Project Quick Start Guide

Agenda

1. Data Sources
2. Data Wrangling
3. Example
4. Project Q&A

Data Sources

- CK is going to give you a list of reputable sources

- Depending on your area of study format will vary

- Some will be in easily workable single file, others full database

- **TIP: Don't go straight for the easiest format, CK wants some real world data wrangling experience so data sets premade for R analysis are not ideal**

Data Wrangling

- Subset data to remove extra columns
- Merge any separate tables into one big table to work with
- Adjust units and formats
- Clean NAs
- Check for and remove duplicates
- Add new columns as necessary

Resources

- F1 Database
  - http://ergast.com/mrd/
- Data Manipulation Tutorial
  - https://www.r-bloggers.com/2020/08/data-manipulation-in-r-using-data-frames-an-extensive-article-of-basics/
- Github Tutorial
  - https://www.youtube.com/watch?v=iv8rSLsi1xo