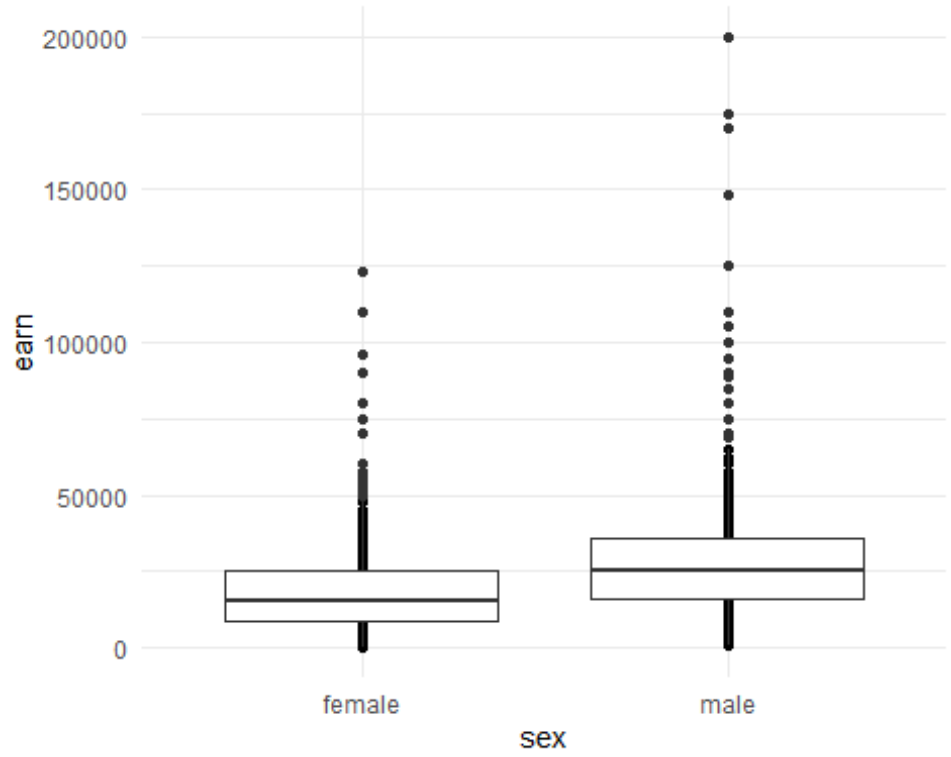
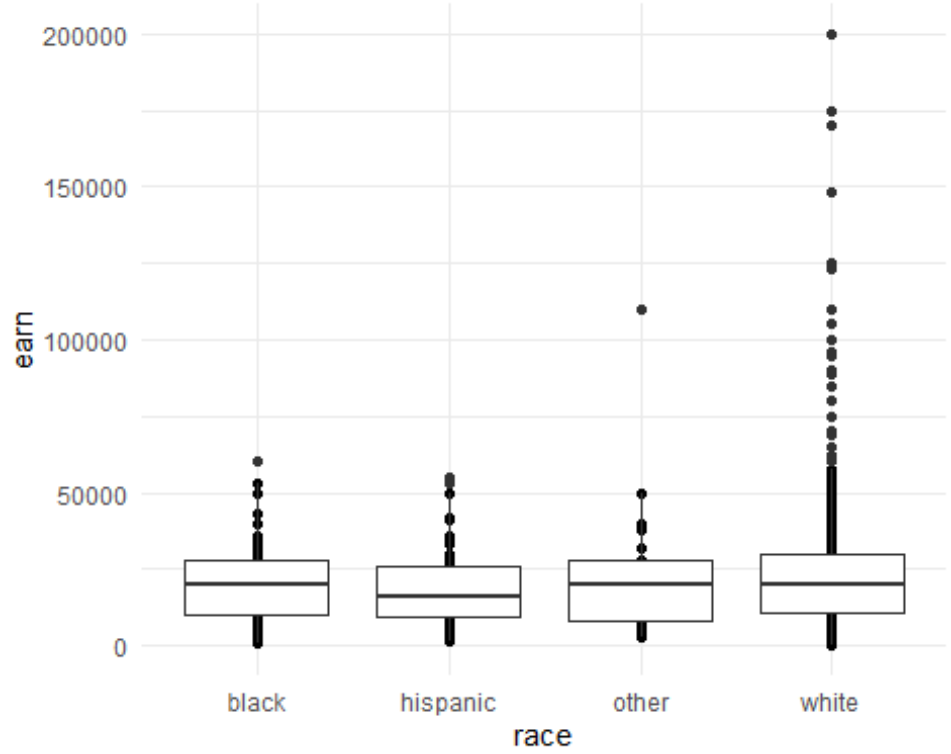
assignment_04_Torres_Gloria.R

2023-04-22

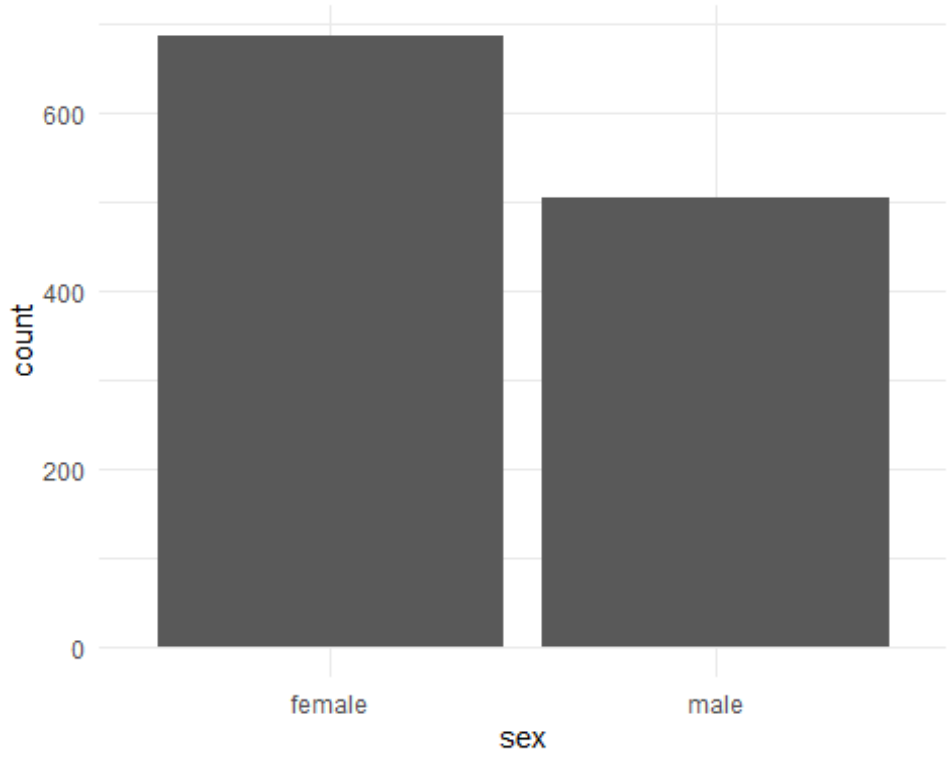
```
# Assignment: ASSIGNMENT 4
# Name: Torres, Gloria
# Date: 2023-04-23



## 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/glori/OneDrive/Documents/Gloria GIT/Gloria_Torres_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()
```



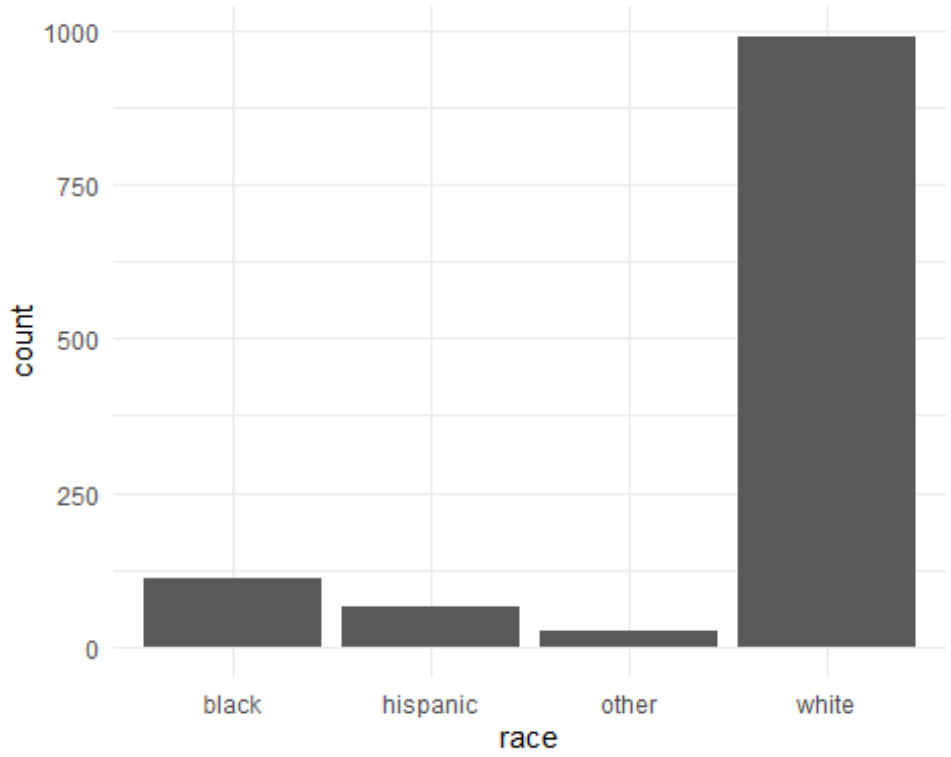

```
## race vs. earn
ggplot(heights_df, aes(x=race, y=earn)) +geom_point() + geom_boxplot()
```



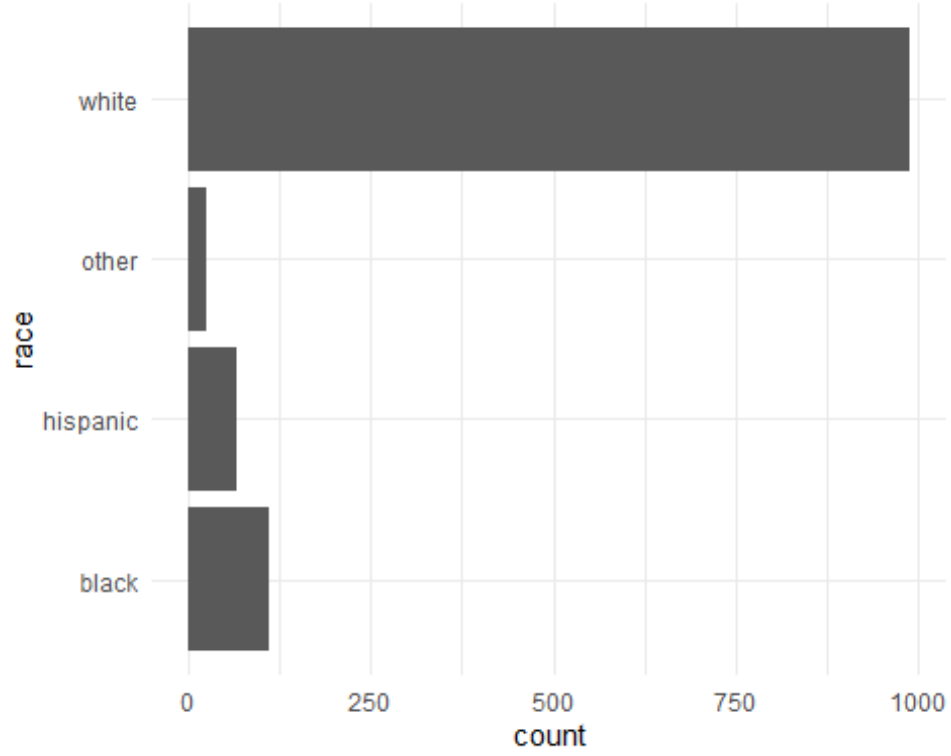
```
## https://ggplot2.tidyverse.org/reference/geom_bar.html
## Using `geom_bar()` plot a bar chart of the number of records for each
`sex`
ggplot(heights_df, aes(sex)) + geom_bar()
```




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



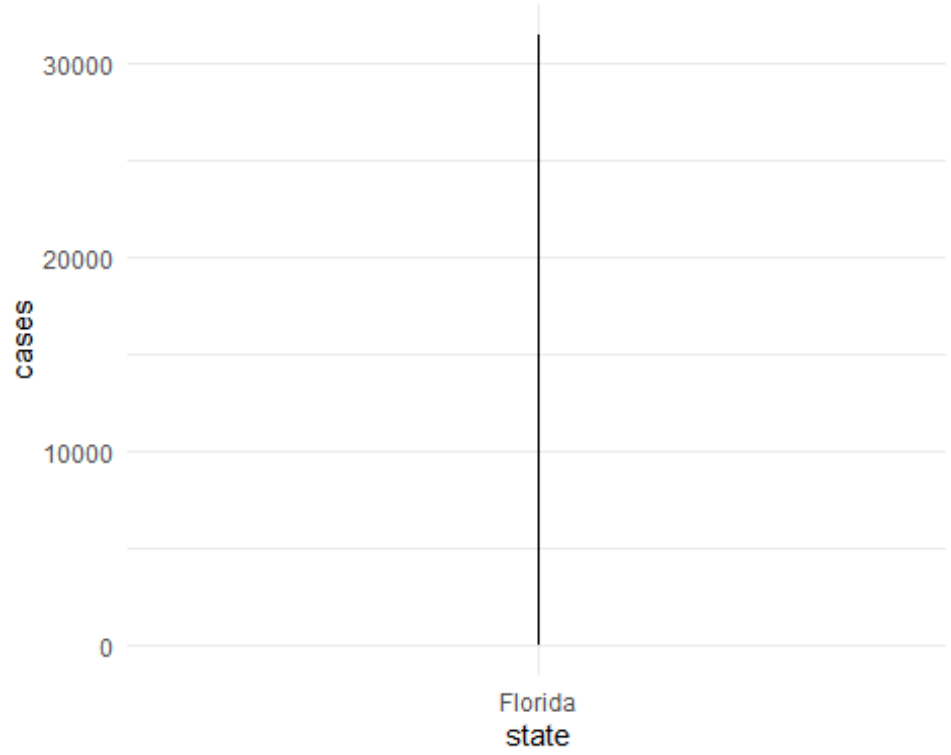

```
## Create a horizontal bar chart by adding `coord_flip()` to the previous plot
ggplot(heights_df, aes(race)) + geom_bar() + coord_flip()
```



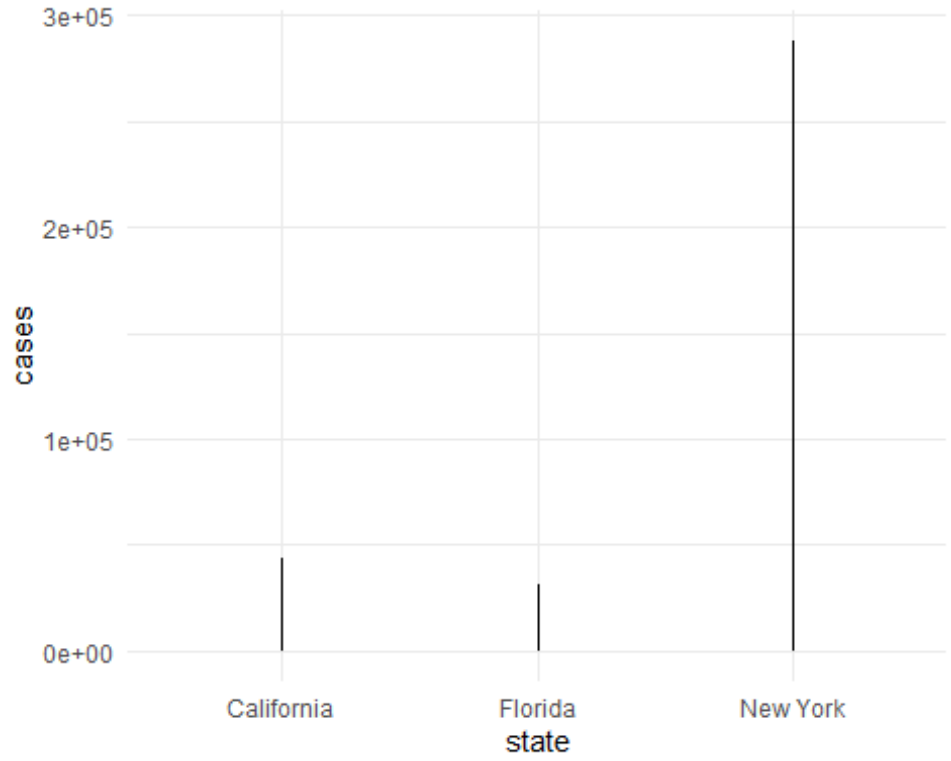
```
##
https://www.rdocumentation.org/packages/ggplot2/versions/3.3.0/topics/geom_path
## Load the file `"data/nytimes/covid-19-data/us-states.csv"` and
## assign it to the `covid_df` dataframe


covid_df <- read.csv("data/nytimes/covid-19-data/us-states.csv")


## Parse the date column using `as.Date()``


covid_df$date <- as.Date(c(covid_df$date))
head(covid_df$date)
```

```
## [1] "2020-01-21" "2020-01-22" "2020-01-23" "2020-01-24" "2020-01-24"
## [6] "2020-01-25"
```

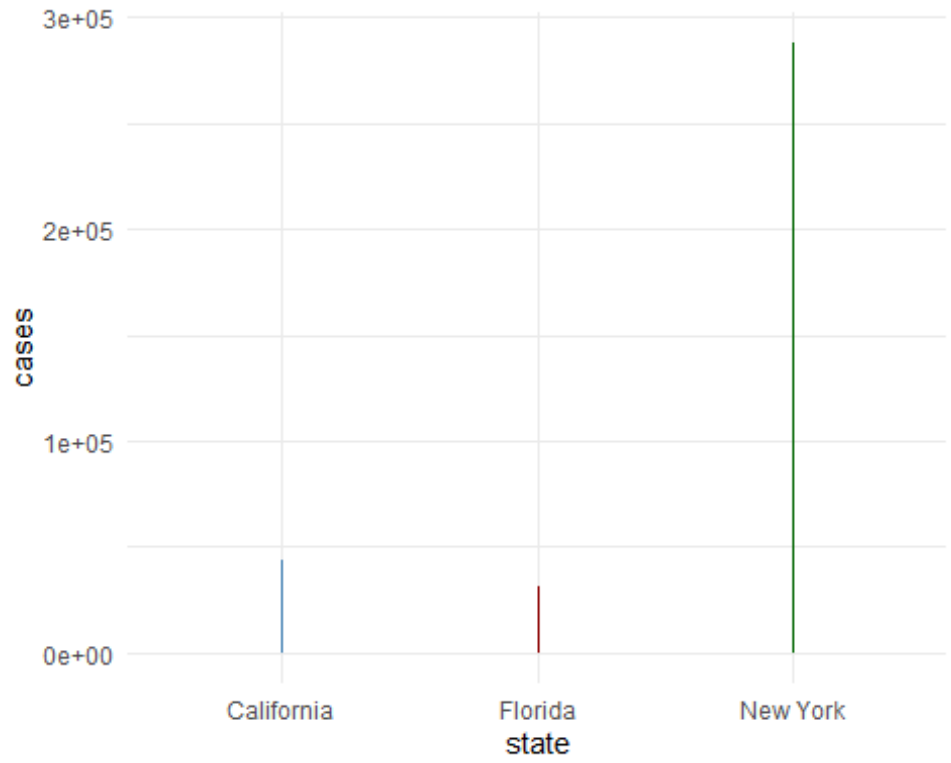
```
## Create three dataframes named `california_df`, `ny_df`, and `florida_df`
## containing the data from California, New York, and Florida
california_df <- covid_df[ which( covid_df$state == "California"), ]
ny_df <- covid_df[ which( covid_df$state == "New York"), ]
florida_df <- covid_df[ which(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

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



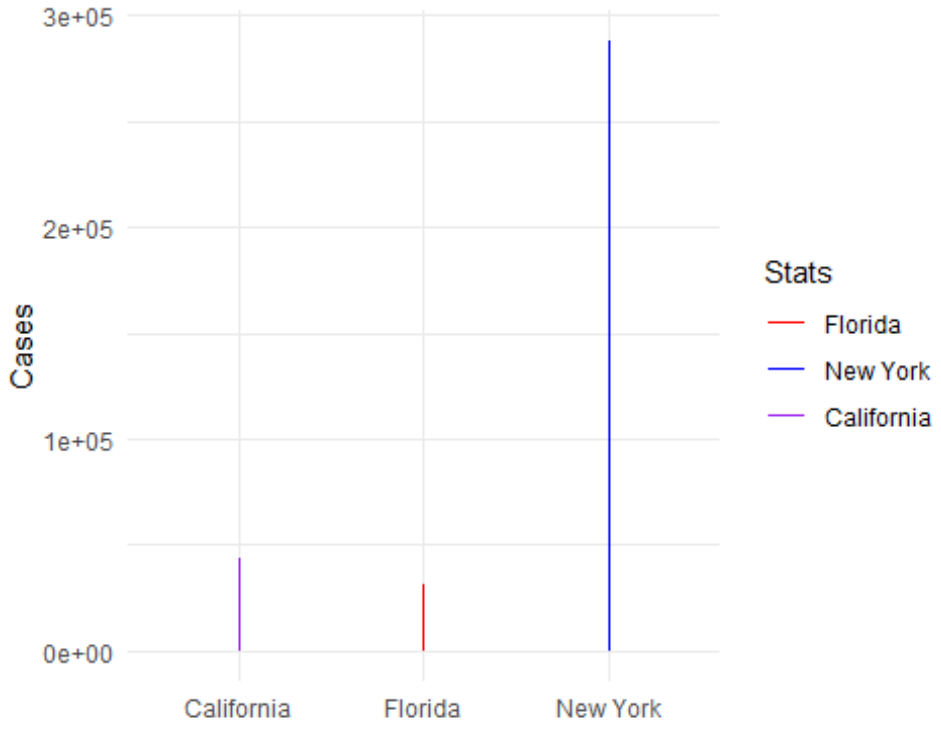

```
## Add lines for New York and California to the plot
ggplot(data=florida_df,aes(x=state, group=1)) +
  geom_line(aes(y = cases)) +
  geom_line(data=ny_df, aes(y = cases)) +
  geom_line(data=california_df, aes(y =cases))
```



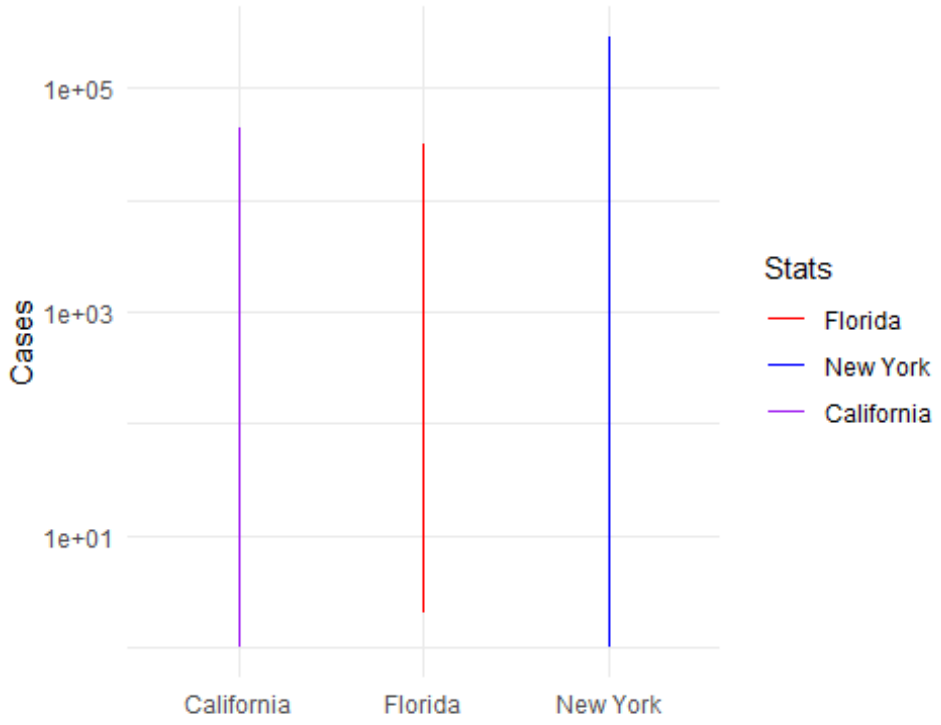

```
## Use the colors "darkred", "darkgreen", and "steelblue" for Florida, New York, and California
ggplot(data=florida_df, aes(x=state, group=1)) +
  geom_line(aes(y = cases), color ="darkred") +
  geom_line(data=ny_df, aes(y =cases), color="darkgreen") +
  geom_line(data=california_df, aes(y =cases), 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
Stats1 <- ggplot(data=florida_df, aes(x=state, group=1)) +
  geom_line(aes(y = cases, color ='Florida')) +
  geom_line(data=ny_df, aes(y = cases, color='New York')) +
  geom_line(data=california_df, aes(y = cases, color='California')) +
  scale_colour_manual(name='Stats',
                      breaks = c("Florida", "New York", "California"),
                      values = c("Florida"="red", "New York"="blue",
"California"="purple")) +
  xlab(" ") + ylab("Cases")
Stats1
```




```
## Scale the y axis using `scale_y_log10()`

Stats2 <- ggplot(data=florida_df, aes(x=state, group=1)) +
  geom_line(aes(y = cases, color = "Florida")) +
  geom_line(data=ny_df, aes(y = cases, color="New York")) +
  geom_line(data=california_df, aes(y = cases, color="California")) +
  scale_colour_manual(name='Stats',
                      breaks = c("Florida", "New York", "California"),
                      values = c("Florida"="red", "New York"="blue",
"California"="purple")) +
  xlab(" ") + ylab("Cases") + scale_y_log10()
Stats2
```


1e+05
1e+03
1e+01
Cases
California
Florida
New York
Stats
Florida
New York
California