Covid19VaccineAnalysisInR

18F-0114_Muhammad Zain Ul Abidin

12/27/2021

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(ggthemes)
```

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

```
library(tidyverse)
```

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

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

```
## v ggplot2 3.3.5      v purrr   0.3.4
## 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
```

```
## 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 dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(colorspace)
library(scales)
```

```
##
## Attaching package: 'scales'
```

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

```
## The following object is masked from 'package:readr':
##
##     col_factor
```

```
library(ggplot2)
library(dplyr)
library(DT)
```

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

```
library(tidyr)
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
```

```
library(lubridate)
```

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

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

```
## The following objects are masked from 'package:data.table':
##
##      hour, isoweek, mday, minute, month, quarter, second, wday, week,
##      yday, year
```

```
## The following objects are masked from 'package:base':
##
##      date, intersect, setdiff, union
```

Including Plots

You can also embed plots, for example:

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

```
##          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/
```

```r
str(info_vac)
```

```
## 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 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>
```

```r
dim (info_vac)
```

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

```r
info_temp_vac <- info_vac[,1:13]
colnames(info_temp_vac)
```

```
##  [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"
```

```
data.frame("Total_NA" = colSums(is.na(info_temp_vac))) %>%
    mutate ("Percentage_of_NA" = (colSums(is.na(info_temp_vac))/dim(info_temp_vac)[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
```

```
info_temp_vac[is.na(info_temp_vac)] = 0
neglect = c('Faeroe Islands','Saint Lucia','Saint Vincent and the Grenadines','England','Northern Irela
info_temp_vac <- info_temp_vac %>%
  filter (!country %in% neglect)
unique(info_temp_vac$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] "Cayman Islands"                 "Central African Republic"
##  [39] "Chad"                           "Chile"
##  [41] "China"                          "Colombia"
##  [43] "Comoros"                        "Congo"
##  [45] "Cook Islands"                   "Costa Rica"
##  [47] "Cote d'Ivoire"                  "Croatia"
##  [49] "Cuba"                           "Curacao"
##  [51] "Cyprus"                         "Czechia"
##  [53] "Democratic Republic of Congo"   "Denmark"
##  [55] "Djibouti"                       "Dominica"
```

```
##  [57] "Dominican Republic"        "Ecuador"
##  [59] "Egypt"                      "El Salvador"
##  [61] "Equatorial Guinea"          "Estonia"
##  [63] "Eswatini"                   "Ethiopia"
##  [65] "Fiji"                       "Finland"
##  [67] "France"                     "French Polynesia"
##  [69] "Gabon"                      "Gambia"
##  [71] "Georgia"                    "Germany"
##  [73] "Ghana"                      "Gibraltar"
##  [75] "Greece"                     "Greenland"
##  [77] "Grenada"                    "Guatemala"
##  [79] "Guernsey"                   "Guinea"
##  [81] "Guinea-Bissau"              "Guyana"
##  [83] "Honduras"                   "Hong Kong"
##  [85] "Hungary"                    "Iceland"
##  [87] "India"                      "Indonesia"
##  [89] "Iran"                       "Iraq"
##  [91] "Ireland"                    "Isle of Man"
##  [93] "Israel"                     "Italy"
##  [95] "Jamaica"                    "Japan"
##  [97] "Jersey"                     "Jordan"
##  [99] "Kazakhstan"                 "Kenya"
## [101] "Kosovo"                     "Kuwait"
## [103] "Kyrgyzstan"                 "Laos"
## [105] "Latvia"                     "Lebanon"
## [107] "Lesotho"                    "Liberia"
## [109] "Libya"                      "Liechtenstein"
## [111] "Lithuania"                  "Luxembourg"
## [113] "Macao"                      "Madagascar"
## [115] "Malawi"                     "Malaysia"
## [117] "Maldives"                   "Mali"
## [119] "Malta"                      "Mauritania"
## [121] "Mauritius"                  "Mexico"
## [123] "Moldova"                    "Monaco"
## [125] "Mongolia"                   "Montenegro"
## [127] "Montserrat"                 "Morocco"
## [129] "Mozambique"                 "Myanmar"
## [131] "Namibia"                    "Nauru"
## [133] "Nepal"                      "Netherlands"
## [135] "New Caledonia"              "New Zealand"
## [137] "Nicaragua"                  "Niger"
## [139] "Nigeria"                    "Niue"
## [141] "North Macedonia"            "Northern Cyprus"
## [143] "Norway"                     "Oman"
## [145] "Pakistan"                   "Palestine"
## [147] "Panama"                     "Papua New Guinea"
## [149] "Paraguay"                   "Peru"
## [151] "Philippines"                "Pitcairn"
## [153] "Poland"                     "Portugal"
## [155] "Qatar"                      "Romania"
## [157] "Russia"                     "Rwanda"
## [159] "Saint Helena"               "Samoa"
## [161] "San Marino"                 "Sao Tome and Principe"
## [163] "Saudi Arabia"               "Senegal"
```

```
## [165] "Serbia"                    "Seychelles"
## [167] "Sierra Leone"              "Singapore"
## [169] "Sint Maarten (Dutch part)"  "Slovakia"
## [171] "Slovenia"                  "Solomon Islands"
## [173] "Somalia"                   "South Africa"
## [175] "South Korea"               "South Sudan"
## [177] "Spain"                     "Sri Lanka"
## [179] "Sudan"                     "Suriname"
## [181] "Sweden"                    "Switzerland"
## [183] "Syria"                     "Taiwan"
## [185] "Tajikistan"                "Thailand"
## [187] "Timor"                     "Togo"
## [189] "Tonga"                     "Trinidad and Tobago"
## [191] "Tunisia"                   "Turkey"
## [193] "Turkmenistan"              "Turks and Caicos Islands"
## [195] "Tuvalu"                    "Uganda"
## [197] "Ukraine"                   "United Arab Emirates"
## [199] "United Kingdom"            "United States"
## [201] "Uruguay"                   "Uzbekistan"
## [203] "Vanuatu"                   "Venezuela"
## [205] "Vietnam"                   "Wallis and Futuna"
## [207] "Yemen"                     "Zambia"
## [209] "Zimbabwe"
```

```r
info_temp_vac$vaccines <- str_replace_all(info_temp_vac$vaccines, " ","")
    usage_vac<- unique(info_temp_vac$vaccines)
    vaccine<- vector()
    for (i in usage_vac){
        for (j in strsplit(i, ",")){
            vaccine<- c(vaccine, j)
        }
    }
    used_vcc<- unique(vaccine)
    used_vcc
```

```
##  [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"
```

```r
    info_vac_val <- data.frame(matrix(ncol = length(used_vcc), nrow = 0))
for (i in info_temp_vac$vaccines){
  info_vac_val<- rbind(info_vac_val, Vectorize(grepl, USE.NAMES = TRUE)(used_vcc, str_replace_all(i," "
}
info_vac_val[info_vac_val == TRUE] = 1
info_vac_val[info_vac_val == FALSE] =0
colnames(info_vac_val) <- paste0(unique(vaccine))

countrywise_vacc<- info_vac_val %>%
mutate(country = info_temp_vac$country)%>%
group_by(country)%>%
```

```
summarise_all(sum)
data <- data.frame("No_of_countries"= apply(countrywise_vacc[-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()
```
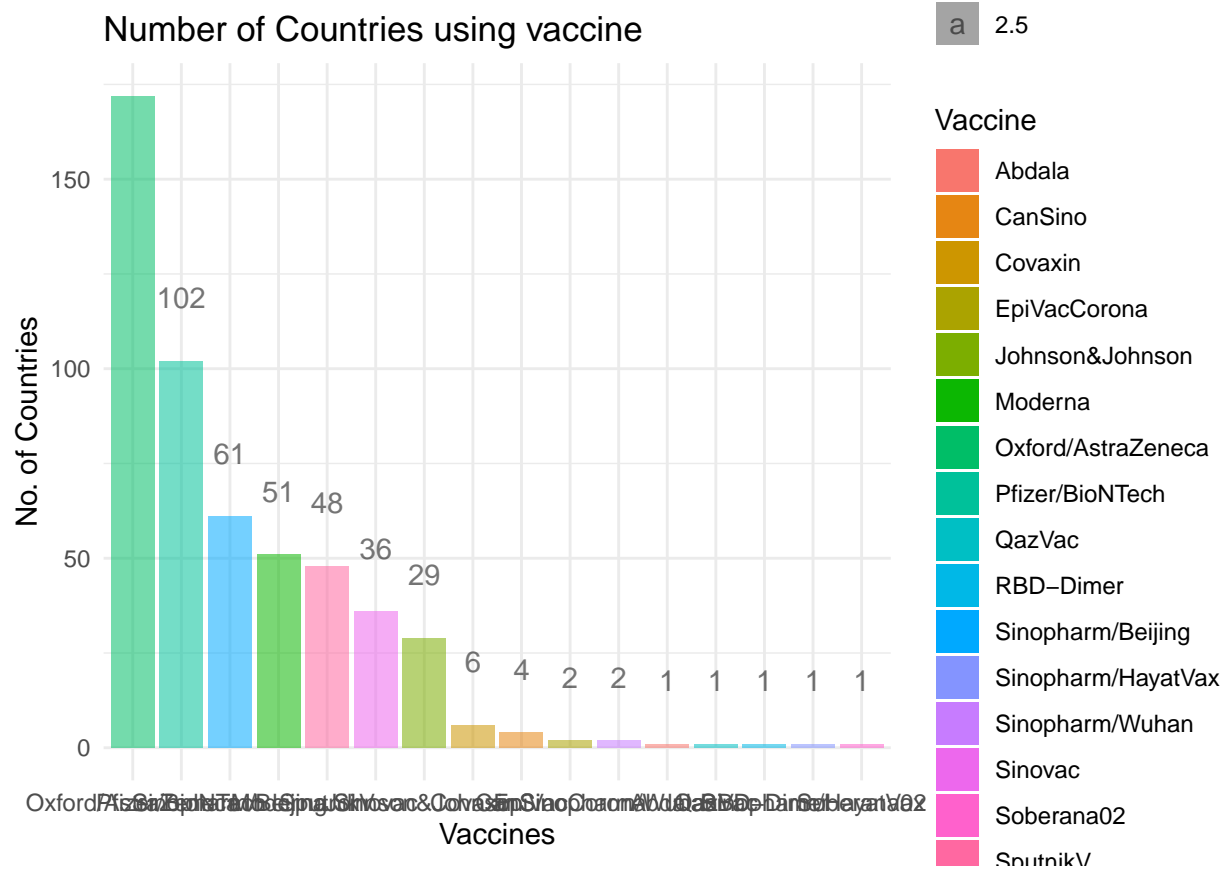

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.