Advanced Data Acquisition

D211

Shantel Johnson

2023

## Section 1: Purpose

The dashboard combines AnyTelecom's Churn dataset provided by WGU as well as an external dataset sourced from Kaggle.com, known as the Cost of Living Index by City (COL). When combined, these two datasets enable executive leaders to determine which locations should receive telecom price adjustments based on local economic factors and demographic characteristics.

## Section 2: Tableau

The dashboard in this analysis was developed using Tableau. Tableau is a closed-source business intelligence tool used for creating data visualizations. It integrates with numerous data sources, such as text files, databases, and cloud storage. Tableau also offers a visual data preparation tool, Tableau Prep Builder, which is discussed further in Section 3. Since Tableau was the chosen business intelligence software used in D210, I decided to build upon my previous experience and use it again for this analysis.

## Section 3: Data Preparation

Tableau Prep Builder was used to clean and prepare the data according to the following steps:

1. Connect Tableau Prep Builder to the datasets:
   1. Download the COL.csv file from Kaggle.com and add it to Tableau Prep Builder as a text file.
   2. Connect to PostgreSQL server and add the location and customer tables in the Churn database.
2. Use the point-and-click interface to filter the data, create additional columns, and rename existing columns.
3. Perform joins:
   1. Perform an inner join on COL and location using matching city names.
   2. Perform an inner join on location and customer using the location\_id.
4. Export the cleaned and combined data as a CSV file.
5. Save the data as a new table in the Churn database.

## Section 4: Dashboard Creation

Tableau was used to create the dashboard according to the following steps:

1. Connect to PostgreSQL server and import the cleaned\_combined.csv data into the cleaned\_combined table within the Churn database.
2. Create individual visualizations (geoplots and bar charts) using data found in the cleaned\_combined table.
3. Compile the visualizations into interactive dashboards, then combine the dashboards into a Tableau Story.
4. Package the dashboard as a twb file.

## Section 5: Results

There were three key takeaways from the analysis:

1. Hawaii and Alaska had the highest cost of living.
2. Texas and California had the highest total monthly revenue.
3. Connecticut had the largest difference between average monthly income and average monthly telecom charges.

Executives can leverage these results to make informed decisions about where to adjust telecom prices. For instance, leadership might consider raising telecom prices in Connecticut since residents in that area have the highest remaining disposable income after factoring in their monthly earnings and existing monthly telecom expenses.

## Section 6: Limitations

WGU's churn database includes data from about 6000 cities in the US, while the COL dataset only has information on 100 US cities. As a result, records without matching cities were dropped from the analysis during the inner join between COL and location (approximately 8000 records). It is likely that the number of dropped records may have impacted analytical results, so I recommend redoing the analysis using an external cost of living dataset that includes a more representative sample of US cities.

## Section 7: Installation Instructions

1. Download and extract the contents of the D211.zip file.
2. In the virtual environment, navigate to and open PgAdmin.
3. Open the Churn database (Servers -> PostgreSql 13 -> Churn).
4. Create the cleaned\_combined table.
   1. Within the Churn database, navigate to Schemas. Right click Schemas, then select Query Tool.
   2. In the Query Tool tab, import the create\_cleaned\_combined.sql file (click the folder icon in the tool ribbon).
   3. Once the script is imported, execute the query (click the play button icon in the tool ribbon).
5. Import data into the cleaned\_combined table.
   1. Within Schemas, navigate to and expand Tables. The cleaned\_combined table should appear in the list of tables.
   2. Right click cleaned\_combined, then select Import/Export.
   3. In the Import/Export menu, move the slider to Import.
   4. Under File Info, click the file browser icon and select the cleaned\_combined.csv file.
   5. Move the slider so that Header is selected.
   6. Scroll to the bottom of the menu and select OK.
6. Connect the data to Tableau.
   1. In the extracted files, double click the D211.twb file to open the dashboard.
   2. If not already there, use the bottom tabs to navigate to the D211 Report. Click the first page of the D211 Report (COL and Revenue) and click Edit Connection.
   3. Enter the following connection details:
      1. Server: localhost
      2. Port: 5432
      3. Database: churn
      4. Username: postgres
      5. Password: Passw0rd!
   4. In the event that Tableau automatically navigates to the Data Source tab, simply navigate back to the D211 Report to view the dashboard.
   5. Use the navigation instructions included in the dashboard to explore the data and visualizations.

## Section 8: Web Sources

Dedb. (2018). *Cost of Living Index by Cities.* Retrieved May 2023, from Kaggle: https://www.kaggle.com/datasets/debdutta/cost-of-living-index-by-country