# SDS315_HW2

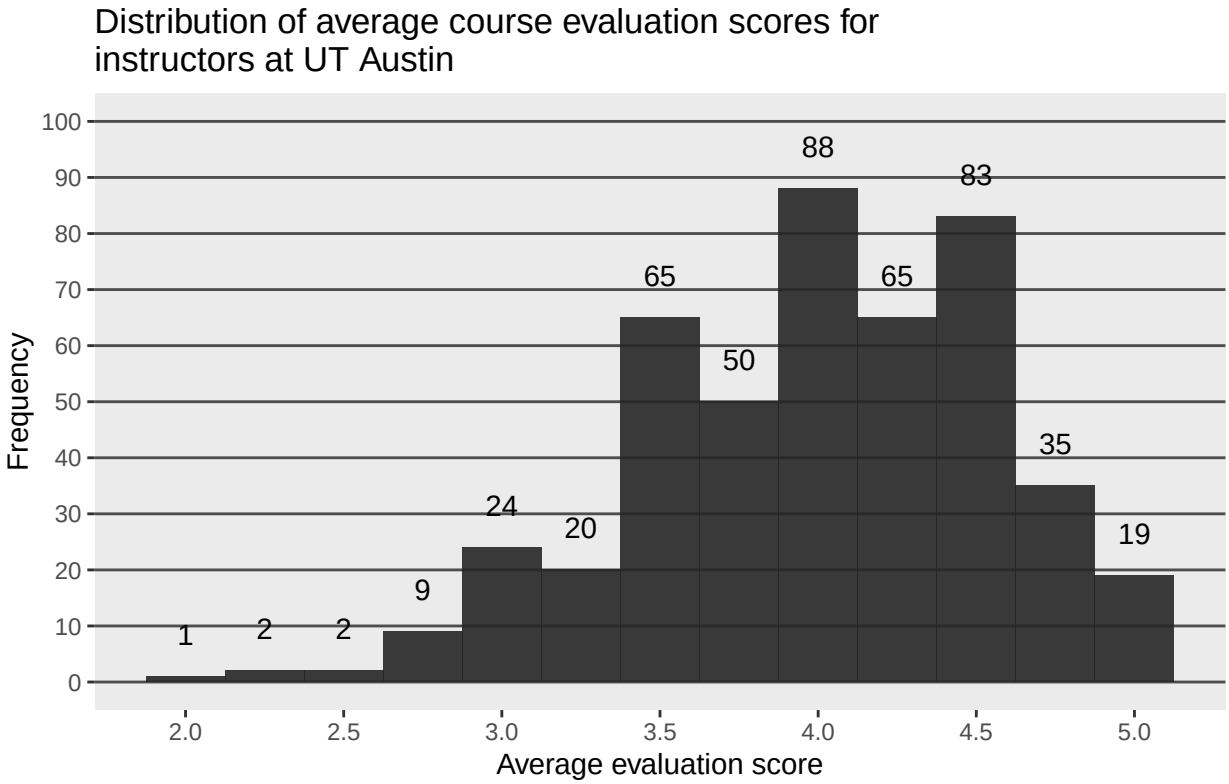
Jaxon Lara

jtl3236

(Github_R_Code)

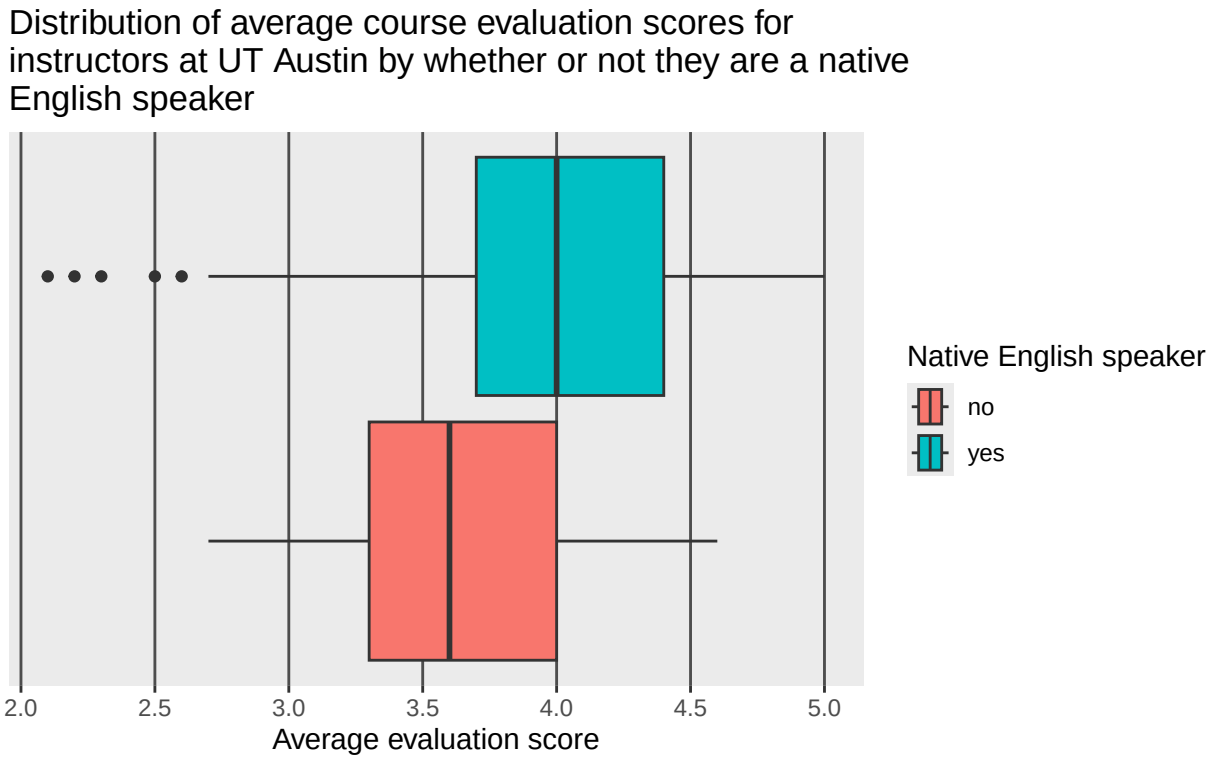
## Problem 1: Beauty, or not, in the classroom

### (A)



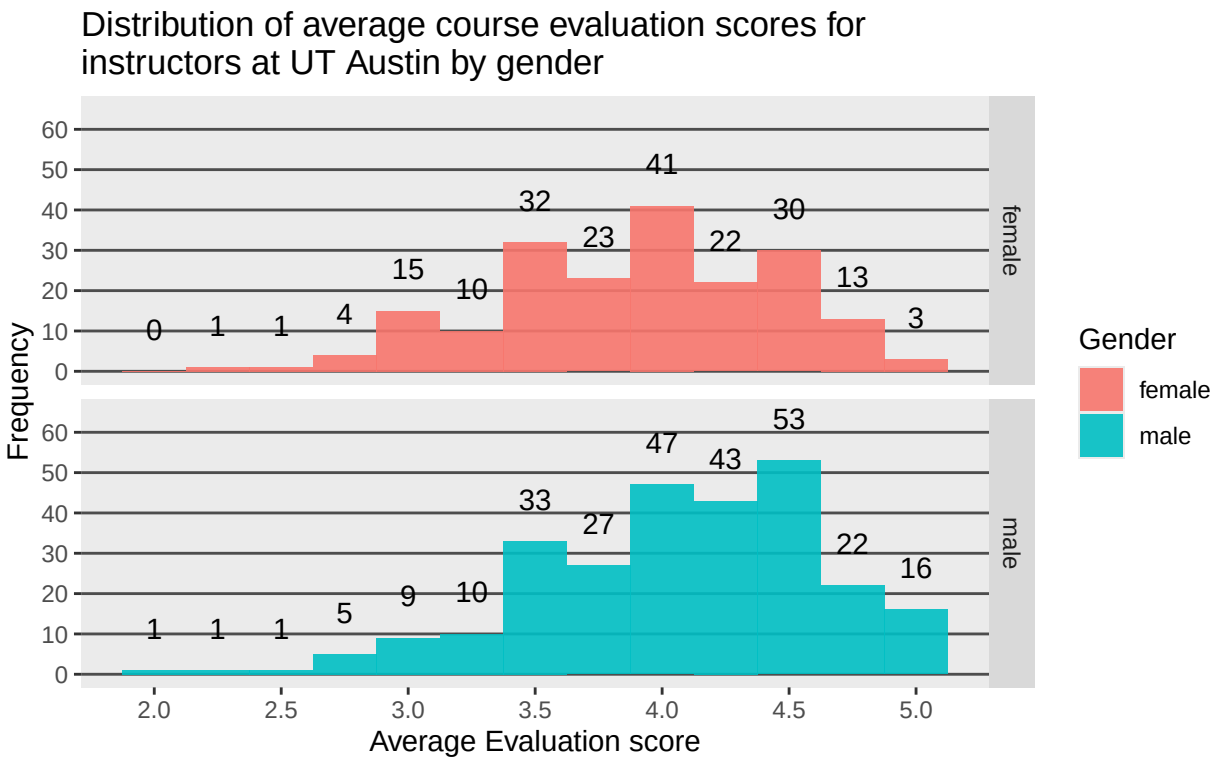

The distribution of average evaluation scores is shown to be slightly skewed right. With the majority of average evaluation scores being between and within the 3.5 and 4.5 score bars.

(B)



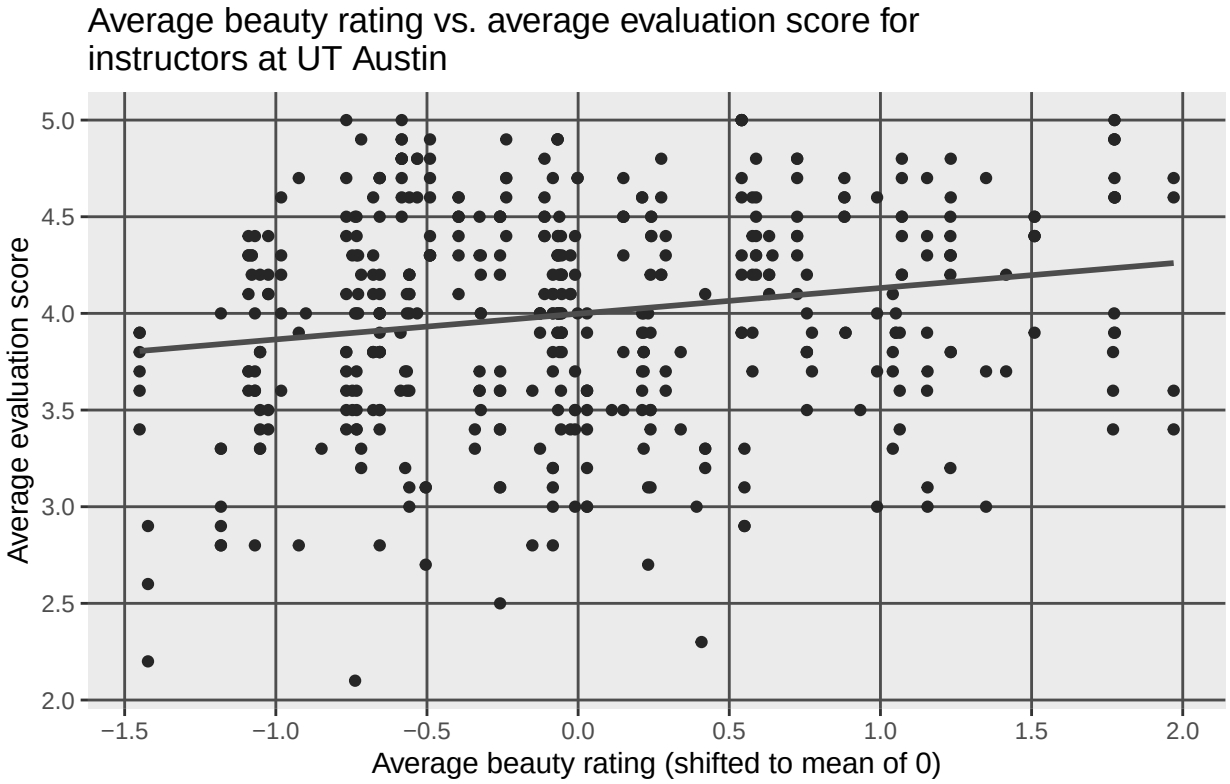

The distribution shows that native English speaking instructors have a higher median average evaluation score closer to 4.0. While the distribution also shows that native English speaking instructors have a wider range of average evaluation scores.

(C)




Both distributions are shown to skew right and have the majority of average evaluation scores between and within the 3.5 and 4.5 score bars. While the male distribution's most frequently occuring score is higher compared to that of the female distribution.

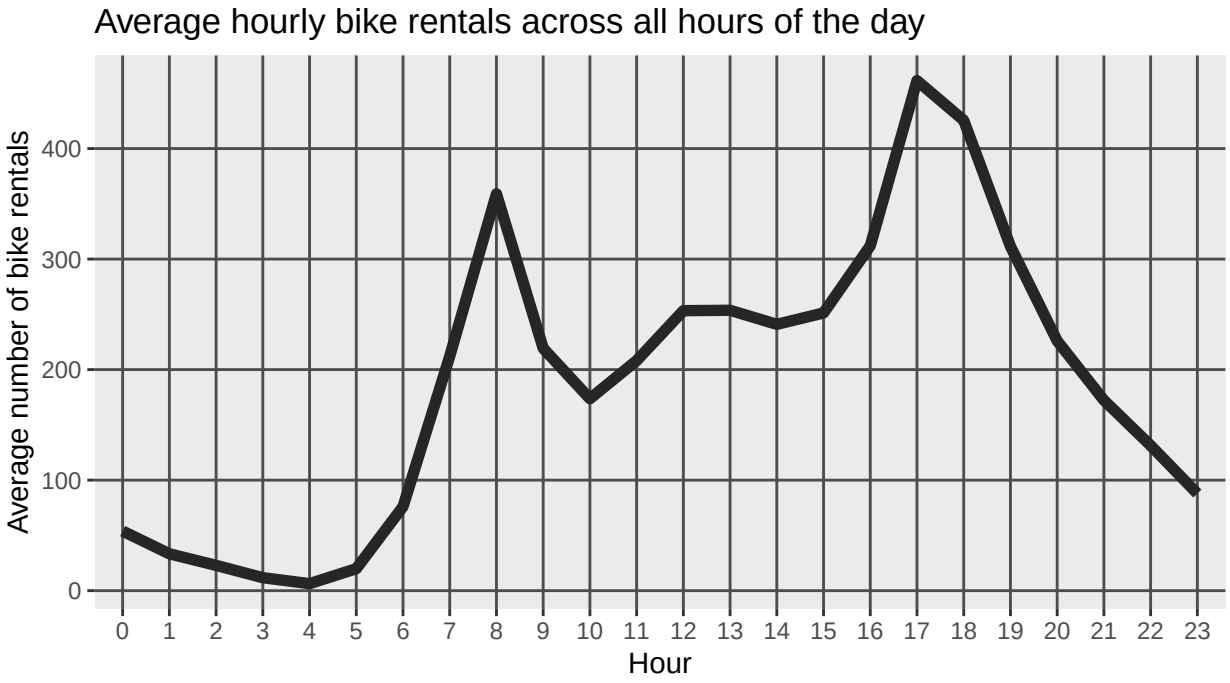
(D)




The distribution shows very little association between an instructor's average beauty rating and their average evaluation score.

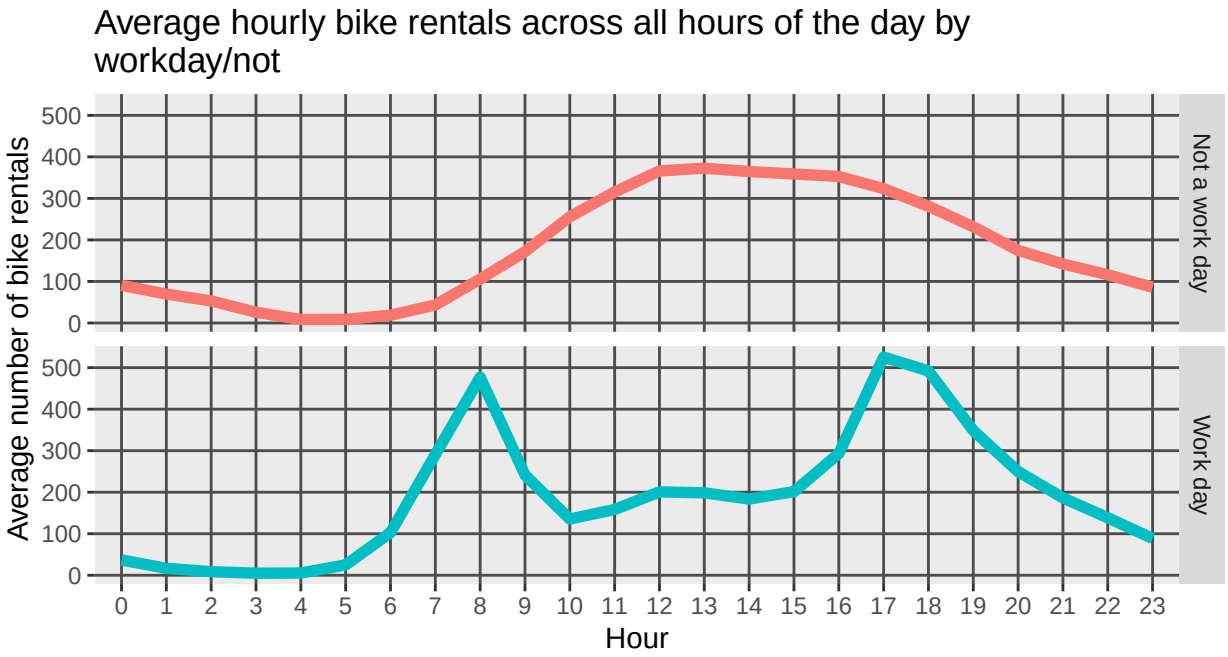
## Problem 2: Bike sharing

### (A)



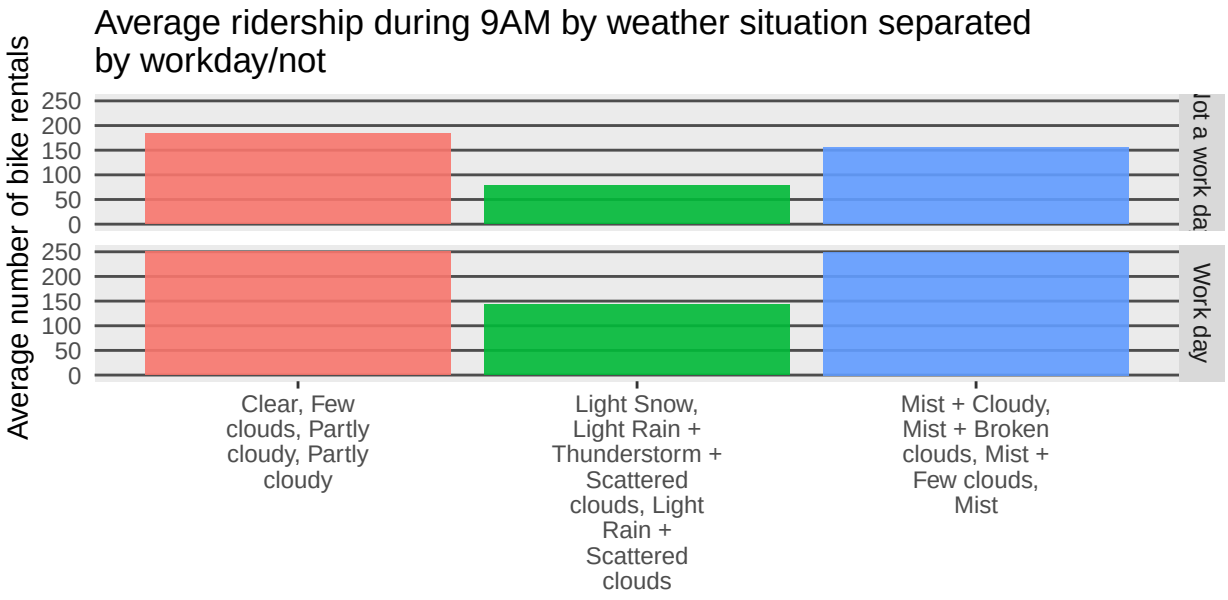

The distribution shows the average number of bike rentals across the 0−23 hours of a day compiled over 2 years of data gathered from Washginton D.C in 2011 and 2012. Over the course of the day two distinct peaks are seen in the plot during the 8th and 17th hour at which the average number of bike rentals reach a relative maximum while the the average number of bike rentals plateaus inbetween the 8th and 17t hour. This shows that there was a distinct average bike rental pattern during a day in Washington d.C from 2011 to 2012.

(B)




The distribution shows the average number of bike rentals across the 0−23 hours of a day compiled over 2 years of data gathered from Washginton D.C in 2011 and 2012 separated by whether it was a work day or not. According to the distributions, there are 2 distinct peaks shown during the 8th and 17th hour in the work day distribution. While the not a work day distribution shows no distinct peaks but rather a hill showing a gradual increase in average bike rentals approaching the 13th hour and decreasing afterwards. This shows the pattern of bike rentals depends heavily on whether or not any given day is a work day.

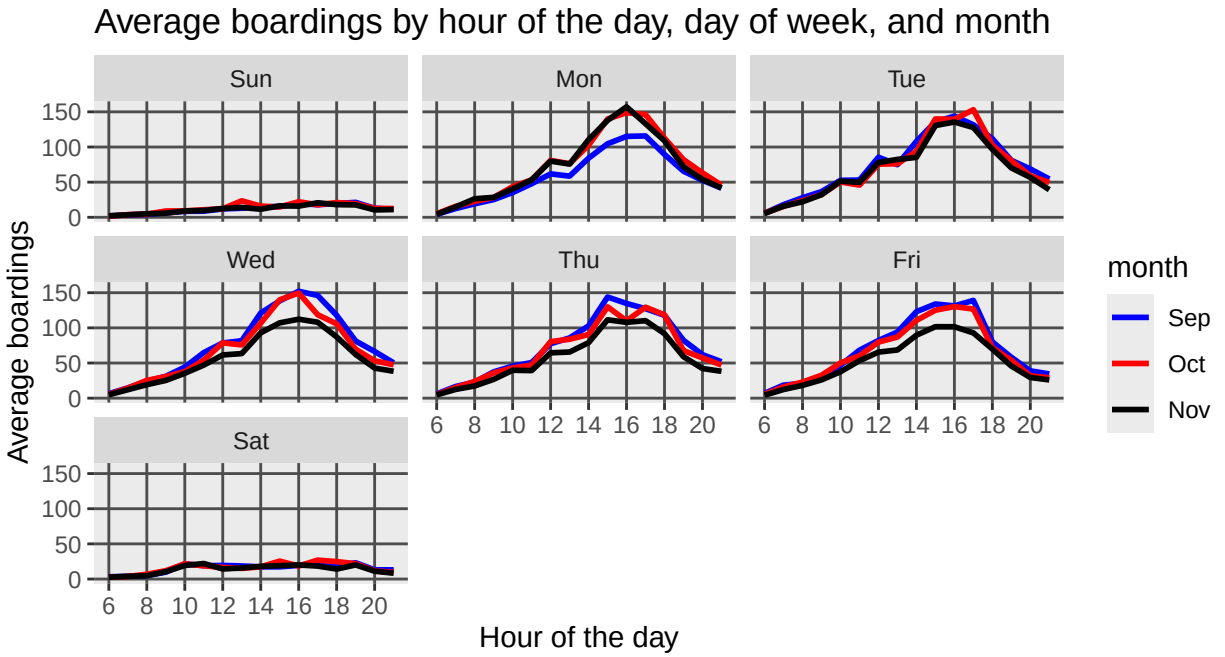
(C)




The distribution shows the average number of bike rentals during 9AM compiled over 2 years of data gathered from Washginton D.C in 2011 and 2012 separated by whether it was a work day or not and grouped by the weather situation. The distribution of both types of day are relatively similar with higher severity weather conditions(green) showing the lowest average bike rentals and the least severe weather conditions showing the highest average bike rentals(red). While the work day distribution shows a general increase in average bike rentals in all weather situation compared to the not a work day distribution. This shows that weather situation also affects the average number of bike rentals along with whether its a work day or not.

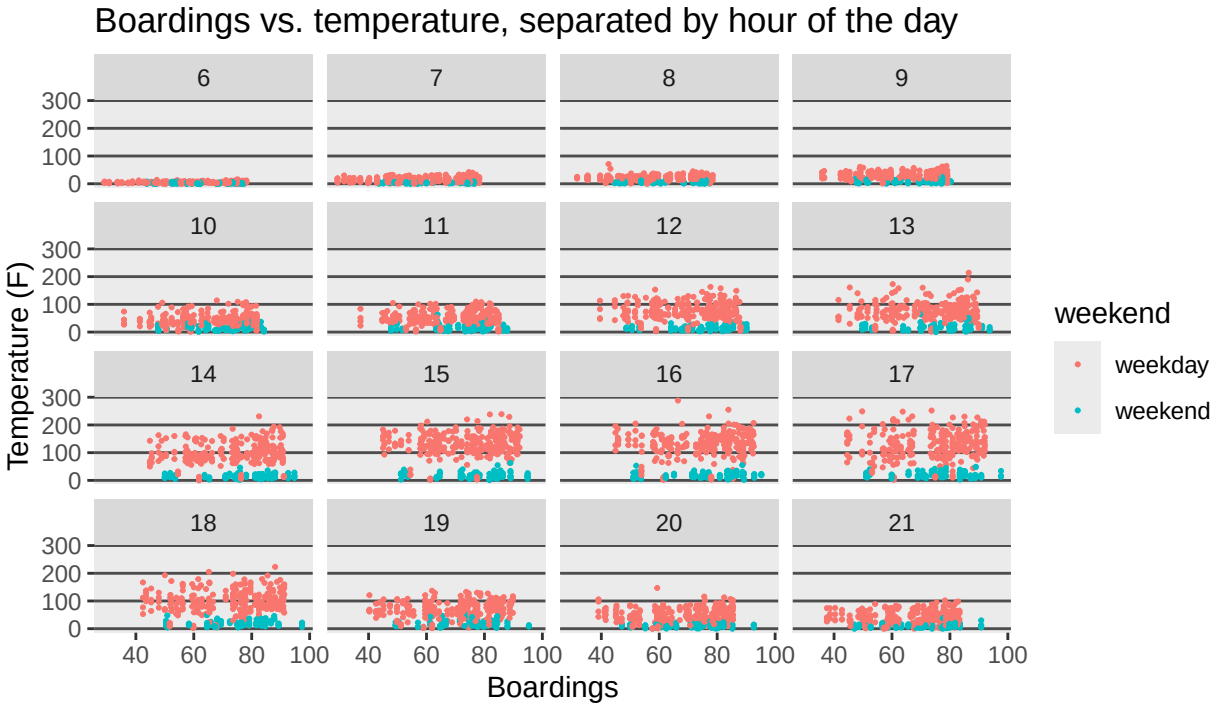
## Problem 3: Capital Metro UT Ridership

### (A)




The distributions show the average number of people who board any Cap Metro Bus on the UT campus by hour of the day separated by day of the week and grouped into the month when the data was recorded. (a) The hour of peak boardings is broadly similar across all days with Monday through Friday showing the most distinct peak boardings during and near the 16th hour while Sunday and Saturday distributions are fairly flat throughout the day. (b) Labor day occurs on a Monday in September which would lead to a smaller average number of boardings since Labor day is a holiday on the UT academic schedule and would mean less activity on or around campus.

(B)




The distributions plot the average number of people who board any Cap Metro Bus on the UT campus vs. the temperature in fahrenheit separated by the hour of the day and grouped by whether or not it was a weekday or on the weekend. (a) no, temperature seems to have no effect on the number of UT students riding the bus because the distribution remains flat as temperature increases instead of a slight skew.

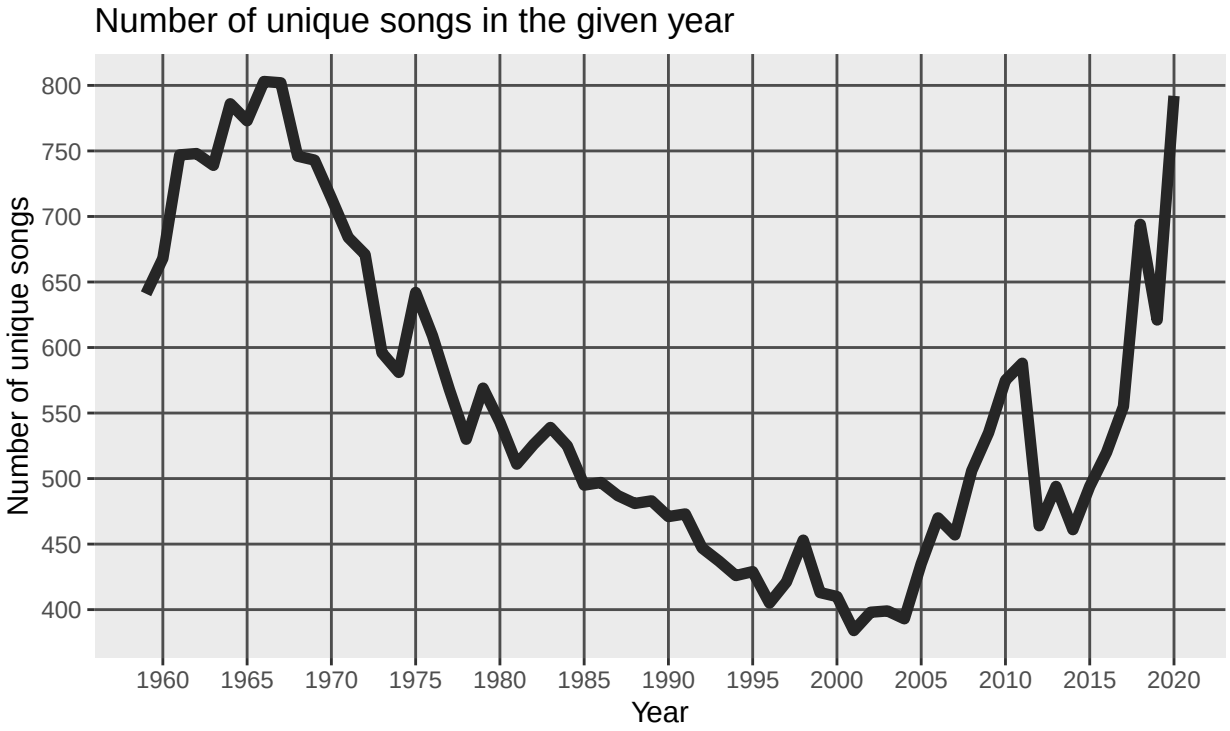
# Problem 4: Wrangling the billboard top 100

## (A)

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

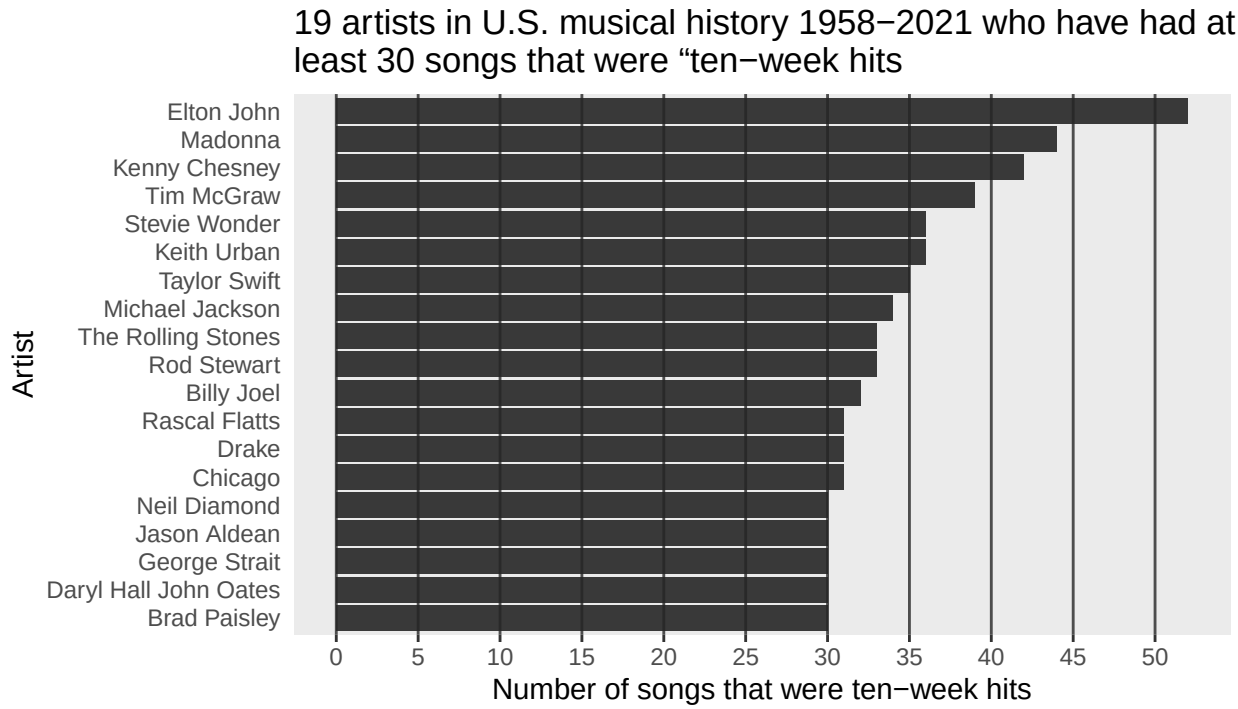
The table shows the top 10 songs measured by the total number of weeks that a song spent on the Billboard Top 100. Radioactive tops the list at a total of 87 weeks on the Billboard Top 100 while Sail and Blinding Lights follow in 2nd and 3rd place respectively.

**(B)**




The distribution shows the number of unique songs that appear in the Billboard Top 100 each year from 1959 to 2020. According to the distribution the number of unique songs peaks around 1966 and begins to rapidly decrease until 2004 when the number of unique songs begins to speak again almost reaching the peak of 1966.

(C)




The distribution shows the 19 artists who had at least 30 songs that each appeared on the Billboard Top 100 for at least 10 weeks and their respective number of songs that were ten−week hits. The distriution is shown to be skewed towards the top 4 artists with the most ten−week hits as there are many more artists with fewer than 37 ten−week hits than those with more.