D211: Advanced Data Acquisition

Devin Belden

The purpose and function of the dashboard created within Tableau is to share with stakeholders the relationships between the Telecom company’s service rates and the utility rates of the other utility companies that service the same state. This is accomplished by first establishing that customer churn is visually correlated with the monthly amount charged to customers. In short, the higher the monthly charge, the more likely a customer will cancel their service. By using the relationship to the local utility rates, it can be determined which states to focus on in order to reduce churn without having to lower prices nationwide.

Tableau was chosen as the business intelligence tool due to its ease of use for the data consumer. Creating a dashboard is also relatively easy, especially when compared to a separate tool like Microsoft Excel. Additionally, there is a version of this tool that is completely free, so a small business can still perform basic data analyses and create storyboards for their stakeholders with no startup cost and very little risk.

In order to integrate the data within the additional utility rates CSV file into the existing Churn database within pgAdmin, it was necessary to do some rudimentary cleaning and aggregation. First, the “states\_raw” table was created within the Churn database to house the utility rate data. Next, the table was populated with data from the raw CSV file. A separate table to house the aggregated data, called “states”, was then created and filled using a SQL SELECT command to aggregate, using the average, all the utility pricing data by state. The state column was set as the primary key. Finally, in order to maintain referential integrity, empty rows for Puerto Rico, District of Columbia, and West Virginia were appended to the “states” table. This allowed for the “location” table to reference the “states” table using the common “state” column as a foreign key.

To create the dashboards within Tableau, a connection was made to the churn database within the SQL server located in pgAdmin. This enables Tableau to pull in all relevant data in one step. The Customer table is dragged into the data connection pane, followed by the Location table, and finally the States table. Location and Customer are joined on the Location ID column, and Location and States are joined on the State column as mentioned before.

For Dashboard 1, the average residential utility rate by state is plotted as a bar chart, next to another bar chart plotting average monthly charge by state. Both of these charts are created by placing the average of each rate to the Columns section, and the State into the Rows section.

For Dashboard 2, a simple visual plotting Churn by Average Monthly Charge is created, showing a clear correlation between the two values. Several measures are then created, including the Average, Standard Deviation, and Z score of the Monthly Charge and Res Rate columns. Next, the difference between the Z scores is calculated as a separate measure as well. This is plotted as a bar chart, separated by state. The Churn vs No Churn bars are stacked to save space, and each bar can be hovered over to view the individual amounts. A filter is added to distinguish between Churn and No Churn sections of both charts on this Dashboard as well.

The results of the data analysis illustrate three main points. First, customers who experience a higher absolute monthly charge are more likely to cancel their service. Second, the telecom company charges largely similar rates across all states, whereas there is a much higher degree of variability in the utility rate prices by state. This illustrates that dynamic pricing based on location is worth investigating. Third, price changes can be lowered by a small amount in certain states and possibly lead to a large decrease in churn for those states.

Perhaps the biggest limitation of this study is that it does not take into account the customer’s overall satisfaction with various aspects of the company’s service. For example, a customer could cancel service based on a poor experience with a company representative, or from experiencing an equipment failure. Additionally, it does not take into account a customer’s income. It can be inferred from each state’s utility rate what a customer is able to pay for services they consider essential, but it would be better to integrate that data directly into the project.

*Note: The extra dataset was taken from The US Department of Energy (Department of Energy - U.S. electric utility companies and rates: Look-up by Zipcode (2023))*

Works Cited

Department of Energy - U.S. electric utility companies and rates: Look-up by Zipcode (2023). Catalog. (n.d.). https://catalog.data.gov/dataset/u-s-electric-utility-companies-and-rates-look-up-by-zipcode-2023/resource/f95f12d2-9e6c-4e85-90eb-7f128745f9ba