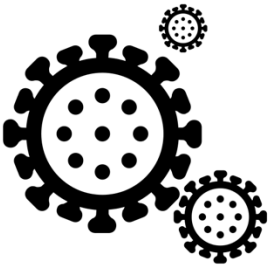


The SICK- COVID-19

Members:
Ameen Alrehn, Shukran Amdeen, Shuyang Ding, Shih-Yuan Yen, Shukran Amdeen, Syed Qavi, Yi-Hsuan Shih

Instructor:
Dr. Eli T. Brown

Table of Contents:

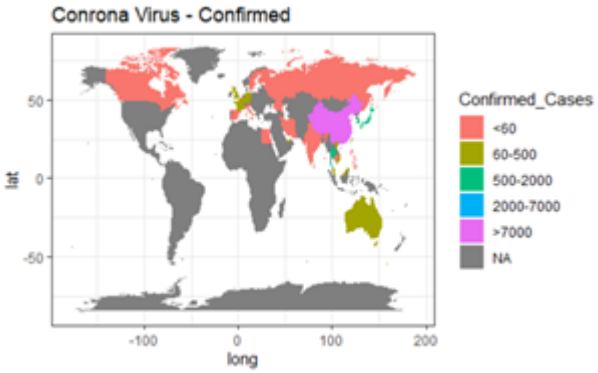
Introduction

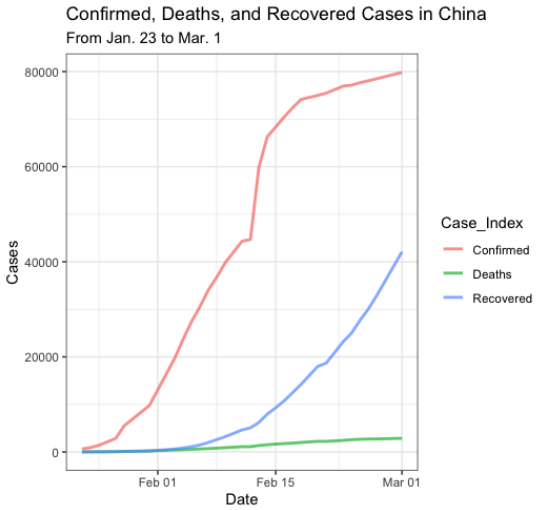
There's a high infectiousness virus that comes from Wuhan, China which is called COVID-19, also known as coronavirus. It is a family of viruses that can cause diseases ranging from the common cold to serious respiratory symptoms, such as SARS or MERS. It started in December 2019. China's government claims that this virus was attached to bats originally; however, some citizens tried to taste the flavor of game meat, and then the virus starts to transmit to people. According to the WHO, the COVID-19 estimates of the incubation period is 14 days. For most infected people, the symptom will show within five to six days. The WHO also states several infected people can be asymptomatic, and it means they will not display any symptoms despite they are the carrier. Finally, the WHO raises the alarm and declares the crisis pandemic on March 11, 2020.

In order to figure out what the influences are to each country, our team collected the data of COVID-19 which comes from Kaggle. The data contains worldwide cases, and we extracted and chose some to be the variables. They are deaths, confirmed and recovered cases, country, provinces, date (1/22/200 - 3/1/2020), gender, age, visiting_Wuhan, from_Wuhan. We conducted those data to do the visualization with maps, bar chart, line plot, scatterplot, circular, violin plot, and mosaic plot. We compared confirmed, death, recovered cases by using bar chart and line plot. The scatter plot gif for showing the rate of spread. For displaying the level of severity, we chose a map to express it on China's provinces and other courtiers and regions. We also made a violin plot and mosaic plot to find out the relationship between gender and age of the confirmed and dead cases. Furthermore, our team compared some historical diseases like SARS and H1N1. This paper explores what we did to kick off our project, what data we chose and visualization we did while the data research process, and how we created our graphs.

Exploratory Analysis

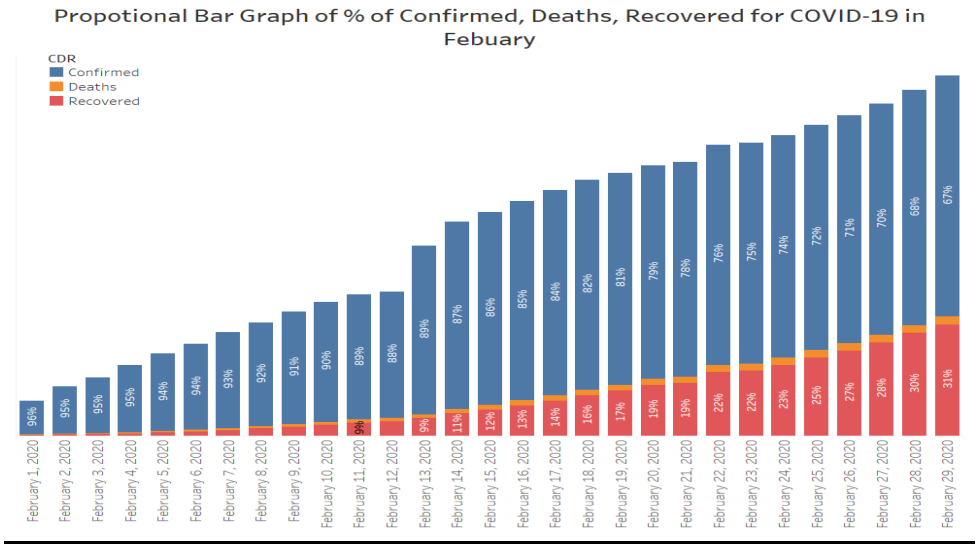
Our group wants to dig in more about COVID-19, which is becoming very serious recently. As the virus spread worldwide, we think plot confirmed, deaths and recovered cases on the world map will be a good way to show how many countries have been infected and how many are still safe so far. As COVID-19 originated from Wuhan, China, which had a much higher amount among all three situations as of 3/2/2020 than other countries and cities do, we think plotting China and world map separately may be a better way to show. Therefore, we mapped the world map and China map and used different colors to represent the cases. From these 2 maps, all colored areas indicate that province/country has COVID-19 confirmed, deaths or recovered. These graphs meet our goal to show cases geographically, but they are not clear to read which area has more cases and to compare. Below are the two example graphs for confirmed cases, we also plotted same thing for death and recovered cases:





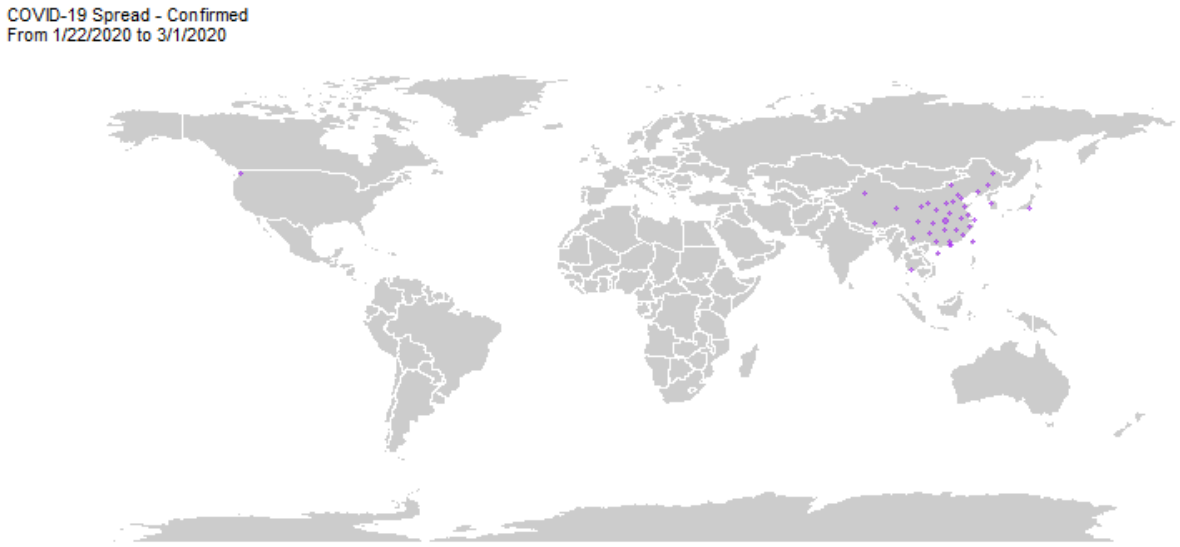
This line graph shows the three types of cases specifically in China. But later on we found out that it would be clearer to compare those with different countries globally.

The variables used for this visualization were confirmed, dead and recovered. We wanted to see the growth rate of these three variables over time. Two datasets COVID-19 and SARS were used to complete the visuals for this exploratory visual. The proportional bar chart was chosen to look at the distribution of growth of these variables. At first a normal proportional bar chart looking at the population growth of confirmed, death and recovered was used as the y variable. Then there was an idea to use percentage as the y variable and see the difference between the percentage growth over time for both SARS and COVID-19 since we were already looking at growth rate in line, scatter and map graphs. But a proportional bar chart for this dataset was a little misleading since confirmed cases included both recovered and death cases so measuring all three plotted on top of each other would not be a good idea and plotting them side by side would just clutter up the graph. This led us to think of a way to measure them separately using another type of graph.



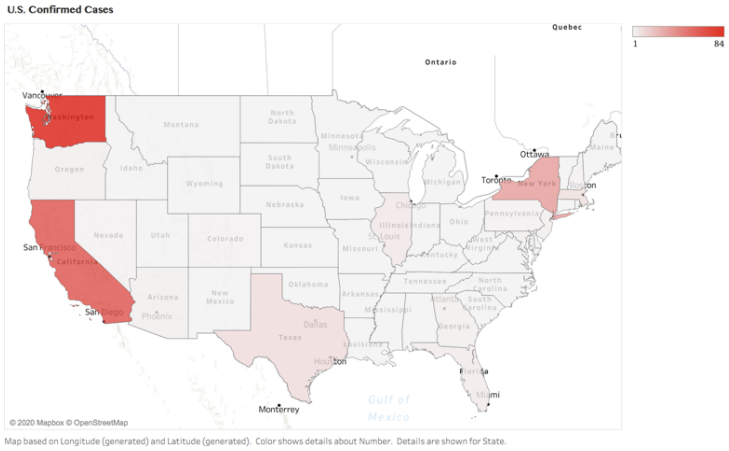
Visualizations

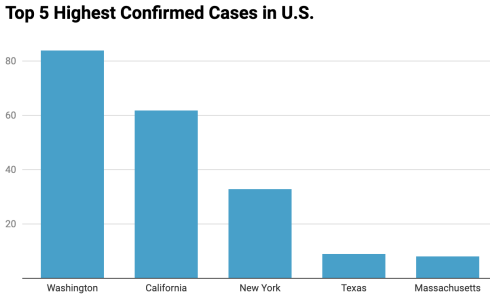
Visualization #1: Map showing the spread of COVID-19 throughout the globe



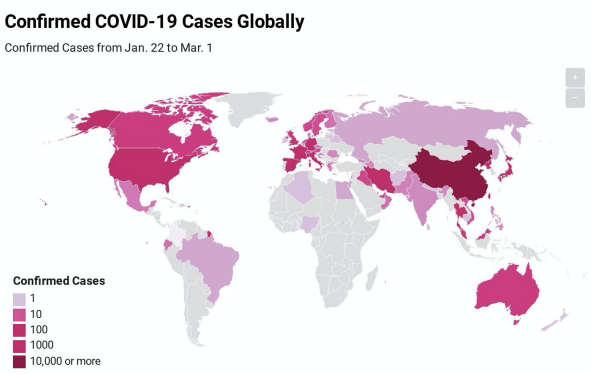
As we would like a visualization to show how COVID-19 confirmed cases spreading and growing throughout the world, we chose to use the world map and dots with increasing size in applied area to reflect both location and visible growth rate. The original dataset has 10 variables including observation date, province, country, latitude, longitude, last update date, confirmed, deaths and recovered. Because confirmed numbers are cumulative amounts until that observation date for every location(time frame of observation date is from 1/22/2020 to 3/2/2020), we modified the original dataset by grouping country and province/state and picked max number from each group. The variables used to map the plot are longitude, latitude, confirmed case amount, country and province/state. At first, we used default size setting from R to plot dots. However, as the virus originated from Wuhan, China, and the growth was an exponential growth, Wuhan had much higher confirmed cases than other locations, making dots that represent other cities can barely be seen. To fix this, we separated confirmed cases into 5 groups and adjusted scale size, so even the countries with 1 or 2 cases can be seen clearly from the map. Then, we decided to use grey background and light-colored dot to make those tiny dots more outstanding. Also, we have decided to remove axis labels and legends as these are not informative and distracting. After basic theme has been settled, we added animation effect to make those dots spreading and growing in the map. From this map, we can see how COVID-19 grow and spread all over the world on daily basis. The virus breakout in China, and then went to other Asian countries, some of them landed in European countries, US and Australia. The last couple seconds of gif shows the virus growing rapidly in

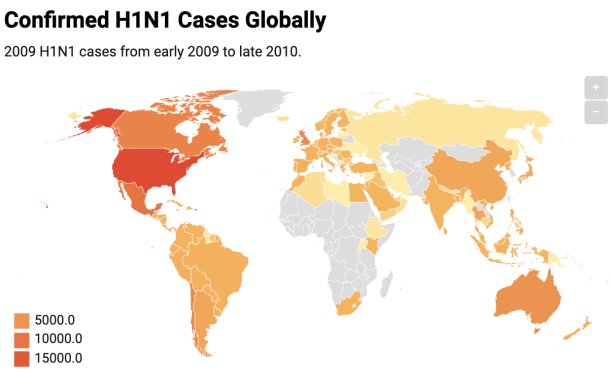
European countries, indicating there's potential outbreak around that area. As of date, from the news, we all know that Italy and Spain have locked down, and the EU has announced travel restrictions to most of the European countries except the UK.



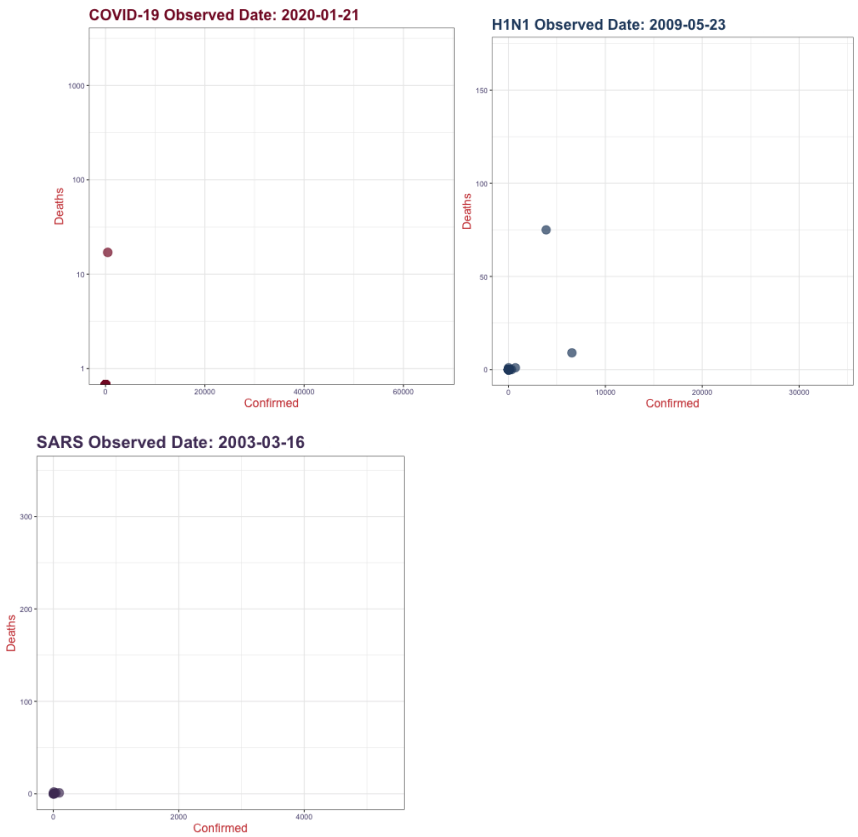


This U.S. map shows the spreads of the Confirmed Cases of COVID-19 over the states. The bar chart next to the map explicitly shows the top 5 highest states that have been tested for the highest number of confirmed cases.



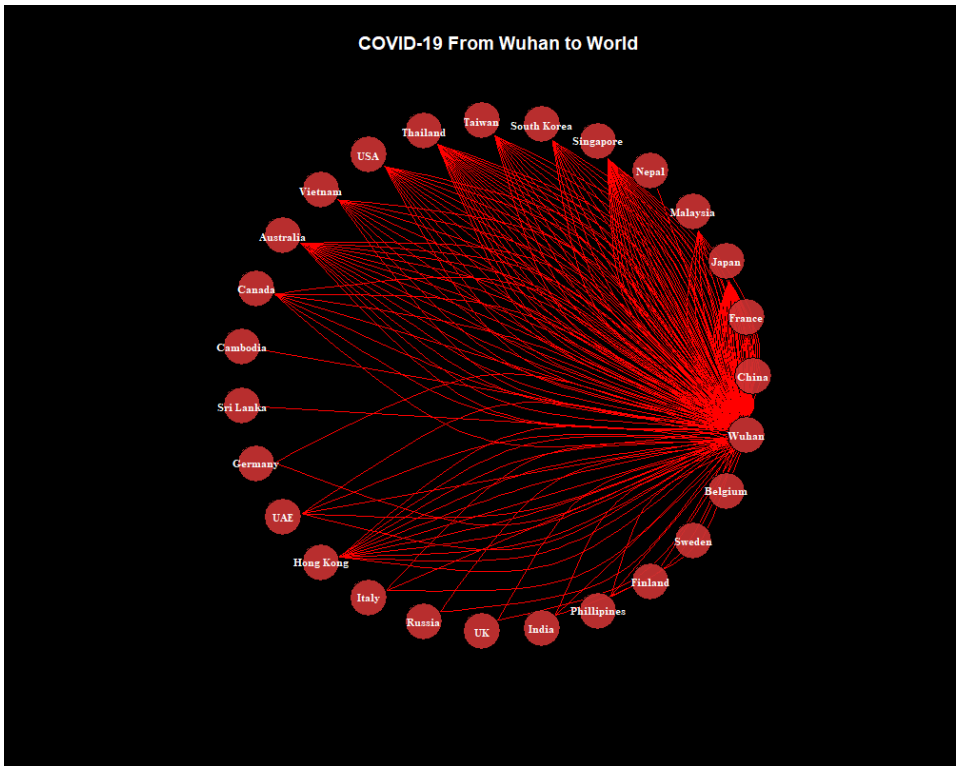


These two choropleths show the confirmed cases globally of COVID-19 and H1N1. On the left is the spreads ofCOVID-19 which we can clearly see that China is the most serious place being affected, while on the left we can see H1N1 affected more seriously in the America areas, especially the U.S. and Canada.



These three interactive charts show the spreads of three different viruses over time. Each dot represents a country, so for instance in COVID-19 graph, the one that stands up uniquely and grows much faster than all the others is apparently China. The date on the top change day by day and the dots show growth increased over the time. The three charts show the difference between the growth of the three viruses. We can observe that for COVID-19, there was one country heavily affected by the virus, which was China, the country where the virus originated, and other countries are still in the beginning stage. But for H1N1 and SARS, we can see multiple countries affected by the virus and, for the last few days of the graph, we see more countries are affected by the virus. This is what differentiated COVID-19 to the other viruses that one country (China) was much ahead of other countries in terms of deaths and confirmed.

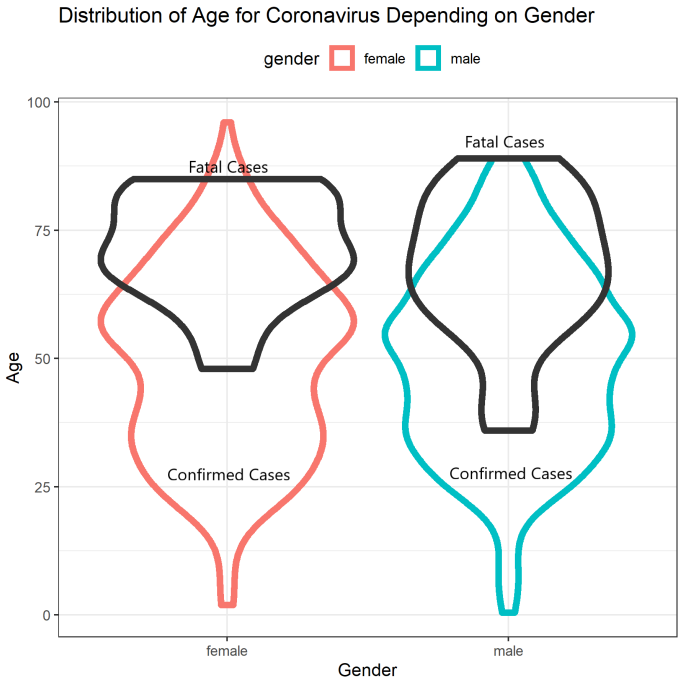
Visualization #2: Network graph showing how the virus was carried from Wuhan



In addition to see how COVID-19 spread and grow worldwide, we also want to visualize to which countries the virus was carried from Wuhan. Therefore, we chose a network graph to present. The variables used to plot network graph are people from Wuhan, people visited Wuhan and country. We have tried to plot network graph in both R and Gephi. The one from R has very density area between Wuhan and China as most of the cases stayed in China. In order to make the dense area viewable, we tried to plot the graph in Gephi with same dataset so we can drag the plot. However, Gephi graphs edges on weight, therefore, for countries with few and more than few confirmed cases, the size of edges have slight difference, making audience hard to read. After comparing two graphs, we went for the one from R.

From the graph, we can tell that most of the confirmed cases are located in China, matching the fact that China had exponential growth for the time period from 1/22/2020 to 3/2/2020. The second most affected countries are Japan and some other Asian countries, who are close to China. From the graph, we can also find the fact that European countries like France, Sweden and Belgium claimed the second place after Aisan countries, which matches the observation from the gif map. Additionally, there are also some cases landed in Western countries like the USA and Canada and also some middle eastern countries like UAE and Sri Lanka.

Visualization #3: Violin plot showing distribution of age by gender



Looking at how COVID-19 spreads is integral when looking to analyze the virus. However, we also wanted to look at any possible patterns that might occur in female versus male patients. The violin plot's goal is two-fold. First, to identify the distributions of age for all cases based on gender. Second, to compare the fatality patterns between the genders.

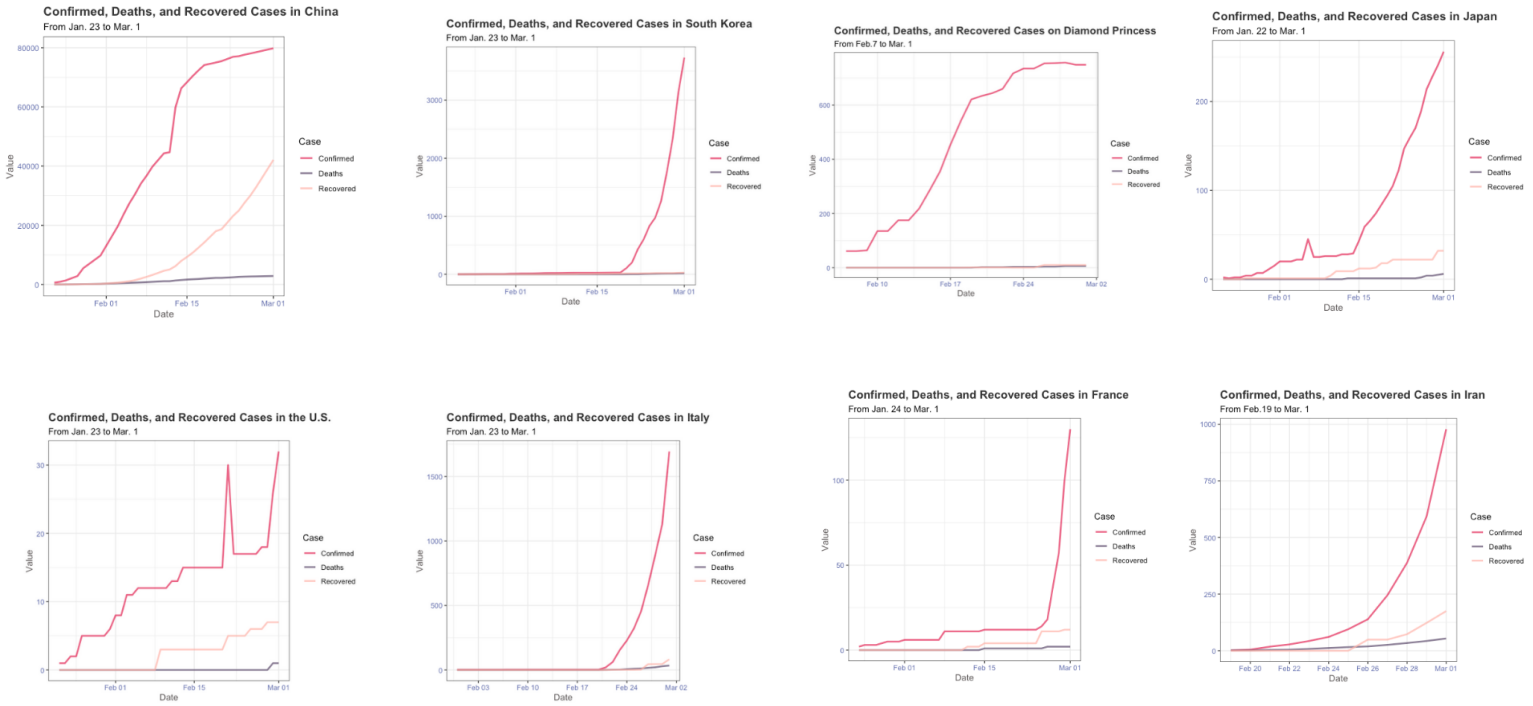
The colored violin plots in the back show the age distributions of male and female confirmed cases. The ranges of the age distributions for the confirmed cases are similar for both genders. While females have slightly older confirmed cases, men have slightly younger confirmed cases. The patterns for the confirmed cases are similar yet not similar enough. Compared to males, females are less likely to be diagnosed with the virus from ages 25 to 50. However, from age 50 to about 85, both males and females are just as likely to be diagnosed with the virus.

The fatal cases' age distributions are very interesting to look at when compared between the genders. It's clear that the age distributions between genders vary. Male's age distribution has a bigger range of fatal cases compared to female's. However, for individuals between the ages of 70 and 85, females are more likely to die due to the virus when compared to men. It's significant to note that when a confirmed case is classified as 'fatal', it does not mean that the other cases have

recovered. Considering how recent the data is, it's important to recognize that there are limitations of this visualization as well.

An artistic choice I made is to print the violin plot as outlines only to allow for clearer understanding for the user. Ensuring that the violin plots in the back that show the distribution of ages for the confirmed cases are preserved and can be analyzed once the fatal cases are stacked on top was an integral part of this visualization. Thus, the outline of the violins guarantee a better read of the data and send a clearer message.

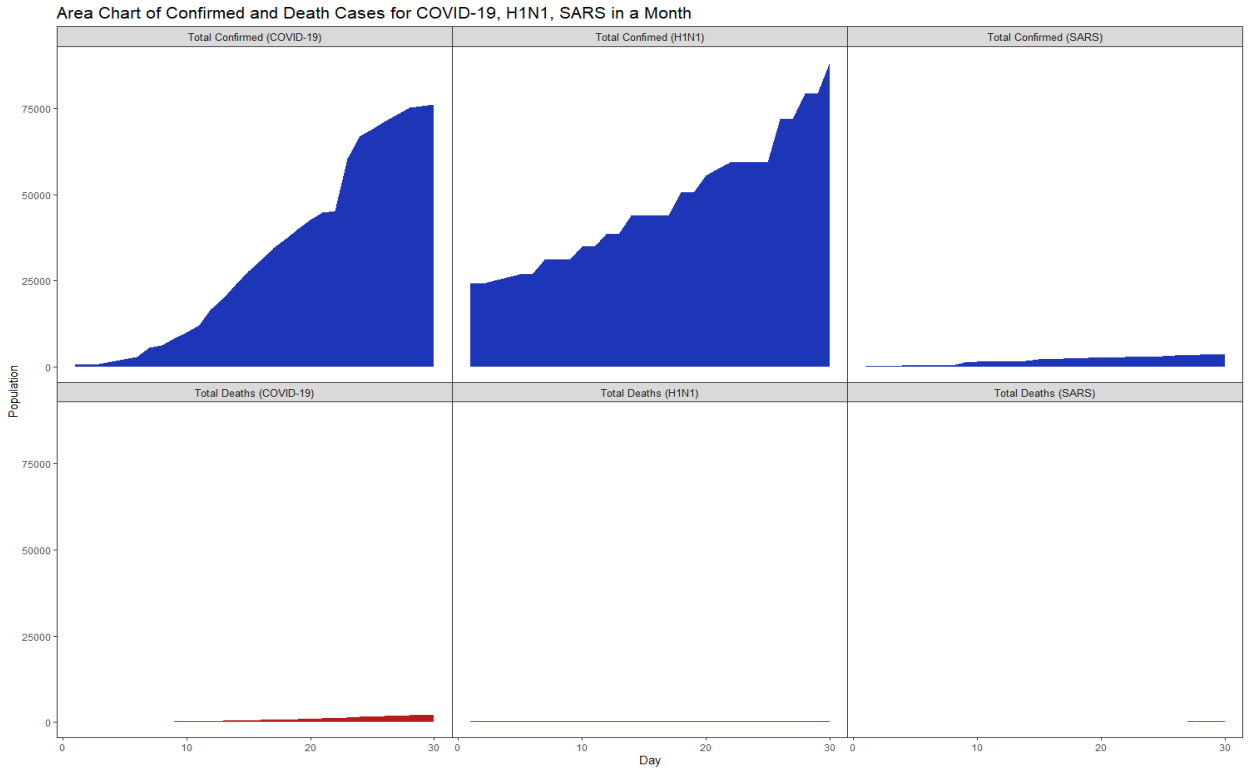
#Visualization 4: Line graph comparing the confirmed, death and recovered patterns between countries



These are combined line graphs comparing confirmed, death, recovered cases among countries. The variables include the date and the value of three categories. By these multiple line charts, it's clear to see that some countries' conditions have increased from the beginning of the outbreak while some remained fine not until the latter half of February. Previously, we created only one line graph to show the relationship among three types of cases over time, specifically in China. But later

on, I found out it would be interesting to compare the same thing with different countries.

#Visualization 5: Small Multiples Area Chart



This is a small multiple area chart that uses variables confirmed and deaths from COVID-19. H1N1 and SARS dataset and plots them using only the first 30 days of the dataset. Each column plots total confirmed and death cases for each of the three viruses side by side so it is easy to compare for the audience. The goal of this visual is to show the audience the difference between the growth of confirmed and deaths for COVID-19. H1N1, and SARS. Before attempting this visual, we went through the process of trying other types of graphs such as the proportional bar graph and scatter plot. Also, tried looking at an area chart that stacked confirmed and deaths cases but this graph was misleading since confirmed also included deaths. So, dividing the area graph individually for confirmed and deaths for three viruses in one plot was a much better choice. The colors such as blue for confirmed, and red for deaths were used to show differences clearly. R was used to make this graph, using the ggplot package. This graph showed that the increase in confirmed cases in COVID-19 was much steeper compared to H1N1 and SARS. Also COVID-19 had a higher death rate in a month compared to H1N1 and SARS.

Analysis and Discussions

We have a dataset showing confirmed, death, and recovered cumulative cases all over the world for COVID-19, H1N1, and SARS. This dataset would be updated on a daily basis for COVID-19. We found many things that could be done with visuals for the datasets. For example, we created choropleth maps showing the total of each of the cases for each country affected. By doing this, we first analyze and organize the data in excel as we only need the latest value (as it shows the cumulative value). Also, we compare the growth by dates from January to March, and we discovered that some countries were seriously affected at an early stage while some were not until almost the end of February. The network graph showed that countries closer to China like Japan, Malaysia, Singapore and many others were the countries where people from Wuhan traveled to but countries from Europe like France, Sweden and Belgium were the destination for a lot of people from Wuhan.

We used the H1N1 and SARS dataset to show cases across countries as well. We compare their distribution with COVID-19 and found that H1N1 was more serious in North and South America areas while SARS spread more seriously in Asia, specifically China, Hong Kong and Taiwan. We represent those comparisons by showing choropleths, interactive charts and area maps. We also measured growth rate over time of confirmed and death cases for the three viruses and we found that COVID-19 had more death cases in the early stage compared to the other viruses. We showed this with an area chart and interactive scatterplot. We also observed from the violin plot that more male were getting diagnosed by COVID-19 than females from ages 25-50 but from ages 50-85, male and female were equally likely to get diagnosed with COVID-19.

APPENDIX

I. Shuyang Ding – Individual Report

Our final project is to create various visualizations for COVID-19 dataset and to compare this dataset to SARS and H1N1. During the first phase, we listed some topics that we would like to express, and then plot graphs based on the topics, so we can have multiple options for each listed topic to pick from or combine different plots together to convey the topic in a better way. Not only limited to the topics we listed, as our dataset kept updating, we also created plots on other topics that we are interested to show.

The COVID-19 dataset includes six sheets. The first group is three sheets about death, confirm and recovered count respectively. In each of the sheets, there's variables as: Date, Country, Latitude, Longitude and Province. The date columns are from 1/22/2020 to 3/2/2020, and each column represents the cumulative number of cases reported from each country or region till that date. These are our key variables that could help us to see how COVID-19 spread and outbreak over time. In addition to the three-individual sheet, there's also a summary sheet that have confirmed, deaths and recovered count together. The second group is two sheets with all other detailed information like Age, Gender, City, Date_Confirmation, Symptoms, Visited_Wuhan , Lives_in _Wuhan, etc. These variables could help us figure how COVID-19 distributed differently on age and gender and what's common COVID-19 symptoms if applicable.

I created a gif map to show how COVID-19 spreading and growing in amount worldwide for confirmed, deaths and recovered(**SD1**) separately. Dots on the map represent either confirmed cases, deaths cases or recovered cases. The size of the dots implies the amount of cases. I used a grey map as a background with purple dots on top of it, so the audience could see those tiny dots more clearly. It's easy to see how viruses grow on a daily basis from timeline graphs, but a gif map could give the audience an idea of how viruses spread and grow geographically. In addition to that, I also plotted an interactive map showing detailed information like country, province and case amount when the viewer hovering mouse cursor over a plot(**SD2**). This plot is meant for those who would like to see where the viruses are on a map and summarized information.

I also plotted a network graph in both R (**SD3**) and Gephi(**SD4**) to show how viruses were carried from Wuhan to the world. After comparing 2 graphs, I went for the one from R as the graph from Gephi was not easy to read and compare. The dataset has 2 columns stating whether the person is from Wuhan and whether the person has visited Wuhan with no double counting. If a person is categorized from Wuhan, he/she will not be included in the Wuhan group. I used those 2 columns and confirmed country columns to build the network graph. Each line connected to Wuhan

represents one case. From the graph, it's obvious to see that most of the cases stayed in China, and some Asian countries, but there are some that went to Eastern countries and outbreak there.

From the project, I found the most difficult thing about data visualization is to determine what topic or content I want to show, and which visual technique is the most appropriate way to realize my idea. A good visualization needs a lot of revision. For example, to make the gif map clear to be seen, I have tried different color and size combinations for both background maps and dots that represent cases. Also, legend and axis labels are another important element to support information I want to convey.

II. Syed Qavi - Individual Report

In the group, my role was to help the group in any way I could. If the group needed visuals for the exploratory or the final explanatory part, I would create one and see if it was helpful to the project. I was responsible for creating a visual which looked at the confirmed, recovered, and death cases for COVID-19, H1N1, and SARS. I created two proportional bar charts as an exploratory visual (SQ1, SQ2). The proportional bar chart looked at the distribution of recovered, death, and confirmed cases of COVID-19. I changed the bar chart to look at the percentage difference between confirmed, death and recovered cases over time for both COVID-19 and SARS as an explanatory visual to compare the distribution but confirmed included both death and recovered cases so this was a misleading visual.

I also created a small multiple area chart as an explanatory visual. This visual compared the death and confirmed cases of COVID-19, H1N1 or swine flu and SARS for the first 30 days of the dataset. I wanted to create a visual that would look at all three side by side and compare which would allow the audience to see how the confirmed and death cases progressed over time. I also tried many other visuals which didn't do a good job explaining the dataset so I decided not to use them. I also looked up additional datasets that would help add to the final project specifically the H1N1 dataset.

I learned a lot about data visualization in the final project. I learned how to use Tableau and R from watching and reading the tutorials made by professor and other outside sources. From Tableau, I learned how to manipulate data, how to join two or more datasets and use them to create visuals. I also learned how to do the table calculations. In R, I was able to use ggplot2 extensively to create many different visuals covered in the course. Since our dataset was limited in terms of variables, I had to try many different graphs with the same variables to see which would give the

best results. Also, I learned to use libraries such as tidyverse, plotly and other libraries used to manipulate date, time and other interactive libraries.

III. Yi-Hsuan Shih Individual Report

In this group, I was responsible for creating several charts including line graphs, maps, bar charts and interactive graphs. I did the combined line graphs(YS1) comparing confirmed, death, and recovered cases all over the world. The line graphs are more comprehensive to show the growing trends over time. I also made choropleths (YS2)comparing different diseases across countries, like COVID-19 and H1N1. For the COVID-19 choropleth, I have to set the value range in legend on my own as China has far more cases than all the other countries. I cannot simply put the data and let Tableau automatically make the map, or the map would look ineffective like China would be the only country that has the darkest color and all other countries are the same color. I also made a map(YS3) to see the overall spreads in the U.S., with the support of a bar chart (YS4) to see which are the top 5 highest confirmed cases in the nation. To put a map along with a bar chart is easier for audience to compare and view the difference. I also made the animation (YS5)comparing the growth of three different diseases. I use R to build the line graphs and interactive charts among different viruses.

In this final project, I learn a lot regarding using R to program visualizations and Tableau to simply create visuals. This final project also gives me the opportunity to do more research on making visualizations in a most effective way. Professor's tutorials both text and video parts are pretty helpful, with only a few lines that I could make a visual that could tell a story.

IV. Ameen Alrehn - Individual Report

In the group, I was responsible for creating multiple charts such as bar charts and butterfly chart. I focused on these charts on other countries other than China to show the impact of Covid-19 on the other countries. I made a bar chart(AA-1) using tableau that focuses on the confirmed cases in the top 20 countries excluding China to make the confirmed cases in different countries more apparent. I have also made a butterfly chart(AA-2) that looks at the age and gender distribution in the top 20 countries.

The charts I made looked at how countries other than China had a big impact from the virus and the ages and gender of people with confirmed cases. The charts I made, we did not end up using them because we had a lot of other charts and we filtered them to pick the best ones out of them.

In the final project, I have learned how to freely look at the data, understand it and use an appropriate chart to display to the audience what the message I want to convey. I have learned how to research and learn more about the visualizations and how to use them to tell a story. I found the professor's recorded tutorials and texts to be helpful in creating the final project. I have also learned how to effectively use tableau such as creating table calculations and how to make charts that needed some additional steps to make such as butterfly chart.

V. Shukran Amdeen - Individual report:

I wanted to focus on any patterns that might exist between gender and the Coronavirus. The data used for this visualization include all cases documented within the COVID19_line_list_data.csv file within the Novel Corona Virus 2019 Dataset from Kaggle. This file includes more data about each case documented within the main covid_19_data.csv file. While not all observations have a defined gender, more than 1,000 cases have an observed gender.

At first, I was trying to build an interactive dashboard on Tableau. The goal of this dashboard was to communicate how genders impact fatality rates, confirmed cases and recovery rates. However, The boxplots I built on Tableau did not communicate the message I was aiming for successfully. I built several boxplots to visualize the distribution of fatal cases and confirmed cases based on genders (SA3). However, knowing that the message wasn't clear in this dashboard, I looked into building violin plots in R.

I had to create 2 separate datasets for the fatal cases and another for the confirmed cases. I built a python script, *clean.py*, to scroll through every observation and determine if it's a fatality or simply, a confirmed case. I created a separate text file, *fatalities.txt*, that contained 58 fatal cases out of the original dataset. Using R, I built two separate violin plots. The first was a violin plot showing the distribution of age depending on gender for fatal cases (SA1). The second was a violin plot showing the distribution of age depending on gender for confirmed cases (SA2). After consolidating the two violin plots into a single plot, I color coded the mappings for a clearer message. Considering that I couldn't find a way to differentiate between the confirmed case mappings and the fatal case mappings, I used a simple Paint application to add labels.

When approaching this project, I expected to build an intricate yet simple visualization that would perfectly show a relationship that exists in our dataset. However, I quickly learned that our data didn't provide enough information for me to find a 'hidden' relationship that could exist between the variables. The biggest lesson I learned throughout this project is that visualizations ought to communicate a clear message from the data. However, the message is rarely determined before the data exploration stage.

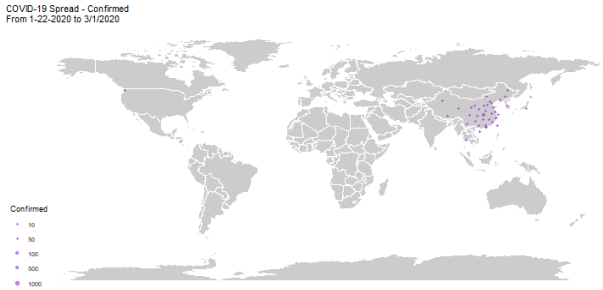
VI. Shih-Yuan Yen individual report:

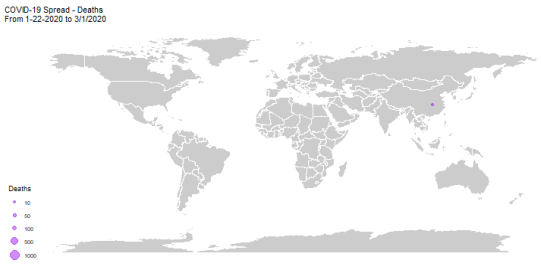
In the group, I made some graphs based on a team's need such as the mosaic plot and butterfly chart to display the connection between gender, age, visiting_Wuhan, and from_Wuhan. The mosaic plot (SY1) is for showing the deaths of the relationship between gender, visiting_Wuhan, and from_Wuhan. The second mosaic plot (SY2) explains the gender and age of the deaths, as we can see the most infected people is male, and 60-80 years old people have a high risk of death. I chose visiting_Wuhan and death to be the variables in order to show worldwide cases by using the butterfly chart (SY3), and it shows China has a notable bar chart. For those graphs, I also tried other ways to display the data; however, those graphs cannot be explored with more information or have no meaning.

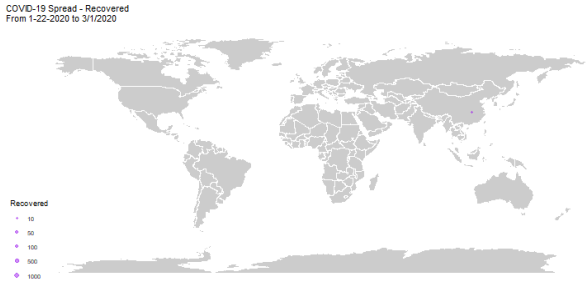
After 10 weeks, I learned how to distinguish different data visualization and know how to use data to build graphs by using Tableau. Also, R is still a challenge for me because I think it has to spend much time practicing in order to display the data well. Just like other computer languages, it takes time. Although R is a new language for me, Prosser Brown's materials for R is very helpful. As an HCI student, I do some research, such as an interview or card sort, etc. This class teaches me how to use the data more appropriately and display an excellent graph to analyze data. It will be a great skill for my future work.

Plots and Graphics:

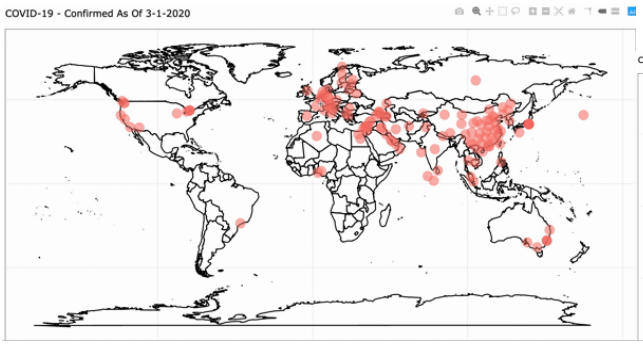
SD1 :

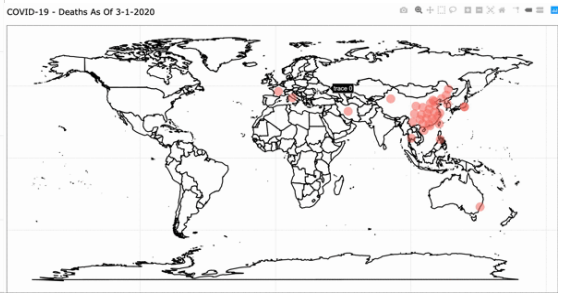


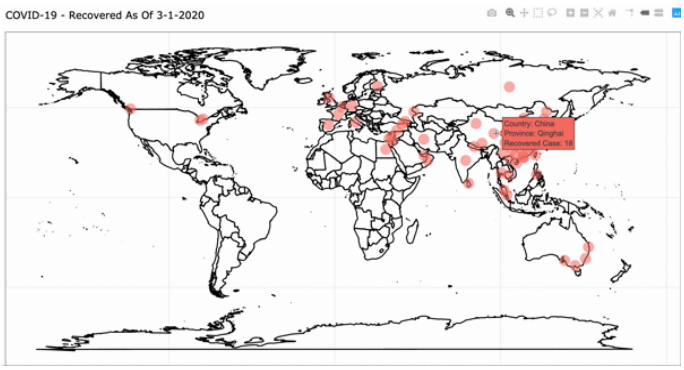




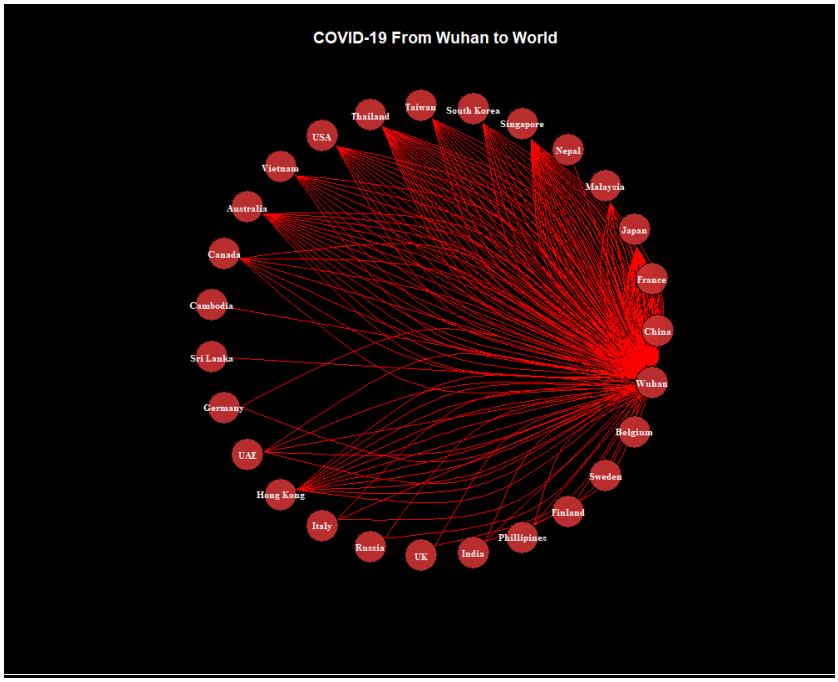
SD2:



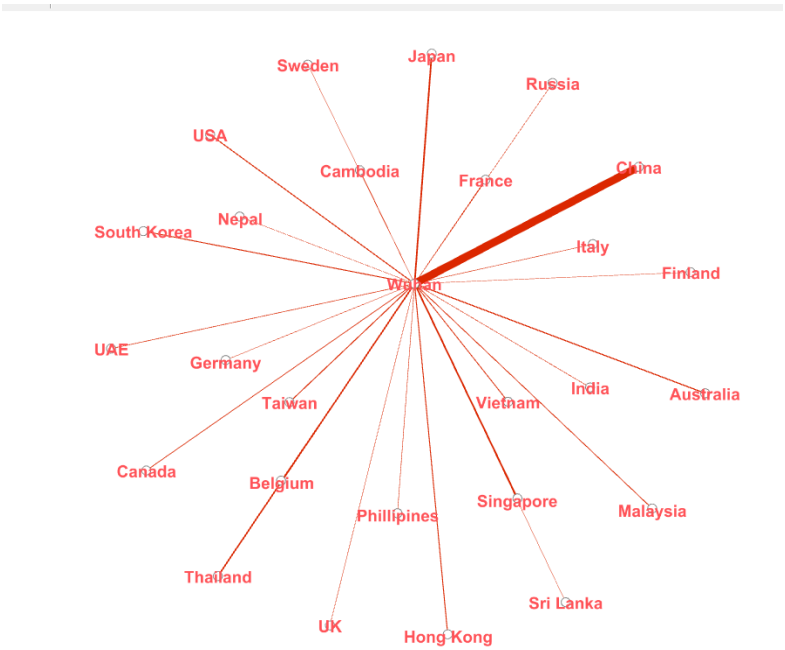



COVID-19 - Recovered As Of 3-1-2020
Country: China
Province: Qinghai
Recovered Case: 18

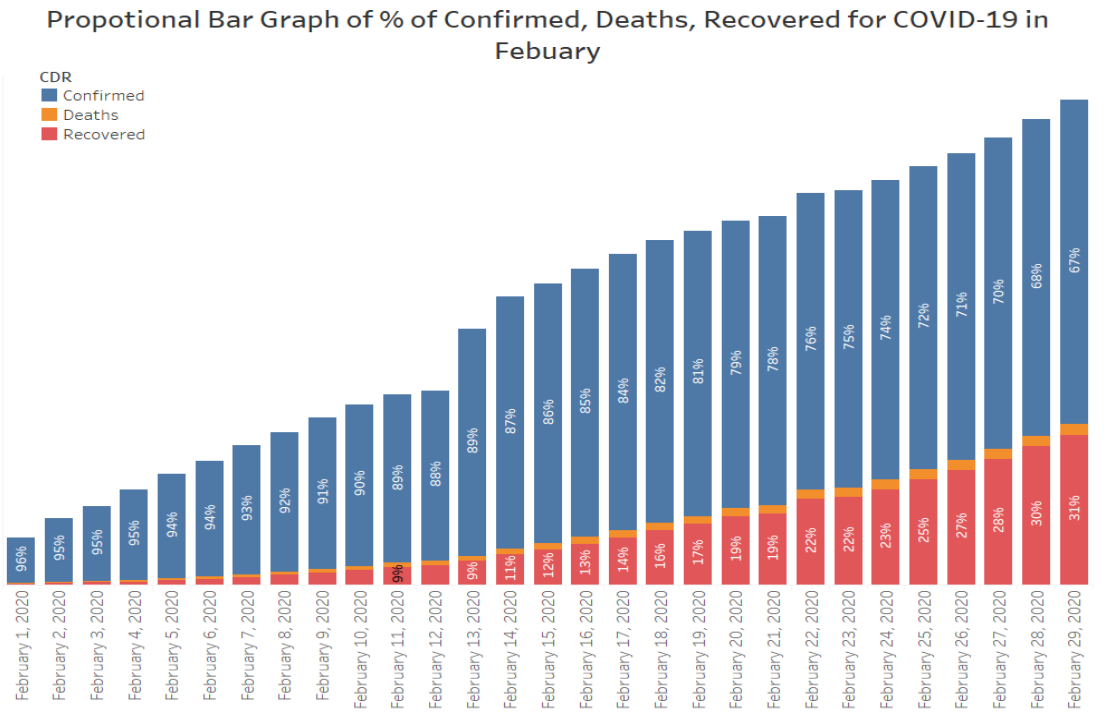
SD3:


COVID-19 From Wuhan to World
Thailand
Taiwan
South Korea
Singapore
USA
Nepal
Vietnam
Malaysia
Australia
Japan
Canada
France
Cambodia
China
Sri Lanka
Wuhan
Germany
Belgium
UAE
Sweden
Hong Kong
Finland
Italy
Phillipines
Russia
UK
India

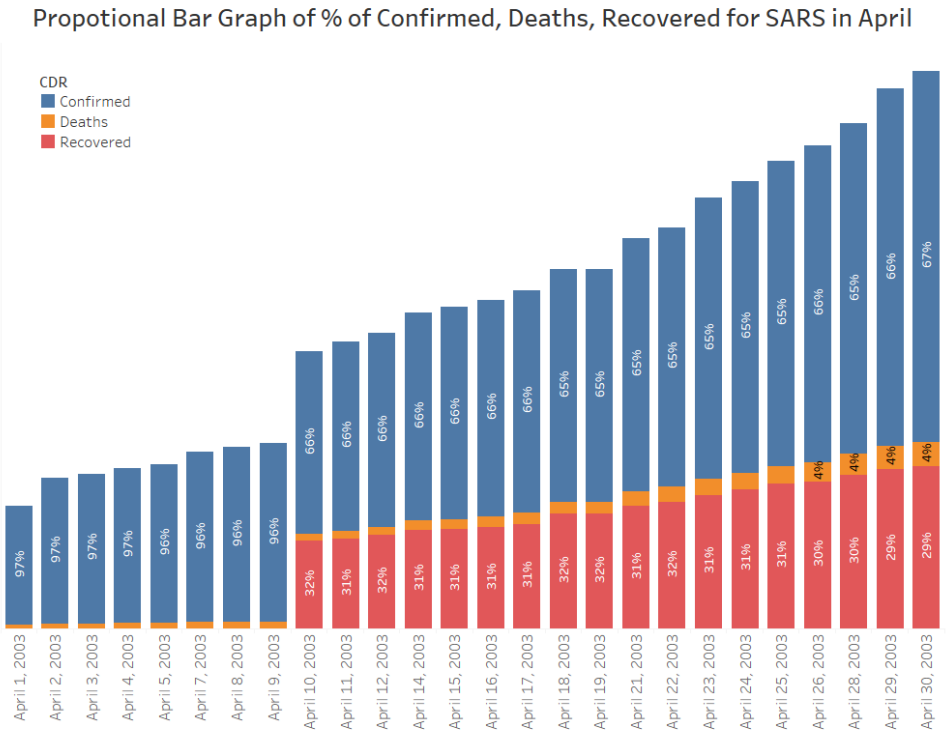
SD4:


Sweden
Japan
Russia
USA
Cambodia
France
China
Nepal
South Korea
Italy
Finland
Wuhan
UAE
Germany
India
Australia
Taiwan
Vietnam
Canada
Belgium
Singapore
Malaysia
Phillipines
Thailand
Sri Lanka
UK
Hong Kong

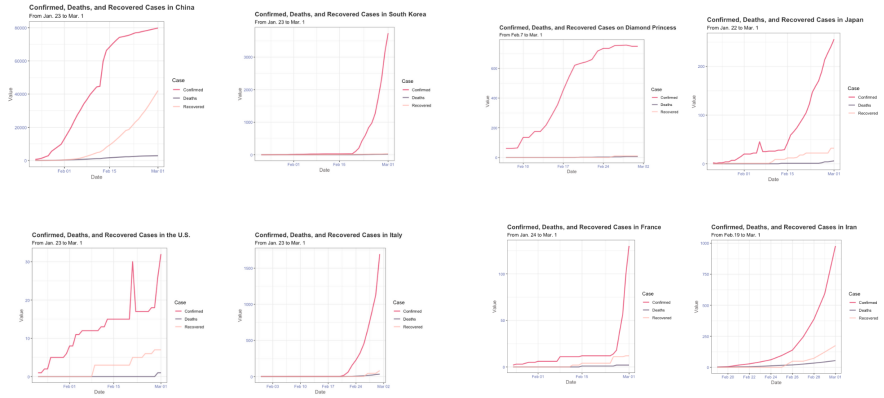
SQ1:


Propotional Bar Graph of % of Confirmed, Deaths, Recovered for COVID-19 in Febuary
CDR
Confirmed
Deaths
Recovered
96%
95%
95%
95%
94%
94%
93%
92%
91%
90%
89%
9%
88%
89%
9%
87%
11%
86%
12%
85%
13%
84%
14%
82%
16%
81%
17%
79%
19%
78%
19%
76%
22%
75%
22%
74%
23%
72%
25%
71%
27%
70%
28%
68%
30%
67%
31%
February 1, 2020
February 2, 2020
February 3, 2020
February 4, 2020
February 5, 2020
February 6, 2020
February 7, 2020
February 8, 2020
February 9, 2020
February 10, 2020
February 11, 2020
February 12, 2020
February 13, 2020
February 14, 2020
February 15, 2020
February 16, 2020
February 17, 2020
February 18, 2020
February 19, 2020
February 20, 2020
February 21, 2020
February 22, 2020
February 23, 2020
February 24, 2020
February 25, 2020
February 26, 2020
February 27, 2020
February 28, 2020
February 29, 2020

SQ2:


Propotional Bar Graph of % of Confirmed, Deaths, Recovered for SARS in April
CDR
Confirmed
Deaths
Recovered
97%
97%
97%
97%
96%
96%
96%
96%
66%
66%
66%
66%
66%
66%
66%
65%
65%
65%
65%
65%
65%
66%
65%
66%
67%
4%
4%
4%
4%
4%
32%
31%
32%
31%
31%
31%
31%
32%
32%
31%
32%
31%
31%
30%
30%
29%
29%
April 1, 2003
April 2, 2003
April 3, 2003
April 4, 2003
April 5, 2003
April 7, 2003
April 8, 2003
April 9, 2003
April 10, 2003
April 11, 2003
April 12, 2003
April 14, 2003
April 15, 2003
April 16, 2003
April 17, 2003
April 18, 2003
April 19, 2003
April 21, 2003
April 22, 2003
April 23, 2003
April 24, 2003
April 25, 2003
April 26, 2003
April 28, 2003
April 29, 2003
April 30, 2003

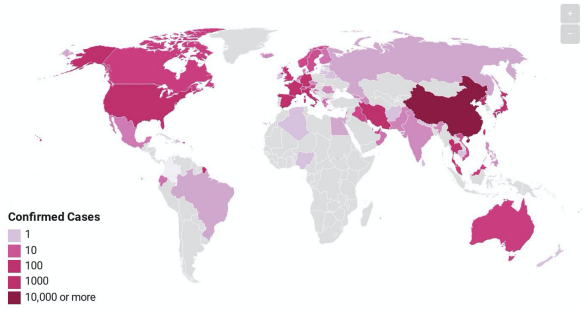
(YS1)


Confirmed, Deaths, and Recovered Cases in China
Confirmed, Deaths, and Recovered Cases in South Korea
Confirmed, Deaths, and Recovered Cases on Diamond Princess
Confirmed, Deaths, and Recovered Cases in Japan
Confirmed, Deaths, and Recovered Cases in the U.S.
Confirmed, Deaths, and Recovered Cases in Italy
Confirmed, Deaths, and Recovered Cases in France
Confirmed, Deaths, and Recovered Cases in Iran

(YS2)

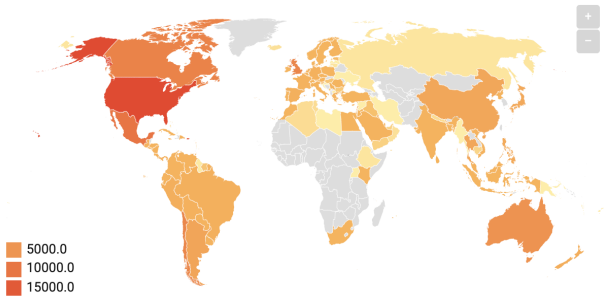
Confirmed COVID-19 Cases Globally

Confirmed Cases from Jan. 22 to Mar. 1



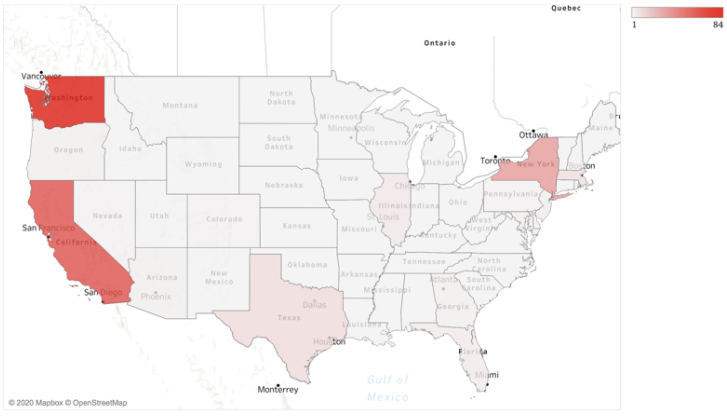
Confirmed H1N1 Cases Globally

2009 H1N1 cases from early 2009 to late 2010.



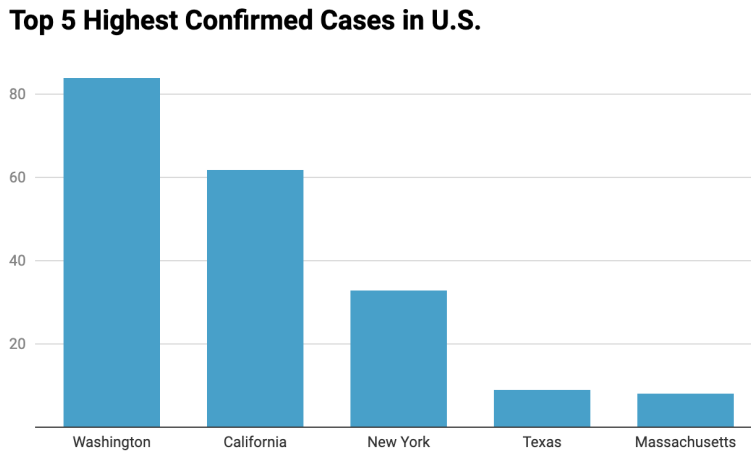
(YS3)

U.S. Confirmed Cases

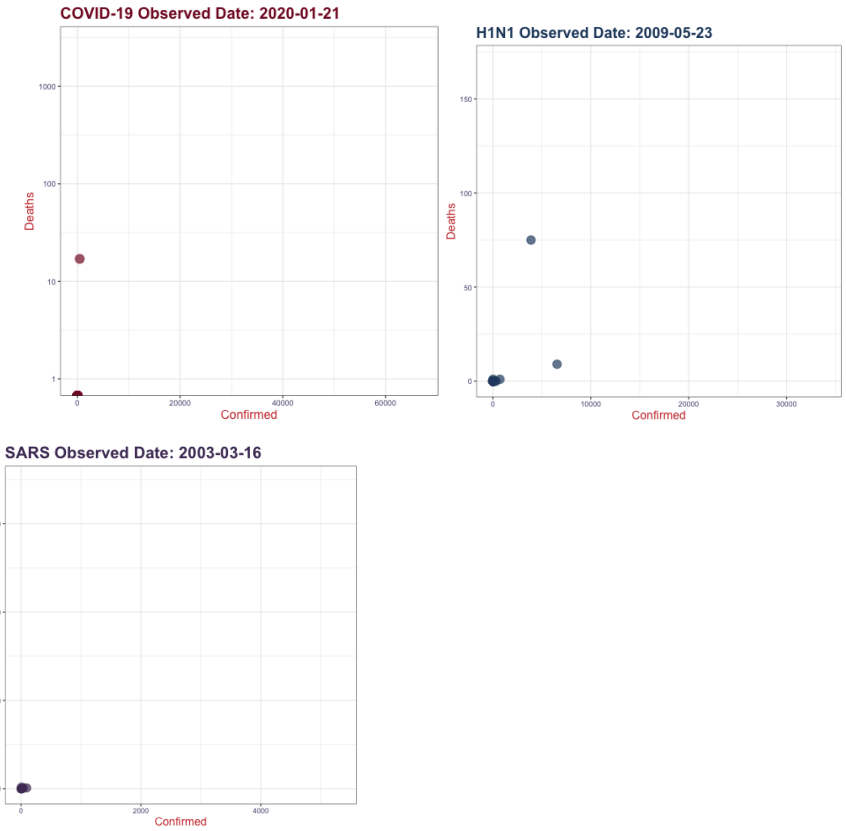


Map based on Longitude (generated) and Latitude (generated). Color shows details about Number. Details are shown for State.

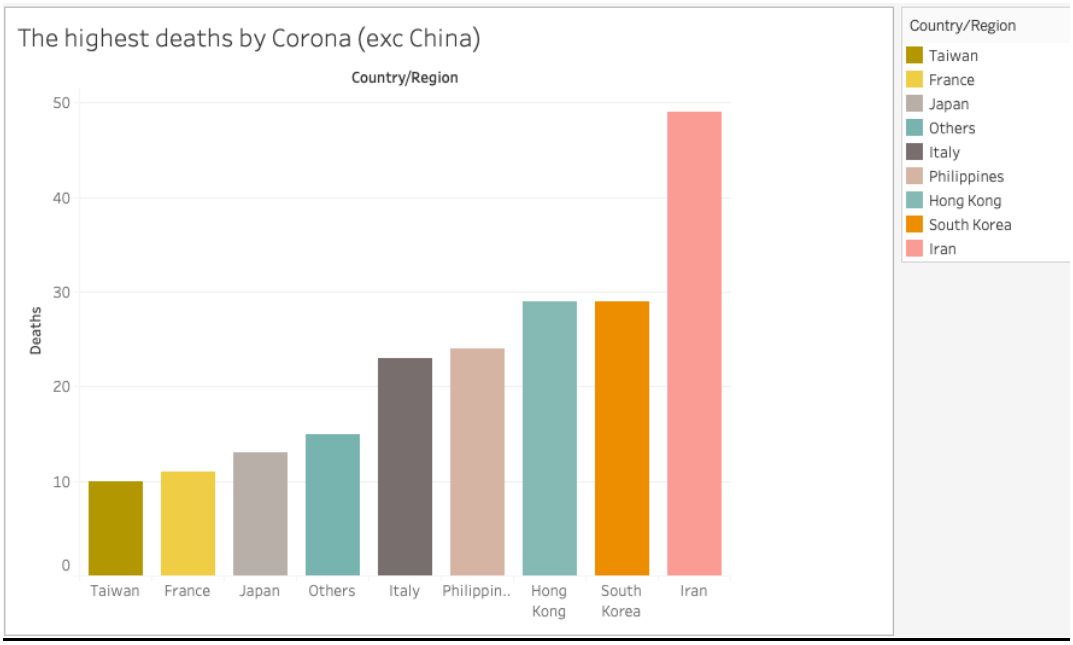
(YS4)


Top 5 Highest Confirmed Cases in U.S.
80
60
40
20
Washington
California
New York
Texas
Massachusetts

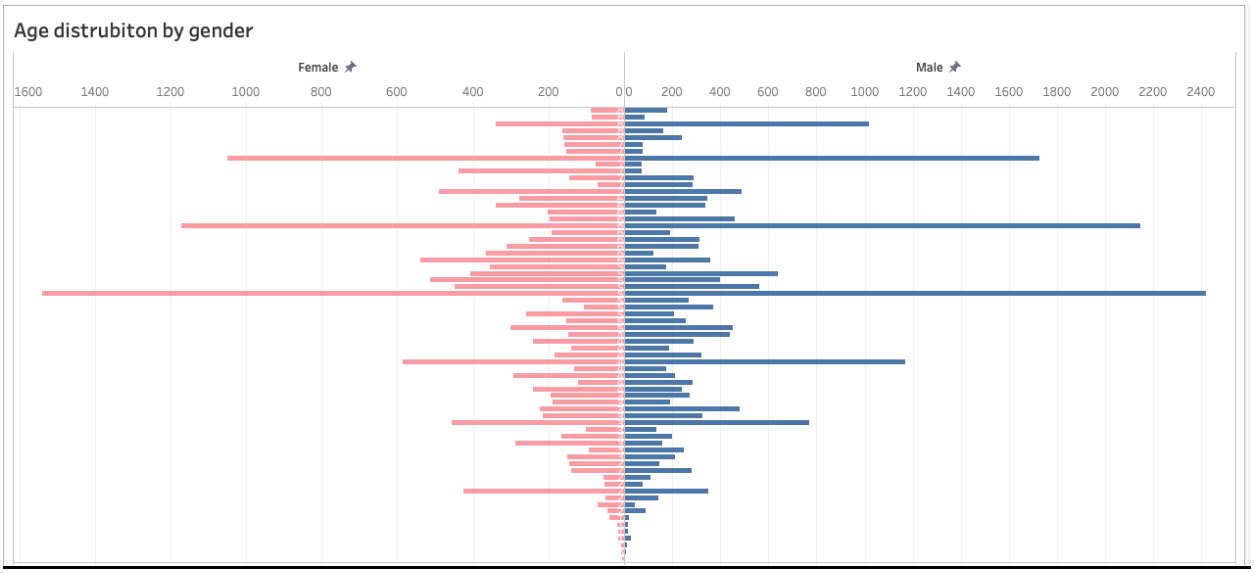
(YS5)


COVID-19 Observed Date: 2020-01-21
Deaths
Confirmed
H1N1 Observed Date: 2009-05-23
Deaths
Confirmed
SARS Observed Date: 2003-03-16
Deaths
Confirmed

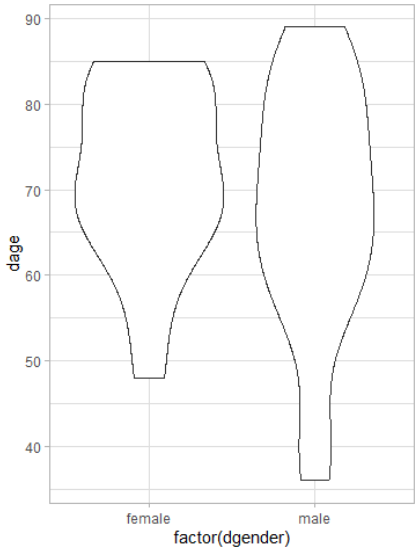
AA1:


The highest deaths by Corona (exc China)
Country/Region
Deaths
50
40
30
20
10
0
Taiwan
France
Japan
Others
Italy
Philippin..
Hong Kong
South Korea
Iran
Country/Region
Taiwan
France
Japan
Others
Italy
Philippines
Hong Kong
South Korea
Iran

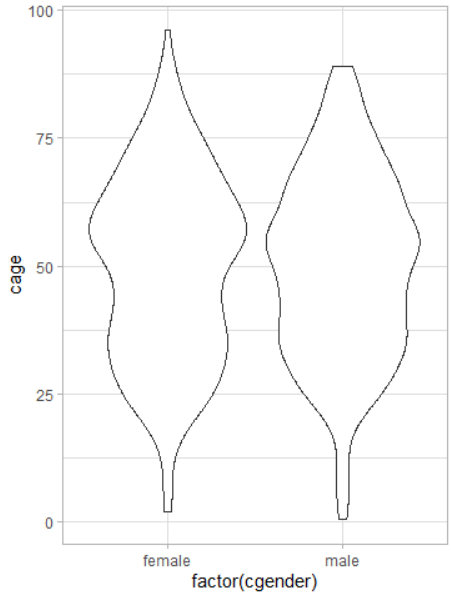
AA2:


Age distrubiton by gender
Female
Male
1600 1400 1200 1000 800 600 400 200 0
0 200 400 600 800 1000 1200 1400 1600 1800 2000 2200 2400

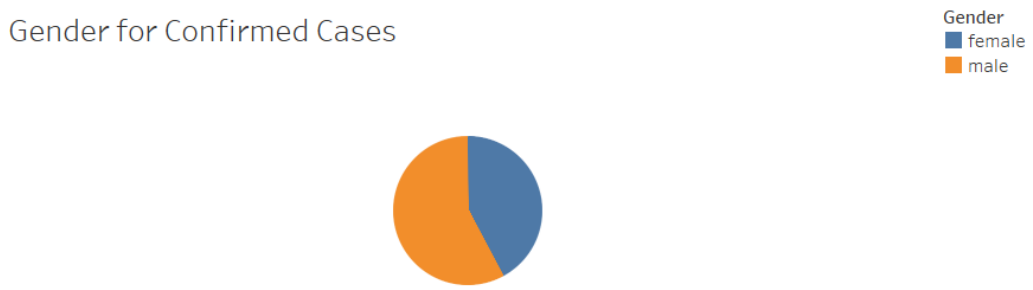
SA1:

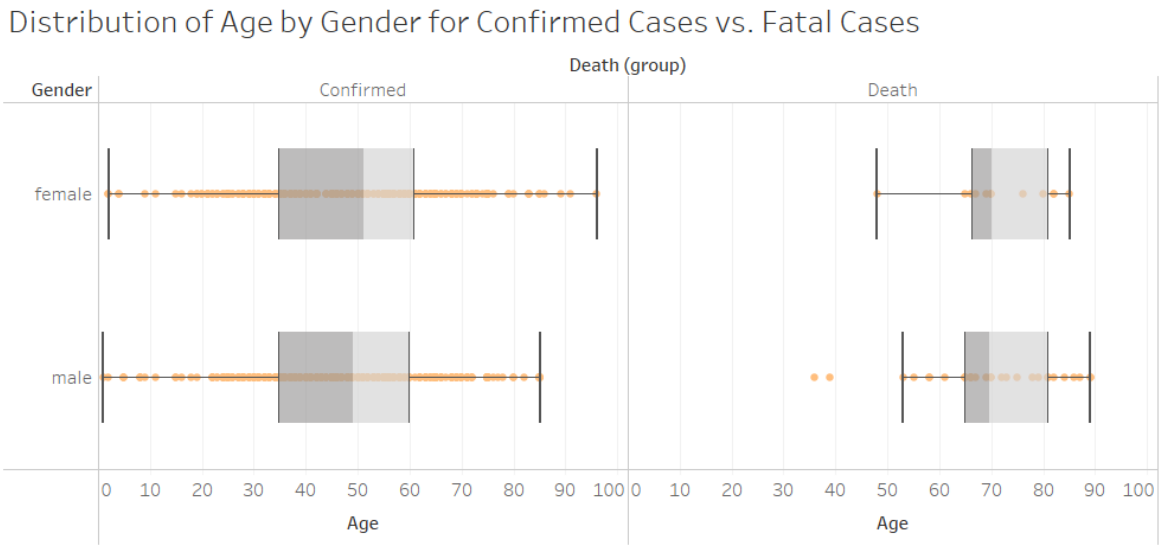

90
80
70
60
50
40
dage
female
male
factor(dgender)

SA2:

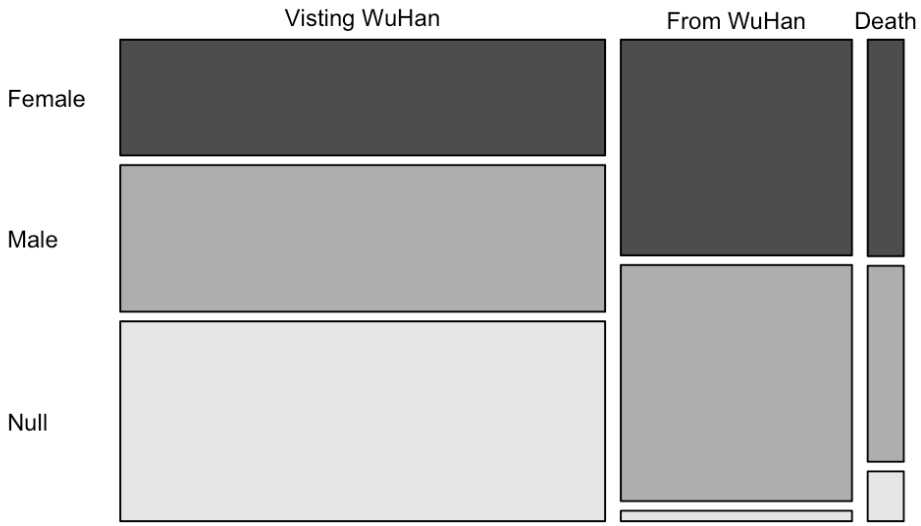

100
75
50
25
0
cage
female
male
factor(cgender)

SA3:

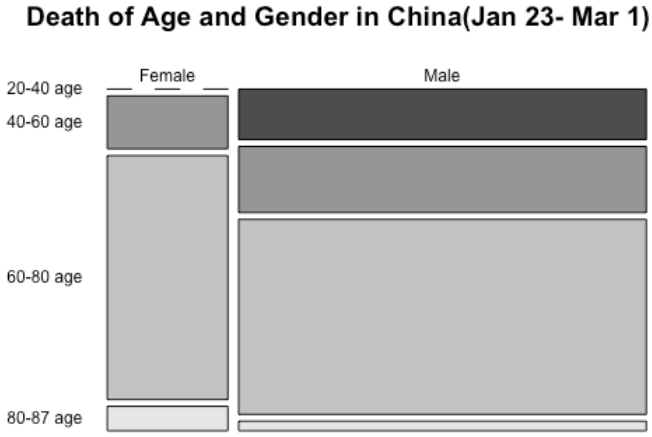

Gender for Confirmed Cases
Gender
female
male


Distribution of Age by Gender for Confirmed Cases vs. Fatal Cases
Death (group)
Gender
Confirmed
Death
female
male
0 10 20 30 40 50 60 70 80 90 100
Age
0 10 20 30 40 50 60 70 80 90 100
Age

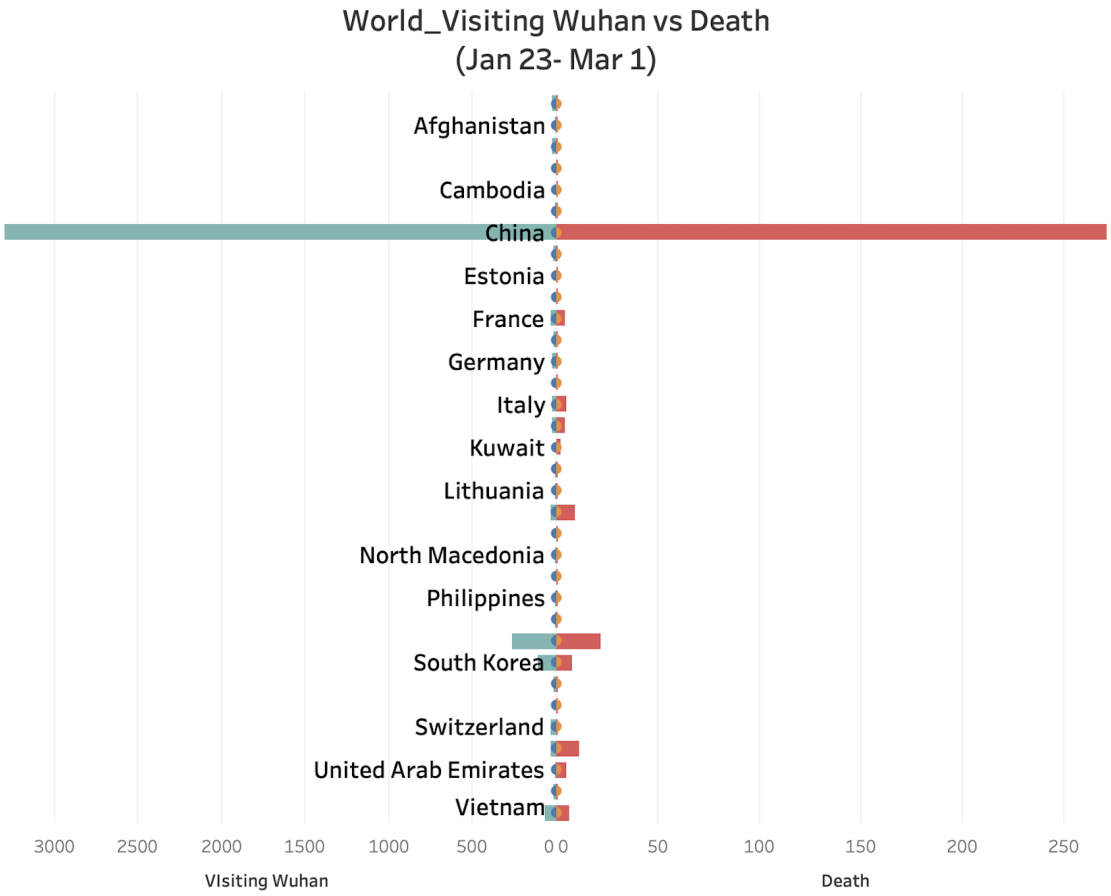
SY1:


Visting WuHan
From WuHan
Death
Female
Male
Null

SY2:


Death of Age and Gender in China(Jan 23- Mar 1)
Female
Male
20-40 age
40-60 age
60-80 age
80-87 age

SY3:



Visiting Wuhan, Zero axis, death and Zero axis for each country (COVID19_open_line_list.csv). Color shows details about visiting Wuhan, Zero axis, death and Zero axis. For pane Sum of Zero axis: The marks are labeled by country (COVID19_open_line_list.csv).