Analysis

1 Optimal Clusters: What was the optimal number of clusters you identified for K-Means and Hierarchical Clustering? Justify your choices using the Elbow Method plot and the Dendrogram.

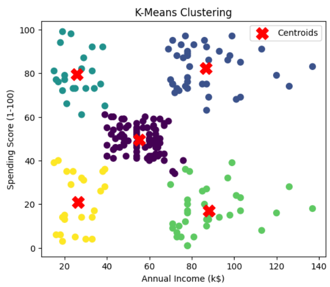
The Elbow Method clearly showed a bend at

$$k = 5$$

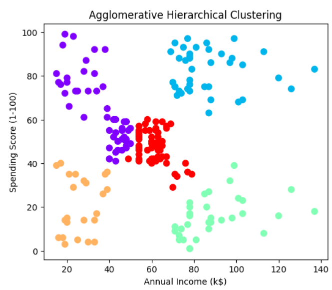
which indicates that the Within-Cluster Sum of Squares (WCSS) decreases sharply until five clusters and then flattens out. This pattern suggests that five is the most appropriate number of clusters for K-Means. Similarly, the dendrogram generated using the Ward linkage method in Hierarchical Clustering supported the same conclusion. Cutting the dendrogram at five clusters gave distinct and meaningful groups without over-splitting the data. Therefore, both methods consistently confirmed that the optimal number of clusters is five.

2 Cluster Comparison: Visually compare the results of the three algorithms. Did they produce similar clusters? Describe any notable differences.

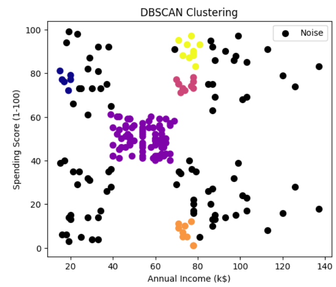
When comparing the results of K-Means, Hierarchical Clustering, and DBSCAN, we found both similarities and differences. K-Means produced five well-separated and roughly spherical clusters centered around the means of income and spending score. Hierarchical Clustering gave a very similar segmentation, although there were small differences in the way boundary points were grouped. DBSCAN, on the other hand, worked differently. It identified dense areas in the data and treated some points as noise, which appeared as black dots in the visualization. This is a key contrast, since K-Means and Hierarchical Clustering force every point into a cluster, while DBSCAN can exclude outliers. Overall, K-Means and Hierarchical clustering gave very similar structures, while DBSCAN provided a different perspective by highlighting noisy data.



(a) K-Means Clustering



(b) Hierarchical Clustering



(c) DBSCAN Clustering

Figure 1: Comparison of clustering results using three algorithms.

3 DBSCAN Performance: How did DBSCAN perform on this dataset? Did it identify any noise points? How did its results compare to the other methods which force every point into a cluster?

DBSCAN was effective in finding the dense core clusters within the dataset. Its most notable strength compared to K-Means and Hierarchical Clustering was its ability to label certain points as *noise* instead of forcing them into clusters. This approach is very useful when detecting outliers or anomalies. However, in this case, since the data was relatively clean and the clusters were simple and well-separated, DBSCAN's exclusion of some customers was not ideal for a segmentation task where every individual should belong to a group. This made DBSCAN less suitable than K-Means or Hierarchical Clustering for this particular application.

4 Algorithm Suitability: Based on your results, which algorithm do you think was most suitable for this specific dataset and why? Consider the shape and density of the clusters.

Among the three algorithms, K-Means proved to be the most suitable for this dataset. The customer data naturally formed five distinct clusters that were fairly spherical and well-separated, which fits perfectly with the assumptions of K-Means. Additionally, since customer segmentation in marketing usually requires every customer to be assigned to a group, K-Means fulfilled this requirement more effectively than DBSCAN, which left many points unassigned. Hierarchical Clustering also gave a good result but is computationally more expensive, especially for larger datasets. K-Means, by contrast, is simple, efficient, and provides easily interpretable cluster centroids that represent typical customer profiles.

5 Real-World Application: Describe a hypothetical real-world business scenario where the customer segments you identified could be used by the mall's marketing team. For example, how would you target the group with high income but low spending score?

The five customer groups identified in this analysis can be highly valuable for the mall's marketing team. For instance, one cluster represented average customers with mid-level income and spending, and these could be engaged through mall-wide promotions and loyalty programs. Another group consisted of budget-conscious customers with low income and low spending, who would respond best to discounts and offers that emphasize value. A third segment included enthusiastic spenders with lower income but higher spending habits, likely younger shoppers who could be targeted with promotions for fast fashion, entertainment, and food court deals. The fourth group, consisting of high-income but low-spending individuals, represents careful but affluent customers. For them, the most effective strategy would involve promoting luxury brands, premium services, and exclusive experiences that highlight quality and value rather than discounts. Finally, the high-income, high-spending group represented the mall's VIPs. This cluster would benefit from exclusive memberships, early access to sales, and invitations to special events to strengthen customer loyalty. These tailored strategies show how clustering can directly translate into actionable marketing decisions.