

Homework 2

Trent Latz

Contents

Problem 1: Beauty, or not, in the classroom	2
Histogram for Overall Course Evaluation Scores	2
Side-by-Side Boxplots by Native English Speaker Status	3
Faceted Histogram by Instructor Gender	4
Scatterplot of Attractiveness vs. Evaluation Scores	5
Problem 2: Bike Sharing	6
Average hourly bike rentals across all hours of the day	6
Faceted line graph of average hourly bike rentals, grouped by working day	7
Average ridership during 9 AM by weather situation, faceted by working day	8
Problem 3: Capital Metro UT Ridership	9
1: Faceted Line Graph: Average Boardings by Hour of Day, Day of Week, and Month	9
2: Faceted Scatter Plot: Boardings vs. Temperature by Hour of Day and Weekend Status	10
Problem 4	10
Table of Top 10 Most Popular Songs	10
Musical Diversity Over Time	11
Ten-Week Hits by Artist	12

Name: Trent Latz UT EID: tj12597 GitHub: GitHub Repo Link

```
knitr::opts_chunk$set(echo = FALSE,
                       eval = TRUE,
                       warning = FALSE,
                       message = FALSE,
                       fig.align = "center",
                       R.options = list(max.print=50))

library(tidyverse)

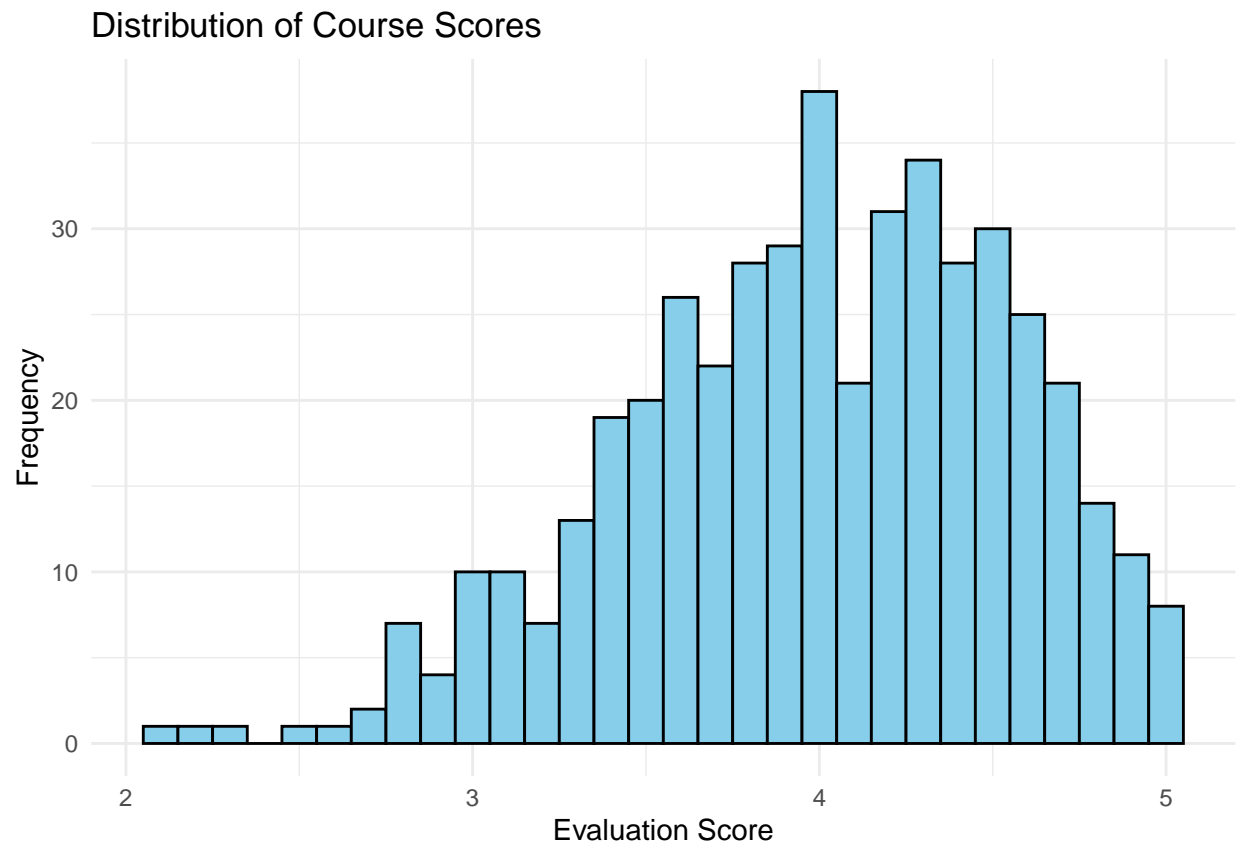
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2    3.5.1      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.1
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

bikeshare <- read.csv("bikeshare.csv")
billboard <- read.csv("billboard.csv")
```

```
metro <- read.csv("capmetro_UT.csv")
profs <- read.csv("profs.csv")
```

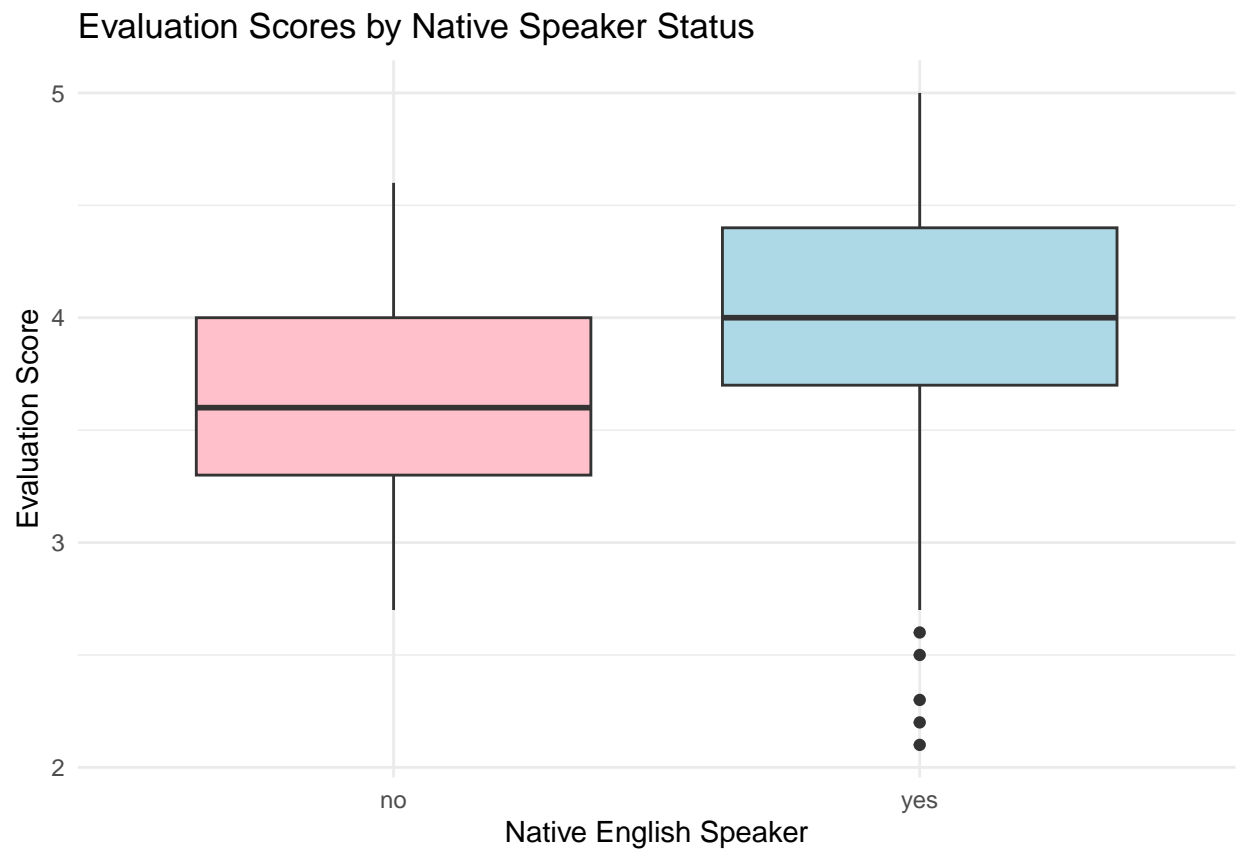
Problem 1: Beauty, or not, in the classroom

Histogram for Overall Course Evaluation Scores



The histogram has roughly a normal distribution centered at about 4. This suggests that the vast majority of students give their courses good scores.

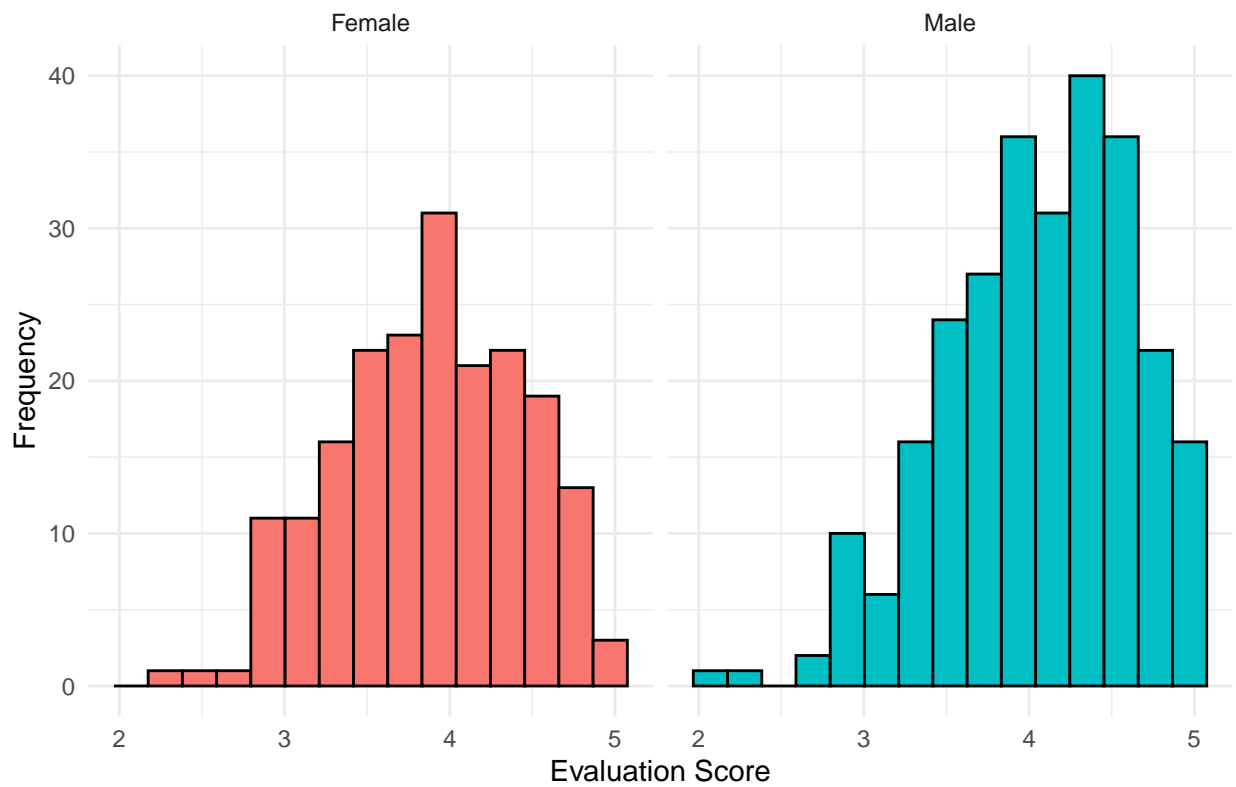
Side-by-Side Boxplots by Native English Speaker Status



The graph above shows that professors who are native English speakers on average get an evaluation about .5 above those who are non-native. This suggests that a foreign accent negatively impacts course evaluation scores.

Faceted Histogram by Instructor Gender

Distribution of Course Scores by Gender



The male professors tend to get higher course evaluation scores than female professors. They have a similar distribution, but the male graph is shifted further right.

Scatterplot of Attractiveness vs. Evaluation Scores

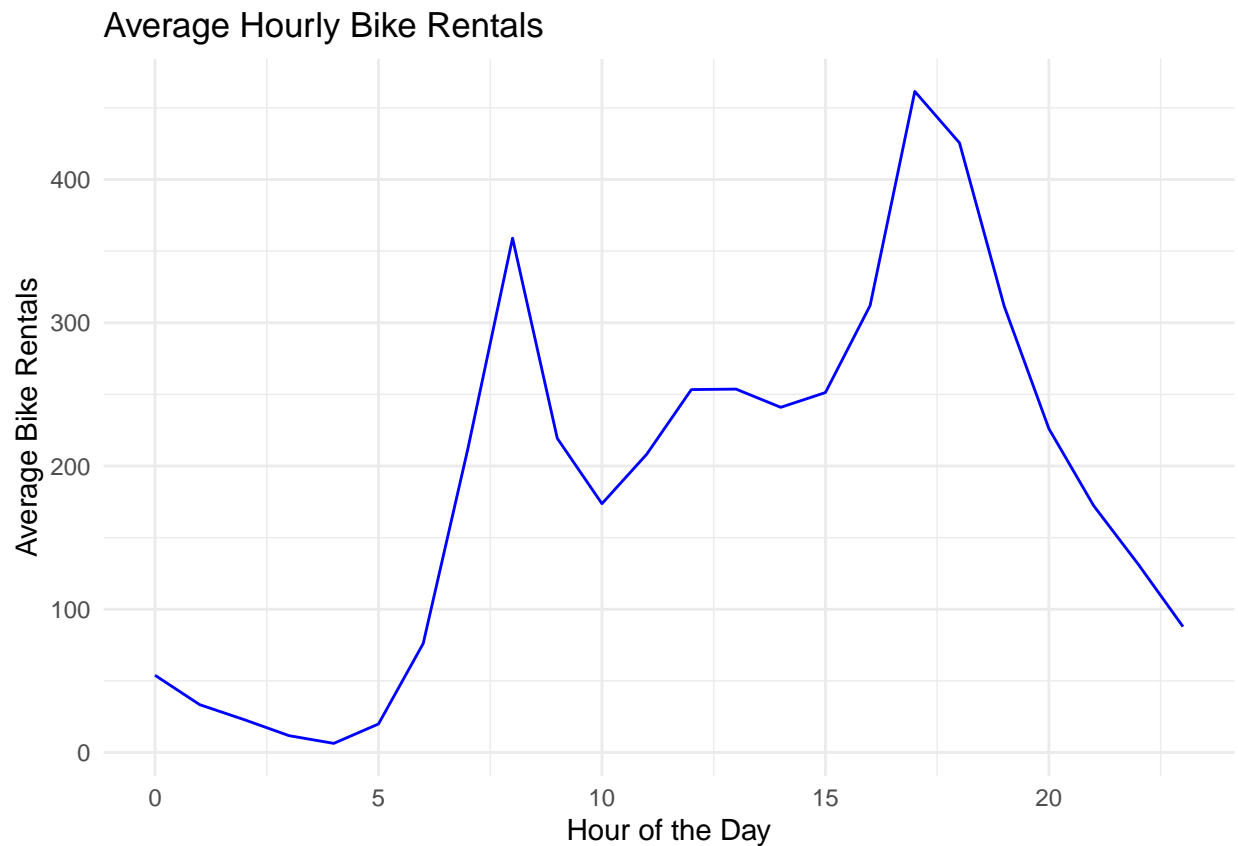


```
## [1] 0.1890391
```

The plot above shows the very small correlation between the professor's attractiveness and the course evaluation score. This suggests that the attractiveness of the professor is a small factor in the student's decision.

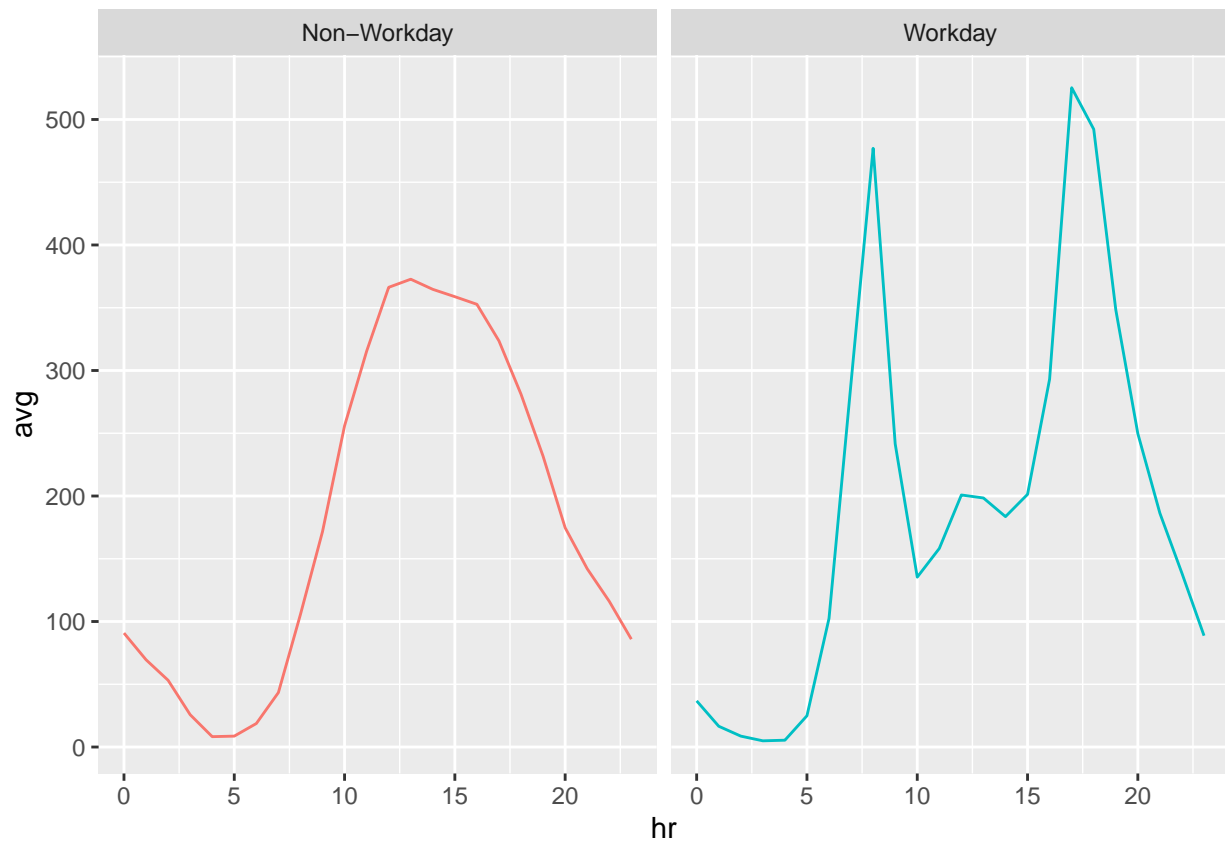
Problem 2: Bike Sharing

Average hourly bike rentals across all hours of the day



This graph shows the average number of bike rentals across all hours of the day, from 0 to 23. The x-axis is the hours of the day, while the y-axis shows the average number of bike rentals (units: total rentals). There are spikes in average bike rentals at about 8 AM and again at about 5 PM. This trend provides insight when bike rental demand is highest, which is useful for managing availability. The data shows that bike rentals follow the typical daily pattern of someone working a 9-5, with peaks when going to or leaving from work.

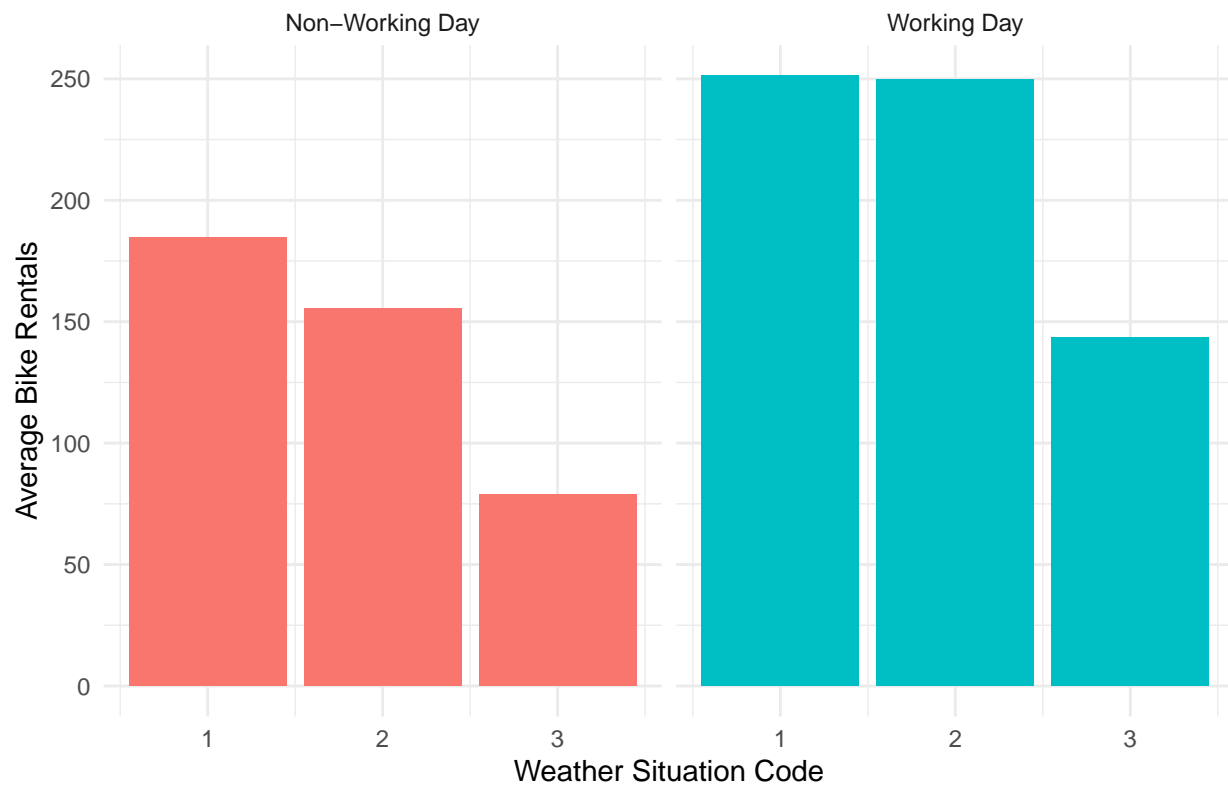
Faceted line graph of average hourly bike rentals, grouped by working day



This faceted line graph illustrates the average bike rentals by hour of the day, separated by where or not it is a working day. The x-axis represents the hours of the day, while the y-axis shows the average number of bike rentals (units: total rentals). These graphs show what was shown in the prior graph, where there are spikes corresponding to the workday. On non-workdays, the max is also around 5 pm. Ridership is significantly higher on working days, especially around 8 AM and 5 PM.

Average ridership during 9 AM by weather situation, faceted by working day

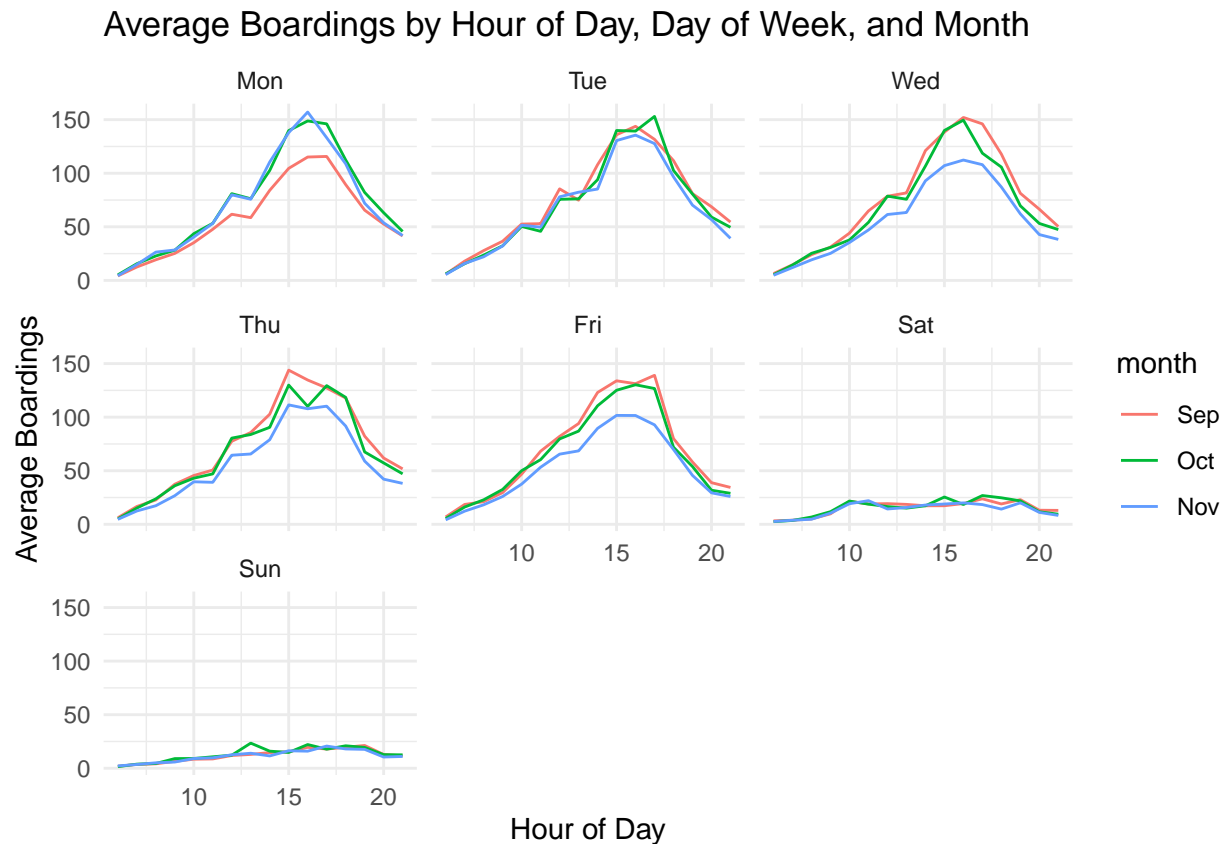
Average Ridership by Weather Situation at 9 AM



This faceted bar plot shows the average number of bike rentals at 9 AM, separated by weather situation codes and whether or not it is a working day. The plot being broken up by the working day variable allows us to compare the patterns during days when people are working versus when they are not. The graphs above show that under the same weather conditions, more bike rentals occur on working days. It also shows that as the weather conditions worsen, the average number of bike rentals decrease.

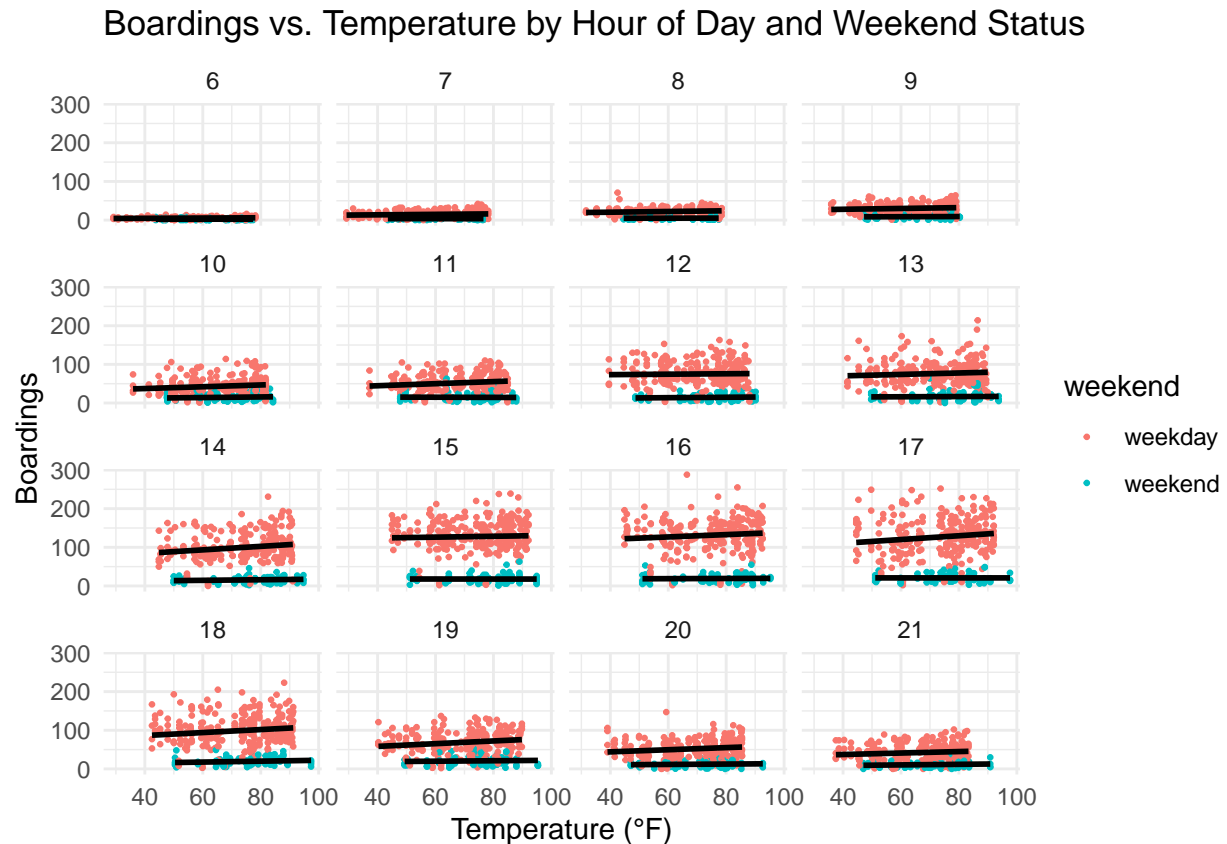
Problem 3: Capital Metro UT Ridership

1: Faceted Line Graph: Average Boardings by Hour of Day, Day of Week, and Month



This faceted line graph shows the average number of boardings at each hour of the day (x-axis) for each day of the week (facets). The lines represent different months (September, October, November) and are distinguished by color. The plot suggests that peak boardings occur around 5 PM, except for Sunday. The 1 PM peak on Sunday could be due to church. The average boardings on Mondays in September appear lower possibly due to reduced demand at the beginning of the semester. Boardings on Wednesdays, Thursdays, and Fridays in November are lower most likely due to Thanksgiving break.

2: Faceted Scatter Plot: Boardings vs. Temperature by Hour of Day and Weekend Status



The scatterplot shows the relationship between temperature (x-axis) and boardings (y-axis) for each hour day, faceted by the hour of the day. Points are colored according to whether it is a weekday or weekend. When weekday status and hour of the day are kept constant, the graphs show that there is extremely low positive correlation between temperature and boardings. I expected the boardings to go up more when temperature increases. The ridership appears to be slightly influenced by temperature.

Problem 4

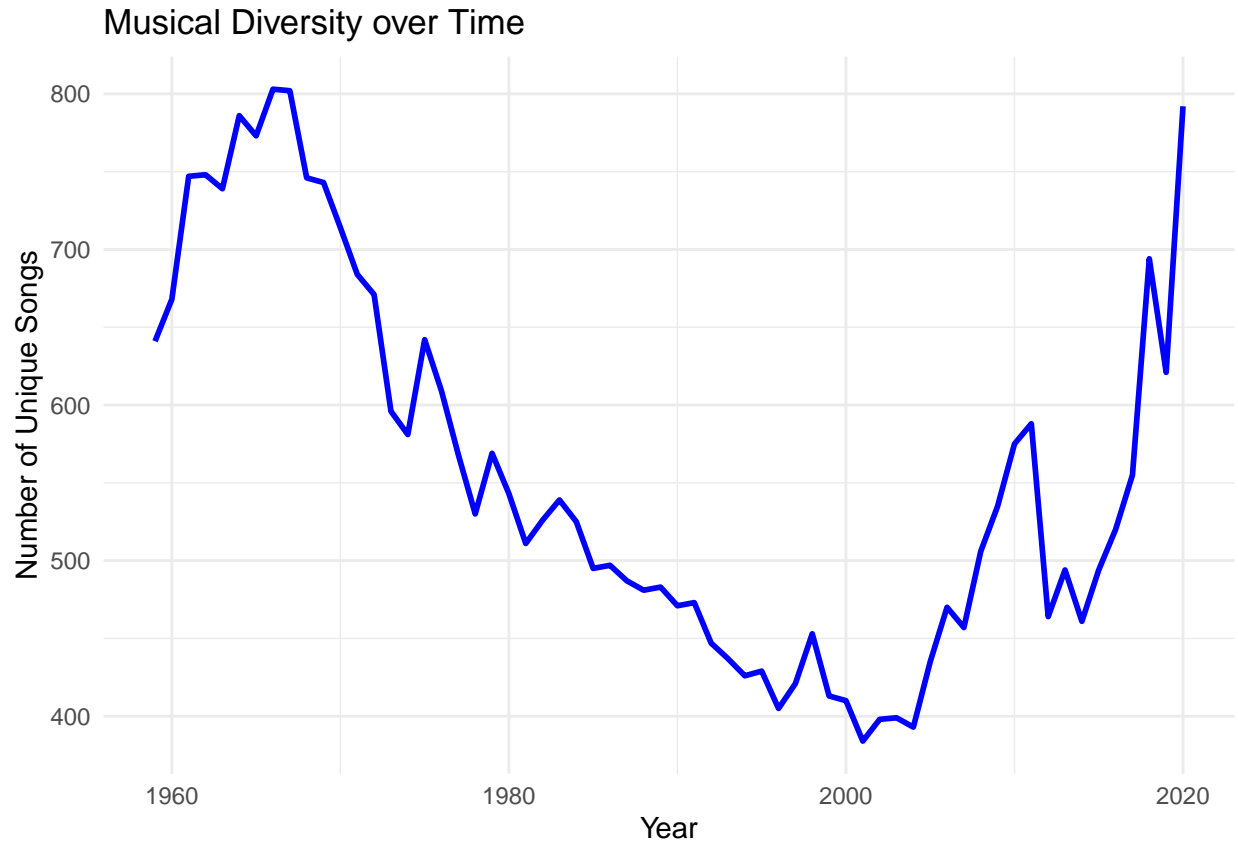
Table of Top 10 Most Popular Songs

```
## # A tibble: 10 x 3
```

##	song	performer	count
##	<chr>	<chr>	<int>
## 1	Radioactive	Imagine Dragons	87
## 2	Sail	AWOLNATION	79
## 3	Blinding Lights	The Weeknd	76
## 4	I'm Yours	Jason Mraz	76
## 5	How Do I Live	LeAnn Rimes	69
## 6	Counting Stars	OneRepublic	68
## 7	Party Rock Anthem	LMFAO Featuring Lauren Bennett & G~	68
## 8	Foolish Games/You Were Meant For Me	Jewel	65
## 9	Rolling In The Deep	Adele	65
## 10	Before He Cheats	Carrie Underwood	64

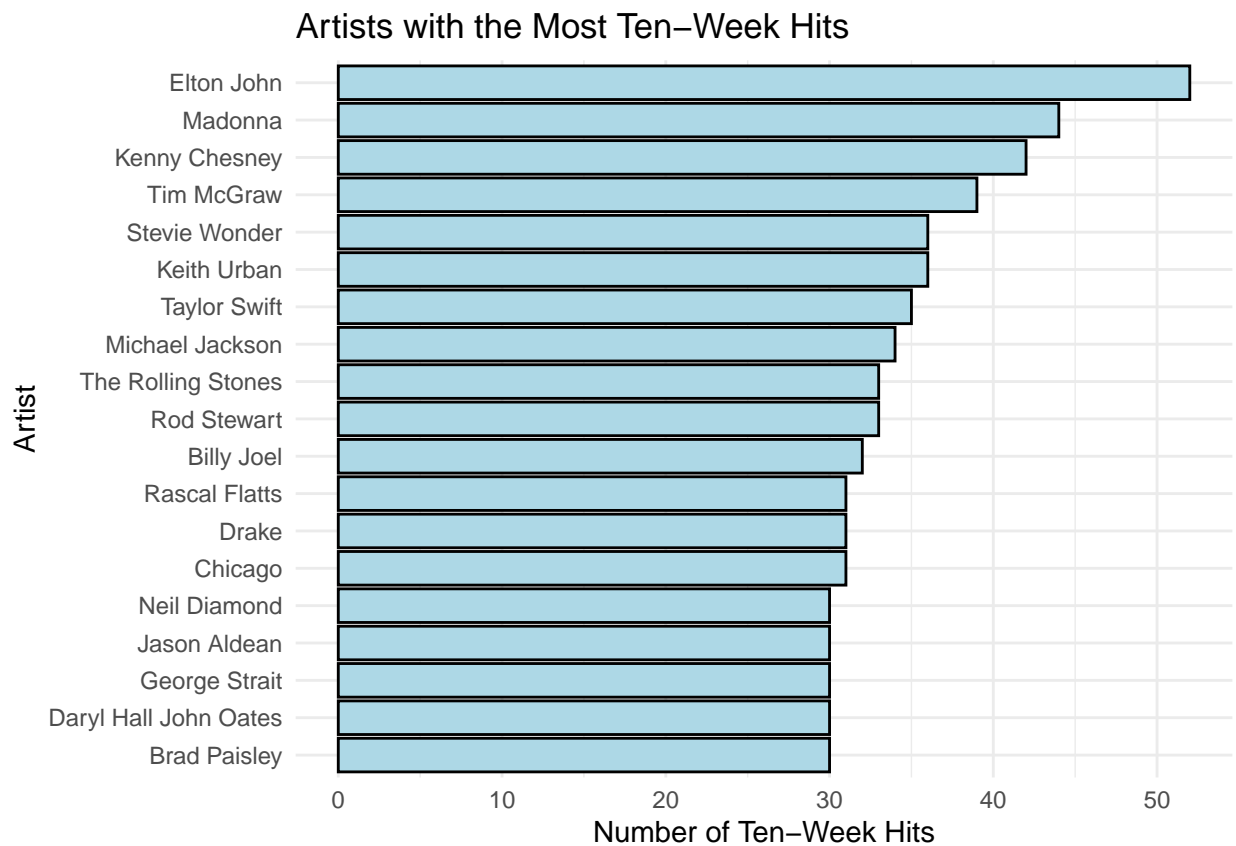
The table shows that the longest a song stayed on the Billboard Top 100. This song was Radioactive by Imagine Dragons. The top tens song ranged from 64 to 87 weeks on the top.

Musical Diversity Over Time



The line graph portrays a max in 1966 followed by 1967. There were more songs charting during those years than any other. This is closely followed by 2020. Interestingly, there was a sharp decline after 2011.

Ten-Week Hits by Artist



This bar plot highlights the 19 artists who have had at least 30 songs appear on the Billboard Top 100 for at least 10 weeks. The x-axis lists the artists, and the y-axis shows the number of ten-week hits. The plot is ordered by the number of hits, emphasizing the most successful artist in terms of time. They range from 30 to 52. Elton John has the longest run.