Homework 2

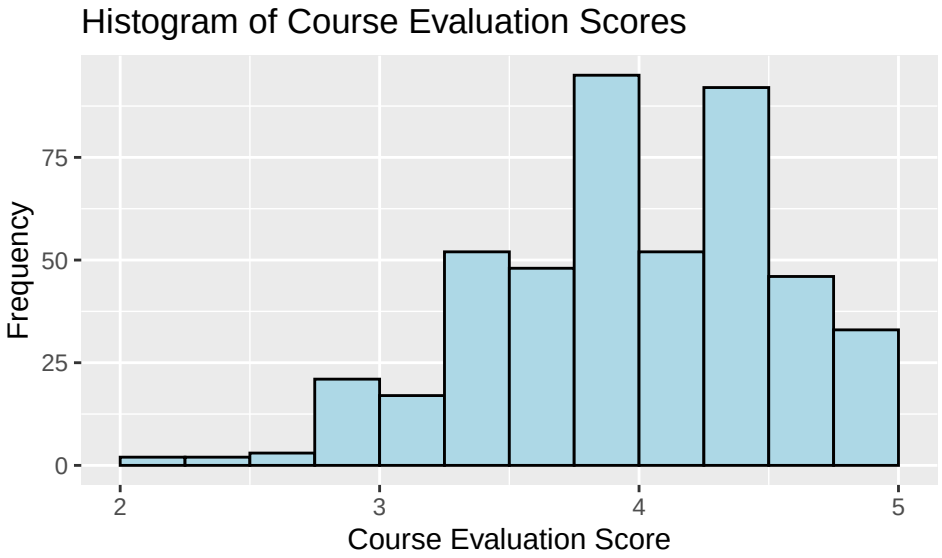
Isaiah Benny EID: ieb357

2024-01-30

Github

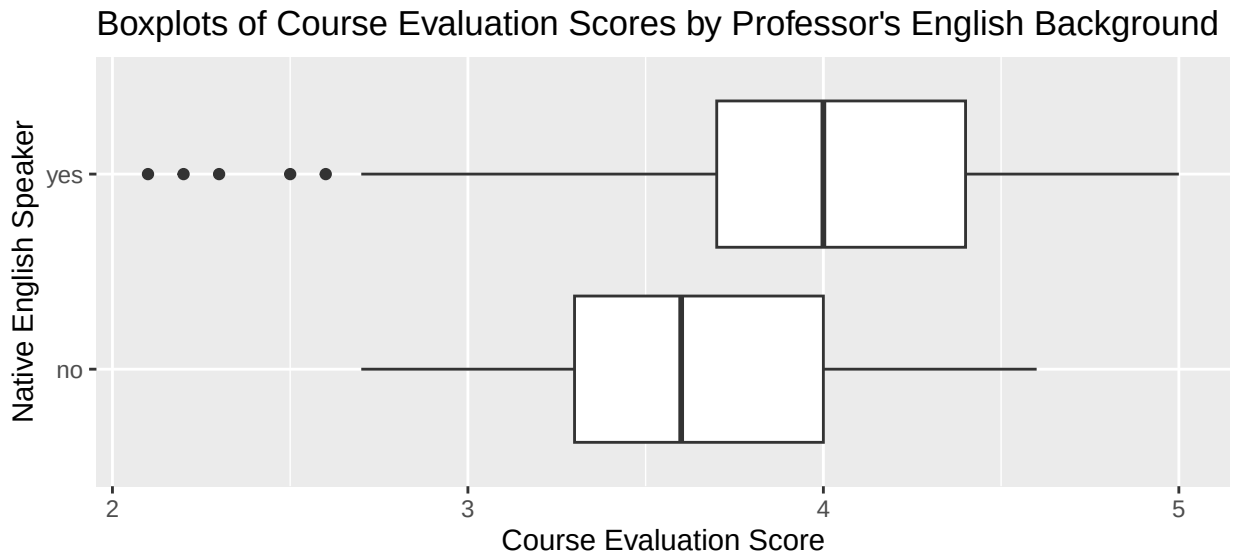
Problem 1

Part A



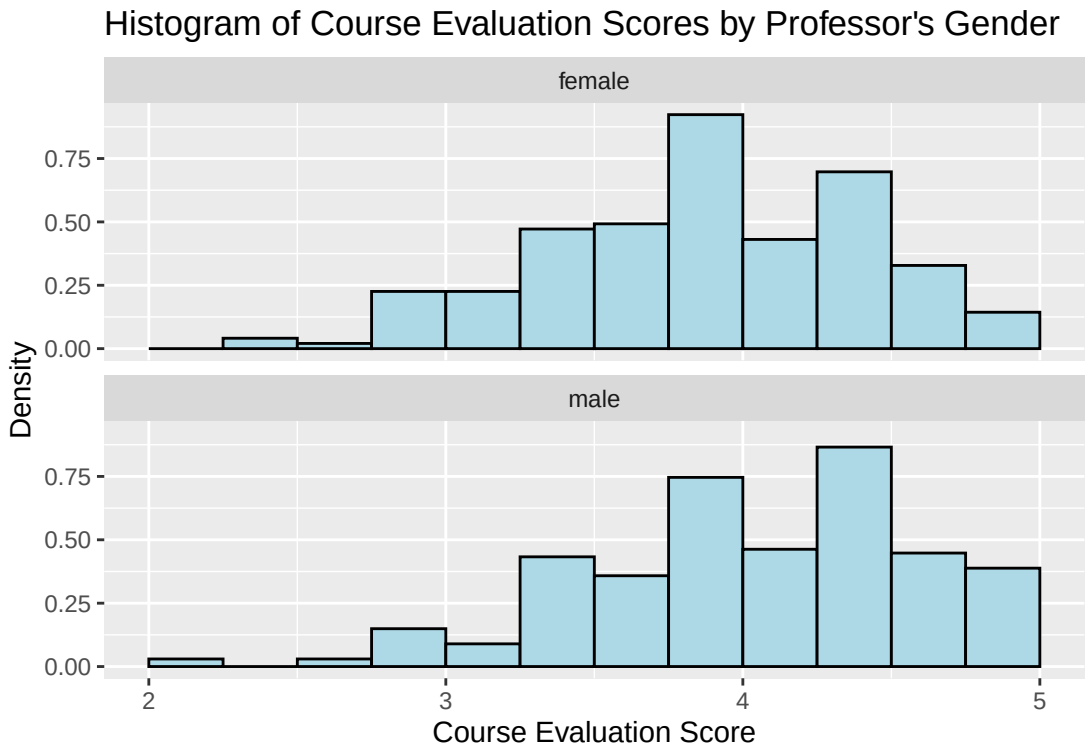
The distribution of course evaluations scores appears to be positively skewed with a median of 4 points. The interquartile range of this distribution is 0.8 points, and no professor in the sample has received an average rating below 2.1 points. Additionally, this histogram seems to show that most professors don't receive "low" scores, in that only a small proportion of this sample of professors had average ratings below a 3.

Part B



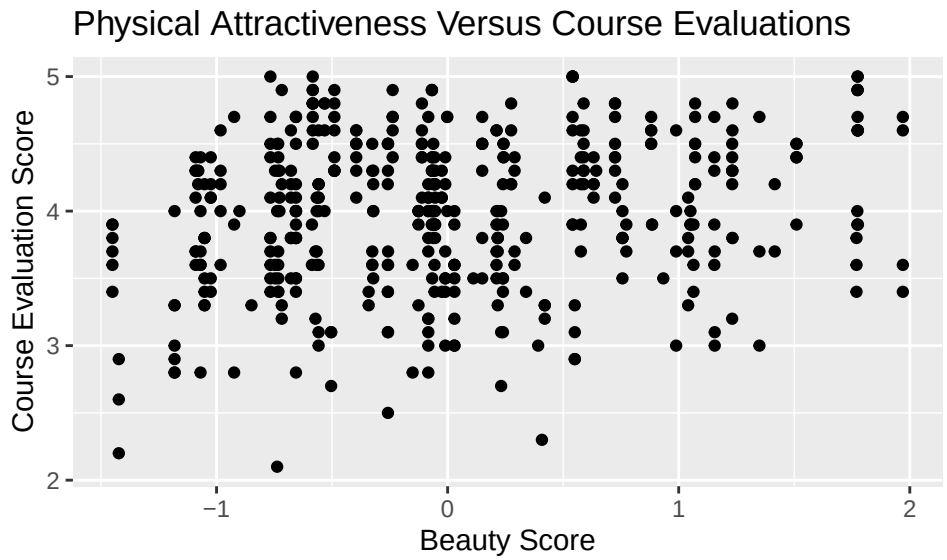
The boxplot above indicates that there may be a difference in course evaluations based on whether the professor is a native english speaker. Professors who are native speakers have a median average rating of 4 points, while non-native speakers have a median average rating of 3.6 points. The course evaluations for professors who are native English speakers and non-native speakers have the same IQR of 0.7 points.

Part C



Since there are 195 female professors and 268 male professors, the histograms of their course evaluation scores have been normalized. This helps illustrate that male professors generally have slightly higher average course evaluations, with a median of 4.15 points and an IQR of 0.8 points, compared to female professors' median of 3.9 points and IQR of 0.7 points.

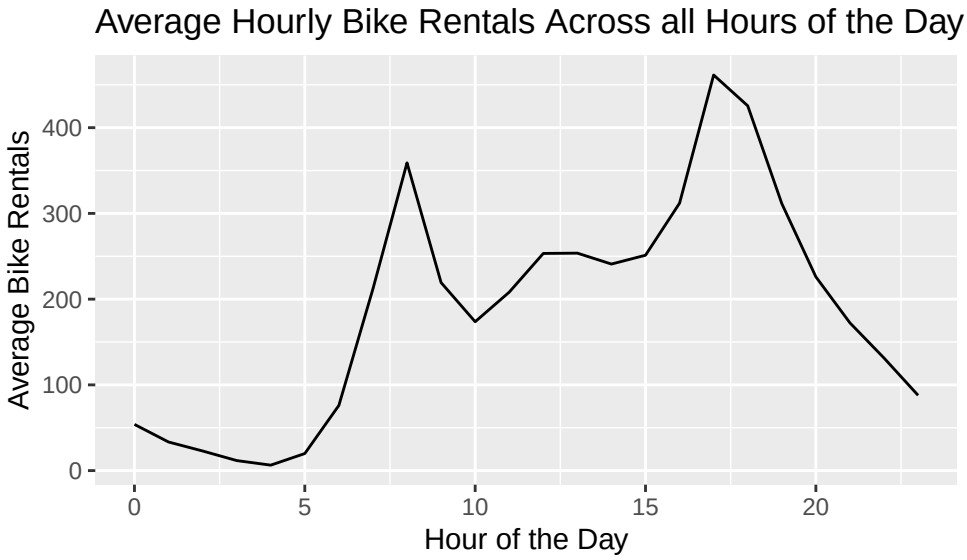
Part D



Judging from the visual appearance of the scatterplot, there does not seem to be a strong association between a professor's beauty score and their average course evaluation score. This is confirmed by the correlation coefficient of these two variables, which is 0.19, indicating a weak positive correlation.

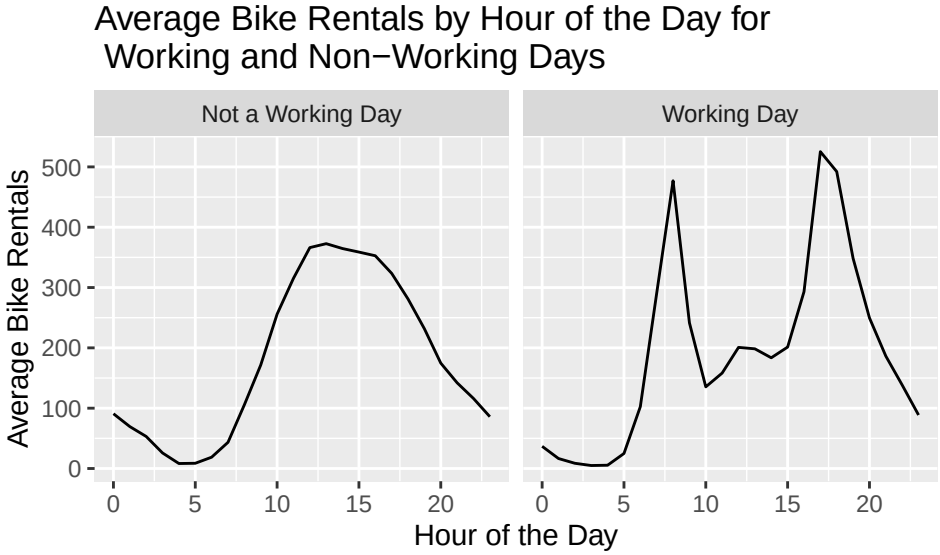
Problem 2

Plot A



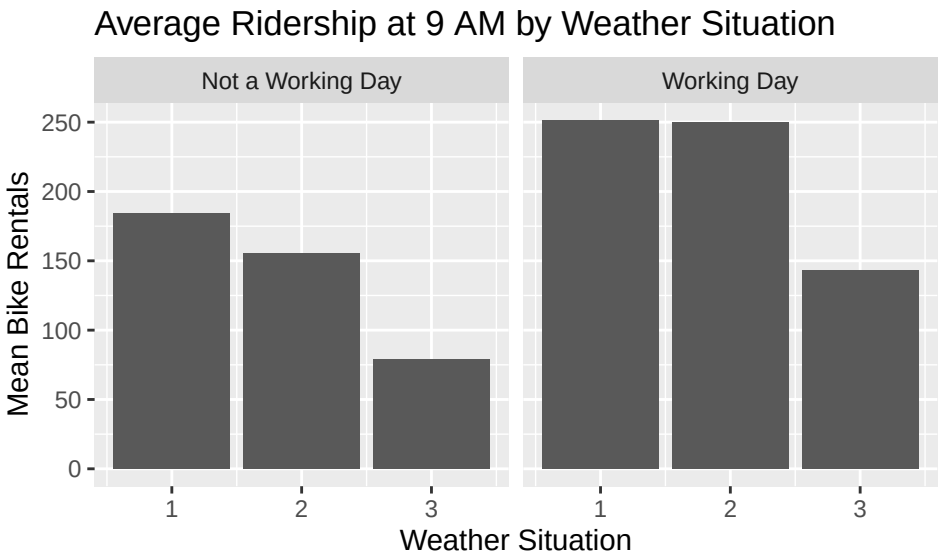
This graph shows the relationship between the hour of the day (0-23) and the average number of bike rentals for that given hour. From this graph, we can ascertain that, on average, bike rentals tend to peak at around 8 AM and 5 PM. This could be because these hours generally coincide with the start and end times of many people's jobs. Additionally, there seems to be a moderate amount of bike rentals during the middle of the day.

Plot B



These line graphs represent the same data as the previous graph, but they are split according to whether the day is a working day. When it is not a working day, we can see that there are no longer any peaks in bike rentals at 8 AM or 5 PM. Instead, there is a less steep peak during the middle of the day. As for working days, the peaks at 8 AM and 5 PM have become more exaggerated. This gives further plausibility to the idea that the peaks at 8 AM and 5 PM are caused by those going to work.

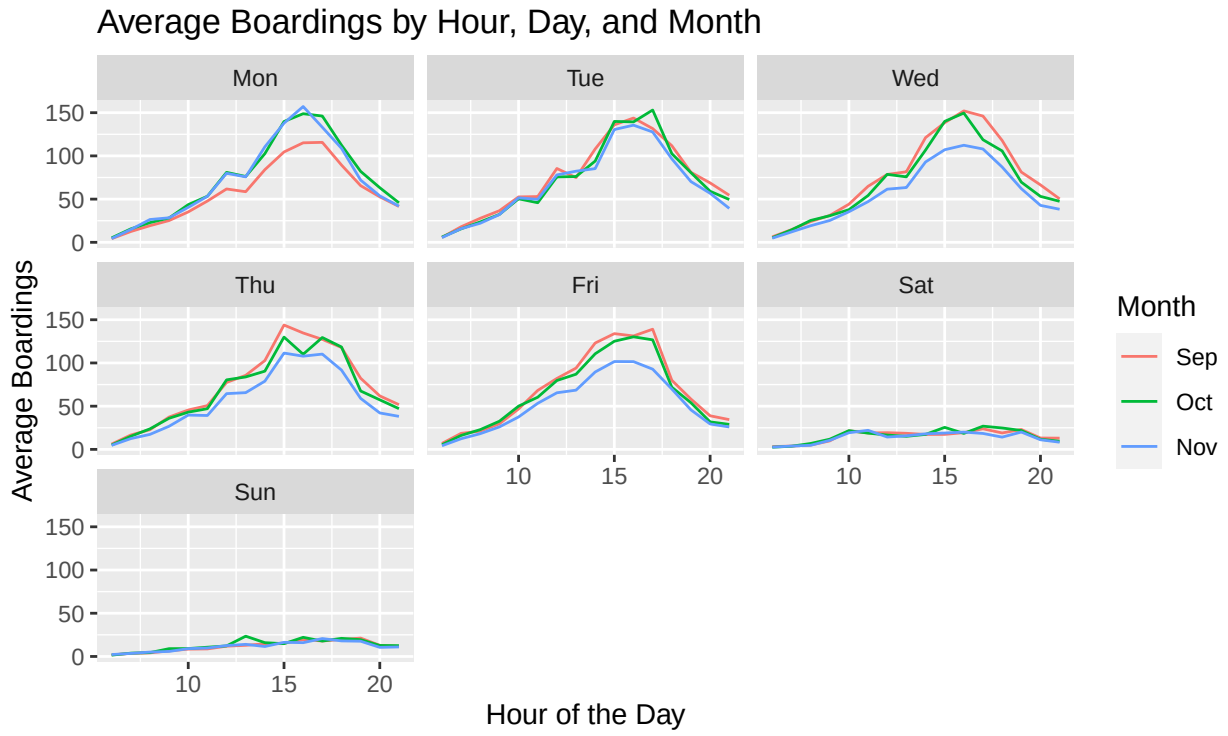
Plot C



These bar graphs show the average bike rentals for different weather situations. The data is also split based on if the day is a working day. A weather code of 1 indicates generally clear conditions, 2 indicates mist with some clouds, and 3 indicates light snow, light rain, or thunderstorms. These two graphs show that unfavorable weather does have some negative association with bike rentals. This makes sense, as it is generally less safe to ride a bike when there is rain or snow. However, this negative association is less pronounced on workdays, especially when comparing weather codes 1 and 2. This could reflect that a decent number of people rely on renting a bike to get to work, regardless of the weather.

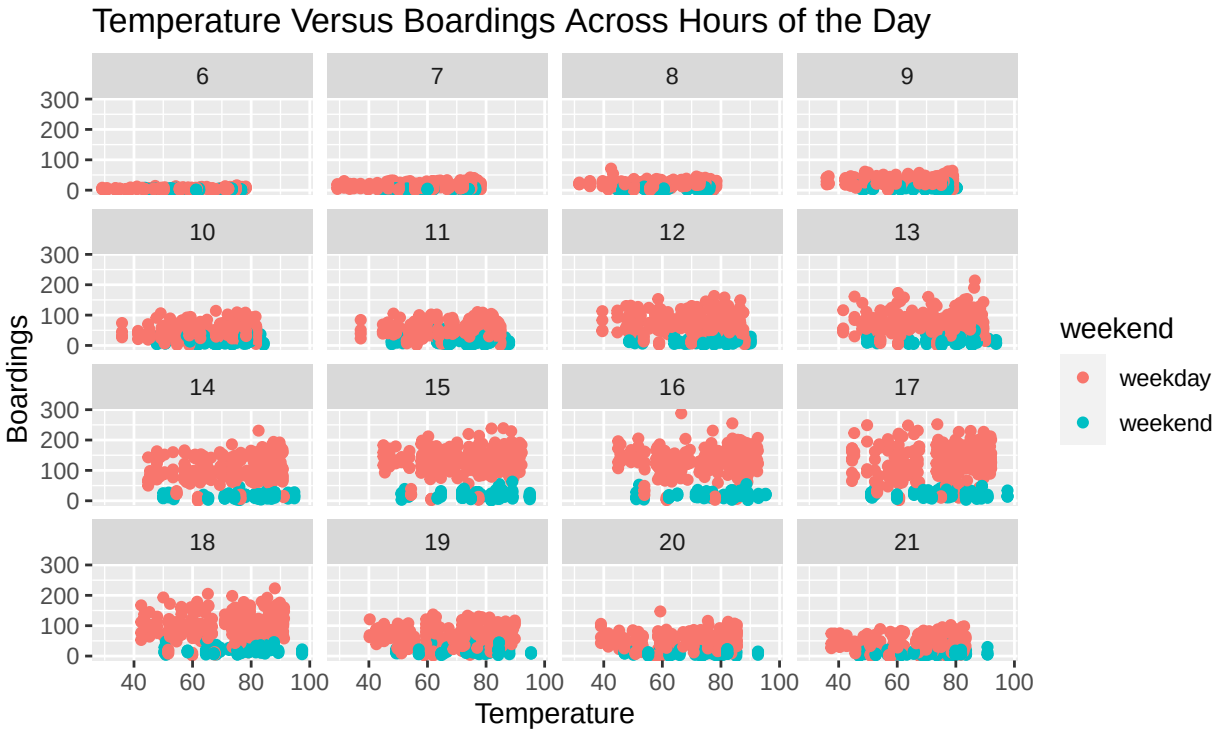
Problem 3

Plot 1



This figure shows 7 different line graphs (one for each day of the week) and displays the relationship between the hour of the day and the average boardings on any Capital Metro bus at the UT campus. There are 3 lines on each graph showing the differences between the months of September, October, and November. Among weekdays, the hour of peak boardings seems to be broadly similar, since all 3 lines for each weekday peak at around 17 hours (5 PM). However, for the weekends, there is no distinct peak throughout the day. One reason why there might be a lower amount of average boardings on Mondays in September is because Labor Day occurs on a Monday in September. During Labor Day, there are no classes held, meaning that there would be a decrease in average boardings. Similarly, one reason why there might be a lower amount of boardings on Wednesdays, Thursdays, and Fridays in November could be because of Thanksgiving break. In past years, UT started Thanksgiving break on the second half of Thanksgiving week, which would cause a decrease in boardings due to the lack of classes being held.

Plot 2



These scatterplots show the relationship between temperature and the number of boardings for each hour of the day. When holding the hour of the day and weekend status to be constant, it appears as if temperature does not have a noticable effect on the number of UT students riding the bus. For any given hour of the day and weekend status, an increase in temperature does not seem to show an increase in boardings. Instead, as the temperature increases, the number of boardings seems to stay the same.

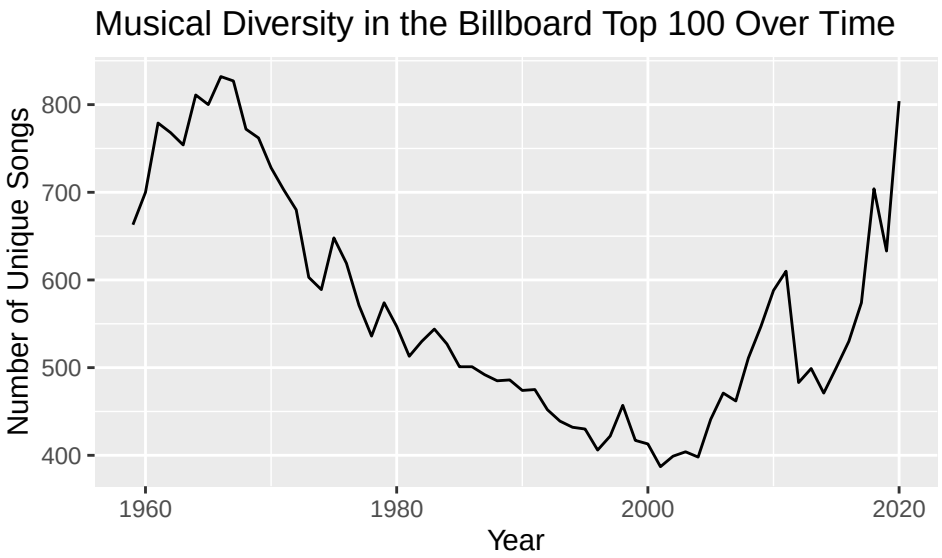
Problem 4

Part A

Performer	Song	Number of Weeks in Billboard Top 100
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

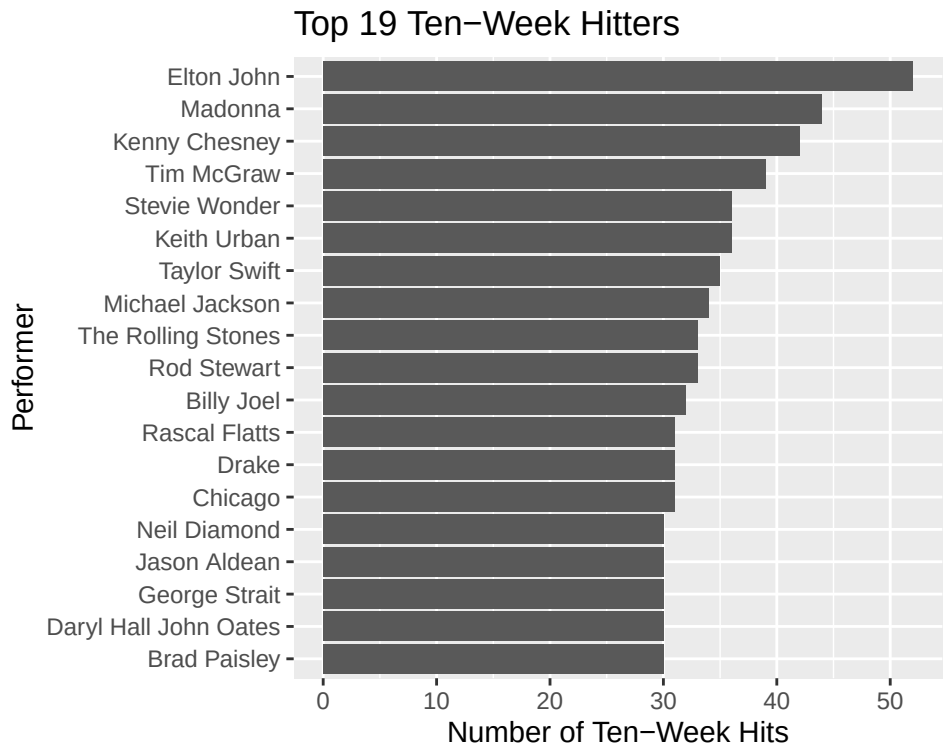
The table above shows the top 10 most popular songs since 1958. Popularity is measured by the total number of weeks that a song spent on the Billboard Top 100.

Part B



This graph shows how the number of unique songs that appear on the Billboard Top 100 changed over time. From the late 1960's to the early 2000's, there appears to be a noticable decrease in the "musical diversity" of the Billboard Top 100. This trend seems to reverse from the early 2000's onwards, but there is one sharp decrease in diversity around 2011.

Part C



This bar graph shows the 19 artists who have had at least 30 songs on the Billboard Top 100 that stayed there for at least 10 weeks.