



# Understanding and Predicting Bank Customer Attrition

Brody Hill







## The Problem:

**In the current higher-for-longer rate environment, more customers than usual are withdrawing balances from the bank. It must be understood what customer attributes most influence attrition to develop targeted strategies to address them.**



## **The Findings Thus Far:**

**A customers member status, their gender, and their geography are the most positive predictors of attrition.**



## The Data:

**The dataset I used for this project contains data on 10,000 customers of a European Multinational Bank, with customers across Germany, France and Spain.**

# Data Wrangling:

Checked for duplicates and missing values (there weren't any) and explored class distribution. 79.6% of customers had not churned, 20.4% had.

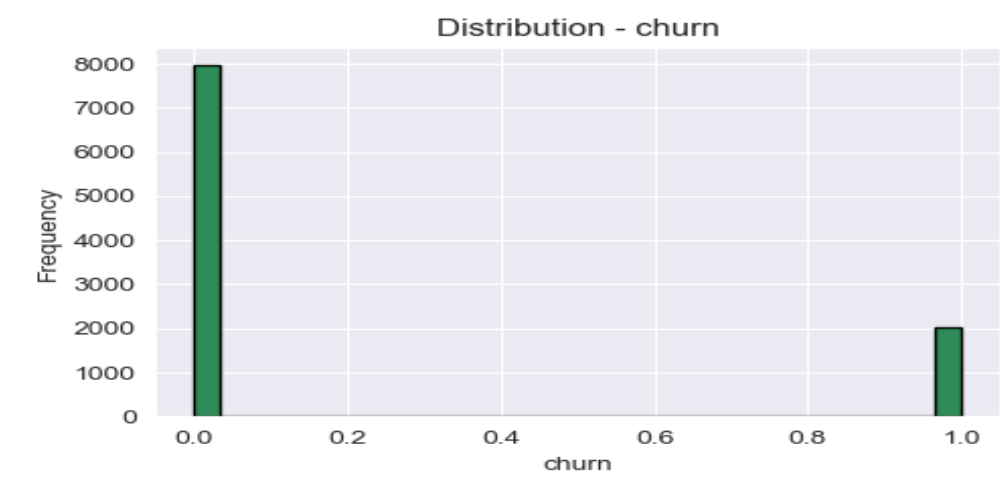
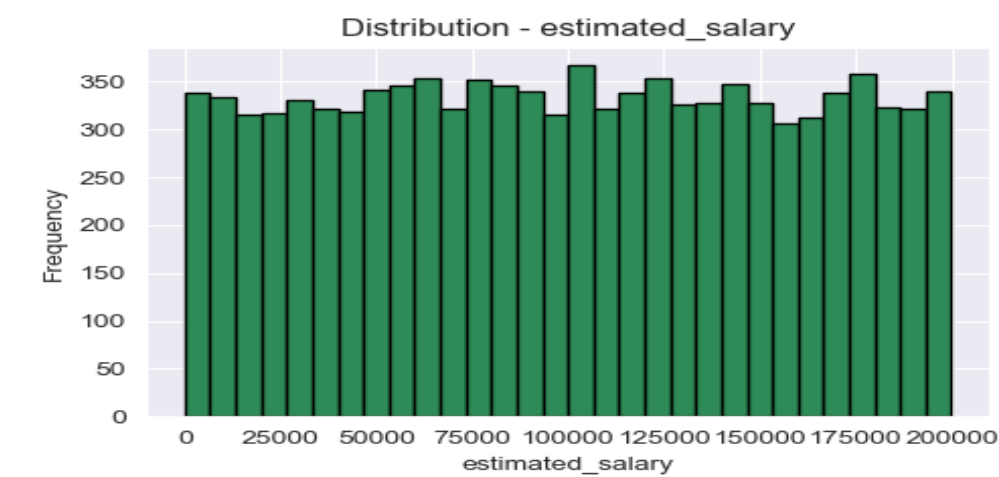
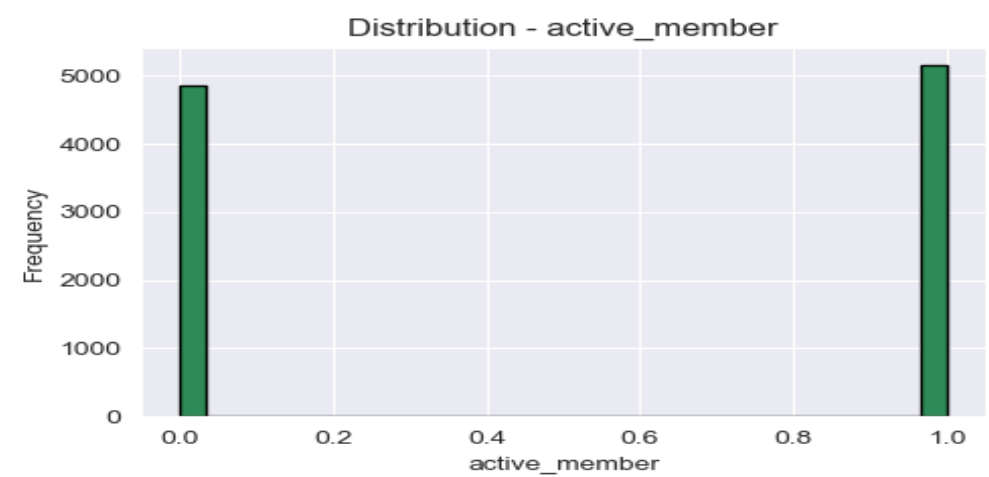
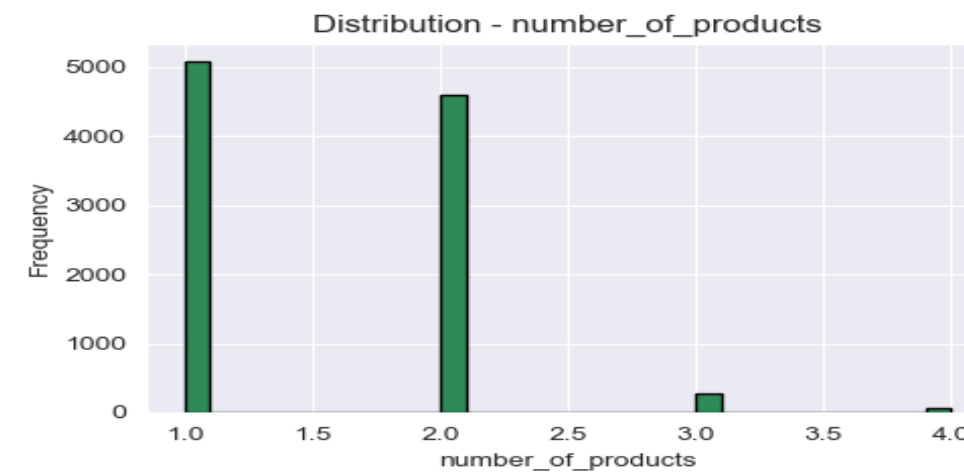
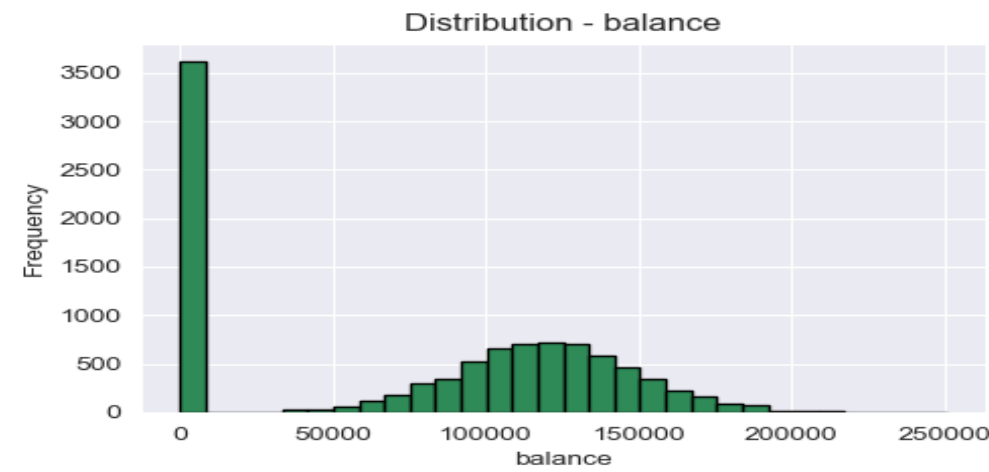
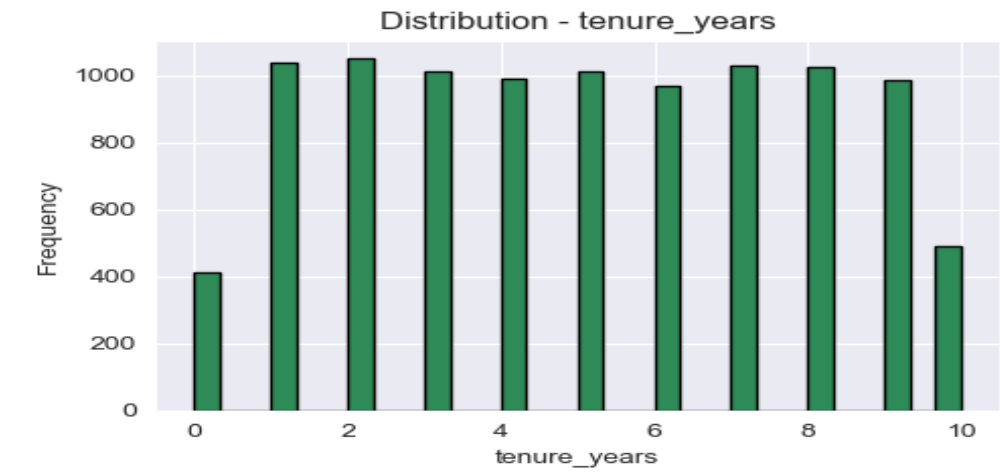
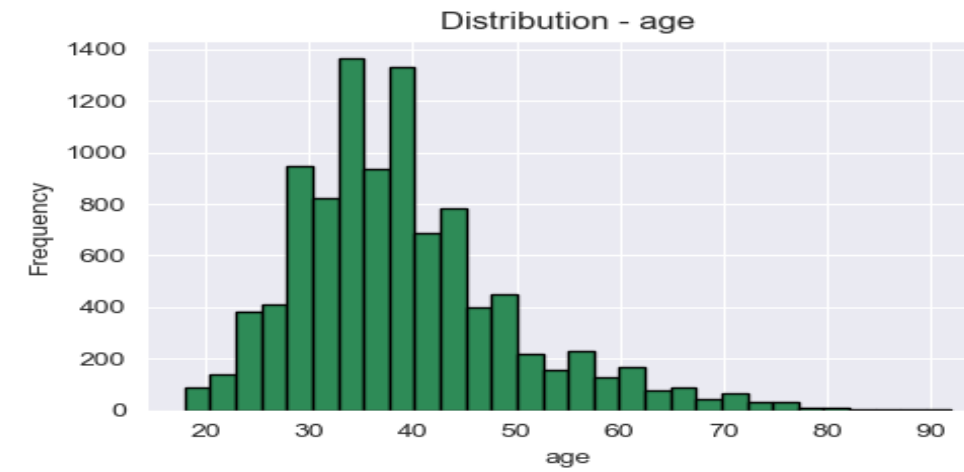
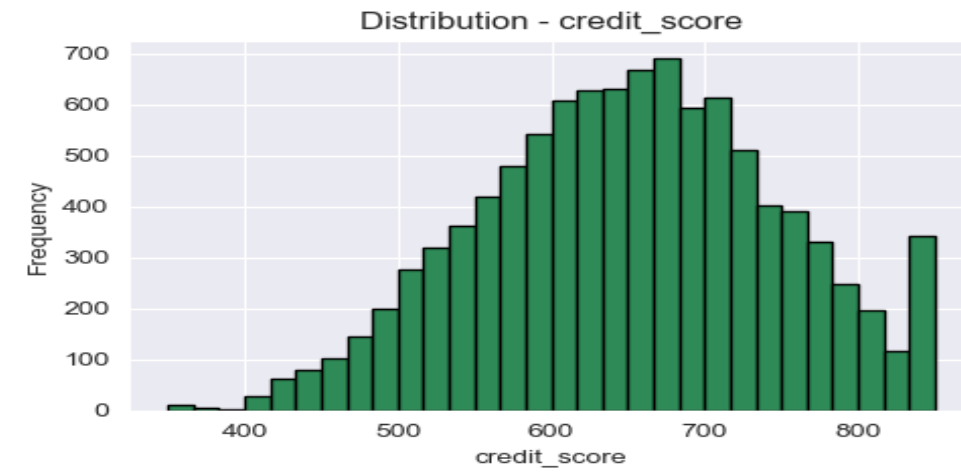
Ensured every columns data type made sense for the data it contained.

Renamed two columns to better describe what they contained.

Dropped the customer id column knowing it would not be useful for modeling.

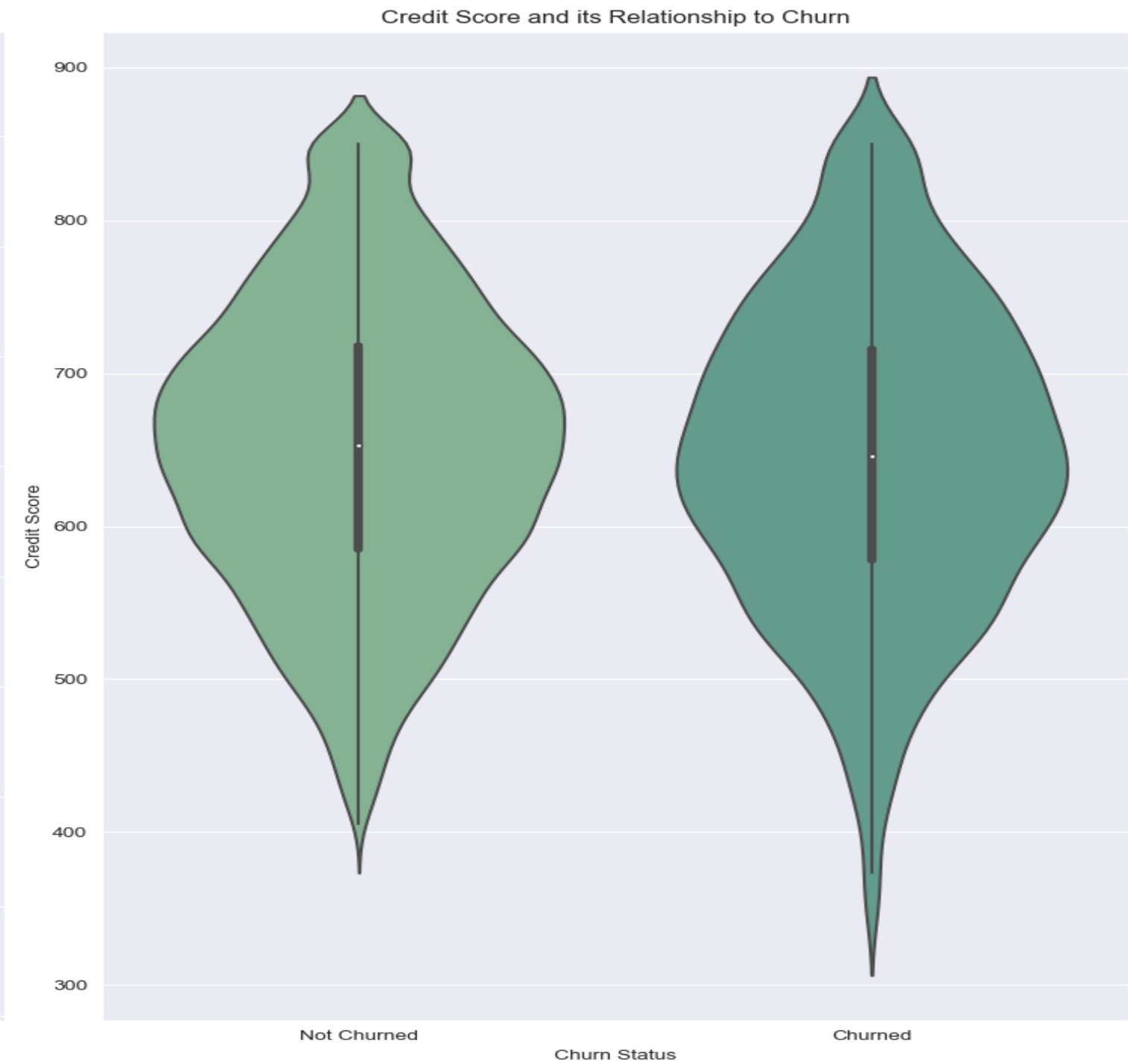
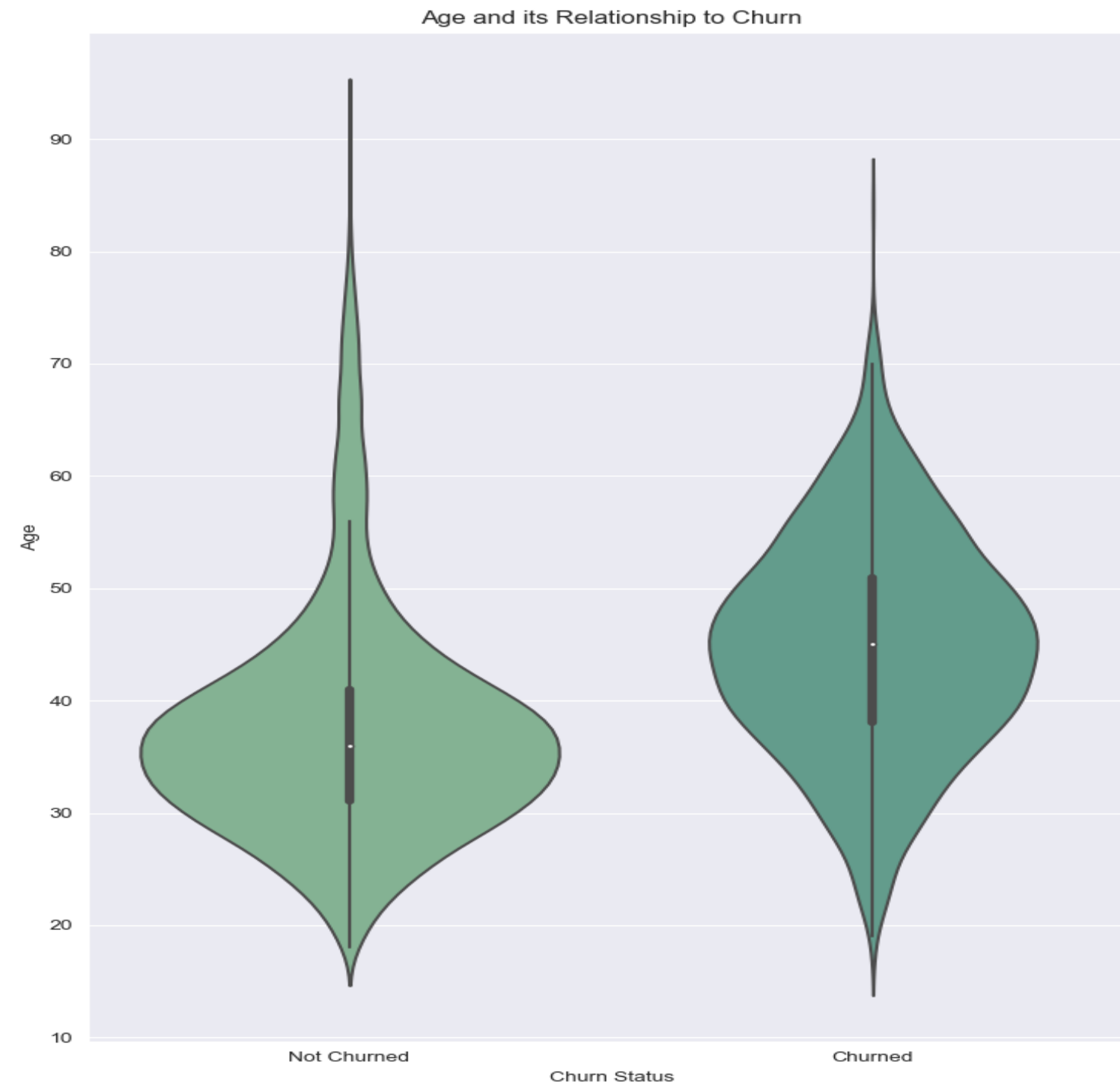


# EDA Part 1:



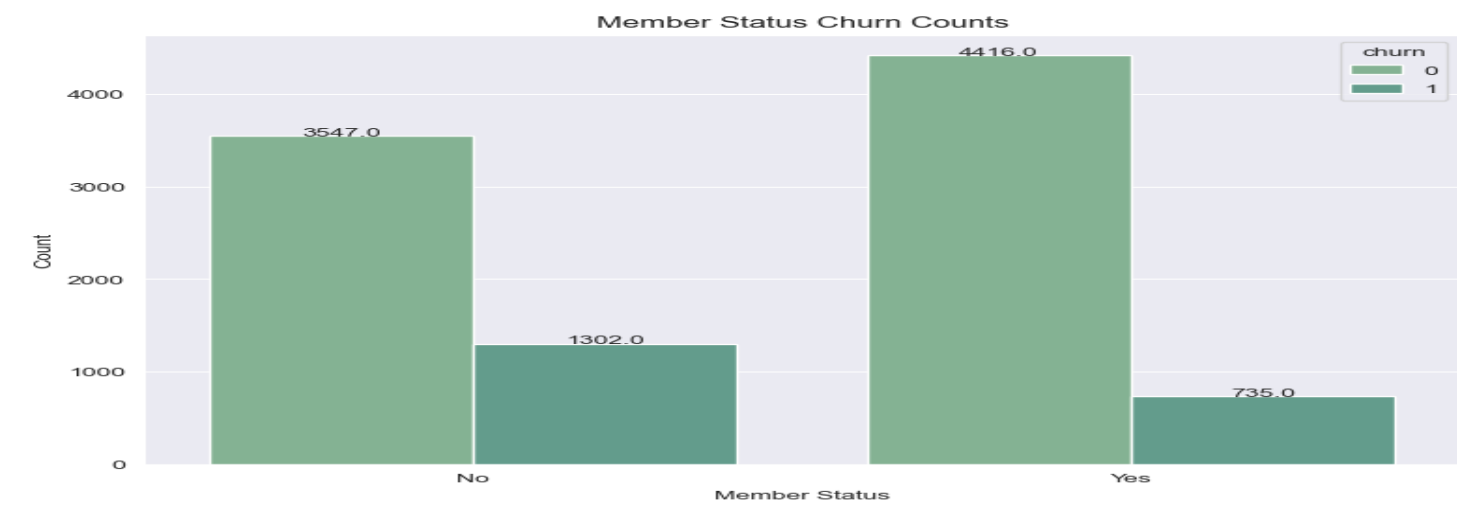
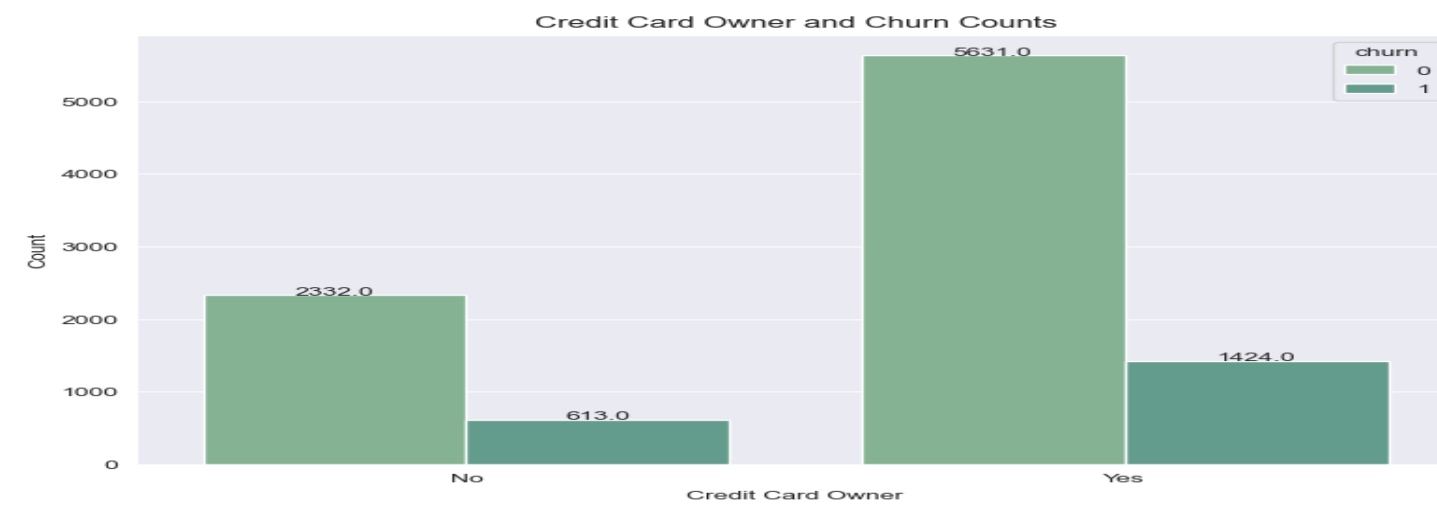
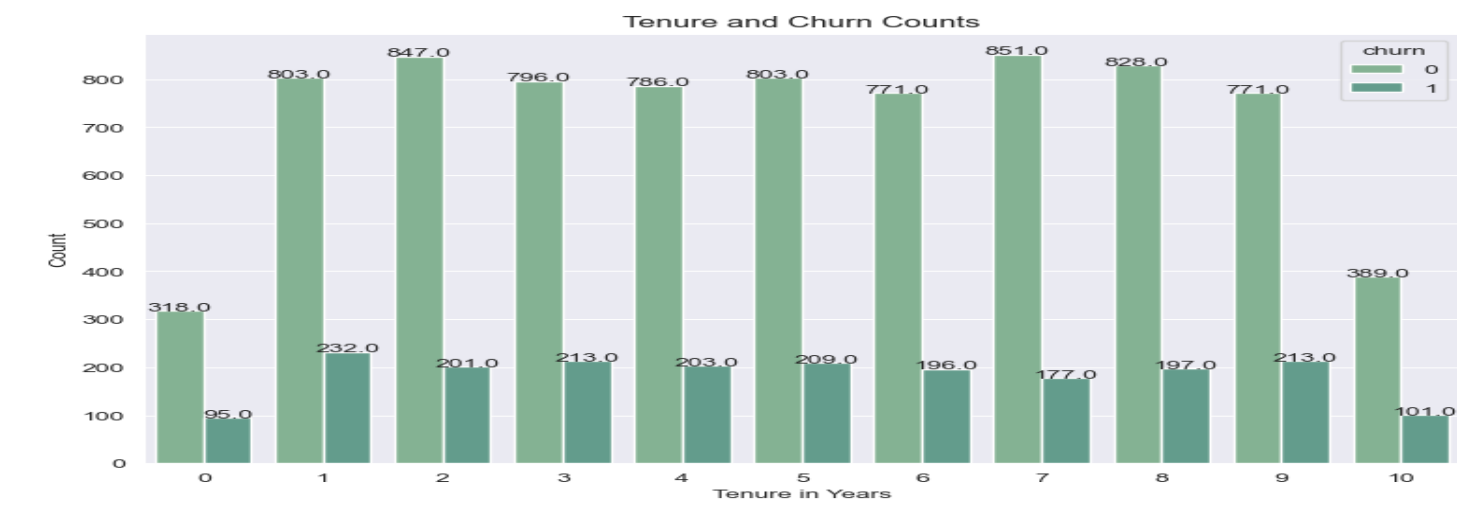
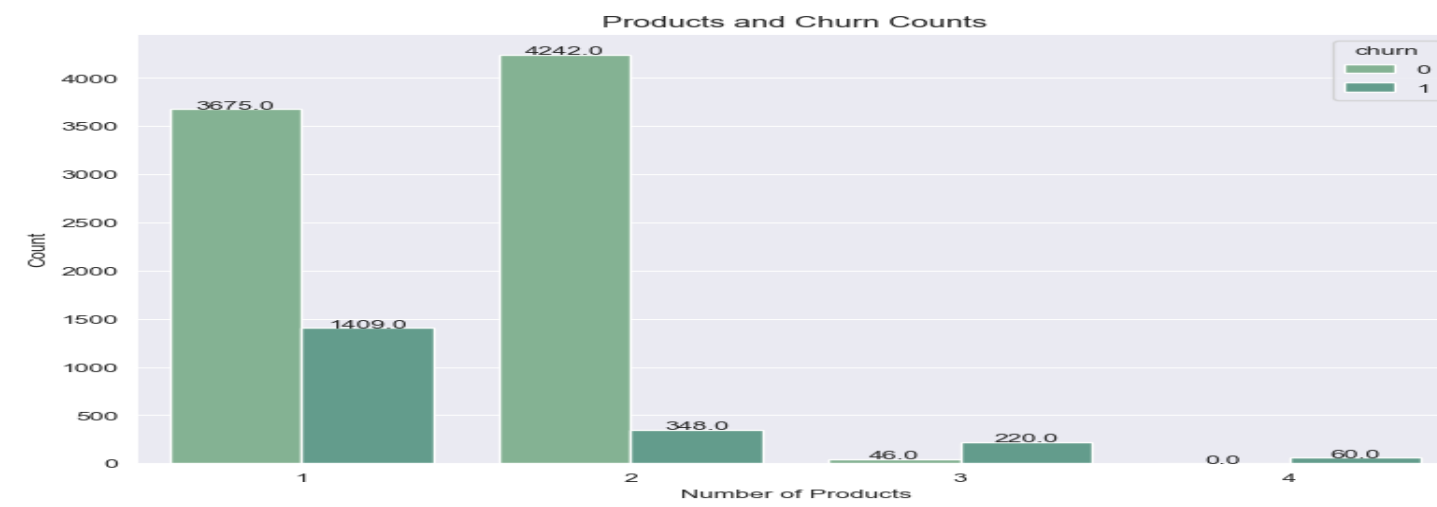
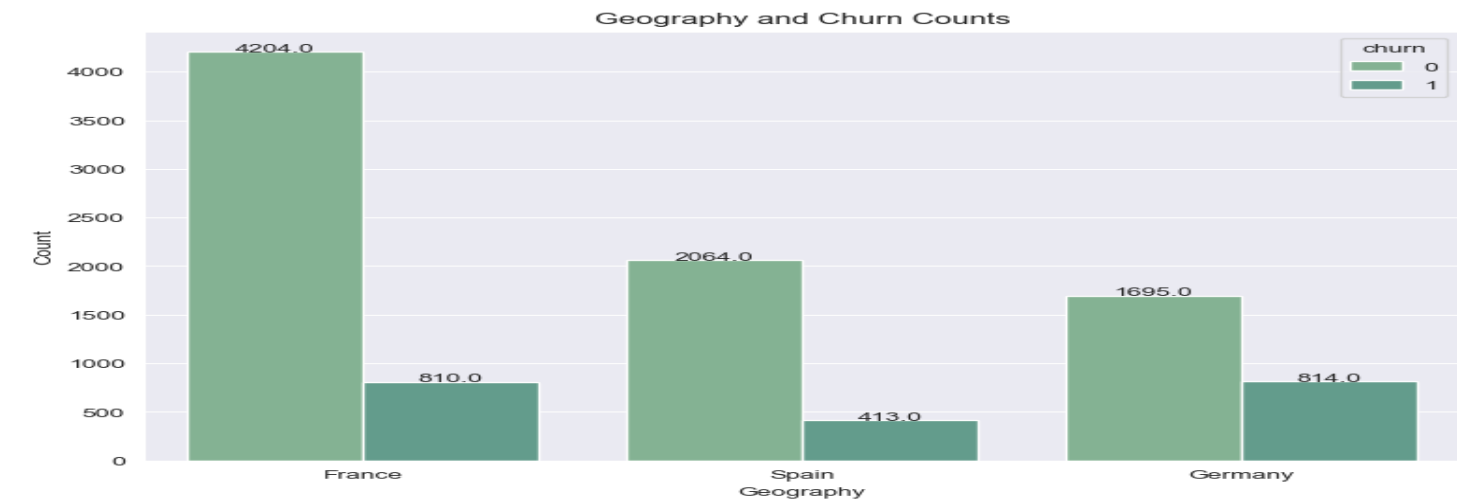
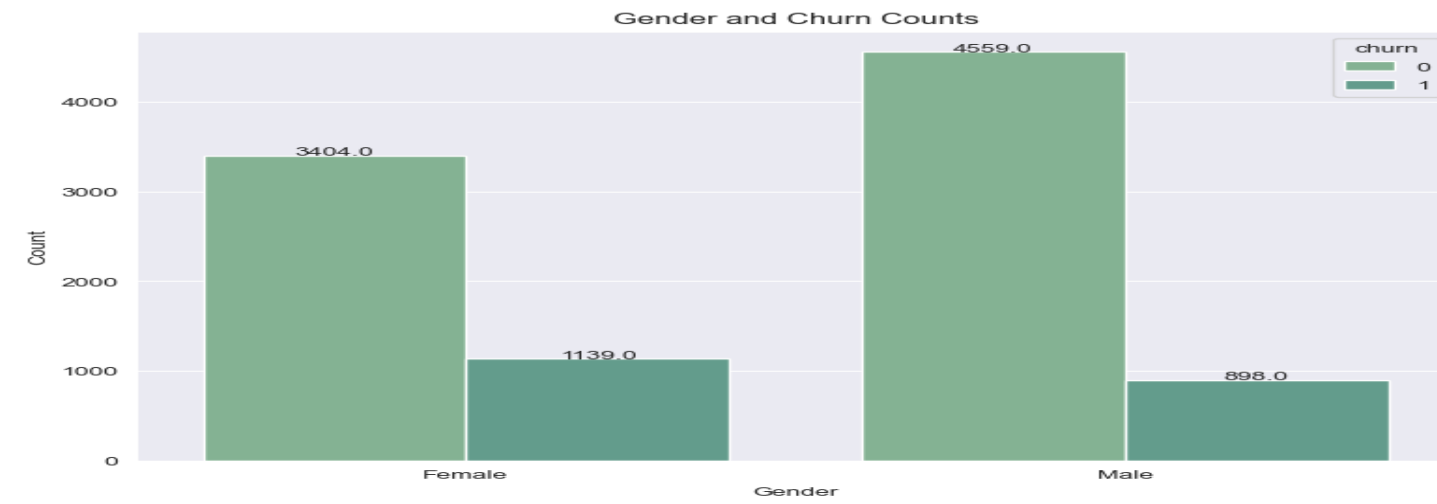


## EDA Part 2:





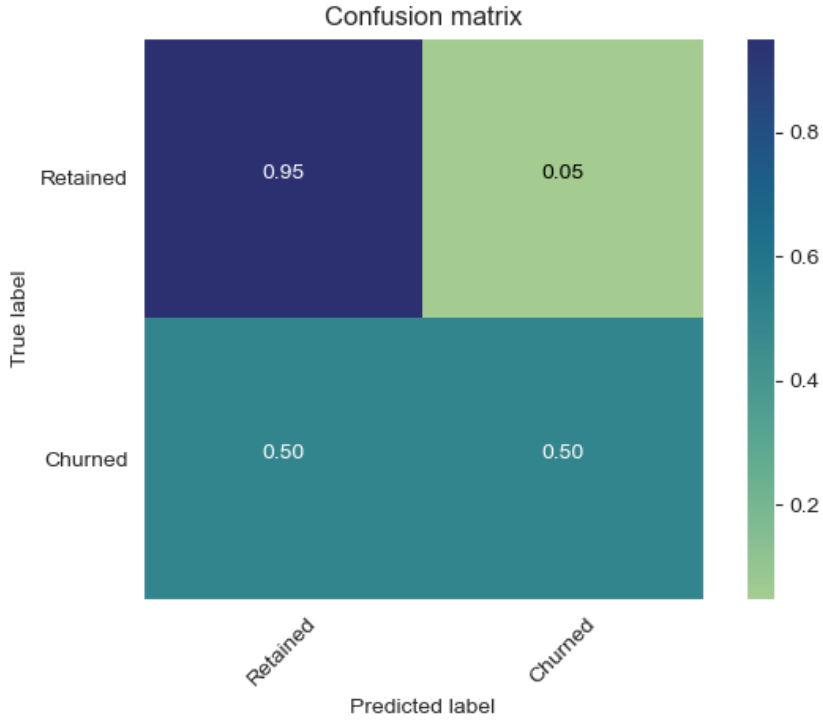
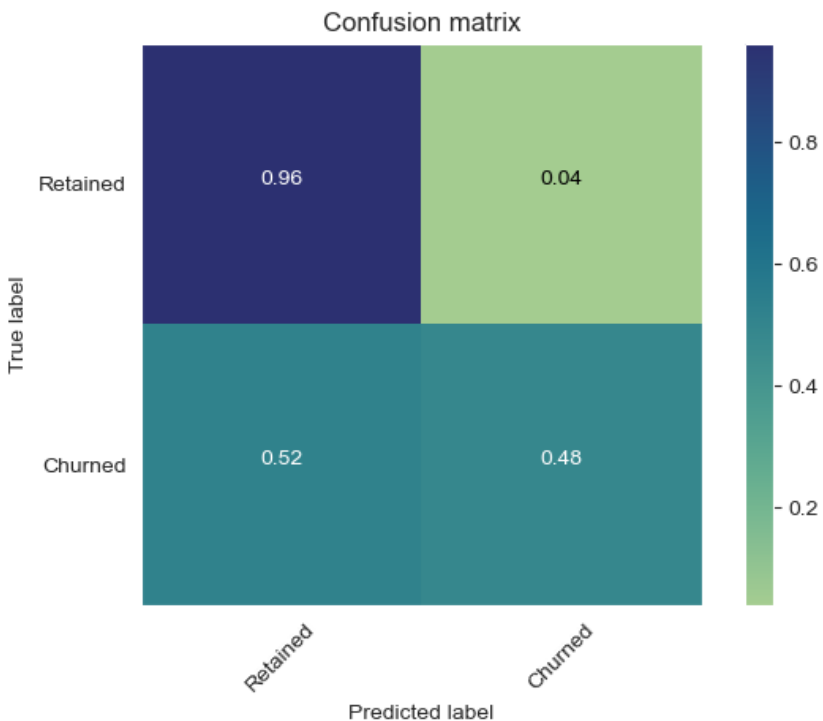
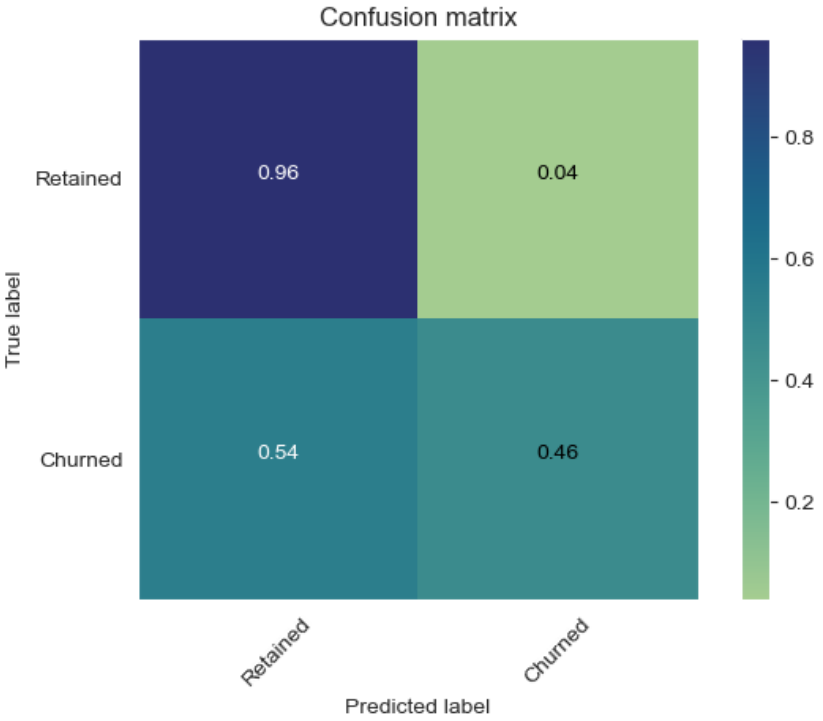
## EDA Part 3:





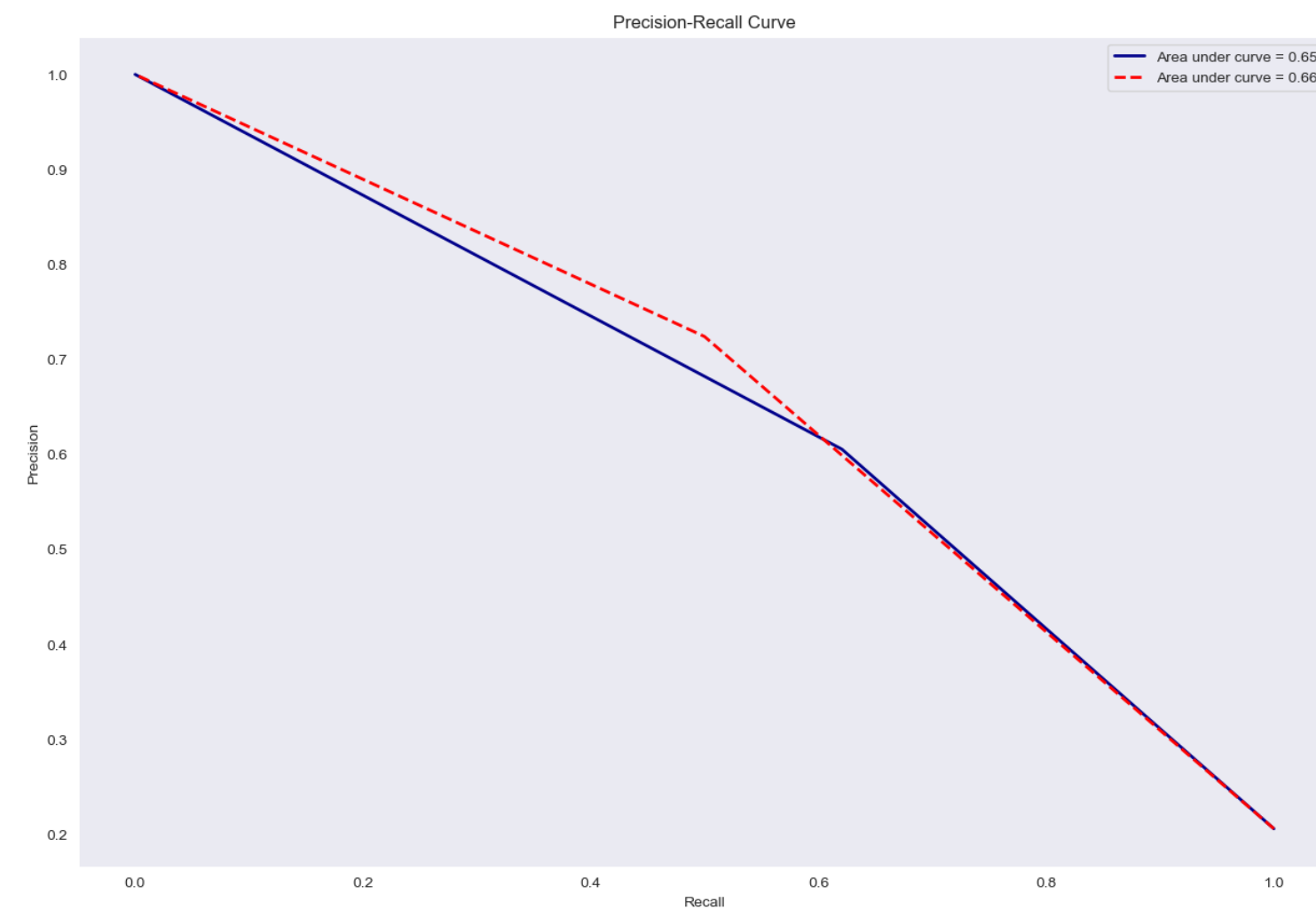
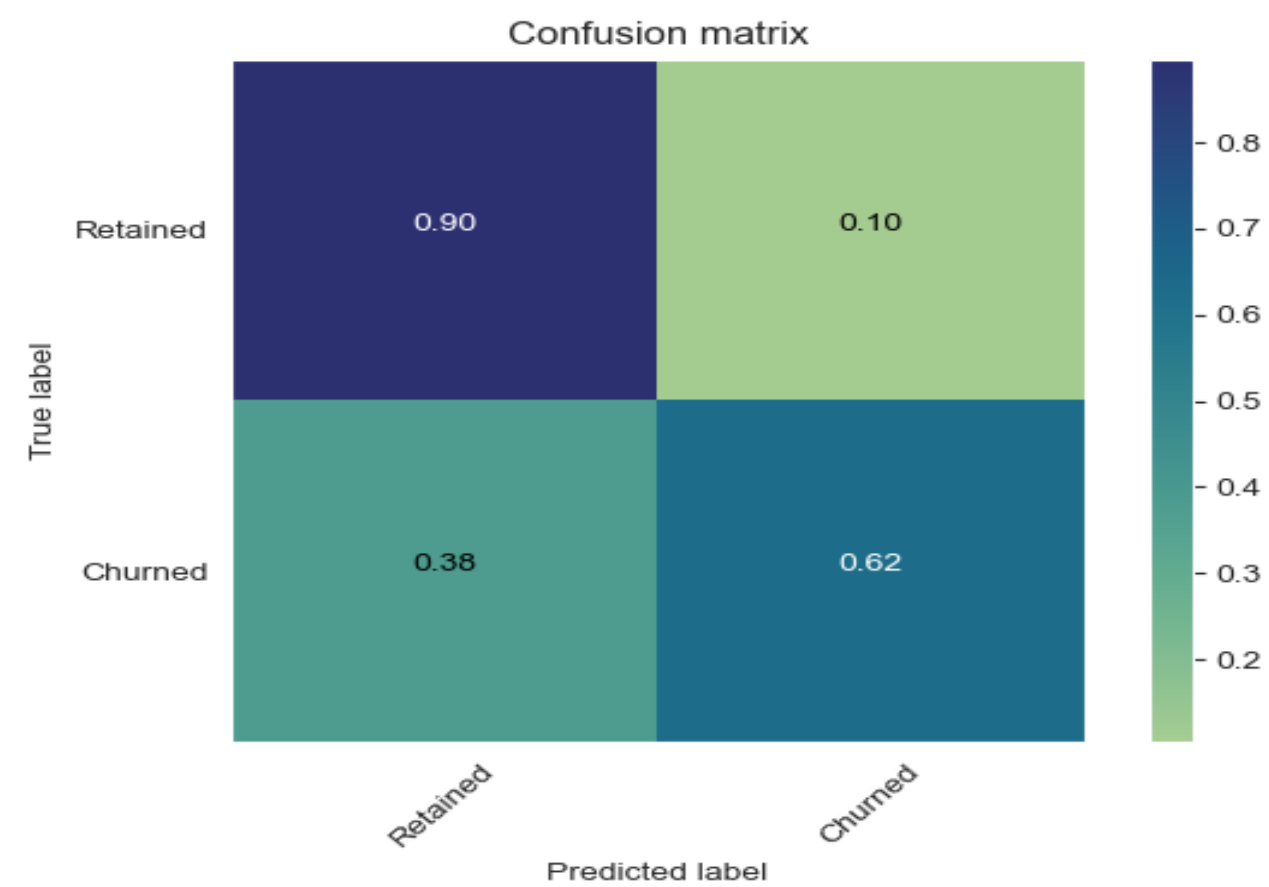


# Modeling Part 1: Baselines



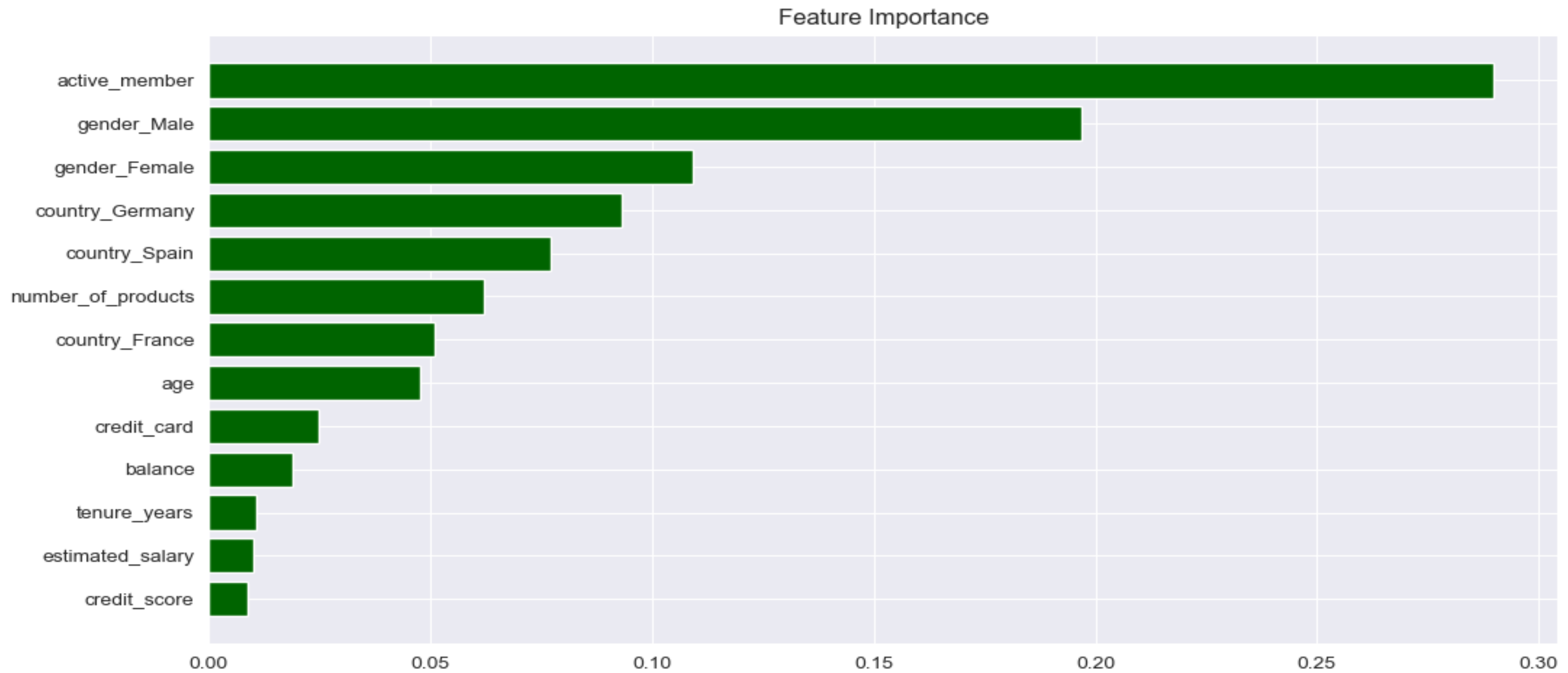


## Modeling Part 2: Adjusting



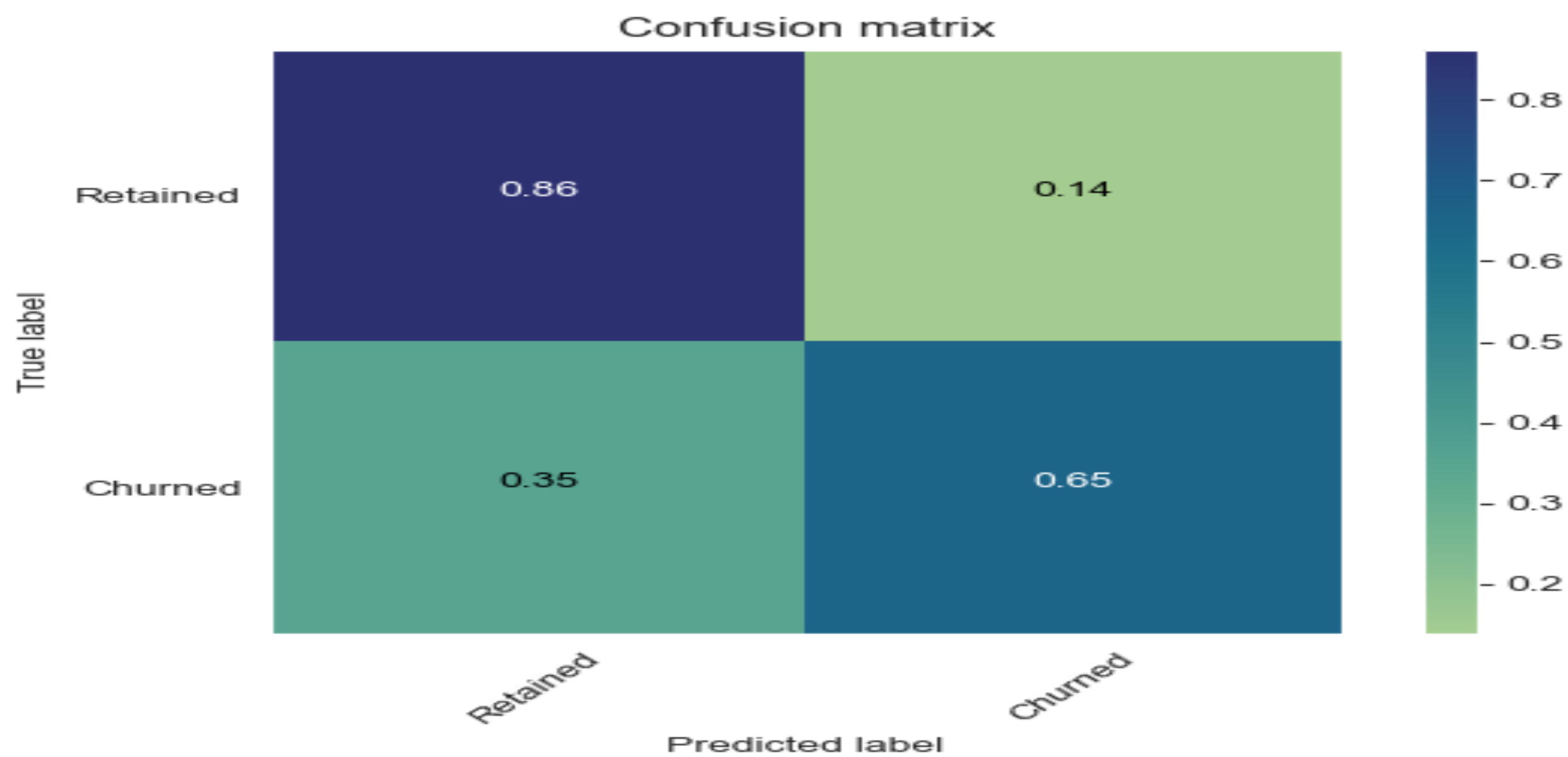


## Modeling Part 2: Adjusting Continued





## Modeling Part 3: Further Refined





### McNemar Test Results:

**P-value of  $1.4249678757951758e-17$ . Extremely Small.**

**95% CI: (12.8% – 15.6%)**





## **Key Takeaways:**

**1. Active member status seems the best indicator of churn.**

**2. Gender and Geography play significant roles.**

**3. Further testing of each feature and their impact on model performance is needed.**



## **Moving Forward:**

- 1. Test each feature and it's impact on model performance.**
- 2. Gather more data on female customers , their member status and geography specifically to see if initial assumptions extend to a broader population.**
- 3. Continue modeling with different methods to improve results.**

***Any Questions?***