Hw_2

Pranita

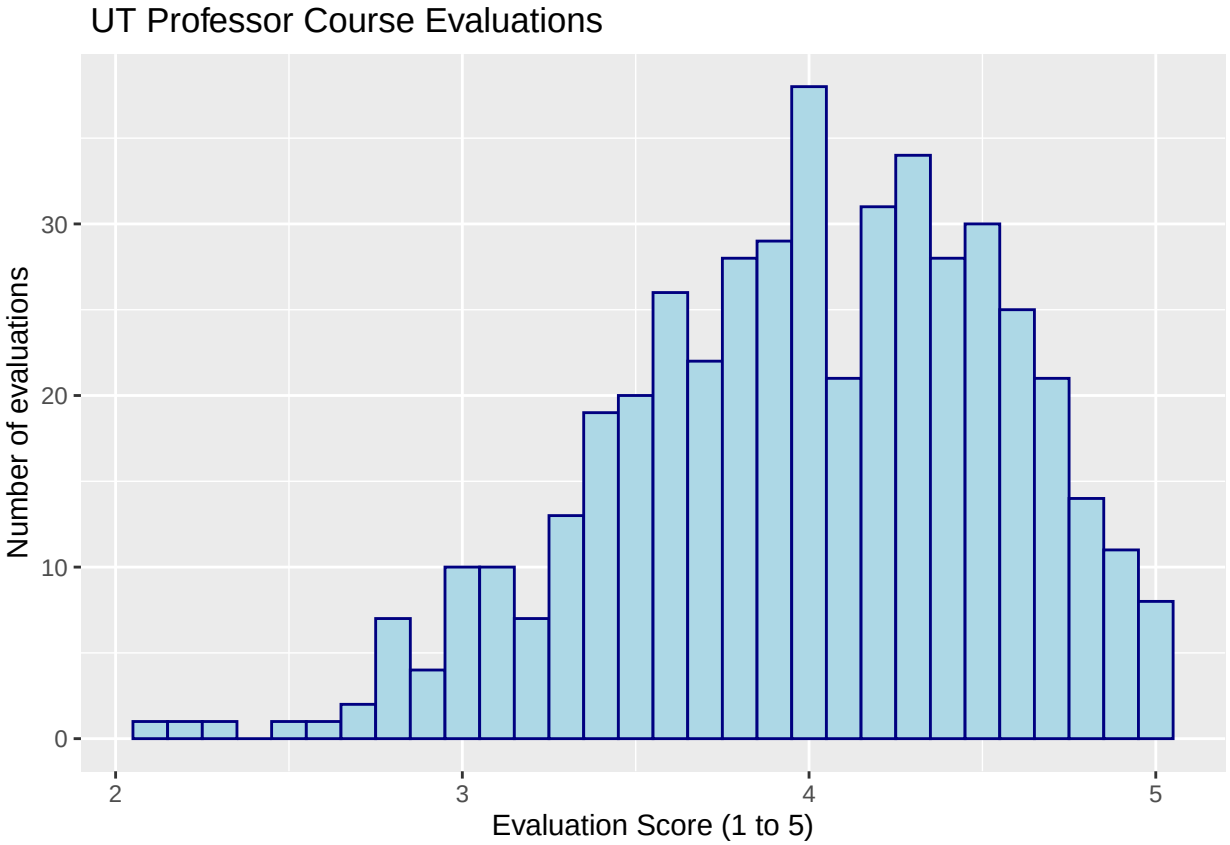
2025-01-27

Name: Pranita Chaudhury

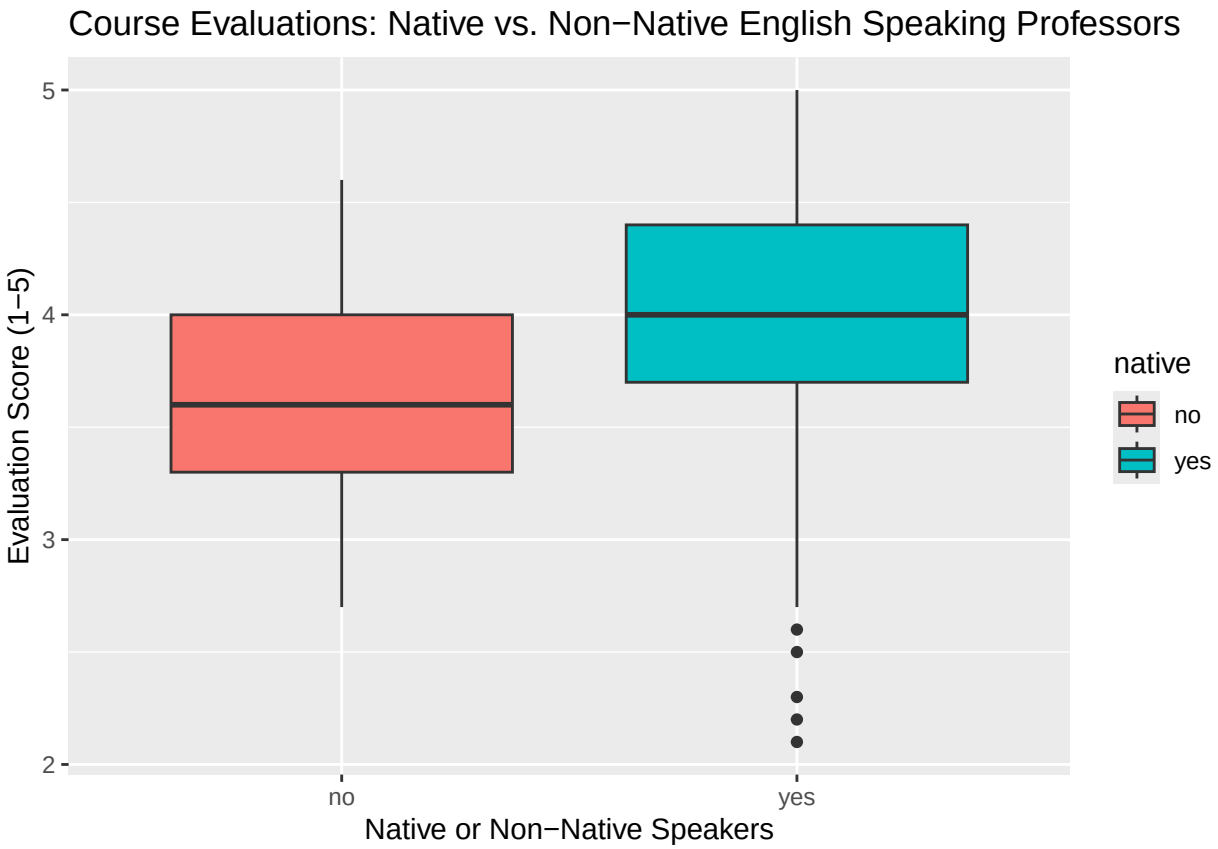
UT EID: pc28377

Click https://github.com/PranitaChau/Hw_2 for my Github

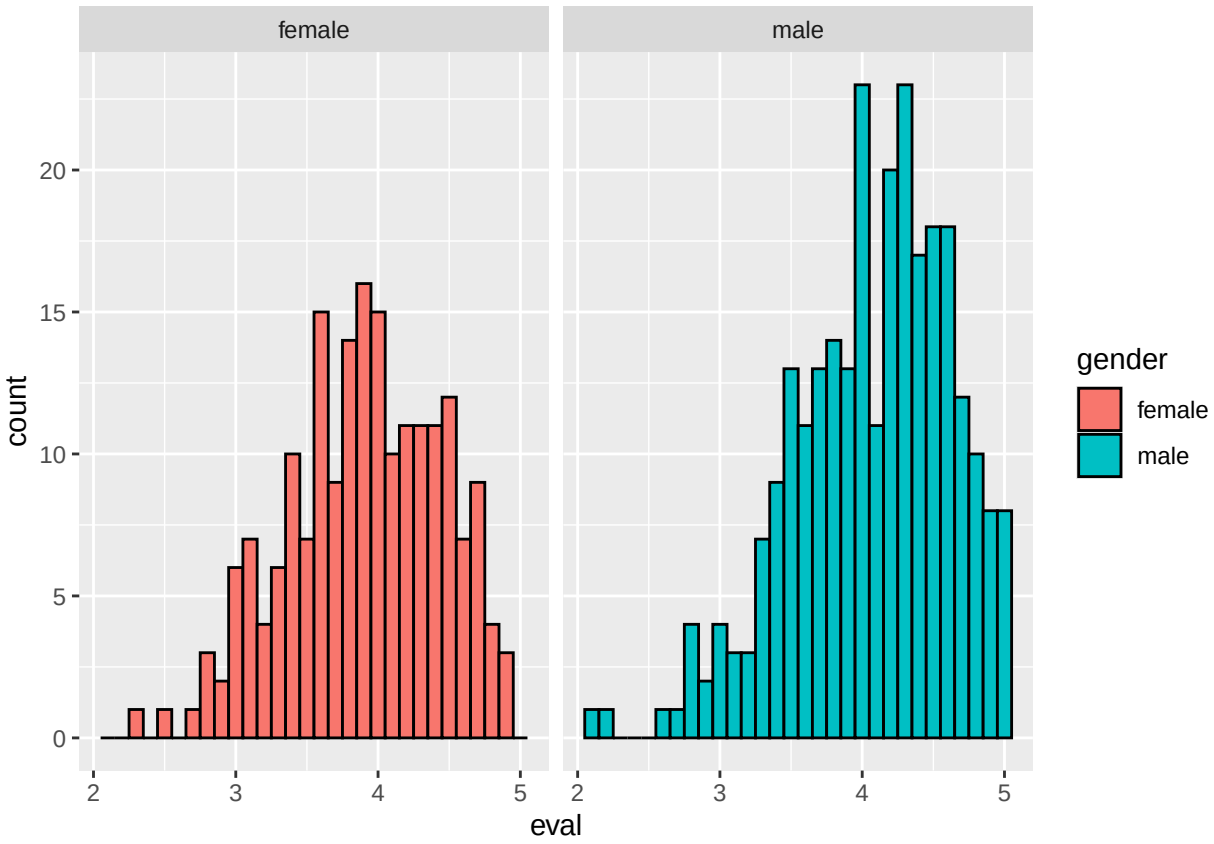
Problem 1



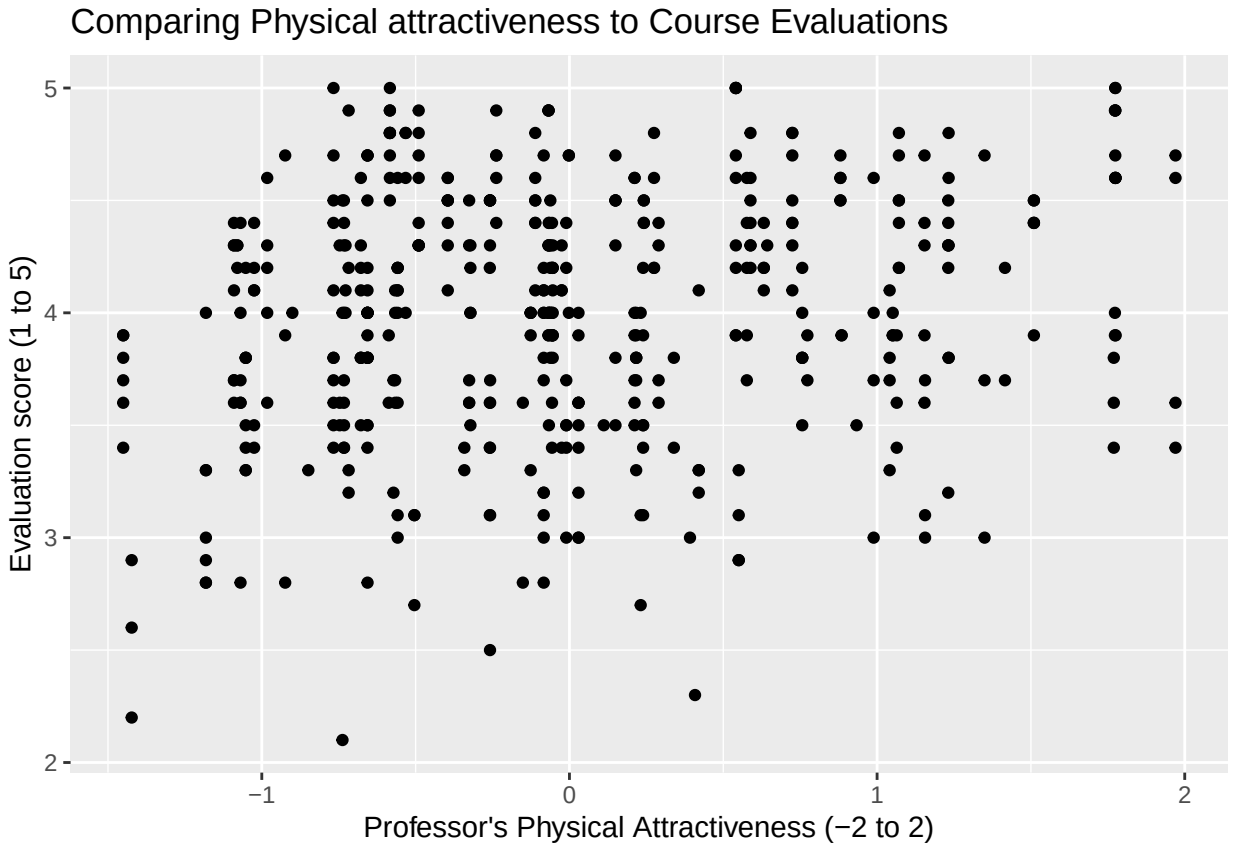
This graph shows the distribution of course evaluations of UT Austin professors from a scale of 1 to 5. Additionally, the majority of course evaluations for UT Austin professors are above 3.6 out of 5.0, with the median landing at 4.0. From this we can takeaway that most professors at UT Austin are above average based on their course ratings.



This boxplot shows the evaluation scores for native English speaking professors compared to non-native English speaking professors. From this it can be concluded that Professors who are native English speakers earn, on average, a higher course evaluation score than non-native English speakers. The median on non-native English speakers is lower than Q1 for native English speaking professors. This suggests that native English speaking professors are either better professors, or it is easier to understand them.

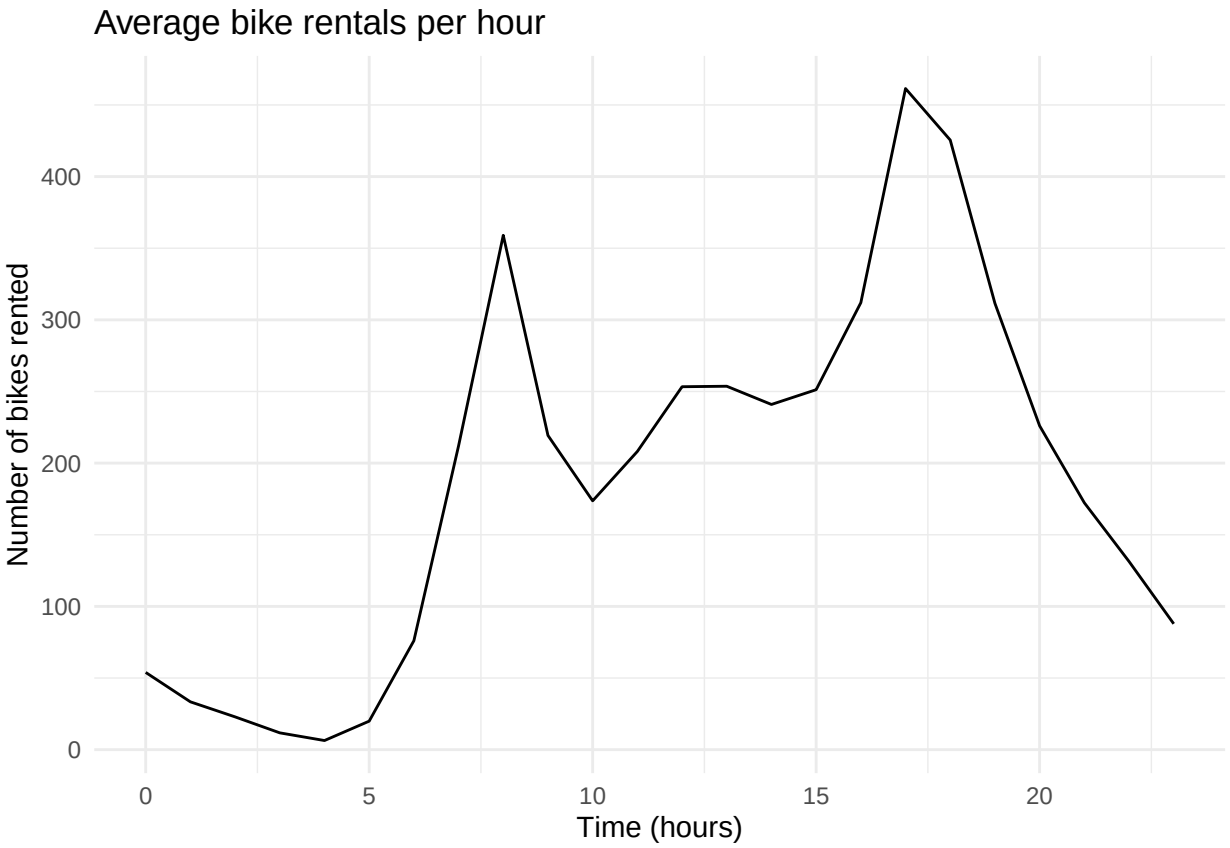


This graph compares UT Austin male professor's course evaluations to female professors. There is a peak for the female professors at the higher end of the 3.0 scores, while the peak for male professors is at the lower end of the 4.0 scores. From this we can take away that the male professors have higher evaluation scores on average when compared to female prefessors at UT Austin.

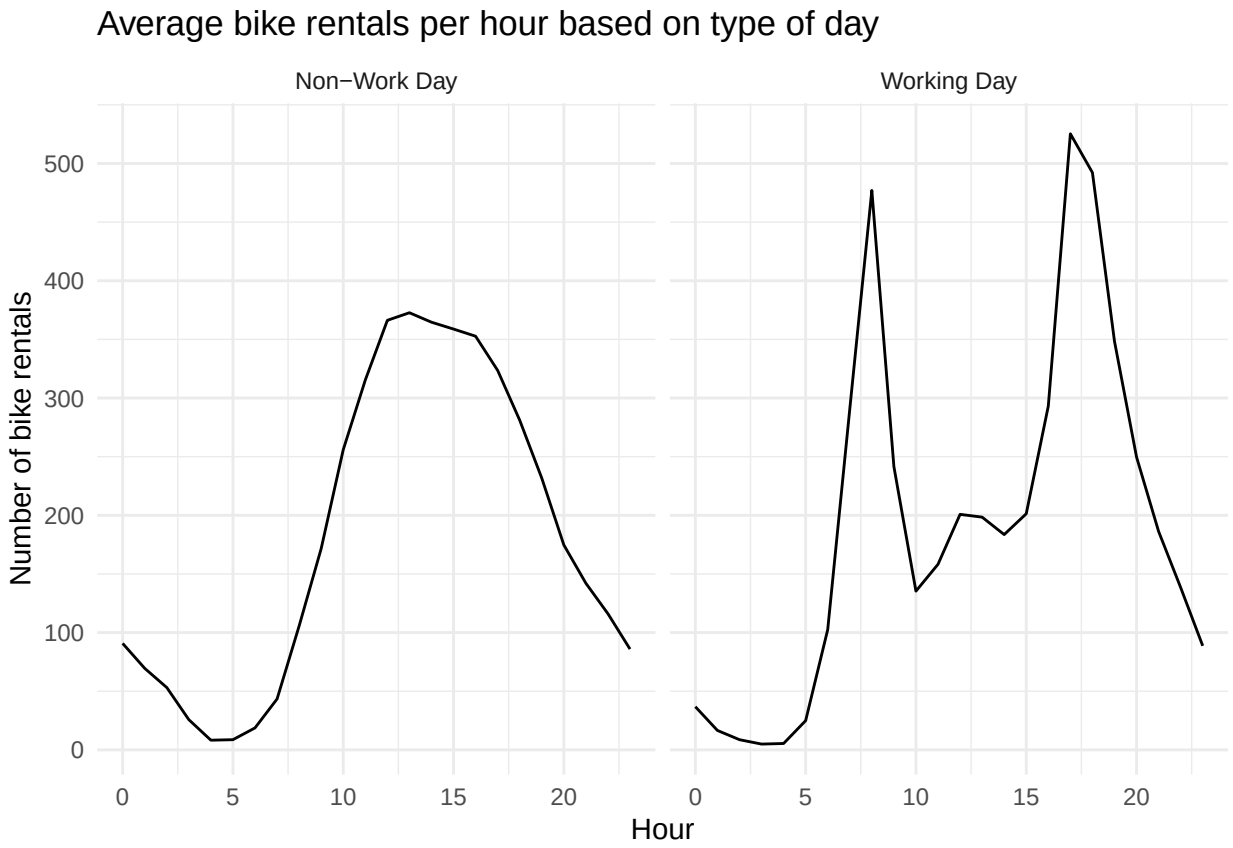


Based on the graph there is no obvious correlation between a professor's physical attractiveness and their course evaluation. This can be backed up by looking at the correlation coefficient, which is 0.1890391, indicating a very weak positive linear correlation between the two variables at hand. This means that the physical attractiveness of a professor has a weak positive impact on their course evaluations.

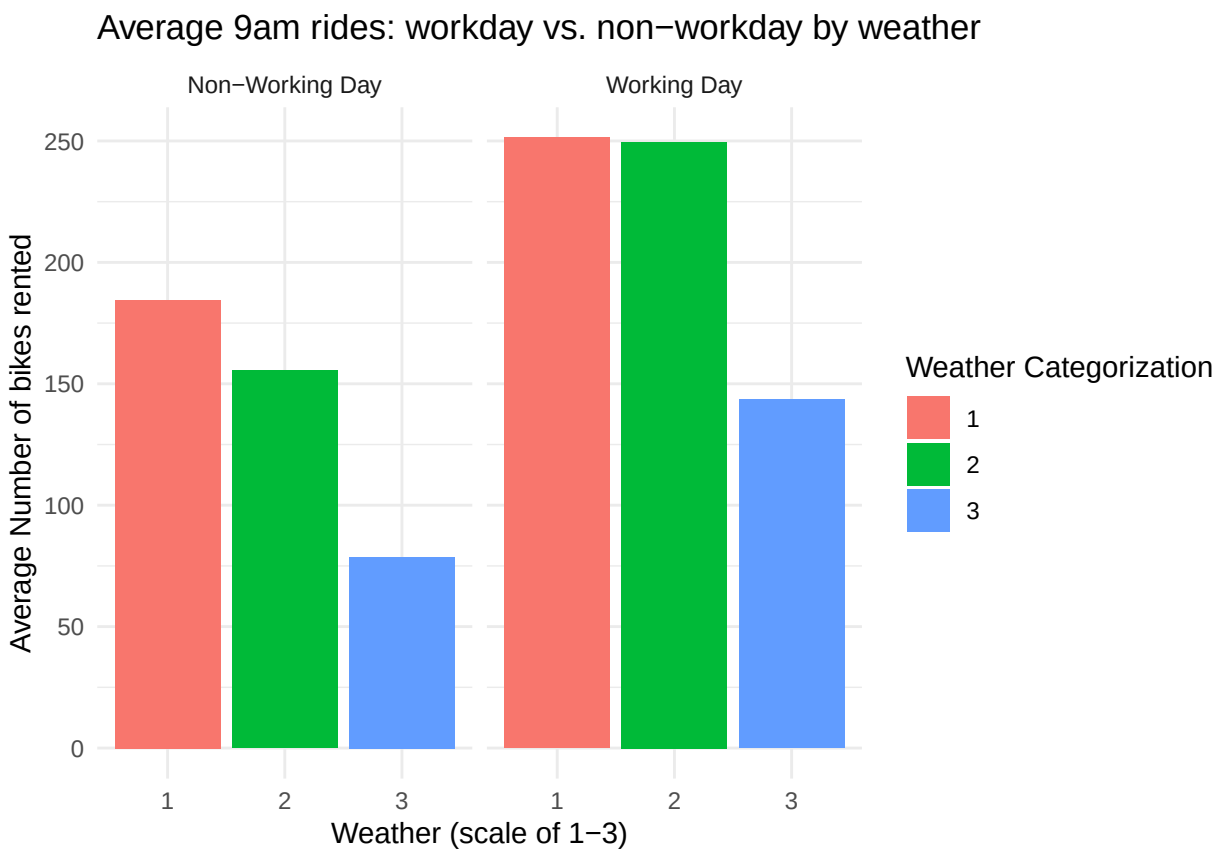
Problem 2



This graph represents the average number of public bike rentals per hour in any given day, with the x-axis representing the hours in the day and the y-axis representing the number of bike rentals (on average) at that time of the day. There is a spike at 8 (8am) and another larger one at 17 (5pm), indicating a sudden increase in bike rental users. This could be a result of people using it to get to work and back home afterwards, as it correlates with the usual work hours of 9 to 5.

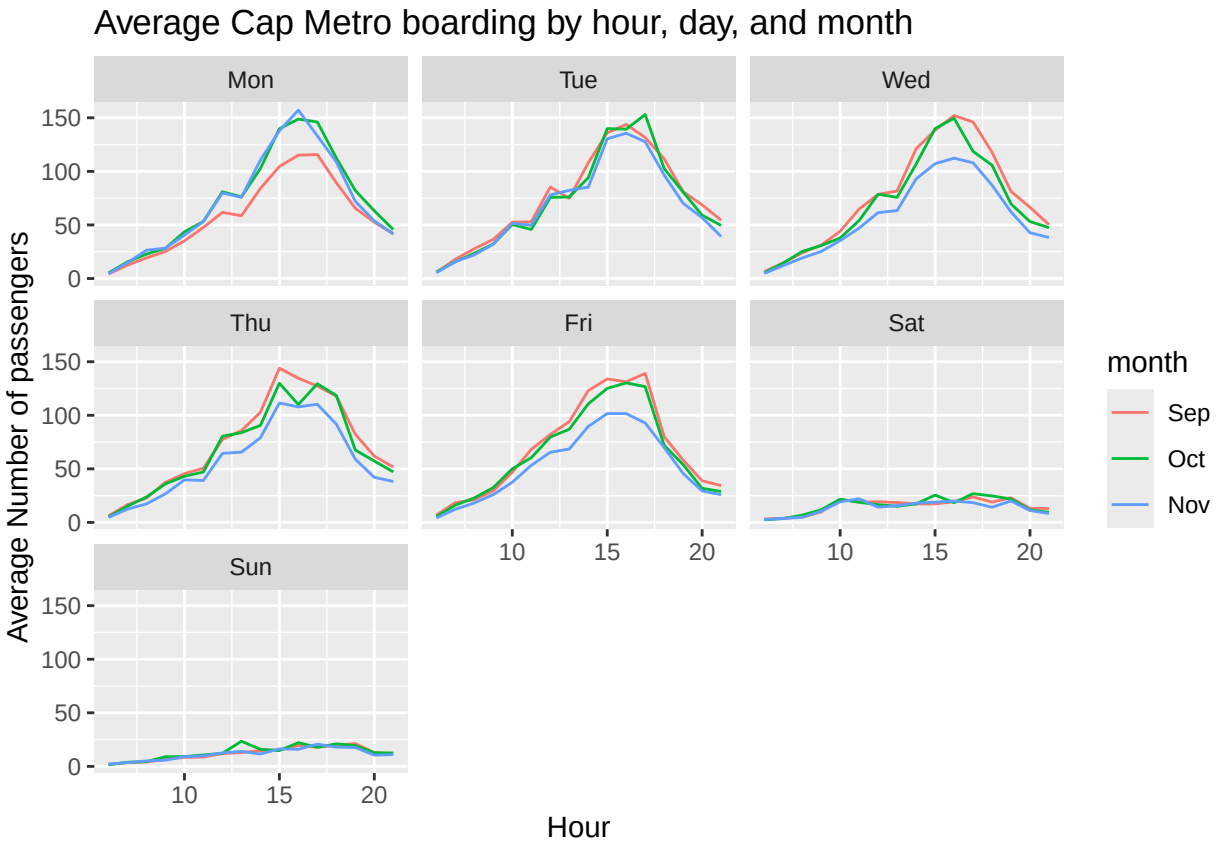


The first graph represents the average number of public bike rentals per hour in a non-work day, while the second graphs refers to average bike rentals per hour on a weekday, with the x-asex representing the hours in the day and the y-axes representing the number of bike rentals (on average) at that time of the day. The first graph shows a curve from around noon to 5pm, while the second graph has more rapid sharp spikes at 9am and 5pm. The bike rentals from first graph are due to people traveling for recreational and errand purposes, which is why are they more spread out and not all at the same time, but the working day graph is due to people traveling to and from work which causes the spikes as everyone needs them at a similar time.

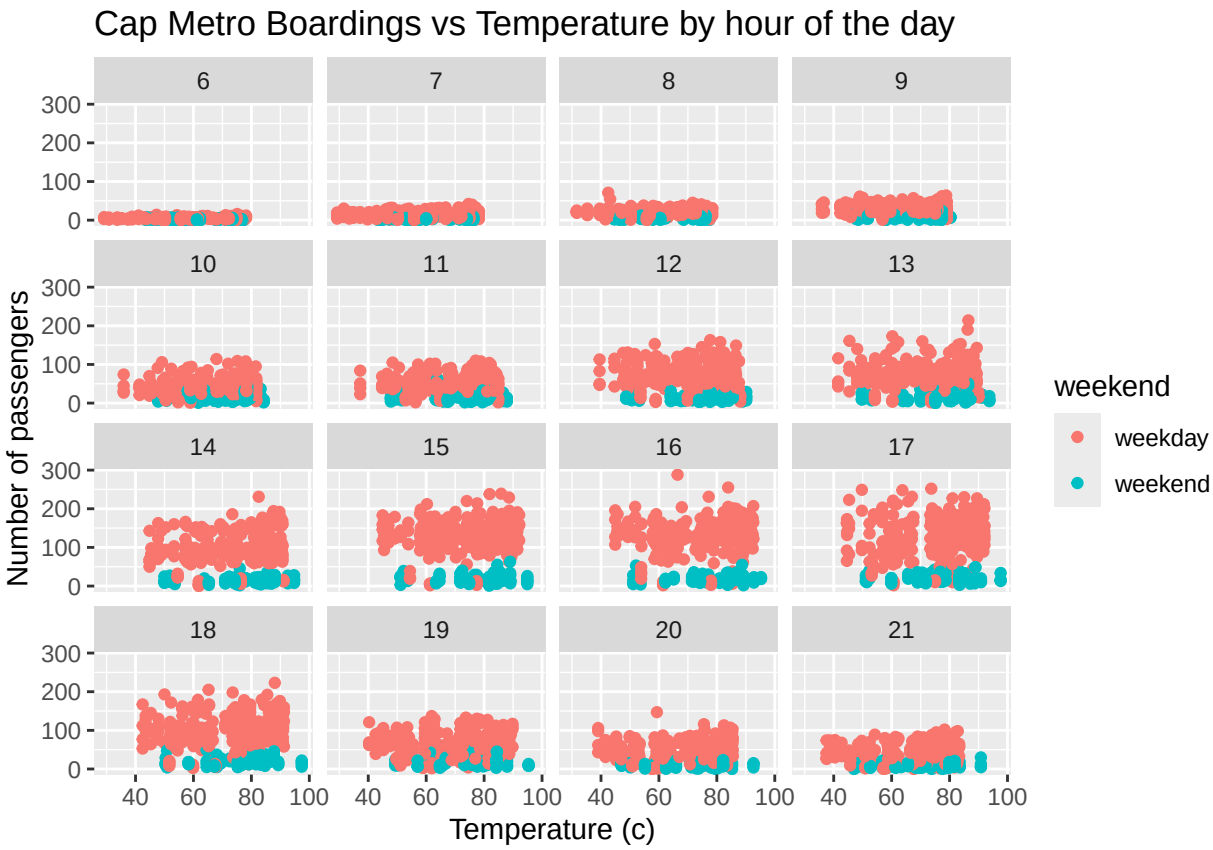


This figure shows the number of average rides taken by public bike rental users at 9am segregated by weather as well as if it is a working day or not. The x-axes represents the weather, and the higher the weather categorization the harsher the weather is. The y-axes represents the average number of bike rentals at 9am during that weather condition, and the left graph shows this data on non-working days while the right graph shows the data on work days. Through this we can conclude that the weather does make an impact on the rides taken at 9am on a non-working day, but it does not impact the number of rides taken at that same time on a working day unless the weather gets really bad (categorized as a 3 or higher). This is due to the fact that people have to use the bikes to get to work at 9am, which a mist will not interfere with however heavy rain or a thunderstorm would cause the amount of people going to work to decrease on that given day.

Problem 3



This graph shows the average number of passengers boarding Capital Metro's bus network during the months of September, October, and November, segregated by day of the week, which show the average number of passengers boarding by each hour of the day. From these it can be inferred that the weekends have less than 25 people board per hour while the weekdays see a lot more traffic. The weekdays see a similar traffic pattern, with a peak between 3-5pm around the end of work hours in time for people to get back home. Additionally, on average the month of November saw the least traffic in passengers other than Mondays, where the month of September saw the least numbers of passengers. I think that this is due to the fact that people are just coming back from summer, and are still getting used to that start of the week (Mondays) and have not yet set into their routines so they are not using the busses daily People may travel less on Wednesday, Thursday, and Fridays due to a combination of cold weather and the holiday season approaching so they are trying to save up for that. In conclusion, Cap Metro can think about decreasing their routes on the weekends and between midnight and 5am since that is where they see the least amount of passengers taking rides, while they should keep all routes running during the day on weekdays as that is when the maximum amount of people board their busses.

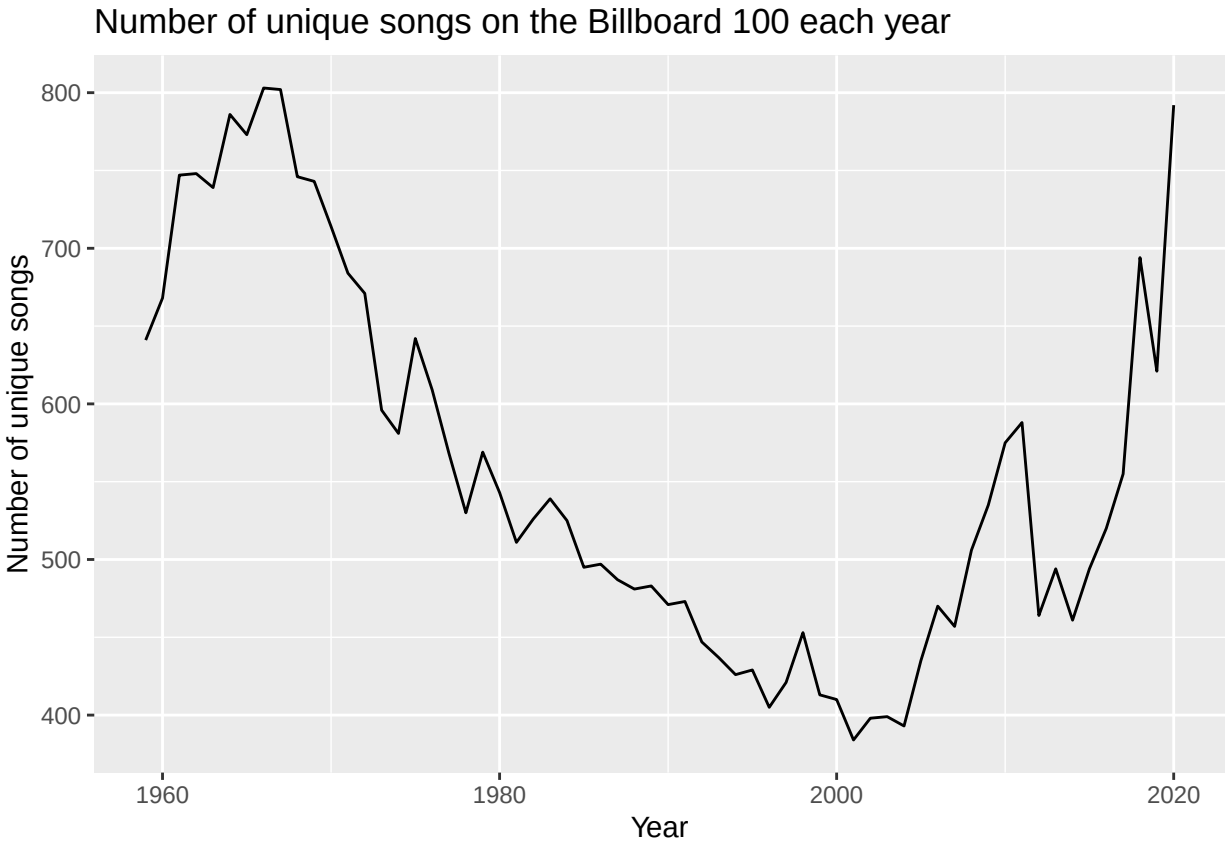


The figure shows a comparison of people boarding Capital Metro busses to the temperature in the day, and this graph is segregated by the hours of the day as well as if it is a working or a non-working day. When holding the hour of the day and weekend status constant, higher temperatures seem to increase the amount of students traveling on the bus. This can be seen as there are more blue dots on the right half of the graph (indicating increased temperatures) compared to the left which has lower temperatures.

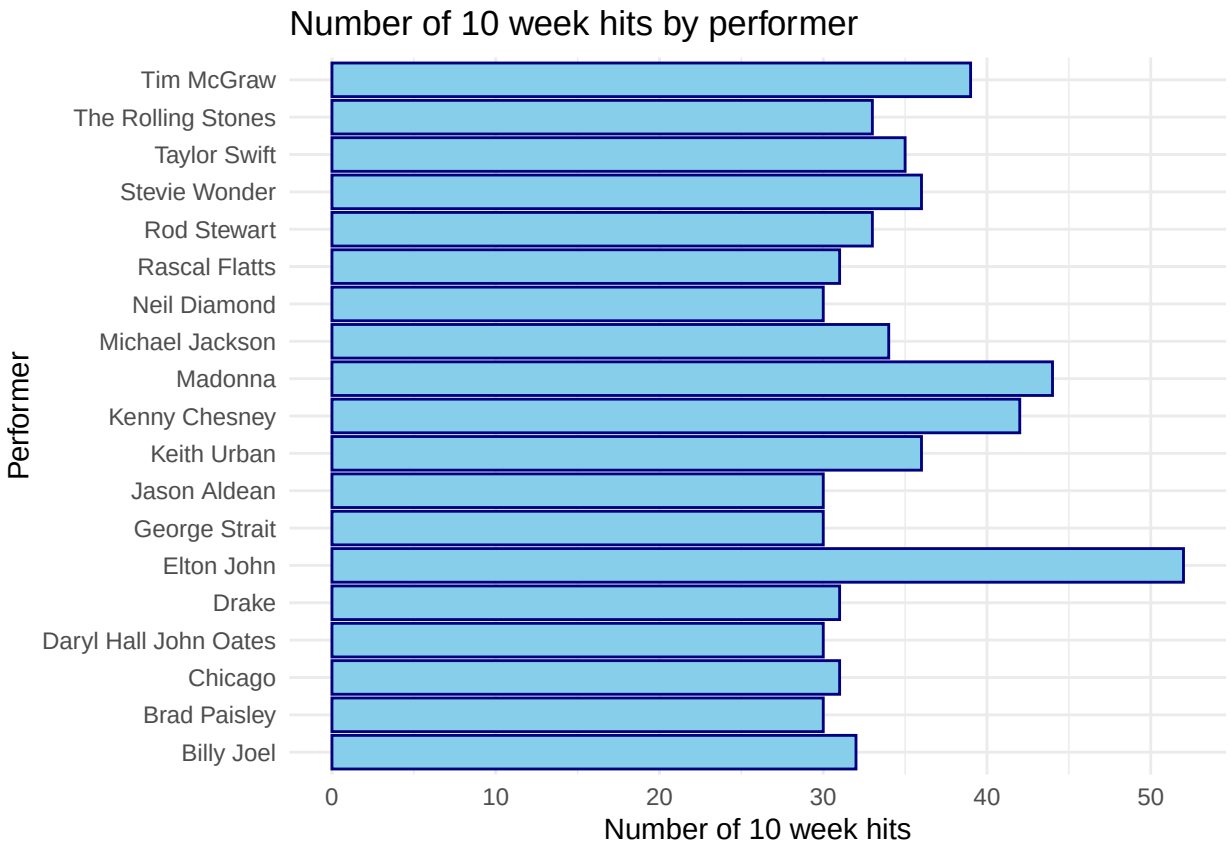
Problem 4

```
## # 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 LeAnn Rimes                              How Do I Live                    69
##  5 LMFAO Featuring Lauren Bennett & GoonRock Party Rock Anthem               68
##  6 OneRepublic                              Counting Stars                   68
##  7 Adele                                    Rolling In The Deep              65
##  8 Jewel                                    Foolish Games/You Were Meant~    65
##  9 Carrie Underwood                         Before He Cheats                 64
## 10 Lifehouse                                You And Me                       62
```

This table shows the top 10 more popular songs from the Billboard 100 from 1958 to 2020 (inclusive), as well as the performer and the count, which is the number of weeks spent on the Billboard 100. The song that spent the most time on the Billboard 100 was Radioactive by Imagine Dragons which spent 87 weeks on the Billboard.



This figure shows the number of unique songs which were on the Billboard 100 each year from 1958 to 2021 (non-inclusive). The year with the highest unique songs on the Billboard was 1966, and the numbers plumeted from then. I thought this was an interesting trend, and could be due to the main stream media and music, which helped make certain songs more popular and sharebale, as we can see with an all time low in 2001. From then it rapidly climbed up again, which I think could be credited to the increase in social media. Therefore, I think that based on the Billboard 100 the "musical diversity" decreased from 1958 to 2001, but after the 2000's the diversity started to increase again.



A 10 week hit is when the song stays on the Billboard for 10 or more weeks. The figure shows the 19 performers that have had over thirty 10 week hits on the Billboard 100, as well as the number of 10 week hits that they have had. Most of these performers had 30-40 such hits, and the performer with the most 10 week hits was Elton John, with 52 such hits on the Billboard.