D211 - Advanced Data Acquisition

Richard M. Flores[1]

Abstract
The telecommunications industry is an ever-evolving and increasingly technologically sophisticated industry and as Borje Ekholm once said, “The telecom industry reaches into every corner of our economies, societies, and private lives, and it is one of the greatest drivers of economic growth and human equality the world has ever seen”. We are currently on the cusp of a new communications revolution with the launch of 5G networks that will reach speeds and distances we never imagined possible.

Keywords
Tableau — Telecom — Churn

[1] *Master of Science, Data Analytics, Western Governors University, United States*
*__Contact Author__: rflo147@wgu.edu

Contents

Introduction

For accessibility and to simplify grading, this Assessment Essay is broken down into 6 subsections which correlates to the 6 sections listed on the grading rubric under Part 3: Reflection Essay.

1. Assessment

1.1 Purpose and Function

The assignment given for the Master of Science, Data Analyst study is to place yourself into the shoes of an analyst for a well-known telecom company that services customers in the United States. In this scenario, the business goal of the analysis is to study customer trends and patterns to gain an understanding of common characteristics in order to reduce churn and to predict customers at high risk of churn. The Executive Dashboard created for this study focuses on key metrics of customer churn highlighting key characteristics such as tenure, revenue generated, data consumption, and geolocation. This information is presented in an easy-to-read format to help shareholders and decision makers gain insight from customer patterns and make wise revenue-generating choices.

1.2 Business Intelligence Tools

For this assessment a combination of three business intelligence tools were used to create the analysis and Executive Dashboard. The professional programs utilized included PGAdmin, Tableau Prep Builder, and Tableau. The programs were provided by WGU on a Hosted Virtual Machine with the applications pre-installed. PGAdmin is a powerful Guided User Interface for accessing and manipulating the PostgreSQL database. PGAdmin allows users to create query tools with colored syntax highlights, a datagrid for displaying and entering data, and an ERD Tool for creating and documenting schemas. In this analysis, PGAdmin was utilized to access the WGU PostgreSQL database and create a new schema for the supplementary competitor database. Tableau Prep Builder is included in the Tableau Suite of products and is used for data preparation. Tableau Prep can be used to combine, shape, and clean data. In this assessment, Tableau Prep was utilized mainly for data cleaning and to verify missing or null values. Data preparation is an essential and crucial step in the process of data analysis and the use of Tableau Prep allows users to easily drag and drop values and create flow steps to filter, rename, pivot, join and union data saving much time and effort versus creating Python or SQL scripts which have the same results. Tableau 2021 was the final program used in the analysis and also used to create the Executive Dashboard. Tableau allows functionality for importing the PGAdmin created PostgreSQL databases and to quickly observe and manipulate data. Tableau allows for drag and drop functionality to create simple SQL JOIN and Union scripts. Tableau also allows creation of customized workbook scripts which can be combined into dynamic and powerful data illustrations on our Executive Dashboard.

1.3 Data Preparation Steps

Cleaning and preparing the data for analysis included taking the following actions. Importing the supplementary dataset into Tableau Prep Builder. Verifying descriptions of the dataset including structure and data types. Review summary statistics of the supplementary dataset. Check records in the dataset for missing or null data and impute values with central tendency measures i.e., mean, median, mode. Remove outliers which are outside the bounds of standard deviations above the calculated mean. View visualizations for univariate and bivariate data categories. Import the dataset from Tableau Prep into PostgreSQL database and create logical relationships between the WGU dataset and Competitor dataset to secure referential integrity by setting primary and foreign keys.

1.4 Dashboard Creation

Creating an Executive Dashboard is a straight forward process once the data cleaning and preparation phase has been completed. The first step in the process is importing both the WGU and Competitor datasets into Tableau from PGAdmin connected to the PostgreSQL database. Once the data has been collected, we then verify referential integrity by checking primary and foreign key associations. Once verified we can create our SQL scripts casting data types, creating unions, and creating necessary Full, Inner, and Outer joins. Once the data is set, the next step is to create the necessary analysis and worksheets. Creating the worksheets includes pages for interactive controls and menus, summary metrics, and data representations. Once the worksheets are completed, we can compile the information and visualizations into the Executive Dashboard. The final step is to upload the Executive Dashboard to the Tableau Public Server enabling easy access for audiences and a simple way to share the dashboard without the need to install the Tableau program on the user's computer.

1.5 Data Analysis Results

The WGU Dataset consists of 10,000 customers and 50 columns or variables with information detailing customer churn, additional service subscriptions, account information, and demographics. The supplementary Dataset is provided by IBM Cognos Analytics which is a fictional telco company that provides home phone and Internet services to 7043 customers. Multiple important demographics are included for each customer, as well as a Satisfaction Score, Churn Score, and Customer Lifetime Value (CLTV) index. Analysis of this dataset allows us to compare and contrast key characteristics and trends of customer churn. From the analysis we gained some valuable actionable insights. One insight is that during the Exploratory Analysis, a trend common to both WGU and Competitor datasets is customer churn tends to happen most often after a one-year period. Another valuable insight, in both the WGU and Competitor Data, we find that customers who contribute to churn have an average bill 22 percent higher than loyal customers. Finally, our greatest actionable insight into customer churn is one that is again common to both WGU and the competitor's customers. The competitor's customers specifically state, and we can infer the same from WGU surveys, the greatest contributing factor to customer churn is the customer service experience. The biggest reason customers choose to discontinue services is a direct result from a poor customer service experience. By studying a competitor Telcom companies' data and comparing to the WGU data, we can turn this understanding into key and actionable insights that will allow our company to produce revenue protecting and generating policies.

1.6 Limitations

The greatest limitation of the data analysis is the limited size and scope of the provided and publicly available datasets. In a real-world scenario, a telecommunications company would have millions or at the least hundreds of thousands of customers. While we can gain meaningful insights from available data, not having a dataset with a realistic size of customers limits the analysis and data techniques available such as regression, classification, and machine learning models. However, the advanced data acquisition assignment provides meaningful and excellent practice in database management and data manipulation using industry approved data analytics software. Another limitation of the experience is the inability to communicate with the stakeholders at the Telcom company or any subset such as data engineers, executives, or even customer representatives. Without communication with Telcom employees it is hard to gain clarification about data or to provide targeted insights into churn metrics. Access to personnel and meetings with decision makers would provide a better alignment for reaching data analytics goals and providing the most benefit from actionable insights.

Acknowledgments

Special thanks to Course Instructor William Sewell for his invaluable resources in completing this assessment, and to my Mentor Emil Stoica for his never-ending support and encouragement. [1, 2].

References

[1] About Tableau Prep. Online help. (tableau help). *Retrieved February 22, 2022*, from:https://help.tableau.com/current/pro/desktop/en–us/storybestpractices.htm, 2022.

[2] PGAdmin Features. *Page, D.* Retrieved February 11, 2022, 2022 edition, from https://www.pgadmin.org/features/.