COVID-19 Deaths By Vaccination Status

Jorge Valderrama

2023-12-23

##Importing Data

#read data  
setwd("~/Downloads")  
USACOVID <- read.csv ("COVID\_Data\_R2.csv")

## Loading Libraries

library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.2 ✔ readr 2.1.4  
## ✔ forcats 1.0.0 ✔ stringr 1.5.0  
## ✔ ggplot2 3.4.2 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.2 ✔ tidyr 1.3.0  
## ✔ purrr 1.0.1   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(ggplot2)  
library(ggthemes)  
library(gganimate)

##Demonstrating Untidy Data

USACOVID

## Month Vaccinated.with.updated.booster  
## 1 8/1/2022 NA  
## 2 8/14/2022 NA  
## 3 8/21/2022 NA  
## 4 8/28/2022 NA  
## 5 9/4/2022 NA  
## 6 9/11/2022 NA  
## 7 9/18/2022 0.00  
## 8 9/25/2022 0.21  
## 9 10/2/2022 0.12  
## 10 10/9/2022 0.11  
## 11 10/16/2022 0.22  
## 12 10/23/2022 0.18  
## 13 10/30/2022 0.23  
## 14 11/6/2022 0.20  
## 15 11/13/2022 0.29  
## 16 11/20/2022 0.31  
## 17 11/27/2022 0.37  
## 18 12/4/2022 0.43  
## 19 12/11/2022 0.40  
## 20 12/18/2022 0.50  
## 21 12/25/2022 0.53  
## 22 1/1/2023 0.44  
## 23 1/8/2023 0.39  
## 24 1/15/2023 0.33  
## 25 1/22/2023 0.30  
## 26 1/29/2023 0.31  
## 27 2/5/2023 0.26  
## 28 2/12/2023 0.28  
## 29 2/19/2023 0.31  
## 30 2/26/2023 0.23  
## 31 3/5/2023 0.21  
## 32 3/12/2023 0.16  
## 33 3/19/2023 0.17  
## 34 3/26/2023 0.17  
## 35 NA  
## 36 NA  
## 37 NA  
## 38 NA  
## 39 NA  
## 40 NA  
## 41 NA  
## 42 NA  
## Vaccinated.without.updated.booster Unvaccinated  
## 1 0.67 3.69  
## 2 0.60 3.35  
## 3 0.59 3.06  
## 4 0.61 3.03  
## 5 0.55 2.79  
## 6 0.51 2.47  
## 7 0.46 2.40  
## 8 0.44 2.51  
## 9 0.48 2.41  
## 10 0.49 2.48  
## 11 0.56 2.46  
## 12 0.59 2.85  
## 13 0.60 2.85  
## 14 0.62 2.71  
## 15 0.63 2.86  
## 16 0.73 3.24  
## 17 0.88 4.08  
## 18 0.90 4.01  
## 19 0.86 3.83  
## 20 0.93 4.03  
## 21 1.01 4.38  
## 22 0.87 3.84  
## 23 0.67 3.29  
## 24 0.55 2.32  
## 25 0.50 2.02  
## 26 0.41 2.10  
## 27 0.42 1.67  
## 28 0.39 1.38  
## 29 0.32 1.60  
## 30 0.30 1.26  
## 31 0.23 1.33  
## 32 0.21 0.95  
## 33 0.16 0.84  
## 34 0.14 0.61  
## 35 NA NA  
## 36 NA NA  
## 37 NA NA  
## 38 NA NA  
## 39 NA NA  
## 40 NA NA  
## 41 NA NA  
## 42 NA NA

##Making Data Tidy

tidyUSACOVID <-USACOVID %>%   
 pivot\_longer(cols = -Month)

##Demonstrating Tidy Data

tidyUSACOVID

## # A tibble: 126 × 3  
## Month name value  
## <chr> <chr> <dbl>  
## 1 8/1/2022 Vaccinated.with.updated.booster NA   
## 2 8/1/2022 Vaccinated.without.updated.booster 0.67  
## 3 8/1/2022 Unvaccinated 3.69  
## 4 8/14/2022 Vaccinated.with.updated.booster NA   
## 5 8/14/2022 Vaccinated.without.updated.booster 0.6   
## 6 8/14/2022 Unvaccinated 3.35  
## 7 8/21/2022 Vaccinated.with.updated.booster NA   
## 8 8/21/2022 Vaccinated.without.updated.booster 0.59  
## 9 8/21/2022 Unvaccinated 3.06  
## 10 8/28/2022 Vaccinated.with.updated.booster NA   
## # ℹ 116 more rows

##Loading lubridate package to change the “Month” column from character type variable to a date type variable

library(lubridate)

##Changing variable type for “Month” variable

tidyUSACOVID$Month <- mdy(tidyUSACOVID$Month)

##Demonstrating change in variable type

tidyUSACOVID

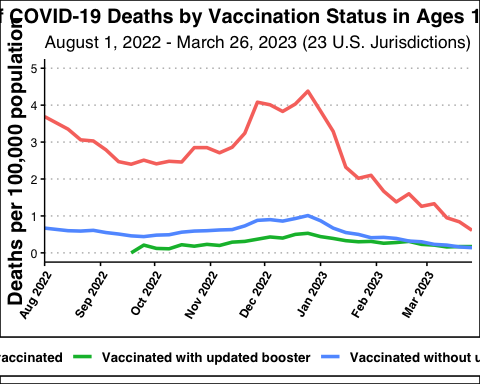
## # A tibble: 126 × 3  
## Month name value  
## <date> <chr> <dbl>  
## 1 2022-08-01 Vaccinated.with.updated.booster NA   
## 2 2022-08-01 Vaccinated.without.updated.booster 0.67  
## 3 2022-08-01 Unvaccinated 3.69  
## 4 2022-08-14 Vaccinated.with.updated.booster NA   
## 5 2022-08-14 Vaccinated.without.updated.booster 0.6   
## 6 2022-08-14 Unvaccinated 3.35  
## 7 2022-08-21 Vaccinated.with.updated.booster NA   
## 8 2022-08-21 Vaccinated.without.updated.booster 0.59  
## 9 2022-08-21 Unvaccinated 3.06  
## 10 2022-08-28 Vaccinated.with.updated.booster NA   
## # ℹ 116 more rows

##Creating a line graph to demonstrate differences in COVID-19 death rates among fully vaccinated and unvaccinated individuals over time.

ggplot(data=tidyUSACOVID, aes(x=Month, y=value, group=name, color=name))+  
 geom\_line(size = 1.25)+  
 ggtitle("Rates of COVID-19 Deaths by Vaccination Status in Ages 18 and Older")+  
 theme\_clean()+  
 theme(axis.text.x=element\_text(angle=60, hjust=1, face = "bold"))+  
 theme(axis.title.y=element\_text(size=15,face="bold"))+  
 theme(plot.title = element\_text(hjust = 0.5, lineheight = 0.9))+  
 theme(plot.title = element\_text(face = "bold", size = 15))+  
 theme(axis.title.x = element\_blank())+  
 theme(legend.title = element\_blank())+  
 theme(legend.text = element\_text(face = "bold", size = 10))+  
 theme(legend.position = "bottom")+  
 scale\_y\_continuous(limits=c(0,5))+  
 labs(y = "Deaths per 100,000 population", subtitle = "August 1, 2022 - March 26, 2023 (23 U.S. Jurisdictions)")+  
 theme(plot.subtitle = element\_text(hjust=0.5))+  
 scale\_color\_hue(labels=c('Unvaccinated', 'Vaccinated with updated booster', 'Vaccinated without updated booster'))+  
 scale\_x\_date(date\_breaks= "1 month", date\_labels = "%b %Y", limit=as.Date(c('2022-08-01', '2023-03-26')), expand=c(0,0))

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
## ℹ Please use `linewidth` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.

## Warning: Removed 30 rows containing missing values (`geom\_line()`).

 ##Creating an animated plot

tidyUSACOVID <- na.omit(tidyUSACOVID)  
  
ggplot(data=tidyUSACOVID, aes(x=Month, y=value, group=name, color=name))+  
 geom\_line(size = 1.25)+  
 ggtitle("Rates of COVID-19 Deaths by Vaccination Status in Ages 18 and Older")+  
 theme\_clean()+  
 theme(axis.text.x=element\_text(angle=60, hjust=1, face = "bold"))+  
 theme(axis.title.y=element\_text(size=15,face="bold"))+  
 theme(plot.title = element\_text(hjust = 0.5, lineheight = 0.9))+  
 theme(plot.title = element\_text(face = "bold", size = 15))+  
 theme(axis.title.x = element\_blank())+  
 theme(legend.title = element\_blank())+  
 theme(legend.text = element\_text(face = "bold", size = 10))+  
 theme(legend.position = "bottom")+  
 scale\_y\_continuous(limits=c(0,5))+  
 labs(y = "Deaths per 100,000 population", subtitle = "August 1, 2022 - March 26, 2023 (23 U.S. Jurisdictions)")+  
 theme(plot.subtitle = element\_text(hjust=0.5))+  
 scale\_color\_hue(labels=c('Unvaccinated', 'Vaccinated with updated booster', 'Vaccinated without updated booster'))+  
 scale\_x\_date(date\_breaks= "1 month", date\_labels = "%b %Y", limit=as.Date(c('2022-08-01', '2023-03-26')), expand=c(0,0))+  
 transition\_reveal(Month)

## `geom\_line()`: Each group consists of only one observation.  
## ℹ Do you need to adjust the group aesthetic?  
## `geom\_line()`: Each group consists of only one observation.  
## ℹ Do you need to adjust the group aesthetic?

