Saemi Ramirez

011926418

D210 Representation and Reporting

8/14/2024

WGU

1. Interactive Data Dashboard
   1. Provided 5 dataset. The external dataset is demographic data and questionnaire data from NHANES (National Health and Nutrition Examination Survey) 2017 – 2018 (<https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017>)
      1. final - combined with WGU and external dataset (final\_combined.csv)
      2. raw demographic dataset (demographic.csv)
      3. raw questionnaire dataset (questionnaire.csv)
      4. cleaned external dataset (final\_ext\_data.csv)
      5. cleaned WGU dataset (final\_medical\_clean.csv)
   2. Dashboard installation is not needed for the clients. Simply click [here](https://public.tableau.com/app/profile/saemi.ramirez/viz/WGUD210DataDashboardStorytelling/PatientsDemographic) to the dashboard.
   3. Upon clicking the provided link, the client will be directed to a dashboard displaying a comparative analysis of patients’ demographics between WGU and NHANES. The first two charts illustrate the comparison between these two sources in terms of age (5-year increment) and marital status. The pie chart on the bottom left represents the gender distribution. Three interactive controls (gender, age, and marital status) are available for clients to adjust as needed.  
      In the next tab, titled ‘Illness Tracker’, clients can view three sets of charts detailing various illness types, with comparisons drawn between WGU and NHANES data on arthritis, asthma, diabetes, high blood pressure, and stroke. The far left-hand section chart name starts with “Ratio of …” displays the percentage of patients diagnosed with a particular condition in comparison to WGU and NHANES (ex. Stroke patients in WGU / (stroke patients in WGU & NHANES) \* 100). The second chart starts with “(illness) in Total” shows the percentile of the diagnosed patients in total (ex. Stroke patients in WGU or NHANES / all patients of WGU and NHANES). The central graphs depict the number of diagnosed patients within each age group, categorized in 5-year increments. An age and source controls are also available, allowing clients to modify the graphs accordingly.
2. Storytelling with Data

Panopto Link: <https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=62f67947-385f-40d5-99ce-b1cc0154b6f9>

1. Reflection Paper
   1. The purpose of the demographic dashboard is identifying the trends by revealing patterns in the distribution of patients across different age groups, marital statuses, and genders, which may be relevant or understanding the population’s health dynamics.   
      For the illness tracker dashboard, it allows for a direct comparison of the prevalence of these chronic conditions (arthritis, asthma, diabetes, high blood pressure, stroke) conditions between WGU and NHANES. This can help in identifying any significant disparities in health outcomes.
   2. The NHANES dataset contained several variables that aligned with those in the WGU medical dataset. Specifically, from the ‘questionnaire.csv’ file, the variables BPD035 (high blood pressure), MCQ025 (asthma), MCQ160F (stroke), MCQ160A (arthritis), and DID040 (diabetes) were utilized. These variables correspond to the question, “At what age were the patient diagnosed with (illness)?” For analysis, the responses were categorized as “Yes” if a value was present and “No” if the field was null.   
      Additionally, three demographic variables for the ‘demographic.csv’ file were included: RIDAGEYR for age, RIAGENDR for gender, and DMDMARTL for marital status. The marital status categories were coded as follows: 1 for married, 2 for widowed, 3 for divorced, 4 for separated, 5 for never married, 6 for living with a partner, and a period (.) for never married because the ages for this category were under 18.
   3. Both data representations I would like to discuss are under ‘Illness Tracker’. First, the highlighted tables were used to show the number of patients diagnosed to either arthritis, asthma, diabetes, high blood pressure, or stroke. The age group was divided into 5-year, and darker the colors for higher numbers. The second illustration is the text table, which simply shows the percentage of the patients diagnosed into particular illness out of total number of patients in both WGU and NHANES.   
      With these charts, the executive leaders can allocate resources more effectively by identifying age groups with higher diagnosis rates such as increasing staffing, specialized equipment, or targeted outreach in those area. Also, if a certain age groups are more affected, leaders can develop and implement specific intervention programs tailored to the needs of those populations, potentially reducing the incidence or severity of the illness. Lastly, the executive leaders can use this information to create or adjust of healthcare policies. For example, if younger age groups are disproportionately affected, policies could focus on preventive measures like education and early screening.
   4. The dashboards offer three primary interactive controls that clients can adjust: the age, marital status, and gender of the patients. Modifying these controls will dynamically update all of the graphs and charts in both patients’ demographic and illness tracker dashboards.
   5. For each graph and chart, I ensured that all colors were chosen from the colorblind-friendly palette to accommodate individuals with color vision deficiencies. Since red and green are particularly challenging for most people with color blindness, I selected a palette that includes orange, gray, blue, and various shades of blue to enhance the accessibility and clarity of the visualization.
   6. I have been focusing on patients diagnosed with high blood pressure. The high blood pressure-related charts under the ‘Illness Tracker’ provide a strong comparison of patient numbers across different age backets between WGU and NHANES. An interesting observation from this chart is that high blood pressure patients in the WGU dataset are fairly evenly distributed across ages 20 to 85. In contrast, the NHANES dataset shows a gradual increase in high blood pressure cases starting from age 15, peaking in the 60-64 age groups.   
      Additionally, the ‘WGU vs. NHANES in Age’ graph under ‘Patients’ Demographics’ reveals that the NHANES dataset includes data on minor (ages 0 to 14), whereas the WGU dataset only covers young adults beginning at age 15. This broader age range in the NHANES dataset provides more comprehensive insights into the distribution of asthma, arthritis, diabetes, high blood pressure, and stroke across different age groups.
   7. This presentation is for the executive leadership to support their decision-making. The dashboard should be used to highlight key trends and patterns in 5 illnesses across different age groups. The presentation needs to be focused on summarizing the most significant findings such as which age groups are most at risk and how these trends compare across the two datasets. This visualization presentation also can help the executive leaderships to align with the strategic goals and priorities of the organization such as focusing on improving patient outcomes, identify the most risk age groups on certain illness, and specific preventions.
   8. The dashboards on Tableau Public are designed for easy access by all clients without the need to install any software on their computers or laptops. Clients can access these dashboards via the internet on any mobile device, such as a phone, tablet, or laptop. This ensures a seamless experience without technical issues related to software compatibility. To accommodate individuals with color blindness, the dashboards do not use red or green; instead, alternative colors have been selected. Additionally, the Panopto presentation includes both a narrative and audio components, catering to those with hearing impairments as well as providing an accessible option for individuals with visual impairments.
   9. According to Pierce, data storytelling consists of three key components: the story, the data, and the visuals. I utilized data provided by WGU, along with external data from NHANES, to compare patient demographics (age, gender, and marital status) and the distribution of illness, including arthritis, asthma, diabetes, high blood pressure, and stroke.  
      The dashboards were designed to be simple and intuitive, ensuring that clients can easily navigate, understand and interact with the information. Specifically, the two charts in the ‘Illness Tracker’ illustrate the percentage of patients diagnosed with each illness relative to both the total diagnosed patients and the overall recorded population in the dataset. These insights will assist executive leadership in comparing WGU patients with those in the NHANES dataset, which examines the nutritional and health status of both adults and children.
2. Acknowledge Sources
   1. *Change the order of a Pandas DataFrame columns in Python*. GeeksforGeeks. (December 5, 2023). <https://www.geeksforgeeks.org/change-the-order-of-a-pandas-dataframe-columns-in-python/>.
   2. Pierce, Pamela. *9 Data Storytelling Tips for More Effective Presentations An Expert Guide*. Oracle NetSuite. (June 8, 2023). <https://www.netsuite.com/portal/resource/articles/data-warehouse/data-storytelling-tips.shtml>.