**D210 Performance Assessment**

**DATA DASHBOARD AND STORYTELLING**

Fahim A. Akbar Student ID 001434895 Masters Data Analytics (January 1, 2021) Program Mentor: Lea Yoakem (877) 435-7948x6422 [fakbar3@wgu.edu](mailto:fakbar3@wgu.edu)

**Part 3: Reflection Paper**

One of the most common chronic health conditions is diabetes. According to the CDC, over 37.3 million Americans have diabetes. (CDC, 2022). By understanding which groups of people are most affected by or most likely to have diabetes, we can better plan for preventative measures that can target the most at-risk groups.

**1. Explain how the purpose and function of your dashboard align with the needs outlined in the data dictionary associated with your chosen data set.**

For this investigation, I chose to focus on diabetes as a point of interest medical condition. The purpose and function of the executive dashboard is to showcase which demographics are most at-risk for diabetes based on the existing patient records. This will help the hospital staff and stakeholders better predict and anticipate which patient groups are most likely at risk for diabetes. From there, they can plan interventions and preventative treatments for target groups.

**2. Explain how the variables in the additional data set enhance the insights that can be drawn from the data set you chose from the provided options.**

The additional data set helps enhance the insights by providing an additional line of comparison for the variables we represented here. The additional data set, titled 1diabetic\_data, is comprised of data collected from 130 US hospitals from the years 1999-2008. (Brandão, 2017). The dataset was obtained from the Center for Machine Learning and Intelligent Systems at University of California, Irvine. (UCI, 2014). The 1diabetic\_data set for DataSet 2 had a much larger sample size of over 100,000 entries as opposed to the 10,000 entries of medical\_clean for DataSet 1. This helped give us additional context to which gender and age groups seemed to be most affected by diabetes. It also showed us that the older age groups were more likely to have diabetic patients than younger groups. The income and geographic data visualizations from DataSet 1 showed us that lower income brackets were more affected by diabetes, and that the disease is more present in states with higher populations.

**3. Explain two different data representations from your dashboard and how executive leaders can use them to support decision-making.**

The two different data representations show us which groups of patients appear to have diabetes. Both tables for Readmission rates show us that the majority of patients with diabetes were female. The tables for DataSet 1 show an almost even split between males and females with diabetes who were readmitted to the hospital. DataSet 2 shows a much higher number of female patients with diabetes being readmitted to the hospital as opposed to male or non-specific genders. Both bar charts for the age groups showed that the older age groups tend to have higher rates of diabetes as opposed to the younger age groups. Hospital executives and stakeholders can use these to help with decision making in regards to how they can lower readmission rates of patients with diabetes. Seeing which groups are the highest at risk could hospitals plan for better initial treatments. They can also use this data to predict future patient trends and plan preventative treatment and educational measures ahead of time. Research by the CDC shows that nearly 1 in 5 Americans with diabetes don’t even know that they have it. (CDC, 2022). By creating targeted educational campaigns, they can further reduce potential readmissions since at risk patients would be better informed and can be treated early or begin preventative measures.

**4. Explain two interactive controls in your dashboard and how each enables the user to modify the presentation of the data.**

One of my dashboard’s interactive controls is the range of values filters for the genders of the patients. This allows user to select or deselect any specific gender for each data set and compare them to one another, such as male or female. The users also have the ability to select and deselect if a patient has diabetes for each respective data sheet. Furthermore, the diabetes filter for DataSet 1 also applies to the income and map visualizations. This can help users see which income levels are most likely to be affected by diabetes. The map visualization allows for users to choose a specific state and see how many people have diabetes for the selected state.

**5. Describe how you built your dashboard to be accessible for individuals with colorblindness.**

To make sure that the dashboard was accessible for individuals with colorblindness, I made sure to use colors from the Tableau Color Blind 10 palette. I also avoided using reds and greens together since those can be problematic for color blind people to distinguish. For my readmission’s tables, those did not require color visualizations. For my age and income graphs, I used dark blue and orange colors form the Color Blind 10 palette since those colors are a good contrast and can be distinguished by color blind people. (Shaffer 2016). For the map graphic, it needed a range of colors for the range of values and did not have a colorblind palette. To compensate for that, I chose a palette of blues that had contrasting dark and light colors. The article “5 Tips on Designing Colorblind-Friendly Visualizations” by Jeffery Shaffer was especially helpful in guiding my decision making for the visualizations to be accessible to colorblind viewers.

**6. Explain how two data representations in your presentation support the story you wanted to tell.**

The readmission rate tables and the bar charts both show that the majority of patients do not have diabetes. The data representations show us that majority of patients who do have diabetes were female. This could be partially due to women also experiencing gestational diabetes during pregnancy along with type 1 and type 2 diabetes. (CDC, 2021). Furthermore, the age visualizations for both datasets show that the majority of patients with diabetes were older of age, between the age groups of 40s-80s. For DataSet 1 the majority of patients with diabetes were in the 50s age group whereas in DataSet 2the majority of patients with diabetes were in the 70s age group. Furthermore, the income graphic for DataSet 1 showed that the majority of patients with diabetes were in the lower income brackets, with people in $10,000 salary range having the highest likelihood. This could be due to people in poverty having a harder time accessing healthier food options or being able to afford proper treatment. Lastly, the interactive map showed that the state with the most amount of diabetic patients was Pennsylvania, followed by New York, California, and Texas. These states are highly populated so it would make sense that they would most likely have higher rates of diabetes as opposed to states with lower populations.

**7. Explain how you used audience analysis to adapt the message in your presentation.**

The target audience for the presentation is hospital staff, hospital owners, and their stakeholders. I chose to base my story on diabetes because this is a common occurrence throughout the country and the people viewing my presentation have a high chance of knowing someone in their lives that are affected. For example, in case a viewer of my presentation is female and of the higher age brackets more susceptible to diabetes, she may be inspired to get a checkup. I also chose to feature readmission as one of the categories because hospitals have a vested interest in reducing readmission rates. Hospitals can be penalized if their readmission rates are too high, so showing the rates of readmission amongst diabetic patients can help them create new protocols with the aim of reducing the readmittance of these patients.

**8. Describe how you designed your presentation for universal access by all audiences.**

My presentation was intentionally designed to be universally accessible by all audiences and not just people in the medical field. The Tableau dashboard was made public and accessible for anyone with internet access to view for free. The dashboard was designed to be simple to understand and does not rely on medical terminology. The colors for the visualizations were chosen so that even people with color blindness would be able to understand and distinguish the visuals.

**9. Explain two elements of effective storytelling that you implemented in your presentation and how each element was intended to engage the audience.**

To better tell my story with the data I first engaged with the audience by introducing myself and setting a narrative they could relate to. Good story telling begins with capturing your audiences attention, and this can be achieved by peaking their interests and invoking an emotion reaction. (Knaflic 2015) By bringing up how 1 in 10 Americans have diabetes, as well as mentioning my own personal experience with diabetes affecting my relative, I hoped to invoke an emotional response for my audience and capture their attention right before presenting. Everybody is familiar with diabetes and many people have a loved one or someone they know with this condition. I also bring up how 1 in 5 Americans who have diabetes don’t even realize that they do. While my intent isn’t to scare or alarm people, perhaps individuals who view my presentation and realize that they fall within the most likely groups of people to diabetes can schedule doctor visits and determine their risk factors as well. Showing how common the disease helps to further contextualize to the audience the need to understand who is most likely impacted. For my dashboard presentation, I decided to go with the “Factors” data story type and focused on showing what categories appeared to be the most associated with diabetes (Tableau 2022). The results showed us that women and older age groups were the most likely to have diabetes. Showcasing the interactive charts and map of the dashboard helped better contextualize these factors by providing adaptive visuals for the audience. They can specifically focus on female and older age groups using the filters and see for themselves how more of the patients within those categories had diabetes. I chose not to rely on color to convey my message since I wanted this presentation to be accessible to people with color blindness. Instead, I made sure to use high contrasting colors for my visualizations. By creating an engaging introduction and using visualize to better contextualize how women and older age groups were more affected by diabetes, I was able to tell the story I wanted with the data presented.

**References**

Brandão, H. (2017) *Diabetes 130 US hospitals for years 1999-2008*, *Kaggle*. Available at: <https://www.kaggle.com/datasets/brandao/diabetes>

CDC (2021) *Gestational diabetes, Centers for Disease Control and Prevention.* Centers for Disease Control and Prevention. Available at: <https://www.cdc.gov/diabetes/basics/gestational.html>

CDC (2022) The facts, stats, and impacts of diabetes, Centers for Disease Control and Prevention. Centers for Disease Control and Prevention. Available at: <https://www.cdc.gov/diabetes/library/spotlights/diabetes-facts-stats.html#:~:text=37.3%20million%20Americans%E2%80%94about%201,t%20know%20they%20have%20it>.

CMS (2022). *Hospital readmissions reduction program (HRRP).*Centers for Medicare & Medicaid Services. Available at: <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Readmissions-Reduction-Program>

Jeffrey Shaffer COO and VP of Information Technology and Analytics (2016) *5 tips on designing colorblind-friendly visualizations*, *Tableau*. Available at: <https://www.tableau.com/blog/examining-data-viz-rules-dont-use-red-green-together>.

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals.* Wiley . (Chapter 1, Chapters 3-5, Chapters 7 - 8) <https://ebookcentral.proquest.com/lib/westerngovernors-ebooks/reader.action?docID=4187267&ppg=1>

Tableau. *Best practices for telling great stories* Available at: <https://help.tableau.com/current/pro/desktop/en-us/story_best_practices.htm>

UCI Machine Learning Repository (2014). *Diabetes 130-US hospitals for years 1999-2008 data set.* Available at: <https://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for+years+1999-2008>