Customer Segmentation Report

1. Introduction

The objective of this project is to segment customers of an e-commerce platform based on their behavior. The dataset includes customer interactions, purchase patterns, and browsing activities. By identifying distinct customer groups, the platform can implement targeted marketing strategies to enhance customer experience and boost sales.

2. Dataset Overview

The dataset consists of the following key attributes:

- **customer_id**: Unique identifier for each customer.
- **total_purchases**: Total number of purchases made by the customer.
- **avg_cart_value**: Average value of items in the customer's cart.
- **total_time_spent**: Total time spent on the platform (in minutes).
- **product_click**: Number of products viewed by the customer.
- **discount_count**: Number of times the customer used a discount code.

3. Exploratory Data Analysis (EDA)

- **Missing Values**: The dataset was checked for missing values, and no missing values were detected.
- **Feature Distributions**: Most features displayed a right-skewed distribution, indicating a majority of customers exhibit lower values for purchases and engagement metrics.
- **Correlations**: A correlation matrix was generated to examine relationships between features. No strong correlations were observed, confirming that each feature provides unique information.

4. Data Preprocessing

- **Handling Missing Values**: Since no missing values were found, no imputation was required.
- **Standardization**: StandardScaler was used to scale all numerical features to ensure uniformity across the dataset.

5. Model Selection

- **KMeans Clustering**: The KMeans algorithm was selected due to its efficiency and effectiveness for customer segmentation.
- **Elbow Method**: The optimal number of clusters was determined using the Elbow Method. The plot showed a clear elbow at **k=3**, indicating that three clusters provide the best segmentation.

6. Clustering Results

Customers were grouped into three distinct clusters:

Bargain Hunters

- High total purchases and frequent use of discounts.
- Low average cart value.
- Moderate time spent and product views.

High Spenders

- Moderate number of purchases.
- High average cart value.
- Low usage of discounts.

Window Shoppers

- Low number of purchases.
- High engagement in browsing (high time spent and product views).
- Low use of discounts.

7. Visualizations

- **Pairplot**: A pairplot was created to visualize feature relationships, color-coded by cluster assignment.
- **Boxplot**: Boxplots were generated to compare feature distributions across clusters, highlighting key behavioral differences.

8. Conclusion

The clustering analysis effectively identified three key customer segments:

- **Bargain Hunters**: Customers who frequently purchase low-value items and utilize discounts.
- **High Spenders**: Customers who make high-value purchases and are less influenced by discounts.
- **Window Shoppers**: Customers who spend a lot of time browsing but make very few purchases.

These insights can be leveraged to develop tailored marketing strategies, such as:

- Personalized discount offers for **Bargain Hunters**.
- Exclusive premium product recommendations for **High Spenders**.
- Targeted promotions to convert **Window Shoppers** into active buyers.

By using customer segmentation, the e-commerce platform can enhance customer satisfaction, improve retention, and optimize revenue generation.