

As The World Churns

Rahul Gupta, Michael Koffie, Brandon May, Tushar Muley, and Edris Safari

Introduction

As our economy and corporations begin to operate in a global context, there have been increasing efforts to retain customers. Frequent acquisition and loss of customers is defined as customer churn and has been a particular area of focus in data science. It is important to businesses as “it is directly tied to firm profitability” [4]. The costs of keeping a customer are usually less than the costs of recruiting new customers [21]. This is why it is becoming increasingly important to use data science techniques and advanced analytics to predict which customers are vulnerable to leaving. It can be difficult to differentiate between customers who will respond to interventions and those who will not [4]. In addition, excessive customer turnover can be a sign of potential fraudulent activity. This is complicated by the fact that technology can serve two purposes to become closer to customers as well as alienate them [21].

The risk can be more than financial; in certain insurance industries, customer churn can signify loss of critical healthcare coverage and can significantly impact a person’s health. In fact, data science technique and predictive analytics in particular are being applied to treat cancer and impact healthcare outcomes [17,20]. Therefore, it benefits us all both economically and personally to obtain further insight into customer churn, its prediction, and its avoidance (if at possible). This project aims to critically evaluate the current state of customer churn and customer behavior in the financial and insurance industry, propose a data science framework and algorithm to ascertain customer churn, and reflect on the future direction of this field.

Deliverables

Main Goal: Reduce Customer Churn by 50%

Goals divided into 3 main states:

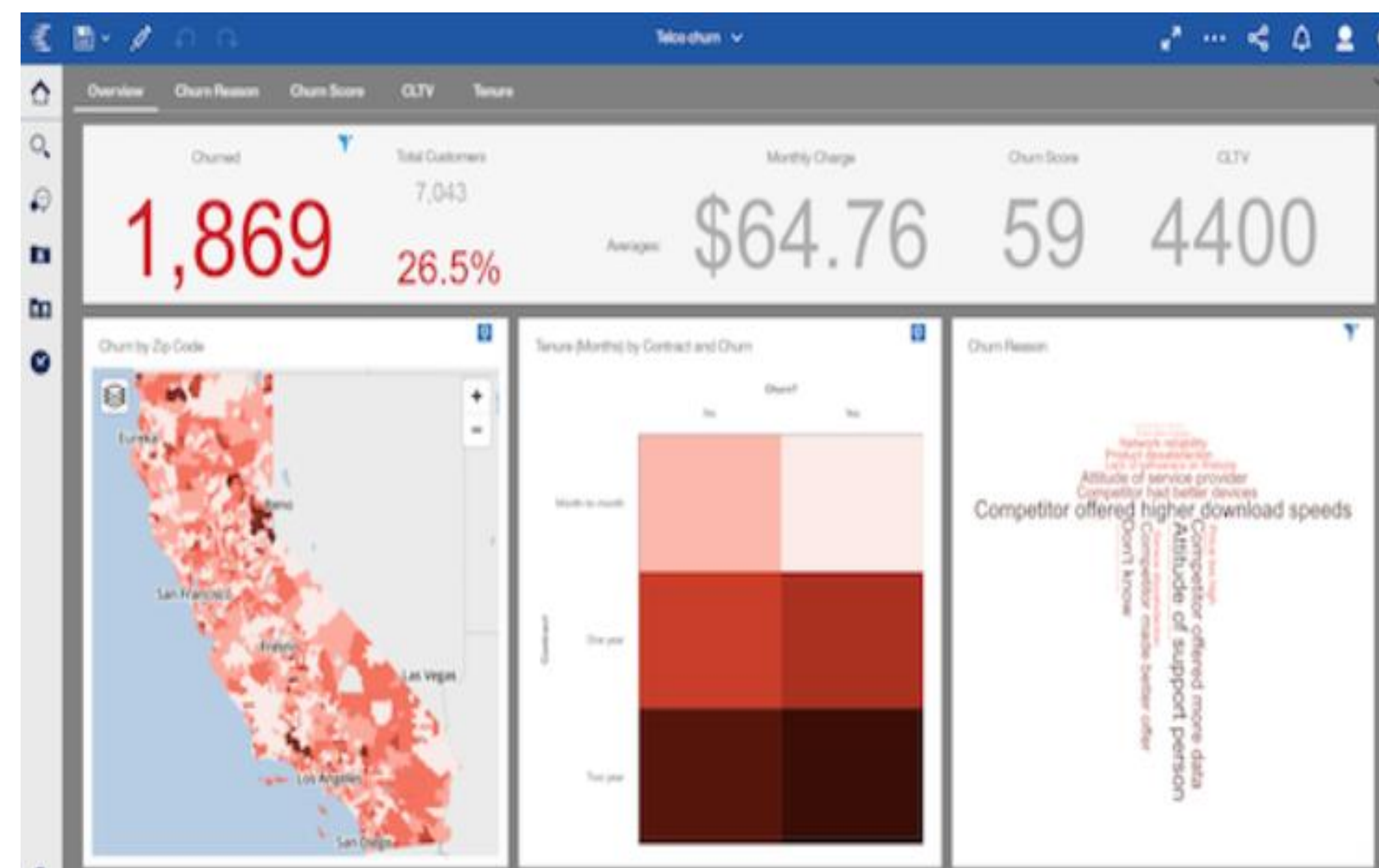
- Short Term Goal to reduce churn by 20%
- Medium Term Goal is to reduce churn by additional 20%
- Long-Term Goal to reduce churn by additional 10% for a total of 50%

Build models using various machine learning algorithms to retain and reduce the rate of customer churn including:

- Neural Networks
- Decision Tree Analysis
- K-Means Clustering
- Random Forest
- Logistical Regression



Sample Customer Churn Dashboard



Create dashboard and visualizations to analyze customer data with TABLEAU to make explanation easy to stakeholders.

Why Is This Data Science?

"Data is the new oil for all industries, and data science is the power that drives the industry."

Data Science transforms raw data into useful information. Industries need data to help them make careful decisions and is used in almost every industry including health, finance, and banking. Companies use the data to analyze their marketing strategies and create better ads. The industry needs data scientists to help them make smarter decisions that make financial sense.

Technology giants such as Facebook, Amazon and Google are constantly working in the field of machine learning and data science. Data science encompasses processes such as purging, processing, and analyzing data. A data scientist collects data from multiple sources, e.g. from surveys and physical data plots. Then, data is passed through strict algorithms to extract important information from the data and create a record. This record could also be used to parse algorithms to make more sense.

According to DOMO research, "More than 2.5 billion bytes of data are created every day and will only grow from there, and by 2020 an estimated 1.7 MB of data will be generated per second for every human being on Earth."

Predicting customer attrition and churn is just another way for data science to flex its capabilities in a modern world.

Conclusions

Customer churn in the financial and insurance sector is high. Companies struggle to identify customers who are likely to leave before they have left. Surveys are infrequent and a poor service might not show up on a survey. To increase customer lifetime value (CLTV), organizations need to understand the correct behavioral attributes and build predictive models using new and traditional data science techniques like k-means or spatio-temporal algorithms [14]. This helps in selecting the correct behavioral traits based on transactions and other demographic behaviors to identify customer churn and determine if a customer is a good candidate to be retained. Once identified, measures can be undertaken to prevent customer churn before it is too late.

Acknowledgements: Thank you to our DSC 500 Colleagues and Professor Shankar Parajulee