Medical Data Dashboard Analysis

Javier Lopez

D211 - Advance Data Acquisition

Table of Contents

C. Reporting

1. Purpose and Function

Our dashboard informs stakeholders of hospital metrics, patient demographics, and health trends. The data dictionary states that the Senior Vice President (SVP) of Hospital Operations is “responsible for developing new initiatives to improve patient outcomes based on observed trends.” Our dashboard lets them see a quick overview of patients’ complication risks and the services rendered upon their initial hospitalization. The data dictionary also states that the SVP “is interested in the broad categorization of patient treatments and outcomes as a function of demographics, and how these trends play out across regions.” We addressed this need by including filters in our dashboard, allowing stakeholders to filter the data from a broad overview nationally and by region (Midwest, Northeast, South, and West) while allowing for a more in-depth investigation into specific States and Counties. To ensure that we met all the needs of the SVP, we made our dashboard filterable by geographic location, patient demographics, and patient health data.

Regarding the Vice President (VP) of Research, we saw that “a key focus of this VP is to oversee research initiatives to identify patterns in patient care and drive improvements in patient outcomes through strategic initiatives.” To meet the needs of the VP of Research, we ensure that patient demographics, health data, and hospital metrics work in tandem to give the best possible overview of how each variable interacts with the others. We achieved this by including a patient age and gender filter comparing demographic and geographic data to network performance. When filtered by

patient age or gender, the VP will be able to see what treatments were performed most on patients and their readmission rate. Additionally, they can compare the abovementioned variables and their effects in each region, State, or County.

The last set of stakeholders, the Panel of Regional Vice Presidents (Regional VPs), are “responsible for executing policies and managing operations in conjunction with the SVP.” We gathered that each Regional VP would be interested in seeing their region’s performance and how it compares to the others and that of the nation. To best assist the Regional VPs, we ensured that the entire dashboard was filterable by region, State, and County. The two most powerful tools we included for Regional VPs were the Performance Overview and the Geographic Analysis. The performance overview allows all stakeholders to see the network’s performance compared to the comparative metrics. It goes one step further by breaking down these key performance indicators (KPIs) by region and offering access to a network map where Regional VPs can dive deeper into their metrics to identify opportunities. A Regional VP can visualize their region using the map to determine the top and bottom performing States. The Counties visualization beside the network map can be filtered by clicking on any State on the map, where they can identify specific counties affecting the State’s metrics.

2. Justification: Business Intelligence Tool

We used pgAdmin 4 to preprocess our data before loading it onto Tableau Public to create our dashboard. Using pgAdmin 4, we took advantage of SQL’s ease of use with structured data to clean and join our two data sets in a manner that made visualization a simple task. Using SQL on pgAdmin 4, we filtered and aggregated data

on two tables before joining them. pgAdmin 4 and Tableau Public also have user-friendly interfaces that make managing data and creating interactive visualizations straightforward. By combining an open-source tool like pgAdmin 4 and Tableau Public, a free version of Tableau, we created a cost-effective solution for data preprocessing and visualization. Additionally, both pgAdmin 4 and Tableau Public allow us to scale our datasets, which are built to handle large amounts of data.

In summary, both pgAdmin 4 and Tableau Public are free tools that provide ease of use even for individuals with limited expertise and can handle large datasets. These solutions make both tools the best-suited to preprocess and visualize our data.

3. Data Preparation

Our objective with the medical data and CMS dataset was to compare the network's performance to the comparative. We began by preprocessing our data on pgAdmin 4 using SQL to achieve that objective. The first step was to create the database that would store our tables, which we called *Medical.* In the Medical database, we created two tables, *medical* and *comp.* We defined each table's columns, data types, and constraints before importing the datasets into their respective tables. The first hurdle after importing the data was finding the best way to combine our tables, which we determined would be using the *State* column.

We hit a roadblock on the *State* column as the medical data had the State name fully spelled out, and the CMS data had just the State's abbreviation. To solve this, we created a third table, *state_lookup,* with two columns, *abbr,* and *name,* where we stored each State's abbreviation and name. We then created the *StateAbbr* column on the

medical table and used the *state_lookup* table to populate the *StateAbbr* with the respective abbreviation for each State.

Now that we had a column to join our tables, we began to work on removing extraneous columns from the *medical* table that would not add value to our analysis or visualizations. We did the same on the *comp* table and removed missing values that appeared as text rather than integers. After removing the missing values, we changed the data types from text to integers, aggregated the data by State, and stored it in the *readmissions* table. Our final step was to join the *medical* and *readmissions* table on the *StateAbbr* and *State* columns and save the resulting table onto a CSV file.

4. Dashboard Creation

The first step in creating our dashboard was importing data from the *readmissions.csv* to Tableau Public. We then created multiple calculated fields that helped us segment our data, such as regions and age groups. Some other calculated fields included calculating the performance variance for the network and the comparative and the network's readmission rate. Once we had all of the calculated fields we needed, we began creating our visualizations for the dashboard, starting with the metrics visualizations.

The metrics visualizations were created on four individual sheets where we compared the readmission rates from the network and comparative to the expected readmission rate and their performance variance. These were all visualized on bar graphs to best show the difference in performance. We then created the performance map, using the longitude and latitude to plot the data geographically and color the map

with the network's performance variance. We leveraged the tooltip to display the readmission rates and performance variance for the network and comparative in each State. The second geographic visualization was the County Performance, which we used on the dashboard to display each region or state's filtered county performance.

The last visualizations we created were the demographic pie charts where we plotted the percent of patients for each complication risk, initial admission, services, age, and gender. Once all visualizations were created, we built our dashboard by dragging them into our desired layout and set all visualizations, except for performance graphs, as filters. We used the *Show/Hide* button to display navigation instructions, add a second layer to our dashboard, and reduce crowding. On the second layer, we set the geographic overview, including the map and county chart. The last step in fine-tuning our dashboard was to set the filter settings to filter all graphs and charts within the dashboard.

5. Results and Executive Decision-Making

Our analysis showed that most patients are sixty-five and over and have the highest readmission rates of all age groups. We also found that the Northeast is the best-performing region with a -18.70% variance, which is not great. Overall, the network is significantly over the nationally expected readmission rate, 19.73%, while the nation averages 0.69% below the expected rate. Our most notable observation was the readmission rate of nonbinary patients compared to male or female patients. Nonbinary individuals have a significantly higher readmission rate in the West, Midwest, and

Northeast. However, in the South, nonbinary patients have about 11% fewer readmissions than male or female patients.

These results support executive decision-making by prompting Regional VPs to dig deeper into how demographics play a role in the regional performance and develop action plans to ensure that all patients receive the same level of care across all regions. The resulting data also alarm all executives that significant work must be done to ensure that the network is doing all it can to align itself with the national average readmission rate. Steps to achieve that goal can take many forms, such as sharing best practices across regions, creating S.M.A.R.T. goals for individual States or Counties, or reviewing protocols for emergency admissions, which comprise about half of all initial admissions.

6. Limitations

The most significant limitation in our data analysis is not having time series data where we can see trends in performance. By having a dataset that does not specify dates, we can assume that all admissions took place around the same time, making it difficult to assess progress. The only other limitation we encountered when analyzing our data was not having the reason for readmission, which is the one factor impacting the purpose of readmission tracking performed by the Center for Medicare and Medicaid Services overall.

D. Sources

1. *Hospital Readmissions Reduction Program*. Center for Medicare & Medicaid Services. (2023, January 17). Retrieved May 1, 2023, from https://data.cms.gov/provider-data/dataset/9n3s-kdb3