



Data Analytics Bootcamp - CAPSTONE PROJECT

Farzana M Kadir

“

**The goal is to turn data into
information and information into
insight.**

Carly Fiorina, Former CEO, Hewlett Packard

Table of Contents

- ❏ BUSINESS PROBLEM
- ❏ DATASET PREPARATION
- ❏ DATA ANALYSIS - TABLEAU
- ❏ INSIGHTS AND RECOMMENDATIONS

Business Problem



Business Problem

A bank manager is concerned about more customers closing their accounts. Using the available data, we want to identify customers at high risk of leaving, so the bank can take steps to keep them.

Dataset Preparation



Dataset Preparation

It contains data of 10,000 bank accounts with the following 14 columns.

Datasheet Columns

- | | |
|---------------|-------------------|
| ❑ RowNumber | ❑ Tenure |
| ❑ CustomerId | ❑ Balance |
| ❑ Surname | ❑ NumOfProducts |
| ❑ CreditScore | ❑ HasCrCard |
| ❑ Geography | ❑ IsActiveMember |
| ❑ Gender | ❑ EstimatedSalary |
| ❑ Age | ❑ Exited |

Data Source: <https://www.kaggle.com/datasets/santoshd3/bank-customers/data>

Dataset Preparation

Variable

Description

❑ RowNumber	Line item number
❑ CustomerId	Customer ID is the unique identifier for the customer holding the bank account
❑ Surname	Surname of the customer
❑ CreditScore	Credit score assigned to the customer based on his creditworthiness (350 to 850)
❑ Geography	Demographic variable - Spain, France, Germany
❑ Gender	Demographic variable - Female, Male
❑ Age	Demographic variable - Customer's age in Years
❑ Tenure	Number of years the customer has been with the bank
❑ Balance	Amount of balance held in the bank by the customer
❑ NumOfProducts	Number of products (cards, etc) held by the customer - 1, 2, 3, 4
❑ HasCrCard	Whether the customer holds a credit card - "1" is Yes, "0" is No
❑ IsActiveMember	Whether the customer makes credit card transactions - "1" is Yes, "0" is No
❑ EstimatedSalary	Demographic variable - Customer's monthly salary
❑ Exited	Whether the customer withdrew his account - "1" is Yes, "0" is No

Dataset Preparation

- ❏ After getting the dataset, I first checked the data to see if there is any error in data values (such as typo error in words, etc). Data was mostly accurate (i.e., no blanks, no duplicates in the Customer ID). Minor discrepancies (e.g., Surname of customer with ID 15592389 has “?” in it) does not have any impact on the analysis performed.
- ❏ Next, I imported the data to Tableau to create and organize the data for better data analysis and visualization.

Data Analysis - Tableau

3



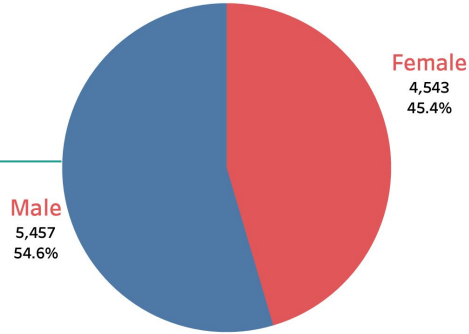
Customer Demographics

Understanding customers demographics may reveal valuable insights on retention trends.

Customers' Demographics - Part 1

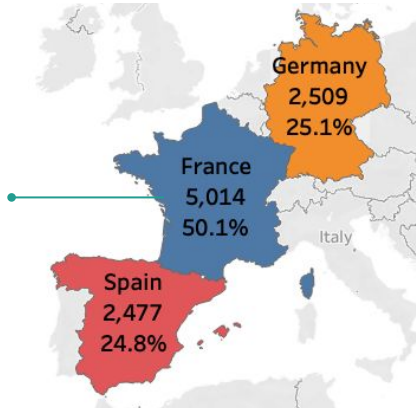
Gender

Males make up more than half of the customer base.



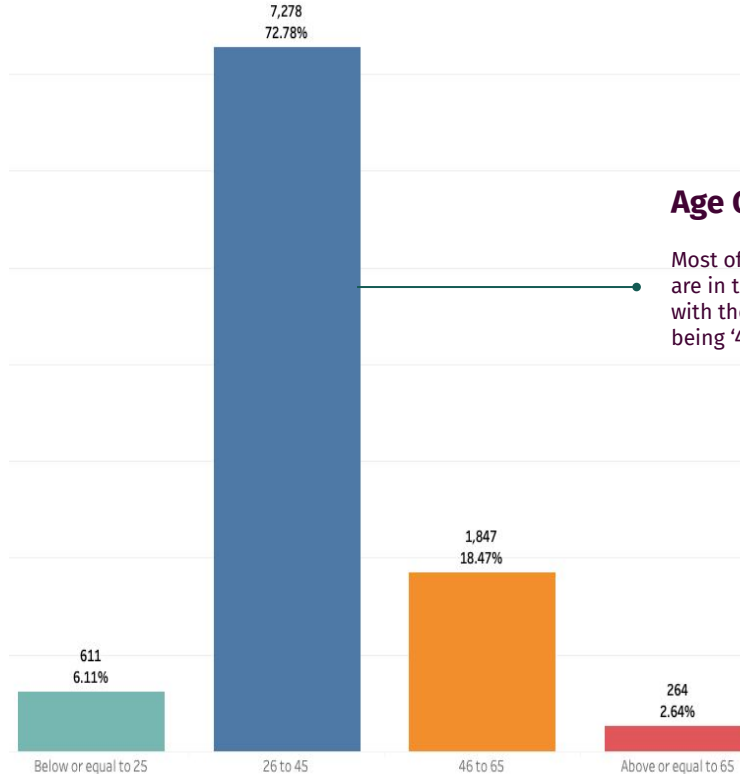
Geography

France makes up more than half of the customer base. Spain and Germany have been split relatively equally.



Age Group

Most of the customers (> 70%) are in the '26 to 45' age group, with the second largest group being '46 to 55'.



Customers' Demographics - Part 2

Gender vs Age, Geography & Estimated Salary



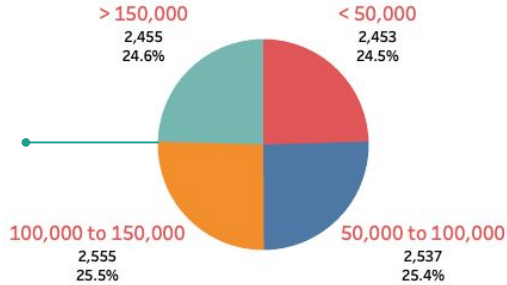
Observations

The divide between gender, estimated salary and age are quite consistent in each country - France, Germany and Spain.

Customers' Demographics - Part 3

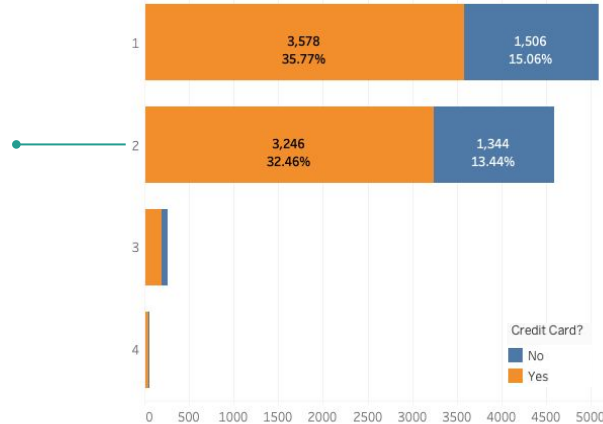
Estimated Salary

Estimated Salary has been categorized into 4 categories. The customer base is evenly split among the different estimated salary bands.

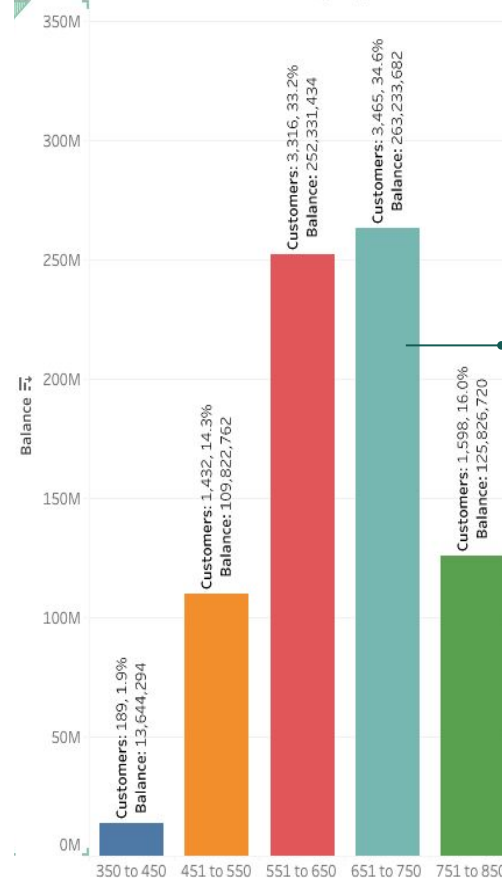


Products

Customers hold 1 to 4 products - majority hold only 1 product (out of which 36% are credit card holders).



Credit Score (group)



Credit Score

Credit Score has been categorized into 5 categories. The higher the credit score, the higher the creditworthiness of the customer. We see that the majority (>50%) of the customers have credit score of 651 and higher. This might indicate the bank's risk-averse approach in engaging in more creditworthy customers.



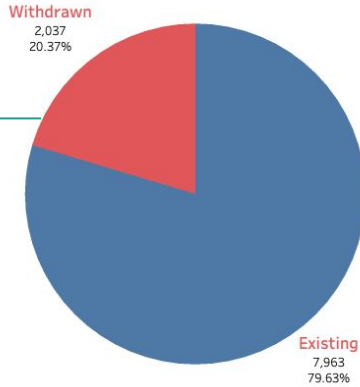
Customer Withdrawals

Deeper analysis into the trends in the
withdrawals

Withdrawal Trends - Part 1

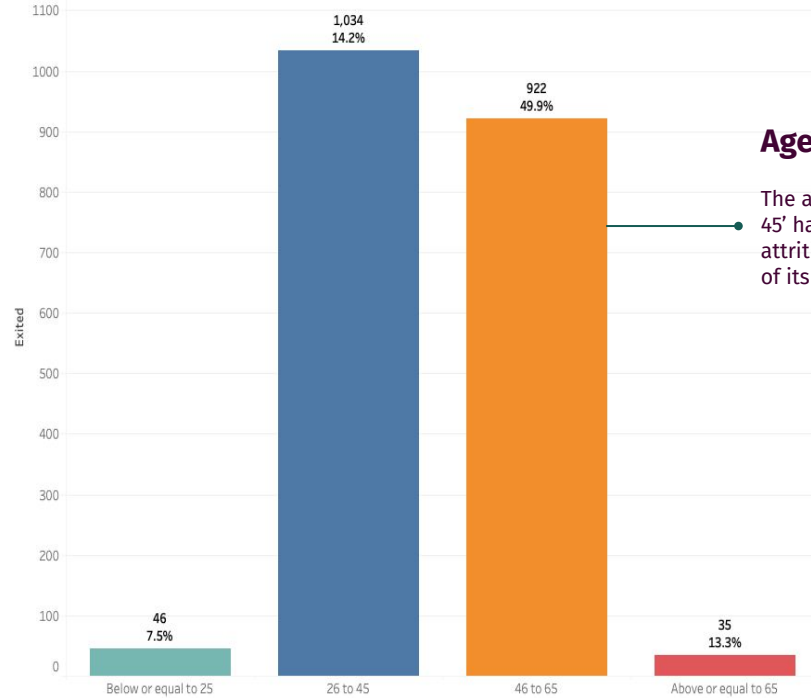
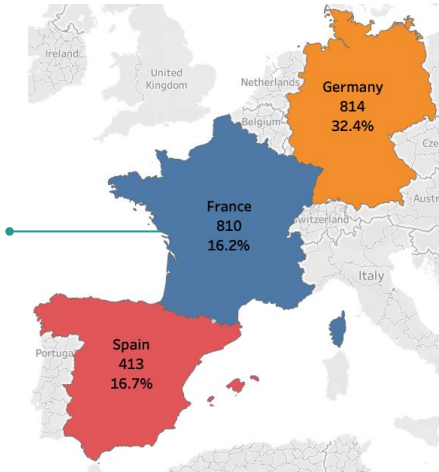
Withdrawals

20% of the customers have withdrawn from the bank. We need to understand the reason as this is a high number.



Geography

I calculated the withdrawals as a % of the total customers in each respective country. Germany has the highest attrition of 32.4% (as a % of its customer base).



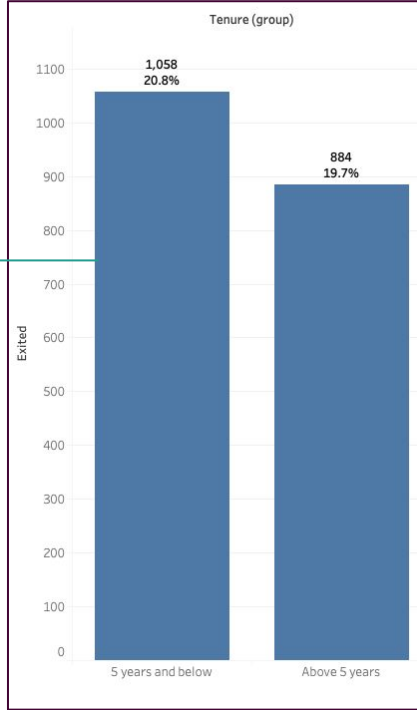
Age Group

The age group '26 to 45' has the highest attrition (49.9%) as a % of its customer base.

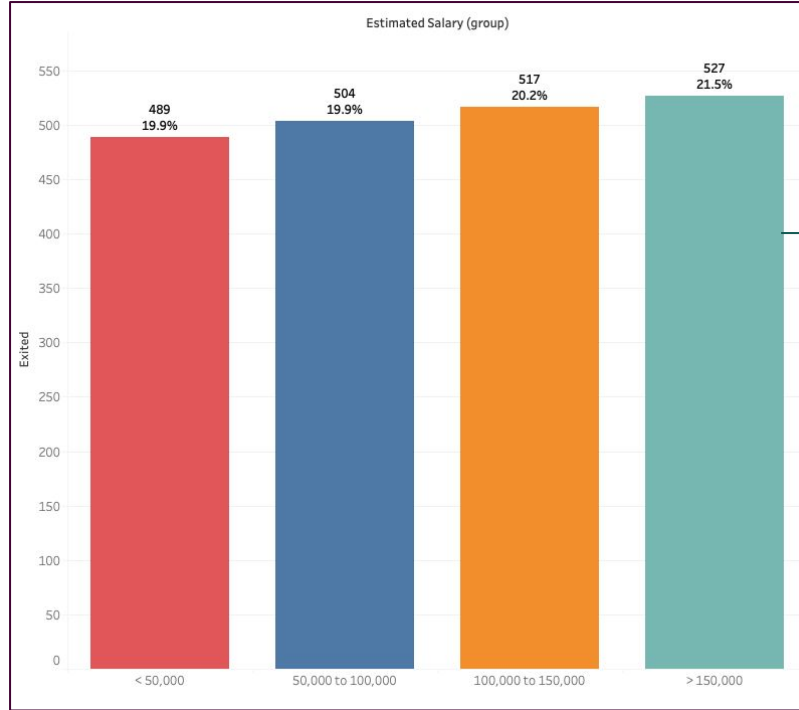
Withdrawal Trends - Part 2

Tenure

Tenure has been divided into 2 categories. No actual trend is observable based on the tenure.



Estimated Salary (group)



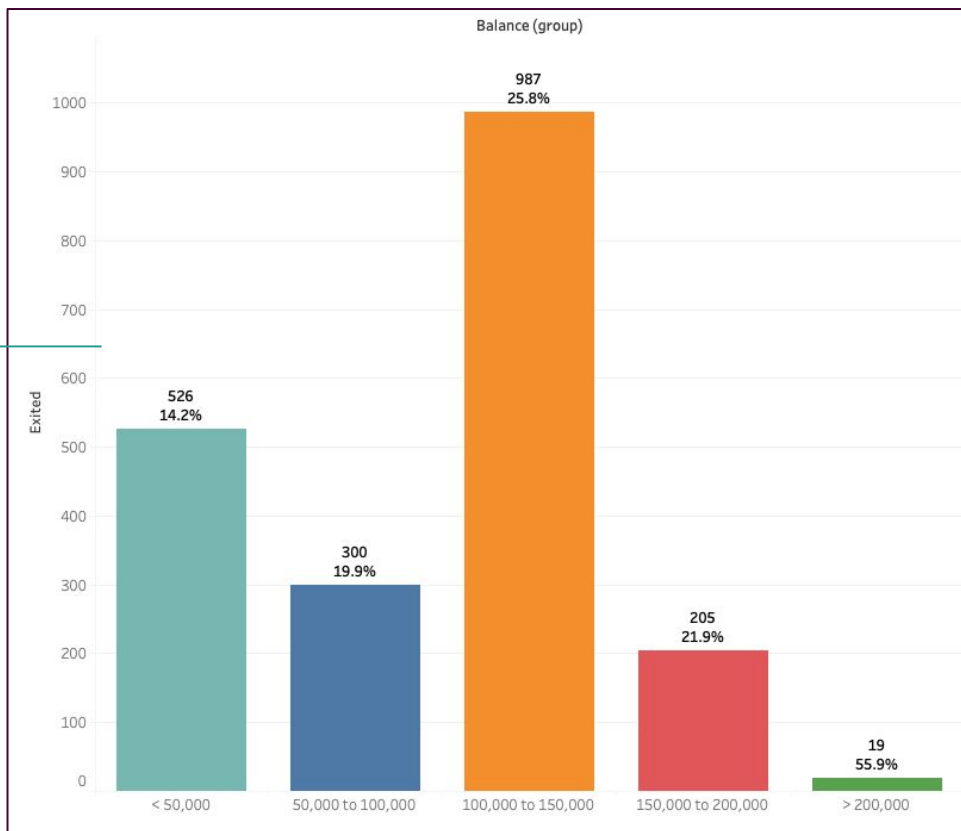
Estimated Salary

No actual trend in estimated salary is observable.

Withdrawal Trends - Part 3

Account Balance

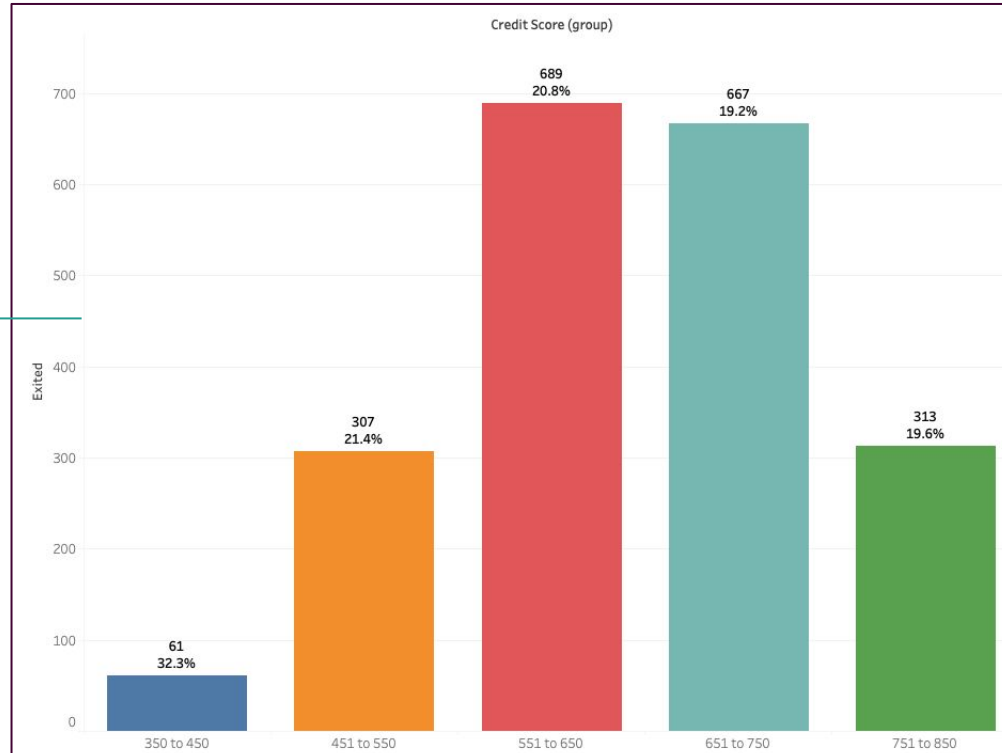
Account balance has been divided into 5 categories. The '>200,000' category has the highest withdrawals (55.9%) as a % of its customer base.



Withdrawal Trends - Part 4

Credit Score

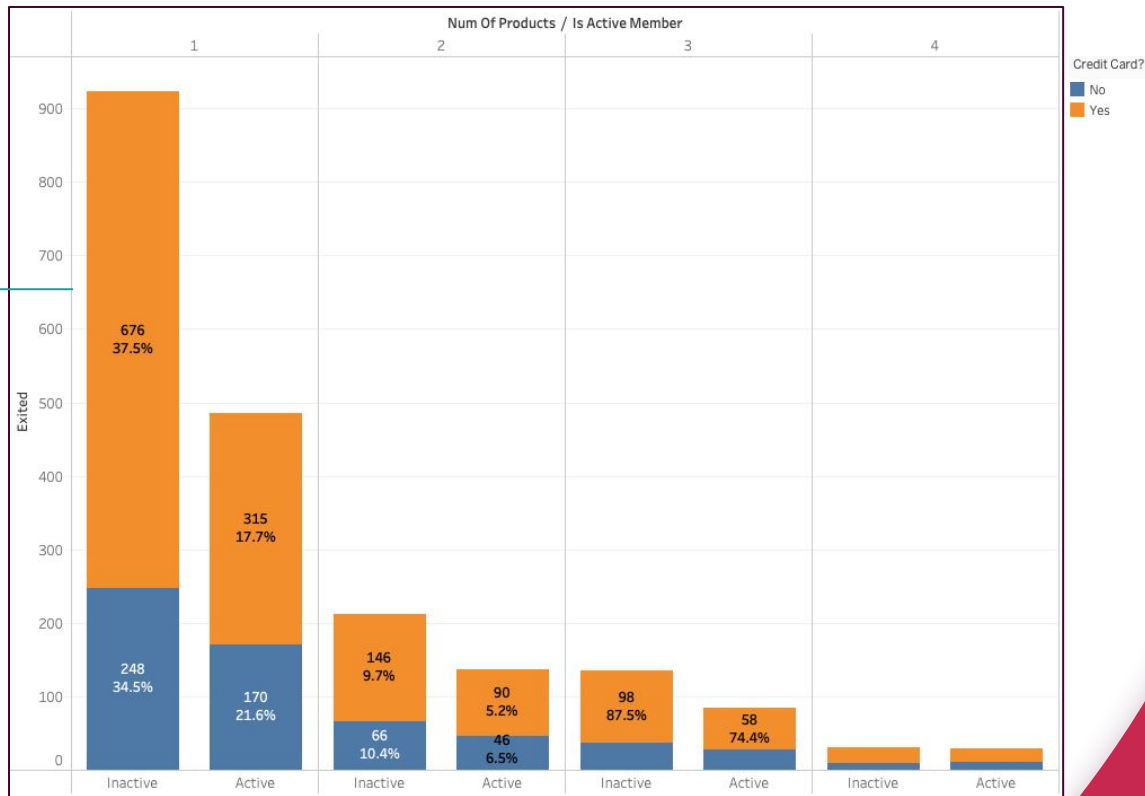
The customers having credit scores between 350 and 450 are the ones with highest withdrawals (32.3%) as a % of their customer base.



Withdrawal Trends - Part 5

Product Type

Based on the number of products held by the customers, whether they held a credit card and whether they had active transactions using the bank's products, I was able to observe that the highest withdrawals (37.5% for customers with credit card, 34.5% for customers without credit card) were for customers with 1 product and inactive transactions.



Insights & Recommendations



Insights & Recommendations - Geography

Insight

Germany has the highest withdrawals.

Recommendations

The manager should try to understand the reason for this. Some factors that he can consider are country-specific economic conditions, marketing campaigns and products offered.

Insights & Recommendations - Age

Insight

The '46 to 65' age group has the highest attrition.

Recommendations

The manager should try to understand the reason for this. Some factors that he can consider are suitability of products offered to this age group and recent government policies affecting the group.

Insights & Recommendations - Balance

Insight

The customers with account balance more than 200,000 have the highest withdrawals.

Recommendations

Given that withdrawals of huge amounts could lead to liquidity issues, the manager should try to address this issue by considering factors such as a better product offering by competitors.

Insights & Recommendations - Credit Score

Insight

Customers with 1 product and inactive transactions have the highest withdrawals, regardless of whether they hold a credit card or not.

Recommendations

The manager should try to understand why the customer is not actively using the bank's products. He can consider factors such as tech-savviness and convenience of the products. He can also try to offer more products to the customers.

Insights & Recommendations - Product

Insight

Customers with credit scores between 350 and 450 have the highest withdrawals.

Recommendations

This need not be worrying, as the bank is losing the less creditworthy customers. Nevertheless, the manager should consider factors such as stringency of the credit scoring methods.



\$185,588,095

That's a lot of money withdrawn

2,037 customers

And a lot of customers lost

Zero customer withdrawals

Target!

Final thoughts..

With this initial set of data analysis, the manager should deep dive to understand the reasons for high withdrawals in certain customer segments and implement measures to retain the customers.



Thank you!

Presentation by Farzana M Kadir

Special thanks to the instructors (Weizhi & Jordan) for their guidance and everyone who released these awesome resources for free:

- Presentation template by [SlidesCarnival](#)
- Photographs by [Unsplash](#)