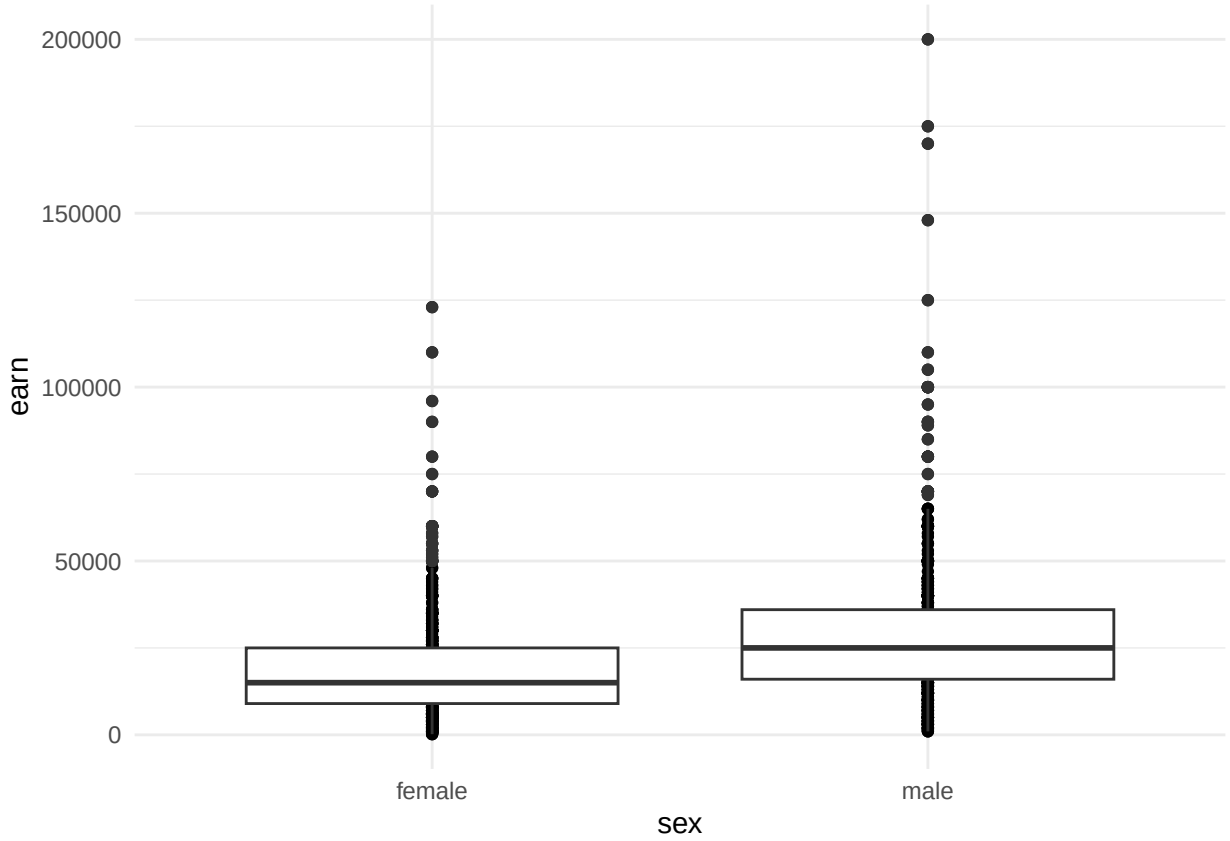
6.2Exercises.R

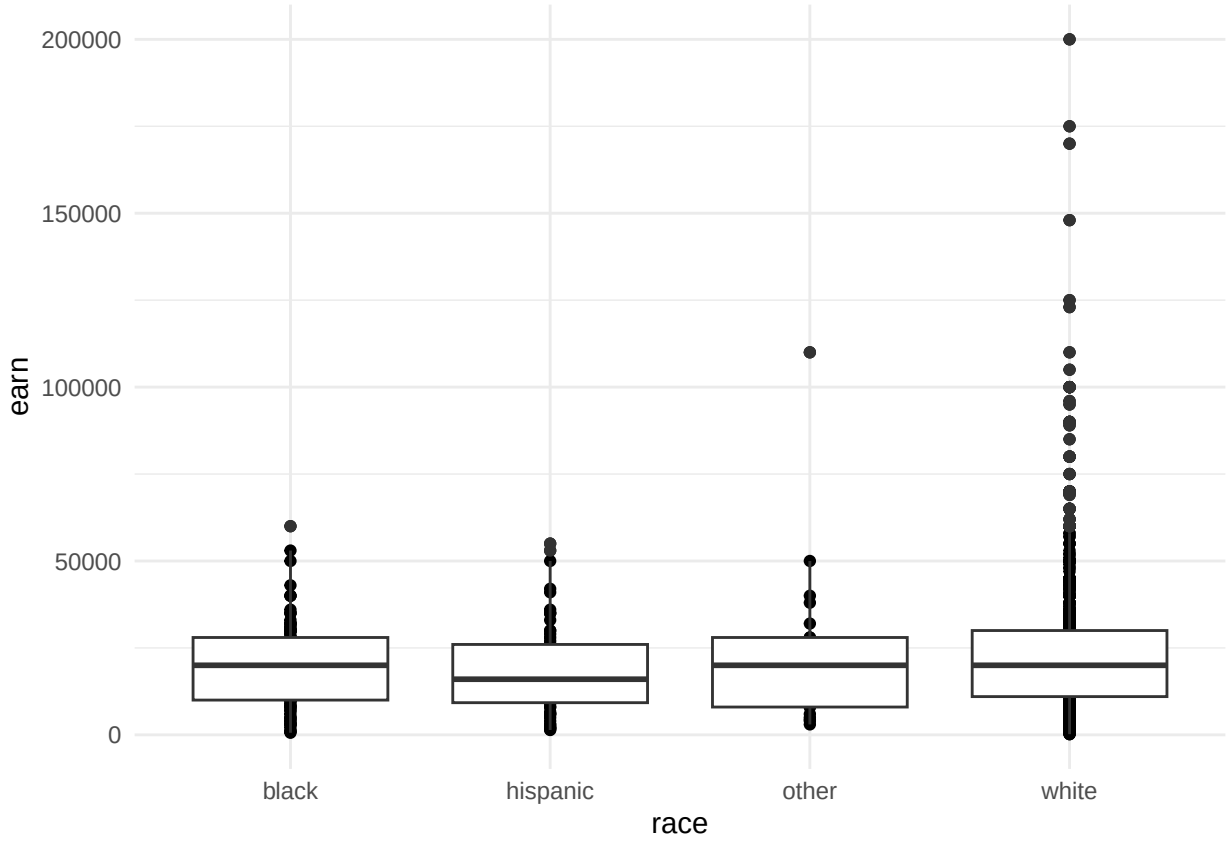
Rahul Rajeev

2023-01-17

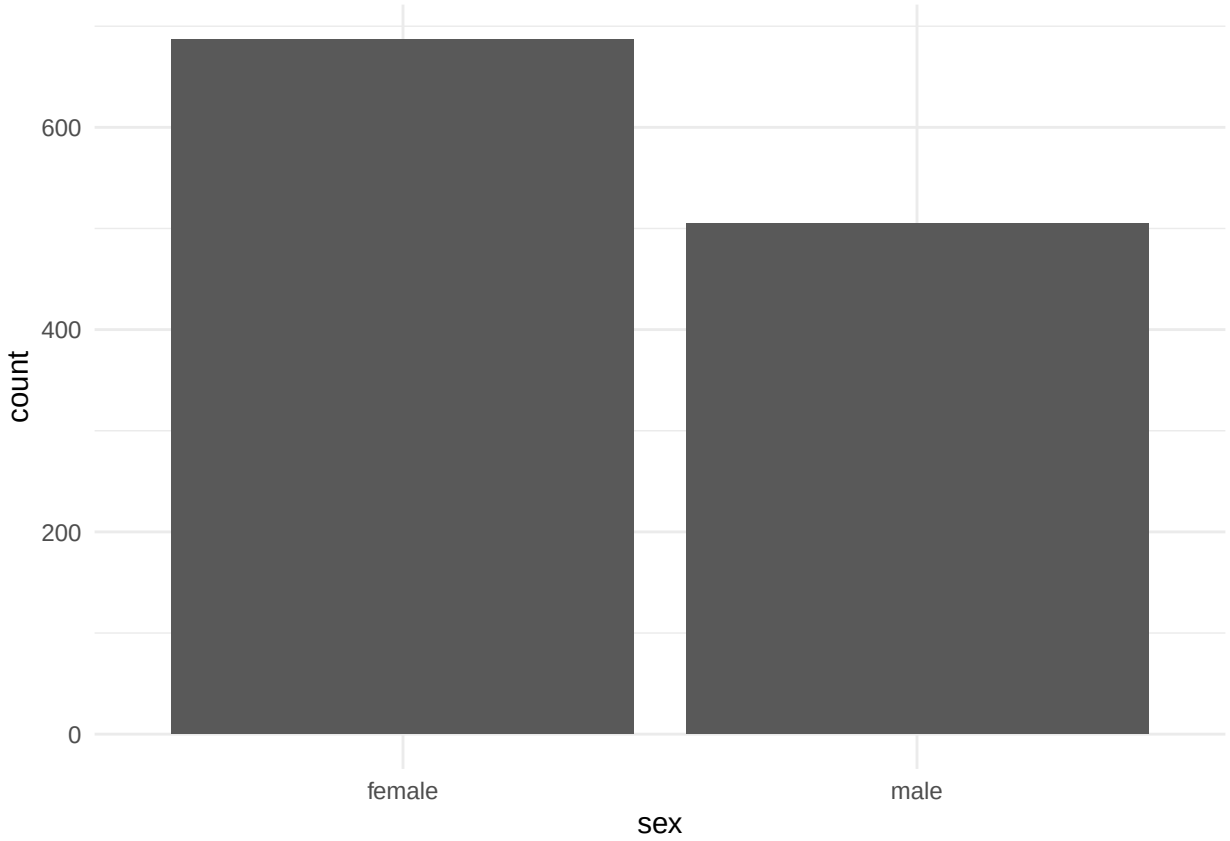
```
# Assignment: 6.2 Exercises
# Name: Rajeev, Rahul
# Date: 2023-01-16

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

## Set the working directory to the root of your DSC 520 directory
setwd("C:/Users/rahul/Documents/Bellevue/DSC 520")

## 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()
```



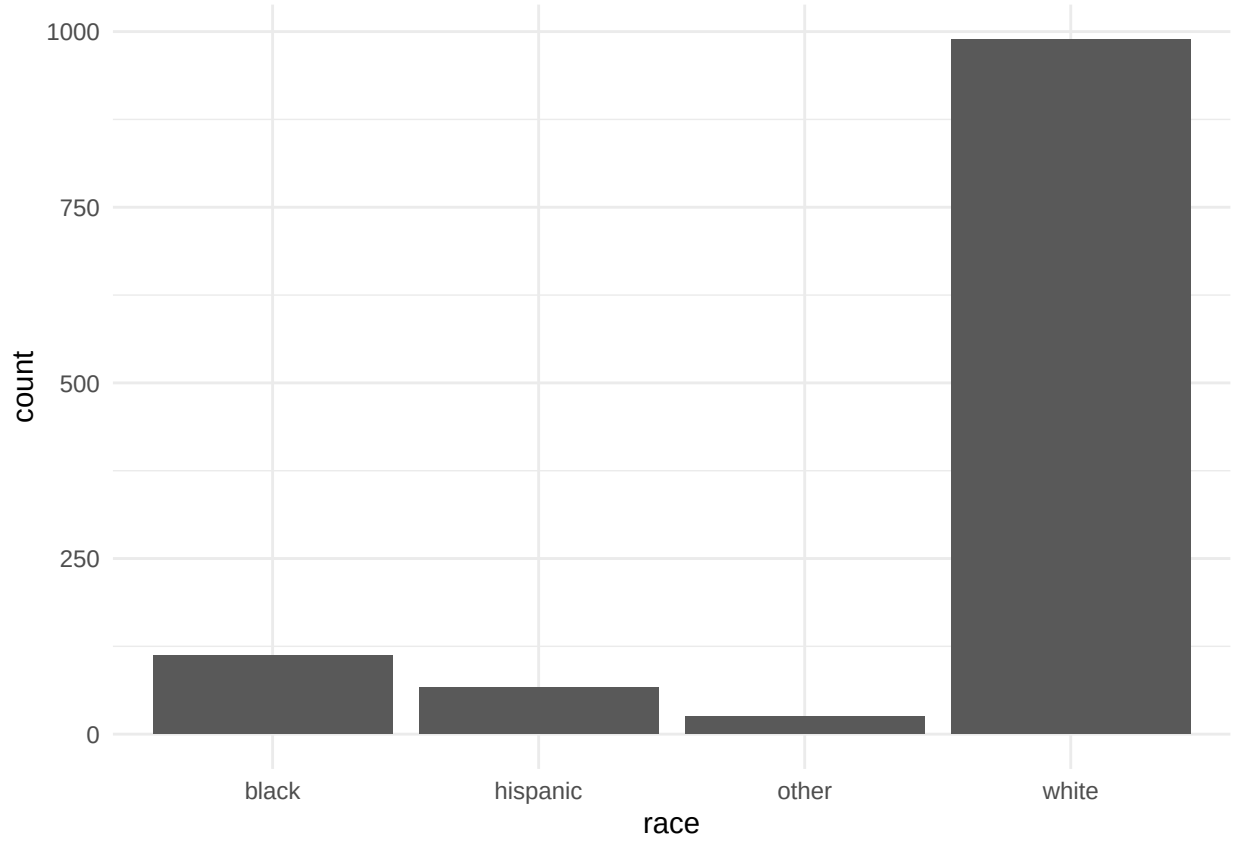

```
## race vs. earn
ggplot(heights_df, aes(x=race, y=earn)) + geom_point()+ geom_boxplot()
```



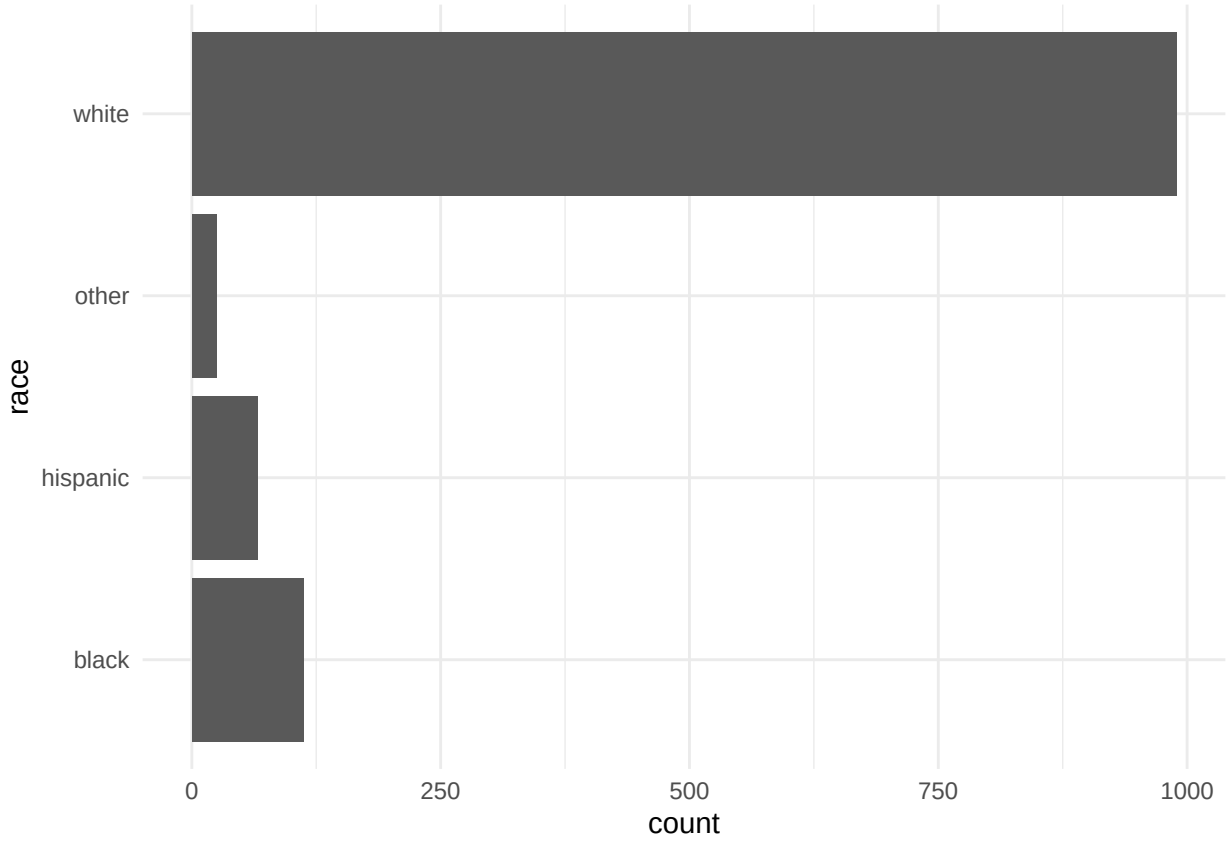

```
# https://ggplot2.tidyverse.org/reference/geom_bar.html
## Using `geom_bar()` plot a bar chart of the number of records for each `sex`
ggplot(heights_df, aes(sex)) + geom_bar()
```



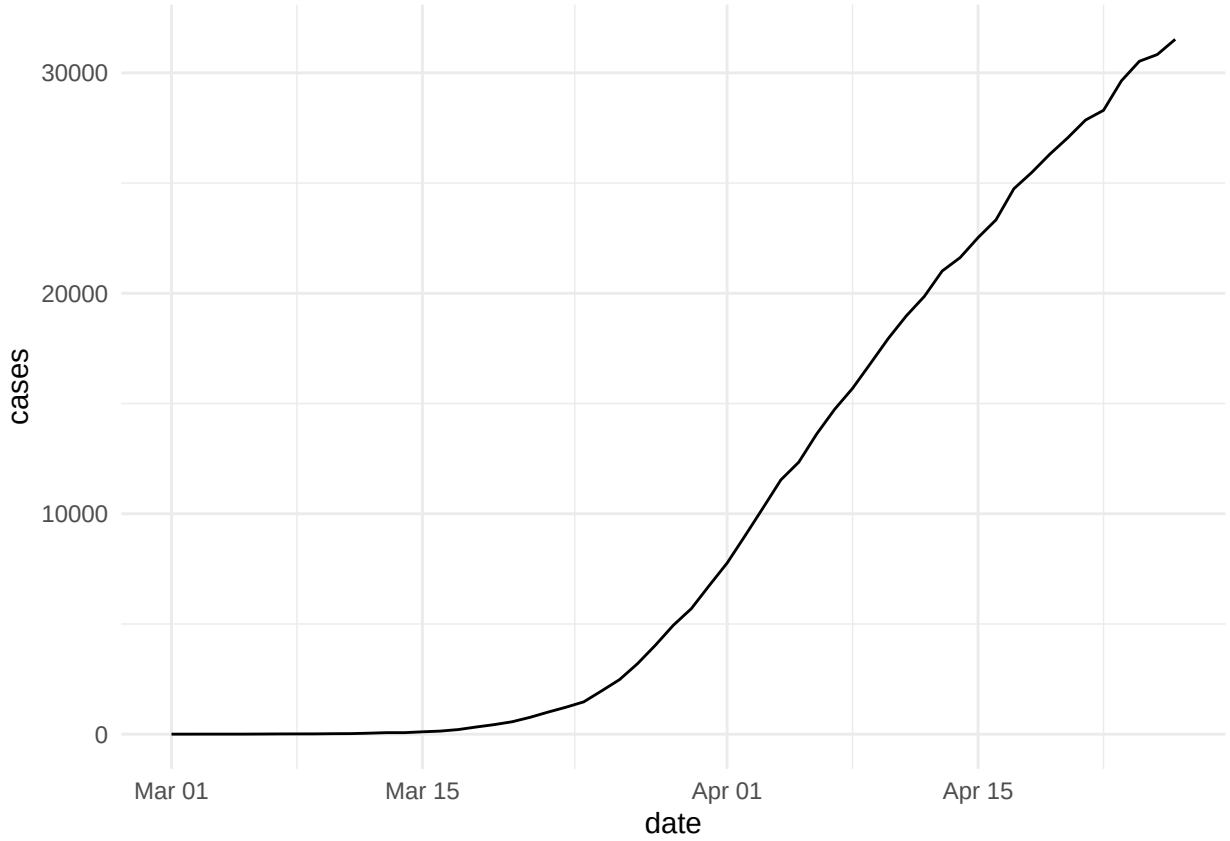

```
## Using `geom_bar()` plot a bar chart of the number of records for each race
ggplot(heights_df, aes(race)) + geom_bar()
```



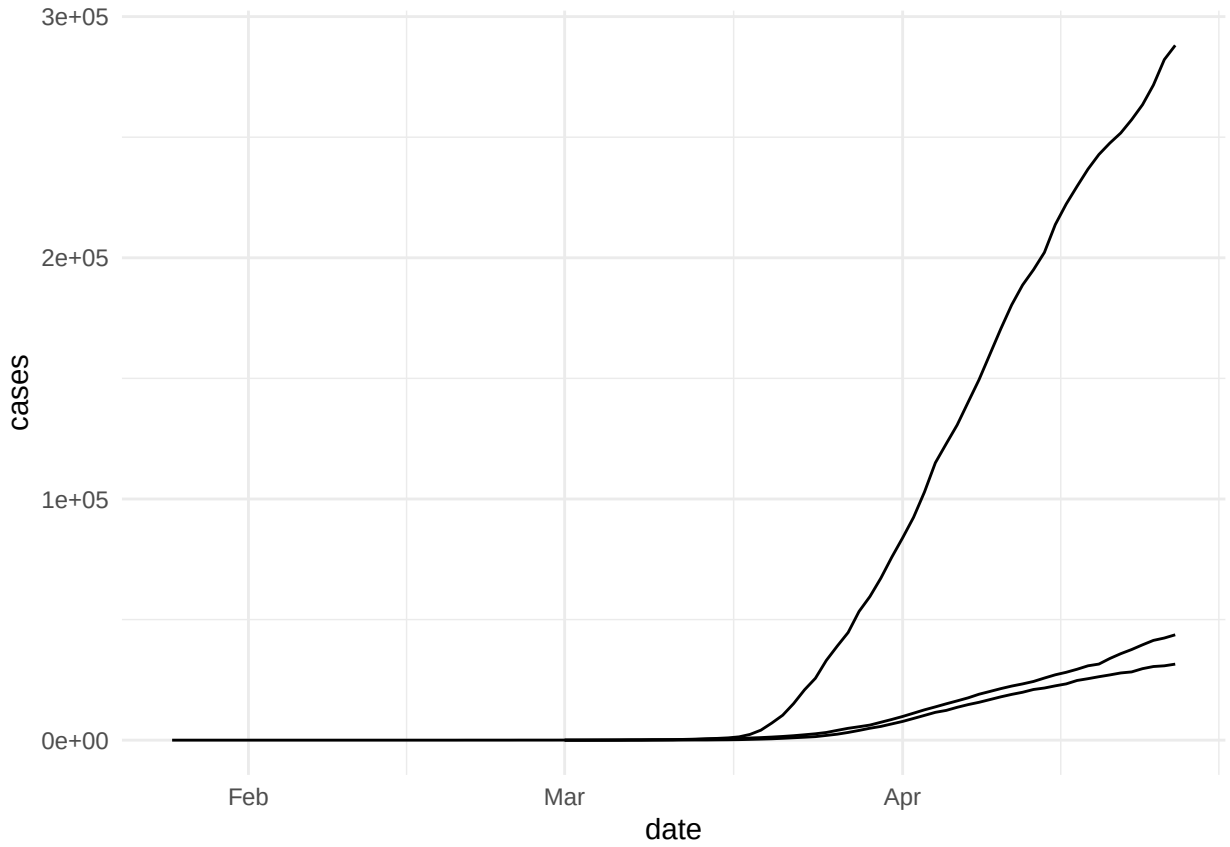

```
## Create a horizontal bar chart by adding `coord_flip()` to the previous plot
ggplot(heights_df, aes(race)) + geom_bar() + coord_flip()
```



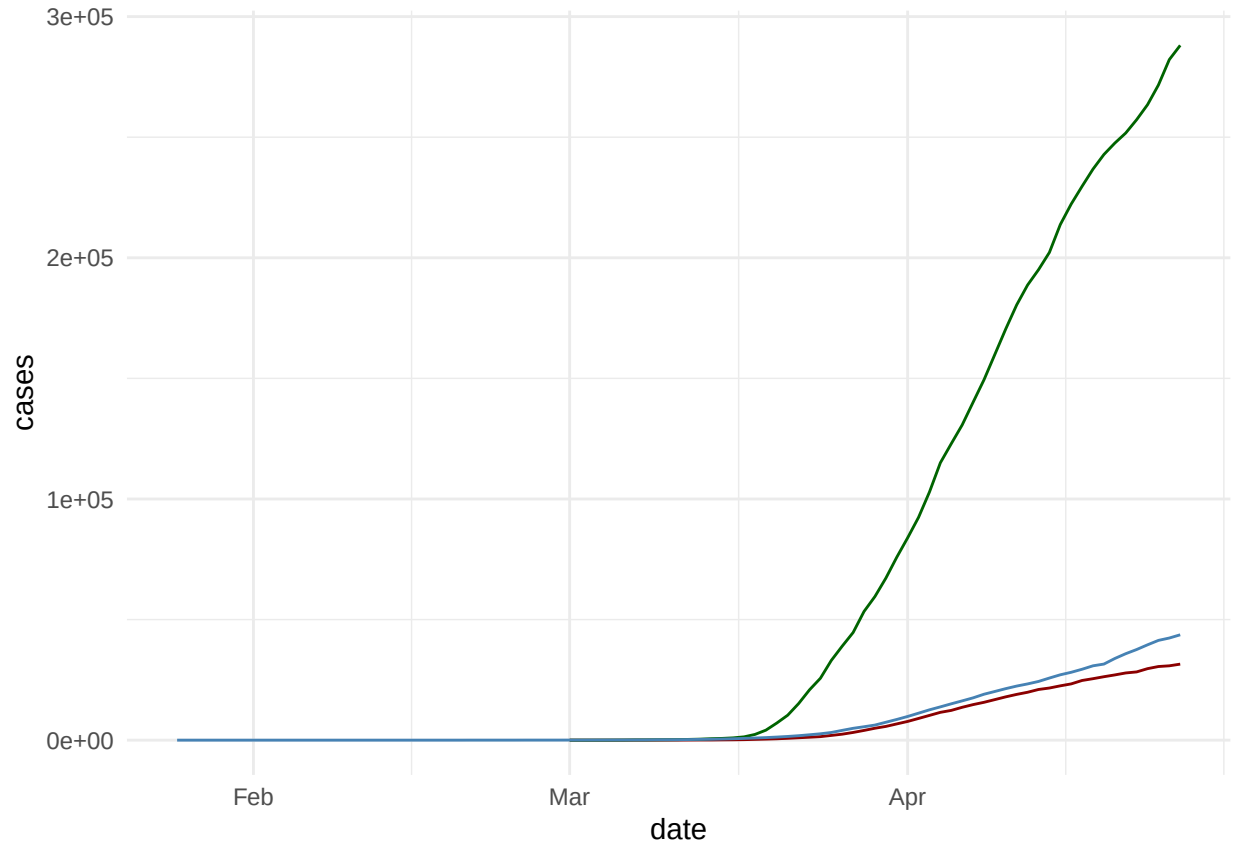

```
# https://www.rdocumentation.org/packages/ggplot2/versions/3.3.0/topics/geom_path
## Load the file `"data/nytimes/covid-19-data/us-states.csv"` and
## assign it to the `covid_df` dataframe
covid_df <- read.csv("data/nytimes/covid-19-data/us-states.csv")

## 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()
```



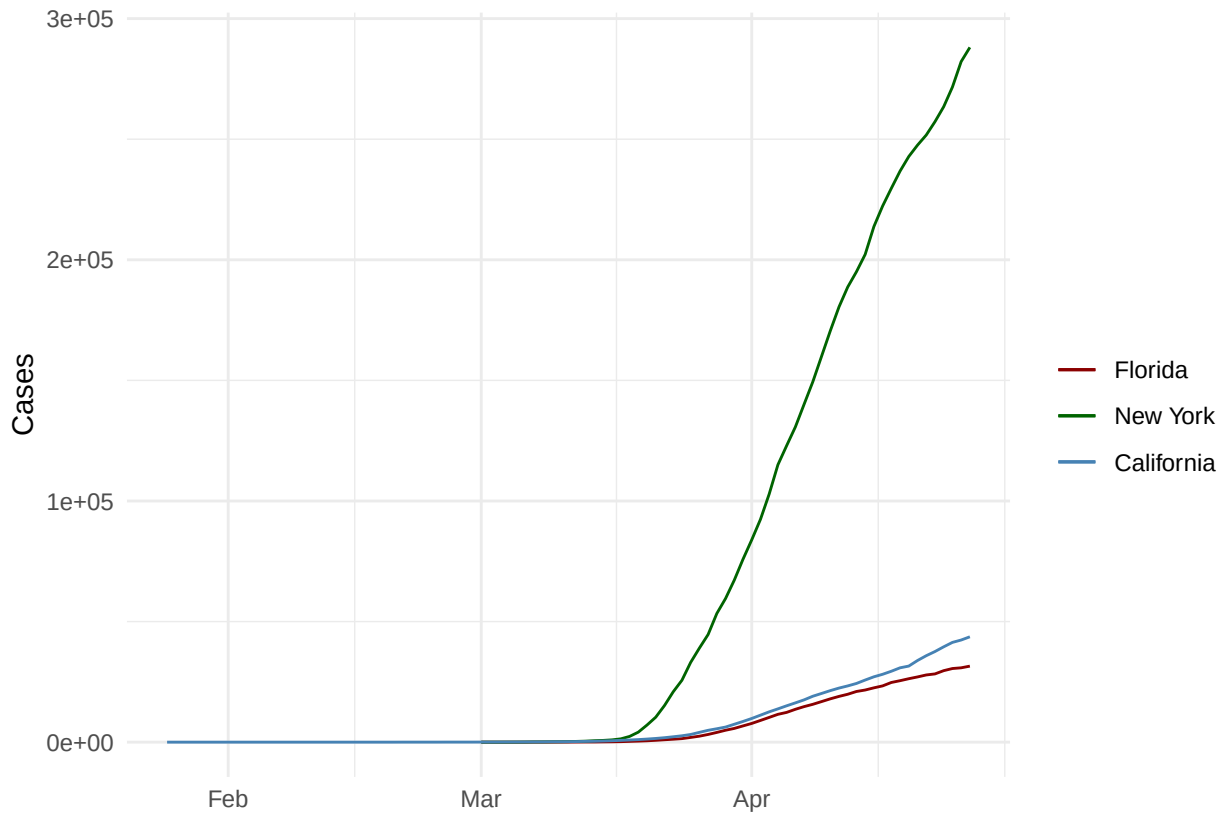

```
## Add lines for New York and California to the plot
ggplot(data=florida_df, aes(x=date, group=1)) +
  geom_line(aes(y = cases)) +
  geom_line(data=ny_df, aes(y = cases)) +
  geom_line(data=california_df, aes(y = cases))
```



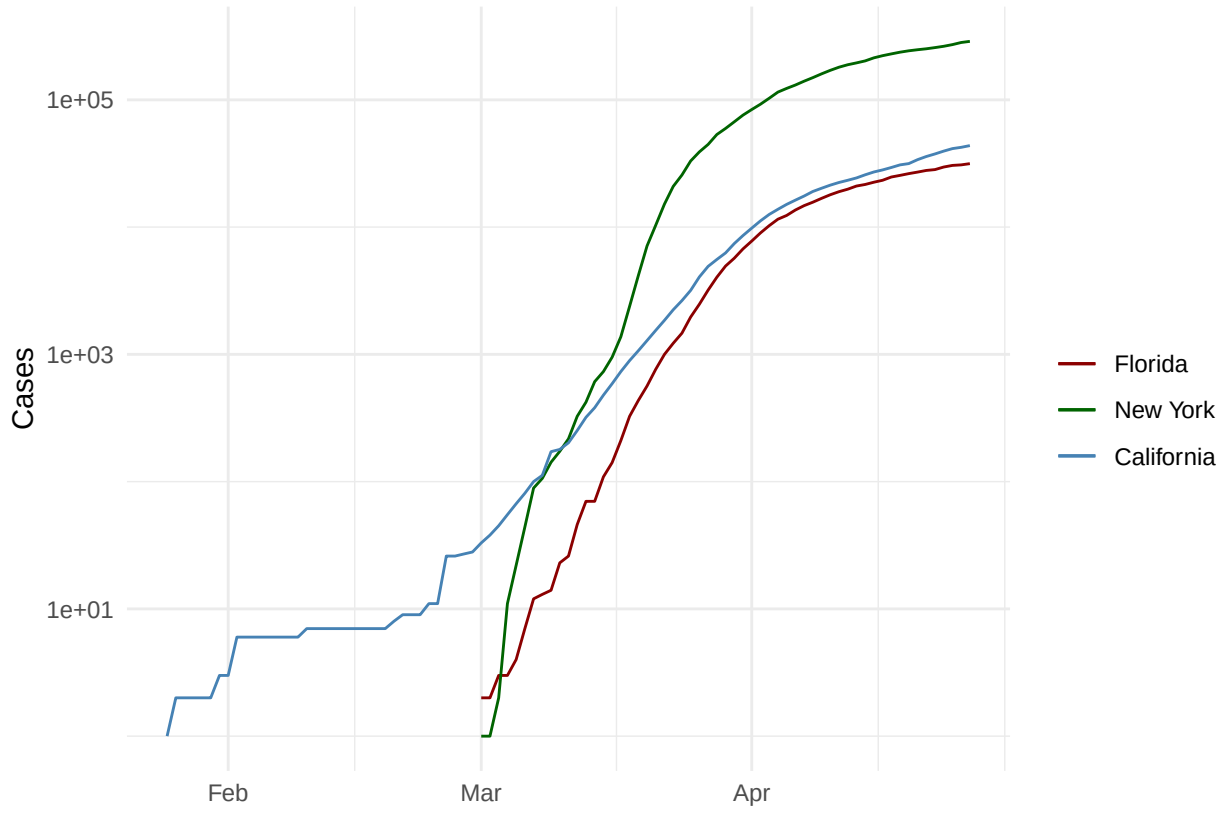

```
## Use the colors "darkred", "darkgreen", and "steelblue" for Florida, New York, and California
ggplot(data=florida_df, aes(x=date, group=1)) +
  geom_line(aes(y = cases), color = 'darkred') +
  geom_line(data=ny_df, aes(y = cases), color = 'darkgreen') +
  geom_line(data=california_df, aes(y = cases),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


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

6.2 Markdown

Rahul Rajeev

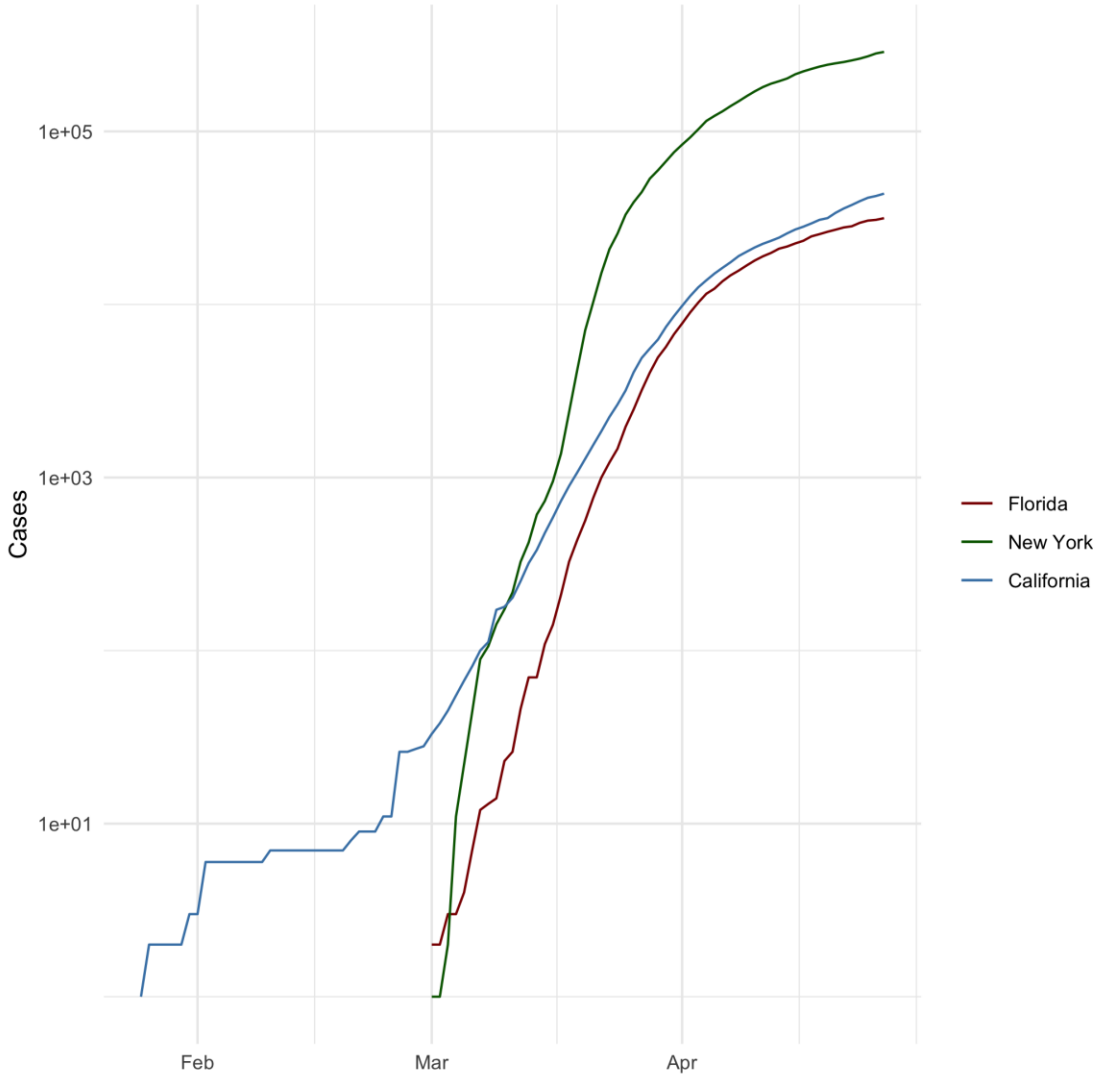
2023-01-16

Markdown Basics

Favorite Foods

1. Falafel
2. Tacos
3. Orange Chicken

Images



All Cases (Log Plot)

Add a Quote

> "I am inevitable." — Thanos

Add an Equation

Mass Energy Equivalence : $E = mc^2$

Add a Footnote

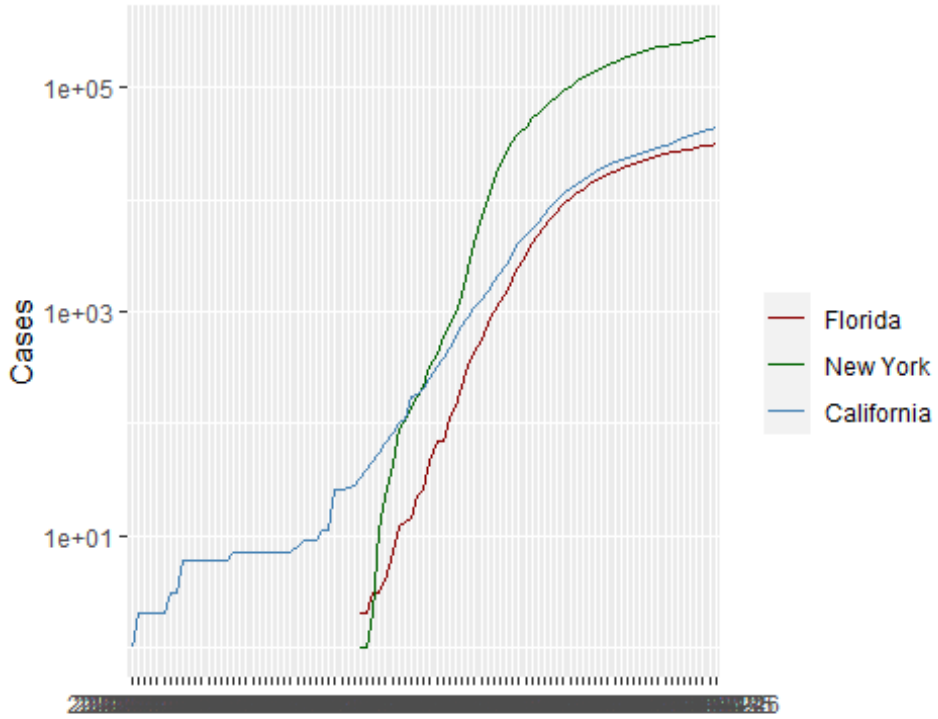
This is a footnote[1]

Add Citations

- R for Everyone (Lander 2014)
- Discovering Statistics Using R (Field, Miles, and Field 2012)

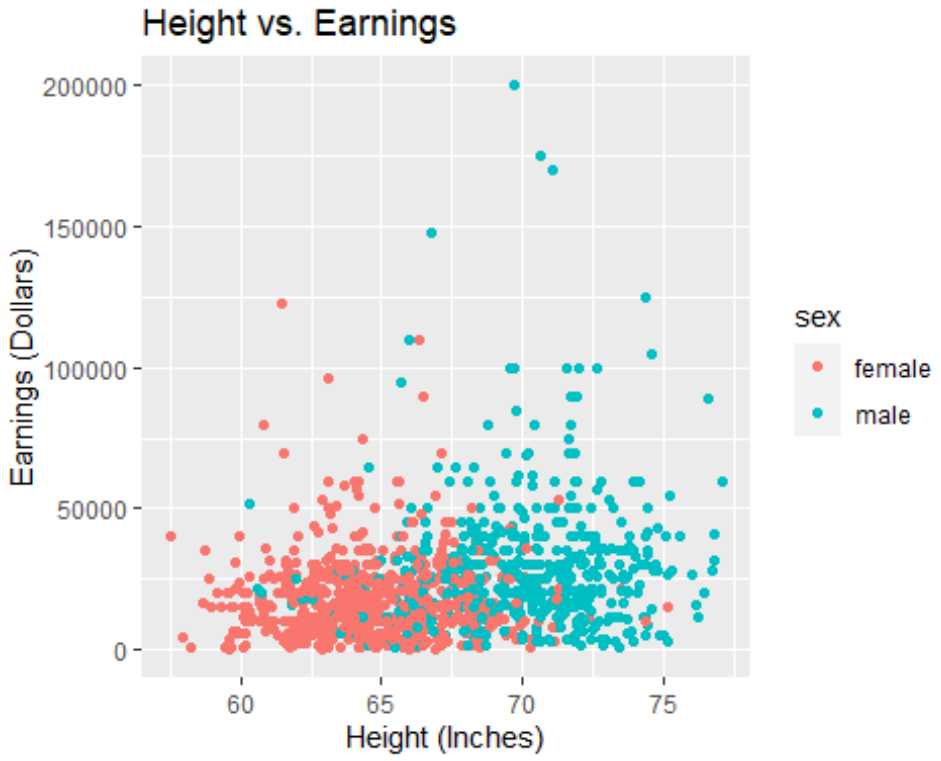
Inline Code

NY Times COVID-19 Data



[1] Footnotes are great!

R4DS Height vs Earnings



Tables

Knitr Table with Kable

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

Name	Race	In Fellowship?	Is Ring Bearer?	Age
Aragon	Men	Yes	No	88

Name	Race	In Fellowship?	Is Ring Bearer?	Age
Bilbo	Hobbit	No	Yes	129
Frodo	Hobbit	Yes	Yes	51
Sam	Hobbit	Yes	Yes	36
Sauron	Maia	No	Yes	7052

References

Field, A., J. Miles, and Z. Field. 2012. *Discovering Statistics Using r*. SAGE Publications. https://books.google.com/books?id=wd2K2zC3swIC.

Lander, J. P. 2014. *R for Everyone: Advanced Analytics and Graphics*. Addison-Wesley Data and Analytics Series. Addison-Wesley. https://books.google.com/books?id=3eBVAgAAQBAJ.