

SDS 315 Homework2

Daniel Wu (EID: djw3627)

2025-01-22

To access my GitHub repository, click here: <https://github.com/DanielWu3627/SDS315/tree/main>

Problem 1

Part A

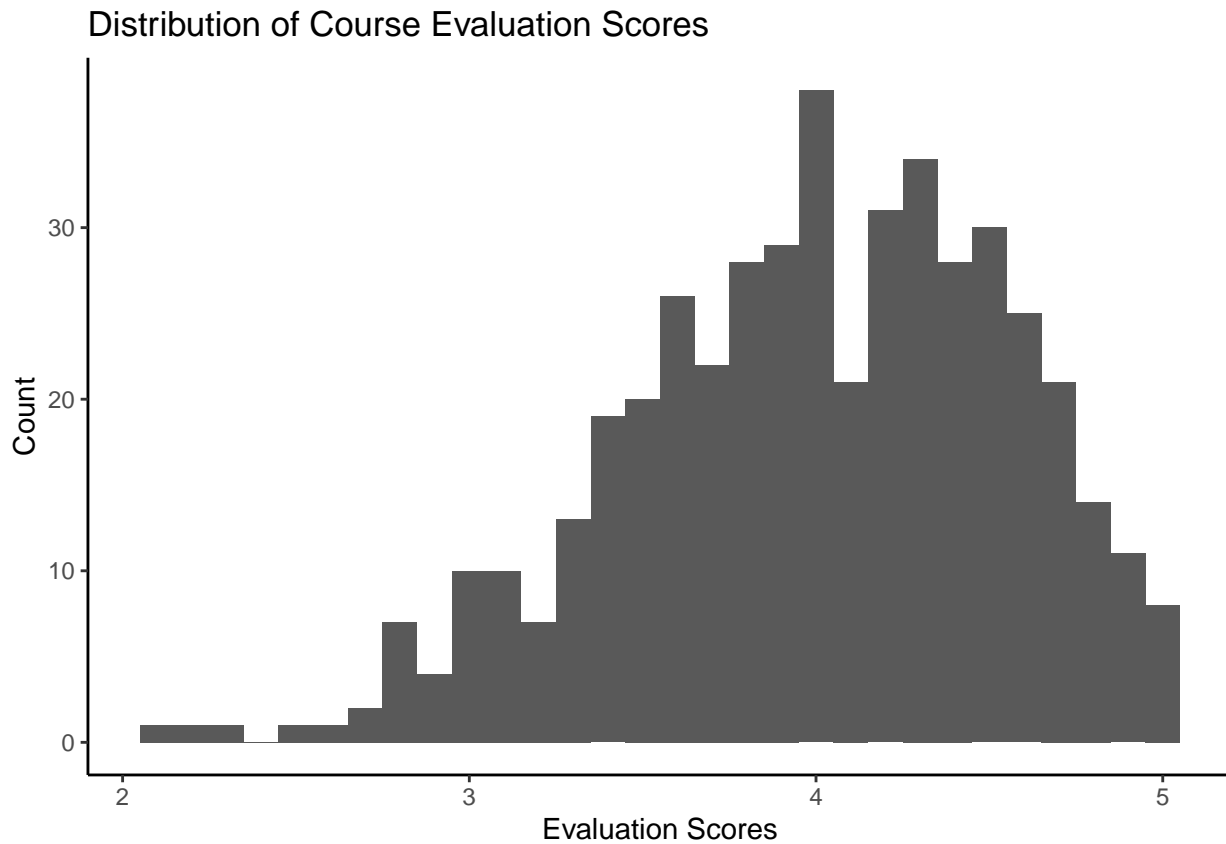


Figure 1a. This histogram shows the distribution of course evaluation scores received by UT professors. The distribution is centered around 4 on a scale of 1 to 5 points, suggesting that many students were quite satisfied with their professors. However, the distribution is left skewed due a few low scores between 2 and 3.

Part B

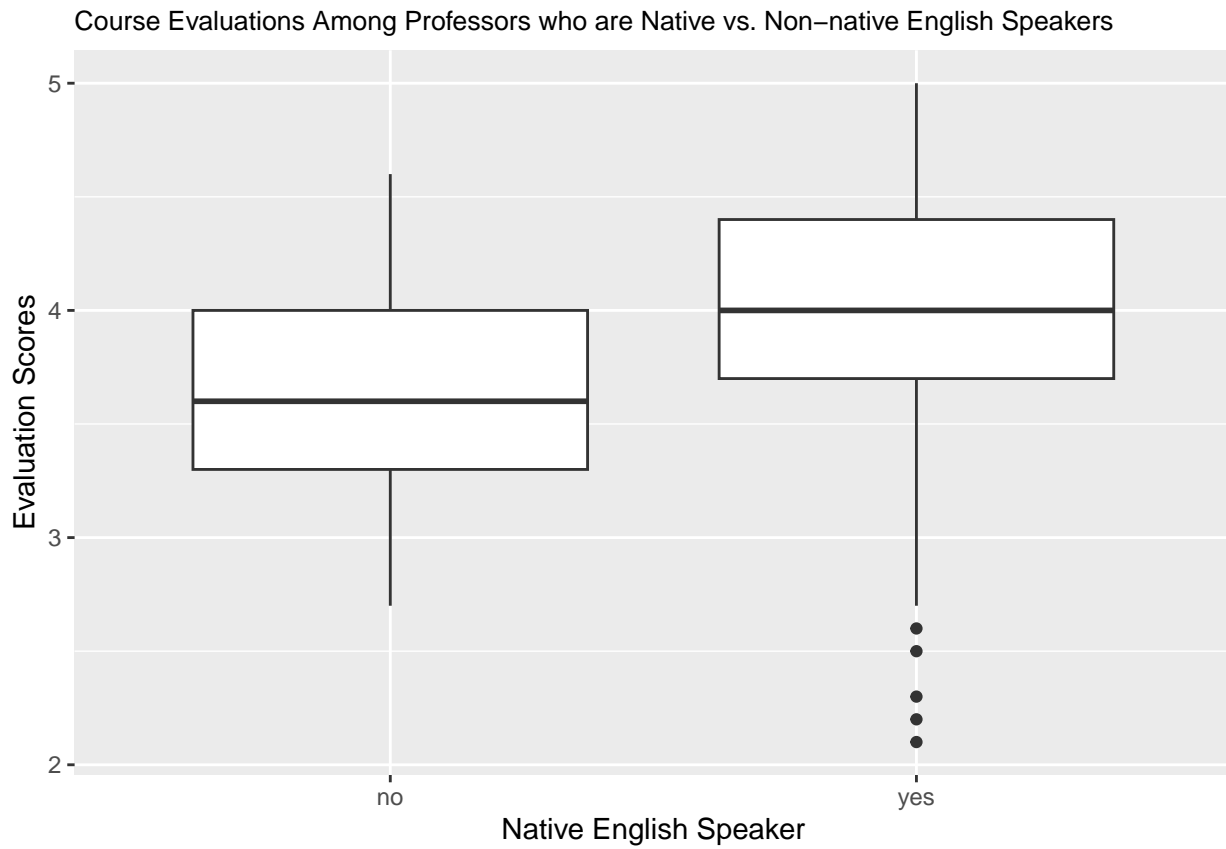


Figure 1b. Distribution of course evaluation scores among professors who are native (yes) or non-native (no) English speakers. According to the box plots, the median rating of professors who are non-native English speakers is lower than that of professors who are native English speakers. However, there are outliers in which students gave low ratings to some professors who are native speakers, making the distribution of native English speaker professors having a wider range of evaluation scores compared to non-native speakers.

Part C

Distributions of Course Evaluation Scores Among Professors of Different Genders

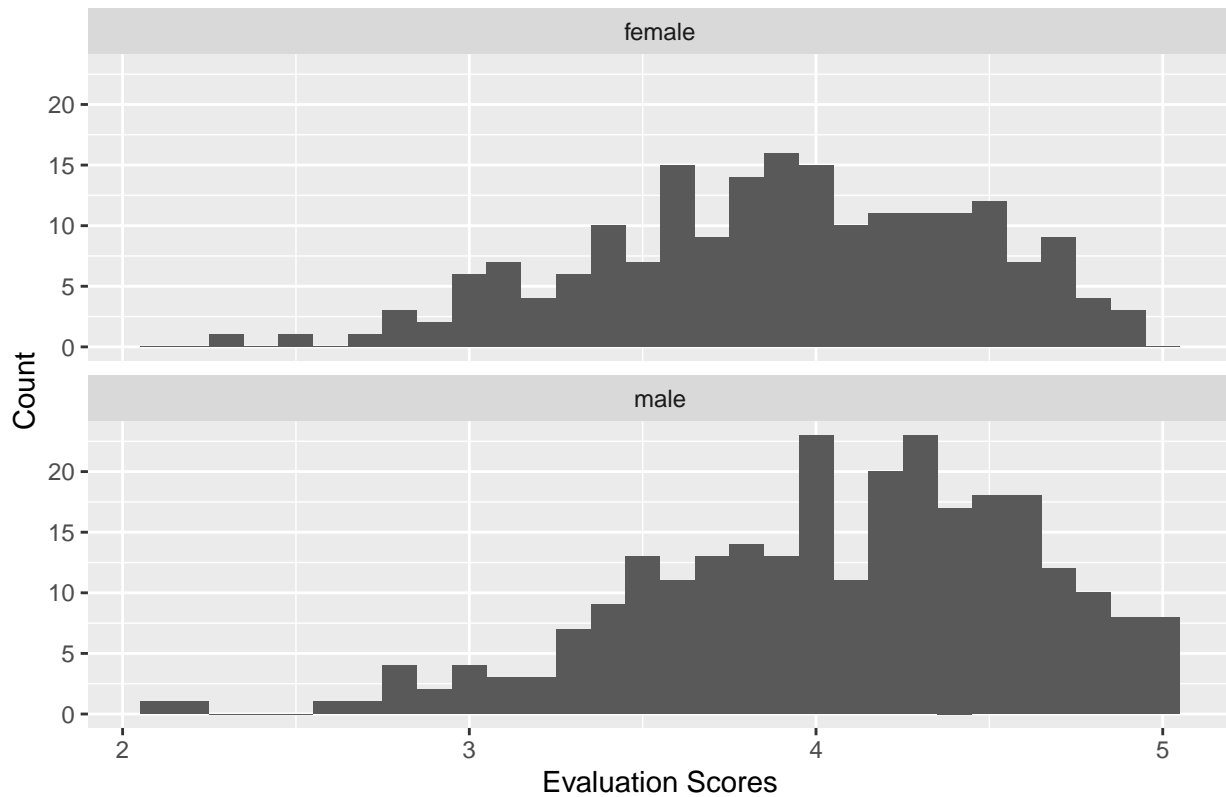


Figure 1c. Distribution of course evaluation scores among female vs. male professors. According to the two histograms, the distribution of course evaluations among female and male professors look very similar. Both distributions are left skewed by a few lower scores between 2 and 3, and centered around 4. However, the median evaluation scores received by female professors are a little lower than 4, while the median scores received by male professors are a little above 4.

Part D

Relationship Between Physical Attractiveness and Course Evaluations

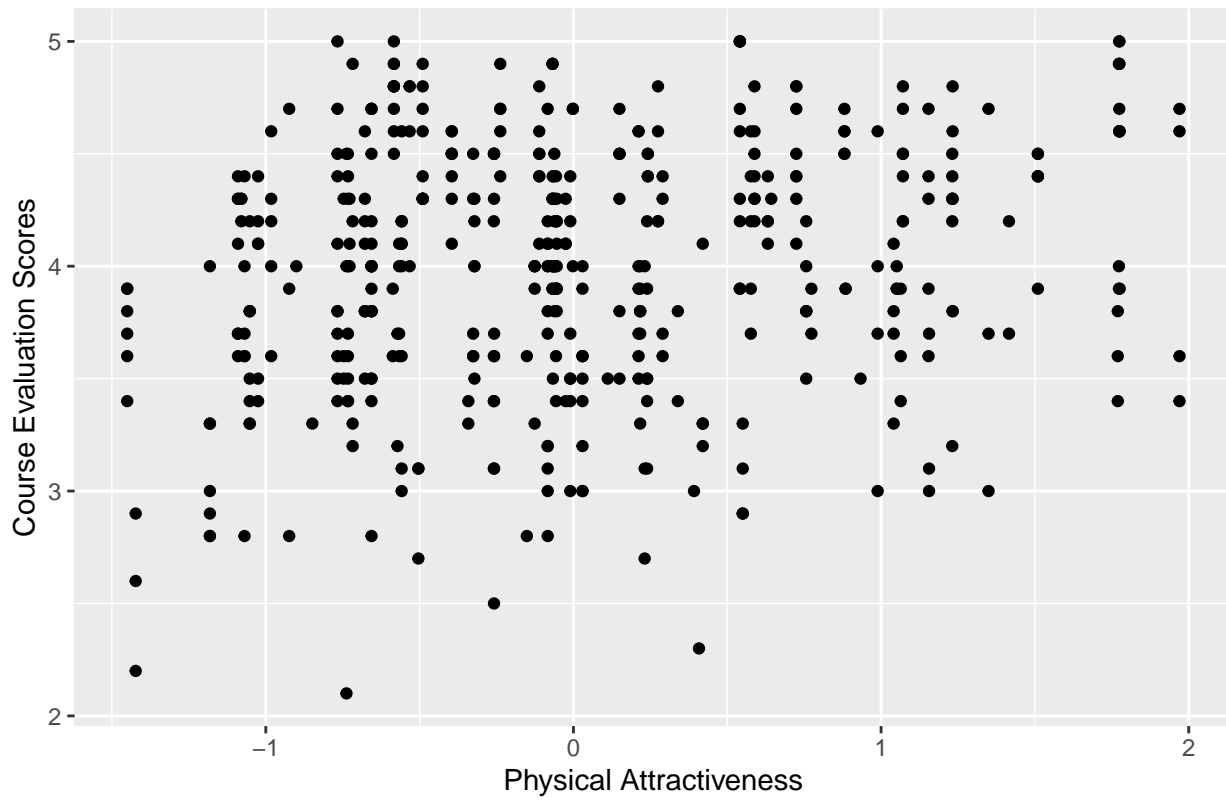


Figure 1d. Relationship between physical attractiveness and the evaluation score of professors. The correlation coefficient is about 0.19. This suggests that there is weak but positive relationship between physical attractiveness and course evaluation scores of professors.

Problem 2

Plot A

Average Number of Bike Rentals Throughout Different Hours of the Day

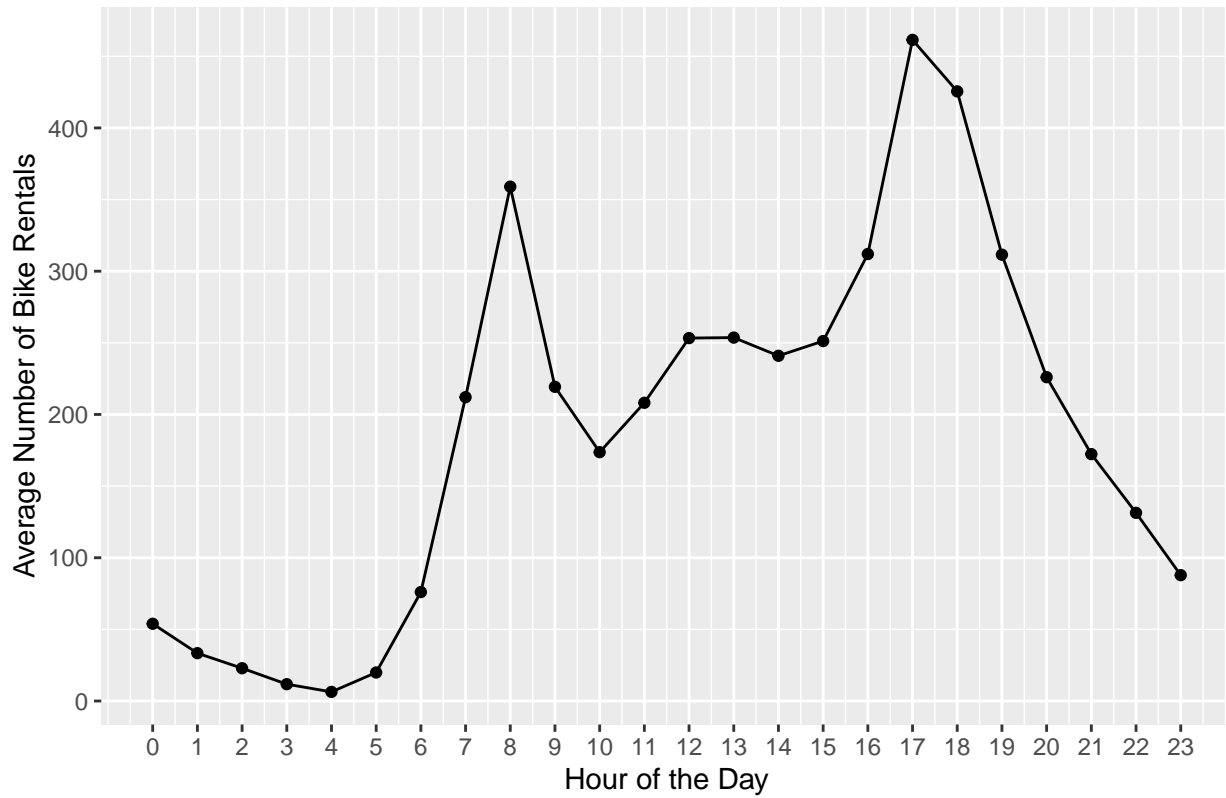


Figure 2a. The average number of bike rentals across different hours of the day. This line graph shows that in the early hours of the day, there are not a lot of bike rentals. However, there is an increase in bike rentals since 6 am, and it peaks at around 8 am, and starts to decrease again. At around 15 hours (3 pm), bike rentals increase again, reaching a peak at around 5 pm. After that time, bike rentals decrease again. This trend shows that bike rentals have a high demand during the morning and evening rush hours.

Plot B

Average Number of Bike Rentals Throughout Different Hours of the Day

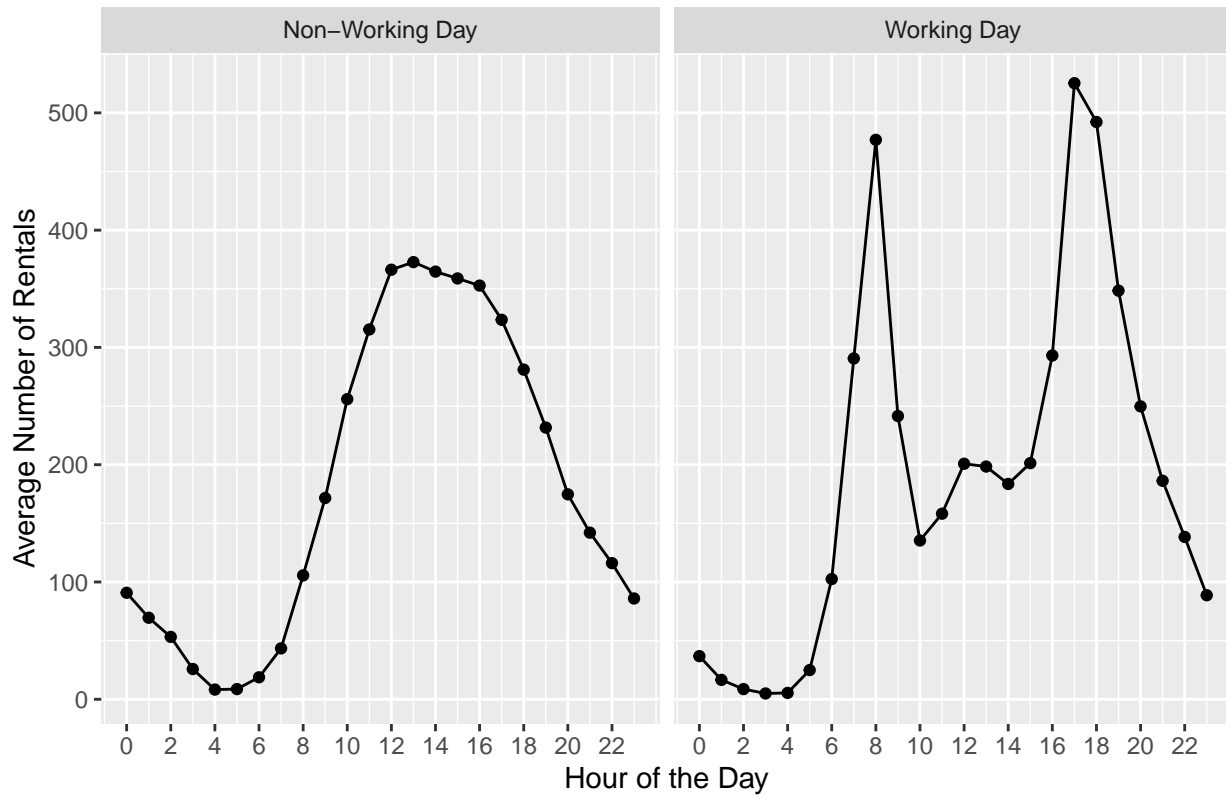


Figure 2b. The average number of bike rentals during the different hours of the day, according to whether it is a working day. It shows that during both working days and non-working days, the average number of bike rentals are low in the early morning and late night hours. However, during working days, there are sharp peaks in average bike rentals around 8 am and 5 pm, which are morning and evening rush hours. In contrast, during non-working days, average bike rentals have a wide plateau during the middle of the day with higher demand from noon to late afternoon.

Plot C

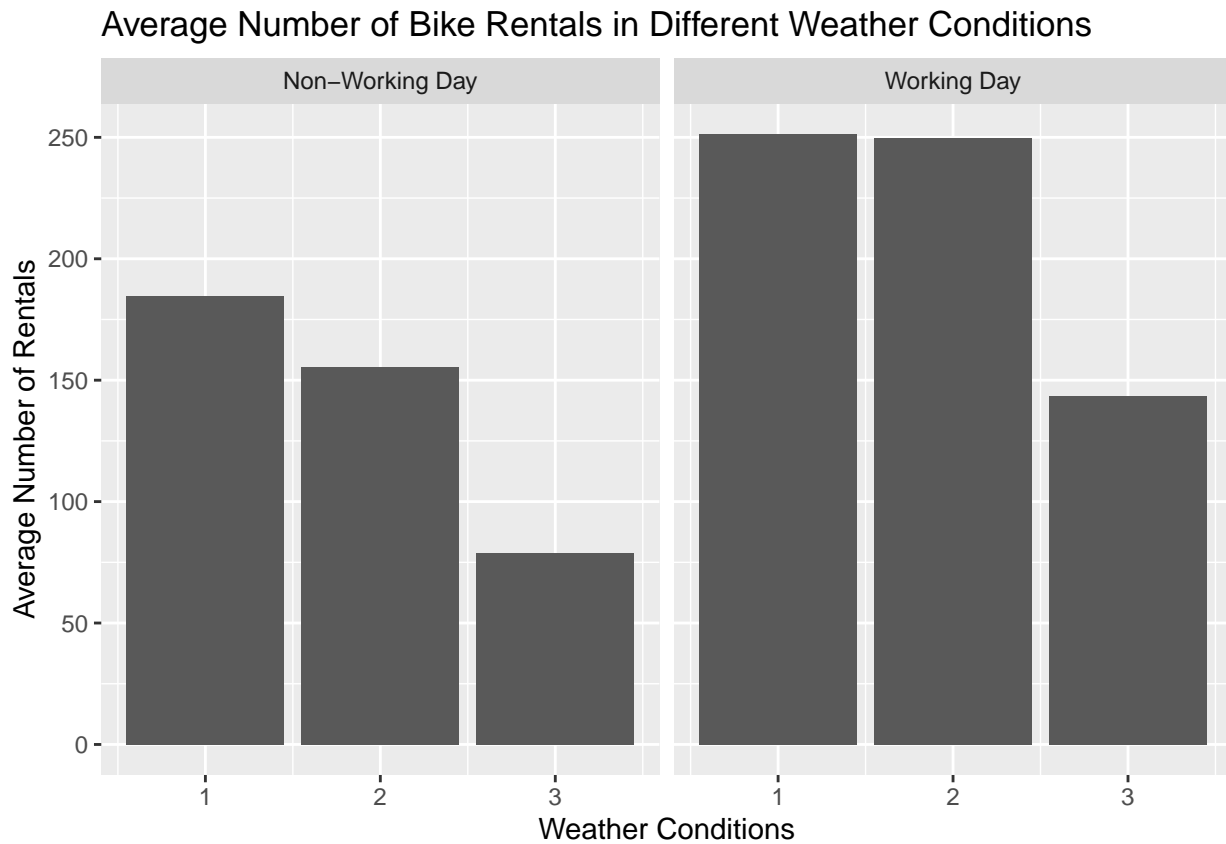


Figure 2c. The average number of bike rentals during the 9 am hour in different weather conditions (1-clear, partly cloudy, 2- mist/cloudy, 3 - lights snow/rain/thunderstorms) according to whether it is a working day. According to the bar plots, weather seems to have an effect on bike rentals. When the weather is clear, average bike rentals are higher than that in rainy/snowy days. regardless of whether it is a working day or not. However, the overall number of bike rentals are higher during working days than that in non-working days, suggesting that some people primarily rely on bikeshare services to commute to work.

Problem 3

Part 1

Average Number of Boardings at Different Hours of the Day in Different Mo

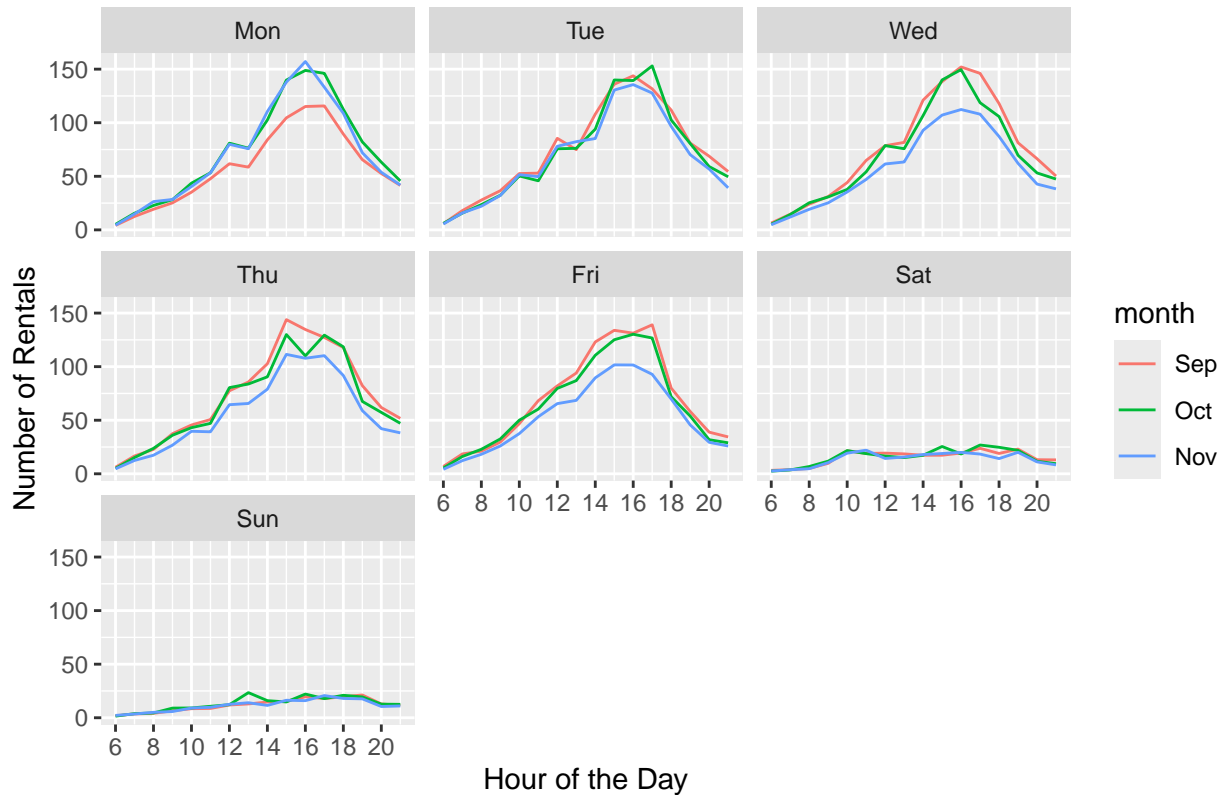


Figure 3a. Average bus boardings in September, October, and November during the different hours of the day, faceted by day of the week. The hour of peak boardings is usually between 3 and 5 pm (15 and 17) during weekdays, suggesting that routine travel patterns (eg. school/work-related travel) peak at those hours. During weekends, there is no obvious peak boarding hours. The average boardings on Mondays in September look lower, compared to other days and months, this can be caused by holidays (eg. Labor Day), which potentially reduces Monday traveling. Similarly, the average boardings on Wednesdays, Thursdays, and Fridays in November look lower, possibly due to Thanksgiving, which can lead to reduced local commuting.

Part 2

Bus Boardings vs. Temperatures During Different Hours of the Day

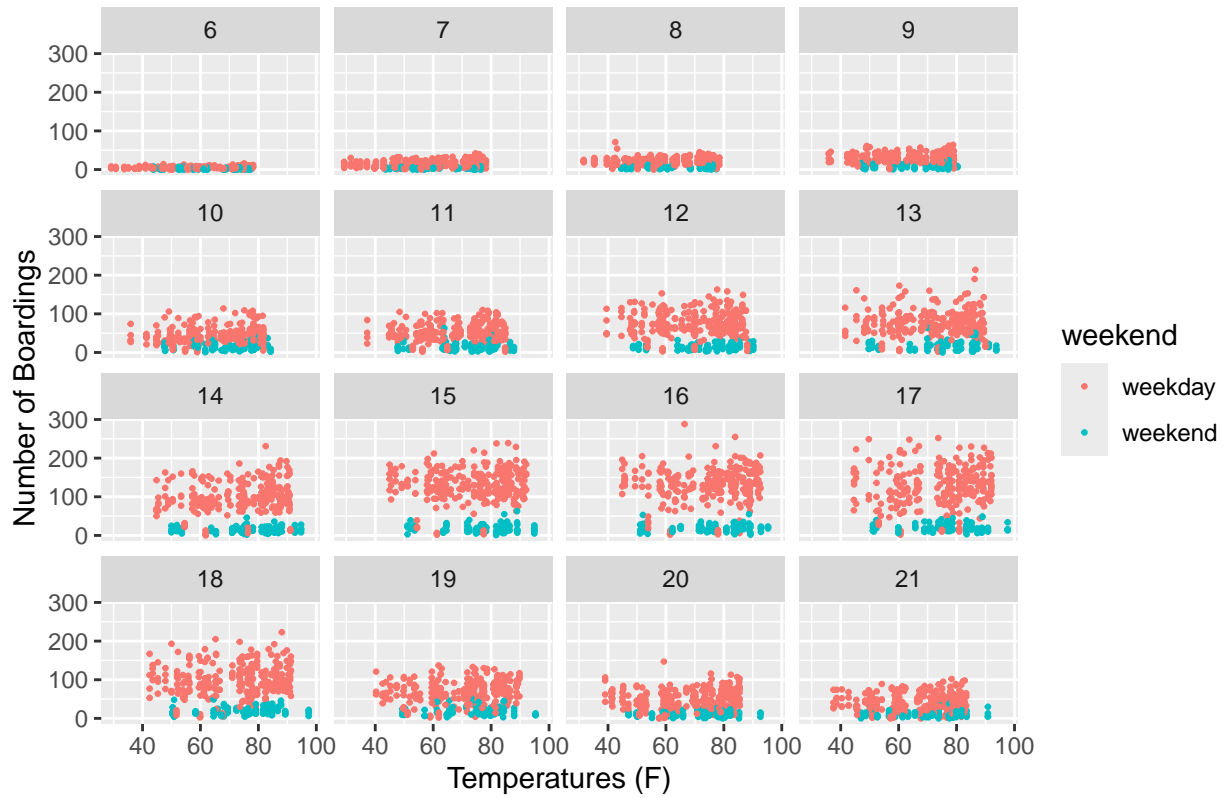


Figure 3b. The scatter plot shows the relationship between the number of boardings and temperature across different hours of the day. Each facet shows a specific hour and the data points are colored based on whether it is a weekday (red) or weekend (blue). When weekend status and hour of day are both held constant, there are few people boarding the bus during extreme temperatures. However, for typical temperatures, they do not have effects on the number of UT students riding the bus. Boardings remained relatively stable regardless of the temperature.

Problem 4

Part A

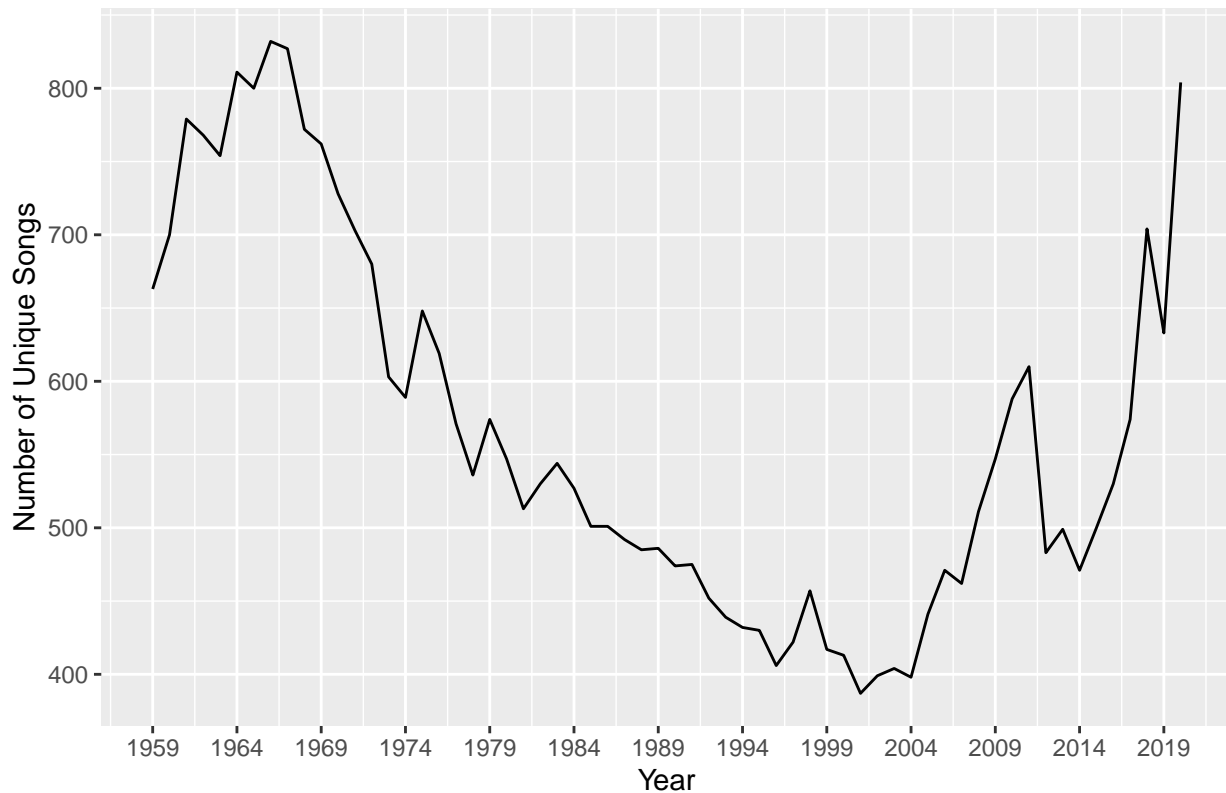
Table 1: Top 10 Songs

performer	song	count
Imagine Dragons	Radioactive	87
AWOLNATION	Sail	79
Jason Mraz	I'm Yours	76
The Weeknd	Blinding Lights	76
LeAnn Rimes	How Do I Live	69
LMFAO Featuring Lauren Bennett & GoonRock	Party Rock Anthem	68
OneRepublic	Counting Stars	68
Adele	Rolling In The Deep	65
Jewel	Foolish Games/You Were Meant For Me	65
Carrie Underwood	Before He Cheats	64

This table shows the top 10 most popular songs since 1958 on Billboard Top 100. The column “count” represents the total number of weeks a song appeared on the Billboard Top 100. The entries are sorted in descending order according to the number of weeks a song appears, with the most popular song appearing at the top. Note. The most popular song in 2021 onwards does not have up-to-the-minute data.

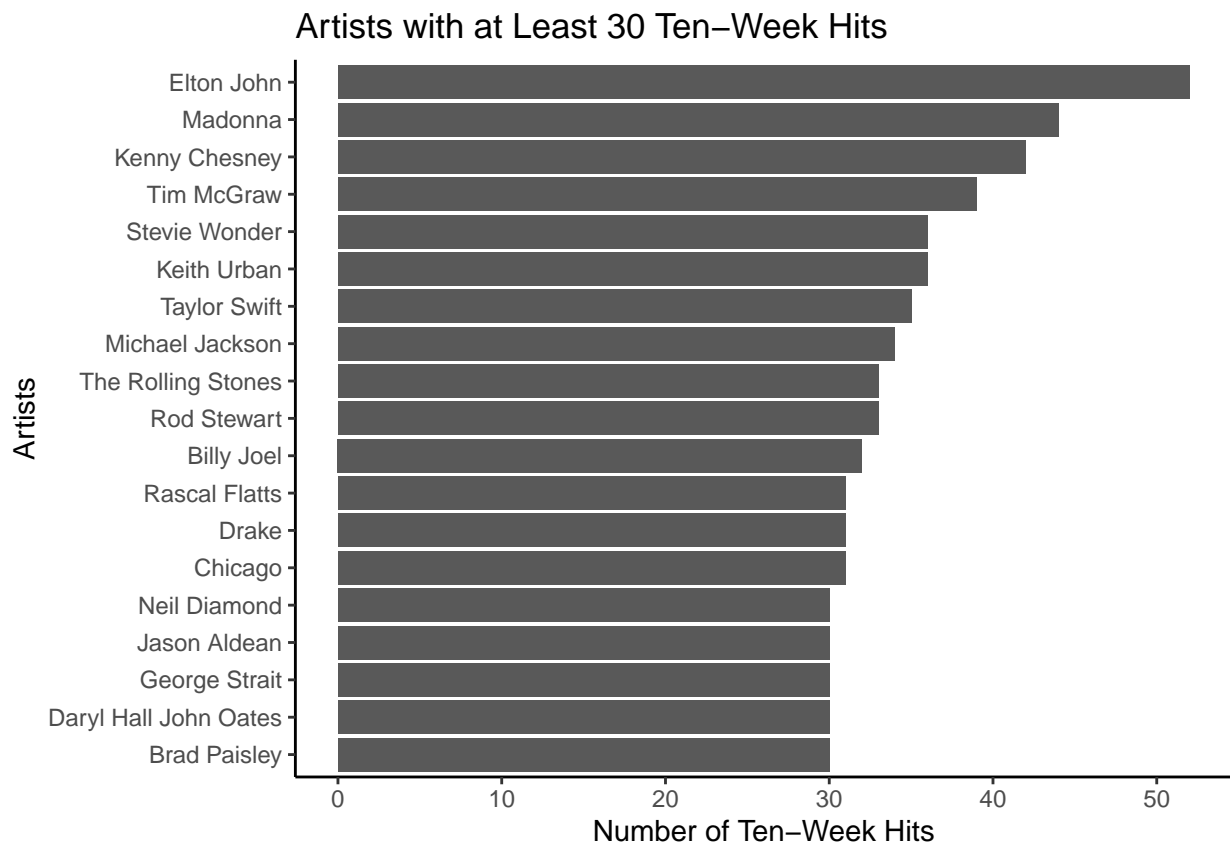
Part B

The Number of Unique Songs in the Billboard Top 100 Over the Years



Number of unique songs appearing at any position on the Billboard Top 100 chart in any week from 1959 to 2020. According to this line graph, the number of unique songs on the Billboard Top 100 peaked in the mid-1960s, and then rapidly declined until around 2004. Then, it started to increase again, reaching a level near the mid-60s peak in 2020. Since the number of unique songs reflects musical diversity in a given year, this indicates that music diversity reached a peak level in the mid-60s before gradually declining, reaching its lowest point in the early 2000s, but in recent years, there has been an increase in music diversity.

Part C



The bar plot shows the number of ten-week hits each artist had in their career, for the artists with at least 30 ten-week hits. The artists are listed according to their number of ten-week hits, with Elton John who has the highest number of ten-week hits shown on the top.