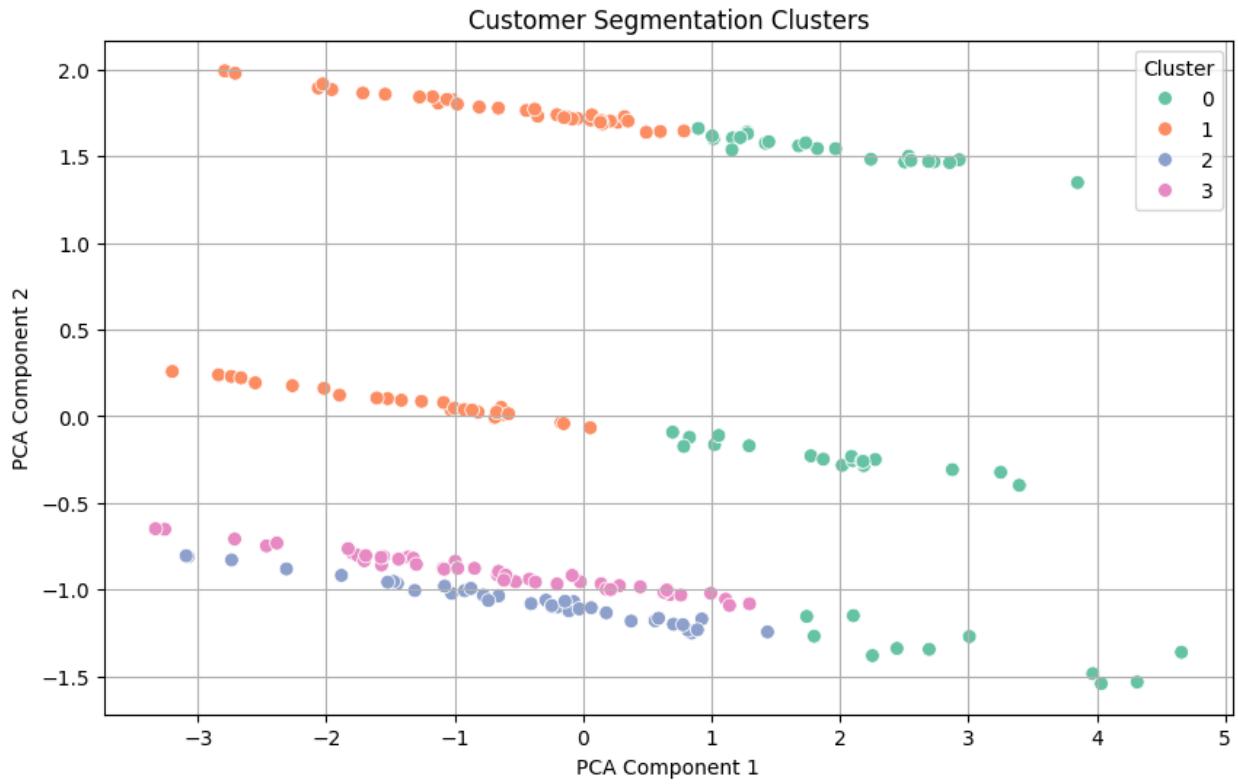
By: Abhishek Kumar

1. Optimal Number of Clusters (Elbow Method)



- The first chart demonstrates the Elbow Method, which helps identify the optimal number of clusters.
- The graph plots the number of clusters against the inertia (sum of squared distances of points to their nearest cluster center).
- A noticeable "elbow" occurs at **4 clusters**, indicating that using 4 clusters balances simplicity and accuracy for segmentation.

2. Customer Segmentation Results



- The second chart shows the customer segmentation after applying clustering.
- PCA was used to reduce the data to 2 dimensions for easy visualization.
- The clusters are represented by distinct colors (e.g., orange, green, purple, blue), each signifying a group of customers with similar profiles and transaction behaviors.
- **Cluster 0 (e.g., orange):** Likely represents a group of customers with similar purchasing habits or demographics.
- **Cluster 1 (e.g., green):** Another distinct customer group, which might differ in their transaction frequency or profile.
- **Cluster 2 and 3 (e.g., purple and blue):** Other unique segments showing varying characteristics from the first two clusters.
- The separation between clusters is relatively distinct, suggesting good segmentation.

3. DB Index (Davies-Bouldin Index):

- Though not directly shown in the charts, the Davies-Bouldin Index (DB Index) would be calculated during the process.
- A lower DB Index implies better-defined clusters. If the calculated DB Index is low, it confirms that the clusters are compact and well-separated.

4. Business Applications:

- These clusters can be used for targeted marketing campaigns, personalized product recommendations, and better customer service.
- For example, you could tailor promotions specifically for customers in Cluster 0 if they represent high-value customers.