Homework 05

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: `hw05cookg.Rmd`. If you collaborate, add all names in the YAML code of your markdown file.

Relevant Modules:

- project management
- data frame manipulation and wrangling
- the grammar of graphics
- coordinates, axes, and position scales
- statistical transformations

This homework is an exercise designed to challenge you to fix x and y axis scales of plots. The plot replication involves inspecting a plot, wrangling data and creating a data subset to prepare the data for plotting, 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 file, 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 you 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

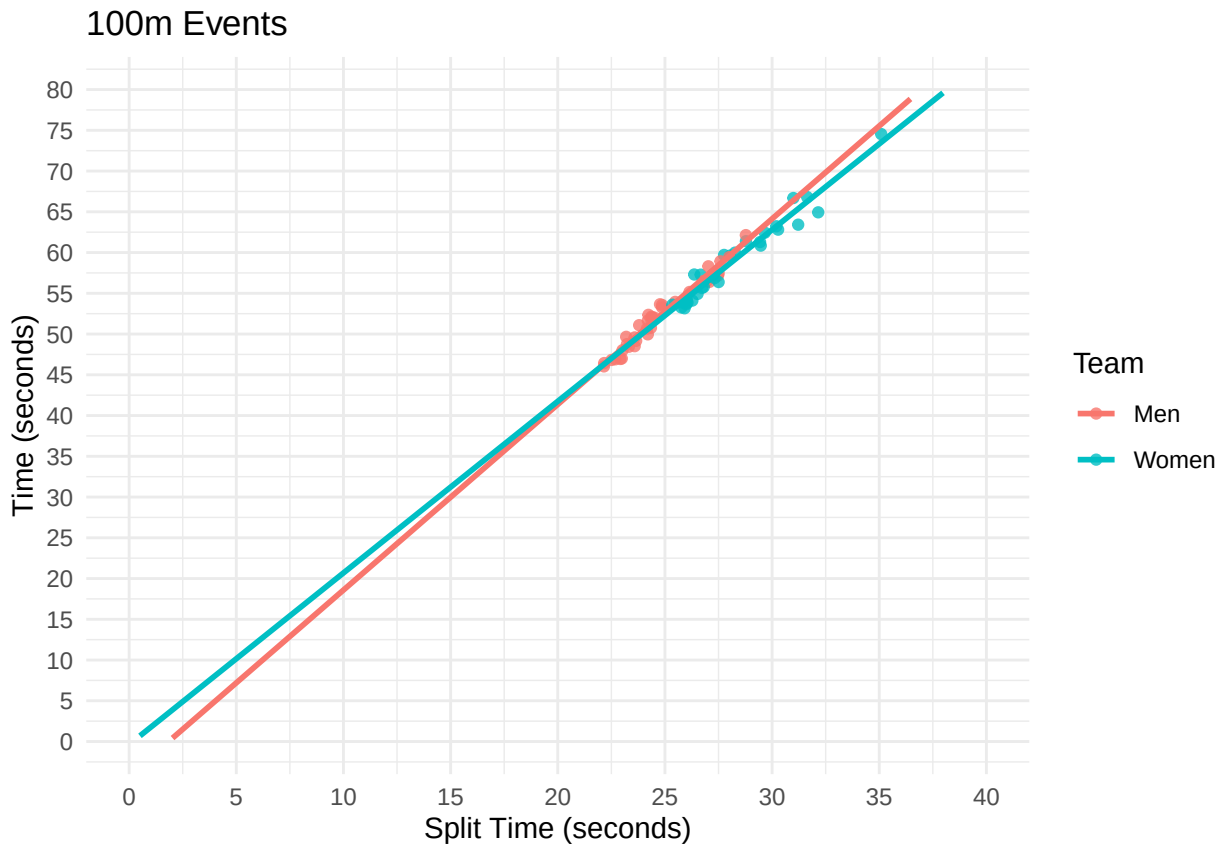
2. Replicating a Visualization

Getting the Data:

```
readr::read_csv(
  "https://raw.githubusercontent.com/slicesofdata/dataviz24/main/data/processed/
  cleaned-2023-cms-invite.csv")
```

The Plot

Feel free to collaborate. Consider the following description of the data to help. *For all events in the season, the split times and event completion times appear linearly related both both male and female swimmers.* 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. Pay attention to issues related to recent topics.



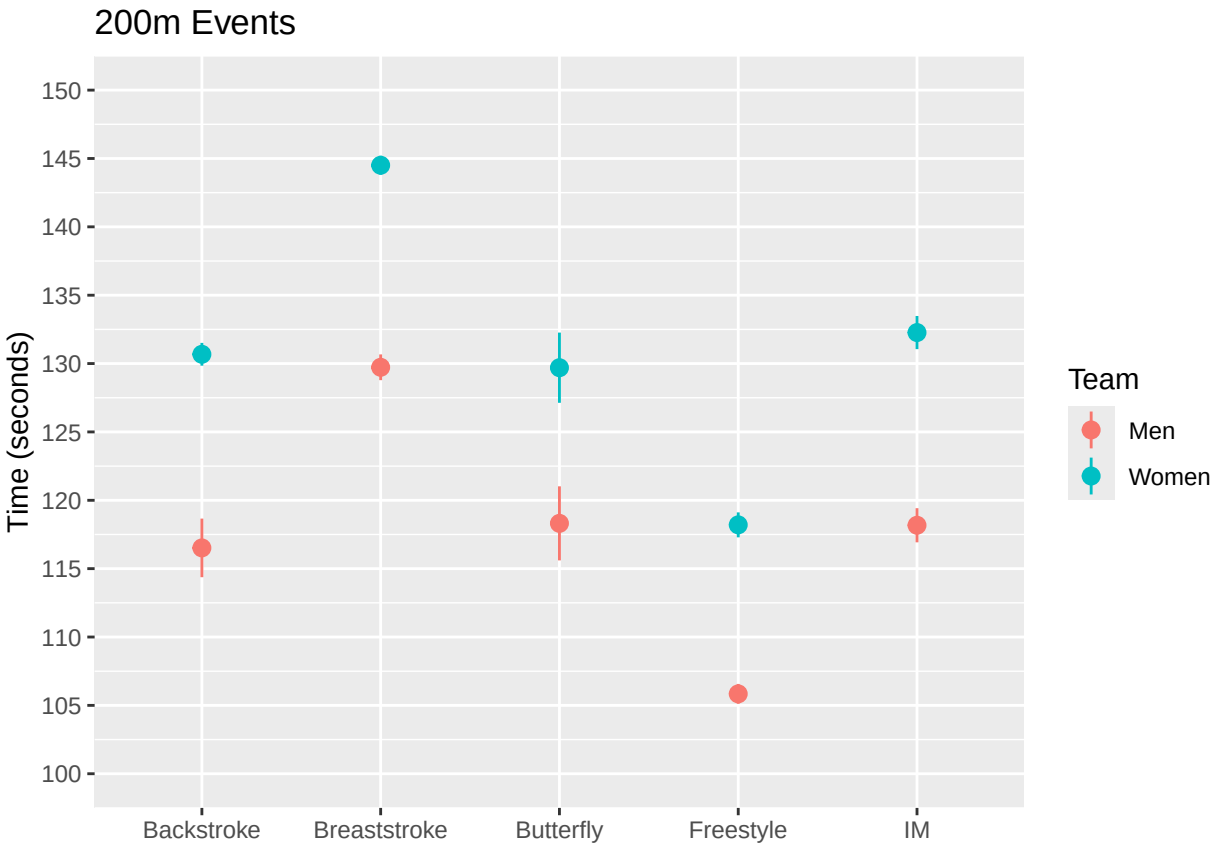
After replicating the plot, add a `coord_equal()` to make the unit change on y be equal to that on x. Does this change how you perceive the data?

BONUS

3. Replicating a Visualization

The Plot

Consider the following very general description of the data to help. *For all events in the season, the mean and standard deviations reveal more variability in butterfly event times for both men and women.* 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.



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