# Greeshma Chowdary Peddineni

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

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

library(ggplot2)  
library(lubridate)

##   
## Attaching package: 'lubridate'

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

library(readr)  
library(forcats)

new\_now\_covid\_22 <- readRDS("C:/Users/Dell/Downloads/covid22.rds")

Let’s find how many different vaccine manufacturers are there?

uniq\_manufacturers <- unique(new\_now\_covid\_22$VAX\_MANU)  
num\_manufacturers <- length(uniq\_manufacturers)  
  
cat("There are", num\_manufacturers, "different COVID-19 vaccine manufacturers:\n")

## There are 5 different COVID-19 vaccine manufacturers:

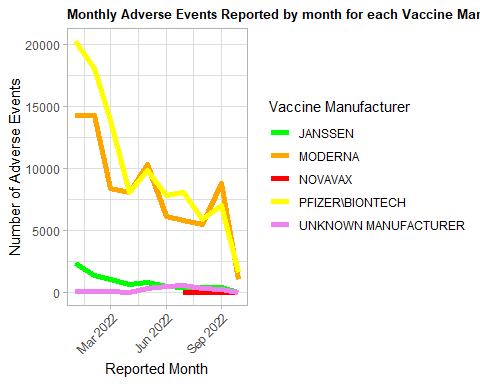
cat(paste("", uniq\_manufacturers), sep = "\t")

## JANSSEN PFIZER\BIONTECH UNKNOWN MANUFACTURER MODERNA NOVAVAX

Create a plot to show adverse events by month for each manufacturer

new\_now\_covid\_22 <- new\_now\_covid\_22 %>%  
 mutate(RECVDATE = as.Date(RECVDATE, format="%Y-%m-%d"),  
 Month = floor\_date(RECVDATE, "month"))   
  
monthly\_events <- new\_now\_covid\_22 %>%  
 group\_by(Month, VAX\_MANU) %>%  
 summarise(Event\_Count = n(), .groups = 'drop')  
  
ggplot(monthly\_events, aes(x = Month, y = Event\_Count, color = VAX\_MANU)) +  
 geom\_line(size = 2) +   
 labs(title = "Monthly Adverse Events Reported by month for each Vaccine Manufacturer",  
 x = "Reported Month",  
 y = "Number of Adverse Events",  
 color = "Vaccine Manufacturer") +  
 theme\_light() +   
 scale\_x\_date(date\_labels = "%b %Y", date\_breaks = "3 months") +   
 scale\_color\_manual(values = c("MODERNA" = "orange",  
 "PFIZER\\BIONTECH" = "yellow",   
 "JANSSEN" = "green",  
 "NOVAVAX" = "red",  
 "UNKNOWN MANUFACTURER" = "violet"),  
 na.translate = FALSE) +   
 theme(plot.title = element\_text(face = "bold", size = 10),   
 axis.text.x = element\_text(angle = 45, hjust = 1))

## 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.



Let’s see which state had the highest number of events?

new\_now\_covid\_22 <- new\_now\_covid\_22 %>%  
 mutate(  
 STATE = as.character(STATE),   
 STATE = trimws(STATE)   
 )  
  
invisible(unique(new\_now\_covid\_22$STATE))  
  
state\_event\_counts <- new\_now\_covid\_22 %>%  
 filter(!is.na(STATE) & STATE != "") %>%   
 group\_by(STATE) %>%  
 summarise(Event\_Count = n(), .groups = 'drop') %>%  
 arrange(desc(Event\_Count))   
  
highest\_state <- state\_event\_counts[1, ]  
  
cat("The state with the highest number of reported events is", highest\_state$STATE,   
 "with", highest\_state$Event\_Count, "events.\n")

## The state with the highest number of reported events is CA with 16444 events.

For Idaho, Texas, and Michigan, show number of events by gender?

selected\_states <- c("TX", "MI", "ID")  
filtered\_data <- new\_now\_covid\_22 %>% filter(STATE %in% selected\_states)  
  
  
gender\_counts <- filtered\_data %>%  
 group\_by(STATE, SEX) %>%  
 summarise(Event\_Count = n(), .groups = 'drop')  
  
gender\_counts

## # A tibble: 9 × 3  
## STATE SEX Event\_Count  
## <chr> <fct> <int>  
## 1 ID F 545  
## 2 ID M 388  
## 3 ID U 26  
## 4 MI F 6012  
## 5 MI M 4608  
## 6 MI U 143  
## 7 TX F 4802  
## 8 TX M 3099  
## 9 TX U 187

Let’s find how many ER/ED visits were reported?

er\_ed\_visits\_count <- new\_now\_covid\_22 %>%  
 filter(!is.na(ER\_ED\_VISIT) & ER\_ED\_VISIT != "") %>%  
 nrow()  
  
cat("Total ER/ED visits reported:", er\_ed\_visits\_count, "\n")

## Total ER/ED visits reported: 21669

Which manufacturer is associated with the lowest number of reported ER/ED visits?

manufacturer\_er\_ed\_counts <- new\_now\_covid\_22 %>%  
 filter(ER\_ED\_VISIT == TRUE | ER\_ED\_VISIT == "Y") %>%  
 group\_by(VAX\_MANU) %>%  
 summarise(ER\_ED\_Visit\_Count = n(), .groups = 'drop')  
  
  
low\_manufacturer <- manufacturer\_er\_ed\_counts %>%  
 arrange(ER\_ED\_Visit\_Count) %>%  
 slice(1)  
  
cat(low\_manufacturer$VAX\_MANU, "manufacturer is associated with the lowest number of reported ER/ED with", low\_manufacturer$ER\_ED\_Visit\_Count, "visits.")

## NOVAVAX manufacturer is associated with the lowest number of reported ER/ED with 9 visits.