The world needs bettwe coordination to deal with pandemics

Project Members: Mamayaa Opoku, Yoshimune Satake, Zeyd Khalil

Introduction

A **Coordinated response** on a global scale is essential to facing this pandemic and any other pandemic to come in the future. Our data analysis of COVID-19 shows, with reasonable confidence, that the world with a coordinated response would have dealth mcuh bettwe with COVID-19 by now.

Abstract

In December 2019, the world was hit by a global pandemic known as **COVID-19**, which stands for Corona Virus Disease 2019. As of today, over **1.5 million people around the world lost their lives**, 250,000 of which where Americans. A lot of people also lost their jobs as economies started to suffer, and schools and universities were left shutdown, leaving students feel isolated from their friends and normal life.

Some countries suffered more than others from COVID-19, and some countries appeared to have little or no impact from this pandemic. The ones that suffered had a **stricter measures** such as nationwide lockdowns, mass testing, and early measures. The countries that suffered more appeared to have followed different measures.

There was no coordinated response, which is what resulted in mass lockdowns, and economies to shutdown.

Looking at this situation, we wanted to evaluate what the outcome would have been, had we globally acted in a coordinated fashion.

For that, we selected a group of "reference" countries where the case and death rates are low. These countries include New Zealand, China, Japan and South Korea. **What if we applied the measures taken in these countries onto badly affected countries?**

Applying those strict measures onto the rest of world, we saw astonishing results, which included lower cases and deaths rates.

We studied 4 different countries for this project:

1. China
2. USA
3. New Zealand
4. South Korea

```
covid_stats <- read_csv("/Users/zkhalil.19/Desktop/owid-covid-data-actual.csv")
```

```
covid_stats_ty <-
  covid_stats %>%
  filter(location == "United States"
         | location == "New Zealand"
         | location == "South Korea"
         | location == "China") %>%
  select(location, continent,
         date, total_cases,
         new_cases, total_deaths,
         new_deaths, total_tests,
         population)

covid_stats_ty <-
  covid_stats_ty %>%
  rename("country" = "location")
```

What makes a country have a "good" response?

We based our analysis on the following metrics:

- Early shutdown date, within a week from March 15, which was the day COVID-19 was declared a global pandemic
- Imposed restrictions - mandatory mask laws, total lockdown, contact tracing.
- What the numbers look like today - death rate and new cases rate.

Let's take a look at some of the countries that meet those metrics:

- New Zealand
- China
- Japan
- South Korea

According to the National News (https://www.thenationalnews.com/uae/health/coronavirus-uae-ranks-high-for-covid-19-response-in-global-league-table-1.1032167), South Korea, China, Japan and New Zealand were ranked 3rd, 5th, 6th and 9th (respectively) in the top 20 countries safest countries in the world from COVID-19.

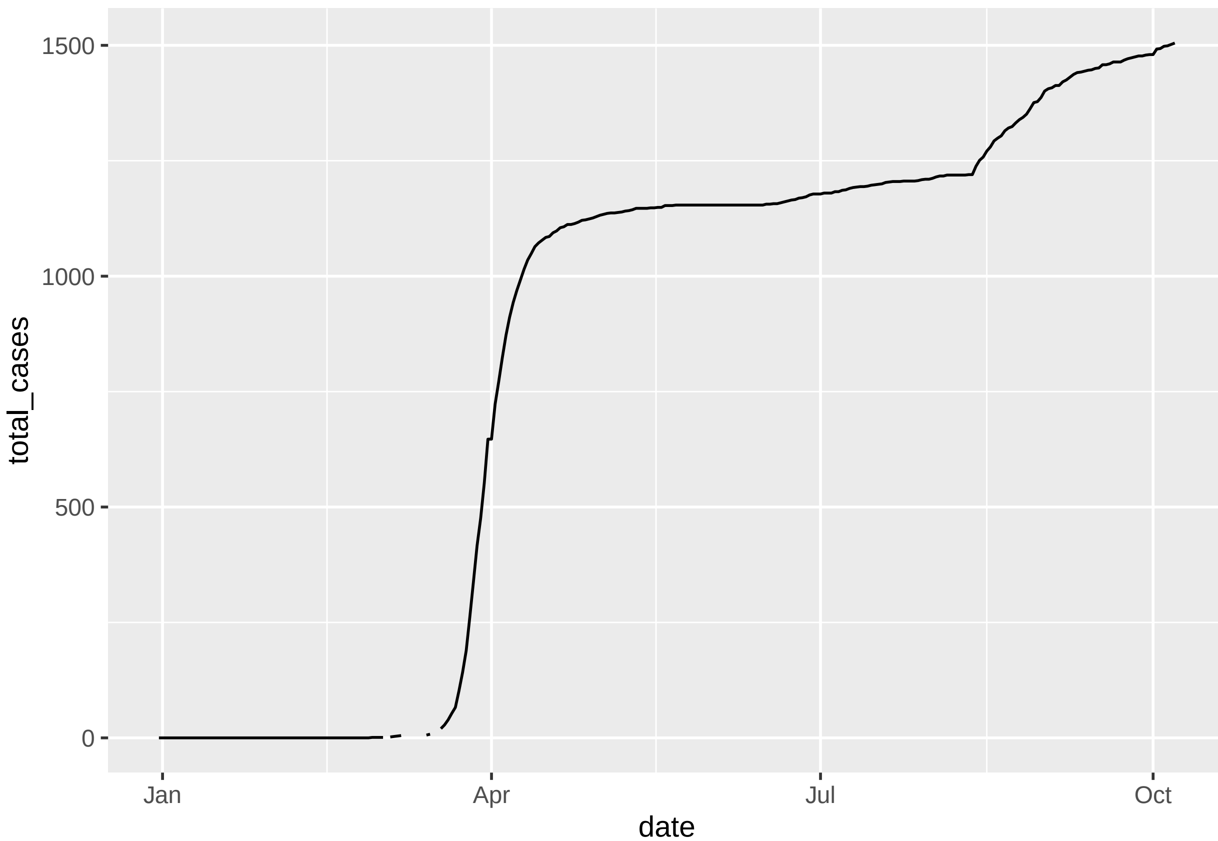
New Zealand

This information can be found in the following website: https://www.contagionlive.com/view/how-did-new-zealand-control-covid19

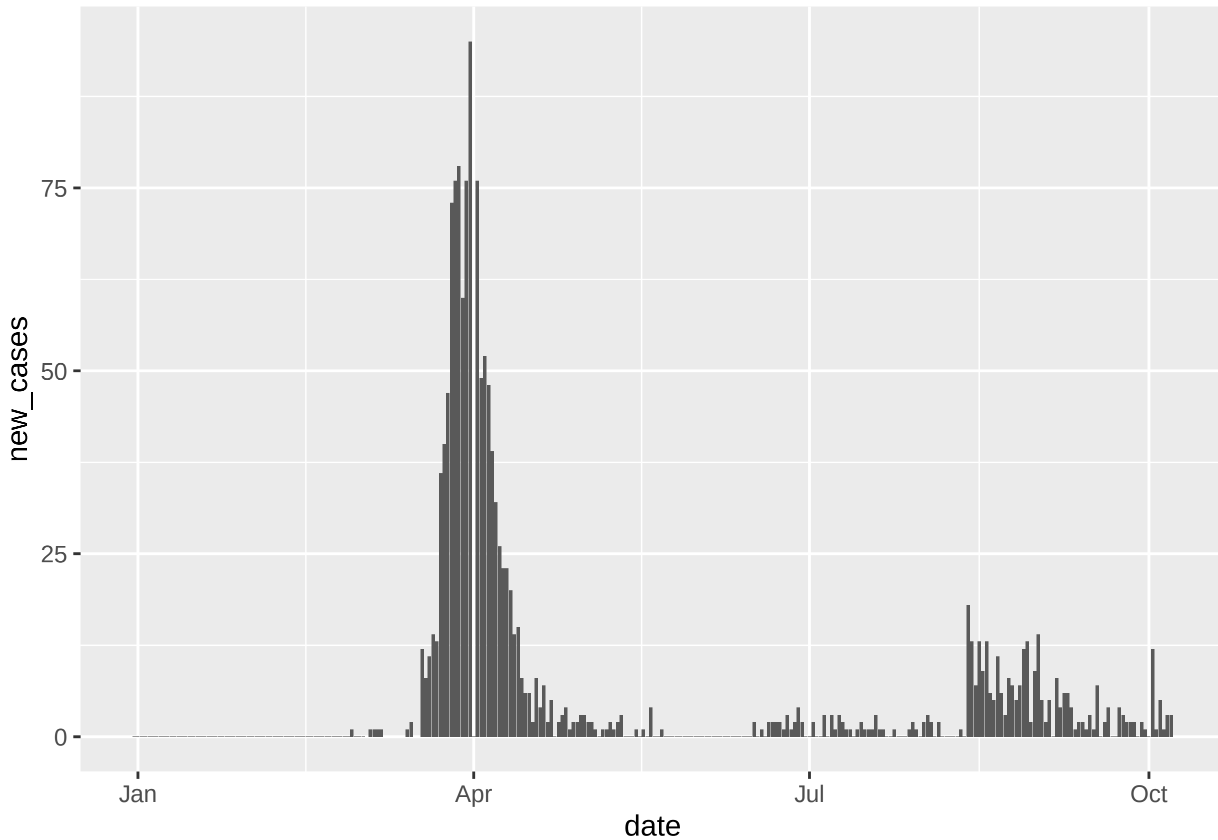
New Zealand is an isolated island nation in the South Pacific, comprised of a little under 5 million people, which is about the same population as the state of Alabama. This small country has become an emblematic champion of proper prevention and response to COVID-19.

```r
newzealand <- covid_stats_ty %>% filter(country == "New Zealand")
```

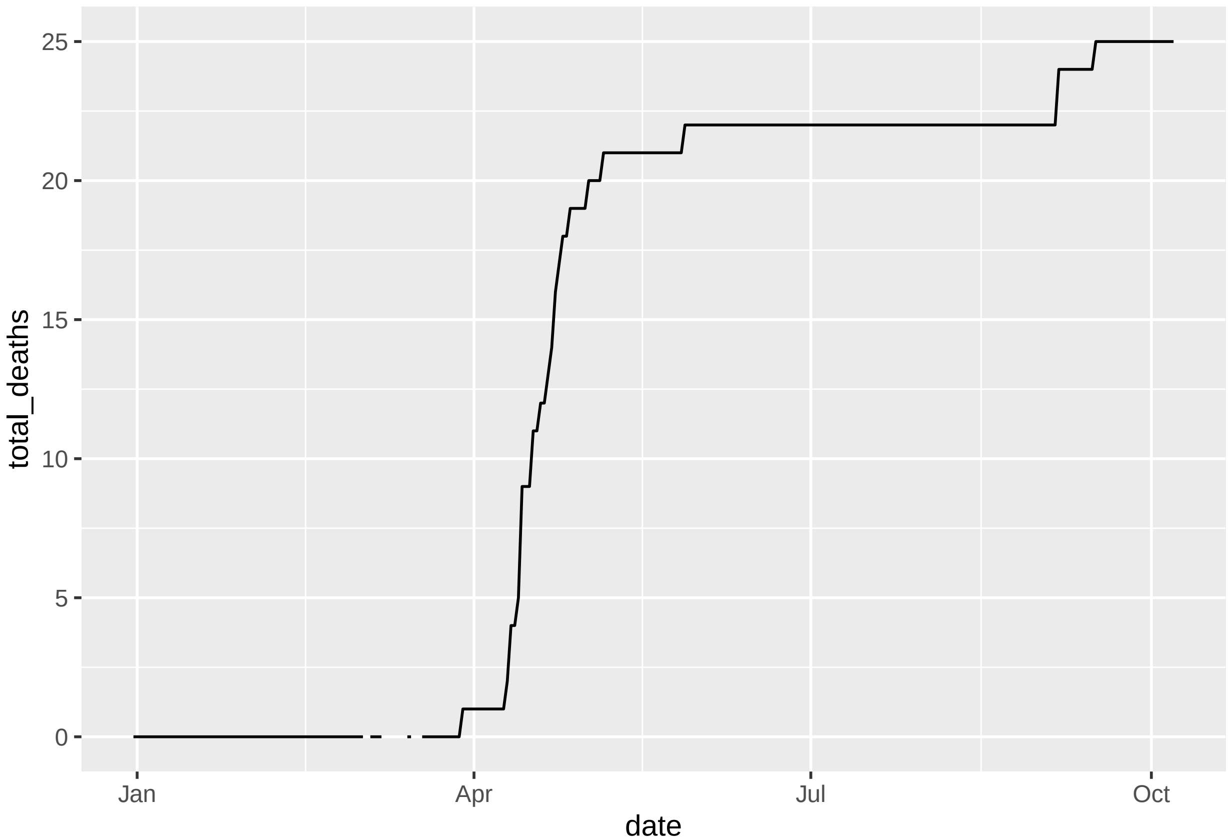
```r
library(ggplot2)

newzealand %>% ggplot(mapping = aes(date, total_cases)) + geom_line()
```



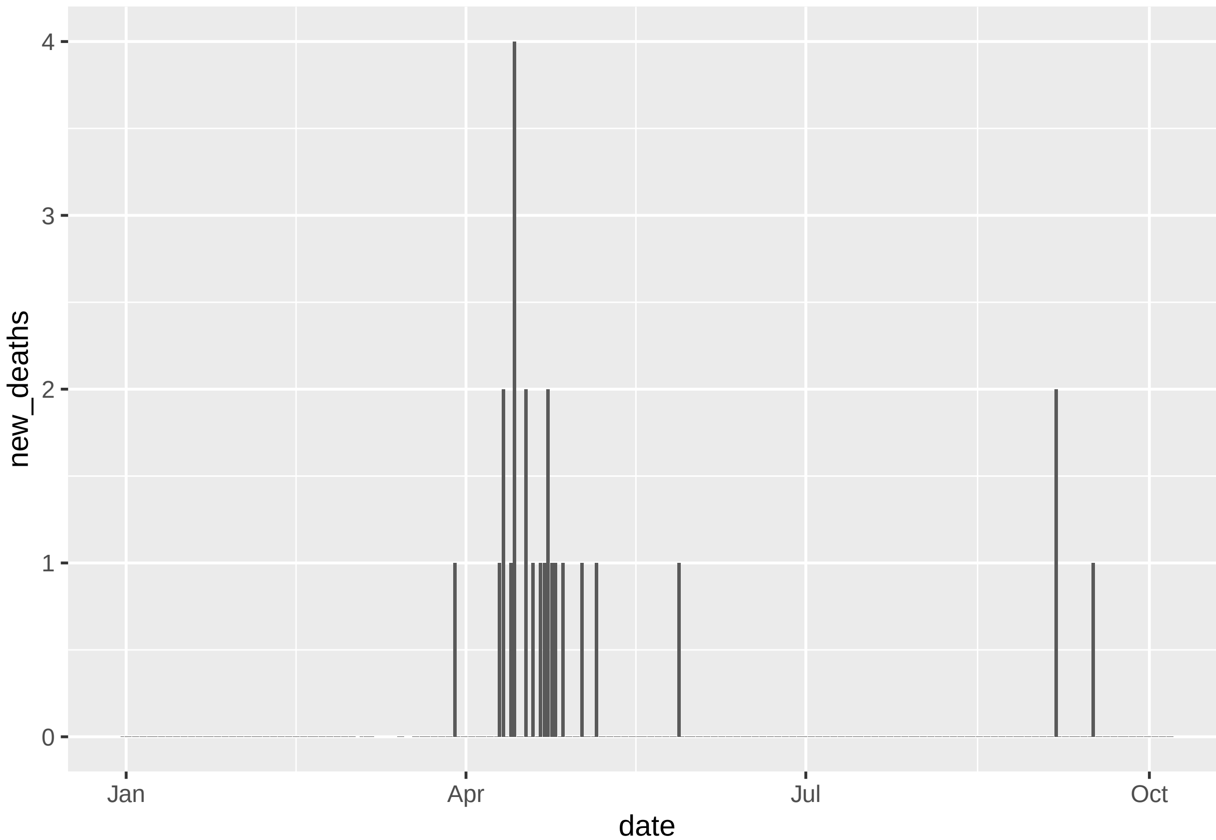

```r
ggplot(newzealand, aes(x = date, y = new_cases)) + geom_bar(stat = "identity")
```




```
newzealand %>% ggplot(mapping = aes(date, total_deaths)) + geom_line()
```




```r
ggplot(newzealand, aes(x = date, y = new_deaths)) + geom_bar(stat = "identity")
```



The first visualization shows us the total cases count in the country from the start of the pandemic. There seems to have been a sharp rise in the total cases count in March and April, but after that, the cases started to rise at a much lower rate, hence, the curve looks flatter.

The second visualization shows us that the months that saw the highest spike in new cases per day tended to be in March and April as well.

The third and fourth visualizations use the same metrics for the recorded COVID-19 deaths in New Zealand. It shows that only 25 people died.

New Zealand only saw a little over 2,000 total cases since February 26, which was the day the country saw its first case. A countrywide lockdown was imposed exactly a month later, on March 26, after having seen that there was a lack of testing and contact-tracing capability. The lockdown included a stay-at-home order, and unless it was for essential purposes, nobody was allowed out of their homes.

After 5 weeks, New Zealand started seeing its new cases rate decline rapidly, which resulted them moving from Alert 4 to Alert 3, which are ways to measure how critical their measures need to be, given the numbers. The country only imposed an extra 2 weeks of lockdown after that.

In early May, the last observed COVID-19 case was identified in New Zealand. The patient was isolated, and eventually recovered. By June 8, 103 days after their first case, they had moved to Alert 1 and declared the pandemic over in the country.

```r
newzealand %>% filter(new_cases == max(new_cases, na.rm = TRUE))
```

```
## # A tibble: 1 x 9
##   country continent date       total_cases new_cases total_deaths new_deaths
##   <chr>   <chr>     <date>           <dbl>     <dbl>        <dbl>      <dbl>
## 1 New Ze~ Oceania   2020-03-31         647        95            1          0
## # ... with 2 more variables: total_tests <lgl>, population <dbl>
```

> This shows the specific date where New Zealand saw the highest number of new cases in a day, which was March 31st, and that number is 95, which, compared to other countries, is extremely low.
>
> As previously mentioned, New Zealand imposed a nationwide lockdown on the 26th of March, and only saw a rise during the first 5 days of the lockdown, before the infection rate started decreasing.
>
> It is important to note that the infection rate kept on decreasing consistently because the people abided by the rules and stayed in their homes.

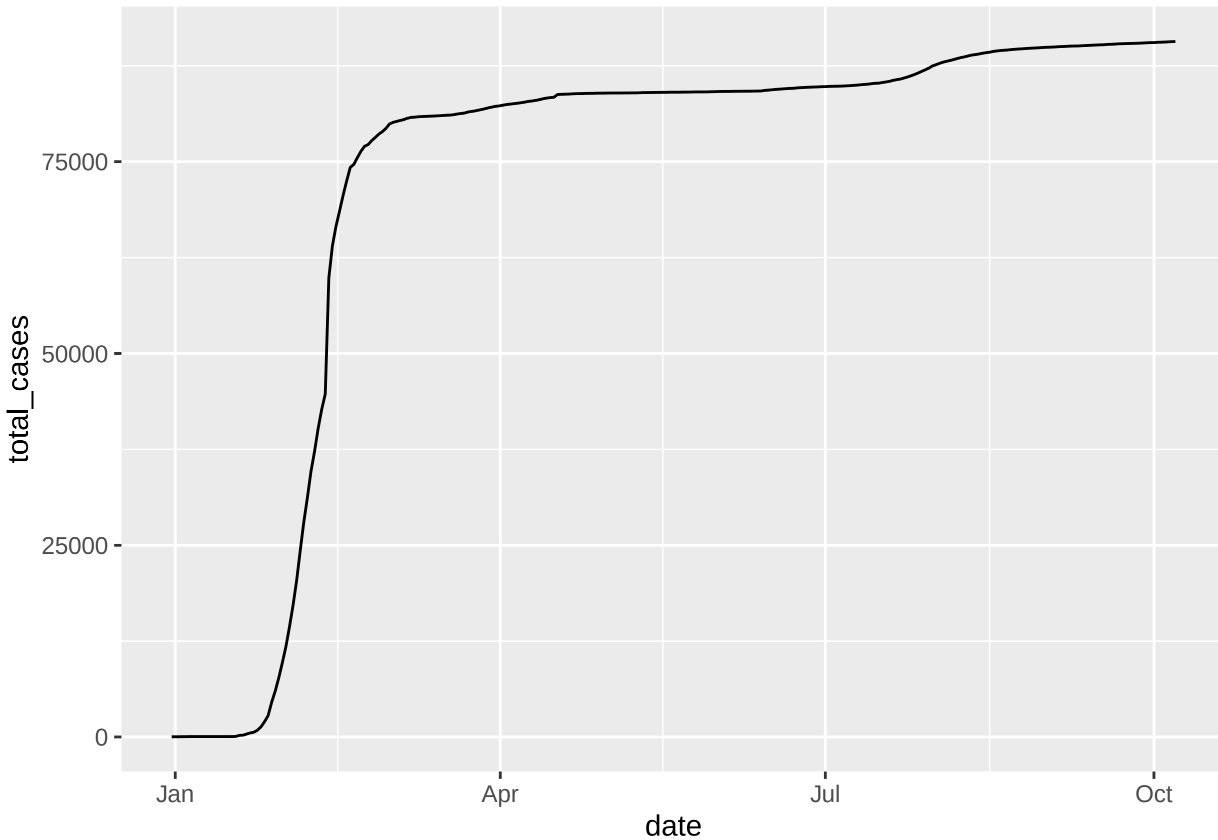
China

This information can be found in https://www.usatoday.com/story/news/world/2020/04/01/coronavirus-covid-19-china-radical-measures-lockdowns-mass-quarantines/2938374001/

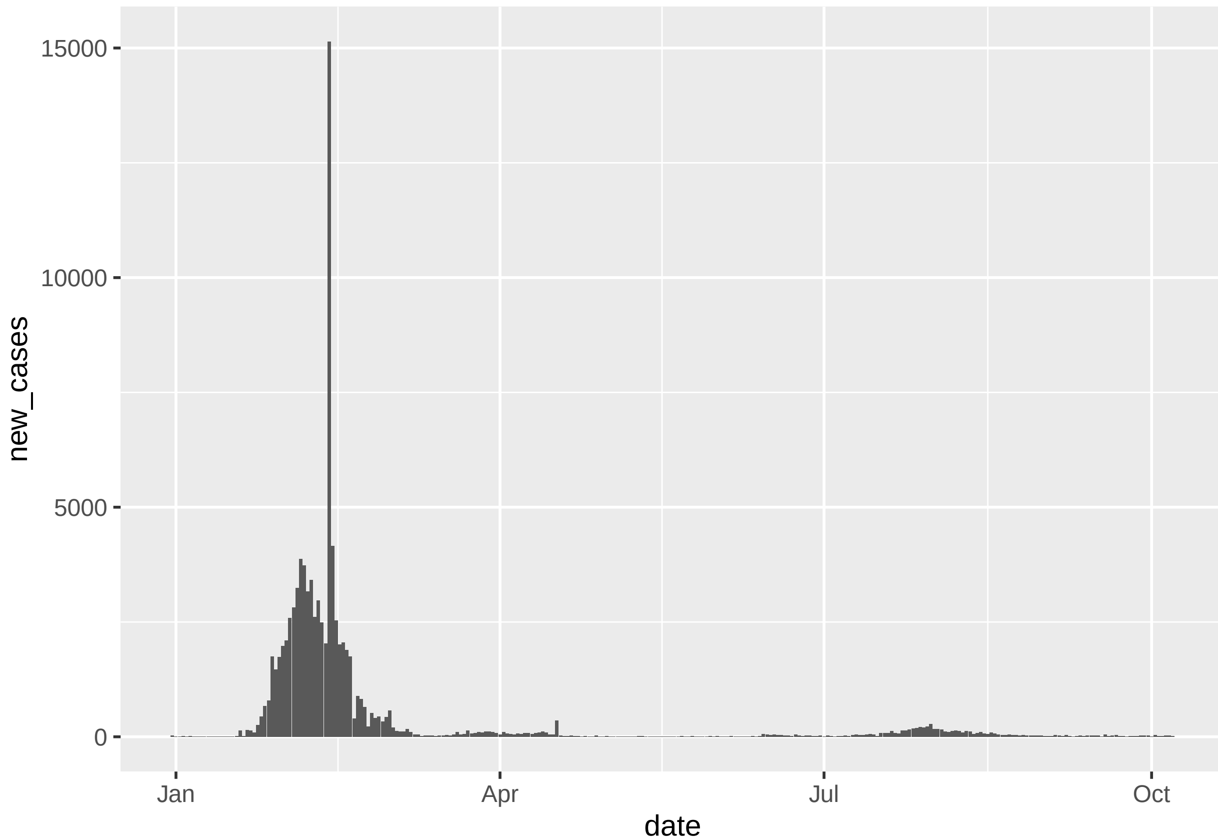
> China is, without a doubt, where COVID-19 originated.
>
> The thing that shocked the rest of the world the most is that China was able to flatten its COVID-19 curve before other countries even started to get badly affected.

```r
china <- covid_stats_ty %>% filter(country == "China")
```

```r
library(ggplot2)

china %>% ggplot(mapping = aes(date, total_cases)) + geom_line()
```




```r
ggplot(china, aes(x = date, y = new_cases)) + geom_bar(stat = "identity")
```




```r
china %>% filter(new_cases == max(new_cases))
```

```
## # A tibble: 1 x 9
##   country continent date       total_cases new_cases total_deaths new_deaths
##   <chr>   <chr>     <date>           <dbl>     <dbl>        <dbl>      <dbl>
## 1 China   Asia      2020-02-13       59865     15141         1368        254
## # ... with 2 more variables: total_tests <lgl>, population <dbl>
```

In Wuhan, China, the city where the virus was born, a mass lockdown was also imposed. This happened in late February, before the rest of the world got badly affected. Authorities went door-to-door for health checks - forcibly isolating every resident displaying even the mildest possible symptoms. Drones could be seen hovering around the city, yelling at people to get inside their homes. There were mandatory phone apps that were able to color code people based on their contagion risk.

The rest of China followed the same measures imposed in Wuhan, which is just what resulted in the flattening of the curve. "China's response to the outbreak was truly a nationwide response: systematic, comprehensive and coordinated."

The visualizations above show some very interesting trends. First off, on February 13, China saw 15,141 new cases in a day. That is highest number of cases recorded in a single day in China. However, if you look at the rest of the second visualization, you can see that after March, the country was practicallu COVID free.

The city of Wuhan and the entire province of Hubei, where Wuhan is located, found itself in absolute lockdown mode from January 23rd to April 8. Nobody was allowed out of their homes other than for essential trips like grocery shopping or medical trips.

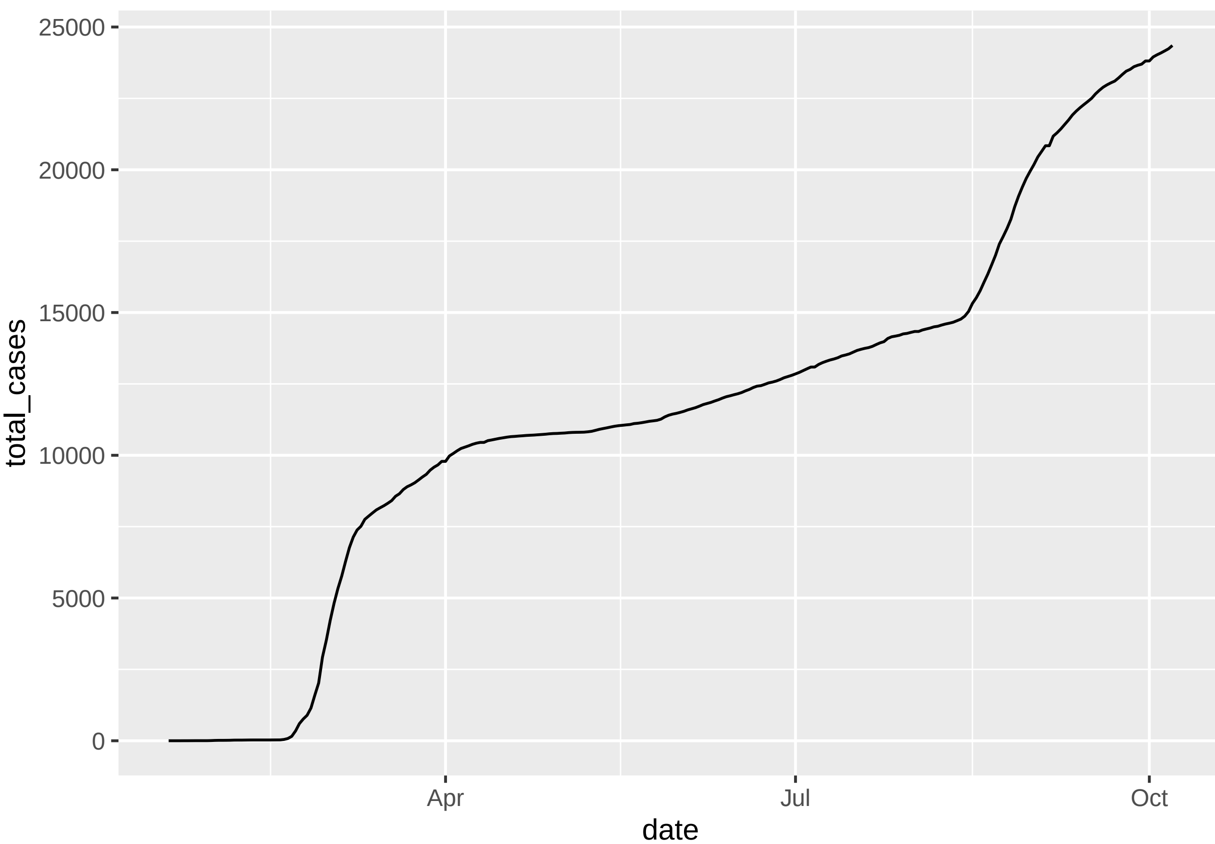
11 million people in Wuhan were on lockdown, along with an additional 57 million across China. The Chinese government took some of the strictest measures out of any other countries in the world, and that is why after March, there was an astonishing decrease in daily cases, until none were able to be seen anymore.

South Korea

This information can be found at https://english.alarabiya.net/en/features/2020/04/03/South-Korea-conquered-coronavirus-without-a-lockdown-a-model-to-follow-

South Korea is a country that had its first case the same day as the US.

```r
southkorea <-
  covid_stats_ty %>%
  filter(country == "South Korea",
         total_cases != 0)

ggplot(southkorea, aes(date, total_cases)) +
  geom_line()
```



The visualization above shows us the total cases in relation with time. Currently, there seems to be a second yet, smaller wave of cases. Between March and September, however, the curve seems to have been flattened as a result of the measures South Korea has taken.

Initially, South Korea was seen as a potential epicenter of the virus throughout the year, just like Italy or Spain, but the country had a strategy to prevent the spread of COVID-19.

Just one after January 20, the day South Korea recorded its first case, the government ordered factories to start producing coronavirys test kits en masse. Within two weeks, the country was producing more than 100,000 test kits a day.

South Korea's rapid response to the virus by mass testing is one of the things that helped the country have very high control over the prevention of its spread. The mass manufacturing of tests allowed the government to successfully test significant parts of the population for coronavirus, rolling out over 600 testing centers and making testing easy and available. This mass of data allowed authorities to monitor the spread of the virus and treat those with it.

Another reason why South Korea did not end suffering nearly as much as other countries is because it applied lessons learned from the 2015 MERS outbreak, which killed 36 people. Rather than rely on people with symptoms volunteering to be tested, authorities took a proactive approach which helped identify cases without symptoms which may have otherwise gone on to spread the virus further.

The government of South Korea made information about the spread of the virus very public, which concerned the population, which allowed both governments and citizens to act effectively and responsibly to slow the spread of the virus – without authorities having to enforce overbearing nationwide regulations or shut down the economy.

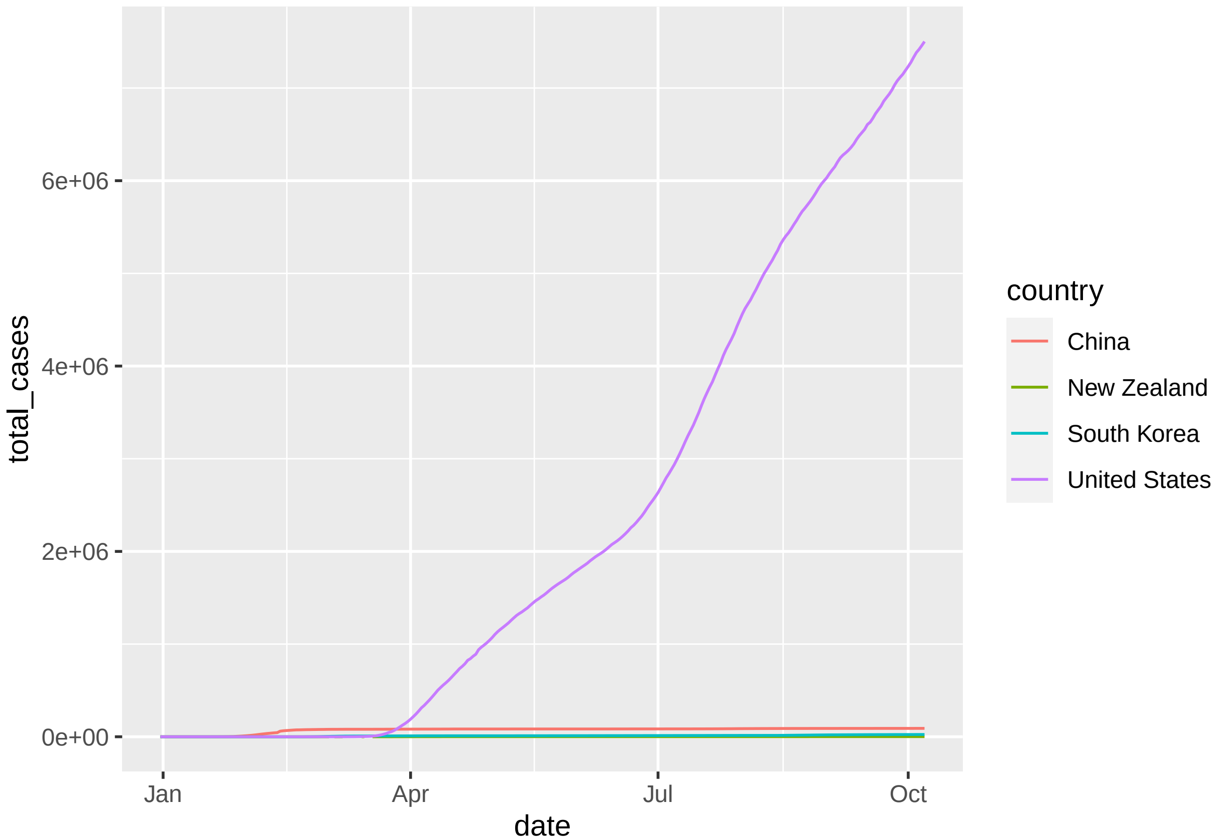
According to the source, residents had access to a live map where they could see all the information on a live tracking system. They could track people via their credit card and CCTV to see exactly where they have been.

This is why South Korea did not end up having a nationwide lockdown, and as a result, it was able to avoid an economic crash.

Contrast that to the US

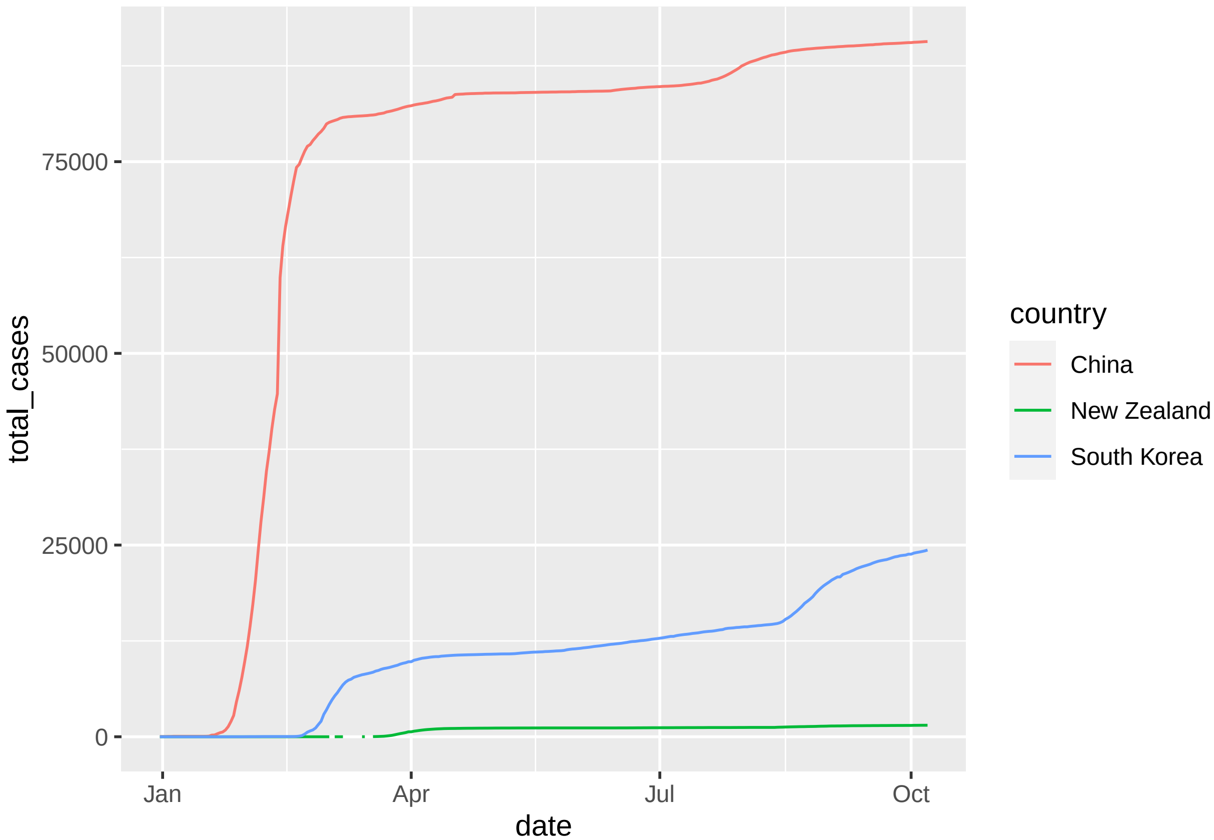
This is a visualization that compares some of the least affected countries with the United States, based on actual data so far.

```
badly_affected <-
  covid_stats_ty %>%
  filter(country == "New Zealand"
         | country == "United States"
         | country == "China"
         | country == "South Korea")

ggplot(data = badly_affected) + geom_line(mapping = aes(x = date, y = total_cases, color = country))
```



Now, let's take a look at the examples of the 3 countries that followed roughly the same measues.

```
good_countries <-
  covid_stats_ty %>%
  filter(country == "New Zealand"
         | country == "China"
         | country == "South Korea")

ggplot(data = good_countries) + geom_line(mapping = aes(x = date, y = total_cases, color = country))
```



We see that as a result of firm and early actions, each country appears to see similar results.

The United States and South Korea both saw their first cases on the same day, January 20. However, South Korea imposed strict measures much earlier and quicker than the US did. The earliest US States to have had a confirmed case are Washington, New York and California.

Applying the infection rate in South Korea to those initial states, would we have seen a better impact in the US? That was the basis for our next analysis:

- Calculated the ratio of South Korea for cases per population on the 30th and 120th day.
- Applied that ratio to the 3 states where the first cases were recorded in the US.
- Compared our actual findings ("As-Is"), with our applied findings ("What-If").

Let's import the US COVID dataset

```
US_by_state <-
  read_csv("https://raw.githubusercontent.com/nytimes/covid-19-data/master/us-counties.csv")

US <-
  US_by_state %>%
  select(date, state, cases, deaths) %>%
  group_by(date, state) %>%
  summarize(total_cases = sum(cases))
```

Here, we made new datasets that show the number of cases in South Korea, New Zealand, China, New York, Washington and California during the first 120 days since each of their first recorded cases.

```
southkorea_fourmonths <-
  southkorea %>%
  filter(total_cases != 0) %>%
  head(120)

newzealand_fourmonths <-
  newzealand %>%
  filter(total_cases != 0)
  head(120)
```

```
## [1] 120
```

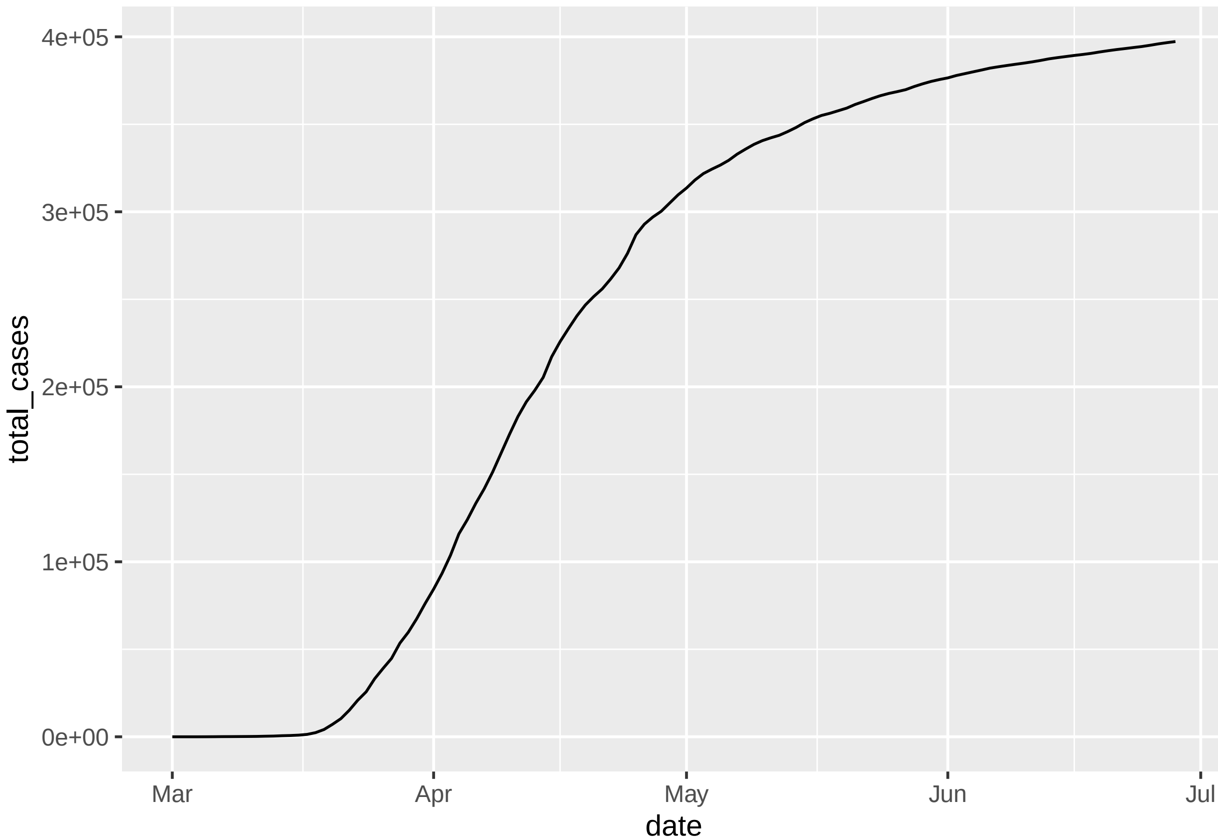
```
china_fourmonths <-
  head(china, 120)
```

```
newyork <-
  US %>%
  filter(state == "New York")

washington <-
  US %>%
  filter(state == "Washington")

california <-
  US %>%
  filter(state == "California")
```

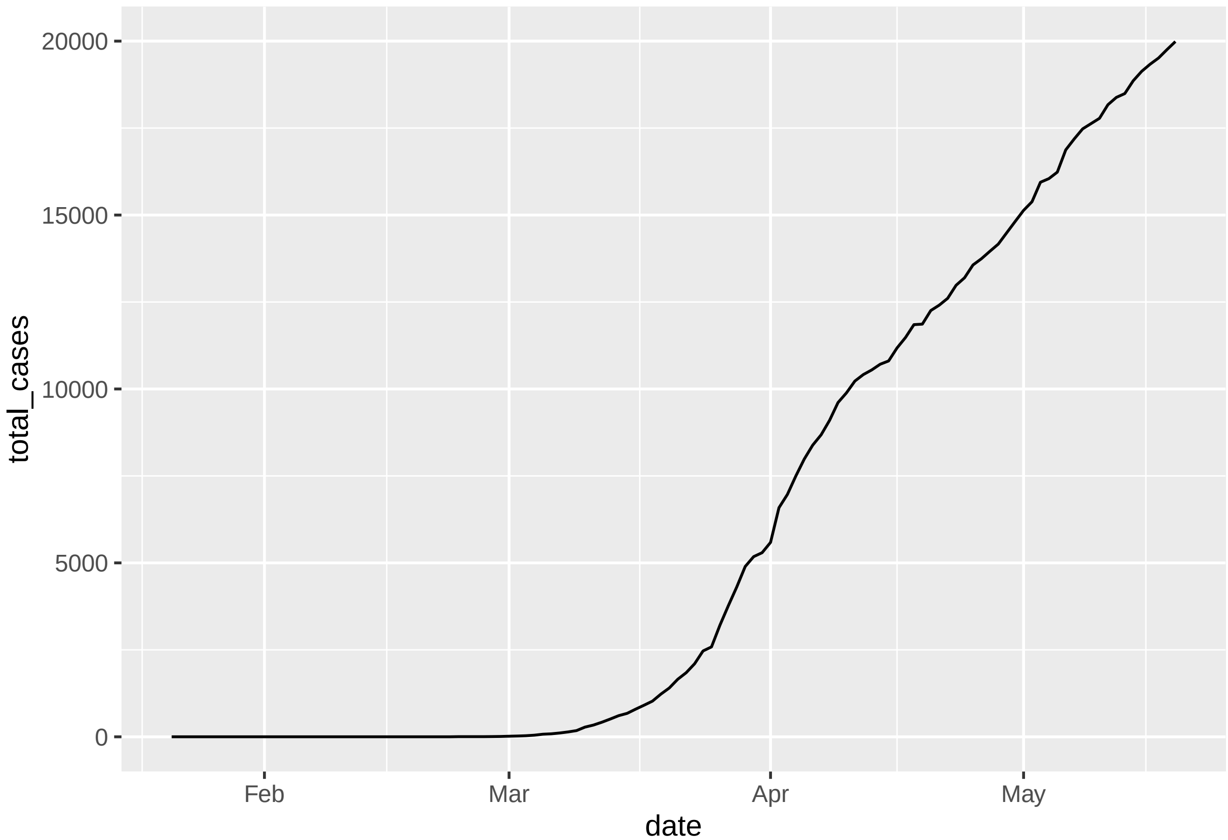
```
newyork <-
  newyork %>%
  head(120)

washington <-
  washington %>%
  head(120)

california <-
  california %>%
  head(120)
```

Now that we have our New York, Washington and California dataset, let's make some visualizations from those datasets.

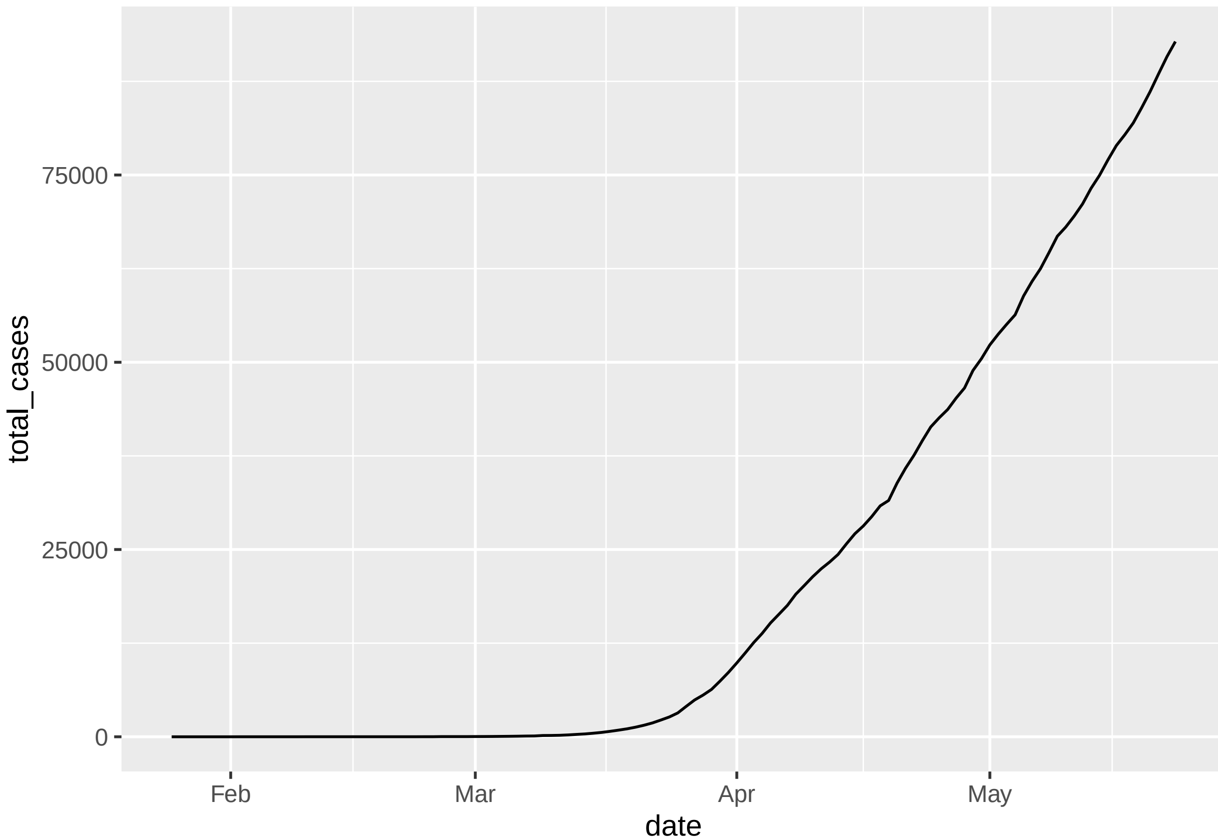
```
newyork %>%
  ggplot(aes(date, total_cases)) +
  geom_line()
```




```
washington %>%
  ggplot(aes(date, total_cases)) +
  geom_line()
```




```
california %>%
  ggplot(aes(date, total_cases)) +
  geom_line()
```



Here's what the numbers looked like in South Korea on its 30th and 120th:

```
State_Country = c("South Korea")
Population = c(max(southkorea$population))
Day_30_cases <- c(southkorea[30,]$total_cases)
Per_population_30 <- c((Day_30_cases / Population) * 100)
Day_120_cases <- c(southkorea[120,]$total_cases)
Per_population_120 <- c((Day_120_cases / Population) * 100)

as_is_korea <- tibble(Day_30_cases,
                Per_population_30,
                Day_120_cases,
                Per_population_120)

as_is_korea
```

```
## # A tibble: 1 x 4
##   Day_30_cases Per_population_30 Day_120_cases Per_population_120
##          <dbl>             <dbl>         <dbl>              <dbl>
## 1           31         0.0000605         11065             0.0216
```

The Per_population 30 and 120 represent the total number of cases on the 30th and 120th days since first recorded case divided by total population, in South Korea.

Assuming New York, Washington and California each took the same exact measures as South Korea and ended up having the same cases to population ratio, what would have their numbers looked like on their 30th and 120th day?

South Korea measures applied in New York

```
State = c("New York")
Pop = c(19450000)
D30_cases <- c(newyork[30,]$total_cases)
pp_30 <- c((D30_cases / Pop) * 100)
D120_cases <- c(newyork[120,]$total_cases)
pp_120 <- c((D120_cases / Pop) * 100)
D30_wif <- (as_is_korea$Per_population_30 * Pop) / 100
D120_wif <- (as_is_korea$Per_population_120 * Pop) / 100

as_is_whatif_ny <- tibble(D30_cases,
                pp_30,
                D120_cases,
                pp_120,
                D30_wif,
                D120_wif)

as_is_whatif_ny
```

```
## # A tibble: 1 x 6
##   D30_cases pp_30 D120_cases pp_120 D30_wif D120_wif
##       <dbl> <dbl>      <dbl>  <dbl>   <dbl>    <dbl>
## 1     67504 0.347     397293   2.04    11.8    4198.
```

> This data shows us that had the state of New York taken the same measures as South Korea, there only would've been 4079 cases by the 120th day since New York saw its first case.

South Korea measures applied in California

```
State_Country = c("California")
Population = c(39510000)
Day_30_cases <- c(california[30,]$total_cases)
Per_population_30 <- c((Day_30_cases / Population) * 100)
Day_120_cases <- c(california[120,]$total_cases)
Per_population_120 <- c((Day_120_cases / Population) * 100)
Day_30_whatif <- (as_is_korea$Per_population_30 * Population) / 100
Day_120_whatif <- (as_is_korea$Per_population_120 * Population) / 100

as_is_whatif_ca <- tibble(Day_30_cases,
                 Per_population_30,
                 Day_120_cases,
                 Per_population_120,
                 Day_30_whatif,
                 Day_120_whatif)

as_is_whatif_ca
```

```
## # A tibble: 1 x 6
##   Day_30_cases Per_population_~ Day_120_cases Per_population_~ Day_30_whatif
##          <dbl>            <dbl>         <dbl>            <dbl>         <dbl>
## 1            9        0.0000228         92815            0.235          23.9
## # ... with 1 more variable: Day_120_whatif <dbl>
```

This data shows us that had the state of California taken the same measures as South Korea, there only would've been 8286 cases by the 120th day since California saw its first case.

South Korea measures applied in Washington

```
State_Country = c("Washington")
Population = c(7615000)
Day_30_cases <- c(washington[30,]$total_cases)
Per_population_30 <- c((Day_30_cases / Population) * 100)
Day_120_cases <- c(washington[120,]$total_cases)
Per_population_120 <- c((Day_120_cases / Population) * 100)
Day_30_whatif <- (as_is_korea$Per_population_30 * Population) / 100
Day_120_whatif <- (as_is_korea$Per_population_120 * Population) / 100

as_is_whatif_wa <- tibble(Day_30_cases,
                 Per_population_30,
                 Day_120_cases,
                 Per_population_120,
                 Day_30_whatif,
                 Day_120_whatif)

as_is_whatif_wa
```

```
## # A tibble: 1 x 6
##   Day_30_cases Per_population_~ Day_120_cases Per_population_~ Day_30_whatif
##          <dbl>            <dbl>         <dbl>            <dbl>         <dbl>
## 1            1        0.0000131         19988            0.262          4.60
## # ... with 1 more variable: Day_120_whatif <dbl>
```

This data shows us that had the state of Washington taken the same measures as South Korea, there only would've been 1597 cases by the 120th day since Washington saw its first case.

Since New York, Washington and California were some of the first major states that recorded cases in the country and are used as travel hubs to travel across the US, let's assume that their numbers determine the case count across the US.

```
US_ <- US %>% group_by(date) %>% summarize(cases_that_day = sum(total_cases)) %>% head(120)
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```r
State_Country = c("United States")
Population = c(328200000)
Day_30_cases <- c(US_[30,]$cases_that_day)
Per_population_30 <- c((Day_30_cases / Population) * 100)
Day_120_cases <- c(US_[120,]$cases_that_day)
Per_population_120 <- c((Day_120_cases / Population) * 100)
Day_30_whatif <- (as_is_korea$Per_population_30 * Population) / 100
Day_120_whatif <- (as_is_korea$Per_population_120 * Population) / 100

as_is_whatif_us <- tibble(State_Country,
                 Day_30_cases,
                 Per_population_30,
                 Day_120_cases,
                 Per_population_120,
                 Day_30_whatif,
                 Day_120_whatif)

as_is_whatif_us
```

```
## # A tibble: 1 x 7
##   State_Country Day_30_cases Per_population_~ Day_120_cases Per_population_~
##   <chr>                <dbl>            <dbl>         <dbl>            <dbl>
## 1 United States           25       0.00000762       1536570            0.468
## # ... with 2 more variables: Day_30_whatif <dbl>, Day_120_whatif <dbl>
```