

## OVERVIEW

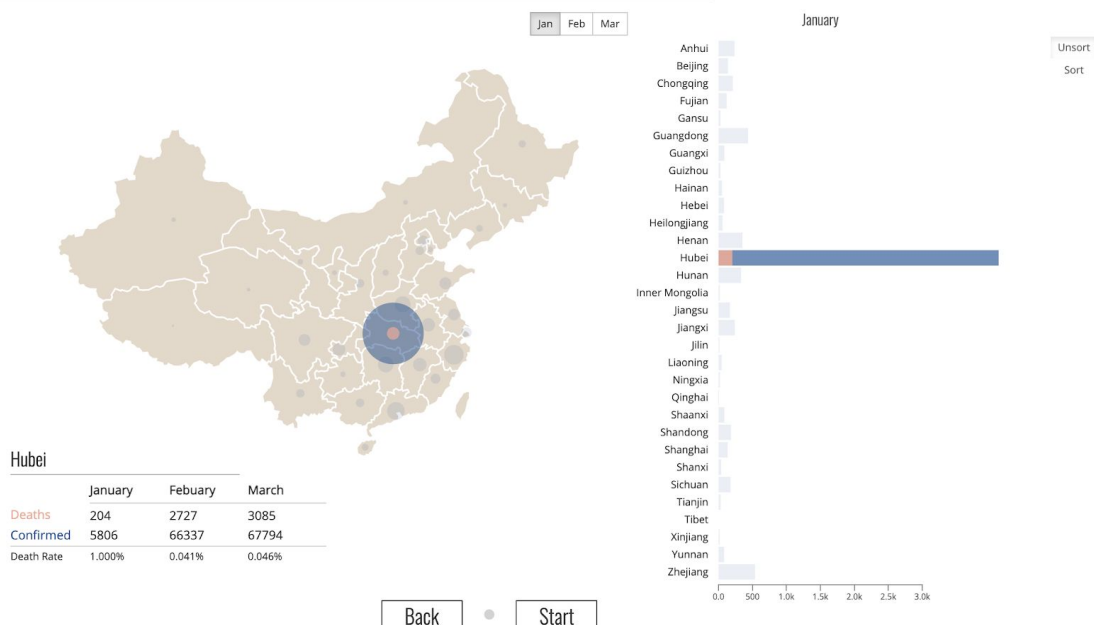
### China COVID-19 Map with Confirmed and Deaths Cases

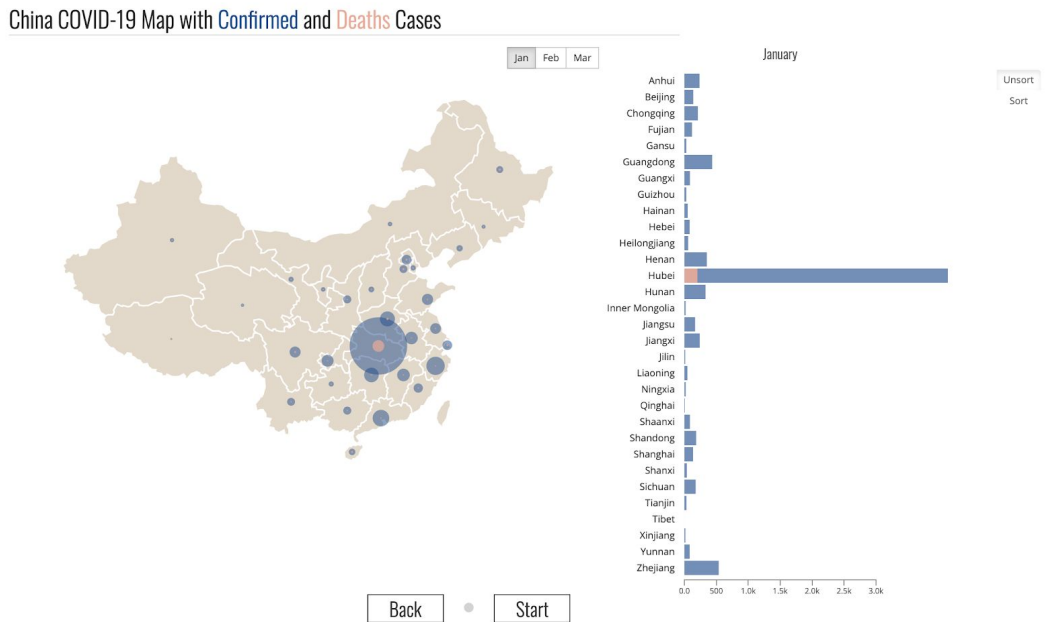
COVID-19 has become the center of contention since the occurrence of a cluster of viral pneumonia cases in Wuhan, Hubei Province, since December 2019. According to an investigation by the Mainland health authorities, a novel coronavirus is found to be the causative agent. Therefore, this project is designed to visualize the COVID-19 confirmed/recovered/death cases from January to March in 31 districts across China. It aims to help people better understand the spread of the pandemic across the country and how well different provinces respond to this disease outbreak.

Back

Start

### China COVID-19 Map with Confirmed and Deaths Cases





The project is designed to visualize the COVID-19 confirmed and death cases from January to March in 31 districts across China. It aims to help people better understand the spread of the pandemic across the country and how well different provinces respond to this disease outbreak.

Two views will be displayed to visualize the confirmed and death cases. A map of China filled with small points, indicating the number of cases with its size in the province. A bar chart on the side with linked highlighting to the map describing the number of cases in the province. The pinkish color denotes the death cases and the blueish color means the confirmed cases. There are also three radioboxes (Jan/Feb/Mar), allowing the user to switch between cases from different months. Another two radioboxes control sorting of the bar charts. Tooltip of the provinces is displayed at the bottom left corner.

## DATA

The original dataset “China COVID-19 Data” can be found on Kaggle. It contains categorical attributes: “Province/State”, “Country/Region”, quantitative attributes: “# Confirmed”, “# Deaths”, “# Recovered”, “Latitude”, “Longitude”, and ordinal attributes: “Observation Date”. The dataset type is tables with a total of 1666 records.

Link: <https://www.kaggle.com/aestheteam01/china-covid19-data>

In this project, the column “Country/Region” were be dropped due to all cells contain the same value “China”. The attributes “Latitude”, “Longitude” and “Province/State” were used as

geographical variables. The number of confirmed/recovered/death cases were recalculated on a monthly basis instead of a daily basis as in the original dataset to derive new variables. Given the size of the dataset, the calculation was done using excel.

## GOALS AND TASKS

Samantha is an epidemiology student who wishes to learn about the early stage of the COVID-19 outbreak. She wants to be able to see when and where the disease first occurred, how it spread throughout the country and how well different provinces were responding to this pandemic.

When Samantha logs on this system, she will see a map of China which is filled with small points, indicating the number of cases in the province. There will also be a bar chart on the side linked to the map, illustrating the number of cases in the province. There are radioboxes by which Samantha can select the specific month (Jan, Feb, Mar) and filter out cases at other times.

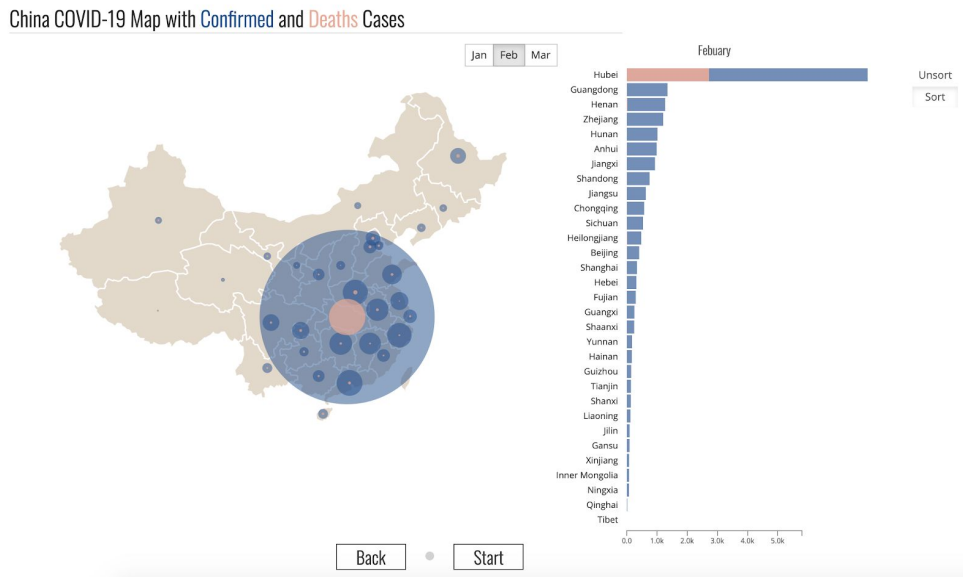
Moreover, there are two radioboxes that allow Samantha to view either the sorted barchart or the original one. She'll also be able to see a tooltip on the bottom left corner showing the rates.

The size of the points on the map is in proportion to the number of cases at the location. The pinkish circles (denoting the death cases) and the blueish circles (denoting the confirmed cases) are overlapped so that viewers can easily make comparisons between them. The same reason explains for the overlapped bars in the bar chart. The bar chart and the map are bidirectionally linked also to make the comparisons clearer. There are three radio boxes (Jan/Feb/Mar), allowing the user to see either the cases in different months, so the spread of the disease can be observed.

## VISUALIZATION

The two views will be displayed to visualize the confirmed/death cases. A map of China filled with small points, indicating the number of cases in the province. A bar chart on the side with linked highlighting to the map describing the number of cases in the province. A radio button group will control the month of cases displayed in the graph. The size of the points on the map view is in direct proportion to the number of cases under each condition. The death section of the bar chart is shown with the confirmed statistics as the background so that the percentage of deaths over confirmed cases is also shown.

The barchart view is designed to be interactive with another widget: one could choose between sorted and unsorted barcharts to see the rank of each province. Below is the sorted barchart with the data of February.



The map view and bar chart view are linked. When clicking on a point/bar on one view, the corresponding item will light up on the other view. The screenshots in the overview section has shown how the views are linked.

## REFLECTION

In our initial proposal, we designed to visualize the data of recovered cases. However, when we dive more into the dataset later, we realized that the difference between the number of confirmed cases and recovered cases is not large enough so that we could visualize it clearly and display valuable information to people. In our final product, the percentage of deaths and confirmed cases is clearly shown. Instead of choosing between recovered cases and deaths, we design a new widget to show the ranking of cases in different provinces.

Our technical goals of visualizing the data have slightly changed. When developing our visualization of the COVID-19 China data, the large number of cases in Hubei comes as a problem for us. The statistics of Hubei province is ten times larger than any other province and the difference could be clearly seen in two views. The size of points on the map view is adjusted so that the maximum size is not exceeding the scope of the entire map. The scale of the barchart between months are not the same. For January, the maximum value is 3.0k and for February and March, the maximum value is 5.0k. That is for the purpose of showing the data of provinces other than Hubei more clearly. The bar of Hubei exceeds the x-domain on the barchart each month but the change in the number of cases and the rank of other provinces is demonstrated with this design.