DATA PROVIDED

During the regular season, each of the 28 NSL teams plays a total of 34 games: 476 games

Game summary fields

Field	Description	Ex1	Ex2
game_id	Unique game ID	game_2023_13	game_2023_75
HomeTeam	Home team city code	SJ	HOU
AwayTeam	Away team city code	ATX	DAL
HomeScore	Total Goals for home team	2	2
AwayScore	Total Goals for away team	2	2
Home_xG	xG total for home team, expected goals for shots	2.50	2.42
Away_xG	xG total for away team, expected goals for shots	1.76	1.73
Home_shots	shots total for home team	17	19
Away_shot	shots total for away team	9	13
Home_PK_shots	Penalty kicks awarded to home team	0	0
Away_PK_shots	Penalty kicks awarded to away team	1	0
Home_PK_goal	Goals from Penalty kicks by home team	0	0
Away_PK_goal	Goals from Penalty kicks by away team	0	0
Home_Corner	Corner kick awarded to home team	5	6
Away_Corner	Corner kick awarded to away team	3	5
Home_ToP	Time of possession of HomeTeam (fraction of total)	0.53	0.48

Team fields

Field	Description	Ex1	Ex2
name	Full name of team	Augusta Arrows FC	Charleston Mariners FC
city	Home City of team	Augusta	Charleston
state	Home state of team	Georgia	South Carolina
Team_ID	3-letter team name	AUG	CHM
conference	NASL conference of team	Eastern	Eastern
timezone_UTC	Time zone of home stadium	-4	-4

SOCCER OVERVIEW: PERIPHERAL VARIABLES

Game outcomes are determined by goals, but peripheral variables can provide a more holistic view of performance and contribute to a better-informed decision-making process. Analyzing these factors alongside the core performance metrics can lead to more accurate predictions, improved player development strategies, and a deeper understanding of the overall dynamics within a sports environment.

- Shots
- xG (Expected Goals)
- Corner kicks
- Time of Possession
- Penalty Kicks

In the following slides each of the NSL Data Science Competition variables will be defined and their utility will be discussed, *NSL Variables.*

SOCCER OVERVIEW: PERIPHERAL VARIABLES

game_id	HomeTeam	AwayTeam	HomeScore	AwayScore	Home_xG	Away_xG	Home_shots	Away_shots	Home_corner	Away_corner	Home_PK_Goal	Away_PK_Goal	Home_PK_shots	Away_PK_shots	Home_ToP
game_2023_27	FAR	ALB	4	0	1.79	0.48	12	10	9	5	0	0	0	0	0.69
game_2023_134	SPR	ALB	2	1	1.99	1.51	19	16	6	0	0	0	0	0	0.5
game_2023_204	SFS	ALB	4	1	1.71	0.73	19	14	3	1	0	0	0	0	0.54
game_2023_80	EUG	ALB	2	0	0.39	0.43	16	7	7	6	0	0	0	0	0.5
game_2023_94	TUC	ALB	1	2	0.61	0.61	9	16	2	3	0	0	0	0	0.53
game_2023_308	BAK	ALB	2	1	1.02	0.66	11	15	1	2	0	0	0	0	0.48
game_2023_241	OAK	ALB	0	1	0.95	0.5	13	8	2	4	0	0	0	0	0.47
game_2023_50	SAS	ALB	2	2	2.25	1.17	17	16	6	6	0	0	1	2	0.49
game_2023_187	PRO	ALB	4	2	0.99	0.44	11	15	1	6	0	0	1	0	0.52
game_2023_442	CHM	ALB	0	0	0.33	0.05	12	7	2	5	0	0	0	0	0.52
game_2023_120	ANC	ALB	0	0	1.14	0.79	14	13	4	6	0	0	0	0	0.52
game_2023_162	WIC	ALB	1	2	1.1	1.12	16	11	6	6	0	0	0	0	0.6
game_2023_470	BOI	ALB	1	0	0.61	0.51	10	6	7	4	0	0	0	0	0.48
game_2023_389	FOR	ALB	1	0	0.75	0.27	10	13	7	5	0	0	0	0	0.71
game_2023_366	TAC	ALB	0	2	1.6	0.16	10	9	7	0	0	0	0	0	0.56
game_2023_258	REN	ALB	0	2	0.82	0.29	15	8	6	3	0	0	0	0	0.57
game_2023_341	MOB	ALB	1	0	1.4	1.34	20	14	2	7	0	0	0	0	0.56
game_2023_87	BAK	ANC	4	4	1.9	1.27	25	8	9	5	0	0	0	1	0.51
game_2023_383	OAK	ANC	2	0	0.79	0.57	16	7	1	4	0	0	1	0	0.43
game_2023_313	SPR	ANC	3	0	3.06	0.86	23	13	7	0	0	0	0	0	0.49
game_2023_217	ALB	ANC	2	1	2.52	0.56	15	12	4	0	0	0	0	0	0.53
game_2023_129	WIC	ANC	1	0	1.95	0.8	21	12	2	1	0	0	0	0	0.59
game_2023_349	EUG	ANC	0	0	1.49	0.16	19	7	5	7	0	0	1	0	0.42
game_2023_464	FOR	ANC	0	1	1.24	0.52	10	10	5	5	0	0	0	0	0.58
game_2023_370	DOV	ANC	1	0	1.86	0.22	18	7	7	4	0	0	0	0	0.59
game_2023_431	SFS	ANC	2	0	1.25	0.24	10	3	9	2	0	0	0	0	0.52
game_2023_289	REN	ANC	0	1	1.33	0.3	22	9	7	2	0	0	0	0	0.54
game_2023_4	FAR	ANC	1	0	1.38	0.73	16	12	5	3	0	0	0	0	0.51
game_2023_42	TUC	ANC	2	2	1.29	0.92	20	10	4	7	0	0	1	0	0.58
game_2023_75	SAS	ANC	2	2	0.76	0.91	18	12	6	8	0	0	0	0	0.35
game_2023_156	TAC	ANC	0	1	1.77	0.71	10	2	6	10	0	0	0	0	0.4

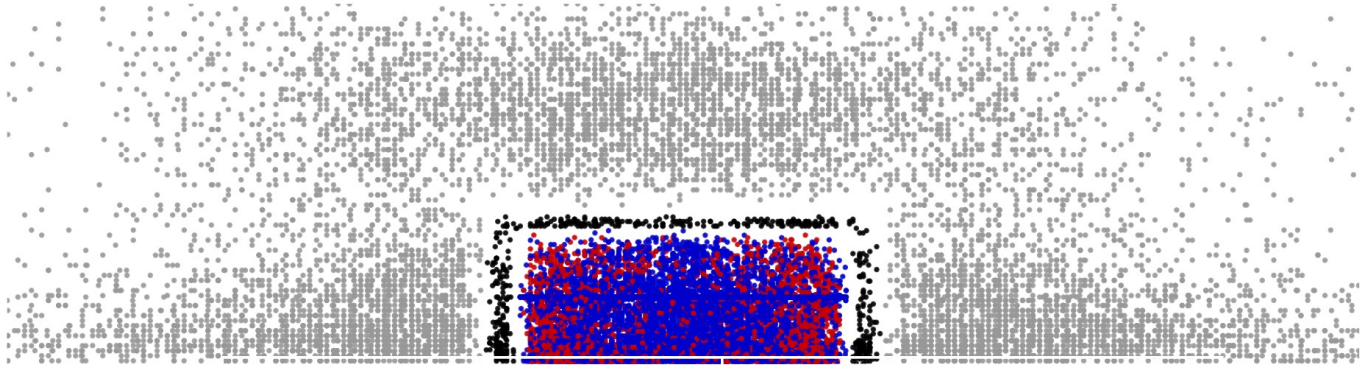
SOCCER VARIABLE: SHOTS

Shot are attempts by a player to propel the ball toward the opposing team's goal with the intent of scoring of goal. A shot can be taken using any part of the player's body, typically the feet or head, and may occur from various distances and angles on the field. The primary objective of a shot is to breach the opposing goalkeeper's defenses and result in a goal.

NSL Variables: HomeScore; AwayScore

Game Impact: Goals come from shots; shots are a standard box score stat.
Frequency: around 6 - 19 per game, per team
Success rate: around 10%; highly variable by type
Because the probability of shot success varies widely, the simple count of shots is a stat with limited utility.

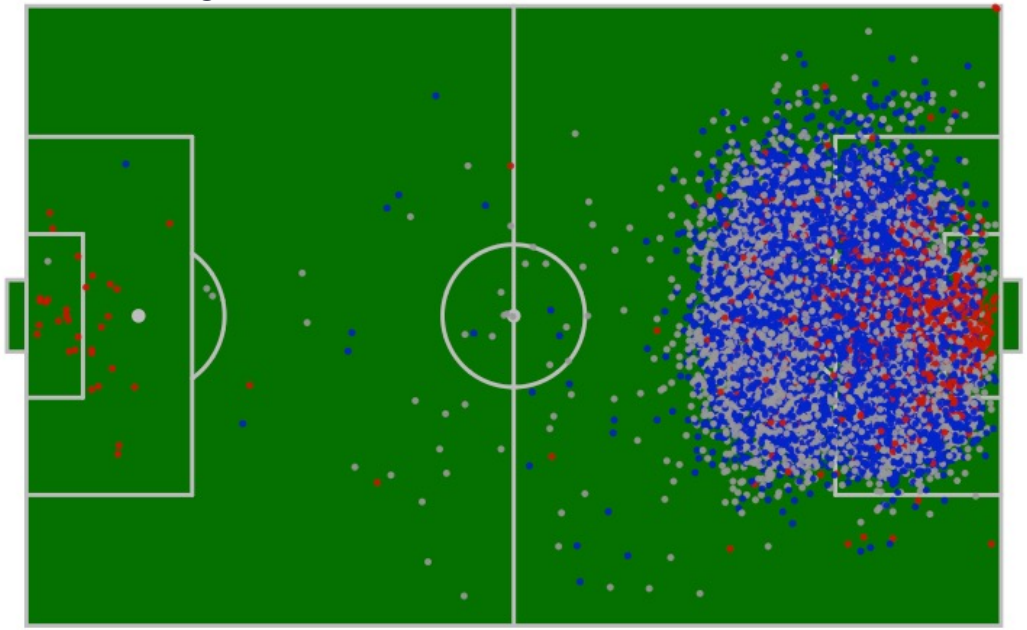
The end location of 25k shots



Grey: Miss **Red: Goal**

Black: Miss that the post **Blue: Saved by goalie**

The starting location of 25k shots



SOCCER VARIABLE: EXPECTED GOALS

Expected Goals (xG) is a statistical metric that quantifies the likelihood of a goal being scored from a particular scoring opportunity. The xG value represents the expected value (in goals) of a shot at the time the ball is struck. In this case, our variables, *Home_xG* & *Away_xG* refer to the total sum of expected goals accumulated from a team's shots in that game. xG provides a more nuanced evaluation of team performance than simply looking at goals scored. xG assess the quality of scoring opportunities a team creates or concedes, providing a more nuanced understanding of offensive and defensive performances.

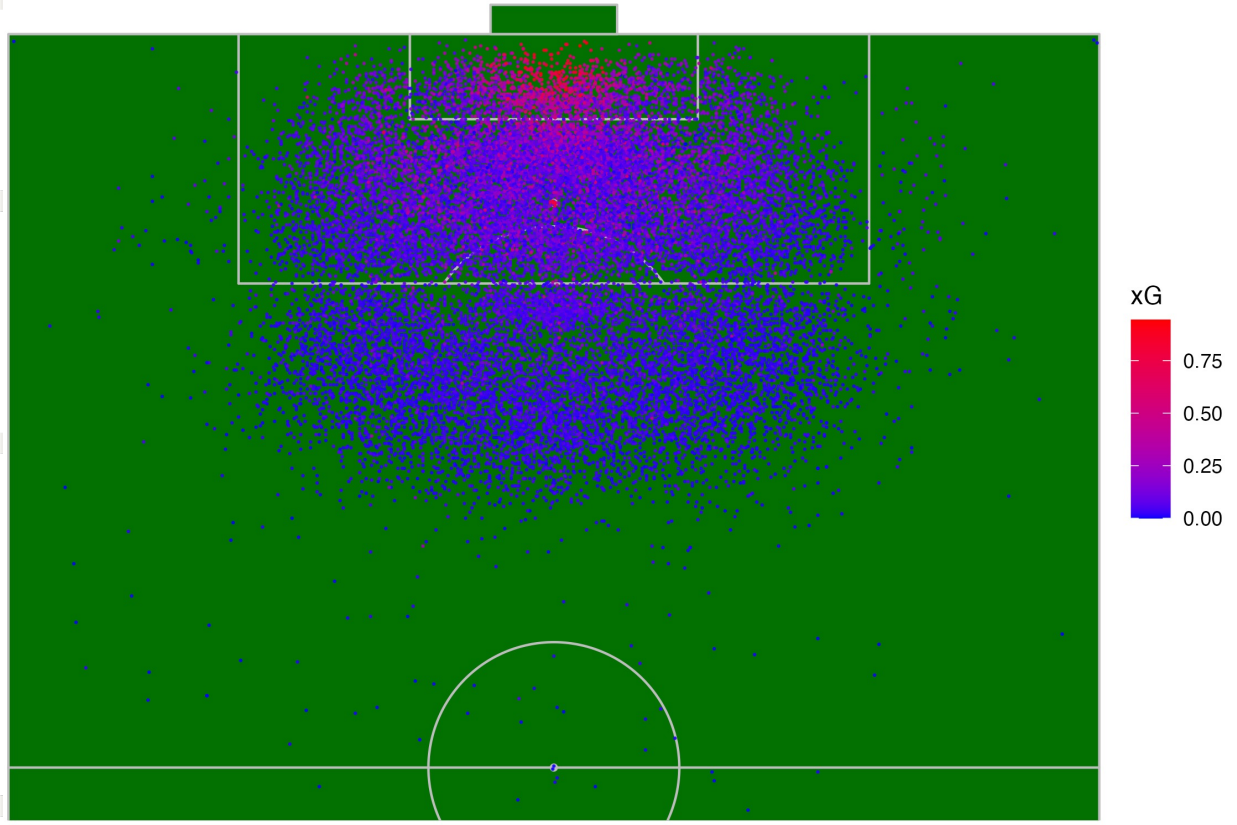
NSL Variables: Home_xG; Away_xG

xG is a black box, a model that is complex and opaque; it is a function that likely includes:

- x and y of shooter,
- x and y of goalie,
- type of play leading to the chance,
- number of defenders goal side,
- body part of the shot,
- nearest defender

The figure at right shows the xG of shots taken from different positions across the field. Note:

- Shots taken closer to the goal have a higher xG.
- xG values vary between 0 and 1 goals.
- Most shots have very low xG.



SOCCER VARIABLE: EXPECTED GOALS, CONTINUED

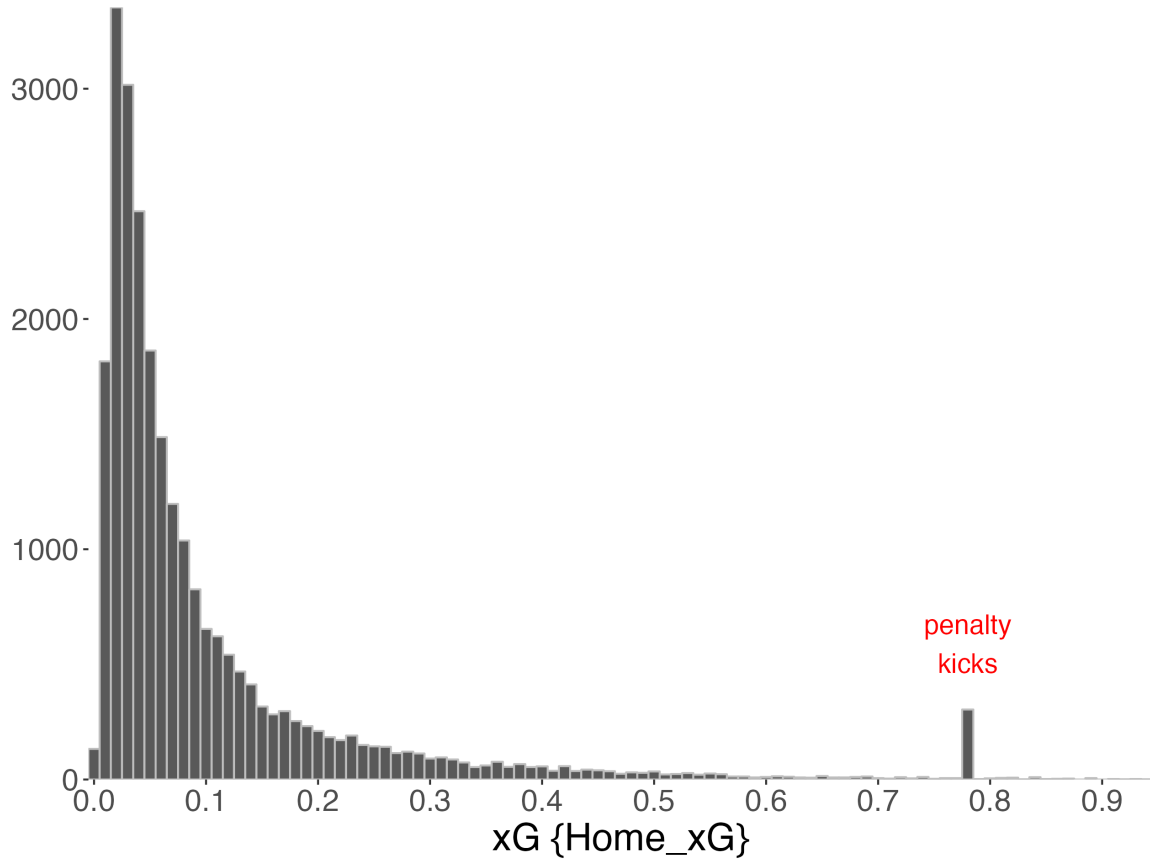
Expected Goals (xG) is a statistical metric that quantifies the likelihood of a goal being scored from a particular scoring opportunity. The xG value represents the expected value (in goals) of a shot at the time the ball is struck. In this case, our variables, *Home_xG* & *Away_xG* refer to the total sum of expected goals accumulated from a team's shots in that game. xG provides a more nuanced evaluation of team performance than simply looking at goals scored. xG assess the quality of scoring opportunities a team creates or concedes, providing a more nuanced understanding of offensive and defensive performances.

NSL Variables: Home_xG; Away_xG

Game Impact: high **-** more predictive of goals than the count of shots.
Frequency: xG is determine for each of the around 6 - 19 shots per game, per team

The figure at right shows xG values from one season of shots

- xG values vary between 0 and 1 goals.
- Most shots have very low xG;
- Median xG = 0.053
- Penalty kicks have high xG value.

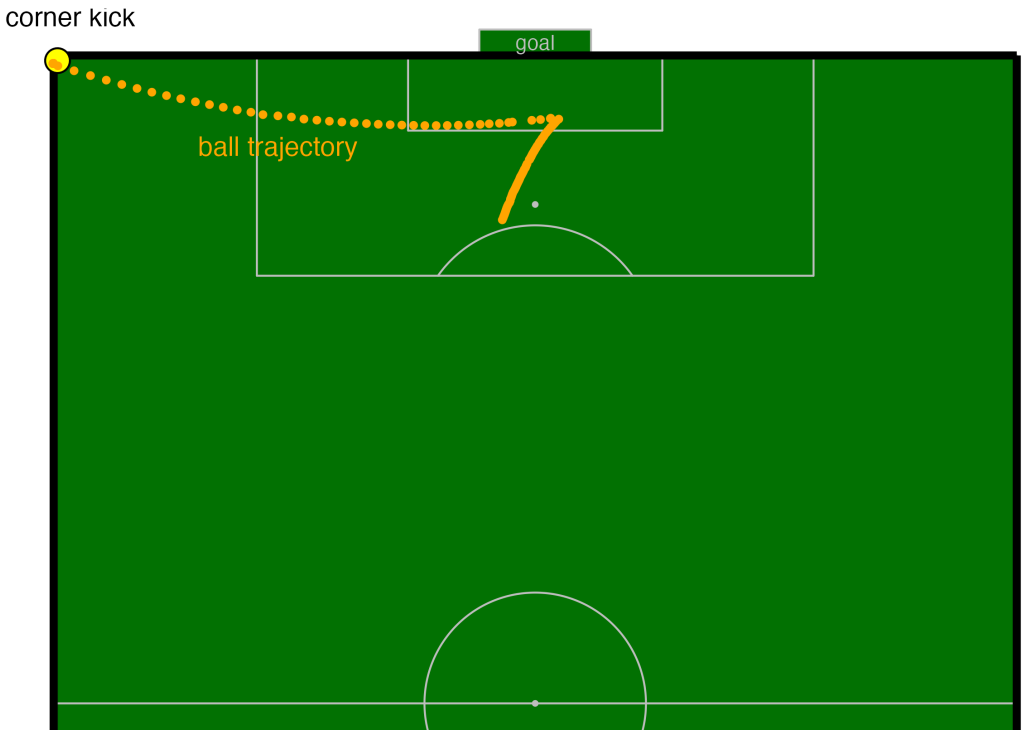


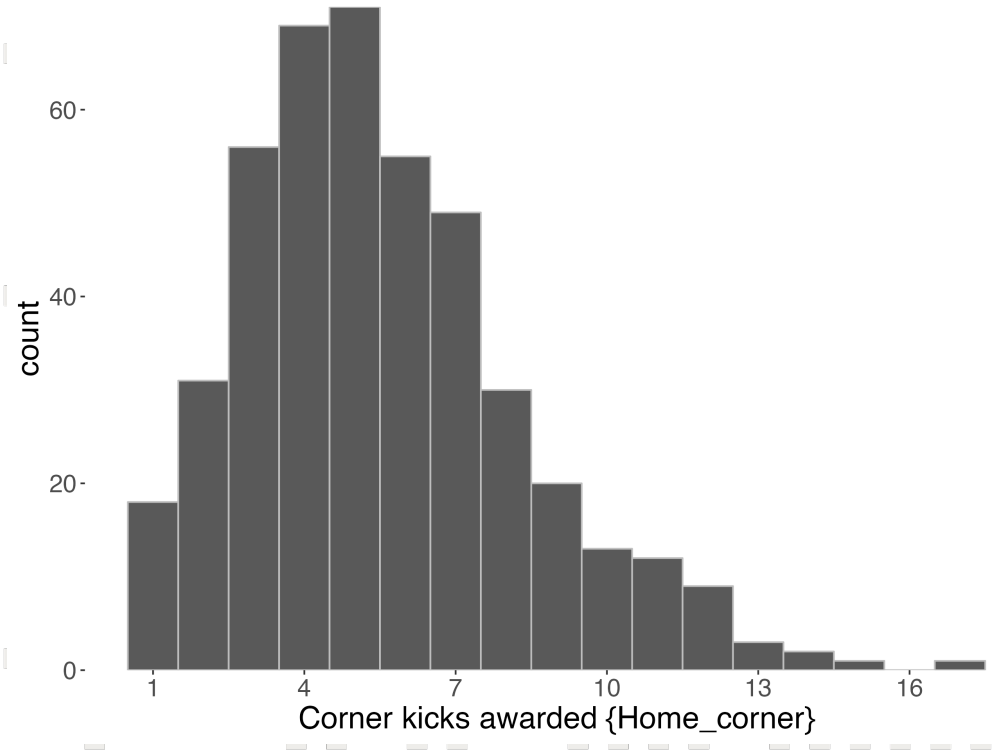
SOCCER VARIABLE: CORNER KICKS

Corner kicks occur when the ball crosses the goal line, the end of the field of play, after touching a defending player, resulting in an attacking team awarded a set-piece opportunity, a restart of play from a stoppage. This takes place from the corner of the field, where the touchline, sideline, meets the goal line. During a corner kick, the objective is to create goal-scoring opportunities by exploiting the height and heading abilities of players. Defensively, teams aim to clear the ball away from their goal.

NSL Variables: Home_corner; Away_corner

Game Impact: low - goals from corner kicks are infrequent
Frequency: around 2 - 8 per game, per team
Success rate: low, around 3%.





SOCCER VARIABLE: TIME OF POSSESSION

Time of possession (ToP) refers to the amount of time a team spends in control of the ball during a match. ToP can be calculated in a few different ways;

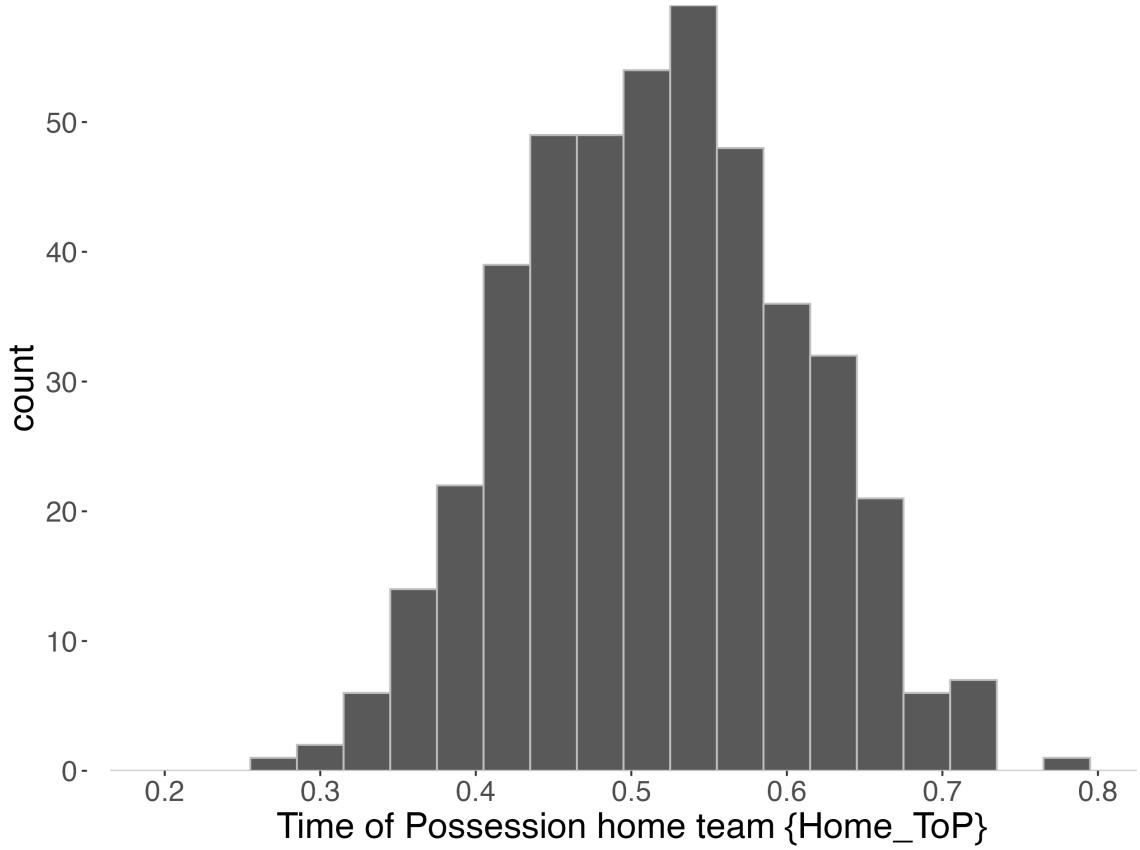
- Stop-watch based, counting time that team holds the ball, or
- Event-based (fraction of total passes, or total completed passes),

All methods have problems; soccer has frequent stoppages and the pace of play is inconsistent.

ToP is probably more indicative of a team's style of play, how a team organizes defensively and attacks. Some teams emphasize possession and short passes, while other teams press the ball forward and setup counter-attacks. Both strategies can be successful and results in different ToP.

NSL Variables: Home_ToP;
ToP_Away can be calculated, (1 - ToP_Home)

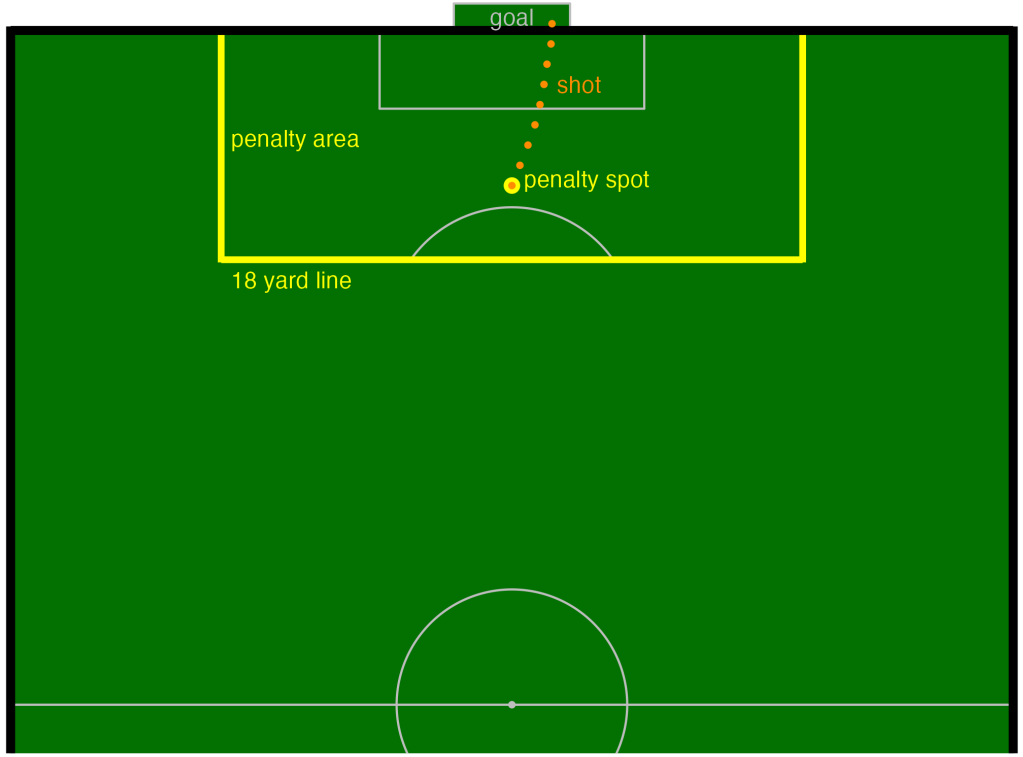
Game Impact: unclear.
ToP appears to be coinvent variable; each team has a game level score for every game. Higher ToP would appear to be better, but is ToP predicted of game outcomes?



SOCCER VARIABLE: PENALTY KICKS

A penalty kick is awarded following a foul committed by a defending player inside their own penalty area, near the attacking goal. The ball is placed on the penalty spot, 12 yards from the goal. The penalty kick is a direct free-kick, the ball can be kicked directly into the goal. Any player on the offensive team can take the penalty kick, so typically on each team the best penalty shooter takes all the penalty shots awarded.

NSL Variables: Home_PK_Goal; Away_PK_Goal; Home_PK_shots; Away_PK_shots



Game Impact: Rare but hugely impactful
Frequency: 0.3 per game, per team
Success rate: around 78%.

For more information see the Module: Soccer Penalty Kicks