Visualizing Amounts Improvement

Md Easin Hasan

3/7/2021

# Import the dataset:

data<-read.csv(file="owid-covid-data.csv",header=TRUE)  
dim(data)

## [1] 73384 59

We have obtained the data-set on COVID-19 (coronavirus) by Our World in Data. They have up-to-date data on confirmed cases, deaths, hospitalizations, testing, and vaccinations, throughout the duration of the COVID-19 pandemic. The dimension of this dataset is 73384 by 59.

data1 <- data[data$iso\_code == "USA", ]   
data3 <- data[69885:69962,]  
data3 <-data3[,c(1,4,35,36,37,38,39,40,41,42,43)]  
colnames(data3)

## [1] "iso\_code"   
## [2] "date"   
## [3] "total\_vaccinations"   
## [4] "people\_vaccinated"   
## [5] "people\_fully\_vaccinated"   
## [6] "new\_vaccinations"   
## [7] "new\_vaccinations\_smoothed"   
## [8] "total\_vaccinations\_per\_hundred"   
## [9] "people\_vaccinated\_per\_hundred"   
## [10] "people\_fully\_vaccinated\_per\_hundred"   
## [11] "new\_vaccinations\_smoothed\_per\_million"

Since we will visualize the number of people vaccinated in USA, here we kept the observations only on USA. The covid-19 vaccinations start from Dec 20, 2020 so we considered the observations from Dec 20, 2020 to Mar 07, 2021 in USA.

miss.info<- function(dat, filename=NULL){   
   
 vnames <- colnames(dat); vnames   
   
 n <- nrow(dat)   
   
 out <- NULL   
   
 for (j in 1: ncol(dat)){   
   
 vname <- colnames(dat)[j]   
   
 x <- as.vector(dat[,j])   
   
 n1 <- sum(is.na(x), na.rm=T)   
   
 n2 <- sum(x=="NA", na.rm=T)   
   
 n3 <- sum(x=="", na.rm=T)   
   
 nmiss <- n1 + n2 + n3   
   
 ncomplete <- n-nmiss   
   
 out <- rbind(out, c(col.number=j, vname=vname, mode=mode(x), n.levels=length(unique(x)), ncomplete=ncomplete, miss.perc=nmiss/n)) }   
   
 out <- as.data.frame(out)   
   
 row.names(out) <- NULL   
   
 return(out)   
   
}   
  
miss.info(data3)

## col.number vname mode n.levels  
## 1 1 iso\_code character 2  
## 2 2 date character 78  
## 3 3 total\_vaccinations numeric 64  
## 4 4 people\_vaccinated numeric 63  
## 5 5 people\_fully\_vaccinated numeric 50  
## 6 6 new\_vaccinations numeric 55  
## 7 7 new\_vaccinations\_smoothed numeric 78  
## 8 8 total\_vaccinations\_per\_hundred numeric 64  
## 9 9 people\_vaccinated\_per\_hundred numeric 63  
## 10 10 people\_fully\_vaccinated\_per\_hundred numeric 50  
## 11 11 new\_vaccinations\_smoothed\_per\_million numeric 78  
## ncomplete miss.perc  
## 1 78 0  
## 2 78 0  
## 3 63 0.192307692307692  
## 4 62 0.205128205128205  
## 5 49 0.371794871794872  
## 6 54 0.307692307692308  
## 7 77 0.0128205128205128  
## 8 63 0.192307692307692  
## 9 62 0.205128205128205  
## 10 49 0.371794871794872  
## 11 77 0.0128205128205128

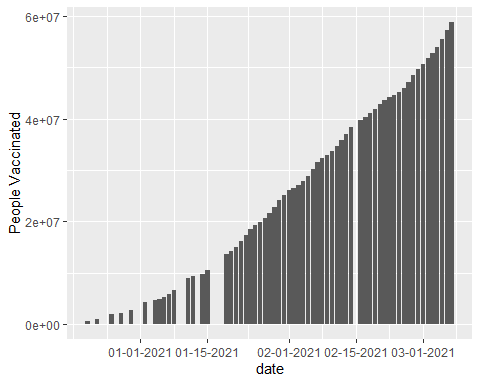
Here we checked the missing value in the selected observations. Since we would like to see in which date there is no vaccination reported, we didn’t remove the missing values.

# people vaccinated

library(ggplot2)  
library(reshape2)

## Warning: package 'reshape2' was built under R version 4.0.4

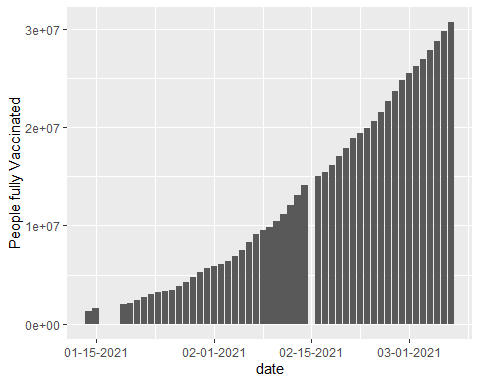
library(scales)  
  
ggplot(data3, aes(x = as.Date(date), y = people\_vaccinated)) +  
 xlab("date") + ylab("People Vaccinated")+  
 geom\_bar(stat="identity")+  
 scale\_x\_date(labels = date\_format("%m-%d-%Y"))



From this graph we can see that number of people vaccinated is increasing gradually.

# people fully vaccinated

ggplot(data3, aes(x = as.Date(date), y = people\_fully\_vaccinated)) +  
 xlab("date") + ylab("People fully Vaccinated")+  
 geom\_bar(stat="identity")+  
 scale\_x\_date(labels = date\_format("%m-%d-%Y"))



The number of people fully vaccinated also increasing gradually.

# Visualizing Amounts Improvement

dfm <- melt(data3[,c('date','people\_vaccinated','people\_fully\_vaccinated')],id.vars = 1)  
ggplot(dfm,aes(x = as.Date(date),y = value)) +   
 geom\_bar(aes(fill = variable),stat = "identity") +   
 scale\_x\_date(labels = date\_format("%m-%d-%Y"))+  
 labs(x = "date")+  
 labs(y = "Number of people vaccinated vs fully vaccinated")

Chart, bar chart, histogram

Description automatically generated

From this visualization we can very easily compare the number of people vaccinated vs number of people fully vaccinated each day in USA.