ASSIGNMENT 4

Stephen Smitshoek

2022-04-18

Markdown Basics

Favorite Foods

1. Pizza
2. Cous Cous
3. Nachos

Images

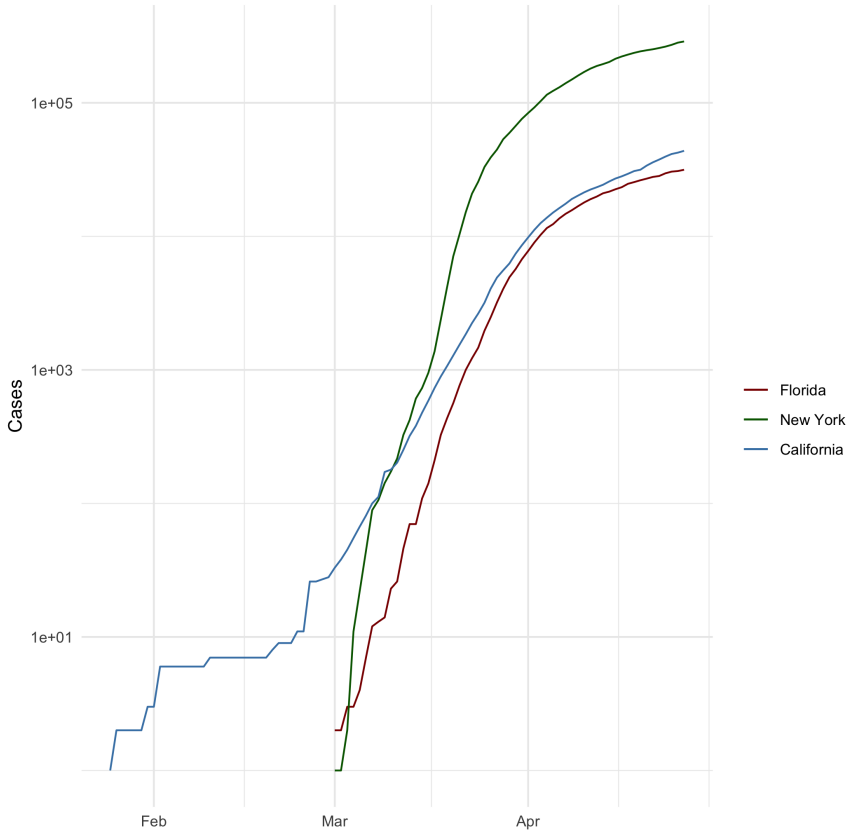


Figure 1: All Cases (Log Plot)

Add a Quote

> Tis but a flesh wound

Add an Equation

$$PV = nRT$$

Add a Footnote

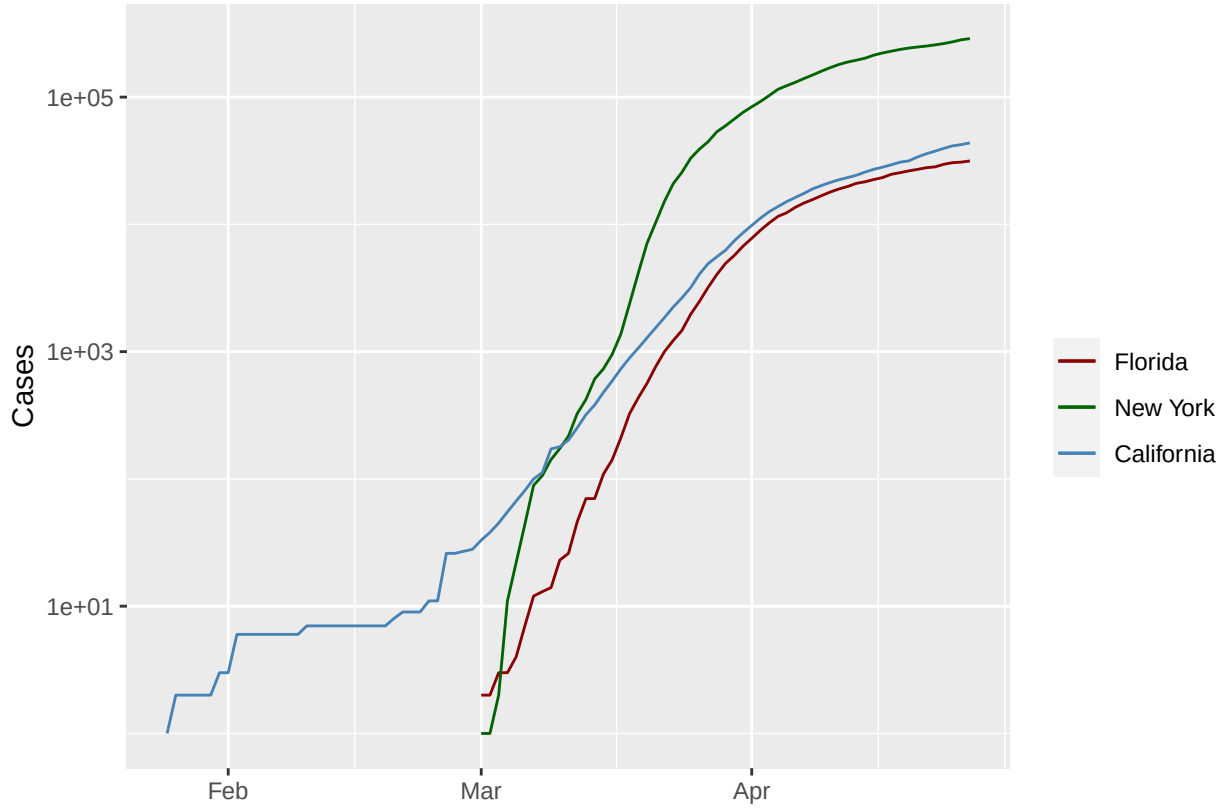
[1]

Add Citations

- Lander, Jared (2021). *R for Everyone*. Addison-Wesley.
- Field, Andy (2012). *Discovering Statistics Using R*. SAGE Publications Inc.

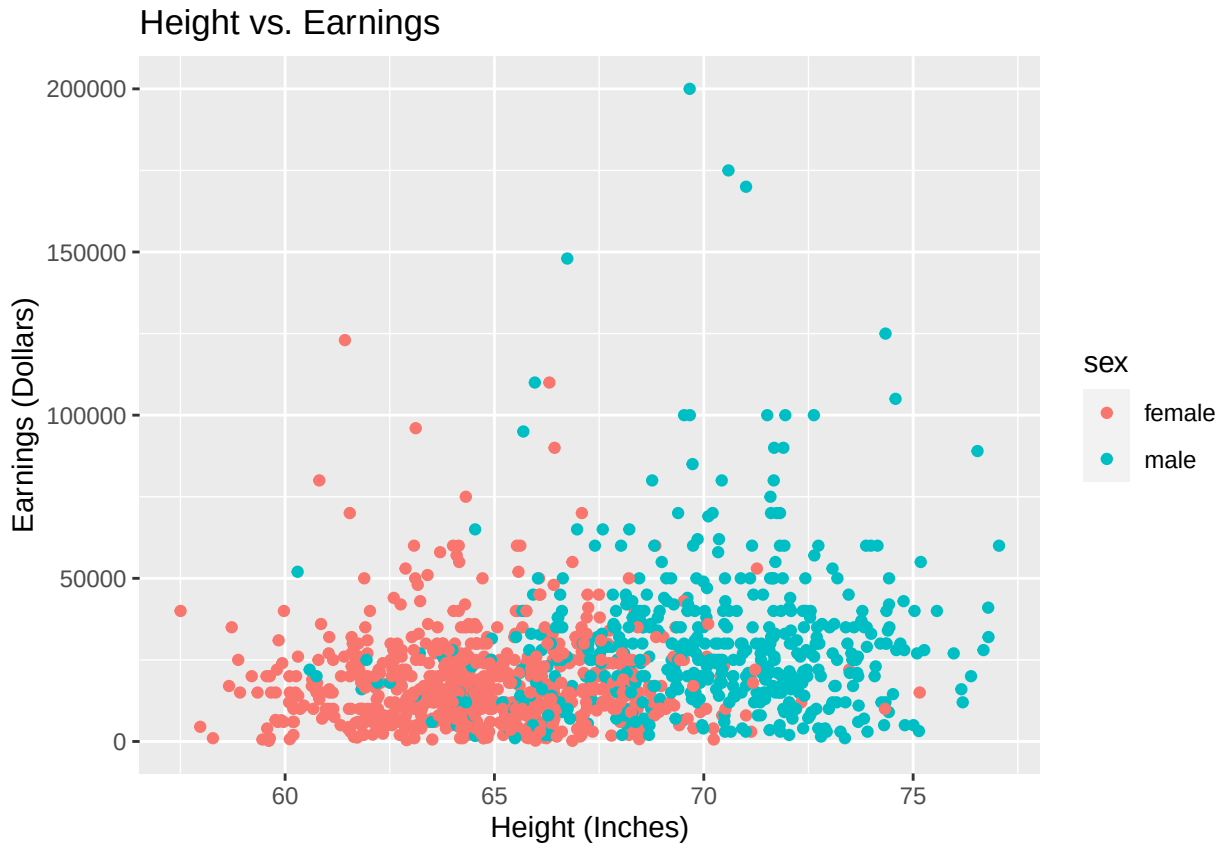
Inline Code

NY Times COVID-19 Data



[1] This is a footnote

R4DS Height vs Earnings



Tables

Knitr Table with Kable

Table 1: One Ring to Rule Them All

name	race	in_fellowship	ring_bearer	age
Aragon	Men	TRUE	FALSE	88
Bilbo	Hobbit	FALSE	TRUE	129
Frodo	Hobbit	TRUE	TRUE	51
Galadriel	Elf	FALSE	FALSE	7000
Sam	Hobbit	TRUE	TRUE	36
Gandalf	Maia	TRUE	TRUE	2019
Legolas	Elf	TRUE	FALSE	2931
Sauron	Maia	FALSE	TRUE	7052
Gollum	Hobbit	FALSE	TRUE	589

Pandoc Table

```
## Warning: package 'pander' was built under R version 4.1.3
```

```
##
##
## +-----------+--------+---------------+-------------+------+
## |   name    |  race  | in_fellowship | ring_bearer | age  |
## +===========+========+===============+=============+======+
## |  Aragon   |  Men   |     TRUE      |    FALSE    |  88  |
## +-----------+--------+---------------+-------------+------+
## |   Bilbo   | Hobbit |     FALSE     |    TRUE     | 129  |
## +-----------+--------+---------------+-------------+------+
## |   Frodo   | Hobbit |     TRUE      |    TRUE     |  51  |
## +-----------+--------+---------------+-------------+------+
## | Galadriel |  Elf   |     FALSE     |    FALSE    | 7000 |
## +-----------+--------+---------------+-------------+------+
## |    Sam    | Hobbit |     TRUE      |    TRUE     |  36  |
## +-----------+--------+---------------+-------------+------+
## |  Gandalf  |  Maia  |     TRUE      |    TRUE     | 2019 |
## +-----------+--------+---------------+-------------+------+
## |  Legolas  |  Elf   |     TRUE      |    FALSE    | 2931 |
## +-----------+--------+---------------+-------------+------+
## |  Sauron   |  Maia  |     FALSE     |    TRUE     | 7052 |
## +-----------+--------+---------------+-------------+------+
## |  Gollum   | Hobbit |     FALSE     |    TRUE     | 589  |
## +-----------+--------+---------------+-------------+------+
```

References

Assignment 04 Code

````
---
title: "ASSIGNMENT 4"
author: "Stephen Smitshoek"
date: '2022-04-20'
output:
  pdf_document: default
  html_document: default
  word_document: default
bibliography: bibliography.bib
---

# Markdown Basics

## Favorite Foods
1. Pizza
1. Cous Cous
1. Nachos

## Images
![All Cases (Log Plot)](C:/Users/sksmi/PeytoAccess/Personal/Bellevue/DSC520/
dsc520/completed/assignment04/plots/10-all-cases-log.png){height=50%}

## Add a Quote
> Tis but a flesh wound

## Add an Equation
$$
PV=nRT
$$

## Add a Footnote

^[This is a footnote]

## Add Citations

* Lander, Jared (2021). _R for Everyone_. Addison-Wesley.
* Field, Andy (2012). _Discovering Statistics Using R_. SAGE Publications
Inc.

# Inline Code
```{r include=FALSE}
setwd('C:/Users/sksmi/PeytoAccess/Personal/Bellevue/DSC520/dsc520')
library(ggplot2)
covid_df <- read.csv("data/nytimes/covid-19-data/us-states.csv")
covid_df$date <- as.Date(covid_df$date)
california_df <- covid_df[ which( covid_df$state == "California"), ]
ny_df <- covid_df[ which( covid_df$state == "New York"), ]
florida_df <- covid_df[ which( covid_df$state == "Florida"), ]
```

## NY Times COVID-19 Data
```{r echo=FALSE}
ggplot(data=florida_df, aes(x=date, group=1)) +
  geom_line(aes(y = cases, colour = "Florida")) +
````
````

```
  geom_line(data=ny_df, aes(y = cases,colour="New York")) +
  geom_line(data=california_df, aes(y = cases, colour="California")) +
  scale_colour_manual("",
                      breaks = c("Florida", "New York", "California"),
                      values = c('darkred', 'darkgreen', 'steelblue')) +
  xlab(" ") + ylab("Cases") + scale_y_log10()
```

R4DS Height vs Earnings
```{r echo=FALSE}
setwd('C:/Users/sksmi/PeytoAccess/Personal/Bellevue/DSC520/dsc520')
library(ggplot2)
heights_df <- read.csv("data\\r4ds\\heights.csv")
ggplot(heights_df, aes(x=height, y=earn, col=sex)) + geom_point() +
ggtitle("Height vs. Earnings") + xlab("Height (Inches)") + ylab("Earnings
(Dollars)")
```

Tables

Knitr Table with Kable
```{r echo=FALSE}
name <- c("Aragon", "Bilbo", "Frodo", "Galadriel", "Sam", "Gandalf",
"Legolas", "Sauron", "Gollum")
race <- c("Men", "Hobbit", "Hobbit", "Elf", "Hobbit", "Maia", "Elf", "Maia",
"Hobbit")
in_fellowship <- c(TRUE, FALSE, TRUE, FALSE, TRUE, TRUE, TRUE, FALSE, FALSE)
ring_bearer <- c(FALSE, TRUE, TRUE, FALSE, TRUE, TRUE, FALSE, TRUE, TRUE)
age <- c(88, 129, 51, 7000, 36, 2019, 2931, 7052, 589)

characters_df <- data.frame(name, race, in_fellowship, ring_bearer, age)


knitr::kable(characters_df, caption = 'One Ring to Rule Them All')
```
Pandoc Table
```{r echo=FALSE}
library(pander)
pandoc.table(characters_df, style='grid')
```

References

Assignment 04 Backup

```
# Assignment: ASSIGNMENT 4
# Name: Smitshoek, Stephen
# Date: 2022-04-18

## Load the ggplot2 package
library(ggplot2)
theme_set(theme_minimal())

## Load the `data/r4ds/heights.csv` to
heights_df <- read.csv("data/r4ds/heights.csv")

# https://ggplot2.tidyverse.org/reference/geom_boxplot.html
## Create boxplots of sex vs. earn and race vs. earn using `geom_point()` and
`geom_boxplot()`
## sex vs. earn
ggplot(heights_df, aes(x=sex, y=earn)) + geom_point() + geom_boxplot()
## race vs. earn
ggplot(heights_df, aes(x=race, y=earn)) + geom_point() + geom_boxplot()

# https://ggplot2.tidyverse.org/reference/geom_bar.html
## Using `geom_bar()` plot a bar chart of the number of records for each
`sex`
ggplot(heights_df, aes(sex)) + geom_bar()

## Using `geom_bar()` plot a bar chart of the number of records for each race
ggplot(heights_df, aes(race)) + geom_bar()

## Create a horizontal bar chart by adding `coord_flip()` to the previous
plot
ggplot(heights_df, aes(race)) + geom_bar() + coord_flip()

# https://www.rdocumentation.org/packages/ggplot2/versions/3.3.0/topics/
geom_path
## Load the file `"data/nytimes/covid-19-data/us-states.csv"` and
## assign it to the `covid_df` dataframe
covid_df <- read.csv("data/nytimes/covid-19-data/us-states.csv")

## Parse the date column using `as.Date()``
covid_df$date <- as.Date(covid_df$date)

## Create three dataframes named `california_df`, `ny_df`, and `florida_df`
## containing the data from California, New York, and Florida
california_df <- covid_df[ which( covid_df$state == "California"), ]
ny_df <- covid_df[ which( covid_df$state == "New York"), ]
florida_df <- covid_df[ which( covid_df$state == "Florida"), ]

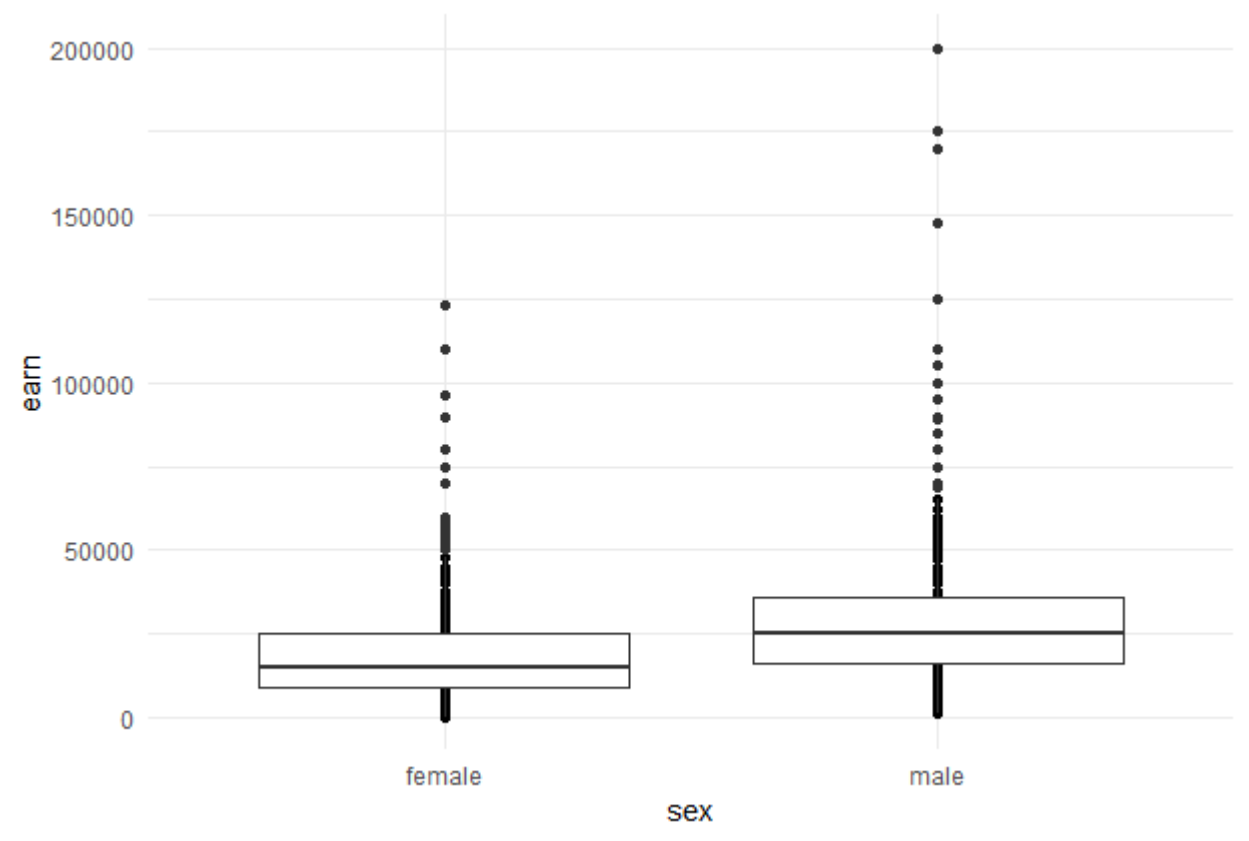
## Plot the number of cases in Florida using `geom_line()`
ggplot(data=florida_df, aes(x=date, y=cases, group=1)) + geom_line()

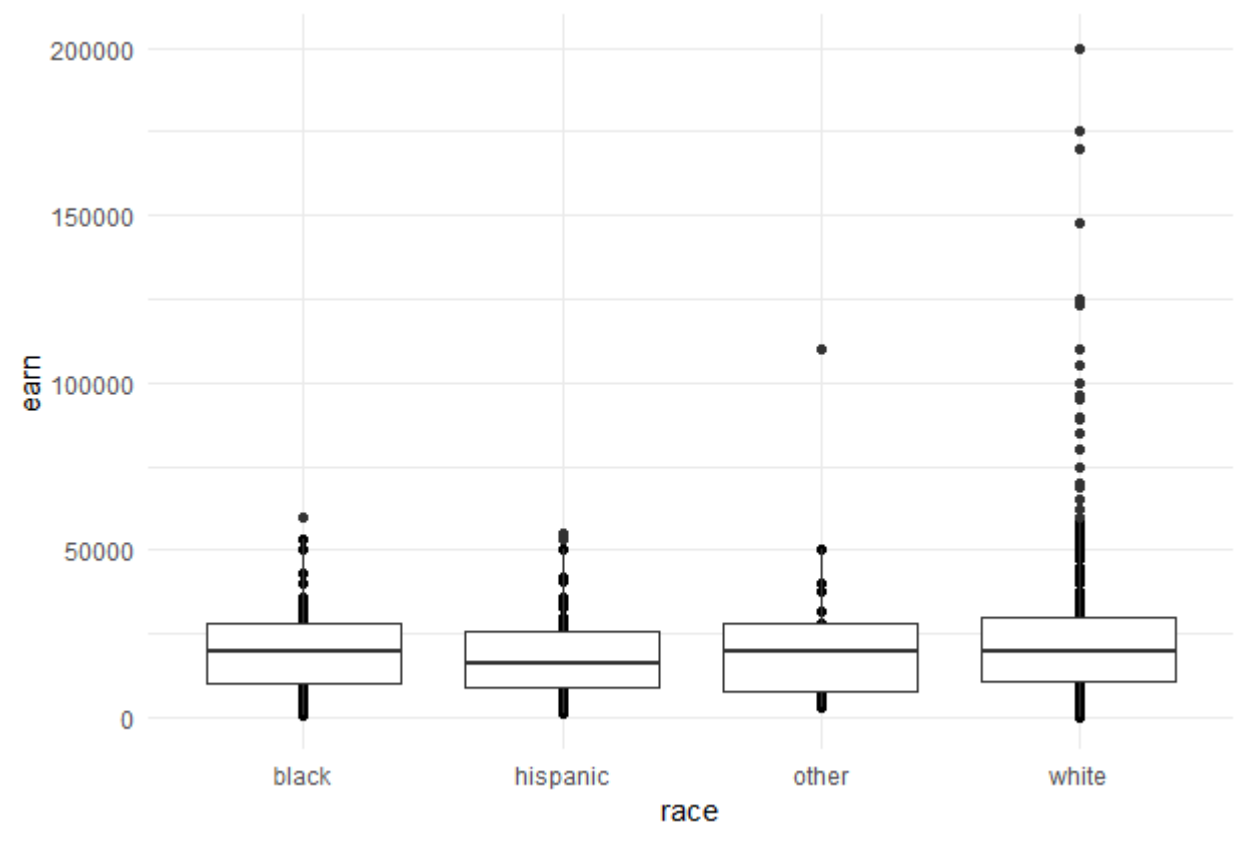
## Add lines for New York and California to the plot
ggplot(data=florida_df, aes(x=date, group=1)) +
  geom_line(aes(y = cases)) +
  geom_line(data=ny_df, aes(y = cases)) +
  geom_line(data=california_df, aes(y = cases))
```

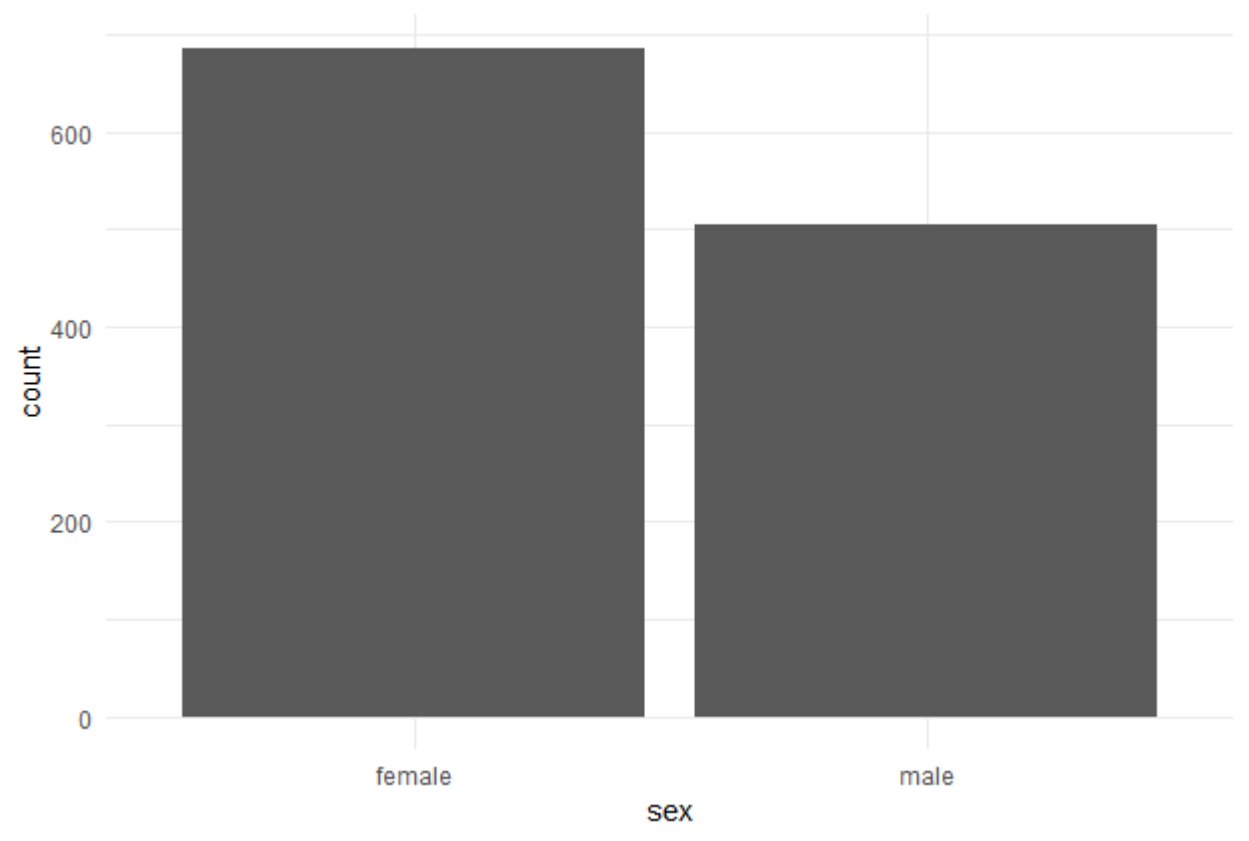
```
## Use the colors "darkred", "darkgreen", and "steelblue" for Florida, New
York, and California
ggplot(data=florida_df, aes(x=date, group=1)) +
  geom_line(aes(y = cases), color = 'darkred') +
  geom_line(data=ny_df, aes(y = cases), color='darkgreen') +
  geom_line(data=california_df, aes(y = cases), color='steelblue')

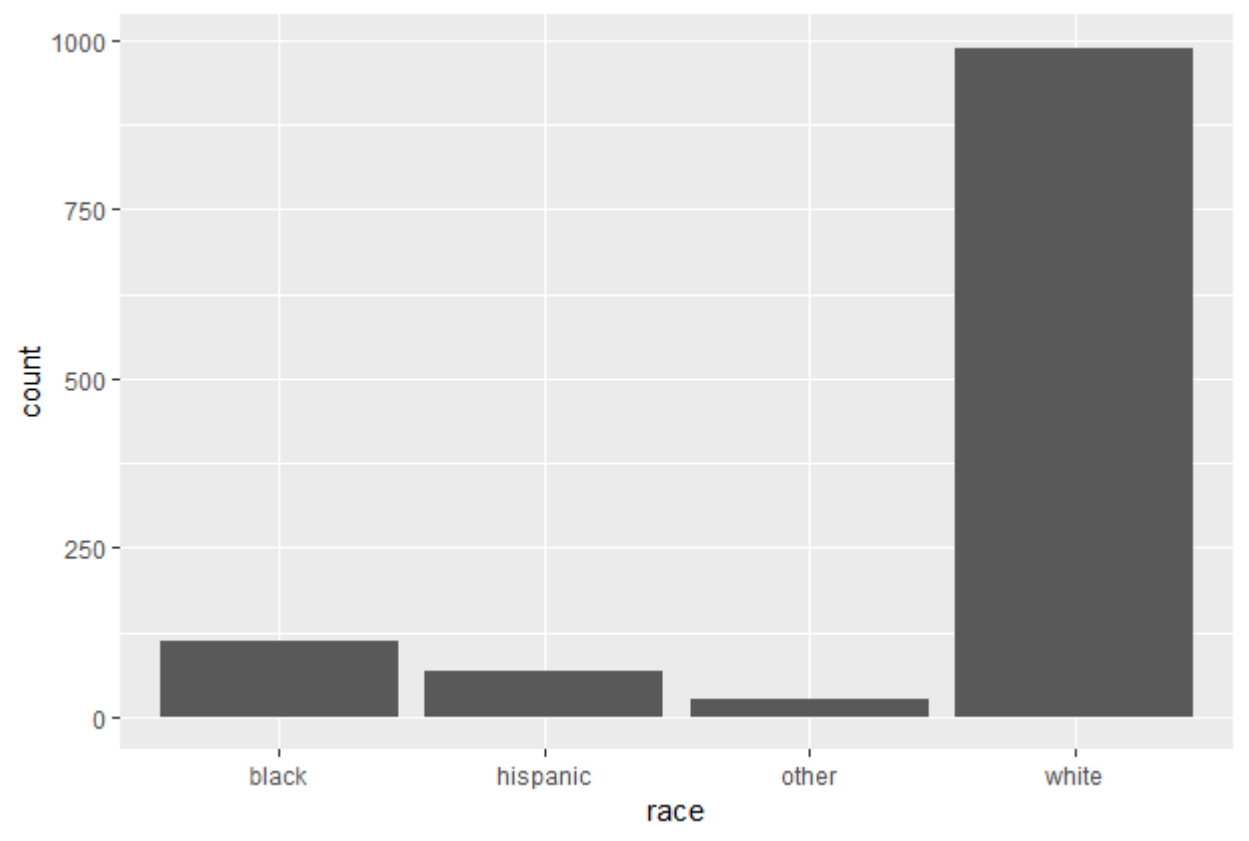
## Add a legend to the plot using `scale_colour_manual`
## Add a blank (" ") label to the x-axis and the label "Cases" to the y axis
ggplot(data=florida_df, aes(x=date, group=1)) +
  geom_line(aes(y = cases, colour = "Florida")) +
  geom_line(data=ny_df, aes(y = cases,colour="New York")) +
  geom_line(data=california_df, aes(y = cases, colour="California")) +
  scale_colour_manual("",
                      breaks = c("Florida", "New York", "California"),
                      values = c('darkred', 'darkgreen', 'steelblue')) +
  xlab(" ") + ylab("Cases")

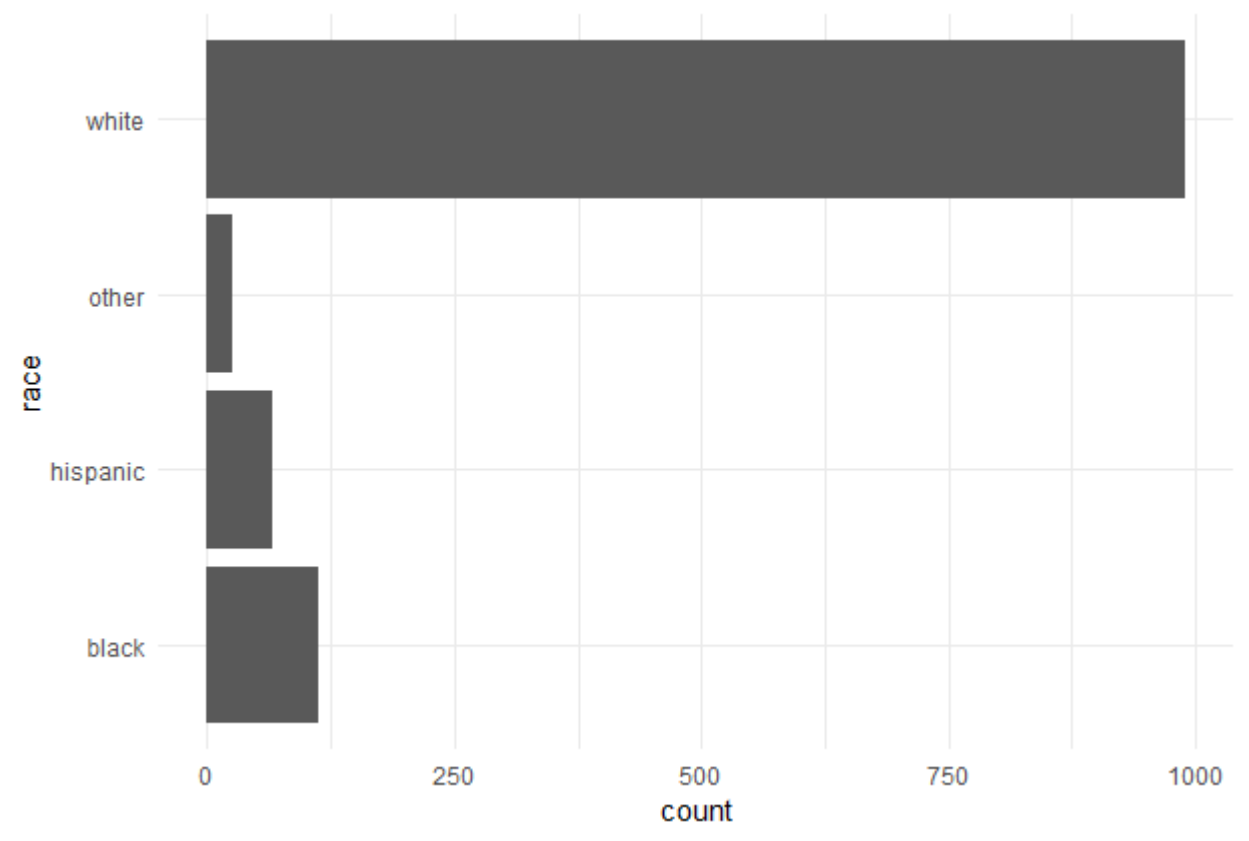
## Scale the y axis using `scale_y_log10()`
ggplot(data=florida_df, aes(x=date, group=1)) +
  geom_line(aes(y = cases, colour = "Florida")) +
  geom_line(data=ny_df, aes(y = cases,colour="New York")) +
  geom_line(data=california_df, aes(y = cases, colour="California")) +
  scale_colour_manual("",
                      breaks = c("Florida", "New York", "California"),
                      values = c('darkred', 'darkgreen', 'steelblue')) +
  xlab(" ") + ylab("Cases") + scale_y_log10()
```

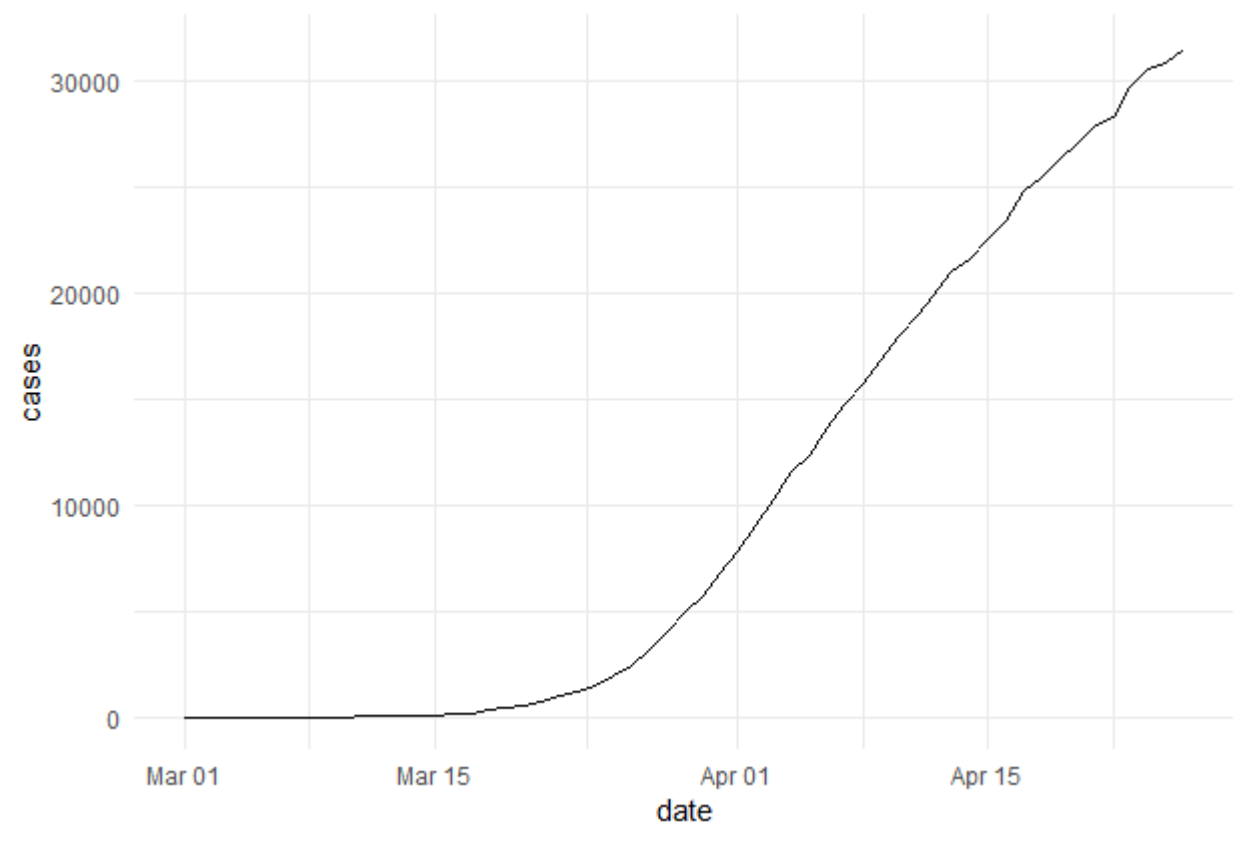

200000
150000
100000
50000
0
earn
female
male
sex

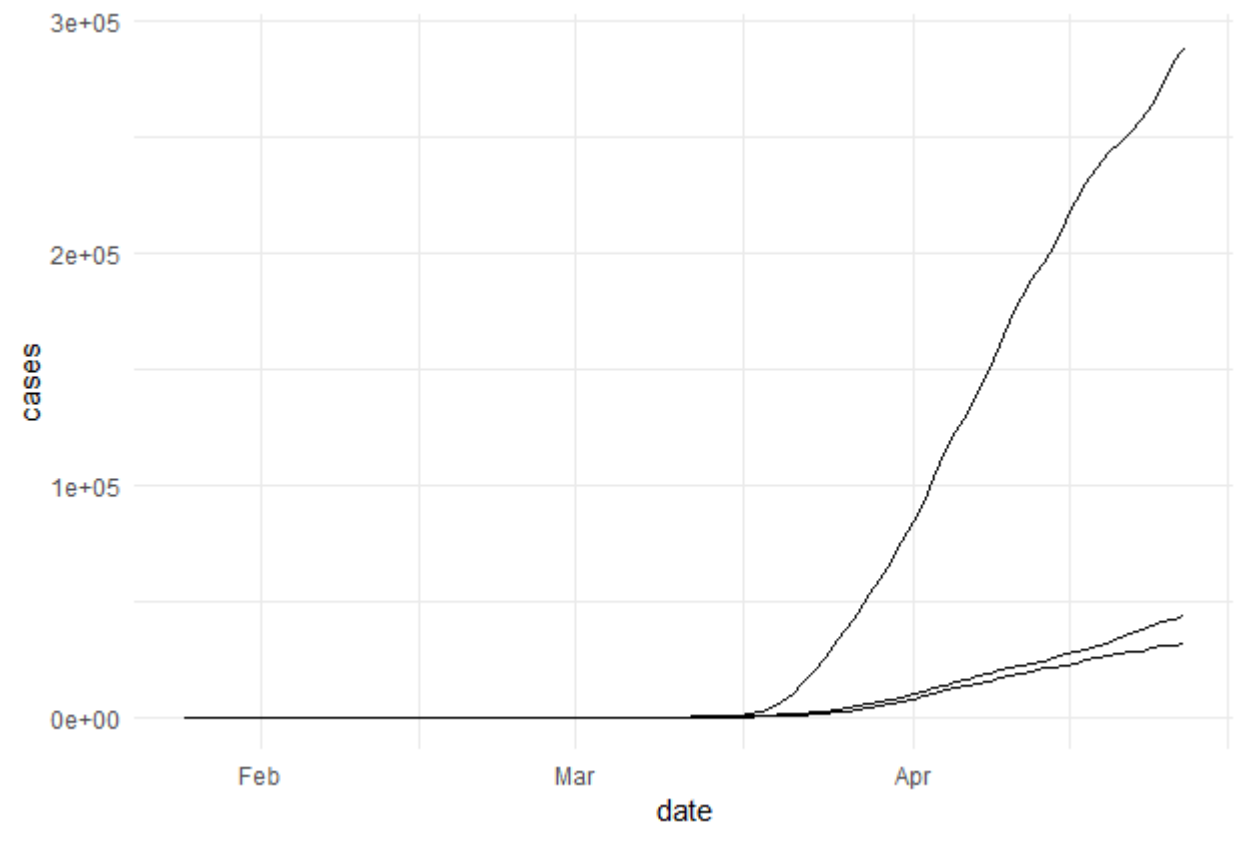

200000
150000
100000
50000
0
earn
black
hispanic
other
white
race

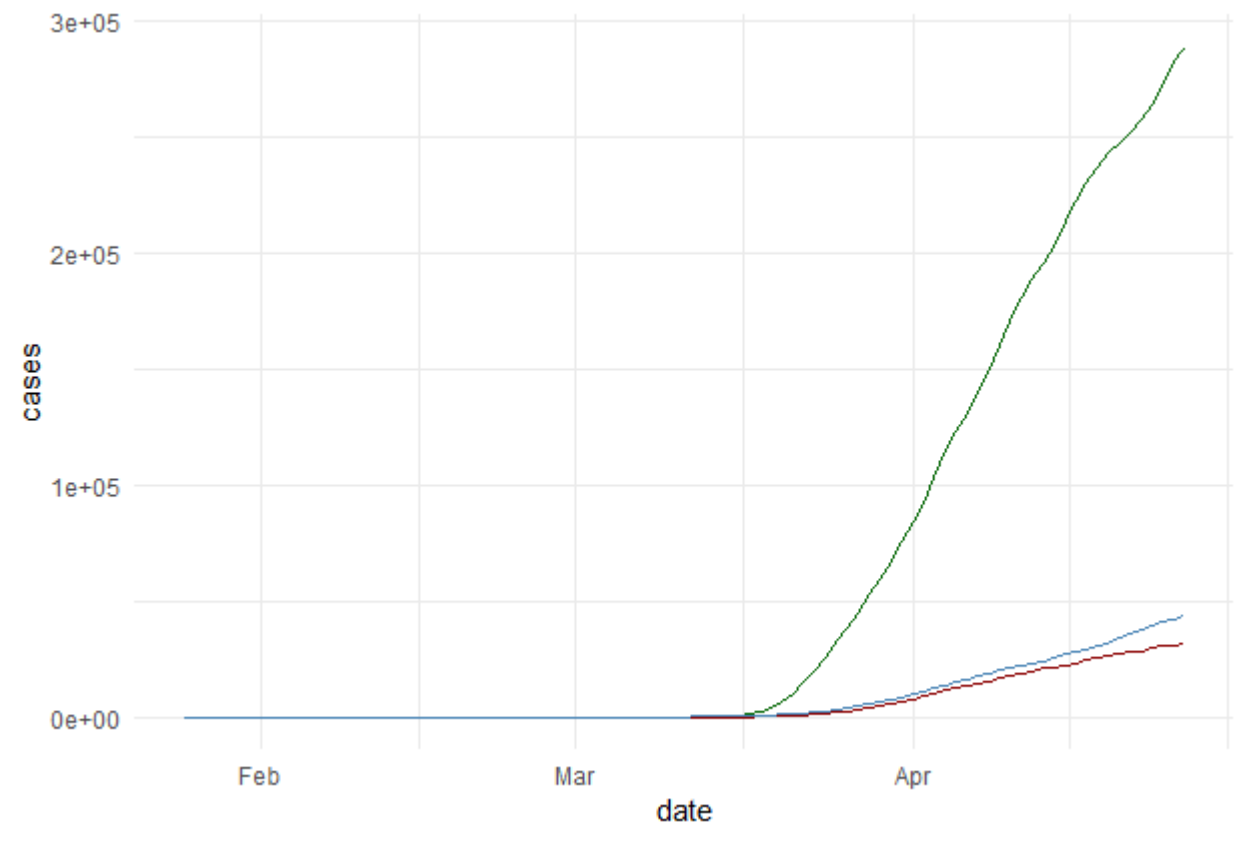

600
400
200
0
count
female
male
sex

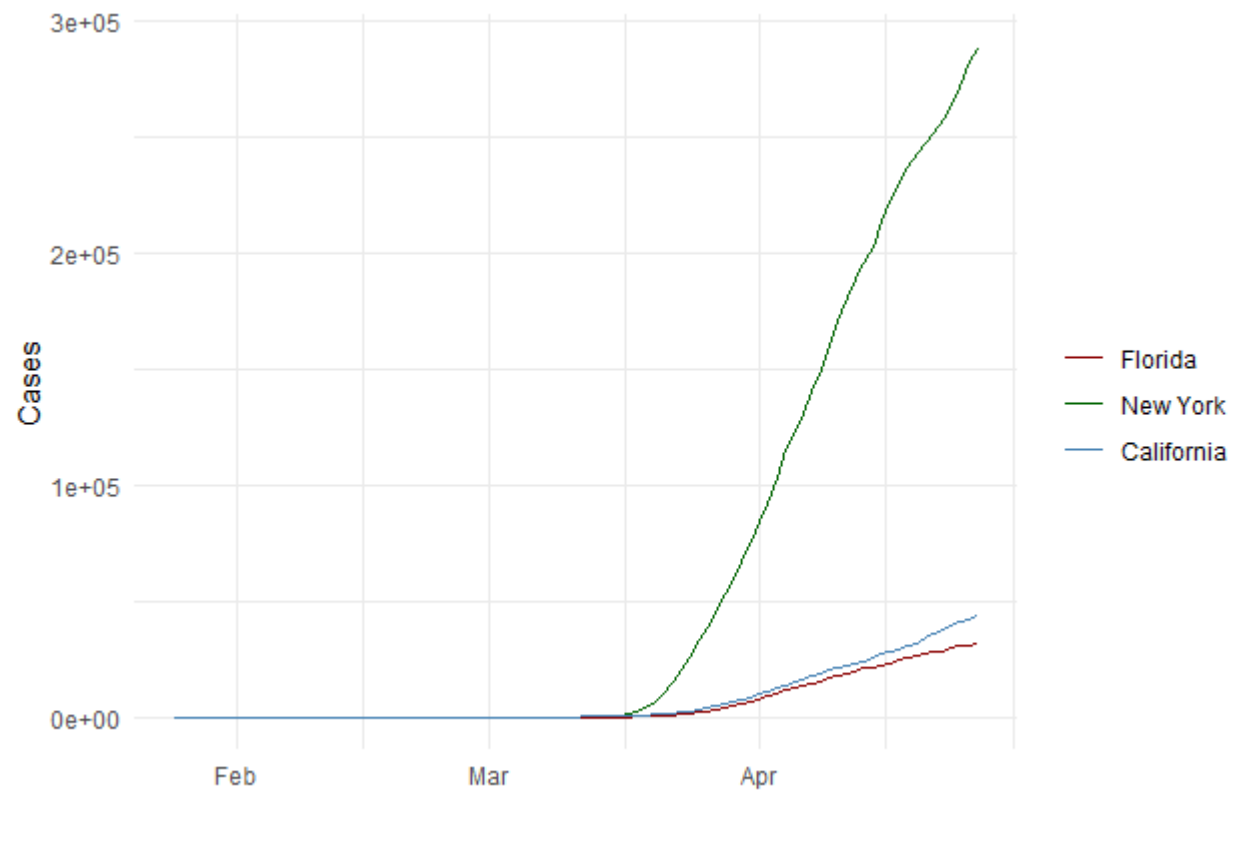

1000
750
500
250
0
count
black
hispanic
other
white
race

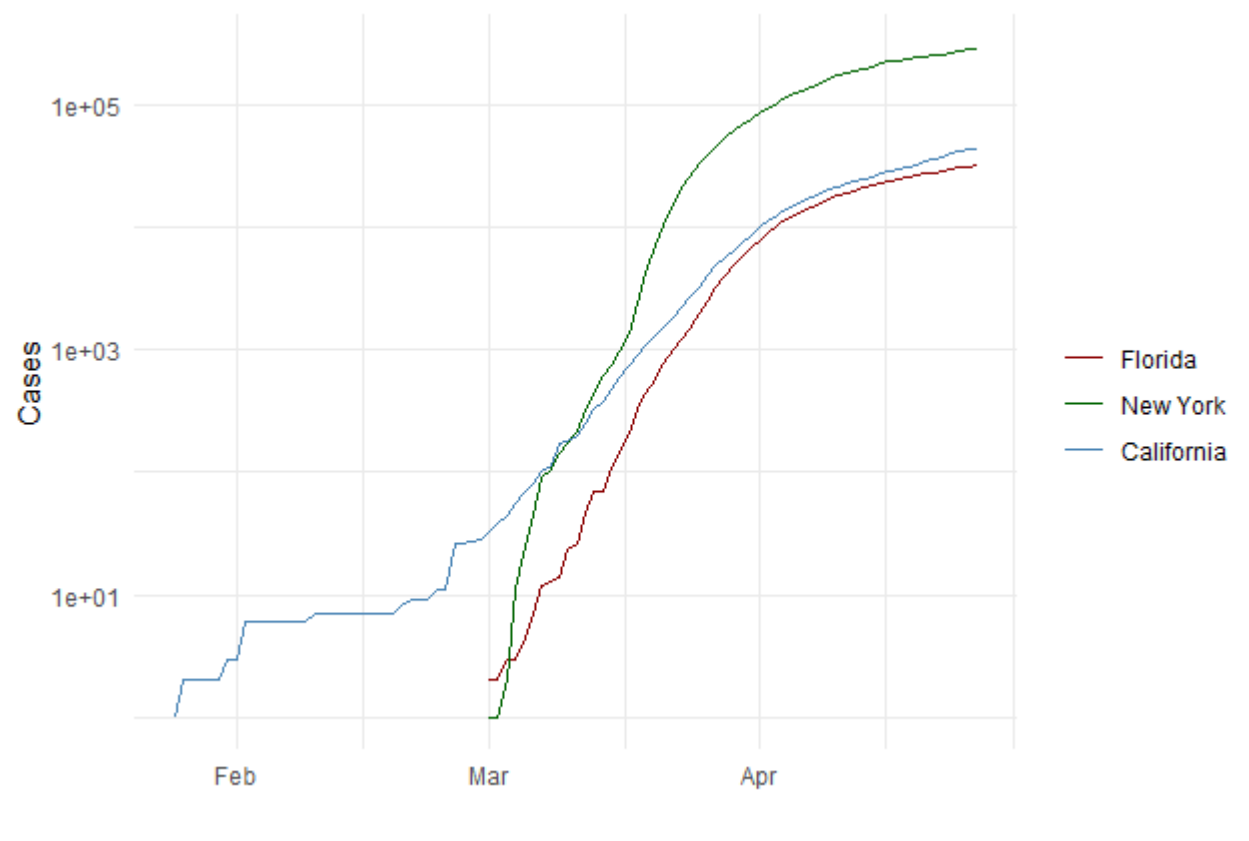

white
other
hispanic
black
race
0
250
500
750
1000
count


30000
20000
10000
0
cases
Mar 01
Mar 15
Apr 01
Apr 15
date


3e+05
2e+05
1e+05
0e+00
cases
Feb
Mar
Apr
date


3e+05
2e+05
1e+05
0e+00
cases
Feb
Mar
Apr
date


3e+05
2e+05
1e+05
0e+00
Cases
Feb
Mar
Apr
Florida
New York
California


1e+05
1e+03
1e+01
Cases
Feb
Mar
Apr
Florida
New York
California