Briana Churchill
D211
Instructor Sewell
3 February 2024

DATA DASHBOARD AND STORYTELLING REFLECTION

A2. Dashboard installation instructions

Background knowledge in data analytics is not required to be able to download and view the dashboard. All one needs to do is download and install Tableau Reader on their computer. The dashboard is provided with this assessment as a packaged file (.twbx). The dashboard can be viewed for free within the Tableau Reader desktop application. Tableau Reader keeps the dashboard fully interactive and allows the viewer to use the filters and navigation aspects of the dashboard described in the Panopto video, and section A3 below.

To download the Tableau Reader desktop application, the viewer can type the following URL into their web browser:
https://www.tableau.com/products/reader
"After going to the web address above, the viewer would then click the "Download Tableau Reader" button to the left. A new page would appear requiring the user to fill out contact information before clicking a button at the bottom of the page that says "DOWNLOAD." After filling out the required information and clicking the download button, the user would then follow the installation instructions provided by the developer (usually a pop-up message after attempting to "open" the application). Additionally, Tableau Reader may require the user to create a login using the information provided prior to downloading, and the user should look for messages indicating so. Once Tableau Reader is installed, the viewer can open the dashboard file (.twbx) provided to them by simply saving the file to their computer, and double clicking on the file." (Churchill, 2024)

A3. Dashboard navigation

Like the dashboard provided for the D10 assessment, I created a story with two dashboards related to the customers, an introduction page, and a summary page. Once the viewer has the file opened within Tableau Reader, the introduction page will already be open and the first thing the viewer sees. At the top is the story title, "Telecom Customer Insights" and four boxes depicting the pages within the story. Underneath the boxes, are instructions for how the viewer can navigate through the pages, which states "Click one of the boxes or arrows above to navigate through the story." In the center of the page is the introduction information, indicating that the dashboard is a comparison of senior-citizen customers for WGUco and Telco in California. "WGUco" is a fictional name I created for the WGU data; Telco is the company name in the additional data set used. Lastly, at the bottom of the page I introduce myself.

The next page, which can be opened using the arrows or clicking the applicable box, is titled "Customer Demographics." In this dashboard are three visualizations. On the left, is a packed bubbles graph which depicts the number of customers for both companies with or without a partner/significant other. To the right are two more visualizations. On the top right, a

horizontal bar graph showing the average tenure of customers measured in months. Below this graph, are pie charts to show the percentage of customer's reported gender identity. Only male and female are being reviewed. The viewer then can use the boxes on the top of the page, or the arrow – to navigate to the next worksheet.

The next page is about customer KPIs, which does include filters. To the top left is a visualization of a treemap showing the percentage of customers within a specific contract length, either month-to month, one-year, or two-year. Underneath this viz is a filter, with instructions beside it and allows the user to toggle between stats for customers with a specific contract length. The filter regarding the "contract length" applies to all four visualizations on the page. To the right of this filter, is another titled "Source" and applies to all four visualizations. Additionally, instructions for the filter are found beside it. This filter is directly below the second visualization found on the top right of the page. This viz is titled "Average Customer Charge" and is a horizontal bar graph depicting the dollar amount of the average charge per customer, per company. On the bottom left are more horizontal bar graphs, a visualization titled "Payment Method" that depicts the different payment methods for each company. To the right of this graph, is the final visualization on the page titled "Customer Churn" which gives the total for customers that have and have not churned.

The last page is titled "Summary" and can be viewed by clicking the last box or clicking the right arrow next to the boxes at the top. Like the introduction page, the summary page is a text box. Within this text box are two main points, key takeaways, and recommended course of action.

A4. Provide a copy of all SQL code

```
--Change yes/no values to true/false
ALTER TABLE public.customer
ALTER churn TYPE bool USING CASE WHEN churn ='No' THEN FALSE ELSE TRUE
END;
ALTER TABLE public.customer
ALTER COLUMN churn SET DEFAULT FALSE;
--Rename Monthly_Charge column for consistency
ALTER TABLE public.customer
RENAME COLUMN monthly_charge TO charge;
--Rename children column for consistency
ALTER TABLE public.customer
ADD Dependents BOOLEAN;
UPDATE public.customer
SET Dependents = CASE WHEN children > 0 THEN TRUE ELSE FALSE END;
ALTER TABLE public.customer
DROP children;
--Change marital status if married to true partner, all else false - for consistency
ALTER TABLE public.customer
ADD partner boolean;
UPDATE public.customer
SET partner = CASE WHEN marital = 'Married' THEN TRUE ELSE FALSE END;
ALTER TABLE public.customer
```

```
DROP marital;
--Rename payment_type column for consistency
ALTER TABLE public.payment
RENAME COLUMN payment_type TO payment_method;
--Change state to bool true/false to differentiate between customers in CA
ALTER TABLE public.location
ALTER state TYPE bool USING CASE WHEN state ='CA' THEN TRUE ELSE FALSE END;
-- Create a column for senior-citizens
ALTER TABLE public.customer
ADD COLUMN senior_citizen BOOLEAN;
UPDATE public.customer
SET senior_citizen = TRUE
WHERE age >= 65;
--Drop null values
DELETE FROM public.customer
WHERE senior_citizen IS NULL;
--Drop false values
DELETE FROM public.customer
WHERE senior_citizen IS FALSE;
--Rename duration column for consistency
ALTER TABLE public.contract
RENAME COLUMN duration TO contract;
-- Update names in Contract column for easier understanding
UPDATE public.contract
SET contract = REPLACE(contract, 'Month-to-month', 'Month');
UPDATE public.contract
SET contract = REPLACE(contract, 'Two Year', 'Two');
UPDATE public.contract
SET contract = REPLACE(contract, 'One year', 'One');
--Add source column
ALTER TABLE public.customer
ADD COLUMN source text;
UPDATE public.customer
SET source = 'WGUco';
--Change numerical columns to decimal and limit decimal points to 2
ALTER TABLE public.customer
ALTER COLUMN tenure TYPE DECIMAL (10,2);
ALTER TABLE public.customer
ALTER COLUMN charge TYPE DECIMAL (10,2);

--Create telco table to bring in additional dataset
CREATE TABLE telco_customers
(
Customer_id text, count text, country text, state text, city text,
zip_code text, lat_long text, latitude text, longitude text, Gender text,
Senior_citizen text, Partner text, Dependents text,
```

```
Tenure_Months text, Phone_Service text, Multiple_Lines text,
Internet_Service text, Online_Security text, Online_Backup text,
Device_Protection text, Tech_Support text, Streaming_TV text,
Streaming_Movies text, Contract text, Paperless_Billing text,
Payment_Method text, Monthly_Charges text, Total_Charges text,Churn_Label text,
Churn_Value text, Churn_Score text, CLTV text, Churn_Reason text,
CONSTRAINT "customer_id_fkey" PRIMARY KEY (customer_id)
);
ALTER TABLE telco_customers
OWNER to postgres;
--Import Telco data
COPY telco_customers
FROM 'C:\Users\Public\Downloads\telco.csv'
DELIMITER ','
CSV HEADER;
--rename applicable columns for consistency
ALTER TABLE telco_customers
RENAME COLUMN monthly_charges TO charge;
ALTER TABLE telco_customers
RENAME COLUMN churn_label TO churn;
ALTER TABLE telco_customers
RENAME COLUMN tenure_months TO tenure;
 --Change yes/no values to true/false
ALTER TABLE telco_customers
ALTER senior_citizen TYPE bool USING CASE WHEN senior_citizen ='No' THEN FALSE
ELSE TRUE END;
ALTER TABLE telco_customers
ALTER COLUMN senior_citizen SET DEFAULT FALSE;
ALTER TABLE telco_customers
ALTER partner TYPE bool USING CASE WHEN partner ='No' THEN FALSE ELSE TRUE
END;
ALTER TABLE telco_customers
ALTER COLUMN partner SET DEFAULT FALSE;
ALTER TABLE telco_customers
ALTER dependents TYPE bool USING CASE WHEN dependents ='No' THEN FALSE ELSE
TRUE END;
ALTER TABLE telco_customers
ALTER COLUMN dependents SET DEFAULT FALSE;
ALTER TABLE telco_customers
ALTER churn TYPE bool USING CASE WHEN churn ='No' THEN FALSE ELSE TRUE
END;
ALTER TABLE telco_customers
ALTER COLUMN churn SET DEFAULT FALSE;
-- Update names in Contract column for easier understanding
UPDATE telco_customers
SET contract = REPLACE(contract, 'Month-to-month', 'Month');
```

```
UPDATE telco_customers
SET contract = REPLACE(contract, 'Two Year', 'Two');
UPDATE telco_customers
SET contract = REPLACE(contract, 'One year', 'One');
-- Change numeric columns to decimal and limit decimal points to 2
ALTER TABLE telco_customers
ALTER COLUMN tenure TYPE DECIMAL (10,2) USING tenure::DECIMAL(10,2);
ALTER TABLE telco_customers
ALTER COLUMN charge TYPE DECIMAL (10,2) USING charge::DECIMAL(10,2);
--Add source column
ALTER TABLE telco_customers ADD COLUMN source text;
UPDATE telco_customers SET source = 'Telco';
--Remove customers that are not senior citizens
DELETE FROM telco_customers
WHERE senior_citizen = FALSE;
--Create second table for organization and consistency
CREATE TABLE telco_service AS
SELECT
churn,
tenure,
charge,
contract,
payment_method,
customer_ID
FROM
telco_customers;
ALTER TABLE telco_customers
DROP churn,
DROP tenure,
DROP contract,
DROP payment_method;
SELECT state
FROM telco_customers;
--Change state to bool true/false to differentiate between customers in CA
ALTER TABLE telco_customers
ALTER state TYPE bool USING CASE WHEN state ='California' THEN TRUE ELSE FALSE
END;
-- Create column with unique characters for PK for service table
ALTER TABLE telco_service
ADD COLUMN unique_id SERIAL PRIMARY KEY;
UPDATE telco_service
SET unique_id = DEFAULT;
--Create FK for service table
ALTER TABLE telco_service
ADD CONSTRAINT fk_constraint_name
FOREIGN KEY (customer_id) REFERENCES telco_customers (customer_id);
```

```
--Create new WGU and telco tables to join existing tables, and then UNION them all
CREATE TABLE wgu_table AS (
SELECT cust.churn, cust.tenure, cust.charge,cust.gender,
cust.dependents, cust.partner, cust.senior_citizen,
cust.source, cust.customer_id, pmt.payment_method, con.contract, loc.state
FROM public.customer AS cust
JOIN public.payment AS pmt ON cust.payment_id = pmt.payment_id
JOIN public.contract AS con ON cust.contract_id = con.contract_id
JOIN public.location AS loc ON cust.location_id = loc.location_id);
CREATE TABLE telco_table AS (
SELECT tserv.churn, tserv.tenure, tserv.charge, tcust.gender,
tcust.dependents, tcust.partner, tcust.senior_citizen, tcust.source,
tcust.customer_id, tserv.payment_method, tserv.contract, tcust.state
FROM telco_service AS tserv
INNER JOIN telco_customers AS tcust
ON tserv.customer_id = tcust.customer_id);
CREATE TABLE both_df AS (
SELECT *
FROM wgu_table
UNION ALL
SELECT *
FROM telco_table);
DROP TABLE IF EXISTS wgu_table;
DROP TABLE IF EXISTS telco_table;
--Drop false values for state
DELETE FROM both_df
WHERE state IS FALSE;
--Review data
SELECT * FROM both_df;
```

C1. Purpose and function of dashboard

The dashboards created in Tableau for the purpose of this analysis included visualizations of data from WGU-MSDA D211 coursework and a fictional company named Telco, retrieved from Kaggle. Both datasets include data related to customer churn. Both datasets had similar columns regarding customers demographics, services, and charges. Comparing WGUco data to that of our leading competitor Telco provides the necessary information needed for our senior executives. The Regional VP will benefit from the information provided regarding customers in the state of California. Customer data involving service usage, demographics, and contract specifications is viewable for interpretation for all stakeholders in this dashboard. The visualizations provide detailed information that can be utilized by both the executive vice president of sales, and the senior vice president for customer experience to improve customer retention and recruitment.

C2. Business Intelligence Tool Used

Having signed up for the one-year free student version of Tableau within the last month, it seemed most appropriate to use for this assessment. Additionally, I am using the same additional dataset as previously used in D210, so I am familiar with creating visualizations for the data. Datacamp instructions provided in D210, and D211 for Tableau also made navigating and creating the application very simple. Tableau is very user friendly and makes visually appealing charts and graphs, making the presentation of the dashboard a breeze.

C3. Explain the steps used to clean and prepare the data for analysis

Considering that the WGU dataset used for this assessment is of familiar data used in previous assessments, I knew that some additional cleaning would be necessary. Because I felt comfortable with the WGU data, and even the Telco data from the previous assessment – I went right into cleaning and preparing the datasets. The steps taken are as follows:

For WGU dataset

- Changed yes/no values to true/false (Boolean).
- changed all the column names to be consistent and common between the datasets.
- Changed children column to dependents, making values of >= 0 being false, all else being true.
- Changed marital status column for consistency to partner, if 'married' = true, all else false
- Fixed minimum zip code by casting and front-filling with zeros.
- Created a new column for senior-citizen, where all age >= 65 is true, all else false
- Dropped false values for senior-citizen column
- Dropped null values.
- Updated names in contract column for easier understanding. Rather than "month-to-month" have "month", and so on.
- Dropped customer data if their state was not ‘CA’
- Added a source column to be able to differentiate between WGUco and Telco customers.
- Changed numeric columns to decimal points and limited to 2.

For Telco dataset

- Created a table and imported the Telco data.
- Renamed columns for consistency, to match column names in the WGUco dataset.
- Change yes/no values to true/false (Boolean).
- Dropped customers not identified as senior-citizen, because Telco did not provide an age column, but only whether customers were senior-citizens or not.
- Created a new column for senior-citizen, where all age >= 65 is true, all else false
- Updated zip code column to be same datatype of WGUco column.

- Updated names in contract column for easier understanding. Rather than "month-to-month" have "month", and so on.
- Changed numeric columns to decimal points and limited to 2.
- Dropped customer data if their state was not ‘California’
- Added a source column to be able to differentiate between Telco and WGUco customers.
- Created a new table named "telco_service" which included information regarding customer packages, charges, tenure, etc.
- Moved all applicable columns to service and dropped them from telco_customers.
- Created a column with unique character to be the PK of the service table.
- Created foreign keys for both service table and demographics table, for relational integrity.

UNION of both WGUco and Telco data

Created two new tables for both the WGUco and Telco tables using INNER JOIN function. One table included both demographics and service columns needed for WGUco, and the other had the same for Telco. After I downsized four tables to two, I then combined the new WGUco table and new Telco table by using the UNION ALL function.

C4. Summarize the steps used to create the dashboards

From start to finish, here are the steps I used to create the dashboards:

- Imported the one CSV file of the combined datasets in Tableau.
- Starting with the first worksheet I created a packed bubbles graph using the partner and source columns, as well as the customer count.
    - The graph was labeled with the true/false value, company, and number of customers using the label marks card.
- Created a new worksheet with a horizontal bar graph to depict the average customer tenure in months.
    - Tenure was placed in the columns section, and source as the rows.
    - The Tenure measure was changed to the average rather than the sum.
- Created third worksheet of pie charts for customer genders.
    - Created a filter to only include "female" and "male" due to Telco dataset not having "prefer not to answer" as an option.
    - Added the count to the visualization using the label marks card. Added a quick table calculation to show the percentage of total.
    - Used the color marks card to use colors to depict the different genders, using colors I planned to use in another visual.
- All three of these visualizations were brought into a dashboard.
    - The dashboard was given a title, and I moved on to creating the worksheets for the next dashboard.
- Next, I created a new viz of a treemap that shows the percentage of customers in a particular contract length.

    - Using the labels marks card, I added the source, contract length, and percentage of total into the visualization.
    - I added a filter for the contract lengths and showed the filter.
- Next, I created a bar graph showing the different payment methods, and what number of customers utilize them.
    - In addition to being split by payment method type, the data is also divided between the two companies.
- Using a highlight table for the next visualization, I compared customer churn data between the companies.
    - Using the labels marks card, I again included the sum of customers and whether they have churned or not in the table.
- Lastly, I created one more viz using horizontal bars.
    - This visualization showed the average customer charge per customer, monthly.
    - Using the labels marks card, I used the quick table calculation to find the average.
    - Additionally, I formatted the labeled values to show as a currency.
    - I added a filter to choose between the companies and showed it for the story.
- Once all four of these visualizations were created, I moved them into a new dashboard and named that dashboard as well.
    - Additionally, I added a text box to include instructions underneath the filters mentioned above.
- Upon completion of the two dashboards above, I created a story.
    - After adding in the two dashboards, I created two new pages in the story.
    - An introduction page in the beginning, and a summary page at the end.
    - Using text boxes, I detailed the information regarding the presentation on both of those pages.

C5. Discuss the results of your data analysis and how it supports executive decision-making

Key Takeaways from the story are as follows:

- More customers are likely to utilize a month-to-month plan compared to contracted one-or-two year plans
- The average cost of service for WGUco customers is much higher in comparison to Telco
- Automatic payments take up the majority of payment methods for both WGUco and Telco, although not by much more for WGUco.

The takeaways support executive decision-making by bringing to light information regardinf our senior-citizen customers in California. Seeing that customers in this demographic prefer month-to-month plans, and could benefit from more automatic payment advertising could allow WGUco to bring up changes to address this information. Additionally, seeing how WGUco compares to our leading competitor allows us to see that market in another light. Seeing the market from another perspective also helps WGUco to see what kind of changes could be made to have better customer retention rates.

C6. Discuss the limitations of the data analysis

The largest limitations that I noted are:

- The D211 WGU dataset does not include customer information regarding add-on services.
- The number of customers available within the WGU dataset is extremely limited.

Although the data is limited by quantity of customers, it could have been of great benefit to still compare add-on services (streaming TV, etc.) because Telco had similar columns available. However, this information was not available for this assessment, and so the usable data overall was extremely limited. In this case, only 203 customers for WGUco were utilized in this analysis. In comparison to Telco's 1,142 it is hard to truly compare the companies and that regional area. The dashboard provided explains the data available but would benefit from additional WGUco data before any business decisions are made.

D. Sources

Churchill, Briana. (2024, January 12). Performance Assessment: Data Dashboard And Storytelling (NAM2). Assignment for MS Data Analytics Course D210. Western Governors University.

na. SQL – BOOLEAN. Tutorialspoint. Retrieved from <https://www.tutorialspoint.com/sql/sql-boolean-bit-operator.htm#:~:text=corresponding%20Boolean%20values.-,Boolean%20in%20MS%20SQL%20Server,can%20also%20accept%20NULL%20values.>

na. PostgreSQL RENAME COLUMN: Renaming a column. PostgreSQL Tutorial. Retrieved from <https://www.postgresqltutorial.com/postgresql-tutorial/postgresql-rename-column/>

Agarwal, H. SQL | Join (Inner, Left, Right and Full Joins). Geeksforgeeks. Retrieved from <https://www.geeksforgeeks.org/sql-join-set-1-inner-left-right-and-full-joins/>

Ravikiran, A.S. SQL UNION: The Best Way to Combine SQL Queries. Simplilearn. Retrieved from <https://www.simplilearn.com/tutorials/sql-tutorial/sql-union>

na. PostgresSQL tutorial. PostgreSQL Primary Key. Retrieved from

<https://www.postgresqltutorial.com/postgresql-tutorial/postgresql-primary-key/>

E. References

Churchill, Briana. (2023, October 2). Performance Assessment: Classification Analysis (NVM2). Assignment for MS Data Analytics Course D209. Western Governors University.

Churchill, Briana. (2023, June 21). Performance Assessment: Data Acquisition. (TGM1) Assignment for MS Data Analytics Course D205. Western Governors University.

Datacamp. D210 – Representing and Reporting. Retrieved from <https://app.datacamp.com/learn/custom-tracks/custom-d210-representation-and-reporting>

Datacamp. D211 – Advanced Data Acquisition. Retrieved from <https://app.datacamp.com/learn/custom-tracks/custom-d211-advanced-data-acquisition>

Tableau Reader. Retrieved from https://www.tableau.com/products/reader

"Telco customer churn: IBM dataset." Retrieved from https://www.kaggle.com/datasets/yeanzc/telco-customer-churn-ibm-dataset