Vaccination rates and COVID-19 cases and deaths in California

Carmen Chen

12/2/2021

This is my PM566 Final Project website home. The website is online at https://carmenchenepi.github.io/PM566-Final-Project/ (https://carmenchenepi.github.io/PM566-Final-Project/).

```
knitr::opts_chunk$set(echo = TRUE)
library(data.table)
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.1.1
```

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

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.2     v dplyr   1.0.6
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## Warning: package 'ggplot2' was built under R version 4.1.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::between()   masks data.table::between()
## x dplyr::filter()    masks stats::filter()
## x dplyr::first()     masks data.table::first()
## x dplyr::lag()       masks stats::lag()
## x dplyr::last()      masks data.table::last()
## x purrr::transpose() masks data.table::transpose()
```

```
library(dplyr)
library(ggplot2)
library(gtsummary)
```

```
## Warning: package 'gtsummary' was built under R version 4.1.1
```

```
## #BlackLivesMatter
```

```
library(plotly)
```

```
##
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following object is masked from 'package:graphics':
##
##     layout
```

Introduction

The vaccinations have begun almost a year since its distribution. The fully vaccinated percentage in California was increasing and stagnated around 60%. However, the herd immunity didn't seem to be achieved and the COVID-19 daily new cases is still very high in California. Therefore, this project aims to investigate the relationship of vaccination rates on COVID-19 daily new cases and deaths in California by using the data from California Health & Human Services Agency and US Center for Disease Control and Prevention.

Methods

Data of COVID-19 cases and vaccination rates were acquired from California Health & Human Services Agency and US Center for Disease Control and Prevention, respectively. Data in California and with variables of interest (i.e., date, county, percentage of first dose, percentage of second dose, daily new cases, cumulative cases, daily new deaths, and cumulative deaths) were extracted. The date variables were formatted in both datasets.

```
cases <- read.csv("data/Cases.csv")
vaccinations <- read.csv("data/Vaccinations.csv")
county_cases <- read.csv("data/Cases.csv")
county_vaccinations <- read.csv("data/county_vaccine.csv")
```

```
#subset the data
vaccinations <- vaccinations[vaccinations$Location == "CA",]
vaccinations <- vaccinations[, c("Date", "Administered_Dose1_Pop_Pct", "Series_Complete_Pop_Pct"
)]
cases <- cases[cases$area == "California",]
cases <- cases[, c("date", "cases", "cumulative_cases", "deaths", "cumulative_deaths")]
#rename the data
vaccinations <- rename(vaccinations, date = Date, dose1 = Administered_Dose1_Pop_Pct, dose2 = Se
ries_Complete_Pop_Pct)
#format date
vaccinations$date <- as.Date(vaccinations$date, format = "%m/%d/%Y")
cases$date <- as.Date(cases$date, format = "%Y-%m-%d")
```

```
#subset the data
county_vaccinations <- county_vaccinations[county_vaccinations$Recip_State == "CA",]
county_vaccinations <- county_vaccinations[, c("Recip_County", "Date", "Administered_Dose1_Pop_P
ct", "Series_Complete_Pop_Pct")]
county_cases <- county_cases[county_cases$area_type == "County",]
county_cases <- county_cases[,c("area", "date", "population", "cases", "cumulative_cases", "deat
hs", "cumulative_deaths")]
county_vaccinations <- rename(county_vaccinations, county = Recip_County, date = Date, dose1 = A
dministered_Dose1_Pop_Pct, dose2 = Series_Complete_Pop_Pct)
#create new variable
county_cases$casesper100k <- ((county_cases$cases)/(county_cases$population))*100000
county_cases$deathsper100k <- ((county_cases$deaths)/(county_cases$population))*100000
county_cases$cu_casesper100k <- ((county_cases$cumulative_cases)/(county_cases$population))*1000
00
county_cases$cu_deathsper100k <- ((county_cases$cumulative_deaths)/(county_cases$population))*10
0000
#format date
county_vaccinations$date <- as.Date(county_vaccinations$date, format = "%m/%d/%Y")
county_cases$date <- as.Date(county_cases$date, format = "%Y-%m-%d")
```

Dimensions, headers, and footers of the two datasets were checked. There are 311 observations and 3 variables in the “vaccinations” dataset, as well as 628 observations and 5 variables in the “cases” dataset. Implausible data (e.g., 0 cases increase) was found in the date variable on “2021-10-20” in the “cases” dataset. Considering the 14-

day incubation period of the COVID-19 disease, the data from 2021-10-06 to 2021-10-20 were not the final accurate number of cases and deaths since there are still many cases and deaths were not reported timely. Thus, these data were removed from the “cases” dataset.

```
#check the date
summary(vaccinations$date)
```

```
##         Min.      1st Qu.       Median         Mean      3rd Qu.         Max.
## "2020-12-14" "2021-03-01" "2021-05-18" "2021-05-18" "2021-08-03" "2021-10-20"
```

```
summary(cases$date)
```

```
##         Min.      1st Qu.       Median         Mean      3rd Qu.         Max.
## "2020-02-01" "2020-07-06" "2020-12-10" "2020-12-10" "2021-05-16" "2021-10-20"
##         NA's
##          "1"
```

```
#check for missing value
cases[!complete.cases(cases),]
```

```
##      date cases cumulative_cases deaths cumulative_deaths
## 3774 <NA>   481          4600506    269             70741
```

```
vaccinations[!complete.cases(vaccinations),]
```

```
## [1] date  dose1 dose2
## <0 rows> (or 0-length row.names)
```

```
#remove rows with missing value
cases <- cases[complete.cases(cases),]
#check the dimensions, headers, footers
dim(vaccinations)
```

```
## [1] 311   3
```

```
dim(cases)
```

```
## [1] 628   5
```

```
head(vaccinations)
```

```
##           date dose1 dose2
## 49  2021-10-20  73.5  60.5
## 115 2021-10-19  73.4  60.4
## 160 2021-10-18  73.4  60.4
## 253 2021-10-17  73.3  60.3
## 260 2021-10-16  73.2  60.2
## 384 2021-10-15  73.1  60.2
```

```
cases <- cases[order(cases$date, decreasing = TRUE),]
head(cases)
```

```
##            date cases cumulative_cases deaths cumulative_deaths
## 3773 2021-10-20     0          4600025      0              70472
## 3772 2021-10-19   604          4600025      1              70472
## 3771 2021-10-18  3269          4599421     10              70471
## 3770 2021-10-17  1944          4596152     16              70461
## 3769 2021-10-16  2408          4594208     21              70445
## 3768 2021-10-15  4418          4591800     13              70424
```

```
tail(vaccinations)
```

```
##             date dose1 dose2
## 19527 2020-12-19     0     0
## 19608 2020-12-18     0     0
## 19627 2020-12-17     0     0
## 19737 2020-12-16     0     0
## 19762 2020-12-15     0     0
## 19821 2020-12-14     0     0
```

```
tail(cases)
```

```
##            date cases cumulative_cases deaths cumulative_deaths
## 3151 2020-02-06     6               50      1                 1
## 3150 2020-02-05     4               44      0                 0
## 3149 2020-02-04     1               40      0                 0
## 3148 2020-02-03     5               39      0                 0
## 3147 2020-02-02     7               34      0                 0
## 3146 2020-02-01    27               27      0                 0
```

```
#remove the data from 2021-10-06 to 2021-10-20
cases <- cases[!(cases$date >= "2021-10-06"),]
#take a look at the variables
str(vaccinations)
```

```
## 'data.frame':    311 obs. of  3 variables:
##  $ date : Date, format: "2021-10-20" "2021-10-19" ...
##  $ dose1: num  73.5 73.4 73.4 73.3 73.2 73.1 73 72.9 72.9 72.8 ...
##  $ dose2: num  60.5 60.4 60.4 60.3 60.2 60.2 60.1 60 59.9 59.9 ...
```

```
summary(vaccinations$dose1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   16.20   53.60   42.30   65.25   73.50
```

```
summary(vaccinations$dose2)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    0.00   39.00   31.47   53.15   60.50
```

```
str(cases)
```

```
## 'data.frame':    613 obs. of  5 variables:
##  $ date             : Date, format: "2021-10-05" "2021-10-04" ...
##  $ cases            : num  5969 6552 3207 3518 5503 ...
##  $ cumulative_cases : num  4546821 4540852 4534300 4531093 4527575 ...
##  $ deaths           : num  51 48 50 69 61 53 65 66 75 69 ...
##  $ cumulative_deaths: num  70118 70067 70019 69969 69900 ...
```

```
summary(cases$cases)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       1    1650    3706    7417    8725   60094
```

```
summary(cases$cumulative_cases)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      27  340258 1457974 2001598 3661841 4546821
```

```
summary(cases$deaths)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0    29.0    71.0   114.4   114.0   707.0
```

```
summary(cases$cumulative_deaths)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0    7199   21429   33168   62592   70118
```

The two datasets were merged into one dataset by date variable. Final dataset has 296 observations and 7 variables. Exploratory data analysis was conducted in the merged dataset. No missing value, implaussible vaule or data error was found. The data includes COVID-19 partial and fully vaccination rates, daily new cases, cumulative cases, daily new deaths, as well cumulative deaths from 2020/12/14 to 2021/10/05. Both univariate and bivariate summary statistics was analyzed. Exploratory graphs were generated between vaccination rates and cases and deaths.

```
#combine the dataset
covid <- merge(vaccinations, cases, by = "date")
#exploratory analysis
dim(covid)
```

```
## [1] 296   7
```

```
head(covid)
```

```
##         date dose1 dose2 cases cumulative_cases deaths cumulative_deaths
## 1 2020-12-14     0     0 49338          1838139    359             24436
## 2 2020-12-15     0     0 48468          1886607    382             24818
## 3 2020-12-16     0     0 45848          1932455    378             25196
## 4 2020-12-17     0     0 44541          1976996    393             25589
## 5 2020-12-18     0     0 41588          2018584    463             26052
## 6 2020-12-19     0     0 29690          2048274    467             26519
```

```
tail(covid)
```

```
##           date dose1 dose2 cases cumulative_cases deaths cumulative_deaths
## 291 2021-09-30  71.7  59.0  5657          4522072     53             69839
## 292 2021-10-01  71.8  59.0  5503          4527575     61             69900
## 293 2021-10-02  71.9  59.1  3518          4531093     69             69969
## 294 2021-10-03  72.1  59.3  3207          4534300     50             70019
## 295 2021-10-04  72.2  59.4  6552          4540852     48             70067
## 296 2021-10-05  72.2  59.4  5969          4546821     51             70118
```

```
str(covid)
```

```
## 'data.frame':    296 obs. of  7 variables:
##  $ date             : Date, format: "2020-12-14" "2020-12-15" ...
##  $ dose1            : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ dose2            : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ cases            : num  49338 48468 45848 44541 41588 ...
##  $ cumulative_cases : num  1838139 1886607 1932455 1976996 2018584 ...
##  $ deaths           : num  359 382 378 393 463 467 442 494 532 503 ...
##  $ cumulative_deaths: num  24436 24818 25196 25589 26052 ...
```

```
summary(covid$date)
```

```
##         Min.      1st Qu.       Median         Mean      3rd Qu.         Max.
## "2020-12-14" "2021-02-25" "2021-05-10" "2021-05-10" "2021-07-23" "2021-10-05"
```

```
summary(covid$dose1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   14.72   51.70   40.75   63.83   72.20
```

```
summary(covid$dose2)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    0.00   35.65   30.02   52.12   59.40
```

```
summary(covid$cases)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     582    1810    4275    9318   11520   60094
```

```
summary(covid$cumulative_cases)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 1838139 3512726 3669089 3631341 3833183 4546821
```

```
summary(covid$deaths)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    5.00   22.75   63.50  155.54  158.75  707.00
```

```
summary(covid$cumulative_cases)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 1838139 3512726 3669089 3631341 3833183 4546821
```

```
covid[!complete.cases(covid),]
```

```
## [1] date              dose1             dose2             cases
## [5] cumulative_cases  deaths            cumulative_deaths
## <0 rows> (or 0-length row.names)
```

Preliminary Results

Table 1 presented the summary statistics of the data, including range of the date and median (IQR) of first dose, second dose, daily new cases, cumulative cases, daily new deaths, and cumulative deaths. There are a total of 296 observations collected from 2020-12-14 to 2021-10-05. The maximum partial and fully vaccination rates in

California are 52% and 64%, respectively. The medians (IQRs) of the daily new cases and deaths are 4275 (1810, 11520) and 64 (23, 159), respectively.

```
#Table 1. Characteristics of the COVID-19 data
covid %>%
  select("Date" = date,
         "First dose, %" = dose1,
         "Second dose, %" = dose2,
         "Daily new cases" = cases,
         "Cumulative cases" = cumulative_cases,
         "Daily new deaths" = deaths,
         "Cumulative deaths" = cumulative_deaths) %>%
  tbl_summary() %>%
  modify_caption ("**Table 1. Characteristics of the COVID-19 data**")
```

Table 1. Characteristics of the COVID-19 data

Characteristic	**N = 296**[1]
Date	2020-12-14 to 2021-10-05
First dose, %	52 (15, 64)
Second dose, %	36 (0, 52)
Daily new cases	4,275 (1,810, 11,520)
Cumulative cases	3,669,089 (3,512,726, 3,833,183)
Daily new deaths	64 (23, 159)
Cumulative deaths	62,710 (58,891, 63,905)

[1] Range; Median (IQR)

Table 2 presented the correlation coefficients of vaccination rates with daily new cases and deaths. Vaccination rates were negatively associated with daily new cases and deaths. Such negative association was stronger in the correlation between vaccination rates and daily new deaths (R: -0.78 for first dose, -0.68 for second dose).

```
#First dose correlation
Cases = cor(covid$dose1, covid$cases, use = "complete")
Deaths = cor(covid$dose1, covid$deaths, use = "complete")
`First dose` = rbind(Cases, Deaths)
colnames(`First dose`) <- "First dose"
#Second dose correlation
Cases = cor(covid$dose2, covid$cases, use = "complete")
Deaths = cor(covid$dose2, covid$deaths, use = "complete")
`Second dose` = rbind(Cases, Deaths)
colnames(`Second dose`) <- "Second dose"
#Combine the table
cbind(`First dose`, `Second dose`) %>%
  knitr::kable(caption = "<strong>Table 2. Correlation coefficients of vacciantion rates and cas
es/deaths</strong>")
```

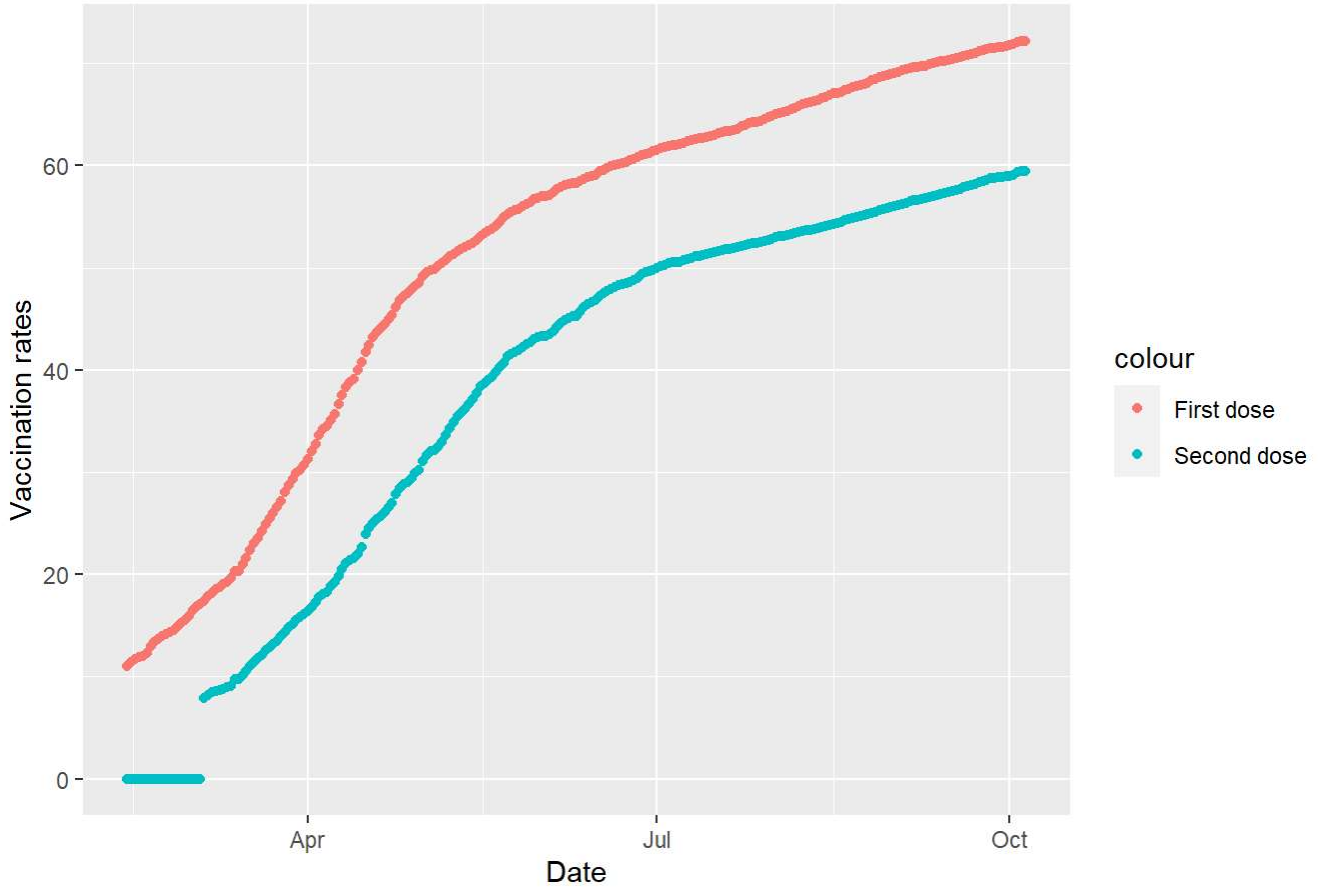
Table 2. Correlation coefficients of vacciantion rates and cases/deaths

	First dose	Second dose
Cases	-0.4990338	-0.3931054
Deaths	-0.7814445	-0.6816794

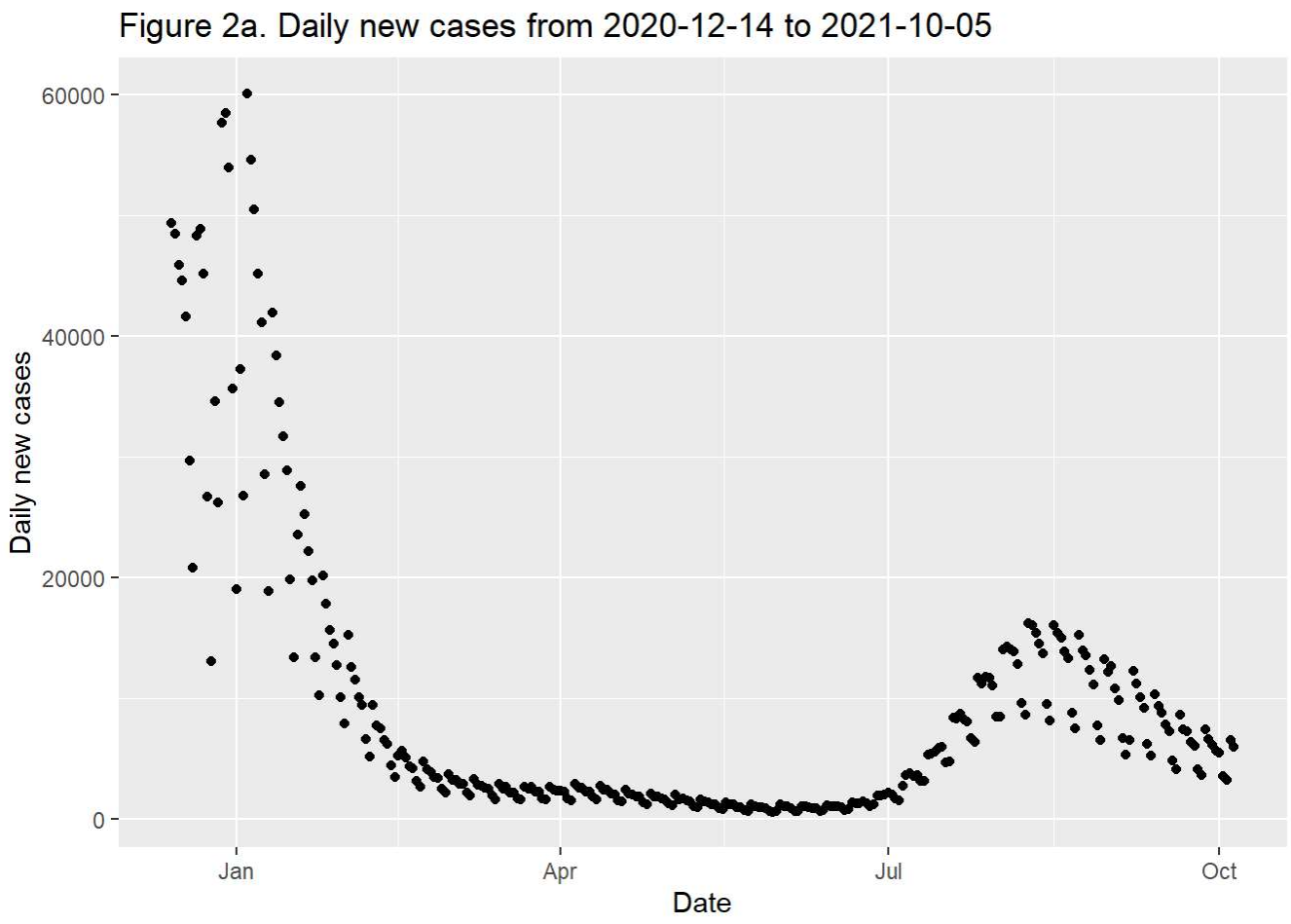
Exploratory graphs were presented in figure 1 and figure 2a-2e. Vaccinations in California started in February and the rates continued to increase. The increase of vaccination rates became more slowly when it achieved around 60% of first dose vaccination rate. The daily new cases started to decrease drastically around February and the cases remained in a stable small number until July. A small break out in daily new cases occurred in July and it achieved its peak in September. The trend pattern of the daily new deaths is similar to daily new cases.

```
#Vaccination rates
covid[covid$dose1>0,] %>%
ggplot() +
  geom_point(mapping = aes(x = date, y = dose1, color = "First dose")) +
  geom_point(mapping = aes(x = date, y = dose2, color = "Second dose")) +
  labs(title = "Figure 1. Vaccination rates from 2020-12-14 to 2021-10-05") +
  labs(x = "Date", y = "Vaccination rates")
```

Figure 1. Vaccination rates from 2020-12-14 to 2021-10-05

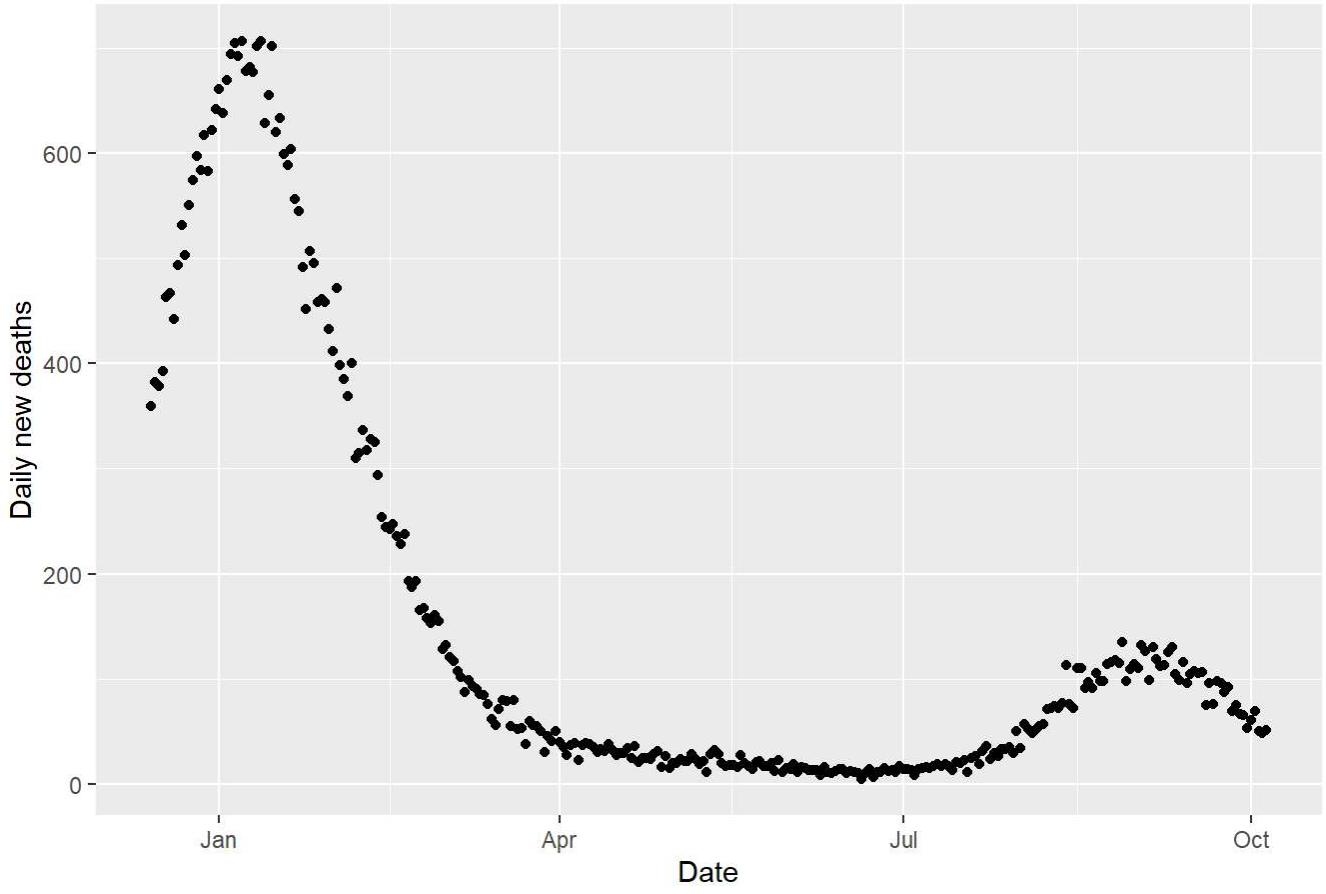



```
#Cases
ggplot(data = covid) +
  geom_point(mapping = aes(x = date, y = cases)) +
  labs(title = "Figure 2a. Daily new cases from 2020-12-14 to 2021-10-05") +
  labs(x = "Date", y = "Daily new cases")
```



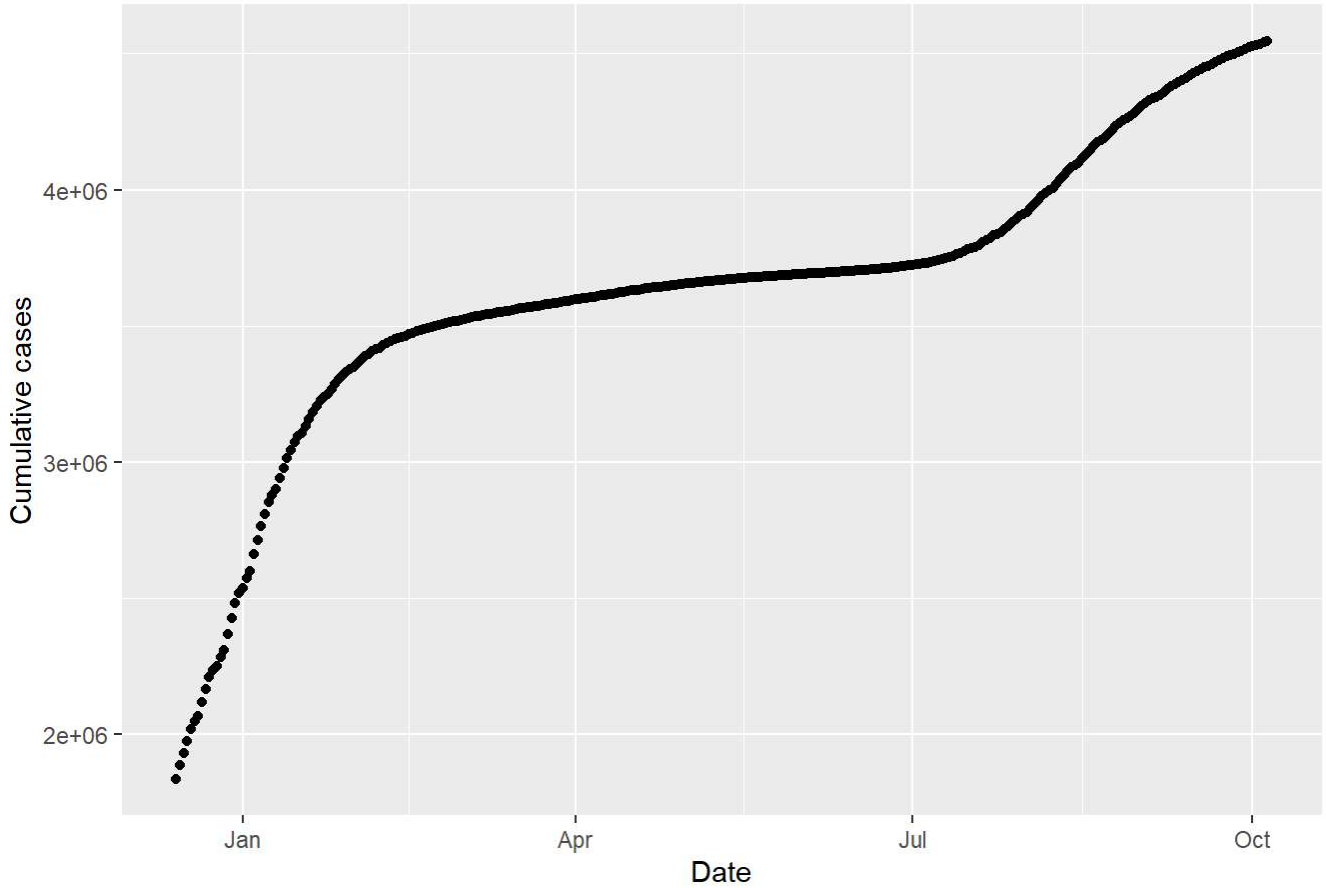

```
#Deaths
ggplot(data = covid) +
  geom_point(mapping = aes(x = date, y = deaths)) +
  labs(title = "Figure 2b. Daily new deaths from 2020-12-14 to 2021-10-05") +
  labs(x = "Date", y = "Daily new deaths")
```

Figure 2b. Daily new deaths from 2020-12-14 to 2021-10-05

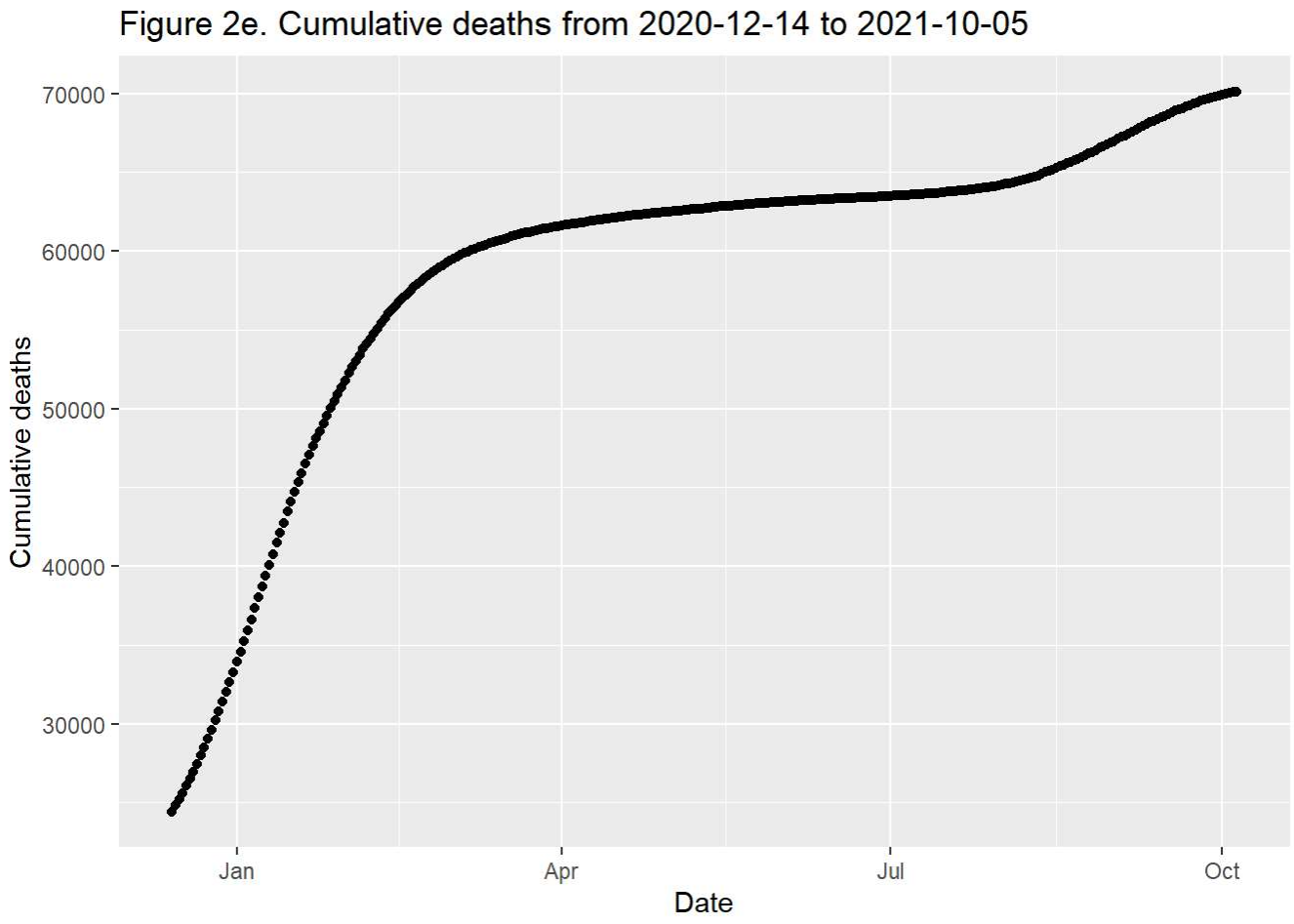



```
#Cumulative cases
ggplot(data = covid) +
  geom_point(mapping = aes(x = date, y = cumulative_cases)) +
  labs(title = "Figure 2c. Cumulative cases from 2020-12-14 to 2021-10-05") +
  labs(x = "Date", y = "Cumulative cases")
```

Figure 2c. Cumulative cases from 2020-12-14 to 2021-10-05



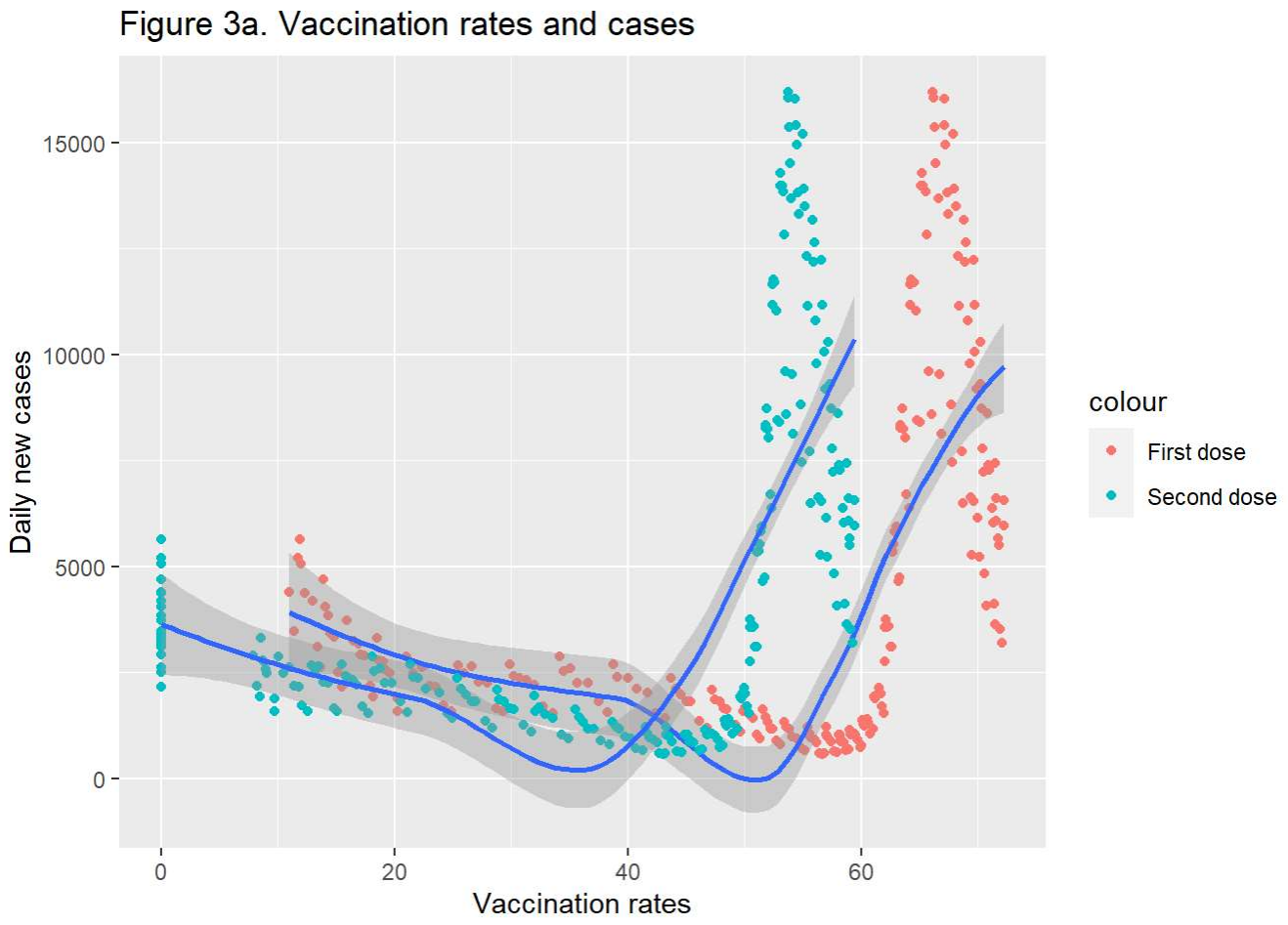

```
#Cumulative deaths
ggplot(data = covid) +
  geom_point(mapping = aes(x = date, y = cumulative_deaths)) +
  labs(title = "Figure 2e. Cumulative deaths from 2020-12-14 to 2021-10-05") +
  labs(x = "Date", y = "Cumulative deaths")
```



Data visualization graphs were generated in Figure 3a & 3b to visualize the relationship between vaccination rates and daily new cases and deaths. There was an non-symmetrical inverse pattern in the relationship of vaccination rates and daily new cases in the beginning. Daily new cases was negatively associated with vaccination rates until the first dose vaccination rates achieved around 50%. After that, daily new cases increase drastically with the increase of vaccination rates. This same pattern was also found in the relationship of vaccination rates and daily new deaths. However, a stronger negative association was observed in the daily new deaths figure in the beginning of the data.

```
#Vaccination rates and cases
covid[covid$dose1 > 0,] %>%
ggplot() +
  geom_point(mapping = aes(x = dose1, y = cases, color = "First dose")) +
  geom_smooth(mapping = aes(x = dose1, y = cases)) +
  geom_point(mapping = aes(x = dose2, y = cases, color = "Second dose")) +
  geom_smooth(mapping = aes(x = dose2, y = cases)) +
  labs(title = "Figure 3a. Vaccination rates and cases") +
  labs(x = "Vaccination rates", y = "Daily new cases")
```

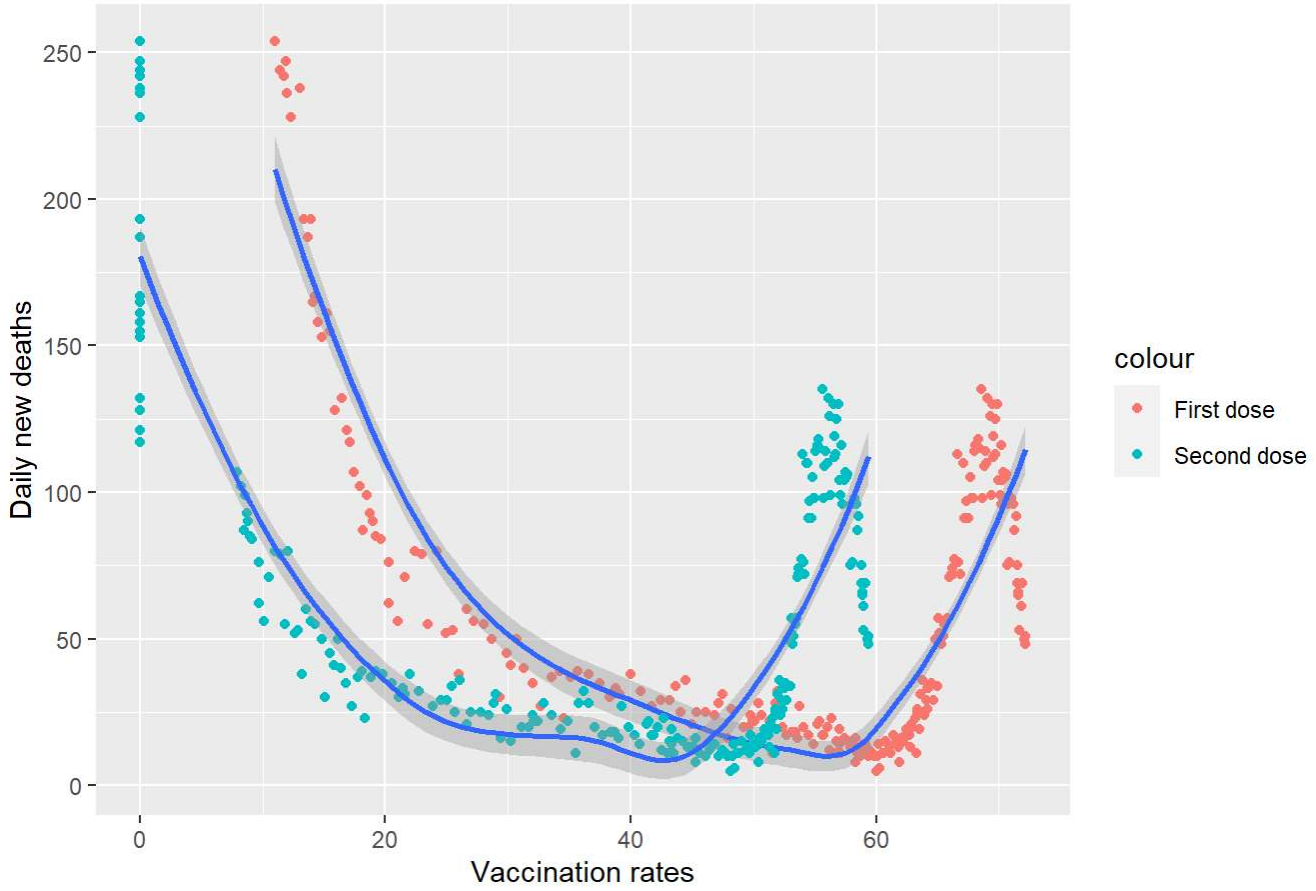
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
#Vaccination rates and deaths
covid[covid$dose1 > 0,] %>%
ggplot() +
  geom_point(mapping = aes(x = dose1, y = deaths, color = "First dose")) +
  geom_smooth(mapping = aes(x = dose1, y = deaths)) +
  geom_point(mapping = aes(x = dose2, y = deaths, color = "Second dose")) +
  geom_smooth(mapping = aes(x = dose2, y = deaths)) +
  labs(title = "Figure 3b. Vaccination rates and deaths") +
  labs(x = "Vaccination rates", y = "Daily new deaths")
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

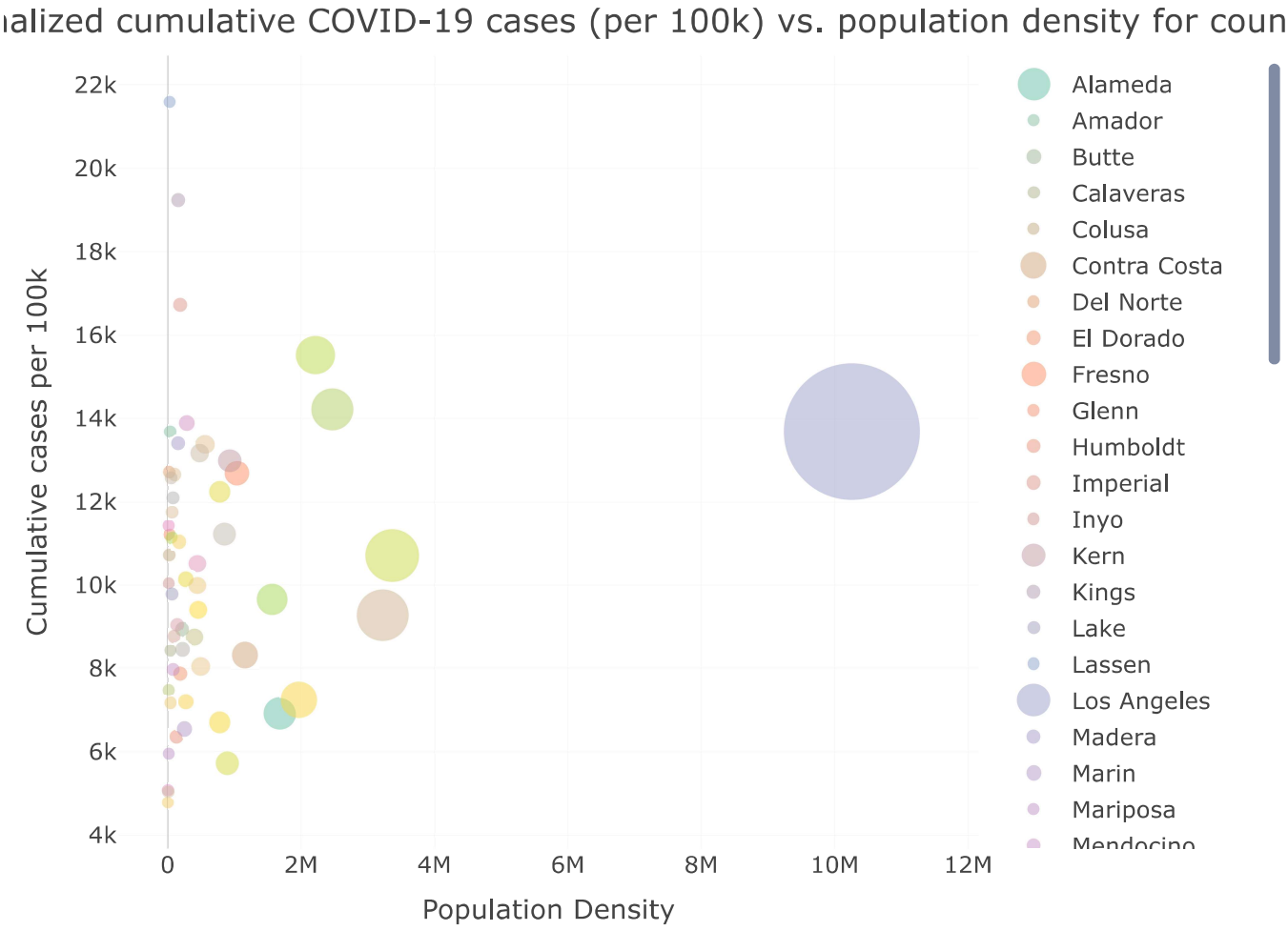
Figure 3b. Vaccination rates and deaths



Interactive visualization plots were created to observe the cumulative COVID-19 cases and deaths, as well as vaccination rates across different counties in California. Accourding to the interactive plot, there is no association between the cumulative COVID-19 cases per 100k population and population density observed. Lassen, Kings, and Imperial county are the three counties with the most cumulative COVID-19 cases per 100k population, following with San Berndino, Riverside, and Los Angeles. Imperial, San Berdino, and Los Angeles are the three counties with the most cumulative COVID-19 deaths per 100k population, following with San Joaquin, Inyo, and Riverside. Lassen, Tehama, and Del Norte are the three counties with the least vaccination rates of less than 40% fully vaccinated people, following with Kings and Shasta.

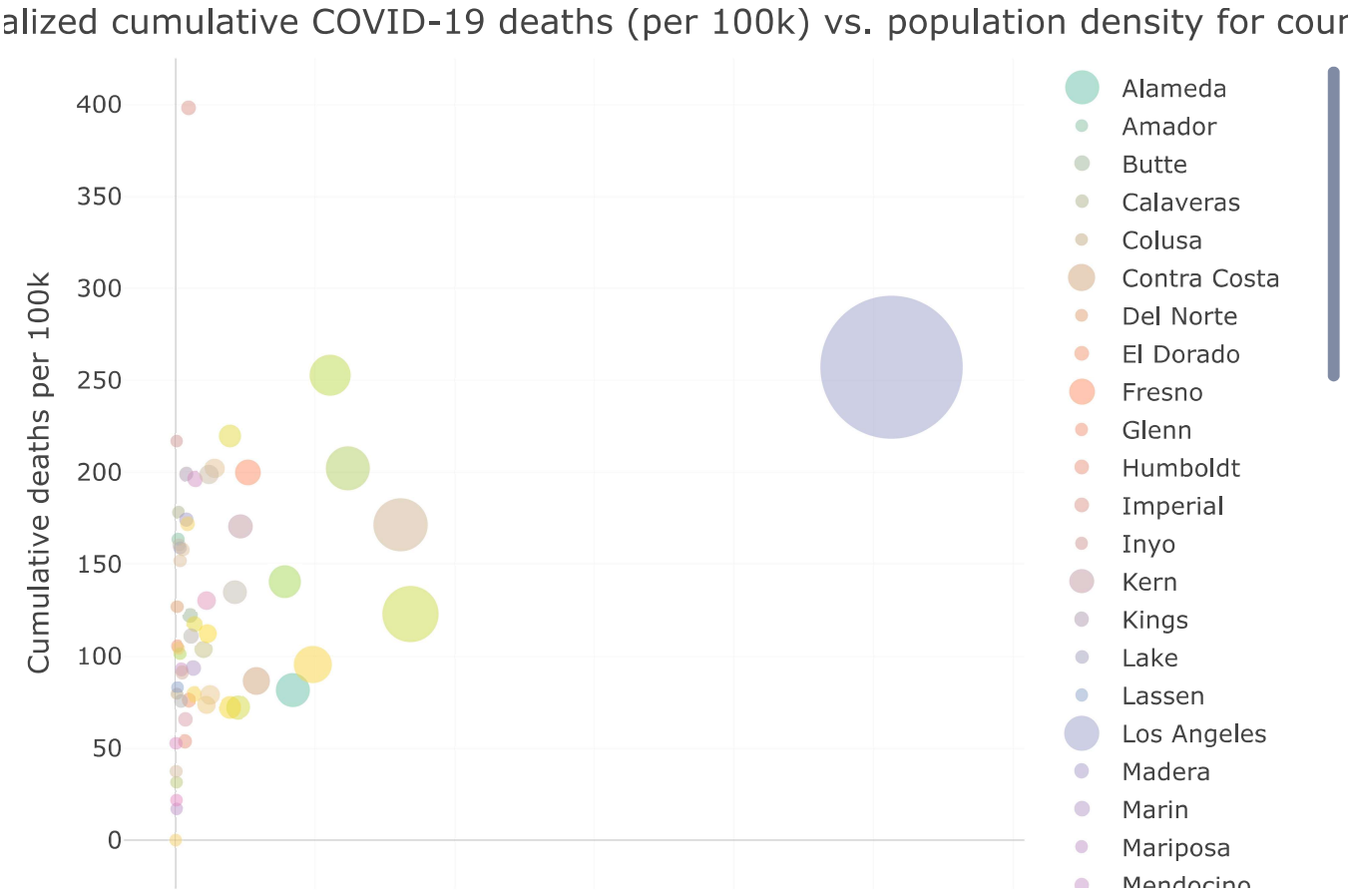
```
#Interactive plot 1: Cumulative Covid-19 cases across counties
county_cases %>% filter(cases>0 & date == "2021-10-05") %>%
  plot_ly(x = ~population, y = ~cu_casesper100k,
          type = 'scatter', mode = 'markers', color = ~area,
          size = ~population, sizes = c(5, 70), marker = list(sizemode = 'diameter', opacity =
0.5)) %>%
  layout(title = "Population-normalized cumulative COVID-19 cases (per 100k) vs. population dens
ity for counties in California", yaxis = list(title = "Cumulative cases per 100k"), xaxis = list
(title = "Population Density"),
         hovermode = "compare")
```

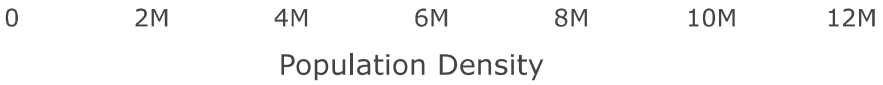
```
## Warning in RColorBrewer::brewer.pal(N, "Set2"): n too large, allowed maximum for palette Set2
is 8
## Returning the palette you asked for with that many colors

## Warning in RColorBrewer::brewer.pal(N, "Set2"): n too large, allowed maximum for palette Set2
is 8
## Returning the palette you asked for with that many colors
```



```
#Interactive plot 2: Cumulative Covid-19 deaths across counties
county_cases %>% filter(cases>0 & date == "2021-10-05") %>%
  plot_ly(x = ~population, y = ~cu_deathsper100k,
          type = 'scatter', mode = 'markers', color = ~area,
          size = ~population, sizes = c(5, 70), marker = list(sizemode = 'diameter', opacity =
0.5),
          hoverinfo = 'text',
          text = ~paste( paste(area, ":", sep = ""), paste(" Cumulative cases per 100k:", cu_cas
esper100k, sep = ""), paste(" Cumulative deaths per 100k:", cu_deathsper100k, sep = ""), sep = "
<br>")) %>%
  layout(title = "Population-normalized cumulative COVID-19 deaths (per 100k) vs. population den
sity for counties in California", yaxis = list(title = "Cumulative deaths per 100k"), xaxis = li
st(title = "Population Density"),
         hovermode = "compare")
```

```
## Warning in RColorBrewer::brewer.pal(N, "Set2"): n too large, allowed maximum for palette Set2
is 8
## Returning the palette you asked for with that many colors

## Warning in RColorBrewer::brewer.pal(N, "Set2"): n too large, allowed maximum for palette Set2
is 8
## Returning the palette you asked for with that many colors
```

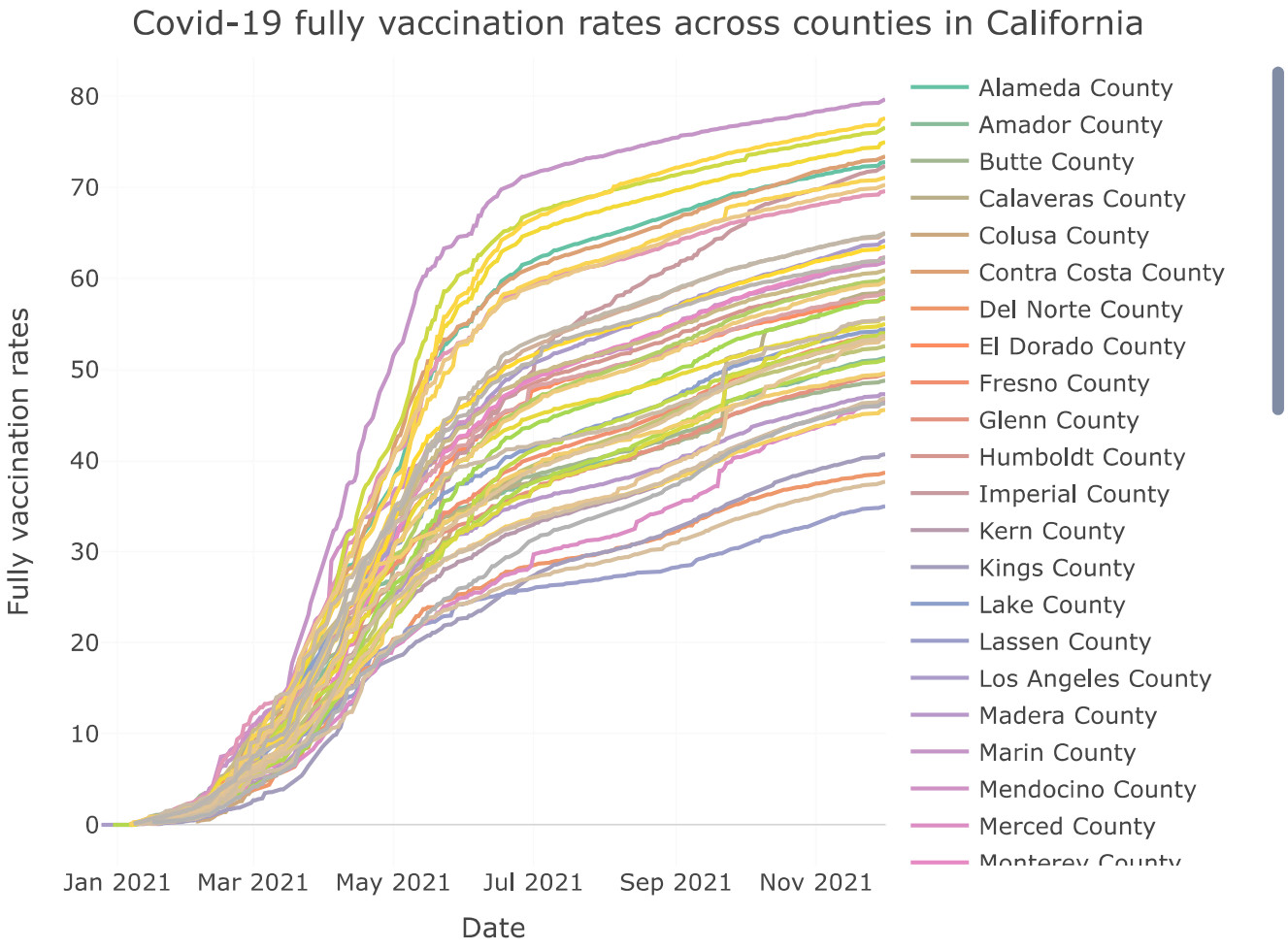




Population Density

```
#Interactive plot 3: Covid-19 vaccination rates across counties
county_vaccinations %>% filter(dose1>0) %>%
  plot_ly(x=~date, y=~dose2, color=~county, type="scatter", mode="lines",
          hoverinfo = 'text',
          text = ~paste(paste("Date: ", date, sep = ""), paste(county, ":", sep=""), paste("Vacc
ination rate: ", dose2, sep=""))) %>%
  layout(title = "Covid-19 fully vaccination rates across counties in California", yaxis = list
(title = "Fully vaccination rates"), xaxis = list(title = "Date"),
         hovermode = "compare")
```

```
## Warning in RColorBrewer::brewer.pal(N, "Set2"): n too large, allowed maximum for palette Set2
is 8
## Returning the palette you asked for with that many colors

## Warning in RColorBrewer::brewer.pal(N, "Set2"): n too large, allowed maximum for palette Set2
is 8
## Returning the palette you asked for with that many colors
```



Conclusion

There is a positive association of vaccination rates and daily new cases and deaths when the first dose of vaccination rate achieved around 50%. This may be due to the re-opening of the economic and lift of mask mandate during that time. Overall, we could see the protective effect of vaccine towards infection and death according to the data in the beginning. A stronger negative association in the beginning of vaccination rates and daily new deaths compared to daily new cases may be due to a stronger efficacy of the vaccine towards preventing mortality.