Western Governors University

School of Technology, College of IT

Master of Science, Data Analytics

D211 Performance Assessment

Tyson Biegler

Student ID: 012170282

D211 - Advanced Data Acquisition

08, 17, 2024

**Installation steps:**

1. Extract the files from the tb.zip file.
2. Open the D211.sql file in the query tool in pgadmin.
3. Run the SQL code.
4. Lastly, In Tableau, open the Dashboard\_D211\_LOD4.twbx Tableau packaged workbook file.

**C1. Explanation of functionality:**

The performance assessment scenario section suggests that telecom companies can experience churn rates of up to 25%. (WGU, n.d). Amy Gallo of Harvard Business Review: *“...acquiring a new customer is anywhere from five to 25 times more expensive than retaining an existing one” (****Gallo, 2014****).* So, identifying the customers and the factors that lead to churn is essential for a business to continue operating. This dashboard will allow the executives to filter demographic information based on churn, targeting specific individuals that match a profile or target a particular area that may be contributing to churn.

**2. Justification of the tools:**

This dashboard was built using Tableau. Tableau allows users to connect directly to a database or manually to individual files. In this project, I used both of these methods.

**3. Explanation of the steps taken to prepare the data:**

The external CSV file I downloaded from Kaggle contains census data by state and does not require any cleaning. To use calculated fields based on the churn rate, I needed to convert ‘Churn’ to binary and then to numeric. I did so using SQL in pgAdmin.

I used The following SQL code to convert ‘Churn’ to binary and then to numeric type.

-- converting Churn to binary

UPDATE customer

SET Churn = CASE

WHEN Churn = 'Yes' THEN 1

WHEN Churn = 'No' THEN 0

END;

-- converting churn to integer

ALTER TABLE customer

ALTER COLUMN Churn TYPE INTEGER USING Churn::integer;

After the conversions, I checked if any NULL values were introduced and found none were.

-- Checking if any null values were introduced

SELECT COUNT(\*)

FROM customer

WHERE Churn IS NULL;

**4. Explanation of the steps taken to create the dashboard:**

This dashboard consists of 6 charts and 3 KPIs. The charts were created using various measures, dimensions, and calculated fields.

**Calculated fields:**

Market penetration represents the ratio between the total population of a state that was retrieved from the external CSV file and the count of customers

**KPIs:**

1. The churn rate was calculated by getting the sum of ‘Churn’ and dividing it by the count of customers using the customer ID. The results were then multiplied by 100 to get the number as a percentage.
   1. SUM([Churn])/COUNT([Customer Id])\*100
2. Average tenure was calculated using a measure (average) of tenure.
3. Similarly, the average monthly charge was calculated using a measure (average) of monthly charges.

**Charts:**

1. Bar chart of average churn based on income (bins).
   1. I overlaid the total customers in each bin using dual axis.
2. Stacked bar chart of age (bins) separated by churn status.
3. A stacked bar chart displays the number of customers with each number of children. This is also separated by churn.
   1. This chart counts how many customers have 0 children, one child, two children, etc., and then distinguishes between their churn status.
4. Map chart displaying each state's churn rate, market penetration and total customers. This chart will allow an executive to see the churn rate and market penetration by state.
5. The last two graphs display the top 10 and bottom 10 states based on market penetration.

**5. Discuss the results of your data analysis and how it supported the purpose and function of your dashboard.**

California has the lowest number of customers based on state population. Iowa has the most customers in relation to its population. Customers whos income is greater than 120K have a higher than average churn rate. However, customers in this income bin only represent a small proportion of the customers. Most customers make less than $60k. The churn rate across ages appears to be uniform. Lastly, the monthly charge of customers who have not churned is $163.01 while the churned customers have an average monthly charge of $199.30.

**6. Explanation of limitations:**

One major limitation, and perhaps the most detrimental regarding analysis, is the lack of time data, specifically yearly or quarterly dates. Because of this, the dashboard can not accurately display revenue changes over time or the impact of specific services on revenue.

The database is more of a list of facts about the customers rather than a financial dashboard. However, the dictionary describes a relationship between a company's revenue and customers' churn. Retaining customers safeguard against revenue loss far better than earning new customers. However, without knowing the data date ranges and creating new tables for calculations across time, any monthly revenue calculations would have to make arbitrary assumptions about the year or quarter in which the data was gathered.

**D. Web sources**

1. **Kaggle.com.** (n.d.). *us-population-by-state* [Data set]. Retrieved August 8, 2024, from https://www.kaggle.com/datasets/alexandrepetit881234/us-population-by-state

**E. Additional Sources**

1. **Gallo, A.** (2014, October 29). The Value of Keeping the Right Customers. Harvard Business Review. Retrieved August 10, 2024, from https://hbr.org/2014/10/the-value-of-keeping-the-right-customers
2. **Western Governors University.** (n.d.). *Advanced Data Acquisition - D211*. WGU. Retrieved August 8, 2024, from https://tasks.wgu.edu/student/012170282/course/34320018/task/4303/overview

**F. Panopto video:**

https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=b53e9162-b6db-4e71-b8f3-b1cf012f5f81