

Lab 4

Exploring Tableau and Applying Tableau Concepts



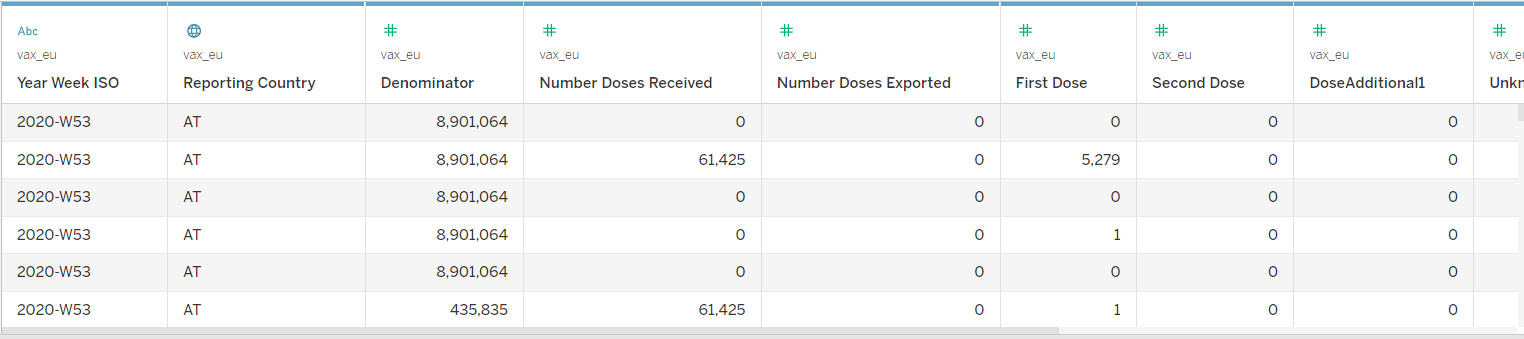
January 26, 2022

William jackson, julie nguyen, matt wehnert

CSI 475

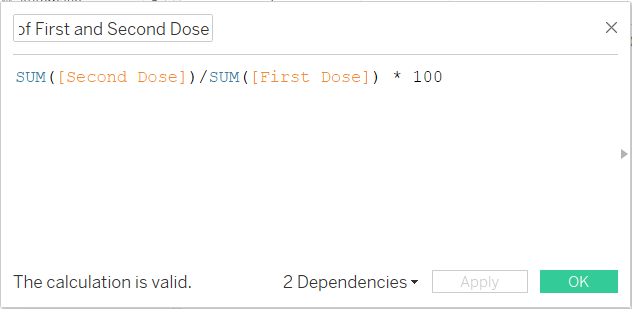
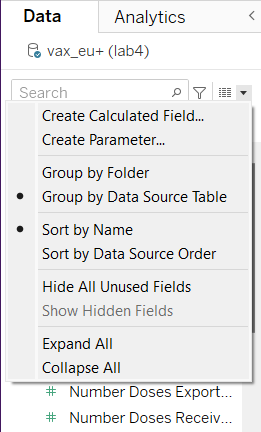
Importing to Tableau Procedure

1. Boot up MySQL and login. Boot up Tableau next and click on MySQL at the start page of Tableau. Login with your localhost server and login details.
2. Click on databases, select your correct database schema (lab3, but due to issues lab4 was made to mitigate the errors) and then select your datasets.
3. Drag the two tables and connect them using the fields: DoseAdditional1 & Additional Doses Unk Manuf. Then click update now. You should see a table:

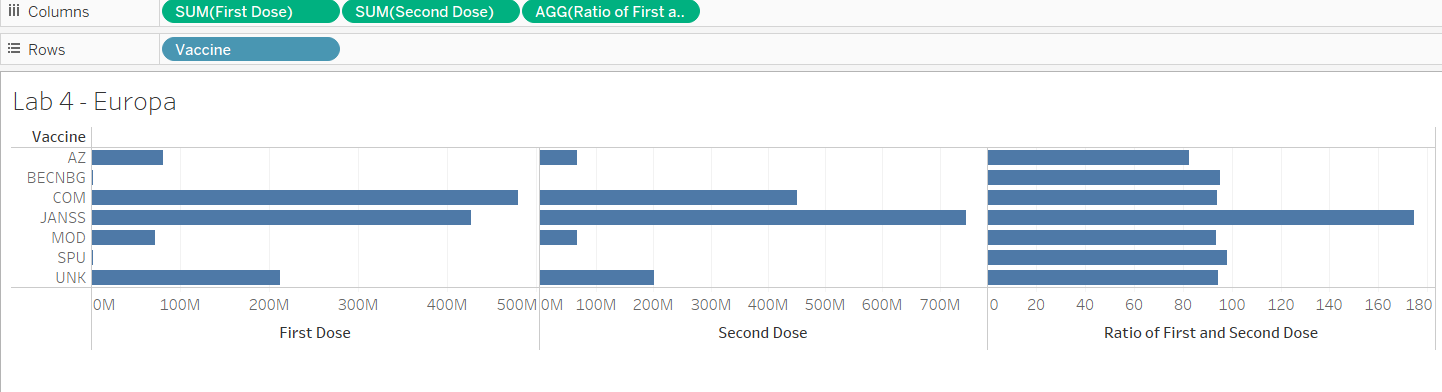


Statistics in Tableau Procedure

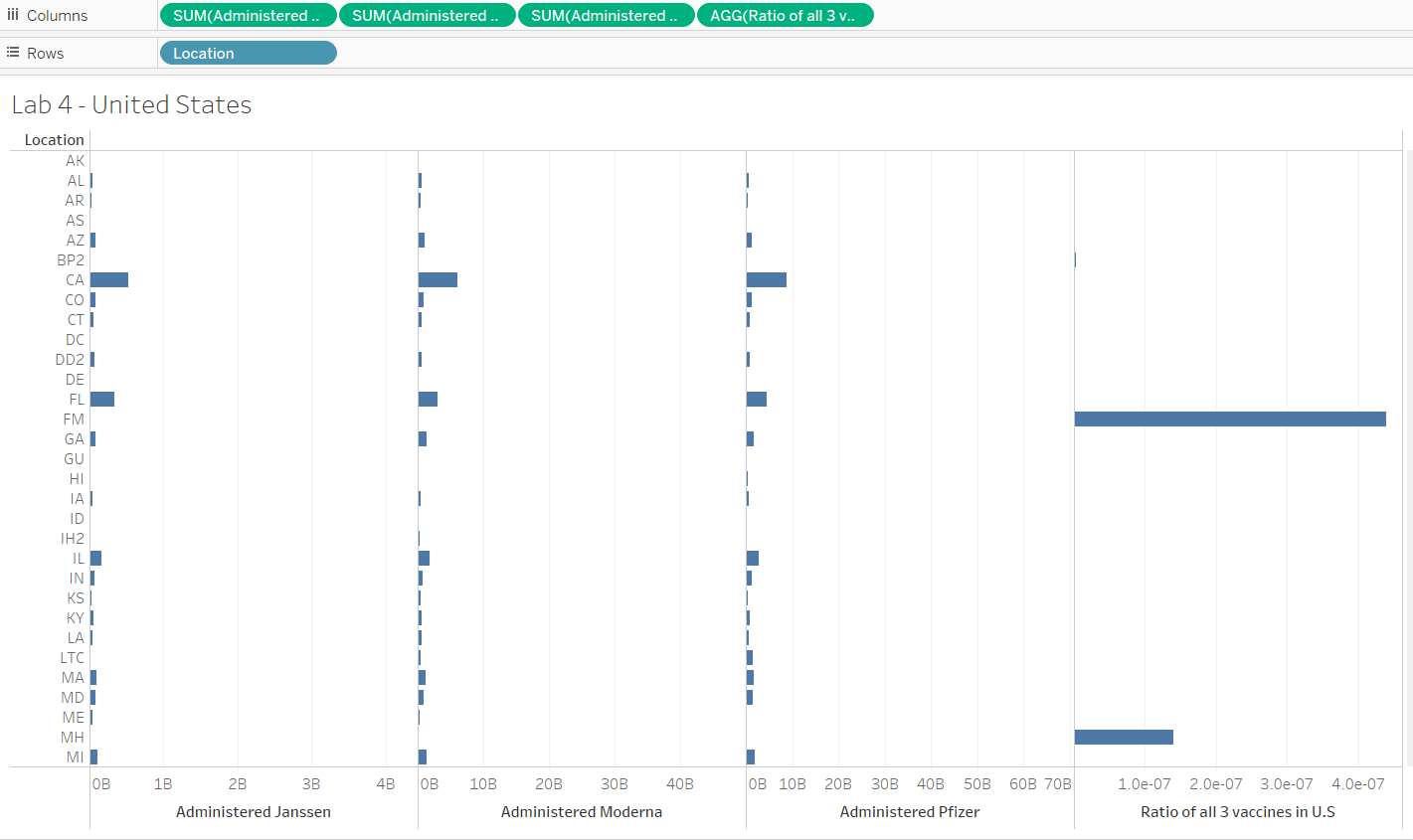
1. Go to sheet 1 on the lower left corner.
2. In sheet 1, drag and drop First Dose and Second Dose as the columns from the Europa dataset. Then for the row, drop Vaccine.
3. Click on the little arrow to display a list of options. From the options, click on “Create Calculated Field”. A special aggregated function is created by typing into the field of the pop-up window. Call this aggregate calculation “Ratio of First and Second Dose”. Please see the following figures below:



1. Click OK once the expression has been created, the \*100 will give a percentage.
2. Repeat step 6 and step 7 to make another aggregate calculation called “Ratio of all 3 vaccines in the United States” with the following calculation: SUM([Administered Janssen])/SUM([Administered Moderna])/SUM([Administered Pfizer]). This calculation will be used later in sheet 2.
3. Going back to sheet 1, put the “Ratio of First and Second Dose” into the columns alongside First and Second Dose. See figure below:

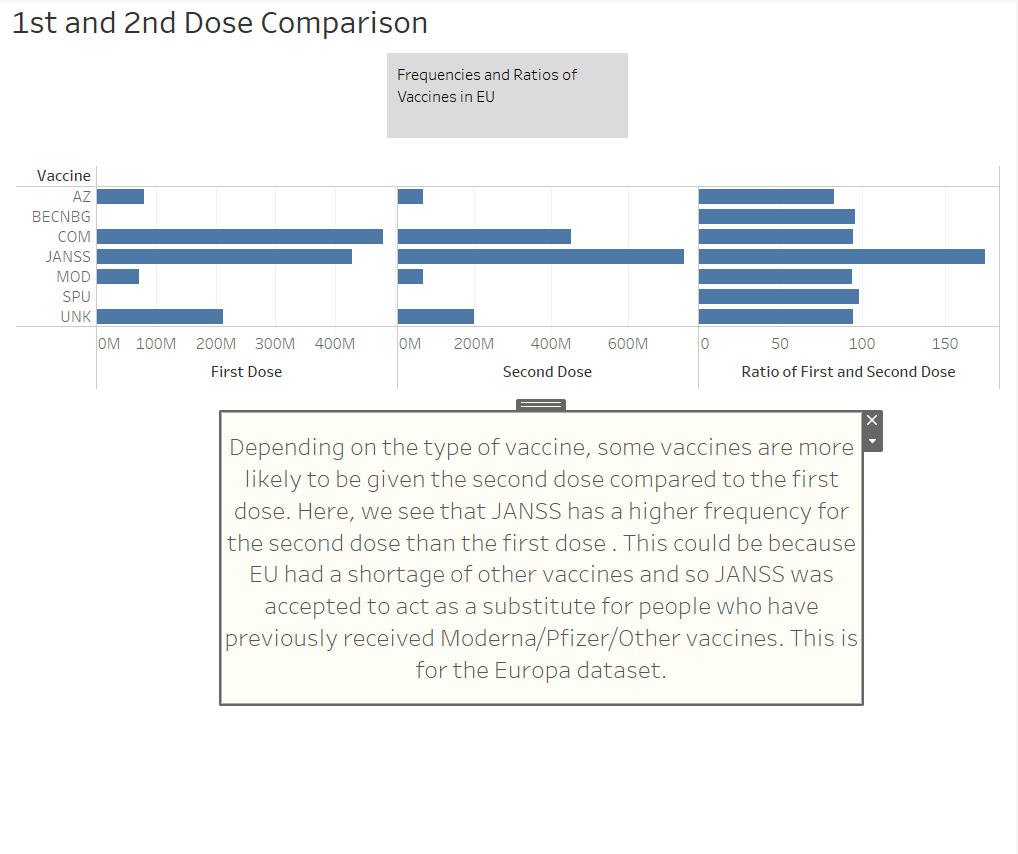


1. Make another sheet, this will be sheet 2. The following columns used are: Administered Janssen, Administered Moderna, and Administered Pfizer. The row will only have the location. Put the aggregate calculation that was made in step 8 into the column. See figure below:

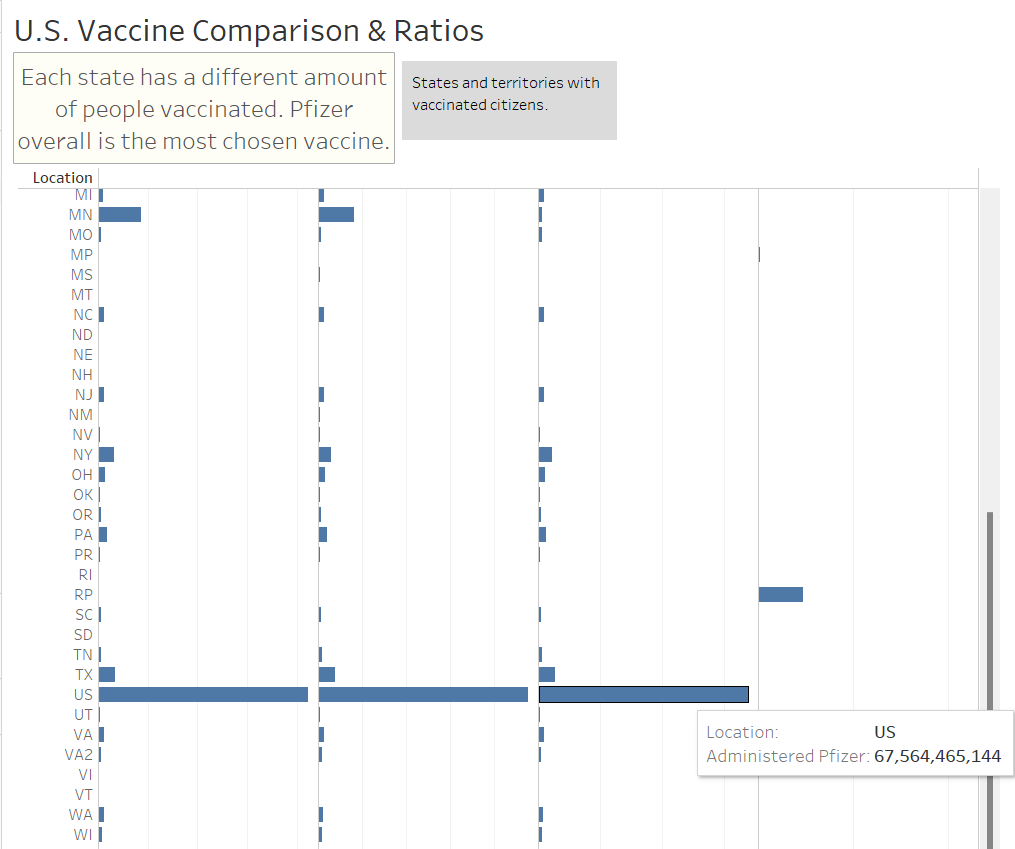


Stories in Tableau Procedure

1. On the toolbar at the top, click Story and “New Story”. This new story is called “Story 1”. From this page you can edit the title, caption, and add a textbox to provide context to the data.
2. Click Story 1 and edit the title to “1st and 2nd Dose Comparison”. Then drag and drop Sheet 1 into the dashed box that says “Drag a sheet here”. Here the data is given in the story.
3. Click on the gray box that says “Add a caption”. In this example we provide a general statement about the data such as “Frequencies and Ratios of Vaccines in EU”. Click ok once your statement is finished.
4. Drag and drop the textbox from the left panel that says “Drag to add text” then move the textbox to the center of the story workspace. Resize and readjust as needed depending on the description See figure below:



1. Repeat steps 11-15 to complete story 2 under a different title as “U.S. Vaccine Comparison & Ratios”, with a different caption, and a different textbox, any readjustments and resizing can be applied anywhere as it can be styled in various ways. see figure below:



1. Save your work as lab4.twb

**Conclusion**

My group and I chose use SUM aggregate function in the calculated field and finding the ratio as a way of getting familiar with one of the techniques. However, we have tried using DISTINCT but found SUM to be easier to interpret as it is the default. We also found out that the second dose of JANSS vaccine is more commonly taken in the EU. During the time of the vaccine shortage in the EU, JANSS was authorized as second dose for the people who have previously taken Moderna/Pfizer/Other vaccines. We believe this could be why we see an increase over 100 percent for a second dose of a one dose vaccine. Another interesting information is that Pfizer vaccine is more commonly taken than the other vaccines in the United States according to the data.

# **Challenges**

1. The queries would take hours to execute because of the mixing and matching of different column fields from both datasets.
2. Had to reclean the USA dataset because of the commas that interfered with calculations.
3. Cross referencing different columns resulted in incorrect formatting of graphs due to different formats from the datasets that were not compatible.

# **Insights & Resolution**

1. Lesson learned: to avoid waiting for a long time for the queue to execute, do not mix and match column fields from two different datasets.
2. By getting rid of the commas in the USA dataset and setting the numeric format, there were no more compatibility issues with the interpretation for Tableau.
3. We were able to look at the data and interpret it clearly by making sure the columns belong to their respective datasets before executing calculations.