



TABLE: game_events						
game_str	play_id	at_bat	play_per_game	timestamp	player_position	event_code
<p>Fields are structured <math>y\{year\}_d\{day\}_{\{away\}_{\{home\}}}</math></p> <p>Each game string identifies the year and the sequential day within the season.</p>	<p>A play is defined as a situation where the ball is live.</p> <p>Play ids generally include at least three events (identified by <code>event_code</code>): the pitch, what happens to the ball, and the end of the play.</p> <p>Play ids begin at 1 and are listed consecutively for each game.</p>	<p>An at-bat corresponds to a batter at the plate. They are listed consecutively for each game.</p> <p>An at-bat includes at least one <code>play_id</code> and one <code>play_per_game</code>.</p>	<p>Plays-per-game are comparable to play ids.</p> <p>Some plays (e.g. pickoff throws) occur during an at-bat, but are not associated with the at-bat.</p>	<p>Times are shown in <i>milliseconds</i>, or</p> <p>1 sec = 1000 ms</p> <p>Timestamps start at the beginning of each game, with the time of the initial play beginning at <math>t &lt; 60</math> seconds.</p>	<p>Corresponding position numbers are given in the <b>Glossary</b>.</p> <p><i>NOTE: Each player position corresponds to a <code>game_info</code> column, and those column names are connected through the <i>Glossary</i>.</i></p>	<p>Numbers corresponding to ball events are given in the <b>Glossary</b>.</p>

TABLE: ball_pos					
game_str	play_id	timestamp	ball_position_x	ball_position_y	ball_position_z
See description in <code>game_events</code> .	See description in <code>game_events</code> .	See description in <code>game_events</code> .	$x = 0$ is defined at the line between home plate and second base, with $x > 0$ towards first base.	$y = 0$ is at the back of home plate, with $y > 0$ towards second base.	$z$ is the height with respect to $z=0$ in real-world coordinates. $z=0$ does not necessarily correspond to the ground, which is neither flat nor level.
Dimensions are given in <i>feet</i> .					

TABLE: player_pos					
game_str	play_id	timestamp	player_position	field_x	field_y
See description in <code>game_events</code> .	See description in <code>game_events</code> .	See description in <code>game_events</code> .	See description in <code>game_events</code> .	Coordinates are the same as <code>ball_pos::ball_position_x</code> .	Coordinates are the same as <code>ball_pos::ball_position_y</code> .
Dimensions are given in <i>feet</i> .					

TABLE: rosters				
DYE	OXG	QEA	RZQ	YJD
<p>Player IDs for the five "known" teams. Player IDs are structured <code>{TEAM}-####</code> where <code>{TEAM}</code> is a 3-letter designation and <code>####</code> is a unique 4-digit number, starting with a leading zero.</p> <p>"Known" teams/rosters may appear in <code>game_info</code> as either <code>home_team</code> or <code>away_team</code> variables, and they are longitudinally consistent over the entire database. The remaining teams/rosters appear only as <code>away_team</code>, and those teams/rosters are unique to a single 3-4 game series. (For example, "Known" Team A could play Team B in multiple series, but Team B and its roster would be redefined as a completely new team/roster for each.)</p>				

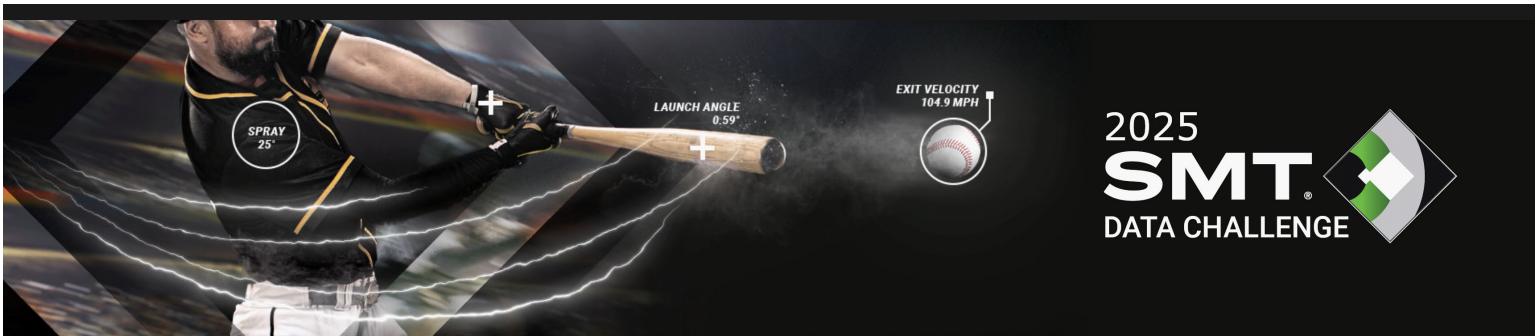
TABLE: game_info												
game_str		home_team		away_team		at_bat		play_per_game		top_bottom_inning		
See description in game_events.		3-letter team designation. All possible home teams appear in rosters.		3-letter team designation. Away teams can either appear in rosters or be unique to a 3-4 game series.		See description in game_events.		See description in game_events.		The away team bats in the top of the inning. The home team bats in the bottom.		

player_ position	pitcher		catcher		first_base		second_base		third_base		shortstop		left_field		center_field		right_field		batter		first_baserunner		second_baserunner		third_baserunner	
	1		2		3		4		5		6		7		8		9		10		11		12		13	
	Each field is populated by a {TEAM} -#### Player ID as described in rosters.																									
	Top of the inning: home team Bottom of the inning: away team																		Top of the inning: away team Bottom of the inning: home team							
	With complete data, these fields are always populated.																				Fields are populated only when runners are on base.					

## GLOSSARY

player_position	CODE	DEFINITION
	1	pitcher
	2	catcher
	3	first baseman
	4	second baseman
	5	third baseman
	6	shortstop
	7	left field
	8	center field
	9	right field
	10	batter
	11	runner on first base
	12	runner on second base
	13	runner on third base
	255	ball event with no player (e.g., ball bounce)
	14	home plate umpire
	15-17	field umpire
	18	first base coach
	19	third base coach

event_code	CODE	DEFINITION
	1	pitch
	2	ball acquired
	3	throw (ball-in-play)
	4	ball hit into play
	5	end of play
	6	pickoff throw
	7	ball acquired - unknown field position
	8	throw (ball-in-play) - unknown field position
	9	ball deflection
	10	ball deflection off of wall
	11	home run
	16	ball bounce



# Relationship between tables (ERD)

