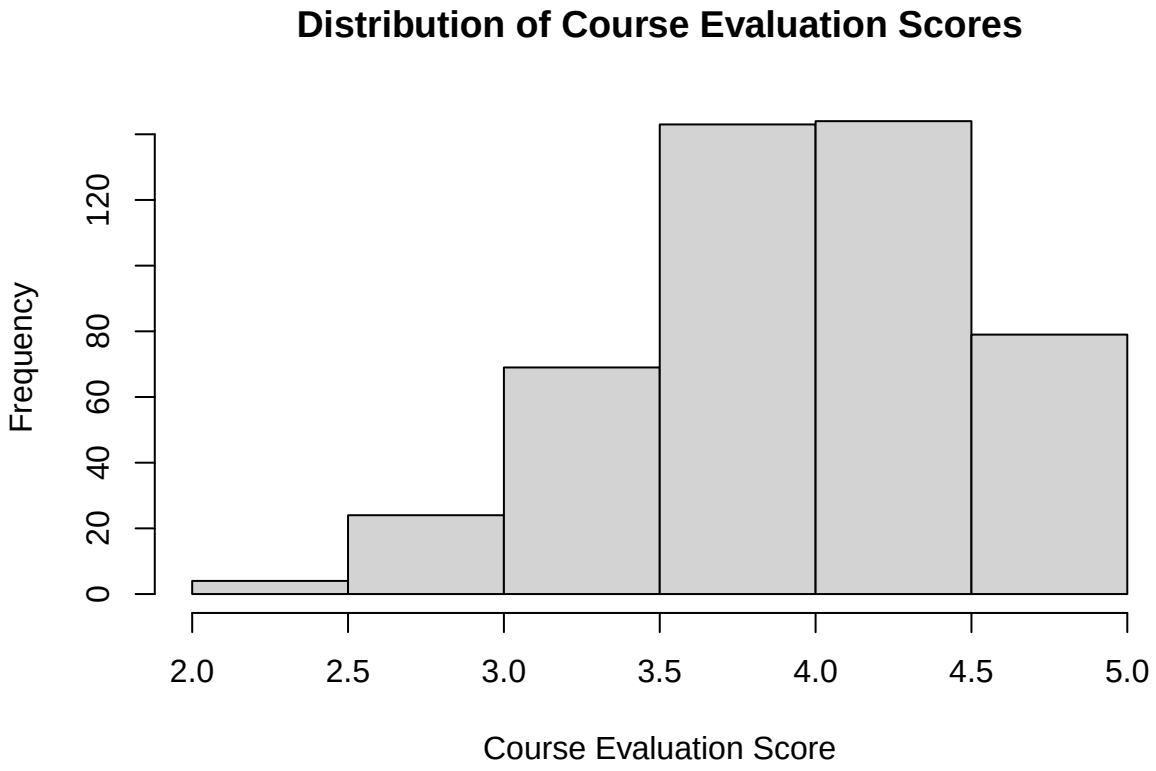
SDS 315 HW2

2024-01-21

Corinne Ward cw42962 https://github.com/corinneward/HW2

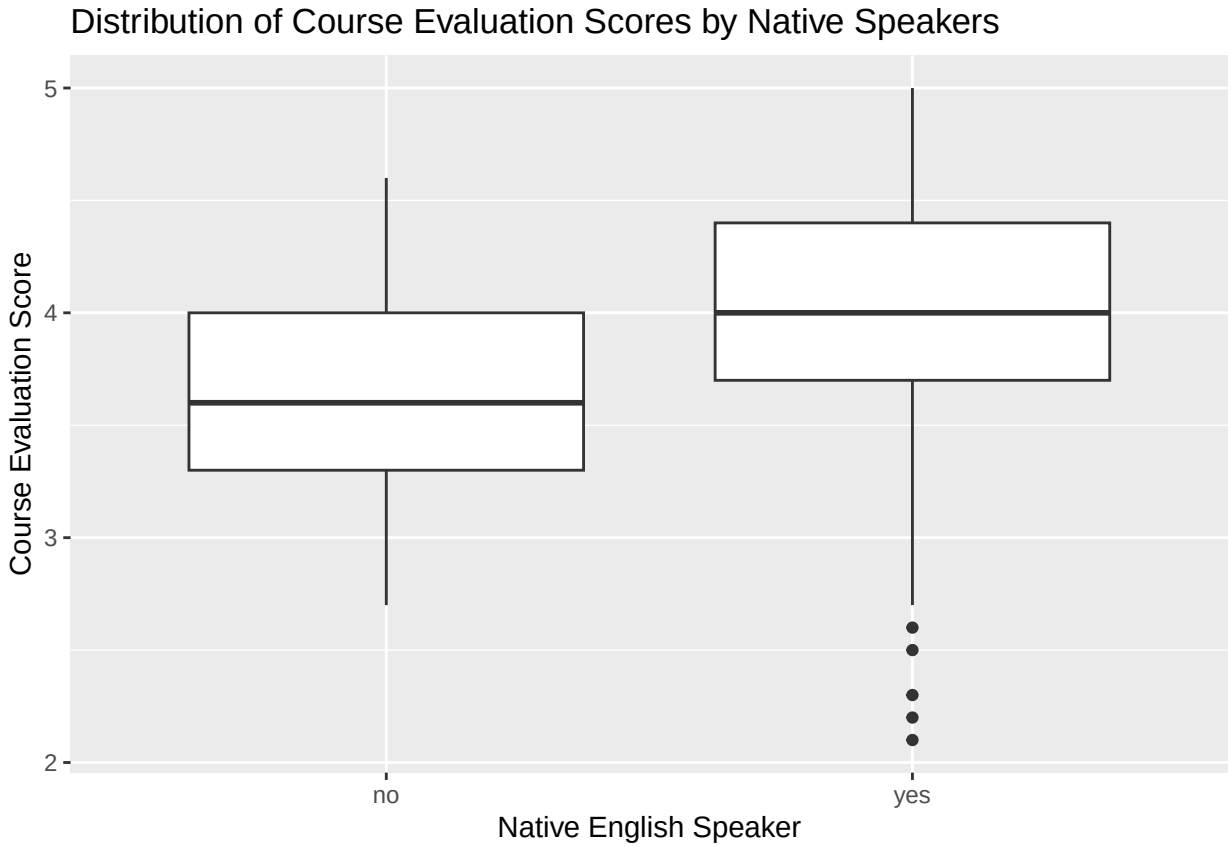
Problem 1

Part A:

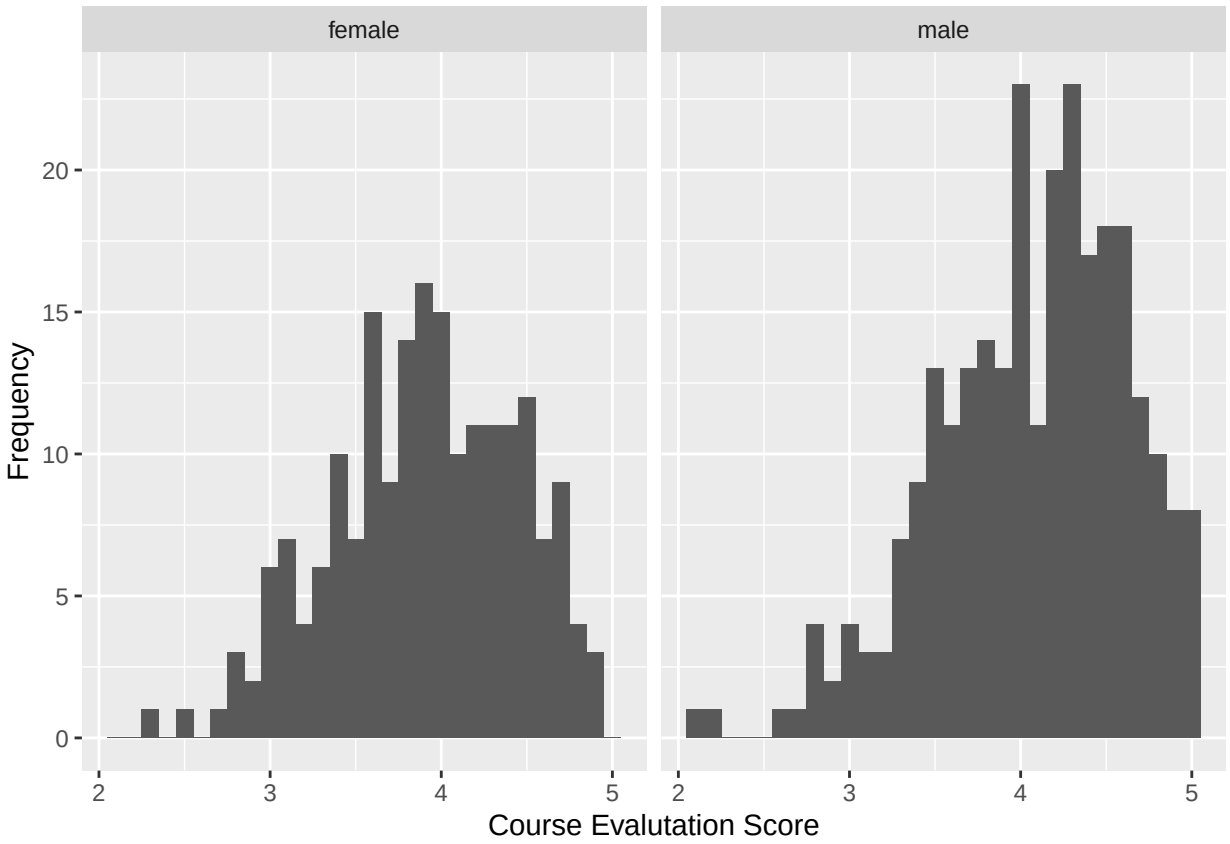


This is a negatively skewed histogram because the values are more concentrated on the right side of the graph and the mean, 3.998 is less than the median of 4.

Part B



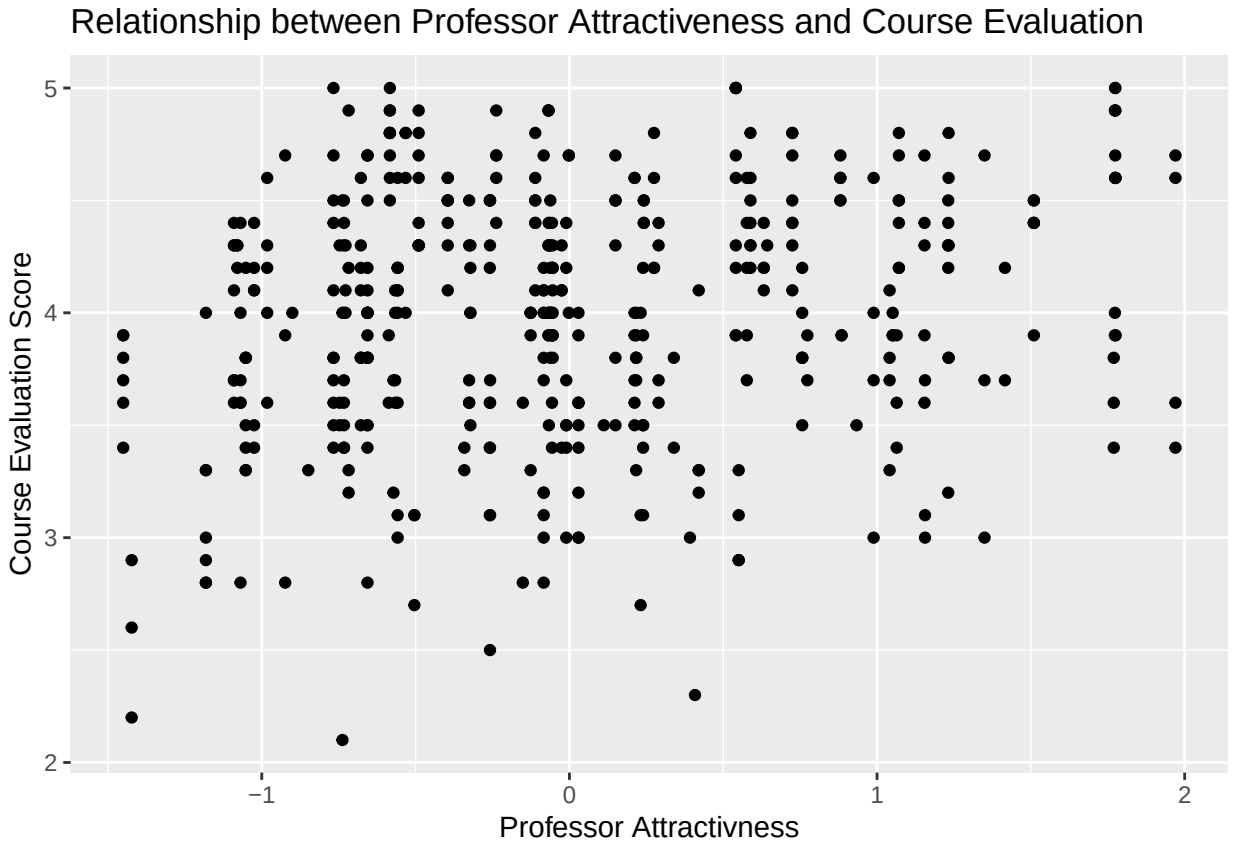
The mean course evaluation score for native English speaking professors is 4.018 while the mean course evaluation score for non-native English speakers is 3.689. The minimum course evaluation score for native English speakering professors was 2.1 and the maximum score was 5. The minimum course evaluation score for non-native English speaking professors was 2.7 while the maximum score was 4.6



Part C

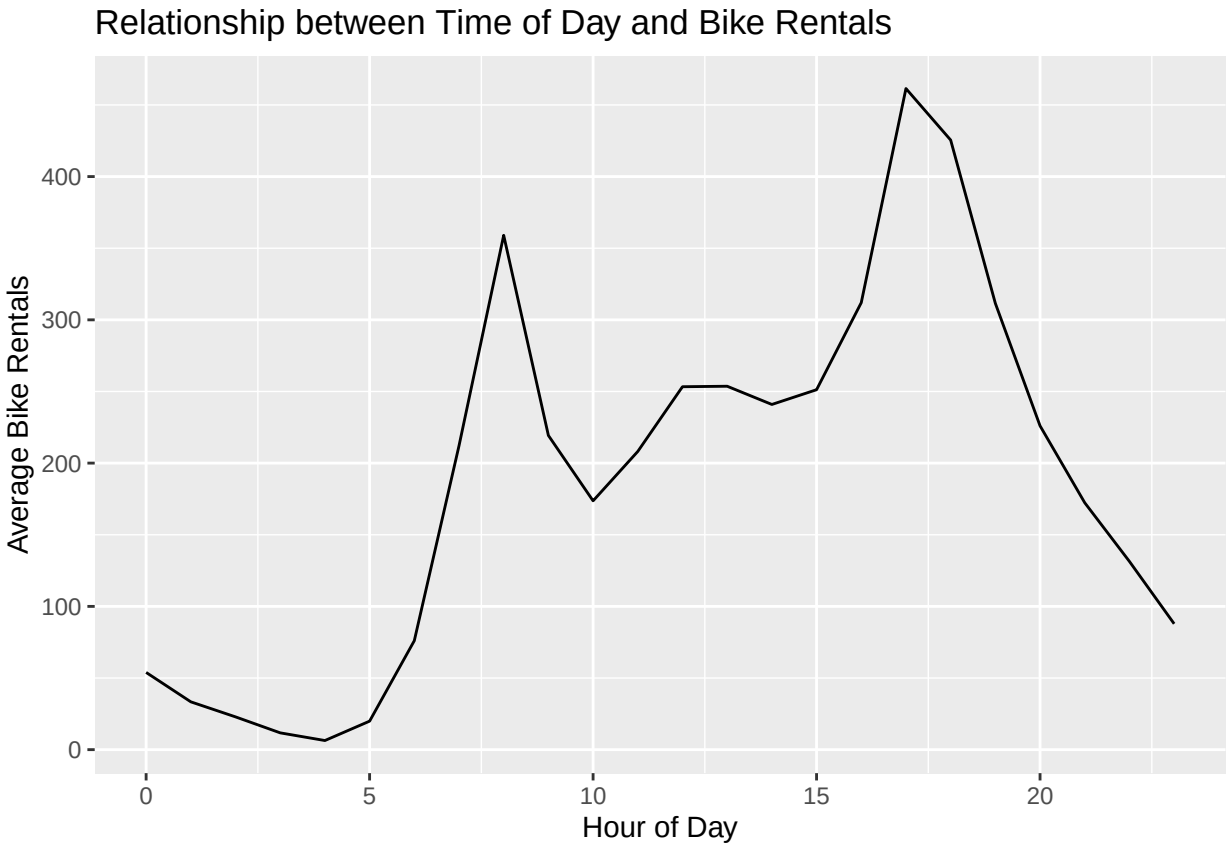
Both female and male professors have negatively skewed course evaluation scores. The mean course evaluation score for female professors is 3.901 and the median course evaluation score is 3.9. The mean course evaluation score for male professors is 4.069 and the median course evaluation score is 4.15.

Part D



The correlation between professor attractiveness and course evaluation score is 0.189. This means there is little to no correlation between a professors attractiveness and their course evaluation skills.

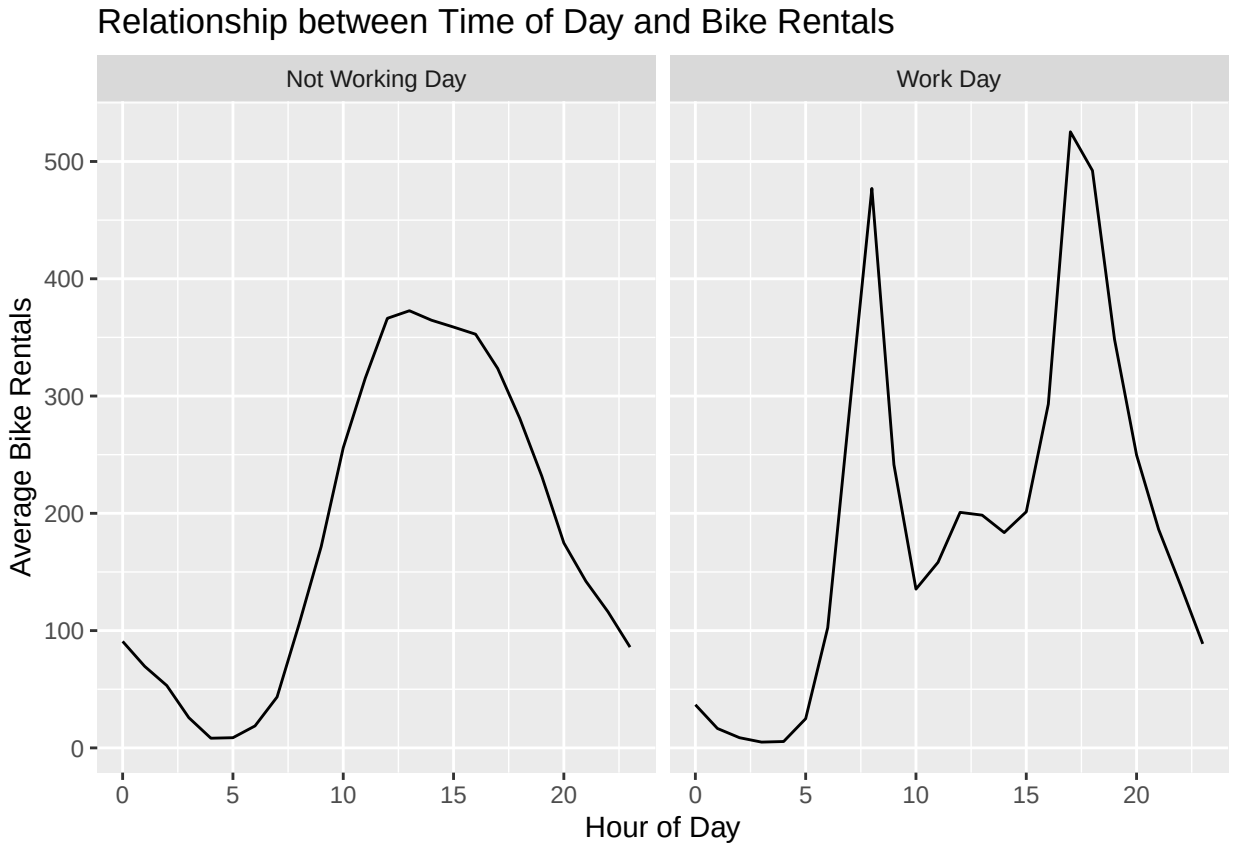
Problem 2



Plot A

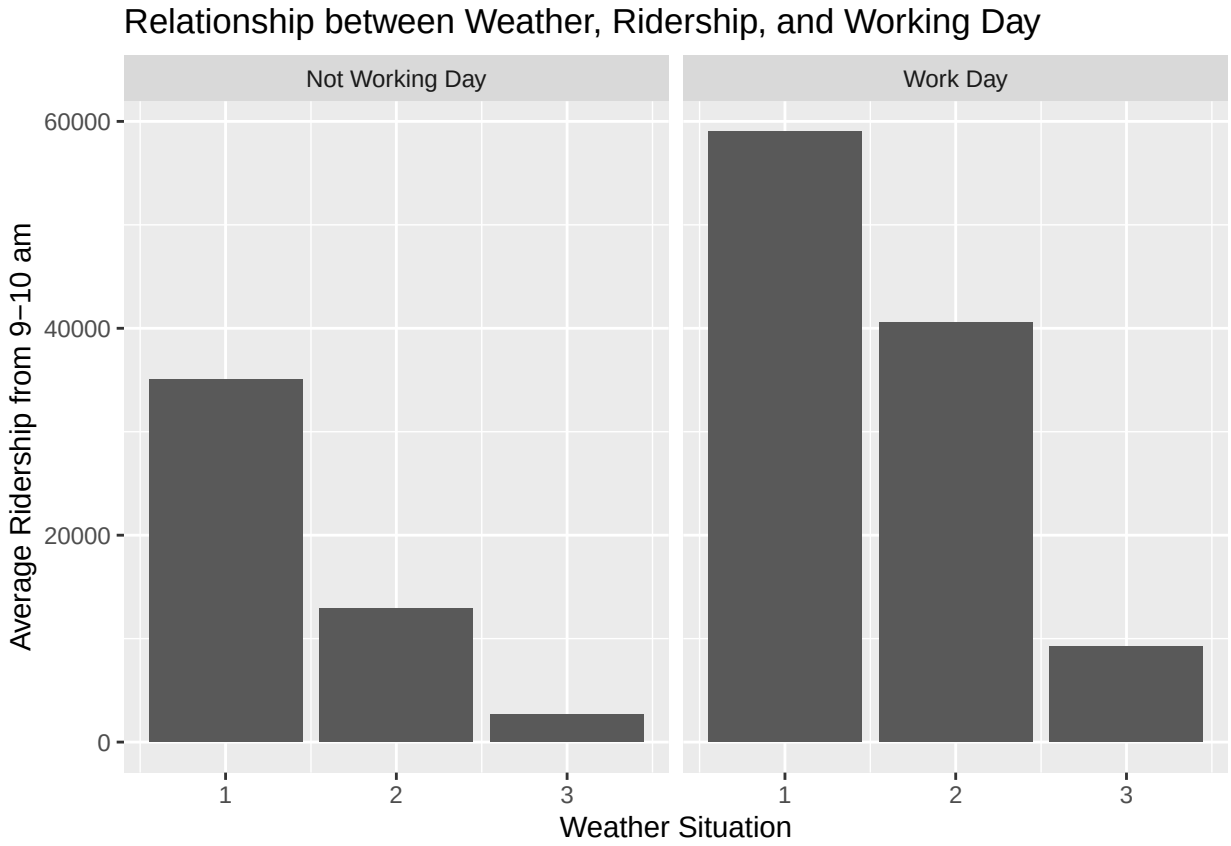
This plot shows the relationship between the time of day it is and the average amount of bike rentals at that time. Through looking at this plot, there are two spikes in the average bike rentals, one at around eight in the morning and another between 4 and 5 in the afternoon. This spike is probably due to people going to work and coming home from work. In addition to those spikes, it is shown that the average bike rentals is at its lowest between 2 and 5 in the morning probably due to sleeping hours. Average bike rentals have peaks around eight am and 5pm and gradually go down after 5.

Part B



This plot shows that the average amount of bike rentals on a non-working day is more gradual and constant than on a working day. This is probably due to a more relaxed start to the day. Additional, on a non-working day there are more average bike rentals in the early hours of the day, around midnight to 3am which is probably due to the fact that night life is more popular on non-working days. Ridership is more gradual and constant on non-working days while it has harsh peaks on working days.

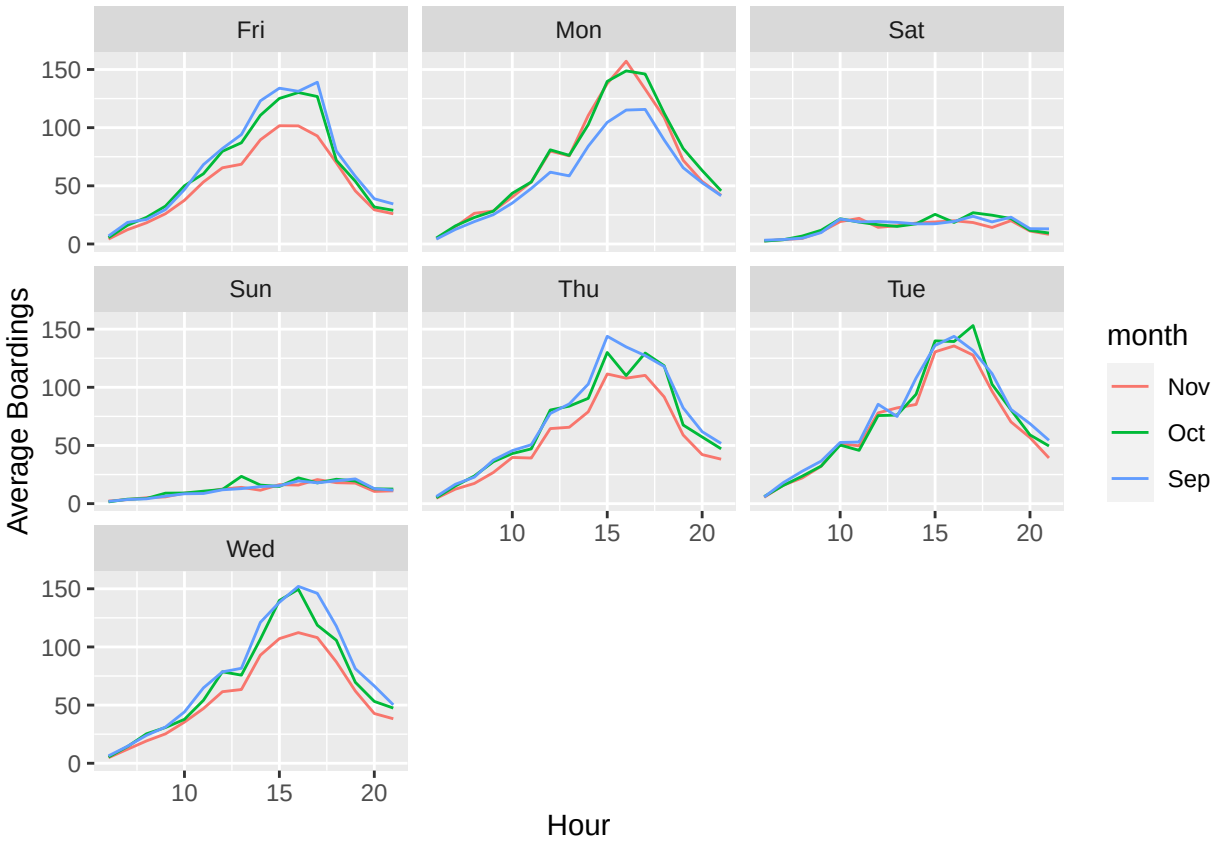
Part C



The weather situation for a one represents a clear to cloudy day, a two represents a misty day, a three represents a light rain or a light thunderstorm, while a four represents a heavy rain and heavy thunderstorm day. This graph shows that people are most likely to ride bikes on a clear work day however as the weather gets worse the likelihood of people riding bikes decreases. Additionally, this graph shows the importance a work day has on ridership between 9 and ten am. The weather has less of an impact on ridership on working days than on non-working days.

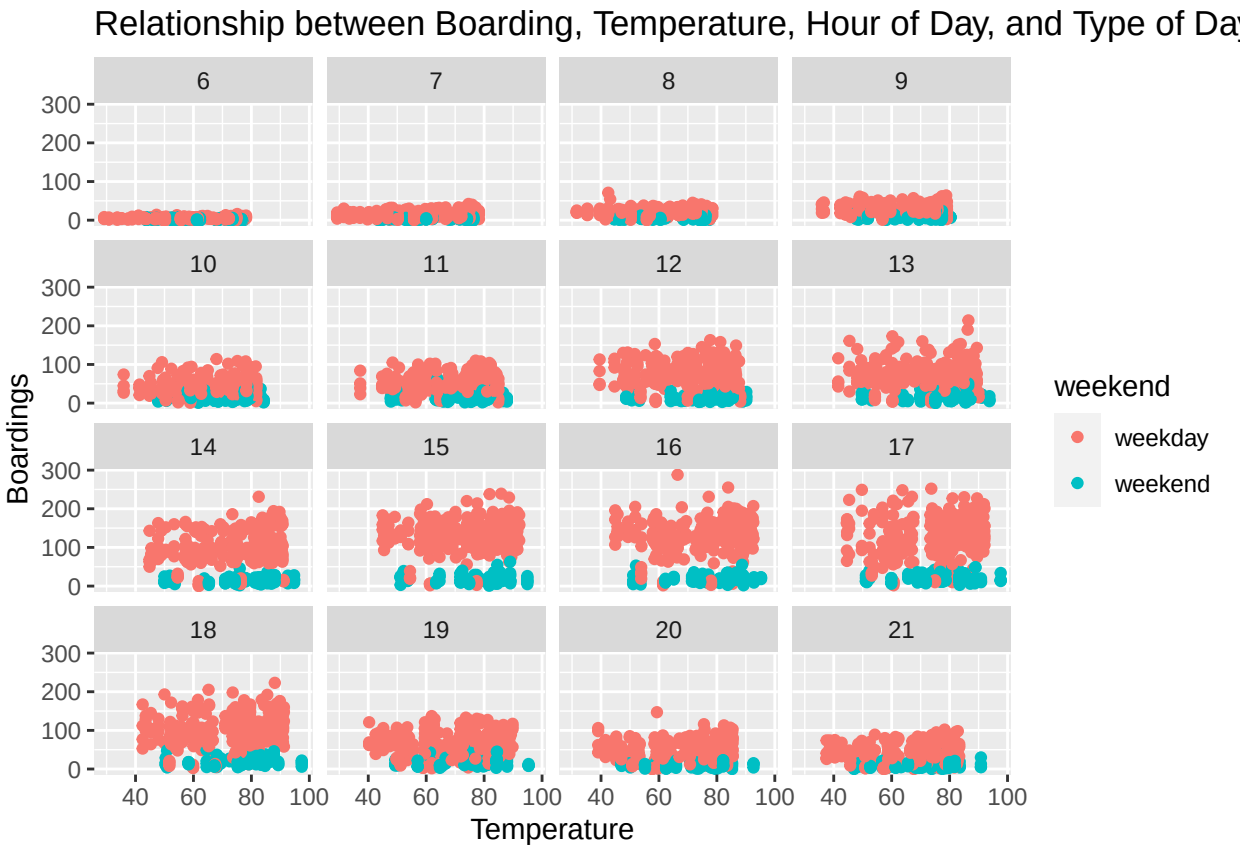
Problem 3

Part 1:



This figure shows the average boardings for the University of Texas bus service each day in the months November, October, and September. The hour of peak boarding is around the same time each weekday but significantly lower on the weekends. On Mondays in September, the average boarding is lower than the other months which is probably a result of no classes on labor day. A similar effect happens to Wednesday, Thursday, and Friday in November as a result of Thanksgiving break and no classes. This graph shows that there is higher average ridership on class days.

Part 2:



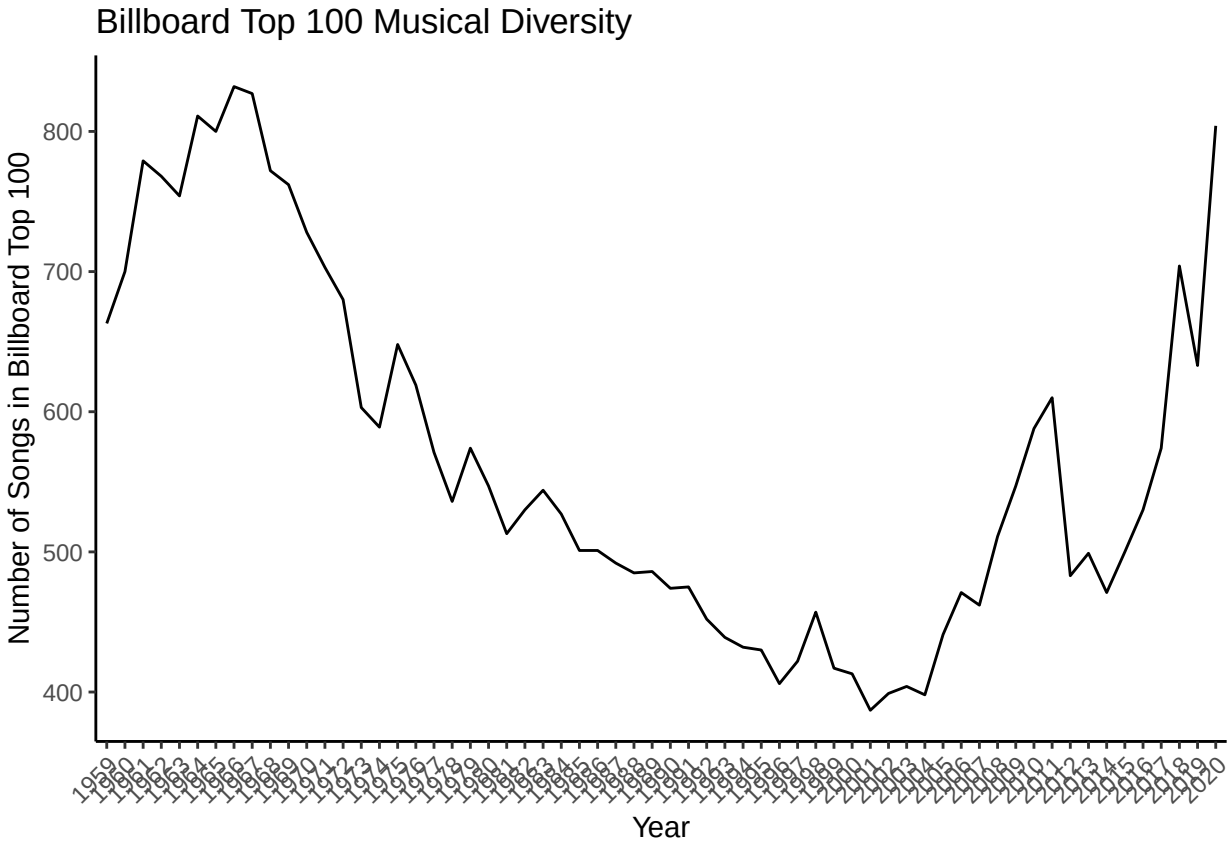
This figure shows the relationship between temperature, boarding, time of day, and type of day. Through this plot, it is clear that ridership, on average, is highest between three and six pm on a weekday, and it shows there is not a noticeable effect of temperature on ridership.

Problem 4

Part A

```
## # A tibble: 10 x 3
## # Groups:   performer [10]
##    performer                                      song                    total_weeks
##    <chr>                                          <chr>                         <int>
##  1 Imagine Dragons                                Radioactive                    3828
##  2 AWOLNATION                                     Sail                           3160
##  3 Jason Mraz                                     I'm Yours                      2926
##  4 The Weeknd                                     Blinding Lights                2926
##  5 LeAnn Rimes                                    How Do I Live                  2415
##  6 LMFAO Featuring Lauren Bennett & GoonRock      Party Rock Anthem              2346
##  7 OneRepublic                                    Counting Stars                 2346
##  8 Adele                                          Rolling In The Deep            2145
##  9 Jewel                                          Foolish Games/You Were~        2145
## 10 Carrie Underwood                               Before He Cheats               2080
```

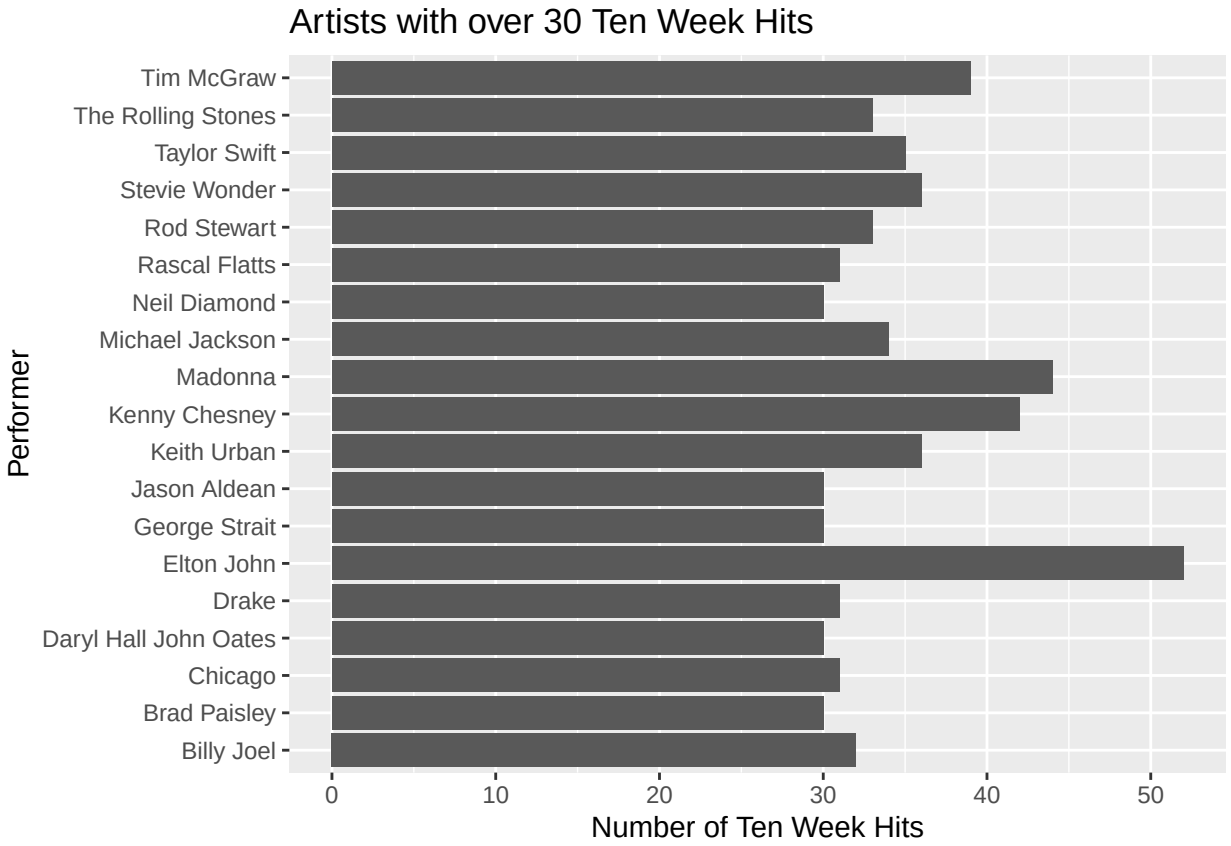
This table shows that the songs with the greatest number of weeks on the charts. The song that has spent the most weeks on the charts is “Radioactive” which is performed by Imagine Dragons.



Part B

This plot shows musical diversity throughout the years. Musical diversity is classified as the number of songs that appeared in the Billboard Top 100 that year. The plot shows there was more diversity around 1966 and then reached a low at and around 2001 and is beginning to increase again. Overall, musical diversity was at its lowest in the late 1990s to the early 2000s.

Part C



This plot shows artists who have more than thirty ten week hit songs. The most popular artist appears to be Elton John who has over fifty ten week hit songs.