Covid 19 Vaccine Analysis

18F0330 Haleema Sadia

12/20/2021

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
library(ggplot2)
library(scales)
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.1.2
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1
## v purrr   0.3.4
```

```
## Warning: package 'tidyr' was built under R version 4.1.2
```

```
## Warning: package 'readr' was built under R version 4.1.2
```

```
## Warning: package 'purrr' was built under R version 4.1.2
```

```
## Warning: package 'dplyr' was built under R version 4.1.2
```

```
## Warning: package 'forcats' was built under R version 4.1.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x readr::col_factor() masks scales::col_factor()
## x purrr::discard()    masks scales::discard()
## x dplyr::filter()     masks stats::filter()
## x dplyr::lag()        masks stats::lag()
```

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 4.1.2
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

```
## The following object is masked from 'package:purrr':
##
##     transpose
```

```
library(stringr)
library(summarytools)
```

```
## Warning: package 'summarytools' was built under R version 4.1.2
```

```
##
## Attaching package: 'summarytools'
```

```
## The following object is masked from 'package:tibble':
##
##     view
```

```
vaccine_data<-fread(file = "country_vaccinations.csv")
head(vaccine_data)
```

```
##        country iso_code       date total_vaccinations people_vaccinated
## 1: Afghanistan      AFG 2021-02-22                  0                 0
## 2: Afghanistan      AFG 2021-02-23                 NA                NA
## 3: Afghanistan      AFG 2021-02-24                 NA                NA
## 4: Afghanistan      AFG 2021-02-25                 NA                NA
## 5: Afghanistan      AFG 2021-02-26                 NA                NA
## 6: Afghanistan      AFG 2021-02-27                 NA                NA
##    people_fully_vaccinated daily_vaccinations_raw daily_vaccinations
## 1:                      NA                     NA                 NA
## 2:                      NA                     NA               1367
## 3:                      NA                     NA               1367
## 4:                      NA                     NA               1367
## 5:                      NA                     NA               1367
## 6:                      NA                     NA               1367
##    total_vaccinations_per_hundred people_vaccinated_per_hundred
## 1:                              0                             0
## 2:                             NA                            NA
## 3:                             NA                            NA
## 4:                             NA                            NA
## 5:                             NA                            NA
```

```
## 6:                              NA                              NA
##    people_fully_vaccinated_per_hundred daily_vaccinations_per_million
## 1:                                  NA                              NA
## 2:                                  NA                              35
## 3:                                  NA                              35
## 4:                                  NA                              35
## 5:                                  NA                              35
## 6:                                  NA                              35
##                                                                vaccines
## 1: Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing
## 2: Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing
## 3: Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing
## 4: Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing
## 5: Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing
## 6: Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, Sinopharm/Beijing
##                  source_name           source_website
## 1: World Health Organization https://covid19.who.int/
## 2: World Health Organization https://covid19.who.int/
## 3: World Health Organization https://covid19.who.int/
## 4: World Health Organization https://covid19.who.int/
## 5: World Health Organization https://covid19.who.int/
## 6: World Health Organization https://covid19.who.int/
```

```
str(vaccine_data)
```

```
## Classes 'data.table' and 'data.frame':   31240 obs. of  15 variables:
##  $ country                             : chr  "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan"
##  $ iso_code                            : chr  "AFG" "AFG" "AFG" "AFG" ...
##  $ date                                : IDate, format: "2021-02-22" "2021-02-23" ...
##  $ total_vaccinations                  : num  0 NA NA NA NA NA 8200 NA NA NA ...
##  $ people_vaccinated                   : num  0 NA NA NA NA NA 8200 NA NA NA ...
##  $ people_fully_vaccinated             : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ daily_vaccinations_raw              : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ daily_vaccinations                  : num  NA 1367 1367 1367 1367 ...
##  $ total_vaccinations_per_hundred      : num  0 NA NA NA NA NA 0.02 NA NA NA ...
##  $ people_vaccinated_per_hundred       : num  0 NA NA NA NA NA 0.02 NA NA NA ...
##  $ people_fully_vaccinated_per_hundred: num  NA NA NA NA NA NA NA NA NA NA ...
##  $ daily_vaccinations_per_million      : num  NA 35 35 35 35 35 35 41 46 52 ...
##  $ vaccines                            : chr  "Johnson&Johnson, Oxford/AstraZeneca, Pfizer/BioNTech, S
##  $ source_name                         : chr  "World Health Organization" "World Health Organization"
##  $ source_website                      : chr  "https://covid19.who.int/" "https://covid19.who.int/" "h
##  - attr(*, ".internal.selfref")=<externalptr>
```

```
dim (vaccine_data)
```

```
## [1] 31240    15
```

```
vaccine_data_copy <- vaccine_data[,1:13]
colnames(vaccine_data_copy)
```

```
##  [1] "country"                             "iso_code"
##  [3] "date"                                "total_vaccinations"
```

```
##  [5] "people_vaccinated"                   "people_fully_vaccinated"
##  [7] "daily_vaccinations_raw"              "daily_vaccinations"
##  [9] "total_vaccinations_per_hundred"      "people_vaccinated_per_hundred"
## [11] "people_fully_vaccinated_per_hundred" "daily_vaccinations_per_million"
## [13] "vaccines"
```

Including Plots

You can also embed plots, for example:

```
data.frame("Total_NA" = colSums(is.na(vaccine_data_copy))) %>%
    mutate ("Percentage_of_NA" = (colSums(is.na(vaccine_data_copy))/dim(vaccine_data_copy)[1]) %>%
             round (3) * 100)
```

```
##                                     Total_NA Percentage_of_NA
## country                                    0              0.0
## iso_code                                   0              0.0
## date                                       0              0.0
## total_vaccinations                     13789             44.1
## people_vaccinated                      14686             47.0
## people_fully_vaccinated                17445             55.8
## daily_vaccinations_raw                 16819             53.8
## daily_vaccinations                       292              0.9
## total_vaccinations_per_hundred         13789             44.1
## people_vaccinated_per_hundred          14686             47.0
## people_fully_vaccinated_per_hundred    17445             55.8
## daily_vaccinations_per_million           292              0.9
## vaccines                                   0              0.0
```

```
vaccine_data_copy[is.na(vaccine_data_copy)] = 0

remove_countries = c('England','Northern Ireland','Scotland','Wales','Falkland Islands','Faeroe Islands

vaccine_data_copy <- vaccine_data_copy %>%
  filter (!country %in% remove_countries)

unique(vaccine_data_copy$country)
```

```
##   [1] "Afghanistan"                     "Albania"
##   [3] "Algeria"                         "Andorra"
##   [5] "Angola"                          "Anguilla"
##   [7] "Antigua and Barbuda"             "Argentina"
##   [9] "Armenia"                         "Aruba"
##  [11] "Australia"                       "Austria"
##  [13] "Azerbaijan"                      "Bahamas"
##  [15] "Bahrain"                         "Bangladesh"
##  [17] "Barbados"                        "Belarus"
##  [19] "Belgium"                         "Belize"
##  [21] "Benin"                           "Bermuda"
##  [23] "Bhutan"                          "Bolivia"
##  [25] "Bonaire Sint Eustatius and Saba" "Bosnia and Herzegovina"
##  [27] "Botswana"                        "Brazil"
```

```
##  [29] "British Virgin Islands"          "Brunei"
##  [31] "Bulgaria"                        "Burkina Faso"
##  [33] "Cambodia"                        "Cameroon"
##  [35] "Canada"                          "Cape Verde"
##  [37] "Central African Republic"        "Chad"
##  [39] "Chile"                           "China"
##  [41] "Colombia"                        "Comoros"
##  [43] "Congo"                           "Cook Islands"
##  [45] "Costa Rica"                      "Cote d'Ivoire"
##  [47] "Croatia"                         "Cuba"
##  [49] "Curacao"                         "Cyprus"
##  [51] "Czechia"                         "Democratic Republic of Congo"
##  [53] "Denmark"                         "Djibouti"
##  [55] "Dominica"                        "Dominican Republic"
##  [57] "Ecuador"                         "Egypt"
##  [59] "El Salvador"                     "Equatorial Guinea"
##  [61] "Estonia"                         "Eswatini"
##  [63] "Ethiopia"                        "Fiji"
##  [65] "Finland"                         "France"
##  [67] "French Polynesia"                "Gabon"
##  [69] "Gambia"                          "Georgia"
##  [71] "Germany"                         "Ghana"
##  [73] "Gibraltar"                       "Greece"
##  [75] "Greenland"                       "Grenada"
##  [77] "Guatemala"                       "Guernsey"
##  [79] "Guinea"                          "Guinea-Bissau"
##  [81] "Guyana"                          "Honduras"
##  [83] "Hong Kong"                       "Hungary"
##  [85] "Iceland"                         "India"
##  [87] "Indonesia"                       "Iran"
##  [89] "Iraq"                            "Ireland"
##  [91] "Israel"                          "Italy"
##  [93] "Jamaica"                         "Japan"
##  [95] "Jersey"                          "Jordan"
##  [97] "Kazakhstan"                      "Kenya"
##  [99] "Kosovo"                          "Kuwait"
## [101] "Kyrgyzstan"                      "Laos"
## [103] "Latvia"                          "Lebanon"
## [105] "Lesotho"                         "Liberia"
## [107] "Libya"                           "Liechtenstein"
## [109] "Lithuania"                       "Luxembourg"
## [111] "Macao"                           "Madagascar"
## [113] "Malawi"                          "Malaysia"
## [115] "Maldives"                        "Mali"
## [117] "Malta"                           "Mauritania"
## [119] "Mauritius"                       "Mexico"
## [121] "Moldova"                         "Monaco"
## [123] "Mongolia"                        "Montenegro"
## [125] "Montserrat"                      "Morocco"
## [127] "Mozambique"                      "Myanmar"
## [129] "Namibia"                         "Nauru"
## [131] "Nepal"                           "Netherlands"
## [133] "New Caledonia"                   "New Zealand"
## [135] "Nicaragua"                       "Niger"
```

```
## [137] "Nigeria"                          "Niue"
## [139] "North Macedonia"                  "Northern Cyprus"
## [141] "Norway"                           "Oman"
## [143] "Pakistan"                         "Palestine"
## [145] "Panama"                           "Papua New Guinea"
## [147] "Paraguay"                         "Peru"
## [149] "Philippines"                      "Pitcairn"
## [151] "Poland"                           "Portugal"
## [153] "Qatar"                            "Romania"
## [155] "Russia"                           "Rwanda"
## [157] "Samoa"                            "San Marino"
## [159] "Sao Tome and Principe"            "Saudi Arabia"
## [161] "Senegal"                          "Serbia"
## [163] "Seychelles"                       "Sierra Leone"
## [165] "Singapore"                        "Sint Maarten (Dutch part)"
## [167] "Slovakia"                         "Slovenia"
## [169] "Solomon Islands"                  "Somalia"
## [171] "South Africa"                     "South Korea"
## [173] "South Sudan"                      "Spain"
## [175] "Sri Lanka"                        "Sudan"
## [177] "Suriname"                         "Sweden"
## [179] "Switzerland"                      "Syria"
## [181] "Taiwan"                           "Tajikistan"
## [183] "Thailand"                         "Timor"
## [185] "Togo"                             "Tonga"
## [187] "Trinidad and Tobago"              "Tunisia"
## [189] "Turkey"                           "Turkmenistan"
## [191] "Turks and Caicos Islands"         "Tuvalu"
## [193] "Uganda"                           "Ukraine"
## [195] "United Arab Emirates"             "United Kingdom"
## [197] "United States"                    "Uruguay"
## [199] "Uzbekistan"                       "Vanuatu"
## [201] "Venezuela"                        "Vietnam"
## [203] "Wallis and Futuna"                "Yemen"
## [205] "Zambia"                           "Zimbabwe"
```

```
vaccine_data_copy$vaccines <- str_replace_all(vaccine_data_copy$vaccines, " ","")
    # remove all spaces in between
    vaccine_val<- unique(vaccine_data_copy$vaccines)
    vaccine<- vector()

    for (i in vaccine_val){
        for (j in strsplit(i, ",")){
            vaccine<- c(vaccine, j)
        }
    }
    vaccine_used<- unique(vaccine)
    vaccine_used
```

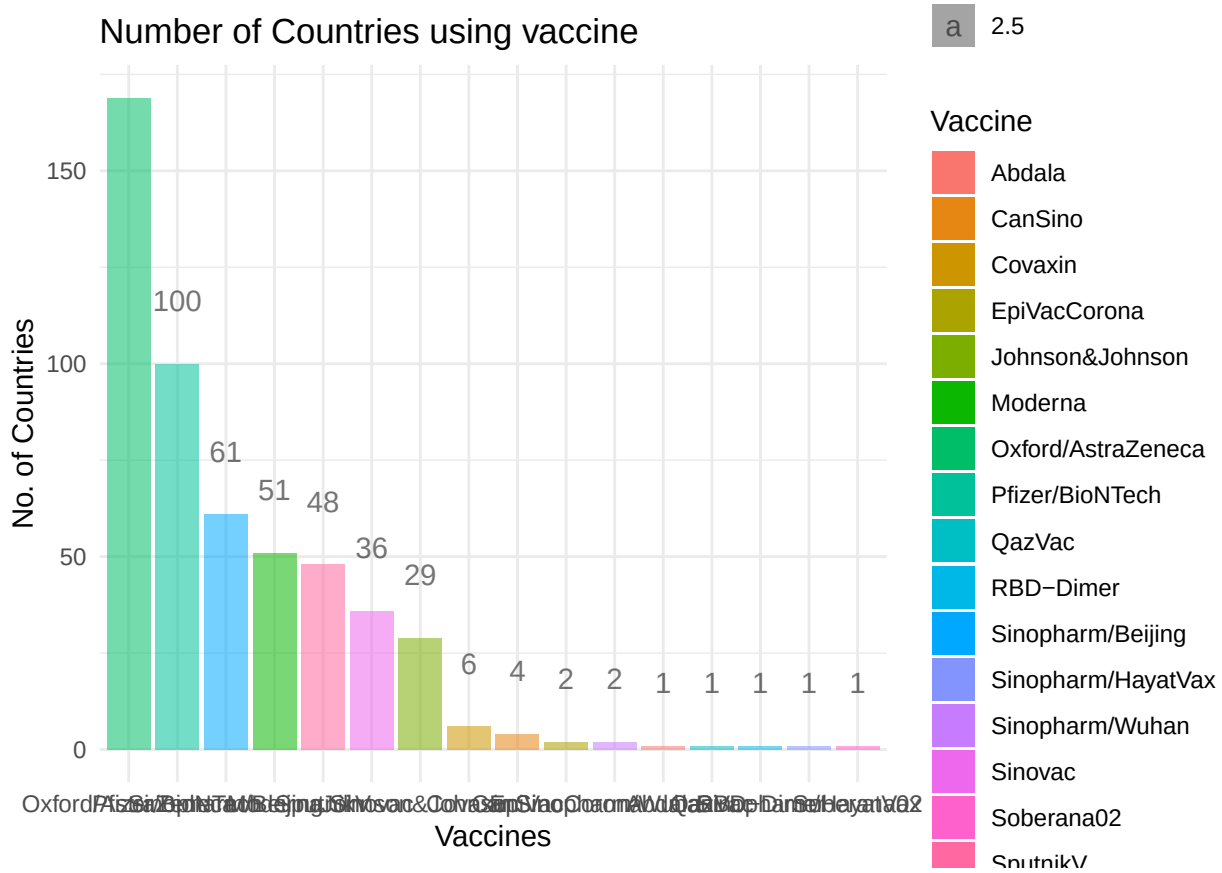
```
##  [1] "Johnson&Johnson"    "Oxford/AstraZeneca" "Pfizer/BioNTech"
##  [4] "Sinopharm/Beijing"  "Sinovac"            "SputnikV"
##  [7] "Moderna"            "Covaxin"            "CanSino"
## [10] "Sinopharm/Wuhan"    "Abdala"             "Soberana02"
## [13] "QazVac"             "Sinopharm/HayatVax" "EpiVacCorona"
```

```
## [16] "RBD-Dimer"
```

```
    vaccine_data_val <- data.frame(matrix(ncol = length(vaccine_used), nrow = 0))
for (i in vaccine_data_copy$vaccines){
  vaccine_data_val<- rbind(vaccine_data_val, Vectorize(grepl, USE.NAMES = TRUE)(vaccine_used, str_repla
}
vaccine_data_val[vaccine_data_val == TRUE] = 1
vaccine_data_val[vaccine_data_val == FALSE] =0
colnames(vaccine_data_val) <- paste0(unique(vaccine))
```

```
vaccine_in_countries<- vaccine_data_val %>%
mutate(country = vaccine_data_copy$country)%>%
group_by(country)%>%
summarise_all(sum)

data <- data.frame("No_of_countries"= apply(vaccine_in_countries[-1],2, function(c)sum(c!=0)))
cbind("Vaccine"=row.names(data),data) %>%
ggplot(mapping=aes(x=reorder(Vaccine, -No_of_countries), y=No_of_countries, fill = Vaccine, alpha=2.5))+
geom_col() +
labs(x = "Vaccines", y = "No. of Countries", title  = "Number of Countries using vaccine")+
geom_text(aes(label = No_of_countries), vjust=-2.5)+
theme_minimal()
```



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.