

Homework 04

For this homework, create an **R Markdown** (e.g., `.Rmd`) file and save it to your `/report` project directory on your computer with the name and a prefix `hw05-Lastname-FirstInitial`. Full name example: `hw04cookg.Rmd`. If you collaborate, add all names in the `YAML` code of your markdown file.

Relevant Modules:

- project management
- data frame manipulation and wrangling
- data subsets and summaries
- the grammar of graphics
- spatial positioning and adjustment
- visualizing amounts
- color scales and palettes

This exercise involves inspecting a plot, wrangling data to prepare for plotting, creating a data subset, selecting geoms, adding plot layers, and applying colors to specific data sub groups.

Problems:

1. Project Progress Updates and Git

We have reached the point in the semester by which everyone is familiar with `R`, `{dplyr}` functions to wrangle data, `{ggplot2}` functions for creating plots, and some `Git` commands for staging, committing, and pushing to a remote repository. Each homework from this point will include weekly `Git` commands (e.g., `add`, `commit`, and `push`) to move your files from a local repository to the remote on `GitHub`. Although you have practiced this process for various exercises, making this process routine will strengthen the behavior and increase your confidence.

The project mid-term presentation is less than three weeks away. General data cleaning including naming or renaming variables, mutating relevant variables, converting character variables to numeric, removing duplicate rows, and combining/saving a cleaned `.Rds` file that preserves your vector data types should be mostly complete. Modifications here and there related to items overlooked, creation of new data subset files for specific plots, or additional project goals is expected.

1. stage your project related files that you have been working on;
2. commit them individually with specific commit messages;
3. push all the commits (with individual pushes after each add or with a single push after all commits)
4. evidence of progress will be considered

2. Replicating a Visualization

Getting the Data:

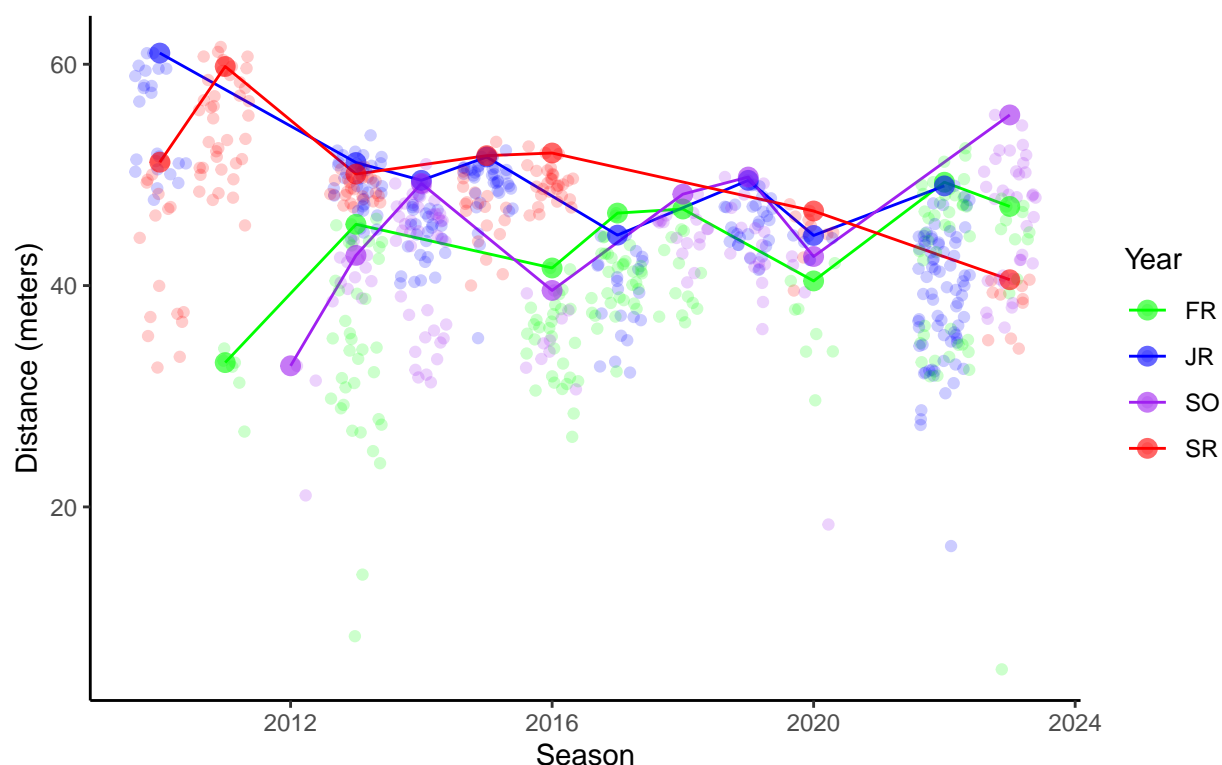
```
readr::read_csv(  
  "https://raw.githubusercontent.com/slicesofdata/dataviz24/main/data/tfrrs/ht-cleaned.csv")
```

The Plot and Data:

Feel free to collaborate. Consider a description for this plot is something like *The longest distance thrown at SCIAC Championship is about the best performance for the season overall*. Look at the data frame and study the plot to determine how you might go about replicating it. Replicate this plot as best as you are able.

- You will notice that the legend keys are in alphabetical order, which seems odd for a plot but that's something we will address in later topics. However, you will want to ensure that the class rank and the colors would remain consistent, for example, if a plot did not include a sub group, like seniors.
- The colors are selected not for aesthetics but to make this easier for you to replicate.

The Colors:



Finally, knit an `html` file and upload to: <https://ln5.sync.com/dl/a038628f0/wwfifjxk-f7rfshin-rkedi3y8-77f9zaii>