Homework 2
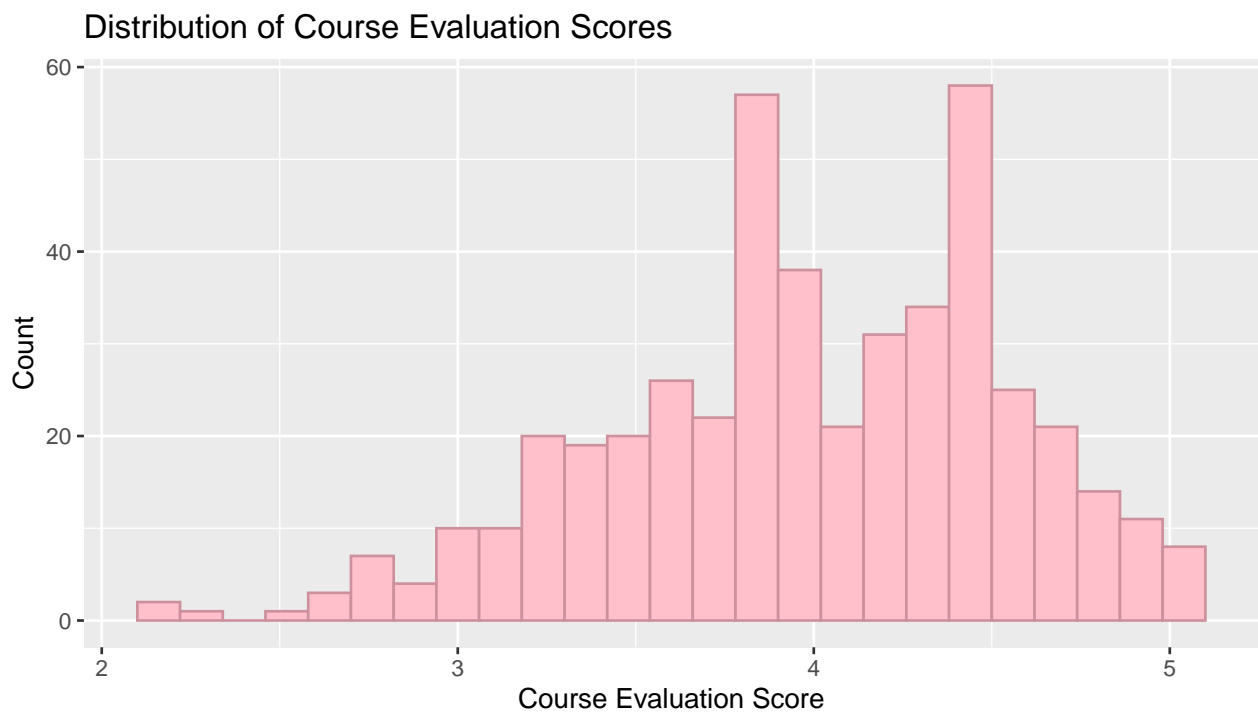
Juhi Malwade - jm97555

1/25/2024

Github Link

Problem 1: Beauty, or not, in the classroom

Part A

Distribution of Course Evaluation Scores

The distribution of course evaluation scores from a sample of 463 UT Austin courses is skewed left and bimodal. The mean evaluation score is 3.998 and the median evaluation score is 4. Overall, UT students gave fairly good ratings to their professors.

Part B

Course evaluations for Native vs Non–native English speakers
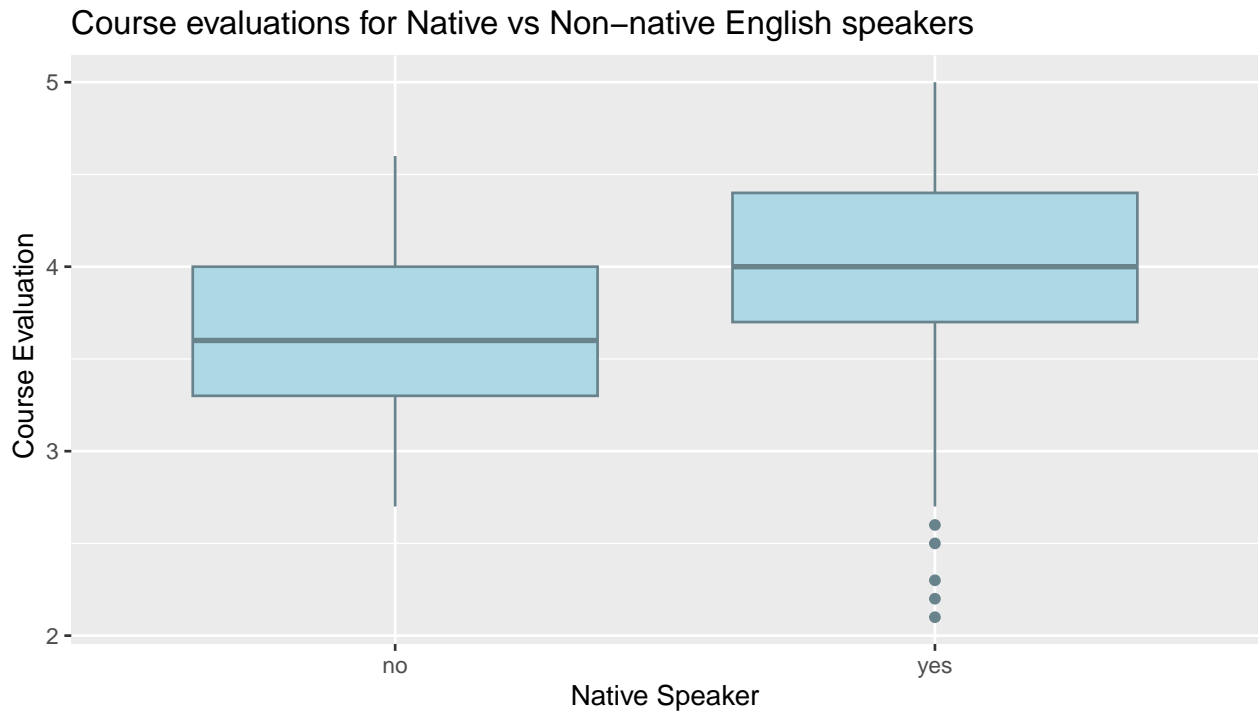
The distribution of course evaluations for non-native English speakers is fairly symmetric with no outliers while the distribution of course evaluations for native English speakers is skewed left with low outliers. The mean non-native speakers is 3.689 and the median is 3.6, while the mean for native speakers is 4.018 and the median is 4. Overall, native speakers had higher course evaluations than non-native speakers.

Part C

Course evaluations for Male vs Female UT Professors
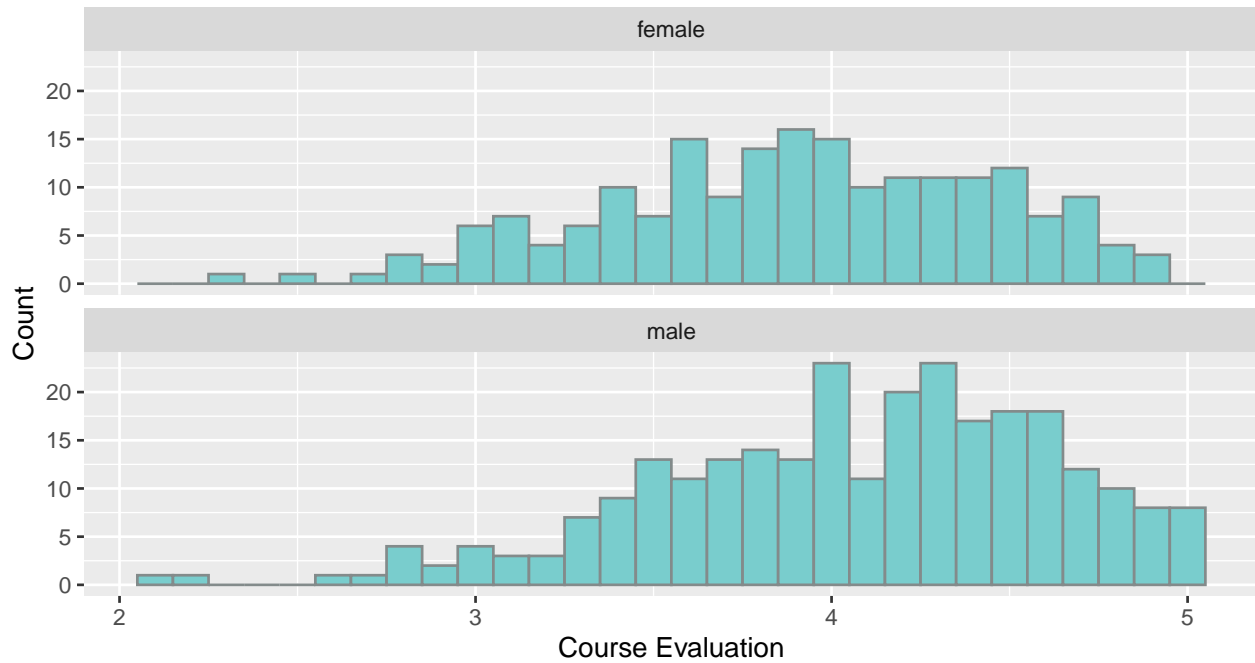
The distribution of course evaluations for both male and female instructors is skewed left. The mean evaluation for females is 3.901 and the median is 3.9, while the mean for males is 4.069 and the median is 4.15. Overall, male instructors had higher course evaluations than females.
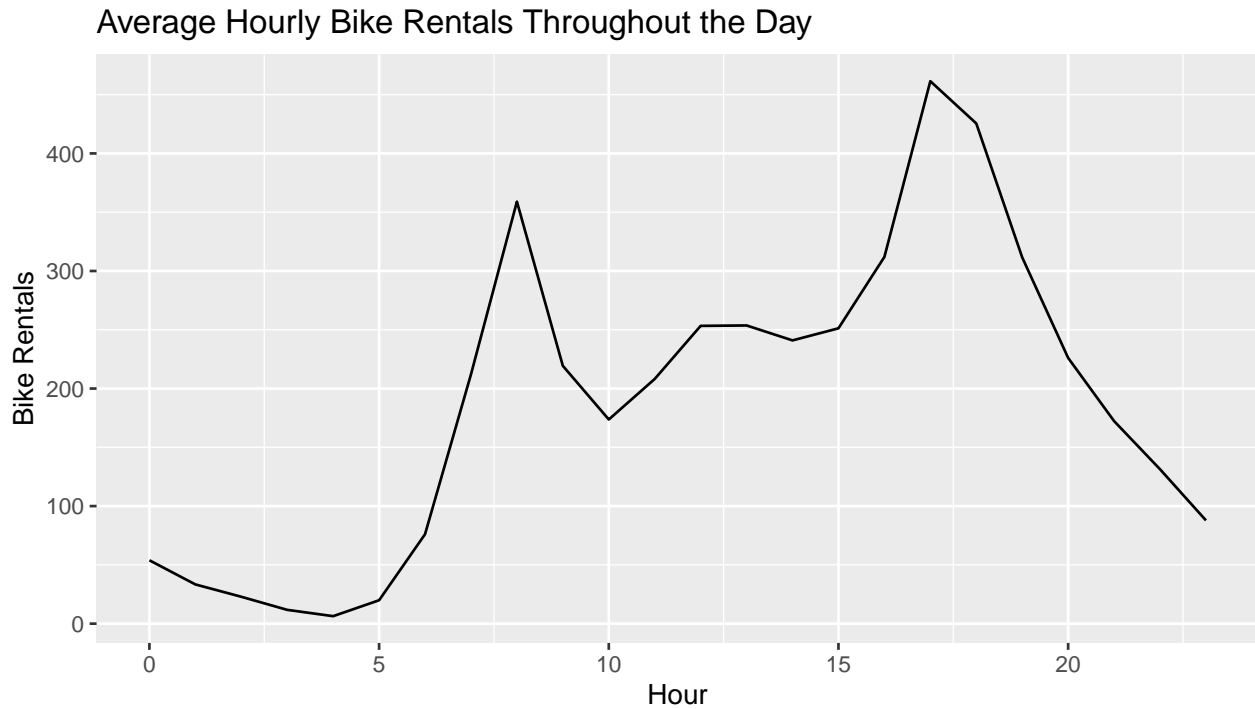
Part D

Physical Attractiveness vs Course Evaluation
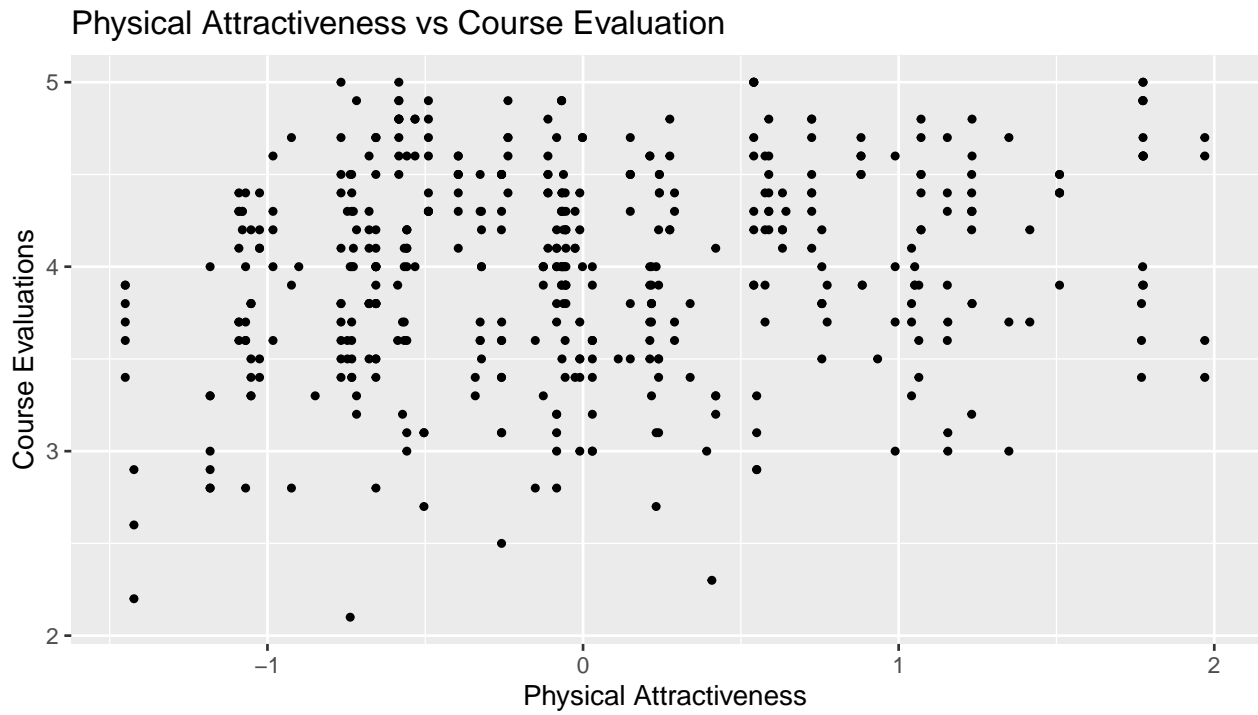
There is a very weak, positive correlation between a professor's physical attractiveness vs course evaluation because the correlation coefficient for the two variables is 0.19.

Problem 2: Bike sharing

Part A

## Average Hourly Bike Rentals Throughout the Day

The plot shows the average hourly bike rentals across all hours of the day. The x-axis shows the hour of the day from 0-23 and the y-axis shows the average bike rentals per hour. The line graph shows that the number of bike rentals is fairly low at the beginning and end of the 24-hour day, when it is dark outside. The minimum average bike rental occurs at hour 4 (4:00AM - 4:59AM) with 6 bikes per hour. Then, throughout the day there are two peaks - one at hour 8 (8:00AM - 8:59AM) with 359.01 bikes per hour and another at hour 17 (5:00PM - 5:59PM) with 462 bikes per hour. Overall, bike ridership is low during the night and high in the morning and late afternoon.

Part B

Average Hourly Bike Rentals Throughout the Day
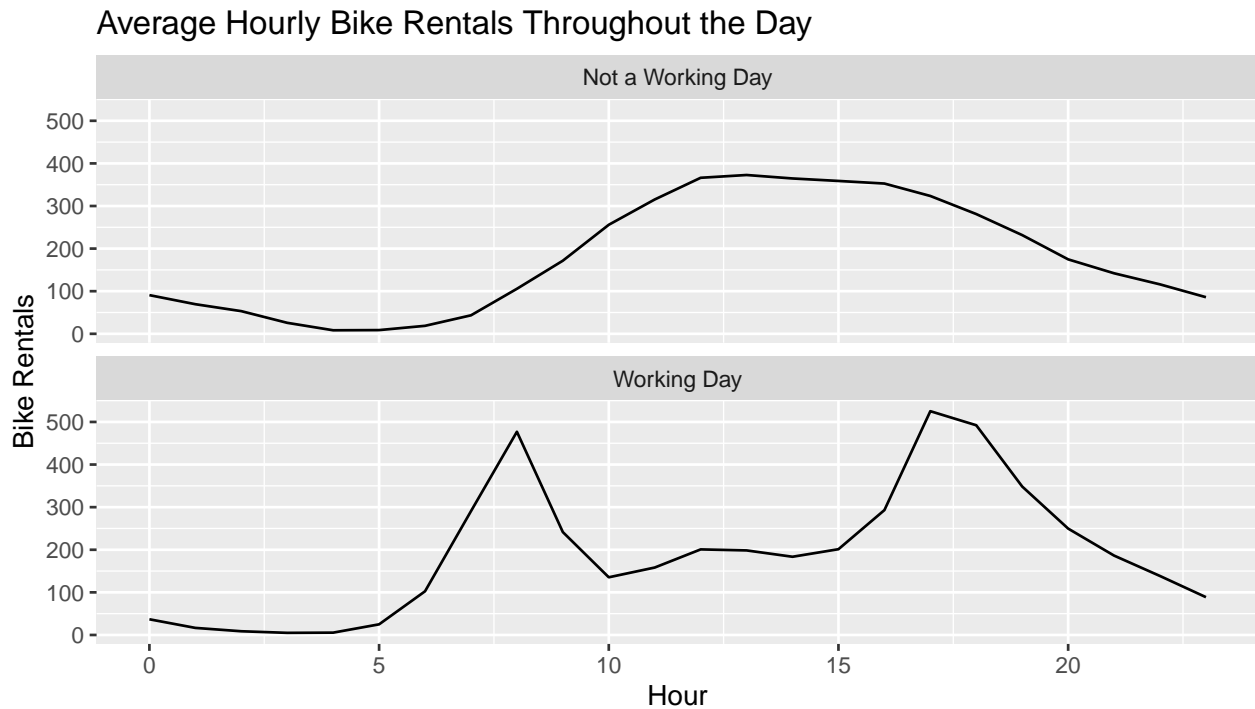
The top graph shows the average hourly bike rentals every hour for a non-working day. The average bike rentals is very low in the early morning and late at night. There is one peak at hour 13 with 373 bike rentals; however the bike rentals plateaus and remains consistently between 300 and 400 between hour 11-17. The bottom graph shows the average hourly bike rentals for a working day. It has more fluctuations than the non-working day graph with two peaks: one at hour 8 with 477 bike rentals and hour 17 with 525 bike rentals. This peaks represent peak commute times on a working day: 8:00AM-8:59AM and 5:00PM-5:59PM. During the dark hours of the day, bike rentals are fairly low. Overall, bike rentals peak and stay consistently high in the afternoon for non-working days and peak in the morning and late afternoon for working days.

Part C

Average Bike rentals during 9 AM hour
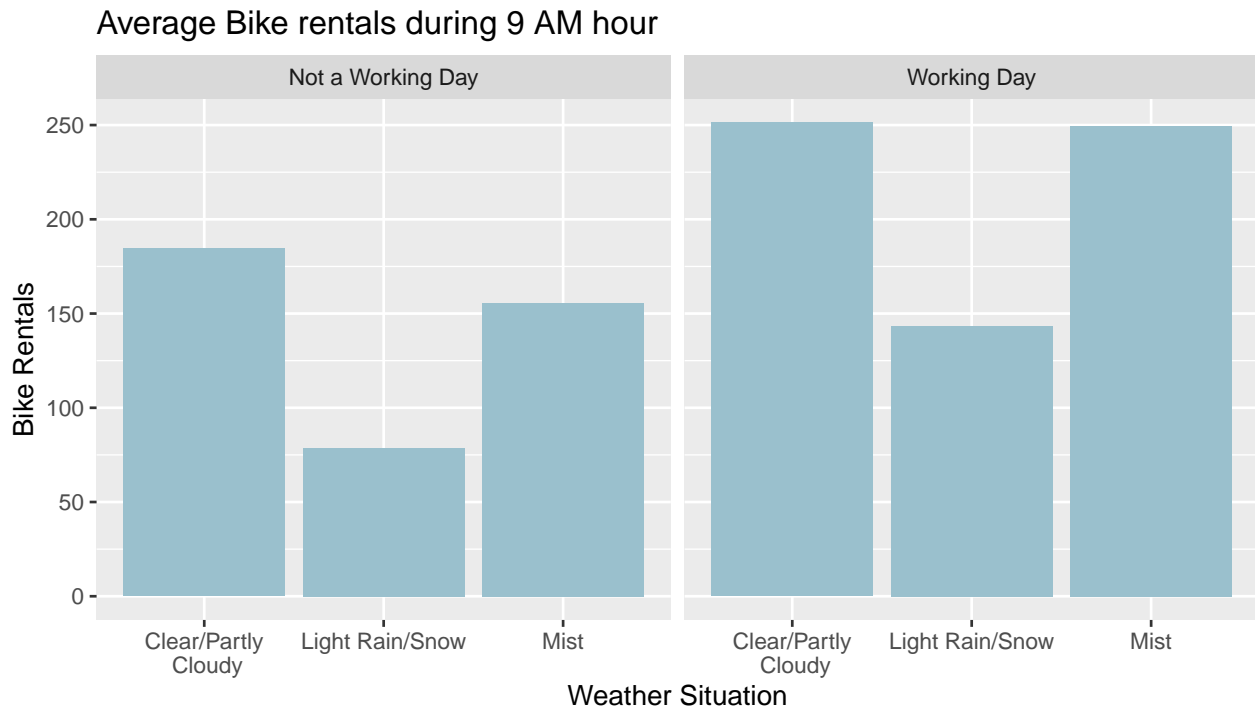
The graph on the left shows ridership at 9AM by weather situation on a non-working day. The graph on the right shows ridership at 9AM by weather situation on a working day. Overall, the ridership is higher on working days compared to non-working days at 9AM, despite the weather situation. For both non-working days and working days, ridership is the highest on clear/partly cloudy days and lowest on days there is light rain/snow. However, the difference between bike rentals on clear/partly cloudy days and misty days is significantly smaller for working days compared to non-working days. This shows that riders are more resistant to riding a bike in misty conditions on non-working days compared to working days. In summary, weather does play a significant factor in ridership at 9AM for both non-working and working days, as ridership was low during rainy weather conditions and high during clear weather conditions.

Problem 3: Capital Metro UT Ridership

Figure 1

Average Boardings by Hour, Day of the Week, and Month
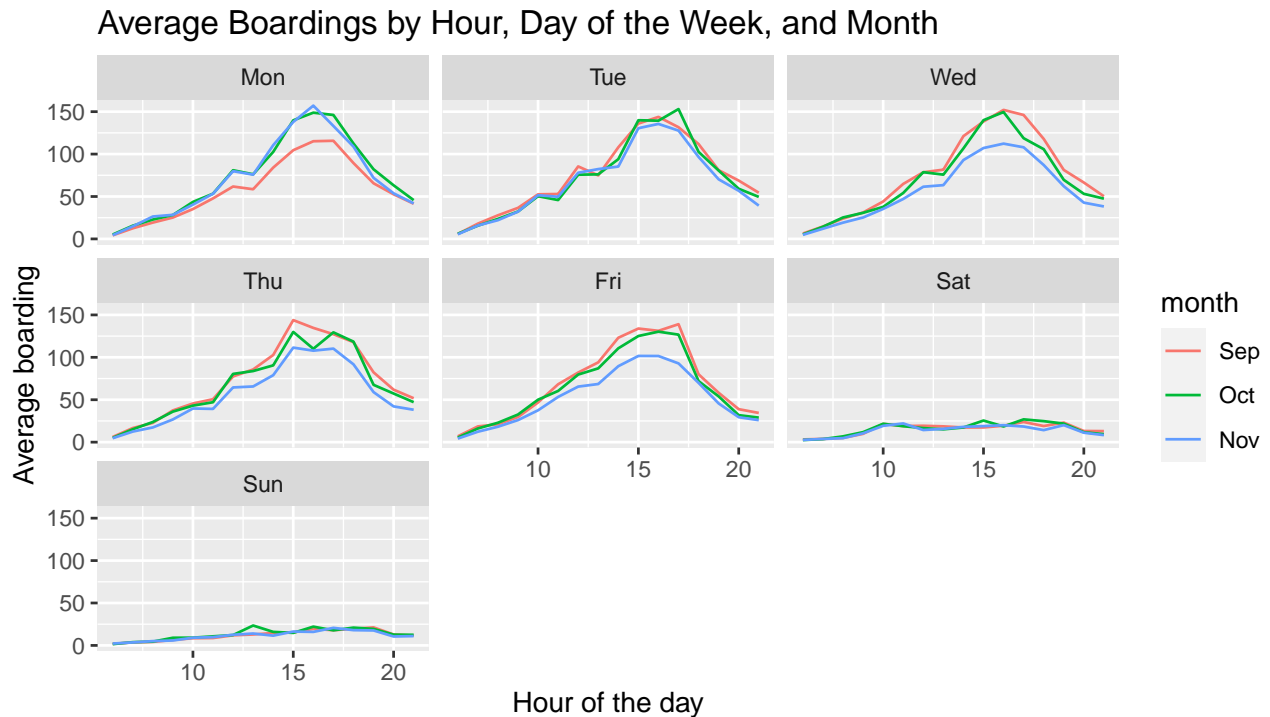
The figure shows the average boardings on the UT capital metro bus for each day of the week, in which the colors represent the month of the year. For all 3 months, there is a significantly smaller number of boarders on Saturday and Sunday. This is because there are no classes on the weekends, and thus less students commuting on campus. On weekdays, the average boarding typically peaks in the mid-afternoon while there is no obvious peaks on weekends. In addition, the average boardings on Mondays in Septembers (represented by the red line) appear lower compared to other days and months. This trend occurs because UT no classes on a Monday in September due to labor day weekend; thus, less students commute on the shuttle on that day. The average boardings on Wednesday, Thursday, and Friday in November (represented by the blue line) also appear lower because UT's November Thanksgiving break cancels classes on these days of the week. Overall, the average boarding is typically dependent on the days UT has classes - when there are no classes, the number of boarders is lower and when there are classes, the number of boarders is higher.

Figure 2
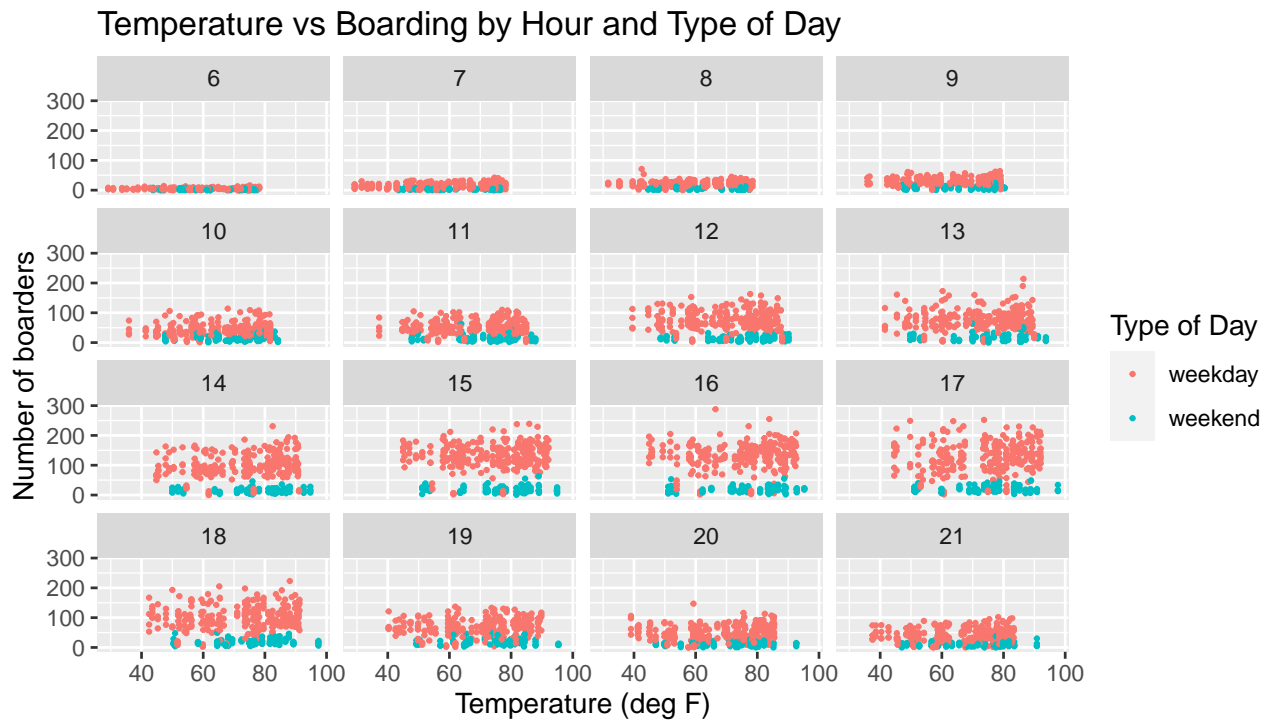
The figure compares temperature vs number of boarders on the UT shuttle. It is faceted by the hour of day and its points are colored according to whether it is a weekday or weekend. Overall, there are typically more boarders on weekdays (represented by red dots) than weekends (represented by blue dots). In addition, boarding is highest in the afternoon/early evening compared to the morning. When the hour of day and weekend status is held constant, the temperature does not seem to have a noticeable effect on the number of UT students riding the bus. Students will typically board the bus regardless of the temperature outside, as they must attend their classes. In summary, temperature does not have a significant effect on the number of boarders; instead, factors like the type of day and hour of day have a larger effect.

Problem 4: Wrangling the Billboard Top 100

Part A

| 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

The table shows the top 10 most popular songs since 1958, as measured by the total number of weeks that a song spent on the Billboard Top 100. The most popular song was "Radioactive" by Imagine Dragons, spending 87 weeks on the Billboard chart. The second most popular song was "Sail" by AWOLNATION, spending 79 weeks on the chart.

Part B
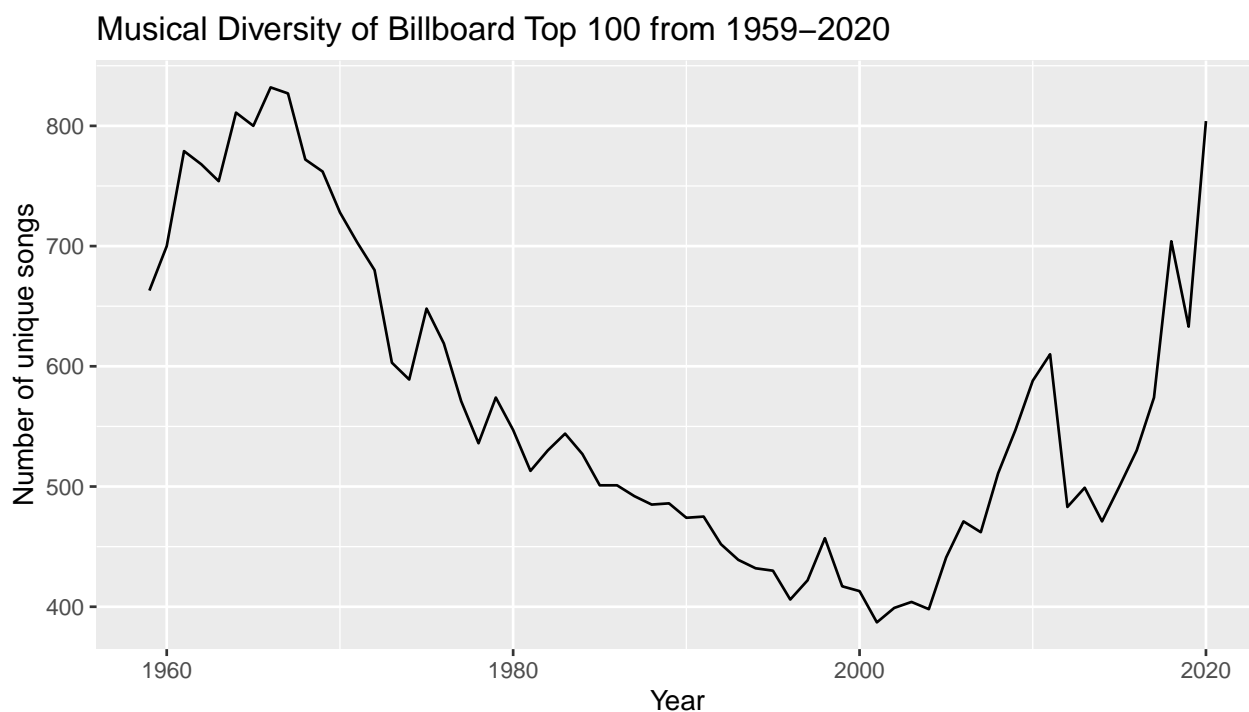
The line graph shows how the number of unique songs in the Billboard Top 100 changed from 1959-2020. The number of unique songs peaked in 1966 with 832 songs. Then, the number of unique songs declined and hit a minimum in 2001 with 387 songs. Since 2001, the number of unique songs has been increasing overall, with there being 804 unique songs in 2020.

Part C

Number of Billboard Top 100 Ten−week Hits for Top Artists
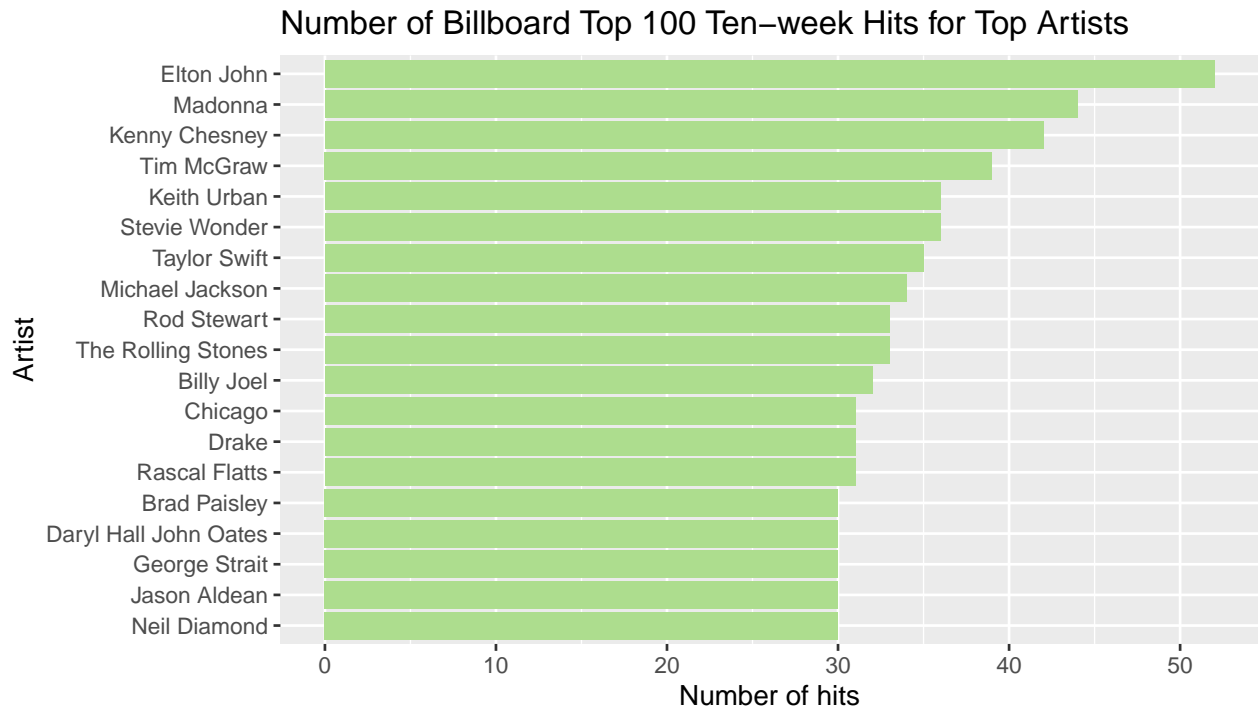
The bar graph shows the 19 artists in U.S. musical history since 1958 who have at least 30 songs appear on the Billboard Top 100 chart for at least 10 weeks. Elton John has had the most songs appear for at least 10 weeks, with 52 hits. Madonna comes in second with 44 hits and Kenny Chesney comes in third with 42 hits.