Challenge-11

```
library(httr)
library(jsonlite)
library(tidyverse)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
✔ dplyr     1.1.0     ✔ readr     2.1.4
✔ forcats   1.0.0     ✔ stringr   1.5.0
✔ ggplot2   3.4.3     ✔ tibble    3.1.8
✔ lubridate 1.9.2     ✔ tidyr     1.3.0
✔ purrr     1.0.1
── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
✖ dplyr::filter()  masks stats::filter()
✖ purrr::flatten() masks jsonlite::flatten()
✖ dplyr::lag()     masks stats::lag()
ℹ Use the ]8;;http://conflicted.r-lib.org/ conflicted package]8;;  to force all
conflicts to become errors
```

```
historic_state_data_url <- "https://api.covidactnow.org/v2/states.timeseries.json?a
raw_data <- GET(historic_state_data_url)

data <- fromJSON(rawToChar(raw_data$content))
glimpse(data)
```

```
Rows: 53
Columns: 25
$ fips                   <chr> "02", "01", "05", "04", "06", "08", "09…
$ country                <chr> "US", "US", "US", "US", "US", "US", "US…
$ state                  <chr> "AK", "AL", "AR", "AZ", "CA", "CO", "CT…
$ county                 <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,…
$ hsa                    <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,…
$ hsaName                <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,…
$ level                  <chr> "state", "state", "state", "state", "st…
$ lat                    <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,…
$ locationId             <chr> "iso1:us#iso2:us-ak", "iso1:us#iso2:us-…
$ long                   <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,…
$ population             <int> 731545, 4903185, 3017804, 7278717, 3951…
$ hsaPopulation          <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,…
$ metrics                <df[,14]> <data.frame[26 x 14]>
$ riskLevels             <df[,6]> <data.frame[26 x 6]>
$ cdcTransmissionLevel   <int> 2, 4, 3, 3, 1, 4, 4, 1, 4, 4, 2, 3,…
$ communityLevels        <df[,2]> <data.frame[26 x 2]>
$ actuals                <df[,19]> <data.frame[26 x 19]>
$ annotations            <df[,30]> <data.frame[26 x 30]>
$ lastUpdatedDate        <chr> "2023-10-30", "2023-10-30", "2023-10…
$ url                    <chr> "https://covidactnow.org/us/alaska-ak",…
$ metricsTimeseries      <list> [<data.frame[1334 x 14]>], [<data.fr…
$ actualsTimeseries      <list> [<data.frame[1334 x 20]>], [<data.f…
$ riskLevelsTimeseries   <list> [<data.frame[1334 x 3]>], [<data.fr…
```

```
$ cdcTransmissionLevelTimeseries <list> [<data.frame[1334 x 2]>], [<data.frame[…
$ communityLevelsTimeseries      <list> [<data.frame[1334 x 3]>], [<data.frame[…
```

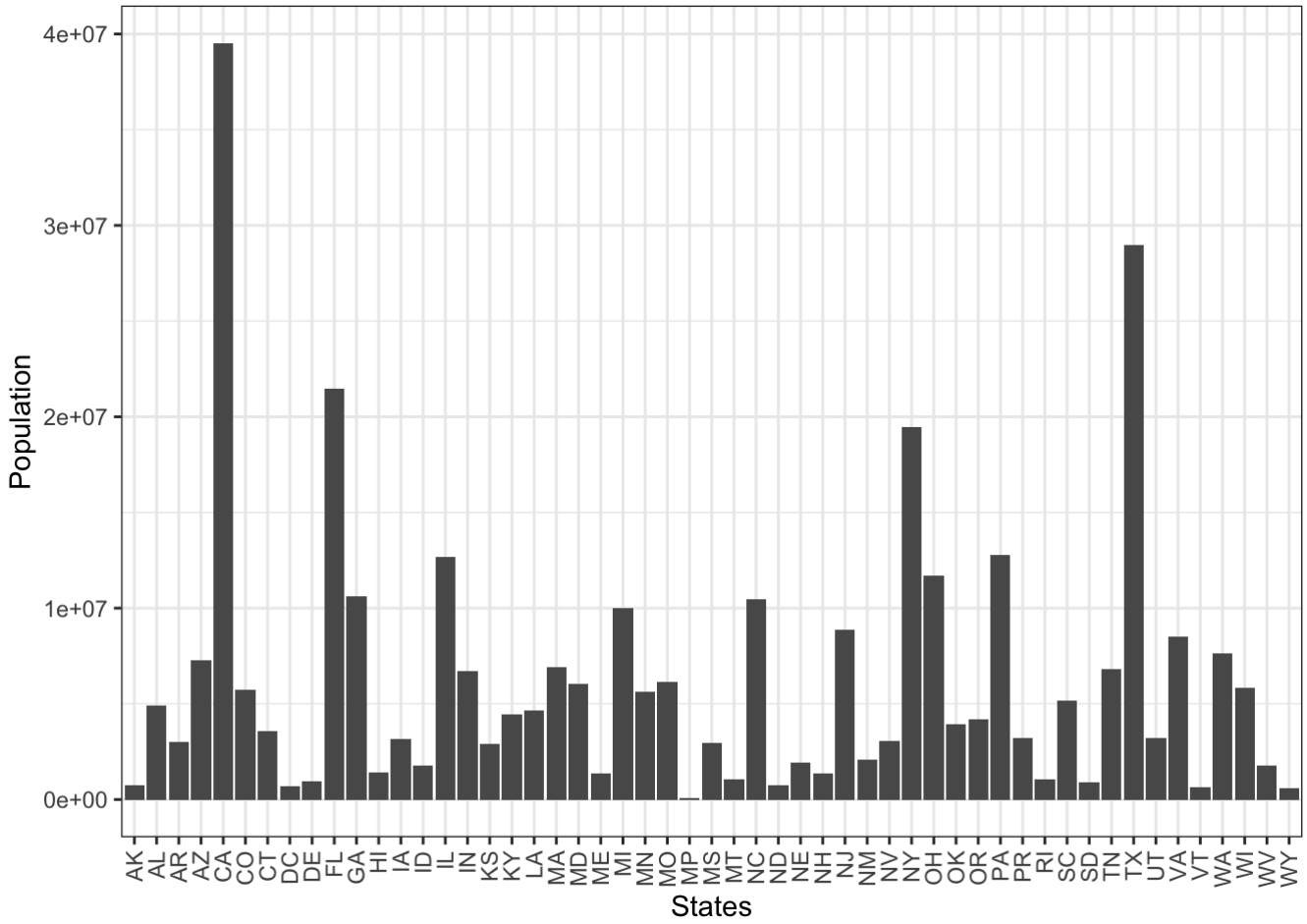
```
time_series <- data %>%
  unnest(actualsTimeseries)

# creating new dataframe with relevant data,
time_series_transmission <- tibble(Date=time_series$cdcTransmissionLevelTimeseries[[wh
time_series_transmission$Alaska <- time_series$cdcTransmissionLevelTimeseries[[which(d
time_series_transmission$California <- time_series$cdcTransmissionLevelTimeseries[[whi
time_series_transmission$New_Jersey <- time_series$cdcTransmissionLevelTimeseries[[whi
time_series_transmission$Tennessee <- time_series$cdcTransmissionLevelTimeseries[[whic
time_series_transmission$District_of_Columbia <- time_series$cdcTransmissionLevelTimes

print(head(time_series_transmission))
```

```
# A tibble: 6 × 6
  Date       Alaska California New_Jersey Tennessee District_of_Columbia
  <chr>       <int>      <int>      <int>     <int>                <int>
1 2020-03-01      0          0          0         0                    0
2 2020-03-02      0          0          0         0                    0
3 2020-03-03      0          0          0         0                    0
4 2020-03-04      0          0          0         0                    0
5 2020-03-05      0          0          0         0                    0
6 2020-03-06      0          0          0         0                    0
```

```
# creating new dataframe with dates,
time_series_cases <- list(Alaska = time_series %>% filter(state=="AK") %>% select(date
time_series_cases$California <- time_series %>% filter(state=="CA") %>% select(date,ca
time_series_cases$New_Jersey <- time_series %>% filter(state=="NJ") %>% select(date,ca
time_series_cases$Tennessee <- time_series %>% filter(state=="TN") %>% select(date,cas
time_series_cases$District_of_Columbia <- time_series %>% filter(state=="DC") %>% sele
```

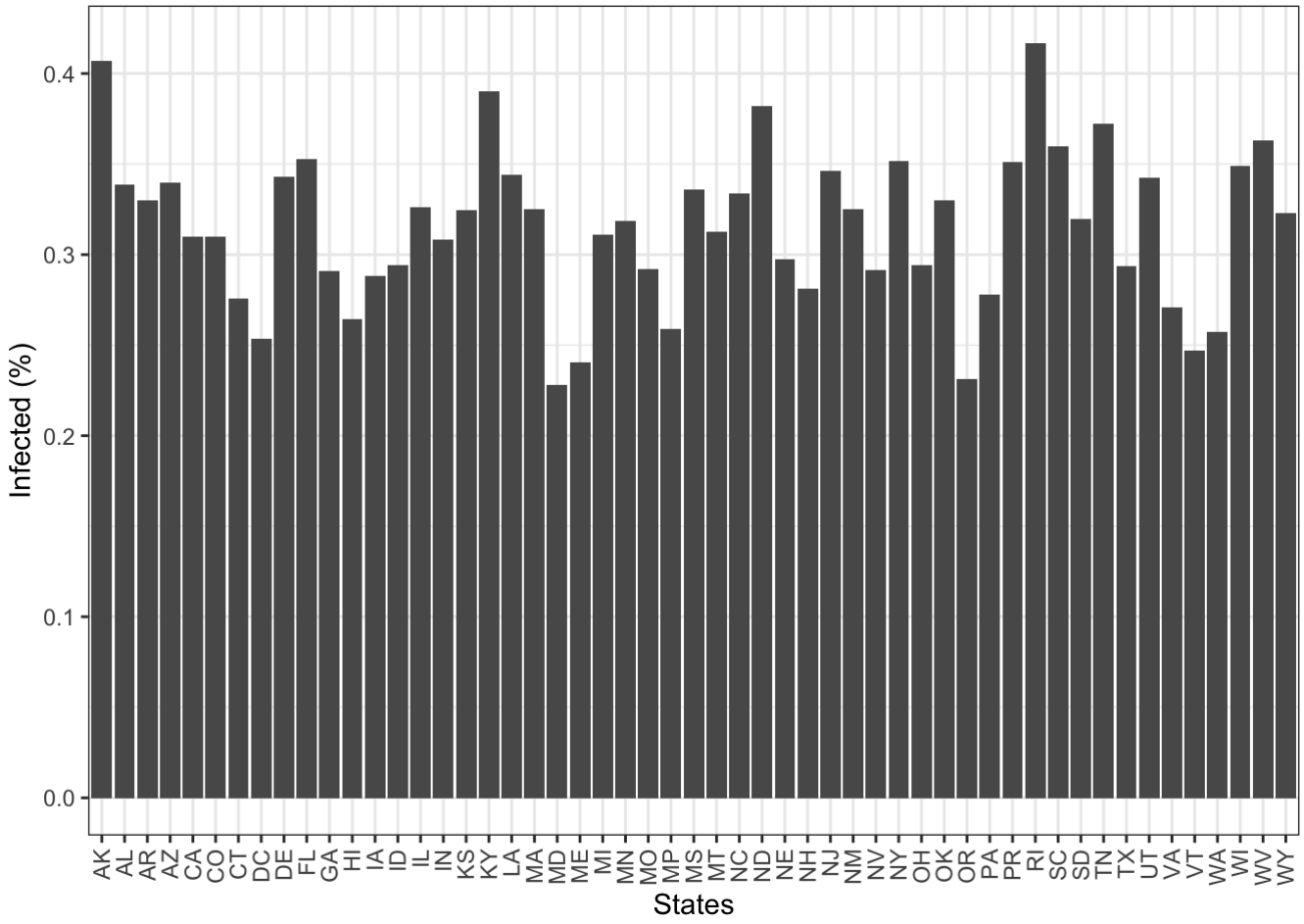
i. What is the population in the various states of U.S.A?

```
library(ggplot2)
ggplot(data, aes(x=state,y=population)) +
  geom_bar(stat="identity") +
  labs(x="States",y="Population") +
  theme_bw() +
  scale_x_discrete(guide = guide_axis(angle = 90))
```



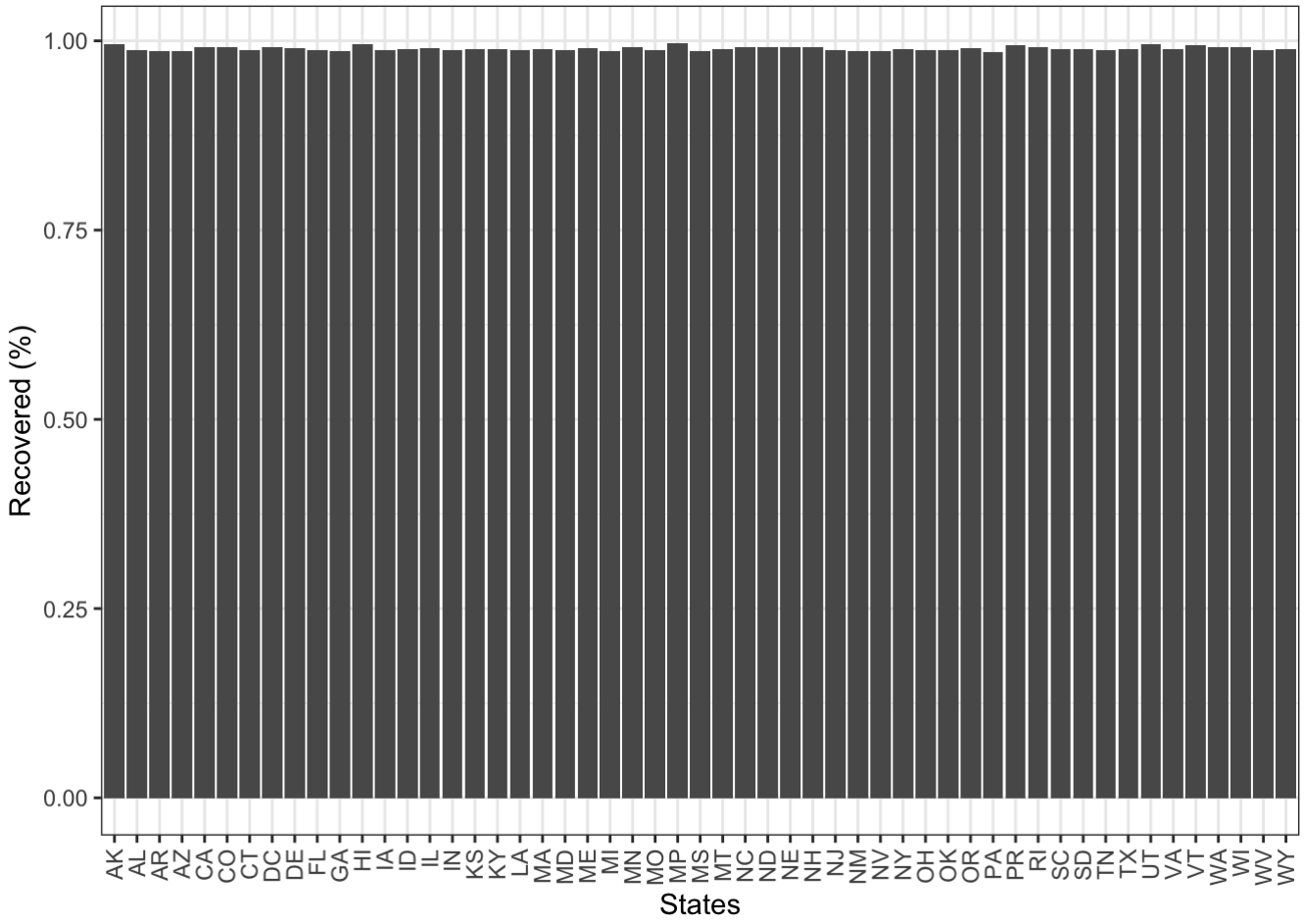
ii. What fraction of the population was infected?

```
ggplot(data, aes(x=state,y=(data$actuals$cases/population))) +
  geom_bar(stat="identity") +
  labs(x="States",y="Infected (%)") +
  theme_bw() +
  scale_x_discrete(guide = guide_axis(angle = 90))
```



iii. What fraction of infected persons recovered?

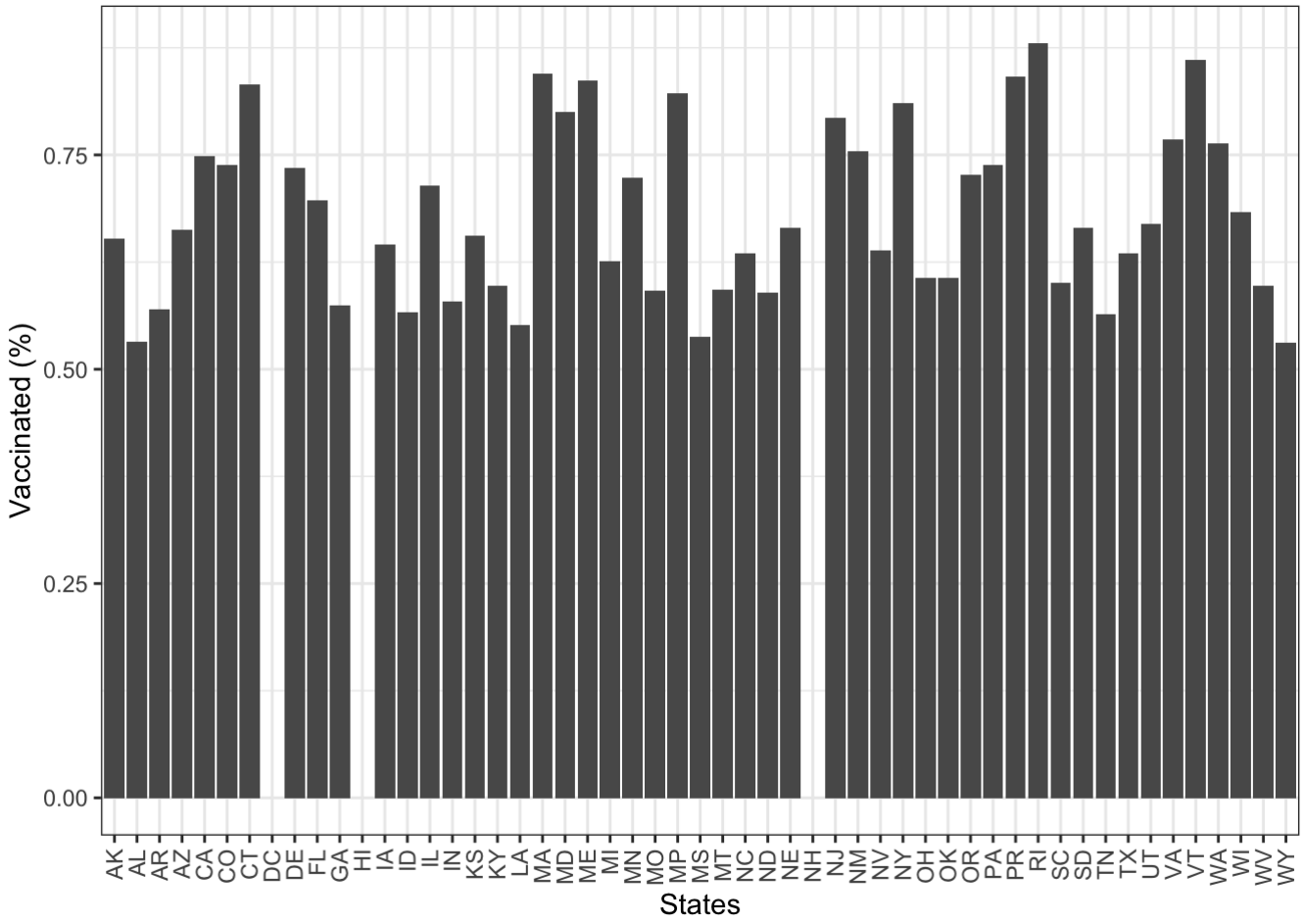
```
ggplot(data, aes(x = state, y = (data$actuals$cases - data$actuals$deaths) / actuals$c
  geom_bar(stat="identity") +
  labs(x="States",y="Recovered (%)") +
  theme_bw() +
  scale_x_discrete(guide = guide_axis(angle = 90))
```



iv. What fraction of the population is currently vaccinated?

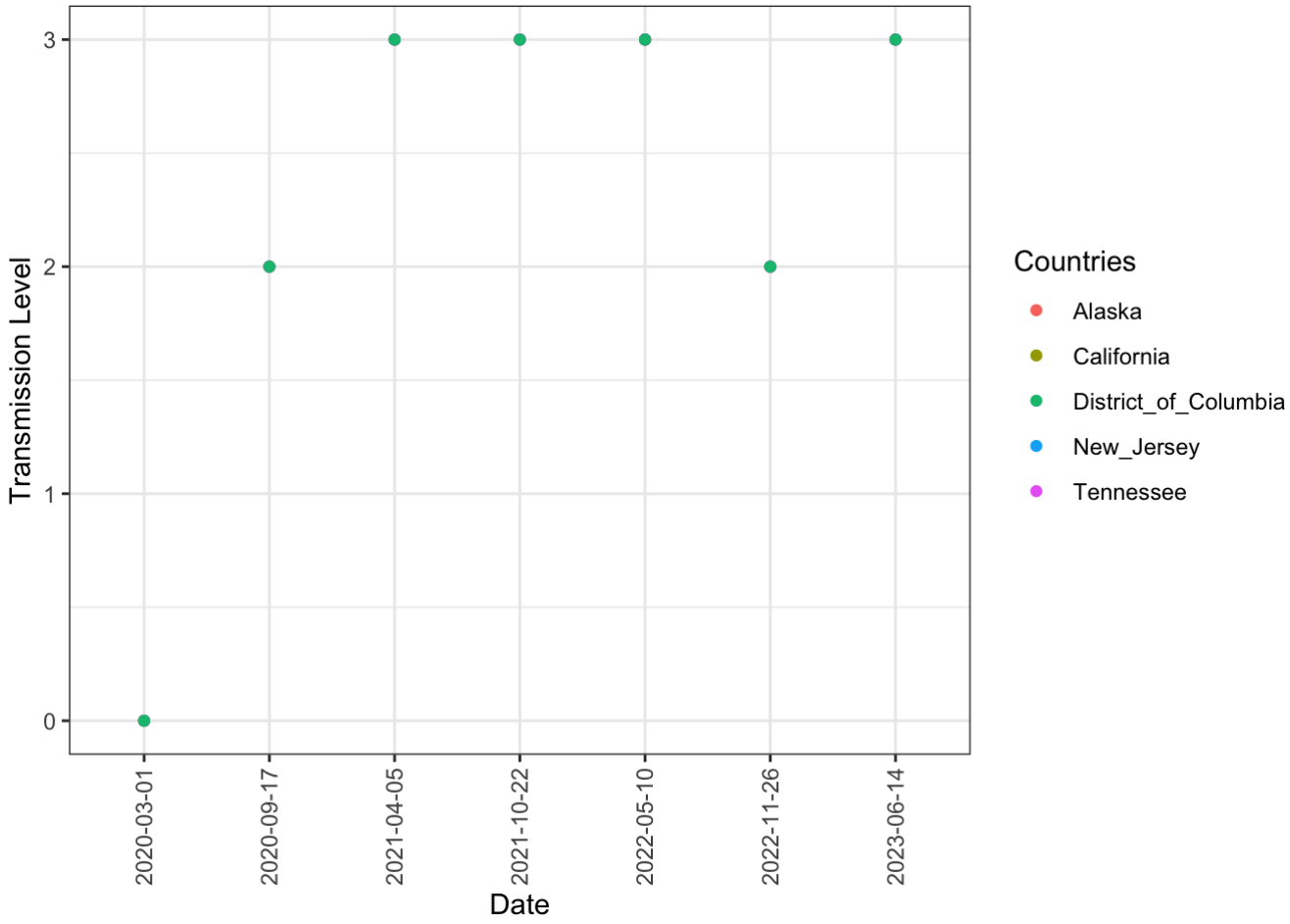
```
ggplot(data, aes(x=state,y=(data$actuals$vaccinationsCompleted/population))) +
  geom_bar(stat="identity") +
  labs(x="States",y="Vaccinated (%)") +
  theme_bw() +
  scale_x_discrete(guide = guide_axis(angle = 90))
```

```
Warning: Removed 3 rows containing missing values (`position_stack()`).
```



v. What was the transmission like in the various states?

```
time_series_transmission[seq(1,1334,by=200),]%>%
pivot_longer(cols=Alaska:District_of_Columbia,names_to="Countries",values_to="Transmis
ggplot(aes(x=Date,y=Transmission,colour=Countries,group=Countries)) +
geom_point(show.legend=TRUE) + labs(x="Date",y="Transmission Level")+theme_bw() + scal
```



vi. How did the disease progress since it started?

```
data_to_plot <- tibble(Date_Alaska = time_series_cases$Alaska$date[seq(1,1300,by=100)]
Cases_Alaska = time_series_cases$Alaska$cases[seq(1,1300,by=100)],
Date_California = time_series_cases$California$date[seq(1,1300,by=100)],
Cases_California = time_series_cases$California$cases[seq(1,1300,by=100)],
Date_New_Jersey = time_series_cases$New_Jersey$date[seq(1,1300,by=100)],
Cases_New_Jersey = time_series_cases$New_Jersey$cases[seq(1,1300,by=100)],
Date_Tennessee = time_series_cases$Tennessee$date[seq(1,1300,by=100)],
Cases_Tennessee = time_series_cases$Tennessee$cases[seq(1,1300,by=100)],
Date_District_of_Columbia = time_series_cases$District_of_Columbia$date[seq(1,1300,by=
Cases_District_of_Columbia = time_series_cases$District_of_Columbia$cases[seq(1,1300,b

data_to_plot
```

```
# A tibble: 13 × 10
   Date_Alaska Cases_Alaska Date_California Cases_California Date_New_Jersey
   <chr>              <int> <chr>                      <int> <chr>
 1 2020-03-01            NA 2020-01-25                     1 2020-03-01
 2 2020-06-09           620 2020-05-04                 56333 2020-06-09
 3 2020-09-17          7413 2020-08-12                595097 2020-09-17
 4 2020-12-26         45247 2020-11-20               1096427 2020-12-26
 5 2021-04-05         63486 2021-02-28               3569578 2021-04-05
 6 2021-07-14         71539 2021-06-08               3798225 2021-07-14
 7 2021-10-22        132393 2021-09-16               4629146 2021-10-22
 8 2022-01-30        211117 2021-12-25               5291605 2022-01-30
 9 2022-05-10        252847 2022-04-04               9110544 2022-05-10
```

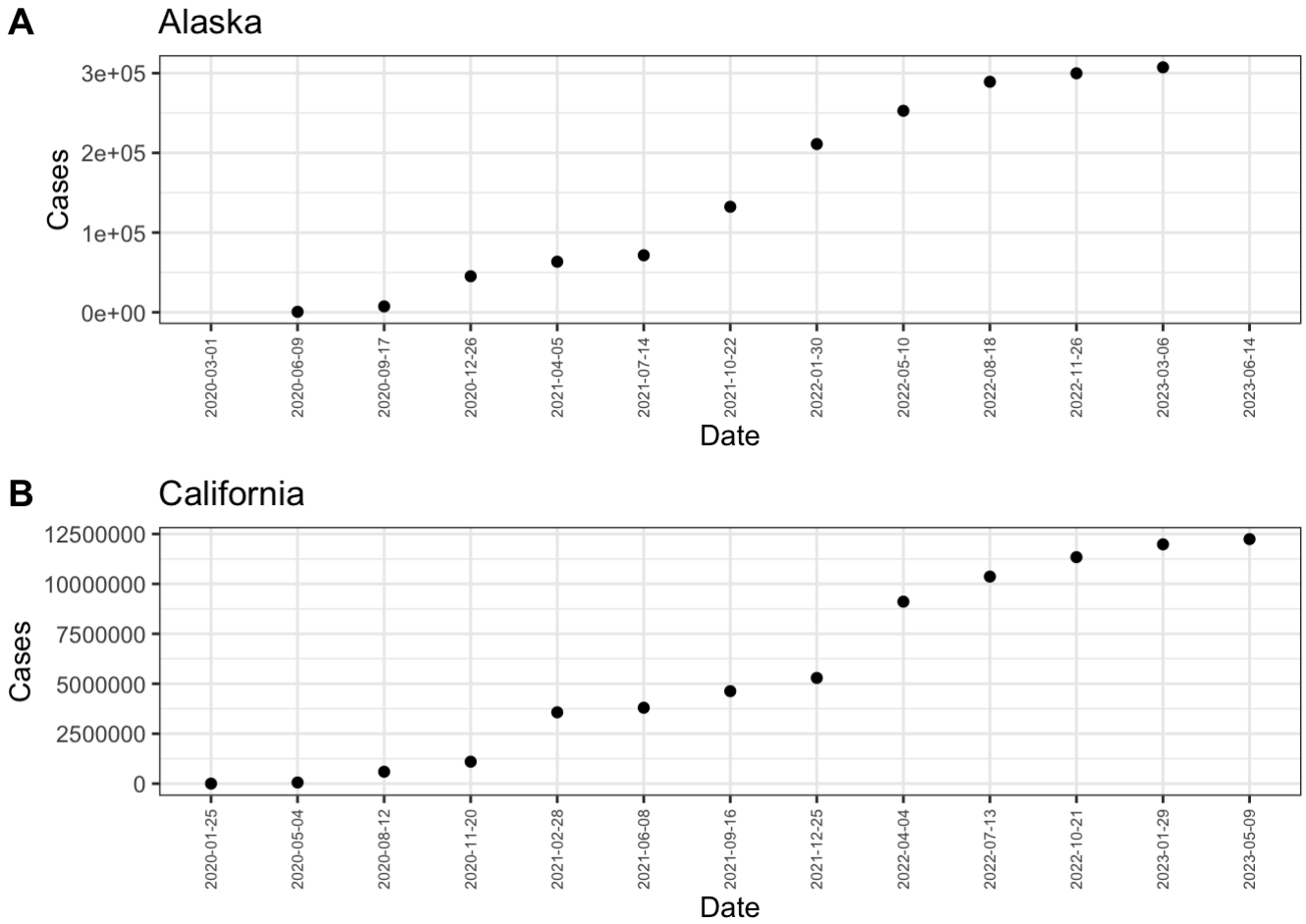
```
10 2022-08-18        289203 2022-07-13             10365785 2022-08-18
11 2022-11-26        299841 2022-10-21             11338846 2022-11-26
12 2023-03-06        307377 2023-01-29             11980312 2023-03-06
13 2023-06-14            NA 2023-05-09             12242634 2023-06-14
# ℹ 5 more variables: Cases_New_Jersey <int>, Date_Tennessee <chr>,
#   Cases_Tennessee <int>, Date_District_of_Columbia <chr>,
#   Cases_District_of_Columbia <int>
```

```
library(cowplot)
```

```
Attaching package: 'cowplot'

The following object is masked from 'package:lubridate':

    stamp
```

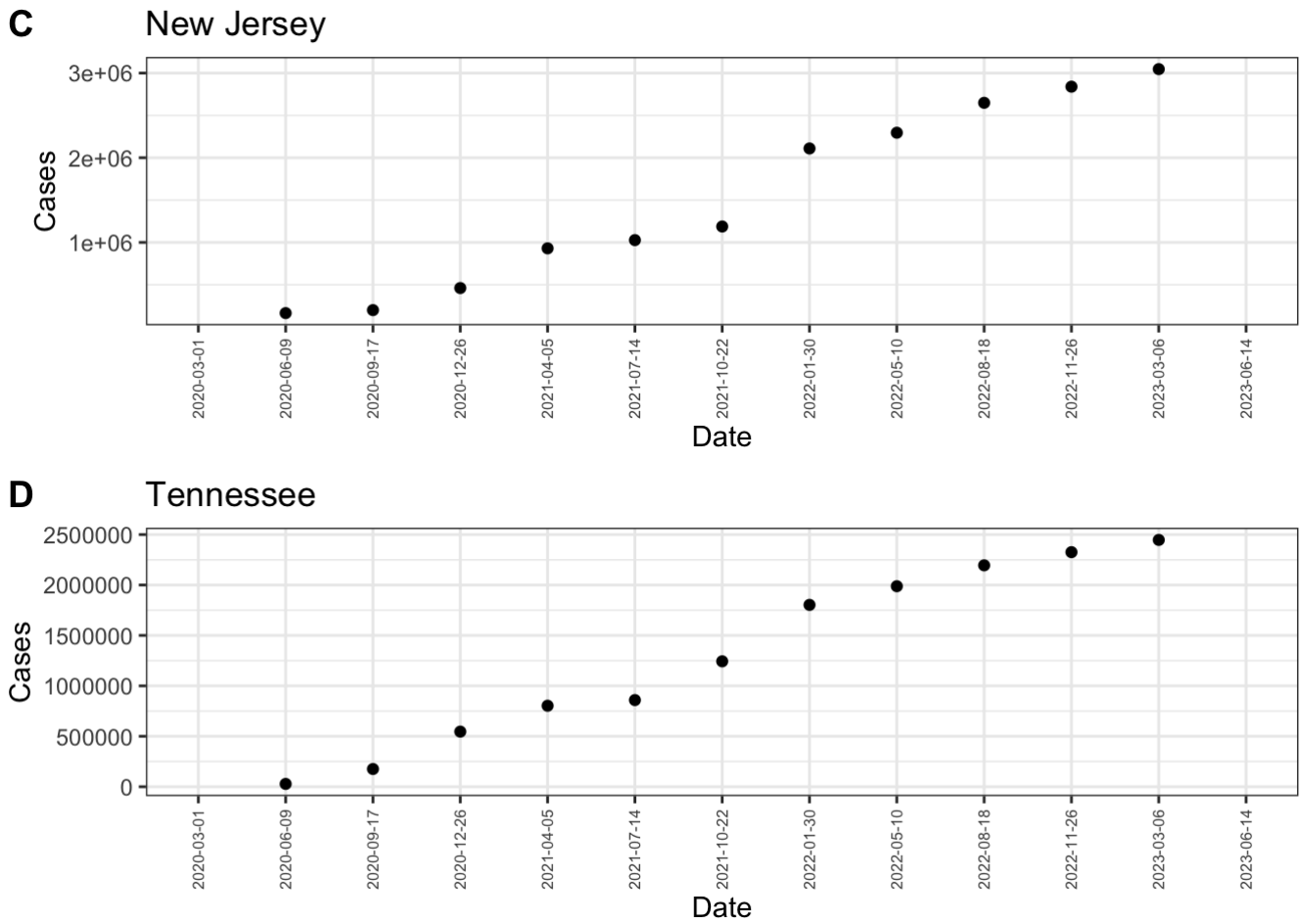
```
fig1<- ggplot(data_to_plot, aes(x=Date_Alaska,y=Cases_Alaska)) +
geom_point() + labs(x="Date",y="Cases", title="Alaska") + theme_bw() + scale_x_discret
fig2<- ggplot(data_to_plot, aes(x=Date_California,y=Cases_California)) +
geom_point() + labs(x="Date",y="Cases", title="California") + theme_bw() + scale_x_dis
fig3<- ggplot(data_to_plot, aes(x=Date_New_Jersey,y=Cases_New_Jersey)) +
geom_point() + labs(x="Date",y="Cases", title="New Jersey") + theme_bw() + scale_x_dis
fig4<- ggplot(data_to_plot, aes(x=Date_Tennessee,y=Cases_Tennessee)) +
geom_point() + labs(x="Date",y="Cases", title="Tennessee") + theme_bw() + scale_x_disc
fig5<- ggplot(data_to_plot, aes(x=Date_District_of_Columbia,y=Cases_District_of_Columb
geom_point() + labs(x="Date",y="Cases", title="District of Columbia") + theme_bw() + s

plot_grid(
fig1 + theme(legend.justification = c(0,1), axis.text.x = element_text(size = 6)),
fig2 + theme(legend.justification = c(1,0), axis.text.x = element_text(size = 6)),
fig3 + theme(legend.justification = c(0,1), axis.text.x = element_text(size = 6)),
fig4 + theme(legend.justification = c(1,0), axis.text.x = element_text(size = 6)),
fig5 + theme(legend.justification = c(0,1), axis.text.x = element_text(size = 6)),
align = "v", axis = "lr",
nrow = 2, ncol = 1,
labels = LETTERS[1:5],
rel_heights = c(1, 1))
```

```
Warning: Removed 2 rows containing missing values (`geom_point()`).
Removed 2 rows containing missing values (`geom_point()`).
Removed 2 rows containing missing values (`geom_point()`).
Removed 2 rows containing missing values (`geom_point()`).
```




```
plot_grid(
fig3 + theme(legend.justification = c(0,1), axis.text.x = element_text(size = 6)),
fig4 + theme(legend.justification = c(1,0), axis.text.x = element_text(size = 6)),
align = "v", axis = "lr",
nrow = 2, ncol = 1,
labels = LETTERS[3:4],
rel_heights = c(1, 1))
```

```
Warning: Removed 2 rows containing missing values (`geom_point()`).
Removed 2 rows containing missing values (`geom_point()`).
```




```
plot_grid(
fig5 + theme(legend.justification = c(0,1), axis.text.x = element_text(size = 6)),
align = "v", axis = "lr",
nrow = 2, ncol = 1,
labels = LETTERS[5],
rel_heights = c(1, 1))
```

```
Warning: Removed 2 rows containing missing values (`geom_point()`).
```

E District of Columbia