Telecom Churn Prediction Using ML

Team Members : Karunakaran R and Dharun RR

Email id : rkarunakaranraja@gmail.com

dharunrio7@gmail.com

Batch : 12^{th} May CDA

Project id : PRCL-0017

Project team id : PTID-CDA-SEP-25-722

Time period : 18^{th} Sep to 22^{th} Sep

Introduction

This project aims to predict customer churn in a telecom company using Machine Learning models. Churn prediction is critical for telecom operators as retaining existing customers is more cost-effective than acquiring new ones.

Key Highlights

- Data pre-processed with feature engineering.
- Models tested: Logistic Regression, Random Forest.
- Best model achieved ~85% accuracy.
- Deployed insights into Power BI dashboard for business monitoring.

Business Problem

- Customer churn is when a subscriber leaves the telecom service provider.
- **Industry Issue**: High churn leads to revenue loss.
- **Objective**: Identify customers likely to churn in advance so retention campaigns can be targeted effectively.

Business Goals

1. Understand drivers of churn.
2. Create a Churn Risk Score for each customer.
3. Introduce a CHURN-FLAG variable (Yes=1, No=0) to support marketing campaigns.

Dataset details

The dataset contains 1000 customer data**.** Features are State, Account Length, Area Code, Phone, International Plan, VMail Plan, VMail Message, Day Mins, Day Calls, Day Charge, Eve Mins, Eve Calls, Eve Charge, Night Mins, Night Calls, Night Charge, International Mins, International calls, International Charge, CustServ Calls, Churn.

Tools & methodology

1. **Data Preprocessing (SQL)**

   Handled missing values, checks duplicates, checks churn rate vs international, checks churn vs high usage customers and low usages customers.

2. **Exploratory Data Analysis (Excel)**

   Used pivot tables and charts for visualizations.

   Customers with International Plan & High CustServ Calls (>3) churn more.

   Higher Day Minutes → higher chance of churn.

   States with poor service → high churn clusters.

3. **Model Building (Python)**

   Models tested: Logistic Regression (Baseline).Random Forest Classifier.

   Metrics used: Accuracy, Precision, Recall, F1-Score.

4. **Real time prediction using ML (Hugging face spaces)**

   Introduced CHURN-FLAG (Yes=1, No=0) columns using churn columns.

   Make prediction using customers data and found churn risk score.

5. **Interactive dashboards (power BI)**

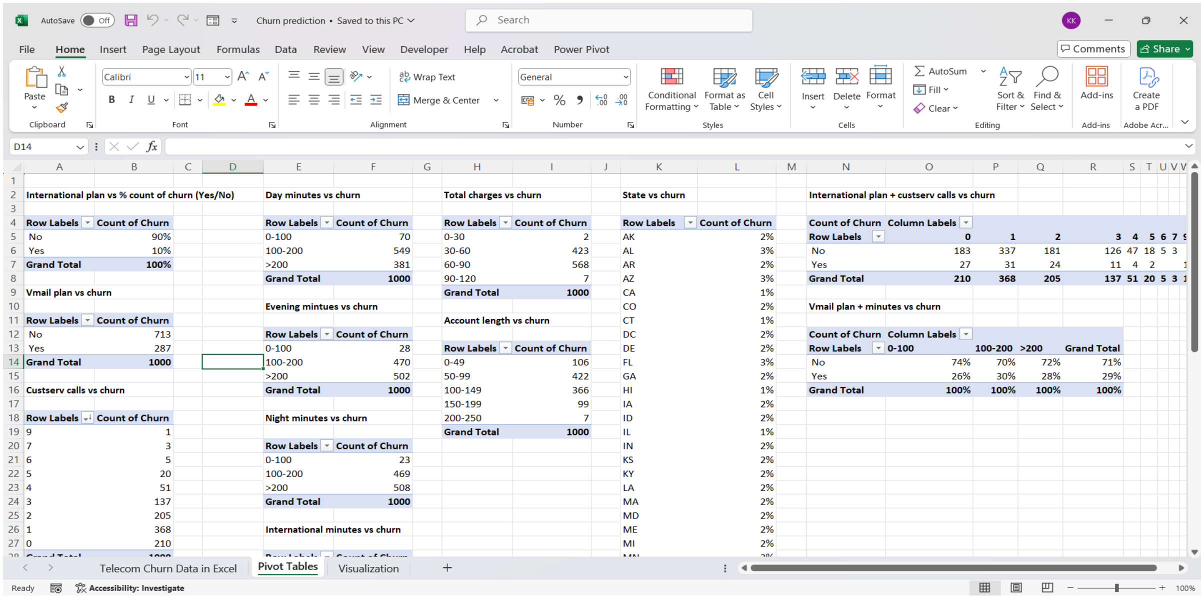
   Create 3 pages dashboard report for telecom churn prediction.

   Churn drivers analysis, Risk score dashboard, Campaign targeting (CHURN-FLAG)**.**

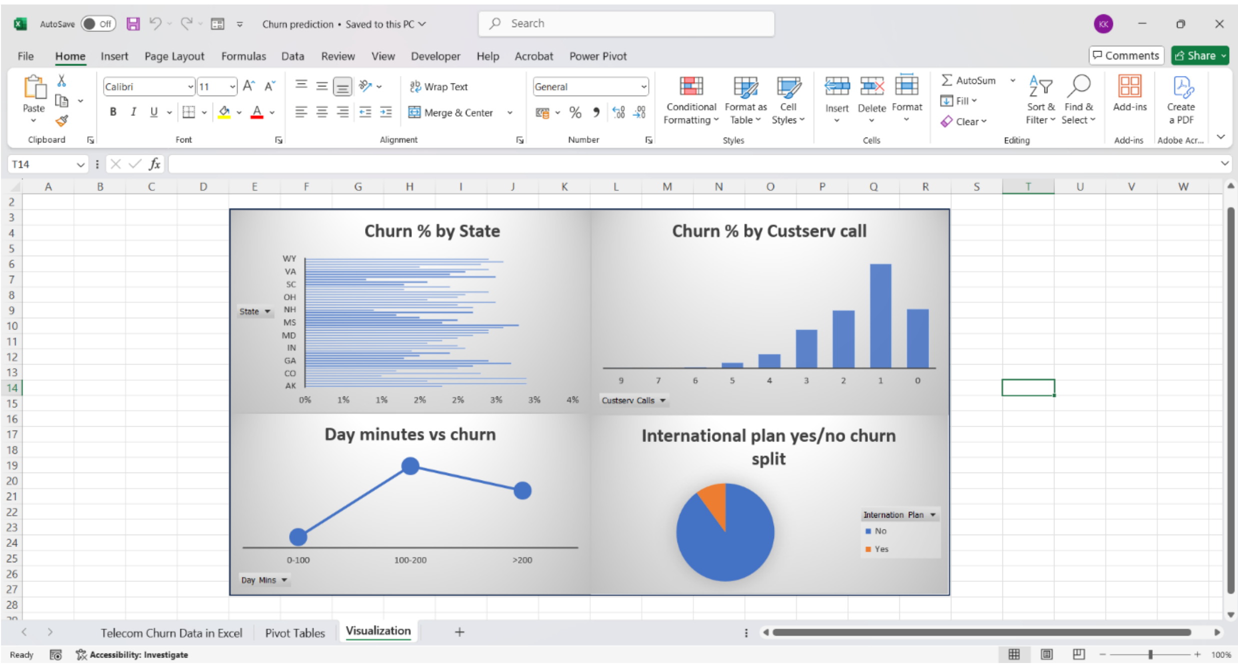
Excel results

We have found top churn drivers using the pivot table in SQL and which is useful for quick aggregation and validation.

We have create pivot table for all features vs churn to checks which makes more numbers of customers to churn.



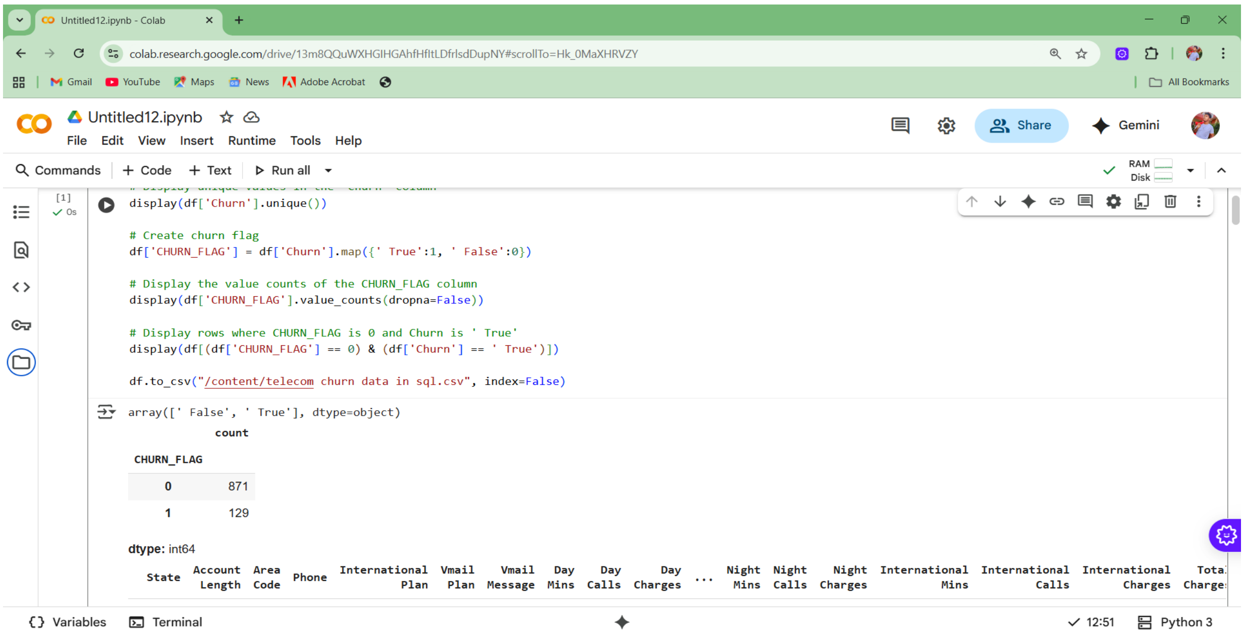
Pivot Tables



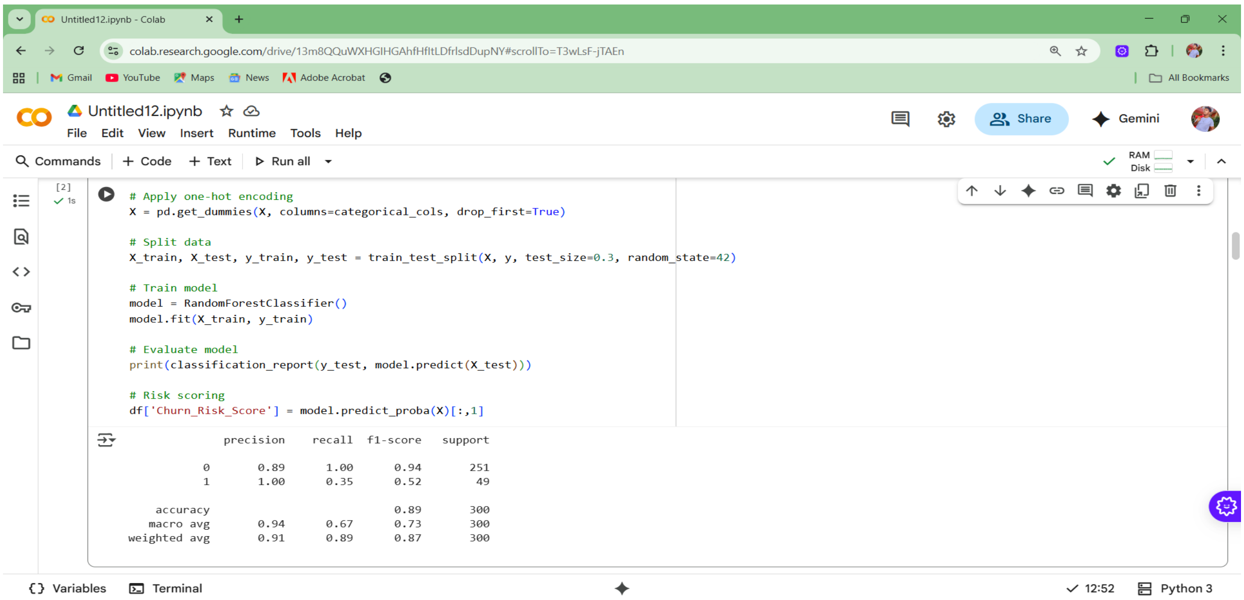
Top Drivers Makes to Churn

Python Results

We have tested model using logistic regression, Random Forest classifier. And the metrics we have used Accuracy, Precision, Recall, F1 Score.



Count of CHURN-FLAG

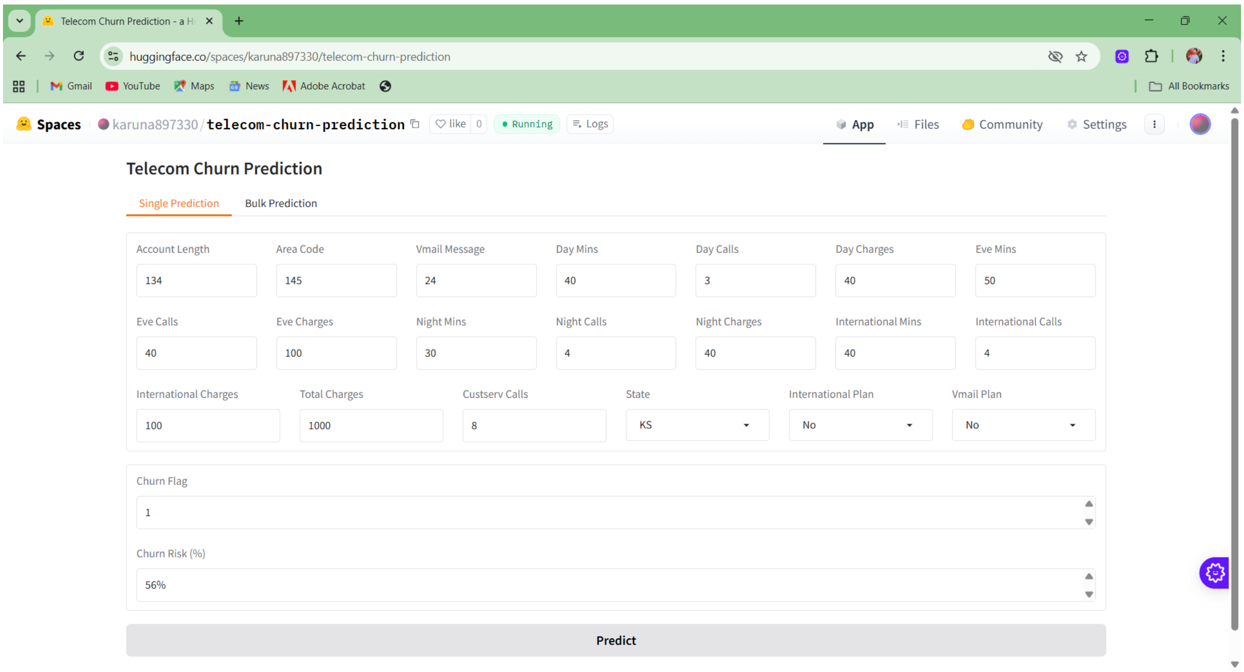


Metrics for Model Testing

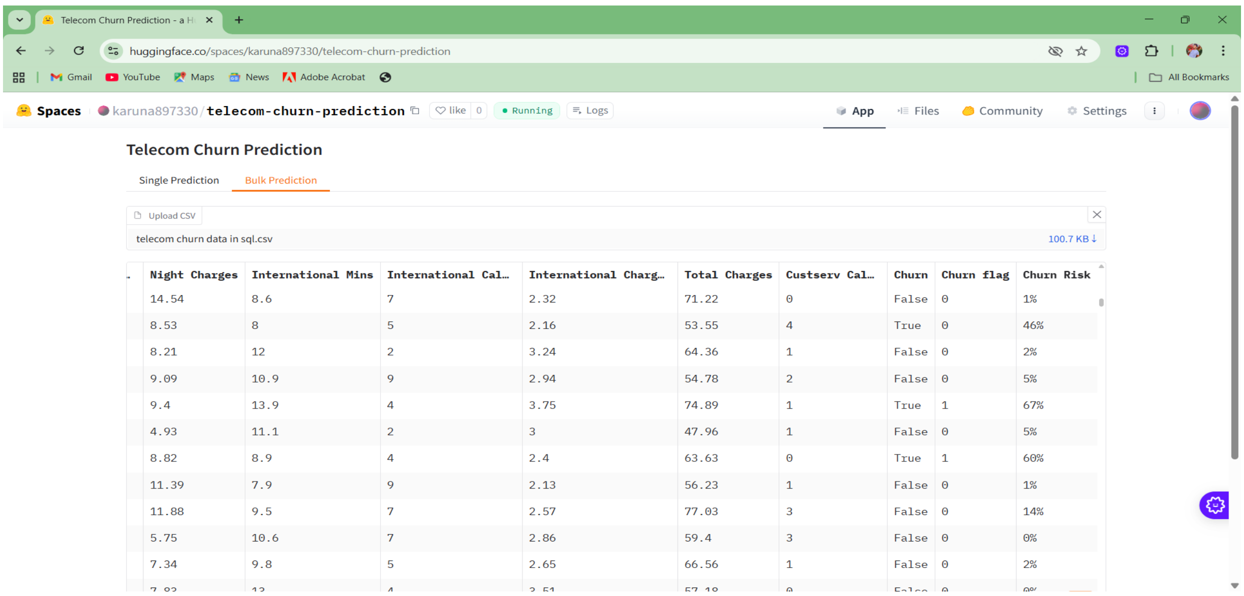
ML prediction (Hugging face spaces)

We have created new columns CHURN-FLAG (Yes=1, No=0) using customer columns in hugging face spaces.

Then using ML it predict churn risk score using CHURN-FLAG and it will used for real time telecom churn prediction.



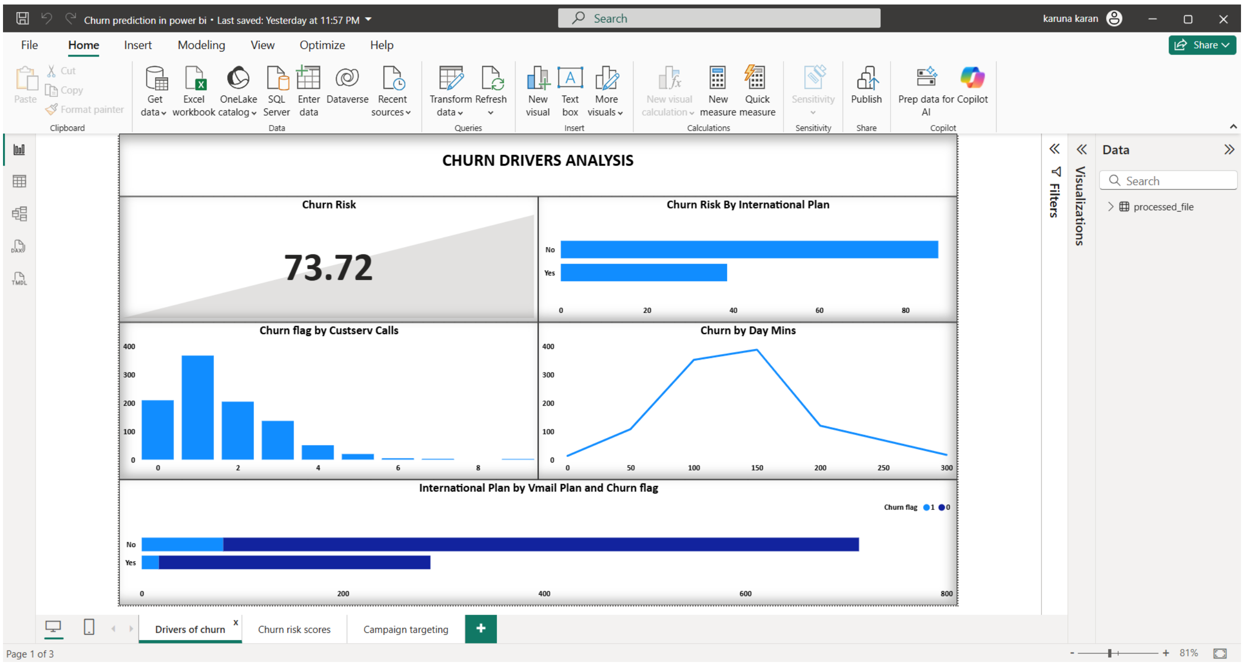
Single Customer Prediction



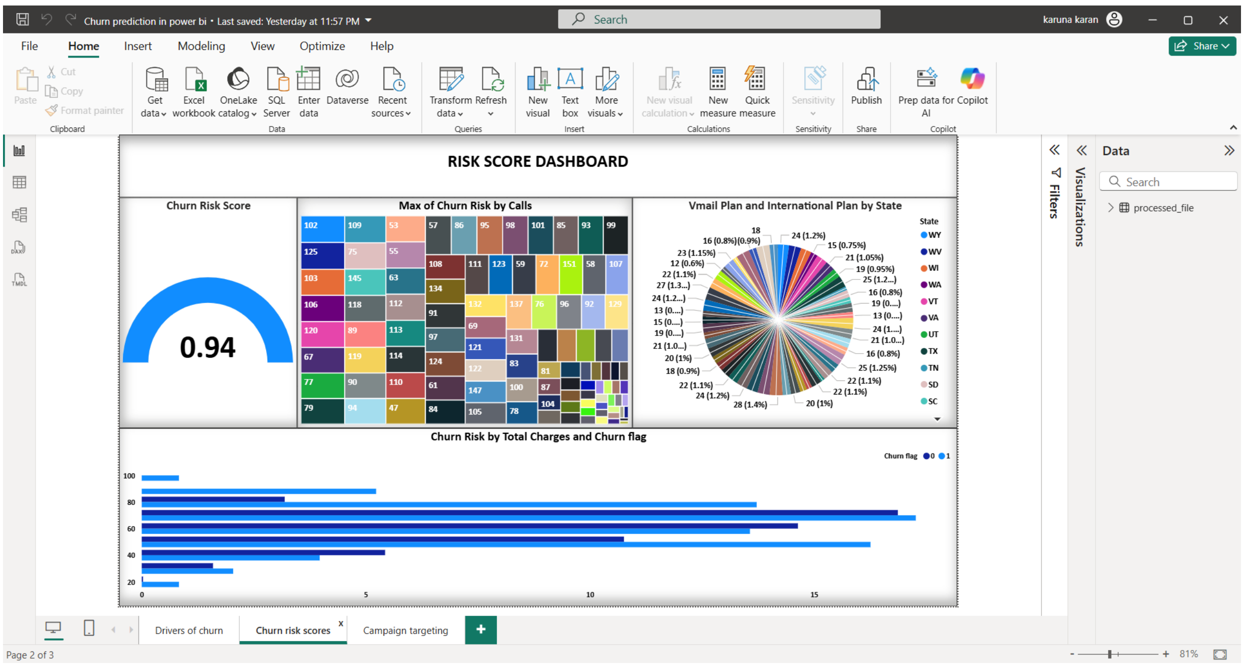
Bulk Customer Prediction

Power BI Dashboards

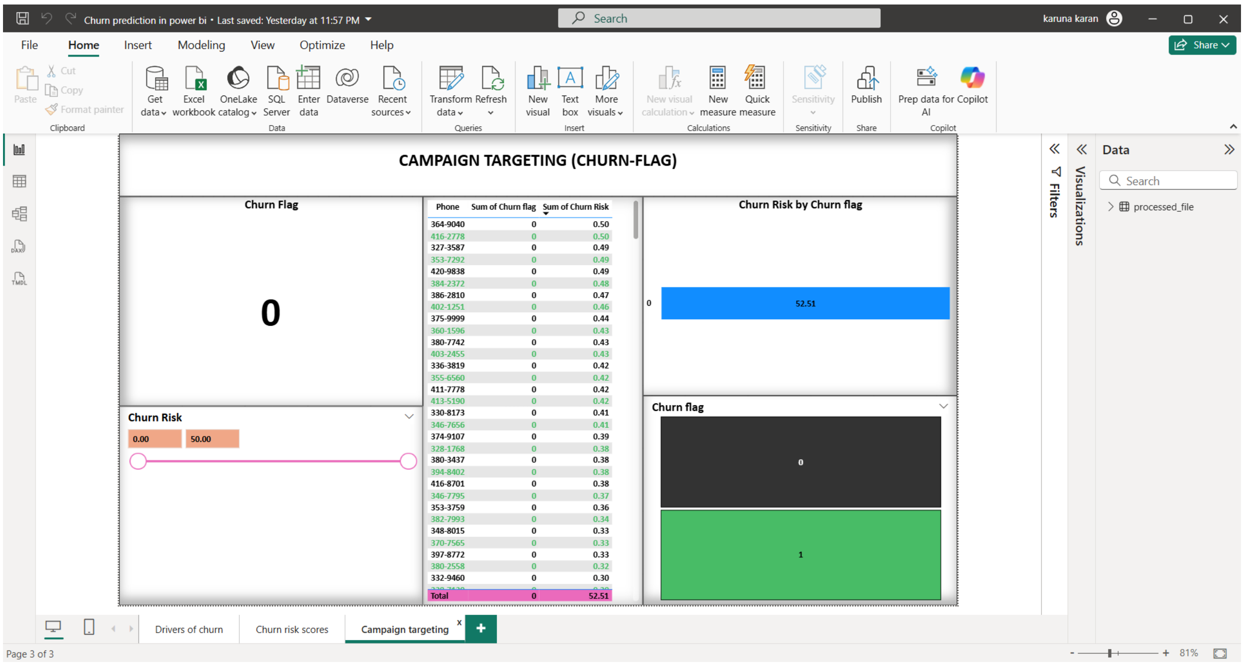
We have creates dashboards for this given three goals. Which will be in interactive manner and understandable improves decision making for business growth.



GOAL 1 – Drivers which makes to churn



GOAL 2 – Finding Churn Risk Score



GOAL 3 – Adding Prediction Variable CHURN – FLAG (Yes=1, No=0)

Insights

- International Plan customers are 3x more likely to churn.
- Customers making >3 customer service calls are at high risk.
- Higher Day Mins and Charges correlate with churn.

Recommendations

1. Customer Retention Strategies

Target high-risk customers identified by the model with retention offers such as discounts, free add-ons, or loyalty rewards.

Provide special retention plans for customers with International Plans (since they churn more).

2. Improve Customer Support

Customers making >3 customer service calls are highly likely to churn.

Train customer support agents to resolve issues within the first two calls.

Introduce VIP support for high-value customers.

3. **Usage-Based Offers**

Heavy users of Day Minutes and Charges are at risk of churn.

Provide customized day-time packages to reduce bill shock.

Introduce flexible billing or rollover minutes.

4. **Proactive Communication**

Send proactive alerts for billing, network outages, or plan expiry to reduce dissatisfaction.

Use email/SMS campaigns targeted to churn-flagged customers.

5. **Data-Driven Monitoring**

Continuously monitor churn with the Power BI dashboard.

Refresh churn predictions monthly and track churn reduction trends.

Conclusion

The project successfully built an ML solution to predict telecom customer churn. The model can guide marketing campaigns and reduce churn. Integrating the solution into BI dashboards ensures real-time monitoring and business adoption.