

Specialization: Data Science

Business Focus: Financial Services

Tool: Python

Optimizing Retail Banking Strategies Through RFM-Based Customer Segmentation

Project Learning Opportunities

Through this project, you have the opportunity to gain understanding and hands on experience on how to carryout RFM analysis(recency, frequency and monetary) using python, how to assign RFM score and proceed by implementing and evaluating unsupervised learning technique for customer profiling on RFM data

Learning Skills:

- Statistical Analysis
- RFM analysis with python
- Assigning RFM score
- Customer segmentation
- Unsupervised machine learning with python

Business Introduction

Imagine a bank that knows you as well as your favorite online store knows your shopping preferences. Welcome to BankTrust, where we're revolutionizing retail banking by taking inspiration from the world of e-commerce.

At BankTrust, we've been serving our customers for over four decades, and our commitment to exceptional service has been our hallmark. However, in today's rapidly evolving banking landscape, we face a unique challenge – how can we provide the same level of personalized service that you enjoy when shopping online?

In this case study, we'll explore how BankTrust is leveraging cutting-edge data analytics and customer segmentation techniques, inspired by e-commerce strategies, to enhance your banking experience. Join us on this journey as we uncover the secrets to understanding your banking needs better, offering tailored solutions, and ultimately making your financial life more convenient and rewarding.

Business Problem

- **Customer Churn:** Customer retention is a significant challenge in the banking industry. Retaining existing customers is often more cost-effective than acquiring new ones. High customer churn rates can lead to a loss of deposits and revenue.
- **Personalization:** BankTrust must provide personalized experiences to meet the diverse needs of their customers. Failing to do so can result in customers seeking alternative banks that offer tailored services and products.
- **Marketing Efficiency:** Inefficient marketing campaigns can lead to high customer acquisition costs and low return on investment. BankTrust need to optimize their marketing strategies to target the right customers with relevant offers.

Why RFM analysis?

RFM stands for **R**ecency, **F**requency and **M**onetary. It is a data-driven technique used to segment customers based on their past behaviours.

- RFM analysis can help BankTrust segment their customer base effectively. By categorizing customers based on their transaction **Recency** (how recently they engaged with the bank), **Frequency** (how often they conduct transactions), and **Monetary** value (the total value of their transactions), BankTrust can identify different customer segments. This segmentation allows for more targeted marketing and service strategies.
- RFM analysis can help BankTrust identify customers who have become less active (lower frequency) or have not engaged with the bank recently (lower recency). Recognizing these signs early will enable BankTrust to implement retention strategies to prevent customer churn, such as offering personalized promotions or contacting customers with special offers.
- RFM analysis helps BankTrust optimize their marketing efforts. Instead of using a one-size-fits-all approach, BankTrust can tailor their marketing campaigns to target specific customer segments with offers that are most likely to resonate. This reduces marketing costs and increases the return on investment (ROI) for marketing campaigns.

The primary objective of this project is to enhance the banking industry's customer relationship management by implementing an RFM-based customer segmentation approach.

This entails leveraging transaction history, recency, frequency, and monetary value as critical customer behavior attributes to categorize clients into meaningful segments. Furthermore, the project aims to employ advanced unsupervised learning techniques to uncover hidden patterns and structures within the customer data, ultimately yielding a comprehensive set of distinct customer clusters.

Through the application of RFM analysis and unsupervised learning, the project aims to unlock valuable insights that will empower the bank to offer tailored solutions, strengthen customer relationships, and drive sustainable growth in a competitive market environment.

Aim of Project

PROJECT ENHANCEMENT

To enhance the RFM-based customer segmentation project, a **Streamlit dashboard** will be developed as an interactive, user-friendly tool. It will enable BankTrust teams to **explore customer segments, visualize RFM metrics, and simulate marketing strategies**—all without technical expertise.

This real-time dashboard empowers data-driven decisions to boost retention, personalize offers, and optimize campaigns. Users can filter segments, view individual profiles, and test “what-if” scenarios (e.g., increasing engagement for low-frequency customers).

Deployed via **Streamlit Cloud**, the dashboard ensures broad adoption across marketing, CRM, and executive teams—turning RFM insights into measurable business outcomes.

Data Description

- **TransactionID:** A unique identifier for each individual transaction, allowing for precise tracking and reference.
- **CustomerID:** A distinct identifier assigned to each customer, enabling the association of transactions with specific individuals.
- **CustomerDOB:** The date of birth of the customer, providing insights into their age and potentially influencing financial behavior.
- **CustGender:** The gender of the customer.
- **CustLocation:** The location or geographical information associated with the customer, which can offer insights into regional trends and behavior.
- **CustAccountBalance:** The current balance in the customer's account, serving as a critical financial indicator for analysis.
- **TransactionDate:** The date on which a transaction occurred.
- **TransactionTime:** A Unix timestamp representing the precise time of a transaction.
- **TransactionAmount:** The monetary value of the transaction, indicating the financial magnitude of each transaction and facilitating analysis of spending or deposit habits.

Tech Stack

1. Programming language – Python

Libraries

- Numpy: For performing mathematical operations over data
- Pandas: For Data Analysis and Manipulation
- Matplotlib.pyplot: For Data Visualization
- Seaborn: For Data Visualization
- Scikit-learn: For Machine Learning



Data Science Project Scope

Exploratory Data Analysis

the data is thoroughly explored and analysed to gain insights and understand its characteristics. This involves statistical analysis, data visualization, and other exploratory techniques to identify patterns, correlations, anomalies, and potential issues in the data.



Feature Engineering

Feature engineering is the process of creating or selecting relevant features (input variables) from the available data that will be used for modeling. It may involve feature extraction, transformation, scaling, or the creation of new features to improve the predictive power of the models.



Model Development

Various modeling techniques are applied to the prepared data in this stage to build predictive or descriptive models. This can include machine learning algorithms, statistical models, or other analytical approaches. The models are trained, validated, and fine-tuned to optimize their performance.



Model Evaluation and Selection

Once the models are developed, they are evaluated using appropriate evaluation metrics and validation techniques. This involves assessing their performance, generalizability, and robustness. The best-performing model(s) are selected for further deployment or refinement.