**D211: Advanced Data Acquisition**

**Performance Assessment**

Logan Rosemeyer

Western Governor's University

D211: Advanced Data Acquisition

Dr. David Gagner

February 13, 2023

# Contents

# A: Data Dashboards

My data dashboards are attached in the submission.

# A1: Both Data Sets

Both of my data sources are attached in the submission.

# A2: Dashboard Installation

### Smoke Everyday Worksheet

1. Drag State1 onto the worksheet.
2. Click show me and select maps.
3. Drag smoke everyday to color on marks.
4. Click color, edit colors, and select Red-Green Diverging, click reversed and click ok.
5. Drag State1 to filters and deselect District of Columbia, Guam, Puerto Rico, and the two national values.

### Readmission Worksheet

1. Drag State onto the worksheet.
2. Click show me and select maps.
3. Drag Readmis to rows.
4. Right click on No and select hide.
5. Drag Readmis to color.
6. Right click on Readmis, hover over measure, and select count
7. Again right click on Readmis, click on add table calculation, select percent of total, compute using specific dimensions, and select only readmis.
8. Click color, edit colors, and select Red-Green Diverging, click reversed and click ok.
9. Right click on the legend and select edit colors. Click advanced, click center, and enter .367, which is the mean.

### Non-Smoking Rank Worksheet

1. Click analysis and select create calculated field. Change title to Don't Smoke. The calculation is [Never smoked] + [Former smoker].
2. Drag Don't smoke to columns and State1 to rows.
3. Click analysis and select create calculated field. Change title to Non-Smoking Rank. The calculation is RANK(AVG([Don't Smoke]), 'desc').
4. Drag State1 to filters and deselect District of Columbia, Guam, Puerto Rico, and the two national values.
5. Drag Non-Smoking Rank to filters, then right click and click show filter.
6. Right click on State1 in rows and select sort. Sort by nested, select descending, field name is Don't Smoke, aggregation is sum.
7. Drag Don't Smoke to label. Right click don't smoke and select format. Select pane, under numbers select number (custom), change decimal places to 1 and make % the suffix.
8. Right click the x-axis and select edit axis. Deselect include zero.

9. Click on the analytics tab and add a reference line to the table. Make the line orange to stand out from the bars.
10. Right click the line in the worksheet and select format. Make the opacity of the line 30% and change the font color of the line label to the same orange as the line.

## Highest Readmission Worksheet

1. Click analysis and select create calculated field. Change title to Readmin %. The calculation is COUNT([Readmis]) / TOTAL(COUNT([Readmis])) * 100.
2. Click analysis and select create calculated field. Change title to Readmin % Rank. The calculation is RANK([Readmin %], 'desc').
3. Drag Readmis to color. In the legend, right click No and select hide.
4. Click on the x-axis and select sort descending.
5. Drag Readmin % to label, then right click and select compute using readmis. After that, right click and select format. Select pane, under numbers select number (custom), change decimal places to 1 and make % the suffix. Change alignment to center align.
6. Drag State to filters and deselect PR and DC.
7. Drag Readmin % Rank to filters, then right click on it and select show filter.
8. Right click on the color legend and select hide card.
9. Drag Don't Smoke to label. Right click don't smoke and select format. Select pane, under numbers select number (custom), change decimal places to 1 and make % the suffix.

## Smoking vs. Readmission Dashboard

1. Drag the smoke everyday worksheet to the top left, non-smoking rank to top right, readmission to bottom left, highest readmission to bottom right.
2. Drag the two rank sliders to beneath the chart that they go with. Change the size of the slider pane to as small as you can while still being able to use the slider.
3. Make the color legends floating.
4. Delete the empty pane on the right side of the dashboard.
5. Drag the middle partition to the right to give the maps more room than the bar charts.
6. Drag the color legends over the maps and position in white space where they will be easy to see.

# A3: Dashboard Navigation

The top left visualization shows the percent of people in each state who smoke every day with red being higher and green being lower. The bottom left visualization shows the readmission rate by state with red having a higher likelihood of readmission and green being lower. The top right visualization shows the rank of states who have the highest percent of non-smokers (never smoked and former smokers). This also has a slider to adjust which rankings you want to see. The bottom right visualization shows the rank of states with the highest readmission. This also has a slider to adjust which rankings you want to see.

# A4: SQL Code

SELECT CAST("patient"."readmis" AS TEXT) AS "readmis",

CAST("location"."state" AS TEXT) AS "state"

```sql
FROM "public"."patient" "patient"
  INNER JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")
GROUP BY 1,
  2


SELECT COUNT(CAST("patient"."readmis" AS TEXT)) AS "cnt:readmis:ok",
  CAST("patient"."readmis" AS TEXT) AS "readmis",
  CAST("location"."state" AS TEXT) AS "state"
FROM "public"."patient" "patient"
  LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")
GROUP BY 2,
  3


SELECT COUNT(CAST("patient"."readmis" AS TEXT)) AS "cnt:readmis:ok",
  CAST("location"."state" AS TEXT) AS "state"
FROM "public"."patient" "patient"
  LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")
GROUP BY 2


SELECT CAST("patient"."readmis" AS TEXT) AS "readmis",
  CAST("location"."state" AS TEXT) AS "state (states.csv)",
  CAST("location"."state" AS TEXT) AS "state"
FROM "public"."patient" "patient"
  INNER JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")
WHERE (((CAST("location"."state" AS TEXT) >= 'AK') AND (CAST("location"."state" AS TEXT) <= 'CT')) OR
((CAST("location"."state" AS TEXT) >= 'DE') AND (CAST("location"."state" AS TEXT) <= 'PA')) OR
((CAST("location"."state" AS TEXT) >= 'RI') AND (CAST("location"."state" AS TEXT) <= 'WY')))


SELECT "patient"."location_id" AS "location_id",
```

```sql
  CAST("location"."state" AS TEXT) AS "state"
FROM "public"."patient" "patient"
 LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")


 SELECT COUNT(CAST("patient"."readmis" AS TEXT)) AS "TEMP(TC_)(4184331154)(0)",
  CAST("location"."state" AS TEXT) AS "state"
FROM "public"."patient" "patient"
 LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")
 WHERE (((CAST("location"."state" AS TEXT) >= 'AK') AND (CAST("location"."state" AS TEXT) <= 'CT')) OR
((CAST("location"."state" AS TEXT) >= 'DE') AND (CAST("location"."state" AS TEXT) <= 'PA')) OR
((CAST("location"."state" AS TEXT) >= 'RI') AND (CAST("location"."state" AS TEXT) <= 'WY')))
 GROUP BY 2


 SELECT COUNT(CAST("patient"."readmis" AS TEXT)) AS "cnt:readmis:ok",
  CAST("patient"."readmis" AS TEXT) AS "readmis",
  CAST("location"."state" AS TEXT) AS "state"
FROM "public"."patient" "patient"
 LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")
GROUP BY 2,
 3


 SELECT CAST("patient"."readmis" AS TEXT) AS "readmis",
  CAST("location"."state" AS TEXT) AS "state"
FROM "public"."patient" "patient"
 INNER JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")
GROUP BY 1,
 2


SELECT CAST("patient"."readmis" AS TEXT) AS "readmis",
```

```sql
    CAST("location"."state" AS TEXT) AS "state (states.csv)",

    CAST("location"."state" AS TEXT) AS "state"

  FROM "public"."patient" "patient"

    INNER JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")

  WHERE (((CAST("location"."state" AS TEXT) >= 'AK') AND (CAST("location"."state" AS TEXT) <= 'CT')) OR
((CAST("location"."state" AS TEXT) >= 'DE') AND (CAST("location"."state" AS TEXT) <= 'PA')) OR
((CAST("location"."state" AS TEXT) >= 'RI') AND (CAST("location"."state" AS TEXT) <= 'WY')))


  SELECT "patient"."location_id" AS "location_id",

    CAST("location"."state" AS TEXT) AS "state"

  FROM "public"."patient" "patient"

    LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")


  SELECT CAST("location"."state" AS TEXT) AS "state"

  FROM "public"."location" "location"

  GROUP BY 1


  SELECT COUNT(CAST("patient"."readmis" AS TEXT)) AS "cnt:readmis:ok",

    CAST("location"."state" AS TEXT) AS "state"

  FROM "public"."patient" "patient"

    LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")

  GROUP BY 2


  SELECT COUNT(CAST("patient"."readmis" AS TEXT)) AS "TEMP(TC_)(4184331154)(0)",

    CAST("location"."state" AS TEXT) AS "state"

  FROM "public"."patient" "patient"

    LEFT JOIN "public"."location" "location" ON ("patient"."location_id" = "location"."location_id")

  WHERE (((CAST("location"."state" AS TEXT) >= 'AK') AND (CAST("location"."state" AS TEXT) <= 'CT')) OR
((CAST("location"."state" AS TEXT) >= 'DE') AND (CAST("location"."state" AS TEXT) <= 'PA')) OR
((CAST("location"."state" AS TEXT) >= 'RI') AND (CAST("location"."state" AS TEXT) <= 'WY')))
```

## B: Panopto Presentation

My Panopto presentation is included in my submission and can also be found using the link below.

https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=6f5ede9e-8e48-4f8a-8396-afb000dc5a6d

## C1: Dashboard Alignment

The needs of the company are to reduce the readmission of patients to the hospital. My dashboard aligns with that because I have two visualizations that show readmission percent of every state, as well as the states with the highest readmission percents.

## C2: Business Intelligence Tool

I decided to use Tableau because it provides quick calculations within the visualizations. I also chose Tableau because it allowed me to easily join my tables together so that I could quickly connect my external data set to my given data set. (Mitisha_A, 2021)

## C3: Data Cleaning

In order to make sure the data was ready to be analyzed, I had to know that each patient had an associated smoke everyday, smoke some days, former smoker, and never smoked value. There were no nulls in any of these 4 columns, which means that each patient did in fact have a value for all 4 variables.

## C4: Dashboard Creation

I included four sheets in my dashboard. The first was a map of states and the percent of people in each state that smoke. This is in the top left of my dashboard. Next is a map of states with their readmission percents. This is in the bottom left of the dashboard. Then, I made a bar chart of the top non-smoking states. This takes the sum of the percent of people who have never smoked and former smokers. This is in the top right of the dashboard. Finally, I made a bar chart showing states with the highest readmission percents. This is in the bottom right of the dashboard.

## C5: Data Analysis Results

I believe the results of my analysis are inconclusive. The two maps don't really match up besides a few states, but there aren't nearly enough to say that states with a higher percent of smokers also have a higher readmission rate. Similarly, there is only one state in the top 10 of highest percent of smokers and top 10 of highest readmission rate. If our organization truly wants to determine if smoking can lead to readmission, we need to see whether our patients are smokers as well as whether they readmitted.

## C6: Analysis Limitations

The major limitation with my analysis is that I do not know whether each patient is a smoker or not. Without knowing whether each patient is a smoker, it is extremely hard to really tell whether smoking is a factor in whether a patient is readmitted or not.

## D: Web Sources

Mitisha_A. *4 Reasons Why Tableau Is The Best Tool For Data Visualization.* Zehntech. July 31, 2021. Retrieved February 20, 2023. https://www.zehntech.com/4-reasons-why-tableau-is-the-best-tool-for-data-visualization/

## E: Sources

Ganiyu, Isola Saheed. *PgAdmin Import CSV: How to Import CSV Into PostreSQL?.* Hev0. February 15, 2022. Retrieved February 18, 2023. https://hevodata.com/learn/postgresql-import-csv/

*Viewing Underlying SQL Queries*. Tableau. October 22, 2013. Retrieved February 16, 2023. https://kb.tableau.com/articles/howto/viewing-underlying-sql-queries-desktop