# Customer Segmentation Analysis Report

## How We Did the Analysis

We used K-means clustering on dataset of 199 customers based on their shopping behaviour. Here are the features we looked at:

- **Transaction-based:** Total Spend, Average Transaction Value, Frequency of purchases
- **Product-based:** Number of unique products bought; total quantity purchased
- **Time-related:** How recent their last purchase was and how long they've been a customer
- **Location:** Customer regions (converted into numerical format)

## Steps We Followed

1. Calculated detailed customer metrics.
2. Standardized the numbers to keep things fair.
3. Converted regions into a format the model could understand.
4. Used PCA (a method to simplify data) to help visualize the results.

# How Many Clusters Did We Use?

We tested different options and found 3 clusters worked best using:

1. **Davies-Bouldin Index (DB Index):** The best score (1.258) showed 3 clusters were optimal.
2. **Elbow Method:** Confirmed this choice with visual evidence.

# What Are the Clusters Like?

**Cluster 0: High-Value, Active Shoppers (77 people, 38.7%)**

- Spend the most and shop the most often.
- Buy lots of items and a wide variety of products.
- **Who they are:** Regular and loyal customers who spend big.

**Cluster 1: Low-Engagement Customers (13 people, 6.5%)**

- Spend and shop much less than others.
- Engage with fewer products and don't buy often.
- **Who they are:** Inactive customers who rarely shop.

**Cluster 2: Moderate Customers (109 people, 54.8%)**

- Spend a little less than average and shop occasionally.
- Buy a fair number of products but not in large quantities.

- **Who they are:** Casual shoppers with steady but moderate habits.

## What the Visuals Show

- **PCA Visualization:** Clearly shows 3 Clusters with distinct behaviors.
- **Spending vs. Frequency Plot:** Highlights differences in how often and how much people shop.
- **Transaction Value vs. Recency Plot:** Shows patterns in when customers shop and how much they spend.

## What This Means for the Business

1. **Better Marketing Strategies**
    - **Cluster 0:** Focus on keeping these customers loyal with exclusive deals and premium products.
    - **Cluster 1:** Run campaigns to re-engage them and spark their interest.
    - **Cluster 2:** Encourage them to shop more often or spend more per transaction.
2. **Smarter Resource Allocation**
    - Invest in retaining Cluster 0 customers since they bring in the most revenue.
    - Create strategies to move Cluster 2 customers into Cluster 0.
3. **Improved Customer Experience**
    - Personalize offers and messages based on each Cluster's habits.
    - Design specific programs to match what each Cluster needs.

## Key Metrics

- The DB Index score of 1.258 shows the model performed well.
- The clusters are well-separated, as seen in the PCA visualization.
- The Clusters are fairly balanced, except Cluster 1, which is smaller.