COVID19_HW

Bilal Gilani

4/8/2020

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.3
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## Warning: package 'tidyr' was built under R version 3.6.3
```

```
## Warning: package 'purrr' was built under R version 3.6.3
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
theme_set(theme_bw())

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

```
## Parsed with column specification:
## cols(
##   date = col_date(format = ""),
##   county = col_character(),
##   state = col_character(),
##   fips = col_character(),
##   cases = col_double(),
##   deaths = col_double()
## )
```

```
deaths_by_state <- dat %>%
  count(state,date, wt = deaths) %>%
  filter(date == max(date)) %>%
  rename(total_deaths = n) %>%
  arrange(desc(total_deaths))
```

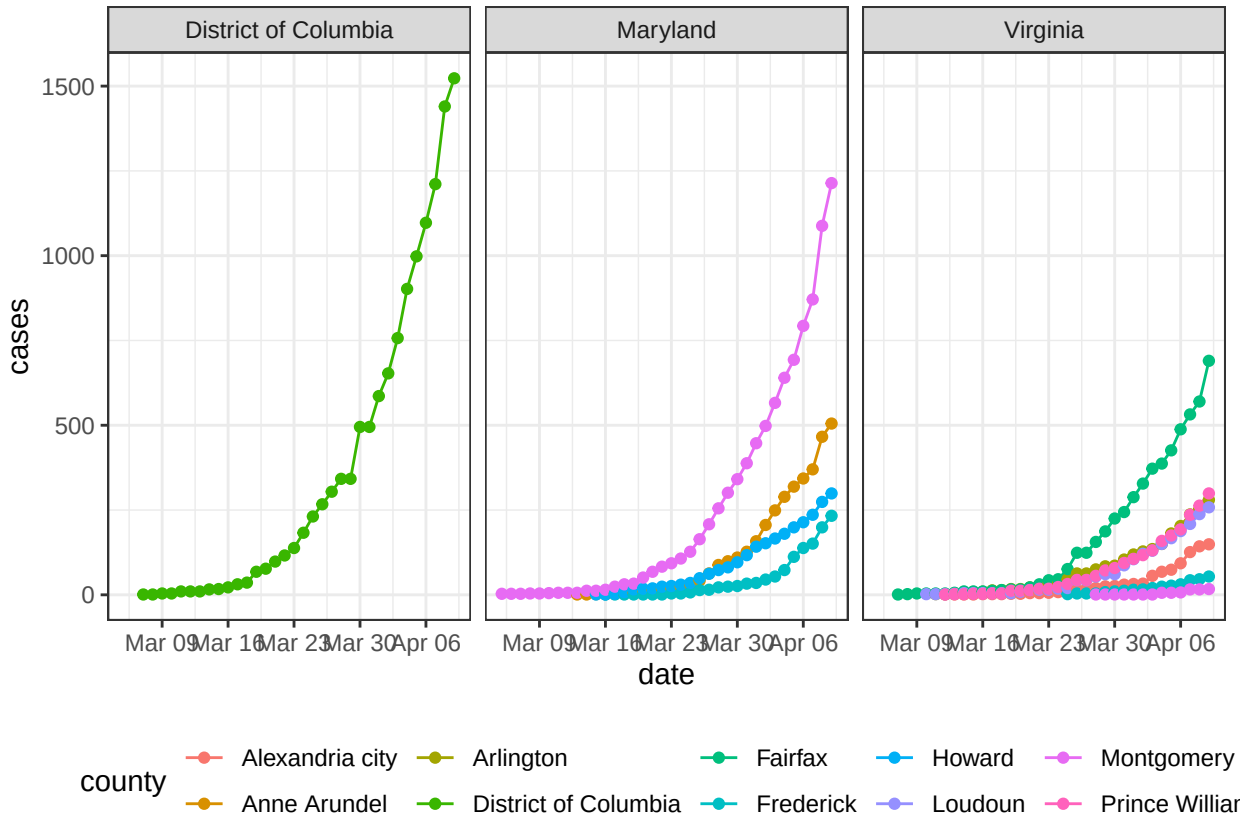
1.

```
deaths_by_state2 <- dat %>%
  count(state,date, wt = deaths) %>%
  filter(date == date) %>%
  rename(total_deaths = n) %>%
  arrange(desc(total_deaths))


cases_by_state <- dat %>%
  count(state,date, wt = cases) %>%
  filter(date == date) %>%
  rename(total_cases = n) %>%
  arrange(desc(total_cases))

total_deaths <- select(deaths_by_state2, total_deaths)
cases_by_state <- cbind(cases_by_state, total_deaths)

cases_by_state <- cases_by_state[c(1,2,4,3)]

head(cases_by_state, n = 10)
```

```
##       state       date total_deaths total_cases
## 1  New York 2020-04-09         7067      159937
## 2  New York 2020-04-08         6268      149401
## 3  New York 2020-04-07         5563      140081
## 4  New York 2020-04-06         5505      130703
## 5  New York 2020-04-05         4161      122911
## 6  New York 2020-04-04         3568      114996
## 7  New York 2020-04-03         2935      102945
## 8  New York 2020-04-02         1958       92770
## 9  New York 2020-04-01         1706       83890
## 10 New York 2020-03-31         1652       75832
```

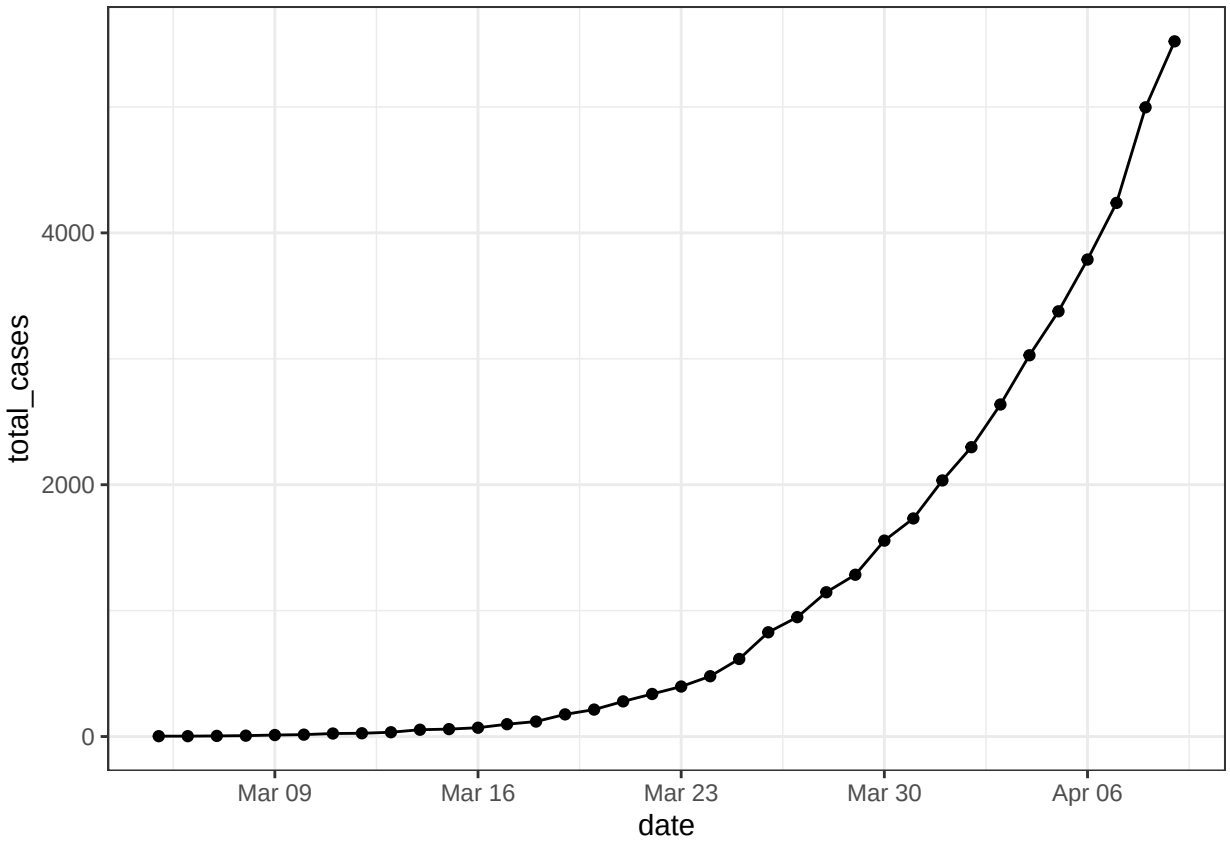
2.

```
dat_small <-
  dat %>%
  filter(state %in% c("District of Columbia", "Maryland", "Virginia")) %>%
  mutate(county = factor(county))

counties = c("Anne Arundel", "Charles city", "Frederick", "Howard", "Montgomery", "Prince Georges", "Al

dat_small %>%
  filter(county %in% counties) %>%
  ggplot(aes(x = date, y = cases, color = county)) +
  geom_line() +
  geom_point() +
  theme(legend.position = "bottom") +
  facet_grid(~ state)
```

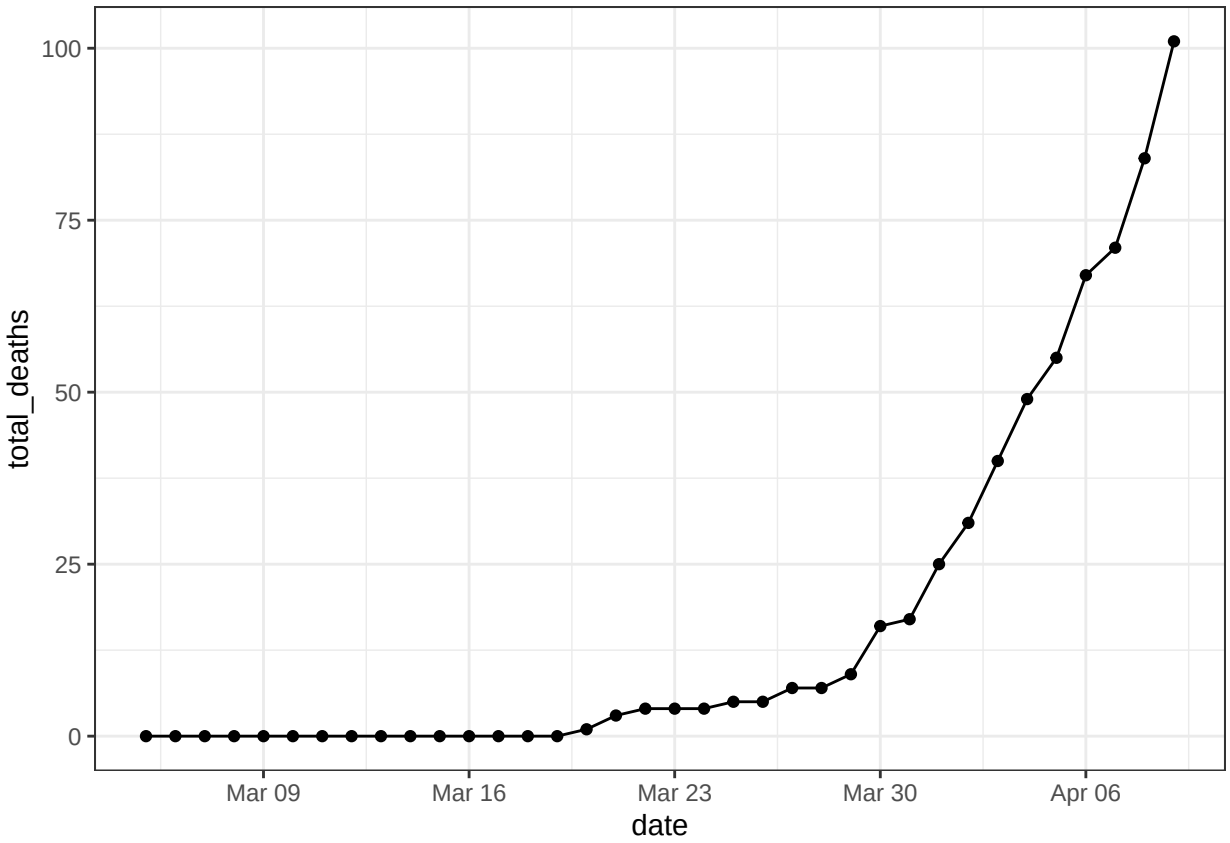


3.

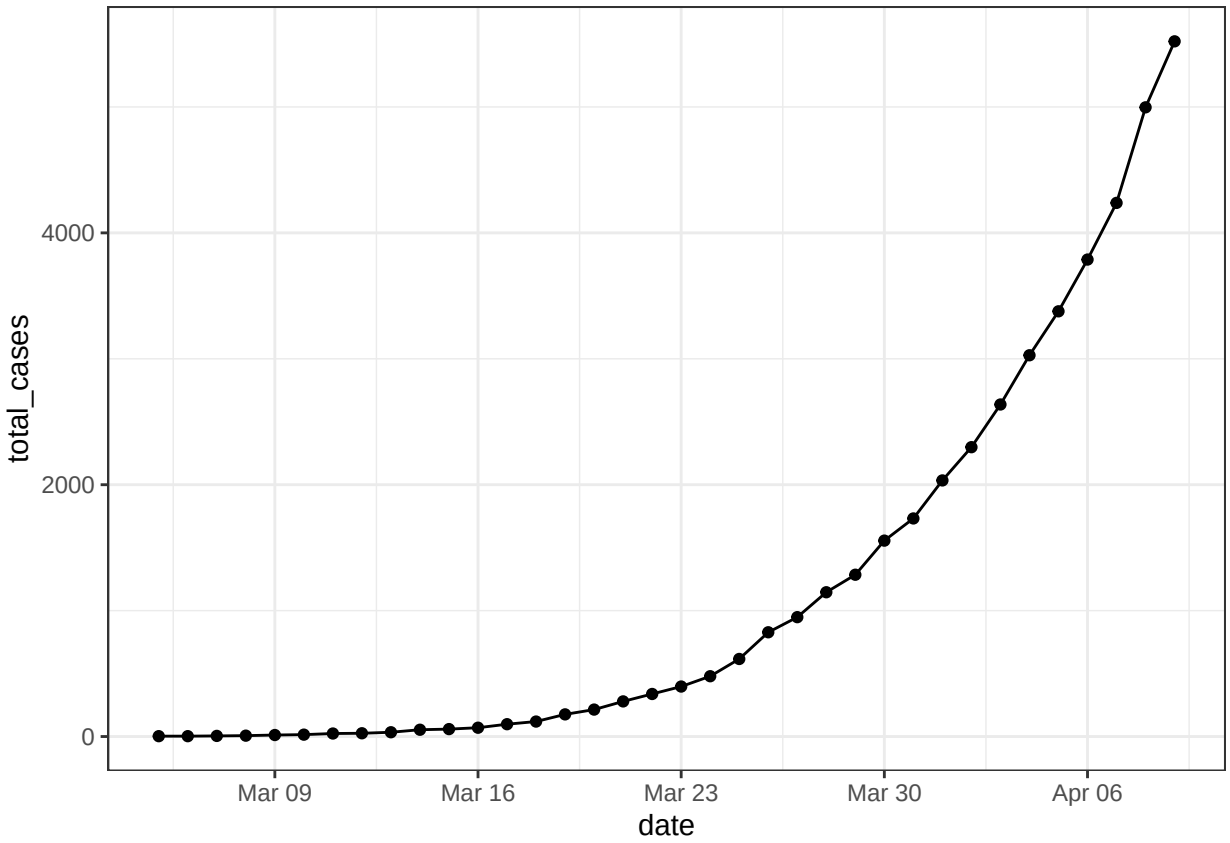
```
counties = c("Anne Arundel", "Charles city", "Frederick", "Howard", "Montgomery", "Prince Georges", "Al

dat_small %>%
  filter(county %in% counties) %>%
  group_by(date) %>%
  summarise(total_cases = sum(cases)) %>%
  ggplot(aes(x = date, y = total_cases)) +
  geom_point() +
  geom_line()
```




```
dat_small %>%
  filter(county %in% counties) %>%
  group_by(date) %>%
  summarise(total_deaths = sum(deaths)) %>%
  ggplot(aes(x = date, y = total_deaths)) +
  geom_point() +
  geom_line()
```

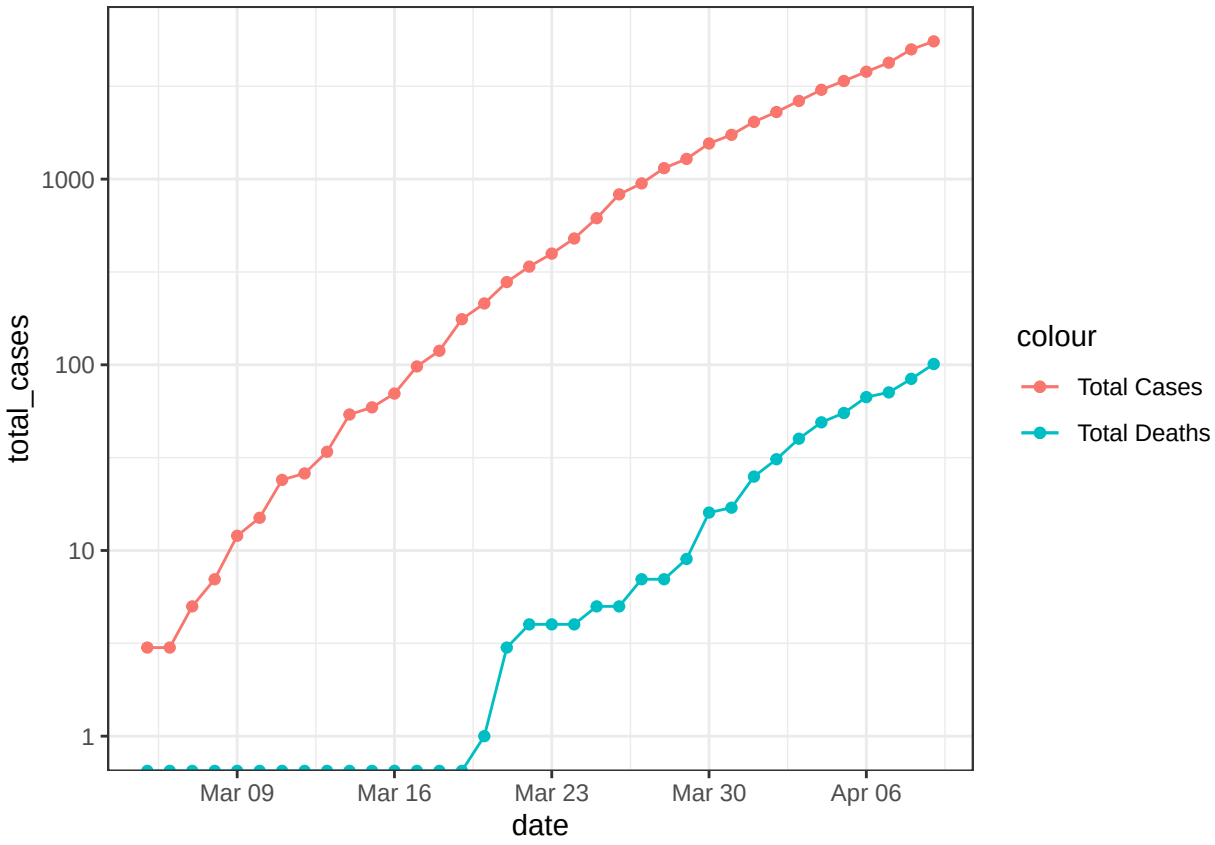



```
dat_small %>%
  filter(county %in% counties) %>%
  group_by(date) %>%
  summarise(total_cases = sum(cases)) %>%
  ggplot(aes(x = date, y = total_cases)) +
  geom_point() +
  geom_line()
```




```
dat_small %>%
  filter(county %in% counties) %>%
  group_by(date) %>%
  summarise(total_cases = sum(cases),
            total_deaths = sum(deaths)) -> dat_DMV

dat_DMV %>%
  ggplot(aes(x = date, y = total_cases, color)) +
  geom_line(aes(y = total_deaths, color = "Total Deaths")) +
  geom_line(aes(y = total_cases, color = "Total Cases")) +
  geom_point(aes(y = total_deaths, color = "Total Deaths")) +
  geom_point(aes(y = total_cases, color = "Total Cases")) +
  theme(legend.position = "right") +
  scale_y_log10()
```

```
## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous y-axis
```



4.

```
dat2 <- read_csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/archived_data/archive
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   `Province/State` = col_character(),
##   `Country/Region` = col_character()
## )

## See spec(...) for full column specifications.
```

```
Country_State <- paste(dat2$'Country/Region', dat2$'Province/State', sep = "_")
Country_State <- gsub("_NA", "    ", Country_State)
dat2 <- cbind(dat2, Country_State)

dat2 <- dat2[c(67, 1:66)]

Country_Cases <- gather(dat2, Date, Confirmed_Cases, 6:67)
Country_Cases <- Country_Cases %>%
  rename("Country/State" = Country_State)
```

```
Country_Cases$"Country/State" <- as.character(Country_Cases$"Country/State")
Country_Cases$Date <- as.Date(Country_Cases$Date, format = "%m/%d/%y")

class(Country_Cases$"Country/State")
```

```
## [1] "character"
```

```
head(filter(Country_Cases, Lat == 15.000, Long == 101.0000), n = 7)
```

```
##   Country/State Province/State Country/Region Lat Long       Date
## 1   Thailand              <NA>       Thailand  15  101 2020-01-22
## 2   Thailand              <NA>       Thailand  15  101 2020-01-23
## 3   Thailand              <NA>       Thailand  15  101 2020-01-24
## 4   Thailand              <NA>       Thailand  15  101 2020-01-25
## 5   Thailand              <NA>       Thailand  15  101 2020-01-26
## 6   Thailand              <NA>       Thailand  15  101 2020-01-27
## 7   Thailand              <NA>       Thailand  15  101 2020-01-28
##   Confirmed_Cases
## 1               2
## 2               3
## 3               5
## 4               7
## 5               8
## 6               8
## 7              14
```

5.

```
library(countrycode)
```

```
## Warning: package 'countrycode' was built under R version 3.6.3
```

```
continent <- countrycode(sourcevar = Country_Cases$"Country/Region", origin = "country.name", destinati
```

```
## Warning in countrycode(sourcevar = Country_Cases$"Country/Region", origin = "country.name", : Some va
```

```
Country_Cases <- cbind(Country_Cases, continent)

no_continent <- Country_Cases[is.na(Country_Cases$continent),]
unique(no_continent)
```

```
##                       Country/State   Province/State Country/Region     Lat
## 166   Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 433                         Kosovo               <NA>         Kosovo 42.6026
## 667   Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 934                         Kosovo               <NA>         Kosovo 42.6026
## 1168  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 1435                        Kosovo               <NA>         Kosovo 42.6026
```

```
## 1669  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 1936                      Kosovo                <NA>         Kosovo 42.6026
## 2170  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 2437                      Kosovo                <NA>         Kosovo 42.6026
## 2671  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 2938                      Kosovo                <NA>         Kosovo 42.6026
## 3172  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 3439                      Kosovo                <NA>         Kosovo 42.6026
## 3673  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 3940                      Kosovo                <NA>         Kosovo 42.6026
## 4174  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 4441                      Kosovo                <NA>         Kosovo 42.6026
## 4675  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 4942                      Kosovo                <NA>         Kosovo 42.6026
## 5176  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 5443                      Kosovo                <NA>         Kosovo 42.6026
## 5677  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 5944                      Kosovo                <NA>         Kosovo 42.6026
## 6178  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 6445                      Kosovo                <NA>         Kosovo 42.6026
## 6679  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 6946                      Kosovo                <NA>         Kosovo 42.6026
## 7180  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 7447                      Kosovo                <NA>         Kosovo 42.6026
## 7681  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 7948                      Kosovo                <NA>         Kosovo 42.6026
## 8182  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 8449                      Kosovo                <NA>         Kosovo 42.6026
## 8683  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 8950                      Kosovo                <NA>         Kosovo 42.6026
## 9184  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 9451                      Kosovo                <NA>         Kosovo 42.6026
## 9685  Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 9952                      Kosovo                <NA>         Kosovo 42.6026
## 10186 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 10453                     Kosovo                <NA>         Kosovo 42.6026
## 10687 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 10954                     Kosovo                <NA>         Kosovo 42.6026
## 11188 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 11455                     Kosovo                <NA>         Kosovo 42.6026
## 11689 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 11956                     Kosovo                <NA>         Kosovo 42.6026
## 12190 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 12457                     Kosovo                <NA>         Kosovo 42.6026
## 12691 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 12958                     Kosovo                <NA>         Kosovo 42.6026
## 13192 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 13459                     Kosovo                <NA>         Kosovo 42.6026
## 13693 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 13960                     Kosovo                <NA>         Kosovo 42.6026
## 14194 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 14461                     Kosovo                <NA>         Kosovo 42.6026
## 14695 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 14962                     Kosovo                <NA>         Kosovo 42.6026
```

```
## 15196 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 15463                     Kosovo                <NA>         Kosovo 42.6026
## 15697 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 15964                     Kosovo                <NA>         Kosovo 42.6026
## 16198 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 16465                     Kosovo                <NA>         Kosovo 42.6026
## 16699 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 16966                     Kosovo                <NA>         Kosovo 42.6026
## 17200 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 17467                     Kosovo                <NA>         Kosovo 42.6026
## 17701 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 17968                     Kosovo                <NA>         Kosovo 42.6026
## 18202 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 18469                     Kosovo                <NA>         Kosovo 42.6026
## 18703 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 18970                     Kosovo                <NA>         Kosovo 42.6026
## 19204 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 19471                     Kosovo                <NA>         Kosovo 42.6026
## 19705 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 19972                     Kosovo                <NA>         Kosovo 42.6026
## 20206 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 20473                     Kosovo                <NA>         Kosovo 42.6026
## 20707 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 20974                     Kosovo                <NA>         Kosovo 42.6026
## 21208 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 21475                     Kosovo                <NA>         Kosovo 42.6026
## 21709 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 21976                     Kosovo                <NA>         Kosovo 42.6026
## 22210 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 22477                     Kosovo                <NA>         Kosovo 42.6026
## 22711 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 22978                     Kosovo                <NA>         Kosovo 42.6026
## 23212 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 23479                     Kosovo                <NA>         Kosovo 42.6026
## 23713 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 23980                     Kosovo                <NA>         Kosovo 42.6026
## 24214 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 24481                     Kosovo                <NA>         Kosovo 42.6026
## 24715 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 24982                     Kosovo                <NA>         Kosovo 42.6026
## 25216 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 25483                     Kosovo                <NA>         Kosovo 42.6026
## 25717 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 25984                     Kosovo                <NA>         Kosovo 42.6026
## 26218 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 26485                     Kosovo                <NA>         Kosovo 42.6026
## 26719 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 26986                     Kosovo                <NA>         Kosovo 42.6026
## 27220 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 27487                     Kosovo                <NA>         Kosovo 42.6026
## 27721 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 27988                     Kosovo                <NA>         Kosovo 42.6026
## 28222 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 28489                     Kosovo                <NA>         Kosovo 42.6026
```

```
## 28723 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 28990                     Kosovo               <NA>         Kosovo 42.6026
## 29224 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 29491                     Kosovo               <NA>         Kosovo 42.6026
## 29725 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 29992                     Kosovo               <NA>         Kosovo 42.6026
## 30226 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 30493                     Kosovo               <NA>         Kosovo 42.6026
## 30727 Cruise Ship_Diamond Princess Diamond Princess    Cruise Ship 35.4437
## 30994                     Kosovo               <NA>         Kosovo 42.6026
##          Long       Date Confirmed_Cases continent
## 166   139.638 2020-01-22               0      <NA>
## 433    20.903 2020-01-22               0      <NA>
## 667   139.638 2020-01-23               0      <NA>
## 934    20.903 2020-01-23               0      <NA>
## 1168  139.638 2020-01-24               0      <NA>
## 1435   20.903 2020-01-24               0      <NA>
## 1669  139.638 2020-01-25               0      <NA>
## 1936   20.903 2020-01-25               0      <NA>
## 2170  139.638 2020-01-26               0      <NA>
## 2437   20.903 2020-01-26               0      <NA>
## 2671  139.638 2020-01-27               0      <NA>
## 2938   20.903 2020-01-27               0      <NA>
## 3172  139.638 2020-01-28               0      <NA>
## 3439   20.903 2020-01-28               0      <NA>
## 3673  139.638 2020-01-29               0      <NA>
## 3940   20.903 2020-01-29               0      <NA>
## 4174  139.638 2020-01-30               0      <NA>
## 4441   20.903 2020-01-30               0      <NA>
## 4675  139.638 2020-01-31               0      <NA>
## 4942   20.903 2020-01-31               0      <NA>
## 5176  139.638 2020-02-01               0      <NA>
## 5443   20.903 2020-02-01               0      <NA>
## 5677  139.638 2020-02-02               0      <NA>
## 5944   20.903 2020-02-02               0      <NA>
## 6178  139.638 2020-02-03               0      <NA>
## 6445   20.903 2020-02-03               0      <NA>
## 6679  139.638 2020-02-04               0      <NA>
## 6946   20.903 2020-02-04               0      <NA>
## 7180  139.638 2020-02-05               0      <NA>
## 7447   20.903 2020-02-05               0      <NA>
## 7681  139.638 2020-02-06               0      <NA>
## 7948   20.903 2020-02-06               0      <NA>
## 8182  139.638 2020-02-07              61      <NA>
## 8449   20.903 2020-02-07               0      <NA>
## 8683  139.638 2020-02-08              61      <NA>
## 8950   20.903 2020-02-08               0      <NA>
## 9184  139.638 2020-02-09              64      <NA>
## 9451   20.903 2020-02-09               0      <NA>
## 9685  139.638 2020-02-10             135      <NA>
## 9952   20.903 2020-02-10               0      <NA>
## 10186 139.638 2020-02-11             135      <NA>
## 10453  20.903 2020-02-11               0      <NA>
## 10687 139.638 2020-02-12             175      <NA>
```

```
## 10954  20.903 2020-02-12              0      <NA>
## 11188 139.638 2020-02-13            175      <NA>
## 11455  20.903 2020-02-13              0      <NA>
## 11689 139.638 2020-02-14            218      <NA>
## 11956  20.903 2020-02-14              0      <NA>
## 12190 139.638 2020-02-15            285      <NA>
## 12457  20.903 2020-02-15              0      <NA>
## 12691 139.638 2020-02-16            355      <NA>
## 12958  20.903 2020-02-16              0      <NA>
## 13192 139.638 2020-02-17            454      <NA>
## 13459  20.903 2020-02-17              0      <NA>
## 13693 139.638 2020-02-18            542      <NA>
## 13960  20.903 2020-02-18              0      <NA>
## 14194 139.638 2020-02-19            621      <NA>
## 14461  20.903 2020-02-19              0      <NA>
## 14695 139.638 2020-02-20            634      <NA>
## 14962  20.903 2020-02-20              0      <NA>
## 15196 139.638 2020-02-21            634      <NA>
## 15463  20.903 2020-02-21              0      <NA>
## 15697 139.638 2020-02-22            634      <NA>
## 15964  20.903 2020-02-22              0      <NA>
## 16198 139.638 2020-02-23            691      <NA>
## 16465  20.903 2020-02-23              0      <NA>
## 16699 139.638 2020-02-24            691      <NA>
## 16966  20.903 2020-02-24              0      <NA>
## 17200 139.638 2020-02-25            691      <NA>
## 17467  20.903 2020-02-25              0      <NA>
## 17701 139.638 2020-02-26            705      <NA>
## 17968  20.903 2020-02-26              0      <NA>
## 18202 139.638 2020-02-27            705      <NA>
## 18469  20.903 2020-02-27              0      <NA>
## 18703 139.638 2020-02-28            705      <NA>
## 18970  20.903 2020-02-28              0      <NA>
## 19204 139.638 2020-02-29            705      <NA>
## 19471  20.903 2020-02-29              0      <NA>
## 19705 139.638 2020-03-01            705      <NA>
## 19972  20.903 2020-03-01              0      <NA>
## 20206 139.638 2020-03-02            705      <NA>
## 20473  20.903 2020-03-02              0      <NA>
## 20707 139.638 2020-03-03            706      <NA>
## 20974  20.903 2020-03-03              0      <NA>
## 21208 139.638 2020-03-04            706      <NA>
## 21475  20.903 2020-03-04              0      <NA>
## 21709 139.638 2020-03-05            706      <NA>
## 21976  20.903 2020-03-05              0      <NA>
## 22210 139.638 2020-03-06            696      <NA>
## 22477  20.903 2020-03-06              0      <NA>
## 22711 139.638 2020-03-07            696      <NA>
## 22978  20.903 2020-03-07              0      <NA>
## 23212 139.638 2020-03-08            696      <NA>
## 23479  20.903 2020-03-08              0      <NA>
## 23713 139.638 2020-03-09            696      <NA>
## 23980  20.903 2020-03-09              0      <NA>
## 24214 139.638 2020-03-10            696      <NA>
```

```
## 24481  20.903 2020-03-10              0      <NA>
## 24715 139.638 2020-03-11            696      <NA>
## 24982  20.903 2020-03-11              0      <NA>
## 25216 139.638 2020-03-12            696      <NA>
## 25483  20.903 2020-03-12              0      <NA>
## 25717 139.638 2020-03-13            696      <NA>
## 25984  20.903 2020-03-13              0      <NA>
## 26218 139.638 2020-03-14            696      <NA>
## 26485  20.903 2020-03-14              0      <NA>
## 26719 139.638 2020-03-15            696      <NA>
## 26986  20.903 2020-03-15              2      <NA>
## 27220 139.638 2020-03-16            696      <NA>
## 27487  20.903 2020-03-16              2      <NA>
## 27721 139.638 2020-03-17            696      <NA>
## 27988  20.903 2020-03-17              2      <NA>
## 28222 139.638 2020-03-18            712      <NA>
## 28489  20.903 2020-03-18              2      <NA>
## 28723 139.638 2020-03-19            712      <NA>
## 28990  20.903 2020-03-19              2      <NA>
## 29224 139.638 2020-03-20            712      <NA>
## 29491  20.903 2020-03-20              2      <NA>
## 29725 139.638 2020-03-21            712      <NA>
## 29992  20.903 2020-03-21              2      <NA>
## 30226 139.638 2020-03-22            712      <NA>
## 30493  20.903 2020-03-22              2      <NA>
## 30727 139.638 2020-03-23            712      <NA>
## 30994  20.903 2020-03-23              2      <NA>
```

```
## Looking through the dataframe created the two NAs in the column continent are associated with the
## Cruise Ship "Diamond Princess" as well as the country of Kosovo. Kosovo is in Europe and this
## issue can easily be remedied. In order to find where the Diamond Princess was docked on the dates
## provided, we can use the Latitude and Longitudes given which places the ship in the city of Yokohama
## in the Kanegawa Prefecture of Japan. Therefore, the ship is in Asia.

Country_Cases$continent[which(Country_Cases$"Country/State" == "Kosovo   ")] <- "Europe"
Country_Cases$continent[which(Country_Cases$"Country/State" == "Cruise Ship_Diamond Princess")] <- "Asi

sum(is.na(Country_Cases$continent))
```

```
## [1] 0
```

6.

```
Top_25 <- Country_Cases %>%
          arrange(desc(Confirmed_Cases)) %>%
          arrange(desc(Date)) %>%
          head(n = 25) %>%
          select("Country/State", "Country/Region", continent, Confirmed_Cases)

names(Top_25) <- c("Country/State", "Country/Region", "continent", "ttl")
Top_25$continent <- as.character(Top_25$continent)
```

```
Top_25
```

```
##                     Country/State Country/Region continent   ttl
## 1                     China_Hubei          China      Asia 67800
## 2                         Italy             Italy    Europe 59138
## 3                         Spain             Spain    Europe 28768
## 4                       Germany           Germany    Europe 24873
## 5                          Iran              Iran      Asia 21638
## 6                   France_France          France    Europe 16018
## 7                     US_New York              US  Americas 15793
## 8                 Korea, South      Korea, South      Asia  8897
## 9                   Switzerland       Switzerland    Europe  7245
## 10 United Kingdom_United Kingdom United Kingdom    Europe  5683
## 11       Netherlands_Netherlands    Netherlands    Europe  4204
## 12                       Belgium           Belgium    Europe  3401
## 13                       Austria           Austria    Europe  3244
## 14                        Norway            Norway    Europe  2383
## 15                  US_Washington              US  Americas  1996
## 16                        Sweden            Sweden    Europe  1934
## 17                  US_New Jersey              US  Americas  1914
## 18                  US_California              US  Americas  1642
## 19                     Portugal          Portugal    Europe  1600
## 20                       Brazil            Brazil  Americas  1593
## 21                 China_Guangdong           China      Asia  1413
## 22                 Denmark_Denmark        Denmark    Europe  1395
## 23                      Malaysia          Malaysia      Asia  1306
## 24                     China_Henan           China      Asia  1274
## 25                   China_Zhejiang           China      Asia  1238
```

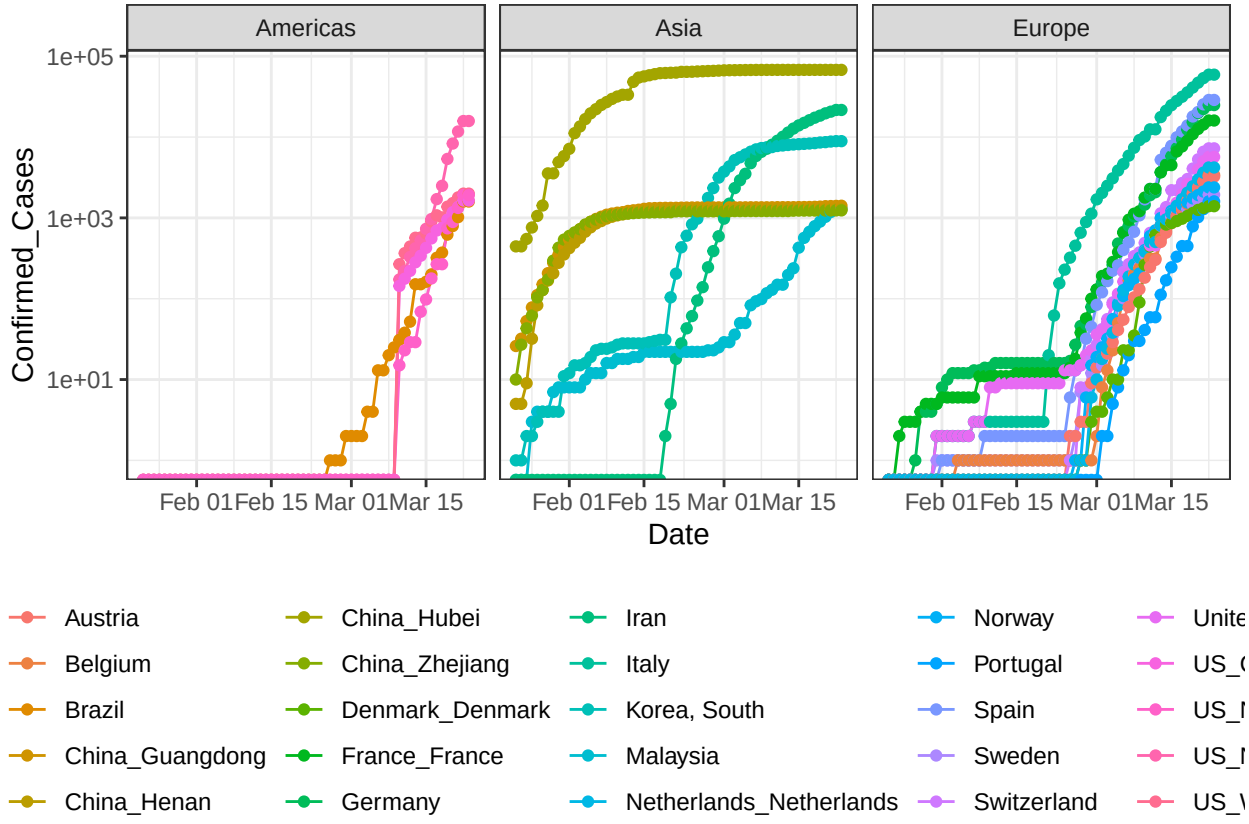
7.

```
Cases_Small <-
  Country_Cases %>%
  filter(`Country/Region` %in% c("China", "Italy", "Spain", "Germany", "Iran", "France",
                                 "US", "Korea, South", "Switzerland", "United Kingdom",
                                 "Netherlands", "Belgium", "Austria", "Norway", "Sweden",
                                 "Portugal", "Brazil", "Denmark", "Malaysia")) %>%
  mutate(`Country/State` = factor(`Country/State`))

countries = c("China_Hubei", "Italy   ", "Spain   ", "Germany   ", "Iran   ", "France_France", "US_New
              "Korea, South   ", "Switzerland   ", "United Kingdom_United Kingdom", "Netherlands_Neth
              "Belgium   ", "Austria   ", "Norway   ", "US_Washington", "Sweden   ", "US_New Jersey",
              "Portugal   ", "Brazil   ", "China_Guangdong", "Denmark_Denmark", "Malaysia   ", "China
              "China_Zhejiang")

WorldGraph <- Cases_Small %>%
  filter(`Country/State` %in% countries) %>%
  ggplot(aes(x = Date, y = Confirmed_Cases, color = `Country/State`)) +
  geom_line() +
  geom_point() +
  theme(legend.position = "bottom") +
```

```
  facet_grid(~ continent) +
  scale_y_log10()

WorldGraph
```

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

```
## Warning: Transformation introduced infinite values in continuous y-axis
```



8.

```
library(lubridate)
```

```
## Warning: package 'lubridate' was built under R version 3.6.2
```

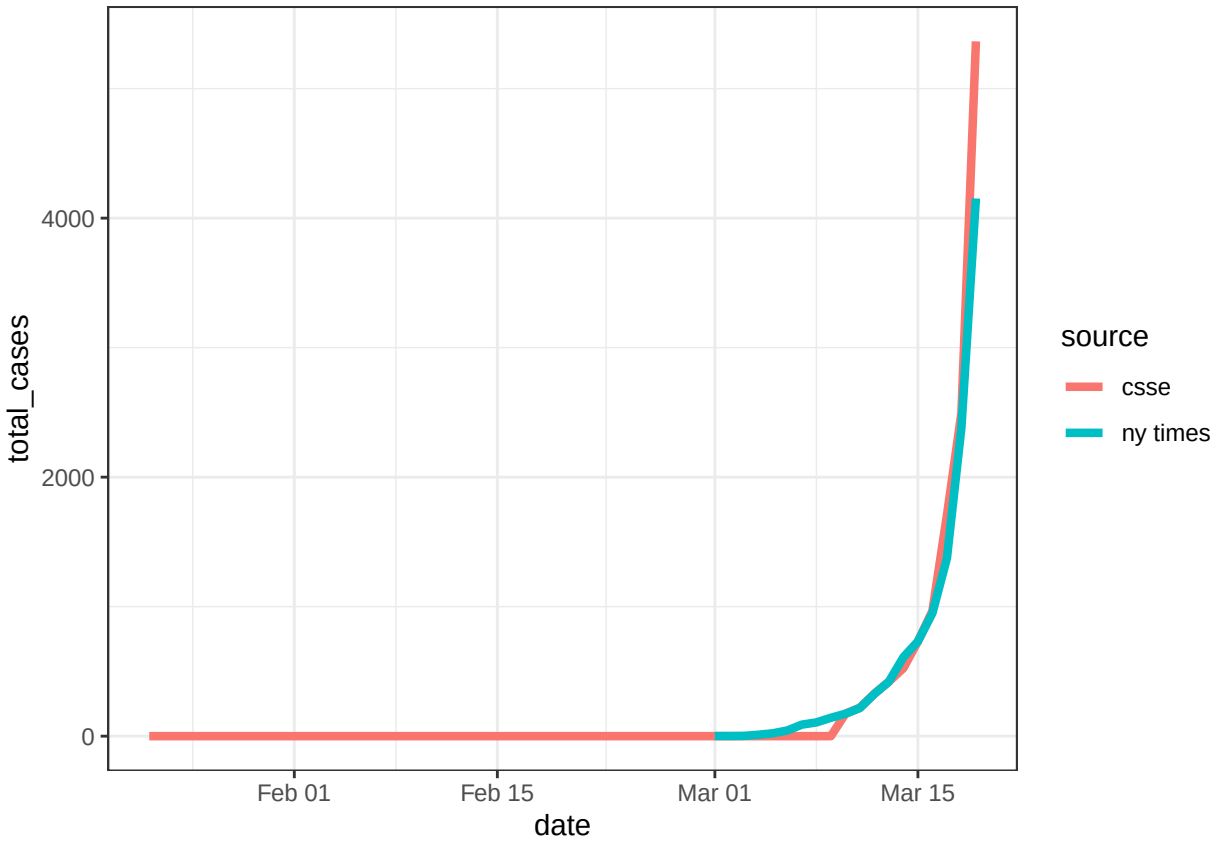
```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
dat %>%
  filter(state == "New York") %>%
  select(cases, date) %>%
  group_by(date) %>%
  summarize(total_cases = sum(cases)) %>%
  mutate(source = "ny times") ->
  ny_ny

Country_Cases %>%
  filter(`Province/State` == "New York") %>%
  select(Confirmed_Cases, Date) %>%
  group_by(Date) %>%
  summarise(total_cases = sum(Confirmed_Cases)) %>%
  mutate(source = "csse") %>%
  rename(date = Date)->
  csse_ny

csse_ny %>%
  full_join(ny_ny, Date = date) ->
  full_ny
```

```
## Joining, by = c("date", "total_cases", "source")
```

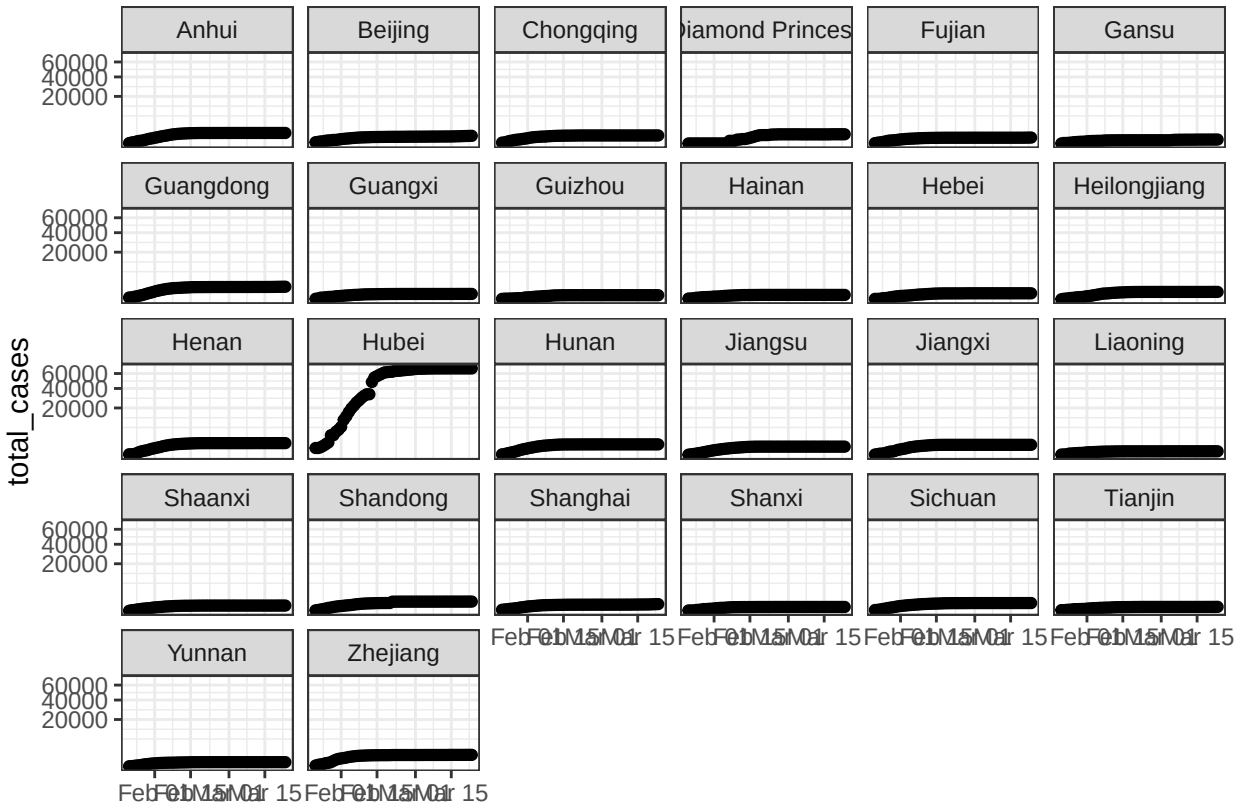
```
full_ny %>%
  filter(date < ymd("2020-03-20")) %>%
  ggplot(aes(date, total_cases, color = source)) +
  geom_line(size = 1.5)
```



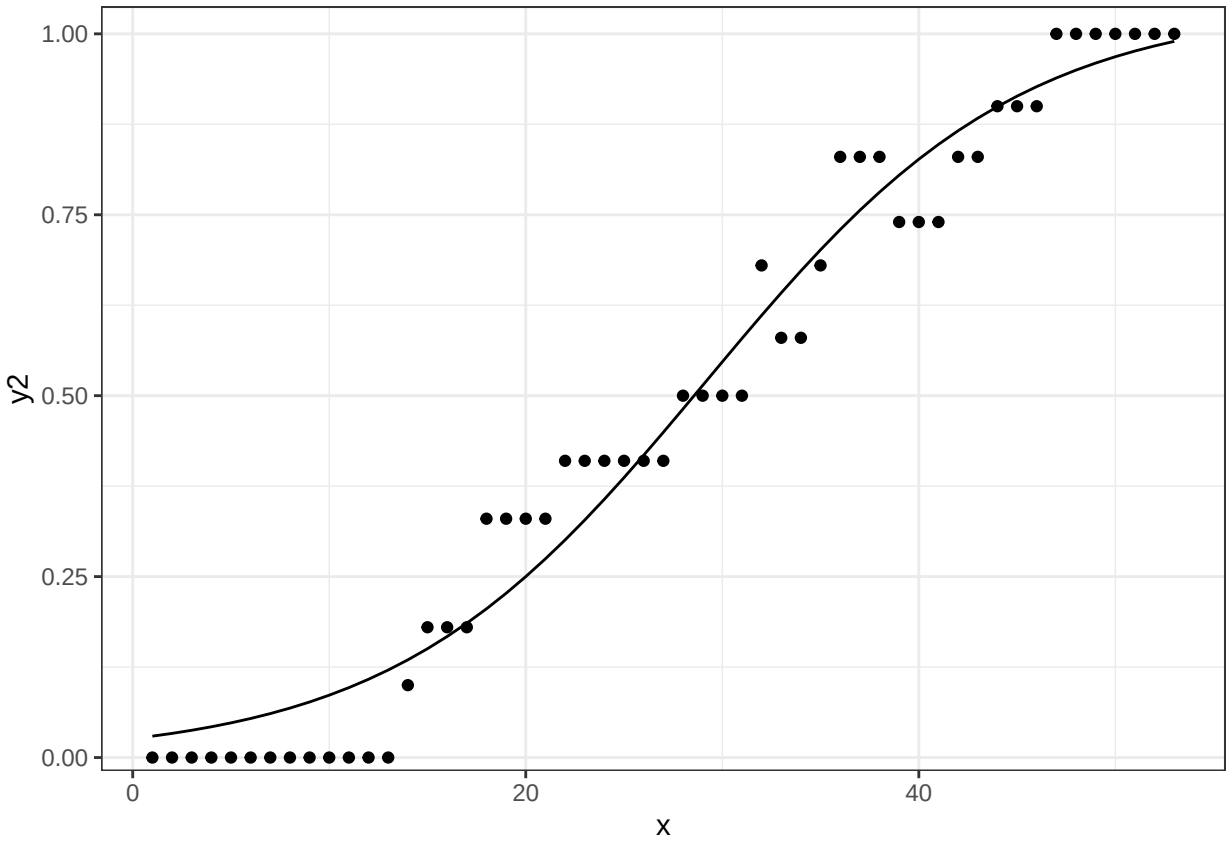
The differences can be found from the time of March 5 to March 15.

9.

```
Country_Cases %>%
  filter(`Country/Region` == "China") %>%
  group_by(`Province/State`) %>%
  summarise(total_cases = sum(Confirmed_Cases)) %>%
  ungroup() %>%
  arrange(desc(total_cases)) %>%
  slice(1:25) %>%
  pull(`Province/State`) ->
  province_list

Country_Cases %>%
  filter(`Province/State` %in% province_list | `Province/State` == "Diamond Princess") %>%
  group_by(Date, `Province/State`) %>%
  summarise(total_cases = sum(Confirmed_Cases)) %>%
  ggplot(aes(Date, total_cases)) +
  geom_point() +
  facet_wrap(~ `Province/State`) +
  scale_y_sqrt()
```




```
sigmoid = function(x, params) {
  params[1] / (1 + exp(-params[2] * (x - params[3])))
}

x = 1:53
y = c(0,0,0,0,0,0,0,0,0,0,0,0,0,0.1,0.18,0.18,0.18,0.33,0.33,0.33,0.33,0.41,
  0.41,0.41,0.41,0.41,0.41,0.5,0.5,0.5,0.5,0.68,0.58,0.58,0.68,0.83,0.83,0.83,
  0.74,0.74,0.74,0.83,0.83,0.9,0.9,0.9,1,1,1,1,1,1,1)
df <- tibble(x = x, y = y)
# fitting code
fitmodel <- nls(y ~ a /(1 + exp(-b * (x - c))), data = df,
                start = list(a = 1, b = 0.5, c = 25))


# get the coefficients using the coef function
params=coef(fitmodel)

df$y2 <- sigmoid(x, params)
df %>% ggplot(aes(x, y2)) + geom_line() + geom_point(y = y)
```

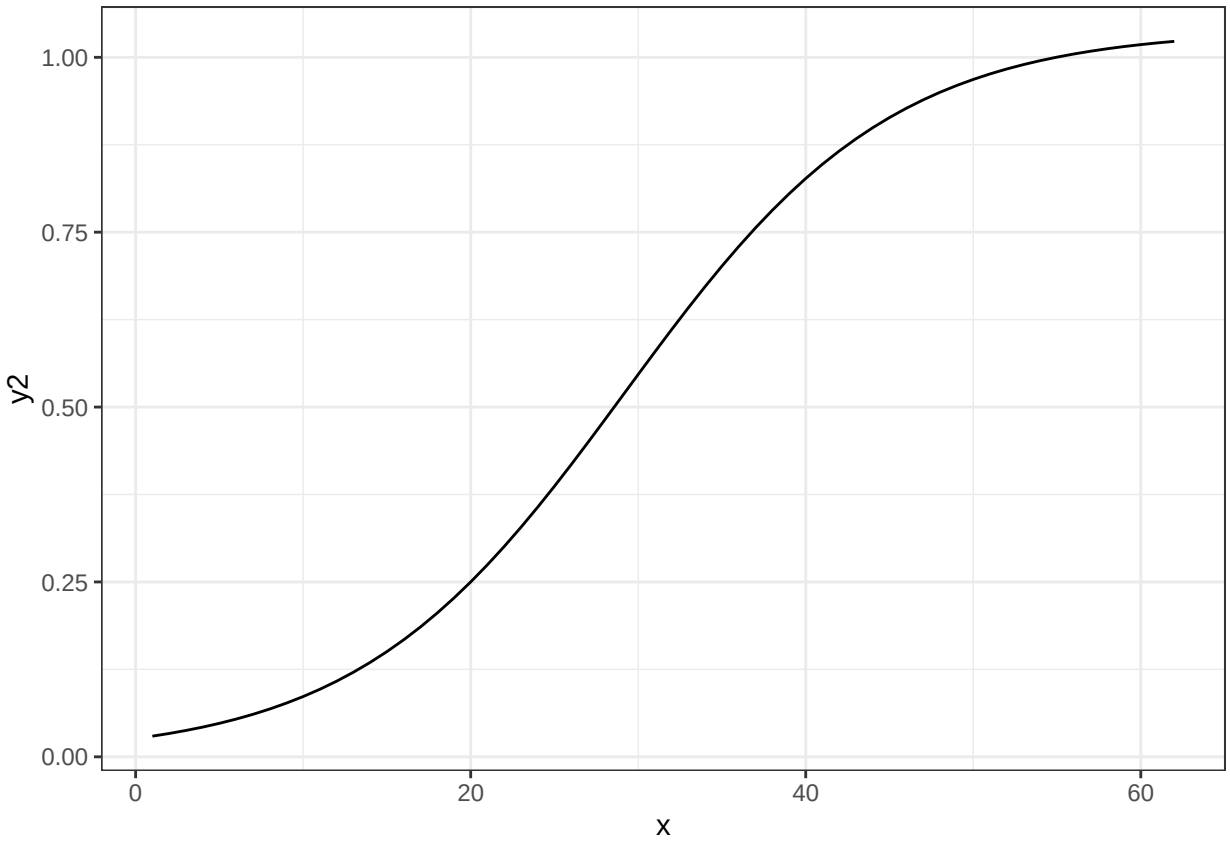


Exercise 10

```
Country_Cases %>%
  filter(`Province/State` == "Hubei") %>%
  group_by(`Date`) %>%
  summarise(total_cases = sum(Confirmed_Cases)) ->
  hubei

x = seq(length(hubei$Date))
y = hubei$total_cases
df <- tibble(x = x, y = y)
# fitting code
fitmodel <- nls(y ~ a /(1 + exp(-b * (x - c))), data = df,
                start = list(a = 1, b = 0.5, c = 25))
```

```
## Error in nls(y ~ a/(1 + exp(-b * (x - c))), data = df, start = list(a = 1, : singular gradient
```

```
# visualization code
# get the coefficients using the coef function
params=coef(fitmodel)

df$y2 <- sigmoid(x, params)
df %>% ggplot(aes(x, y2)) + geom_line() + geom_point(y = y)
```


1.00
0.75
0.50
0.25
0.00
y2
0
20
40
60
x