**Final Report**

**Customer Segmentation & Product Recommendation System for GTBank**

**1. Introduction**

GTBank aims to enhance its customer experience by offering personalized banking services. The bank currently lacks an effective system for segmenting customers and recommending suitable products. This project addresses this issue by analyzing customer data, segmenting customers using clustering techniques, and developing a product recommendation system.

**2. Data Cleaning and Exploration**

2.1 Data Cleaning Process

* Handling Missing Values: Used statistical imputation for missing numerical data and mode replacement for categorical values.
* Outlier Detection: Identified outliers using the IQR method and applied capping techniques.
* Data Formatting: Standardized date formats, removed duplicate entries, and ensured consistency in categorical data.

2.2 Exploratory Data Analysis (EDA)

* Customer Demographics: Age distribution, gender ratio, and geographical presence.
* Transaction Patterns: Frequency, amount, and types of transactions (withdrawals, deposits, loan payments, etc.).
* Account Balances: Identified high-value, mid-range, and low-balance accounts.

**3. Customer Segmentation Analysis**

3.1 Clustering Technique Used

We employed K-Means Clustering due to its efficiency in handling large datasets and ease of interpretation. The optimal number of clusters was determined using the Elbow Method, which indicated that five clusters provided the best segmentation.

3.2 Customer Segments Identified

* Business Owners – High transaction frequency, large deposits, and withdrawals.
* High-Value Customers – Large account balances and investments in premium banking products.
* Low-Value Customers – Small balances and low transaction frequency.
* Loan Seekers – Customers with active loans or frequent loan applications.
* Regular Customers – Moderate balances and standard banking usage.

**4. Product Recommendation System**

4.1 Recommendation Approach

* Business Owners → Business Loans, Corporate Credit Cards.
* High-Value Customers → Premium Savings, Investment Plans.
* Low-Value Customers → Basic Savings Accounts, Cashback Debit Cards.
* Loan Seekers → Loan Offers, Debt Consolidation Plans.
* Regular Customers → Mobile Banking Services, Standard Savings Accounts.

4.2 Implementation

The recommendation system was implemented using Python (Pandas, Sklearn). The model assigns customers to their respective segments and suggests relevant banking products based on their profile.

**5. Challenges & Solutions**

5.1 Key Challenges

* Data Quality Issues: Missing values and inconsistencies.
* Optimal Clustering Selection: Choosing the right number of clusters.
* Interpreting Results for Business Use: Ensuring recommendations align with banking strategies.

5.2 Solutions Applied

* Preprocessed Data Thoroughly: Cleaned data to improve model accuracy.
* Used the Elbow Method for Clusters: Ensured a well-balanced segmentation approach.
* Mapped Recommendations to Business Goals: Validated results with domain experts.

**6. Business Impact & Insights**

* Increased Customer Engagement: Personalized services improve retention.
* Higher Revenue Potential: Tailored product recommendations boost sales.
* Better Customer Understanding: Helps in strategic decision-making.

**7. Conclusion & Future Work**

This project successfully implemented a data-driven approach to customer segmentation and product recommendations. Future enhancements include:

* Real-time Customer Segmentation using machine learning.
* Integration with GTBank’s Mobile App for automated recommendations.
* More Behavioral Data Analysis to improve segmentation accuracy.