**Capstone Project Document**

# Project Title:

**Customer Retention and Churn Prediction Dashboard**

# 1. Introduction

Customer retention is vital for the sustainability and profitability of any business, especially for those operating in subscription-based models. This project leverages SQL, Python, and Power BI to analyze customer behavior, predict churn, and suggest data-driven strategies to improve customer retention.

# 2. Business Problem

The company has observed a growing number of customers leaving its services. There is no centralized system that analyzes churn causes or tracks customer behavior in real time. Business decisions are reactive rather than proactive.

**Problem Statement:** > How can we identify customers who are likely to churn and what actionable insights can be derived to reduce customer attrition?

# 3. Goals

| Goal | Description |
| --- | --- |
| G1 | Identify patterns in customer data that correlate with churn |
| G2 | Build a predictive churn model using Python |
| G3 | Use SQL for data extraction, transformation, and insights |
| G4 | Create an interactive dashboard in Power BI to visualize trends and KPIs |
| G5 | Propose business strategies to improve retention rates |

# 4. Tools & Technologies

| Tool | Purpose |
| --- | --- |
| SQL | Data extraction, joins, aggregation, filtering |
| Python | Data preprocessing, machine learning, model evaluation |
| Power BI | Visualization and reporting |

# 5. Dataset Description

The dataset simulates customer data for a telecom company. Key columns include: - CustomerID - Gender, Age, Geography - Tenure - MonthlyCharges, TotalCharges - Contract Type - Internet Service - Customer Support Calls - Churn (Yes/No)

# 6. Tasks Breakdown

## 6.1 SQL Tasks

* Create and populate customer database
* Extract customer segments based on churn and tenure
* Analyze monthly churn rates
* Aggregate metrics for visualization (e.g., average charges by churn)

## 6.2 Python Tasks

* Load and clean dataset
* Perform EDA (correlations, visual trends)
* Encode categorical variables
* Train/test ML models (e.g., Logistic Regression, Random Forest)
* Evaluate models (accuracy, recall, F1-score)

## 6.3 Power BI Tasks

* Design dashboard with:
  + Churn rate over time
  + Churn by contract type and support calls
  + Customer value segmentation
  + At-risk customer profiles from Python model
* Add filters for region, tenure, etc.

# 7. Insights & Solutions

## Identified Problems

* Month-to-month customers have higher churn rates
* Customers with multiple support calls are at greater risk
* Younger customers show lower engagement and higher churn

## Proposed Solutions

* Provide incentives for annual contracts
* Improve customer service quality and first-call resolution
* Launch loyalty programs targeting younger demographics

# 8. Deliverables

* SQL scripts (schema, queries, ETL)
* Python notebooks (EDA, model building, evaluation)
* Power BI dashboard (export or publish online)
* Final report summarizing findings, models, and business recommendations

# 9. Conclusion

This capstone project demonstrates the power of combining SQL, Python, and Power BI to create a comprehensive business intelligence solution. It not only helps predict customer churn but also drives actionable strategies to boost customer retention and overall business health.

*Prepared by: [Ritesh Salunkhe]*