D210 – Representation and Reporting

WGU M.S. Data Analytics

Lyssa Kline

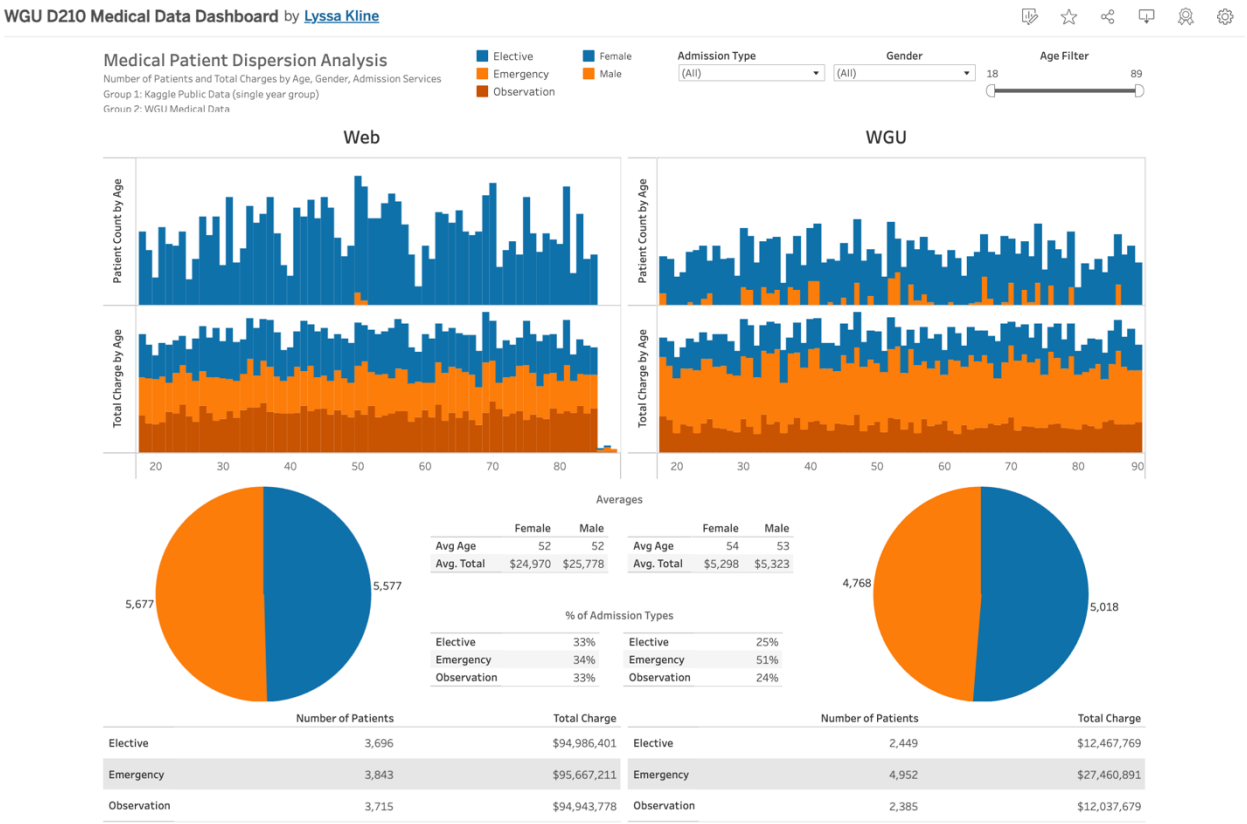
January 15, 2025

A, Part 1: Tableau Report

A1, Dashboard

The tableau dashboard for this assignment includes a Dispersion Analysis performed on medical data. This report consists of a side-by-side comparison using two different datasets, four different data representations, two interactive controls, and two KPIs computed using the data from both data sets.

The tableau dashboard has been published publicly and can be found at WGU D210 Medical Patient Dispersion Analysis.



A2, Data Sources

Two datasets were used for this dashboard analysis. See below.

Dataset #1

- WGU Provided Medical Clean Data
- CSV File: WGU Medical Clean Data
- The dataset consisted of key patient demographics and services performed while admitted into a hospital.

Dataset #2

- Kaggle Open-Source Public Medical Data

- CSV File: Kaggle Medical History Data
- Open-Source Link: Kaggle Dataset Link
- Data has been created to serve as a resource for data science, this data contains key patient demographics and services performed while admitted into a hospital. Data was chosen due to its similarity to the WGU-provided data.

A3, Installation Instructions

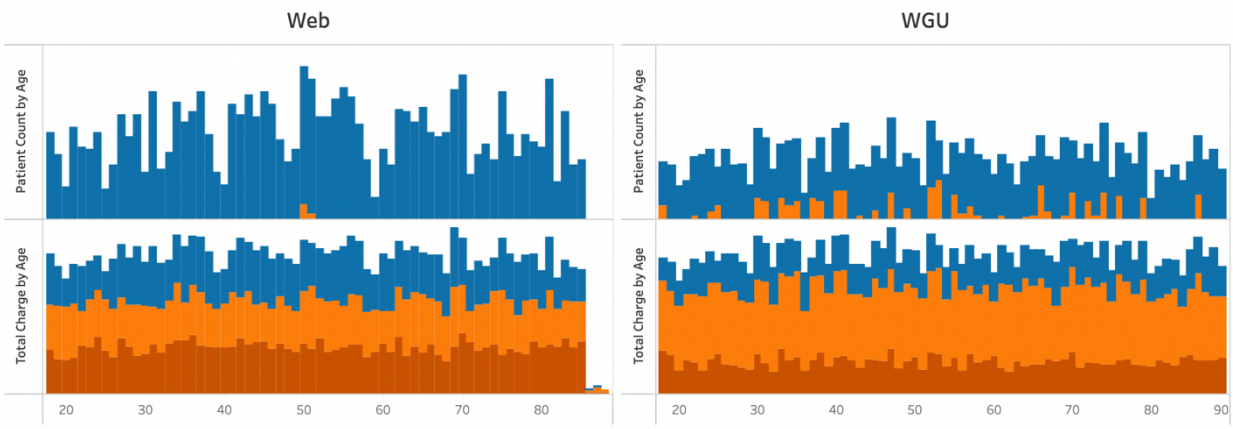
The created dashboard for this assignment has been published to Tableau Public; therefore, no installation is necessary to view it. Users can access the dashboard by clicking here, the provided link will automatically open the dashboard on a web browser and is accessible to anyone.

A2, Navigation Instructions

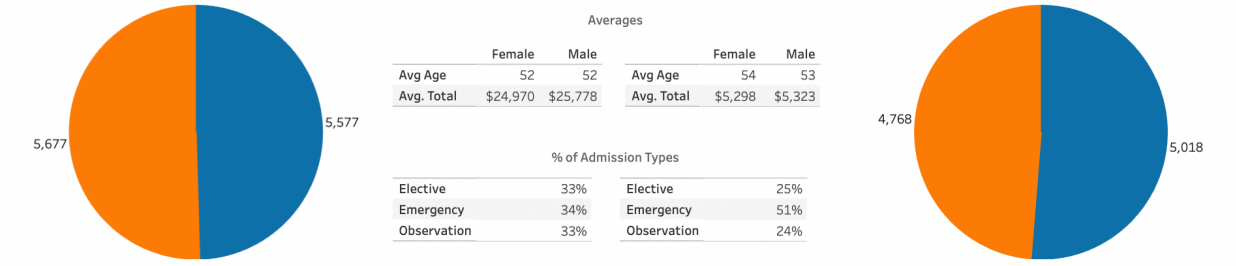
Navigating the dashboard on Tableau Public is simple; the provided link above will allow you to open the dashboard and begin navigating its interactive controls. Due to the two datasets' similarities, the data inside the dashboard has been joined together. By doing so, users can view the same metrics between both datasets side by side with the filters working for both outputs.

The section at the top that looks like a header is where you will find the title, color legends, variable filters, and a note section outlining key data details. The dashboard is an analysis of medical data, therefore the title of 'Medical Patient Dispersion Analysis' has been chosen. This indicates that the dashboard is showing a dispersion of medical patient data. Within this section, you get an overview of the two different datasets contained, one of which being the WGU-provided Medical Data, the other being a publicly posted Kaggle source. Next to this information, you can see the color legends selected for the various variables that are used throughout the dashboard. Lastly, in this section, there are three filters, admission type, gender, and age group. These filters allow users to interact with the dashboard and select a certain control group for viewing.

The dashboard is split in a side-by-side comparison with the Web (Kaggle) data source on the left-hand side and the WGU (provided) data source on the right-hand side. The first set of charts shows you a patient count dispersion and a total charge dispersion by age, the height of each bar on the bar chart indicates the patient count or the total charged to the patient. An additional variable was dropped on top of these bar charts to show you how the total charge/patient count is impacted by the admission types. The admission types are 'Elective', 'Emergency', and 'Observation' the legend for which color corresponds to which variable is in the header at the top.



Below the bar graphs, you can see two pie charts. These pie charts show the count of patients within each gender group. This visual allows you to see what the patient's makeup is based on gender. The color legend for the male and female gender groups can be found in the header section at the top.



The bottom portion of the dashboard shows specific numbers within a chart, indicating the total number of patients and the total charge for each of the admission types (elective, emergency, and observation). This section guides users into a high-level view of numbers, allowing users to have a better grasp of how many patients are included within each dataset and how much the total charge is.

	Number of Patients	Total Charge
Elective	3,696	$94,986,401
Emergency	3,843	$95,667,211
Observation	3,715	$94,943,778

	Number of Patients	Total Charge
Elective	2,449	$12,467,769
Emergency	4,952	$27,460,891
Observation	2,385	$12,037,679

The last part of this dashboard is the middle section which contains the computed KPIs. The average age and average total spent were computed for each of the genders. Additionally, the percent of admission type is displayed below. These KPIs were computed using tableau functions on the datasets.

Averages

	Female	Male
Avg Age	52	52
Avg. Total	$24,970	$25,778

	Female	Male
Avg Age	54	53
Avg. Total	$5,298	$5,323

% of Admission Types

Elective	33%
Emergency	34%
Observation	33%

Elective	25%
Emergency	51%
Observation	24%

B, Part 2: Multimedia Presentation

The video recording for this assignment includes a vocalized demonstration of the dashboard, the dashboard being used, and a discussion of the charts created within it.

The video recording for this project can be found inside the Panopto drop box titled “D210-Kline”.

Panopto video link: https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=fd2ee35e-9d85-4df7-9e49-b2660139da1d

C, Part 3: Reflection Paper

C1, Purpose and Function

The chosen dataset provides medical data that enables analysts to predict hospital readmissions however it contains a variety of fields allowing analysts to pull insights from different directions. The direction taken for this analysis was broad in scope, I did not use readmission as used previously and instead decided to look at the overall customer base, reason for admission, and total charges.

The primary purpose of this dashboard is to provide executive leaders with an interactive tool for analyzing patient data across different demographics, admission types, and other variables. It is designed to support data-driven decision-making by presenting clear, actionable insights about patient distribution and healthcare costs. This dashboard combines data from two sources (a Kaggle dataset and a WGU-provided dataset) to allow comparisons and uncover trends that can guide operational strategies, research initiatives, and resource allocation.

The side-by-side comparison of "Web" and "WGU" datasets allows users to identify similarities, differences, and trends across data sources. This dual-dataset approach enables benchmarking and validation of insights.

The function of this dashboard aligns with the needs of the organization utilizing it. For the Senior VP of Hospital Operations, the dashboard helps identify demographic and admission trends that inform operational strategies, resource allocation, and staffing. For instance, identifying high costs for a specific demographic can lead to targeted cost-reduction measures. For the VP of

Research, the dashboard aids in spotting patterns that warrant further investigation, such as trends in healthcare utilization or specific demographics requiring additional focus. Additionally, for the Regional VPs, the dashboard allows comparisons across regions or datasets, helping benchmark performance and standardize practices across the organization. Lastly, for the Data Analytics Team, it serves as a foundation for generating insights, improving data storytelling, and supporting evidence-based recommendations.

In summary, this dashboard provides a comprehensive, interactive, and user-friendly platform to explore patient data, ensuring leaders make informed decisions that improve operational efficiency and patient outcomes.

C2, Variable Driven Insights

The additional dataset from Kaggle was chosen due to its similarities to the WGU dataset. The selected dataset contained the same base fields used in this dashboard (i.e. gender, admission type, total costs, patient counts, etc.). Due to the similarities, the two datasets were joined together and then used in the same visuals. In pulling a dataset with the same variables we can perform a side-by-side comparison of "Web" and "WGU" datasets which allows users to identify similarities, differences, and trends across data sources. This dual-dataset approach enables benchmarking and validation of insights.

By comparing the two datasets, users can benchmark their organization's performance against broader, external trends. In doing so we can start to ask questions like are the age groups or admission types with the highest costs similar across both datasets? Or is the gender distribution consistent or do significant differences emerge that require further analysis? The charts displayed for both datasets enable direct visual comparisons allowing users to spot differences in patient demographics and cost patterns.

The comparison allows users to detect trends that are common across both datasets such as age groups with high healthcare utilization, and similar admission type distributions. These trends can confirm broader insights that are reliable for strategic planning and decision-making. Additionally, the comparison may allow for discrepancies to be identified, if the WGU dataset shows a significantly higher proportion of emergency admissions compared to the Web dataset, it may indicate systemic differences in care delivery, patient demographics, or reporting standards. Such differences prompt deeper investigation into whether these variations are due to operational practices, regional demographics, or other factors.

In conclusion, the side-by-side comparison of the "Web" and "WGU" datasets offers a robust framework for benchmarking, validating insights, and identifying unique trends or discrepancies. This feature empowers executive leaders to make data-driven decisions based on a broader context, ensuring that strategies are well-informed and aligned with both internal performance and external industry standards.

C3, Data Representations

One representation from the dashboard shows the distribution of patient counts and total charges segments by age, broken down further by admission type. This can be used by a few parties to draw some insights into the patient population. The SVP of Hospital Operations could

use this data to identify which age groups are most frequent in hospital visits and their associated costs. This insight helps in resource allocation, such as staffing and specialized care services for specific demographics. Additionally, the VP of Research can use these trends to investigate the underlying causes of higher hospital visits or charges for specific age groups. For example, if younger patients show high total charges, it could highlight areas requiring further research into acute conditions affecting them.

Another representation from the dashboard includes a pie chart that displays the percentage breakdown of the gender groups indicating the proportion of male and female patients in each dataset. This chart could be used by the Senior VP of Hospital Operations to ensure equitable resource allocation and service planning based on gender representation. For example, if one gender significantly dominates the patient count it might highlight a need for tailored programs or services (e.g. women's health initiatives or men's preventative care. Additionally, the VP of Research can analyze these gender-based patient trends to guide research into specific healthcare needs. For instance, if female patients represent a high proportion of the total count, research could focus on understanding the conditions driving this trend and how outcomes can be improved.

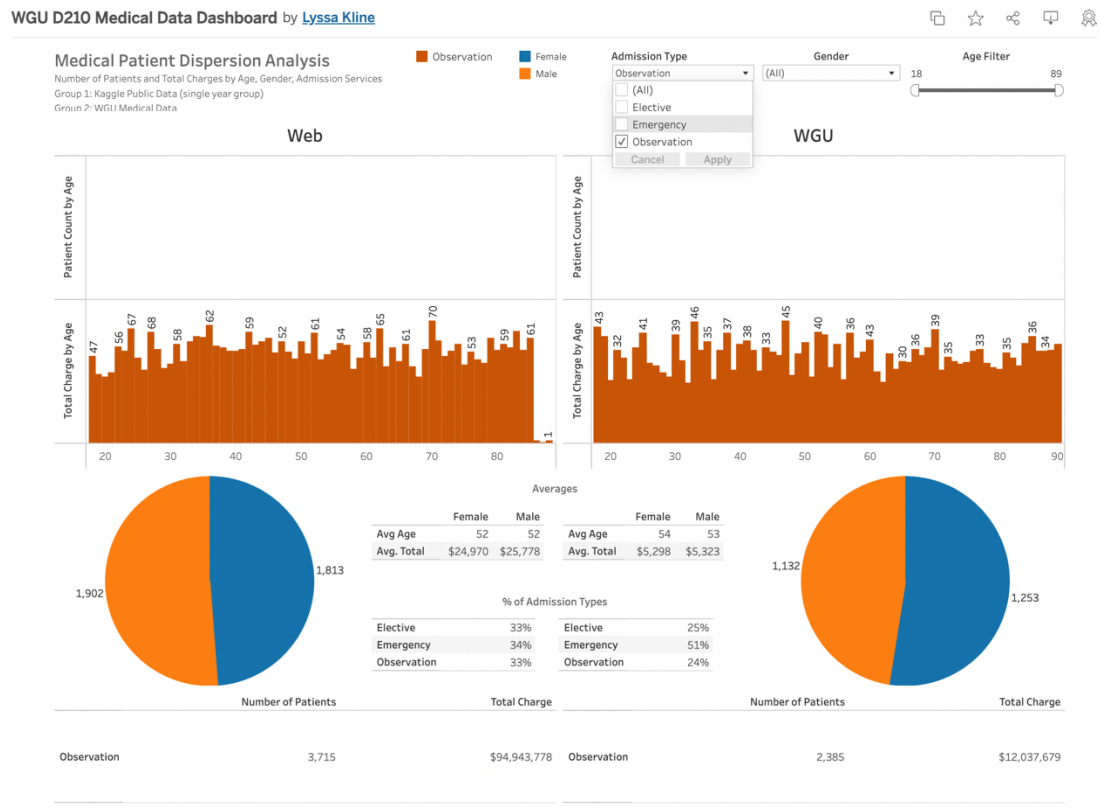
These representations align with the interest of the executive audience and support strategic initiatives by focusing on patient demographics and admission trends.

C4, Interactive Controls

The interactive controls displayed at the top of the dashboard allow users to navigate and filter the dashboard on a few different metrics. There are two filters at the top displayed as dropdowns, the first being admission type and the second being gender.

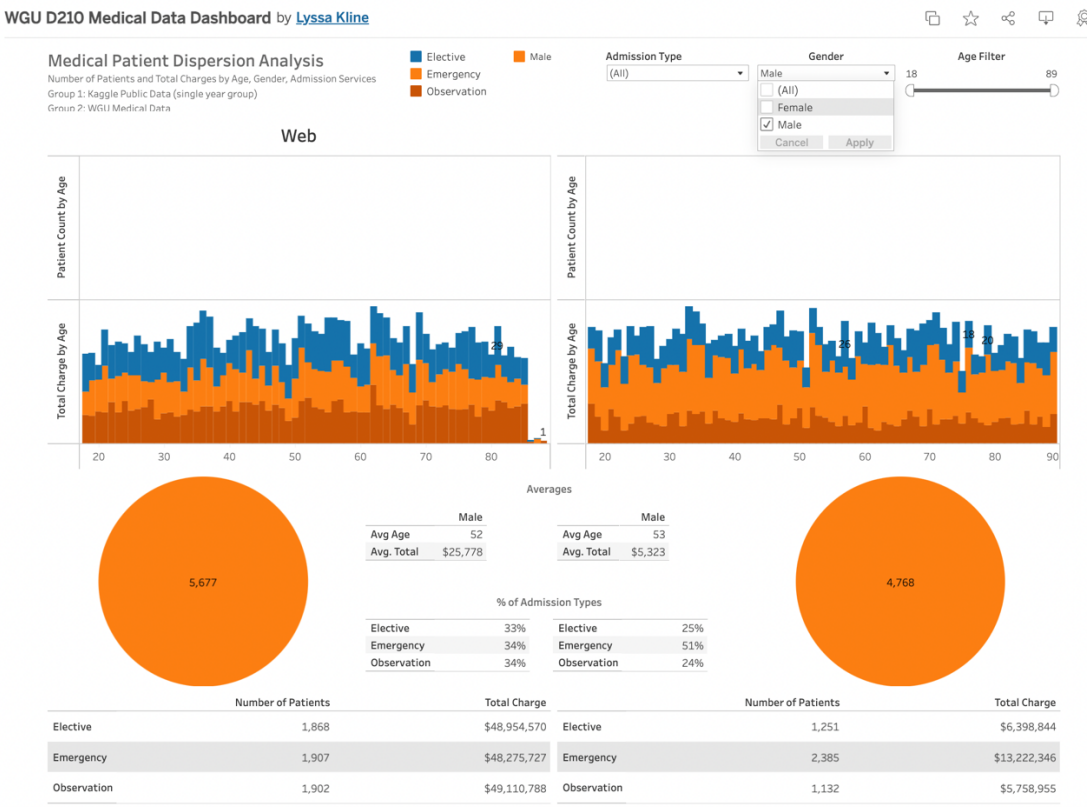
The admission type control allows users to filter the dashboard's visualizations by specific admission types, electives, emergencies, or observations. When a user selects a particular admission type, the graphs and metrics update dynamically to display data related only to the chosen type. For instance, selecting 'emergency' will show patient counts, total charges, and gender distribution specific to emergency admissions. This helps users focus on a specific category of patient care and analyze trends within that subset. Users can use this filter to examine operational efficiencies and resource usage for each type of admission, enabling more informed decisions on staff allocation and facility readiness.

Example of filtering using the admission type control:



The gender control enables users to filter the data by gender, displaying insights exclusively for male or female patients. Applying this filter updates all visualizations (e.g. bar charts, pie charts) to show patient counts, total charges, and admission-type proportions for the selected gender. This helps users explore gender-specific trends and disparities. Users can use this filter to identify gender-specific healthcare needs, ensuring that resources, facilities, and services are equitably distributed.

Example of filtering using the gender control:



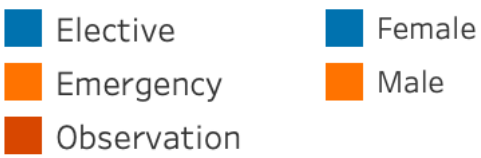
These interactive controls empower executive leaders to customize their view of the data, focusing on specific areas of interest to support data-driven decision-making.

C5, Colorblindness Accessible

Creating a Tableau dashboard that is accessible for individuals with colorblindness involves a combination of design principles, data visualization techniques, and thoughtful adjustments to ensure that the information is conveyed effectively to users with various types of color vision deficiencies.

Tableau offers several built-in color palettes. One palette offered by Tableau is optimized for colorblindness, this palette is called "Colorblind 10" (and is a set of 10 distinguishable colors). For this dashboard, I chose only 3 colors and decided to select these from Tableau's colorblind pallet.

A legend with the selected colors and corresponding metrics is laid out at the top of the dashboard, like the following.



The colorblind palette uses shades and hues that are easily distinguishable by people with various forms of colorblindness. The use of blues and oranges (the colors selected) allows all viewers to distinguish between gender or other categories effectively.

Using a colorblind palette ensures inclusivity and accessibility for executive leaders or stakeholders who may have color vision deficiencies. It reduces the risk of misinterpretation and ensures that critical insights are clear to everyone reviewing the data.

C6, Data Story

The bulk of the story can be told using the bar graphs and the pie charts. The bar graphs visually represent patient distribution and healthcare costs across different age groups, further segmented by admission type. This data highlights key demographic trends, such as which age groups account for the highest patient volume and associated charges. The story emphasizes understanding demographics and financial drivers in healthcare. These bar graphs help communicate which groups (e.g., middle-aged or elderly patients) are the primary users of healthcare services and where the most significant costs are incurred. This representation also aligns with the narrative that healthcare organizations must allocate resources effectively based on demographic patterns. For example, the high charges for older patients may signal a need for additional geriatric care resources or preventive measures to reduce costs.

By focusing on age-related trends, the story becomes relatable to executive leaders, emphasizing operational priorities (for the Senior VP of Hospital Operations) and research opportunities (for the VP of Research). It provides actionable insights into how specific age groups impact the organization's operations and finances.

The pie charts illustrate the proportion of male and female patients and the breakdown of admission types (Elective, Emergency, and Observation). The gender distribution chart supports the story of equitable resource allocation by showing whether there is a balanced utilization of healthcare services between male and female patients. Any imbalances may prompt further exploration into gender-specific healthcare needs or disparities. The admission-type pie chart highlights the importance of understanding care delivery trends. For example, a high proportion of emergency admissions may indicate a need for better preventive care programs to reduce emergency visits.

These visualizations simplify complex data into digestible insights, helping the audience quickly identify key trends. They support the story by reinforcing the need for strategic initiatives, such as improving emergency response efficiency or tailoring services to gender-specific health trends.

These data representations were carefully chosen to enhance the narrative by visually highlighting the most critical aspects of patient demographics, healthcare costs, and care delivery. They support the story by making the data both engaging and actionable, aligning perfectly with the goals of the audience (executive leaders and data analysts).

C7, Audience Analysis

The dashboard message was designed to align with the needs, priorities, and expertise levels of the target audience.

For the Executive Leaders (SVP of Hospital Operations and VP of Research), insights are presented at a high level, focusing on key metrics (e.g., patient count, total charges, admission type proportions) to facilitate strategic decision-making without overwhelming them with technical details. Visuals such as pie charts and bar graphs make complex data easily digestible, helping them quickly grasp trends and actionable insights. The interactive filters allow these leaders to explore data relevant to their specific interests, such as gender-based trends or admission types.

For the Regional VPs, comparisons between datasets provide benchmarking opportunities for identifying performance gaps or regional differences. Emphasis is placed on actionable insights to help VPs implement operational changes across their regions.

For Data Analytics Peers, more granular data and interactive features cater to the analytical mindset of this audience. They can dive deeper into patterns and validate findings, supporting their need for methodological rigor.

This audience analysis ensured that the presentation strikes a balance between simplicity for decision-makers and depth for analysts, making it relevant and accessible to all.

C8, Presentation Design

The presentation was designed with universal accessibility in mind to ensure inclusivity. Specifically, a colorblind palette was used to ensure that visual elements were distinguishable by viewers with color vision deficiencies. For instance, blue and orange were chosen to represent gender, avoiding problematic color combinations like red and green.

Additionally, filters for gender and admission type allow users to customize their view of the data, ensuring relevance to their specific needs. Charts, graphs, and pie charts are easy to interpret, minimizing cognitive load and enabling all audiences to understand the data without requiring technical expertise.

The dashboard also was designed by using clear fonts, adequate spacing, and concise labeling to enhance readability and comprehension for users of all abilities. Lastly, side-by-side views make it easier for audiences to identify trends and discrepancies without needing advanced analytical skills.

These features ensure the presentation is usable and meaningful for a diverse audience, regardless of their technical background or accessibility needs.

C9, Storytelling Elements

The dashboard tells a logical story by breaking the data into key sections (e.g., patient demographics, admission types, and costs). Each section builds upon the previous one, guiding the audience from broad overviews (e.g., age and gender distributions) to actionable insights (e.g.,

admission type breakdowns, financial trends). This structure ensures that the audience can follow the analysis easily, keeping them engaged and focused on the key message without being overwhelmed.

By using both the WGU dataset (internal data) and the Web dataset (external benchmark), the presentation creates a relatable context for the audience. Leaders can see how their organization's performance compares to broader industry trends, making the insights more relevant and impactful. This dual-dataset approach allows the audience to see the "big picture" while also connecting the data to their specific responsibilities, keeping them invested in the outcomes.

These storytelling elements were designed to capture attention, make the data relatable, and drive engagement, ensuring that the insights presented resonate with the audience and inspire informed decision-making.

D, Web Sources

Healthcare Data: Dummy Data with a Multi-Category Classification Problem on Kaggle is the dataset used for this project. The Kaggle site listed key information about the demographics and contents of the dataset. This site was utilized to gain an understanding of the dataset's variables.

E, Acknowledge the Sources

I acknowledge that no segments of third-party sources were directly stated or copied from the web into this report.