

Statcast Data Dictionary

Jake Coleman

August 19, 2016

This dataset is pitch-level data for all Major League Baseball teams in the first week of the 2016 season, recorded by Statcast tracking technology. Using high-resolution cameras and radar equipment, Statcast can calculate variables describing the trajectory of the pitch (e.g. speed, movement, pitch type), as well as batted-ball information (e.g. distance/angle/speed the ball was hit). The data comes from baseballsavant.com, which nightly downloads and organizes Statcast data from MLB Advanced Media, L.P. This dataset was put together by a search performed on baseballsavant.com for all pitches thrown in the first week of the 2016 season, and then cleaned and trimmed for ease of exploration and analysis. The dataset contains 9159 rows and 19 columns.

Variable List

- **pitcher**: Pitcher name
- **batter_ID**: Unique batter ID
- **pitch_type**: Type of pitch thrown. “FF” denotes fastball, “CU” denotes curveball, “SL” denotes slider, “CH” denotes changeup
- **description**: Description of pitch outcome
- **ab_outcome**: Description of eventual at-bat outcome
- **pitcher_hand**: Handedness of pitcher. “R” denotes right-handed pitcher, “L” denotes left-handed pitcher.
- **home_team**: 3-digit abbreviation of home team. Full team names can be found [here](#).
- **away_team**: 3-digit abbreviation of away team. Full team names can be found [here](#).
- **inning**: Inning of the pitch. Ranges from 1-11
- **outs**: Number of outs before the pitch. Ranges from 0-2
- **balls**: Number of balls before the pitch. Ranges from 0-3
- **strikes**: Number of strikes before the pitch. Ranges from 0-2
- **speed**: Speed (in mph) of the pitch thrown. Strictly positive.
- **x_location**: Horizontal location of the pitch, from the catcher’s point of view. 0 in the middle of the strike zone
- **z_location**: Vertical location of the pitch. 0 is the ground at home plate
- **strike_zone_top**: Top of the strike zone for the current batter, in feet from the ground. Unlike the width of the strike zone, this changes between batters.
- **strike_zone_bottom**: Bottom of the strike zone for the current batter, in feet from the ground. Unlike the width of the strike zone, this changes between batters.
- **x_movement**: Movement in the x-axis, in inches, from the catcher’s point of view. Positive means the pitch moved left-to-right
- **z_movement**: Movement in the z-axis, in inches, removing the effect of gravity. Positive means this pitch would have moved upwards without the effect of gravity.
- **hit_distance**: Distance, in feet, of the resulting hit ball. NA if not hit or not in play. Strictly positive.
- **hit_speed**: Speed, in mph, of resulting hit ball. NA if not hit. Strictly positive.