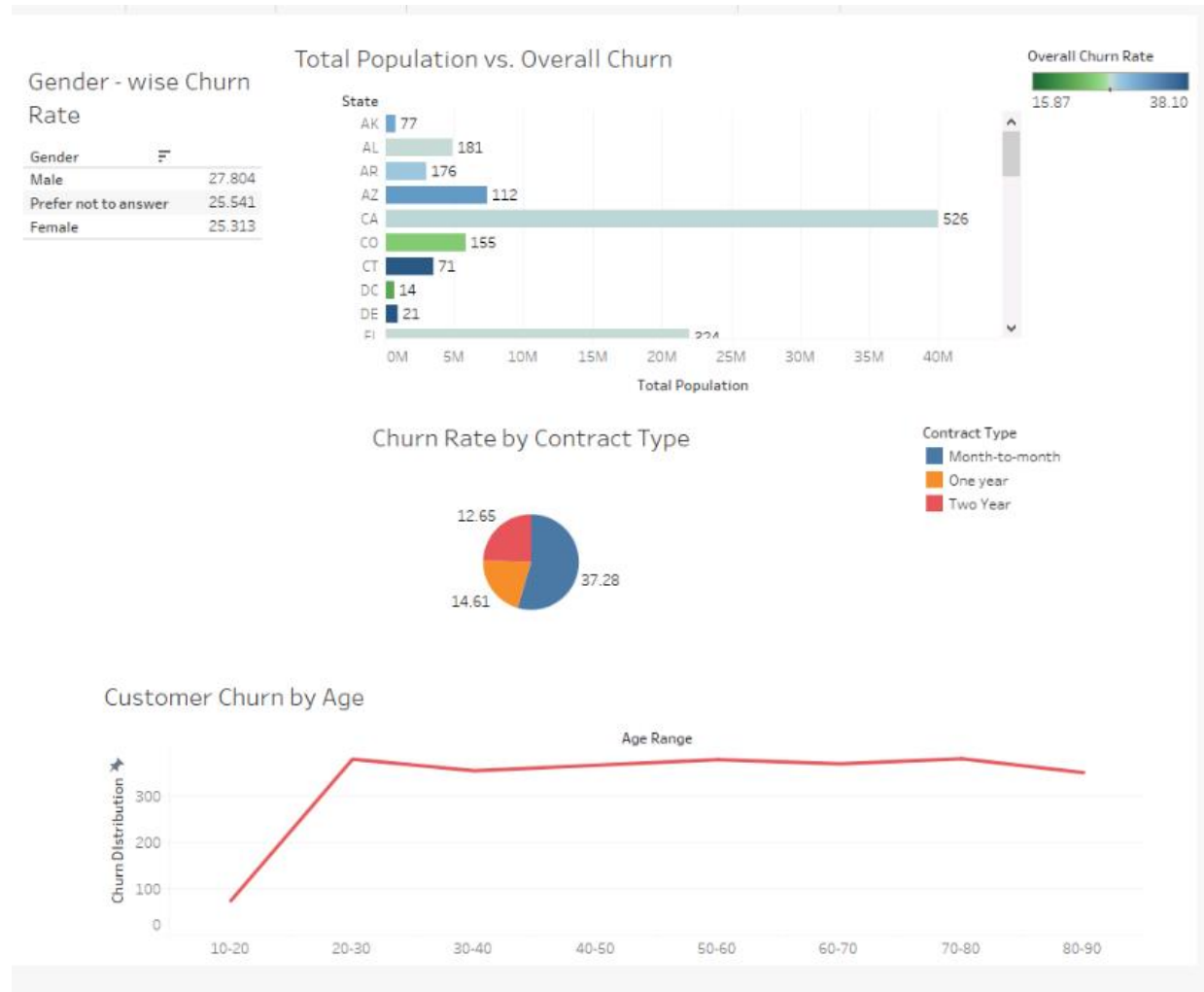


D211: Advanced Data Acquisition

Western Governors University

Oviya Selvaraj

## Part 1: Data Dashboard that supports executive decision making:



### A1. The datasets used for the analysis are:

- 1) WGU churn dataset, provided by WGU.
- 2) The external dataset containing total population of each state in United States.

The external dataset has been included in the zip file.

### A2.

#### Create PostgreSQL Database:

1. Open the virtual machine and on the desktop double click pgAdmin to open

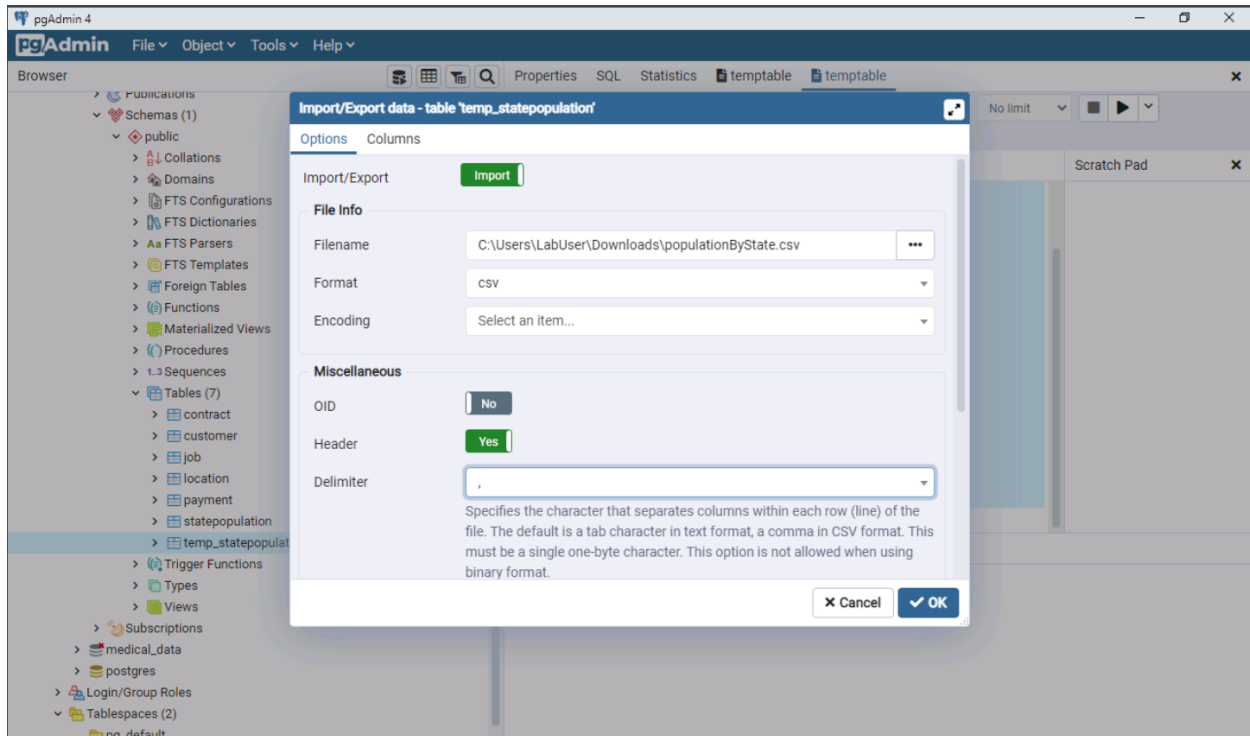
the database.

2. On the left pane, navigate to churn database one level below databases.
3. Right click on Query tool.

To create the temporary table to store external CSV file, paste the SQL query below.

```
CREATE TABLE temp_statepopulation (  
  rank INTEGER,  
  state TEXT ,  
  pop INTEGER,  
  growth FLOAT,  
  pop2018 INTEGER,  
  pop2010 INTEGER,  
  growthSince2010 FLOAT,  
  percent FLOAT,  
  density FLOAT  
)
```

4. Import the external CSV file: Right-click on the newly created table temp\_statepopulation and click 'Import/Export'. A dialog box appears, select import on the toggle button, select the file 'populationByState.csv' from C:\Users:\LabUser\Downloads\, toggle yes for header and select comma for delimiter. It should look something like this:



5. Select OK. The data should be imported, and a success message appears.

6. Now, create the target table using the SQL code below:

```
CREATE TABLE statepopulation(  
state TEXT PRIMARY KEY,  
pop INTEGER  
)
```

7. Using the INSERT command copy the values 'state' and 'pop' from the temp\_statepopulation into the statepopulation table. Use the SQL command below:

```
INSERT INTO statepopulation(state, pop)  
SELECT state, pop from temp_statepopulation;
```

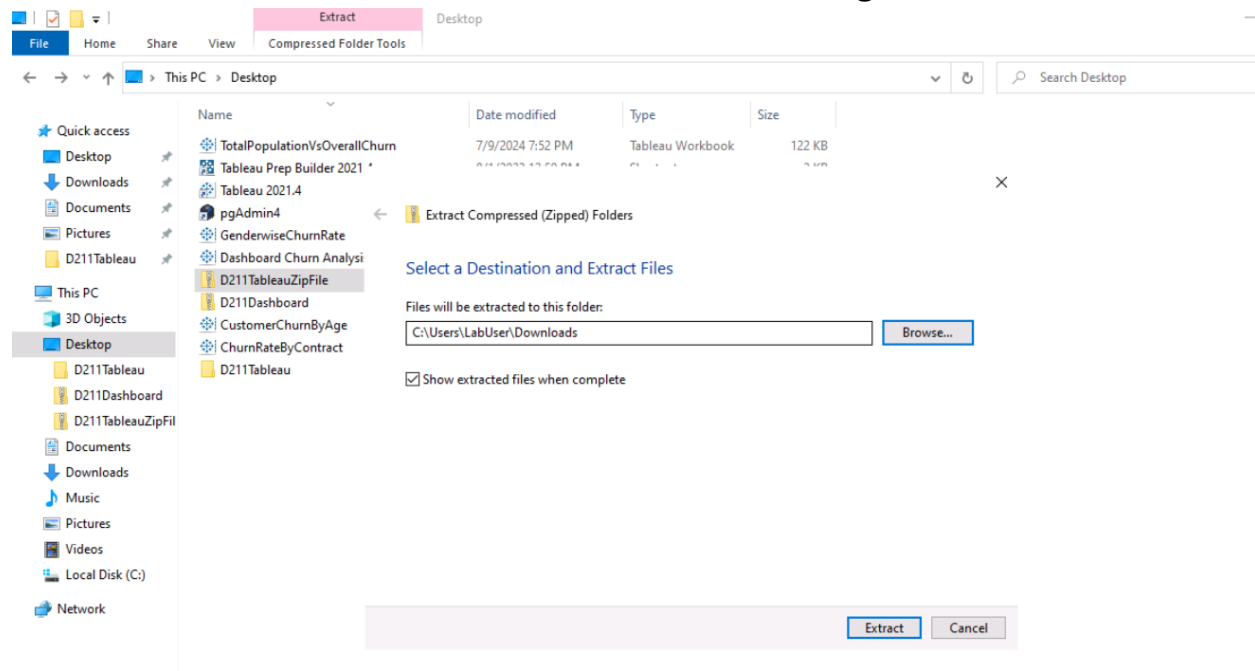
Values have now been inserted into the target table.

8. Delete the temporary table using the 'Drop' command.

```
DROP TABLE 'temp_statepopulation'
```

## View Tableau Dashboard

1. Right click on the D211TableauZipFile and select extract all to extract the files to the downloads folder. It should look something like this:



2. Double click Tableau 2021.4 to open Tableau Desktop.
3. At the top left click File>Open and navigate to downloads file. Select the 'D211TableauZipFile'>'DashboardChurnAnalysisD211'.
4. When prompted enter the username 'postgres' and password 'Passw0rd!'.
5. The dashboard should open and can be interacted with.

### A3. Dashboard Usage

After opening the dashboard navigate to the Story 1 tab. There is a brief description of the analysis performed.

The Churn Analysis Dashboard is an interactive dashboard containing analysis done on the customers based on various factors like Age, total State population, contract type and gender.

**Total population vs. Overall Churn** is a bar chart representing the customer count in each state against the total population. The churn is indicated based on the color gradient on the top left corner. Hover over the bars to view details such as state, count of customer, total population and overall churn rate.

**Customer Churn by Age** contains bins of the age distribution of the customer data. The line graph represents the drastic increase of churn after the age of 20+. Hover over the line graph to view the age range and sum of churn

indicator.

**Gender-wise Churn Rate** is a tabular representation of the overall churn rate based on the various gender types 'male', 'female' and 'prefer not to say'.

Using the sorting symbol, the user can view the tabular data in ascending or descending order. The values can be sorted based on 'Overall Churn Rate withing Gender' or 'Gender within Gender'.

**Churn Rate by Contract Type** is a pie chart representing the overall churn rate based on the contract types. Select the various divisions of the pie chart to view the duration, count of customer and the overall churn rate.

#### A4. SQL code used in my analysis:

##### PgAdmin:

- Create a temporary table to import the external CSV file 'populationByState':

```
CREATE TABLE temp_statepopulation (  
  rank INTEGER,  
  state TEXT ,  
  pop INTEGER,  
  growth FLOAT,  
  pop2018 INTEGER,  
  pop2010 INTEGER,  
  growthSince2010 FLOAT,  
  percent FLOAT,  
  density FLOAT  
)
```

- Create the target table 'statepopulation' containing the two columns 'state' and 'population' used for the visualization.

```
CREATE TABLE statepopulation(  
  state TEXT PRIMARY KEY,
```

pop INTEGER

)

- Using the INSERT copy the values for 'state' and 'population' from the 'temp\_statepopulation'.

```
INSERT INTO statepopulation(state, pop)
SELECT state, pop from temp_statepopulation;
```

- Delete the temporary table 'temp\_statepopulation'

```
DROP TABLE 'temp_statepopulation'
```

### Tableau:

- I used SQL query to create the calculated fields "Overall Churn Rate. This calculated field was used in my 'Gender - wise churn rate' and 'Churn Rate by Contract and Gender' visualization.

```
(SUM(IF [Churn]='Yes' THEN 1 ELSE 0
END)/COUNT([Customer Id])) * 100
```

- I used SQL query to calculate the 'Churn Indicator'. This calculated field was used in my 'Customer Churn by Age' visualization.

```
IF [Churn] ='Yes' THEN 1
ELSEIF [Churn]='No' THEN 0
ELSE NULL
END
```

- I used SQL query to set the range of bins, Age range calculation:

```
IF [Age] >=10 AND [Age]<20 THEN "10-20"
```

```
ELSEIF [Age]>=20 AND [Age]<30 THEN "20-30"
ELSEIF [Age]>=30 AND [Age]<40 THEN "30-40"
ELSEIF [Age]>=40 AND [Age]<50 THEN "40-50"
ELSEIF [Age]>=50 AND [Age]<60 THEN "50-60"
ELSEIF [Age]>=60 AND [Age]<70 THEN "60-70"
ELSEIF [Age]>=70 AND [Age]<80 THEN "70-80"
ELSEIF [Age]>=80 AND [Age]<90 THEN "80-90"
END
```

### **Part 3: Report**

#### **C1. Dashboard Alignment with needs of the Business:**

Churn refers to the percentage of customers who discontinue their service with a provider. It is an essential metric to discover customer dissatisfaction and compare it to preference over competing services. Analyzing the given churn against multiple factors helps companies improve and reduce churn.

The dashboard contains Churn analysis against gender, age, the total population in each state, and contract type. The original dataset provided by WGU contains five tables with a primary table 'customer,' which contains churn, gender, contract ID, and location ID. Using the 'location' table, we can gain the number of customers in each state. The 'contract' table provides the duration of the various contract types. The additional external dataset 'populationBystate' provides the total population of all the States in the USA.

#### **C2. Business Intelligence Tool:**

**Data visualization:** Tableau Desktop was used for interactive data visualization. It helps to convey findings and draw conclusions from the data



analysis.

**Data Cleaning:** PostgreSQL was used to import external data file and create temporary and target tables to store the data.

### **C3. Data Preparation:**

- 1) Create table 'temp\_statepopulation' to import and temporarily store the data from the external CSV 'populationByState'.
- 2) Create the target table 'statepopulation' containing the two columns 'state' and 'population' used for the visualization.
- 3) Copy the 'state' and 'population' values from the 'temp\_statepopulation' to the 'statepopulation' table using the 'INSERT' query.

### **C4. Dashboard Creation:**

Steps to create the dashboard:

#### **Create the calculated field OVERALL INDICATOR:**

- Select the dropdown near the search bar under the 'Data' tab to select the 'Create Calculated Field.'
- In the title, enter 'Overall Indicator' and calculation

```
IF [Churn] ='Yes' THEN 1  
ELSEIF [Churn]='No' THEN 0  
ELSE NULL  
END
```

#### **Create calculated field OVERALL CHURN RATE:**

- Select the dropdown near the search bar under the 'Data' tab to select the 'Create Calculated Field.'

- In the title, enter 'Overall Churn Rate' and enter the calculation

(SUM(IF [Churn]='Yes' THEN 1 ELSE 0  
END)/COUNT([Customer Id])) \* 100

### **Create calculated field AGE RANGE:**

- Select the dropdown near the search bar under the 'Data' tab to select the 'Create Calculated Field.'
- Enter the calculation below to set the age range for the customers,

IF [Age] >=10 AND [Age]<20 THEN "10-20"  
  
ELSEIF [Age]>=20 AND [Age]<30 THEN "20-30"  
  
ELSEIF [Age]>=30 AND [Age]<40 THEN "30-40"  
  
ELSEIF [Age]>=40 AND [Age]<50 THEN "40-50"  
  
ELSEIF [Age]>=50 AND [Age]<60 THEN "50-60"  
  
ELSEIF [Age]>=60 AND [Age]<70 THEN "60-70"  
  
ELSEIF [Age]>=70 AND [Age]<80 THEN "70-80"  
  
ELSEIF [Age]>=80 AND [Age]<90 THEN "80-90"  
  
END

### **Total Population vs. Overall Churn Rate:**

- Drag and drop the 'state' rows from the 'location' table to the rows.
- Drag and drop 'Total Population' to the column.
- Drag and drop customer to label mark and change measure to COUNT.
- Drag and drop 'Overall Churn Rate' to the color mark.
- Select 'Edit color' and choose Green – Blue Diverging.

### **Customer Churn by Age:**

- Drag the 'Age Range' to the columns.
- Drag and drop the 'Churn Indicator' to the rows and change the measure to sum.
- Change the bar color to 'Red'.

#### **Gender-wise Churn Rate:**

- Drag and drop 'Gender' from customer table to rows.
- Drag and drop 'Overall Churn Rate' over the mark label.
- Select the Table type graph.

#### **Churn Rate by Contract Type:**

- Drag and drop 'Duration' from the 'contract' table over the color mark.
- Drag and drop customer to detail mark, change the measure to COUNT.
- Drag and drop the 'Overall Churn Rate' to the text mark.

#### **Dashboard: Churn Analysis D211**

- Import the 'Gender – wise Churn' right click and select 'floating'. Then move it to the top left corner.
- Import the 'Geographical Distribution of Customer Churn' right click and select 'floating'. Then move it to the top right corner.
- Import the 'churn Rate by Contract Type' right click and select "floating". Then move it to the center below the two previous visualizations.
- Import the 'Customer Churn by Age' right-click and select 'floating'. Then move it to the bottom.

#### **C5. Results of Data Analysis:**

The dashboard gives us the summary of churn based on various customer factors such as age, contract type, gender, and total population in each State. Based on these factors, the following conclusions can be drawn:

- **Total population vs. Overall Churn:** Rhode Island has a high churn rate (36.84) with a low customer count (19), suggesting strong competition and service issues.
- **Customer Churn by Age:** It can be observed from this analysis that the churn drastically increases for the ages 20+. This may suggest customers have higher expectations for customer support, service, and value for money.
- **Gender-wise churn Rate:** The data shows that male customers have higher churn. This indicates the need for personalized engagement strategies like tailored communication to meet the preferences of the male customer base.
- **Churn Rate by Contract Type:** Customers with Month-to-month contracts have the highest churn rate, indicating the need to enhance these contract conditions. Some strategies include good onboarding experience, which can help with product adoption, and adding incentives or perks to transitioning to longer-term contracts.

## **C6. Limitations of Data Analysis:**

The Population data used in the external data may not be up to date leading to inaccuracies if the population has changed significantly since the data was collected.