HW 2

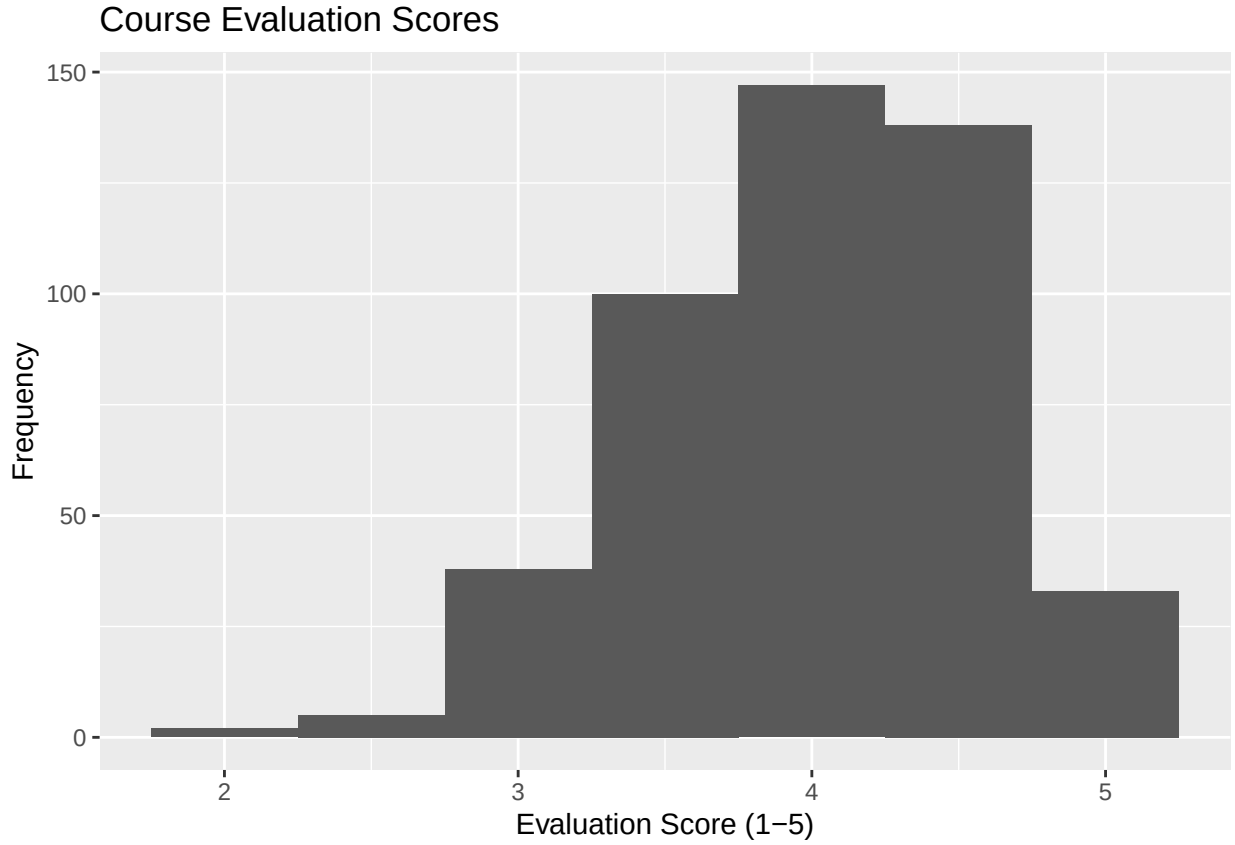
Maggie Ma

2025-01-26

Maggie Ma, mm227339, Github Link: https://github.com/maggiema9966/Homework-2

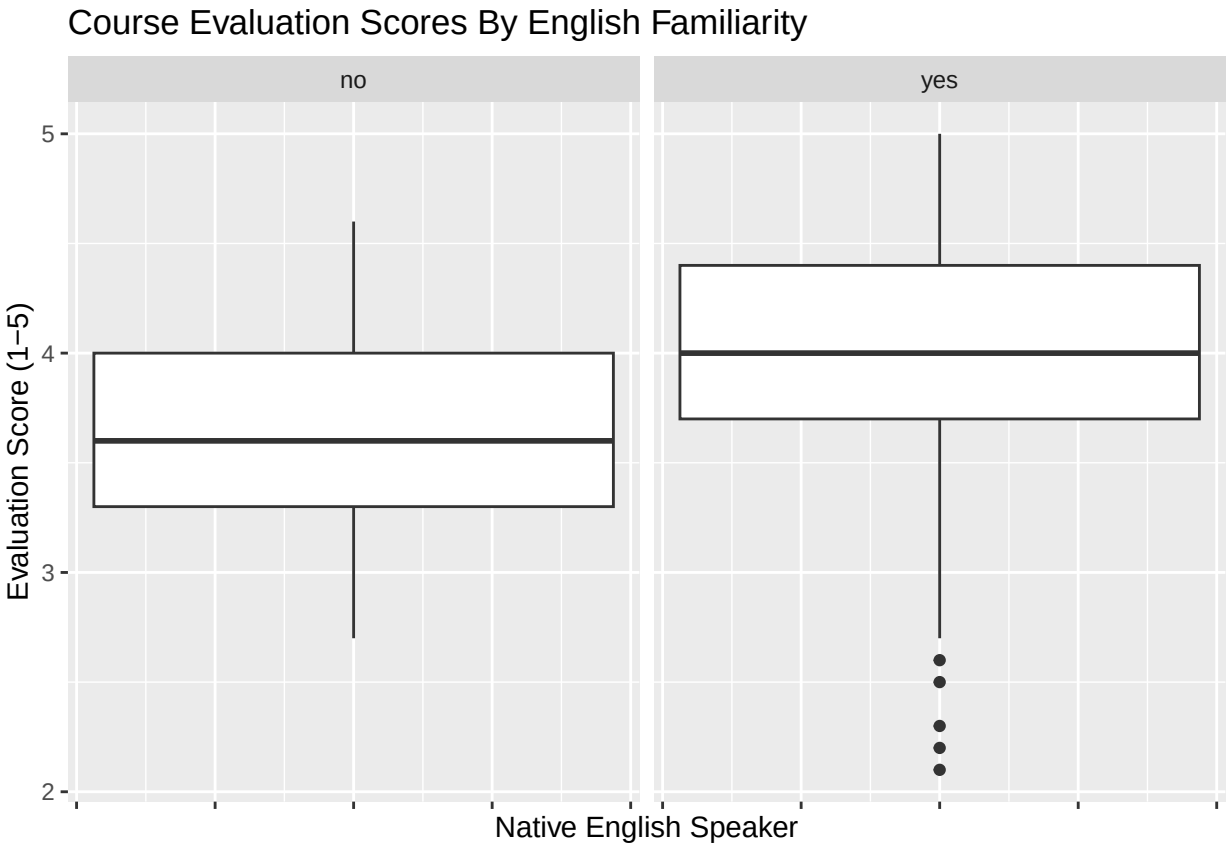
Problem 1: Beauty, or not, in the classroom

Part A



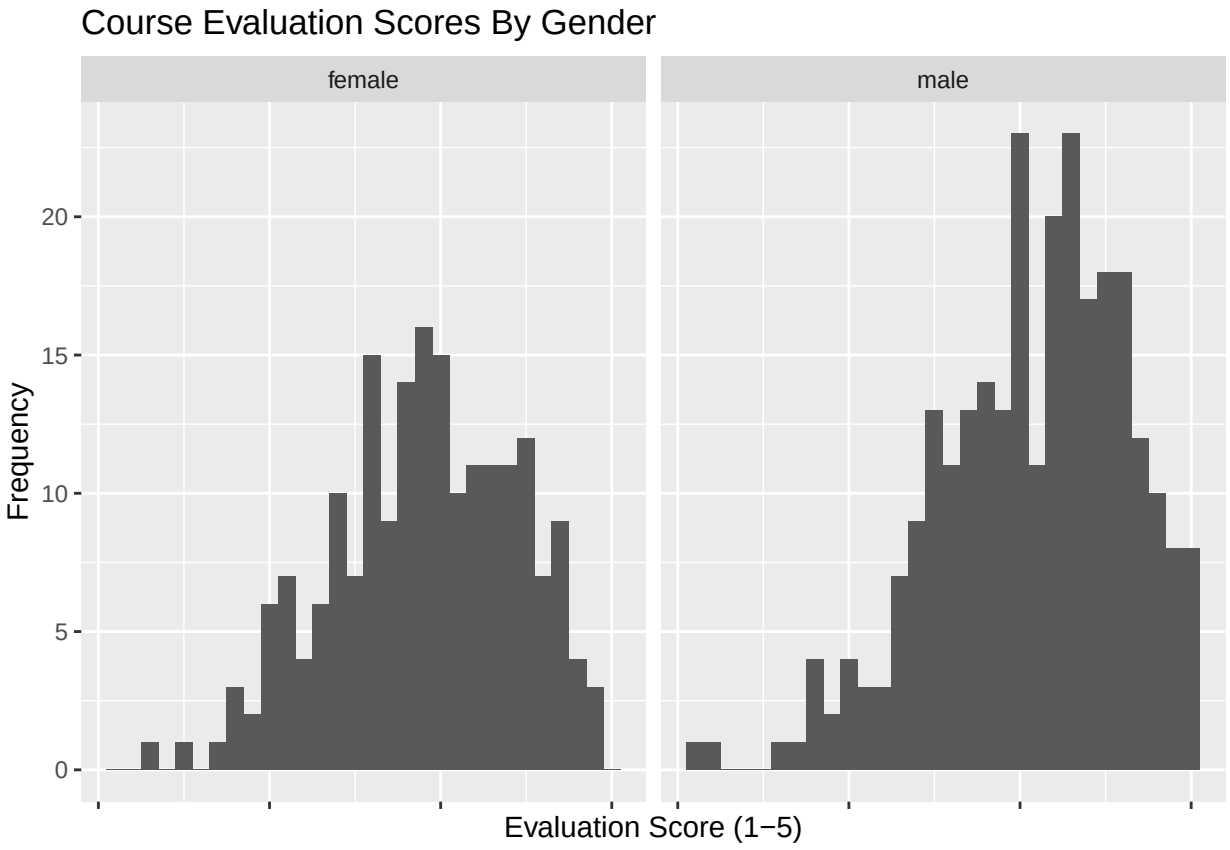
Histogram of the score distribution of how students rated their courses on a scale from 1-5. The scores seem to skew left, with no scores lower than 2.

Part B



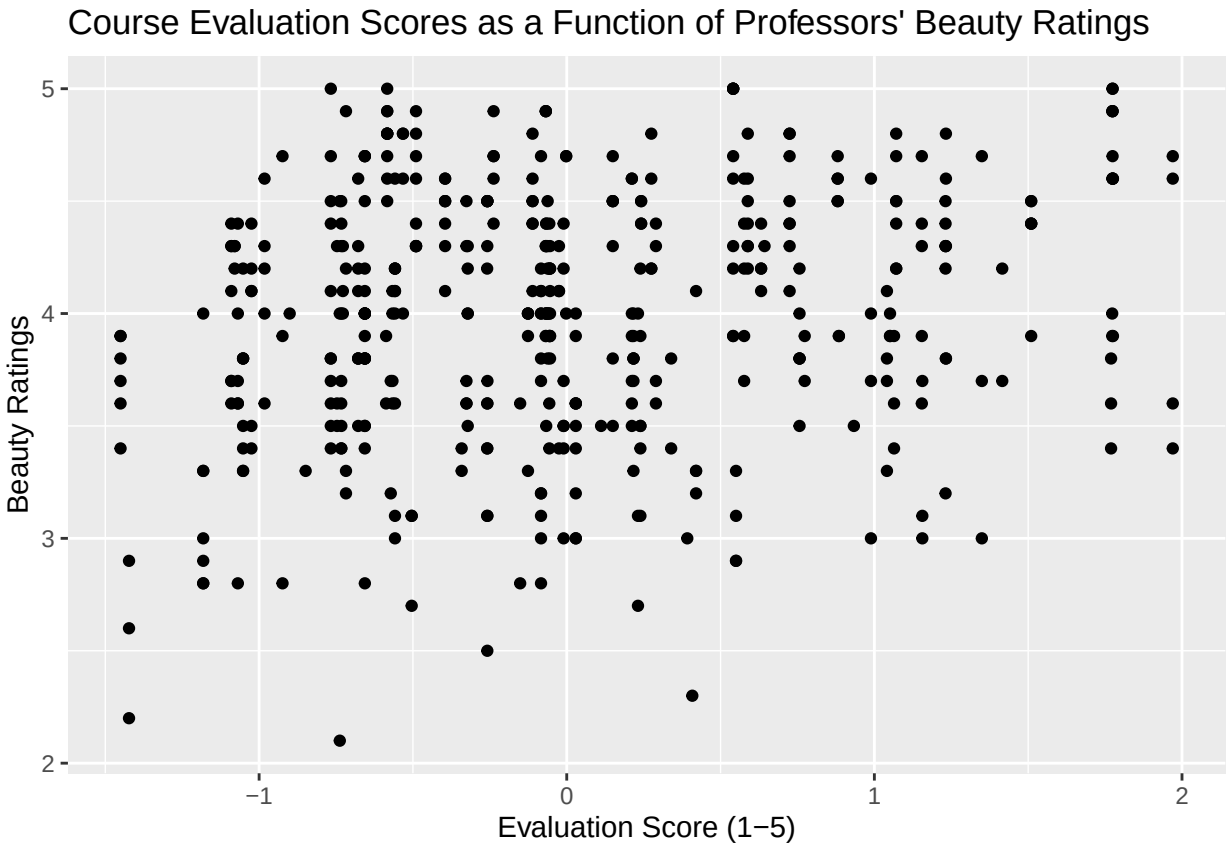
Boxplots of the score distribution of how students rated their courses on a scale from 1-5 based on whether the professor is a native English speaker. The median of the native English-speaking professors was slightly higher than the non-native English-speaking professors.

Part C



Histograms of the score distribution of how students rated their courses on a scale from 1-5 based on a professor's gender. Both distributions are skewed left, but there seems to be a larger sample of male professors than female professors.

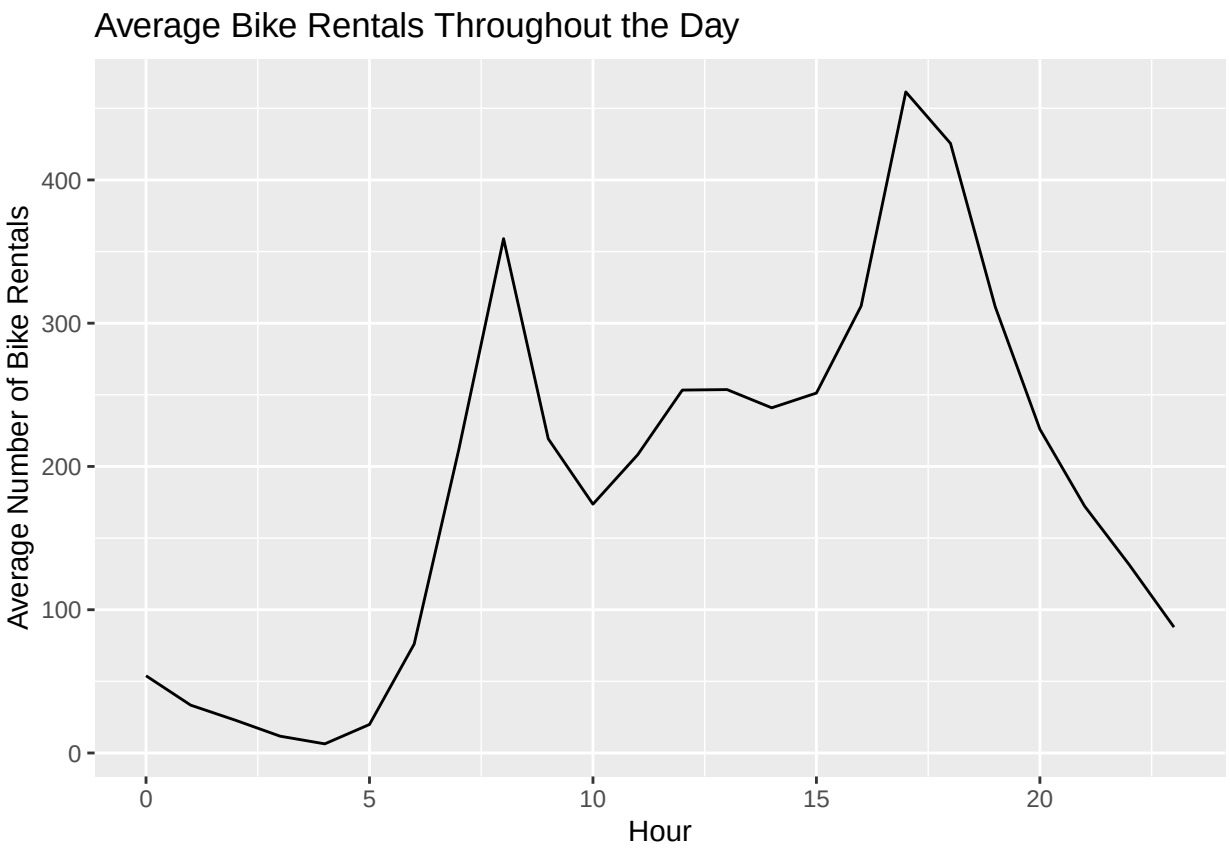
Part D



Scatterplot of the score distribution of how students rated their courses on a scale from 1-5 in relation to the professor's beauty ratings. There is a weak correlation between the two variables.

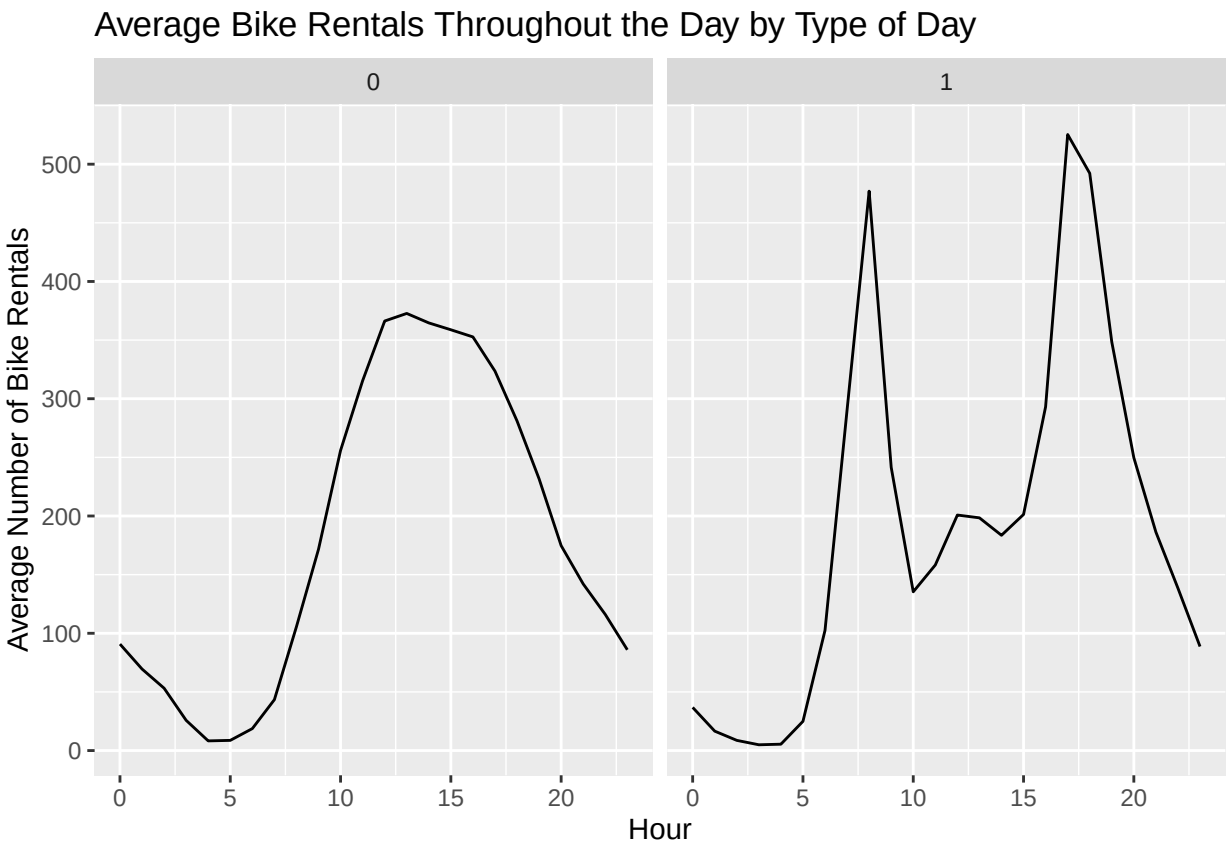
Problem 2: Bike Sharing

Part A



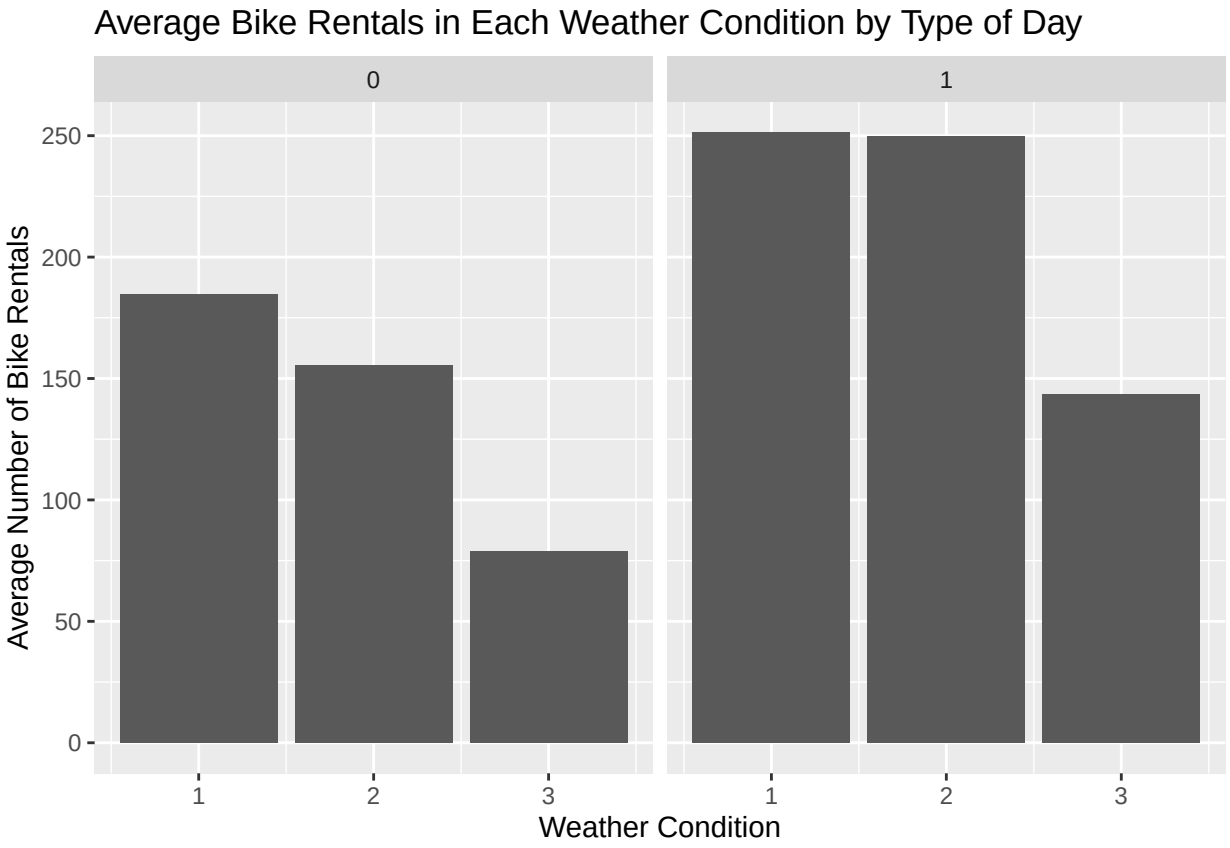
Line graph of the average number of bike rentals by each hour. There are 2 peaks in the number of bike rentals during hours 8 and 17.

Part B



Line graph of the average number of bike rentals by each hour based on whether the day is a work day (1) or not (0). The graphs seem to be almost inverses of each other, where it is usually busy during rush hours in the early morning and late afternoon during work days but around noon on weekends/holdiays.

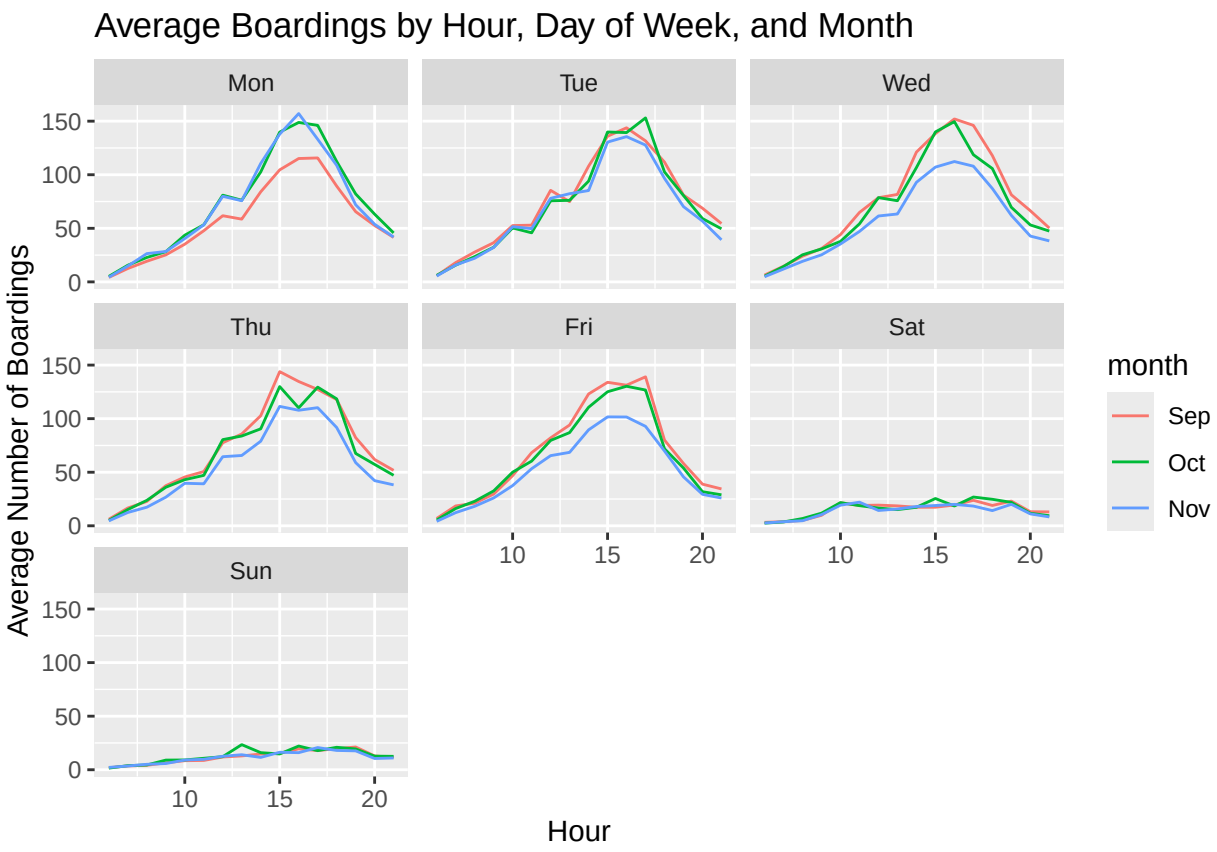
Part C



Bar graph of the average number of bike rentals by each weather condition (1: clear/partly cloudy; 2: misty; 3: light rain; 4: heavy/severe precipitation) based on whether the day is a work day (1) or not (0). There is a significantly less amount of bike rentals when there is precipitation on both work and non-work days.

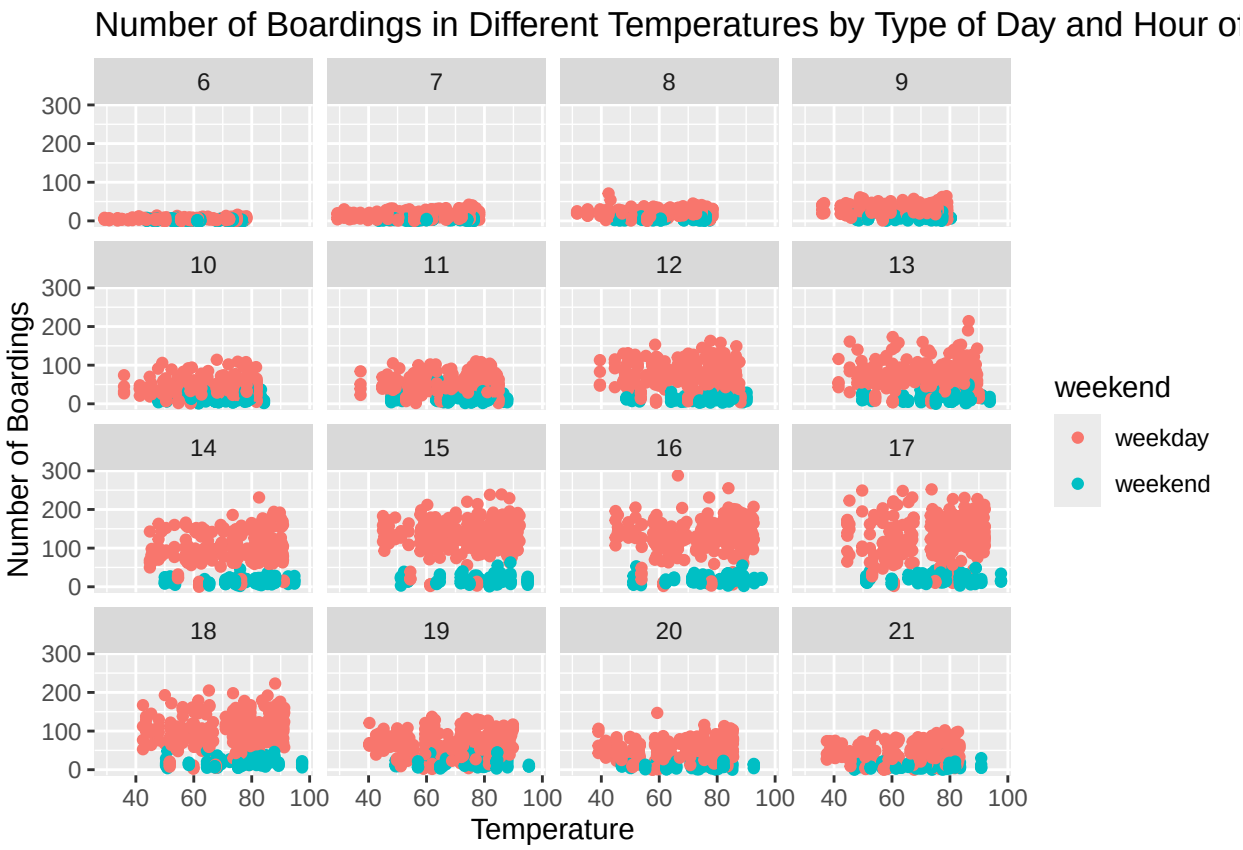
Problem 3: Capital Metro UT Ridership

Part 1



Line graph of the average number of boardings each hour based on each day of the week and the month. The number of boardings are much higher in all months during the weekdays, likely due to the need for transportation to work and school, while there are very few boarding numbers during the weekend. September Mondays have a lower peak, possibly due to the Labor Day holiday, where people do not have work or school. Similarly, November Wednesdays, Thursdays, and Fridays have a lower peak, possibly due to people taking work off for Thanksgiving.

Part 2



Scatterplot of the number of boardings each temperature range based on the type of day (weekday or weekend) and hour of day. The number of boardings are much higher in all hours of the day during the weekdays, and the number of boardings increases as the hours increase and begin decreasing after hour 17. However, when holding the hour of day and weekend status constant, there does not seem to be a significant difference in the number of boardings, which means temperature does not have a huge impact on the number of boardings.

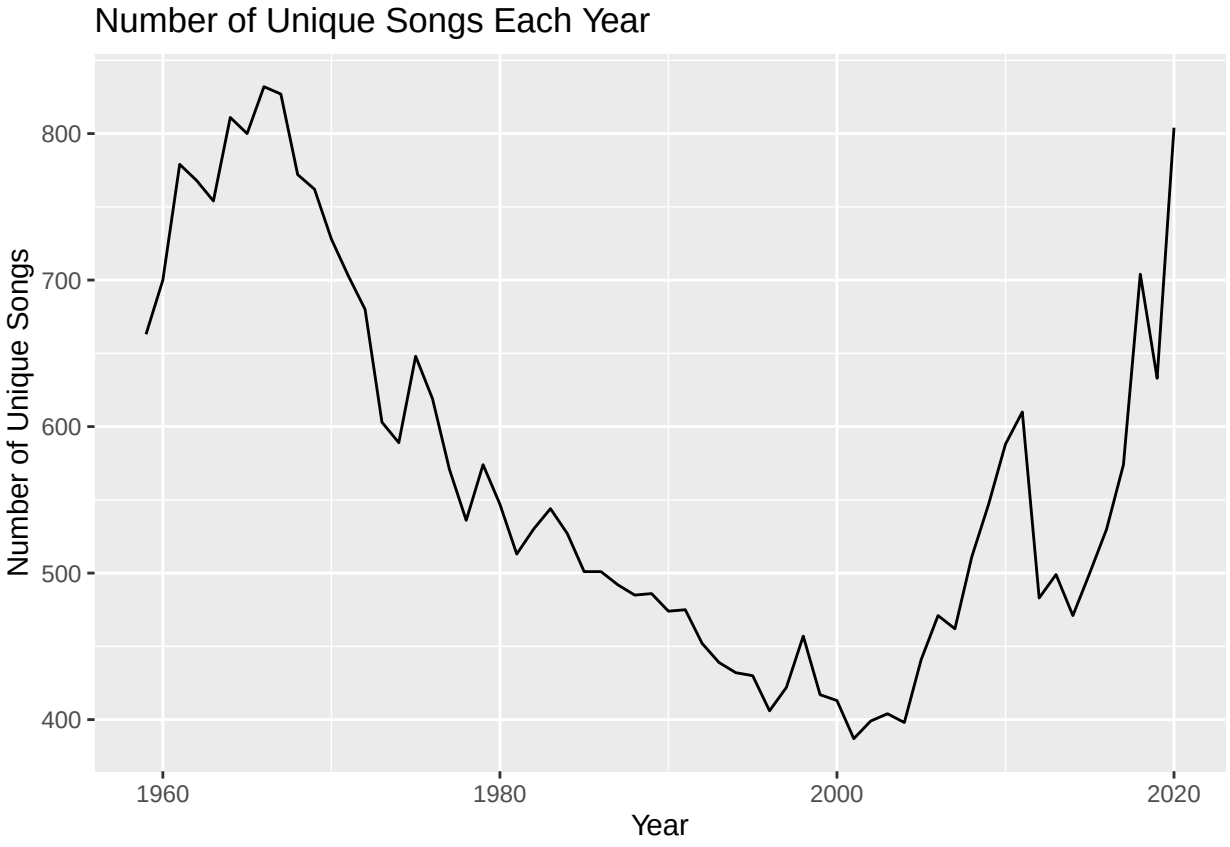
Problem 4: Wrangling the Billboard Top 100

Part A

```
## # A tibble: 10 x 3
## # Groups:   performer [10]
##    performer                                song                          count
##    <chr>                                    <chr>                         <int>
##  1 Imagine Dragons                          Radioactive                      87
##  2 AWOLNATION                               Sail                             79
##  3 Jason Mraz                               I'm Yours                        76
##  4 The Weeknd                               Blinding Lights                  76
##  5 LeAnn Rimes                              How Do I Live                    69
##  6 LMFAO Featuring Lauren Bennett & GoonRock Party Rock Anthem               68
##  7 OneRepublic                              Counting Stars                   68
##  8 Adele                                    Rolling In The Deep              65
##  9 Jewel                                    Foolish Games/You Were Meant~    65
## 10 Carrie Underwood                         Before He Cheats                 64
```

Table of the top 10 most popular songs (based on how many weeks the song was on the Billboard Top 100). The table displays the song title, artist, and the number of weeks the song was on the Billboard Top 100.

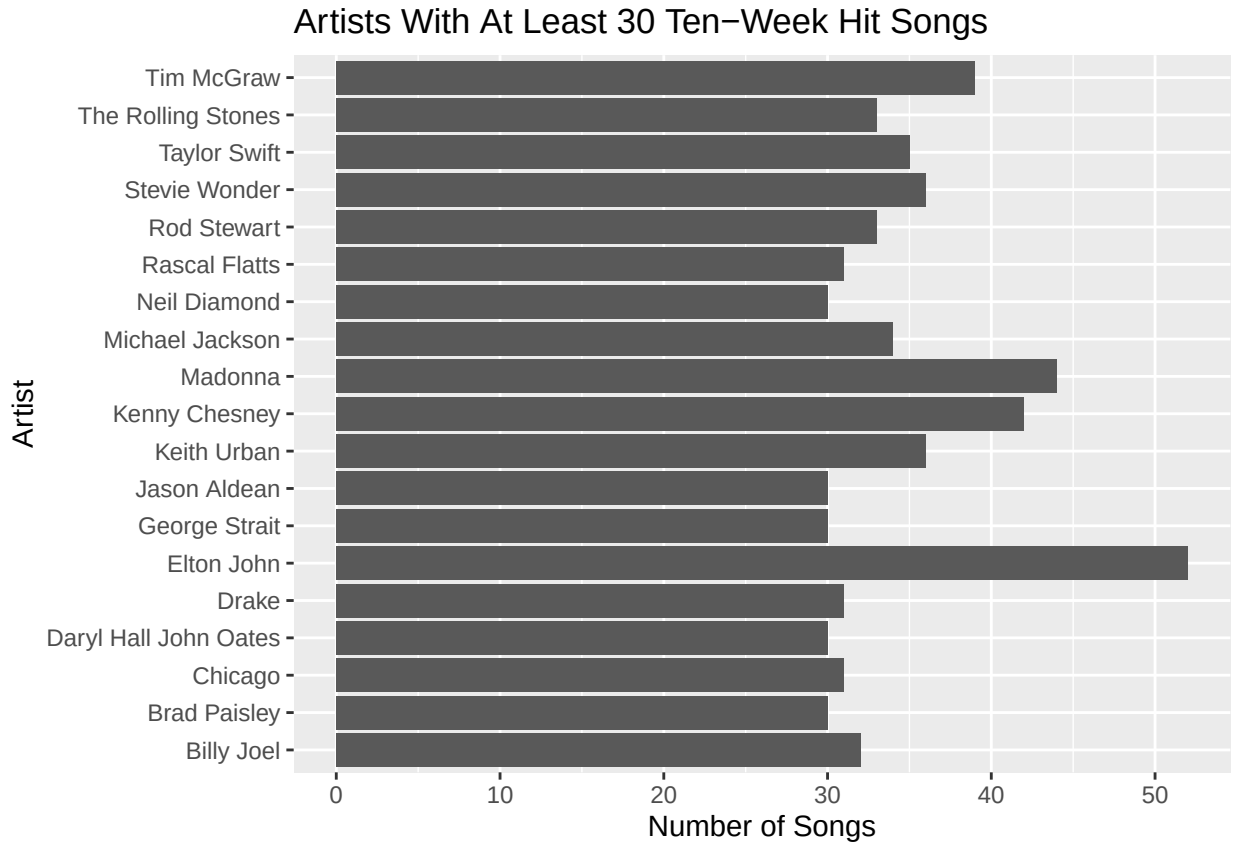
Part B



Line graph of the number of unique songs in each year. After ~1965, the number of unique songs plummeted, with very few unique songs by the end of the 2000s. However, it begins to

rise again after ~2005 and reached a peak of a similar number of unique songs by 2020.

Part C



Bar graph of the various artists that had at least 30 songs that were a ten-week hit. Elton John stands out with the highest number of ten-week hits.