**Performance Assessment**

SLM1 — TASK 1: DATA ANALYSIS

Bader Ale

Department of Information Technology, Western Governors University

D211 Advanced Data Acquisition

March 2024

Table of Contents

[Part I: Purpose and Function 3](#_Toc160369870)

[Part II: Business Intelligence Tool 3](#_Toc160369871)

[Part III: Data Cleaning Steps 3](#_Toc160369872)

# Part I: Purpose and Function

The purpose of the dashboard is to explore the readmission rates for the hospitals and try to uncover the root cause for said readmissions. For this assessment, both internal datasets and a publicly available dataset from the US. Census Bureau was used. Moreover, the datasets were joined, and a tableau dashboard was created to facilitate insights into the readmission rates.   
Our stakeholders encompass key figures within the organization, including the Senior Vice President of Hospital Operations (SVP), Vice President of Research (VP), and a Panel of Regional Vice Presidents (Regional VPs). By thoroughly analyzing the data and identifying the root cause, we can make informed decisions to reduce readmissions and potentially alleviate the financial burdens associated with them.

# Part II: Business Intelligence Tool

The business intelligence tool used was Tableau. Tableau serves as a valuable tool for presenting data insights to stakeholders due to its user-friendly interface, interactive dashboards, and dynamic visualizations, which make complex information easily understandable and accessible. Stakeholders can explore data from various angles in real-time, customize visualizations to their preferences, and collaborate effectively with others. Making the raw data in a format that is easily digestible for stakeholders is key to transforming data into actionable results (Leung, 2021).

# Part III: Data Cleaning Steps

In order to prepare the dataset for the analysis, some data cleaning steps needed to be performed. Firstly, detecting and treating null values is a key step before transferring creating the Tableau dashboard. This ensures that when the visualizations are made a complete picture can be presented to the stakeholders and no NULL values shown. Furthermore, treating duplicated values is also a step paramount to data analysis. If there exists duplicated values in the dataset then the information would be incorrect again depicting erroneous information that could be used by stakeholders for decision making. Several modules within the SciKit learn python package, such as the *sklearn.preprocessing* allow for such transformation of the dataset before finalization (scikit-learn developers, 2024).

# Part IV: Dashboard Creation Steps

In order to create the dashboard, several steps needed to be taken. Firstly, the appropriate database in PostgreSQL was created along with all the tables necessary. The tables created on the local server are shown in the picture below from pgAdmin:

A screenshot of a computer

Description automatically generatedAs shown in the figure above, the appropriate tables were created, namely *admission*, *complication*, *job*, *location*, *patient*, *servicesaddon* and *survey\_responses\_addon*. Next, the links were created between the tables as appropriate to produce the correct data connections. An entity relationship diagram (ERD) was then generated to verify these connections. A screenshot is A screenshot of a computer

Description automatically generatedshown below.

An external dataset was also used for the Tableau dashboard – census population dataset per state from the US Census Department was downloaded and saved for later use. This contained the latest population statistics for every state.

In order to create the Tableau dashboard, both the tables from the local PostgreSQL server and the census table had to be imported. This was easily done via the data source tab as shown below:

A screenshot of a computer

Description automatically generated

The *patient* table was joined with the *location* table on the *location\_id* column using the Custom SQL Query command from within Tableau – the SQL query is shown below:

A screenshot of a computer

Description automatically generated

This query retrieved all the information from the patient table, and also provide me with the zip code, city, state, and county for each patient by matching each patient's location identifier with the location identifier in the location table. The patient table had alias ‘p’ and the location table had alias ‘l’. An inner join was created with the location table on *location\_id.* This created a new table that was used in the creation of the dashboard.

In order to create the dashboard, both the census dataset and the newly created joined table was used. This allowed for population analysis as well as with the initial medical dataset. The dashboard created showed population vs readmissions per state as shown below.

A graph of different colored bars

Description automatically generated

# Part V: Results

The provided Tableau dashboard serves as an analytical tool for visualizing the relationship between hospital readmission rates and the population by state, presenting health care executives with actionable data. The map highlights the geographic spread of readmissions, while the bar chart allows for an at-a-glance comparison of these rates in relation to state populations, facilitating the identification of any disproportionate figures. Through its interactive capabilities, the dashboard permits a granular examination of the data, enabling the leadership to focus efforts on improving patient outcomes and pinpointing where to allocate resources effectively.

Such a resource is pivotal for detecting regional issues swiftly and monitoring the impact of health interventions over time. Hospital administrators can leverage this information to compare current performance to established benchmarks and other pertinent health care indicators. This empowers them to make well-informed choices about enhancing care quality and managing hospital resources adeptly, fostering an initiative-taking stance in health care administration. The data visualization provided by the dashboard lays the groundwork for informed and responsive decision-making in the complex landscape of health care services.

# Part VI: Limitations

The Tableau dashboard comes with its set of challenges that could affect the interpretation of its analysis by its stakeholders. The reliability of its visualizations is contingent on the quality, accuracy, and timeliness of the data input, which if compromised, could lead to incorrect interpretations. It does not include any real-time updates or predictive modeling, which limits its ability to forecast future trends or adapt to ongoing changes. Additionally, the dashboard may not incorporate critical external variables such as economic conditions or healthcare accessibility, which can have a significant impact on readmissions.

Moreover, the dashboard may inadvertently oversimplify complex data, and without proper user expertise in data analysis, there is a risk of misinterpretation, potentially resulting in poor decision-making. It is also focused on quantitative data, possibly neglecting insights that qualitative feedback can offer. The dashboard only shows relationships between variables but not why they occur, which is essential for developing effective interventions.

Works Cited

Leung, K. (2021, September). *How Tableau Helps Your Organization Achieve Greater Data Insights*. Retrieved March 2024, from Datacamp: Rader Analytics: https://www.datacamp.com/blog/how-tableau-helps-your-organization-achieve-greater-data-insights

scikit-learn developers. (2024). *scikit Learn*. Retrieved 2024, from https://scikit-learn.org/stable/modules/preprocessing.html: https://scikit-learn.org/stable/modules/preprocessing.html