Project

Topic

My project for tableau dashboard is about the spread of COVID-19 in the USA. In this project, we are going to show the spread of Covid-19 and the deaths caused by it in the year 2020 by state. We are also going to show the number of recoveries from COVID-19. My dashboard will help users draw insights about the spread of the virus in the US in each state. It will help the users to see the trends in the rate of recovery/deaths and positive cases. It will help the users get an idea about how quickly or slowly the virus is spreading and how deadly it is.

Dataset

The dataset used for this project is from Kaggle.com. It reports cases and deaths in the US caused by COVID-19 on a daily basis from January 22nd, 2020, to December 6th 2020. There are three different files in the dataset but the one we are interested in, is the one that gives us daily record of Covid 19 cases by state. We are going to aggregate the data monthly when creating the dashboard. There are 55 columns and 15,633 rows in the dataset. Out of those 55 columns there are only a few columns that we are interested in, for our dashboard. The columns that we are interested in are:

1. Date
2. State
3. Positive
4. Recovered
5. Death

Dataset link: https://www.kaggle.com/sudalairajkumar/covid19-in-usa

List of questions/queries answered by the dashboard:

This dashboard will answer questions like:

1. What was the recovery rate of COVID-19 in 2020 in each state?
2. What was the death rate of COVID-19 in 2020 in each state?
3. How much did the virus spread in each state?
4. The months with the highest/lowest spread of the virus?
5. Months with the highest/lowest number of deaths/recoveries?
6. Which states have the highest/lowest rate of positive cases/deaths/recoveries?
7. States that have been hit hard by the virus?
8. States that have been affected the least by the virus?

9. Insights drawn by looking at the rate of the spread, rate of recovery and death rate of a particular state.
10. Trends in the spread of the virus.

This dashboard can be used by anyone, who is interested in knowing how the virus is spreading and its effect on the population in the USA. The whole world is worried about the pandemic right now, and everybody is interested in looking at the stats.

The most prominent users of this dashboard will be healthcare workers, researchers, lawmakers (to pass appropriate and timely laws/restrictions to help curb the spread of the virus), CDC and scientists. But as I mentioned earlier, this dashboard can be useful to anyone, given how the pandemic has affected everybody in the world.

Exploratory Dashboard

Exploratory dashboard uses filters and actions to help the user delve deeper into the dataset and explore the dataset. The user can answer a number of different questions about the data with the help of filters and actions provided on the dashboard. Incorporating filters and actions into the dashboard gives the user freedom to choose subsets of data or alter the context of the dashboard by using a selection. That helps the user to explore the data on a deeper level and answer different questions.

For our exploratory dashboard, we have three visualizations, Trend line, Map and Ranking.

The trend line shows us positive cases of Covid-19, deaths caused by it and recovered cases from the month of January to December of 2020, monthly in each state

The map gives us a total number of cases that have been tested positive, deaths caused and recovered cases for the year 2020 by state.

The ranking shows us positive cases, deaths and recovered cases by state in descending order for the year 2020.

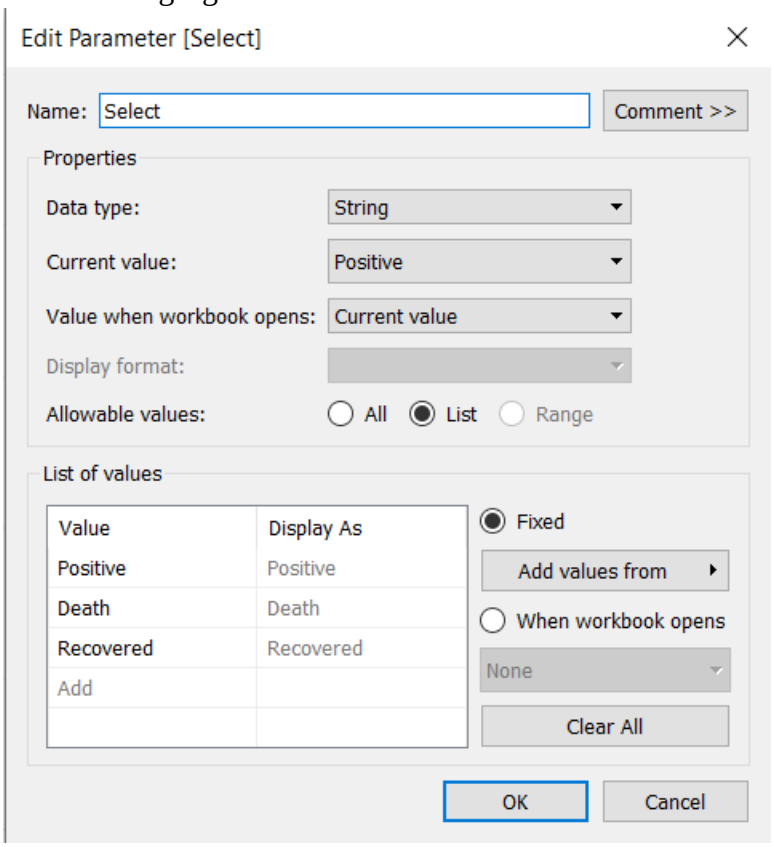
How was the dashboard created?

First, we created all the three visualizations separately in Tableau before adding them to our exploratory dashboard. We loaded the dataset (the one that had daily cases by state) that we downloaded from Kaggle.com to Tableau.

- Launch Tableau
- On its home page in the left hand panel that says Connect to a File, click on More…
- In the dialog box that pops up, select us_states_covid19_daily. This should load the CSV file as a data source into the newly created Tableau workbook.
- We'll create separate worksheets before creating the dashboard.
- The first worksheet we'll create is the Trend line.

Trend Line

- Double click the Sheet 1 tab and rename it as Trend Line.
- Next right-click (control-click on the Mac) in the blank space in the dimensions and measures area.
- From the context menu, select Create Parameter
- Using the dialog box that pops up, define a parameter called Select with values shown in the following figure.

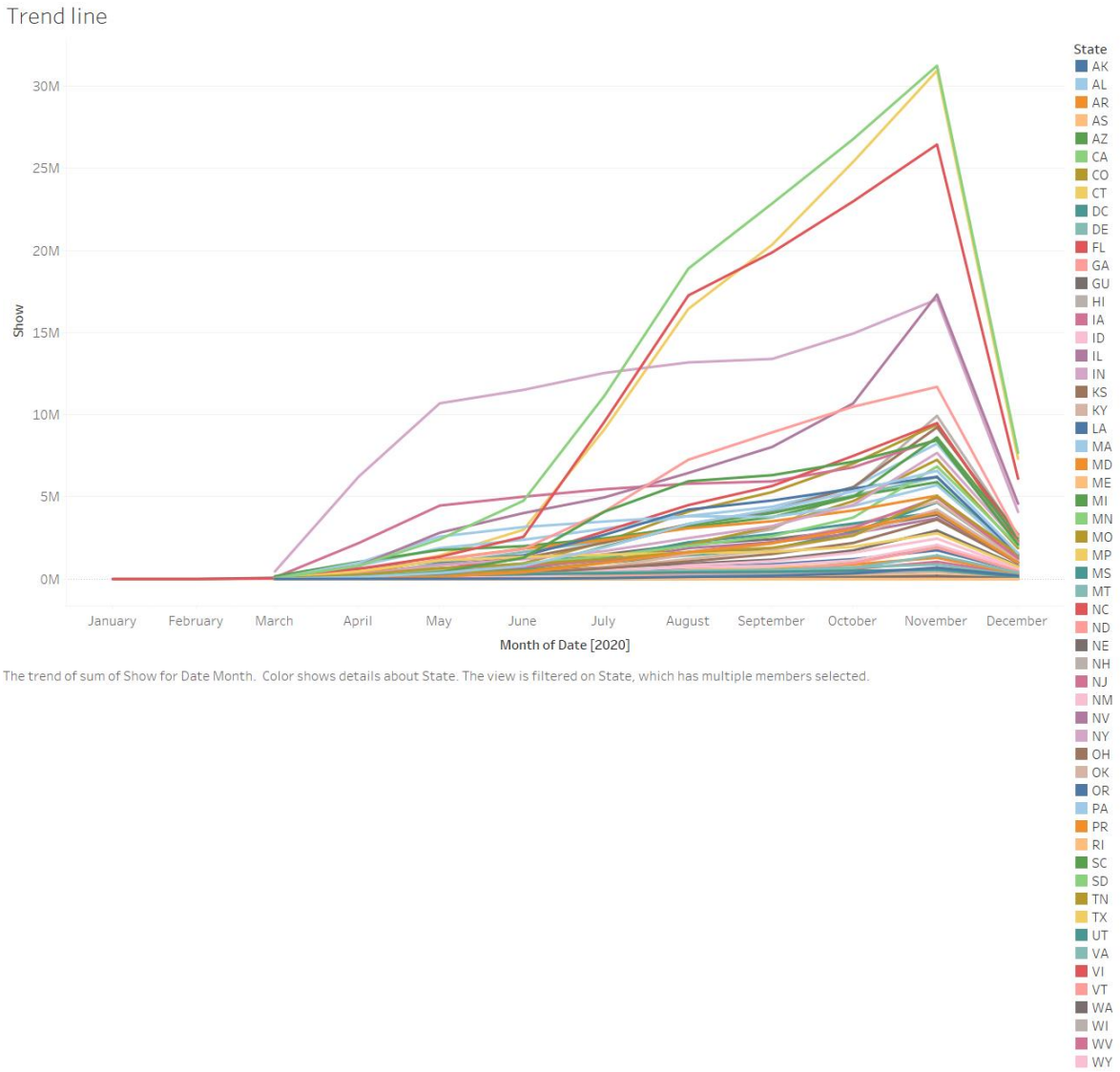


- Next create a calculated field called ‘Show’ by right clicking on an empty space in the dimensions and measure area. From the context menu select, ‘Create calculated field’. In the dialogue box that appears, type the following:

```
CASE [Select]
WHEN "Positive" THEN [Positive]
WHEN "Death" THEN [Death]
WHEN "Recovered" THEN [Recovered]
END
```

- This calculated field Show will be associated with the Select parameter.
- Next, drag Date into columns shelf and set it to Month Year.
- Drag Show (SUM) to the rows shelf.

- Drag state to the filters card and drag state to the colors shelf on the marks card.
- From the context menu of the Select Parameter (when you hover your mouse over Select, you should see a downward pointing arrow for the context menu), select Show Parameter Control.
- We should now have trend line that looks like the following, and we should be able to switch between Positive, Death, Recovered by using the Select parameter control (which is not visible on this image):



We now have a trend line for every state from the month of January to December with a select parameter where we have the option to select if we want to see the trend line for positive cases, deaths or recovered cases.

The trend line will help us see the rate of positive cases/deaths/recoveries for the year 2020 on a monthly basis. People who are interested in knowing how the virus is spreading with time and

affecting the population with time will use the trend line to answer questions related questions. For example the time/month when the spread of the virus slowed down, etc.

Map

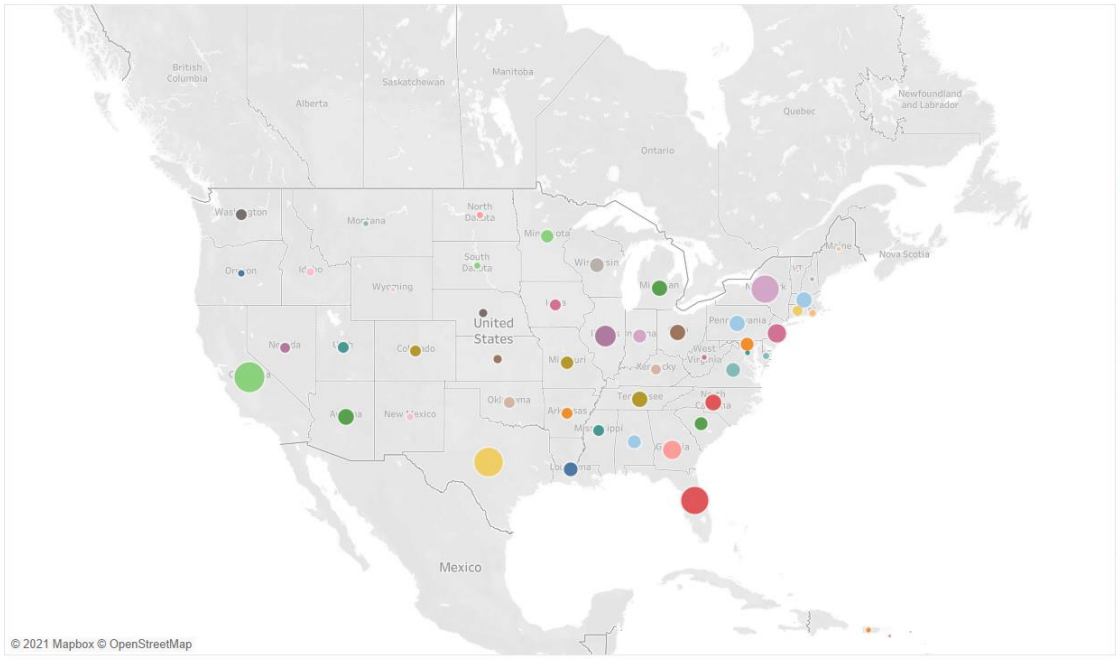
Now we'll create our second visualization/worksheet for the dashboard.

- To create a map, click the New Worksheet icon at the bottom to bring up a new worksheet and rename it Map.
- Drag longitude (generated) to columns shelf and Latitude (generated) to rows shelf.
- Drag state to the Filters shelf.
- Drag state to the colors shelf on marks card, Show (sum) to size shelf on the marks card and Date (set to Year) to the details shelf on the marks card.

We get a map with circles on each state. The size of the circle shows the spread of Covid-19. The bigger the circle, the higher the number of cases or deaths. We have a select parameter, which gives us the option to select if we want to see the number of positive cases in each state, recovered cases or deaths in each state.

We should now have a map that looks like the following, and we should be able to switch between Positive, Death, Recovered by using the Select parameter control (which is not visible on this image):

Map



Map based on Longitude (generated) and Latitude (generated). Color shows details about State. Size shows sum of Show. Details are shown for Date Year. The view is filtered on State, which keeps 56 of 56 members.

Show
0
50,000,000
100,000,000
126,643,296

State
AK
AL
AR
AS
AZ
CA
CO
CT
DC
DE
FL
GA
GU
HI
IA
ID
IL
IN
KS
KY
LA
MA
MD
ME
MI
MN
MO
MP
MS
MT
NC
ND
NE
NH
NJ
NM
NV
NY
OH
OK
OR
PA
PR
RI
SC
SD
TN
TX
UT
VA
VI
VT
WA
WI
WV
WY

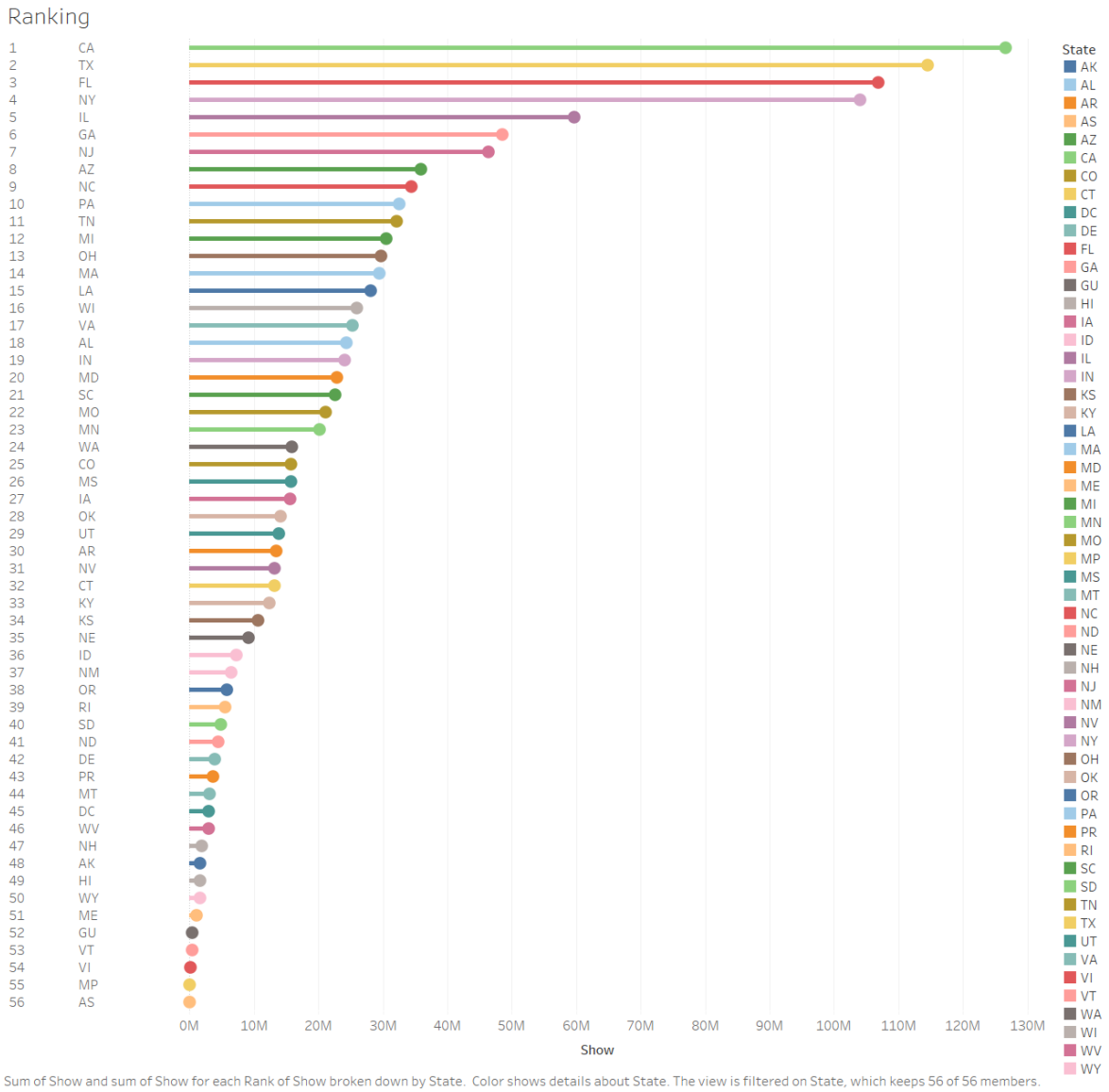
The map will give us a visual overview of all the states affected by COVID-19. We can easily tell from the map the states that have been affected the most and the ones that have been affected the least. The map will help the user to delve deep into geospatial data and answer questions related to it. For example like the region affected the most/least by the virus.

Ranking

The third and last visualization/worksheet for the dashboard is the ranking which is a dual-axis bar-chart (sometimes called the lollipop chart).

- Click the New Worksheet icon at the bottom to bring up a new worksheet and rename it Ranking.
- Drag SUM(Show) to the Columns shelf twice. Notice, how tableau creates two charts that share the same x-axis.
- To superimpose these charts: On the second of the two pills on the Columns shelf, click on the down arrow to bring up its context menu. Select Dual Axis
- On the x-axis showing on the top, right-click (control-click on Mac) and select Synchronize Axis. Right-click again and uncheck Show Header
- You now have a single x-axis and the charts are superimposed.
- Drag State to the rows shelf.
- Notice that the Marks card shows multiple sections, one for each column and one titled All that applies to all the columns. In the All section, drag State to the color shelf.
- Now drag SUM(Show) to Rows shelf.
- On the pill, click on the down arrow to bring up its context menu, select Discrete.
- Move the Show pill to the left of State pill on the Rows shelf.
- From its context menu again, select Add Table Calculation
- In the dialog box that pops up, change the Calculation Type to Rank and Compute Using to Specific Dimensions.
- For Specific Dimensions, make sure you check State before closing the dialog box.
- This should rank individual countries based on the select parameter chosen at the time.
- On the Marks card, click on the top of the two SUM(Show) sections and, from the drop-down list, change its mark type to circle. You may increase the size of the circle by clicking on Size shelf in this section and using the slider.
- Now click on the bottom SUM(Show) section and change its mark type to bar. Again, click on Size and using the slider make the bars skinnier so that the bar and circles together begin to look like a lollipop.
- In the top left section of the chart, right-click (control-click on Mac) the row labels and select Hide Field Labels for Rows.
- Also, right-click anywhere on the chart and select Format.
- In the formatting panel that appears on the left-hand side, click on the borders icon and, on the Sheet tab, change the Pane selection for the Row Divider and Column Divider to None.

We now have the dual-axis bar chart that looks like the following:



With the help of Ranking visualization, the user can find out questions related to ranking (as the name implies). The user can find answers to questions like the top three states with the most positive cases/deaths/recoveries. The top three states that have been affected the least by the virus and so on.

Dashboard

We can now start assembling our dashboard.

- Click on the New Dashboard icon at the bottom. This should bring up a new dashboard. Name this dashboard Exploratory.
- From the Dashboard panel on the left-hand side, drag and drop the worksheet Trend Line on to the dashboard.

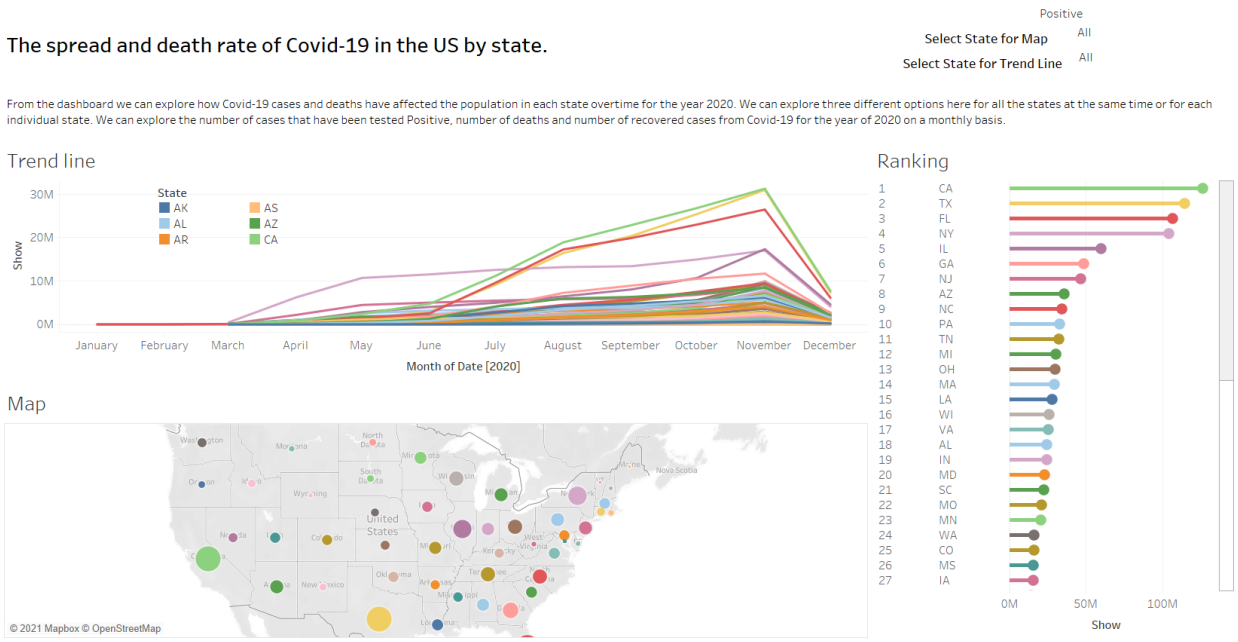
- Next drag and drop the Map on to the dashboard vertically below the Trend Line worksheet.
- Please note, as you drag the Map onto the dashboard, Tableau grays out the region where it will place the Map on the dashboard.
- When you feel you have the correct location, just drop the Map there.
- The Trend Line and Map have legends associated with them that also show up on the dashboard.
- Select the one titled "Show" and remove it by clicking on the x on the top right of its window
- Select the one titled "State" and from its context menu (click on the drop down arrow in the top right corner of its window to bring it up), select floating. Tableau will undock the legend and will allow us to place it wherever we desire. For now, move it some place where it does not interfere with us. Later we will place it in the top right corner of our dashboard.
- Next drag and drop the Ranking chart to the right of the Trend Line and the Map
- Carefully observe the grayed-out area to make sure it spans across both the Trend Line and the Map, and is to their right before you drop the Ranking chart.
- Adjust the size of the chart by dragging out one of its edges.
- Also make the Select parameter control floating. Turn off it's title and place it in the top left corner of the dashboard.
- From the dashboard panel on the left-hand side (right below the worksheets), drag a Text object and place it at the top.
- Again, before you drop the Text object, carefully observe the grayed-out area to make sure it spans horizontally across both the Trend Line and the Ranking chart at the top.
- Once you drop the text object, a dialog box pops up. Enter the following title text (adjust the font type and size) and click OK:

  "The spread and death rate of Covid-19 in the US by state."
- Drag and drop another Text object right below the title text ensuring it spans the length of the title text.
- Enter the following lead-in paragraph text (adjust the font type and size) and click OK:

  "From the dashboard we can explore how Covid-19 cases and deaths have affected the population in each state overtime for the year 2020. We can explore three different options here for all the states at the same time or for each individual state. We can explore the number of cases that have been tested Positive, number of deaths and number of recovered cases from Covid-19 for the year of 2020 on a monthly basis."
- With the Map worksheet selected, go to the Analysis menu, and then select Filters → State. This places State filter on the dashboard.
- From its context menu:
  - make the filter floating
  - turn off its title
  - make it a single value dropdown

    - select Apply to Worksheets → All Using This Data Source
- Place it on the top below the Select parameter.
- Also place a floating Text object labeled “Select State for Map” to the left of it.

- With the Trend Line worksheet selected, go to the Analysis menu, and then select Filters → State. This places State filter on the dashboard.
- From its context menu:
    - make the filter floating
    - turn off its title
    - make it a single value dropdown
    - select Apply to Worksheets → All Using This Data Source
- Place it on the top below the Select State for Map filter.
- Also place a floating Text object labeled “Select State for Trend Line” to the left of it.
- Place the State legend somewhere on the trend line, where it won’t interfere with the visualization.
- Place the State legend somewhere on the trend line, where it won’t interfere with the visualization.
- In the Dashboard panel, set the size of the Dashboard to automatic and press F7 to launch the dashboard (or Option-Return on a Mac**).** [Note**:** you can press ESC to return to Tableau]. You should now see an interactive dashboard much like the image below.



Reasons for choosing the three visualizations for our dashboard

We decided to choose trend line, map and ranking as our visualizations for the dashboard because all three of them serve a different purpose and helps the user in finding answers to various questions.

The trend line helps the user to find answer to questions and queries related to time. It gives the user a timeline of the spread of the virus. The user can see how the virus is spreading each month, if it's slowing down, if it's spreading rapidly by looking at the trend line. The user can find trends in the spread of the virus with the help of trend line (as the name implies). The user can look at the timeline and then check for months where the spread of the virus slowed down. Then the user can investigate if there were any restrictions that were put in place to help curb the spread of the virus. For the months where we see a surge in the spread of the virus, the user can find out from the news/articles the reasons for that. That way the user can find trends in the spread of the virus with the laws/restrictions held in place during that time.

If a user is interested in looking at the regions where the virus has spread the most, or just get a general idea of how the virus is spreading in different regions/states, the map can help the user give a nice overview of that. The map also gives the user information about the total number of positive cases/deaths/recoveries for every state. If the user isn't interested in finding out the spread of the virus every month but in fact is interested in looking at the information of how bad each state has been hit by the virus, the map can help with that.

The ranking chart (dual-axis bar chart) gives the user information about the states that have been affected the most to the least in order or the other way around. Ranking chart will help answer questions related to the ranking (as the name implies) of the states according to the spread of the virus.

As we can see all three visualizations help the user to answer different questions and dive deeper into the data to draw valuable insights from the data.