

rank

ryang273

26/11/2020

```
library(tidyverse)
```

```
## -- Attaching packages -----  
## v ggplot2 3.3.2      v purrr  0.3.4  
## v tibble  3.0.3      v dplyr  1.0.2  
## v tidyr   1.1.2      v stringr 1.4.0  
## v readr   1.4.0      v forcats 0.5.0  
  
## -- Conflicts -----  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()    masks stats::lag()
```

```
library(reshape2)
```

```
##  
## Attaching package: 'reshape2'  
  
## The following object is masked from 'package:tidyr':  
##  
## smiths
```

```
library(ggplot2)  
require(cowplot)
```

```
## Loading required package: cowplot
```

```
Confirmed <- read_csv("/Users/yangruiqin/Desktop/2864/time_series_covid19_confirmed_global.csv")
```

```
##  
## -- Column specification -----  
## cols(  
##   .default = col_double(),  
##   `Province/State` = col_character(),  
##   `Country/Region` = col_character()  
## )  
## i Use `spec()` for the full column specifications.
```

```
Recovered <- read_csv("/Users/yangruiqin/Desktop/2864/time_series_covid19_recovered_global.csv")
```

```
##  
## -- Column specification -----  
## cols(  
##   .default = col_double(),  
##   `Province/State` = col_character(),  
##   `Country/Region` = col_character()  
## )
```

```

## i Use `spec()` for the full column specifications.
Deaths <- read_csv("/Users/yangruiqin/Desktop/2864/time_series_covid19_deaths_global.csv")

##
## -- Column specification -----
## cols(
##   .default = col_double(),
##   `Province/State` = col_character(),
##   `Country/Region` = col_character()
## )
## i Use `spec()` for the full column specifications.
Confirmed_data <- Confirmed %>%
  group_by(`Country/Region`) %>%
  summarise(Confirmed = sum(`11/22/20`, na.rm = TRUE)) %>%
  rename(Region = `Country/Region`)

## `summarise()` ungrouping output (override with `.groups` argument)
Recovered_data <- Recovered %>%
  group_by(`Country/Region`) %>%
  summarise(Recovered = sum(`11/22/20`, na.rm = TRUE)) %>%
  rename(Region = `Country/Region`)

## `summarise()` ungrouping output (override with `.groups` argument)
Deaths_data <- Deaths %>%
  group_by(`Country/Region`) %>%
  summarise(Deaths = sum(`11/22/20`, na.rm = TRUE)) %>%
  rename(Region = `Country/Region`)

## `summarise()` ungrouping output (override with `.groups` argument)
global <- Confirmed_data %>%
  left_join(Recovered_data, by = "Region") %>%
  left_join(Deaths_data, by = "Region")
data.confirmed <- Confirmed_data %>%
  select(Region, Confirmed,) %>%
  mutate(ranking = dense_rank(desc(Confirmed)))
top.countries_confirmed <- data.confirmed %>% filter(ranking <= 10) %>%
  arrange(ranking) %>% pull(Region) %>% as.character()
top.countries_confirmed %>% print()

## [1] "US" "India" "Brazil" "France"
## [5] "Russia" "Spain" "United Kingdom" "Italy"
## [9] "Argentina" "Colombia"

data.recovered <- Recovered_data %>%
  select(Region, Recovered,) %>%
  mutate(ranking = dense_rank(desc(Recovered)))
top.countries_recovered <- data.recovered %>% filter(ranking <= 10) %>%
  arrange(ranking) %>% pull(Region) %>% as.character()
top.countries_recovered %>% print()

## [1] "India" "Brazil" "US" "Russia" "Argentina"
## [6] "Colombia" "Peru" "Mexico" "South Africa" "Germany"

```

```
data.deaths <- Deaths_data %>%
  select(Region, Deaths,) %>%
  mutate(ranking = dense_rank(desc(Deaths)))
top.countries_deaths <- data.deaths %>% filter(ranking <= 10) %>%
  arrange(ranking) %>% pull(Region) %>% as.character()
top.countries_confirmed %>% print()
```

```
## [1] "US"          "India"        "Brazil"       "France"
## [5] "Russia"      "Spain"        "United Kingdom" "Italy"
## [9] "Argentina"   "Colombia"
```

```
rank.countries_confirmed <- global[order(global$Confirmed,decreasing= T),]
top.countries_confirmed <- head(rank.countries_confirmed, 15)
top.countries_confirmed
```

```
## # A tibble: 15 x 4
##   Region      Confirmed Recovered Deaths
##   <chr>      <dbl>    <dbl> <dbl>
## 1 US        12246849  4526513 256782
## 2 India     9139865  8562641 133738
## 3 Brazil    6071401  5457973 169183
## 4 France    2191180  157259  48807
## 5 Russia    2071858  1582768 35838
## 6 Spain     1556730  150376  42619
## 7 United Kingdom 1515802    3322  55120
## 8 Italy      1408868  553098  49823
## 9 Argentina  1370366  1195492 37002
## 10 Colombia  1248417  1150932 35287
## 11 Mexico    1041875  779104 101676
## 12 Peru      948081  879439  35549
## 13 Germany   932367  611627  14159
## 14 Poland    861331  423971  13618
## 15 Iran      854361  603445  44802
```

```
overall <- top.countries_confirmed %>%
  melt(id.vars = "Region",
       measure.vars = c("Confirmed"))
options(scipen=220)
p1 <- ggplot(overall) +
  geom_bar(aes(x = Region, y = value, fill = variable),
           stat = "identity", position="identity") +geom_text(aes(x = Region, y = value,label = value),
  theme(axis.text.x = element_text(angle = 45, hjust=0.5)) +
  labs(title = "Cases in Top 15 Countries", x = "Region", y = "") +
  guides(fill = guide_legend(title = NULL))
overall_deaths <- top.countries_confirmed%>%mutate(deaths.rate=(100 * Deaths/Confirmed) %>% round(1))
overall_deaths
```

```
## # A tibble: 15 x 5
##   Region      Confirmed Recovered Deaths deaths.rate
##   <chr>      <dbl>    <dbl> <dbl>    <dbl>
## 1 US        12246849  4526513 256782     2.1
## 2 India     9139865  8562641 133738     1.5
## 3 Brazil    6071401  5457973 169183     2.8
## 4 France    2191180  157259  48807     2.2
## 5 Russia    2071858  1582768 35838     1.7
```

```
## 6 Spain      1556730    150376  42619      2.7
## 7 United Kingdom 1515802      3322  55120      3.6
## 8 Italy      1408868    553098  49823      3.5
## 9 Argentina  1370366    1195492  37002      2.7
## 10 Colombia  1248417    1150932  35287      2.8
## 11 Mexico    1041875    779104  101676     9.8
## 12 Peru      948081    879439  35549      3.7
## 13 Germany   932367    611627  14159      1.5
## 14 Poland    861331    423971  13618      1.6
## 15 Iran      854361    603445  44802      5.2
```

```
overall_Deaths <- overall_deaths %>%
  melt(id.vars = "Region",
        measure.vars = c("deaths.rate"))
p2 <- ggplot(overall_Deaths) +
  geom_bar(aes(x = Region, y = value, fill = variable),
           stat = "identity", position="identity") + geom_text(aes(x = Region, y = value, label = value),
           theme(axis.text.x = element_text(angle = 45, hjust=0.5)) +
  labs(title = "Deaths rate(%) in Top 15 Countries", x = "Region", y = "") +
  guides(fill = guide_legend(title = NULL)) +
  scale_fill_discrete(labels = c("Deaths rate(%o)"))
new <- top.countries_confirmed%>%mutate(recover.rate=(100 * Recovered/Confirmed) %>% round(1))
new
```

```
## # A tibble: 15 x 5
##   Region      Confirmed Recovered Deaths recover.rate
##   <chr>          <dbl>     <dbl>   <dbl>         <dbl>
## 1 US            12246849   4526513 256782         37
## 2 India          9139865   8562641 133738        93.7
## 3 Brazil         6071401   5457973 169183        89.9
## 4 France         2191180   157259  48807         7.2
## 5 Russia         2071858   1582768  35838        76.4
## 6 Spain          1556730   150376  42619         9.7
## 7 United Kingdom 1515802      3322  55120         0.2
## 8 Italy          1408868   553098  49823        39.3
## 9 Argentina     1370366   1195492  37002        87.2
## 10 Colombia     1248417   1150932  35287        92.2
## 11 Mexico       1041875   779104  101676       74.8
## 12 Peru          948081   879439  35549        92.8
## 13 Germany       932367   611627  14159        65.6
## 14 Poland        861331   423971  13618        49.2
## 15 Iran          854361   603445  44802       70.6
```

```
new_overall <- new %>%
  melt(id.vars = "Region",
        measure.vars = c("recover.rate"))
new_overall
```

```
##           Region      variable value
## 1           US recover.rate  37.0
## 2          India recover.rate  93.7
## 3         Brazil recover.rate  89.9
## 4         France recover.rate   7.2
## 5         Russia recover.rate  76.4
## 6           Spain recover.rate   9.7
## 7 United Kingdom recover.rate   0.2
```

```

## 8          Italy recover.rate 39.3
## 9      Argentina recover.rate 87.2
## 10      Colombia recover.rate 92.2
## 11          Mexico recover.rate 74.8
## 12          Peru recover.rate 92.8
## 13      Germany recover.rate 65.6
## 14      Poland recover.rate 49.2
## 15          Iran recover.rate 70.6

p3 <- ggplot(new_overall) +
  geom_bar(aes(x = Region, y = value , fill=variable),
    stat = "identity", position="identity") +geom_text(aes(x = Region, y = value,label = value),
    theme(axis.text.x = element_text(angle = 45, hjust=0.5)) +
  labs(title = "Recover rate(%) in Top 15 Countries", x = "Region", y = "") +
  guides(fill = guide_legend(title = NULL)) +
  scale_fill_discrete(labels = c("Recover rate(%)"))
multiplot <- function(..., plotlist = NULL, file, cols = 1, layout = NULL) {
  require(grid)

  plots <- c(list(...), plotlist)

  numPlots = length(plots)

  if (is.null(layout)) {
    layout <- matrix(seq(1, cols * ceiling(numPlots/cols)),
      ncol = cols, nrow = ceiling(numPlots/cols))
  }

  if (numPlots == 1) {
    print(plots[[1]])
  } else {
    grid.newpage()
    pushViewport(viewport(layout = grid.layout(nrow(layout), ncol(layout))))

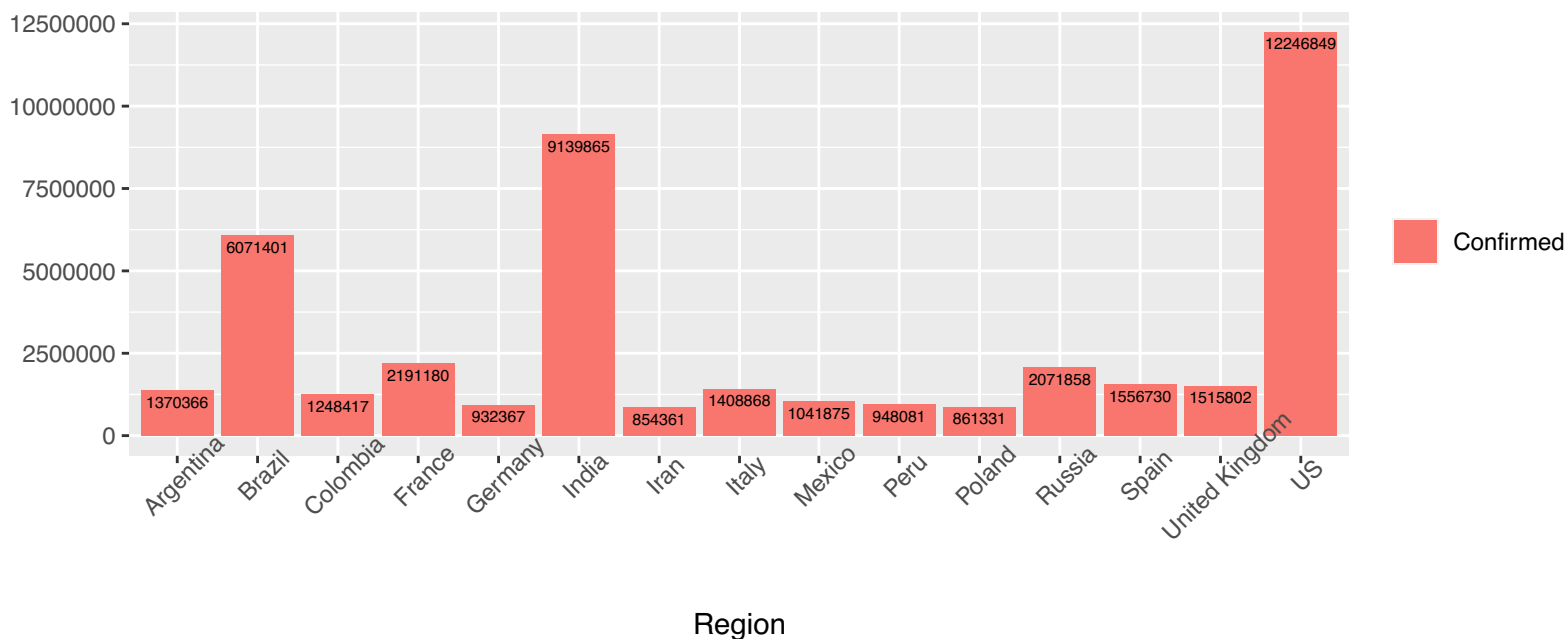
    for (i in 1:numPlots) {
      matchidx <- as.data.frame(which(layout == i, arr.ind = TRUE))

      print(plots[[i]], vp = viewport(layout.pos.row = matchidx$row,
        layout.pos.col = matchidx$col))
    }
  }
}
multiplot(p1, p2,p3, cols = 1)

## Loading required package: grid

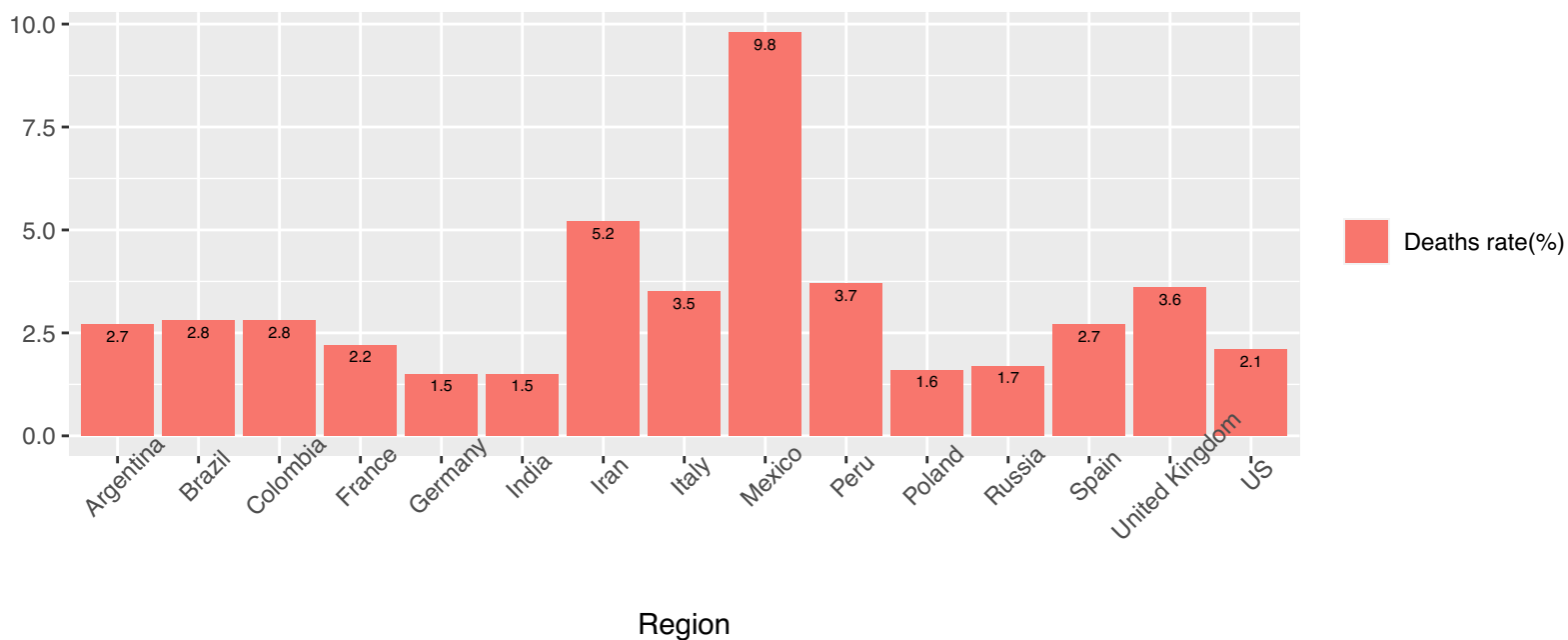
```

### Cases in Top 15 Countries



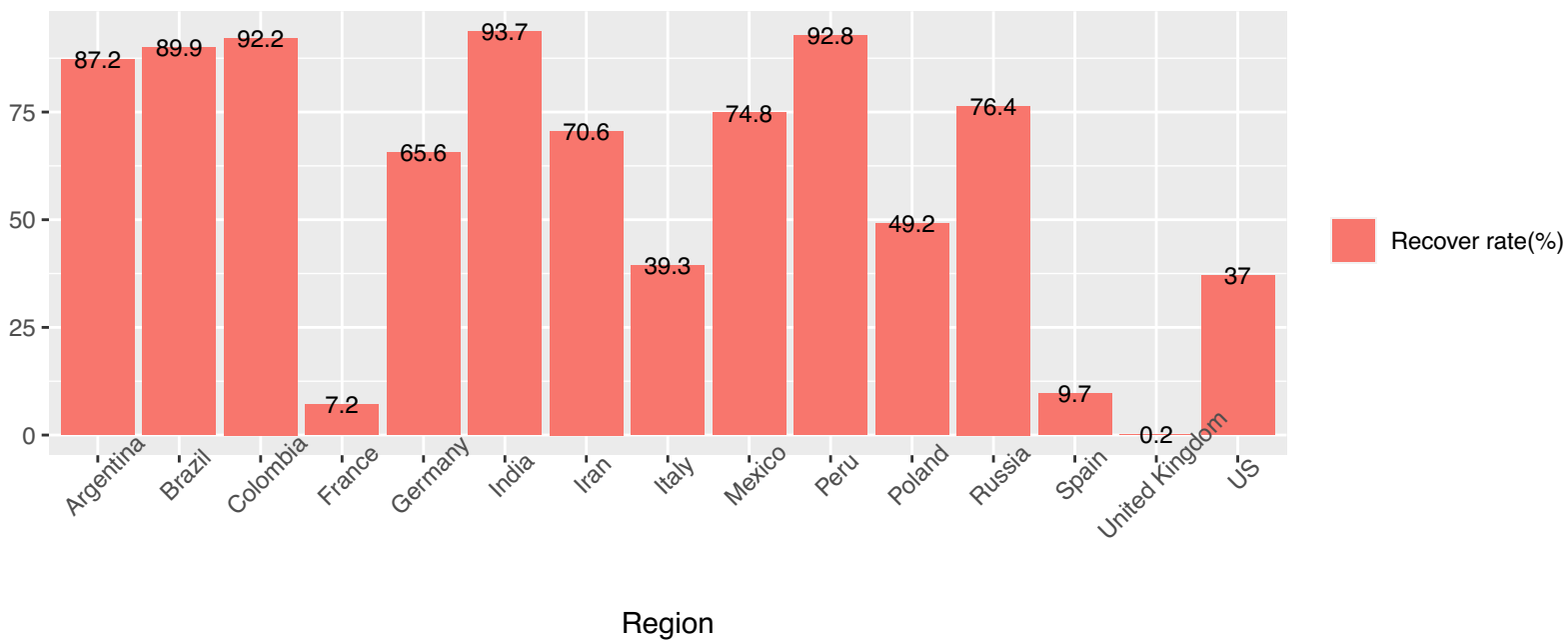
Region

### Deaths rate(%) in Top 15 Countries



Region

### Recover rate(%) in Top 15 Countries



Region