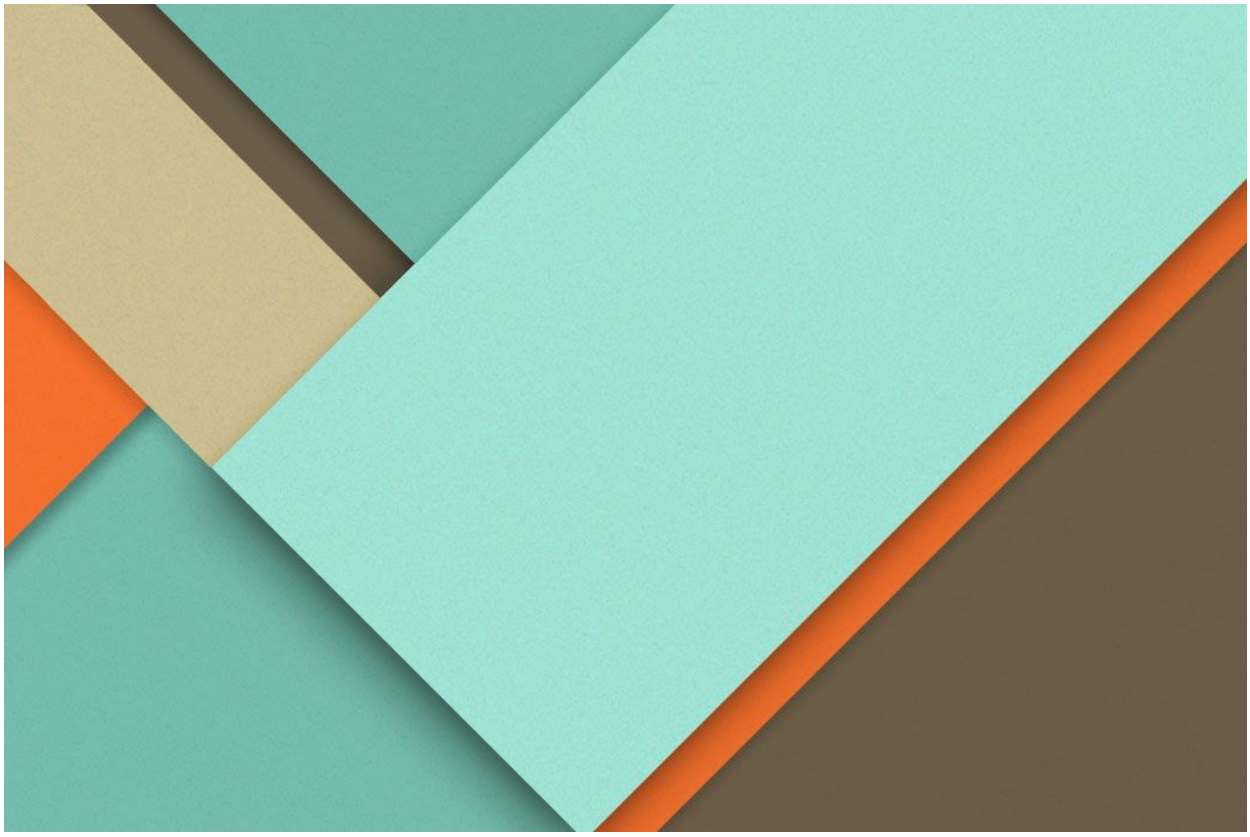


Covid-19 Visualization

SI 649 final project, team 1-3

April 21st, 2020

Team Members:

Shengnan Duan (elenore)

Jin-Seo Bae (jinbae)

Supapitch Kittisarakul (supapitk)

Xiaoshan He (xshe)

Xinhao Liao (theoliao)

Contents

Executive Summary

1) Members in the group

   👩 Shengnan Duan (elenore)
   👩 Jin-Seo Bae (jinbae)
   👩 Supapitch Kittisarakul (supapitk)
   👩 Xiaoshan He (xshe)
   👨 Xinhao Liao (theoliao)

2) Communication goal

   We wanted to tell a story of how COVID-19 has started and has spread all over the world. In the meantime, we also wanted to show and compare how this outbreak is being handled by the governments and the medical systems in the different countries. Some discovered patterns, based on how the medical systems in different countries have been working since day 1, can be used for comparison and future prediction.

   With this visualization, the audience is expected to understand more about the spreading pattern of COVID-19 over the time and have a sense of the conditions of medical systems in different countries.

3) Target audience

   Our expected audiences include following three groups:
   - Individuals who are interested in or have influenced by COVID-19
   - Any public institution that conducts a research on COVID-19
   - Policy makers who track and adjust policy strategies to handle COVID-19

Questionnaire

1) What is your data?

    - Recorded data for number of confirmed cases, deaths, and recoveries of COVID-19 across the world overtime. (https://www.kaggle.com/imdevskp/corona-virus-report)

2) What are the domain tasks or communication goals you want to support? What should someone be able to understand after seeing / using your visualization?

    - Our main domain task is to show the pattern of how the confirmed cases number globally change overtime. With our visualization, we tried to display geographically how many people have been infected worldwide in the past few months. Also, we wanted to show how both the government and the hospitals, in a particular country, have handled COVID-19 compared with the other countries that have major infections, such as China and Italy.

3) How are you encoding the data visually?

    - During Scrolling:

      Map Chart: How COVID-19 spread geographically in global wise

      The world map is to encode the location of the countries. Each country has a circle to indicate more information. The size of the circle is to encode the total confirmed cases and the color represents death rate (death cases/total confirmed cases). The time is encoded as a slider that links all parts of visualization together.

      Line Chart: How health care system responded to COVID-19 in China and Italy

      The line chart shows the ratio of recorded deaths versus recorded recoveries (recorded deaths / recorded recoveries), on Y axis, changing over time, on X axis. Specifically, the X axis represents the number of days after the first 3

recovery cases found. China and Italy will serve as the benchmarks to compare with the selected third country.

- After Scrolling:

  Map Chart: The location of the specific chosen country

  The user can choose a country to view the detailed information by following steps:
  1) select the country on the map, the highlighted color will be used to encode the selected country
  2) hover on the circle to get the statistical information, e.g. accumulated confirmed cases, death and recovered cases
  3) change the time slider to see the overall trend of the selected country.

  Line Chart: Compare a third country with China and Italy

  Changing of the deaths/recoveries ratio of the selected country will be shown by a new red line, with the grey lines of China and Italy as comparable benchmarks.

4) Why is your solution effective?

- **Map Chart**:

  The map can show how COVID-19 has been spread geographically over time.

  - For the encoding of the map chart, using the size of circles to encode the number of confirmed cases can help the audiences see the spreading pattern in a particular country more obviously. To illustrate, the circle in China is bigger when the time is changing. In addition, the size of the circles can somehow reflect the population size in that country.
  - By employing a world map, not only are the viewers able to indicate the location of a specific country, but they also can see the international spreading pattern of COVID-19.
  - Furthermore, the color range of the circles is utilized to indicate the death rate. Since we considered that, in the map, the number of confirmed cases is more important than the death rate, and the death rate varies a little overtime compared to the changing of the confirmed cases. Hence, we decided to display the number of confirmed cases in a more conspicuous way.

- **Line Chart**:

  The line chart can show the trend of deaths/recoveries ratio over time.

  - We used the ratio of recorded deaths to recorded recoveries to show how well the medical system in each country is working. Although the number of recorded recoveries might not be as reliable as the recorded deaths since some people possibly have recovered without being identified, this ratio is at least able to explain whether enough tests and medical services are provided to the patients.
  - Relatively small number of recorded deaths compared to recorded recoveries implies a great number of tests, an effective medical system functioning well, and sufficient other paid-off efforts to save people's lives. On the other hand, a high ratio means not enough tests (to find out recoveries) and a hard time for the local medical system.
  - There is also a threshold here. The line only shows the ratio after the first three recoveries to reduce the impact of initial sharp fluctuations when the overall number is small.

5) How are you using text to support your visualization? Do you have any narrative structure?

- We created a text box with a relevant picture for some particular dates to enhance the viewers to link that major event with our visualization.

- In the tooltip, we used text to provide more detailed information, i.e. accumulated confirmed cases, death and recovered cases, for the selected country at a specific time as a complementary to the encoded circle.

- On the panel, we used text to explain how the ratio of recorded deaths to recorded recoveries (death/recovered) would reflect the performance of medical systems in a specific country.

6) How are you using interactivity (if at all)? Why does it support your task?

- **Select:** in the second phase of our visualization, the viewer can select the country of interest and the system would zoom in to that particular country. And the flow of deaths/recoveries ratio would be included and highlighted in the line chart.

- **Abstract/Elaborate:** the viewer can zoom in and out the map to focus on a selected country or see the trend of the world.

- **Connect:** In the first phase of our visualization, the viewer would scroll to change the date and go through the spread process of COVID-19. As the viewer scrolled, the date shown on the map, the handle position on the time slider, the content of popped up boxes, the line chart and the circle size and color would change accordingly. All data displayed are connected by the date which will help the viewers to understand the flow.

7) What are the limitations of your solution?

- The recovery variable in the database is confusing, since different countries (or regions) don't apply the same standard on tracking the number of recovered cases. According to Duszynski (2020) [1], "Recovery involves biology, epidemiology and a little bit of bureaucracy too".

- For example, to be admitted, the official number of recovered case in the US as CDC requirement, patients need to meet both medical and testing criteria, which are a) they must be fever-free without fever-reducing medications for three consecutive days and b) they must test negative for the coronavirus twice, with the tests taken at least 24 hours apart. [1]

- However, in the state of Michigan, "MDHHS is reviewing vital records statistics to identify any laboratory confirmed COVID-19 cases who are 30 days out from their onset of illness to represent recovery status" [2], and the number is updated weekly, not daily.

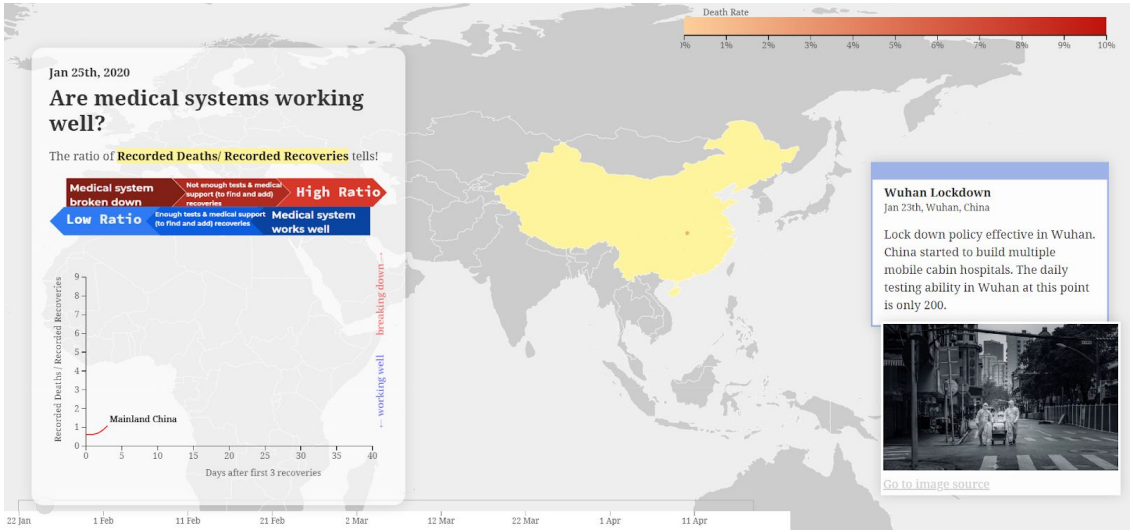
- Therefore, by shifting the idea, we look at the insights of how well the medical system in each country responds to COVID-19 from death/recovery ratio rather than simply compare the ratios between different countries. Since it could be biased.

Implementation / screenshots

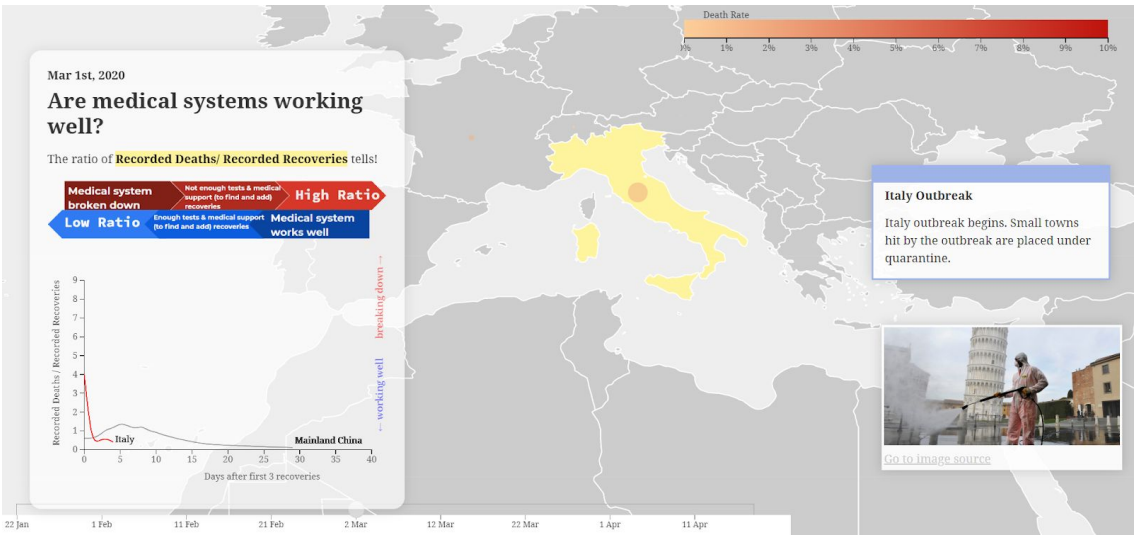
Link to our implementation: https://theoliao1998.github.io/si649-final-vis/
Link to the code: https://github.com/theoliao1998/si649-final-vis

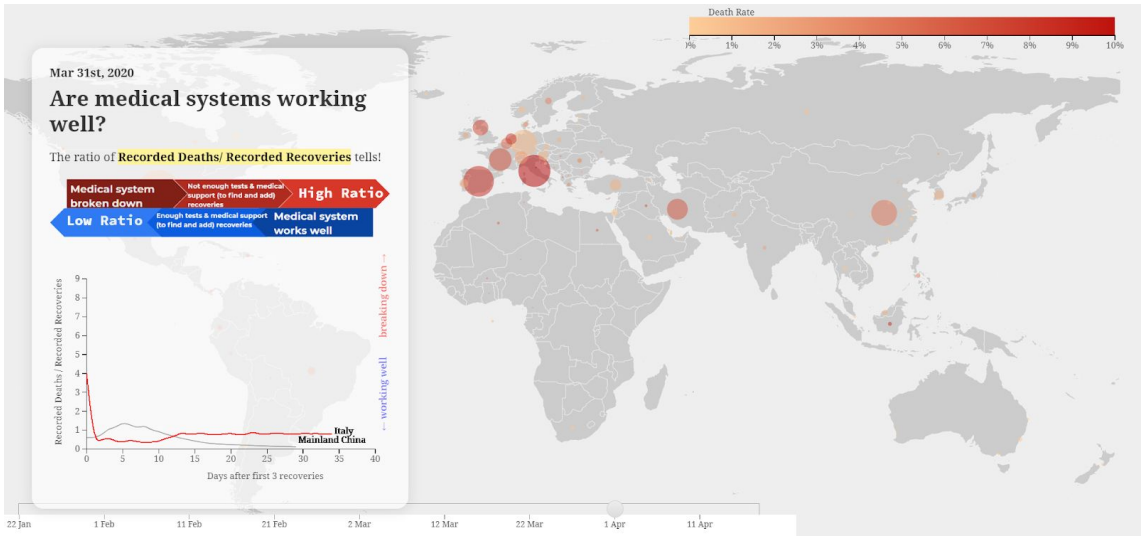
- Phase I: The Spreading Process of COVID-19
    - This visualization is divided into two parts, the first part is author controlled, the system would walk the viewer through the spread process of COVID-19. After entering this page, the viewer would scroll to start the journey. As the viewer scrolled, the date would change accordingly from January 22nd to April 18th.

    - The viewer would be led to China first. The major part of this visualization is a map. There are circles on the map, circle sizes represent the number of accumulated confirmed cases, and the color is encoded as death rate, as shown in this legend. There is also a line chart on this panel which shows how the ratio of recorded deaths to recorded recoveries change overtime. The reason why we think this ratio important is explained in the Questionnaire section. There are also some popped out boxes of important events or news with some corresponding pictures to provide more context to the visualization.



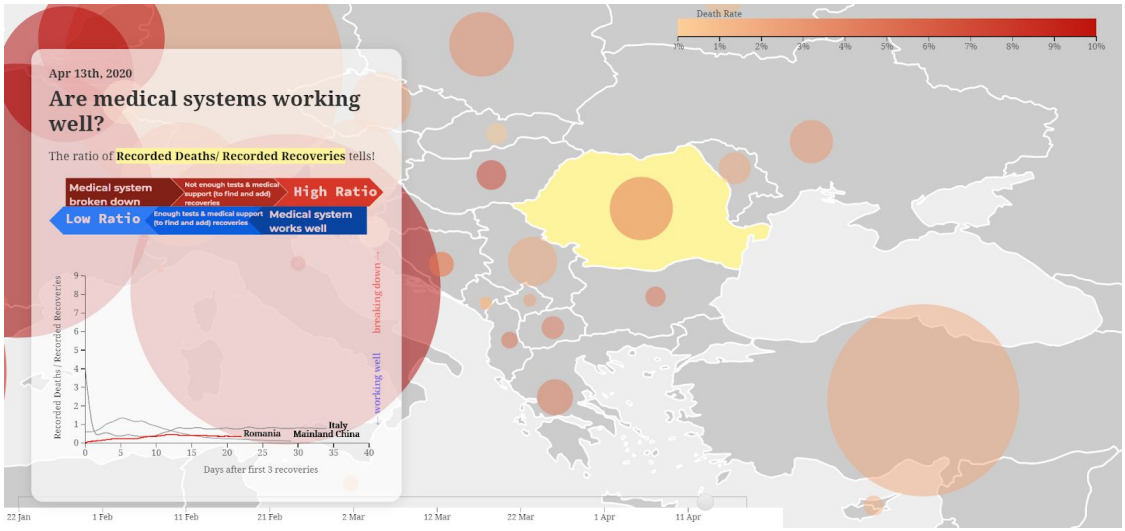
- then the viewer would be led to Italy when the outbreak starts there. After that the viewer would return to the world map.





- Phase II: Viewer Interaction (selecting the country of interest and comparing with China and Italy)

    - This part is more viewer controlled. The viewer can choose the country of interest and click on it. The selected country would be highlighted and zoomed in. By interacting with this time slider, the viewer can see how the virus spread in this country as changing of the circle size . On the line chart, the recorded deaths / recorded recoveries ratio of the selected country is displayed as a red line. The line of Italy and China would be kept as benchmarks. If the viewer hover over the circles, some detailed information would be displayed in the tooltip. The viewer can zoom in and out by scrolling.



Sketches / Experiments

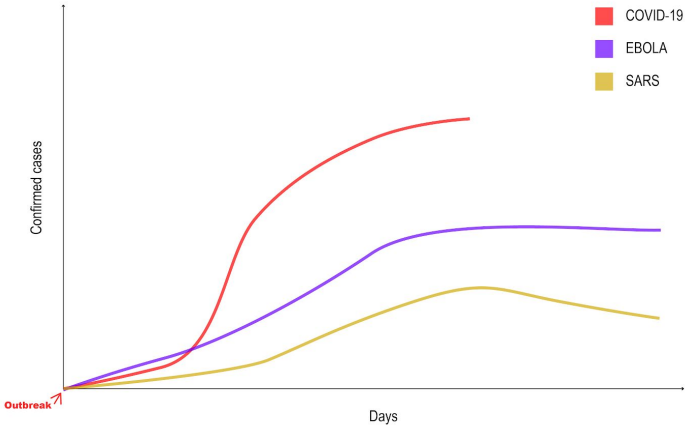
The sketches we provided here are the sketches we created in the mockup phase.

- Explanation

  In the mockup stage, we came up with three ideas and tried to connect them with a time sequence. But later, we found the three ideas are too scattered, and we decided to focus on the last two ideas. One is the performance of medical systems, and another one is a comparison between different countries. We used the ratio of deaths to recoveries to reflect the efficiency of the medical systems instead of the diagnosed time due to database limitations.

- Mockup Part I: Comparison with the previous pandemics.

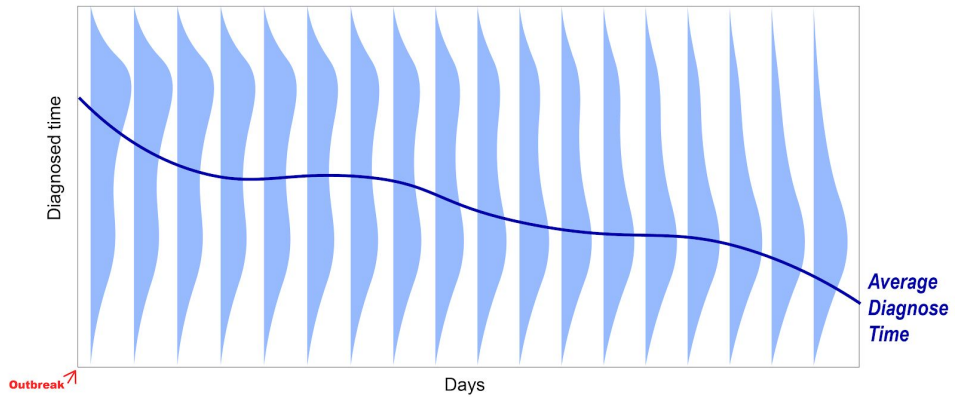
  As there were several outbreaks have occurred since in the past, i.e. SARS, Ebola as well as COVID-19, which are all caused by novel respiratory pathogens. We, therefore, would like to use the overall trends of SARS and EBOLA to compare with the trends of COVID-19 in terms of the confirmed cases as shown in the line chart.



  We firstly cleaned the date type for all dataset, and converted the date to days, because different dataset uses different time scales. The X axis will be ‘days’ and the Y axis will be the cumulative number of confirmed cases. We also used different colors to represent different viruses.

- Mockup Part II: Current Progress

  To see current progress of COVID-19, we calculated the diagnosed time by subtracting the hospital visiting date from the confirmed date. This was to show the distribution of days taken for diagnosis and diagnosed time changes over time.

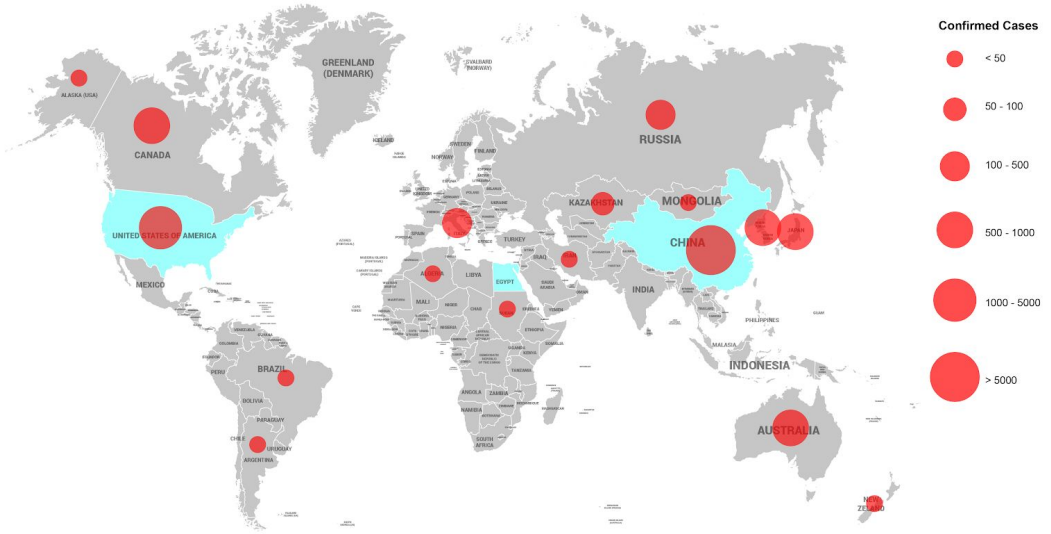


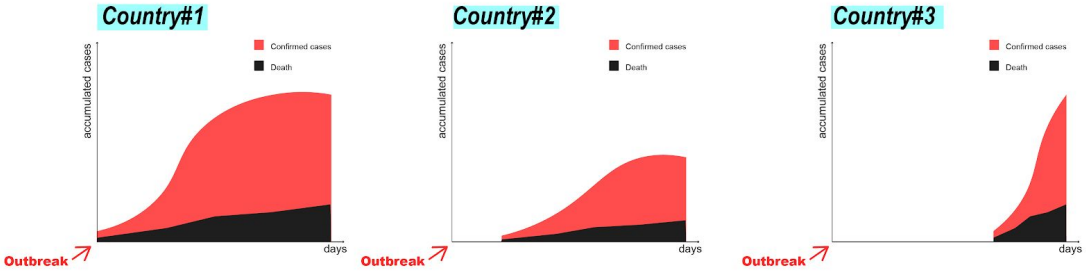
  We used a line chart containing vertical curvy lines to show the distribution of days taken for diagnosis (Each vertical curvy line is a summary of 5 days). Then, we added a dark blue line in the middle to show the overall average of the diagnosed time.

- Mockup Part III: Future Patterns

  After the analysis above, we wonder if it is possible to explore the COVID-19 patterns of the other countries, and predict the future pattern in specific countries. For example, from countries that are now able to control COVID-19 like China, Japan and South Korea, we can assume the possible virus life cycle. While the other countries that are still undergoing the serious virus situation, e.g. US., we can assume the possible prediction in foreseen next week, and validate it soon.

  To fulfill this communication goal, we will use a map chart to display the geographical distribution of virus spreading, and use circle size to show the number of confirmed cases. Users can select regions they are interested to see detailed information. The detailed information will be shown as an area chart, showing the confirmed cases and death cases in different colors. The selected regions will turn blue on the map.


Confirmed Cases
< 50
50 - 100
100 - 500
500 - 1000
1000 - 5000
> 5000
GREENLAND (DENMARK)
CANADA
RUSSIA
KAZAKHSTAN
MONGOLIA
CHINA
UNITED STATES OF AMERICA
MEXICO
INDIA
BRAZIL
INDONESIA
AUSTRALIA


Country#1
Confirmed cases
Death
accumulated cases
days
Outbreak
Country#2
Confirmed cases
Death
accumulated cases
days
Outbreak
Country#3
Confirmed cases
Death
accumulated cases
days
Outbreak

Reference

[1] https://theconversation.com/what-does-recovered-from-coronavirus-mean-4-questions-answered-about-how-some-survive-and-what-happens-next-134883

[2] https://www.michigan.gov/

[3] New York Times COVID-19 visualization https://www.nytimes.com/interactive/2020/03/22/world/coronavirus-spread.html?from=timeline&isappinstalled=0