

Aidan Soares

012042436

D211 – Advanced Data Acquisition

Western Governors University

May 17, 2024

A1 - Datasets: For this assessment I will be utilizing two datasets that focus on the customer demographics, tenure length, and churn status of telecommunications companies. The two datasets used can be found below:

- WGU's churn dataset found within the PgAdmin4 database list in the Lab Environment
- A dataset for a telecommunications company based in California during its 3rd quarter. This dataset contains demographic data similar to WGU's churn information which makes it a great candidate for industry and performance comparison. This dataset was sourced from (Blastchar, 2018): <https://www.kaggle.com/datasets/blastchar/telco-customer-churn>

A2 – Dashboard Installation: To install my dashboard, please follow the instructions below

1. Download the WA_Fn-UseC_-Telco-Customer-Churn.csv file from either the link above in A1 or from the provided dataset. **This file needs to be placed in the C:/Users/Public/Downloads folder as PgAdmin4 does not have permission to read files from the other folders.**
2. Open PgAdmin4.
3. On the Server menu on the left of the window, use the drop-down arrows to expand Servers > PostgreSQL > Databases > Churn.
4. With the Churn database selected click the Query Tool button above the Servers window.
5. Copy the code provided within the pgadmin4 .txt file.
6. Paste and run this code in the Query Tool window by highlighting all the code and clicking the Play button in the top right.
7. You will now be able to find the imported table and joined table by navigating the Server menu by using drop-down arrows on Churn > Schemas > Tables.
8. If the tables do not appear here after running the code, right click the Tables drop-down menu and click Refresh.
9. Open **Demographic and Retention Dashboard.twbx**.
10. The window will prompt you to sign in to the database
11. Insert the following information then click Sign In:
 - a. Username: postgres
 - b. Password: Passw0rd!
12. You have now installed my dashboard.

A3 – Dashboard Navigation: The dashboard is all laid out with the information presented without having to jump between any dashboards. The information present can be manipulated by utilizing the filter on the right side of the screen.

The Payment Id filter isolates the information present within the Churn Status, Gender Breakdown, and Tenure Assessment tables based on the payment form selected, the filter must be confirmed with the Apply button to prevent accidental filtering without the user's decision. This filter makes

the table more dynamic, shifting the information presented to focus on specific features that may give insight into areas of improvement that can yield better customer retention.

Additionally, if you use your mouse and hover over any aspect of a visualization such as a bar or a pie slice, you will be presented with a tooltip window that presents additional information that would otherwise render my visualizations cluttered. Said information can help the user get a more granular look into customer counts and labels that ensure better readability.

A4 – SQL Code:

```
--Create kaggle table with column names from csv
```

```
CREATE TABLE IF NOT EXISTS public.kaggle(
```

```
    customerid TEXT ,  
    gender TEXT ,  
    seniorcitizen INT,  
    partner TEXT ,  
    dependents TEXT ,  
    tenure NUMERIC,  
    phoneservice TEXT ,  
    multiplelines TEXT ,  
    internetervice TEXT ,  
    onlinesecurity TEXT ,  
    onlinebackup TEXT ,  
    deviceprotection TEXT ,  
    techsupport TEXT ,  
    streamingtv TEXT ,  
    streamingmovies TEXT ,  
    contract TEXT ,  
    paperlessbilling TEXT ,  
    paymentmethod TEXT ,  
    monthlycharges NUMERIC,  
    totalcharges TEXT ,
```

```

        churn TEXT ,

        PRIMARY KEY (customerid)

    );

--import kaggle dataset into created table
copy public.kaggle

    (customerid, gender, seniorcitizen, partner,
     dependents, tenure, phoneservice, multiplelines,
     internetservice, onlinesecurity, onlinebackup,
     deviceprotection, techsupport, streamingtv, streamingmovies,
     contract, paperlessbilling, paymentmethod, monthlycharges,
     totalcharges, churn)

    FROM 'C:/Users/Public/Downloads/WA_FN~1.CSV'

    DELIMITER ',' CSV HEADER;

--changing column names only for columns that are relevant for analysis
ALTER TABLE public.kaggle

    RENAME COLUMN customerid to customer_id;

ALTER TABLE public.kaggle

    RENAME COLUMN monthlycharges to monthly_charge;

ALTER TABLE public.kaggle

    RENAME COLUMN paymentmethod to payment_id;

ALTER TABLE public.kaggle

    RENAME COLUMN contract to contract_id;

--re-expressing payment method and contract names to recorded IDs

```

```
UPDATE public.kaggle
```

```
SET contract_id =
```

```
    CASE
```

```
        WHEN contract_id = 'Month-to-month' THEN '1'
```

```
        WHEN contract_id = 'One year' THEN '2'
```

```
        ELSE '3'
```

```
    END;
```

```
UPDATE public.kaggle
```

```
SET payment_id =
```

```
    CASE
```

```
        WHEN payment_id = 'Bank transfer (automatic)' THEN '1'
```

```
        WHEN payment_id = 'Credit card (automatic)' THEN '2'
```

```
        WHEN payment_id = 'Electronic check' THEN '3'
```

```
        ELSE '4'
```

```
    END;
```

```
--altering id columns from the kaggle table to be integers instead of text
```

```
ALTER TABLE public.kaggle
```

```
    ALTER COLUMN payment_id TYPE INT USING payment_id::integer,
```

```
    ALTER COLUMN contract_id TYPE INT USING contract_id::integer;
```

```
--adding a differentiator column for the WGU dataset and its competitor
```

```
ALTER TABLE public.kaggle
```

```
    ADD company TEXT
```

```
    DEFAULT 'Kaggle';
```

```
ALTER TABLE public.customer
```

```

ADD company TEXT

DEFAULT 'WGU';

--rounding tenure and churn in the WGU customer data to match competitor data
UPDATE public.customer

    set tenure = ROUND(tenure, 0);

UPDATE public.customer

    set monthly_charge = ROUND(monthly_charge, 2);

--creating a joined table
CREATE TABLE public.combined AS

    SELECT customer_id, gender, tenure, churn,

    monthly_charge, payment_id, contract_id, company

    FROM public.customer

    UNION ALL

    SELECT customer_id, gender, tenure, churn,

    monthly_charge, payment_id, contract_id, company

    FROM public.kaggle;

```

B - Panopto Video: My Panopto video can be found here:

<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=a309b04d-d643-4fdb-bdd7-b173015e95a7>

C1 – Dashboard Alignment: The objective of my dashboard is to provide stakeholders a high-level overview of key customer demographics information and digestible analysis of the service offerings or features that influence the tenure of customers. These key metrics provide significant insight into the relevant criteria necessary to bolster customer retention in a highly competitive environment, allowing decision makers to determine the contract structure, or payment method that they should strive for customers to utilize. The dashboard provides a direct comparison to an industry

competitor to review industry trends, assess WGU Telecom Co.'s performance in relation, and identify ways to increase retention.

C2 – Tableau Justification: For this assessment I used Tableau to generate my dashboard as it provides easy to navigate visualization building tools with a wide variety of visualization tools, labels, dashboard options, and storytelling features. It also allows for seamless connection to a PostgreSQL server, making the data sourcing stage more efficient through dynamic data updating. As such, Tableau serves as a perfect platform to generate the dashboard I need using the data available to me.

C3 – Cleaning and Preparing the Data: To prepare the data for use in my dashboard, I conducted cleaning and transformation on the datasets provided by WGU and sourced externally in multiple stages. After importing the external dataset to a table within my churn database, I converted the names of the Kaggle dataset to match that of WGU's for later table concatenation. I then re-expressed the Contract and Payment information in the Kaggle dataset to match the ID charts as per WGU's churn database:

Contract ID	
Month-to-month	1
One year	2
Two year	3

Payment ID	
Bank transfer (automatic)	1
Credit card (automatic)	2
Electronic check	3
Mailed check	4

I then changed the datatype for these columns to integers. Next, I added a new column in both the Customer table and Kaggle table representing the data's respective company. The last cleaning stage before combining the two tables together was to round the Customer tables tenure and monthly charge rates so that both tables would match.

Once the above steps were completed, I created a new table to house the concatenation of the two tables. I used only the most relevant columns to my assessment such as information like customer id, gender, tenure, churn, payment type, contract type, monthly charges, and company.

C4 – Creating the Dashboard: The dashboard submitted for this assessment was broken into multiple parts: data source, visualizations, dashboard. To begin, the combined table Data Source must be selected as per the instructions in A2.

Creating the visualizations

1. Click New Worksheet and title it "Customer Demographics"

2. Right click on the Contract Id table from the Tables box on the left side of the screen and click Aliases. Rename the aliases to match the name of the Contract type to match its Id (1 is month-to-month, 2 is One year, 3 is Two year) to achieve appropriately named columns
3. Drag the Contract Id table from the Tables box to Columns, repeat this but dragging Contract Id to Rows
4. Right click on Contract Id under Rows and select Measure > Count
5. Hold Ctrl and drag Contract Id from Columns to the Color box under Marks
6. Hold Ctrl and drag CNT(Contract Id) from Rows to the Label box under Marks
7. Right click on the CNT(Contract Id) Label and click Quick Table Calculation > Percentage of Total
8. Drag Tenure from the Tables box on the left to the Label box under Marks
9. Right click on the Tenure Label and select Measure > Average

10. Repeat steps 1 through 9 for a new worksheet called "Payment Demographics". Using Payment Id instead of Contract Id. For step 2, the aliases are as follows: 1 is Bank Transfer (automatic), 2 is Credit card (automatic), 3 is Electronic check, 4 is Mailed check

11. Create a New Worksheet titled "Churn status"
12. Drag the Churn table to Columns and Rows
13. Right click on the Churn table in Columns and select Measure > Count
14. Drag the Company table to the Label box
15. Hold Ctrl and drag the CNT(Churn) table from Columns to the Label box
16. Right click the CNT(Churn) Label and select Quick Table Calculation > Percent of Total
17. Right click it again and select Edit Table Calculation
18. In the Table Calculation box, select Compute Using Specific Dimensions and deselect Company

19. Create a New Worksheet titled "Gender Breakdown"
20. Drag the Gender Table to both Columns and Rows
21. Right click on the Gender table in Rows and select Measure > Count
22. In the top right, click on Show Me and select Pie Chart
23. Drag the Company Table to Columns
24. Drag the Gender Table to the Label box
25. Right click on the Gender Label and select Measure > Count
26. Right click on it again and select Quick Table Calculation > Percent of Total
27. Right click on it again and select Edit Table Calculation and select Compute Using Table (down) to get percentage by company

28. Create a New Worksheet titled "Tenure Assessment"
29. Drag Tenure table to both Columns and Rows
30. Right click on Tenure in Rows and select Measure > Count
31. Click on Show Me in the top right corner and select Area Chart

32. Drag Company table to the Color box under Marks
33. Drag Payment Id table to the Filters box

Creating the dashboard

1. Click New Dashboard and name it Customer Demographics and Retention Assessment
2. Drag all the created sheets to the dashboard
3. Drag a text object to the top of the dashboard and title it “Customer Demographic and Retention Assessment” with size 12 font
4. Use the drop-down menu for the Company box on the right side of the dashboard, select Floating and drag it over to the Tenure Assessment visualization for a floating legend

C5 – Results of Data Analysis: The primary of my dashboard was to summarize and present the customer demographics and trend towards longer tenures, so that stakeholders of the WGU Telecom company will be able to establish a focus on specific elements (such as contract structure and payment format), that the company can encourage customers to shift towards.

From the dashboard generated, it is noted that customers who have subscribed on a Two Year contract, and utilize some form of automatic payment resulted in the longest tenures within the data acquired. As it currently stands, a strong majority of customers have month-to-month contracts and use checks to pay for their subscriptions. Manually submitting checks can be tedious, both for the consumer and on the corporate processing end, but it is possible that the customers who use this payment method may not be aware of the alternatives that exist. This is further supported when using the filter to isolate customers that only use Bank Transfers; the Tenure Assessment chart then changes to reflect that the competitor yields much more customers at the highest tenure length when they use bank transfers. As such, establishing an informational marketing campaign around alternative payment methods could certainly encourage customers to transition towards a simpler payment system that yields better retention.

In addition, a marketing campaign built around discounted rates for longer contract terms could also strongly encourage customers to change their contract to get a better subscription charge. In doing so, the company will hit the key metrics that indicate longer tenure times, and they will attain a steadier revenue stream.

C6 – Limitations: The analysis presents an opportunity to gain insight into industry norms and trends that exist between both WGU’s telecom company and its competitor. From the Churn Status chart, we can see that both companies have roughly the same churn rates for the rolling period recorded. This would imply that the churn rates that a telecommunication company would see in historical data should match, and that customers happen to just trend in this manner within the industry.

Despite this however, it is also possible that the two companies happen to just operate in the same manner and they both trend uniquely from the industry norm, since this assessment only compares two companies it is unknown for certain. As such, this analysis is limited by the utilization of these two companies rather than an aggregate comparison dataset.

Additionally, the WGU customer dataset also lacks some granular service selection information such as tech support, online backup, phone backup, streaming features, etc... This prevents further analysis to be conducted on service offerings that would result in better customer retention. The limitation here is that executive management will not be able to gain insight into what service offerings would be suitable for a bundled package to boost tenures.

D & E - Sources: Blastchar. (2018, February). *Telco Customer Churn*. Kaggle.
<https://www.kaggle.com/datasets/blastchar/telco-customer-churn>