Data Cleaning

Tate Huffman

1/31/2022

Overview

This is part of a series of RMarkdown files that will break down the code contained in the document at the end of the fall. This file in particular will redo the code that constructs the variables needed to run the O'Flaherty and Siow model, altering some components that may have been inaccurate in the first pass at this model.

Altering the Variables

The first code chunk adds the necessary variables for signal, exits, promotions, etc. for batters, and the second does the same for pitchers. Key changes in this version of the code include limiting the player pool to those who debuted from 1999 to 2009; altered promotion code to account for second-year promotions; playing time filters; and percentile-based performance standards.