**CS-498-Data-Visualization**

**A Visualization Narrative by Shilpesh Pathak (**[**spathak3@illinois.edu**](mailto:spathak3@illinois.edu)**)**

**Messaging**

COVID-19 pandemic has caused unprecedented public health and economic crises across the world and continues to have same impact; This visualization is intended to present point of view on COVID-19 pandemic that has affected united states in particular (From January to July 25th), There are 3 different pages that forms a story; First pages gives you holistic information about how confirmed cases have risen from January to July 25th across US followed by a bubble chart to show how many people have been getting tested across individual states and final page shows list of states having high number of recoveries.

As we know by now that covid-19 is not going anytime soon which means conducting tests becomes more and more important, state government(s) are trying to test as many people as possible so that better preventive actions can be taken and this can be seen on recoveries some of the states have seen lately, for example Texas had over 4 million people tested and it is also the state with highest number of recoveries this presents a view that as you conduct more testing, chances of identifying potential positive cases increases thereby giving doctors better chances to effectively treat people(s) and save their life’s.

**Narrative Structure**

This narrative visualization uses the Interactive Slideshow approach to take you through a small use case of COVID-19 dataset generated and maintained by John Hopkins university.

John Hopkins university (JHU) contains various datasets on covid-19, for this visualization I am using time series summary data to show how US started seeing an increase in confirmed case from January to July 25th; Another dataset that I am using is a “Daily dataset” that gives cumulative numbers (State wise) on Confirmed cases, people tested, number of deaths, active cases amongst others.

This was a very small use case of generating views/trends and there is always a possibility to enhance a visualization to show a trend that would make more sense as the time progresses, Since JHU dataset contains lot of critical information an obvious enhancement to this visualization using D3 could be to generate a trend that shows how mortality rate is affected by hospitalization rate and testing rate across US states, you can also use this dataset to answer questions like “Has the curve started getting flat? And which state is handling it better”

JHU dataset can be downloaded from here - <https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data>

**Scenes**

A homogeneous template was created using CSS layout and javascript to set up the scenes for this visual story. Scene transition follows an intuitive approach of sliding up and down via mouse, touch or key-up/down. Slide navigator, along with tool-tip located at the right side of each slide provides a good navigation visual cue. It also acts as a trigger allowing readers to jump to a specific slide.

The order of the scenes were chosen to first highlight the overall trend based on one of the critical elements in data in scene 1, second to explore additional trend in scene 2 based on number of people getting tested in individual states here user have the option to filter out states based on how many people have been tested, categorically there are five options to choose from and that highlights the state which satisfies the condition and finally in scene 3 show the trend on recoveries seen in different states.

**Annotations**

Annotations are used to highlight a trend in the data, direct the user to further investigate the data, and ask the user to draw a conclusion from the data. The annotations use a consistent template for font size and style

* The annotation in Scene 1 is texts positioned inside the visualization container meant to highlight the primary finding of the visualization, particularly show “Initial confirmed count” and “confirmed cases till date!”. There are also labels and legends visible to give additional context about the visualization.
* Each data point in the bubble chart (scene 2) is clearly labeled with tooltip showing “No of people tested ” and “count” upon hovering over the bubble, there are also legends visible to give additional context about the visualization
* Scene 3 Bar chart is clearly labeled and with legends and various bar charts are equipped to show tool tips containing “State Name” and “Recovered” this gives additional details to user.

**Parameters**

The main parameters of this story are US states, overall confirmed cases, number of people tested and number of people recovered; One of the parameters used in scene2 (Bubble chart) is “People Tested”, here user has the ability to examine the data based on value selected in “People tested” also by hovering over any bubble in the chart user can examine exact count. For bar chart as well user has the ability to examine the data on each bar by hovering over it, this allows user to gain more information about number of recoveries

For line and bar chart visualization, state is pre-defined i.e. they are designed to show a trend based on critical data elements present in the dataset, however for bubble chart initial state is to show all the datapoints and as user selects a new value in drop down the state of visualization changes to reflect selected values

**Triggers**

Triggers are utilized in two ways for this visualization

* There is a slide navigator available on all the pages which serves as a trigger thereby allowing users to jump from one page to another, sliders are labeled so that user can easily navigate between pages.
* “People Tested” selector serves as the mechanism to trigger the highlights of states based on the selected type (e.g. All, Between 0 to 500k, Between 500k to 1 million). Data tooltips appear upon mouse over event.