A2: National City Bank National City Bank EDA & Modeling

Santiago Crane

Optimize the acceptance rate for a new line of credit

- **Develop a customer propensity model** for a new line of credit against a household's used car

- **Identify the top 100 customers** with the highest probability of accepting the offer from a prospective customer list

- Determine **variable importance** using historical data

EDA

About the Data

$ HHuniqueID
$ dataID

Unique Variables

$ Communication
$ LastContactDay
$ LastContactMonth
$ NoOfContacts
$ DaysPassed
$ PrevAttempts
$ past_Outcome
$ headOfhouseholdGender
$ annualDonations
$ EstRace
$ PetsPurchases
$ DigitalHabits_5_AlwaysOn
$ AffluencePurchases
$ Age
$ Job
$ Marital
$ Education
$ DefaultOnRecord
$ RecentBalance
$ HHInsurance
$ CarLoan
$ carMake
$ carModel
$ carYr

Informative Variables

$ Y_AcceptedOffer

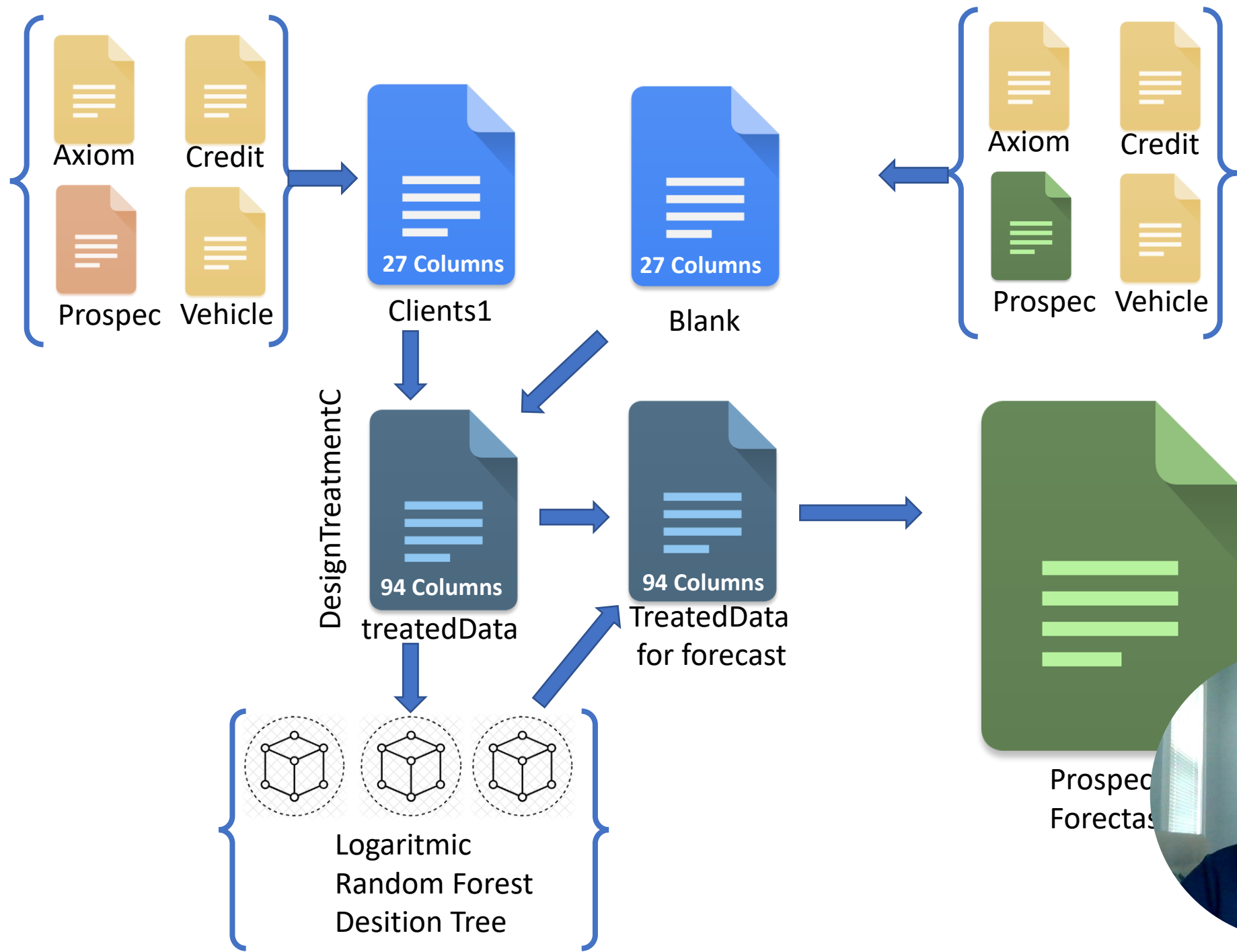Target Variable:
Success Class = "Accepted"

Distribution of Age

Distribution of Age
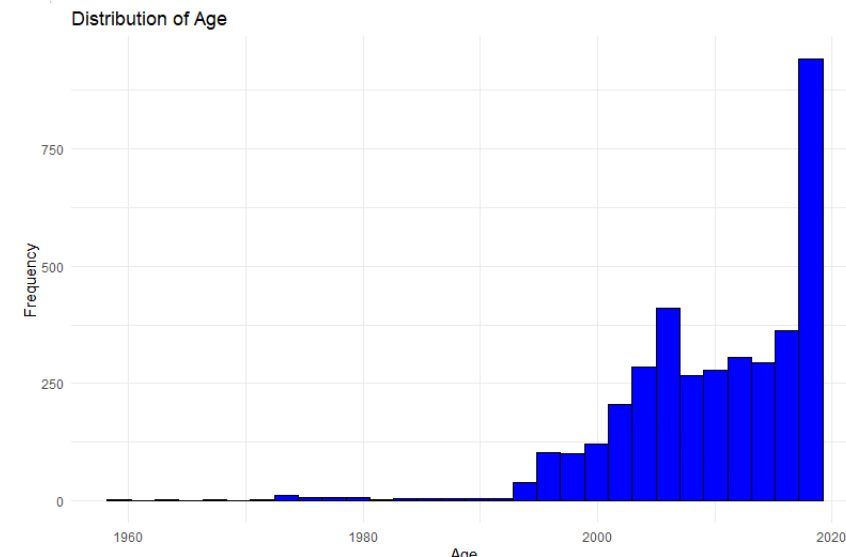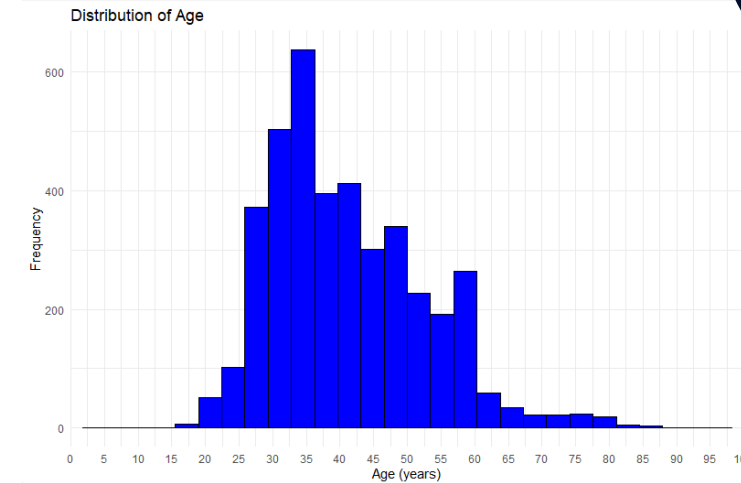
The average in USA is 8 years.(Meyer, 2022)

Model Preparation

Data Structure

Model Building

Logistic Regression	Random Forest	Decision Tree
• **Accuracy: 79%**	• **Accuracy: 80%**	• **Accuracy: 66%**
• Includes : Parsimony fit	• Includes : 50 Trees	• Lowes accuracy, model will not be used further
• Used to get the probabilities of each outcome	• Best accuracy level	
	• Used to get the predictions	

Key Findings

Coefficient Analysis

	Estimate
(Intercept)	0.90960615
Communication_catB	-0.85924410
LastContactMonth_catB	-1.11954611
NoOfContacts	0.04605602
past_Outcome_catB	-0.80544507
annualDonations_catB	-1.71274307
EstRace_catB	-1.01350506
PetsPurchases	-0.17225221
Job_catB	-0.45044982
Marital_catB	-0.86590286
Education_catB	-0.57741427
HHInsurance	0.48719084
CarLoan	0.42963821
carModel_catB	-0.96813027
LastContactMonth_lev_x_apr	-0.47686337
LastContactMonth_lev_x_aug	-0.48633778
LastContactMonth_lev_x_feb	-0.49795357
LastContactMonth_lev_x_jul	-0.89162037
LastContactMonth_lev_x_jun	-1.41984847
LastContactMonth_lev_x_may	-1.14750416
LastContactMonth_lev_x_nov	-0.62581283
LastContactMonth_lev_x_oct	-0.62571150
Job_lev_x_blue_minus_collar	-0.50191231
Education_lev_NA	0.56194006
carMake_lev_x_GMC	0.40957836

Variables that influence the most for accepting the credit:

- Health Insurance

- Car Loan

- No education

- GMC Car brand*

The ones that affect the outcome the most :

- Communication frequency

- No contact on the previous month

- No annual donations from customer

*GMC US Market share held 18% for 2022 and its considered thelargest in the USA, maybe this is the reason that this affects the outcome. (Carlier 2022)

Key Insights – Top 100

- With a 80% of accuracy, these are the top 100 customers that have the highest probabilities of Accept the credit.

  - **80%** of them have not been contacted before to get a credit

  - **52%** Female, **48%** male

  - **59%** have pets

  - Age distribution is scattered, not relevant-

  - Vast majority are Technician, Managers or Executives

  - **51%** Married

Backup Resourses

- Santiago Crane_Bank_Script.R

- top_prospects.csv

References

- Meyer, S. (n.d.). *Survey: Average length of car ownership in America - the zebra*. Retrieved March 22, 2023, from https://www.thezebra.com/resources/driving/average-length-of-car-ownership/

- Carlier, M. (2023, January 11). *U.S.: General Motors' Vehicle Sales Market Share 2022*. Statista. Retrieved March 22, 2023, from https://www.statista.com/statistics/239607/vehicle-sales-market-share-of-general-motors-in-the-united-states/#:~:text=General%20Motors'%20vehicle%20sales%20market%20share%20in%20the%20U.S.%202015%2D2022&text=At%20about%2017.1%20percent%2C%20General,the%20U.S.%20market%20in%202022.

Thank you

Santiago Crane