



A dark slide with a title and author information. The title is "DATA VISUALIZATION WITH TABLEAU OF COVID-19 VACCINE DATA". Below the title is the author's name, "Vaishnavi Arni Lakshmana Rao". The background of the slide is a blue-tinted image of various data charts and graphs.



## Objective

- ❖ COVID-19 vaccine is a vital preventive step that aided to end the COVID-19 pandemic. COVID-19 vaccinations are now widely accessible in the United States, and the CDC advises that all individuals aged 12 and older be immunized against COVID-19. We sought answers on how COVID 19 vaccination affected individuals.



# Data Description

The Primary Dataset has information regarding confirmed COVID19 cases county-wide for the state of Texas. We approached this data to prove five hypotheses using the tool Tableau.

- > The primary dataset County-wise vaccination rate. The sheets are:
  1. About The Data - explains every column and every unique heading in the data to understand the data better
  2. By county - 15 columns that segregates data as per doses administered, health condition, age, health care worker, resident and other uniqueness
  3. By Age, Gender, Race - explains data by gender, age group, race/ethnicity, and doses administered
  4. By Age, Day - number of doses administered per day in all counties by age group
  5. By County, Age - Vaccines administered at county, age and doses number level
  6. By County, Race - Break-up of data by race/ethnicity, county, and dose number
  7. By Vaccination Date - Break-up of data by vaccination date and doses administered



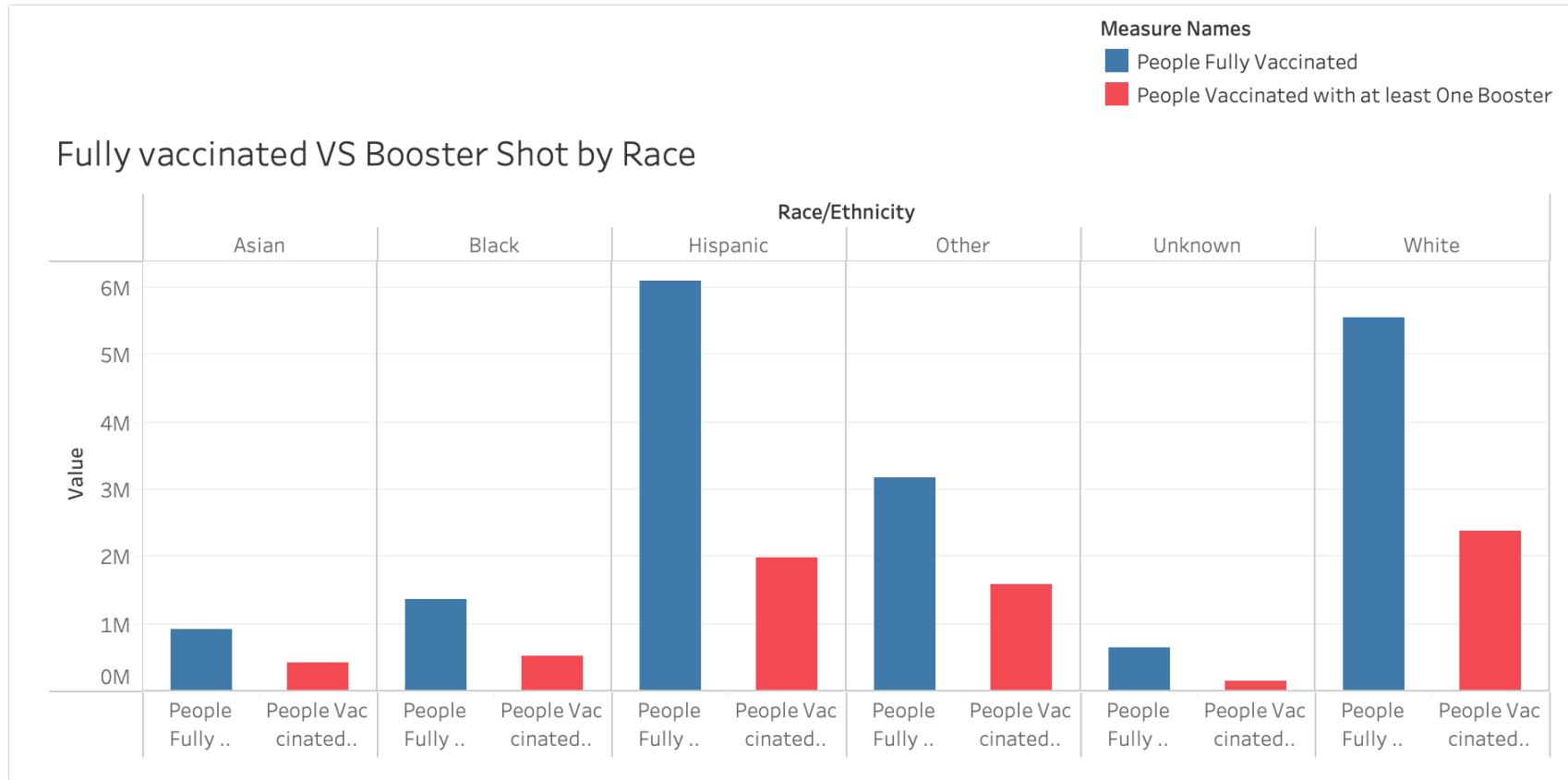
# Data Cleaning

- ❖ The data was cleaned using Excel and Python. Pandas library in python was mainly used for cleaning and Excel's built-in function was used to remove duplicates. Our original primary dataset consists of information on COVID-19 vaccine data by county. We've pre-processed the data and used only the information that was needed for our visualizations and our hypothesis.
- ❖ For the booster vaccination doses records after July 2021 were imputed since the booster shots have started in the month of August,2021. The COVID 19 vaccine data by County dataset records have the count of the people who received booster doses and immunocompromised people who received additional doses.
- ❖ Median value has been substituted wherever appropriate when the values were missing as this would not affect the distribution of the data.

# Hypothesis 1:

People of Black Ethnicity have received the least number of booster shots

- ❖ The visualization has considered 4 major ethnic groups: Asian, Hispanic, Black, and White. With the visualization, we can reject the null hypothesis as we see that the Asian ethnicity has received the least number of booster shots- 427,753. Before we make any further inferences about this visualization about the ethnicities here are a few things to consider:
  1. Asian communities are the least populous group in the counties that we have considered
  2. We need to look at the percentage of people that belong to a particular ethnicity to get a better idea about the data
  3. There are 2 other groups considered which are “others” and “unknown”. These total up to 1.7 million people. This makes 24% of the data. “others” and “unknown” can include people who are biracial, mixed, etc.



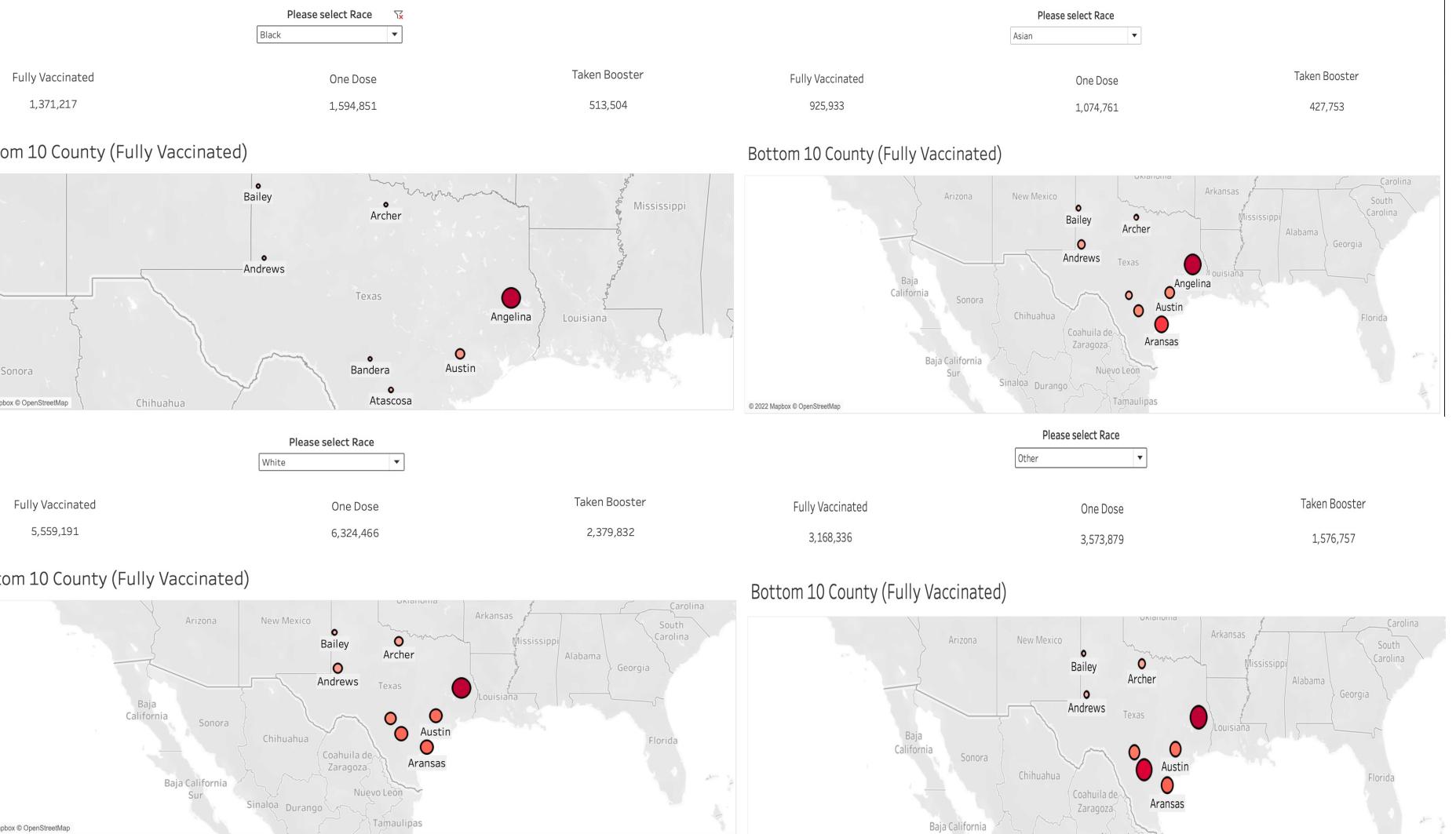


## Hypothesis 2:

There are a few common counties when we see least vaccinated areas by ethnicity.

- ❖ When we divide the data by county and by race, we see that the counties Austin, Angelina, Archer and Bailey have the least full vaccination rates for most ethnicities. Using this information, we can research why these counties have lower vaccination rates. Could socio-economic status, standard of living, vaccine availability, religion, and culture be the reason behind low vaccination rates? This demographic data can give us a good base for further research.

Here, we fail to reject the hypothesis.





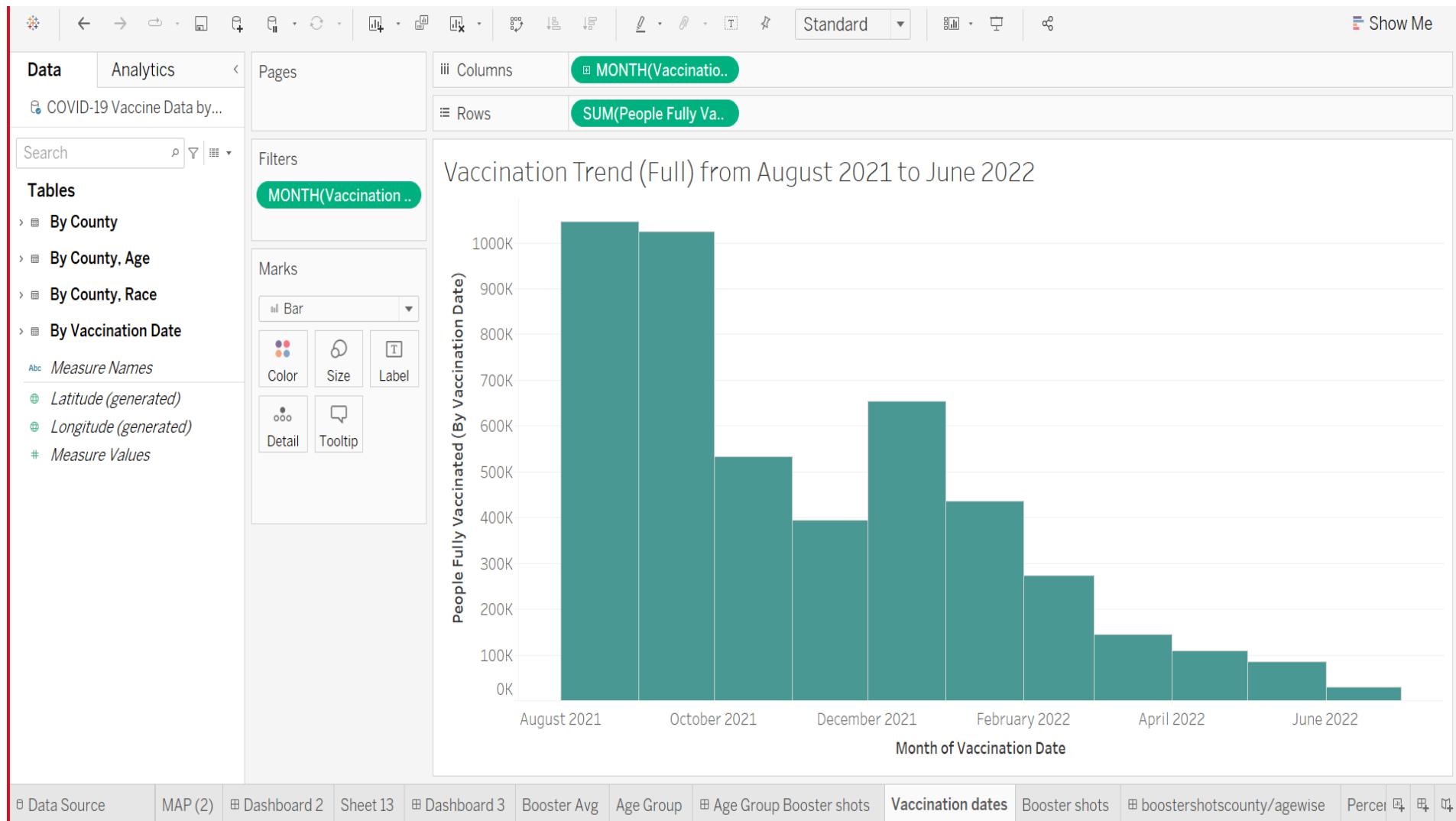
## Hypothesis 3:

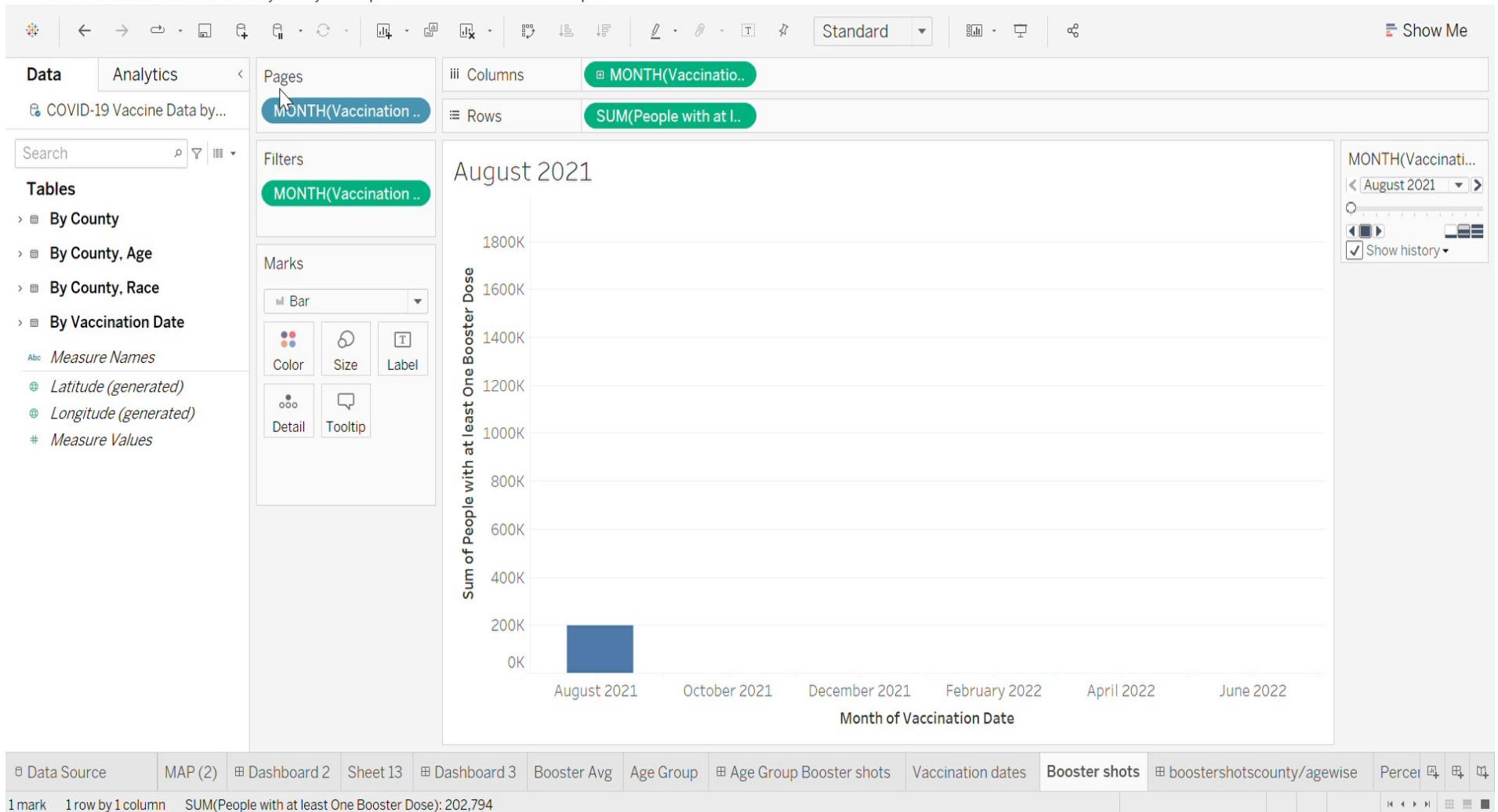
There is an increasing trend in the amount of booster shots taken by individuals from Aug 2021 to Jun 2022

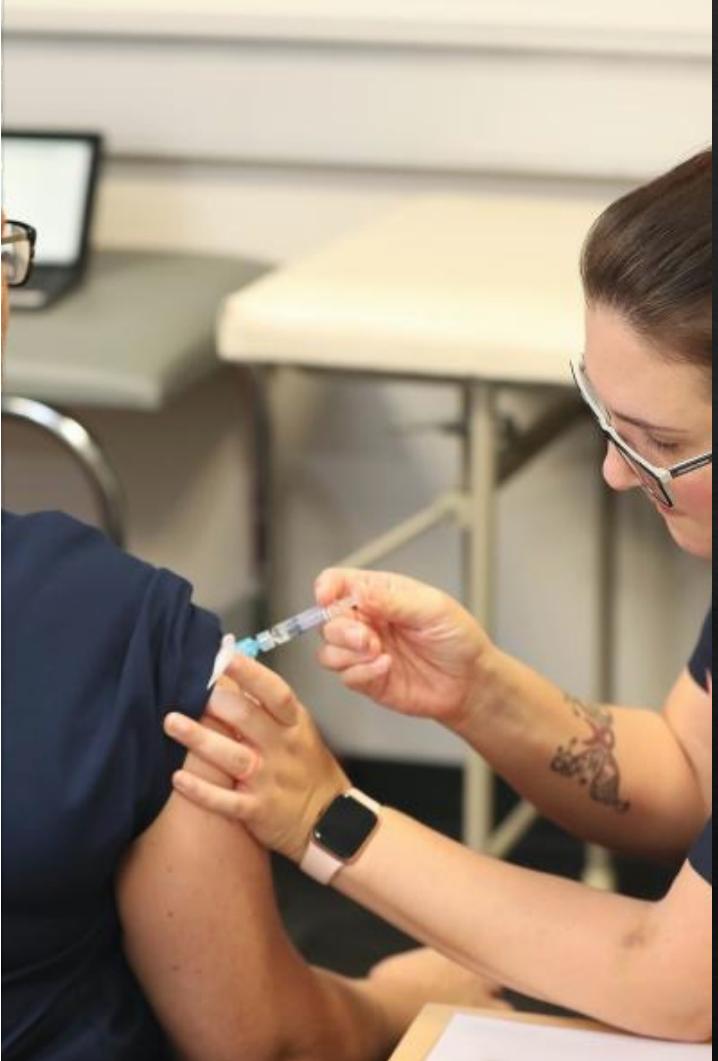
- ❖ With the visualization we can see that when the vaccine boosters were rolled out in Aug. 2021, there was a month over month increase in the boosters administered. However, we see a change in trend in 2022. Since Jan 2022, there has been a decline in the number of boosters administered every month. Therefore, we reject the Null Hypothesis.

❖ We can also see that from August 2021, there is a declining trend in the full vaccine doses administered as well. Using this information, researchers can investigate the reason for this sudden decline which could be:

- a) Decline in the number of COVID cases may have left people confident that they do not require the booster
- ❖ b) People believe that covid boosters may not be effective in reducing the severity of the disease
- ❖ c) most people have already been vaccinated
- ❖ d) Vaccine hesitancy among the people could have left them not wanting to complete their vaccination



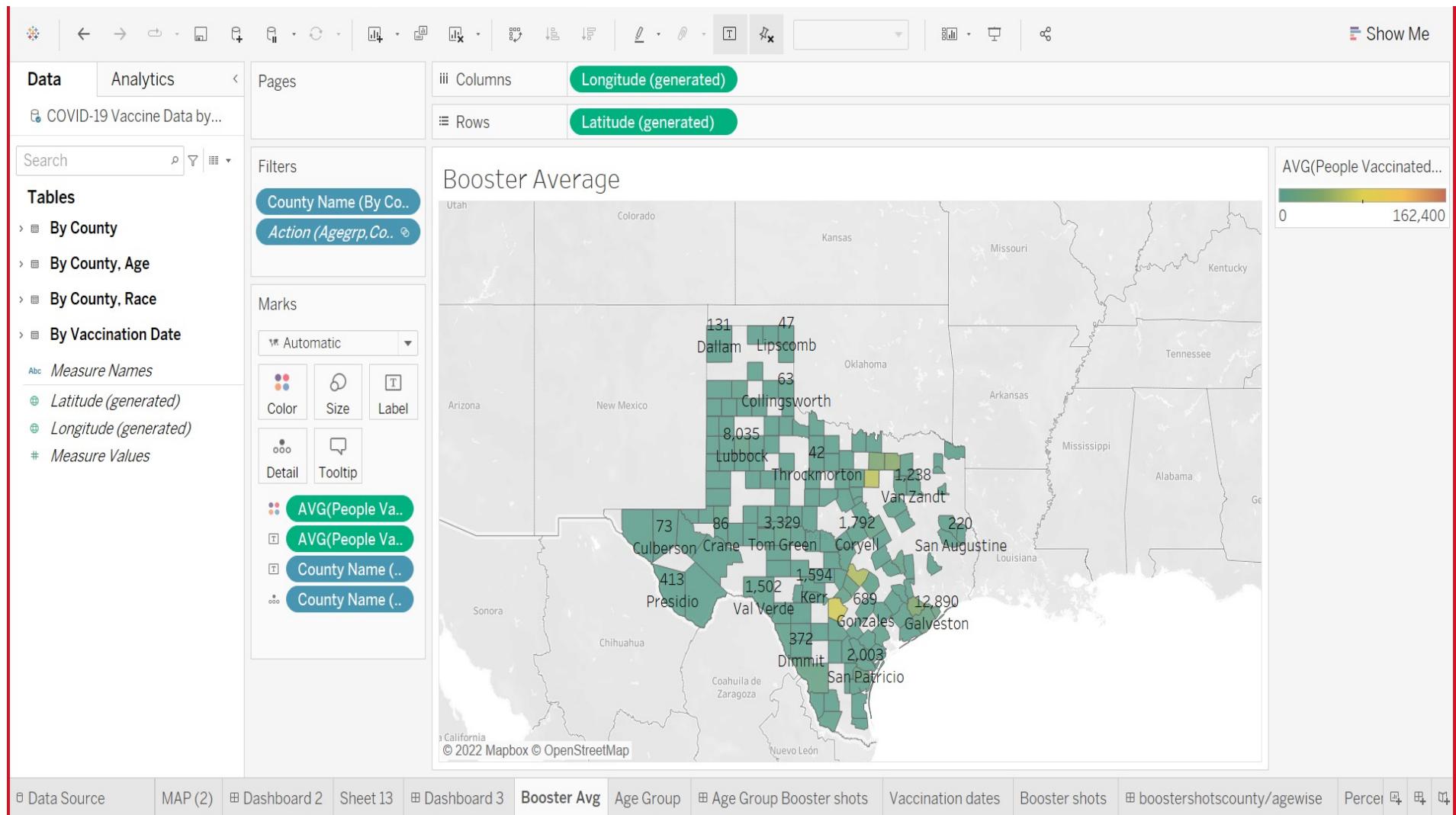


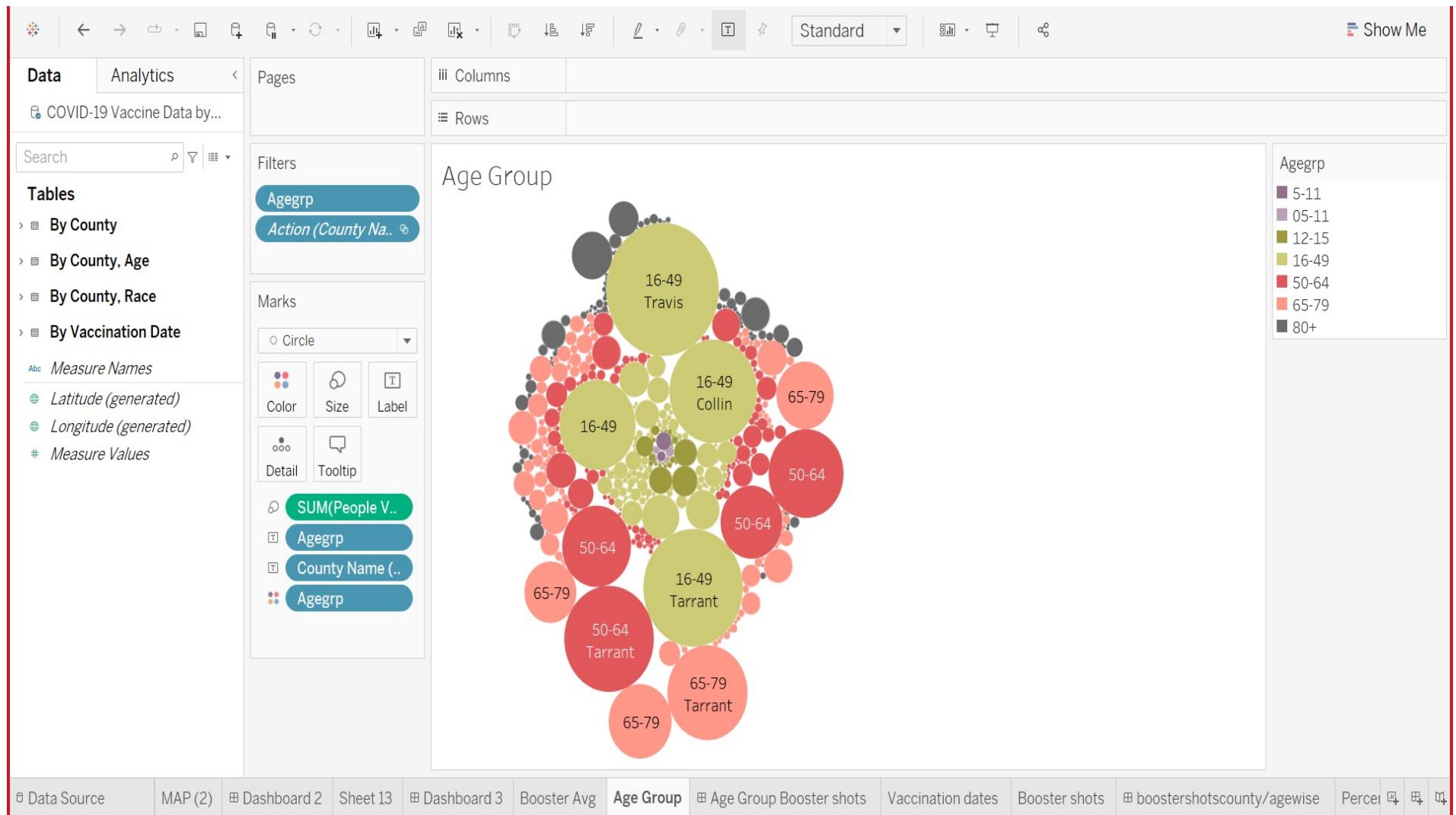


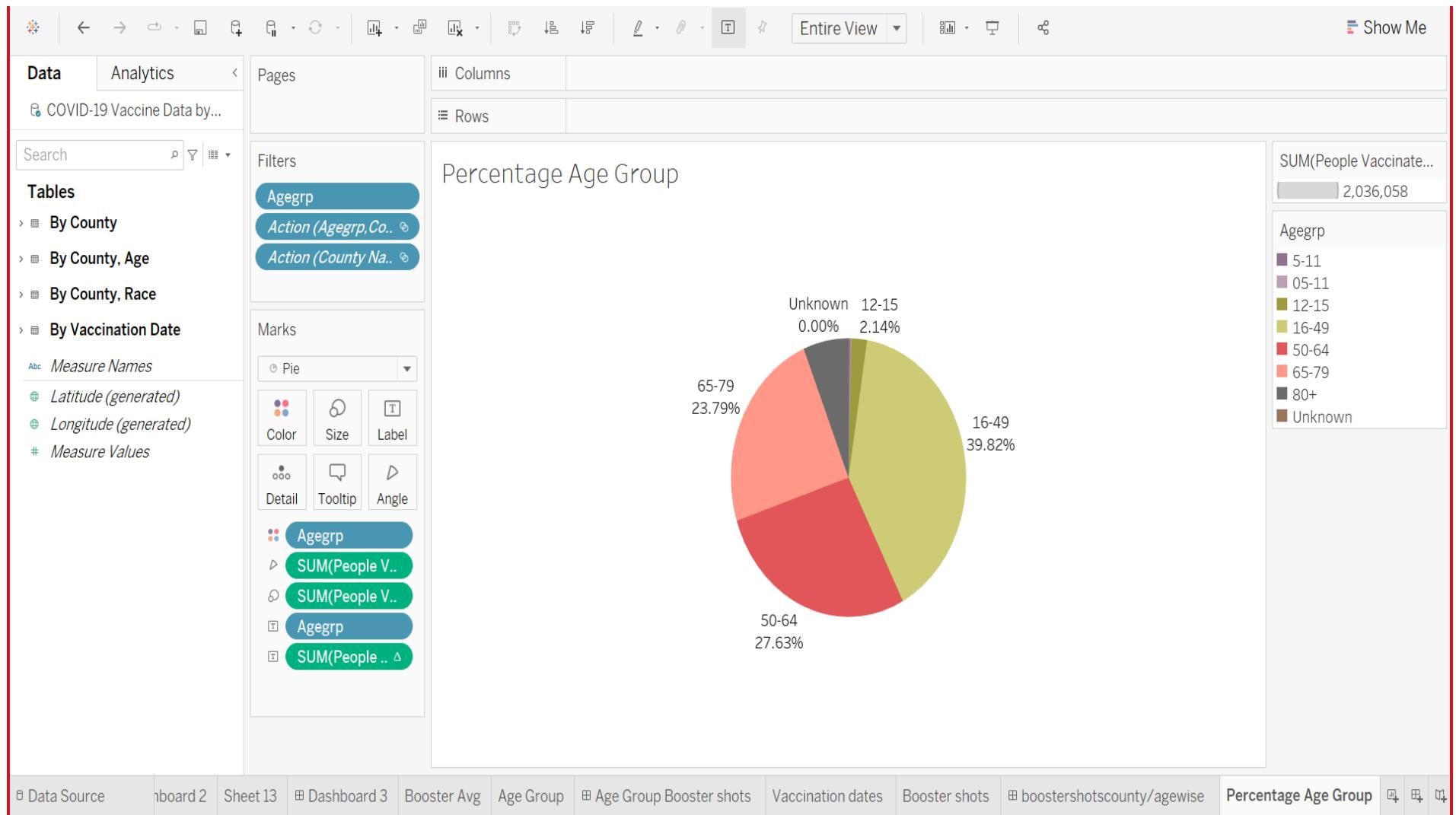
## Hypothesis 4:

Overall in all counties, we are seeing a high booster shot intake among youth and mid-age group

- ❖ As per the visualization 39% of the booster doses that have been administered in all the counties have been to the age group 16-49 which is the age group that we are testing in our visualization. The age group 50-64 is seeing the next highest number of booster dose intake. Among this we see that the counties Harris, Travis, Dallas, and Bexar have taken the highest boosters. This information can be further used to study why we're seeing a lower vaccination rate among other age groups as well as in other counties. We fail to reject the Null Hypothesis.







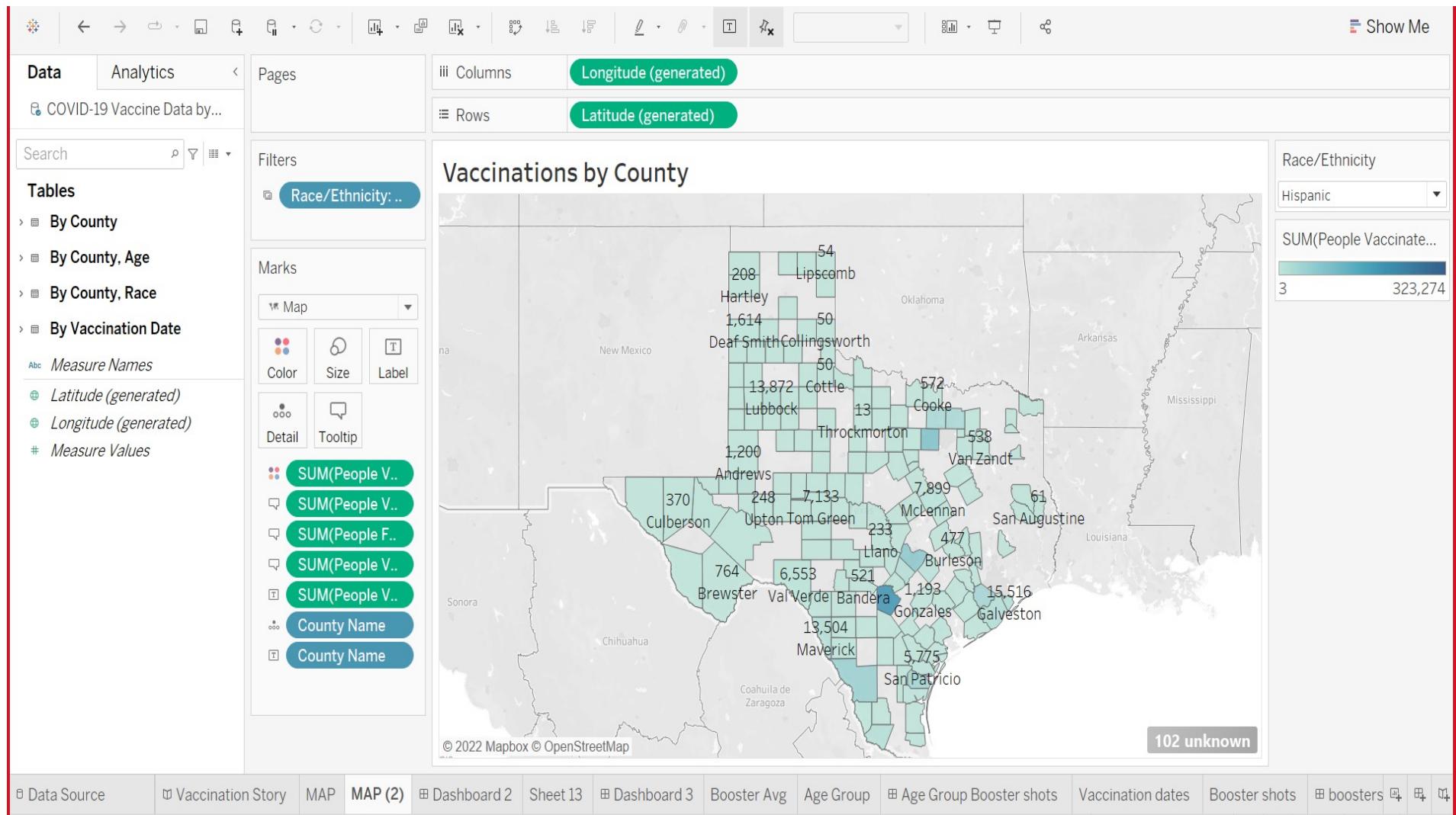


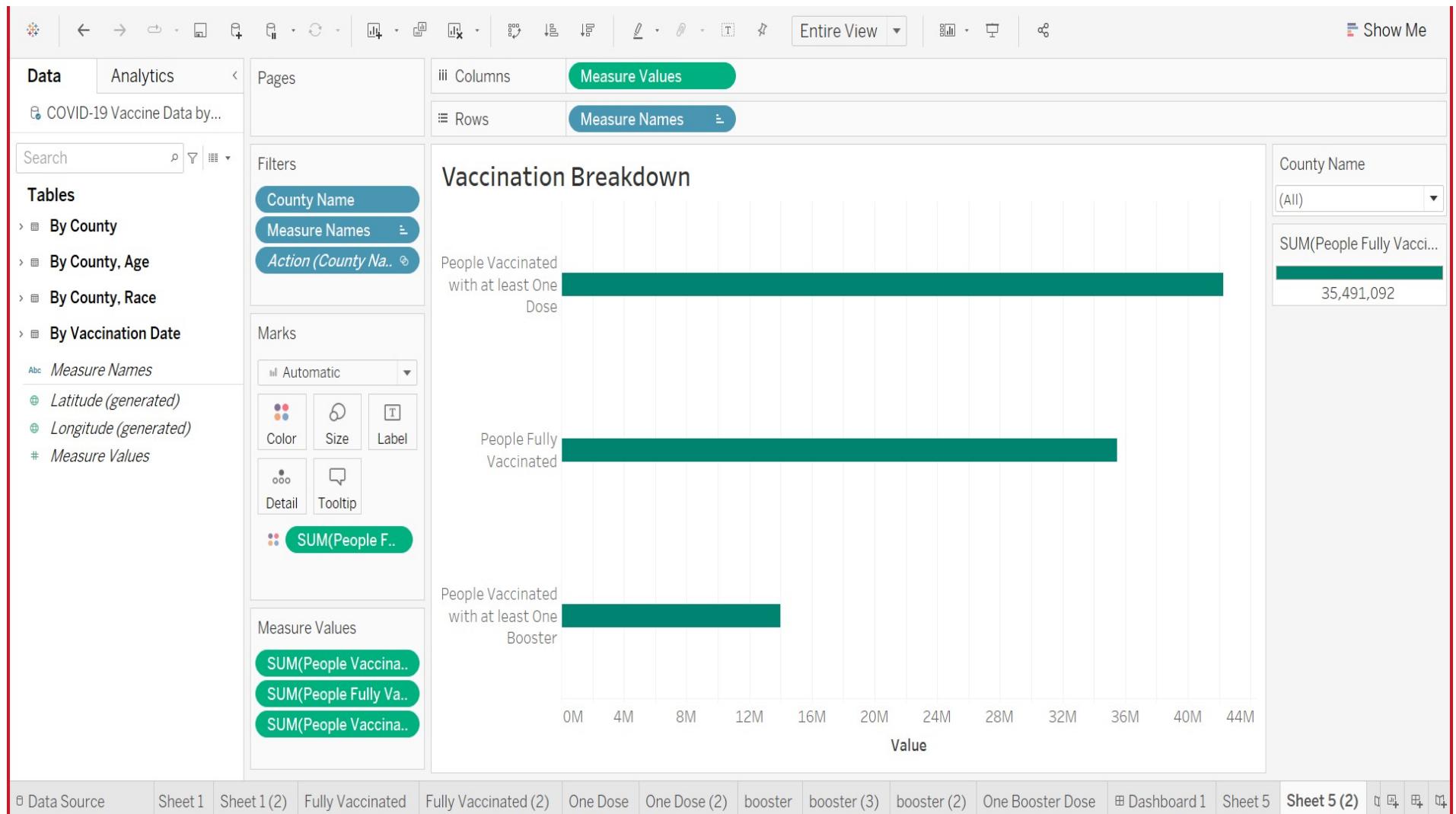
## Hypothesis 5:

Everyone that has taken a single vaccine dose has also taken the second vaccine dose and the booster dose.

We have taken data over 3 years, 2020, 2021, and 2022 to test our hypothesis. We see that there is a downward trend after a single vaccine dose is administered. The total number of first dose vaccines administered is 42.2 million, fully vaccinated is 35.5 million, and the total number of booster doses administered is 14 million. Based on this information we will reject the null hypothesis. These insights can be used to research why people have not followed up with the second and booster doses. Some of this can be:

1. Lack of availability of vaccines
2. Lack of trust in vaccines
3. Post covid 60 day immunity
4. Decline in the number of covid cases







# Conclusion

1. People of Black Ethnicity have received the least number of booster shots:

The visualization has considered 4 major ethnic groups: Asian, Hispanic, Black, and White. From the visualizations, we can draw a conclusion that we can reject the null hypothesis as we see that the Asian ethnicity has received the least number of booster shots- 427,753.

2. There are a few common counties where we see least vaccinated areas by ethnicity:

When we divide the data by county and by race, we see that the counties Austin, Angelina, Archer and Bailey have the least full vaccination rates for most ethnicities. From the visualizations, we can draw to a conclusion that we fail to reject the hypothesis.



# Conclusion

**3. There is an increasing trend in the amount of booster shots taken by individuals from Aug 2021 to Jun 2022:** We can see that when the vaccine boosters were rolled out in Aug. 2021, there was a month over month increase in the vaccines administered. Since Jan 2022 there is a decline in the number of vaccines administered every month. Therefore, we reject the Null Hypothesis.

**4. Overall in all counties, we are seeing a high booster shot intake among youth and mid-age group:** 39% of the booster doses that have been administered in all the counties have been to the age group 16-49 which is the age group that we are testing in our visualization. Hence, we fail to reject the null hypothesis.

**5. Everyone that has taken a single vaccine dose has also taken the second vaccine dose and the booster dose:** We see that there is a downward trend after a single vaccine dose is administered. Based on this information we will reject the null hypothesis.



THANK YOU