

WGU D210 TASK 1 REV 10 - MATTINSON

Dashboard and Storytelling of Telecom Churn Data

Mike Mattinson

Master of Science, Data Analytics, WGU.edu

Task 1: Data Dashboard and Storytelling

March 11, 2022

Abstract

Wgu.edu Data Analytics coursework. Create interactive dashboards using Tableau Public. Combine churn data provided by Wgu.edu with external US state population and region data. Use data from both data sets to create data visualizations and engaging stories. Provide viewer of the dashboard actionable insights regarding the data.

Contents

Part 1:	Interactive Data Dashboard	4
A.	Interactive dashboard	4
A1.	The Dashboard	5
A2.	Key Performance Indicators	8
A3.	The Data	9
A4.	Install (Create) Dashboard using Tableau Desktop	14
A5.	Navigate Dashboard using Tableau Public	14
Part 2:	Storytelling with Data	16
B.	Presentation Video	16
B1.	Introduction	Error! Bookmark not defined.
B2.	Data Summary	Error! Bookmark not defined.
B3.	Outline Key Results	Error! Bookmark not defined.
B4.	Data Representations	Error! Bookmark not defined.
B5.	Actionable Insights	Error! Bookmark not defined.
Part 3:	Reflection Paper	21
C.	Reflection Paper	21
C1.	Explain	Error! Bookmark not defined.
C2.	Explain	Error! Bookmark not defined.
C3.	Explain	Error! Bookmark not defined.
C4.	Explain	Error! Bookmark not defined.
C5.	Describe	Error! Bookmark not defined.
C6.	Explain	Error! Bookmark not defined.
C7.	Explain	Error! Bookmark not defined.
C8.	Describe	Error! Bookmark not defined.
C9.	Explain	Error! Bookmark not defined.

List of Tables

List of Figures

Figure 1 Example from dashboard	6
Figure 2 Lost Revenue Calculated Field	Error! Bookmark not defined.
Figure 3 Revenue Calculated Field	Error! Bookmark not defined.
Figure 4 Data Relationships	13
Figure 5 Top 3 States Lost Revenue	15

Part 1: Interactive Data Dashboard

A. Interactive dashboard

Provide a copy of your interactive Tableau dashboard to support executive decision-making. Your dashboard must be accessible to users with colorblindness, and must include the components in each of the following areas:

- A1. The Interactive Dashboard
- A2. Key Performance Indicators computed from both data sets
- A3. Data integrated from two (2) data sets
- A4. Instructions to install dashboard
- A5. Instructions to navigate dashboard

A1. The Dashboard

Dashboard. For this task, I created multiple interactive dashboards using Tableau Public 2021.4.

- Link to my profile <https://public.tableau.com/app/profile/mike.mattinson>
- Link to the **Density** Dashboard
https://public.tableau.com/app/profile/mike.mattinson/viz/D210_Task1_Density_1/Density
- Link to the **Lost Revenue** Dashboard
https://public.tableau.com/app/profile/mike.mattinson/viz/D210_Task1_Lost_Revenue_45/LostRevenue
- Link to the **Age Histogram** Dashboard
https://public.tableau.com/app/profile/mike.mattinson/viz/d210_task1_Age_6/Age_Histogram
- Link to the **Lost Customer** Dashboard
https://public.tableau.com/app/profile/mike.mattinson/viz/D210_Task1_Lost_Customers_1/LostCustomers
-

The figure below shows an example from the Lost Revenue dashboard (Mattinson, 2022) :

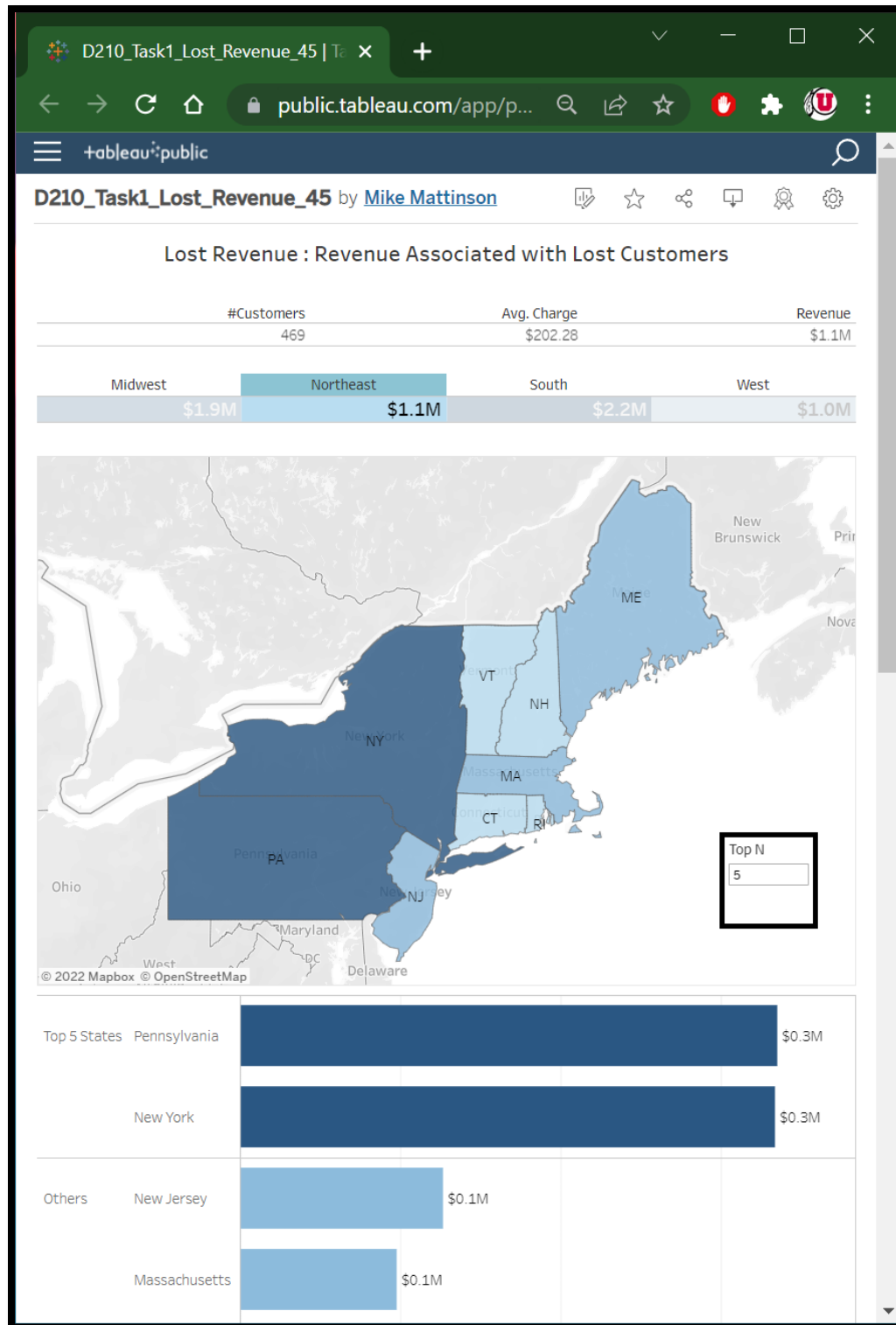


Figure 1 Example from dashboard

INTERACTIVE DASHBOARD SHOWING NUMBER OF LOST CUSTOMERS FROM NORTHWEST REGION. DASHBOARD COMPONENTS ARE AUTOMATICALLY RECALCULATED WHEN DIFFERENT OPTIONS ARE SELECTED BY USER.

Four (4) Representations. My dashboard includes the following data representations to summarize the data or display trends:

Dashboard	Primary Viz	Secondary Viz
Lost Revenue	Heatmap	Barchart
Density	Heatmap	Barchart
Age	Histogram	
Lost Customer	Hexmap	

Two (2) Interactive Controls. My dashboard contains the following interactive controls:

Dashboard	Calculated	Parameter	Filter
Lost Revenue	Revenue		Churn=Yes
Density	Density		Churn=No
Age	Age (bins)	Bin Size	
Lost Customers			

In addition to the primary controls, Tableau has other built-in controls such as:

- All visualizations have tooltips that have appropriate data displayed when user mouses into or over the data elements of the visualization
- When looking at a map, the user can zoom in and out with mouse wheel or by using the map tools at top left of map.
- On the map, the user can select one or more states or regions by selecting the item in the sub-totals, or by Ctrl-clicking on an area of the map.

A2. Key Performance Indicators

Key Performance Indicators (KPI). My dashboard includes the following key performance indicators:

Lost Revenue: Total aggregation of annual revenue lost because of lost customers.

$[\text{Monthly Charge}] * 12$

Revenue: Total aggregation annual revenue based on the loyal customer's monthly charge.

$\text{count}([\text{Customer}]) / \text{sum}([\text{Population}]) * 1000$

Lost Customers: Total number of lost customers.

%Lost Customers: # of lost customers / total customers

A3. The Data

Provide both data sets that serve as the data source for the dashboard.

Data. This task uses one (1) primary internal dataset and two (2) external datasets.

Data set 1: churn_clean.csv. This is the primary data provided by the telecom organization. It consists of 10,000 customer records. The data is broken down into the following attributes:

1. **Customer_id** is categorical (CATEGORICAL): ['K409198' 'S120509' 'K191035' ... 'I243405' 'I641617' 'T38070'].
2. **State** is categorical (CATEGORICAL): ['AK' 'MI' 'OR' 'CA' 'TX' 'GA' 'TN' 'OK' 'FL' 'OH' 'PA' 'PR' 'IA' 'ME' 'IL' 'WI' 'NC' 'AL' 'NM' 'VT' 'MD' 'NY' 'WA' 'CT' 'NJ' 'DC' 'ND' 'LA' 'NE' 'WV' 'AZ' 'MO' 'WY' 'MT' 'VA' 'KY' 'MN' 'KS' 'MA' 'IN' 'SC' 'NH' 'DE' 'MS' 'ID' 'AR' 'SD' 'CO' 'HI' 'UT' 'RI' 'NV'].
3. **Area** is categorical (CATEGORICAL): ['Urban' 'Suburban' 'Rural'].
4. **TimeZone** is categorical (CATEGORICAL): ['America/Sitka' 'America/Detroit' 'America/Los_Angeles' 'America/Chicago' 'America/New_York' 'America/Puerto_Rico' 'America/Denver' 'America/Menominee' 'America/Phoenix' 'America/Indiana/Indianapolis' 'America/Boise' 'America/Kentucky/Louisville' 'Pacific/Honolulu' 'America/Indiana/Petersburg' 'America/Nome' 'America/Anchorage' 'America/Indiana/Knox' 'America/Juneau' 'America/Toronto' 'America/Indiana/Winamac' 'America/Indiana/Vincennes' 'America/North_Dakota/New_Salem' 'America/Indiana/Tell_City' 'America/Indiana/Marengo' 'America/Ojinaga'].
5. **Children** is numerical (CONTINUOUS) - type: int64.
Unique: [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
6. **Age** is numerical (CONTINUOUS) - type: int64.
Unique: [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89]
7. **Income** is numerical (CONTINUOUS) - type: float64.
Min: 348.670 Max: 258900.700 Std: 28199.917
8. **Marital** is categorical (CATEGORICAL): ['Widowed' 'Married' 'Separated' 'Never Married' 'Divorced'].
9. **Gender** is categorical (CATEGORICAL): ['Male' 'Female' 'Nonbinary'].

10. **Churn** is categorical (CATEGORICAL): ['No' 'Yes'].
11. **Outage_sec_perweek** is numerical (CONTINUOUS) - type: float64.
Min: 0.100 Max: 21.207 Std: 2.976
12. **Email** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23]
13. **Contacts** is numerical (CONTINUOUS) - type: int64.
Unique: [0, 1, 2, 3, 4, 5, 6, 7]
14. **Yearly equip failure** is numerical (CONTINUOUS) - type: int64.
Unique: [0, 1, 2, 3, 4, 6]
15. **Techie** is categorical (CATEGORICAL): ['No' 'Yes'].
16. **Contract** is categorical (CATEGORICAL): ['One year' 'Month-to-month' 'Two Year'].
17. **Port_modem** is categorical (CATEGORICAL): ['Yes' 'No'].
18. **Tablet** is categorical (CATEGORICAL): ['Yes' 'No'].
19. **InternetService** is categorical (CATEGORICAL): ['Fiber Optic' 'DSL' 'None'].
20. **Phone** is categorical (CATEGORICAL): ['Yes' 'No'].
21. **Multiple** is categorical (CATEGORICAL): ['No' 'Yes'].
22. **OnlineSecurity** is categorical (CATEGORICAL): ['Yes' 'No'].
23. **OnlineBackup** is categorical (CATEGORICAL): ['Yes' 'No'].
24. **DeviceProtection** is categorical (CATEGORICAL): ['No' 'Yes'].
25. **TechSupport** is categorical (CATEGORICAL): ['No' 'Yes'].
26. **StreamingTV** is categorical (CATEGORICAL): ['No' 'Yes'].
27. **StreamingMovies** is categorical (CATEGORICAL): ['Yes' 'No'].
28. **PaperlessBilling** is categorical (CATEGORICAL): ['Yes' 'No'].

29. **PaymentMethod** is categorical (CATEGORICAL): ['Credit Card (automatic)' 'Bank Transfer(automatic)' 'Mailed Check' 'Electronic Check'].

30. **Tenure** is numerical (CONTINUOUS) - type: float64.
Min: 1.000 Max: 71.999 Std: 26.443

31. **MonthlyCharge** is numerical (CONTINUOUS) - type: float64.
Min: 79.979 Max: 290.160 Std: 42.943

32. **Bandwidth_GB_Year** is numerical (CONTINUOUS) - type: float64.
Min: 155.507 Max: 7158.982 Std: 2185.295

33. **Item1** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7]

34. **Item2** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7]

35. **Item3** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7, 8]

36. **Item4** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7]

37. **Item5** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7]

38. **Item6** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7, 8]

39. **Item7** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7]

40. **Item8** is numerical (CONTINUOUS) - type: int64.
Unique: [1, 2, 3, 4, 5, 6, 7, 8]

Data set 2: population.csv. The original data had a population attribute, but I wanted to include population calculations based on each state's total population. This is an external dataset from United States Census Bureau showing population data for the US states for years 2020 and 2021. The external population data is broken down as follows:

1. **NAME** is categorical (CATEGORICAL): ['United States' 'Northeast Region' 'Midwest Region' 'South Region' 'West Region' 'Oklahoma' 'Nebraska' 'Hawaii' 'South Dakota' 'Tennessee' 'Nevada' 'New Mexico' 'Iowa' 'Kansas' 'District of Columbia' 'Texas' 'Missouri' 'Arkansas' 'Michigan' 'New Hampshire' 'North Carolina' 'Ohio' 'South Carolina' 'Wyoming' 'California' 'North Dakota' 'Louisiana' 'Maryland' 'Delaware' 'Pennsylvania' 'Georgia' 'Oregon' 'Minnesota' 'Colorado' 'New Jersey' 'Kentucky' 'Washington' 'Maine' 'Vermont' 'Idaho' 'Indiana' 'Montana' 'New York' 'Puerto Rico' 'Connecticut' 'Florida' 'Virginia' 'Massachusetts' 'Illinois' 'Mississippi' 'Arizona' 'Utah' 'Wisconsin' 'Alabama' 'West Virginia' 'Rhode Island' 'Alaska'].

2. **POP_2021** is numerical (CONTINUOUS) - type: int64.

Unique: [732673, 29527941, 5039877, 21781128, 6165129, 19835913, 4246155, 1104271, 895376, 1441553, 6984723, 7276316, 57159838, 774948, 2115877, 5707390, 1963692, 4624047, 1782959, 11780017, 5190705, 2934582, 3143991, 9267130, 1095610, 1388992, 645570, 7738692, 3337975, 39237836, 2949965, 10799566, 3986639, 4509394, 1372247, 12964056, 6168187, 3605597, 3263584, 6805985, 670050, 3025891, 68841444, 5812069, 8642274, 5895908, 12671469, 331893745, 127225329, 6975218, 578803, 1900923, 3193079, 1003384, 10551162, 10050811, 78667134]

Source: NST_EST2021_POP Annual Estimates of the Resident Population for the United States, Regions, States, District of Columbia, and Puerto Rico: April 1, 2020 to July 1, 2021 (US Census Bureau, 2022)

Data set 3: states.csv. To link between churn data and the population data, I needed another indexing table. Churn data uses two (2) letter state code and the population data uses the full state name. The index table has both fields and can be used to link between tables. The external states data is broken down as follows:

1. **State** is categorical (CATEGORICAL): ['Alabama' 'Alaska' 'Arizona' 'Arkansas' 'California' 'Colorado' 'Connecticut' 'Delaware' 'District of Columbia' 'Florida' 'Georgia' 'Hawaii' 'Idaho' 'Illinois' 'Indiana' 'Iowa' 'Kansas' 'Kentucky' 'Louisiana' 'Maine' 'Maryland' 'Massachusetts' 'Michigan' 'Minnesota' 'Mississippi' 'Missouri' 'Montana' 'Nebraska' 'Nevada' 'New Hampshire' 'New Jersey' 'New Mexico' 'New York' 'North Carolina' 'North Dakota' 'Ohio' 'Oklahoma' 'Oregon' 'Pennsylvania' 'Rhode Island' 'South Carolina' 'South Dakota' 'Tennessee' 'Texas' 'Utah' 'Vermont' 'Virginia' 'Washington' 'West Virginia' 'Wisconsin' 'Wyoming'].
2. **Code** is categorical (CATEGORICAL): ['AL' 'AK' 'AZ' 'AR' 'CA' 'CO' 'CT' 'DE' 'DC' 'FL' 'GA' 'HI' 'ID' 'IL' 'IN' 'IA' 'KS' 'KY' 'LA' 'ME' 'MD' 'MA' 'MI' 'MN' 'MS' 'MO' 'MT' 'NE' 'NV' 'NH' 'NJ' 'NM' 'NY' 'NC' 'ND' 'OH' 'OK' 'OR' 'PA' 'RI' 'SC' 'SD' 'TN' 'TX' 'UT' 'VT' 'VA' 'WA' 'WV' 'WI' 'WY'].

Source: US States Names and Abbreviations data. (World Population Review, 2022)

The tables will be linked as follows:

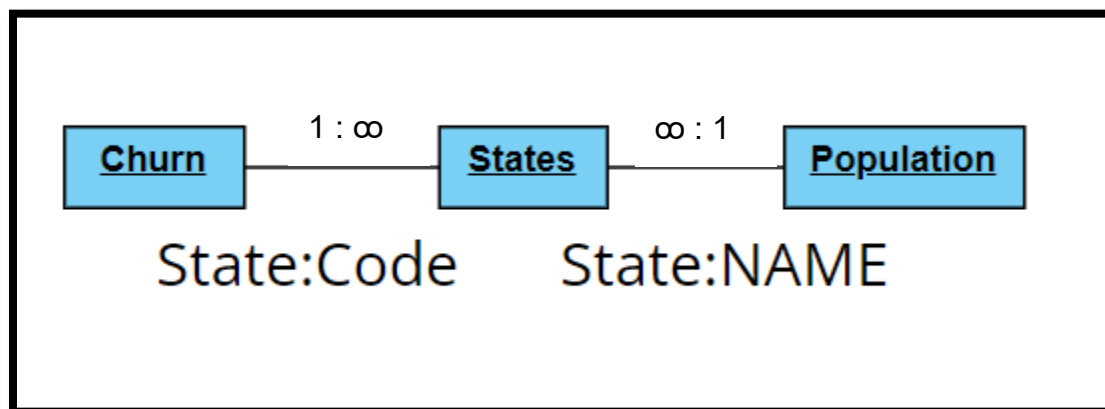


Figure 2 Data Relationships

CHURN: STATE HAS 1:MANY RELATIONSHIP WITH STATES:CODE. POPULATION: NAME HAS 1:MANY RELATIONSHIP WITH STATES:STATE

A4. Install (Create) Dashboard using Tableau Desktop

Provide step-by-step instructions to guide users through the dashboard installation.

I have created a separate document titled, “**Creating the Density Dashboard**” which explains the details of creating the sheets and dashboards.

A5. Navigate Dashboard using Tableau Public

Provide instructions to help users navigate the dashboard. There are numerous ways to navigate the dashboard, the fact that it is an interactive dashboard means the user can change the look and results of each data representation, and thereby, the user can create any number of navigational sequences.

For the purpose of this paper, I will demonstrate a typical navigational sequence by answering the following question:

- i. Step 1. ***What are the top three (3) states contributing to the highest lost revenue?*** *That is, aggregate revenue by state, sort by aggregated revenue descending and only consider lost customers. The results should show a list of states in sorted order, and we want to consider the top 3 states in that list.*
- ii. Step 2. Open the latest version of the ***Lost Revenue*** dashboard
- iii. Step 3. Make sure that you are seeing all of the domestic US states, if there were a selection filter applied, you may only be seeing a portion of the total data. Clear any of the selection filters by clicking on the blue heading to un-select it.
- iv. Step 4. While looking at all states, update the Top N parameter to 3, then click enter key. This will refresh the table and the heading will show “***Top 3 States***”

- v. Step 5. Now click on the lower table where it says “Top 3 States”, the map will refresh and the upper key measures tables will refresh data based on the selection, it should look like this:

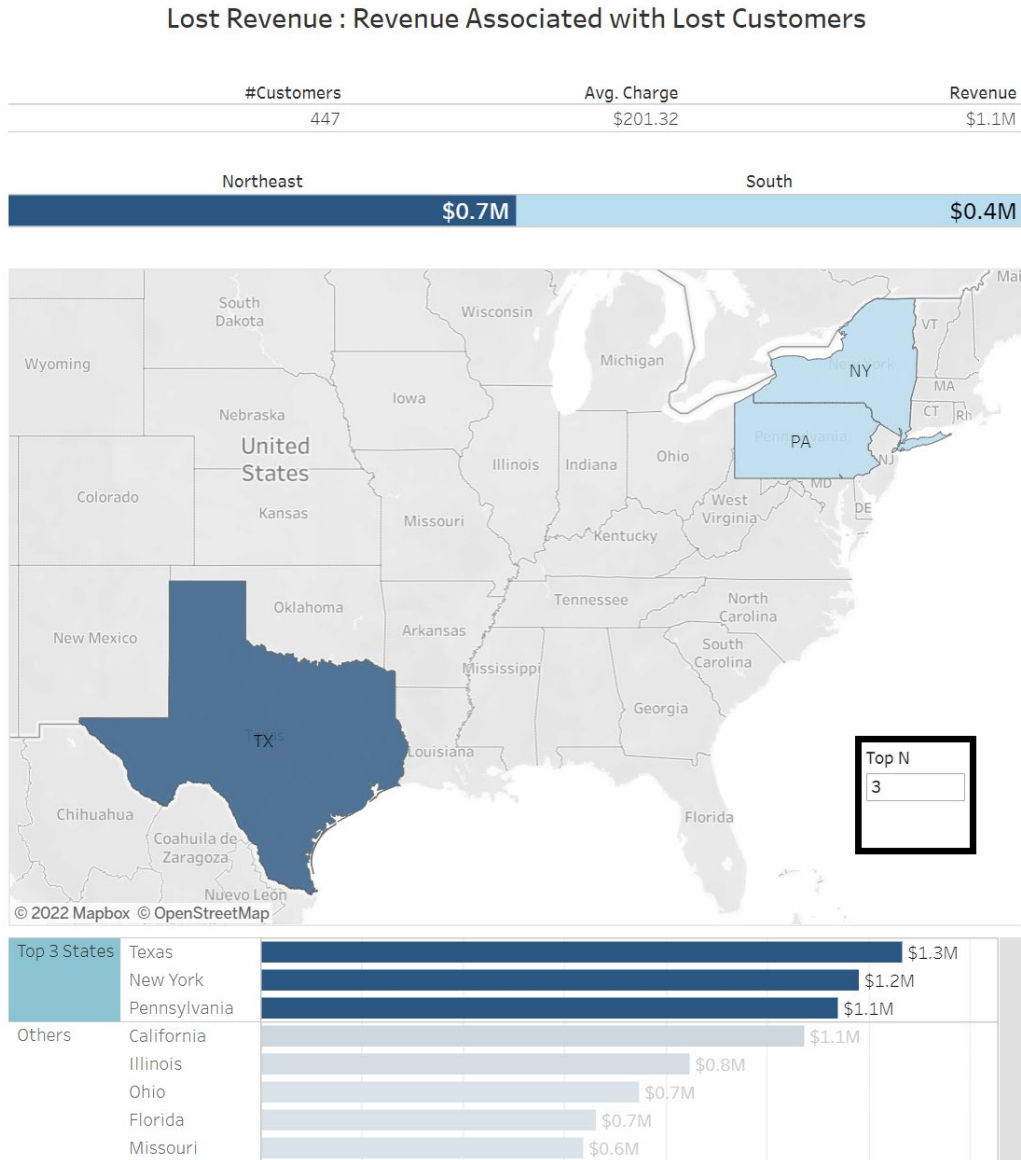


Figure 3 Top 3 States Lost Revenue

THE FIGURE SHOWS TOTAL NUMBER OF LOST CUSTOMERS, 447. THE AVERAGE MONTHLY CHARGE FOR THOSE LOST CUSTOMERS, \$201.32. THE TOTAL AMOUNT OF LOST REVENUE, \$1.1M. THE MAP REFRESHES TO SHOW ONLY THOSE STATES, **TX, PA AND NY**. THE BOTTOM TABLE SHOWS THE THREE STATES RANKED BY HIGHEST LOST REVENUE.

Part 2: Storytelling with Data

B. Presentation Video

B. Provide a link to a Panopto multimedia presentation in which you tell a story about the data to an audience of data analytics peers. Your presentation should implement elements of effective storytelling and include *each* of the following elements:

- an introduction of yourself and your background
- a summary of *both* chosen data sets and the context in which they occur
- an outline of key results from your analysis of the two data sets
- **two** different data representations to serve as supporting evidence for your results
- a summary of actionable insights based on your results

Note: Your two data representations may come directly from your Tableau dashboard created in Part A or you may choose to use another software tool for data representation.



Telecom Customer Data

- In the telecommunications industry, customers can choose from multiple service providers and actively switch from one provider to another. Customer "churn" is defined as the percentage of customers who stopped using a provider's product or service during a certain time frame. In this highly competitive market, some telecommunications industries can experience average annual churn rates as high as 25 percent. Given that it costs 10 times more to acquire a new customer than to retain an existing one, customer retention has now become even more important than customer acquisition.
- For many providers, retaining highly profitable customers is the number one business goal. To reduce customer churn, telecommunications companies need to predict which customers are at high risk of churn.
- You are an analyst on a team of analysts in a popular telecommunications company, which serves customers in all regions of the United States. You have been asked to clean the raw data set in preparation to explore the data, identify trends, and compare key metrics.
- Demographics
 - Age
 - Gender
 - Location
- Churn data
 - Churn (Yes or No)
- Revenue data
 - Monthly Charge
- Population data. Note, there was a population field in the customer data, but I was looking for a more complete state population, so I have included US Census data for my population and density calculations

Churn Data

- 1. Customer_id** is categorical (CATEGORICAL): ['K409198' 'S120509' 'K191035' ... 'I243405' 'I641617' 'T38070'].
- 2. State** is categorical (CATEGORICAL): ['AK' 'MI' 'OR' 'CA' 'TX' 'GA' 'TN' 'OK' 'FL' 'OH' 'PA' 'PR' 'IA' 'ME' 'IL' 'WI' 'NC' 'AL' 'NM' 'VT' 'MD' 'NY' 'WA' 'CT' 'NJ' 'DC' 'ND' 'LA' 'NE' 'WV' 'AZ' 'MO' 'WY' 'MT' 'VA' 'KY' 'MN' 'KS' 'MA' 'IN' 'SC' 'NH' 'DE' 'MS' 'ID' 'AR' 'SD' 'CO' 'HI' 'UT' 'RI' 'NV'].
- 10. Churn** is categorical (CATEGORICAL): ['No' 'Yes'].
- 31. MonthlyCharge** is numerical (CONTINUOUS) - type: float64. Min: 79.979 Max: 290.160 Std: 42.943

Source: Wgu.edu Course Data

US State and Regional Data

- Need to include United States state and regional data to enhance the provided company data
- States data
 - State names
 - State abbreviation
 - State population
- Regional data
 - West, Northeast, South, Midwest
- Census data
 - Population data for year 2021

US States Data

- 1. **State** is categorical (CATEGORICAL):
['Alaska' 'Alabama' 'Arkansas' 'Arizona' 'California' 'Colorado' 'Connecticut' 'District of Columbia' 'Delaware' 'Florida' 'Georgia' 'Hawaii' 'Iowa' 'Idaho' 'Illinois' 'Indiana' 'Kansas' 'Kentucky' 'Louisiana' 'Massachusetts' 'Maryland' 'Maine' 'Michigan' 'Minnesota' 'Missouri' 'Mississippi' 'Montana' 'North Carolina' 'North Dakota' 'Nebraska' 'New Hampshire' 'New Jersey' 'New Mexico' 'Nevada' 'New York' 'Ohio' 'Oklahoma' 'Oregon' 'Pennsylvania' 'Rhode Island' 'South Carolina' 'South Dakota' 'Tennessee' 'Texas' 'Utah' 'Virginia' 'Vermont' 'Washington' 'Wisconsin' 'West Virginia' 'Wyoming'].
- 2. **Code** is categorical (CATEGORICAL):
['AK' 'AL' 'AR' 'AZ' 'CA' 'CO' 'CT' 'DC' 'DE' 'FL' 'GA' 'HI' 'IA' 'ID' 'IL' 'IN' 'KS' 'KY' 'LA' 'MA' 'MD' 'ME' 'MI' 'MN' 'MO' 'MS' 'MT' 'NC' 'ND' 'NE' 'NH' 'NJ' 'NM' 'NV' 'NY' 'OH' 'OK' 'OR' 'PA' 'RI' 'SC' 'SD' 'TN' 'TX' 'UT' 'VA' 'VT' 'WA' 'WI' 'WV' 'WY'].
- 3. **Region** is categorical (CATEGORICAL):
['West' 'South' 'Northeast' 'Midwest'].
- 4. **Division** is categorical (CATEGORICAL):
['Pacific' 'East South Central' 'West South Central' 'Mountain' 'New England' 'South Atlantic' 'West North Central' 'East North Central' 'Middle Atlantic'].

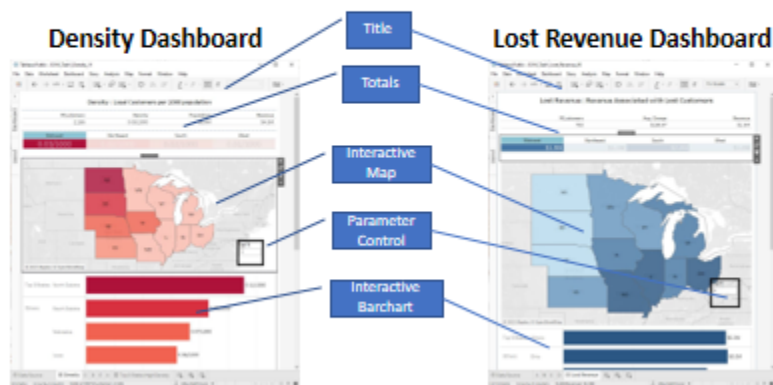
US States Names and Abbreviations data. (World Population Review, 2022)

Population Data

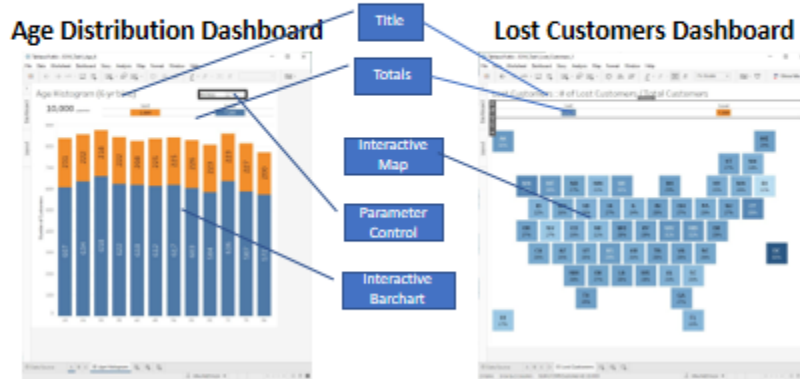
- 1. **NAME** is categorical (CATEGORICAL):
['Oklahoma' 'Nebraska' 'Hawaii' 'South Dakota' 'Tennessee' 'Nevada' 'New Mexico' 'Iowa' 'Kansas' 'District of Columbia' 'Texas' 'Missouri' 'Arkansas' 'Michigan' 'New Hampshire' 'North Carolina' 'Ohio' 'South Carolina' 'Wyoming' 'California' 'North Dakota' 'Louisiana' 'Maryland' 'Delaware' 'Pennsylvania' 'Georgia' 'Oregon' 'Minnesota' 'Colorado' 'New Jersey' 'Kentucky' 'Washington' 'Maine' 'Vermont' 'Idaho' 'Indiana' 'Montana' 'New York' 'Puerto Rico' 'Connecticut' 'Florida' 'Virginia' 'Massachusetts' 'Illinois' 'Mississippi' 'Arizona' 'Utah' 'Wisconsin' 'Alabama' 'West Virginia' 'Rhode Island' 'Alaska'].
- 2. **POP_2021** is numerical (CONTINUOUS):
- type: Int64. Unique: [732673, 29527941, 5039877, 21781128, 6165129, 19835913, 4246155, 1104271, 895376, 1441553, 6984723, 7276316, 774948, 2115877, 1963692, 4624047, 1782959, 11780017, 5190705, 2934582, 3143991, 9267130, 1095610, 1388992, 645570, 7738692, 3337975, 39237836, 2949965, 10799566, 3986639, 10050811, 4509394, 1372247, 12964056, 3605597, 3263584, 6805985, 670050, 3025891, 8642274, 5812069, 5895908, 12671469, 6975218, 578803, 1900923, 3193079, 1003384, 10551162, 6168187, 5707390]

NST_EST2021_POP Annual Estimates of the Resident Population for the United States, Regions, States, District of Columbia, and Puerto Rico: April 1, 2020 to July 1, 2021 (US Census Bureau, 2022)

Data Representations



Data Representations



Part 3: Reflection Paper

C. Reflection Paper

- C. Write a reflection paper to demonstrate your understanding of data representation and reporting by doing the following:
1. Explain how the purpose and function of your dashboard align with the needs outlined in the data dictionary associated with your chosen data set.
 2. Explain how the variables in the additional data set enhance the insights that can be drawn from the data set you chose from the provided options.
 3. Explain **two** different data representations from your dashboard and how executive leaders can use them to support decision-making.
 4. Explain **two** interactive controls in your dashboard and how *each* enables the user to modify the presentation of the data.
 5. Describe how you built your dashboard to be accessible for individuals with colorblindness.
 6. Explain how **two** data representations in your presentation support the story you wanted to tell.
 7. Explain how you used audience analysis to adapt the message in your presentation.
 8. Describe how you designed your presentation for universal access by all audiences.
 9. Explain **two** elements of effective storytelling that you implemented in your presentation and how *each* element was intended to engage the audience.

C1. Explain

The Telecom organization data contains 10,000 customer records “**in all regions of the United States**”. The organization breaks down the United States into four (4) regions and has a management team dedicated to each region.

The purpose of the **Lost Revenue** dashboard is to break down trends for lost revenue by region. The information will be made available to the company’s regional managers and enable them to focus on states with high number of lost customers.

The purpose of the **Density** dashboard is to break down trends for customer density by region. The information will be made available to the company’s regional managers and enable them to focus on states with high or low customer density.

The purpose of the **Age** dashboard is to show distribution of lost and loyal customers by age. The Age Bins is a parameter that will allow the dashboard to be changed to show different age groups.

C2. Explain

The external data for United States regions and population are included with the organization’s customer data. Each customer record has location data (City, State, County, Zip, Lat, Lng) and population data. The customer data will be linked to the United States data using the State attribute. The organization’s data shows the state as a two-letter abbreviation, the external data uses the state’s full name. Instead of using the organization’s Population attribute, showing population within a mile radius of the customer’s residence, the dashboard will include external data from US Census on the population for the entire state.

By incorporating the external data, the dashboard will be able to aggregate customer data by Region, and, in the case of the Density data, will be calculated using the entire state's population.

C3. Explain

The Density and the Lost Revenue dashboard both use a similar layout which include both a map representation and a barchart representation. In addition, there are two (2) areas where totals are show for the current selected situation. All of these components are linked together and refreshed as selections are made by the dashboard user.

The **map representation** shows the domestic United States, AK, HI, PR and DC were intentionally excluded from the data so that the map would be scaled to show just the contiguous United States. Because density is associated with loyal customers, the total number of customers on the density dashboard are the total number of loyal customers. Also, the revenue on the density dashboard is the actual aggregated revenue for those loyal customers. In contrast, the customers and revenue on the lost revenue dashboard are lost customers and lost revenue. Individual states can be selected and de-selected by using the CTRL key with the mouse. When multiple states are selected using this method, click on any state twice will clear the selection and go back to all states being shown.

The **barchart representation** is a list of all the states ranked by either density (for the density dashboard) or by lost revenue (for the lost revenue dashboard). When the user changes the Top N parameter to another value, the barchart is recalculated and changes the display to show the "Top N States". Once that happens, the user can click on the "Top N States" and the map and totals will refresh to show just those top N states. To remove the selection, click again on the "Top N States", which is shown in blue highlight, and the map and total will go back to showing all the states.

The **totals** above the map can also be selected, for example, by clicking on "Midwest", the map and barchart will refresh to show only those states in the Midwest region. Clicking on "Midwest" or the blue highlighted area again will unselect that region and all the states will be shown again.

Regional managers can use both dashboards to interactively change and focus on specific regions or states needing specific action. These dashboards should allow managers and the organization to find regions with high customer churn to find ways to reduce or eliminate that churn and realize the lost revenue attributed to those lost customers as well as highlighting regional areas with high or low density or revenue to increase density and thereby, increasing the overall revenue. Region managers of areas with high density and high revenue can share best practices with the other regions where more focus is required to increase customer and revenue base or curtail the rate at which the organization is losing customers.

C4. Explain

The primary controls for both dashboards are the “Top N” parameter and the built-in Tableau selection filters.

The Top N parameter allows the dashboard user to change the number of states highlighted at the top of the barchart. When the value is changed, the following chain of events takes place:

- The **State Set** is updated which basically divides all of the states into two (2) groups, states in the top N ranked by either Density or Lost Revenue, and states that are not in the Top N states
- The **Subset Density** (or Subset Revenue) is a calculated field that will be either “Top N States” if the state is in the State Set, or “Others” if not.

The Built-in Selection filters are the way Tableau connects the different parts of the dashboard. By selecting “Use as Filter” option, any change to that component will cascade to the other components. In the back

ground, action filters are created. You can get crazy with how these action filters are created and apply to other components. But, for these dashboards, the basic action filter is used. The end result is that when selections are made to the barchart, the map is refreshed, and, when you click on a region in the total area, both the map and barchart refresh, etc.

On the Age dashboard, the Age Bins parameter is used to re-distribute the histogram into varying sized age bins.

C5. Describe

Colorblindness. To make the dashboard more accessible to users with colorblindness, I have chosen to use color-blindness friendly palette of blue and orange. According to Shaffer (Shaffer, 2022), “one color used together in combination with another color is generally fine when one of them is not usually associated with CVD.” Common CVD color combinations are Red/Green or Blue/Yellow, with Red/Green accounting for over 90% being the most common. (Turgut & Karanfil, 2022)

C6. Explain

C7. Explain

C8. Describe

C9. Explain

References

- Mattinson, M. (2022, January 30). *Mike Mattinson*. Retrieved from Tableau Public: <https://public.tableau.com/app/profile/mike.mattinson>
- Shaffer, J. (2022, January 30). *5 tips on designing colorblind-friendly visualizations*. Retrieved from Tableau Blog: <https://www.tableau.com/about/blog/examining-data-viz-rules-dont-use-red-green-together#:~:text=For%20example%2C%20blue%20is,palette%20designed%20by%20Maureen%20Stone>.
- The Investopedia Team. (2022, February 1). *What Are the Best Metrics to Evaluate a Telecommunications Company?* Retrieved from Investopedia.com: <https://www.investopedia.com/ask/answers/122414/what-are-best-metrics-evaluate-telecommunication-company.asp>
- Turgut, B., & Karanfil, F. C. (2022, January 30). *Appropriate terminology in the nomenclature of the color vision deficiency*. Retrieved from Open Access Text: <https://www.oatext.com/appropriate-terminology-in-the-nomenclature-of-the-color-vision-deficiency.php>
- US Census Bureau. (2022, January 31). *NST_EST2021_POP*. Retrieved from data.census.gov: https://data.census.gov/cedsci/table?tid=PEPPOP2021.NST_EST2021_POP
- World Population Review. (2022, January 31). *List of State Abbreviations*. Retrieved from worldpopulationreview.com: <https://worldpopulationreview.com/states/state-abbreviations>