**Part 1: Interactive Data Dashboard**

1. A copy of the tableau dashboard is provided as an attachment to this submission.
2. We provide both the WGU medical dataset and UCI diabetic dataset (Strata et al, 2014) as attachments to this submission. In addition, the UCI dataset, data dictionary, and additional information are available at the UCI Machine Learning Repository.
3. To recreate this dashboard, follow the steps below after downloading the WGU and UCI datasets and ensuring Tableau is properly installed.

Step One: Create a new Tableau Project

1. In the top left corner of Tableau, select “File” and then “New” to create a new project.

Step Two: Load the data.

1. In the bottom left corner of Tableau, select “Data Source”
2. Click the button “Add” in the Connections area of the menu to the left.
3. Select the file for the WGU dataset on your local drive.
4. Select “Add” again and select the UCI dataset on your local drive.
5. Drag the WGU dataset to the open space to the right.
6. Drag the UCI dataset to the open space to the right.
7. Click “OK” in the resulting dialogue box to connect the data sources.

Step Three: Creating Worksheet One

1. Click on the empty worksheet in the bottom bar, immediately to the right of “Data Source”, and click the arrow next to the WGU data set in the “Data” section located to the left of the worksheet’s display area to show the variables available in this dataset.
2. Right click on the Initial\_days variable and select “Bins” from the “Create” submenu.
3. In the dialogue box that appears, set the bin size to one.
4. Drag the newly created Initial days (bin) variable to the columns section of the worksheet, found above the worksheet’s display area.
5. Next, drag the Re Admis variable to the rows section of the worksheet, just below the “Columns” section.
6. Click the arrow on the right of the variable’s bubble and select “Count” from the “Measures” section of the drop-down menu.
7. Drag another instance of the Re Admis variable to the “Color” box in the worksheet’s “Marks” section.
   * 1. By clicking on the box next to the variable’s name, you can set the colors of the grouping to best suit your liking, but we recommend selecting colors from the colorblind palette.

Step Four: Creating Worksheet Two

1. Collapse the WGU dataset by clicking the arrow to the left of the dataset name as you did in Step Three, part 1.
2. Expand the UCI dataset by clicking the arrow to the left of the dataset name as you did in Step Three, part 1.
3. Create a new worksheet by pressing CTRL+M on your keyboard.
4. Create a binned measure for the Time In Hospital variable, following the process in Step Three parts 2-3.
5. Drag this new variable to the “Columns” section of the worksheet.
6. Add the Readmitted variable to the “Rows” section of the worksheet and convert it to a count measurement as we did with Re Admis in Step Three part 6.
7. Set Readmitted as a grouping parameter like you did in Step Three part 7.
   * 1. These colors should be from the same palette as the previous worksheet, and we again recommend using the colorblind palette.
8. Under the Analysis tab, located in the toolbar at the top of the window, select “Percentage of” and set this parameter to “Column” so that each column shows the percentage of patients each group makes up.

Step Five: Creating Worksheet Three

1. Create a new worksheet by pressing CTRL+M on your keyboard.
2. Place the variables Number Diagnoses and Time In Hospital in the “Column” section of the workbook.
3. Place the variable Number Emergency in the “Row” section.
4. Set each of these variables to dimensions by selecting “Dimensions” from the drop-down menu accessible by clicking the arrow on the right of each variable’s bubble in the “Columns” and “Rows” section respectively.
5. Select the option “All” in the “Marks” section of the worksheet.
6. Place the Readmitted variable in the “Color” box of the “Marks” section under the heading “All” to group the data points with the same colors as in the previous worksheet.
7. Add the variable Weight to the “Filter” section and in the resulting pop-up, select all options except the value “?” to eliminate all values from the dataset with an unknown weight.
8. Under the “Analysis” tab, select “Trend Lines” then “Show All Trend Lines” to add the trend lines to the graphs in this worksheet.

Step Six: Creating Worksheet Four

1. Create a new worksheet by pressing CTRL+M on your keyboard.
2. Place the Number Diagnoses variable in the “Columns” section.
3. Place the Time In Hospital variable in the “Rows” section.
4. Repeat parts 4-8 of Step Five in the worksheet four.

Step Seven: Creating Worksheets Five through Fifteen

1. Create a new worksheet by pressing CTRL+M on your keyboard.
2. Right-click in the blank area of the data tab on the left of the worksheet and select “Create Parameter.”
   * 1. If the blank space is too small, collapse the data sources to create more room.
3. In the resulting pop-up, set the data type to “String” and select the option with the value “List” in “Allowable Values”
4. Add the following options in the “List of values” box:
   * 1. Allergic Rhinitis
     2. Anxiety
     3. Arthritis
     4. Asthma
     5. Back Pain
     6. Diabetes
     7. High Blood Pressure
     8. Hyperlipidemia
     9. Overweight
     10. Reflux Esophagitis
     11. Stroke
5. Rename the parameter to your liking and select ok.
6. Under the “Analysis” tab in the top toolbar, select “Create Calculated Field.”
7. Name the calculated field something different from your parameter.
8. Type the name of your created parameter in brackets in the text field of the resulting pop-up.
   * 1. For example, if the parameter was named Select Diagnosis, the text box of the calculated field would contain “[Select Diagnosis]” without quotes.
9. Drag this calculated field to the filters section of the worksheet and click “OK” in the resulting pop-up.
10. Place the Re Admis variable in the “Colors” box of the “Marks” section, as you did in Step Three part 7.
11. Place the Initial days variable in the “Columns” section.
12. Right click on the parameter created earlier, located below the variables in the “Data” section to the left of the worksheet’s display area, and select “Show Parameter”.
13. In the “Analytics” tab, next to the “Data” tab to the left of the worksheet’s display area, click “Average Line” under “Summarize”.
14. Duplicate this worksheet ten times by clicking the “Duplicate Worksheet” button just below the “Analysis” tab.
15. Place the following variables in the “Rows” section of the eleven worksheets, one variable per worksheet.
    * 1. Allergic Rhinitis
      2. Anxiety
      3. Arthritis
      4. Asthma
      5. Back Pain
      6. Diabetes
      7. High Blood Pressure
      8. Hyperlipidemia
      9. Overweight
      10. Reflux Esophagitis
      11. Stroke
16. In Worksheet Five, set the parameter, located to the right of the worksheet’s display, to the value that matches the variable in the “Rows” section.
17. In the calculated field of the worksheet’s Filters section, select “Edit Filter” and select the variable shown in the “Rows” section.
18. Repeat parts 15 and 16 above for worksheets six through fifteen.

Step Eight: Creating the Dashboard

1. Create a new dashboard by clicking on the “Dashboard” tab in the top toolbar and selecting “New Dashboard”.
2. In the “Dashboard” tab to the left of the dashboard’s display area, set the size to “Generic Desktop (1366 x 768)”.
3. At the bottom of the “Dashboard” tab, select the option “Floating” and check the box for “Show dashboard title”.
4. Drag Sheet 1 from the “Sheets” section of the “Dashboard” tab to the display area.
5. Click on Sheet 1 in the display area and navigate to the “Layout” tab to the left of the display area.
6. Deselect the option “Show Title” and change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface, application

Description automatically generated

1. Click on the filter accompanying Sheet 1, titled Re Admis, and change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface, application

Description automatically generated

1. Drag Sheet 2 from the “Sheets” section of the “Dashboard” tab to the display area.
2. Click on Sheet 2 in the display area and navigate to the “Layout” tab to the left of the display area.
3. Deselect the option “Show Title” and change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface

Description automatically generated

1. Click on the filter accompanying Sheet 2, titled Readmitted, and change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface, application

Description automatically generated

1. Drag Sheet 3 from the “Sheets” section of the “Dashboard” tab to the display area.
2. Click on Sheet 3 in the display area and navigate to the “Layout” tab to the left of the display area.
3. Deselect the option “Show Title” and change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface, application

Description automatically generated

1. Drag Sheet 4 from the “Sheets” section of the “Dashboard” tab to the display area.
2. Click on Sheet 4 in the display area and navigate to the “Layout” tab to the left of the display area.
3. Deselect the option “Show Title” and change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface

Description automatically generated

1. Drag Sheet 5 from the “Sheets” section of the “Dashboard” tab to the display area.
2. Click on Sheet 5 in the display area and navigate to the “Layout” tab to the left of the display area.
3. Deselect the option “Show Title” and change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface, application

Description automatically generated

1. Repeat steps 18-20 for sheets 6-15, using the same values for x, y, w, and h as above.
2. Right click Sheet 5, or whichever of the stacked graphs is visible at the moment, and select “Filters” then click on the parameter we completed earlier.
3. In the “layout” tab, change the values of x, y, w, and h to match those in the screenshot below.

Graphical user interface, application

Description automatically generated

1. Double-click on the title of the dashboard and replace the text in the resulting pop-up with “Patient Readmittance and the Importance of Prioritization” without quotes and press “Apply” then “OK”.
2. With the title selected, change the values of x, y, w, and h in the “Layout” tab to match those in the screenshot below.

Graphical user interface, application

Description automatically generated

1. There are three controls users can use to control and interact with the visualizations. First is the Re Admis filter directly below the title. By selecting “Yes” or “No” the data in the two left-most visualizations can be filtered by the patient’s readmission status. A similar filter is present for the remaining visualizations in the top right of the dashboard to select readmission status of patients from the UCI dataset. Finally, the filter above the bottom left visualization allows the user to select how to filter the visualization below. By selecting a different diagnosis in this filter, the below visualization will update to be divided based on the value of that variable per patient.

**Part 2: Storytelling with Data**

1. <https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=5e4d6e58-ed21-46b4-bd2f-adba010cc2b6>

**Part 3: Reflection Paper**

1. Write a reflection paper to demonstrate your understanding of data representation and reporting by doing the following:
2. Our dashboard highlights the characteristics that are the primary cause of patient readmittance in the WGU dataset and compares these statistics with statistics from a related dataset, the UCI dataset. The data dictionary accompanying the WGU dataset describes the need for hospitals to reduce the number of patient readmittances to comply with government regulations at the state and federal level. By highlighting the causes of readmittance, hospitals will be able to make well-informed decisions that will lead to the reduction of patient readmittance.
3. The UCI dataset includes some similar variables to the WGU dataset while also including information on the total number of diagnoses each patient had at time of admittance and the number of emergencies each patient suffered during their admittance. These variables increase the understanding of why patients who spend longer periods in their initial admission and more likely to be readmitted and highlight how hospitals already employ prioritization of patient care, which is a recommendation of our analysis. By further dividing patient readmittance into three categories, to show how soon after discharge a patient was readmitted, the dataset provides a better evaluation of the proportion of patients that are readmitted within 30 days of discharge, allowing for more granular approaches to recommendations of patient prioritization.
4. The distribution of patient admittance length grouped by readmittance, located in the top left of the accompanying dashboard, highlights the trend of patient readmittance based on the patient’s initial length of stay. This infographic helps hospital executives to visualize the greatest factor of patient readmittance, according to the WGU data set, and begin working towards policies that will alleviate the conditions leading to readmittance. In the lower-right corner of the dashboard, a scatterplot of patient length of stay and number of diagnoses with trend lines for each group of patient readmittance is provided. This infographic highlights the relationship between the number of diagnoses a patient has at time of admission and the length of their admittance. This visualization allows hospital executives to understand what patient characteristics increase length of admittance, the primary driver of patient readmittance. This information allows the executives to develop prioritization parameters for patient care that will eventually lead to reduced readmittance rates.
5. The dashboard provides users with two ways to filter the data through interaction. First, there are two panes which allow the grouping parameters of each visualization to be isolated. Above the upper-left distribution are two boxes, one labeled “No” and one labeled “Yes”. By toggling either of these by clicking them, only data points with the specified value in the ReAdmis variable is shown in the graph, updating any average lines as well. Above the visualizations of the UCI data is another set of grouping toggles, this time with three options. These three options correspond to the three categories of readmission in the UCI data set and filter the data from the UCI data set in the same way as the toggles for the WGU data set.

The second tool for interacting with the data corresponds to the bottom right visualization. Above this visualization is a drop-down menu that allows users to control which diagnostic variable is shown in the corresponding visualization. By selecting an option from this menu, the visualization is replaced with a visualization depicting the distribution of the chosen variable and length of patient stay.

1. The color palette chosen for the visualizations is the Tableau color blind palette. Using this palette, and specifically colors from it with high contrasts and various darkness levels, users with color blindness will be able to see and understand the included visuals.
2. The story of our data is one of need. Patients come to the hospital in need of care. Hospital executives need to ensure that patients are not readmitted once discharged. Our dashboard tells this story through the included visualizations. Beginning in the upper left-hand corner, we see what leads to a patient’s readmission, which is reinforced by the center graph. On the right, we see the conditions that lead to increased admission lengths. Together, a story unfolds in which patients with the greatest risk spend the longest time in the hospital and are eventually readmitted.
3. This presentation was tailored for hospital executives, a group of individuals who are scientifically motivated and respond best to logical arguments. Because of this, the presentation was presented as an analysis of facts and the recommendations offered are supported by logical conclusions derived from the data. If the audience had been hospital workers, who perform their duties with motivation from a place of emotional commitment to their patients, the presentation would have been frame differently. In that case, the most effective argument would hinge on the well-being of patients and include emotional appeals to the duty of caregivers and the needs of patients.
4. While the presentation was designed for hospital executives who have high levels of understanding of medical terminology and the inner workings of the medical industry, the presentation remains accessible. The statistics used are simple and organized in a way so that the story of the data tells itself. As viewers move from the left to the right, they see that patients who are admitted for longer periods are readmitted more often, then those patients who have more diagnoses have longer hospital stays. The conclusion that more diagnoses lead to longer stays, which in turn lead to higher chances of readmittance, is told through the visualizations.
5. The first storytelling technique we employ in this presentation is the inclusion of characters. During the presentation, we introduce two characters who are both seeking care from the hospital, and the viewers have to decide which patient is more likely to be readmitted. This decision-making process on the part of the audience is another storytelling technique designed to engage the audience: interactivity. While the presentation was filmed without an audience, under ideal conditions, the audience would answer the questions asked in the presentation, causing them to think about the gravity of the decision when prioritizing one patient over another. Additionally, the audience’s inquiry could deepen the characters in the story and only increase the emotional impact of the decision. These characteristics, especially when combined, deepen engagement by giving the audience a sense of personal stake in the decisions and bringing the theoretical story told by the data to life.
6. Acknowledge sources, using in-text citations and references, for content that is quoted, paraphrased, or summarized.

Beata Strack, Jonathan P. DeShazo, Chris Gennings, Juan L. Olmo, Sebastian Ventura, Krzysztof J. Cios, and John N. Clore, “Impact of HbA1c Measurement on Hospital Readmission Rates: Analysis of 70,000 Clinical Database Patient Records,” BioMed Research International, vol. 2014, Article ID 781670, 11 pages, 2014.