

SDS315HW

Neil Sebastian
2024-01-29

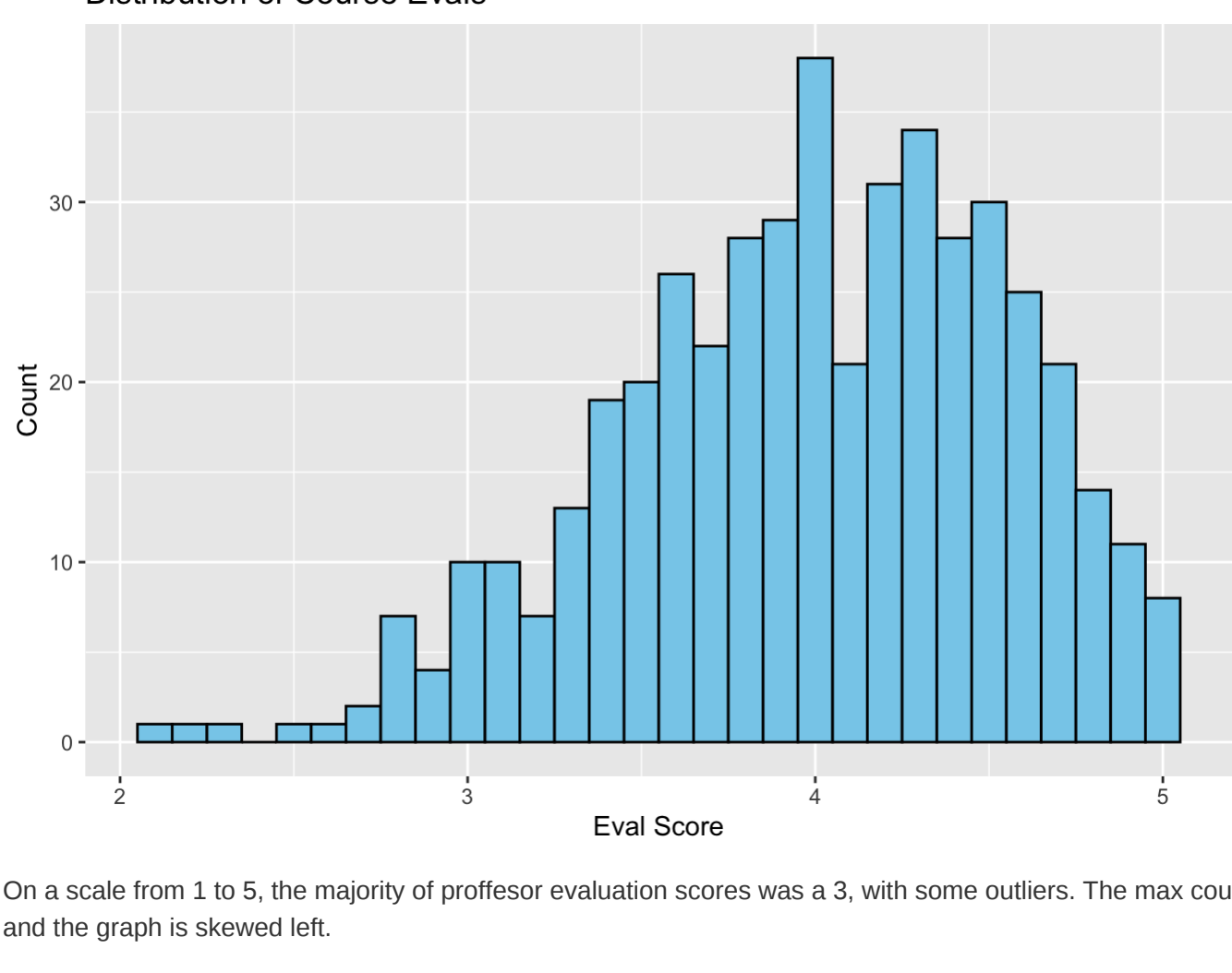
HW2

UT EID: ns36965

[GitHub Link](#)

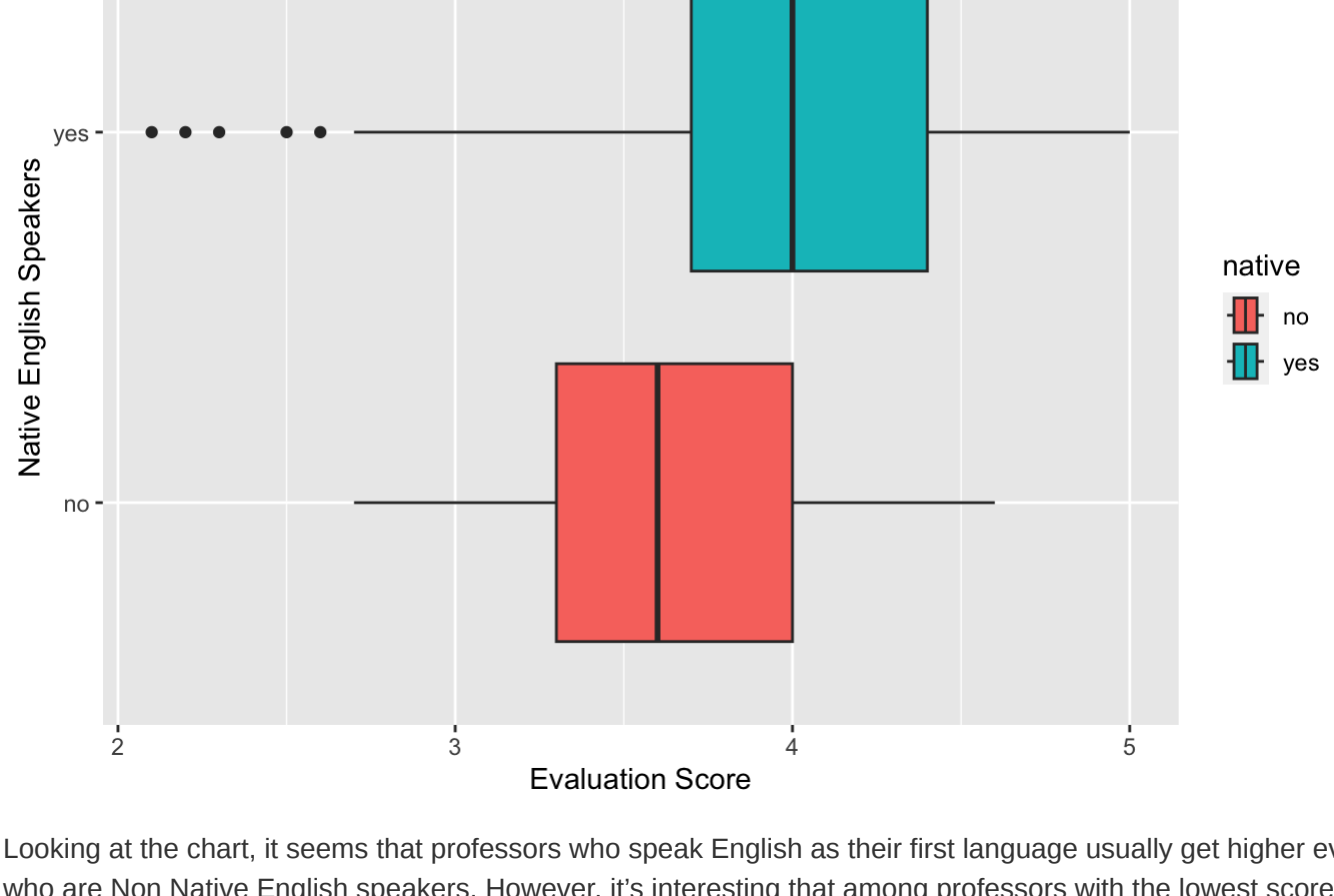
Problem One: Beauty, or not, in the classroom

Part A



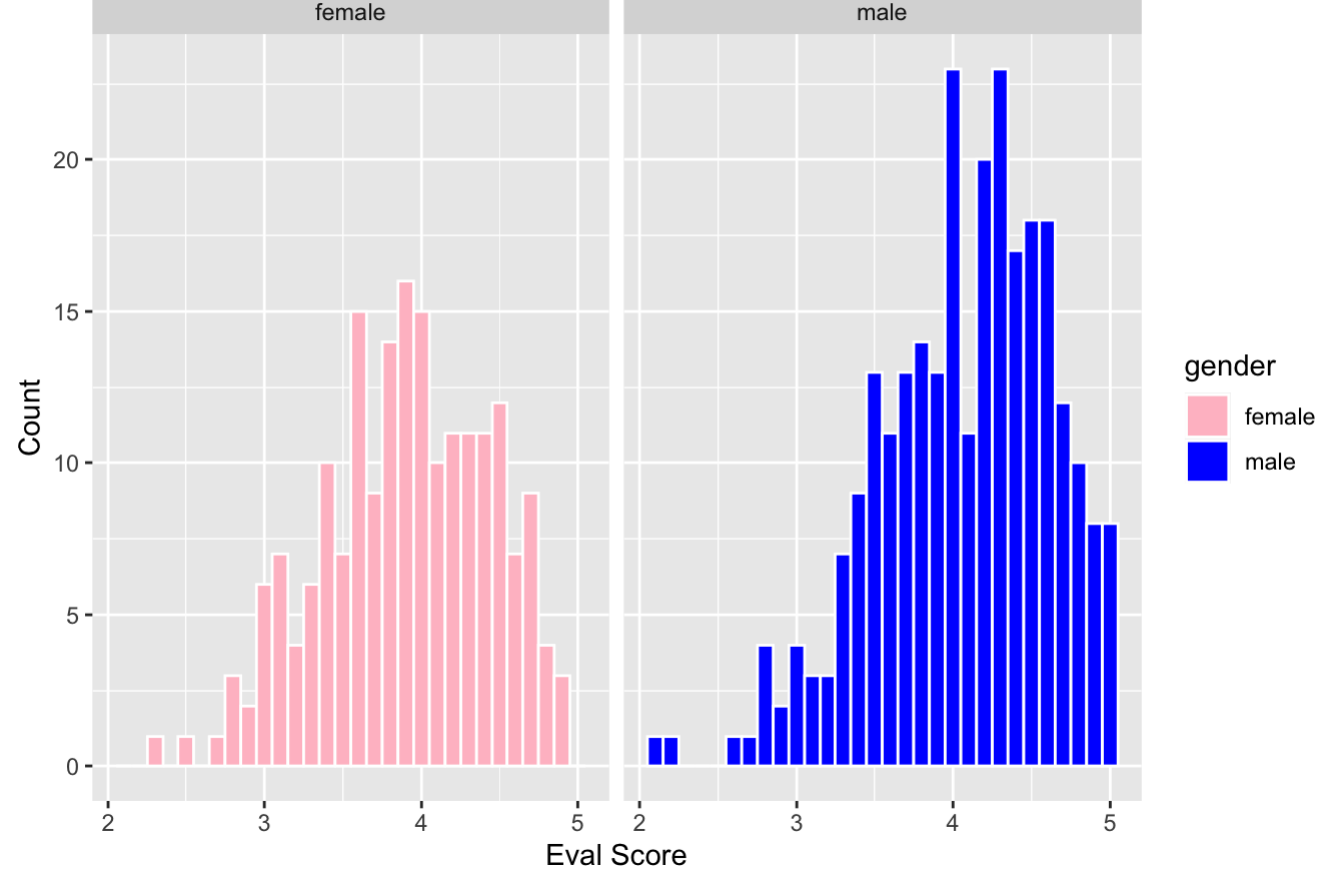
On a scale from 1 to 5, the majority of professor evaluation scores was a 3, with some outliers. The max count of scores for these professors was 4, and the graph is skewed left.

Part B



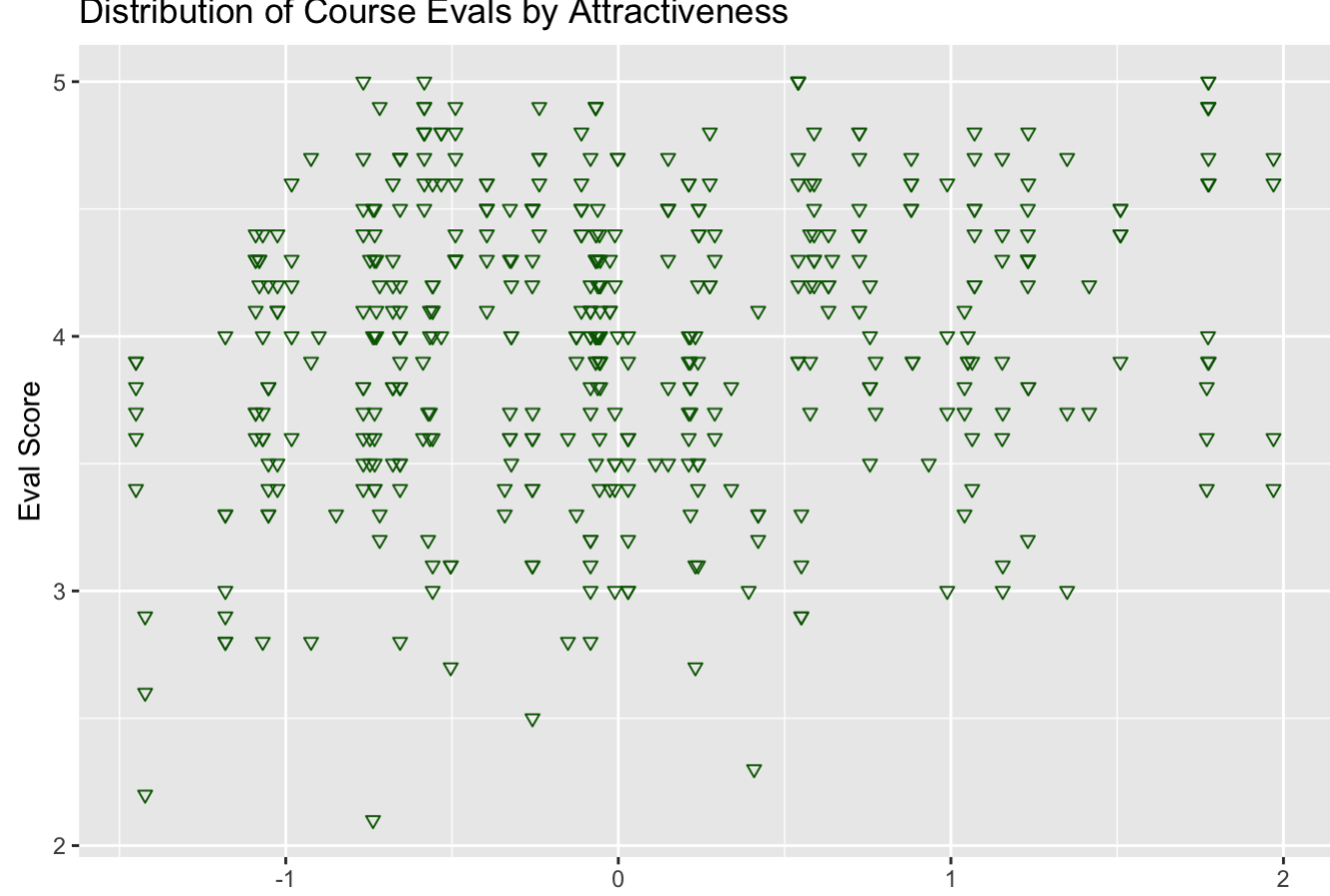
Looking at the chart, it seems that professors who speak English as their first language usually get higher evaluation scores compared to those who are Non Native English speakers. However, it's interesting that among professors with the lowest scores, there are still some native English speakers.

Part C



Although these graphs express a higher count of better evaluation scores for male professors, the total number of male and female professors differ. The general shapes of the graphs are similar, but should be examined in relation to the number of male/female professors.

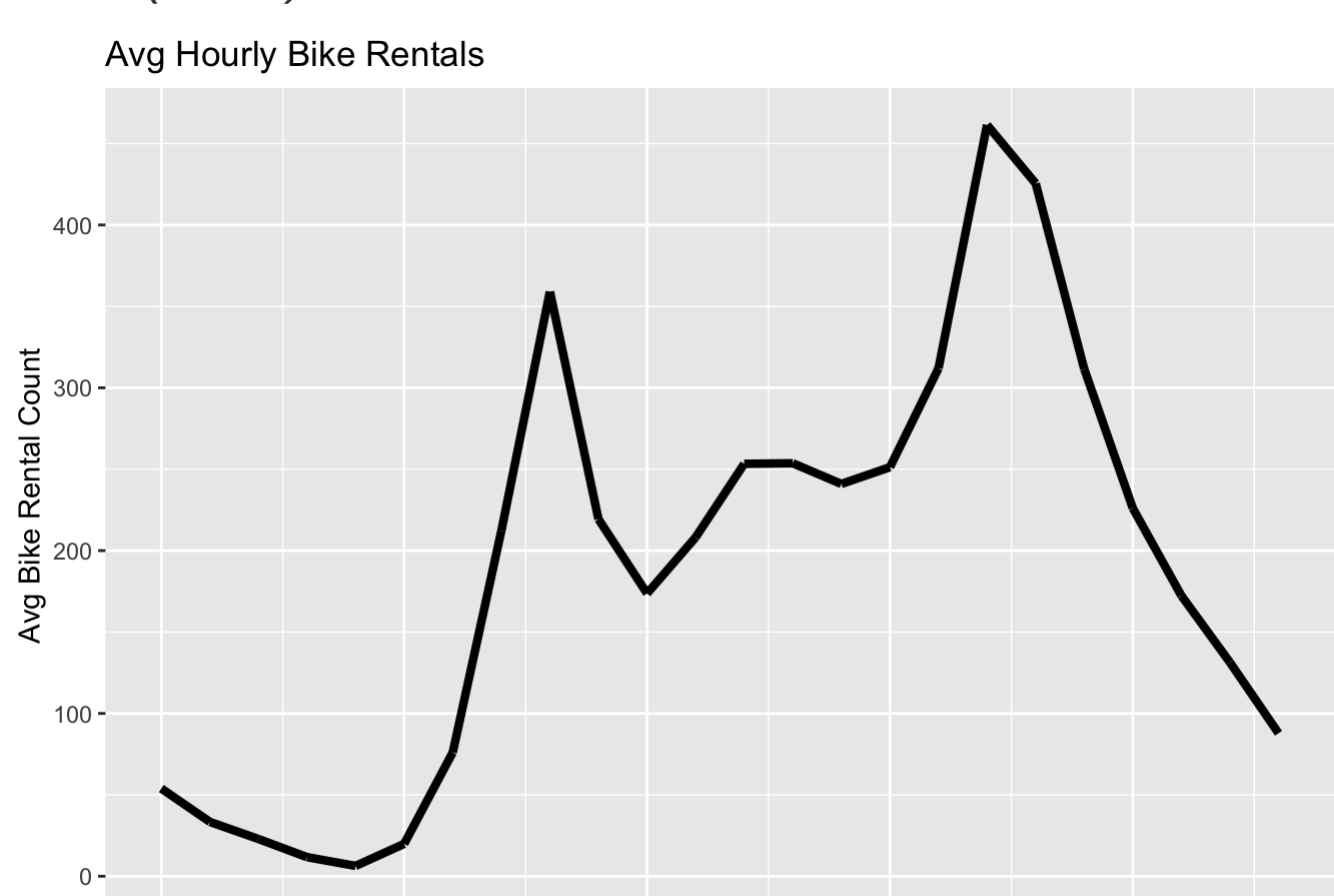
Part D



The graph above shows a slight correlation between the attractiveness of a professor and their evaluation score. With the 'more' attractive professors typically getting higher evaluation scores than 'less' attractive professors, the highest sample of well-scored professors scored below 0 on the attractiveness scale.

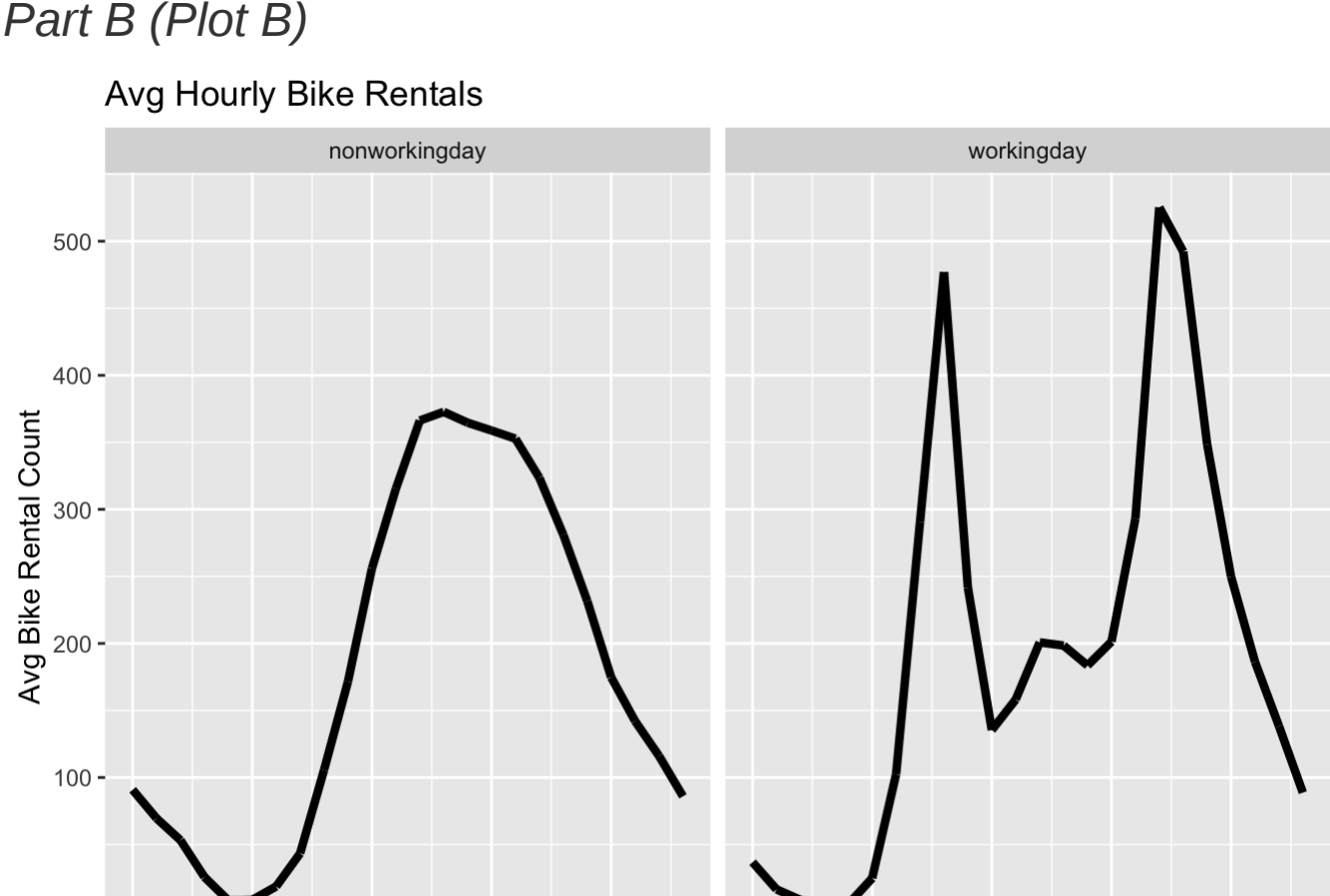
Problem 2: Bike Sharing

Part A (Plot A)



This chart about bike rentals during the day shows that very few bikes are rented in the middle of the night and early morning. The most rentals happen at 8 am and 5-6 pm when people go to and from work. There's also a small increase in rentals around lunchtime (12-1 pm).

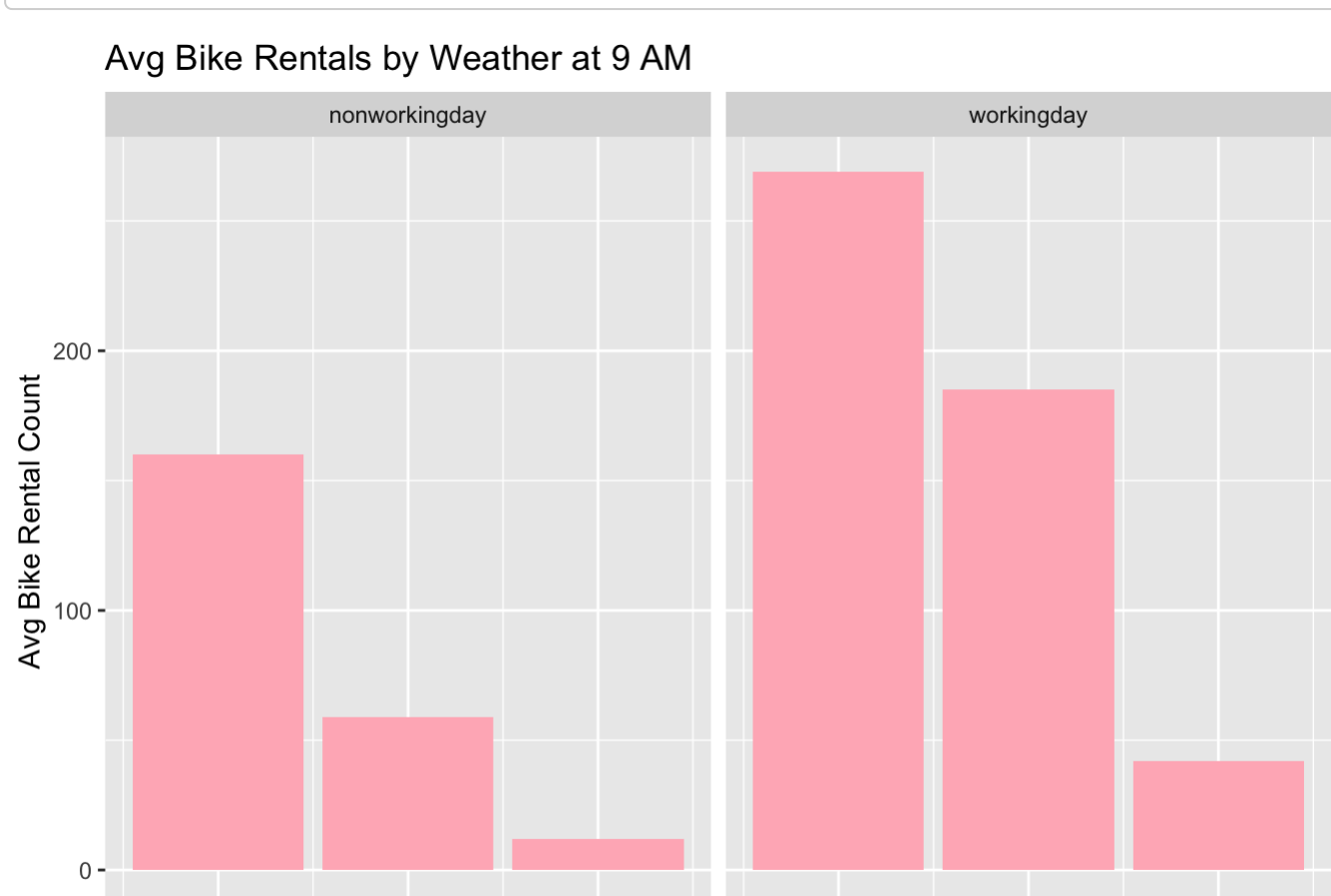
Part B (Plot B)



The working day graph is fairly similar to the previous graph. However, the nonworking day data follows a more 'smooth' curve increase and decrease in activity rather than sudden jumps. The max for the left graph is around 1pm or 2pm, with much less activity around 8am and 5pm which are considered busier times of the day.

Part C (Plot C)

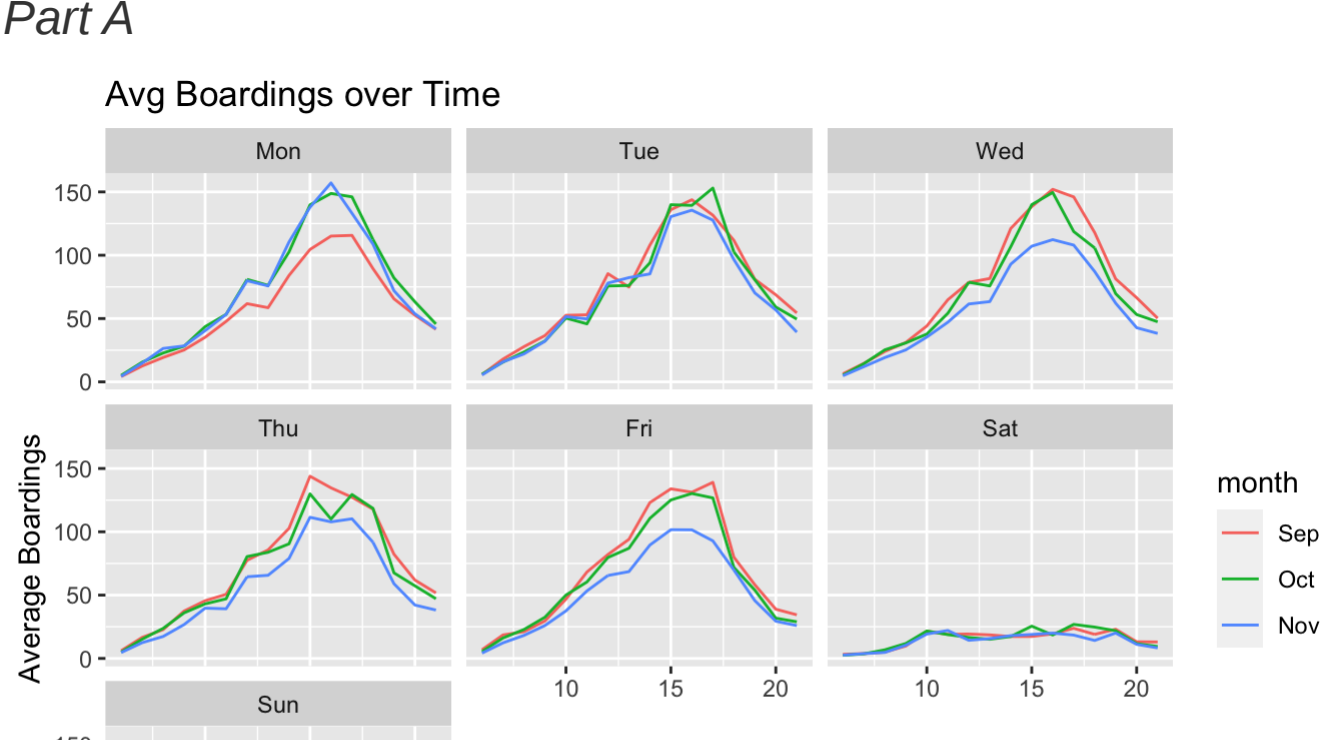
```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



The above graph supports the idea that people are more likely to rent bikes/ride on workdays with bad weather and that on non working days people tend to ride less overall following the same weather pattern habits but at a lower frequency.

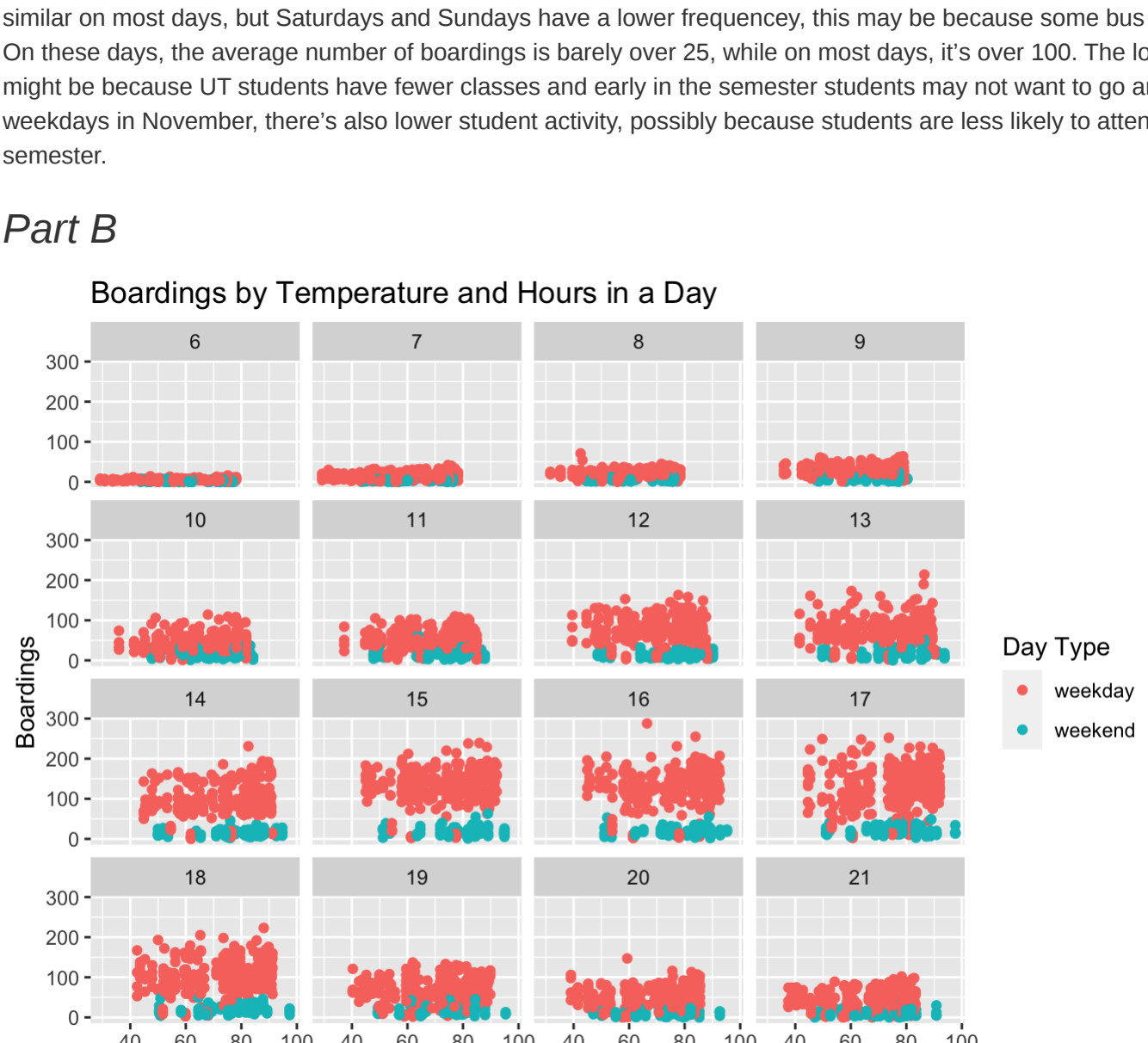
Problem 3: Capital Metro UT Ridership

Part A



These graphs show how many people, on average, board Cap Metro buses in Austin at different times and days of the week. The busiest times are similar on most days, but Saturdays and Sundays have a lower frequency; this may be because some bus routes are not operating on weekends. On these days, the average number of boardings is barely over 25, while on most days, it's over 100. The lower activity on Mondays in September might be because UT students have fewer classes and early in the semester students may not want to go around Austin. However, in later weekdays in November, there's also lower student activity, possibly because students are less likely to attend classes later in the week and semester.

Part B



The graph above is analyzing the total number of boardings throughout the day over multiple recorded months and how the temperature of that hour of the day affects the number of UT students using the buses. Based on the graphs above, one could make an argument that more extreme temperatures have a lower number of UT students traveling at that time, but there are too many unrecognized variables that could affect this data. There could have been very few/no hours in these months with consistent extreme temperatures, which creates the impression of students avoiding the bus due to weather conditions rather than lack of comparable data. A better measure of the noticeable effects of temperature on bus travelers should be conducted based on averages of boardings over hours and temperatures rather than a comparison of the total number of travelers.

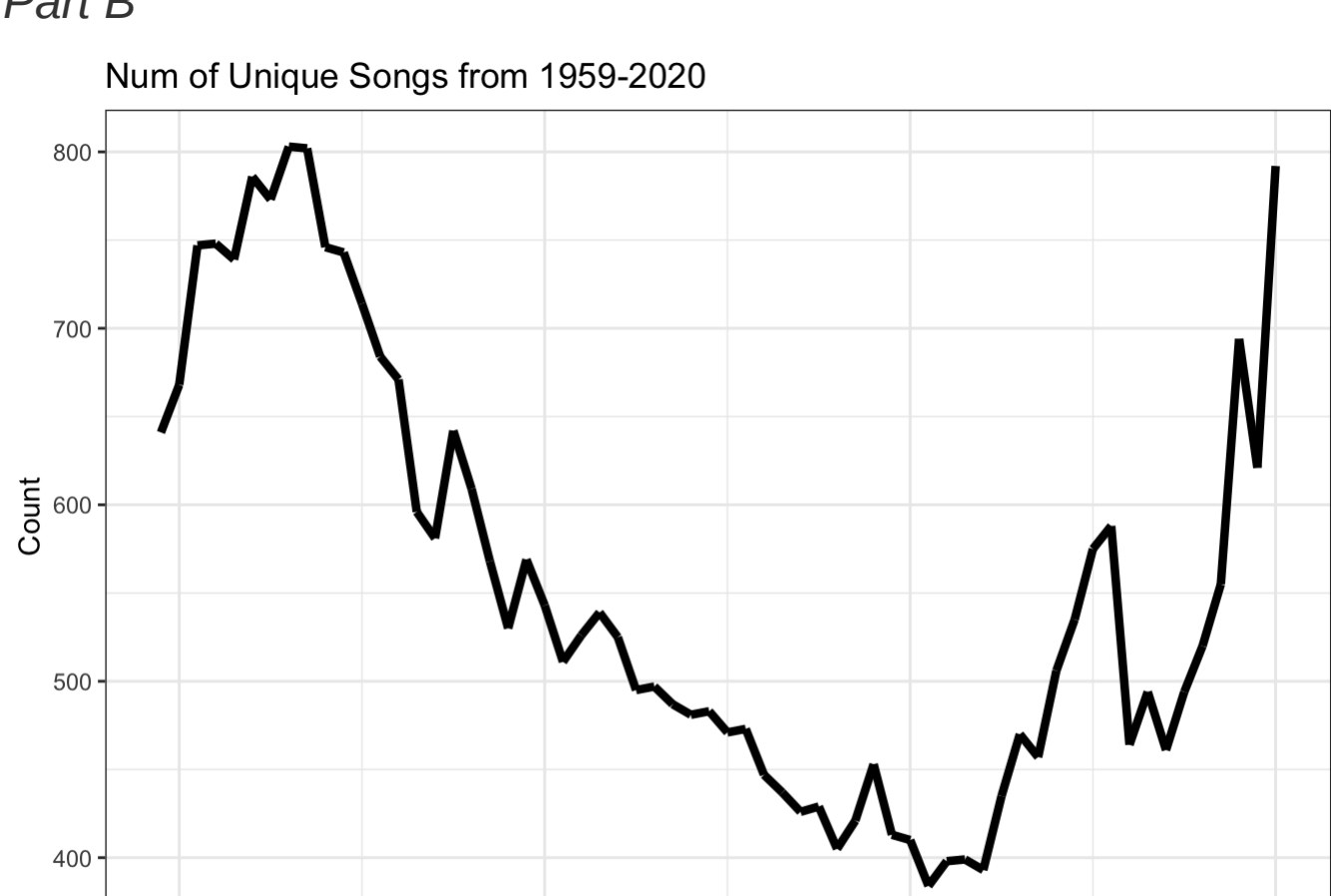
Problem 4: Wrangling the Billboard Top 100

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 above shows the top 10 most popular songs all time since 1958, this is shown by the number of times they appeared on the Billboard Top 10 charts. The most popular song on this table using this measurement is "Radioactive" by Imagine Dragons.

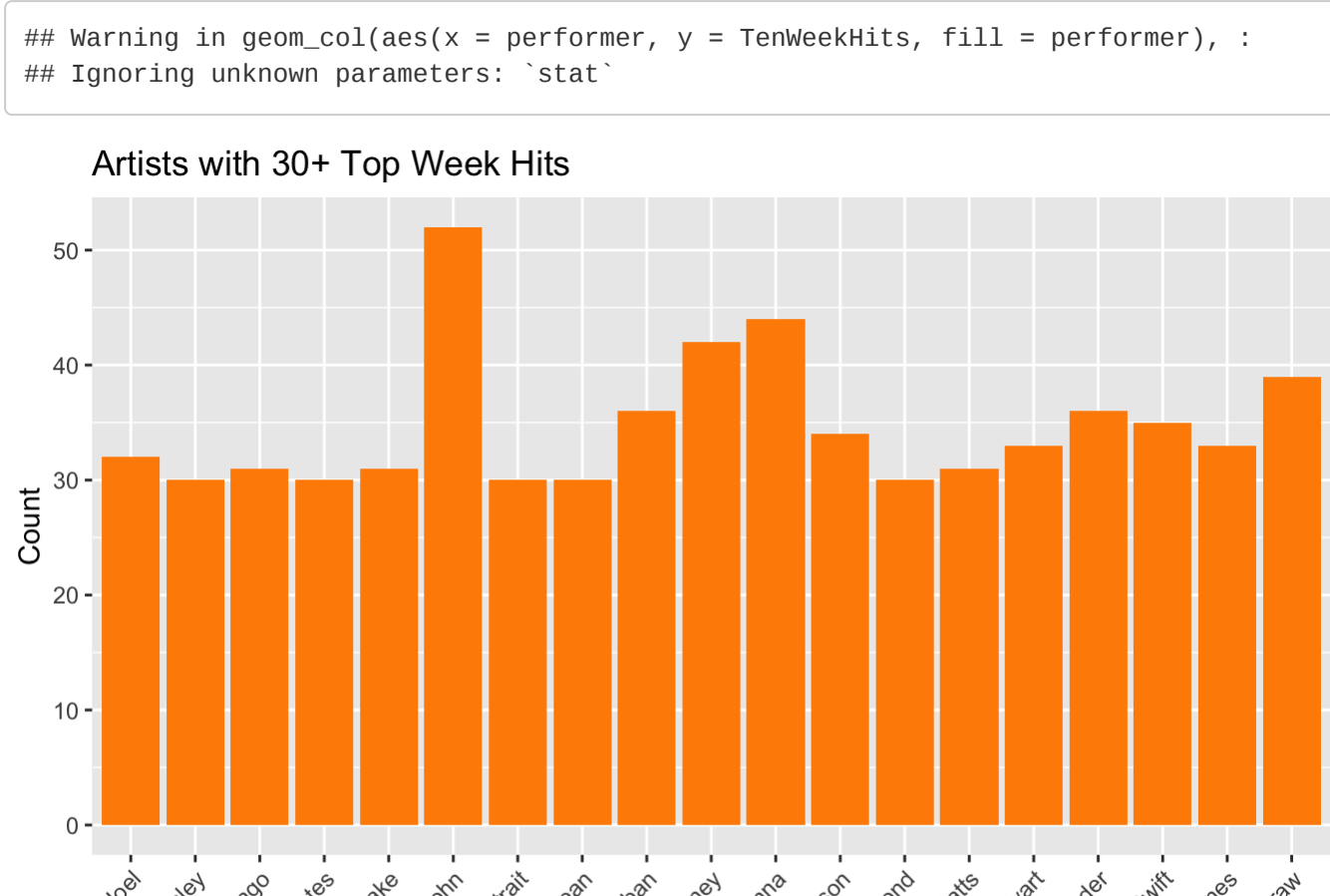
Part B



The line graph above shows the trend of unique songs released between 1958 and 2021 (exclusive of 1958 and 2021). The graph ends approaching a new maximum show last in the 60's. The graph reached it lowest number of unique songs around the 2000's. This came after the max in the 60's meaning we could have the same trend in the future.

Part C

```
## Warning in geom_col(aes(x = performer, y = TenWeekHits, fill = performer), :
## Ignoring unknown parameters: `stat`
```



The plot shows the 19 artists that had over 30 top 100 songs that remained in the Billboard Top 100 for at least 10 weeks. Some popular artists are Drake, Neil Diamond, Taylor Swift and the leader Elton John.