Task 1: Data Analysis – D211

Abhishek Aern

Western Governor University

Contents

Part I - Data Dashboard

Interactive Tableau dashboards are provided (Advanced Data Acquisition - D211 PA.twbx file) with the submission.

A1: Data Sets

I am analyzing different components of the customer churn dataset based on demographic characteristics like gender, differences in churn rates across different geographic regions, customer tenure, customer Income, customer payment method, and different contract periods.

My data Analysis makes use of the following two datasets –

- The WGU provided the Churn dataset from the telecom industry.
- 2019 Census US Population Data by State downloaded from Kaggle. As per Kaggle population data was pulled from the 2019 US Census (Cohen).

The WGU-provided dataset has most of the data required for my analysis and the population dataset will help segment the customers based on the population density of the area they reside in. High-density areas may have different customer behavior and churn patterns than low-density areas.

The WGU-provided dataset has 10000 customer records and the external dataset has a population of 51 states(not including Puerto Rico). I need to add the PR state population manually as we have customer records from this state as well. Also, the external dataset has lat and long columns and we already have that in WGU-provided data, so I am deleting that manually and then loading the data into the database.

The WGU-provided dataset is already loaded in VM and I am providing the external dataset as 2019_Census_US_Population_Data_By_State_Lat_Long.csv and the combined prepared dataset as Prepared combined dataset.csv file with my submission.

A2: Dashboard Installation Instructions

Assuming that the PostgreSQL database using pgAdmin already has all WGU-provided tables loaded and ready to use.

(Sewell, 2023) Here are the steps to load the external dataset "2019_Census_US_Population_Data_By_State_Lat_Long.csv" into a new table called "state_pop" in the PostgreSQL database -

1. Create a new table called state_pop with the following schema:

```
CREATE TABLE state_pop (
    state TEXT NOT NULL,
    pop integer NOT NULL
);
```

This will create a table with two columns: state, and pop.

2. Save the provided external file to the desktop. The file name is "2019_Census_US_Population_Data_By_State_Lat_Long.csv".
3. Open pgAdmin and connect to the PostgreSQL database called "churn" and go to Tables under public schemas.

4. Right-click on the state_pop table and select "Import/Export".

5. In the "Import/Export" dialog box, select "Import".

6. In the "Filename" field, browse to the location of the .csv file containing the population data(file name provided in step 2).

7. In the "Format" field, select "CSV".

8. In the "Delimiter" field, enter the delimiter ',' (comma) as this is a CSV file.

9. In the "Header" field, select "Yes" as the external dataset contains the header record and keep all other fields as default values.

10. Click "OK" to start the import process.

Once the import is complete, you should see the data populated in the state_pop table.

This state_pop table contains state names as full names and WGU provided dataset contains two-letter state codes, so we need to convert the state names to two-letter state codes using the reference table.

Follow the below steps –

- Create a reference table (state_codes) that maps each state name to its corresponding two-letter state code. Here's an SQL statement to create a reference table and load the data -

```
CREATE TABLE state_codes (
    state_name TEXT NOT NULL,
    state_code CHAR(2) NOT NULL
);

INSERT INTO state_codes (state_name, state_code)
VALUES ('Alabama', 'AL'),
       ('Alaska', 'AK'),
       ('Arizona', 'AZ'),
       ('Arkansas', 'AR'),
       ('California', 'CA'),
       ('Colorado', 'CO'),
       ('Connecticut', 'CT'),
       ('Delaware', 'DE'),
       ('District of Columbia', 'DC'),
       ('Florida', 'FL'),
       ('Georgia', 'GA'),
       ('Hawaii', 'HI'),
       ('Idaho', 'ID'),
       ('Illinois', 'IL'),
       ('Indiana', 'IN'),
       ('Iowa', 'IA'),
       ('Kansas', 'KS'),
       ('Kentucky', 'KY'),
       ('Louisiana', 'LA'),
       ('Maine', 'ME'),
       ('Maryland', 'MD'),
       ('Massachusetts', 'MA'),
       ('Michigan', 'MI'),
       ('Minnesota', 'MN'),
       ('Mississippi', 'MS'),
       ('Missouri', 'MO'),
       ('Montana', 'MT'),
       ('Nebraska', 'NE'),
       ('Nevada', 'NV'),
       ('New Hampshire', 'NH'),
       ('New Jersey', 'NJ'),
       ('New Mexico', 'NM'),
       ('New York', 'NY'),
       ('North Carolina', 'NC'),
       ('North Dakota', 'ND'),
       ('Ohio', 'OH'),
       ('Oklahoma', 'OK'),
       ('Oregon', 'OR'),
```

```
        ('Pennsylvania', 'PA'),
        ('Rhode Island', 'RI'),
        ('South Carolina', 'SC'),
        ('South Dakota', 'SD'),
        ('Tennessee', 'TN'),
        ('Texas', 'TX'),
        ('Utah', 'UT'),
        ('Vermont', 'VT'),
        ('Virginia', 'VA'),
        ('Washington', 'WA'),
        ('West Virginia', 'WV'),
        ('Wisconsin', 'WI'),
        ('Wyoming', 'WY');
```

This creates a state_codes table with two columns: state_name and state_code. The INSERT statement populates the table with the corresponding state names and two-letter state codes.

- Now modify the previously created state_pop table to use the two-letter state codes instead of state names by adding a new column for the state codes and populating it using a join with the state_codes table.

Here's the SQL statement to modify the state_pop table:

```
ALTER TABLE state_pop ADD COLUMN state_code CHAR(2);

UPDATE state_pop
SET state_code = sc.state_code
FROM state_codes sc
WHERE state_pop.state = sc.state_name;

ALTER TABLE state_pop DROP COLUMN state;
```

This will add a new column state_code to the state_pop table, populate it with the corresponding two-letter state codes using a join with the state_codes table, and drops the original state column.

- Verify that the state_pop table now uses the two-letter state codes.
- The Puerto Rico (PR) state population is missing so we need to manually add this using the below Insert statement –

```
INSERT INTO state_pop (state_code, pop)
VALUES ('PR', 3263584);
```

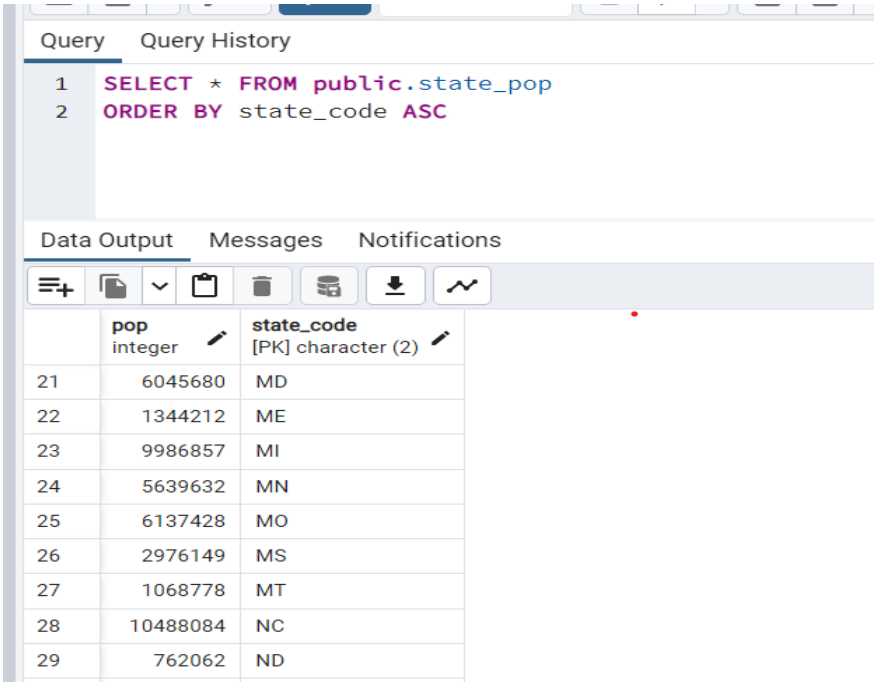
- Adds a primary key constraint to the state_pop table, using the state_code column as the primary key.

```
ALTER TABLE state_pop ADD PRIMARY KEY (state_code);
```

- Also creates a foreign key constraint on the state column of the location table, referencing the state_code column of the state_pop table. This ensures that any value in the state column of the location table must match a value in the state_code column of the state_pop table, preventing the insertion of invalid data.

```
ALTER TABLE location ADD CONSTRAINT location_state_code_fk
    FOREIGN KEY (state)
    REFERENCES state_pop (state_code);
```

- Finally, the state_pop table should look like the below -



This is all the setup needed for the PostgreSQL database using pgAdmin (Sewell, 2023). Next, the following steps are required inside Tableau -

- Users should have Tableau Desktop or Tableau Reader on their computers.
- The packaged workbook is provided so the user can directly access this in Tableau without connecting to a data source to view the dashboard.

- Save the packaged workbook in the local drive and use the open option in Tableau to view the workbook or navigate to the folder and double-click on the workbook file(".twb" or ".twbx") to open it in Tableau.
- Once the workbook is open, click on the "Story - Customer Churn" tab to open my dashboard.

I am also provided below additional instructions in case the evaluator needs to connect to the data source as well for any reason.

1. Connect to the PostgreSQL database in Tableau by selecting "PostgreSQL" as the data source type and entering the necessary server and login information(Already provided by WGU).

2. In the "Connect" pane, select "Custom SQL" and enter below SQL query to join all the necessary tables and select the necessary columns required for my analysis (Tableau Documentation, n.d.).

```
/*
Creates SQL for Tableau
*/
SELECT CAST("customer"."customer_id" AS TEXT) AS "customer_id",
  "customer"."age" AS "age",
  CAST("customer"."churn" AS TEXT) AS "churn",
  "customer"."contacts" AS "contacts",
  "customer"."contract_id" AS "contract_id",
  "contract"."duration" AS "duration",
  CAST("customer"."gender" AS TEXT) AS "gender",
  "customer"."income" AS "income",
  "customer"."lat" AS "lat",
  "customer"."lng" AS "lng",
  "customer"."location_id" AS "location_id",
  CAST("customer"."marital" AS TEXT) AS "marital",
  "customer"."monthly_charge" AS "monthly_charge",
  "customer"."payment_id" AS "payment_id",
  CAST("payment"."payment_type" AS TEXT) AS "payment_type",
  "state_pop"."pop" AS "pop",
  TO_CHAR("state_pop"."pop"::integer, 'fm999G999G999') AS
"Population",
  "location"."state" AS "state",
  "customer"."tenure" AS "tenure",
  "location"."zip" AS "zip"
From "public"."customer" "customer"
  LEFT JOIN "public"."contract" "contract" ON
("customer"."contract_id" = "contract"."contract_id")
```

```
  INNER JOIN "public"."location" "location" ON
("customer"."location_id" = "location"."location_id")
  LEFT JOIN "public"."payment" "payment" ON
("customer"."payment_id" = "payment"."payment_id")
  INNER JOIN "public"."state_pop" "state_pop" ON
("location"."state" = "state_pop"."state_code")
```

3. Click "OK" to close the custom SQL editor and return to the Tableau worksheet.
4. Now click on the "Story - Customer Churn" tab to open my dashboard.

A3: Dashboard Navigation Instructions

The “Story - Customer Churn” board has 2 tabs, and both tabs have various visualizations which provide different KPIs for customer churn in the telecom industry.

The first tab has an interactable “Dashboard1” which provides details about the Number of Customers, Churn rate, and average monthly charges along with the visualizations for state population, gender distribution, and contract type.

Gender, contract type, and state population are interactable and can be used as filtered to adjust other visualizations in the same board. For example, selecting any one particular state will show all the information about that particular state, like the number of customers in the selected state, Churn Rate, Average Monthly charges, contract type along with gender distribution. Similarly, clicking on any gender in the gender visualizations will update all the visualizations to reflect only the selected gender. For example, clicking on the "male" portion of the pie charts will adjust all other information as per the male customers. Clicking on the same filter again will "release" the selected filter and return the visualizations to the original view.

The second tab is another interactable dashboard, providing visualizations for churned customers by payment method, Churn by payment method, and Churn customers by income level along with average monthly charges and customer churn percentage.

The payment method, Tenure, and Income can be used as a filter to analyze KPI for different periods. Selecting the particular payment method will adjust all data and the visualizations for that payment method. The ‘Churned customer by Tenure” visualization can be adjusted using the parameter ‘Tenure range’ with the increment of 10 months and selecting any income group bar will adjust the dashboard according to that group.

A4: SQL Code

All the SQL codes are provided in section A2.

I am loading the external dataset into the PostgreSQL database using pgAdmin and making the necessary changes before connecting it to Tableau.

Here is all the SQL code together needed for the PostgreSQL database using pgAdmin

```
/*
Create a new table state_pop
*/
CREATE TABLE state_pop (
    state TEXT NOT NULL,
    pop integer NOT NULL
);

/*
Create a reference table state_codes
*/
CREATE TABLE state_codes (
    state_name TEXT NOT NULL,
    state_code CHAR(2) NOT NULL
);

/*
load the data into reference table state_codes
*/
INSERT INTO state_codes (state_name, state_code)
VALUES ('Alabama', 'AL'),
       ('Alaska', 'AK'),
       ('Arizona', 'AZ'),
       ('Arkansas', 'AR'),
       ('California', 'CA'),
       ('Colorado', 'CO'),
       ('Connecticut', 'CT'),
       ('Delaware', 'DE'),
       ('District of Columbia', 'DC'),
       ('Florida', 'FL'),
       ('Georgia', 'GA'),
       ('Hawaii', 'HI'),
       ('Idaho', 'ID'),
       ('Illinois', 'IL'),
       ('Indiana', 'IN'),
       ('Iowa', 'IA'),
       ('Kansas', 'KS'),
       ('Kentucky', 'KY'),
```

```
        ('Louisiana', 'LA'),
        ('Maine', 'ME'),
        ('Maryland', 'MD'),
        ('Massachusetts', 'MA'),
        ('Michigan', 'MI'),
        ('Minnesota', 'MN'),
        ('Mississippi', 'MS'),
        ('Missouri', 'MO'),
        ('Montana', 'MT'),
        ('Nebraska', 'NE'),
        ('Nevada', 'NV'),
        ('New Hampshire', 'NH'),
        ('New Jersey', 'NJ'),
        ('New Mexico', 'NM'),
        ('New York', 'NY'),
        ('North Carolina', 'NC'),
        ('North Dakota', 'ND'),
        ('Ohio', 'OH'),
        ('Oklahoma', 'OK'),
        ('Oregon', 'OR'),
        ('Pennsylvania', 'PA'),
        ('Rhode Island', 'RI'),
        ('South Carolina', 'SC'),
        ('South Dakota', 'SD'),
        ('Tennessee', 'TN'),
        ('Texas', 'TX'),
        ('Utah', 'UT'),
        ('Vermont', 'VT'),
        ('Virginia', 'VA'),
        ('Washington', 'WA'),
        ('West Virginia', 'WV'),
        ('Wisconsin', 'WI'),
        ('Wyoming', 'WY');

/*
Add new column state_code into state_pop table to use the two-letter
state code
*/
ALTER TABLE state_pop ADD COLUMN state_code CHAR(2);

/*
Join with the state_codes table to populate the data into the
state_code column of the state_pop table
*/
UPDATE state_pop
SET state_code = sc.state_code
FROM state_codes sc
WHERE state_pop.state = sc.state_name;

/*
Drop the existing state column
```

```
*/
ALTER TABLE state_pop DROP COLUMN state;

/*
Add Puerto Rico (PR) state population to the state_pop table
*/
INSERT INTO state_pop (state_code, pop)
VALUES ('PR', 3263584);

/*
Adds a primary key constraint to the state_pop table
*/

ALTER TABLE state_pop ADD PRIMARY KEY (state_code);

/*
Creates foreign key constraint on the state column of the location
table
*/
ALTER TABLE location ADD CONSTRAINT location_state_code_fk
    FOREIGN KEY (state)
    REFERENCES state_pop (state_code);
```

Here is the SQL code needed for the Tableau custom query –

```
/*
Creates SQL for Tableau
*/
SELECT CAST("customer"."customer_id" AS TEXT) AS "customer_id",
  "customer"."age" AS "age",
  CAST("customer"."churn" AS TEXT) AS "churn",
  "customer"."contacts" AS "contacts",
  "customer"."contract_id" AS "contract_id",
  "contract"."duration" AS "duration",
  CAST("customer"."gender" AS TEXT) AS "gender",
  "customer"."income" AS "income",
  "customer"."lat" AS "lat",
  "customer"."lng" AS "lng",
  "customer"."location_id" AS "location_id",
  CAST("customer"."marital" AS TEXT) AS "marital",
  "customer"."monthly_charge" AS "monthly_charge",
  "customer"."payment_id" AS "payment_id",
  CAST("payment"."payment_type" AS TEXT) AS "payment_type",
  "state_pop"."pop" AS "pop",
  TO_CHAR("state_pop"."pop"::integer, 'fm999G999G999') AS
"Population",
  "location"."state" AS "state",
```

```
  "customer"."tenure" AS "tenure",
  "location"."zip" AS "zip"
FROM "public"."customer" "customer"
  LEFT JOIN "public"."contract" "contract" ON
("customer"."contract_id" = "contract"."contract_id")
  INNER JOIN "public"."location" "location" ON
("customer"."location_id" = "location"."location_id")
  LEFT JOIN "public"."payment" "payment" ON
("customer"."payment_id" = "payment"."payment_id")
  INNER JOIN "public"."state_pop" "state_pop" ON
("location"."state" = "state_pop"."state_code")
```

Part II - Demonstration

My Panopto presentation covers B1 – B7 and the recording link is provided with the submission.

https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=7486db7f-5b7a-4485-a570-b0000119cbff

Part III - Reflection Paper

C1: Dashboard Alignment

The WGU Chun dataset consists of 10000 customers with 50 different variables which contain information on customer behavior and characteristics, as well as details about the services provided by the company. This dataset consists of customer demographic data such as age, gender, location, and income, as well as information on the customer's usage patterns and behavior. This dataset also provides Billing Information like monthly charges, Customer Service Interactions, Tenure along with Churn information.

I am analyzing these different components of the customer churn dataset which can help to identify the key drivers of customer churn, understand the factors that contribute to customer retention, and take action to improve customer satisfaction and reduce churn rates.

My dashboards provide visualizations of churn rates by state, which can help identify areas where customer churn is high and potentially identify factors contributing to the churn. Also, it includes visualizations of churn by tenure, age,

contract type, payment method, Income, and gender, which can help identify customer segments that are more likely to churn and potentially tailor retention strategies accordingly. I am including population density visualizations in the dashboard so the telecom company can better understand how population density may be related to customer churn rates. This can provide important context for interpreting churn data and developing effective retention strategies.

I have also provided various interactive controls in both dashboards to help users with the ability to modify the presentation of the data and filter the information according to specific needs. The gender filter will allow the user to filter the data by selecting a specific gender or multiple genders. The contract type filter allows the user to filter the data by selecting a specific contract type or multiple contract types. The state map provides a visual representation of the population density for different states and the user can select any state to see additional details like the number of customers, churn rate, etc. based on selection. Similarly, the Tenure, Income, and Payment method filter help users drill down into the data to gain more insights.

C2: Business Intelligence Tool

(Shaptunova) Tableau is a popular business intelligence tool that provides a wide range of features and benefits. Some of the key justifications for selecting Tableau include:

Ease of use: Tableau provides a user-friendly interface that makes it easy for even non-technical users to create and customize interactive dashboards and reports. With drag-and-drop functionality, users can easily create visualizations and explore data without needing to write complex code or SQL queries.

Visual appeal: Tableau provides stunning and visually appealing visualizations that are highly customizable. With a variety of chart types, color schemes, and formatting options, Tableau helps users to create engaging and impactful visualizations that can help to drive better decision-making.

Speed and scalability: Tableau is designed to handle large volumes of data and is highly scalable. It can easily connect to a variety of data sources, including databases, spreadsheets, and cloud-based data sources, and can quickly analyze and process data.

Integration with other tools: Tableau integrates with a wide range of other tools, including data preparation tools, ETL tools, and statistical analysis tools. This makes it easy to incorporate Tableau into existing data workflows and use it in conjunction with other data tools.

Active community: Tableau has a large and active user community that provides support, tutorials, and best practices for using the tool. This community helps to ensure that users can easily find answers to their questions and get up to speed quickly with Tableau (Shaptunova).

C3: Data Cleaning

As per my previous coursework in D210, the WGU-provided customer churn dataset is clean and there are no duplicate data, missing values, or any other errors. The customer table has population data, but it doesn't look good as per my previous analysis so the external dataset will provide population per state for my analysis.

My data cleaning and preparation involve the following steps –

- Combine data from different sources into a single source. This step involves loading the external dataset into the PostgreSQL database and converting the state column to 2 character state code.
- The total number of customers is calculated using COUNT([Customer ID])
- The churn variable has a categorical value of Yes/No so that was converted to numerical and stored in the variable 'Churned'.
  IF [Churn] = 'Yes' THEN 1 ELSE 0 END
- Number of Churned Customers using SUM([Churned])
- The churn rate is calculated using the below and converted this field into a percentage.
  [Number of Churned Customers] / [Number of Customers]
- I have grouped the customers with a tenure of 10 months and created the parameter to provide users with the ability to adjust it as per their needs and give more flexibility.
- Create a tenure bin and Income bin for the visualization.

C4: Dashboards Creation

My story has 2 dashboards that contain multiple visualizations. Following are steps to create those visualizations (Aern, 2023) –

Total Customers :

- The number of Customers field was calculated as per section C3.
- Drag that field from the data pane to the "Text" box.
- Hide the title.
- Click on the Text box and go to edit label to add the title ‘Total Customers’.

Churned Rate :

- The churn rate field was already created as per section C3.
- Drag that field from the data pane to the "Text" box.
- Hide the title.
- Click on the Text box and go to the edit label to add the title ‘Churn Rate.
- Right-click in the Churn rate field in the data pane and go to default property then Number format and select percentage with 2 decimal places to show this field as %.

Monthly Average Charges :

- Drag the monthly charges field from the data pane to the "Text" box.
- Click on the field in the Mark section and select measure as average instead of sum.
- Hide the title.
- Click on the Text box and go to edit label to add the title ‘Average Monthly Charges’.

Population density and Customer Churn

- Drag and drop the "State" field from the data pane to the worksheet.
- Change the mark type from "Automatic" to "Map".
- Drag the "State" field from the data pane to Label.
- Drag the "pop" field from the data pane to color shelves to create a state-wise view of the population density.
- Click on the pop field in the Mark section and select measure as average instead of sum.
- Click on the color and go to edit color and choose orange-blue diverging and check the reversed option.

- Drag the number of customers, churned customers, and population fields from the data pane to the tooltip.
- Renamed the header as ‘Population Density and Customer Churn’.

Gender distribution :

- Drag the "Gender" field to the "Color" shelf.
- Drag the calculated "Churned Rate" field to the "Label" shelf.
- Select the Visualization as a Pie.
- Renamed the header as ‘Gender distribution’.

Contract :

- Drag the "Duration" field to the "Label" shelf.
- Drag the calculated "Churn Rate" field to the "Color" shelf.
- Change the chart type to "Tree Map".
- Adjust the size and colors of the boxes to make the view easier to interpret.
- Renamed the header as ‘Contract’.

Churned customers by payment method :

- Drag the "Payment type" field to the "Rows" shelf.
- Drag the calculated field "Number of Churned” to the "Column” shelf.
- Change the chart type to "Horizontal Bar".
- Sort the bars in descending order by clicking on the "Number of Churned customers” axis and selecting "Sort Descending".
- Renamed the header as ‘Churned Customers by Payment Method’.

Churned customers by Tenure :

Create a tenure parameter:

- Go to the "Data" pane and click on "Create Parameter"
- Name the parameter "Tenure Range" and set the data type to "Integer"
- Set the allowable values to "Range" with a Minimum of 0, Maximum of 80 with a step size of 10.
- Click "OK" to create the parameter

Create a calculated field for the tenure bin:

- Go to the "Data" pane and click on Tenure then go to Create option and select Bins.
- Do not change the New filed name and keep it Tenure (bin).
- Choose a Tenure Range for the Size of the bin.
- Click "OK" and it will create the Tenure (bin) in the data pane.

Create a vertical bar chart:

- Drag the "Tenure (bin)" field to the Columns shelf
- Drag the calculated field "Number of Churned Customers" to the Rows shelf
- It will create the vertical bar chart and adjust the size and formatting of the chart
- Renamed the header as ‘Churned Customers by Tenure’
- Right-click on the "Tenure Range" parameter in the "Parameters" pane and select "Show Parameter Control".
- Adjust the parameter to see the chart update based on the selected tenure bin size.

Churned customers by Income :

- Go to the "Data" pane and click on Income then go to Create option and select Bins.
- Do not change the New filed name and keep it Income (bin).
- Chose the size of the bin as 30000.
- Click "OK" and it will create the Income (bin) in the data pane.
- Drag the "Income (bin)" field to the Columns shelf
- Drag the calculated field "Number of Churned Customers" to the Rows shelf
- It will create the vertical bar chart and adjust the size and formatting of the chart
- Renamed the header as ‘Churned Customers by Income’.

Dashboard 1 :

- Import the Total customer, Churn Rate, and Average monthly charges worksheet to the top portion of the page.
- Import Population density and Customer Churn to the left, Gender distribution to the middle right, and Contract to the bottom right.

- Make a gender legend floating and place it left side of the gender pie chart.
- Activate “Use As Filter” for all 3 visualizations.

Dashboard 2 :

- Import the Total customer, Churn Rate, and Average monthly charges worksheet to the top portion of the page.
- Import Churned Customers by Payment Method to the middle portion of the dashboard.
- Import Churned Customers by Tenure worksheet to the bottom left and Churned Customers by Income worksheet to the bottom right side of the dashboard.
- Activate “Use As Filter” for all 3 visualizations.

Story – Customer Churn

- Import dashboard 1 to the first story point and rename the story to State, Gender, and contract distribution.
- Import dashboard 2 to the second story point and rename the story to Tenure, Payment method, and Income distribution.

C5: Data Analysis Results

The different data representations are available on both the dashboards in my story. Based on the analysis of the telecom company's customer churn dataset and external dataset for population by state, several insights and patterns can be identified that support executive decision-making -

Gender: This visualization shows the distribution of male, female, and other customers and how they relate to churn. This can help leaders to understand if gender plays a role in customer churn and whether there are any patterns or trends. For example, the dataset represents that male customers are more likely to churn, the company may want to investigate why this is the case and take steps to address any issues specific to male customers.

Contract Type: This visualization shows the distribution of contract types (e.g., monthly, 1 year, and 2 years) and how they relate to churn. This helps leaders to understand if there are any patterns or trends with different types of contracts. For example, This dataset represents that customers on monthly contracts are more likely

to churn, the company may want to consider changing its pricing strategy or offering incentives to encourage customers to switch to longer-term contracts.

Population Density and Customer Churn: The dataset shows that there is a correlation between population density and churn rate. States with a higher population density tend to have a higher churn rate. This information could be used to target marketing and sales efforts in those states and implement retention strategies for customers.

The second story point represents some of the additional KPIs like the number of Customer churns by Tenure, Payment method, and Income.

Churn customers by payment method: The horizontal bar chart in the dashboard shows the distribution of customer churn based on their payment methods like Electronic checks, mailed checks, bank transfers, or credit cards. Customers using electronic checks as a payment method are more likely to churn compared to other payment methods. Offering discounts or promotions for customers to switch to other payment methods could be a potential solution.

Customer churn by Tenure: This is represented in the vertical bar chart which shows the distribution of customer churn based on their tenure with the telecom company. This shows the number of customers who have churned grouped by the range of their tenure with the company which will help executive leaders understand how customer churn is affected by the length of time a customer has been with the company. For example, in the dataset, a large proportion of customers who have been with the company for less than a year are churning, which may indicate a problem with onboarding or early customer experience. By understanding the relationship between customer tenure and churn, executive leaders can make decisions around improving customer retention strategies and identifying areas for improvement in the customer experience.

Customer Churn by Income - The income vertical bar chart provides valuable insights into the relationship between income levels and customer churn. From the chart, we can see that the highest number of churned customers comes from the income range of $20,000 to $30,000, followed by $30,000 to $40,000. This information is valuable for executive decision-making because it indicates that customers with lower income levels are more likely to churn, which could be due to a variety of reasons such as affordability or satisfaction with the services provided.

All these different data representations in my dashboards will help executive leaders to identify patterns and trends in customer churn and take steps to address any issues. By using these insights to inform decision-making, the company can improve customer retention and increase revenue over time.

C6: Analysis Limitation

Here are some of the limitations of my data analysis -

Limited data: The analysis is based on a limited dataset that only includes information about customer churn and some demographic and account-related factors. There could be other factors that contribute to customer churn that are not captured in the dataset.

Data quality: The accuracy and completeness of the data could be a concern. There could be missing or incorrect data in the dataset that could affect the results of the analysis.

Causality: My analysis is focused on identifying correlations between different factors and customer churn. Correlation does not imply causation, and it is important to keep in mind that there could be other factors that contribute to customer churn that are not captured in the dataset and my analysis.

While Tableau is a powerful and flexible tool for data analysis and visualization, I think there are a few limitations (Shaptunova)–

- Tableau has a steep learning curve for users who are new to the tool.
- Tableau can be expensive, especially for enterprise-level deployments, which may make it difficult for smaller organizations to adapt.
- Tableau has some basic data preparation capabilities, but more advanced data cleaning and preparation tasks may require the use of other tools or manual intervention.

D. References

Aern, A. (2023). D210 PA.
https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=f40b8420-a772-4a26-8f10-aff900cd07c7

Cohen, P. (n.d.). 2019 Census US Population Data By State. https://www.kaggle.com/datasets/peretzcohen/2019-census-us-population-data-by-state?select=2019_Census_US_Population_Data_By_State_Lat_Long.csv

Sewell, D. W. (2023). SQL Sunday. https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=880cdde7-2726-470a-8024-aeb901724385

Shaptunova, Y. (n.d.). Tableau Software Review: Pros and Cons of a BI Solution for Data Visualization. https://www.sam-solutions.com/blog/tableau-software-review-pros-and-cons-of-a-bi-solution-for-data-visualization/

Tableau Documentation. (n.d.). Connect to a Custom SQL Query. https://help.tableau.com/current/pro/desktop/en-us/customsql.htm