Velocity user notes

The data, pitcher_game_logs_2002_2026.csv, consist of FanGraphs pitcher game logs for 2008-2026. The relevant columns are FBv for FanGraphs's FBv game-average fastball velocity; TBF for total batters faced; and SO for strikeouts recorded.

The principal analysis script velocity_vs_hitter_propensity_python_code.txt. It consists of a series of discrete "Jupyter notebook" cells. The cells contain introductory annotations for ease of interpretation. Cell 0, which removes position player pitchers, should be run first, and Cell 1, which runs the binomial logit model, thereafter. At that point, cells that reproduce the analysis and Figures reported in the blog post can be run.

Another analysis script, fbv_pre_2017_adjust.txt, performs the operations used to determine the +0.5 MPH adjustment made for pre-2017 FBv speeds to reflect the change in velocity-measurement point adopted by MLB in 2017.

Consult the blog post "method notes" for additional information.