# Customer Segmentation, Purchase Pattern & Churn Risk Analysis

**Organization:** Alfido Tech
**Domain:** Data Analytics
**Prepared by:** *Sanyam Sharma*
**Tools Used:** Python, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn
**Dataset:** E-commerce Customer Transaction Data

---

## 1.Executive Summary

This report presents a comprehensive analysis of customer transaction data to understand customer behavior, identify meaningful customer segments, analyze purchase and retention patterns, and assess churn risk. Using Exploratory Data Analysis (EDA), RFM (Recency, Frequency, Monetary) analysis, and K-Means clustering, customers were segmented into distinct behavioral groups.

The insights derived from this analysis help Alfido Tech design targeted marketing strategies, improve customer retention, reduce churn, and optimize overall business performance through data-driven decision-making.

---

## 2.Project Objectives

The main objectives of this project are:

- To clean and preprocess customer transaction data
- To engineer meaningful behavioral features using RFM analysis
- To segment customers based on purchasing behavior
- To analyze purchase patterns and retention trends
- To identify churn-prone customer groups
- To provide actionable business recommendations tailored to Alfido Tech

---

## 3.Dataset Overview

The dataset contains customer-level transactional and demographic information, including:

- Customer ID
- Purchase Date
- Product Category
- Quantity and Total Purchase Amount

- Payment Method
- Customer Age and Gender
- Churn indicator (1 = churned, 0 = retained)

Each row represents a transaction made by a customer.

---

# 4.Data Cleaning & Preprocessing

The following steps were performed to ensure data quality:

- Removed duplicate records
- Converted date columns into datetime format
- Verified missing values and ensured consistency
- Ensured numeric columns were in the correct data type

These steps ensured reliable and accurate analysis.

---

# 5.Exploratory Data Analysis (EDA)

EDA was conducted to understand customer behavior and purchasing trends:

### Key Observations:

- A small group of customers contributes a large share of total revenue
- Certain product categories dominate overall sales
- Payment methods show clear customer preferences
- Repeat customers generate significantly higher revenue than one-time buyers

EDA provided a strong foundation for further feature engineering and segmentation.

---

# 6.Feature Engineering: RFM Analysis

RFM analysis was used to quantify customer value:

- **Recency:** Number of days since the customer’s last purchase
- **Frequency:** Number of purchases made by the customer
- **Monetary:** Total amount spent by the customer

Customers were grouped by Customer ID and RFM metrics were calculated to capture purchasing behavior in a structured manner.

---

# 7.RFM Scoring

Each RFM metric was divided into quartiles and assigned scores:

- Higher scores indicate better customer value
- Recency scoring was reversed (lower recency = higher score)

An overall RFM score was created to represent customer engagement and value.

---

# 8.Customer Segmentation Using K-Means Clustering

To create data-driven customer segments, K-Means clustering was applied to scaled RFM values. Customers were grouped into **four distinct clusters**.

## Identified Customer Segments:

1. **High-Value Loyal Customers**
    - Low recency, high frequency, high spending
    - Core revenue contributors
2. **Regular Customers**
    - Moderate engagement and spending
    - Consistent but not top spenders
3. **Occasional / At-Risk Customers**
    - Low frequency and increasing inactivity
    - High risk of churn
4. **Churned / Lost Customers**
    - Very high recency and low engagement
    - Minimal contribution to revenue

---

# 9.Purchase Pattern Analysis

Purchase patterns were analyzed across different dimensions:

- **Product Categories:** Certain categories significantly outperform others
- **Payment Methods:** Customers show strong preference for specific payment modes
- **Age Groups:** Spending behavior varies across age ranges

Age was grouped into meaningful ranges to avoid overcrowded visualizations and to improve interpretability.

---

# 10.Retention Analysis

Retention behavior was studied using:

- Monthly active customers
- Repeat vs one-time customers
- Average purchase frequency by customer segment

### Key Retention Insights:

- Repeat customers are fewer in number but generate higher lifetime value
- High-value and regular customers show strong retention behavior
- At-risk and churned segments show clear signs of disengagement

---

# 11.Churn Risk Analysis

The churn column was used to identify behavioral differences between churned and retained customers.

### Churn Behavior Insights:

- Churned customers have:
    - Higher recency (longer inactivity)
    - Lower frequency
    - Lower monetary value
- RFM metrics strongly correlate with churn risk

Churn rates were significantly higher among **Occasional / At-Risk** and **Churned** customer segments.

---

# 12.Business Insights Summary

- Customer behavior varies significantly across segments
- High recency is a strong indicator of churn
- Loyal customers drive most of the revenue
- Not all customers should be targeted with the same marketing strategy

---

# 13.Actionable Business Recommendations (For Alfido Tech)

### 1. Protect High-Value Customers

- Introduce VIP programs, exclusive deals, and early access

- Focus on personalized communication

### 2. Strengthen Loyalty for Regular Customers

- Implement reward points and milestone-based incentives
- Encourage upselling and cross-selling

### 3. Re-Engage At-Risk Customers

- Launch time-bound discount campaigns
- Send reminder emails and personalized offers

### 4. Optimize Marketing Spend on Churned Customers

- Run limited win-back campaigns
- Avoid long-term paid marketing for permanently churned users

### 5. Implement Predictive Churn Monitoring

- Track inactivity thresholds
- Trigger automated alerts when churn risk increases

---

# 14. Conclusion

This project demonstrates how customer transaction data can be transformed into actionable business insights using data analytics techniques. By applying RFM analysis, clustering, and churn analysis, Alfido Tech can better understand customer behavior, improve retention, reduce churn, and increase customer lifetime value.

The methodology and insights from this analysis can be scaled and integrated into real-time customer analytics systems for continuous business improvement.

---

# 15.Deliverables

- Jupyter Notebooks (EDA, RFM, Segmentation, Churn Analysis)
- PDF/DOC Reports with executive summaries and visualizations
- GitHub repository with clean code and documentation