**Customer Segmentation Using Agglomerative Hierarchical Clustering**

**Name:** Fatima Muhammed  
**Reg No.:** CST/19/COM/00318

**1. Introduction** The dataset used in this study consists of 18 behavioral features representing customer spending patterns and payment behaviors. The objective of this project is to segment customers into meaningful clusters using Agglomerative Hierarchical Clustering. This segmentation can help identify different customer groups for personalized marketing strategies.

**2. Data Preprocessing** Before applying clustering algorithms, the dataset underwent thorough preprocessing:

* Missing Value Imputation: Missing values in the "MINIMUM\_PAYMENTS" and "CREDIT\_LIMIT" columns were replaced with their respective medians.
* Column Selection: Selected the best features for clustering: "BALANCE", "PURCHASES", "CASH\_ADVANCE", "CREDIT\_LIMIT", "PAYMENTS", and "MINIMUM\_PAYMENTS".
* Normalization: Scaled the selected features to ensure each feature contributes equally to the clustering process.
* ID Column Removal: Removed the "CUST\_ID" column as it does not provide meaningful information for clustering.

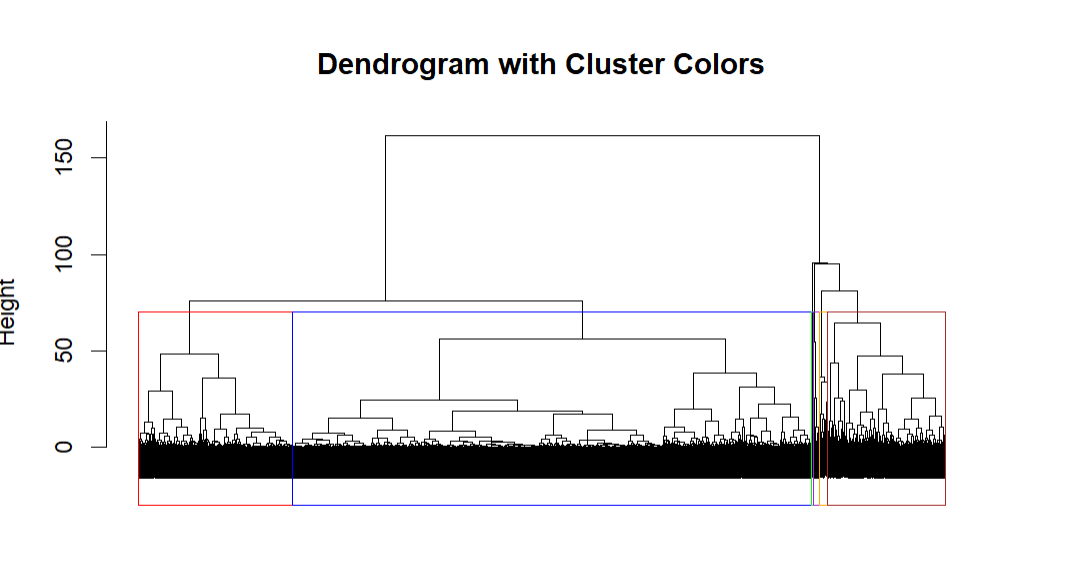
**3. Exploratory Data Analysis (EDA)** EDA was conducted to understand the distribution and relationships between features:

* Heatmap: A heatmap of the correlation matrix revealed strong correlations between some features, indicating potential redundancy.
* Pair Plots: Pair plots were generated to visualize relationships between key features across clusters, providing insights into customer behavior patterns.

**4. Clustering Process**

* PCA: Principal Component Analysis (PCA) was applied to reduce dimensionality and visualize clusters.
* Agglomerative Hierarchical Clustering: Applied Agglomerative Clustering with Ward’s linkage method, experimenting with 5 and 6 clusters.
* Cluster Assignment: Each customer was assigned to a cluster, and the results were added to the dataset.

**5. Generating Dendrogram** A dendrogram was generated to visualize the hierarchical clustering process. Cluster colors were added to better interpret the grouping of customers.



**6. Results and Insights**

* The clustering revealed distinct customer groups with varying spending habits and payment behaviors.
* Customers with high "BALANCE" and "CASH\_ADVANCE" formed separate clusters, suggesting they rely heavily on credit.
* Some clusters had low "MINIMUM\_PAYMENTS" relative to "CREDIT\_LIMIT", indicating a need for targeted financial education or personalized repayment plans.

**7. Marketing Strategy** Based on the clustering results, the following marketing strategies are recommended:

* High Spenders: Offer premium rewards or loyalty programs.
* Cash Advance Users: Introduce lower-interest cash advance options or personalized repayment plans.
* Low Payment Customers: Send reminders for minimum payments and provide financial literacy content.
* Balanced Users: Promote additional products such as savings plans or investment opportunities.

**8. Conclusion** The implementation of Agglomerative Hierarchical Clustering successfully segmented customers into meaningful groups. This segmentation offers valuable insights that can drive personalized marketing strategies and improve customer engagement.