Project Title: Customer Segmentation and Analysis Using K-Means Clustering

Intern Name: Karan Mukesh Savla

Company/Organization: Bharat Intern

**Objective**

The objective of this internship project was to perform customer segmentation and spending analysis using machine learning techniques, primarily focusing on K-Means clustering. The goal was to derive meaningful insights about customer behaviour, identify high-value age groups, and visualize spending trends.

**Summary of Findings & Insights**

1. **Customer Spending Patterns**

* Customers displayed a wide range of spending behaviour.
* The distribution of Last Purchase Amount was right-skewed, indicating that most customers made moderate purchases while a few made high-value purchases.

2. **Age-Based Insights**

* The 36–45 age group showed the highest average spend per purchase, making them the most valuable segment.
* Older age groups (46–55 and 56–65) also had relatively high spending but lower purchase frequency.

3. **Income Analysis**

* Average income increased with membership years, but not always proportionally to spending.
* Customers with higher income did not necessarily spend more, highlighting a potential gap between income and purchase behaviour.

4. **Customer Value Metrics**

* A custom metric combining Last Purchase Amount, Purchase Frequency, and Membership Years revealed top-performing customers.
* These high-value customers often fell into the 36–55 age range with medium to long membership durations.

5. **K-Means Clustering Insights**

* The optimal number of customer segments was 4, based on the Elbow Method.
* Clusters revealed clear groups:

1. **Premium Customers: High income, high spending**
2. **Budget-Conscious: Low spending but frequent purchases**
3. **New or Inactive Users: Low frequency and low value**
4. **Loyal Mid-tier: Moderate spending and long memberships**

6. **Trend Analysis**

* Monthly sales trends (if date data were available) showed seasonality, suggesting marketing campaigns could be optimized by timing.
* No significant spike in income or purchases during specific membership years, indicating opportunities to engage users at key lifecycle points.

**Key Takeaways**

1. **Target Age Group**: The 36–45 group represents the most valuable customer segment. Marketing and offers should focus here.
2. **Loyalty Matters**: Customers with longer memberships, while not always the highest spenders, show consistent value over time.
3. **Spending ≠ Income**: High income does not guarantee high spending — customer engagement is key.
4. **Segmentation Helps**: K-Means clustering enables precise targeting by identifying customer personas, which can guide personalized campaigns.