Possible Project: Runs and Wins of Baseball
UConn–PCS: Data Science, 2022

The percentage of wins obtained by a team over the course of a season is strongly related with the number of runs it scores and allows. The relationship between runs and wins is explored in Chapter 4 of the book by Marchi, Albert, and Baumer (2019). The source code of the book is at `https://github.com/bayesball/bayesball.github.io`. A workshop on baseball analytics at the last two UConn Sports Analytics Symposiums led by Dr. Zhe Wang included this as an example. The training materials of this workshop can be accessed at `https://statds.org/events/ucsas2020/workshops.html#baseball`. In particular, after downloading it, the code for this example is `Runs_and_Wins.R` under the `Code` folder.

1. Run the example code on your own computer line by line to reproduce the analysis.
2. Explain what Pythagorean formula is and how well it fits baseball.
3. Follow your interests to further explore the data and report your findings. Examples topics are: relationship between win percentage and run differential across decades; Pythagorean residuals for stronger and weaker teams. Be creative and don't be limited by the examples.