Lab 02

due January 19th at 11:59 PM

Ayush Jain

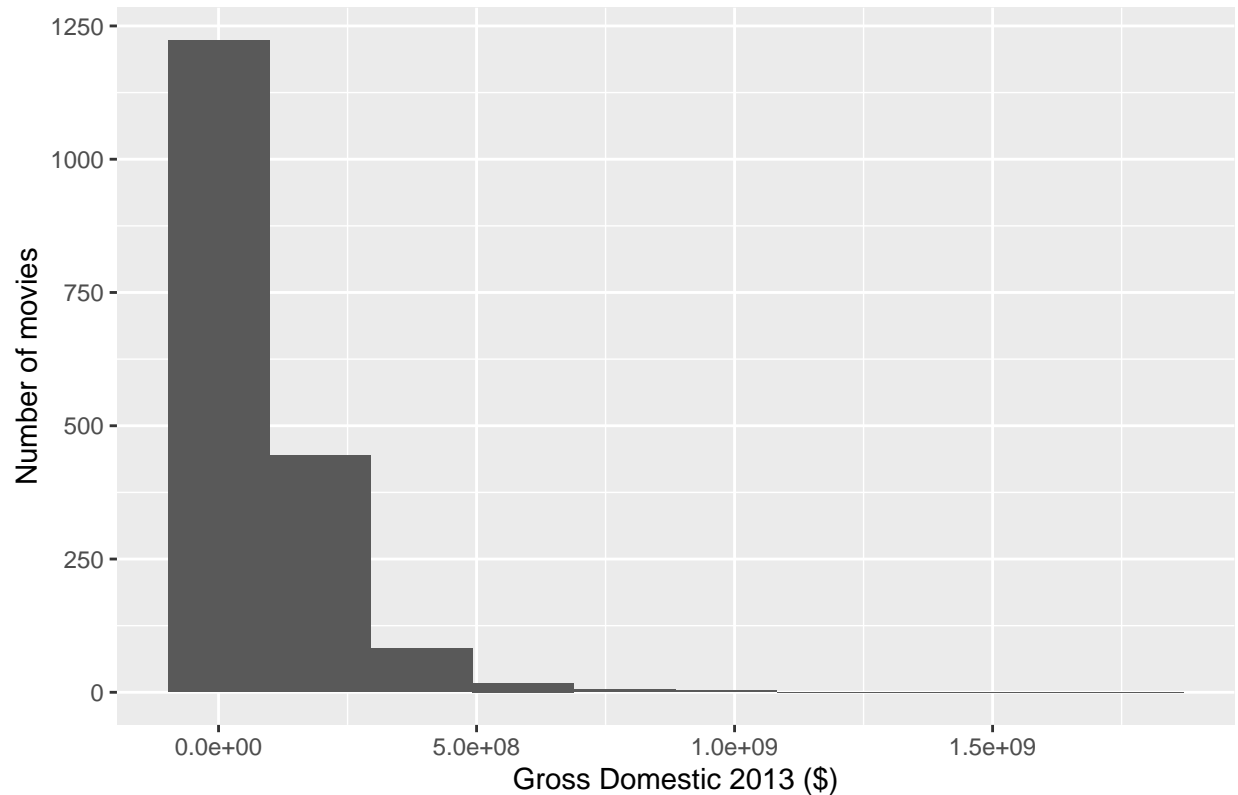01/10

Load Packages

```
library(tidyverse)
library(viridis)
library(fivethirtyeight)
```

Exercise 1

```
ggplot(bechdel, aes(x = domgross_2013))+
  geom_histogram(bins = 10)+
  labs(title = "How much most movies grossed domestically in 2013",
       x = "Gross Domestic 2013 ($)",
       y = "Number of movies")
```

How much most movies grossed domestically in 2013

The above histogram shows the domestic gross of movies in 2013 in dollars. The graph is very clearly skewed, and we can see that as the Gross Domestic increases (X-Axis), the number of movies decreases (Y-axis). All data fits this trend, however, it is hard to identify outliers since it is not very clear on the histogram

Exercise 2

```
ggplot(bechdel, aes(x = budget_2013, y = domgross_2013, color = binary))+
  geom_point()+
  labs(title = "Movie budget vs Domestic Gross for 2013",
       x = "Budget 2013 ($)",
       y = "Gross Domestic 2013 ($)")+
  scale_color_viridis(discrete = TRUE, option = "C", name = "Passed Bechdel Test?")
```

Movie budget vs Domestic Gross for 2013
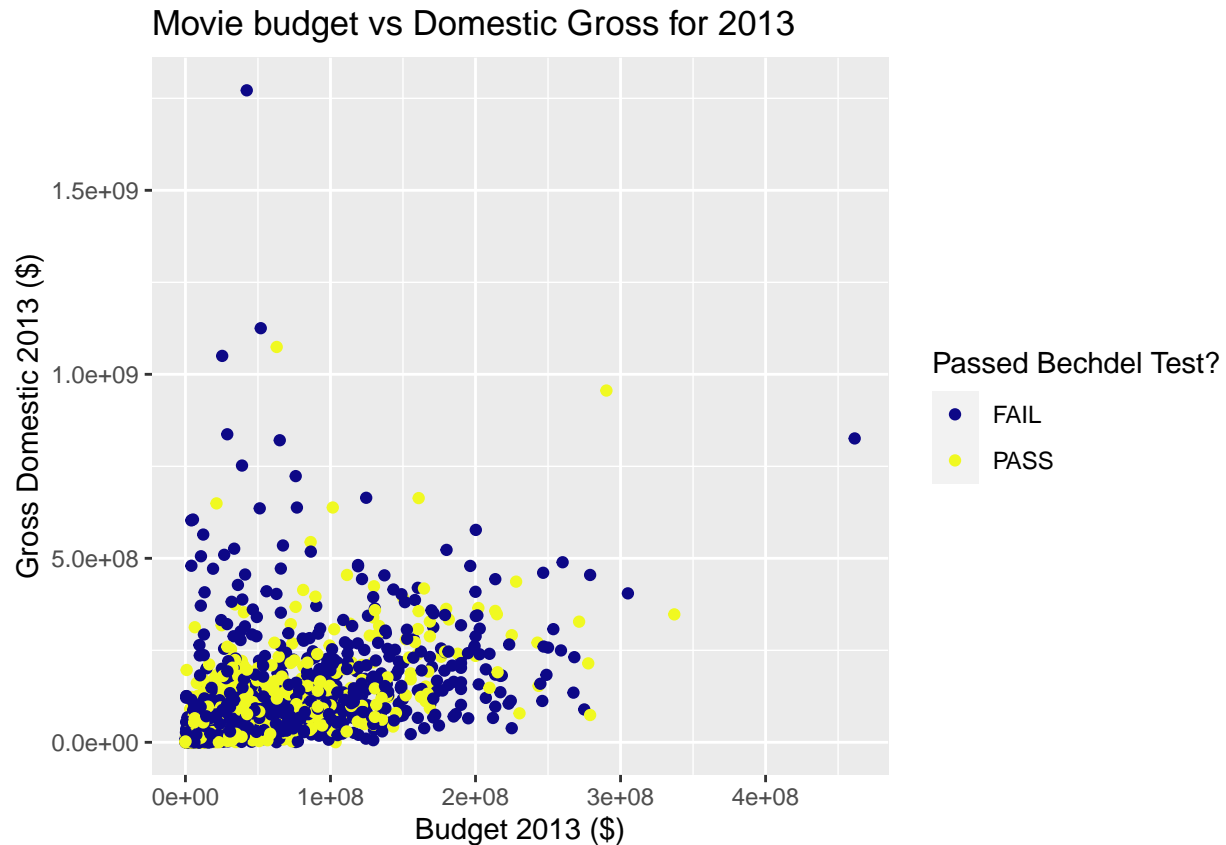
Exercise 3

No, the trend differs for movies that pass and those that don't. Movies that pass the test tend to gross around the same amount even as their budget increases. However, movies that don't pass the test tend to gross lower on average as their budget increases.

Exercise 4

```
bechdel_pass <- bechdel %>%
  filter(binary == "PASS")
bechdel_fail <- bechdel %>%
  filter(binary == "FAIL")
ggplot(bechdel_pass, aes(x = budget_2013, y = domgross_2013))+
  geom_point()+
  geom_smooth(method=lm, se=FALSE)+
  labs(title = "Movie budget vs Domestic Gross for 2013",
       subtitle = "For movies that passed the Bechdel Test",
       x = "Budget 2013 ($)",
       y = "Gross Domestic 2013 ($)")
```

Movie budget vs Domestic Gross for 2013
For movies that passed the Bechdel Test
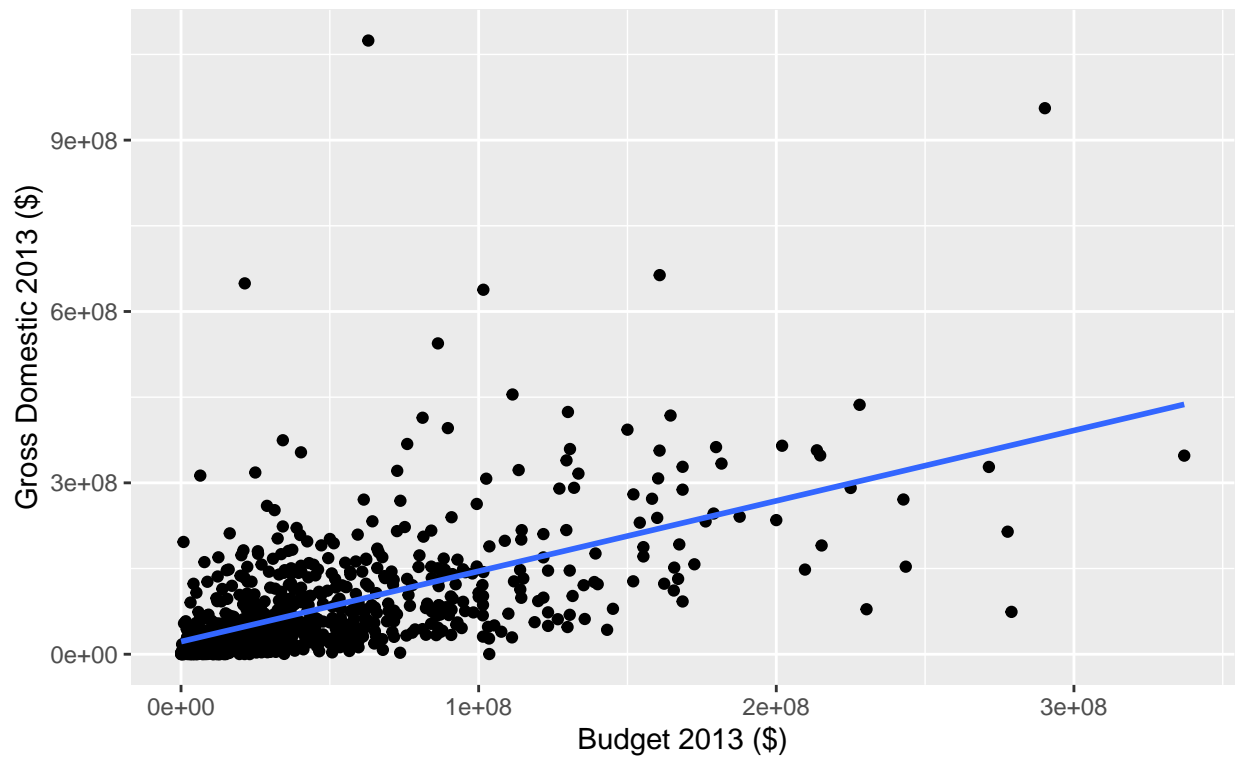

```
ggplot(bechdel_fail, aes(x = budget_2013, y = domgross_2013))+
  geom_point()+
  geom_smooth(method=lm, se=FALSE)+
  labs(title = "Movie budget vs Domestic Gross for 2013",
       subtitle = "For movies that failed the Bechdel Test",
       x = "Budget 2013 ($)",
       y = "Gross Domestic 2013 ($)")
```

Movie budget vs Domestic Gross for 2013
For movies that failed the Bechdel Test
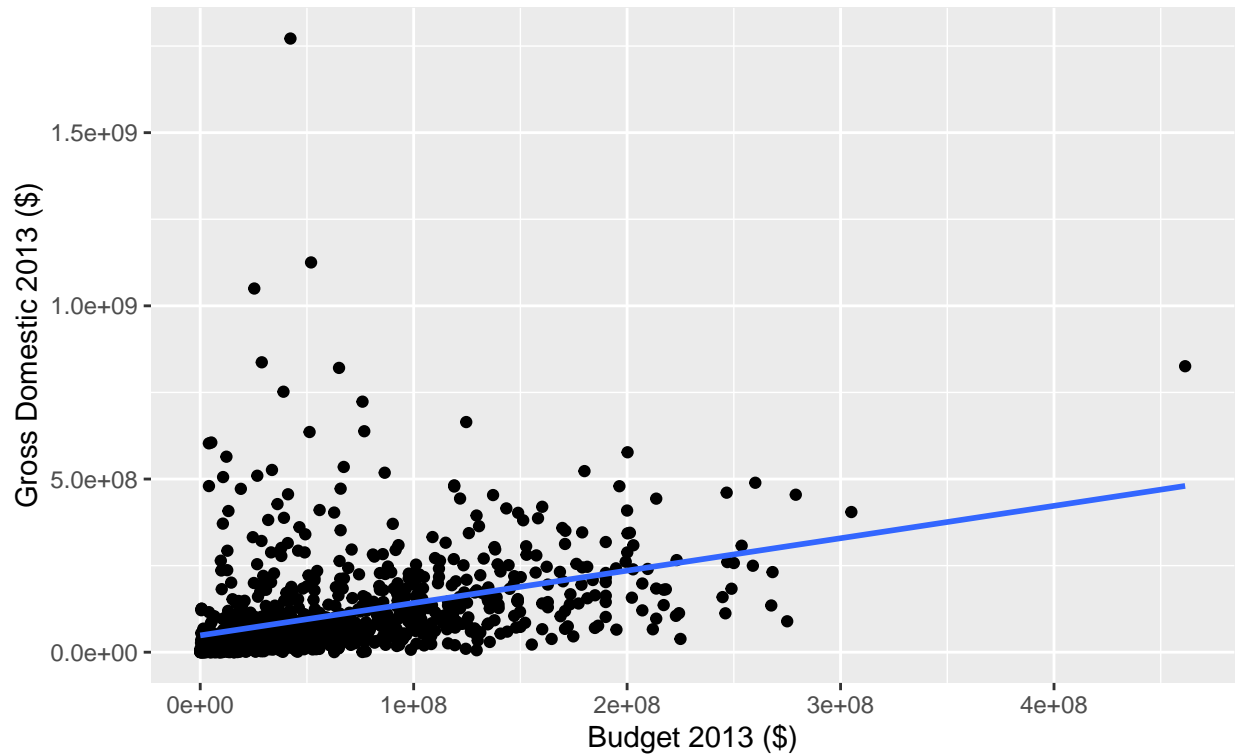
I preferred the previous plot because it allowed us to see the data for movies that passed as well as those that didn't, allowing us to make better comparisons between the two datasets
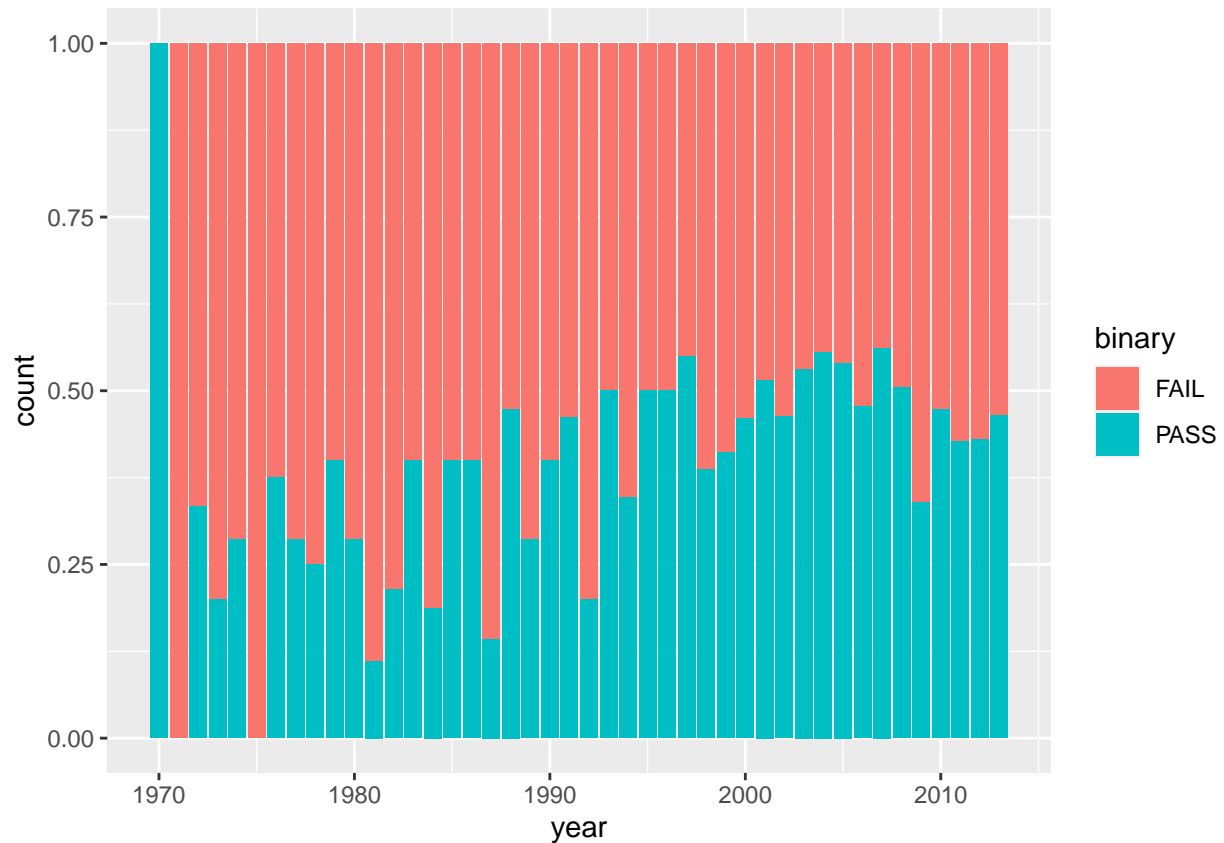
Exercise 5

```
ggplot(data = bechdel, aes(x = factor(clean_test), y = intgross_2013)) +
  geom_boxplot() +
  labs(title = "International gross vs Bechdel result",
       x = "Detailed result",
       y = "International gross or 2013")
```

International gross vs Bechdel result
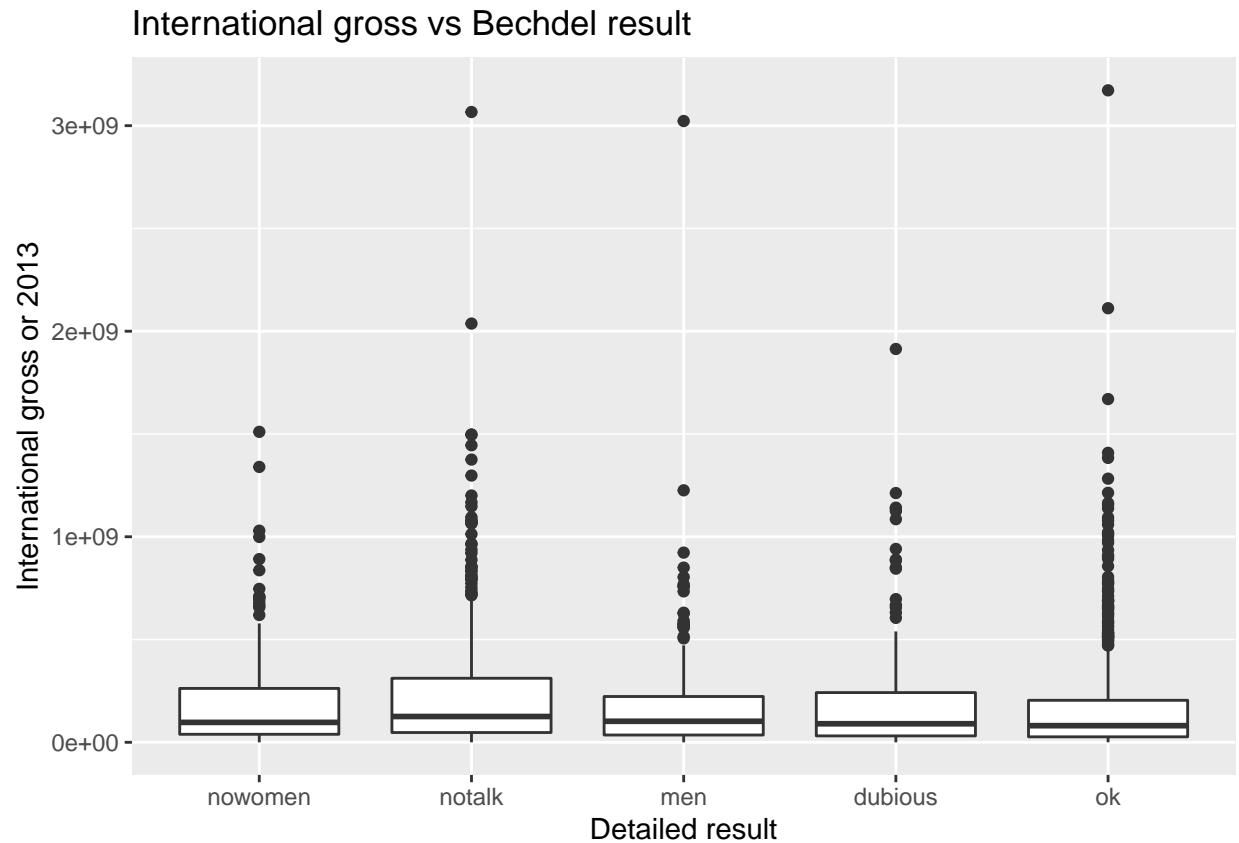
Yes, the international gross does depend on whether the movie passes the Bechdel Test. Movies with no women generally tend to gross the lowest. Only movies in the ok, men, and notalk category have grossed over $3 billion

Exercise 6

```r
ggplot(data = bechdel,aes(x = year, fill = binary)) +
 geom_bar(position = "fill")
```

The percentage of movies that pass the Bechdel Test has increased over time

Exercise 7

```
ggplot(data = bechdel,aes(x = year, y = domgross_2013, color = budget)) +
 geom_point(size = 0.75)+
  labs(title = "Are Movies Making More Over Time",
       subtitle = "faceted by whether movies passed Bechdel Test",
       x = "Year", y = "Domestic Gross", color = "Budget") +
  facet_wrap(~binary) +
  theme_bw()
```

Are Movies Making More Over Time
faceted by whether movies passed Bechdel Test