Customer Segmentation Analysis

Cluster Centers

Table

Cluster	Age	Income	Spend Score
1	42	$79,640	15.4
2	53	$55,118	49.5
3	25	$24,950	81
4	45	$26,304	20.9
5	33	$79,633	81
6	33	$86,538	82.1

ANOVA Results

- **Age**: Significant effect on spending scores ($p < 0.001$).
- **Income**: No significant effect on spending scores ($p = 0.523$).
- **Spend Score**: Significant effect on spending scores ($p < 0.001$).

Cluster Descriptions

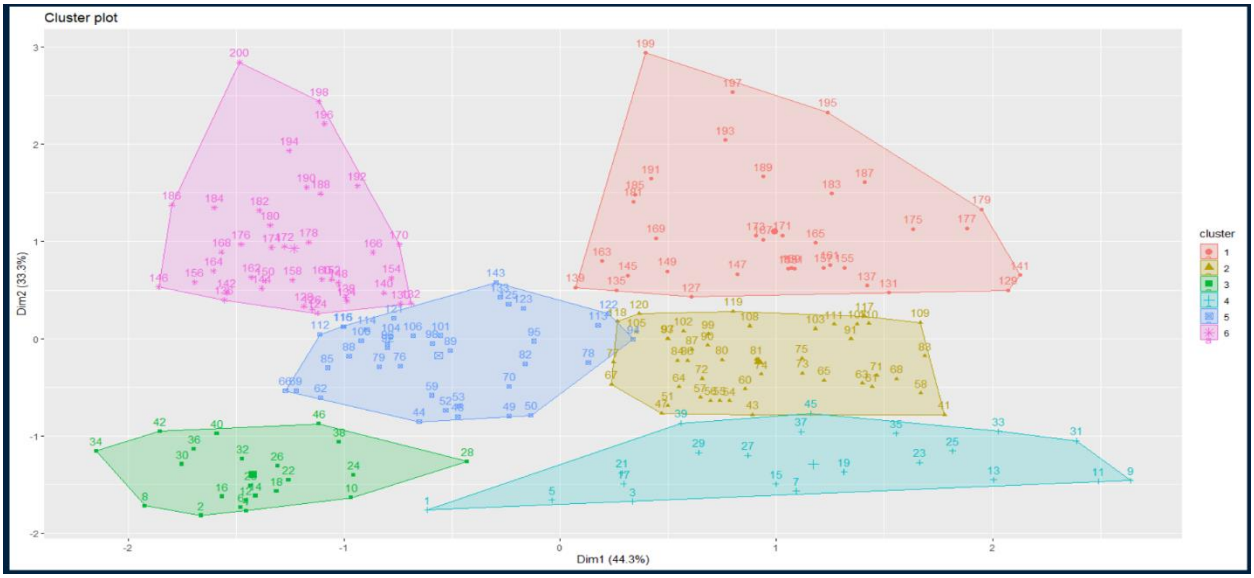
1. **Cluster 1**: Older customers (mean age 42) with high incomes but low spending scores. Likely frugal or conservative spenders.
2. **Cluster 2**: Older customers (mean age 53) with moderate incomes and spending scores. Representing typical spending behavior.
3. **Cluster 3**: Younger customers (mean age 25) with low incomes but very high spending scores. Probably students or young adults.
4. **Cluster 4**: Older customers (mean age 45) with lower incomes and lower spending scores. Potentially retirees or those on fixed incomes.

5. **Cluster 5**: Middle-aged (mean age 33) with high incomes and very high spending scores. Young professionals with disposable income.
6. **Cluster 6**: Middle-aged (mean age 33) with high incomes and high spending scores. Affluent households with discretionary spending.

Explanation of the Code:

The code performs both hierarchical and k-means clustering on the Mall_Customers dataset, using the variables Age, Income, and Spend_Score. It utilizes the Hubert index and the D index to determine the optimal number of clusters, which is found to be 6 based on the majority rule.

The code then profiles the final clusters from the k-means solution, using ANOVA to identify the variables that significantly differentiate the clusters. This allows for a deeper understanding of the characteristics of each cluster.



Descriptive Summary of Clusters:

Cluster 1:

These customers are middle-aged (mean age 41.69)

They have moderate income (mean $88,230)

They have a relatively low spending score (mean 17.29)

They appear to be more conservative spenders

Cluster 2:

These customers are older (mean age 56.16)

They have moderate income (mean $53,380)

They have a moderate spending score (mean 49.09)

They seem to have a balanced approach to spending

Cluster 3:

These customers are younger (mean age 25.27)

They have low income (mean $25,730)

They have a high spending score (mean 79.36)

They likely have an active, high-spending lifestyle

Cluster 4:

These customers are middle-aged (mean age 44.14)

They have low income (mean $25,140)

They have a moderate spending score (mean 19.52)

They appear to be cautious spenders due to limited income

Cluster 5:

These customers are young (mean age 27.00)

They have moderate income (mean $56,660)

They have a moderate spending score (mean 49.13)

They seem to have a balanced approach to spending and savings

Cluster 6:

These customers are middle-aged (mean age 32.69)

They have high income (mean $86,540)

They have a high spending score (mean 82.13)

They likely enjoy a more luxurious lifestyle due to their higher income