Lab 4

```
library(readr)
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────────────────────────────────────── tid
yverse 1.2.1 ──
```

```
## ✔ ggplot2 3.2.1     ✔ purrr   0.3.2
## ✔ tibble  2.1.3     ✔ dplyr   0.8.3
## ✔ tidyr   1.0.0     ✔ stringr 1.4.0
## ✔ ggplot2 3.2.1     ✔ forcats 0.4.0
```

```
## ── Conflicts ────────────────────────────────────────────────────────────────────────── tidyverse
_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

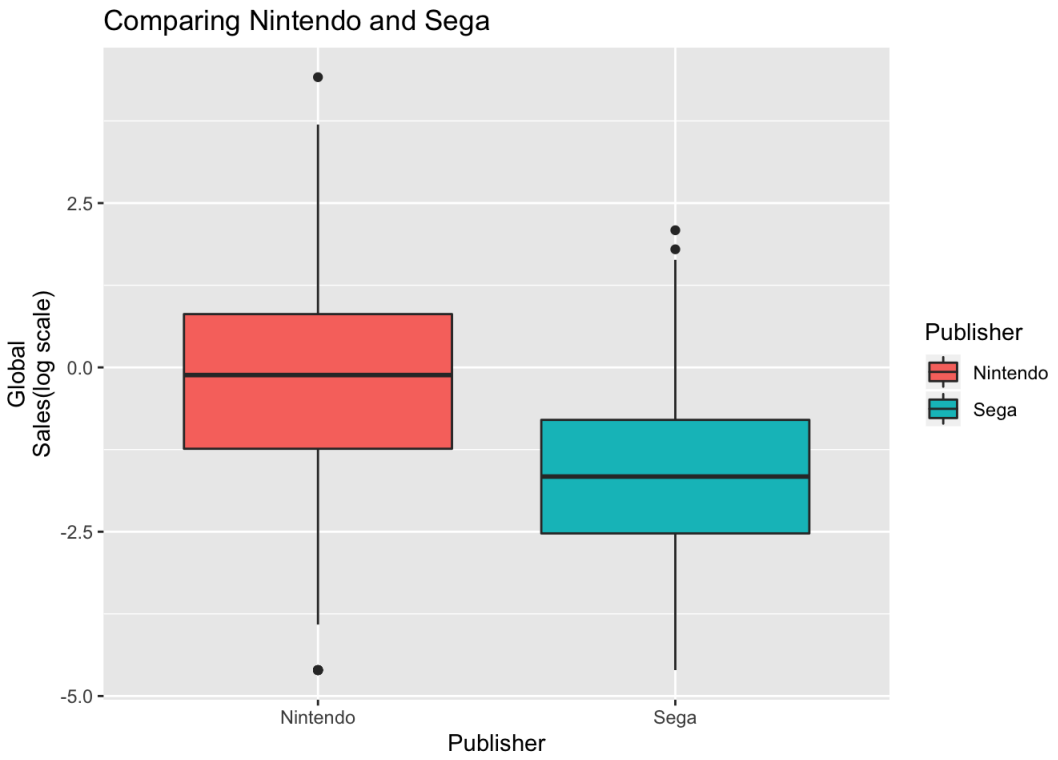
```
vgsales <- read_csv("vgsales.csv")
```

```
## Parsed with column specification:
## cols(
##   Rank = col_double(),
##   Name = col_character(),
##   Platform = col_character(),
##   Year = col_character(),
##   Genre = col_character(),
##   Publisher = col_character(),
##   NA_Sales = col_double(),
##   EU_Sales = col_double(),
##   JP_Sales = col_double(),
##   Other_Sales = col_double(),
##   Global_Sales = col_double()
## )
```

Replicate a plot

```
Nintendo_Sega <- filter(vgsales, Publisher=="Nintendo" | Publisher=="Sega")

ggplot(Nintendo_Sega, aes(x=Publisher, y=log(Global_Sales), fill = Publisher))+
  geom_boxplot() + ggtitle("Comparing Nintendo and Sega") + ylab("Global
Sales(log scale)")
```



Rise in Violent Video Games?

```
old <- 0
new <-0

j <- nrow(vgsales)
for (i in 1:j){
  if(vgsales %>% slice(i) %>% pull(Genre) == "Shooter"){
    if(vgsales %>% slice(i) %>% pull(Year) < 2000)
      old = old + 1
    if(vgsales %>% slice(i) %>% pull(Year) > 1999)
      new = new + 1
  }
}
old
```

```
## [1] 167
```

```
new
```

```
## [1] 1143
```

```
Prior to the 2000s, only 167 of the video games in our dataset were categorized as shooters. After the 2000s
, 1143 of the video games in our dataset were categorized as shooters. Because of this drastic increase, par
ents concern with the rise in violent video games are merited and reasonable.
```

Flex Points

```
vgsales %>%
  select(Name, Publisher, NA_Sales, EU_Sales, JP_Sales) %>%
  pivot_longer(cols = c(NA_Sales, EU_Sales, JP_Sales), names_to = "Location", values_to = "Num_Sales")
```

```
## # A tibble: 49,794 x 4
##    Name              Publisher Location Num_Sales
##    <chr>             <chr>     <chr>        <dbl>
##  1 Wii Sports        Nintendo  NA_Sales     41.5
##  2 Wii Sports        Nintendo  EU_Sales     29.0
##  3 Wii Sports        Nintendo  JP_Sales      3.77
##  4 Super Mario Bros. Nintendo  NA_Sales     29.1
##  5 Super Mario Bros. Nintendo  EU_Sales      3.58
##  6 Super Mario Bros. Nintendo  JP_Sales      6.81
##  7 Mario Kart Wii    Nintendo  NA_Sales     15.8
##  8 Mario Kart Wii    Nintendo  EU_Sales     12.9
##  9 Mario Kart Wii    Nintendo  JP_Sales      3.79
## 10 Wii Sports Resort Nintendo  NA_Sales     15.8
## # … with 49,784 more rows
```