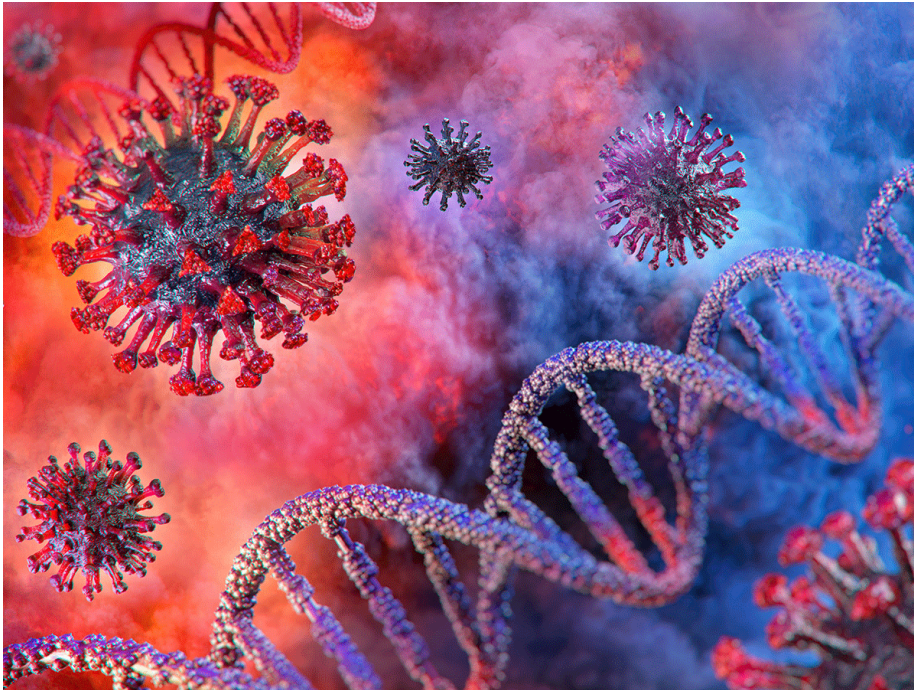


Data Science - COVID-19

Carlos Utrilla Guerrero

05/04/2020



Analysis coronavirus disease (COVID-19).

This is an R Markdown document. It is intended to publicly illustrate how R statistics can help you out to output data science pipeline.

About this data

It changes rapidly

It doesn't include all cases

Confirmed cases aren't all cases. They only include people who tested positive. Testing rules and availability vary by country.

Data Repository: Johns Hopkins University.

```
# This is an analysis report of the Novel Coronavirus (COVID-19)
# Aim for data processing, visualisation and statistics
# Source code: http://yanchang.rdatamining.com/
# set directory
# Data Source: 2019 Data Repository https://github.com/CSSEGISandData/COVID-19
# R Packages:
library(magrittr) # pipeline operations
library(lubridate) # date operation
```

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

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

```
library(tidyverse) # data science pips
```

```
## -- Attaching packages -----
## v ggplot2 3.3.0      v purrr   0.3.3
## v tibble  2.1.3      v dplyr   0.8.5
## v tidyr   1.0.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0
```

```
## -- Conflicts -----
## x lubridate::as.difftime() masks base::as.difftime()
## x lubridate::date()       masks base::date()
## x tidyr::extract()        masks magrittr::extract()
## x dplyr::filter()         masks stats::filter()
## x lubridate::intersect()  masks base::intersect()
## x dplyr::lag()            masks stats::lag()
## x purrr::set_names()      masks magrittr::set_names()
## x lubridate::setdiff()    masks base::setdiff()
## x lubridate::union()      masks base::union()
```

```
library(gridExtra) # grid based plots
```

```
##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##   combine
```

```
library(dplyr)
library(leaflet)
library(ggforce)
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':
##
##   group_rows
```

```
# Loading data
# At first, three CSV files, are downloaded and saved as local files
# and then loaded into R
# source data files changes everytime
filenames <- c('time_series_covid19_confirmed_global.csv',
               'time_series_covid19_deaths_global.csv',
               'time_series_covid19_recovered_global.csv')
url.path <- paste0('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/',
                  'master/csse_covid_19_data/csse_covid_19_time_series/')

#download files to local folder
download <- function(filename) {
  url <- file.path(url.path, filename)
  dest <- file.path('./data', filename)
  download.file(url, dest)
}
bin <- lapply(filenames, download)

# load data into R
data.confirmed.original <- read.csv('./data/time_series_covid19_confirmed_global.csv')
data.deaths.original <- read.csv('./data/time_series_covid19_deaths_global.csv')
data.recovered.original <- read.csv('./data/time_series_covid19_recovered_global.csv')
```

```
# check dimension of data confirmed  
dim(data.confirmed.original)
```

```
## [1] 258 76
```

Below we check the time frame of data set

```
# check time frame of the data  
n.col <- ncol(data.confirmed.original) # 58 variables  
# get dates from column names  
dates <- names(data.confirmed.original)[5:n.col] %>% substr(2,8) %>% mdy()  
range(dates)
```

```
## [1] "2020-01-22" "2020-04-02"
```

```
min.date <- min(dates)  
max.date <- max(dates)  
max.date.txt <- max.date %>% format('%d %b %Y')  
min.date.txt <- min.date %>% format('%d %b Y')  
# last update on 26 March 2020 max.date
```

```
# Data Preparation steps:  
# 1.From wide to long format  
# 2.Aggregate by country  
# 3. merge into a single dataset  
# cleaning and transformation  
cleanData <- function(data) {  
  ## remove some columns  
  data %<>% select(-c(Province.State, Lat, Long)) %>% rename(country=Country.Region)  
  ## convert from wide to long format  
  data %<>% gather(key=date, value=count, -country)  
  ## convert from character to date  
  data %<>% mutate(date = date %>% substr(2,8) %>% mdy())  
  ## aggregate by country  
  data %<>% group_by(country, date) %>% summarise(count=sum(count)) %>% as.data.frame()  
  return(data)  
}  
# clean the three datasets  
data.confirmed <- data.confirmed.original %>% cleanData() %>% rename(confirmed=count)  
data.deaths <- data.deaths.original %>% cleanData() %>% rename(deaths=count)  
data.recovered <- data.recovered.original %>% cleanData() %>% rename(recovered=count)  
  
# merge above 3 datasets into one, by country and date
```

```

data <- data.confirmed %>% merge(data.deaths, all = T) %>% merge(data.recovered, all = T)

# countries/regions with confirmed cases (excl cruise ships)
countries <- data %>% pull(country) %>% setdiff('Cruise Ship')

# last 10 records when it first broke out in Spain
data %>% filter(country == 'Spain') %>% tail(10)

```

```

##      country      date confirmed deaths recovered
## 63  Spain 2020-03-24   39885   2808     3794
## 64  Spain 2020-03-25   49515   3647     5367
## 65  Spain 2020-03-26   57786   4365     7015
## 66  Spain 2020-03-27   65719   5138     9357
## 67  Spain 2020-03-28   73235   5982    12285
## 68  Spain 2020-03-29   80110   6803    14709
## 69  Spain 2020-03-30   87956   7716    16780
## 70  Spain 2020-03-31   95923   8464    19259
## 71  Spain 2020-04-01  104118  9387    22647
## 72  Spain 2020-04-02  112065 10348    26743

```

```

# counts for worldwide
data.world <- data %>% group_by(date) %>%
  summarise(country='World',
            confirmed=sum(confirmed, na.rm = T),
            deaths=sum(deaths, na.rm = T),
            recovered=sum(recovered, na.rm = T))

data %<>% rbind(data.world)

# current confirmed cases
data %<>% mutate(remaining.confirmed = confirmed - deaths - recovered)

```

```

# Visualisation
# After preparing the data, we portrait it in various graphs

# TOP Ten Countries
# ranking by confirmed cases
data.latest.all <- data %>% filter(date == max(date)) %>%
  select(country, date,
         confirmed, confirmed.new, remaining.confirmed, recovered, deaths.new, deaths, death)

# top 20 countries incl 11 World
k<- 20
top.countries <- data.latest.all %>% filter(ranking <= k+1) %>%
  arrange(ranking) %>% pull(country) %>% as.character()
top.countries %>% setdiff('World') %>% print()

```

```
## [1] "US" "Italy" "Spain" "Germany"
## [5] "China" "France" "Iran" "United Kingdom"
## [9] "Switzerland" "Turkey" "Belgium" "Netherlands"
## [13] "Canada" "Austria" "Korea, South" "Portugal"
## [17] "Brazil" "Israel" "Sweden" "Norway"
```

```
names(data.latest.all)
```

```
## [1] "country" "date" "confirmed"
## [4] "confirmed.new" "remaining.confirmed" "recovered"
## [7] "deaths.new" "deaths" "death.rate"
## [10] "ranking"
```

```
## add 'Others'
top.countries %<>% c('Others')
## put all others in a single group of 'Others'
data.latest <- data.latest.all %>% filter(!is.na(country)) %>%
mutate(country=ifelse(ranking <= k + 1, as.character(country), 'Others')) %>%
mutate(country=country %>% factor(levels=c(top.countries)))

data.latest %<>% group_by(country) %>%
  summarise(confirmed=sum(confirmed), confirmed.new=sum(confirmed.new), remaining.confirmed=
  mutate(death.rate=(100*deaths/confirmed) %>% round(1))
data.latest %<>% select(c(country, confirmed, deaths,death.rate, confirmed.new, deaths.new,1
data.latest %>% mutate(death.rate=death.rate %>% format(nsmall=1) %>% paste0('%')) %>% kable
```

Worldmap

```
x <- data.confirmed.original
x$confirmed <- x[, ncol(x)]
x %>% select(c(Country.Region, Province.State, Lat, Long, confirmed)) %>%
  mutate(txt=paste0(Country.Region, '-', Province.State, ':', confirmed))
```

```
## Country.Region Province.State
## 1 Afghanistan
## 2 Albania
## 3 Algeria
## 4 Andorra
## 5 Angola
## 6 Antigua and Barbuda
## 7 Argentina
## 8 Armenia
## 9 Australia Australian Capital Territory
```

Table 1: Cases in Top 20 Countries - 02 Apr 2020.

	country	confirmed	deaths	death.rate	confirmed.new	deaths.new	remaining.confirmed
1	World	1,013,157	52,983	5.2%	80,552	6,174	749,911
2	US	243,453	5,926	2.4%	30,081	1,169	228,526
3	Italy	115,242	13,915	12.1%	4,668	760	83,049
4	Spain	112,065	10,348	9.2%	7,947	961	74,974
5	Germany	84,794	1,107	1.3%	6,922	187	61,247
6	China	82,432	3,322	4.0%	71	6	2,545
7	France	59,929	5,398	9.0%	2,180	1,355	41,983
8	Iran	50,468	3,160	6.3%	2,875	124	30,597
9	United Kingdom	34,173	2,926	8.6%	4,308	569	31,055
10	Switzerland	18,827	536	2.8%	1,059	48	14,278
11	Turkey	18,135	356	2.0%	2,456	79	17,364
12	Belgium	15,348	1,011	6.6%	1,384	183	11,842
13	Netherlands	14,788	1,341	9.1%	1,092	166	13,187
14	Canada	11,284	139	1.2%	1,724	30	9,410
15	Austria	11,129	158	1.4%	418	12	9,222
16	Korea, South	9,976	169	1.7%	89	4	3,979
17	Portugal	9,034	209	2.3%	783	22	8,757
18	Brazil	8,044	324	4.0%	1,208	84	7,593
19	Israel	6,857	36	0.5%	765	10	6,483
20	Sweden	5,568	308	5.5%	621	69	5,157
21	Norway	5,147	50	1.0%	284	6	5,065
22	Others	96,464	2,244	2.3%	9,617	330	83,598

```

## 10      Australia      New South Wales
## 11      Australia      Northern Territory
## 12      Australia      Queensland
## 13      Australia      South Australia
## 14      Australia      Tasmania
## 15      Australia      Victoria
## 16      Australia      Western Australia
## 17      Austria
## 18      Azerbaijan
## 19      Bahamas
## 20      Bahrain
## 21      Bangladesh
## 22      Barbados
## 23      Belarus
## 24      Belgium
## 25      Benin
## 26      Bhutan
## 27      Bolivia
## 28      Bosnia and Herzegovina
## 29      Brazil
## 30      Brunei
## 31      Bulgaria

```

## 32	Burkina Faso	
## 33	Cabo Verde	
## 34	Cambodia	
## 35	Cameroon	
## 36	Canada	Alberta
## 37	Canada	British Columbia
## 38	Canada	Grand Princess
## 39	Canada	Manitoba
## 40	Canada	New Brunswick
## 41	Canada	Newfoundland and Labrador
## 42	Canada	Nova Scotia
## 43	Canada	Ontario
## 44	Canada	Prince Edward Island
## 45	Canada	Quebec
## 46	Canada	Saskatchewan
## 47	Central African Republic	
## 48	Chad	
## 49	Chile	
## 50	China	Anhui
## 51	China	Beijing
## 52	China	Chongqing
## 53	China	Fujian
## 54	China	Gansu
## 55	China	Guangdong
## 56	China	Guangxi
## 57	China	Guizhou
## 58	China	Hainan
## 59	China	Hebei
## 60	China	Heilongjiang
## 61	China	Henan
## 62	China	Hong Kong
## 63	China	Hubei
## 64	China	Hunan
## 65	China	Inner Mongolia
## 66	China	Jiangsu
## 67	China	Jiangxi
## 68	China	Jilin
## 69	China	Liaoning
## 70	China	Macau
## 71	China	Ningxia
## 72	China	Qinghai
## 73	China	Shaanxi
## 74	China	Shandong
## 75	China	Shanghai
## 76	China	Shanxi
## 77	China	Sichuan

## 78	China	Tianjin
## 79	China	Tibet
## 80	China	Xinjiang
## 81	China	Yunnan
## 82	China	Zhejiang
## 83	Colombia	
## 84	Congo (Brazzaville)	
## 85	Congo (Kinshasa)	
## 86	Costa Rica	
## 87	Cote d'Ivoire	
## 88	Croatia	
## 89	Diamond Princess	
## 90	Cuba	
## 91	Cyprus	
## 92	Czechia	
## 93	Denmark	Faroe Islands
## 94	Denmark	Greenland
## 95	Denmark	
## 96	Djibouti	
## 97	Dominican Republic	
## 98	Ecuador	
## 99	Egypt	
## 100	El Salvador	
## 101	Equatorial Guinea	
## 102	Eritrea	
## 103	Estonia	
## 104	Eswatini	
## 105	Ethiopia	
## 106	Fiji	
## 107	Finland	
## 108	France	French Guiana
## 109	France	French Polynesia
## 110	France	Guadeloupe
## 111	France	Mayotte
## 112	France	New Caledonia
## 113	France	Reunion
## 114	France	Saint Barthelemy
## 115	France	St Martin
## 116	France	Martinique
## 117	France	
## 118	Gabon	
## 119	Gambia	
## 120	Georgia	
## 121	Germany	
## 122	Ghana	
## 123	Greece	

## 124	Guatemala	
## 125	Guinea	
## 126	Guyana	
## 127	Haiti	
## 128	Holy See	
## 129	Honduras	
## 130	Hungary	
## 131	Iceland	
## 132	India	
## 133	Indonesia	
## 134	Iran	
## 135	Iraq	
## 136	Ireland	
## 137	Israel	
## 138	Italy	
## 139	Jamaica	
## 140	Japan	
## 141	Jordan	
## 142	Kazakhstan	
## 143	Kenya	
## 144	Korea, South	
## 145	Kuwait	
## 146	Kyrgyzstan	
## 147	Latvia	
## 148	Lebanon	
## 149	Liberia	
## 150	Liechtenstein	
## 151	Lithuania	
## 152	Luxembourg	
## 153	Madagascar	
## 154	Malaysia	
## 155	Maldives	
## 156	Malta	
## 157	Mauritania	
## 158	Mauritius	
## 159	Mexico	
## 160	Moldova	
## 161	Monaco	
## 162	Mongolia	
## 163	Montenegro	
## 164	Morocco	
## 165	Namibia	
## 166	Nepal	
## 167	Netherlands	Aruba
## 168	Netherlands	Curacao
## 169	Netherlands	Sint Maarten

170 Netherlands
171 New Zealand
172 Nicaragua
173 Niger
174 Nigeria
175 North Macedonia
176 Norway
177 Oman
178 Pakistan
179 Panama
180 Papua New Guinea
181 Paraguay
182 Peru
183 Philippines
184 Poland
185 Portugal
186 Qatar
187 Romania
188 Russia
189 Rwanda
190 Saint Lucia
191 Saint Vincent and the Grenadines
192 San Marino
193 Saudi Arabia
194 Senegal
195 Serbia
196 Seychelles
197 Singapore
198 Slovakia
199 Slovenia
200 Somalia
201 South Africa
202 Spain
203 Sri Lanka
204 Sudan
205 Suriname
206 Sweden
207 Switzerland
208 Taiwan*
209 Tanzania
210 Thailand
211 Togo
212 Trinidad and Tobago
213 Tunisia
214 Turkey
215 Uganda

## 216			Ukraine	
## 217			United Arab Emirates	
## 218			United Kingdom	Bermuda
## 219			United Kingdom	Cayman Islands
## 220			United Kingdom	Channel Islands
## 221			United Kingdom	Gibraltar
## 222			United Kingdom	Isle of Man
## 223			United Kingdom	Montserrat
## 224			United Kingdom	
## 225			Uruguay	
## 226			US	
## 227			Uzbekistan	
## 228			Venezuela	
## 229			Vietnam	
## 230			Zambia	
## 231			Zimbabwe	
## 232			Canada	Diamond Princess
## 233			Dominica	
## 234			Grenada	
## 235			Mozambique	
## 236			Syria	
## 237			Timor-Leste	
## 238			Belize	
## 239			Canada	Recovered
## 240			Laos	
## 241			Libya	
## 242			West Bank and Gaza	
## 243			Guinea-Bissau	
## 244			Mali	
## 245			Saint Kitts and Nevis	
## 246			Canada	Northwest Territories
## 247			Canada	Yukon
## 248			Kosovo	
## 249			Burma	
## 250			United Kingdom	Anguilla
## 251			United Kingdom	British Virgin Islands
## 252			United Kingdom	Turks and Caicos Islands
## 253			MS Zaandam	
## 254			Botswana	
## 255			Burundi	
## 256			Sierra Leone	
## 257			Netherlands	Bonaire, Sint Eustatius and Saba
## 258			Malawi	
##	Lat	Long	confirmed	
## 1	33.000000	65.000000	273	
## 2	41.153300	20.168300	277	

## 3	28.033900	1.659600	986
## 4	42.506300	1.521800	428
## 5	-11.202700	17.873900	8
## 6	17.060800	-61.796400	9
## 7	-38.416100	-63.616700	1133
## 8	40.069100	45.038200	663
## 9	-35.473500	149.012400	87
## 10	-33.868800	151.209300	2298
## 11	-12.463400	130.845600	21
## 12	-28.016700	153.400000	835
## 13	-34.928500	138.600700	367
## 14	-41.454500	145.970700	72
## 15	-37.813600	144.963100	1036
## 16	-31.950500	115.860500	400
## 17	47.516200	14.550100	11129
## 18	40.143100	47.576900	400
## 19	25.034300	-77.396300	24
## 20	26.027500	50.550000	643
## 21	23.685000	90.356300	56
## 22	13.193900	-59.543200	46
## 23	53.709800	27.953400	304
## 24	50.833300	4.000000	15348
## 25	9.307700	2.315800	13
## 26	27.514200	90.433600	5
## 27	-16.290200	-63.588700	123
## 28	43.915900	17.679100	533
## 29	-14.235000	-51.925300	8044
## 30	4.535300	114.727700	133
## 31	42.733900	25.485800	457
## 32	12.238300	-1.561600	288
## 33	16.538800	-23.041800	6
## 34	11.550000	104.916700	110
## 35	3.848000	11.502100	306
## 36	53.933300	-116.576500	969
## 37	49.282700	-123.120700	1121
## 38	37.648900	-122.665500	13
## 39	53.760900	-98.813900	167
## 40	46.565300	-66.461900	91
## 41	53.135500	-57.660400	183
## 42	44.682000	-63.744300	193
## 43	51.253800	-85.323200	2793
## 44	46.510700	-63.416800	22
## 45	52.939900	-73.549100	5518
## 46	52.939900	-106.450900	206
## 47	6.611100	20.939400	3
## 48	15.454200	18.732200	8

## 49	-35.675100	-71.543000	3404
## 50	31.825700	117.226400	990
## 51	40.182400	116.414200	582
## 52	30.057200	107.874000	579
## 53	26.078900	117.987400	345
## 54	37.809900	101.058300	138
## 55	23.341700	113.424400	1507
## 56	23.829800	108.788100	254
## 57	26.815400	106.874800	146
## 58	19.195900	109.745300	168
## 59	39.549000	116.130600	325
## 60	47.862000	127.761500	488
## 61	33.882000	113.614000	1276
## 62	22.300000	114.200000	802
## 63	30.975600	112.270700	67802
## 64	27.610400	111.708800	1019
## 65	44.093500	113.944800	117
## 66	32.971100	119.455000	647
## 67	27.614000	115.722100	937
## 68	43.666100	126.192300	98
## 69	41.295600	122.608500	141
## 70	22.166700	113.550000	41
## 71	37.269200	106.165500	75
## 72	35.745200	95.995600	18
## 73	35.191700	108.870100	255
## 74	36.342700	118.149800	775
## 75	31.202000	121.449100	522
## 76	37.577700	112.292200	137
## 77	30.617100	102.710300	554
## 78	39.305400	117.323000	176
## 79	31.692700	88.092400	1
## 80	41.112900	85.240100	76
## 81	24.974000	101.487000	183
## 82	29.183200	120.093400	1258
## 83	4.570900	-74.297300	1161
## 84	-4.038300	21.758700	22
## 85	-4.038300	21.758700	134
## 86	9.748900	-83.753400	396
## 87	7.540000	-5.547100	194
## 88	45.100000	15.200000	1011
## 89	0.000000	0.000000	712
## 90	22.000000	-80.000000	233
## 91	35.126400	33.429900	356
## 92	49.817500	15.473000	3858
## 93	61.892600	-6.911800	177
## 94	71.706900	-42.604300	10

## 95	56.263900	9.501800	3386
## 96	11.825100	42.590300	40
## 97	18.735700	-70.162700	1380
## 98	-1.831200	-78.183400	3163
## 99	26.000000	30.000000	865
## 100	13.794200	-88.896500	41
## 101	1.500000	10.000000	15
## 102	15.179400	39.782300	22
## 103	58.595300	25.013600	858
## 104	-26.522500	31.465900	9
## 105	9.145000	40.489700	29
## 106	-17.713400	178.065000	7
## 107	64.000000	26.000000	1518
## 108	3.933900	-53.125800	51
## 109	-17.679700	149.406800	37
## 110	16.250000	-61.583300	128
## 111	-12.827500	45.166200	116
## 112	-20.904300	165.618000	18
## 113	-21.135100	55.247100	308
## 114	17.900000	-62.833300	6
## 115	18.070800	-63.050100	22
## 116	14.641500	-61.024200	138
## 117	46.227600	2.213700	59105
## 118	-0.803700	11.609400	21
## 119	13.443200	-15.310100	4
## 120	42.315400	43.356900	134
## 121	51.000000	9.000000	84794
## 122	7.946500	-1.023200	204
## 123	39.074200	21.824300	1544
## 124	15.783500	-90.230800	47
## 125	9.945600	-9.696600	52
## 126	5.000000	-58.750000	19
## 127	18.971200	-72.285200	16
## 128	41.902900	12.453400	7
## 129	15.200000	-86.241900	219
## 130	47.162500	19.503300	585
## 131	64.963100	-19.020800	1319
## 132	21.000000	78.000000	2543
## 133	-0.789300	113.921300	1790
## 134	32.000000	53.000000	50468
## 135	33.000000	44.000000	772
## 136	53.142400	-7.692100	3849
## 137	31.000000	35.000000	6857
## 138	43.000000	12.000000	115242
## 139	18.109600	-77.297500	47
## 140	36.000000	138.000000	2495

## 141	31.240000	36.510000	299
## 142	48.019600	66.923700	435
## 143	-0.023600	37.906200	110
## 144	36.000000	128.000000	9976
## 145	29.500000	47.750000	342
## 146	41.204400	74.766100	116
## 147	56.879600	24.603200	458
## 148	33.854700	35.862300	494
## 149	6.428100	-9.429500	6
## 150	47.140000	9.550000	75
## 151	55.169400	23.881300	649
## 152	49.815300	6.129600	2487
## 153	-18.766900	46.869100	59
## 154	2.500000	112.500000	3116
## 155	3.202800	73.220700	19
## 156	35.937500	14.375400	196
## 157	21.007900	10.940800	6
## 158	-20.200000	57.500000	169
## 159	23.634500	-102.552800	1378
## 160	47.411600	28.369900	505
## 161	43.733300	7.416700	60
## 162	46.862500	103.846700	14
## 163	42.500000	19.300000	144
## 164	31.791700	-7.092600	708
## 165	-22.957600	18.490400	14
## 166	28.166700	84.250000	6
## 167	12.518600	-70.035800	60
## 168	12.169600	-68.990000	11
## 169	18.042500	-63.054800	18
## 170	52.132600	5.291300	14697
## 171	-40.900600	174.886000	797
## 172	12.865400	-85.207200	5
## 173	17.607800	8.081700	98
## 174	9.082000	8.675300	184
## 175	41.608600	21.745300	384
## 176	60.472000	8.468900	5147
## 177	21.000000	57.000000	231
## 178	30.375300	69.345100	2421
## 179	8.538000	-80.782100	1317
## 180	-6.315000	143.955500	1
## 181	-23.442500	-58.443800	77
## 182	-9.190000	-75.015200	1414
## 183	13.000000	122.000000	2633
## 184	51.919400	19.145100	2946
## 185	39.399900	-8.224500	9034
## 186	25.354800	51.183900	949

## 187	45.943200	24.966800	2738
## 188	60.000000	90.000000	3548
## 189	-1.940300	29.873900	84
## 190	13.909400	-60.978900	13
## 191	12.984300	-61.287200	2
## 192	43.942400	12.457800	245
## 193	24.000000	45.000000	1885
## 194	14.497400	-14.452400	195
## 195	44.016500	21.005900	1171
## 196	-4.679600	55.492000	10
## 197	1.283300	103.833300	1049
## 198	48.669000	19.699000	426
## 199	46.151200	14.995500	897
## 200	5.152100	46.199600	5
## 201	-30.559500	22.937500	1462
## 202	40.000000	-4.000000	112065
## 203	7.000000	81.000000	151
## 204	12.862800	30.217600	8
## 205	3.919300	-56.027800	10
## 206	63.000000	16.000000	5568
## 207	46.818200	8.227500	18827
## 208	23.700000	121.000000	339
## 209	-6.369000	34.888800	20
## 210	15.000000	101.000000	1875
## 211	8.619500	0.824800	39
## 212	10.691800	-61.222500	94
## 213	34.000000	9.000000	455
## 214	38.963700	35.243300	18135
## 215	1.000000	32.000000	45
## 216	48.379400	31.165600	897
## 217	24.000000	54.000000	1024
## 218	32.307800	-64.750500	35
## 219	19.313300	-81.254600	28
## 220	49.372300	-2.364400	193
## 221	36.140800	-5.353600	88
## 222	54.236100	-4.548100	95
## 223	16.742500	-62.187400	5
## 224	55.378100	-3.436000	33718
## 225	-32.522800	-55.765800	350
## 226	37.090200	-95.712900	243453
## 227	41.377500	64.585300	205
## 228	6.423800	-66.589700	146
## 229	16.000000	108.000000	233
## 230	-15.416700	28.283300	39
## 231	-20.000000	30.000000	9
## 232	0.000000	0.000000	0

## 233	15.415000	-61.371000	12
## 234	12.116500	-61.679000	10
## 235	-18.665695	35.529562	10
## 236	34.802075	38.996815	16
## 237	-8.874217	125.727539	1
## 238	13.193900	-59.543200	3
## 239	0.000000	0.000000	0
## 240	19.856270	102.495496	10
## 241	26.335100	17.228331	11
## 242	31.952200	35.233200	161
## 243	11.803700	-15.180400	9
## 244	17.570692	-3.996166	36
## 245	17.357822	-62.782998	9
## 246	64.825500	-124.845700	2
## 247	64.282300	-135.000000	6
## 248	42.602636	20.902977	125
## 249	21.916200	95.956000	20
## 250	18.220600	-63.068600	3
## 251	18.420700	-64.640000	3
## 252	21.694000	-71.797900	5
## 253	0.000000	0.000000	9
## 254	-22.328500	24.684900	4
## 255	-3.373100	29.918900	3
## 256	8.460555	-11.779889	2
## 257	12.178400	-68.238500	2
## 258	-13.254308	34.301525	3
##			txt
## 1			Afghanistan-:273
## 2			Albania-:277
## 3			Algeria-:986
## 4			Andorra-:428
## 5			Angola-:8
## 6			Antigua and Barbuda-:9
## 7			Argentina-:1133
## 8			Armenia-:663
## 9			Australia-Australian Capital Territory:87
## 10			Australia-New South Wales:2298
## 11			Australia-Northern Territory:21
## 12			Australia-Queensland:835
## 13			Australia-South Australia:367
## 14			Australia-Tasmania:72
## 15			Australia-Victoria:1036
## 16			Australia-Western Australia:400
## 17			Austria-:11129
## 18			Azerbaijan-:400
## 19			Bahamas-:24

20 Bahrain-:643
 ## 21 Bangladesh-:56
 ## 22 Barbados-:46
 ## 23 Belarus-:304
 ## 24 Belgium-:15348
 ## 25 Benin-:13
 ## 26 Bhutan-:5
 ## 27 Bolivia-:123
 ## 28 Bosnia and Herzegovina-:533
 ## 29 Brazil-:8044
 ## 30 Brunei-:133
 ## 31 Bulgaria-:457
 ## 32 Burkina Faso-:288
 ## 33 Cabo Verde-:6
 ## 34 Cambodia-:110
 ## 35 Cameroon-:306
 ## 36 Canada-Alberta:969
 ## 37 Canada-British Columbia:1121
 ## 38 Canada-Grand Princess:13
 ## 39 Canada-Manitoba:167
 ## 40 Canada-New Brunswick:91
 ## 41 Canada-Newfoundland and Labrador:183
 ## 42 Canada-Nova Scotia:193
 ## 43 Canada-Ontario:2793
 ## 44 Canada-Prince Edward Island:22
 ## 45 Canada-Quebec:5518
 ## 46 Canada-Saskatchewan:206
 ## 47 Central African Republic-:3
 ## 48 Chad-:8
 ## 49 Chile-:3404
 ## 50 China-Anhui:990
 ## 51 China-Beijing:582
 ## 52 China-Chongqing:579
 ## 53 China-Fujian:345
 ## 54 China-Gansu:138
 ## 55 China-Guangdong:1507
 ## 56 China-Guangxi:254
 ## 57 China-Guizhou:146
 ## 58 China-Hainan:168
 ## 59 China-Hebei:325
 ## 60 China-Heilongjiang:488
 ## 61 China-Henan:1276
 ## 62 China-Hong Kong:802
 ## 63 China-Hubei:67802
 ## 64 China-Hunan:1019
 ## 65 China-Inner Mongolia:117

66 China-Jiangsu:647
 ## 67 China-Jiangxi:937
 ## 68 China-Jilin:98
 ## 69 China-Liaoning:141
 ## 70 China-Macau:41
 ## 71 China-Ningxia:75
 ## 72 China-Qinghai:18
 ## 73 China-Shaanxi:255
 ## 74 China-Shandong:775
 ## 75 China-Shanghai:522
 ## 76 China-Shanxi:137
 ## 77 China-Sichuan:554
 ## 78 China-Tianjin:176
 ## 79 China-Tibet:1
 ## 80 China-Xinjiang:76
 ## 81 China-Yunnan:183
 ## 82 China-Zhejiang:1258
 ## 83 Colombia-:1161
 ## 84 Congo (Brazzaville)-:22
 ## 85 Congo (Kinshasa)-:134
 ## 86 Costa Rica-:396
 ## 87 Cote d'Ivoire-:194
 ## 88 Croatia-:1011
 ## 89 Diamond Princess-:712
 ## 90 Cuba-:233
 ## 91 Cyprus-:356
 ## 92 Czechia-:3858
 ## 93 Denmark-Faroe Islands:177
 ## 94 Denmark-Greenland:10
 ## 95 Denmark-:3386
 ## 96 Djibouti-:40
 ## 97 Dominican Republic-:1380
 ## 98 Ecuador-:3163
 ## 99 Egypt-:865
 ## 100 El Salvador-:41
 ## 101 Equatorial Guinea-:15
 ## 102 Eritrea-:22
 ## 103 Estonia-:858
 ## 104 Eswatini-:9
 ## 105 Ethiopia-:29
 ## 106 Fiji-:7
 ## 107 Finland-:1518
 ## 108 France-French Guiana:51
 ## 109 France-French Polynesia:37
 ## 110 France-Guadeloupe:128
 ## 111 France-Mayotte:116

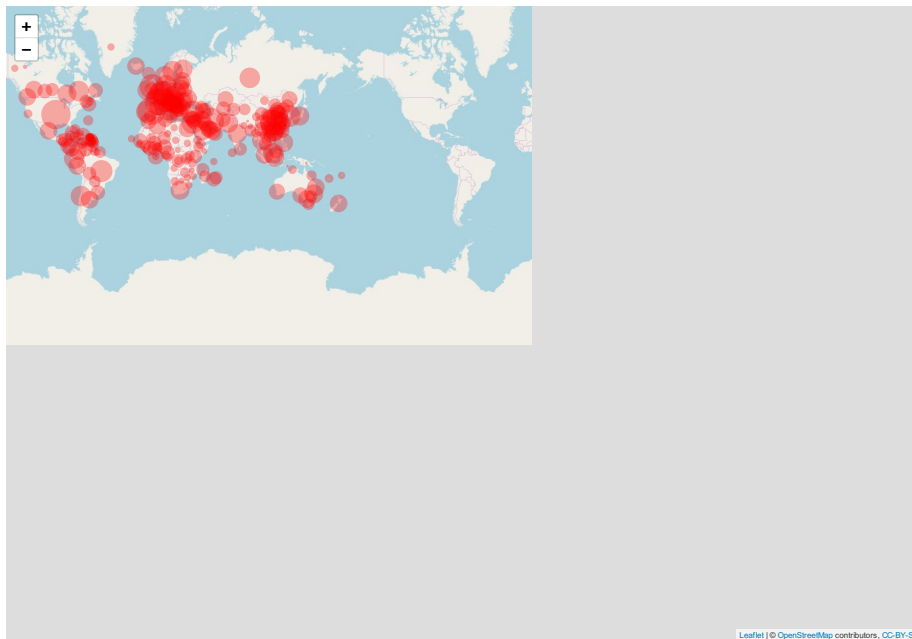
## 112	France-New Caledonia:18
## 113	France-Reunion:308
## 114	France-Saint Barthelemy:6
## 115	France-St Martin:22
## 116	France-Martinique:138
## 117	France-:59105
## 118	Gabon-:21
## 119	Gambia-:4
## 120	Georgia-:134
## 121	Germany-:84794
## 122	Ghana-:204
## 123	Greece-:1544
## 124	Guatemala-:47
## 125	Guinea-:52
## 126	Guyana-:19
## 127	Haiti-:16
## 128	Holy See-:7
## 129	Honduras-:219
## 130	Hungary-:585
## 131	Iceland-:1319
## 132	India-:2543
## 133	Indonesia-:1790
## 134	Iran-:50468
## 135	Iraq-:772
## 136	Ireland-:3849
## 137	Israel-:6857
## 138	Italy-:115242
## 139	Jamaica-:47
## 140	Japan-:2495
## 141	Jordan-:299
## 142	Kazakhstan-:435
## 143	Kenya-:110
## 144	Korea, South-:9976
## 145	Kuwait-:342
## 146	Kyrgyzstan-:116
## 147	Latvia-:458
## 148	Lebanon-:494
## 149	Liberia-:6
## 150	Liechtenstein-:75
## 151	Lithuania-:649
## 152	Luxembourg-:2487
## 153	Madagascar-:59
## 154	Malaysia-:3116
## 155	Maldives-:19
## 156	Malta-:196
## 157	Mauritania-:6

## 158	Mauritius-:169
## 159	Mexico-:1378
## 160	Moldova-:505
## 161	Monaco-:60
## 162	Mongolia-:14
## 163	Montenegro-:144
## 164	Morocco-:708
## 165	Namibia-:14
## 166	Nepal-:6
## 167	Netherlands-Aruba:60
## 168	Netherlands-Curacao:11
## 169	Netherlands-Sint Maarten:18
## 170	Netherlands-:14697
## 171	New Zealand-:797
## 172	Nicaragua-:5
## 173	Niger-:98
## 174	Nigeria-:184
## 175	North Macedonia-:384
## 176	Norway-:5147
## 177	Oman-:231
## 178	Pakistan-:2421
## 179	Panama-:1317
## 180	Papua New Guinea-:1
## 181	Paraguay-:77
## 182	Peru-:1414
## 183	Philippines-:2633
## 184	Poland-:2946
## 185	Portugal-:9034
## 186	Qatar-:949
## 187	Romania-:2738
## 188	Russia-:3548
## 189	Rwanda-:84
## 190	Saint Lucia-:13
## 191	Saint Vincent and the Grenadines-:2
## 192	San Marino-:245
## 193	Saudi Arabia-:1885
## 194	Senegal-:195
## 195	Serbia-:1171
## 196	Seychelles-:10
## 197	Singapore-:1049
## 198	Slovakia-:426
## 199	Slovenia-:897
## 200	Somalia-:5
## 201	South Africa-:1462
## 202	Spain-:112065
## 203	Sri Lanka-:151

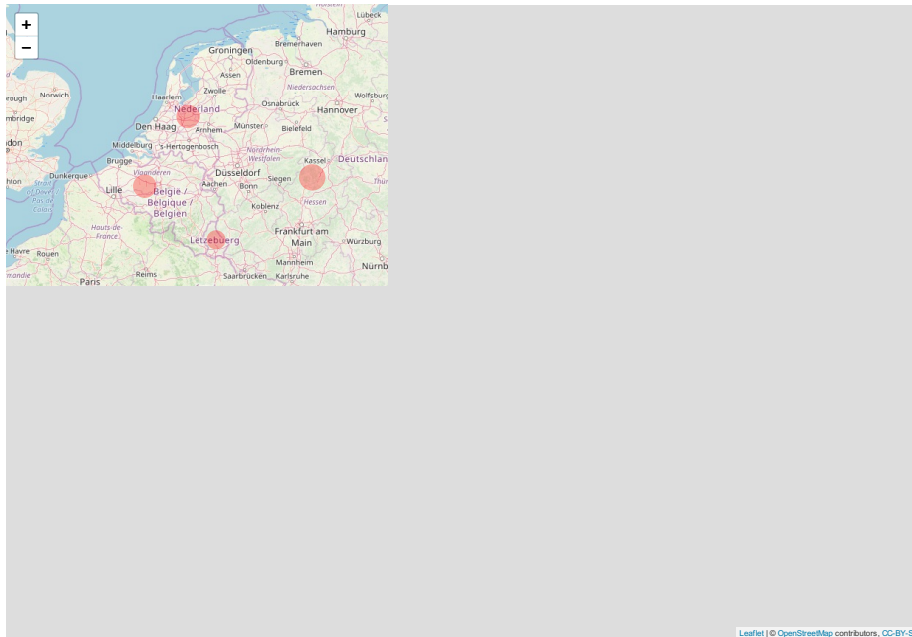
204 Sudan-:8
 ## 205 Suriname-:10
 ## 206 Sweden-:5568
 ## 207 Switzerland-:18827
 ## 208 Taiwan*-:339
 ## 209 Tanzania-:20
 ## 210 Thailand-:1875
 ## 211 Togo-:39
 ## 212 Trinidad and Tobago-:94
 ## 213 Tunisia-:455
 ## 214 Turkey-:18135
 ## 215 Uganda-:45
 ## 216 Ukraine-:897
 ## 217 United Arab Emirates-:1024
 ## 218 United Kingdom-Bermuda:35
 ## 219 United Kingdom-Cayman Islands:28
 ## 220 United Kingdom-Channel Islands:193
 ## 221 United Kingdom-Gibraltar:88
 ## 222 United Kingdom-Isle of Man:95
 ## 223 United Kingdom-Montserrat:5
 ## 224 United Kingdom-:33718
 ## 225 Uruguay-:350
 ## 226 US-:243453
 ## 227 Uzbekistan-:205
 ## 228 Venezuela-:146
 ## 229 Vietnam-:233
 ## 230 Zambia-:39
 ## 231 Zimbabwe-:9
 ## 232 Canada-Diamond Princess:0
 ## 233 Dominica-:12
 ## 234 Grenada-:10
 ## 235 Mozambique-:10
 ## 236 Syria-:16
 ## 237 Timor-Leste-:1
 ## 238 Belize-:3
 ## 239 Canada-Recovered:0
 ## 240 Laos-:10
 ## 241 Libya-:11
 ## 242 West Bank and Gaza-:161
 ## 243 Guinea-Bissau-:9
 ## 244 Mali-:36
 ## 245 Saint Kitts and Nevis-:9
 ## 246 Canada-Northwest Territories:2
 ## 247 Canada-Yukon:6
 ## 248 Kosovo-:125
 ## 249 Burma-:20

```
## 250 United Kingdom-Anguilla:3
## 251 United Kingdom-British Virgin Islands:3
## 252 United Kingdom-Turks and Caicos Islands:5
## 253 MS Zaandam-:9
## 254 Botswana-:4
## 255 Burundi-:3
## 256 Sierra Leone-:2
## 257 Netherlands-Bonaire, Sint Eustatius and Saba:2
## 258 Malawi-:3
```

```
map <- leaflet() %>% addTiles()
#marker
map %<>% addCircleMarkers(x$Long, x$Lat, radius = 2+log2(x$confirmed), stroke = F,
color = 'red', fillOpacity = 0.3, popup = x$txt)
map
```



```
map %>% setView(5, 52, zoom = 6)
```

Number of cases:

```
world.long <- data.long %>% filter(country == 'World') # can be also filtered for different
```

```
# area plot
```

```
plot1 <- world.long %>% filter(type != 'Total Confirmed') %>%
  ggplot(aes(x=date, y=count)) +
  geom_area(aes(fill=type), alpha=0.5) +
  labs(title=paste0('Cases Worldwide - ', max.date.txt)) +
  scale_fill_manual(values=c('red', 'green', 'black')) +
  theme(legend.title=element_blank(), legend.position='bottom',
        plot.title = element_text(size=8),
        axis.title.x=element_blank(),
        axis.title.y=element_blank(),
        legend.key.size=unit(0.2, 'cm'),
        legend.text=element_text(size=6),
        axis.text=element_text(size=7),
        axis.text.x=element_text(angle=45, hjust=1))
```

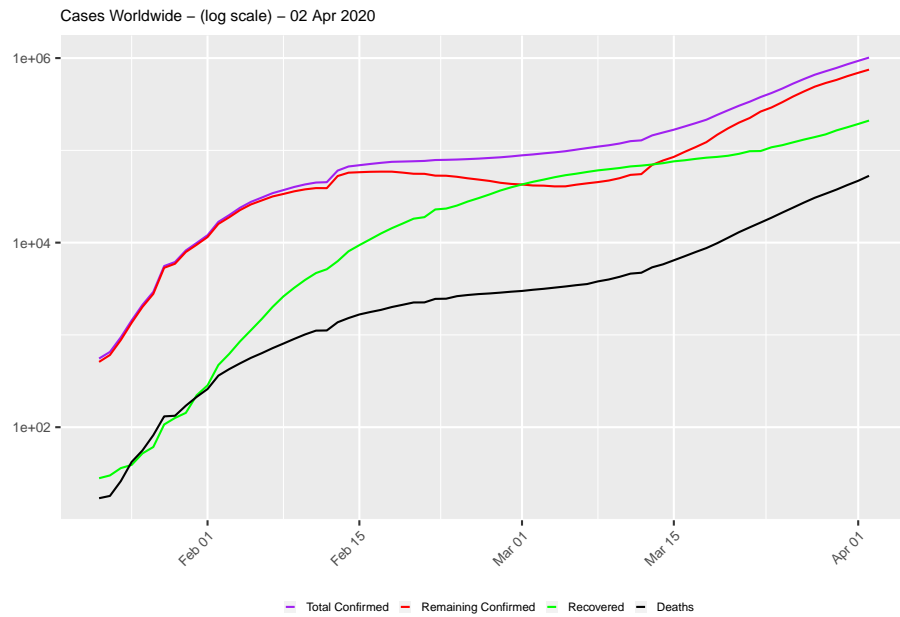
```
plot2 <- world.long %>%
  ggplot(aes(x=date, y=count)) +
  geom_line(aes(color=type)) +
  labs(title = paste0('Cases Worldwide - (log scale) - ', max.date.txt)) +
  scale_color_manual(values=c('purple', 'red', 'green', 'black')) +
```

```

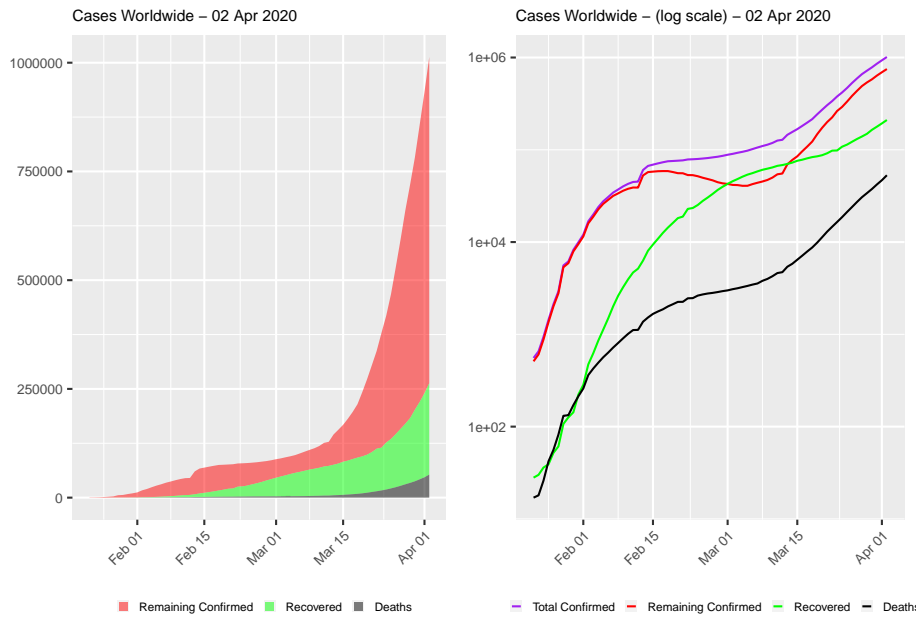
theme(legend.title=element_blank(), legend.position='bottom',
      plot.title = element_text(size =8),
      axis.title.x=element_blank(),
      axis.title.y = element_blank(),
      legend.key.size = unit(0.2, 'cm'),
      legend.text = element_text(size =6),
      axis.text = element_text(size = 7),
      axis.text.x =element_text(angle = 45, hjust = 1)) +
scale_y_continuous(trans = 'log10')

```

plot2



```
grid.arrange(plot1, plot2, ncol=2)
```



Current confirmed Cases:

```
data.world <- data %>% filter(country == 'World')
n <- nrow(data.world)

##current confirmed and daily new confirmed
plot1 <- ggplot(data.world, aes(x=date, y=remaining.confirmed)) +
  geom_point()+geom_smooth()+
  xlab('') + ylab('Count') + labs(title = 'Current Confirmed Cases') +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

plot2 <- ggplot(data.world, aes(x=date, y=confirmed.new))+ geom_point() + geom_smooth() + xlab('') +
  theme(axis.text.x = element_text(angle = 45, hjust=1))
```

```
## List of 1
## $ axis.text.x:List of 11
## ..$ family      : NULL
## ..$ face        : NULL
## ..$ colour      : NULL
## ..$ size        : NULL
## ..$ hjust       : num 1
## ..$ vjust       : NULL
## ..$ angle       : num 45
## ..$ lineheight  : NULL
## ..$ margin      : NULL
```

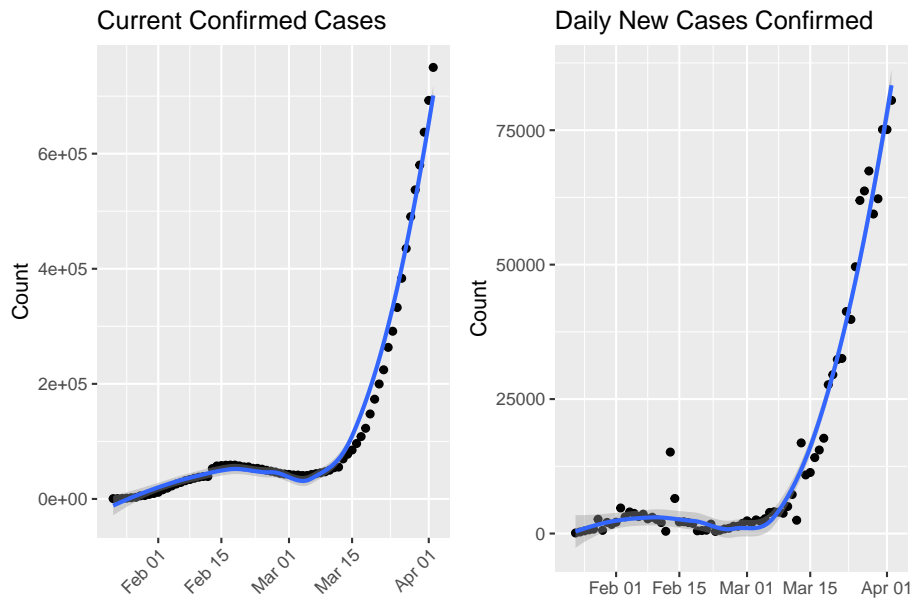
```
## ..$ debug          : NULL
## ..$ inherit.blank: logi FALSE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## - attr(*, "class")= chr [1:2] "theme" "gg"
## - attr(*, "complete")= logi FALSE
## - attr(*, "validate")= logi TRUE
```

```
grid.arrange(plot1, plot2, ncol=2)
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```

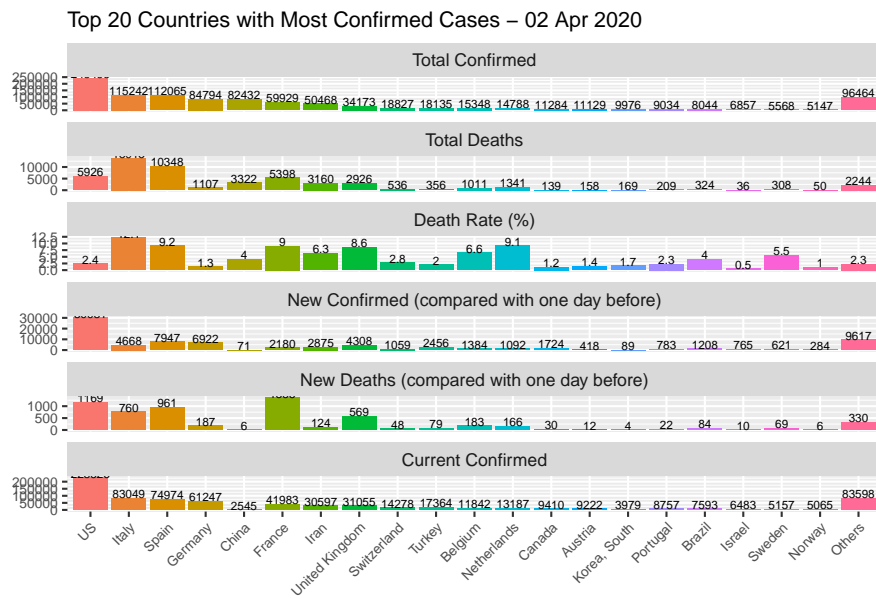


Bar Chart

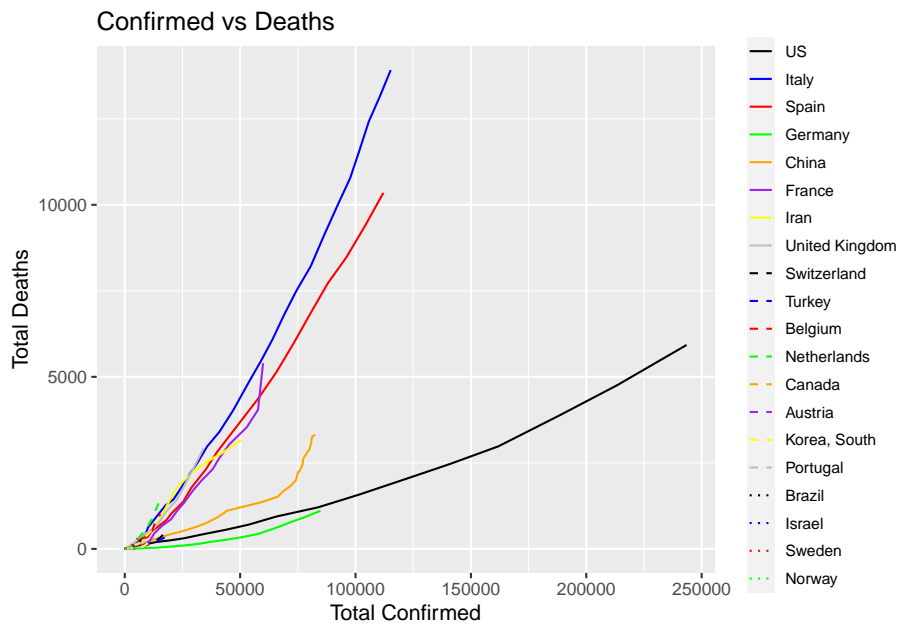
```
data.latest.long <- data.latest %>% filter(country!='World') %>% gather(key=type, value=count)
```

```
data.latest.long %<>% mutate(type=recode_factor(type, confirmed='Total Confirmed', deaths='Total Deaths'))
```

```
## bar chart
data.latest.long %>% ggplot(aes(x=country, y=count, fill=country, group=country)) +
  geom_bar(stat='identity') +
  geom_text(aes(label=count, y=count),size=2, vjust=0) +
  xlab('') + ylab('') +
  labs(title=paste0('Top 20 Countries with Most Confirmed Cases - ', max.date.txt))+ scale_y_log10()
theme(legend.title=element_blank(),
      legend.position='none',
      plot.title=element_text(size=11),axis.text=element_text(size=7), axis.text.x=element_text(size=7))
```



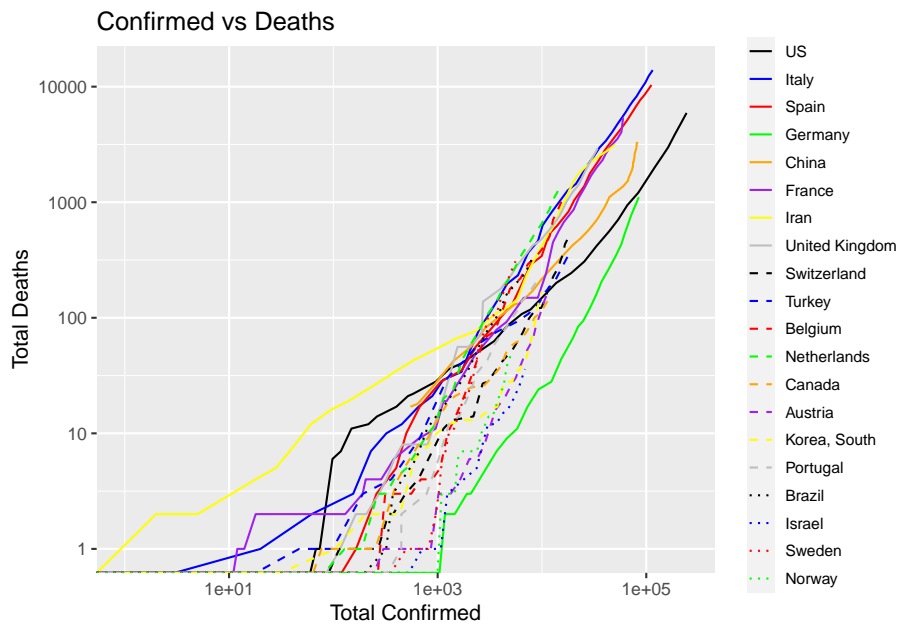
```
# Confirmed versus Deaths
linetypes <- rep(c("solid", "dashed", "dotted"), each=8)
colors <- rep(c('black', 'blue', 'red', 'green', 'orange', 'purple', 'yellow', 'grey'), 3)
df <- data %>% filter(country %in% setdiff(top.countries, c('World', 'Others')) %>%
mutate(country=country %>% factor(levels=c(top.countries)))
vs <- df %>% ggplot(aes(x=confirmed, y=deaths, group=country)) +
  geom_line(aes(color=country, linetype=country)) +
  xlab('Total Confirmed') + ylab('Total Deaths') +
  scale_linetype_manual(values=linetypes) +
  scale_color_manual(values=colors) +
  theme(legend.title=element_blank(),
        legend.text=element_text(size=8),
        legend.key.size=unit(0.5, 'cm')) + ggtitle('Confirmed vs Deaths')
vs
```



```
vs + scale_x_log10() + scale_y_log10()
```

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

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

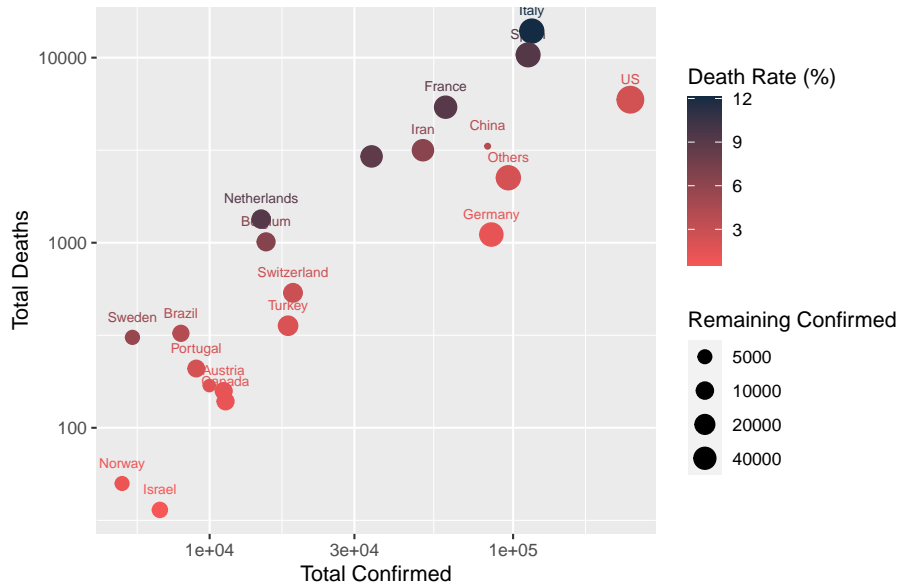


Number of confirmed cases and deaths in top 20 countries.

```
df <- data.latest %>% filter(country %in% setdiff(top.countries, 'World'))

plot1 <- df %>% ggplot(aes(x=confirmed, y=deaths, col=death.rate, size=remaining.confirmed))
  scale_size(name='Remaining Confirmed', trans='log2', breaks=c(1e3, 2e3, 5e3, 1e4, 2e4, 4e4))
  geom_text(aes(label=country), size=2.5, check_overlap=T, vjust=-1.6) +
  geom_point() +
  xlab('Total Confirmed') + ylab('Total Deaths') +
  labs(col="Death Rate (%)") +
  scale_color_gradient(low='#f75656', high='#132B43') +
  scale_x_log10() + scale_y_log10()
plot1
```

Number of confirmed cases and deaths in top 20 countries.

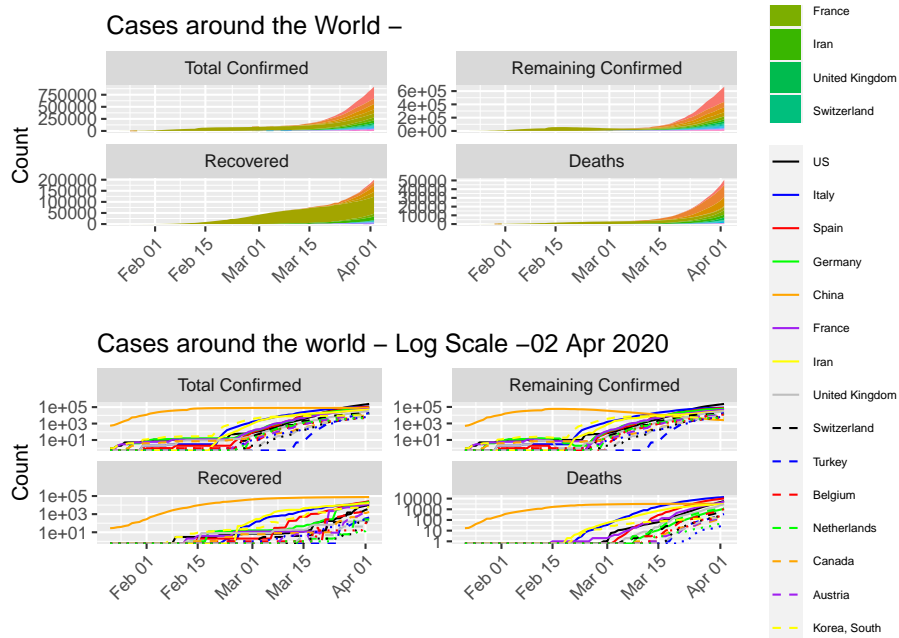


```
df <- data.long %>% filter(country %in% top.countries) %<>% mutate(country=country %>% factor)

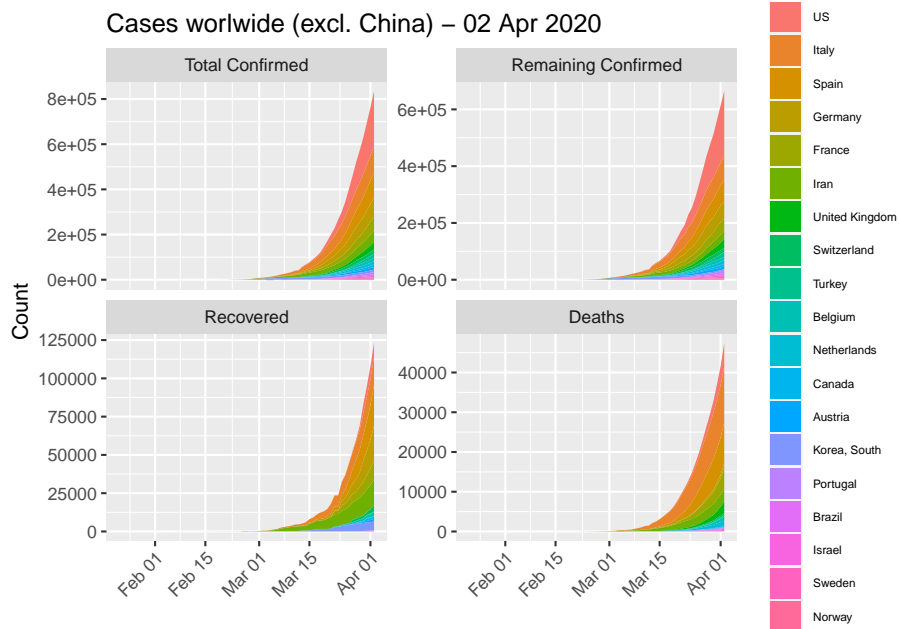
### CASES AROUND WORLD
p <- df %>% filter(country != 'World') %>%
  ggplot(aes(x=date, y=count)) + xlab('') + ylab('Count') +
  theme(legend.title=element_blank(),
        legend.text = element_text(size=6),
        legend.key.size=unit(0.6, 'cm'),
        axis.text.x=element_text(angle = 45, hjust=1)) +
  facet_wrap(~type, ncol = 2, scale='free_y')
# area plot
plot1 <- p + geom_area(aes(fill=country)) +
  labs(title='Cases around the World - ', max.date.txt)

# line plot and in log scale
#linetypes <- rep(c('solid','dashed','dotted'), each=8)
#colors <- rep(c('black','blue','red','green','orange','purple','yellow','grey'), 3)
plot2 <- p + geom_line(aes(color=country, linetype=country)) +
  scale_linetype_manual(values = linetypes) +
  scale_color_manual(values = colors) +
  labs(title =paste0('Cases around the world - Log Scale -', max.date.txt)) +
  scale_y_continuous(trans = 'log10')
grid.arrange(plot1, plot2, ncol=1)
```


Warning: Transformation introduced infinite values in continuous y-axis



```
# Plot: excluding China
p <- df %>% filter(!(country %in% c('World', 'China'))) %>%
  ggplot(aes(x=date, y=count)) + xlab('') + ylab('Count') +
  theme(legend.title=element_blank(),
        legend.text = element_text(size=6),
        legend.key.size=unit(0.6, 'cm'),
        axis.text.x=element_text(angle = 45, hjust=1)) +
  facet_wrap(~type, ncol = 2, scale='free_y')
p + geom_area(aes(fill=country)) +
  labs(title=paste0('Cases worldwide (excl. China) - ', max.date.txt))
```



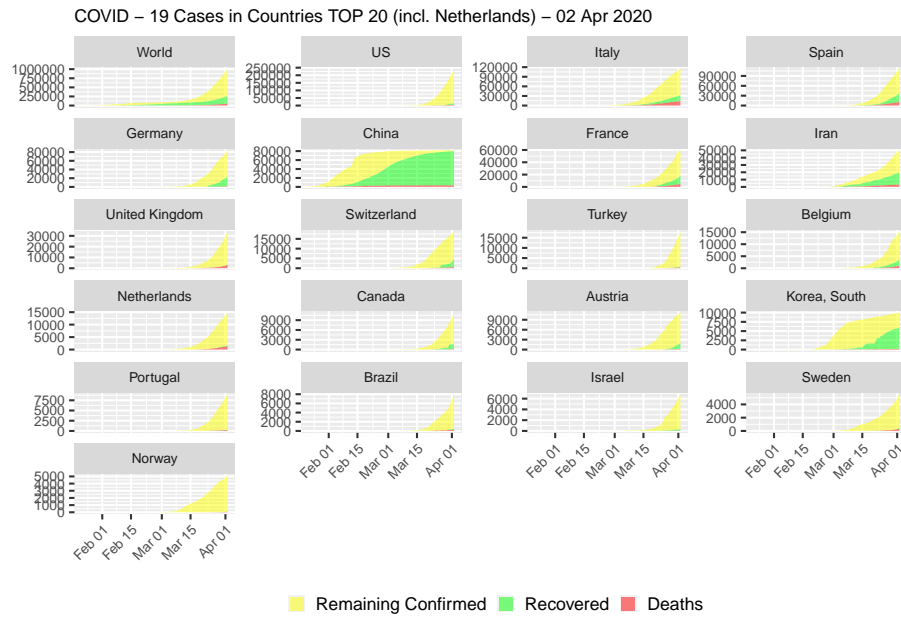
```

### list(countries) == 'Netherlands'

## If The Netherland is not top 20, add it in and remove 'Others'
if(!('Netherlands' %in% top.countries)) {
  top.countries %<>% setdiff('Others') %>% c('Netherlands')
  df <- data.long %>% filter(country %in% top.countries) %>%
    mutate(country=country %>% factor(levels = c(top.countries)))
}

# cases by country - area plot
df %>% filter(type != 'World' & type != 'Total Confirmed') %>%
  ggplot(aes(x=date, y=count, fill=type)) +
  geom_area(alpha=0.5) +
  labs(title = paste0('COVID - 19 Cases in Countries TOP 20 (incl. Netherlands) - ', max.date)) +
  scale_fill_manual(values=c('yellow','green','red')) +
  theme(legend.title=element_blank(), legend.position='bottom',
        plot.title= element_text(size = 9),
        axis.title.x=element_blank(),
        axis.title.y = element_blank(),
        legend.key.size = unit(0.3, 'cm'),
        strip.text.x = element_text(size=7),
        axis.text=element_text(size = 7),
        axis.text.x = element_text(angle=45, hjust=1)) +
  facet_wrap(~country, ncol=4, scale='free_y') + facet_wrap(-country, ncol=4, scales = 'free')

```



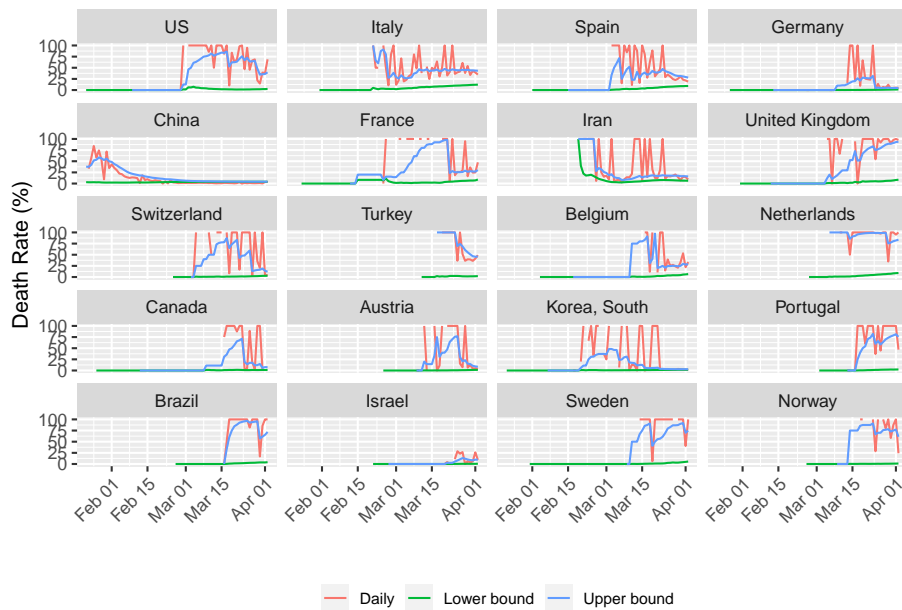
Deaths rate:

```
rate.max <- rates.long$count %>% max(na.rm=T)

df <- rates.long %>% filter(country %in% setdiff(top.countries, 'World')) %>%
  mutate(country=factor(country, levels=top.countries))

df %>% ggplot(aes(x=date, y=count, color=type)) +
  geom_line() +
  xlab('') + ylab('Death Rate (%)') +
  theme(legend.position='bottom', legend.title=element_blank(),
        legend.text=element_text(size=8),
        legend.key.size=unit(0.5, 'cm'),
        axis.text.x=element_text(angle=45, hjust=1)) +
  ylim(c(0, 100)) +
  facet_wrap(~country, ncol=4)
```

Warning: Removed 36 row(s) containing missing values (geom_path).



Countries with Highest Death Rates

```

## sort the latest data by death rate, and if tie, by confirmed
df <- data %>% filter(date == max(date) & country != 'World' & confirmed >= 100) %>%
  select(country, confirmed, confirmed.new, remaining.confirmed,
         recovered, deaths, deaths.new, death.rate=rate.lower) %>%
  arrange(desc(death.rate, confirmed))

df %>% head(20) %>%
  mutate(death.rate=death.rate %>% format(nsmall=1) %>% paste0('%')) %>%
  kable('latex', booktabs=T, row.names=T, align=c('l', rep('r', 7)),
        caption=paste0('Top 20 Countries with Highest Death Rates - ', max.date.txt), format
  kable_styling(font_size=7, latex_options=c('striped', 'hold_position', 'repeat_header'))

```

Note that this is an developing story. Check back for updates.

Table 2: Top 20 Countries with Highest Death Rates - 02 Apr 2020

	country	confirmed	confirmed.new	remaining.confirmed	recovered	deaths	deaths.new	death.rate
1	San Marino	245	9	194	21	30	4	12.2%
2	Italy	115,242	4,668	83,049	18,278	13,915	760	12.1%
3	Congo (Kinshasa)	134	25	118	3	13	4	9.7%
4	Indonesia	1,790	113	1,508	112	170	13	9.5%
5	Spain	112,065	7,947	74,974	26,743	10,348	961	9.2%
6	Netherlands	14,788	1,092	13,187	260	1,341	166	9.1%
7	France	59,929	2,180	41,983	12,548	5,398	1,355	9.0%
8	Algeria	986	139	839	61	86	28	8.7%
9	United Kingdom	34,173	4,308	31,055	192	2,926	569	8.6%
10	Iraq	772	44	516	202	54	2	7.0%
11	Egypt	865	86	606	201	58	6	6.7%
12	Belgium	15,348	1,384	11,842	2,495	1,011	183	6.6%
13	Bolivia	123	8	114	1	8	1	6.5%
14	Honduras	219	47	202	3	14	4	6.4%
15	Iran	50,468	2,875	30,597	16,711	3,160	124	6.3%
16	Morocco	708	54	633	31	44	5	6.2%
17	Albania	277	18	185	76	16	1	5.8%
18	Burkina Faso	288	6	222	50	16	0	5.6%
19	Sweden	5,568	621	5,157	103	308	69	5.5%
20	Dominican Republic	1,380	96	1,304	16	60	3	4.3%