Case Study: Enhancing Customer Strategy at NovaBank with RFM Segmentation

About the Project

This project focuses on improving customer relationship strategies for **NovaBank,** a growing digital-first retail bank. While NovaBank had a wide customer base and increasing transaction volume, it lacked clarity on how different customers were engaging with their services and more importantly, which ones were truly valuable.

To address this, we applied **RFM (Recency, Frequency, Monetary)** analysis using Python, alongside unsupervised machine learning techniques, to segment customers based on their transaction behavior. The aim was to help NovaBank build smarter, more personalized marketing and retention strategies.

Business Challenge

NovaBank faced three core issues:

- **Customer Retention Risk:** Many customers were becoming inactive, and churn rates were starting to climb. Without knowing who was likely to leave, retention was reactive instead of proactive.

- **Lack of Personalization:** Marketing messages were generic. The bank didn't have enough behavioral segmentation to target users based on value or engagement level.

- **Inefficient Resource Allocation:** Marketing and support resources were spread evenly, without insight into which customers were worth the most time and investment.

Project Objectives

This project aimed to:

- Analyze customer transaction data to compute RFM scores.

- Segment customers into meaningful groups using clustering algorithms.

- Identify high-value segments for targeted campaigns.

- Help NovaBank retain its best customers and engage those at risk of churn.

Data Description

The project used anonymized transaction records from NovaBank's retail banking unit.

Key fields included:

- TransactionID: Unique transaction reference

- CustomerID: Unique customer identifier

- TransactionDate: Date of each transaction

- TransactionAmount: Monetary value of transaction

- CustomerDOB, Gender, Location: Demographic details

- CustAccountBalance: Current balance info

- TransactionTime: Unix timestamp

Tech Stack

- **Python**: Core programming language

- **Libraries Used**:

    o Pandas, Numpy – Data cleaning and manipulation

    o Matplotlib, Seaborn – Visualization

    o Scikit-learn – Clustering algorithms (KMeans)

Project Approach

1. **Exploratory Data Analysis (EDA)**
   Cleaned and explored transaction data to identify missing values, outliers, and general behavior patterns.

2. **RFM Scoring**
   For each customer:

    o **Recency:** How recently they transacted

    o **Frequency:** How often they transacted

    o **Monetary:** How much they spent
    Each metric was ranked and scored.

3. **Segmentation with Clustering**
   Applied unsupervised learning (KMeans) to group customers into clusters — from top-tier loyal users to those at risk of churn.

4. **Profile Interpretation**
   Each cluster was profiled to understand typical traits, demographics, and behavior.

5. **Visualization & Insight Generation**
   Generated plots and charts to communicate how segments differ and how they can be acted upon.

Key Learning Outcomes

- How to engineer and score RFM metrics

- Segmenting customers using clustering techniques

- Extracting actionable insights from transaction patterns

- The value of personalized marketing based on behavioral data