**Bank Customer Churn Analysis - Project Documentation**

**1. Introduction**

**1.1 Overview**

The **Bank Customer Churn Analysis** project aims to analyze customer data to predict churn rates and identify key factors influencing customer retention in the banking sector. By leveraging **SQL** for data manipulation and analysis, the project provides actionable insights to inform marketing strategies and improve customer retention. The project focuses on understanding the characteristics of customers who churn versus those who remain, enabling the bank to develop targeted retention strategies.

**1.2** **Significance**

Customer churn is a critical issue in the banking industry, as retaining existing customers is more cost-effective than acquiring new ones. This project is significant because it helps the bank identify at-risk customers, understand the reasons behind churn, and implement strategies to improve customer loyalty. By analyzing customer demographics, product usage, and financial behaviour, the bank can reduce churn rates and enhance customer satisfaction.

**1.3 Scope**

The scope of this project includes:

* Analyzing customer data to identify patterns and trends related to churn.
* Predicting the likelihood of churn based on customer attributes.
* Providing insights into customer demographics, product usage, and financial behavior.
* Developing targeted marketing strategies to reduce churn and improve retention.

**1.4 Stakeholders**

* **Bank Executives**: To make strategic decisions based on churn analysis.
* **Marketing Teams**: To develop targeted retention campaigns.
* **Customer Service Teams**: To improve customer engagement and satisfaction.
* **Data Analysts**: To explore data and generate insights.

**2. Project Objectives**

**2.1 Problem Statements**

* High customer churn rates are impacting the bank’s revenue and customer base.
* Lack of understanding of the key factors driving customer churn.
* Inefficient customer retention strategies due to insufficient data-driven insights.

**2.2 Key Goals**

* **Churn Prediction**: Develop a model to predict which customers are likely to churn.
* **Customer Segmentation**: Identify characteristics of customers who churn versus those who remain.
* **Product Usage Analysis**: Understand how product usage impacts churn rates.
* **Demographic Insights**: Analyze the impact of age, gender, and country on churn rates.
* **Retention Strategies**: Develop targeted marketing strategies to reduce churn.

**3. Data Sources & Preparation**

**3.1 Description of Datasets**

The project utilizes the following datasets:

* **Customer Demographics**: Includes age, gender, country, and estimated salary.
* **Account Information**: Includes account balance, credit score, and tenure.
* **Product Usage**: Includes the number of products held by each customer.
* **Churn Data**: Includes information on whether a customer has churned or not.

**3.3 Data Sources**

* **Internal Databases**: Customer and account data extracted from the bank’s CRM and ERP systems.
* **External Data**: Market trends and competitor analysis data from third-party sources.

**3.4 Data Preparation**

* **Data Cleaning**: Removed duplicates, handled missing values, and corrected inconsistencies.
* **Data Transformation**: Aggregated data to create meaningful metrics such as churn rate, average balance, and product usage.
* **Data Integration**: Combined multiple datasets to create a unified data model for analysis.

**4. Methodology and Approach**

**4.1 Implementation**

The project was implemented using **SQL** for data manipulation and analysis. The implementation involved:

* **Data Querying**: Extracting relevant data from the database using SQL queries.
* **Data Aggregation**: Calculating key metrics such as churn rate, average balance, and product usage.
* **Data Analysis**: Identifying patterns and trends related to churn.

**4.2 Framework**

The project followed the **CRISP-DM (Cross-Industry Standard Process for Data Mining)** framework:

1. **Business Understanding**: Defined project objectives and stakeholder requirements.
2. **Data Understanding**: Explored and cleaned the datasets.
3. **Data Preparation**: Transformed and integrated data for analysis.
4. **Modelling**: Created data models and calculated key metrics.
5. **Evaluation**: Validated the analysis with stakeholders.
6. **Deployment**: Provided actionable insights to stakeholders.

**4.3 Methods Applied**

* **Descriptive Analytics**: Used to summarize historical customer data.
* **Diagnostic Analytics**: Applied to identify reasons behind churn.
* **Predictive Analytics**: Used to predict the likelihood of churn based on customer attributes.

**4.4 Assumptions and Limitations**

* **Assumptions**:
  + Data provided is accurate and up-to-date.
  + Customer behaviour patterns remain consistent over time.
* **Limitations**:
  + Limited by the quality and completeness of the data.
  + Predictive analytics may not account for unforeseen market changes.

**5. Insights & Analysis**

**5.1 Key Findings**

* **Churn Rate**: The overall churn rate is **20.37%**, with higher rates among older customers and those with higher balances.
* **Product Usage**: Customers holding two products show a significantly lower churn rate (**0.0758**), while those with four products experience a high churn rate (**1.00**).
* **Demographics**: Female customers exhibit a higher churn rate (**0.25**) compared to males (**0.16**), with the average age for females being slightly higher (**40.15**).
* **Country Comparison**: Germany has the highest churn rate, particularly among female customers (**0.37%**).
* **Balance Impact**: Customers with low balances exhibit lower churn rates (**0.19**) compared to those with high balances (**0.23**).

**5.2 Business Insights**

* **Targeted Marketing**: Focus on middle-aged females in France and Spain and young males in Germany for tailored marketing campaigns.
* **Product Bundling**: Encourage customers to hold multiple products by offering incentives or discounts for bundled services.
* **Customer Engagement**: Develop personalized engagement strategies for customers aged 40-50 to address their specific needs and concerns.

**6. Challenges & Solutions**

**6.1 Challenges**

* **Data Quality Issues**: Incomplete and inconsistent data.
* **Complex Data Integration**: Combining multiple datasets with different structures.
* **Stakeholder Alignment**: Ensuring all stakeholders agreed on the key metrics and visualizations.

**6.2 Solutions**

* **Data Cleaning**: Used SQL to clean and transform data.
* **Data Modelling**: Created a unified data model to integrate different datasets.
* **Stakeholder Collaboration**: Conducted regular meetings to align on project goals and deliverables.

**7. Conclusion & Future Enhancements**

**7.1 Conclusion**

The **Bank Customer Churn Analysis** project successfully identified key factors influencing customer churn and provided actionable insights to improve retention. By leveraging SQL for data analysis, the project enabled the bank to develop targeted marketing strategies and enhance customer engagement.

**7.2 Future Enhancements**

* **Advanced Predictive Analytics**: Incorporate machine learning models to predict churn more accurately.
* **Real-Time Data Integration**: Integrate real-time data feeds for up-to-the-minute insights.
* **Customer Segmentation**: Add more granular customer segmentation to better understand different customer groups.
* **Customer Feedback**: Conduct regular surveys to gather customer feedback and improve satisfaction.

This documentation provides a comprehensive overview of the **Bank Customer Churn Analysis** project, outlining its objectives, methodology, and key findings. The insights derived from this analysis will help the bank reduce churn rates and improve customer retention.