Understanding
Bank Customer Churn

Contents

Business Problem

Predict customer churn in a bank to identify at-risk customers and implement targeted retention strategies.

Managing customer churn is crucial for maintaining financial health, profitability, and competitive advantage. Banks need to understand the factors driving churn and implement effective strategies to enhance customer retention.

Here are some of the financial implications of the decline in customers for the bank

1. Loss of Revenue

Direct Loss: When customers leave, the bank loses the revenue generated from their accounts, including fees, interest, and other charges.

Future Revenue: Potential future revenue from these customers is also lost, impacting long-term financial stability.

2. Increased Costs

Acquisition Costs: It is generally more expensive to acquire new customers than to retain existing ones. Marketing and promotional expenses to attract new customers can be substantial.

Operational Costs: Handling the churn process, including account closures and customer service interactions, incurs additional operational costs.

3. Impact on Profitability

Reduced Margins: The combination of lost revenue and increased costs can reduce the bank's profit margins.

Investment in Retention: Banks may need to invest more in customer retention programs, loyalty incentives, and improved customer service to reduce churn rates.

4. Brand and Reputation

Negative Perception: High churn rates can damage the bank's reputation, making it harder to attract new customers and retain existing ones.

Customer Trust: Losing customers can indicate underlying issues with customer satisfaction and trust, which can have broader implications for the bank's market position.

5. Competitive Disadvantage

Market Share: High churn rates can lead to a loss of market share to competitors who are better at retaining their customers.

Strategic Focus: The bank may need to shift its strategic focus to address churn, potentially diverting resources from other growth initiatives.

Dataset Story

At the beginning of the search for the dataset, we searched for many datasets, for example, the first one was not suitable because the columns in it were not clear, then we searched more until we reached the appropriate dataset that we will work on during our project

This dataset contains demographic and banking-related information of customers of a bank.

This dataset can be used to understand the customer base of the bank and to predict customer churn. Customer churn is an important topic for banks to understand and address in order to retain customers and reduce customer loss. Analyzing this dataset and developing a machine learning model to predict churned customers based on various factors can be valuable for the bank's customer retention efforts

In this project, we will look at which customers are prone to churn and non-churn, what are their characteristics, and the factors that affect churn and non-churn.

Dataset Overview

The dataset contains information about bank customers, focusing on various attributes such as credit score, country, gender, age, tenure, balance, number of products, credit card ownership, active membership, estimated salary, and churn status.

Key Attributes

Credit Score: Numerical value representing the creditworthiness of the customer.

Country: The country where the customer resides (France, Spain, Germany).

Gender: The gender of the customer (Male, Female).

Age: The age of the customer.

Tenure: The number of years the customer has been with the bank.

Balance: The current balance in the customer's account.

Product Number: The number of products the customer has with the bank.

Credit Card: Indicates whether the customer has a credit card (1 for yes, 0 for no).

Active Member: Indicates whether the customer is an active member (1 for yes, 0 for no).

Estimated Salary: The estimated annual salary of the customer.

Churn: Indicates whether the customer has churned (1 for yes, 0 for no).

Gender Male: 1 indicates to true and 0 to false

Age group: Includes three age groups (17-36, 37-55, and 56-65)

Age seg(A): Incudes (17-36)

Age seg(B): Incudes (37-55)

Age seg(C): Incudes (56-65)

Gender country: Includes type of gender and country

Balance to salary ratio: Demonstrates a client's financial stability

Balance to credit score ratio: High ratio: A high ratio of credit-to-credit score may indicate that the customer has significant financial reserves. However, if the credit score is low, it may indicate that the customer is not managing their credit well, despite having a high balance.

Low ratio: A low ratio may indicate that a customer has a balance more in line with their credit score, indicating balanced financial management.

This ratio helps in understanding the relationship between a customer's available credit and their creditworthiness. It can be a useful metric for financial institutions to assess the risk and financial behavior of their customers.

Initial Findings

Demographics:

Countries Represented: France, Spain, Germany.

Gender Distribution: Both male and female customers.

Age Range: Customers range from 17 to 88 years old.

Financial Information:

Credit Scores: Range from 350 to 850.

Balances: Vary significantly, with some customers having a balance of 0.

Estimated Salaries: Range widely, from 11.58 to 199992.48

Account Details:

Tenure: Ranges from 0 to 10 years.

Number of Products: Customers have between 1 to 4 products.

Credit Card Ownership: Some customers have credit cards, while others do not.

Active Membership: Indicates whether customers are active members.

Churn Analysis:

The churn status is a critical attribute, indicating whether the customer has left the bank. This can be used for further analysis to understand customer behavior and retention.

Data Cleaning and Preparation

- Data Cleaning Process

We had many Outliers in the age column, credit score column when we made a Boxplot of them

The number of outliers in Age was about 350 and the number of outliers in Credit score was 15

They were dealt with by scanning

We also removed the customer_id column because we don't need it for analytics

- Data Transformation

We have created a lot of features and they (Age group, Credit score seg, Age seg(A), Age seg(B), Age seg(C), Balance to salary ratio, Balance to credit score ratio, credit score seg, tenure seg, gender num, Country Spain, Country Germany, Gender country, Gender male, Balance seg)

We used Label Encoder to convert categorical features to numeric values to make machine learning easier

Here are the features we converted

1- Country

2- Gender

3- Age seg

4- Credit score seg

5- Balance seg

6- Tenure seg

This transformation allows machine learning models to effectively interpret categorical data

Exploratory Data Analysis (EDA)

- **Summary Statistics**

-Credit_score

Min = 383

Max = 850

Mean = 650.74

Std = 96.05

-Age

Min = 18

Max = 61

Mean = 37.65

Std = 8.63

-Tenure

Min = 0

Max = 10

Mean = 5.01

Std = 2.88

-Balance

Min = 0

Max = 250898.09

Mean = 76537.28

Std = 62419.71

-Estimated Salary

Min = 11.58

Max = 199992.48

Mean = 100140.94

Std = 57510.32

-Balance Salary Ratio

Min = 0
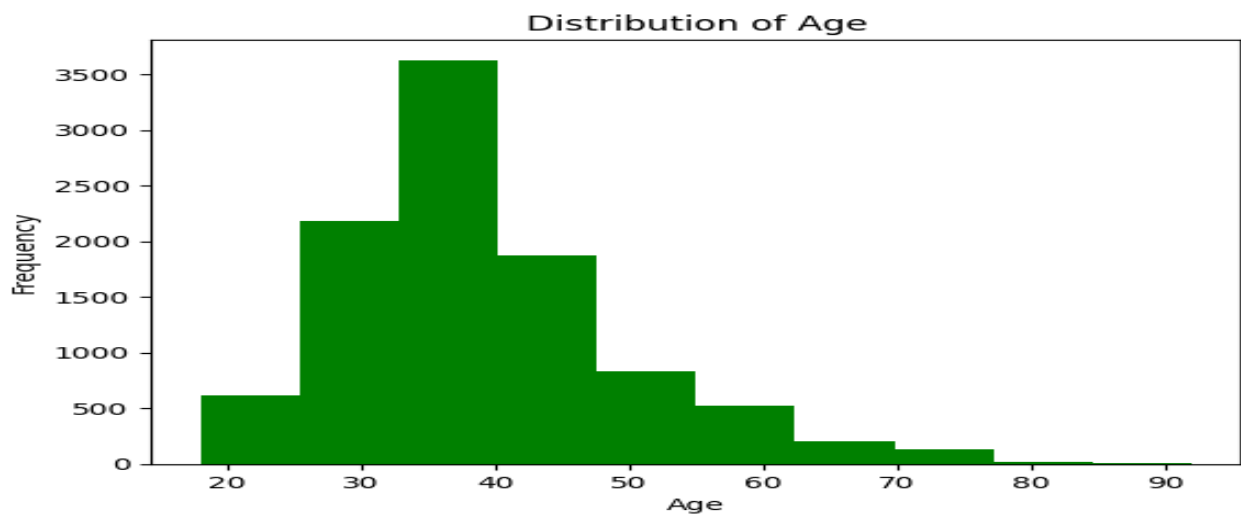
Max = 10614.65

Mean = 3.97

Std = 110.71

-Balance Credit Ratio

Min = 0

Max = 453.03

Mean = 120.23

Std = 101.09

- **Data Distribution**

-Age

Distribution of Age

-Country

Distribution of Country

-Balance

Distribution of Balance

-Credit score
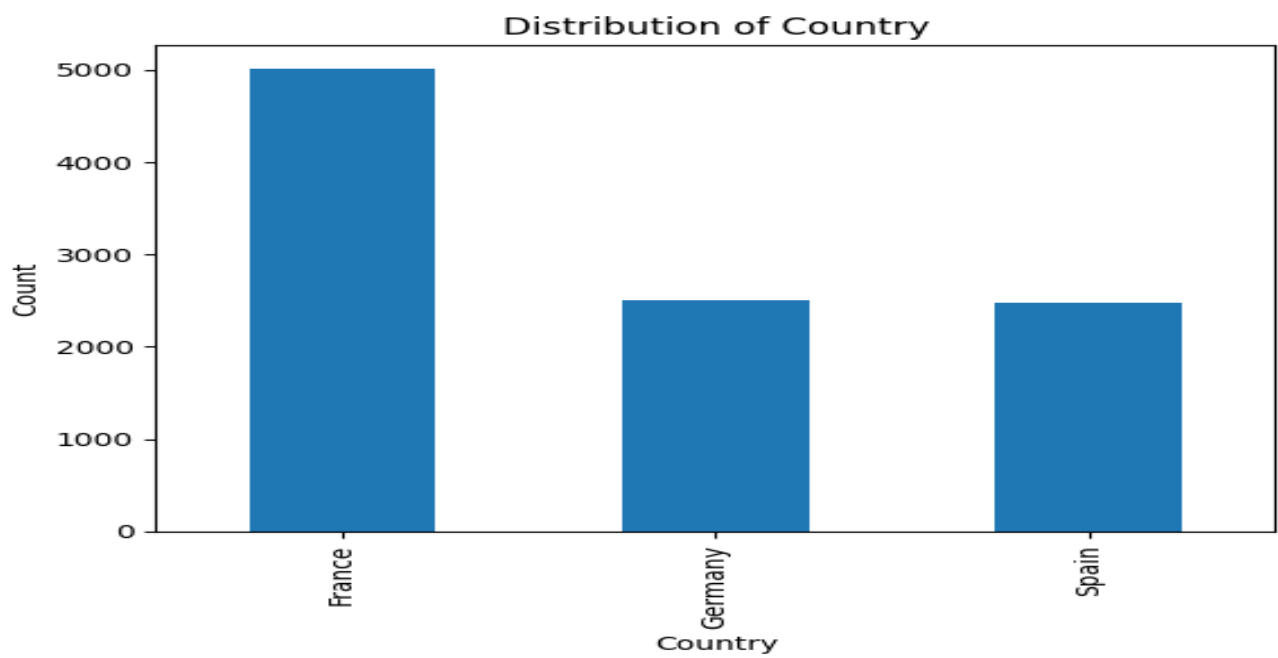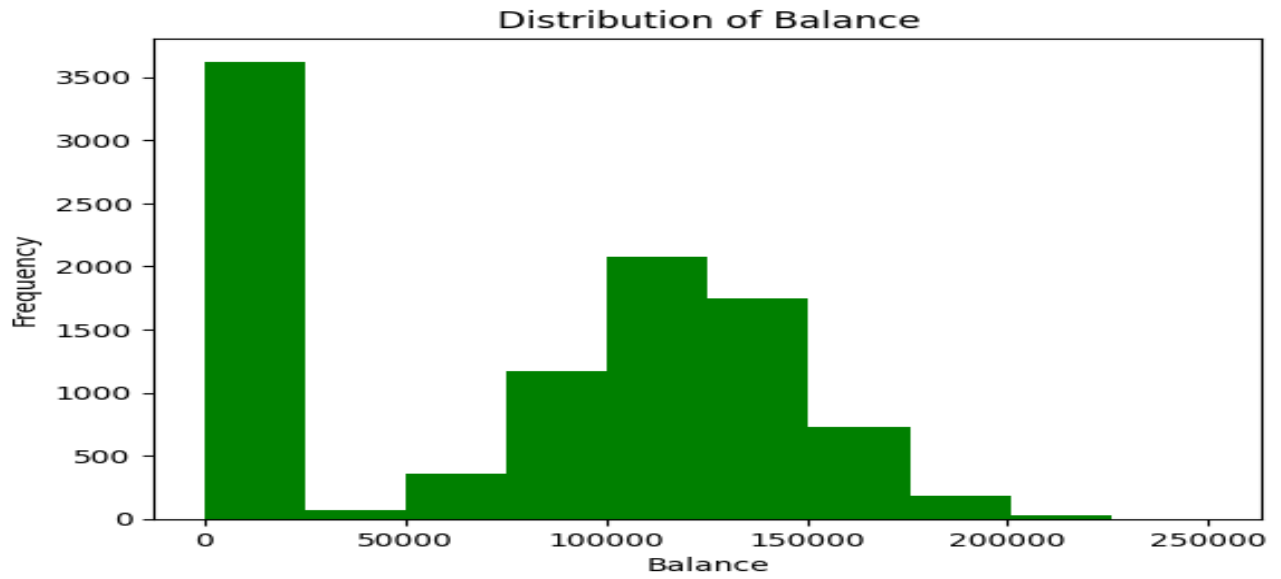
Distribution of Credit Score

-Estimated Salary
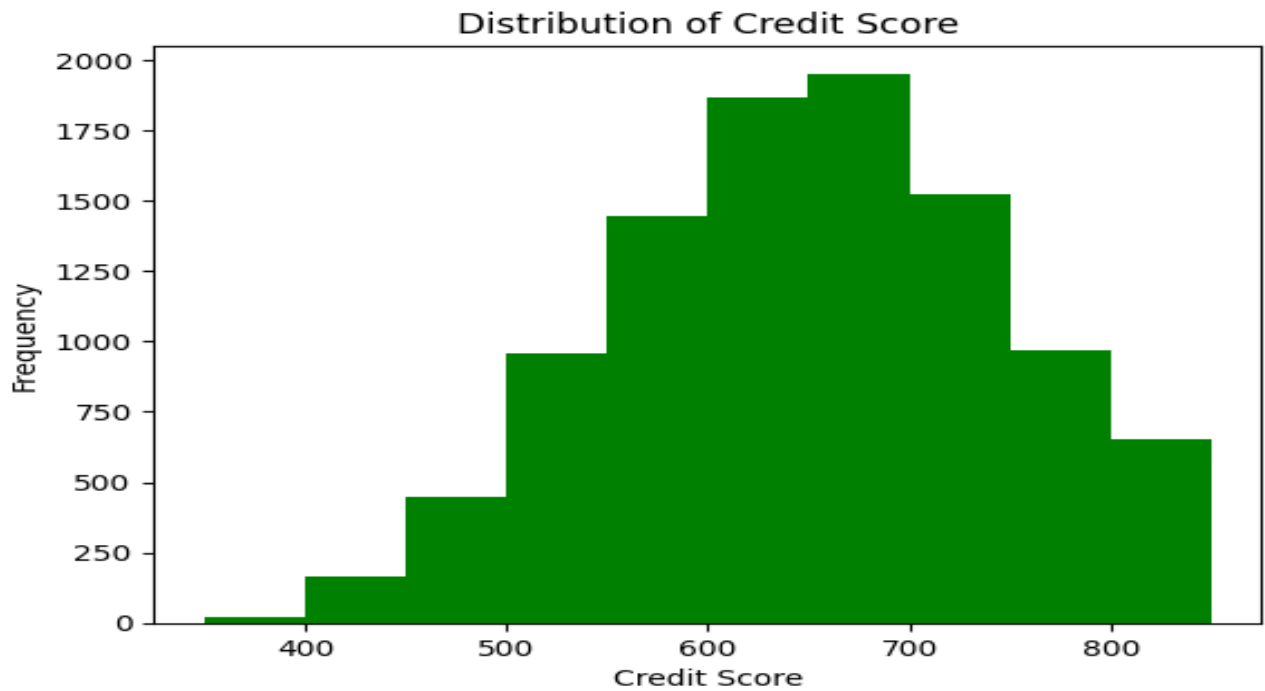
Box Plot of Estimated Salary

-Credit score
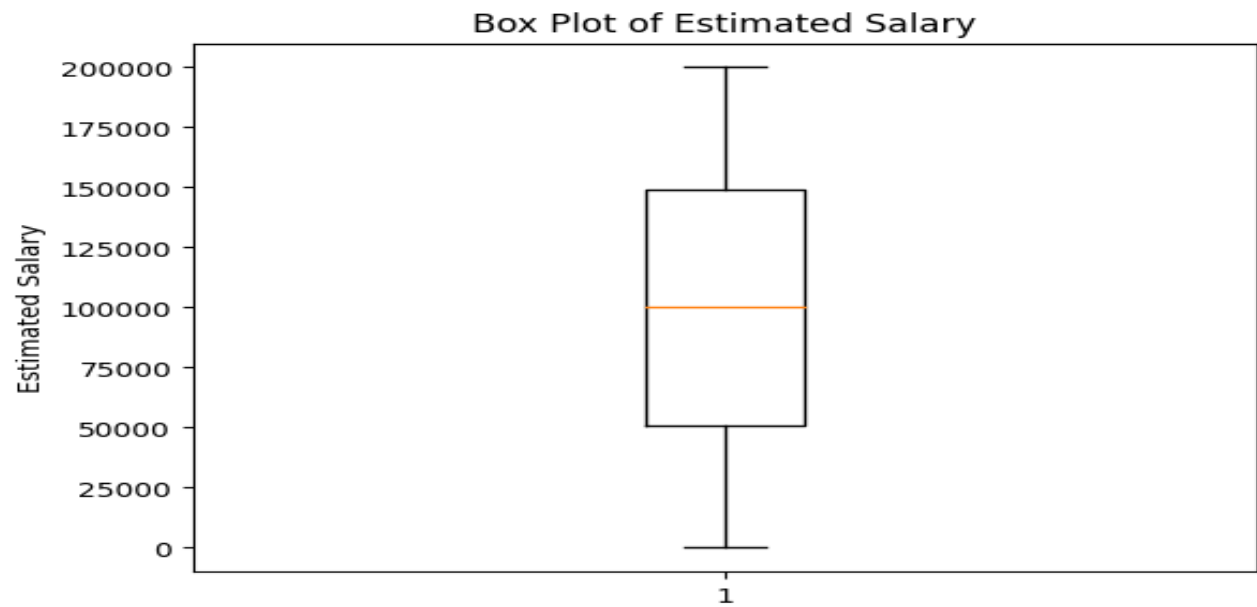
Box Plot of Credit Score
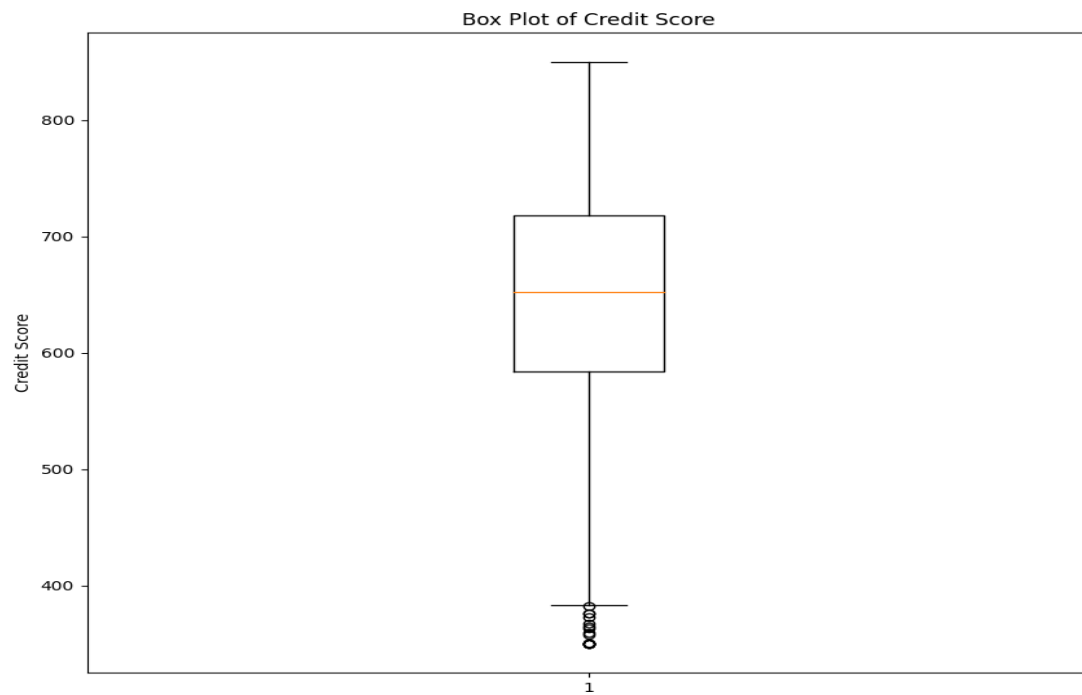
-Correlation Analysis

Correlation Matrix
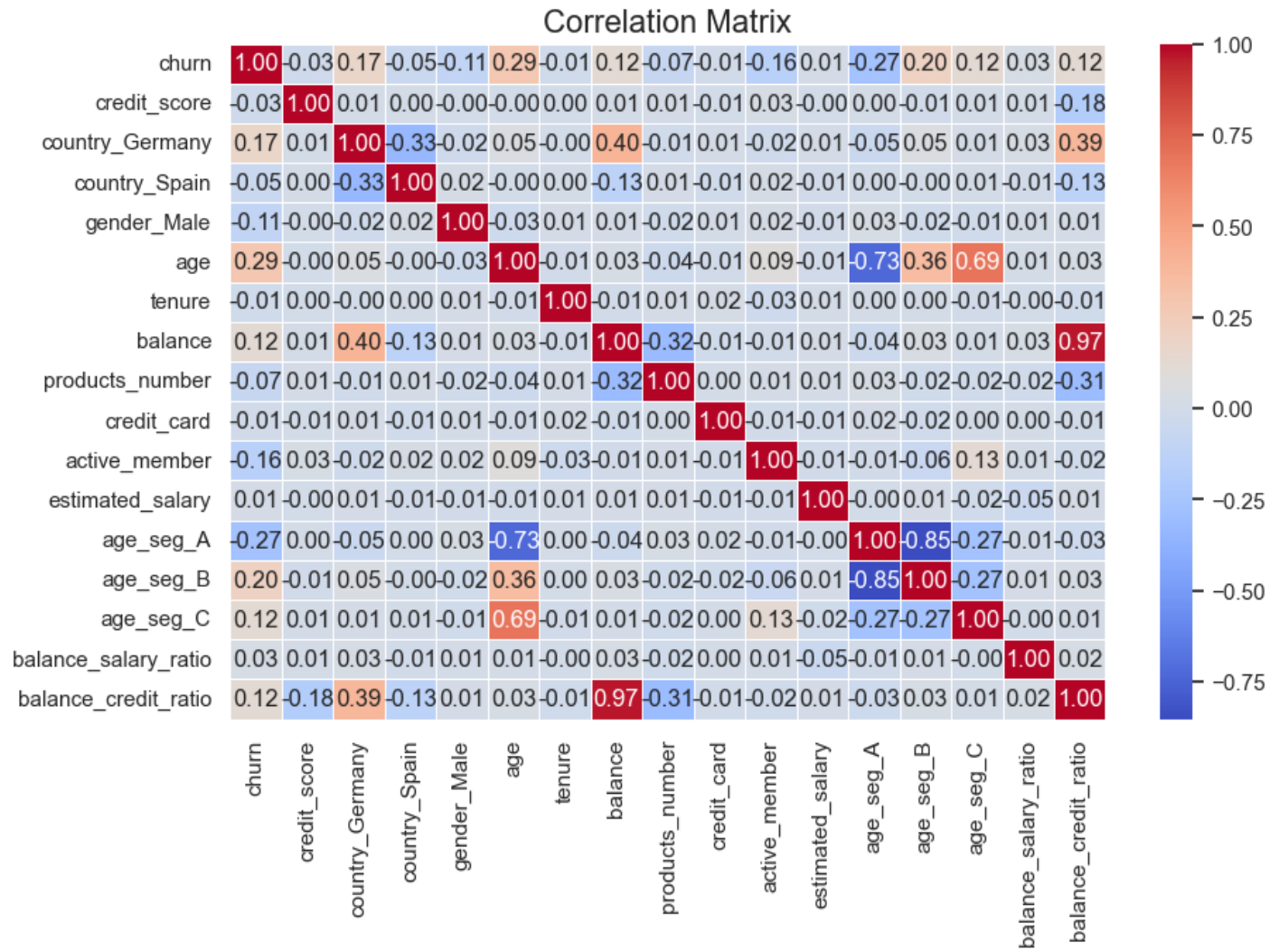
-Categorical Variable Analysis
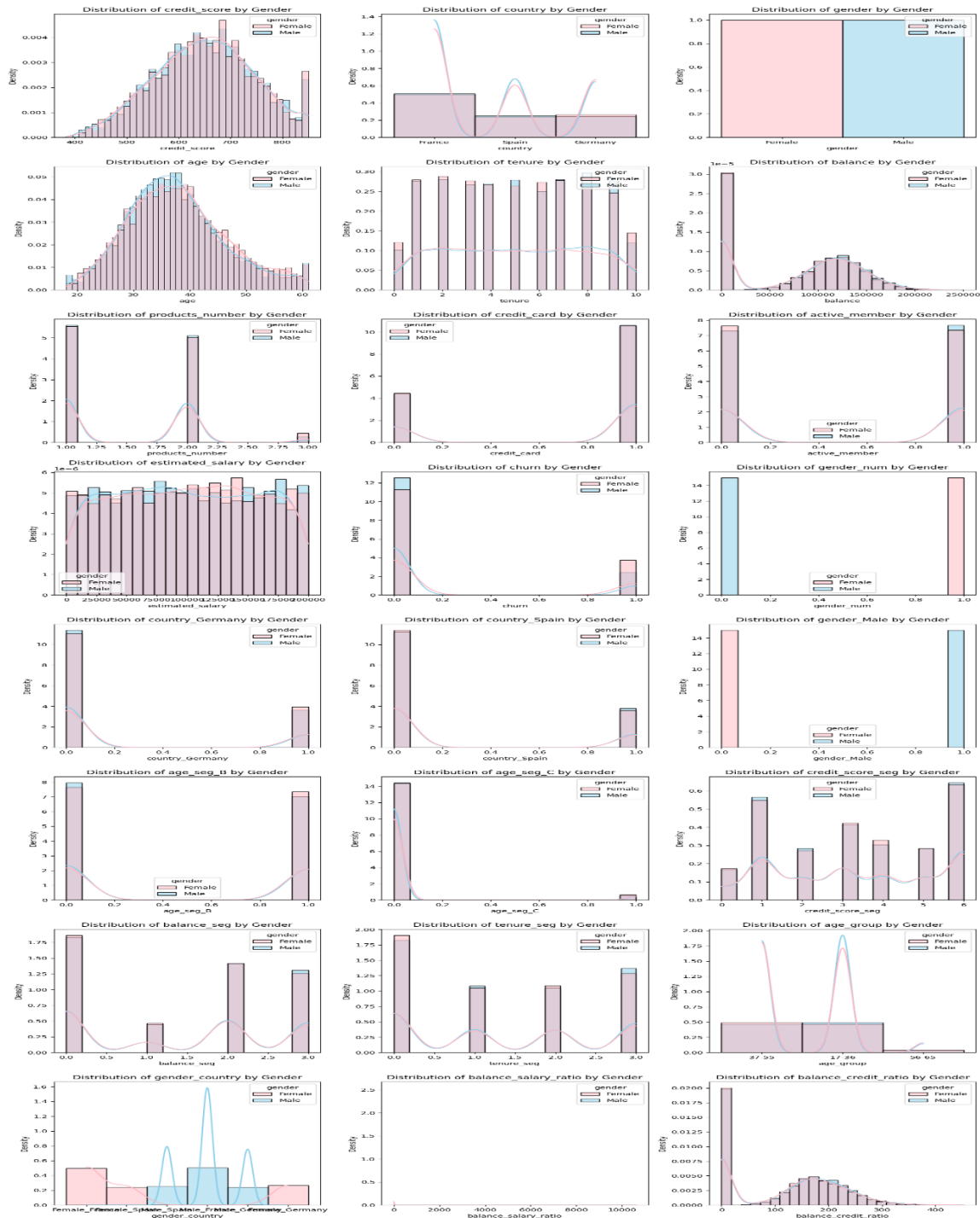
Visualization

- Visualization Techniques

-Pie chart

-Heatmap
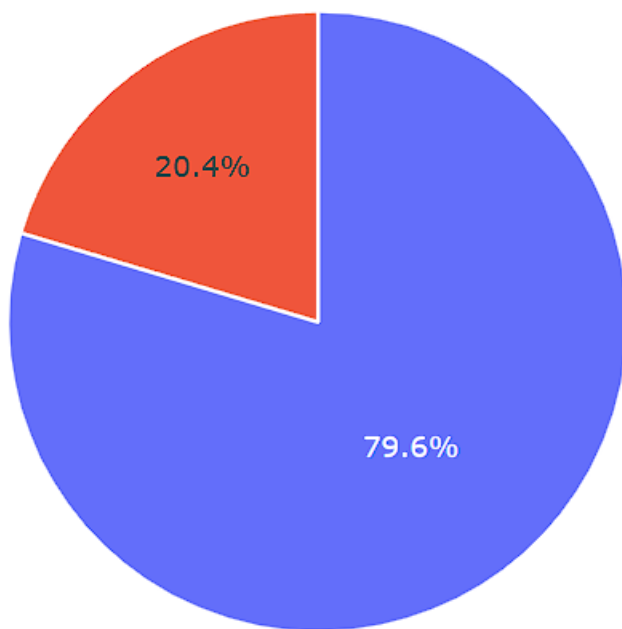
-Histogram

-Scatter plot

-Bar plot

-Box plot

-Plotly

-Folium

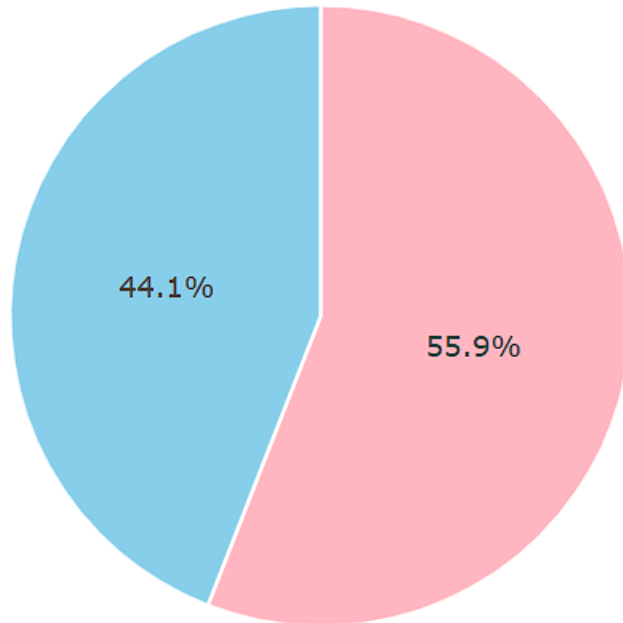- Key findings

-Churn Status for customers

Orange = Churn   Blue = non-churn

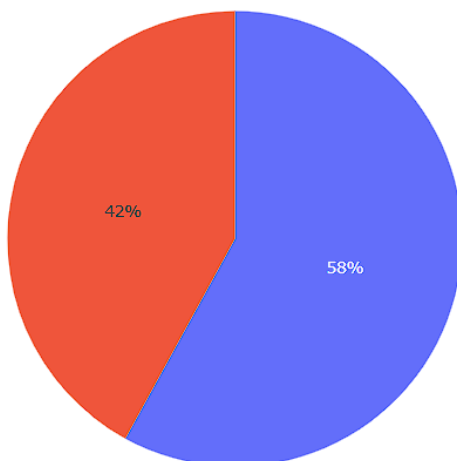-Gender Distribution for **Churn**
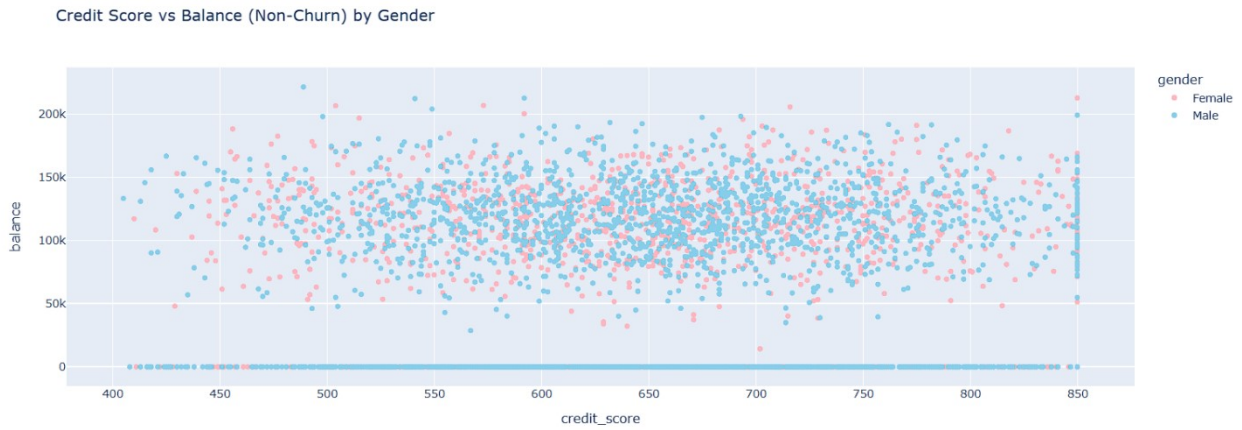
Blue = Male   Pink = Female

-Age (30-40) Gender Distribution for **Churn**

Blue = Male   Orange = Female

-Credit Score vs Balance (Non-Churn) by Gender

Credit Score vs Balance (Non-Churn) by Gender
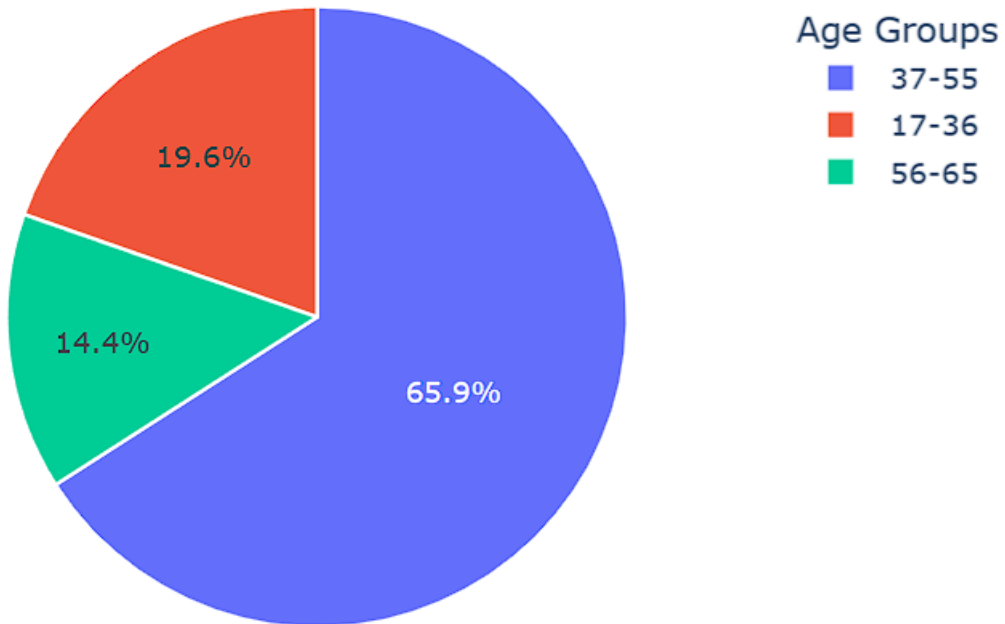
-Estimated Salary vs Balance (Non-Churn) by Gender

Estimated Salary vs Balance (Non-Churn) by Gender

-Average Estimated Salary by Country (Aged 30-40, Non-Churn)

Average Estimated Salary by Country (Aged 30-40, Non-Churn)
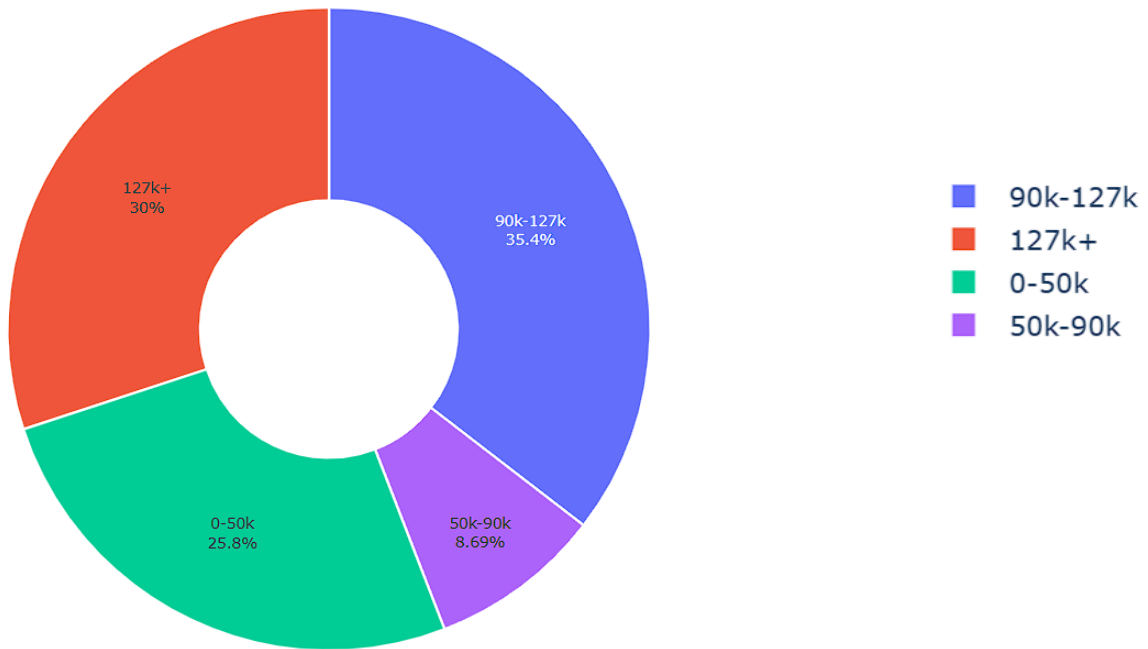
-Average balance by Country (Aged 30-40, Non-Churn)
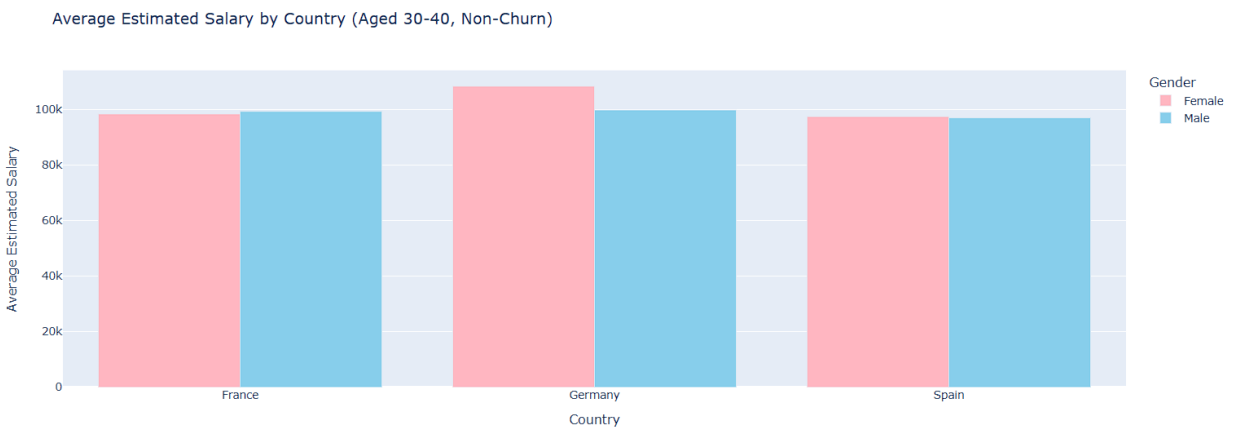
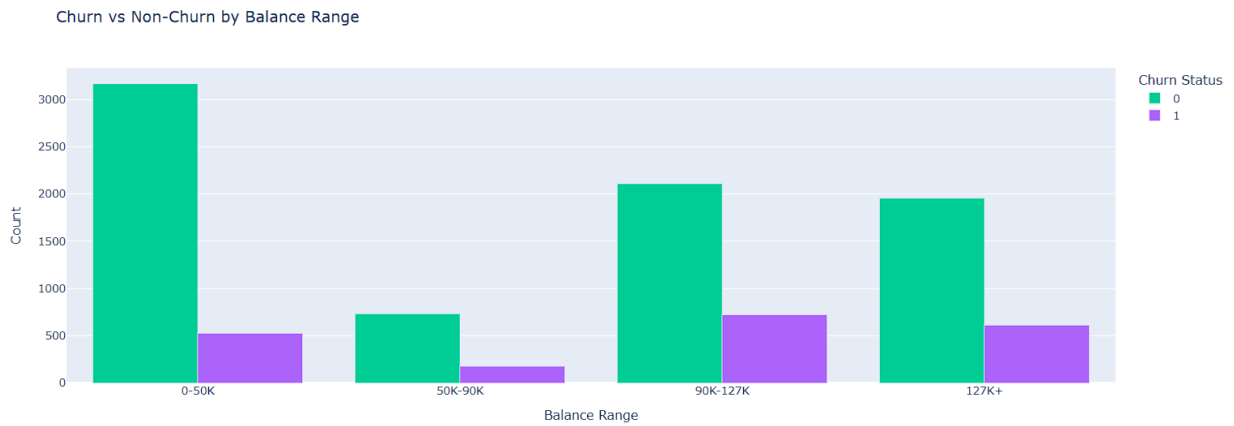Average balance by Country (Aged 30-40, Non-Churn)
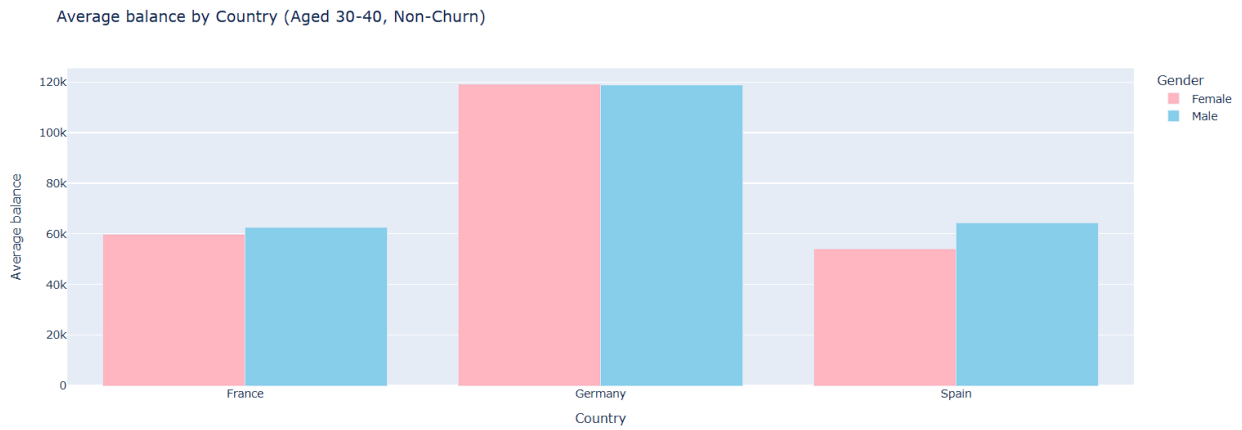
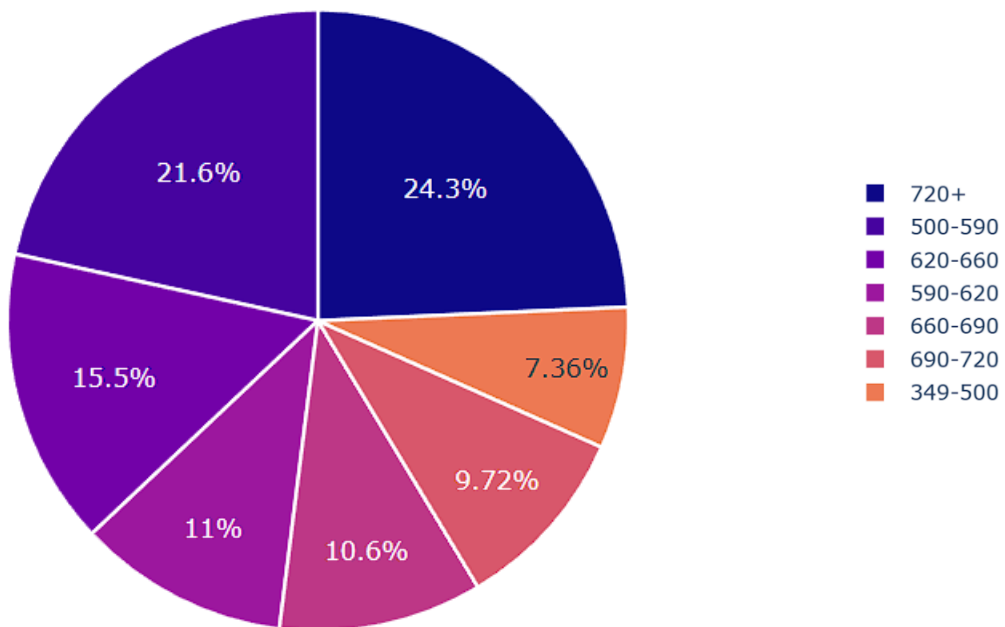-Percentage of Churned Customers by Age Groups

-Churned Customers Distribution by Balance Segmentation

-Churn vs Non-Churn by Balance Range

-Credit Score Distribution for Churned Customers
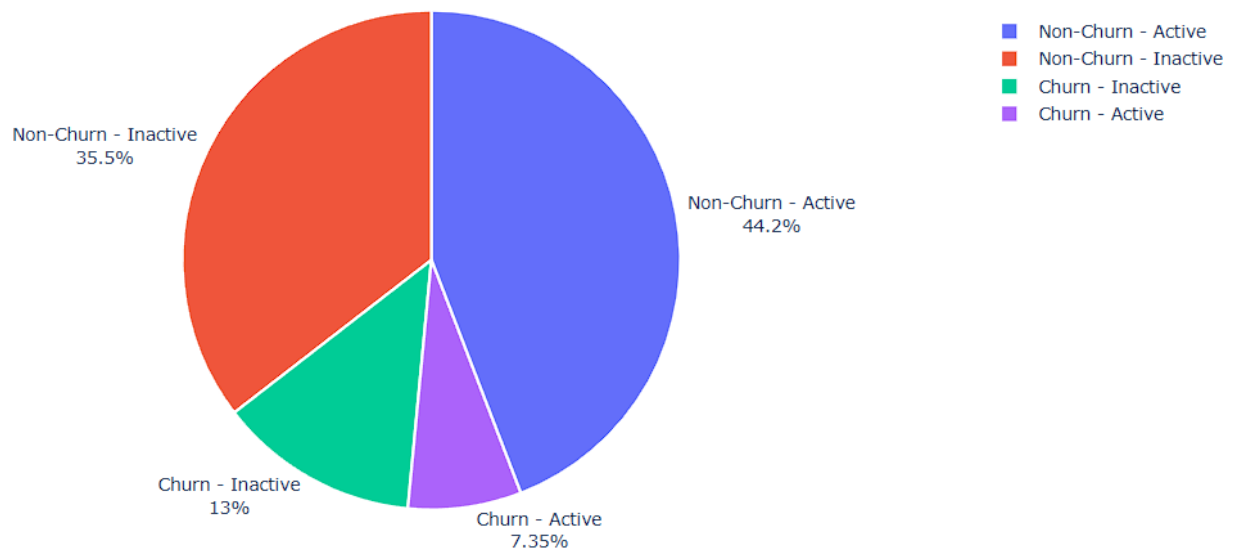

Legend:
- 720+
- 500-590
- 620-660
- 590-620
- 660-690
- 690-720
- 349-500

-Churn Status Distribution by Credit Score

Legend:
- 0
- 1
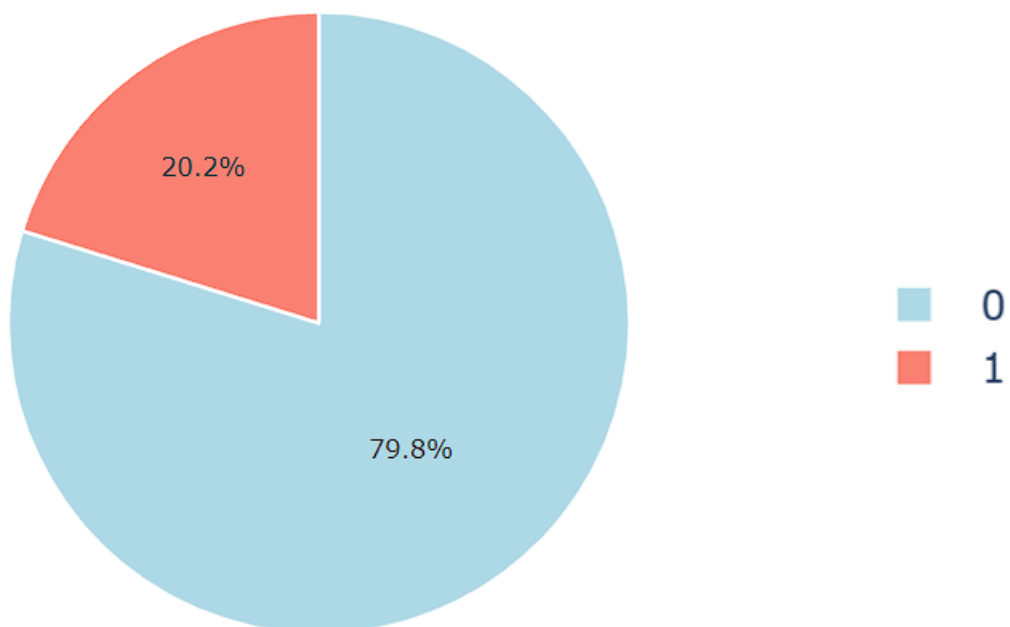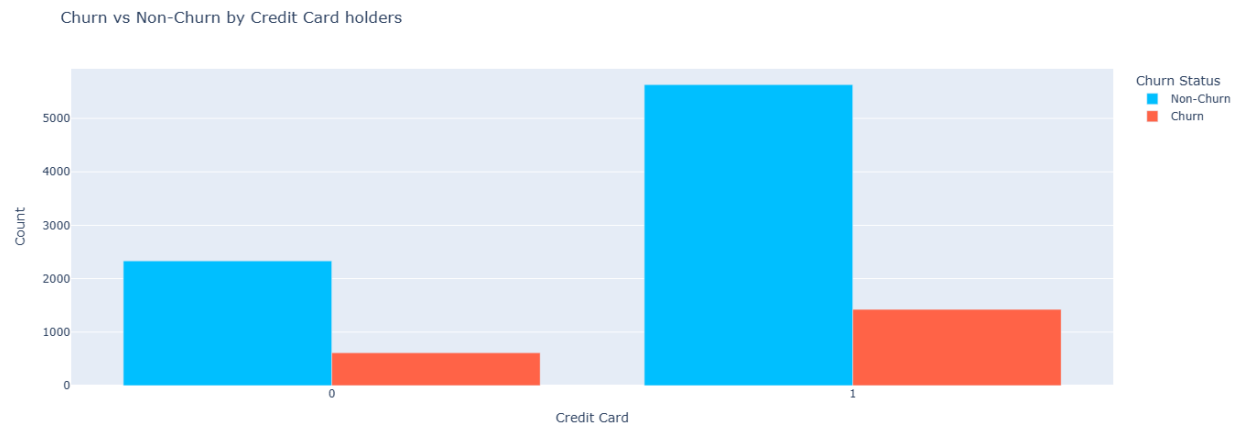
-Churn vs Non-Churn by Active Membership

-Percentage of Customers by Products Number

-Churn vs Non-Churn by Credit Card holders
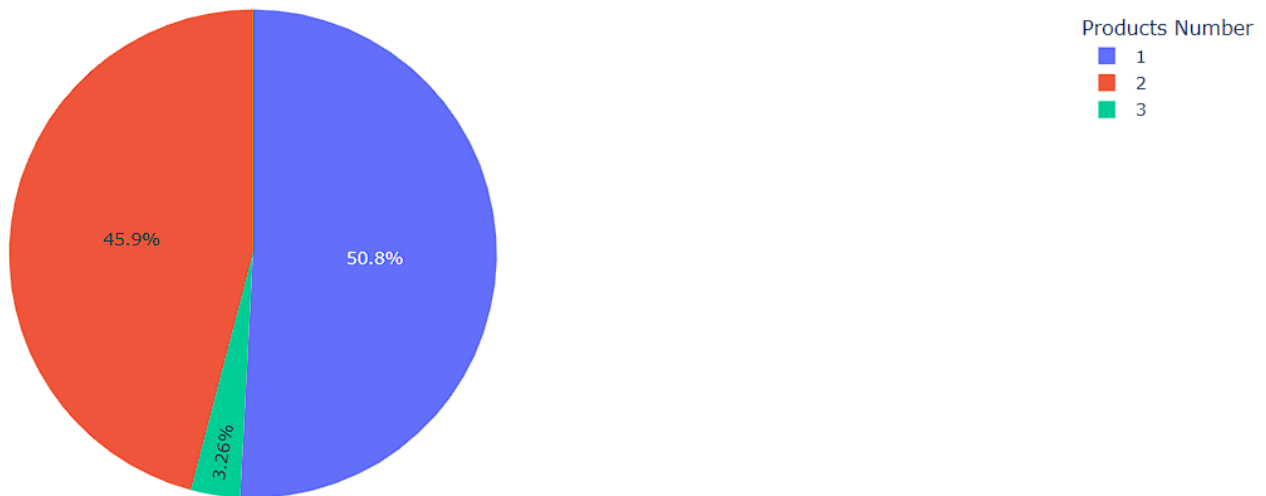
Churn vs Non-Churn by Credit Card holders

-Tenure Subscribed for Customers by Churn Status

Tenure Subscribed for Customers by Churn Status

-Customer Count by Country and Churn Status
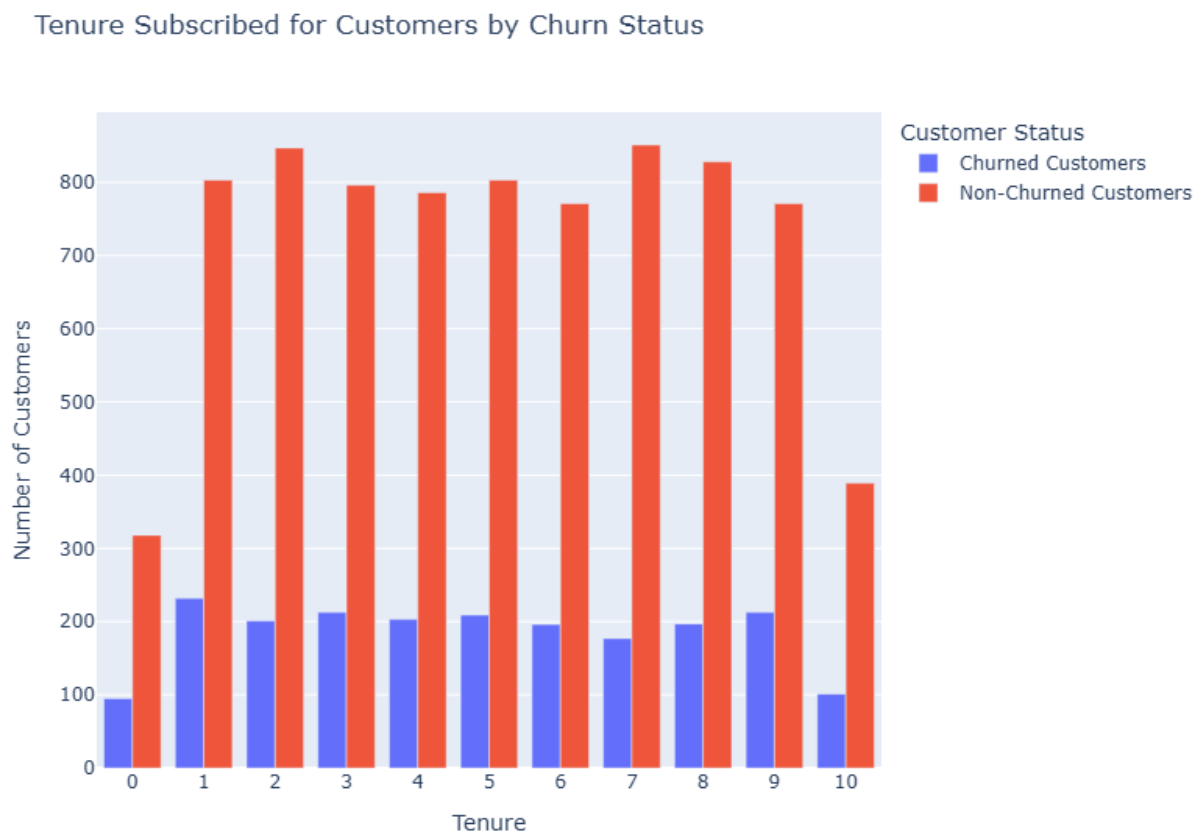
Customer Count by Country and Churn Status
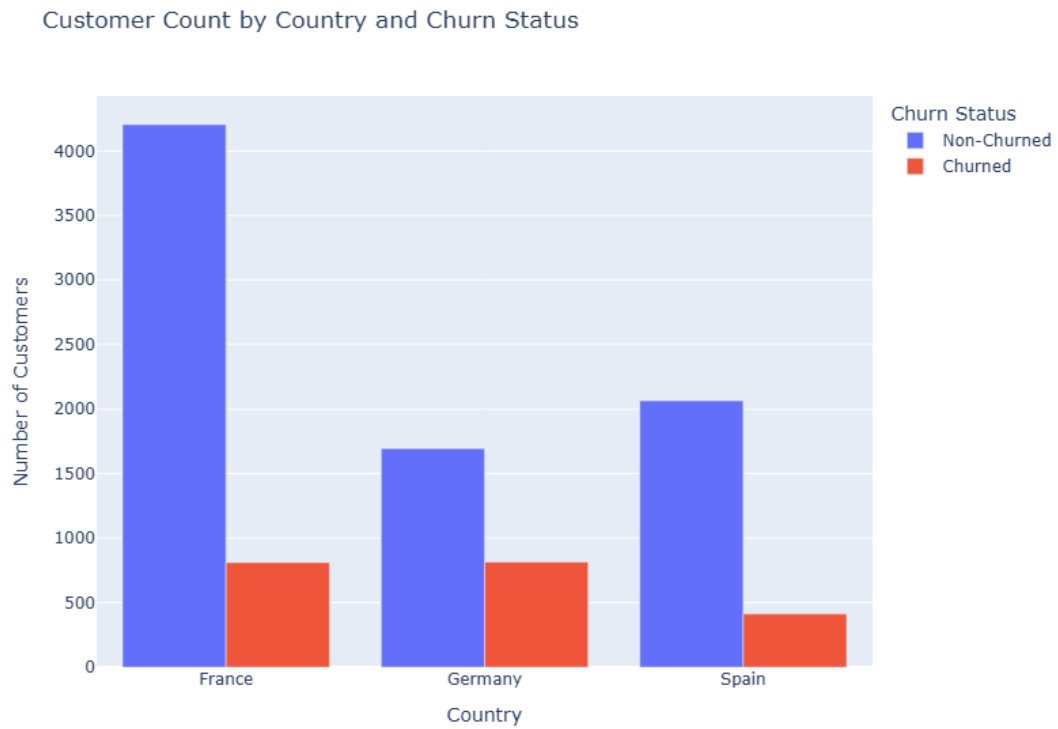
Conclusions & Recommendations

- **Summary of findings**

We studied and analyzed this dataset to understand the bank's customer base and predict customer churn. Customer churn is an important topic that banks must understand and address in order to retain customers and minimize customer loss. Analysis of the dataset revealed that 20.4% of customers identified as likely to churn are in the age group of 37 to 55 years old. These customers predominantly carry credit cards and reside in France and Germany. Their bank account balances range from 90k to 127k, and they have a high credit score. It is worth noting that these customers maintained their accounts with the bank for about a year.

These findings underscore the importance of understanding customer demographics, financial behaviors, and tenure in predicting customer churn. By targeting customers within these specific segments, a bank can implement strategies to enhance customer retention and minimize customer loss.

- **Recommendations**

Understanding the Churn-Prone Customer Profile Based on the analysis, the bank can identify the key characteristics of customers at risk of churning:

-Demographic: Age 37-55, residing in France and Germany

-Financial Behavior: Credit card holders, bank account balances between 90K and 127K, high credit score

-Tenure: Approximately one year with the bank

Targeted Retention Strategies

1-Personalized Offers and Promotions:

Tailor offers and promotions to the specific needs and preferences of this customer segment. For example, exclusive discounts on products or services that align with their financial goals or lifestyle.

2-Enhanced Customer Communication:

Proactively reach out to customers in this segment to address any concerns or questions they may have. Personalize communication efforts by using their names and referring to their specific account information.

3-Loyalty Programs:

Implement loyalty programs that reward customers for their continued business. Consider offering tiered rewards based on customer tenure and spending.

4-Product and Service Enhancements:

Review the bank's product and service offerings to ensure they meet the needs of this customer segment. Consider introducing new products or features that cater specifically to their preferences.

5-Customer Feedback Mechanisms:

Actively seek customer feedback through surveys, focus groups, or one-on-one interactions. Use customer feedback to identify areas for improvement and address any pain points.

Employee Training:

Train bank employees to recognize and address the needs of customers in this segment. Equip them with the knowledge and skills to provide exceptional customer service.

Continuous Monitoring and Analysis:

1-Regular Data Updates:

Ensure that the customer data is regularly updated to reflect changes in customer behavior and preferences.

2-Ongoing Analysis:

Continuously analyze customer data to identify emerging trends and patterns that may impact churn. Refinement of Strategies: Adjust retention strategies based on the latest insights and data analysis.

By implementing these recommendations, the bank can significantly improve customer retention, reduce churn rates, and strengthen its overall customer base.