

# Telecom Churn Analysis

## Introduction to Churn Analysis

In today's highly competitive market, retaining customers is essential for sustained success. Churn analysis is a vital approach to understanding and minimizing customer turnover. This process involves analyzing customer data to uncover patterns and identify the reasons behind their decision to leave. By leveraging advanced analytics and machine learning, businesses can predict customers likely to churn and determine the underlying factors influencing their choices. This insight enables organizations to take proactive measures to enhance customer satisfaction and strengthen loyalty.

## Project Goals

Create an entire ETL process in a database and a power BI dashboard to utilise the customer data and achieve below goals:

1. Analyse customer data at below levels:
  - a. Demographic
  - b. Geographic
  - c. Payment and account info
  - d. Services
2. Study Churner Profile & identify areas for implementing marketing campaigns
3. Identify a method to predict future churners

## Project Overview

- **Objective:** Analyze customer churn for a telecom firm and apply insights across industries.
- **Deliverables:**
  - ETL process in SQL Server.
  - Power BI dashboards for visualization.
  - Churn prediction using Random Forest in Python.

## Data Management with SQL Server

- **Setup:**
  - Use SQL Server Management Studio (SSMS) to create and manage databases.
  - **ETL Process:**
    - Import raw customer data into a staging table (stg\_Churn) via Import Wizard.
    - Perform data cleaning (handling nulls, datatype adjustments).
    - Populate a production table (prod\_Churn) with refined data.
  - **Views for Reporting:**
    - vw\_ChurnData: Focus on churned and stayed customers.
    - vw\_JoinData: Focus on new customers.

## Data Visualization with Power BI

- **Transformations:**
  - Create derived columns (e.g., Churn Status, Monthly Charge Range).
  - Map and sort groups like Age and Tenure.
  - Unpivot service-related columns for better analysis.
- **Measures:**
  - Calculate metrics such as Total Customers, Churn Rate, and New Joiners.
- **Dashboards:**
  - **Summary Page:** Includes demographic and geographic breakdowns, account info, and service usage.
  - **Churn Reason:** Highlights churn reasons and categories.
  - **Churn Prediction:** Visualize churn by internet type and other services.

## Predicting Churn with Machine Learning

- **Setup:**
  - Install Python libraries via Anaconda (e.g., pandas, numpy, scikit-learn).
  - Import data views into Excel for model training.

- **Pre-processing:**
  - Drop irrelevant columns and label-encode categorical data.
  - Split data into training and testing sets (80-20 split).
- **Model Training:**
  - Train a **Random Forest Classifier** on historical churn data.
  - Evaluate performance using a confusion matrix and classification report.
  - Analyze feature importance to identify key churn factors.
- **Prediction:**
  - Apply the trained model to predict churn for new customers (vw\_JoinData).

## Insights and Next Steps

- **Business Impact:**
  - Proactive churn mitigation via targeted campaigns.
  - Data-driven decision-making to improve customer retention.
- **Extensions:**
  - Automate the ETL pipeline.
  - Use advanced ML models for enhanced predictions.
  - Incorporate real-time data streaming for dynamic dashboards.

## Conclusion

This project demonstrates a comprehensive approach to customer churn analysis, leveraging SQL for data management, Power BI for visualization, and Python for predictive modeling. These techniques ensure actionable insights to enhance customer satisfaction and loyalty.