Shrinking rfield user notes

The data files are saved as eitherh .csv or .xlsx files and should be modified appropriately for the statistical program to be used (.dta for Stata).

The analysis script is a Stata .do file but is annotated sufficiently to be translated into another language. The script imports and merges the various data files as necessary to reproduce the analysis featured in the bbcardstats.com blog post. It generates certain figures featured in the blog posts and exports to excel files data necessary to construct the remaining ones.

Note that multi-season models use standardized versions of runs allowed per game, the outcome variable, and of rfield and FIP, the predictors. The use of standardized variables is necessary to remove the bias associated with extraneous variance in the means and standard errors of the model variables across different seasons. For discussion, see Schell, M. J., Baseball's All-Time Best Hitters: How Statistics Can Level the Playing Field (Princeton University Press 2013); Schell, M. J., Baseball's All-Time Best Sluggers: Adjusted Batting Performance from Strikeouts to Home Runs (Princeton University Press 2016). The single season models use the untransformed values for runs allowed, rfield, and FIP.