QUANTIFYING CHURN RISK IN BANKING: A COMPARATIVE STUDY OF PREDICTIVE MODELS

Raheela Charania

Anmol Anchala

Emmanuel Wediko

8/1/2025

BACKGROUND

- What is customer churn?

- Why does it matter in banking?

- Insight from 2019

- Solution Approach

8/1/2025

BUSINESS PROBLEM

- Improve customer retention, financial performance, and bank reputation

- Develop a predictive model to identify banking customers at risk of churning

- Reduce customer churn rate by offering targeted incentives

8/1/2025

DATA DICTIONARY

Exited
- Whether or not the customer left the bank. 0 = customer stays. 1 = customer leaves.

Points Earned
- The points earned by the customer for using credit card

CreditScore
- Can have an effect on customer churn, since a customer with a higher credit score is less likely to leave the bank.

Balance
- Also a very good indicator of customer churn, as people with a higher balance in their accounts are less likely to leave the bank compared to those with lower balances.

Geography
- A customer's location can affect their decision to leave the bank.

Gender
- It's interesting to explore whether gender plays a role in a customer leaving the bank.

NumOfProducts
- Refers to the # of products that a customer has purchased through the bank.

8/1/2025

EDA : DESCRIPTIVE STATS

	Mean	Standard Deviation
⭐ Age	37.74	8.78
Points Earned	605.92	225.87
⭐ Balance	76426.07	62417.82
Credit Score	650.66	96.13
⭐ NumOfProduct	1.52	0.54

8/1/2025

EDA: DESCRIPTIVE STATS

Gender	Has Credit Card	Card Type	Exited
• Male: 55% • Female: 45%	• No (0): 29% • Yes (1): 71%	• Diamond: 25% • Gold: 25% • Platinum: 25% • Silver: 25%	• No (0): 80% • Yes (1): 20%

8/1/2025

EDA: EXPLORATORY STATS
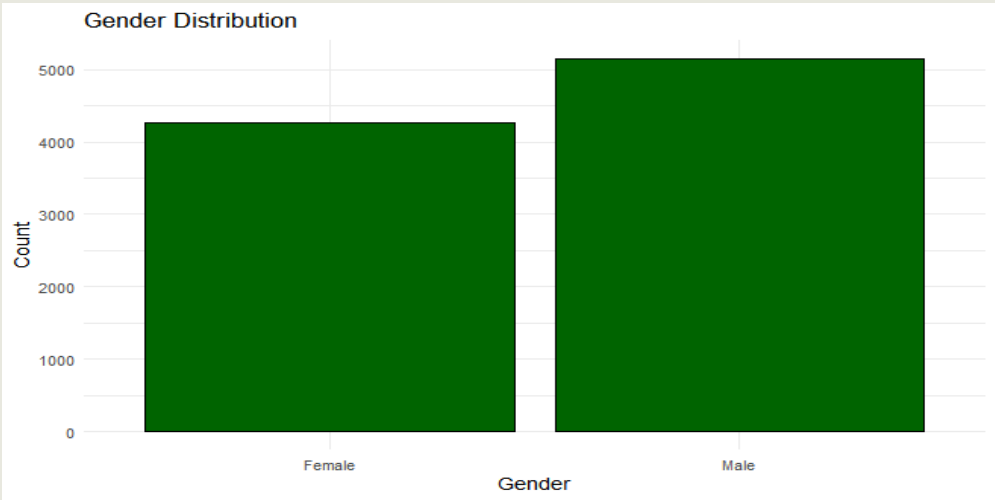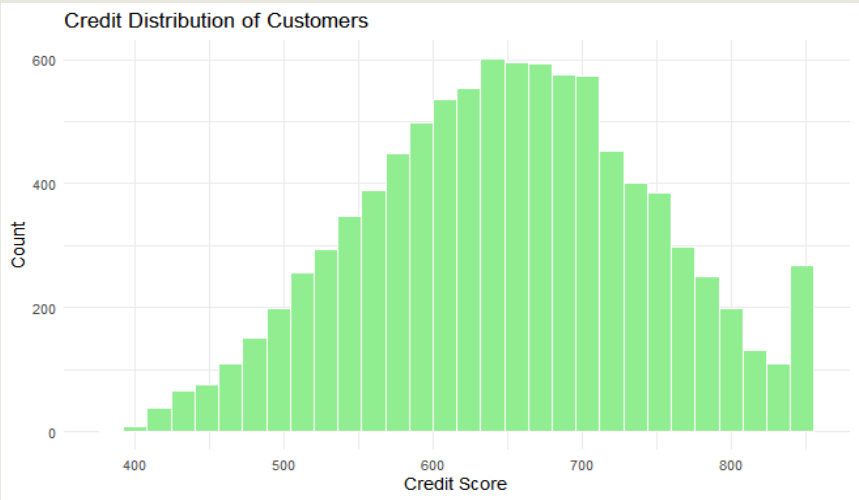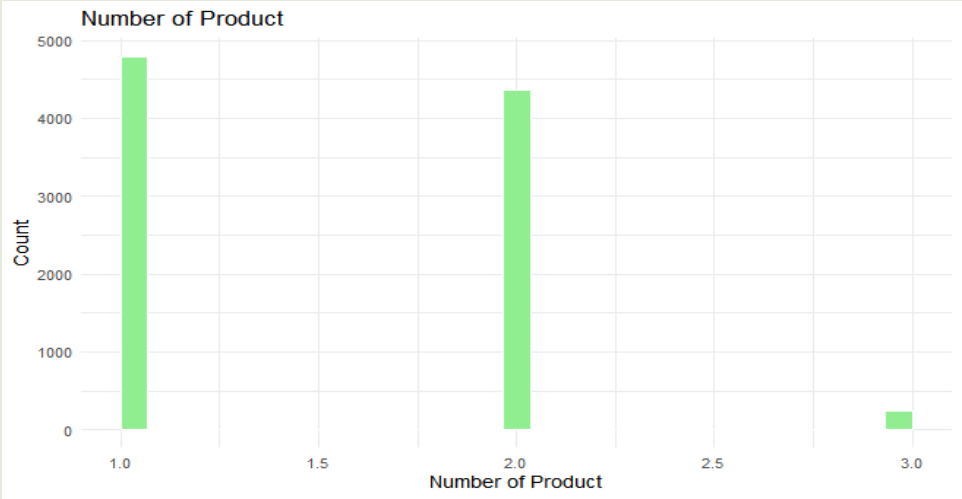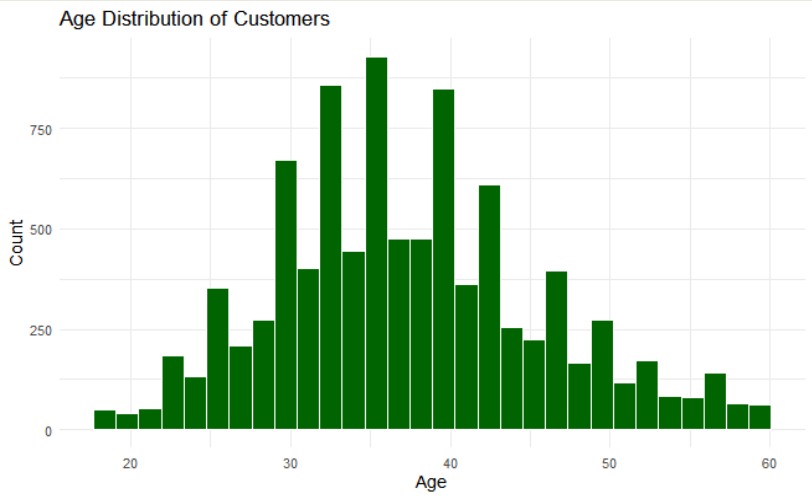
DATA PREPARATION

Data cleaning
- No missing values
- Removed outliers

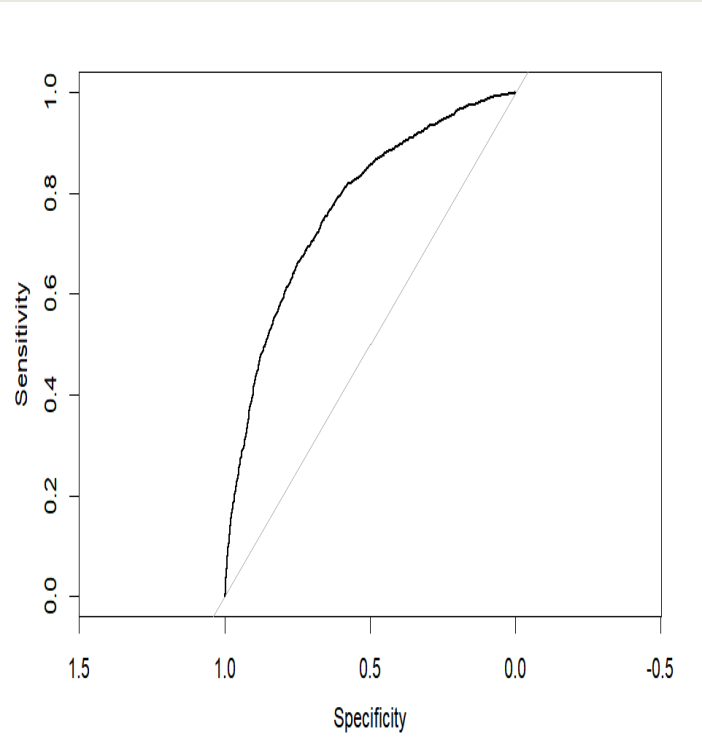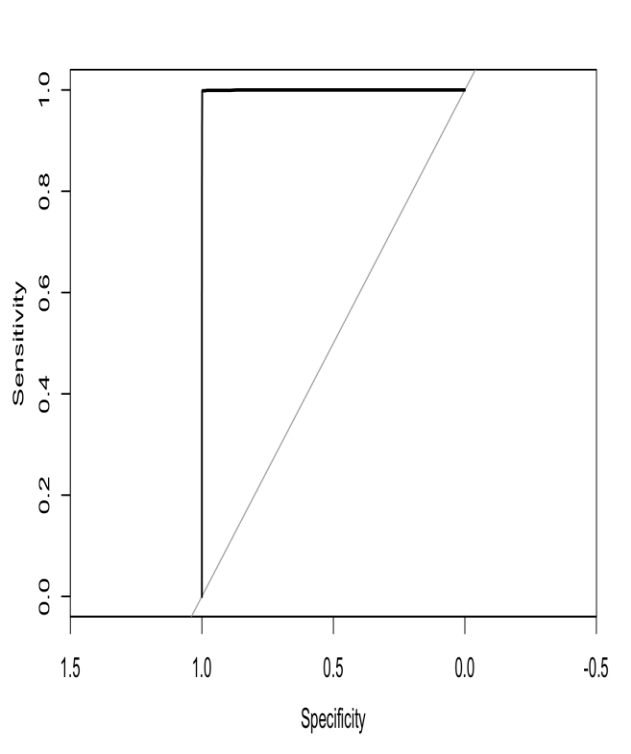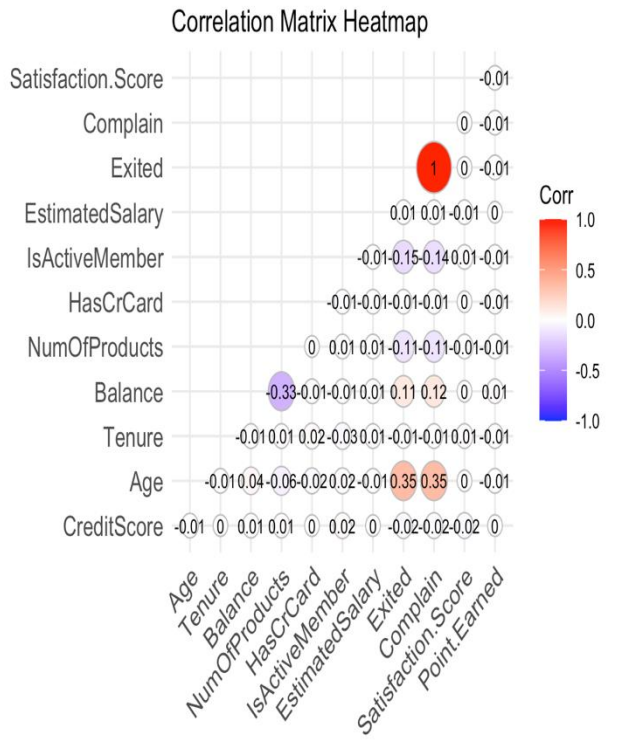**Data partitioning**
- 70% traning
- 30% testing

Standardization
- Z- score

8/1/2025

COMPLAIN

We removed the variable **'complain'** from our analysis as it showed a high correlation with the target variable **'exited'**, which led to overfitting issues.

MODEL STATS

Algorithm	Accuracy	Sensitivity	Specificity	AUC
Logistic	71.7	73.50	69.87	79.59
CART	79.73	81.89	71.21	83.22
⭐ RF	83.8	89.26	62.24	85.48
KNN	83.53	68.41	98.96	83.68

8/1/2025

RESULTS

RF	MeanDecreaseAccuracy	MeanDecreaseGini
Age	109.97	1018.36
Balance	59.36	393.54
NumofProducts	59.67	553.83

Logistic Model	Estimated Coefficient	P-Value
Age	1.092e-01	<2e-16***
Balance	2.761e-06	1.37e-11***
NumOfProducts	-2.691e-01	1.26e-12***

8/1/2025

BUSINESS RECOMMENDATIONS

Age: As age increases the log odds of leaving increases, age is a continuous variable. As age increases the customer is 2.98 times more likely to exit. We can offer restaurant, groceries, and travel rewards to those customers

Balance: As balance increases the log odds of leaving the bank increases. A customer is 15.81 more times likely to exit. To prevent customers from leaving the bank, the bank will could offer extra rewards for every three months that are paid off earlier.

NumofProducts: Customer with a higher number of products through the bank are approximately 6.78 times less likely to churn compared to those without such a plan. We can offer customers to pick industry of their liking to earn double to points for each month.
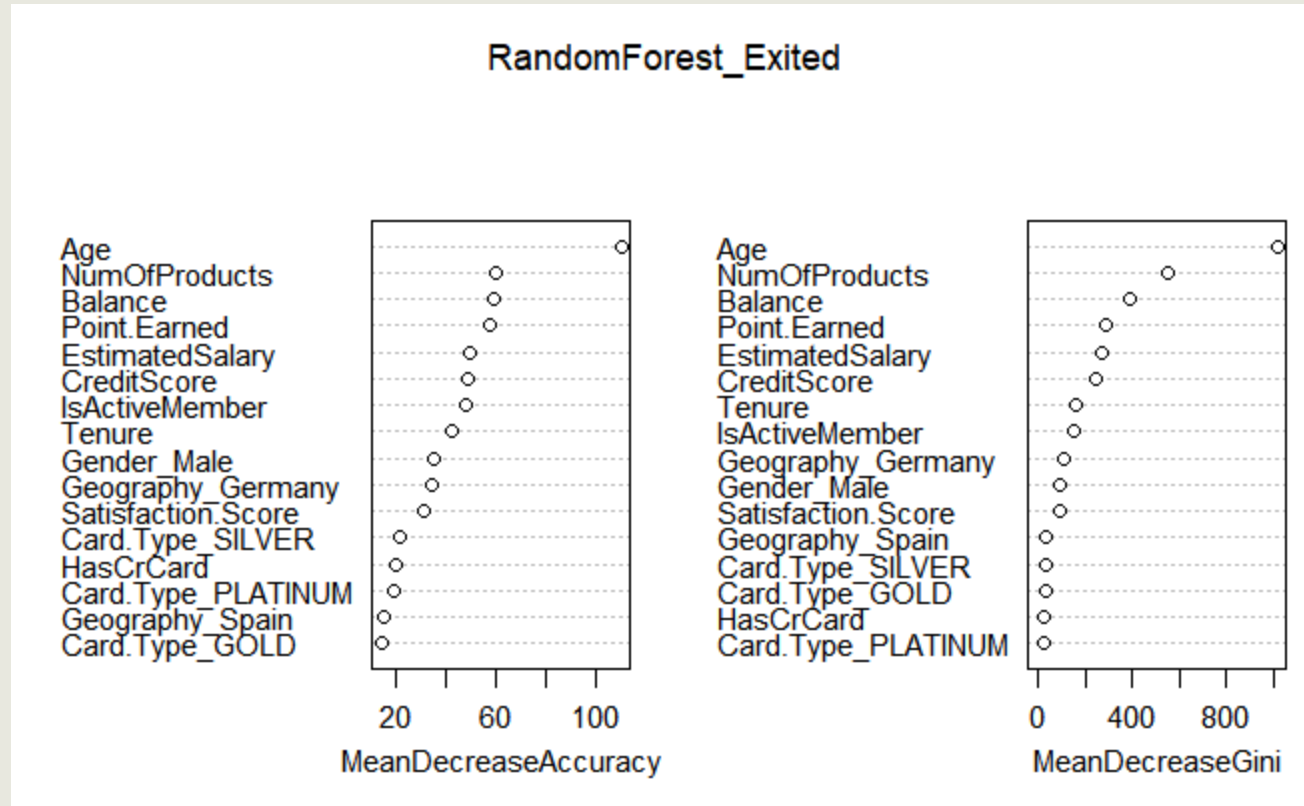
8/1/2025

CONCLUSION

- After running various models, we conclude that the best model is the Random Forest model.
- We reviewed the importance factor to determine the variables that are most important in affecting customer churn.
- We have provided business recommendations accordingly.

8/1/2025

Mercer University

THANK YOU!
QUESTIONS?

8/1/2025

APPENDIX 1: ROC AUC WITH DIFFERENT MODELS
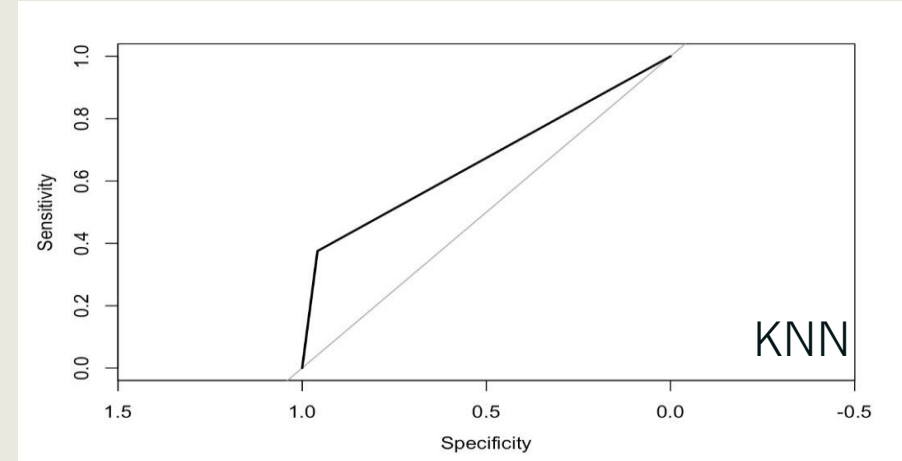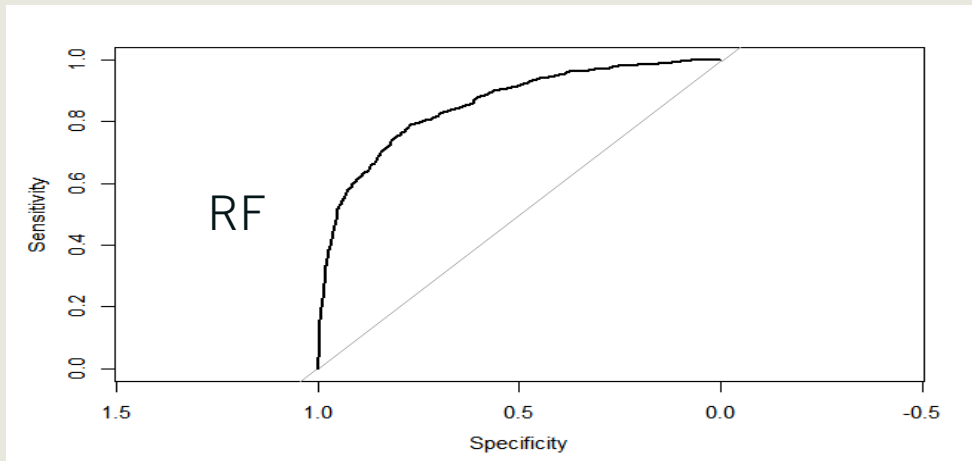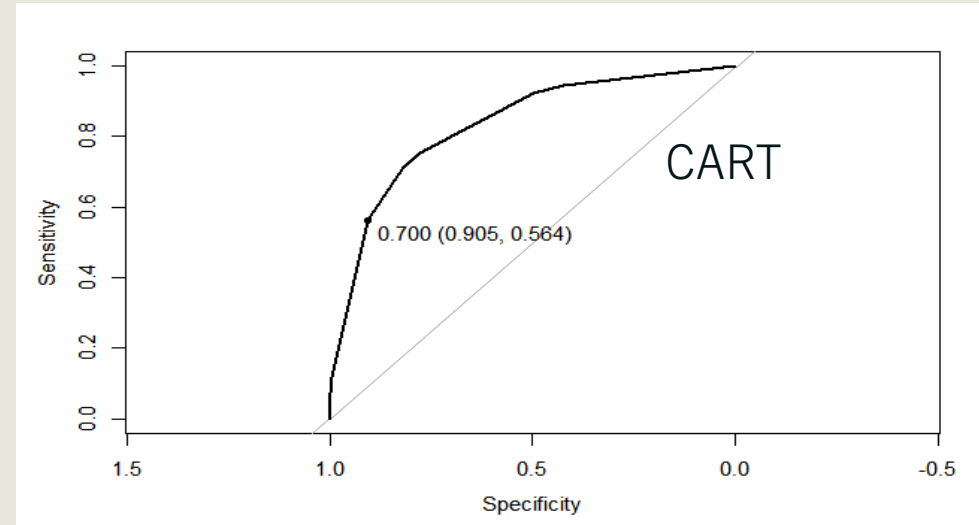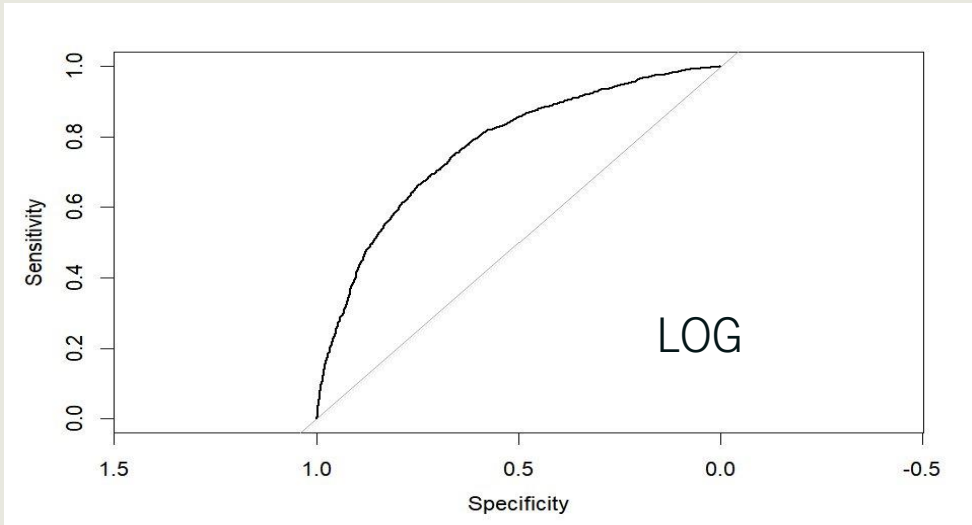
APPENDIX 2: RANDOM FOREST IMPORTANCE

APPENDIX 3: AGE AND PRODUCTS