

```

---
title: "COVID-19 EDA"
output:
pdf_document: default
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse)
library(lubridate)
library(plotly)
library(flexdashboard)

covid <- read_csv(
 "data/covid_19_data.csv",
 col_types = cols(
 ObservationDate = col_date(format="%m/%d/%Y")
)) %>%
group_by(`Country/Region`, ObservationDate) %>%
summarize(
 cases = sum(Confirmed),
 deaths = sum(Deaths),
 recovered = sum(Recovered)
) %>%
mutate(
 first_occurrence = first(ObservationDate),
 days_in_country = ObservationDate - first_occurrence
) %>%
rename("country" = `Country/Region`, "date" = ObservationDate)
```

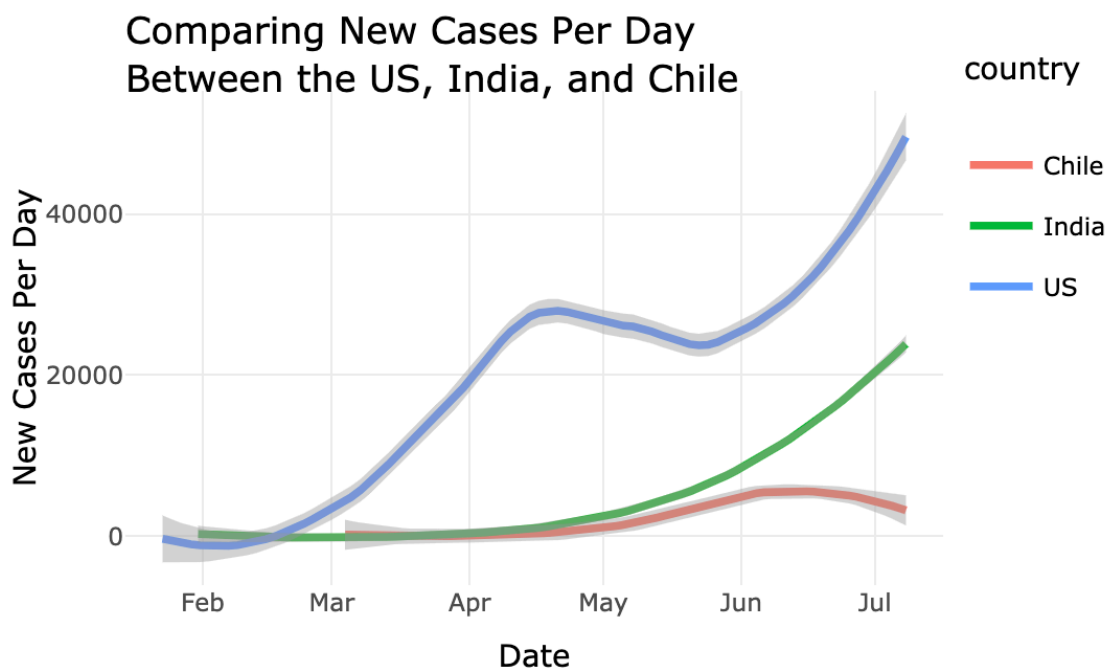
```

```

# Graph 1 - Compares the new cases per day between the US, India, and
Chile
```{r}
graph_1 <- covid %>%
 mutate(new_cases = cases - lag(cases)) %>%
 filter(country == "US" | country == "India" | country == "Chile")
%>%
 ggplot() +
 geom_smooth(mapping = aes(x = date, y = new_cases, color = country))
+
 labs(title = "Comparing New Cases Per Day \nBetween the US, India,
and Chile", x = "Date", y = "New Cases Per Day") +
 theme_minimal()

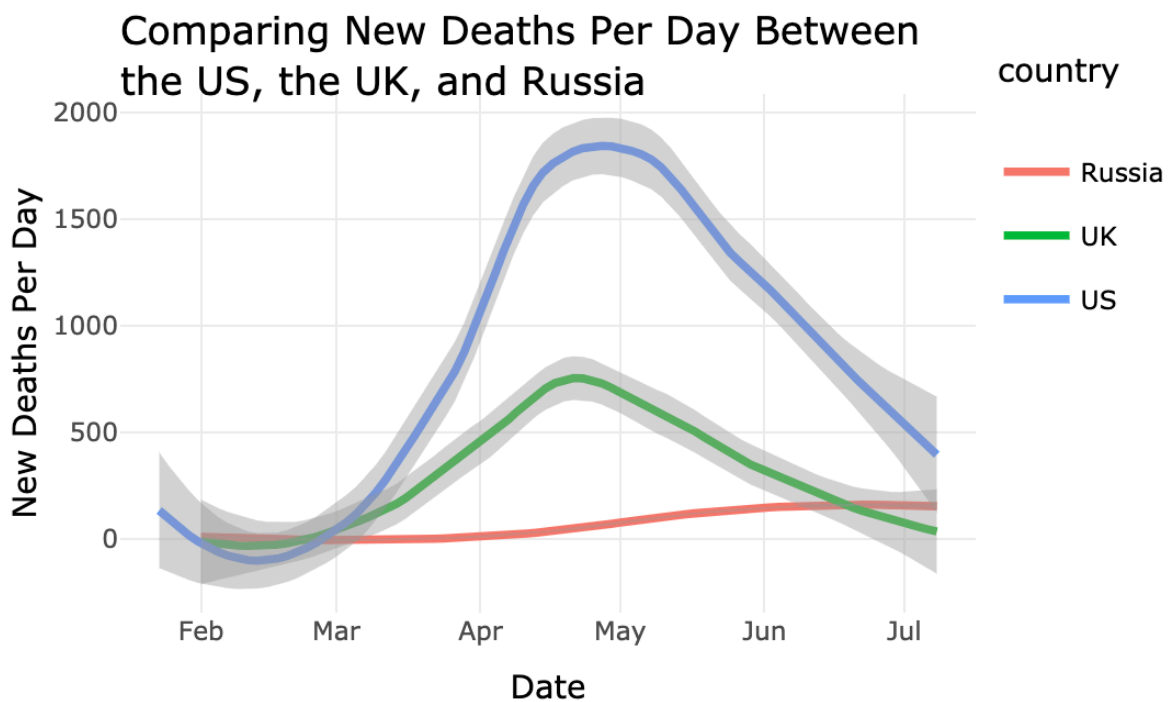
ggplotly(graph_1)
```

```



```
# Graph 2 - Compares the new deaths per day between the US, the UK,
and Russia
```

```
```{r}
save the ggplot graph to a variable:
graph_2 <- covid %>%
 mutate(new_deaths = deaths - lag(deaths)) %>%
 filter(country == "US" | country == "UK" | country == "Russia") %>%
 ggplot() +
 geom_smooth(mapping = aes(x = date, y = new_deaths, color =
country)) +
 labs(title = "Comparing New Deaths Per Day Between \nthe US, the
UK, and Russia", x = "Date", y = "New Deaths Per Day") +
 theme_minimal()
now convert to plotly
ggplotly(graph_2)
```

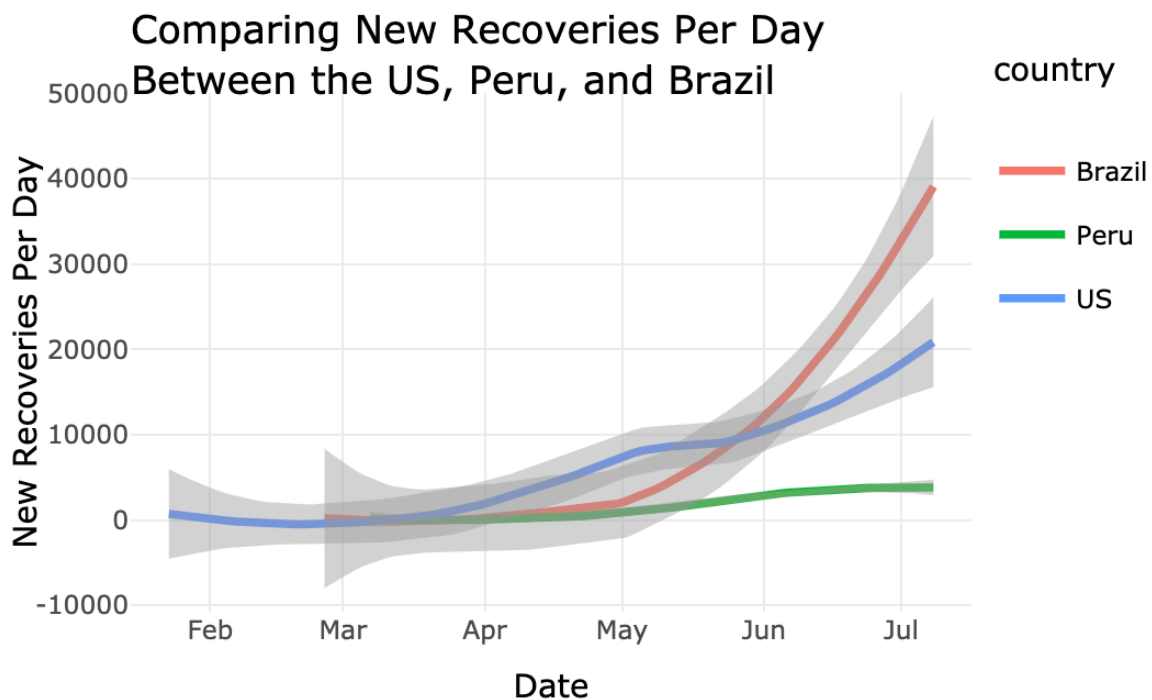


```
```
```

```
# Graph 3 - Compares the new recoveries per day between the US,  
Brazil, and Peru.
```

```
```{r}  
graph_3 <- covid %>%
 mutate(new_recoveries = recovered - lag(recovered)) %>%
 filter(country == "US" | country == "Brazil" | country == "Peru") %>%
 ggplot()+
 geom_smooth(mapping = aes(x = date, y = new_recoveries, color =
country)) +
 labs(title = "Comparing New Recoveries Per Day \nBetween the US,
Peru, and Brazil", x = "Date", y = "New Recoveries Per Day")+
 theme_minimal()

ggplotly(graph_3)
```
```



```

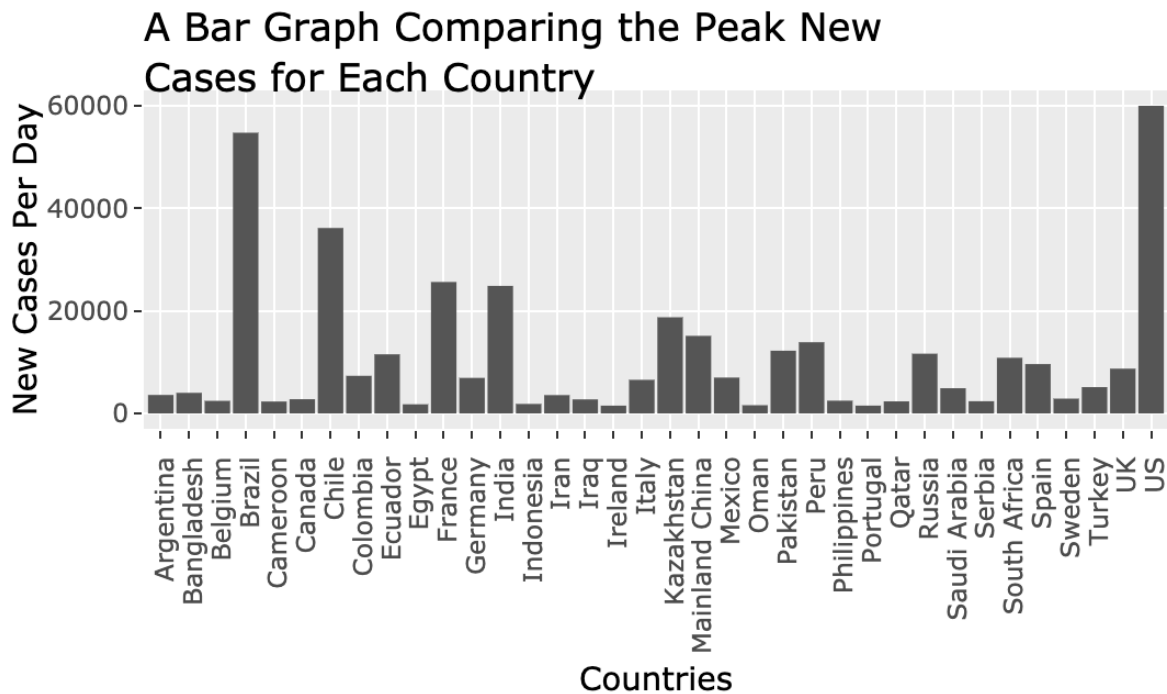
# Graph 4 - Plotting the peak cases for each country
``` {r}
find_new_cases <- covid %>%
 mutate(new_cases = cases - lag(cases))

group_countries_and_find_peak <- find_new_cases %>%
 group_by(country) %>%
 summarize (
 peak_cases = max(new_cases, na.rm=TRUE)
) %>%
 filter(peak_cases > 1500)

group_countries_and_find_peak

ggplotly(ggplot(data = group_countries_and_find_peak)+
 geom_col(aes(x = country, y = peak_cases))+
 labs(title = "A Bar Graph Comparing the Peak New \nCases for Each
Country", x = "Countries", y = "New Cases Per Day") +
 theme(axis.text.x = element_text(angle = 90, vjust = 0.5,
hjust=1)))
```

```



```

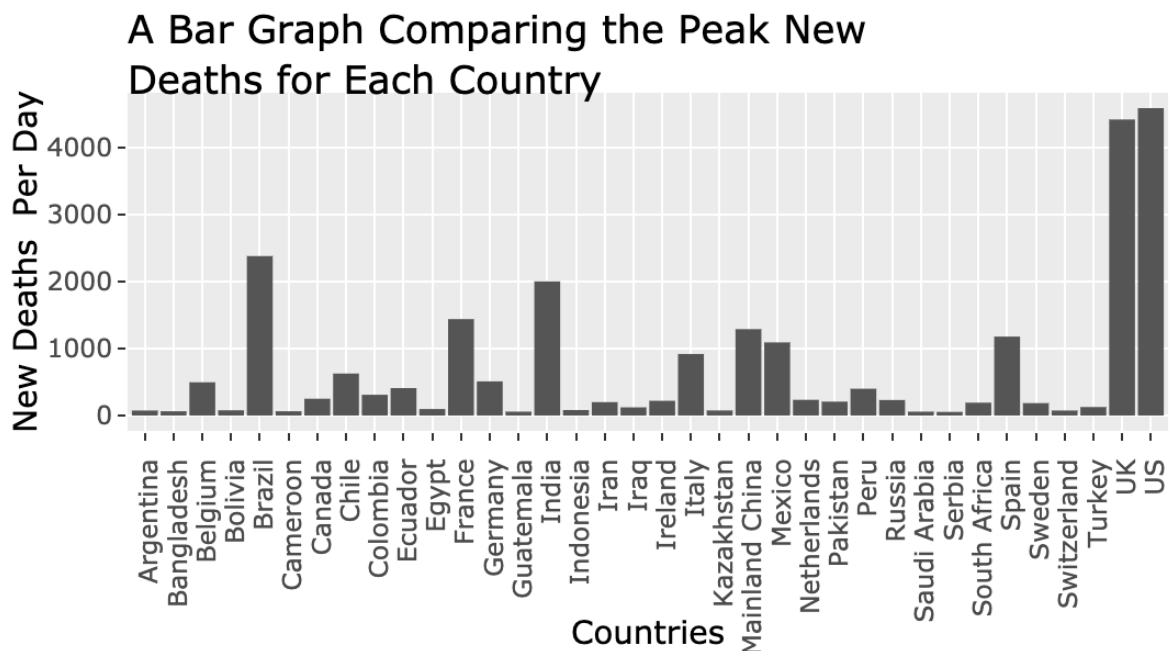
# Graph 5 - Plotting the peak deaths for each country
``` {r}
find_new_deaths <- covid %>%
 mutate(new_deaths = deaths - lag(deaths))

group_countries_and_find_peak <- find_new_deaths %>%
 group_by(country) %>%
 summarize (
 peak_deaths = max(new_deaths, na.rm=TRUE)
) %>%
 filter(peak_deaths > 50)

group_countries_and_find_peak

ggplotly(ggplot(data = group_countries_and_find_peak)+
 geom_col(aes(x = country, y = peak_deaths))+
 labs(title = "A Bar Graph Comparing the Peak New \nDeaths for Each
Country", x = "Countries", y = "New Deaths Per Day") +
 theme(axis.text.x = element_text(angle = 90, vjust = 0.5,
hjust=1)))
```

```



```

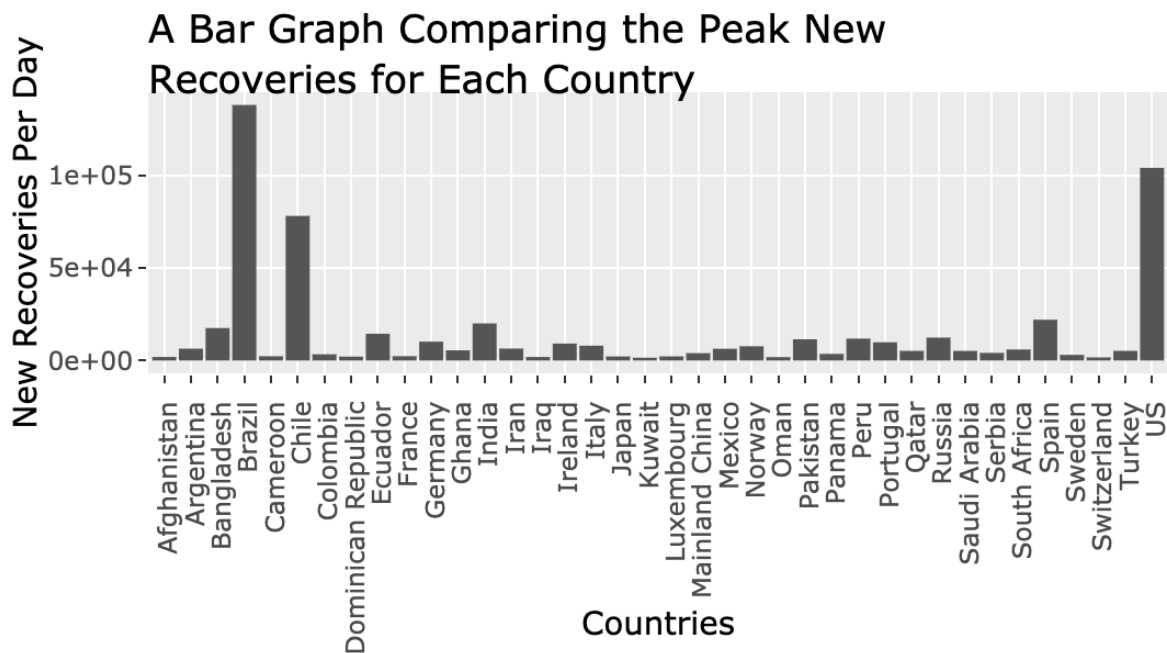
# Graph 6 - Plotting the peak recoveries for each country
``` {r}
find_new_recoveries <- covid %>%
 mutate(new_recoveries = recovered - lag(recovered))

group_countries_and_find_peak <- find_new_recoveries %>%
 group_by(country) %>%
 summarize (
 peak_recoveries = max(new_recoveries, na.rm=TRUE)
) %>%
 filter(peak_recoveries > 1500)

group_countries_and_find_peak

ggplotly(ggplot(data = group_countries_and_find_peak)+
 geom_col(aes(x = country, y = peak_recoveries))+
 labs(title = "A Bar Graph Comparing the Peak New \nRecoveries for
Each Country", x = "Countries", y = "New Recoveries Per Day") +
 theme(axis.text.x = element_text(angle = 90, vjust = 0.5,
hjust=1)))
```

```



Graph 7 - Comparing the New Cases of Countries on the Date of the US's Peak of New Cases

```
```{r}

find_us_peak_cases <- covid %>%
 mutate(new_cases = cases - lag(cases)) %>%
 select(-cases:-days_in_country) %>%
 spread(key = country, value = new_cases) %>%
 arrange(desc(US)) %>%
 head(1) %>%
 gather(`('St. Martin',)`:Zimbabwe, key = "country", value =
"new_cases") %>%
 filter(new_cases > 300) %>%
 ggplot()+
 geom_col(aes(x = country, y = new_cases))+
 labs(title = "A Bar Graph Comparing the New Cases of Countries on
the Date of the US's Peak of New Cases", x = "Countries", y = "New
Cases on July 7th, 2020") +
 theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))

ggplotly(find_us_peak_cases)
```
```

