# The "A" of Statistics III

**Mini Workshop**

Yangsheng Xu

October 15, 2024

## 0.1 Introduction

Please make sure your name shows somewhere on your submission. This is an individual assignment but you are encouraged to work with others or to seek help from them.

## 0.2 The Work

You were given access to the original Rootman data in Part 2 (Individual portion) of the "The A of Consulting". Please use that data or some modified form of it to produce two results (whatever language you wish to use)

1. One data visualization
2. One table of some sort (could be simple counts could be more complex like model output)

These should fit into the context of the problem Rootman is seeking to solve. You could look at the graduate lab example that was shown during the Ruland discussion as an encouragement.

3. Please record yourself (like your UCLA Story video) telling us a story about either your visualization OR your table (what is this result communicating to you and please communicate it to your viewer). Not both, choose one. The story might be really short (e.g., "this is a map of electric vehicle counts by the adult population in Washington state. Vehicle registration is concentrated in the Seattle metro area specifically in Kings County")

## 0.3 What to turn in

1. A single PDF with your name somewhere
2. A video recording (.mp4 or .mov work best)
3. The associated code that produced the PDF. Quarto or LaTeX or RMarkdown is preferred, but it is OK if you used something else, just include it as part of the submission.
4. The PDF should contain the visualization and the table and the story you want to tell using one of them (writing things out will be helpful when having to talk about the result on video).
5. This is due before the next Tuesday meeting when we will see the team's answer.

## 0.4 Reminders

1. It's the story that is most important, the graphic and the table only lend support.
2. Your PDF does not need to be very long, a page or two is about all you need.
3. Your recording does not need to be very long, less than minute, but no penalties if it runs long.
4. It's OK to turn this in late, just upload it as a comment but remember, once I finish grading, it's a zero because reopening a completed assessment is a headache…

```
library(ggplot2)
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
library(readxl)
library(readr)

lighting_data <- read.delim("lighting_tall.tsv", sep="\t")

summary_table <- lighting_data %>%
```

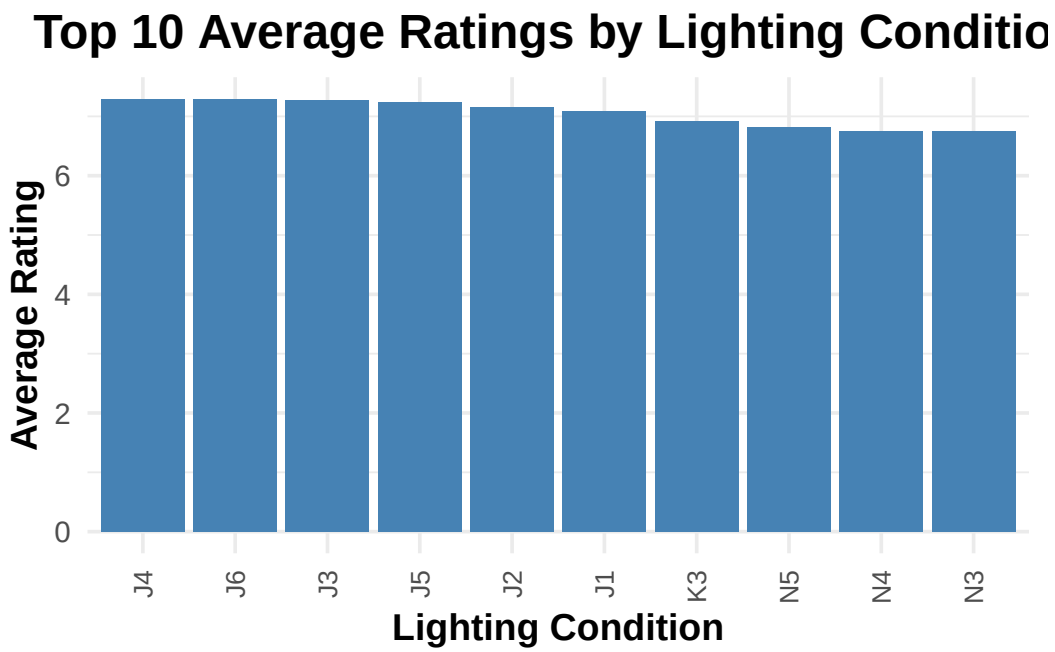
```
  group_by(subjectLighting) %>%
  summarise(average_rating = mean(rating),
            count = n()) %>%
  arrange(desc(average_rating))

summary_table
```

```
# A tibble: 60 x 3
   subjectLighting average_rating count
   <chr>                    <dbl> <int>
 1 J4                        7.30   200
 2 J6                        7.29   200
 3 J3                        7.28   200
 4 J5                        7.24   200
 5 J2                        7.15   200
 6 J1                        7.08   200
 7 K3                        6.92   200
 8 N5                        6.82   200
 9 N4                        6.75   200
10 N3                        6.74   200
# i 50 more rows
```

```
top_summary_table <- summary_table[1:10, ]

ggplot(top_summary_table, aes(x = reorder(subjectLighting, -average_rating), y = average_rati
  geom_bar(stat = "identity", fill = "steelblue") +
  labs(title = "Top 10 Average Ratings by Lighting Condition",
       x = "Lighting Condition",
       y = "Average Rating") +
  theme_minimal(base_size = 14) + # Set a smaller base size
  theme(
    axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5, size = 10), # Rotate label
    plot.title = element_text(hjust = 0.5, size = 18, face = "bold"), # Center and bold title
    axis.title.x = element_text(face = "bold"), # Bold x-axis title
    axis.title.y = element_text(face = "bold")  # Bold y-axis title
  )
```


Top 10 Average Ratings by Lighting Conditio
Average Rating
0
2
4
6
J4
J6
J3
J5
J2
J1
K3
N5
N4
N3
Lighting Condition