Mike Mattinson

Wgu.edu, D211 Advanced Data Acquisition

Task 1, Report

April 22, 2022

A1. Provide both data sets

All the SQL and external data included in .ZIP file.

A2. Provide step by step dashboard installation

Separate file, "**d211_1_12_creating_the_dashboard.pdf**" included in .ZIP file.

A3. Provide clear instructions to help users navigate dashboard

Video demonstration included with the Panapto video.   Here is a summary recap of instructions to help user navigate the dashboard:

1.   Start Tableau, open the "d211_1_12.tbwx" file
2.   All of the data should already be included
3.   Easiest way to view dashboard(s), enter presentation mode
4.   To view customer density by region
     a.   Click on Customer Density dashboard
     b.   In the top summary, select the desired region
     c.   After selecting region, the map and details will refresh
     d.   Look at the total numbers of customers and the calculated customer density for each region or state

A4. Provide a copy of all SQL code and other code supporting the dashboards

All SQL code is located in the text file, "**d211_1_12.sql**", which is included in .ZIP file.

B. Demonstration

The demonstration video is created and available on Wgu.edu Panopto website. Here is the link:

https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=2134aad4-3746-4877-b6e2-ae7f0149a963

C1. Explain how purpose and function of dashboard aligns with data

The primary research attempted to calculate the customer density broken down by US regions. The company provided data was insufficient to make and display such calculations. I found two external data files that provided a way of breaking down each customer into a specific US region. In addition, the population data in the company provided data did not include total population data by state. So, I was able to find another external data (Source: US Census.gov) that includes total population by state. The external data enables me to make the customer density calculations, the customer density a function of the number of loyal customers in a region divided by total population for that region.

C2. Justify tool

I choose to use Tableau because the course suggested it. However, having used Tableau now for two (2) separate courses, I can say that it makes it very easy to create data sources and quickly create very powerful data visualizations. Tableau allows the most basic dashboard designer, to quickly create meaningful visualizations. The basic steps to create the data connection then create the visualization using very straight-forward drag and drop mouse operations.

C3. Explain Data Cleaning

The 10,000 record customer data was ready to go.

The external US states and US census population data needed a bit of cleaning. Simply things like removing unwanted attributes (or columns).

C4. Summarize steps used to create dashboards

The following steps were used to create the dashboards:

- Connect to SQL database

- Connect to external .CSV data

- Create necessary joins

- Create Region, State collection

- Create calculated fields

- **Blend data sources using primary and secondary linking fields**

- Setup Default Properties

    - Color

    - Number Format

- Create detail worksheet

- Create summary worksheet

- Create map worksheet

- Create dashboard by combining all the previous worksheets

The complete, step-by-step, instructions are in the file, "**d211_1_12_creating_the_dashboard.pdf**", which is included with the .ZIP file.

C5. Discuss results

The dashboard(s) give intuitive summary of customer density by US region, which was the design purpose of the dashboard. In addition, another similar dashboard was created to show lost revenue by US region. Both dashboards enable regional managers to quickly access density and revenue issues.

C6. Discuss limitation(s)

The most obvious limitation is simply that the number of company customers is so small compared to the total population of all US states. It makes the customer density calculation harder to process. This was the reason I changed the customer density calculated filed "to multiply by 1000", to make the calculation easier to understand.

D Record web sources

A complete list of textbooks and web sources used are in the file, "**d211_references.pdf**", which is included with the .ZIP file.

E. Acknowledge sources

The references listed in the attached file, "**d211_references.pdf**" indicate how much I relied on internet and textbooks. Many times I would hit errors in Tableau trying to do this or that, then spend any number of hours researching internet to see how others suggested to work-around or resolve the issue.