6414 Group Project - Telecom Customer Churn Modeling

Jared Babcock, Rishi Bubna, Marta Bras

2019-11-19

Contents

1 Introduction

Customer churn (or customer attrition) analysis is one of the vital measures for subscription-based business models such as telecom services, internet providers. This is more important in business sectors that have high competition. Companies can choose to increase engagement with customers that are expected to churn to retain them, which would lead the company to increase revenue.

1.1 Summary (goals and major findings)

1.2 Project goals

1. Building a predictive model for churn rate that best identifies which customers are likely to churn.

2. Building a predictive model for CLTV that best identifies how much a customer is worth for the company.

3. Perform customer segmentation to identify high value customers that are likely to churn.

1.3 Reasons for our study

1.4 A Priori Expectations

2 Data

The IBM Business Analytics Community provides a fictional dataset of over 7,000 customers for a telecom company that contains information about which customers have left, stayed, or signed up for their service. The dataset also contains major demographic information for customers, along with Satisfaction Score, Churn Score, and Customer Lifetime Value (CLTV) index.

The database has data from 7,043 telecom customers, all located in California (USA). The average tenure of the customers is 32 months with an average churn rate of 59% and an average CLTV of 4,400$.
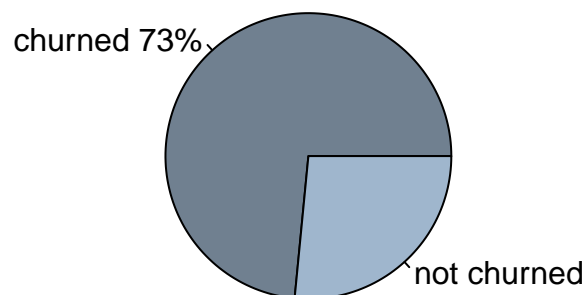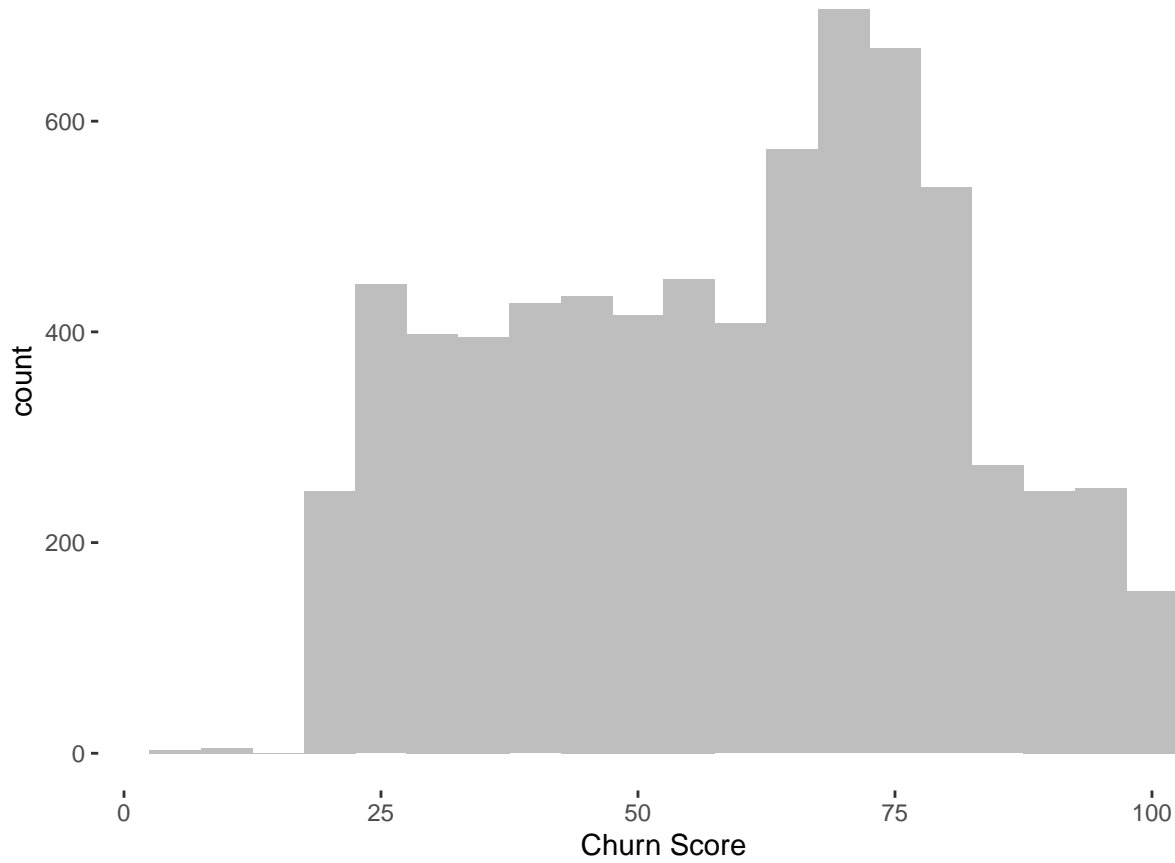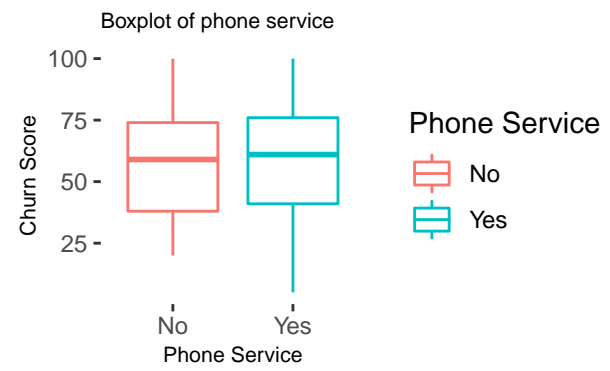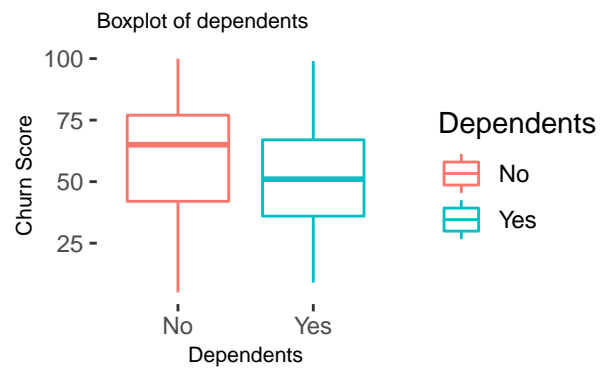
Table 1: Demo Table	
number observations(#)	7043
average tenure (months)	32
min tenure (months)	0
max tenure (months)	72
average churn rate(%)	59
min churn rate(%)	5
max churn rate(%)	100
average CLTV($)	4400
min CLTV($)	2003
max CLTV($)	6500

3 Explanatory Data Analysis

Histogram of Churn Score

Boxplot of Gender

Boxplot of seniority

Boxplot of dependents

Boxplot of phone service
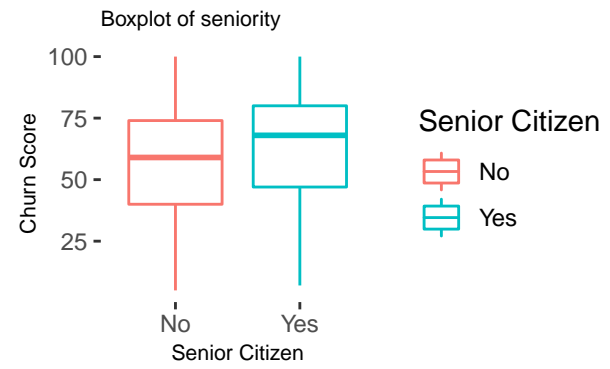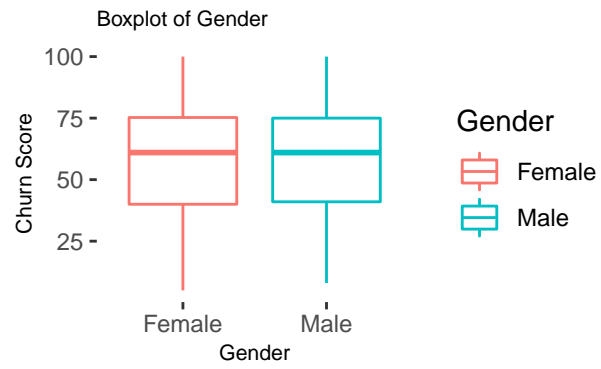
4 Churn Rate : Predictive model

4.1 1. Full model

4.2 2. Model fit

4.3 3. Variable transformation

4.4 4. Re-running the model

4.5 5. Variable selection

4.6 6. Model selection

5 CLTV : Predictive model

5.1 1. Full model

5.2 2. Model fit

5.3 3. Variable transformation

5.4 4. Re-running the model

5.5 5. Variable selection

5.6 6. Model selection

6 Customer segmentation

6.1 1. Identification of high value customers

6.2 2. Targeting of high value customers

7 Conclusions

8 Subject Matter Implications

9 Limitations and next steps

10 Attachments