Homework 6

Ishita Dutta

11/11/2021

#Roadmap ##Current Progress: Question selected, looking over given data set to gain useful variables, finding dataset on daily vaccination doses given.

##Question: How does the administration of vaccines affect new COVID-19 Case counts over time?

##Reasoning for question: We can show through this question that the vaccine is a really affecive preventive measure against the spread of COVID-19, also prompting people who are against getting the vaccine to go get the vaccine.
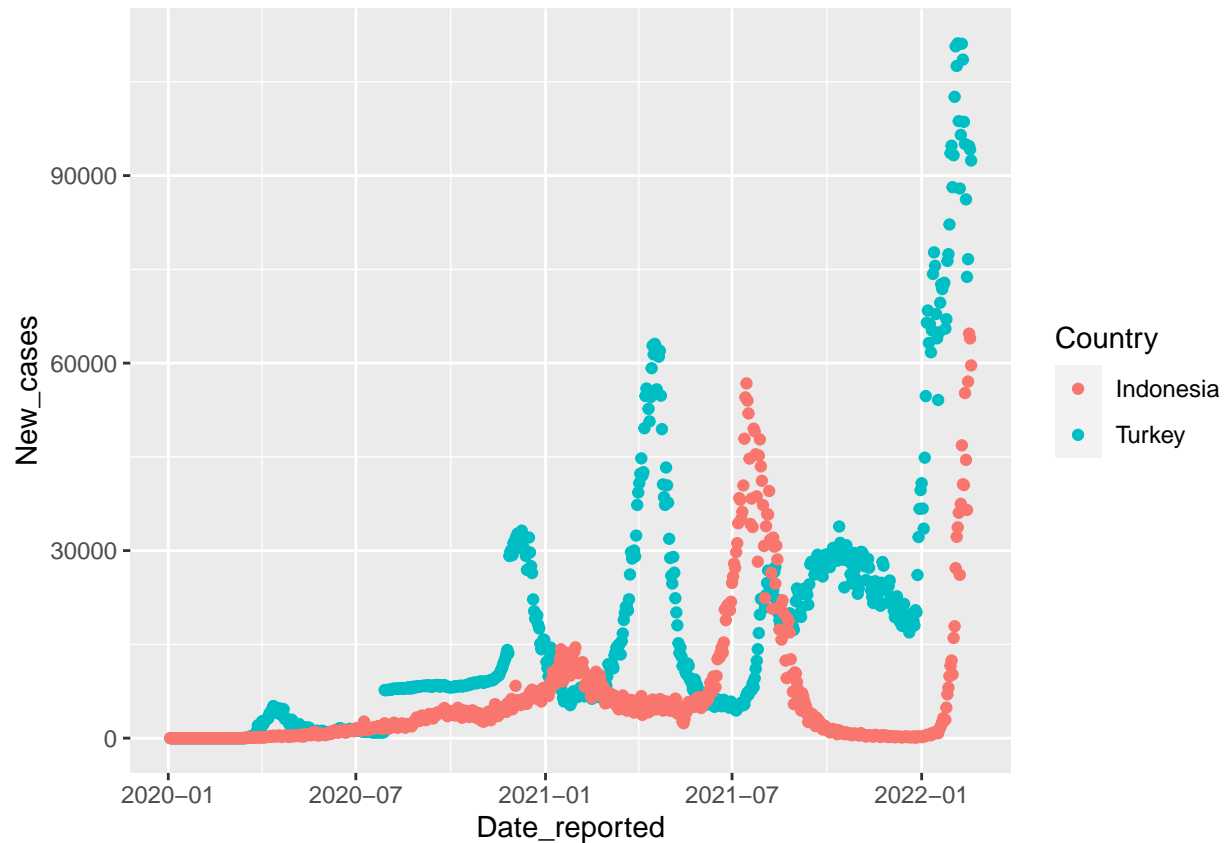
##Given Data and Progress:

The given data had these variables:

Out of these, Cumulative cases and deaths are not as important to my question as the new cases and deaths. These two variables will not be counted in my data:

```
colnames(covid)
```

```
## [1] "Date_reported" "New_cases"     "New_deaths"     "Country"
## [5] "Country_code"
```

Now, to work with the cleaner data more suited to my question. First, checking what the country data looks on its own in this graph comparing cases reported in Turkey and Indonesia:

This was just to check that the data has been subsetted correctly and similar methods can be used to get the data for other countries.

On this note, I also checked what other countries there are by subsetting on a random given date, October 12 2021:

```
#include = FALSE, echo = FALSE
covid_oct122021 = subset(covid, Date_reported == "2021-10-12")
covid_oct122021
```

```
## # A tibble: 235 x 5
##    Date_reported New_cases New_deaths Country            Country_code
##    <date>            <dbl>      <dbl> <chr>              <chr>
##  1 2021-10-12           31          2 Afghanistan        AF
##  2 2021-10-12          195          9 Albania            AL
##  3 2021-10-12           98          2 Algeria            DZ
##  4 2021-10-12            0          0 American Samoa     AS
##  5 2021-10-12           16          0 Andorra            AD
##  6 2021-10-12          202          7 Angola             AO
##  7 2021-10-12            0          0 Anguilla           AI
##  8 2021-10-12            0          0 Antigua and Barbuda AG
##  9 2021-10-12            0          0 Argentina          AR
## 10 2021-10-12          903         43 Armenia            AM
## # ... with 225 more rows
```

This gives me a list of countries so that I can easily sort by countries that I want to compare for a sample.

##Supplemental Data This data is like the given data but only measures the number vaccines administered per day. We can join this into the WHO data for one combined dataset but I need to check this before combining to make sure there is a proper overlap to combine.

```
##           Entity Code        Day new_vaccinations_smoothed_per_million
## 350 Afghanistan  AFG 2022-02-07                                 0.0329
## 351 Afghanistan  AFG 2022-02-08                                 0.0334
## 352 Afghanistan  AFG 2022-02-09                                 0.0339
## 353 Afghanistan  AFG 2022-02-10                                 0.0346
## 354 Afghanistan  AFG 2022-02-11                                 0.0353
## 355 Afghanistan  AFG 2022-02-12                                 0.0360
```

##Timeline: This is all the work I have done so far, so below is a timeline for what I will do:

```
##   duedates                                                    checks
## 1   Feb 21   clean all data and put all tables together for analysis
## 2   Feb 24             first iteration of analysis process complete
## 3   Mar 10                  Analysis iteration process complete
## 4   Mar 13                  Formatting and report writing complete
## 5   Mar 14 Final report submission MANY HOURS before the deadline :)
```

##Abstract

##Introduction

##Background

##Descriptive Analysis

##Inferential Analysis

##Sensitivity Analysis

##Discussion

##Appendix ###Acknowledgements

###Reference

###Code

###Session Info