CUSTOMER CHURN PREDICTION OF TELECOM COMPANY

SASSY PREDICTORS
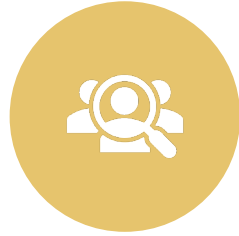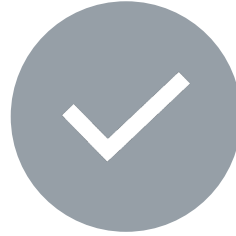
GROUP 8

CONTENTS

SECTION 1: INTRODUCTION

LAYING FOUNDATION OF OUR PROJECT'S VISION, BRIEF INDUSTRY BACKGROUND AND THE GOALS WE AIM TO MEET VIA ANALYSIS

Customer loyalty is the key to profitability in the telecom industry

Analysis to focus on the behavior of telecom customers who are more likely to leave the platform (i.e., Customer churn).

Intend to initially perform EDA to gauge customer behavior

Thereafter, use predictive analytics techniques to determine the customers who are most likely to churn

Lastly, analyze model and results to predict behavior for a customized customer retention programs, which in turn would help in increasing profits

INTRODUCTION

Telecom services such as voice call and internet are a necessity these days.

Telecom Market has become saturated over the years lacking differentiating factors between competing brands.

Acquiring new customers can be as much as 5X expensive than retaining existing ones thus, companies are constantly looking for innovations to retain customers.

TELECOM INDUSTRY BACKGROUND

WHY UNDERSTAND CUSTOMER CHURN?

Customer acquisition is expensive thus, it is important to retain existing customers.

Is all churn bad? Understanding customer lifetime value is important for that.

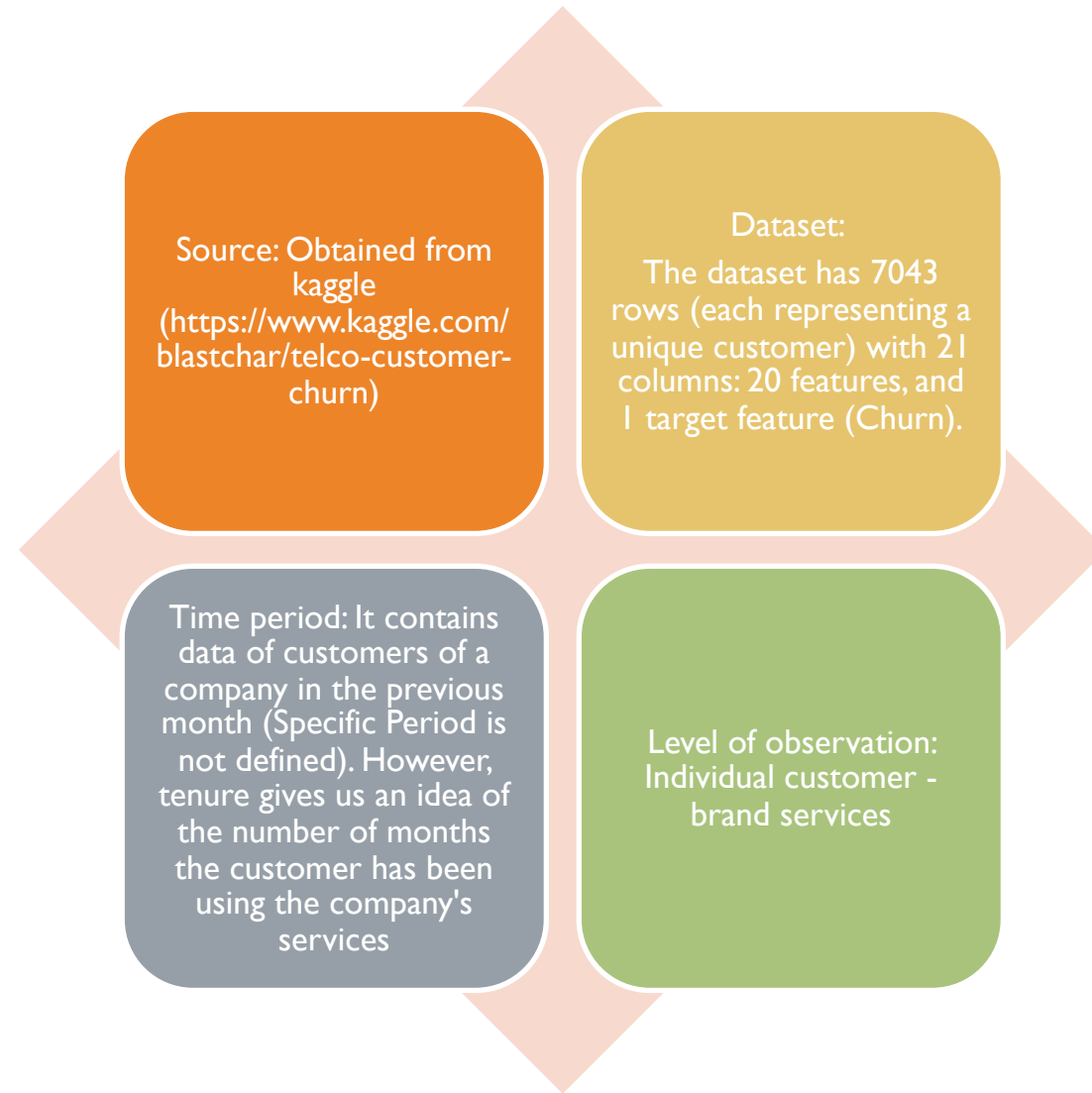Churn is usually result of combination of several customer issues.

OBJECTIVES

- Identify potential factors leading to customer churn

- Understanding demographic information affecting churn

- Analyzing behavior based on the type of contracts and payment methods

- Identifying loyal customers and rewarding them

- Engaging with customers likely to churn

- Develop marketing and promotional strategies
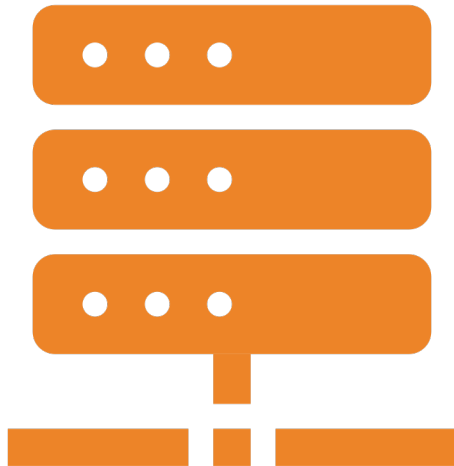
SECTION 2: DATA

DESCRIBING THE VARIABLES AND PARAMETERS IN OUR DATA AND DEMOGRAPHICS DATA EXPLORATION AND VISUALIZATION
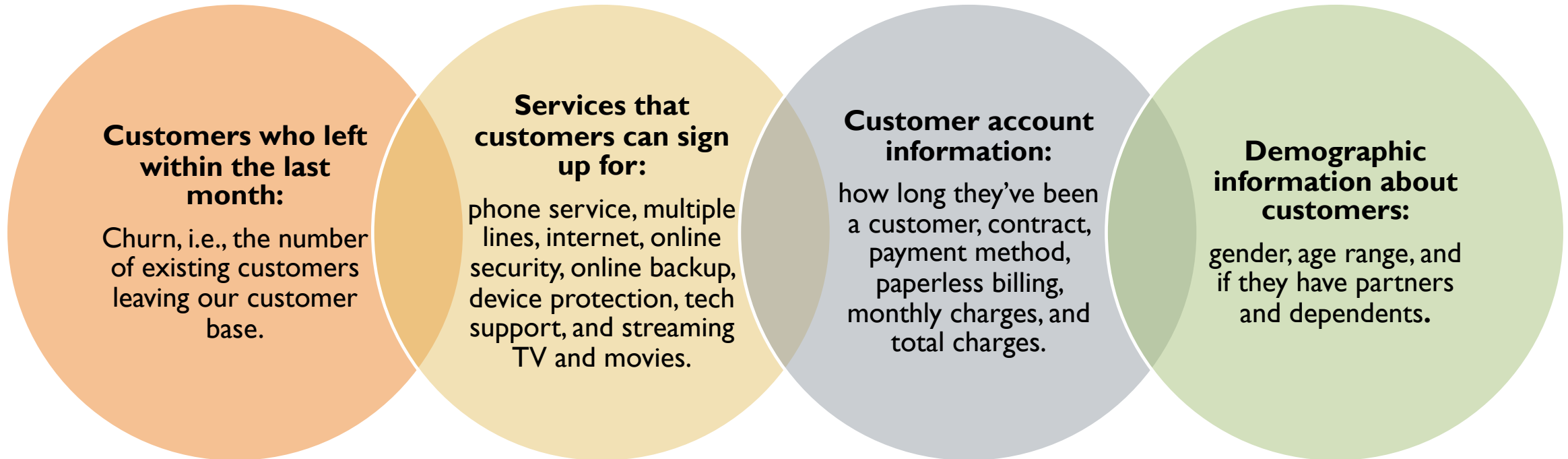
Source: Obtained from kaggle (https://www.kaggle.com/blastchar/telco-customer-churn)

Dataset:
The dataset has 7043 rows (each representing a unique customer) with 21 columns: 20 features, and 1 target feature (Churn).

Time period: It contains data of customers of a company in the previous month (Specific Period is not defined). However, tenure gives us an idea of the number of months the customer has been using the company's services

Level of observation: Individual customer - brand services

DATA OVERVIEW

| Attribute	Description	Type
Customer ID	Each unique identification number of a customer	VARCHAR
Gender	Gender of the customer (Female / Male)	CHAR
Senior Citizen	Age Range to identify customers' age group (0/1)	BINARY
Partner	Whether the customer has a partner (spouse or partner) or NOT (Yes/No)	BOOLEAN
Dependents	Whether the customer has a dependent (for e.g.,: child, spouse, parent, or certain other relative to whom they provide financial support) or NOT (Yes/ No)	BOOLEAN
Tenure	Months for which subscriber has been with company	NUM
Phone Service	Whether the Customer has a phone service or NOT (Yes/No)	BOOLEAN
Multiple Lines	Whether customer has subscribed to multiple lines or NOT (Yes / No / No phone service)	CHAR
Internet Service	Whether the customer has an internet service, if yes - which data service (DSL / Fiber optic / No)	CHAR
Online Security	Whether the customer has online security or NOT (Yes / No / No internet service)	CHAR
Online Backup	Whether the customer has online backup or NOT (Yes / No / No internet service)	CHAR
Device Protection	Whether the customer has device protection or NOT (Yes / No / No internet service)	CHAR
Tech Support	Whether the customer has technological assistance or NOT (Yes / No / No internet service)	CHAR
Streaming TV	Whether the customer has television streaming or NOT (Yes / No / No internet service)	CHAR
Streaming Movies	Whether the customer has movie streaming or NOT (Yes / No / No internet service)	CHAR
Contract	Type of customer's contract term (Month-to-month / One year / Two year)	CHAR
Paperless Billing	Whether the customer has paperless billing or NOT (Yes / No)	BOOLEAN
Payment Method	The type of payment method the customer uses (Electronic check / Mailed check / Bank transfer (automatic) / Credit card (automatic))	CHAR
Monthly Charges	The amount of charges the customer pays monthly	NUM
Total Charges	The total amount of charges the customer pays	NUM
Churn	Whether the customer has left the provider or NOT (Yes / No)	BOOLEAN

DATA DICTIONARY

KEY VARIABLES

Customers who left within the last month:

Churn, i.e., the number of existing customers leaving our customer base.

Services that customers can sign up for:

phone service, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies.

Customer account information:

how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges.

Demographic information about customers:

gender, age range, and if they have partners and dependents.

CONSUMER DEMOGRAPHICS

- We see that large portion of customers who churn have no partners or dependents. (Focusing on better, cheaper family plans can be useful to reduce churn)

- We see considerable amount of churn who have monthly contracts around 23%. (this is because most monthly customers are confused between carriers and don't want to commit to anyone)

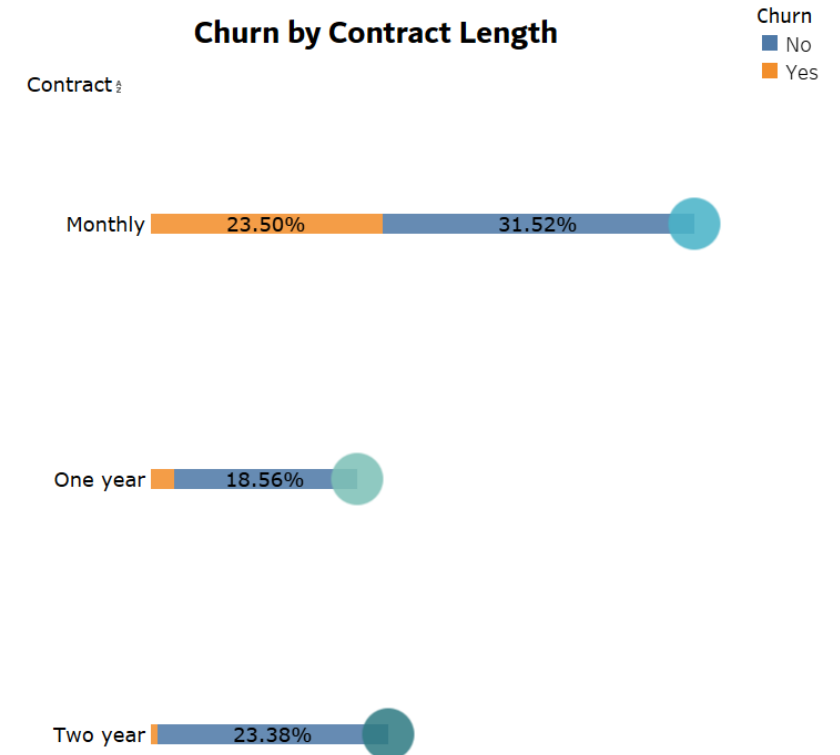- Gender seems to have no significant difference on why a customer leaves the company

CUSTOMER CHURN ANALYSIS

% of Churn by Partner Status

Has a Partner	Single
80.34%	67.04%
19.66%	32.96%

% of Churn by Dependents

No	Yes
68.72%	84.55%
31.28%	15.45%
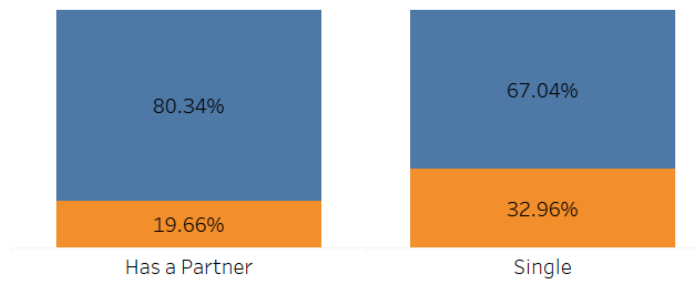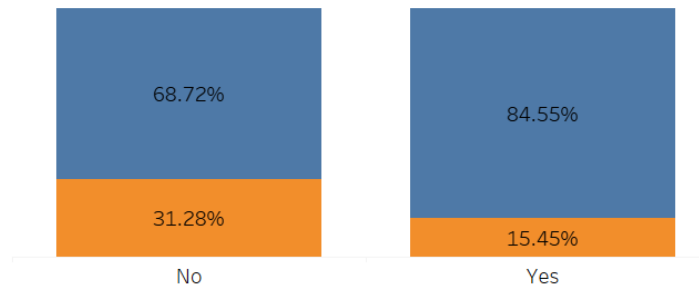
Churn by Contract Length

Contract

Churn
- No
- Yes

Monthly   23.50%   31.52%
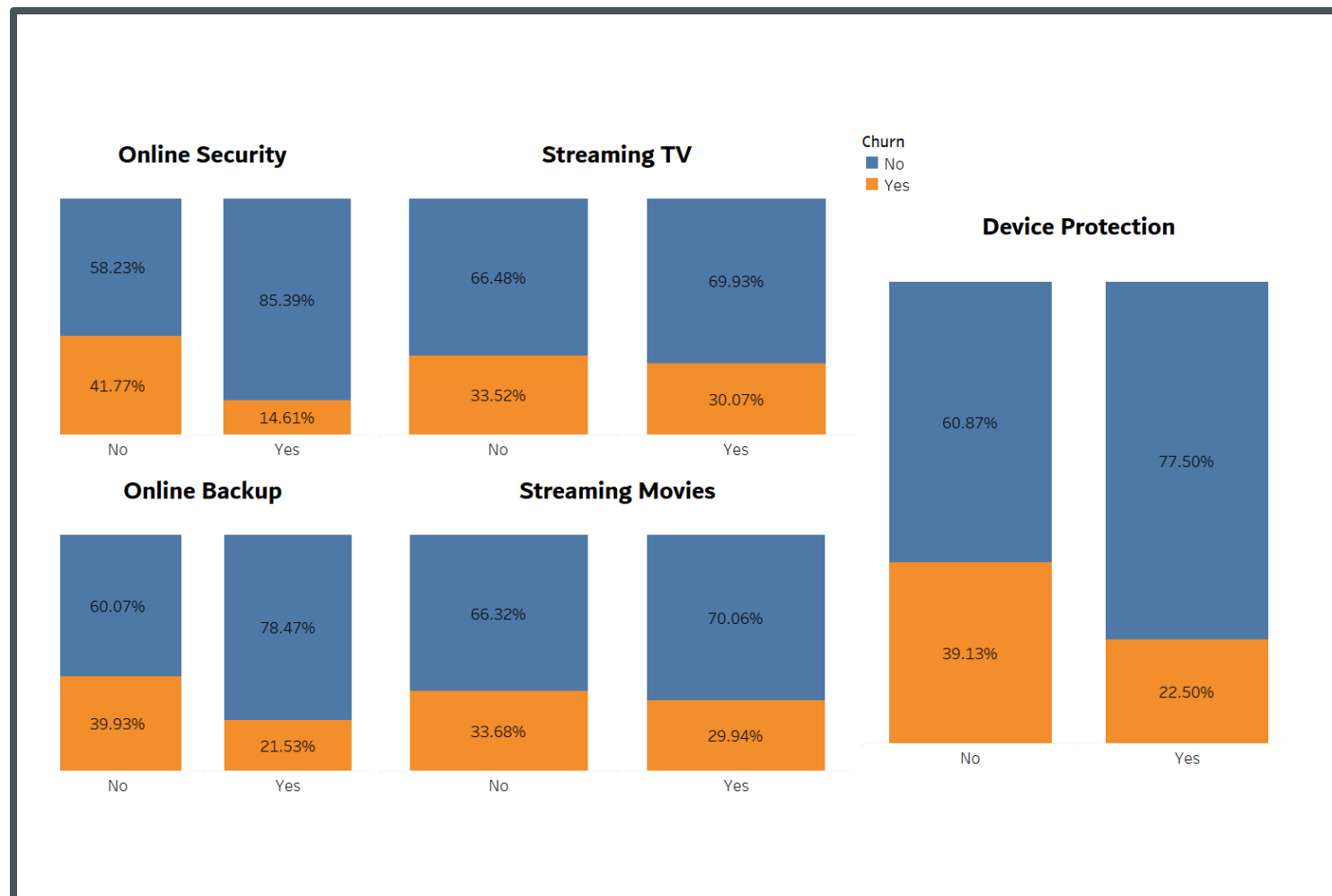
One year   18.56%

Two year   23.38%
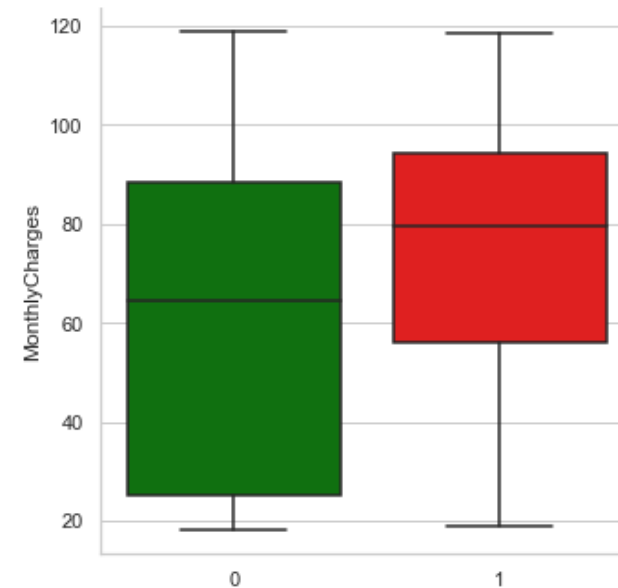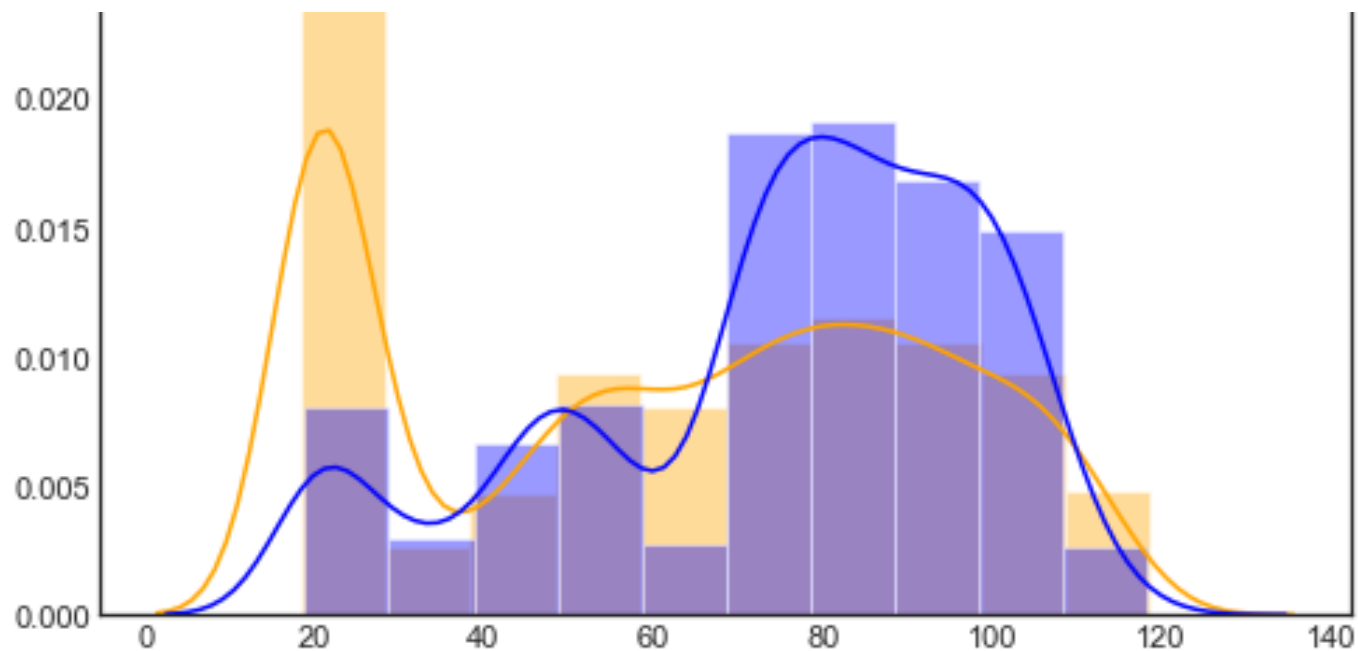
MONTHLY CHARGES & CHURN

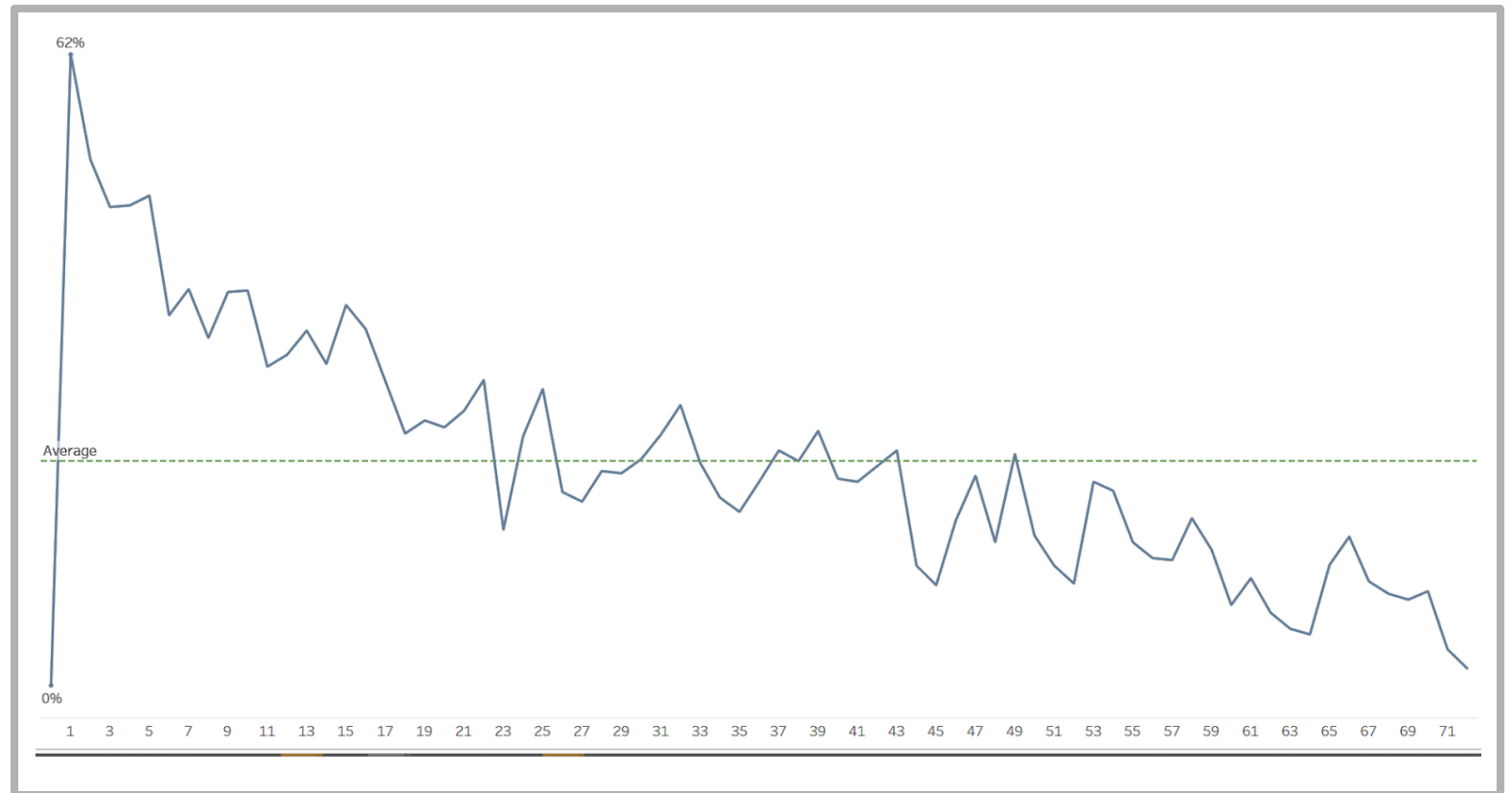*We see the maximum churn amongst our customers at monthly charges around $70-110.*

SERVICES USAGE

- With increasing number of services we see a considerable drop in the churn rate for customers.

- This is more evident in services where there is some barrier to change to other platform unlike Streaming services.
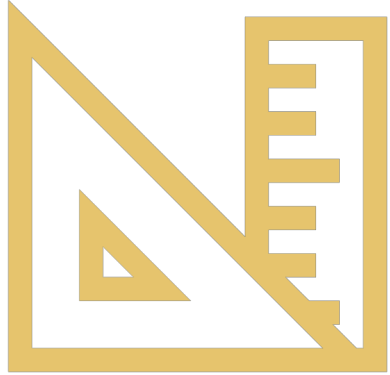
AT WHAT POINT IN TENURE IS THE CHURN HIGHEST?

- *Customers who are with the telecom providers for a shorter duration (usually prepaid customers) usually switch to another provider without a prior notice.*

- *Predicting their churn is tricky as one might wonder if the services are being used on and off (for a short travel every few months), or the customer has churned.*
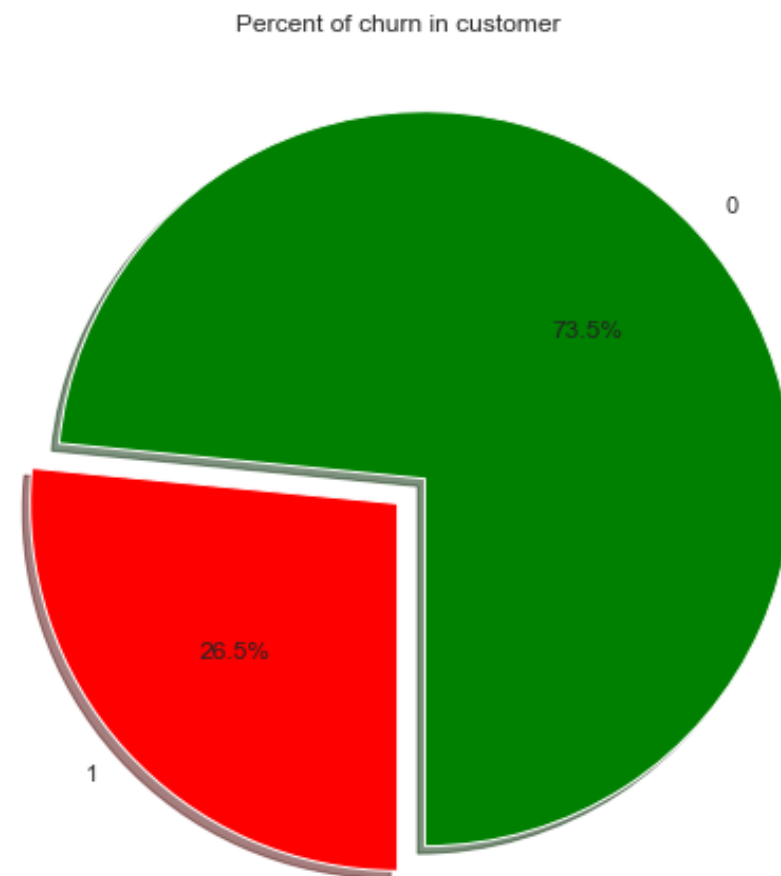
SECTION 3: ANALYSIS

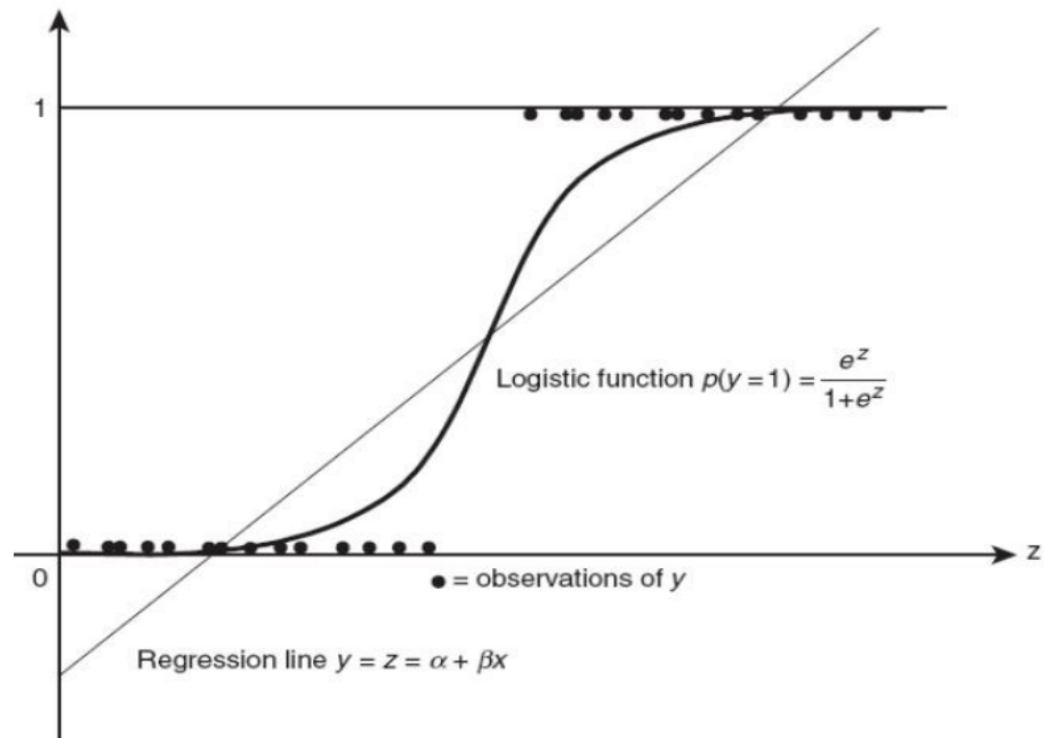MODEL SELECTION, METHODOLOGY AND DETAILED ANALYSIS

MODEL SELECTION

Customer Churn Classification (categorical: yes / no)

Choice of Model: Logistic Regression

Percent of churn in customer

0

73.5%

26.5%

1

WHY LOGISTIC REGRESSION

Logistic function $p(y = 1) = \dfrac{e^z}{1+e^z}$

$\bullet$ = observations of y

Regression line $y = z = \alpha + \beta x$

Logit function vs Linear Probability Model

Probability values between
- [0, 1]
Homoscedasticity - residuals not normal

Odds Ratio : Better interpretation

FEATURE SELECTION & DATA PREPARATION

Based on exploratory analysis and statistical significance, following features are **not considered**:

Gender, Senior Citizen, Online Security, Online Backup, Device Protection, and Tech Support

Monthly Charges has been selected over Total Charges to address **multi-collinearity.**

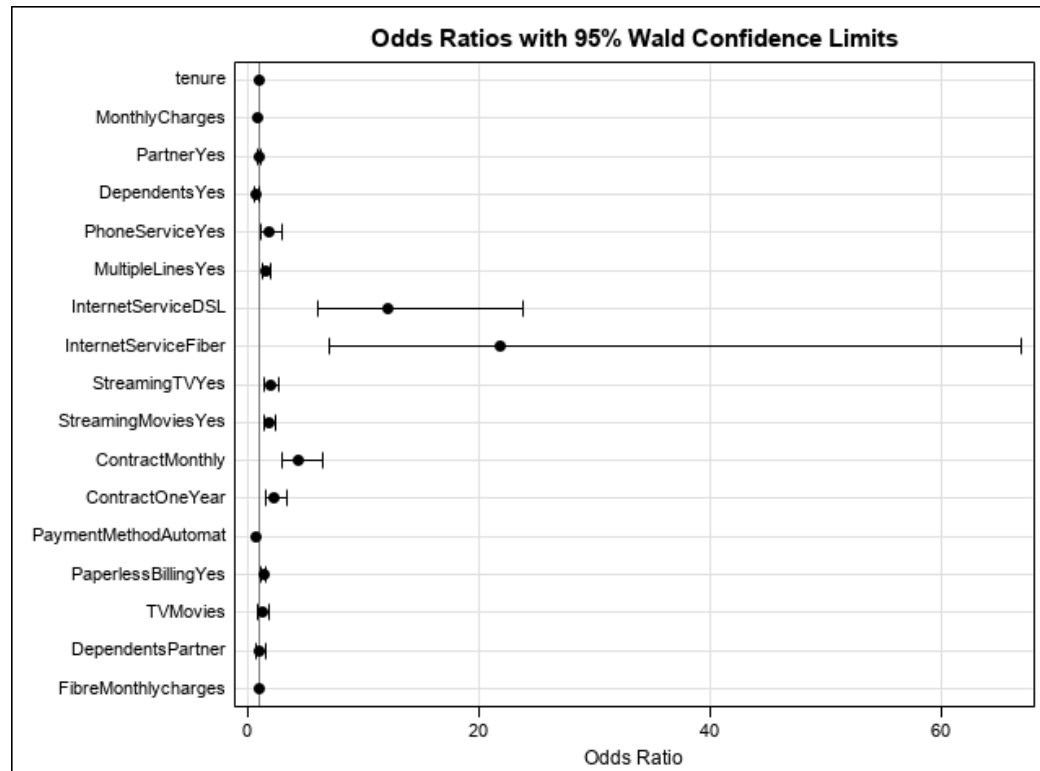Interaction effects included: Fiber Optic*Monthly Charges, Streaming Movies*Streaming TV, and Dependents*Partners

Yes/No ~ 1/0

Categorical ~ dummy

Train-Test split ~ 80:20

19
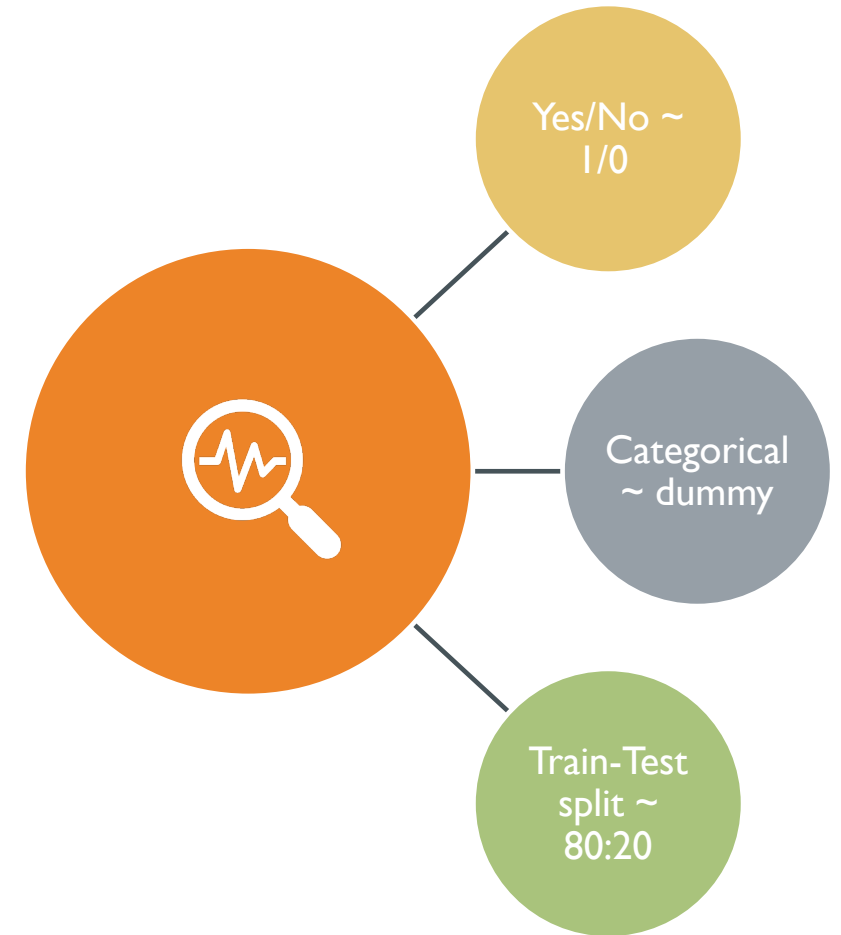
SECTION 4: RESULTS

FINDINGS OF THE ANALYSIS, OUR CONCLUSIONS, MACRO IMPACT AND RECOMMENDATIONS

LOGISTIC REGRESSION MODEL

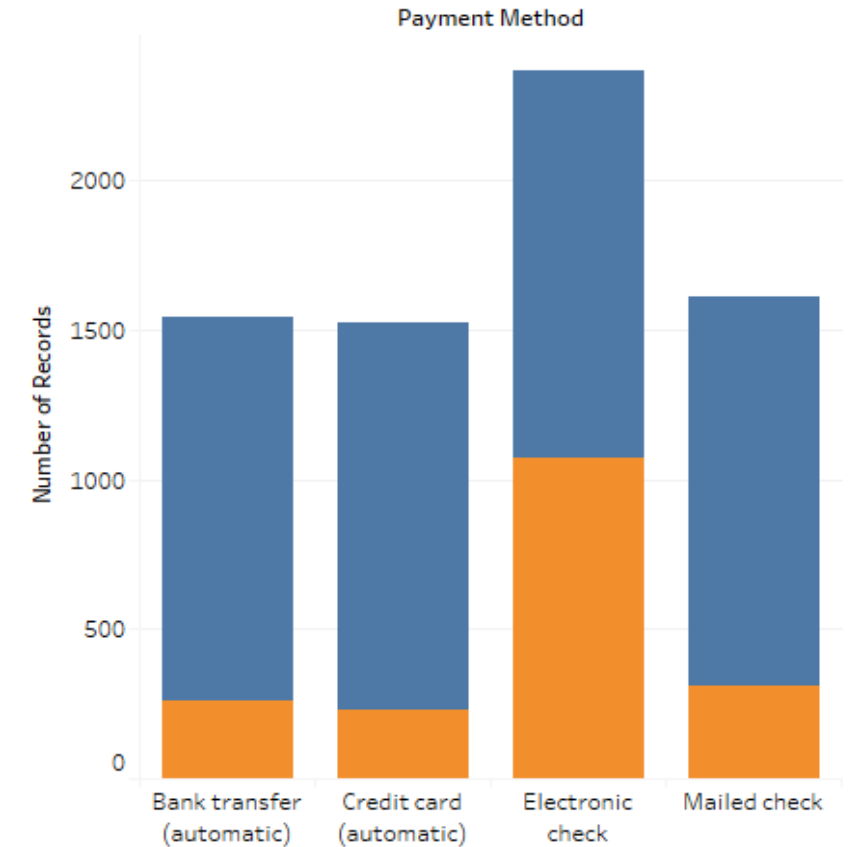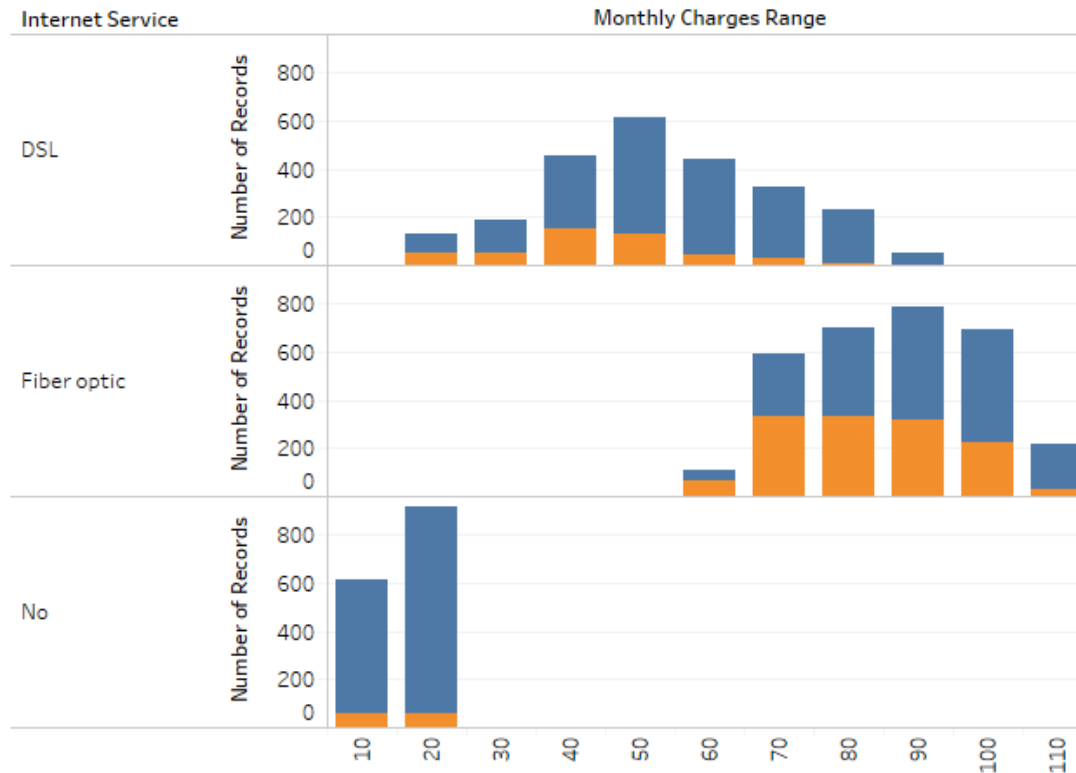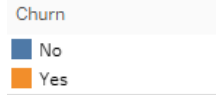Odds Ratios with 95% Wald Confidence Limits

Odds Ratio: variables having higher odds ratio and statistical significance:

- Internet Service with Fiber Optic
- Internet Service with DSL
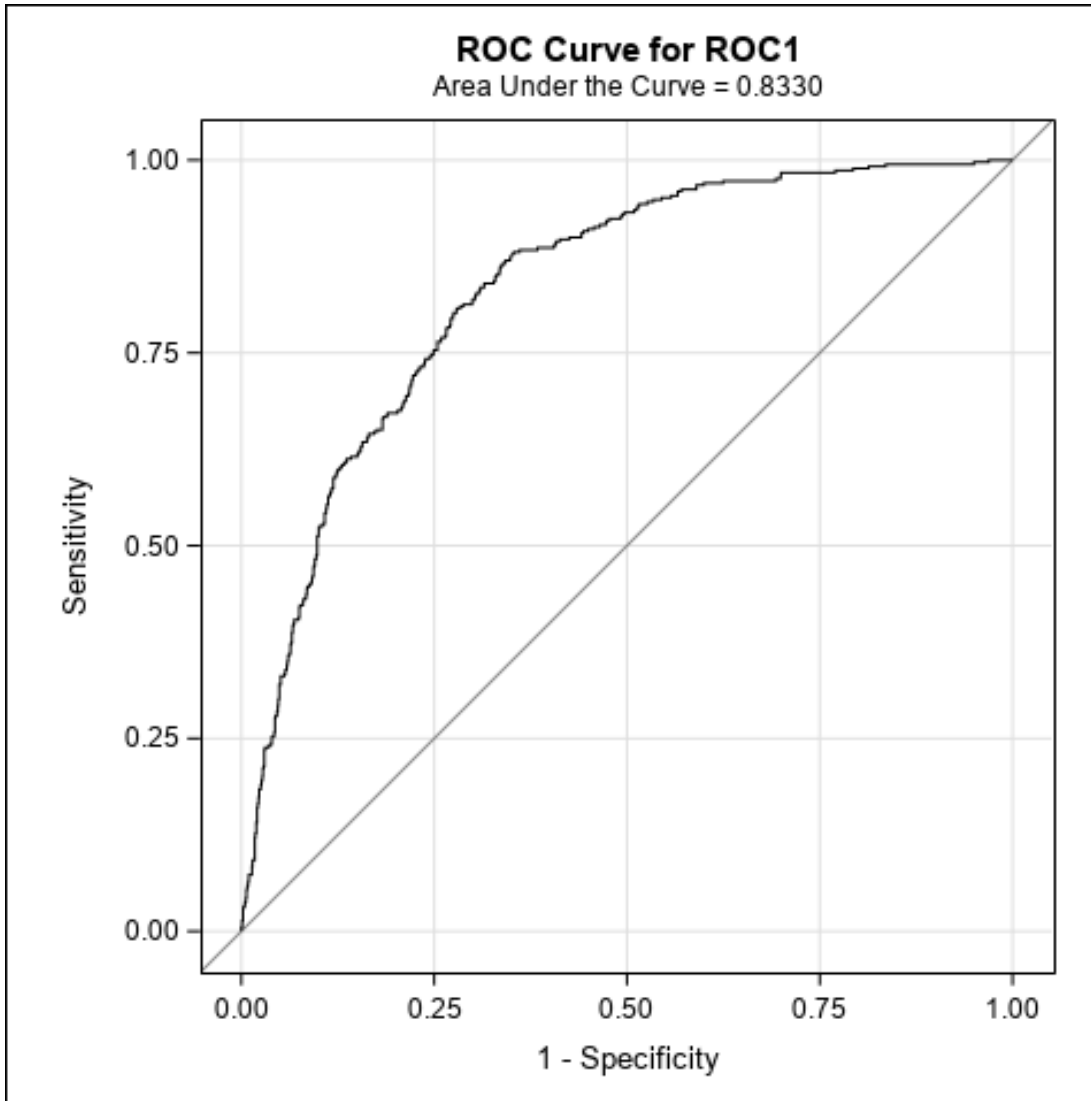- Monthly Contracts
- One Yearly Contracts

Variable having lower odds ratio and statistical significance:

- Payment Method Automatic

EFFECTS OF MONTHLY CHARGES & INTERNET SERVICE, AND PAYMENT METHOD ON CHURN AS OBSERVED IN THE LOGISTIC MODEL

ROC CURVE
AUC = 0.8330

APPLICATION BASED THRESHOLD SELECTION

Promotional offers & Bundle provision

- Cost of predicting false churn (False Positive) - $20
- Cost of not predicting true churn (False Negative) - $80
- Minimum Cost = $10,940
- Threshold Probability = 0.201
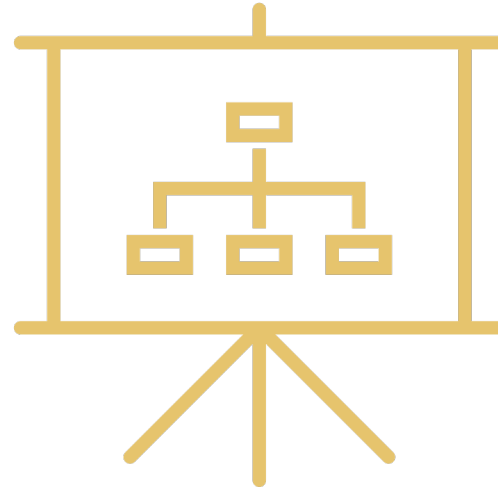
APPLICATION BASED THRESHOLD SELECTION

Engaging with Customers by call / email:

- Cost of predicting false churn (False Positive) - $3
- Cost of not predicting true churn (False Negative) - $10
- Minimum Cost = $1,541
- Threshold Probability = 0.203

SECTION 5: CONCLUSION

DESCRIBING THE VARIABLES AND PARAMETERS IN OUR DATA AND DEMOGRAPHICS DATA EXPLORATION AND VISUALIZATION

1

Optical fiber users were most likely to churn in the previous month.

2

Customer churn is very high among monthly contract users.

3

Autopay enrollment is the best performing feature for retention.

4

The belief that multiple lines using customers are less likely to churn is wrong.

CONCLUSION

RECOMMENDATIONS

Immediate actions required for internet prices, especially for optical fibre.

Try to convert monthly users to long term plans.

Encourage customers for automatic payment enrollment by giving discounts/credits.

Reward customers - least likely to churn and long tenure.

Service bundling for reward/promotion to attract for longer terms

To engage with customers who are likely to churn.

THANK YOU!

WE ARE OPEN FOR QUESTIONS

Arunachalam Ramanathan
Mihir Bansal
Mitansh Patel
Priyanka Agarwal
Rohini Kanungo
Sagar Patel