

# Progetto Lab 2

Martina Crisafulli, Martina Manno & Enrico Grandi

12/12/2021

## ##INTRODUCTION

In order to analyze the dataset these packages are required: ggplot2, dplyr, magrittr. Then we import the dataset. This dataset is about Covid-19 pandemic going from 31/12/2019 to 14/12/2020 and has 12 variables: date, day, month, year, cases, deaths, country, code (of the country), population, continent, cases\_cum (cumulative cases) and deaths\_cum (cumulative deaths).

```
knitr::opts_chunk$set(echo = TRUE)
library(ggplot2)
library(dplyr)
library(magrittr)

df <- read.delim("/Users/martina/Desktop/covid.csv", sep=";")
```

## DATA ANALYSIS

This analysis will focus on registered deaths and cases of Covid-19 in five continents: Europe, Asia, Africa, Oceania and America. Since it is a time series, the “date” variable needs to be transformed into the correct format so that the visualization of data is correct.

```
ord_df<- df[order(as.Date(df$date, format="%d/%m/%Y")),]
ord_df$date <- as.Date(ord_df$date,format = "%d/%m/%Y")

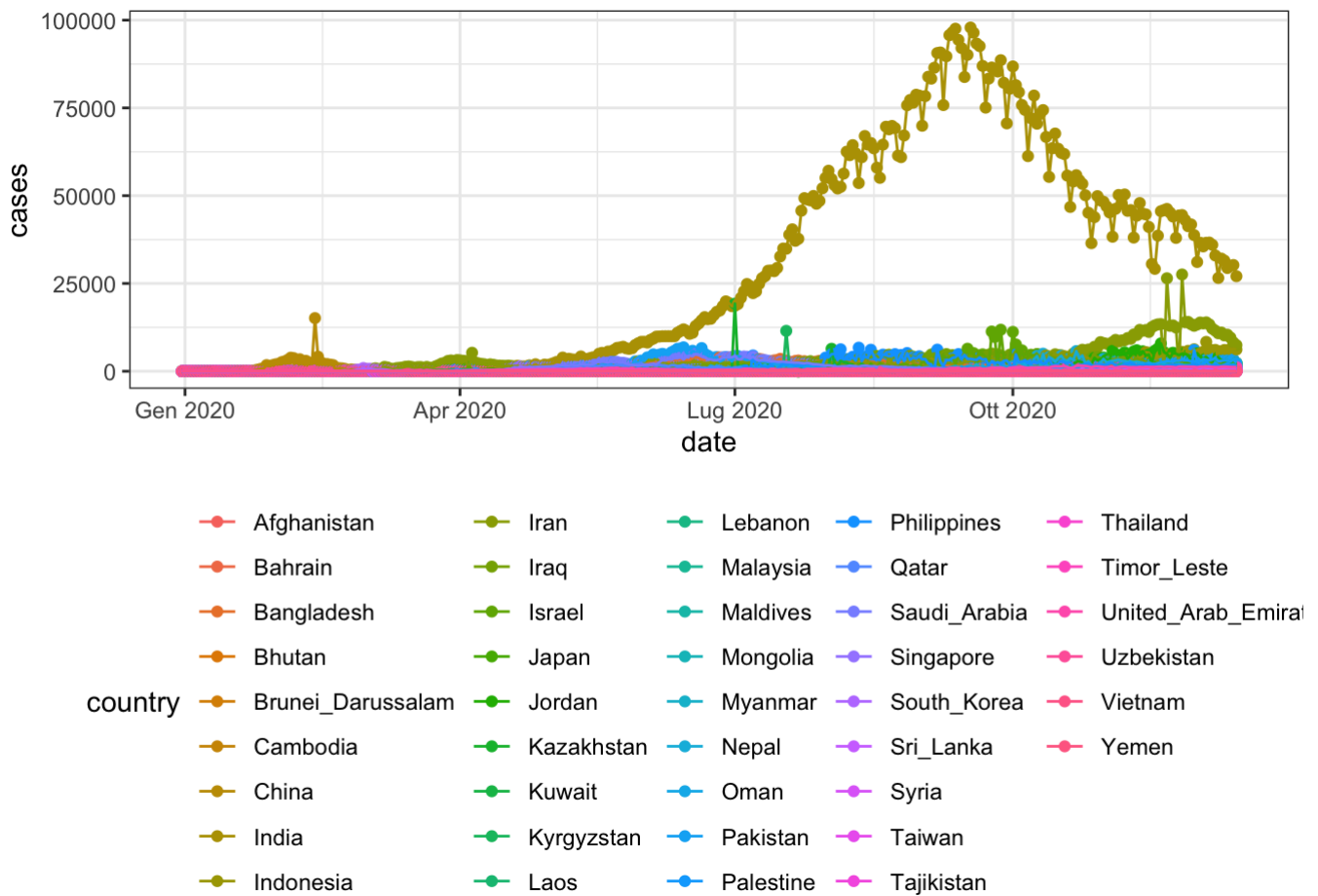
format(as.Date(ord_df$date), "%m/%Y")
```

## DATA VISUALIZATION AND MANIPULATION

### ##Cases of Covid-19

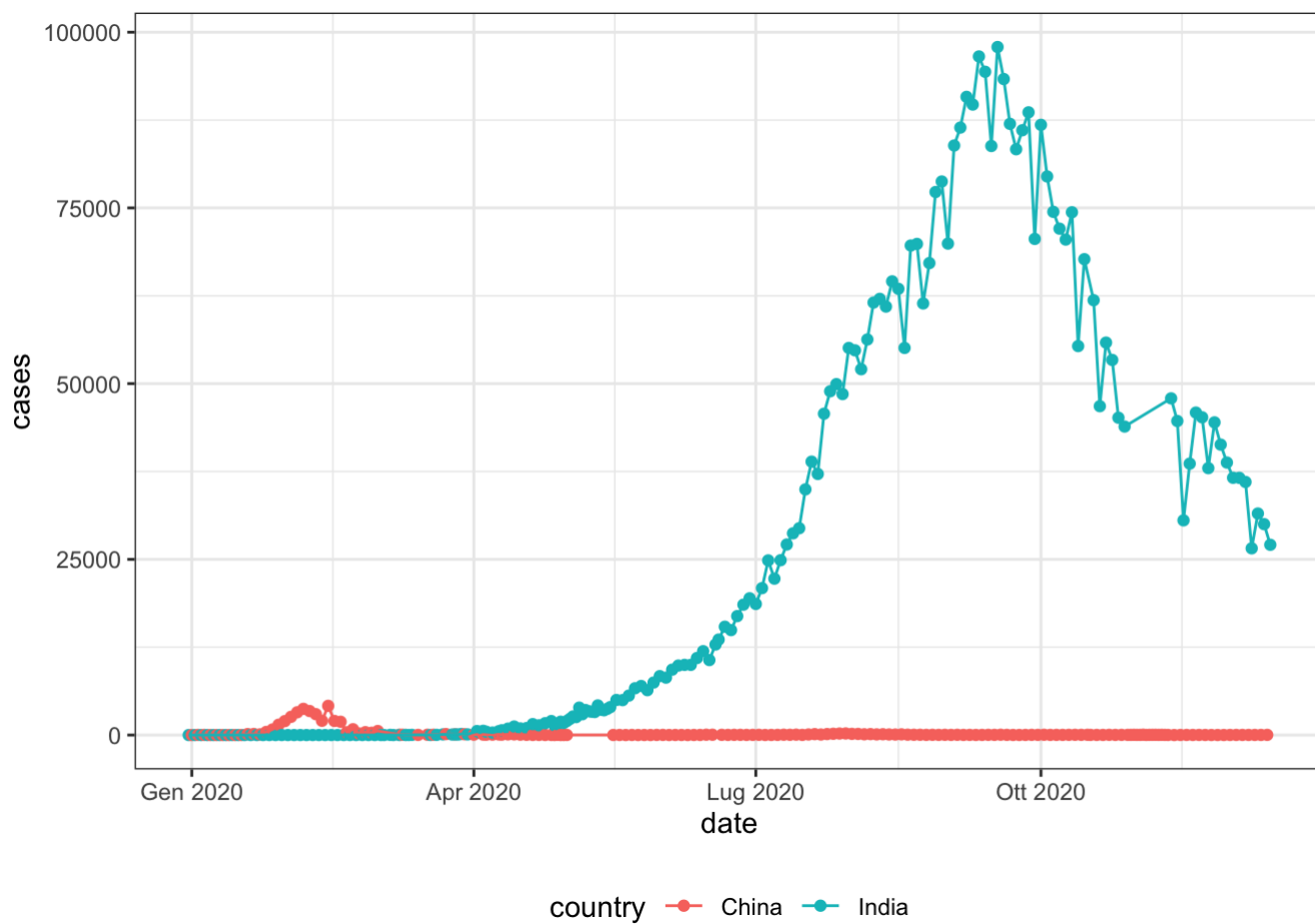
The visualization process starts at continent-level and then focuses on the most relevant and affected countries.

```
#plot of Asia 2019-2020 daily cases, on y cases, on x months
ord_df %>% filter(continent == "Asia") %>%
  ggplot(aes(x = date, y = cases, color = country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
```

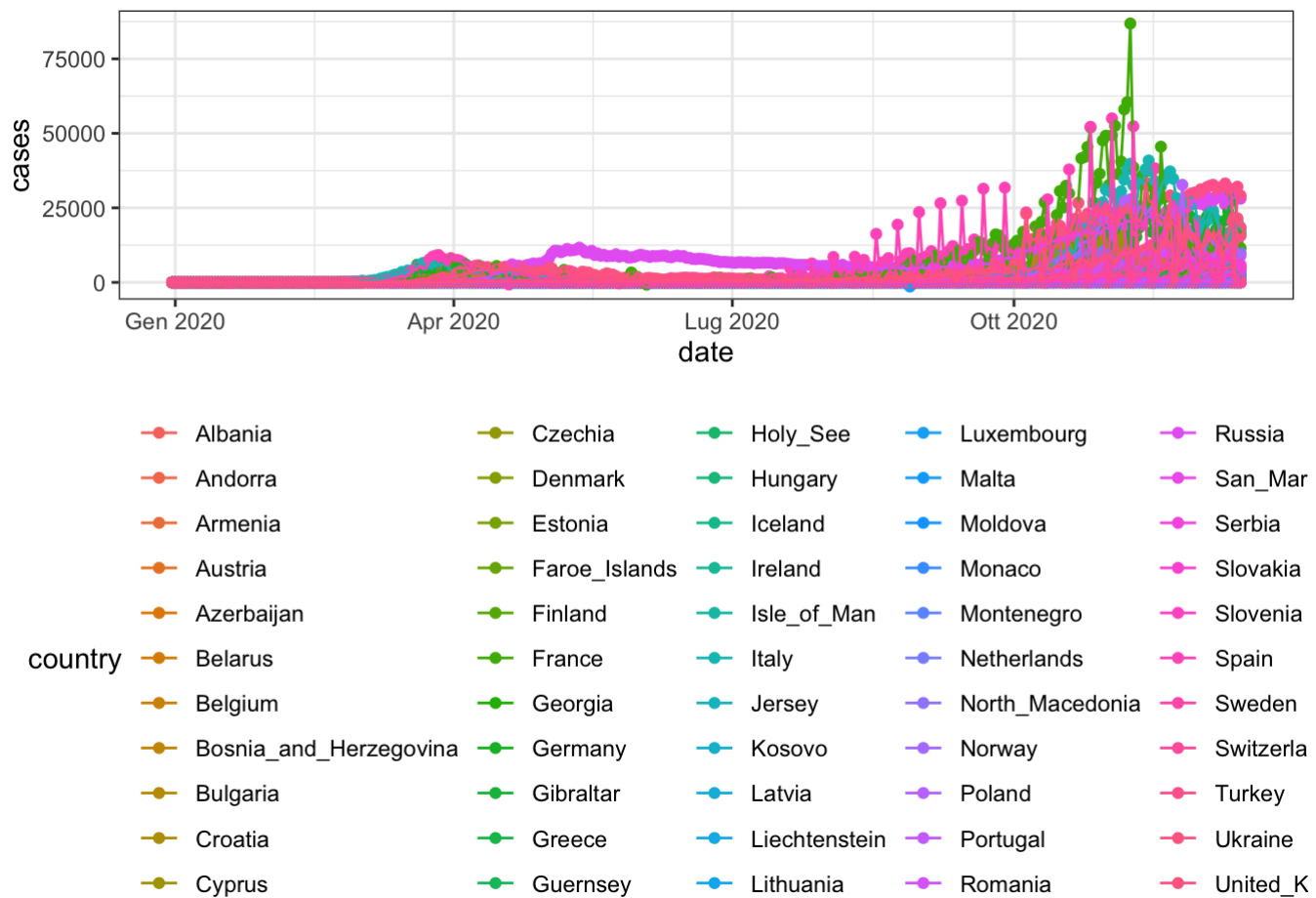


```
#we focus on China and India
```

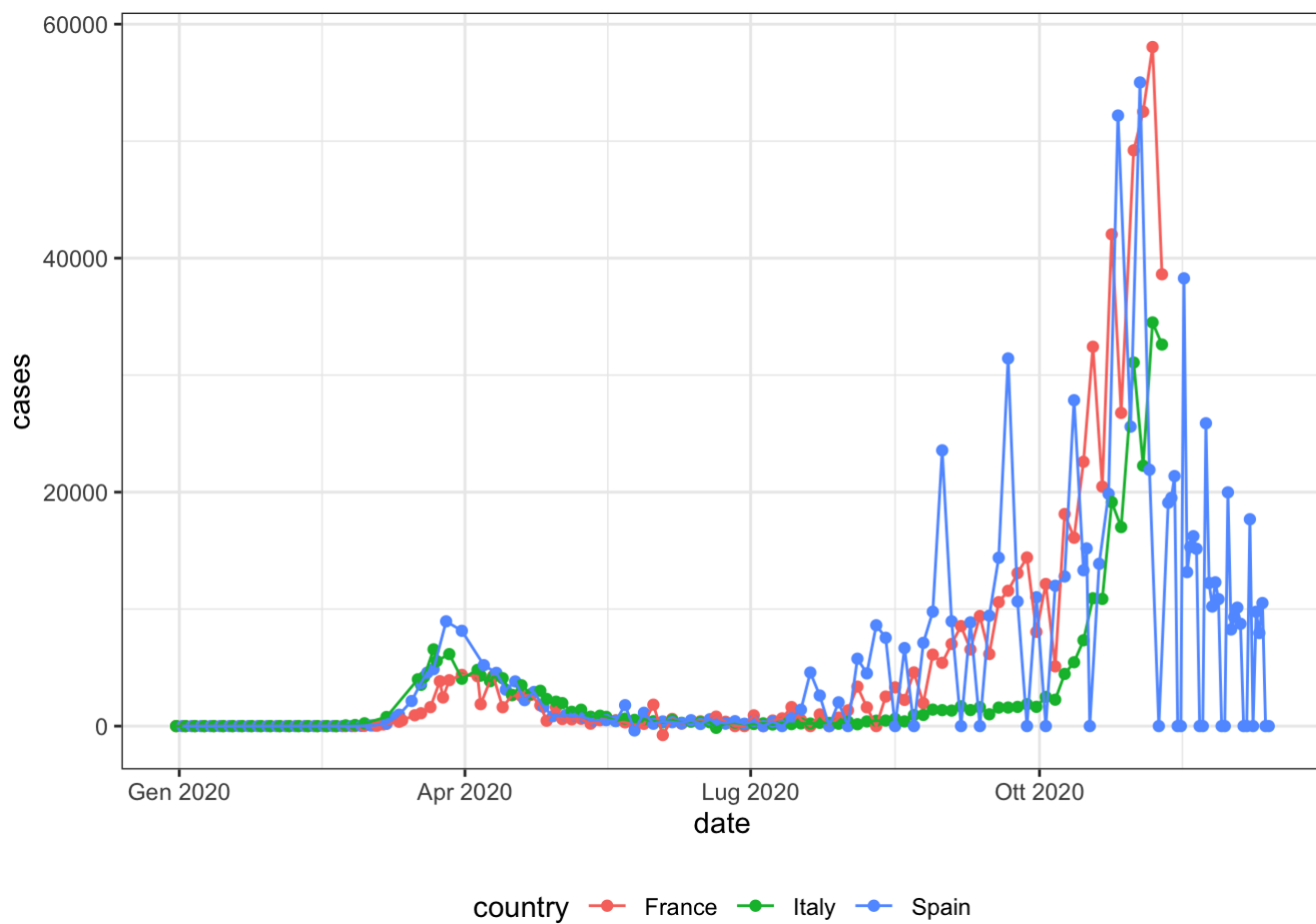
```
c<- ord_df %>% filter(country == c("China","India")) %>%
  ggplot(aes(x = date, y = cases, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
c
```



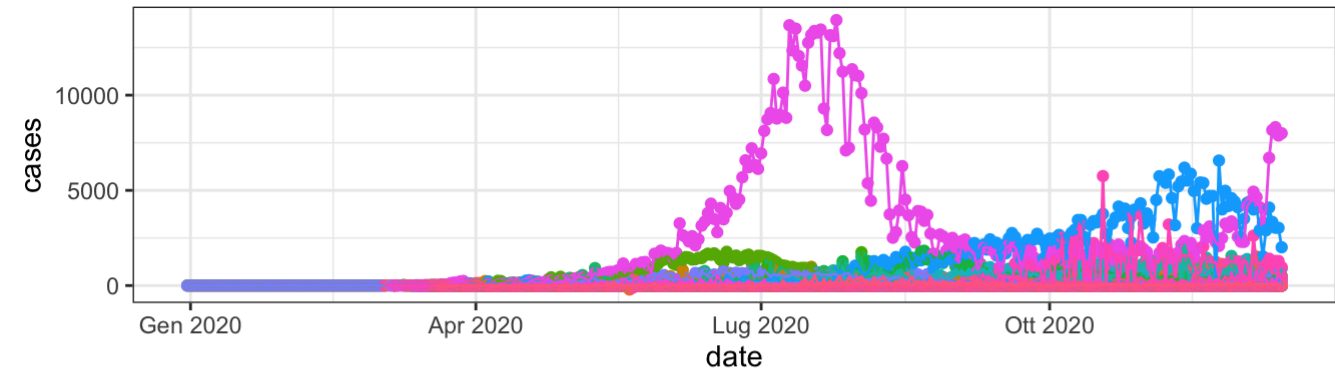
```
#plot of Europe  
ord_df %>% filter(continent == "Europe") %>%  
  ggplot(aes(x = date, y = cases, color = country)) + geom_point() + geom_line() +  
  theme_bw() +  
  theme(legend.position = "bottom")
```



```
#plot of France, Italy and Spain
e<- ord_df %>% filter(country == c("Italy", "France", "Spain"))%>%
  ggplot(aes(x = date, y = cases, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
e
```

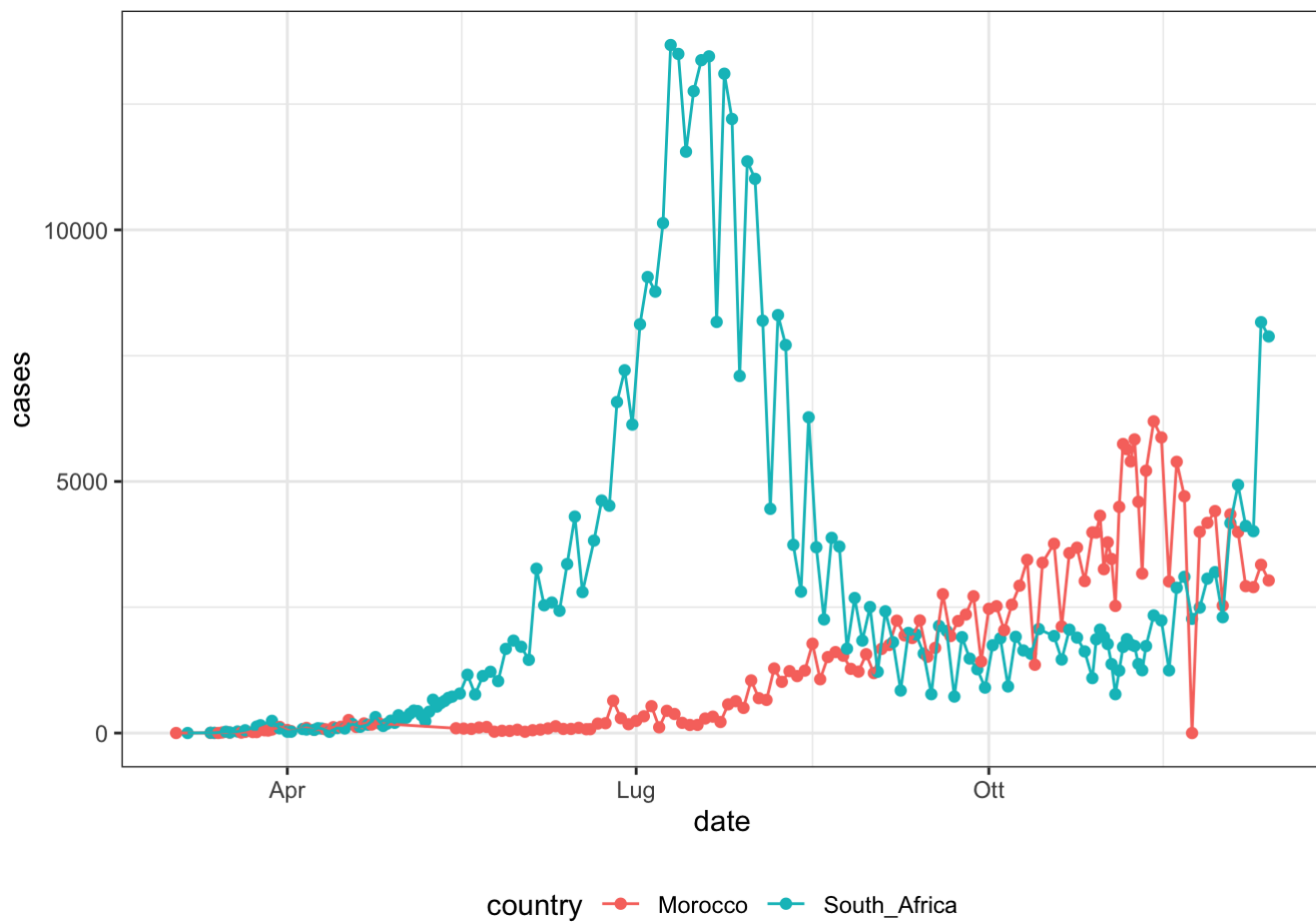


```
#plot of Africa  
ord_df %>% filter(continent == "Africa") %>%  
  ggplot(aes(x = date, y = cases, color = country)) + geom_point() + geom_line() +  
  theme_bw() +  
  theme(legend.position = "bottom")
```

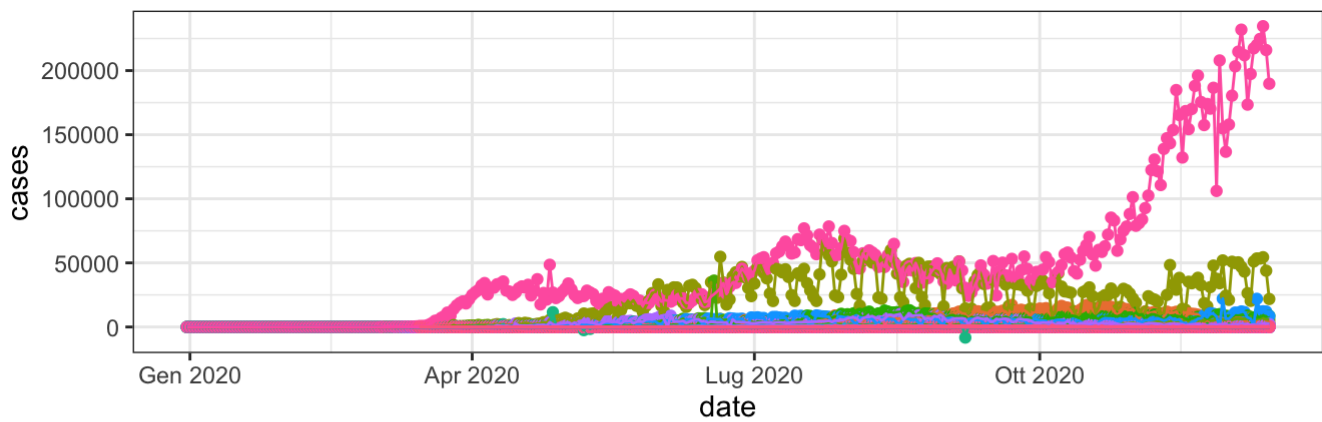


igeria	Congo	Ghana	Mauritius	
gola	Cote_d'Ivoire	Guinea	Morocco	
nin	Democratic_Republic_of_the_Congo	Guinea_Bissau	Mozambique	
tswana	Djibouti	Kenya	Namibia	
rkina_Faso	Egypt	Lesotho	Niger	
rundi	Equatorial_Guinea	Liberia	Nigeria	
meroon	Eritrea	Libya	Rwanda	
pe_Verde	Eswatini	Madagascar	Sao_Tome_and_Principe	
ntral_African_Republic	Ethiopia	Malawi	Senegal	
ad	Gabon	Mali	Seychelles	
moros	Gambia	Mauritania	Sierra_Leone	

```
#plot of South Africa
a<- ord_df %>% filter(country == c("South_Africa", "Morocco")) %>%
  ggplot(aes(x = date, y = cases, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
a
```



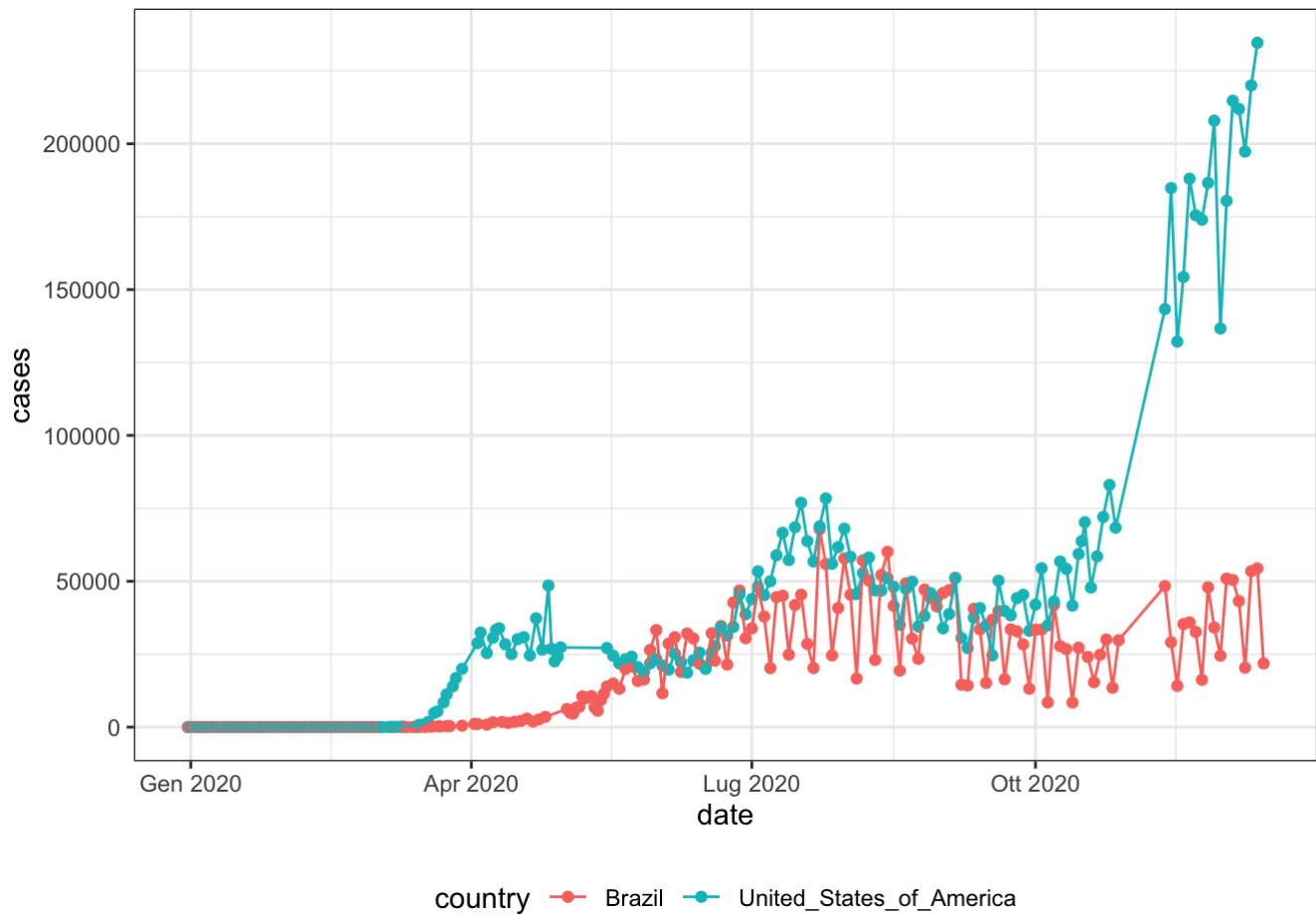
```
#plot of America  
ord_df %>% filter(continent == "America") %>%  
  ggplot(aes(x = date, y = cases, color = country)) + geom_point() + geom_line() +  
  theme_bw() +  
  theme(legend.position = "bottom")
```



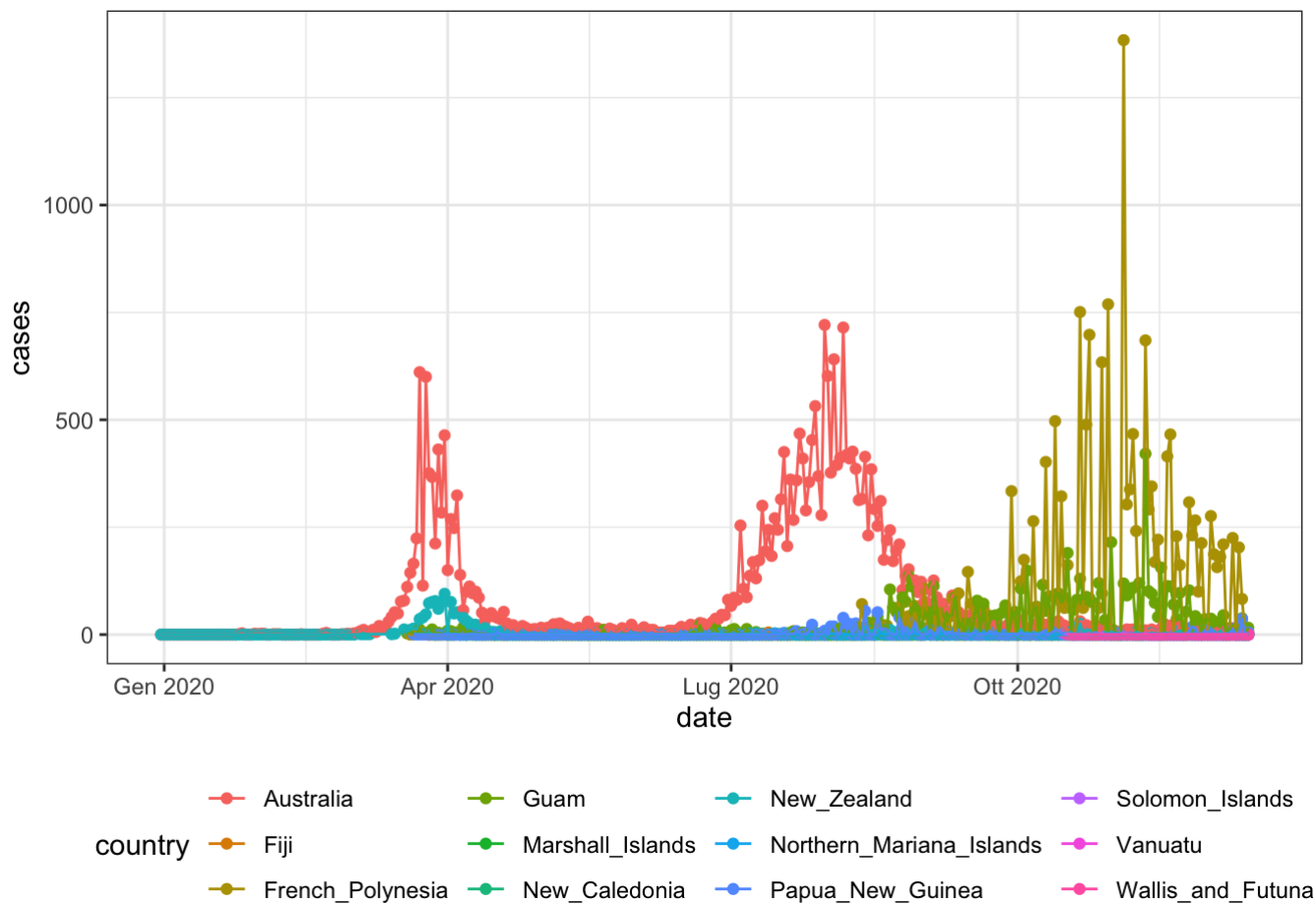
	Brazil	Dominican_Republic	Jamaica	
and_Barbuda	British_Virgin_Islands	Ecuador	Mexico	
na	Canada	El_Salvador	Montserrat	
	Cayman_Islands	Falkland_Islands_(Malvinas)	Nicaragua	
as	Chile	Greenland	Panama	
os	Colombia	Grenada	Paraguay	
	Costa_Rica	Guatemala	Peru	
la	Cuba	Guyana	Puerto_Rico	
	Curaçao	Haiti	Saint_Kitts_and_Nevis	
, Saint Eustatius and Saba	Dominica	Honduras	Saint_Lucia	

```
#plot United States of America and Brazil
us<- ord_df %>% filter(country == c("United_States_of_America", "Brazil")) %>%
  ggplot(aes(x = date, y = cases, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
us
```

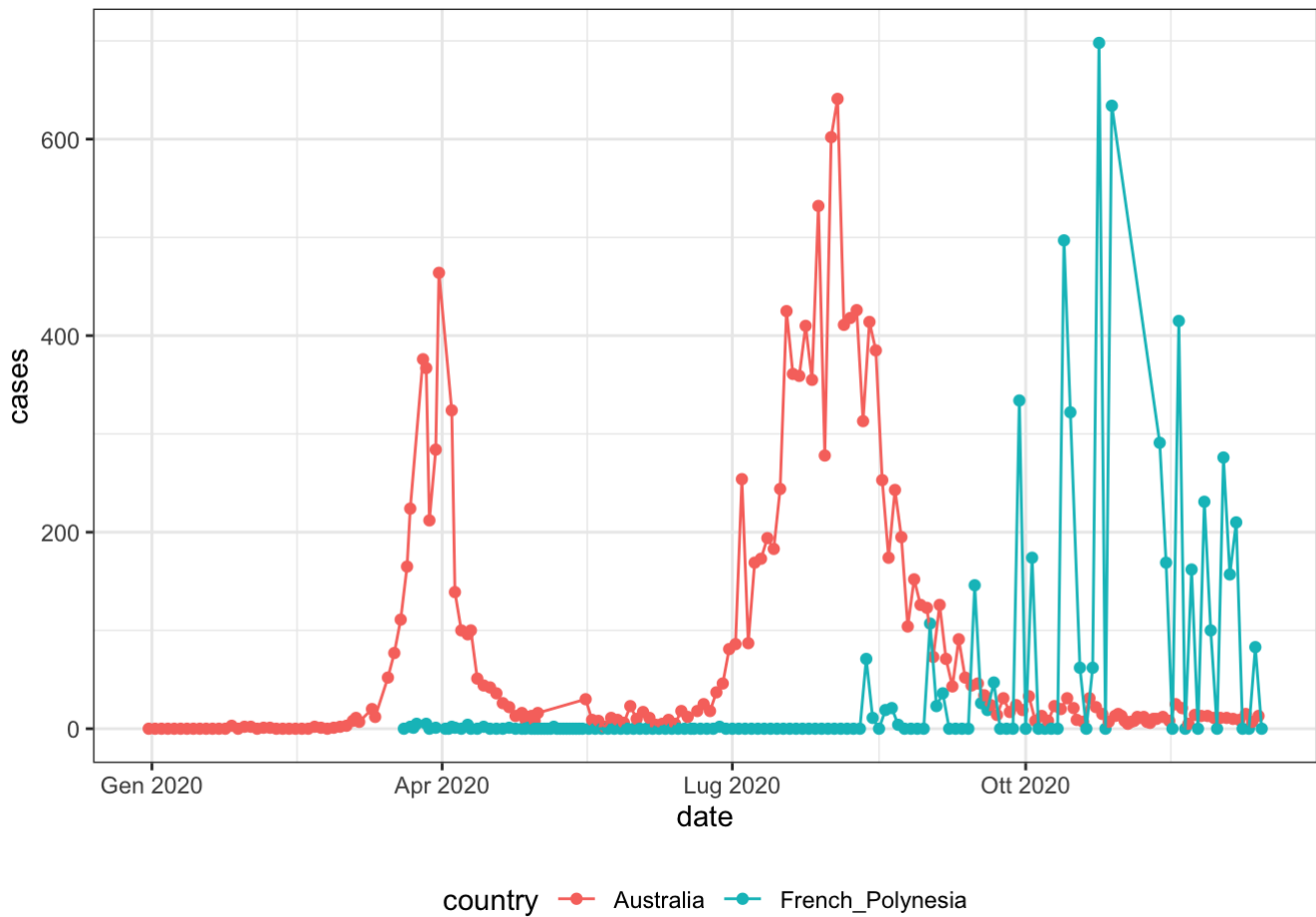




```
#plot of Oceania
ord_df %>% filter(continent == "Oceania") %>%
  ggplot(aes(x = date, y = cases, color = country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
```



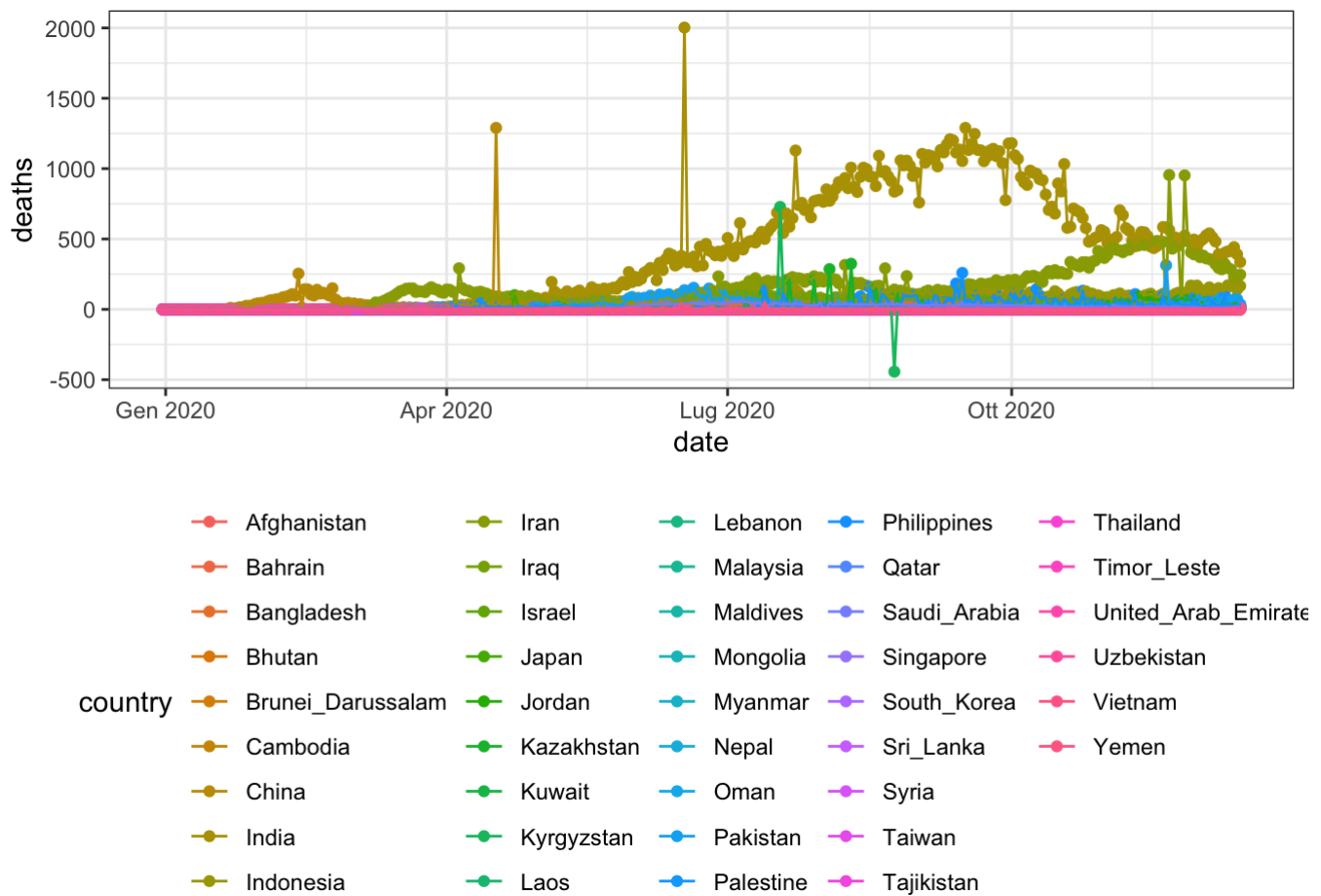
```
#plot of Australia and French Polynesia
o<- ord_df %>% filter(country == c("French_Polynesia", "Australia")) %>%
  ggplot(aes(x = date, y = cases, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
o
```



## ##Deaths

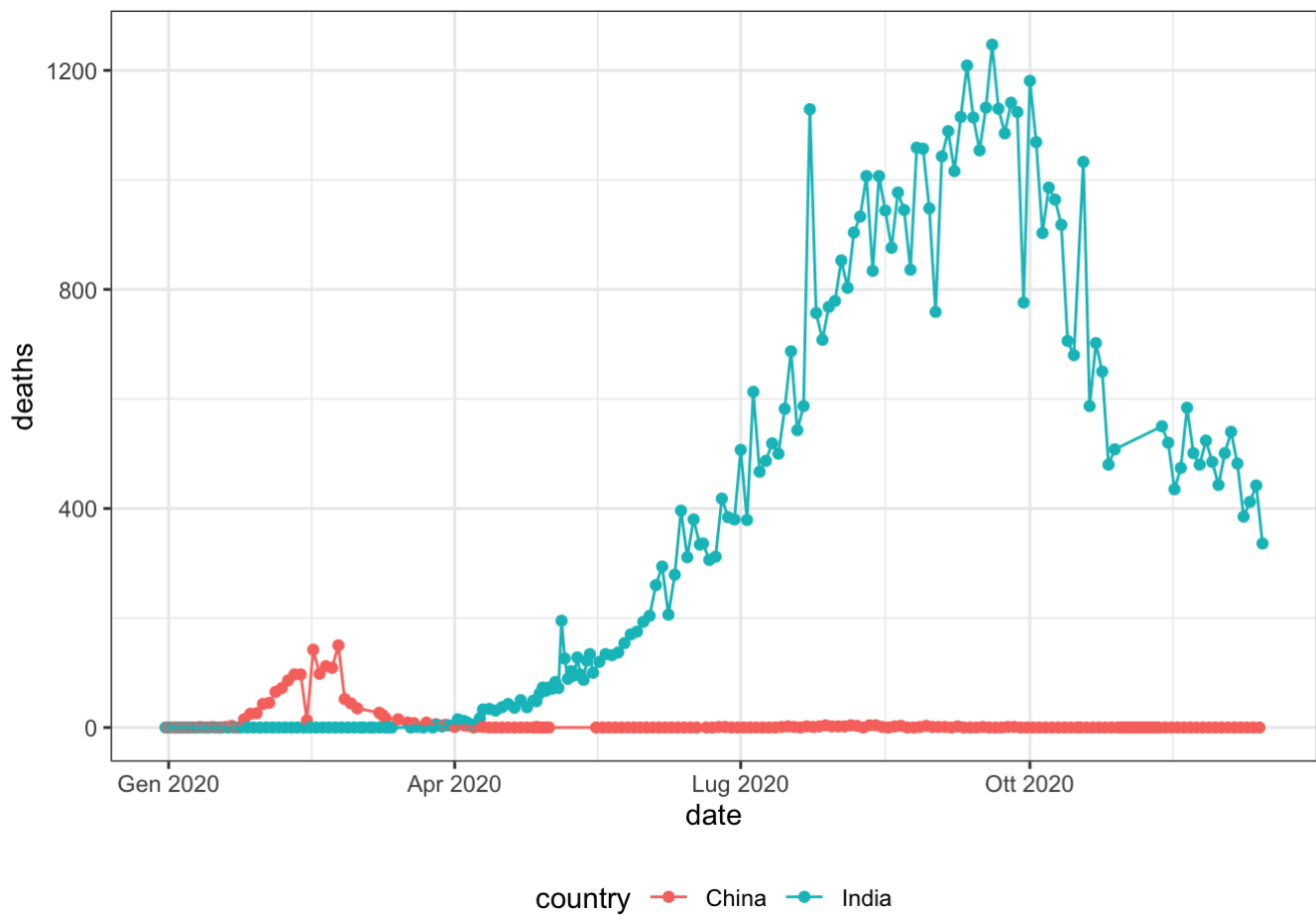
Moreover these plots show the number of deaths at continent- and country-level. Notice that one of the observation of Spain deaths is negative so there should be an error in the dataset.

```
ord_df %>% filter(continent == "Asia") %>%
  ggplot(aes(x = date, y = deaths, color = country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
```

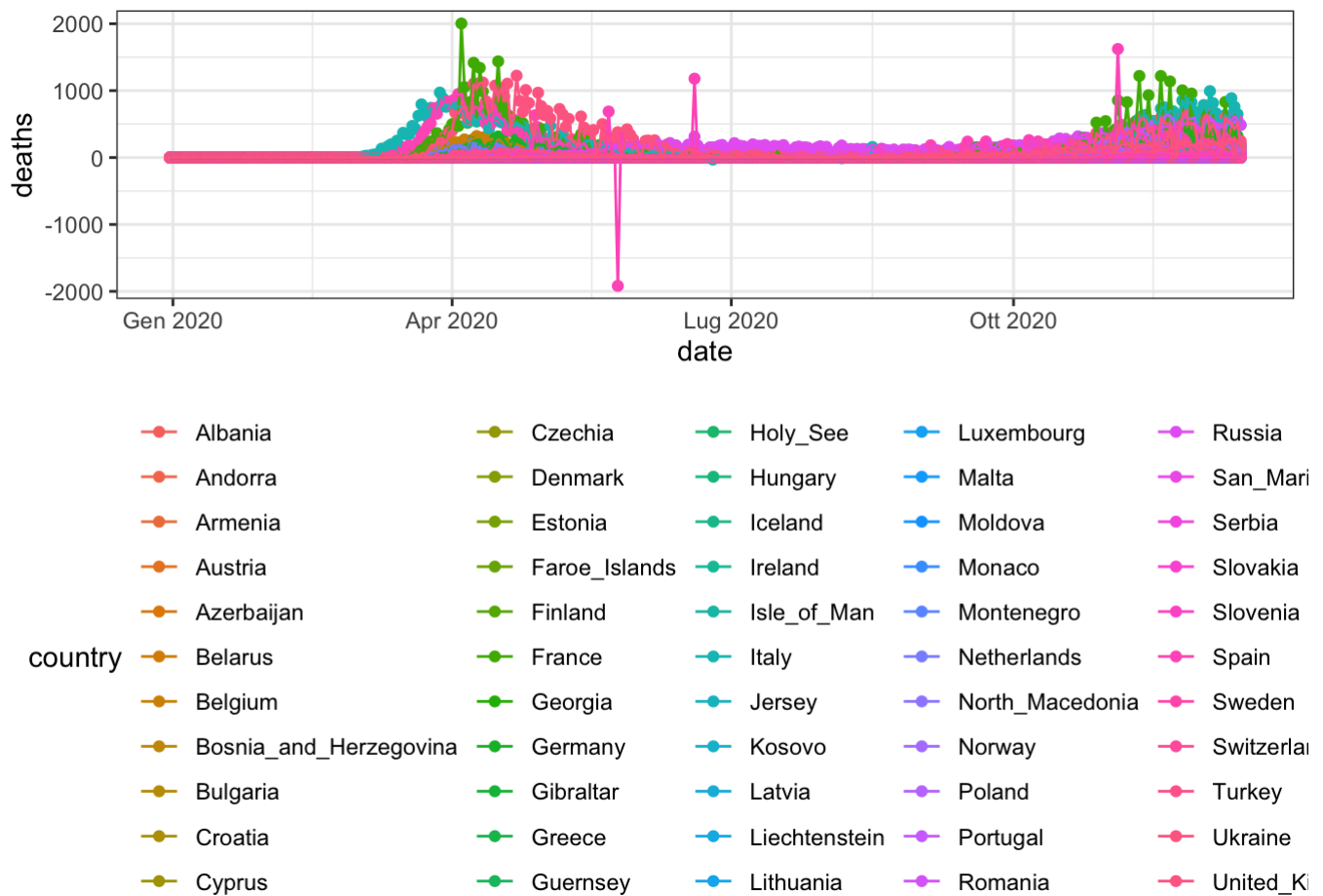


```
#we focus on China and India
```

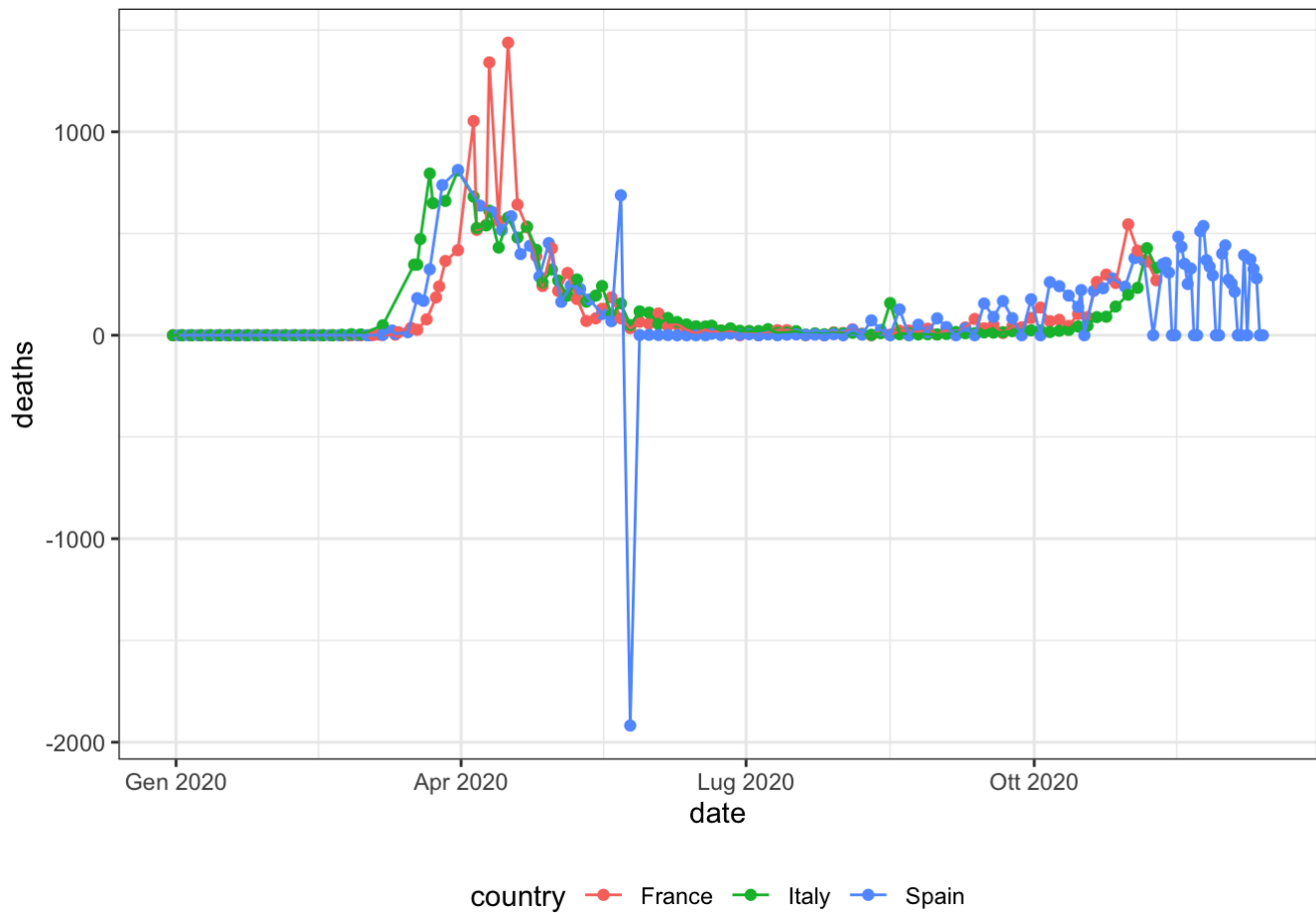
```
c1<- ord_df %>% filter(country == c("China", "India")) %>%
  ggplot(aes(x = date, y = deaths, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
c1
```



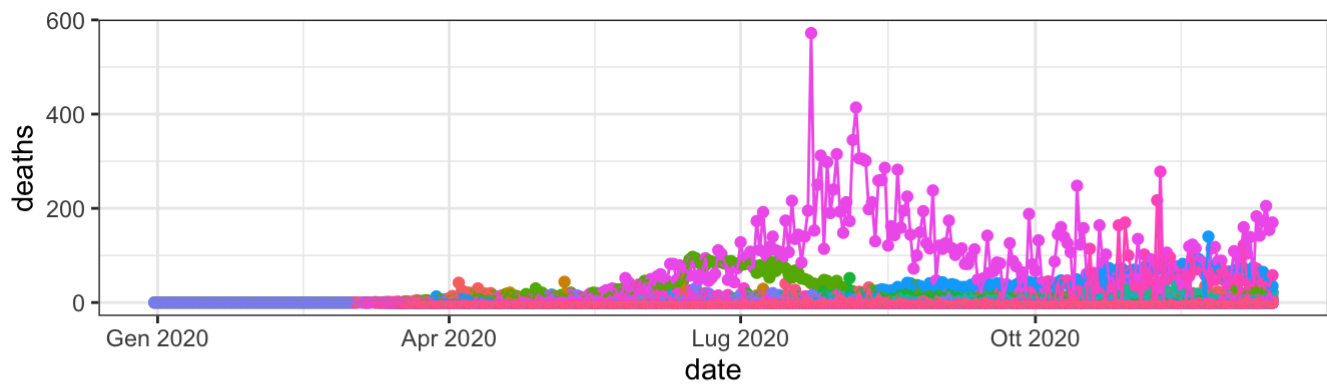
```
#plot of Europe
ord_df %>% filter(continent == "Europe") %>%
  ggplot(aes(x = date, y = deaths, color = country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
```



```
#plot of France, Italy and Spain
e1<- ord_df %>% filter(country == c("Italy", "France", "Spain"))%>%
  ggplot(aes(x = date, y = deaths, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
e1
```



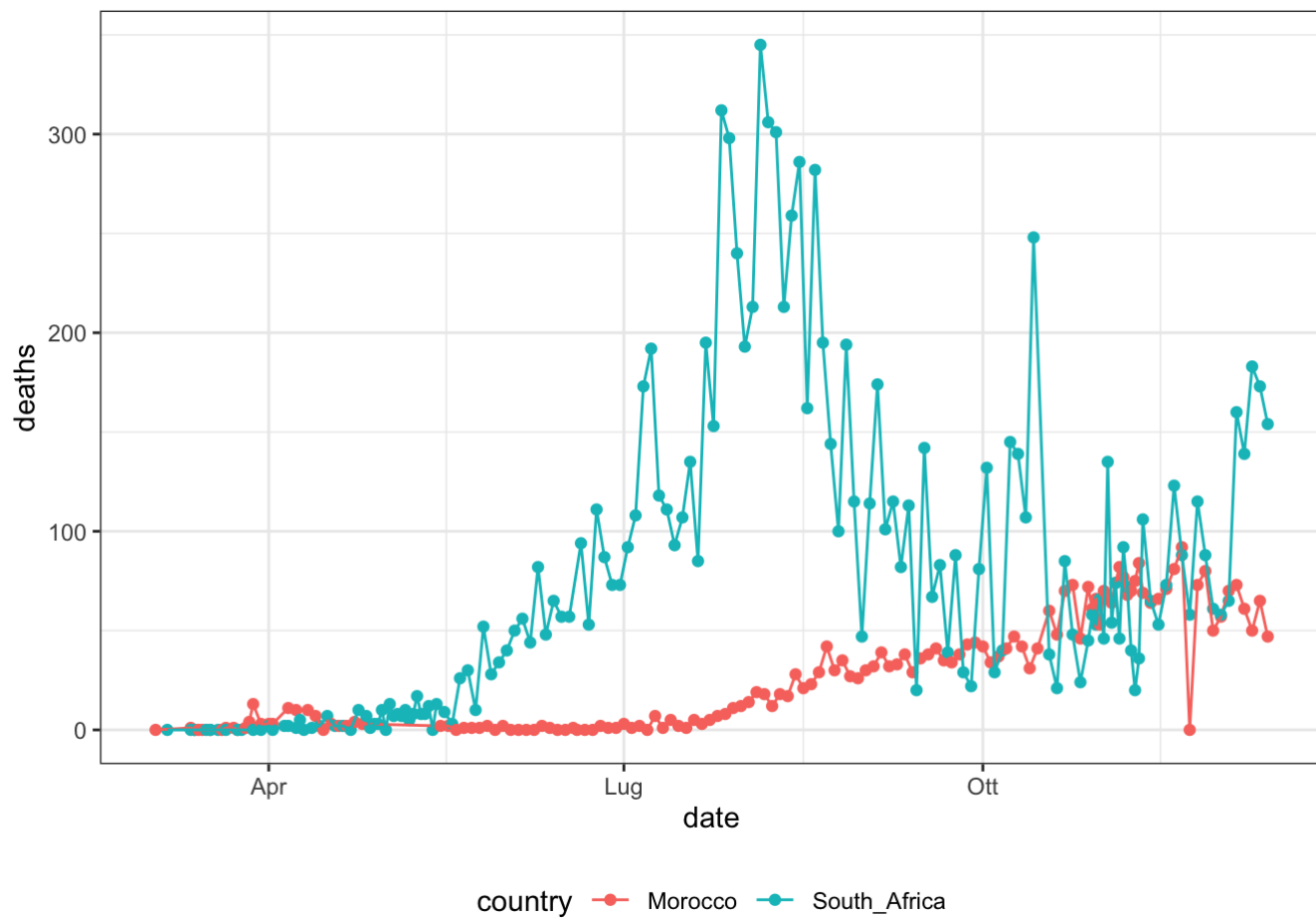
```
#plot of Africa
ord_df %>% filter(continent == "Africa") %>%
  ggplot(aes(x = date, y = deaths, color = country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
```



eria	Congo	Ghana	Mauritius	S
ola	Cote_d'Ivoire	Guinea	Morocco	S
in	Democratic_Republic_of_the_Congo	Guinea_Bissau	Mozambique	S
swana	Djibouti	Kenya	Namibia	S
ina_Faso	Egypt	Lesotho	Niger	T
ndi	Equatorial_Guinea	Liberia	Nigeria	T
eroon	Eritrea	Libya	Rwanda	U
e_Verde	Eswatini	Madagascar	Sao_Tome_and_Principe	U
tral_African_Republic	Ethiopia	Malawi	Senegal	V
d	Gabon	Mali	Seychelles	Z
oros	Gambia	Mauritania	Sierra_Leone	Z

```
#plot of South Africa
al<- ord_df %>% filter(country == c("South_Africa", "Morocco")) %>%
  ggplot(aes(x = date, y = deaths, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
al
```



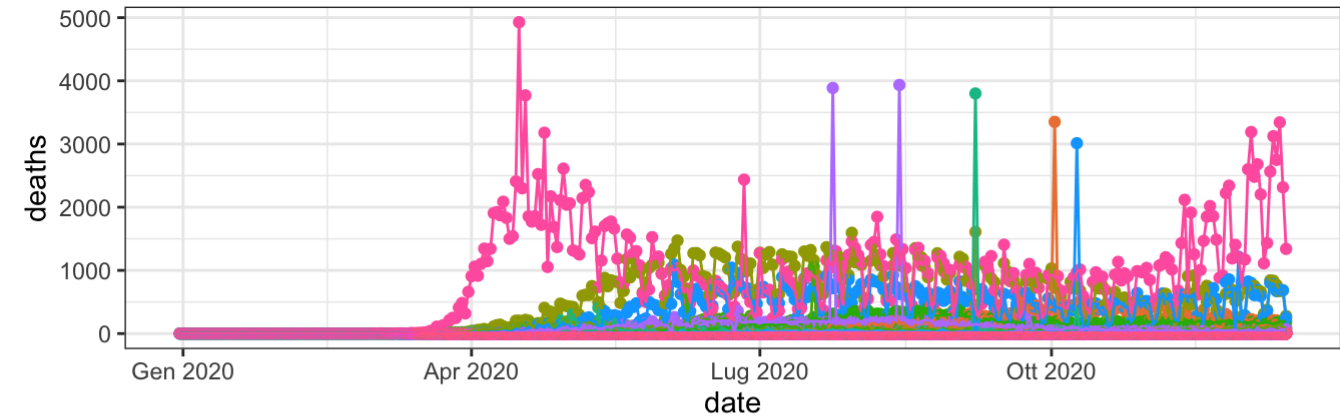


```
#plot of America
```

```
ord_df %>% filter(continent == "America") %>%
```

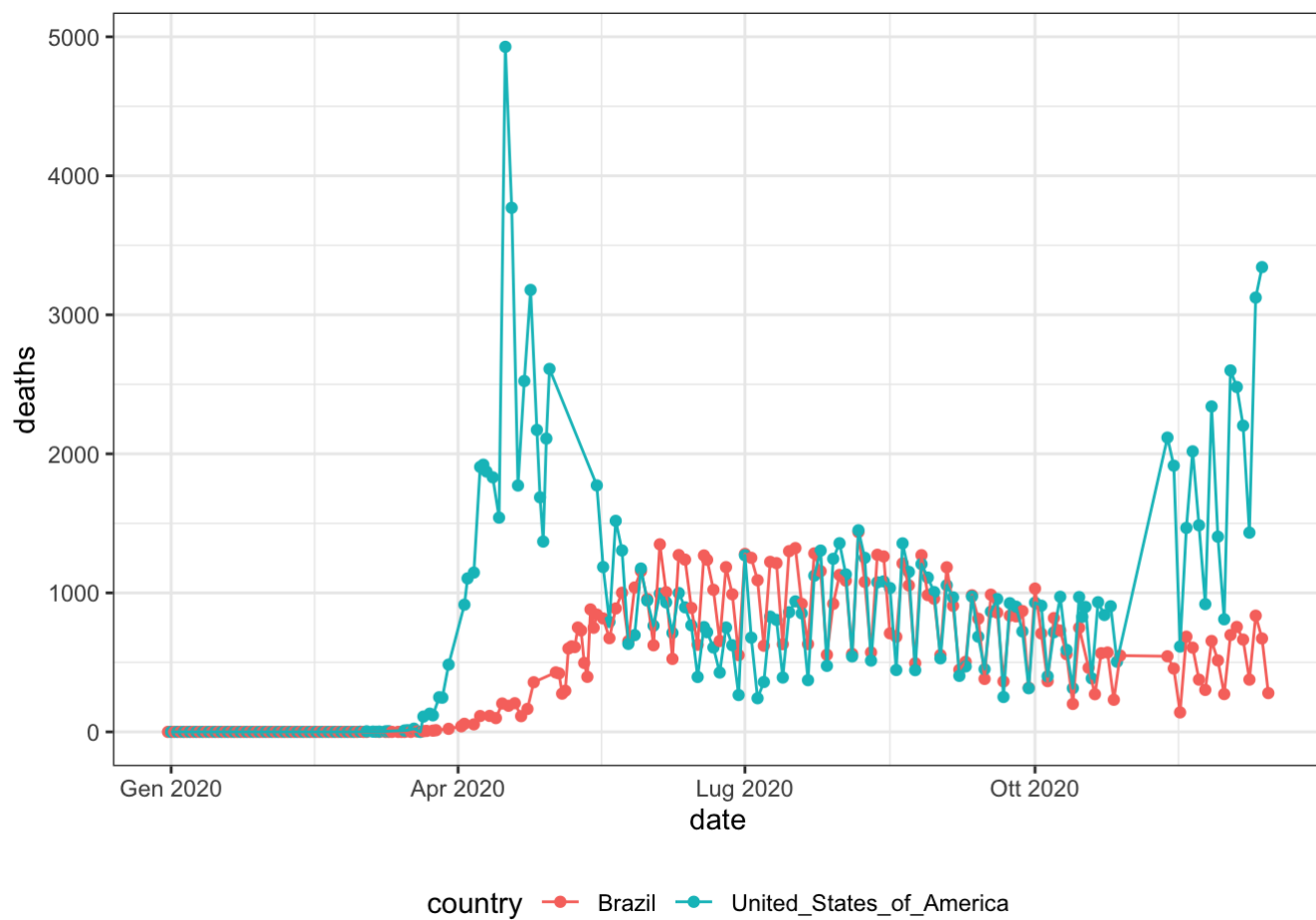
```
  ggplot(aes(x = date, y = deaths, color = country)) + geom_point() + geom_line() +  
  theme_bw() +
```

```
  theme(legend.position = "bottom")
```

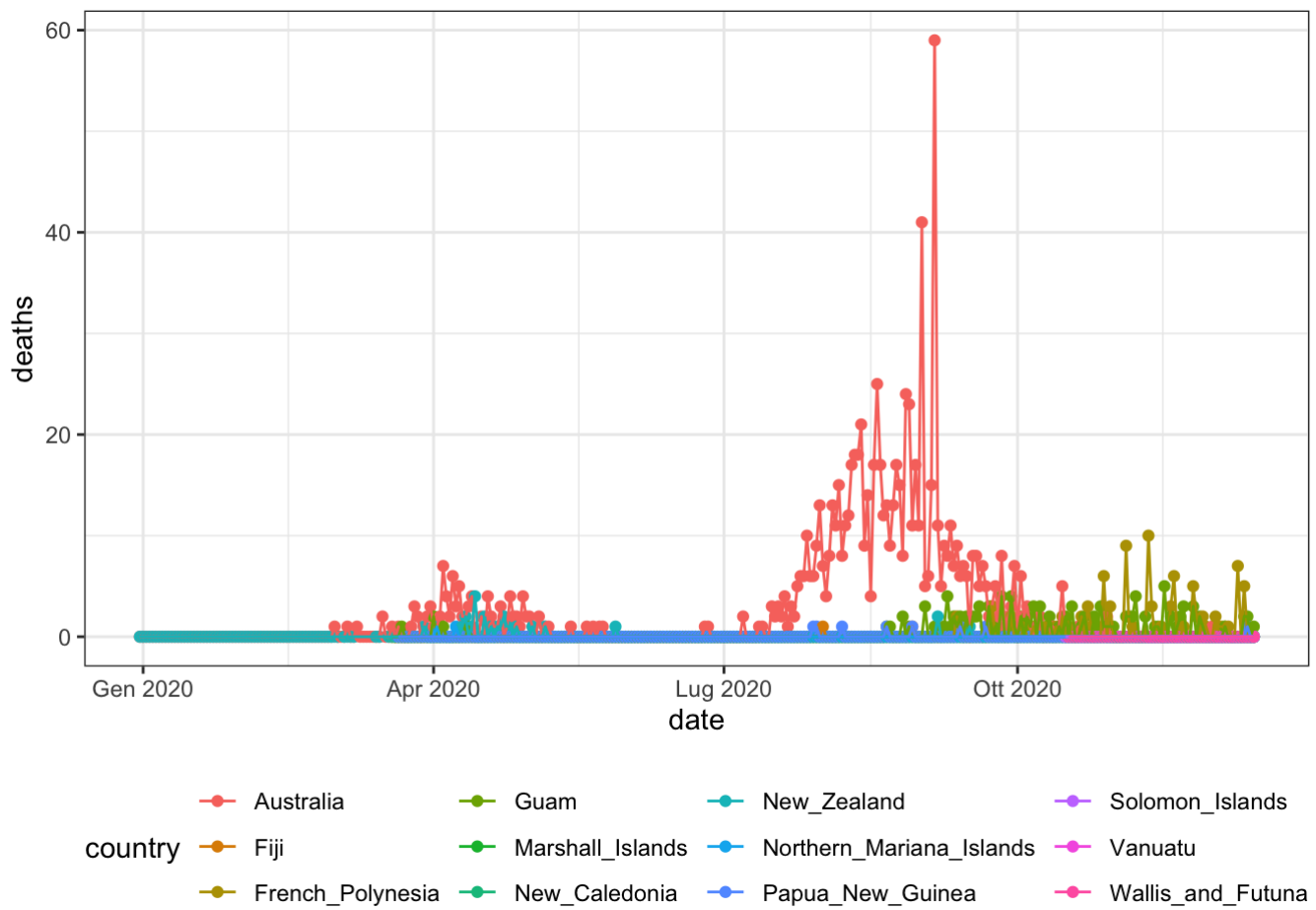


	<span style="color: #800000;">●</span> Brazil	<span style="color: #008000;">●</span> Dominican_Republic	<span style="color: #0000FF;">●</span> Jamaica	<span style="color: #FF00FF;">●</span> S
.and_Barbuda	<span style="color: #800000;">●</span> British_Virgin_Islands	<span style="color: #008000;">●</span> Ecuador	<span style="color: #0000FF;">●</span> Mexico	<span style="color: #FF00FF;">●</span> S
a	<span style="color: #800000;">●</span> Canada	<span style="color: #008000;">●</span> El_Salvador	<span style="color: #0000FF;">●</span> Montserrat	<span style="color: #FF00FF;">●</span> S
s	<span style="color: #800000;">●</span> Cayman_Islands	<span style="color: #008000;">●</span> Falkland_Islands_(Malvinas)	<span style="color: #0000FF;">●</span> Nicaragua	<span style="color: #FF00FF;">●</span> Tl
s	<span style="color: #800000;">●</span> Chile	<span style="color: #008000;">●</span> Greenland	<span style="color: #0000FF;">●</span> Panama	<span style="color: #FF00FF;">●</span> Tl
s	<span style="color: #800000;">●</span> Colombia	<span style="color: #008000;">●</span> Grenada	<span style="color: #0000FF;">●</span> Paraguay	<span style="color: #FF00FF;">●</span> U
	<span style="color: #800000;">●</span> Costa_Rica	<span style="color: #008000;">●</span> Guatemala	<span style="color: #0000FF;">●</span> Peru	<span style="color: #FF00FF;">●</span> U
i	<span style="color: #800000;">●</span> Cuba	<span style="color: #008000;">●</span> Guyana	<span style="color: #0000FF;">●</span> Puerto_Rico	<span style="color: #FF00FF;">●</span> U
	<span style="color: #800000;">●</span> Curaçao	<span style="color: #008000;">●</span> Haiti	<span style="color: #0000FF;">●</span> Saint_Kitts_and_Nevis	<span style="color: #FF00FF;">●</span> V
Saint Eustatius and Saba	<span style="color: #800000;">●</span> Dominica	<span style="color: #008000;">●</span> Honduras	<span style="color: #0000FF;">●</span> Saint_Lucia	

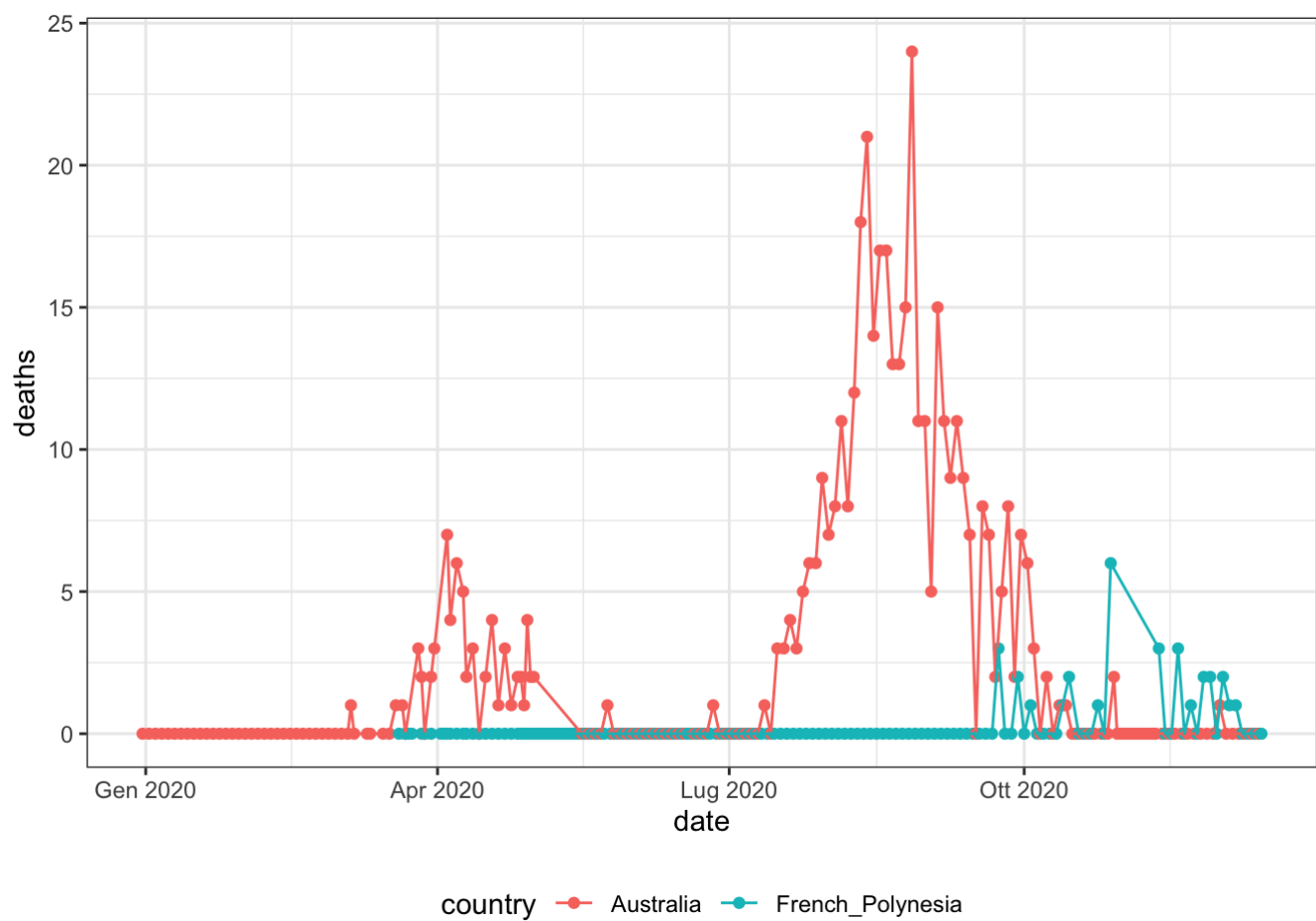
```
#plot United States of America and Brazil
us1<- ord_df %>% filter(country == c("United_States_of_America", "Brazil")) %>%
  ggplot(aes(x = date, y = deaths, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
us1
```



```
#plot of Oceania
ord_df %>% filter(continent == "Oceania") %>%
  ggplot(aes(x = date, y = deaths, color = country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
```



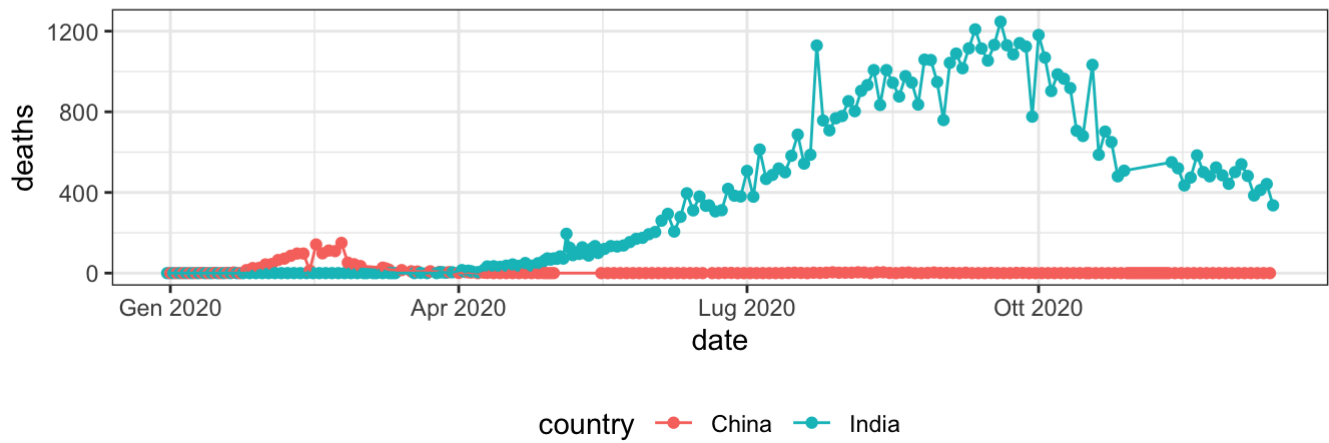
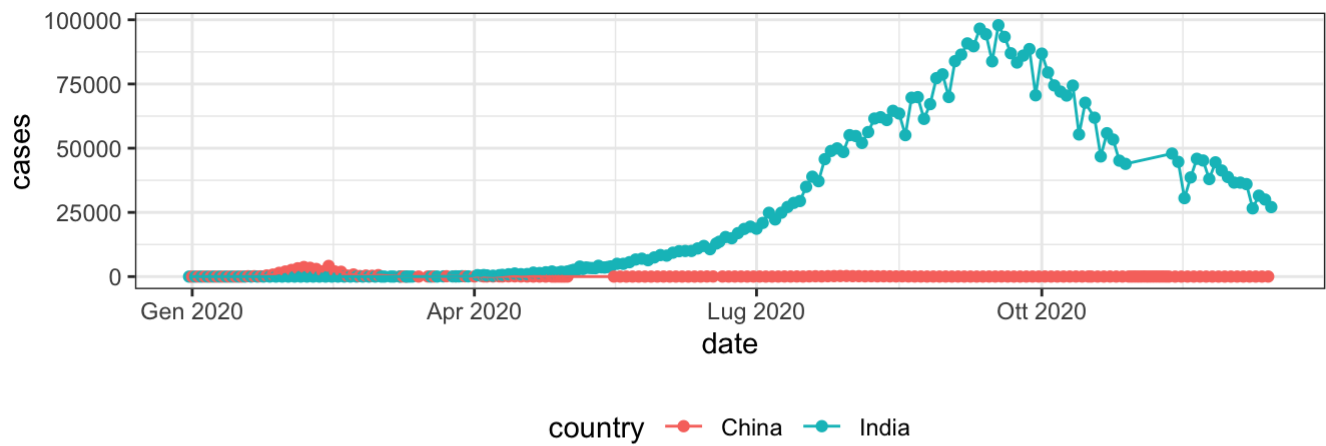
```
#plot of Australia and French Polynesia
o1<- ord_df %>% filter(country == c("French_Polynesia", "Australia")) %>%
  ggplot(aes(x = date, y = deaths, color=country)) + geom_point() + geom_line() +
  theme_bw() +
  theme(legend.position = "bottom")
o1
```



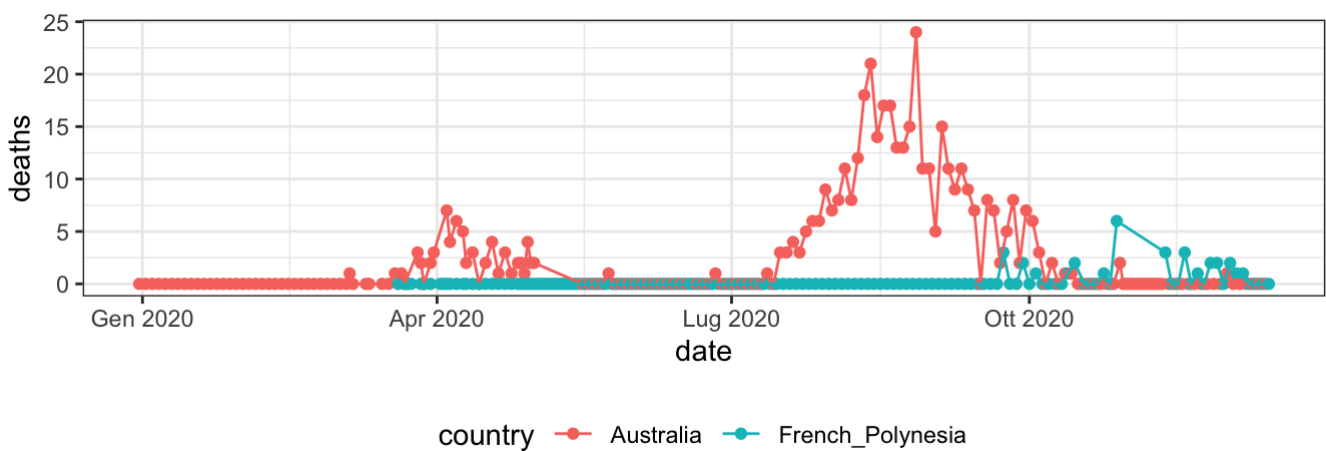
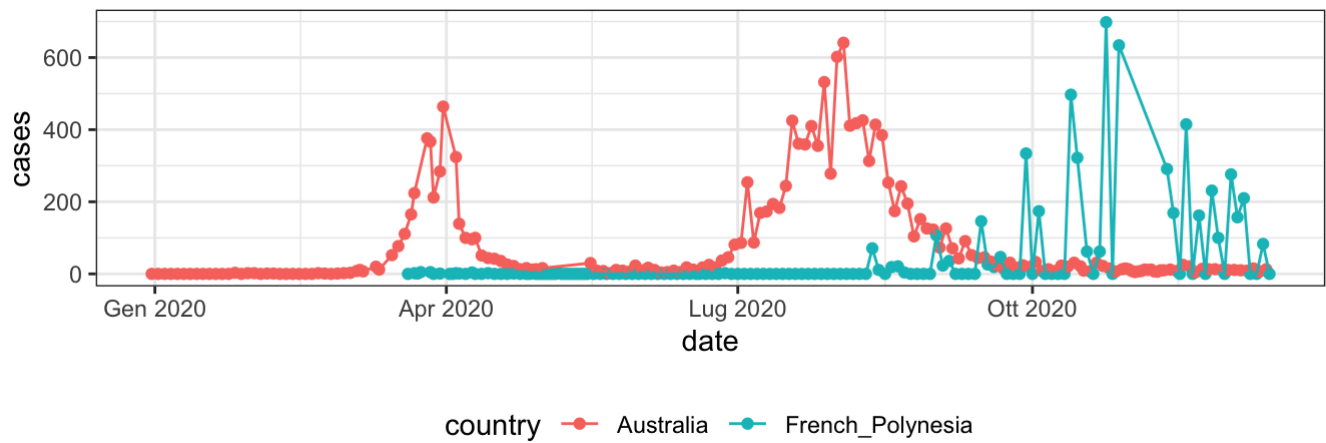
### ###COMPARISON DEATHS AND CASES FOR COUNTRIES

The following plots compare deaths and cases in the most affected countries.

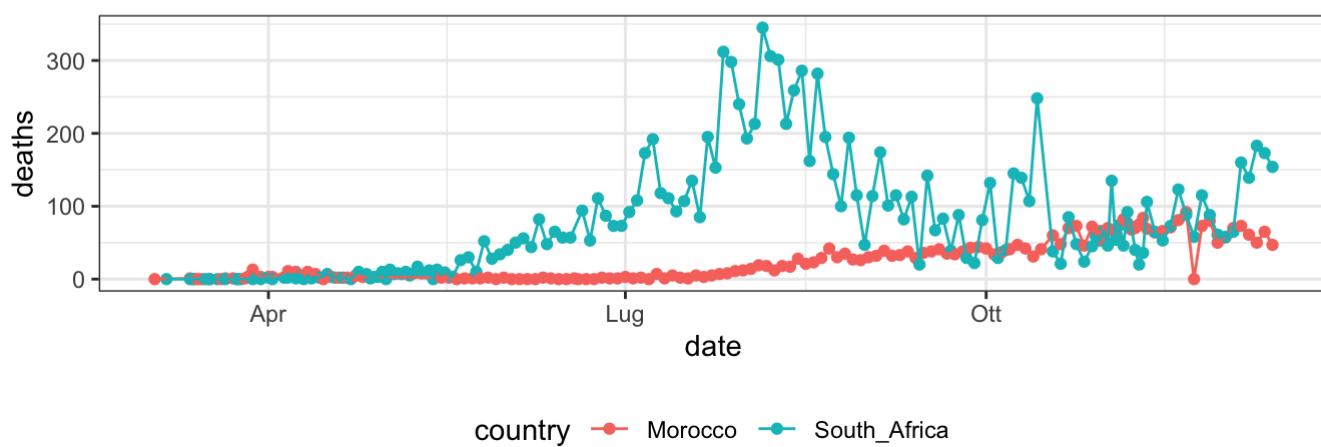
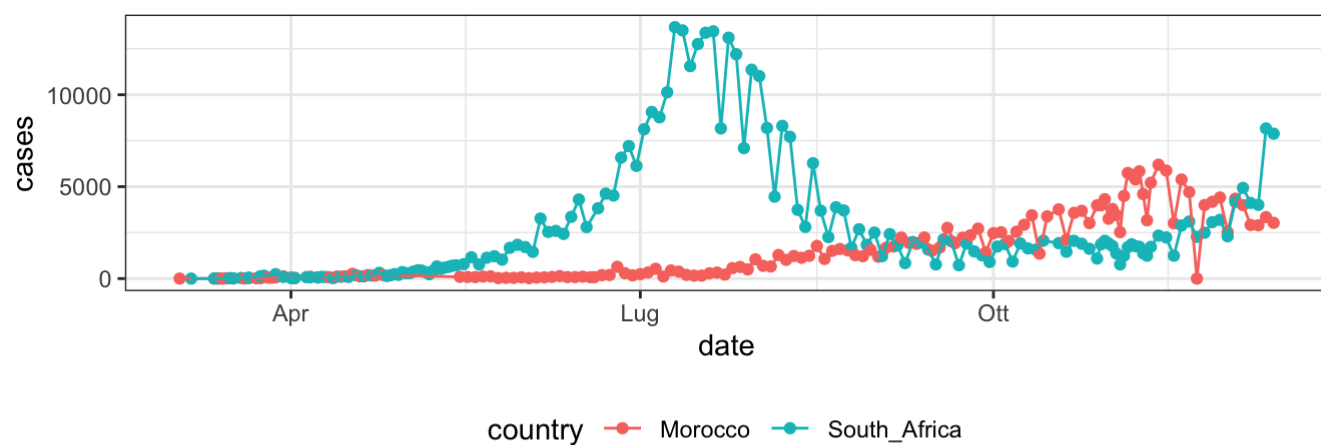
```
cowplot:: plot_grid(c,c1, nrow=2, ncol=1)
```



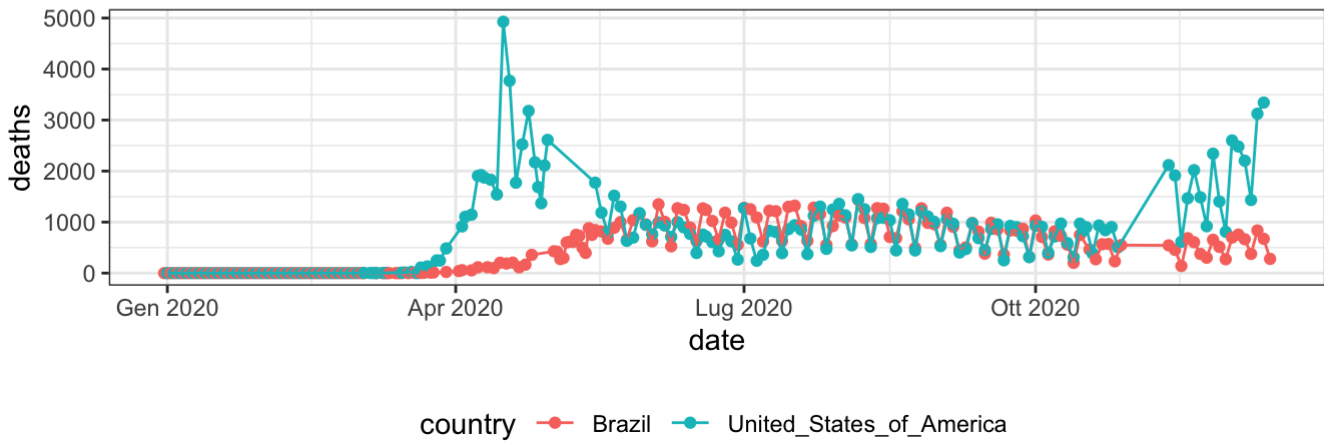
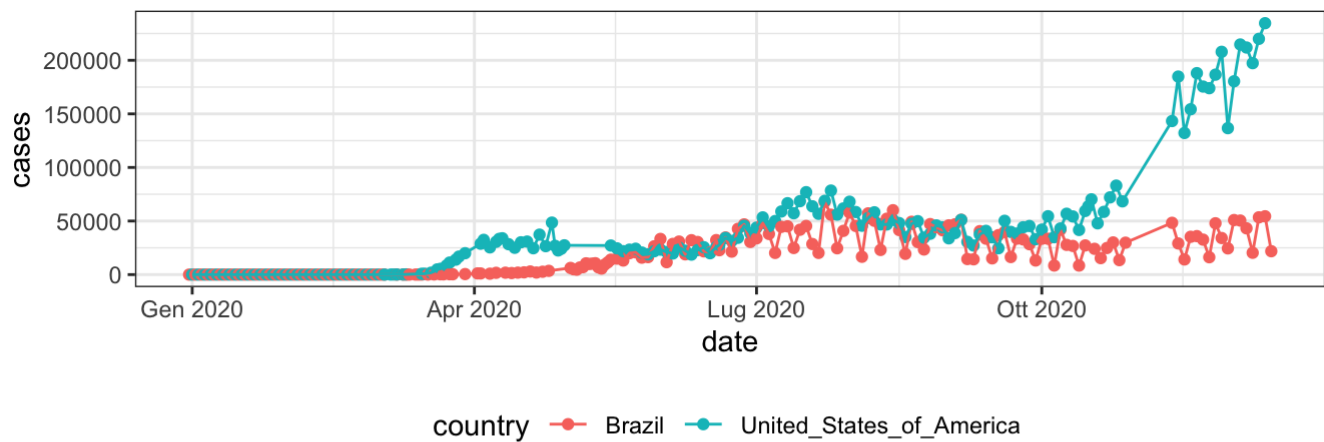
```
cowplot:: plot_grid(o,o1, nrow=2, ncol=1)
```



```
cowplot:: plot_grid(a,a1, nrow=2, ncol=1)
```



```
cowplot:: plot_grid(us,us1, nrow=2, ncol=1)
```



```
cowplot:: plot_grid(e,e1, nrow=2, ncol=1)
```

