Statistical Thinking HW 2

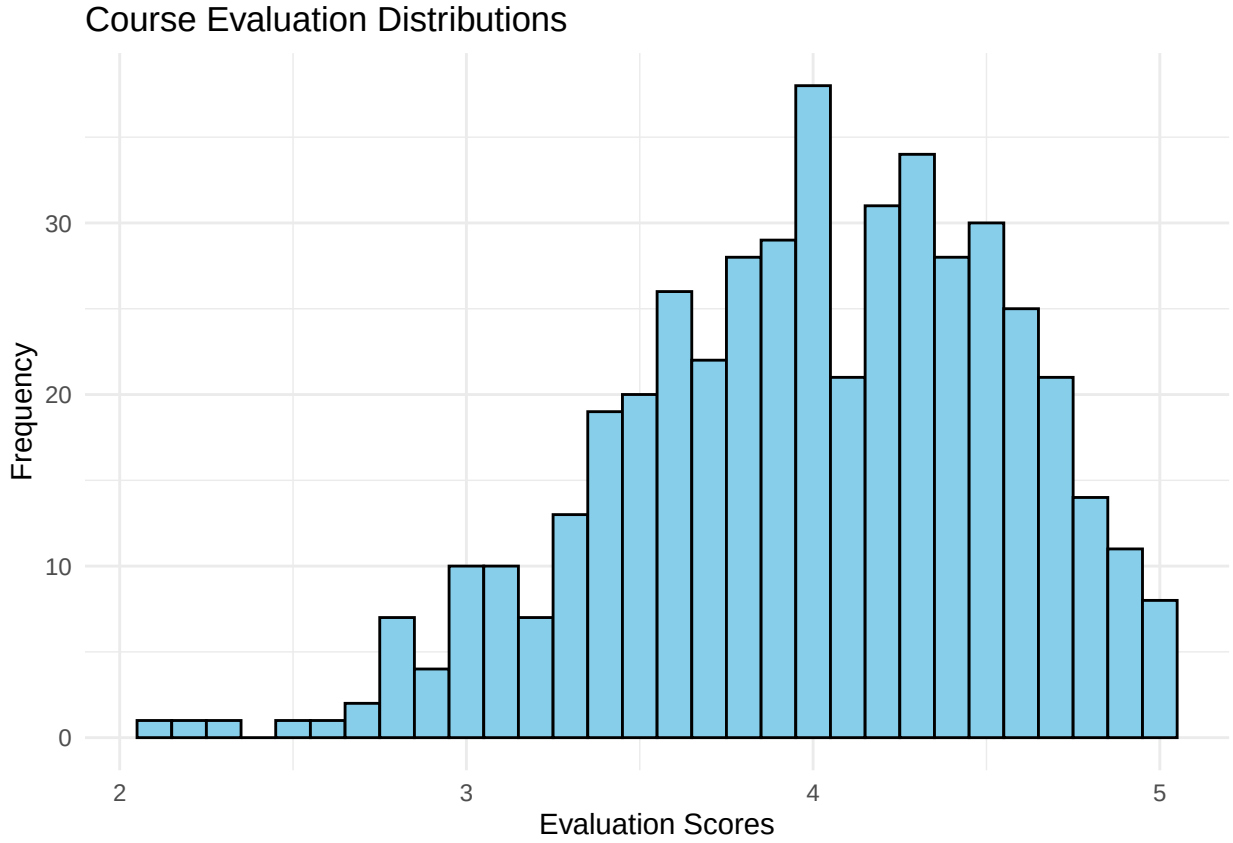
Estrella Fuentes eaf2758

2025-01-25

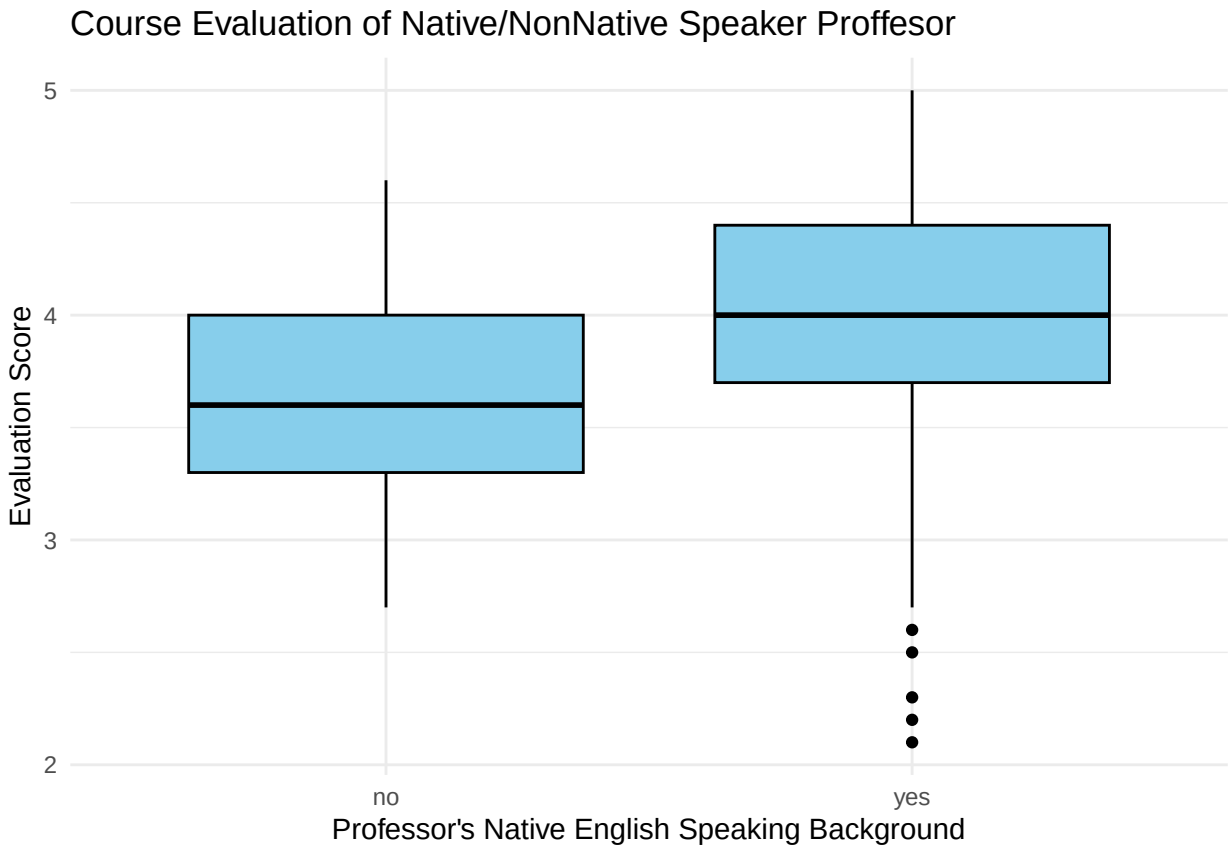
Contents

Problem 1: Beauty, or not, in the classroom 1
Problem 2: Bike Sharing 5
Problem 3: Capital Metro UT Ridership 8
Problem 4: Wrangling the Billboard Top 100 9

Problem 1: Beauty, or not, in the classroom

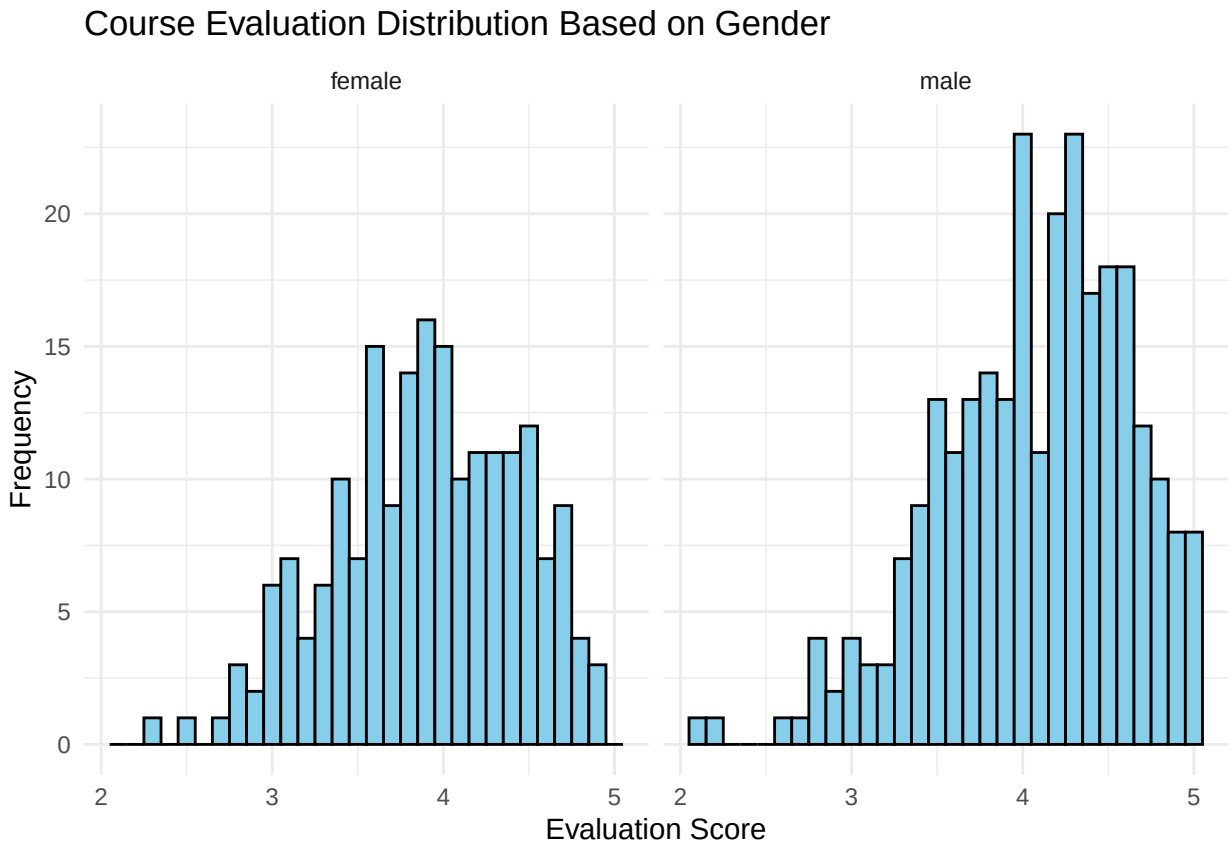


Part A -
The histogram below shows the distribution of the evaluation scores given to the professors. We notice the highest score is a 4 and the graph appears to be more skewed to the left than right indicating the professors likeness.



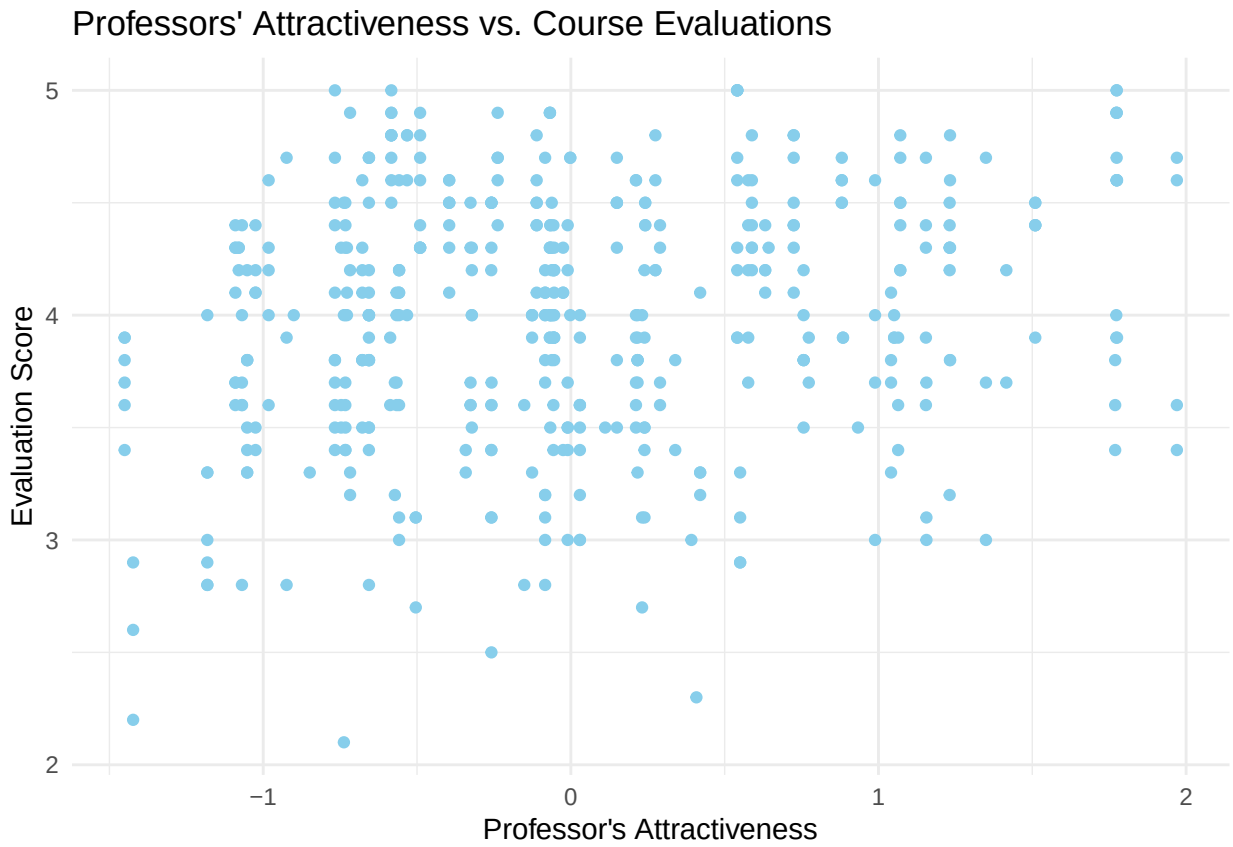
Part B -

The box plots below demonstrate the course evaluation distributions as well as the outliers for native and non native English speakers. By analyzing the box plots we can see native English speakers have a higher score on their course evaluations while also having many outlier scores compared to the non native English speaking professors.



Part C -

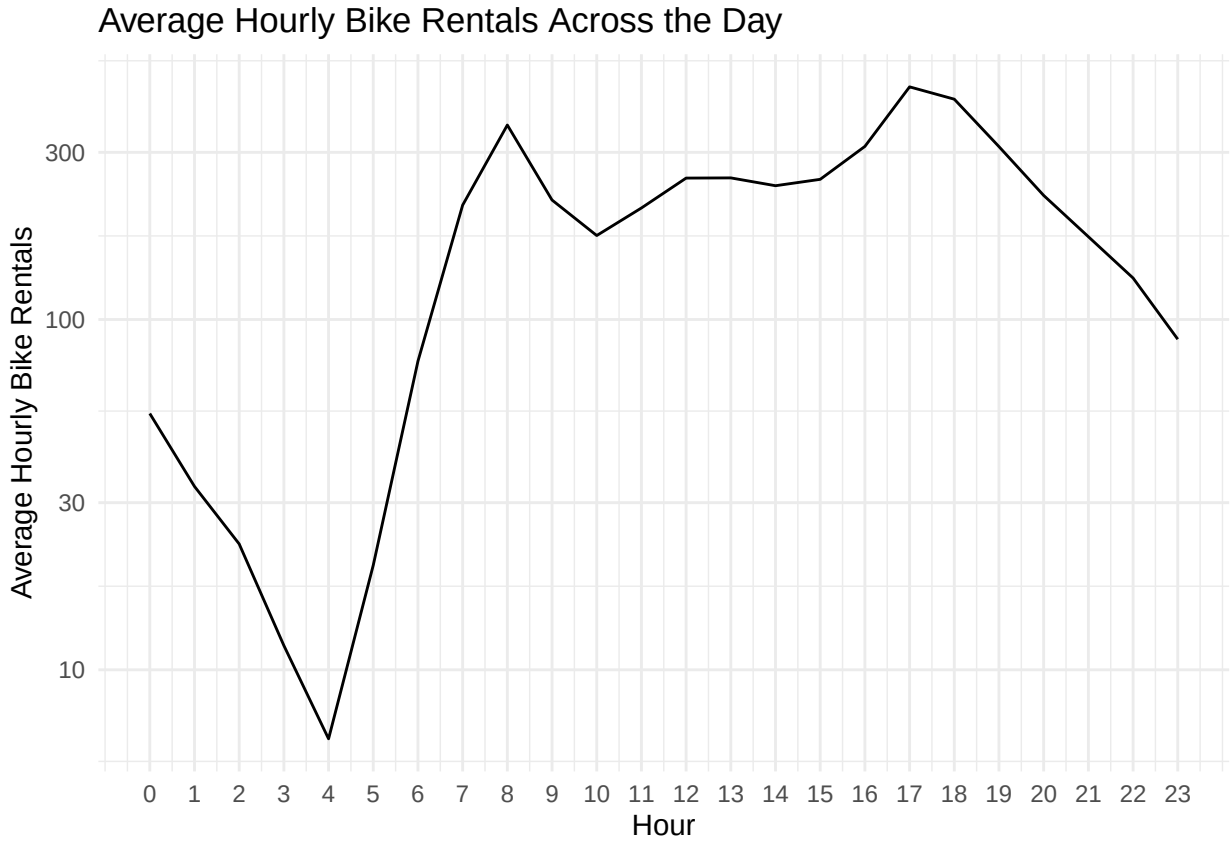
The histograms below demonstrate the course evaluation distribution based on gender. I notice that male professors have a more rightly skewed histogram but this could be due to the male to female professor ration. Other than the high peaks in the male professor histogram both graphs seem pretty similar.



Part D -

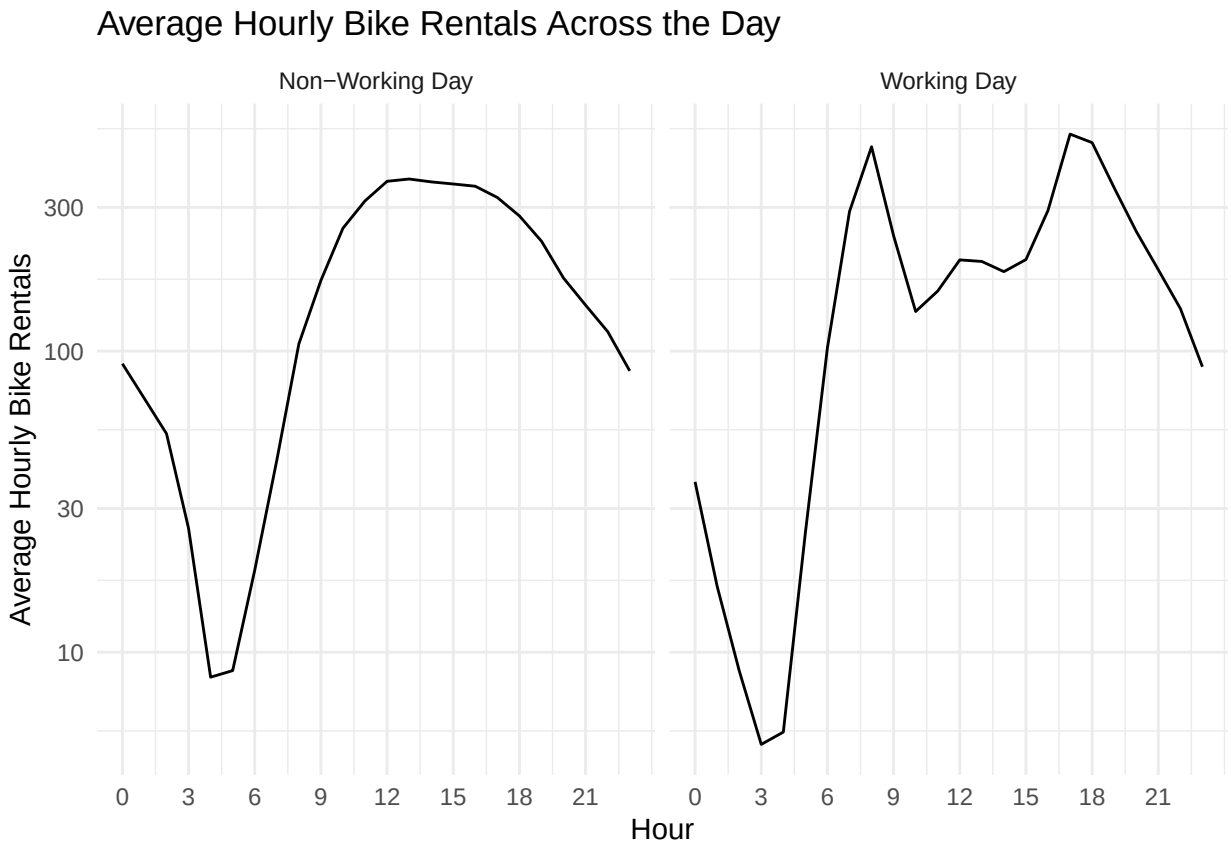
The scatter plot below show the correlation between the professors course evaluation and their attractiveness. By the look of the scatter plots their seems to not be a relevant correlation because of the lack of dot grouping or any display of a pattern.

Problem 2: Bike Sharing



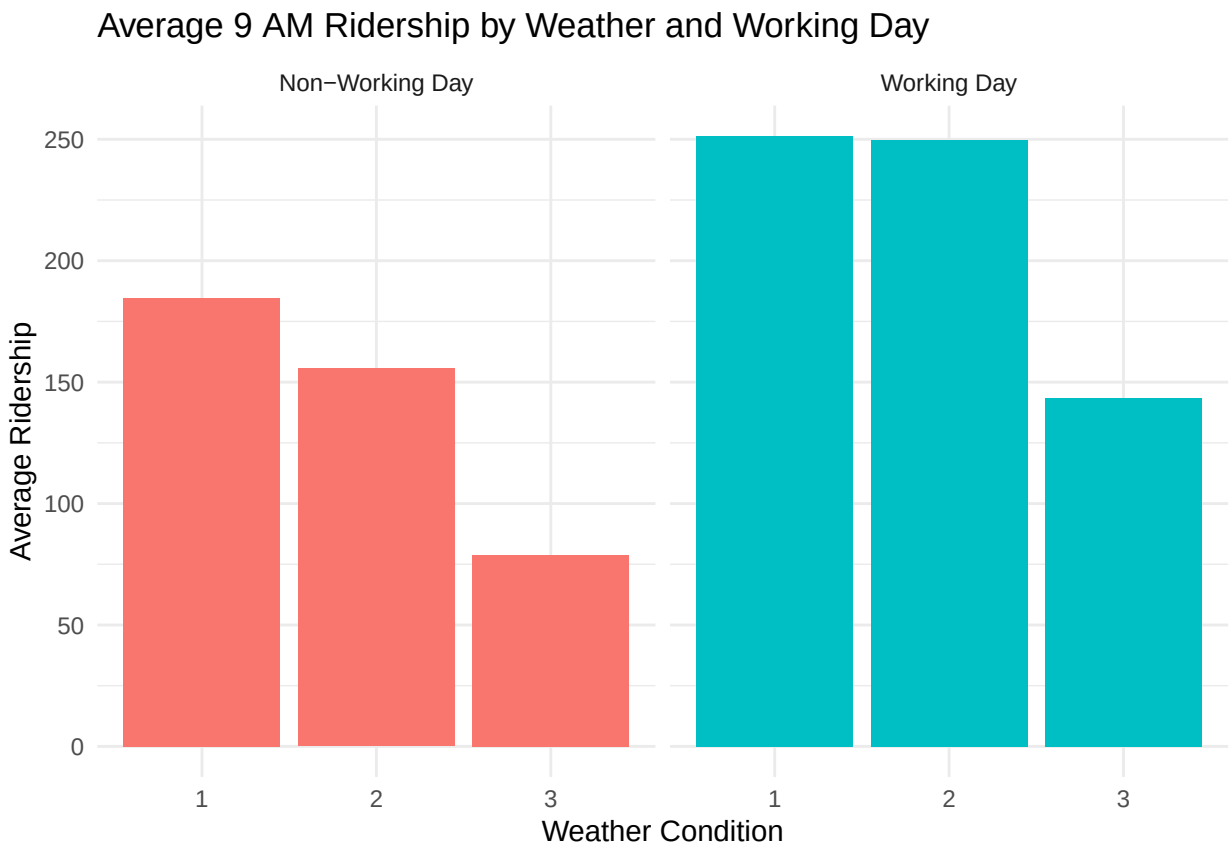
Part A -

The line graph below demonstrate the average hourly bike rentals across all hours of the day. The high point of the graph demonstrate that bikes are used the most during that time and when the point is at its lowest then it demonstrates the time the bikes are the least used.



Part B -

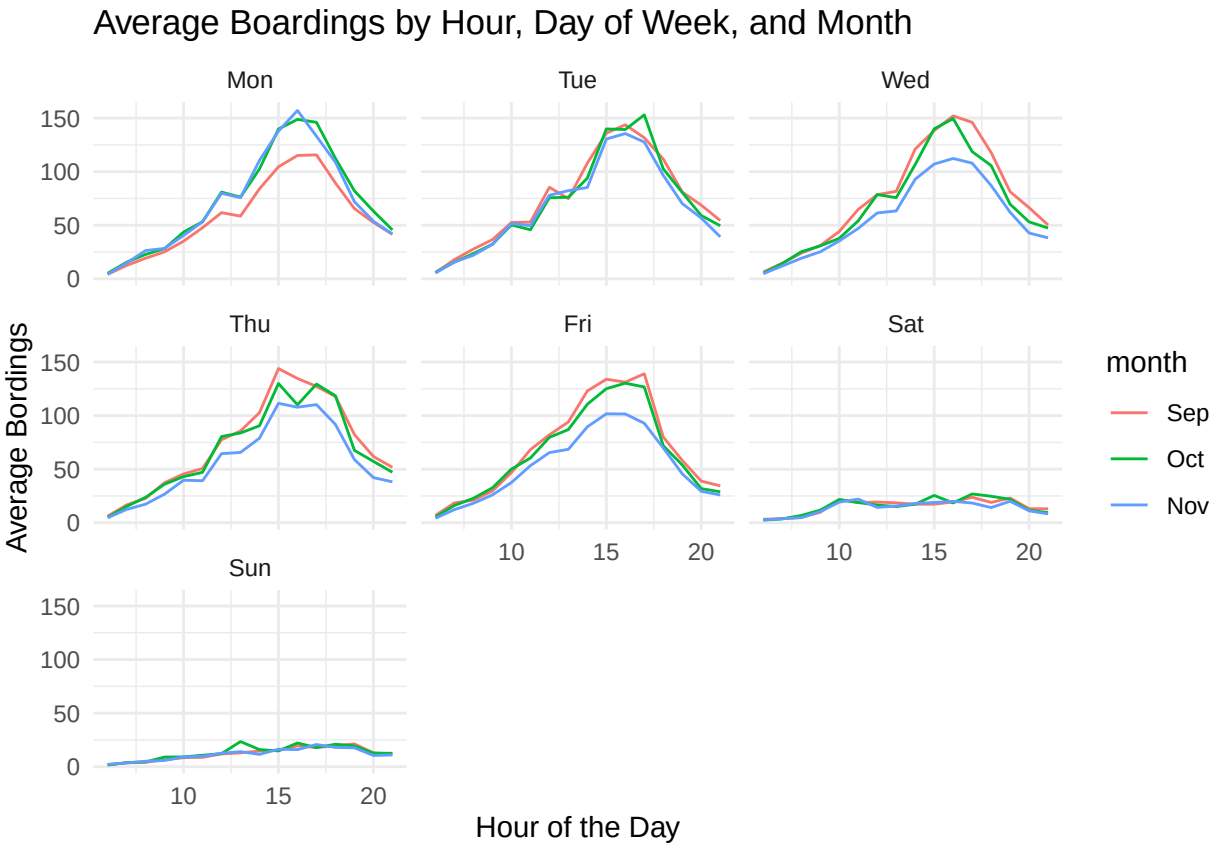
The line graphs below are similar to the ones above, they both demonstrate the average amount of hours bike rental are used across the day. The graphs below are different than the one above because these graphs are faced on whether whether the day was a weekend or holiday. Graph 1 shows the days the days that were neither a holiday and weekend while graph 0 show the weekends and holidays.



Part C -

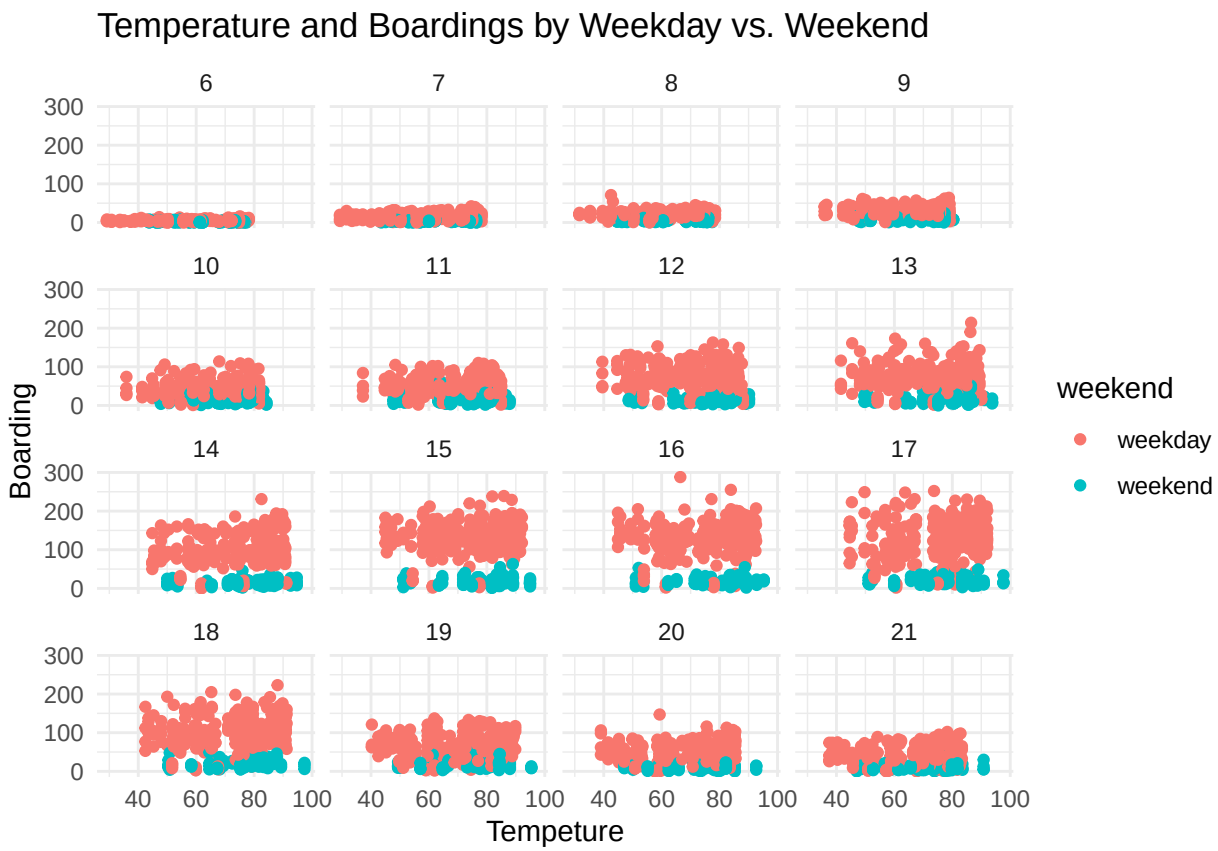
The bar plots below show the average ridership at 9 am on different weather condition days, and the two graphs are split on whether it was a working day or weekend/holiday. By the looks of the graphs their is more ridership on working days throughout all of the weather conditions than on nonworking days.

Problem 3: Capital Metro UT Ridership



Part A -

The figures above are faceted by day of the week, the three lines demonstrate the 3 different months and their average of the boarding per hour throughout that day of the week. The hour of peak boarding changes slight when it comes to day to day, the weekdays have similar peak out mostly likely because its the end of the school and work day. Weekends on the other hand appear to having minimal peak but the average boarding does pick up in the evenings similar to the weekdays. The average boarding on Weds/Thurs/Fri are lower probably because of the holidays and the fact that students are on break so the amount of boarding diminish because student and the majority of the population are at home. On September the average boarding could have been lower than the other months because of a couple of reasons like labor day happening and more significantly it is the begging of the school year so a lot of student could be getting situated to the campus rather then venturing out.



Part B -

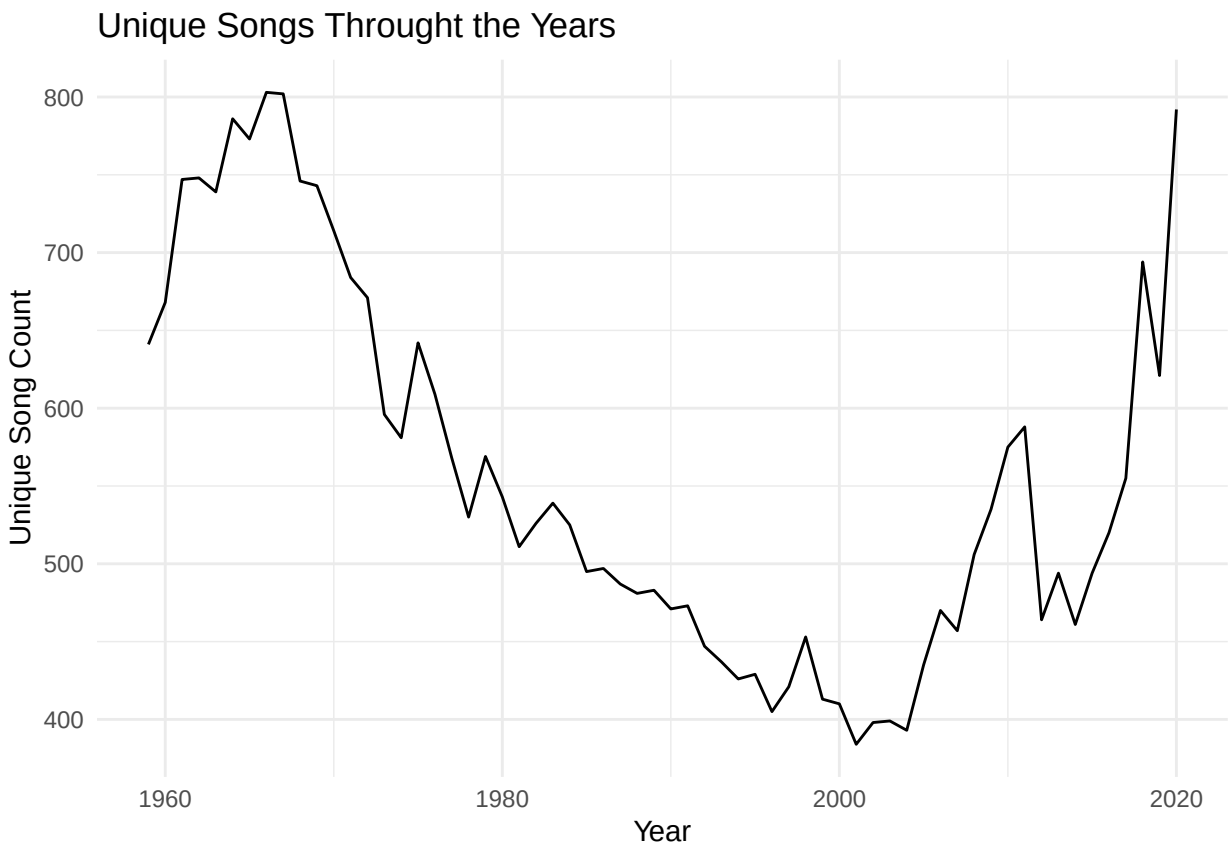
The many scatter plots below are faceted by hour of the day and demonstrate the relationship between the temperature and the amount of boardings that happened and every dot is colored deepening on whether it was a weekend or weekday. While observing the graphs it appears that temperature doesn't affect the amount of boarding as much as the time of day does. The boxy shape that the dots form indicate the constant boardings regardless of the temperature.

Problem 4: Wrangling the Billboard Top 100

Part A -

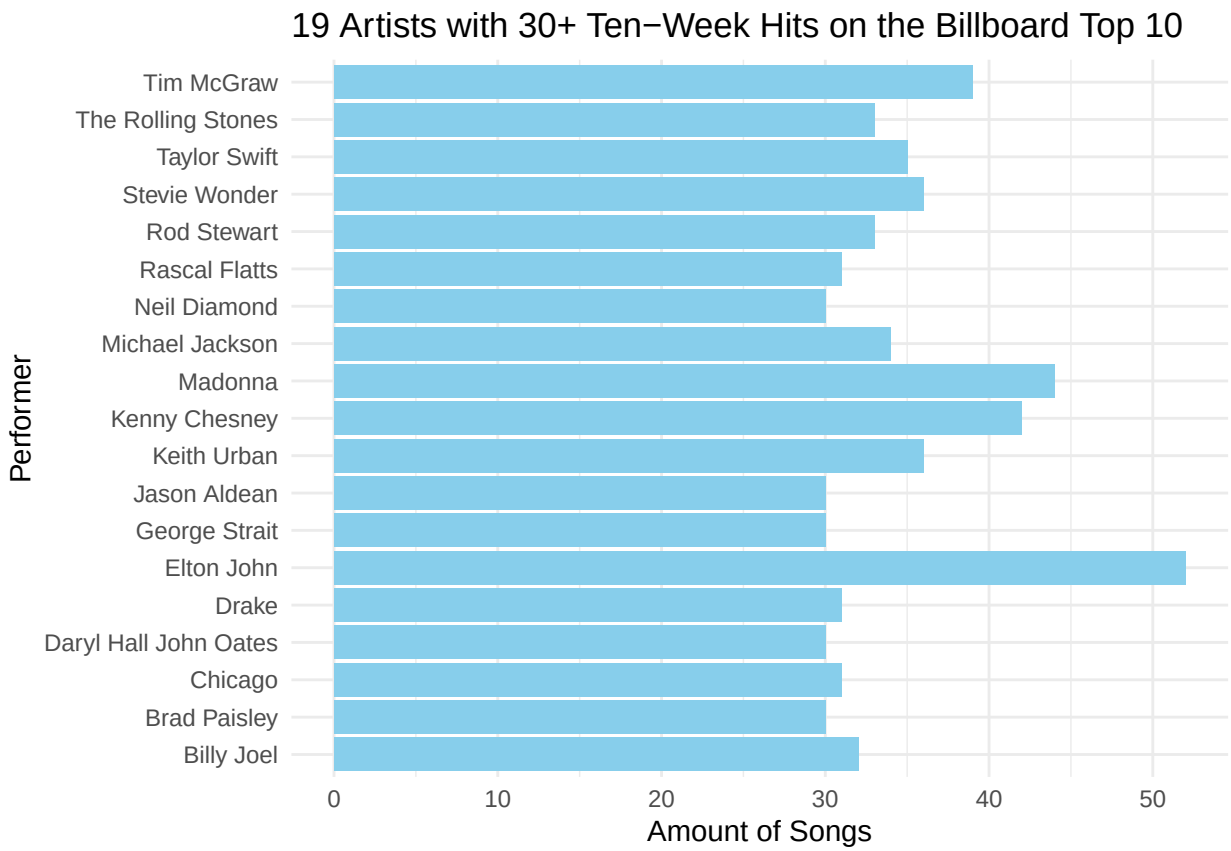
```
## # A tibble: 10 x 3
## # Groups:   performer [10]
##    performer                                song                  total_weeks
##    <chr>                                    <chr>                       <int>
##  1 Imagine Dragons                          Radioactive                  3828
##  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
```

The table above show the songs with the most weeks on the Top 100 Billboards.



Part B -

The line graph below demonstrates the count unique songs that appeared on the Billboard 100 per year.



Part C -

The bar plot shows the 19 artist who had at least 30 songs appear on the Top 100 Billboard for more than 10 weeks.