■ Customer Segmentation Report

1. Objective

The goal of this project was to segment customers into distinct groups using Principal Component Analysis (PCA) for dimensionality reduction and KMeans clustering. This helps identify customer profiles and supports targeted marketing strategies.

2. Data Overview

Dataset included demographic and transaction-related features. Key features: Age, Education, Family Size, Children, Living Status, Kidhome, Teenhome, Is_Parent, and Total Spend.

3. Methodology

1. Preprocessing: Cleaned data, encoded categorical variables, scaled numerical features. 2. PCA: Reduced dimensionality, retaining ~85% variance. 3. KMeans: Optimal number of clusters found to be 4 using Elbow Method and Silhouette Score.

4. Results & Findings

Cluster 0: High Income, High Spend (Luxury-oriented). Cluster 1: Young Families, Budget-Conscious (bulk buyers). Cluster 2: Students/Young Adults (low income & spend). Cluster 3: Middle-Class Professionals (average spenders). Cluster Distribution: C0=18%, C1=32%, C2=25%, C3=25%.

5. Business Insights

Cluster 0 → Premium products & loyalty programs. Cluster 1 → Family bundles & seasonal discounts. Cluster 2 → Student discounts, entry-level offers. Cluster 3 → Upsell mid-range products.

6. Conclusion

PCA + KMeans successfully segmented customers into 4 groups. These insights support data-driven marketing, product positioning, and customer retention strategies.

Cluster	Description	Percentage
0	High Income, High Spend	18%
1	Young Families, Budget-Conscious	32%
2	Students/Young Adults	25%
3	Middle-Class Professionals	25%