

# CUSTOMER CHURN ANALYSIS

FOR A EUROPEAN BANK

**Developer: Tian Fu**

# HIGHLIGHTS

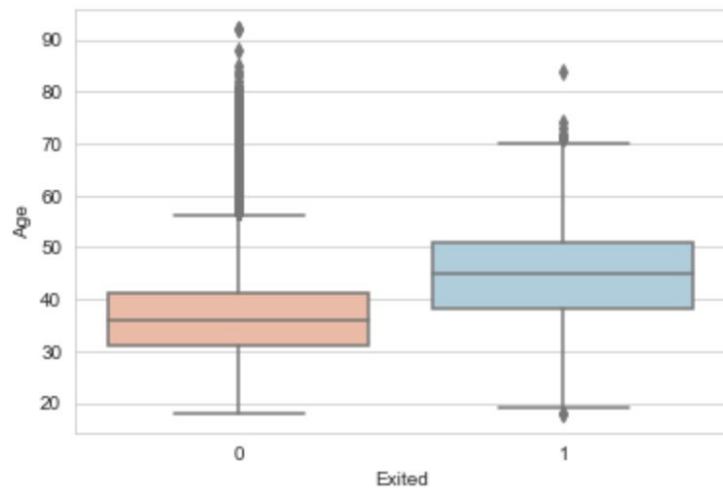
- ❖ An overall understanding of the dataset including features and target variable through various visualizations and summary statistics.
  - Most features have discrete values.
- ❖ Explore potentially important variables.
  - Most customers are only using 1 or 2 bank products.
- ❖ Three different machine learning algorithms (logistic regression, random forest, and xgboost) to predict whether a customer churned based on selected features. Based on evaluation metrics - AUC, xgboost model was the best.

# REVIEW PROGRESS

During this sprint, I completed the following stories:

1. Three stories in Epic #1:
  - a. Overall distribution of features.
  - b. Distribution of features across churned/not-churned customers.
  - c. Distribution of features across demographic groups.
2. Three stories in Epic #2:
  - a. Feature selection.
  - b. Modeling.
  - c. Hyper-parameter tuning.

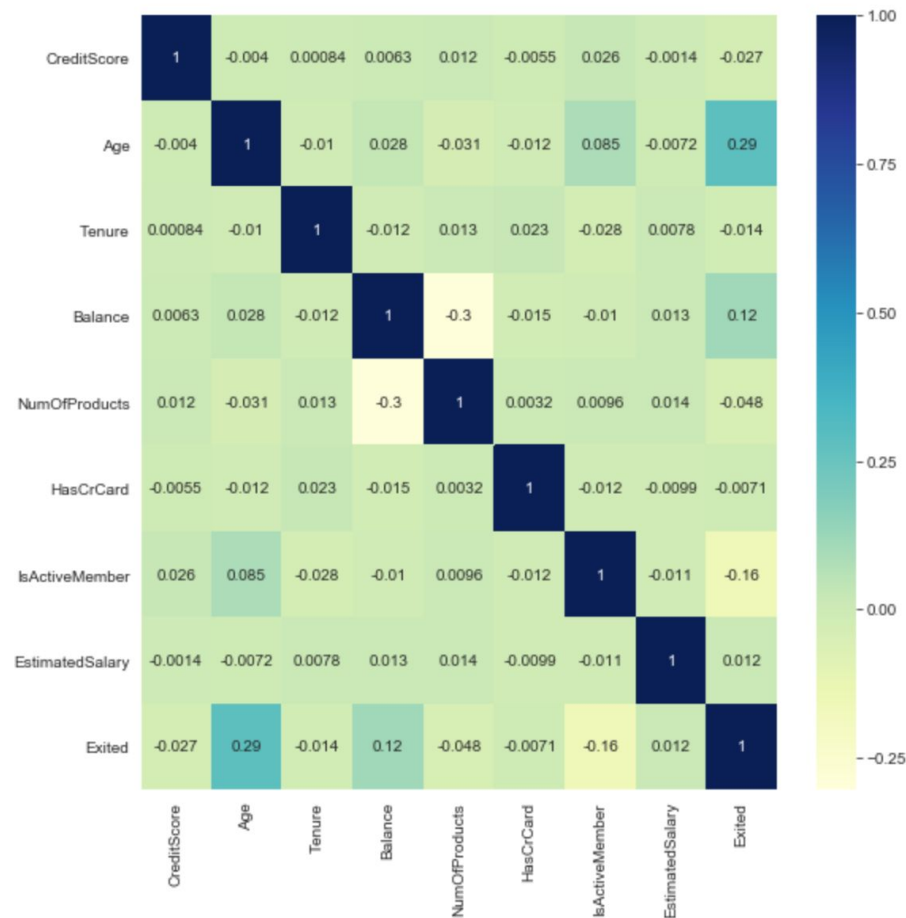
# VISUAL ANALYSIS



Distribution of Age across Two Groups

churned customers are mostly around 40-50 years old who need to support the whole family and think about retirement

## Confusion Matrix



# LESSONS LEARNED

about this bank

- ❖ This bank should be relatively small regional European bank operating only in Germany, Spain, and France. There might be only around 4-5 products at the bank, offering customers few choices. Surprisingly, churned customers tried a variety of products and decide to quit, which is indicative of quality of this bank's products.
- ❖ Among churned customers, some are new customers but some are old customers with tenure of 9 years. The bank should consider talking to current customers to seek for ways to improve their customer service and conduct some competitive analysis to find out if they are falling behind in the market.
- ❖ More French and Germans churned. The bank might consider customizing its marketing strategies and products by regions.

# RECOMMENDATIONS

The following stories should be finished in next sprint:

- ❖ Enable user input options and predict whether the customer will churn or not based on user input.
- ❖ Construct a pipeline for this project, create database and load dataset on AWS.
- ❖ Model deployment on AWS.

THANK YOU!