Optimal Batting Approach & Season Projections

By Nick Lukowsky

Project Overview

- Batters skills in the MLB (Major League Baseball) naturally change as they age over their careers
- With new advanced metrics being tracked each year, batters should attempt to best tweak their approach to converge to ideal values of these metrics
- In conjunction with knowing how their game will naturally change year-over-year, batters can extend their careers and maximize success through the use of metrics
- Findings can not only be used by batters and hitting coaches, but for front offices to build their roster with batters who excel in these areas

Goals & Predictions

Batting Approach

Use optimization to find ideal batting approach metrics

Predictions

High exit velocity, pull% and a disciplined approach will be recommended

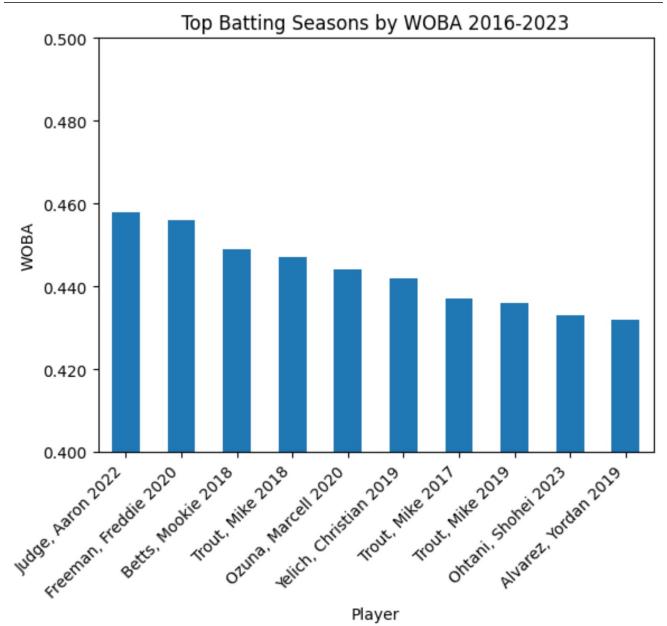
Batting Projections

Use previous player season data and aging curve to project next season

Predictions

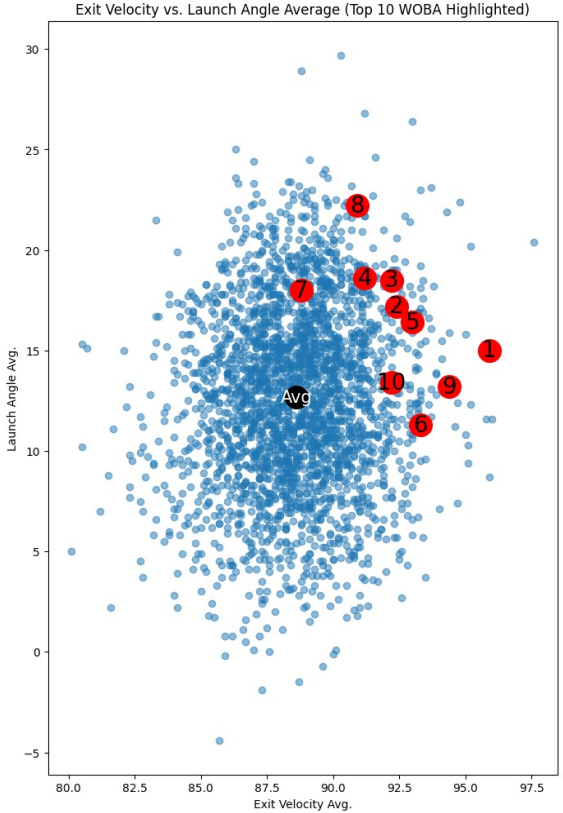
Projections will be based off a weighted average of prior seasons and a peak age will be approached for each statistic

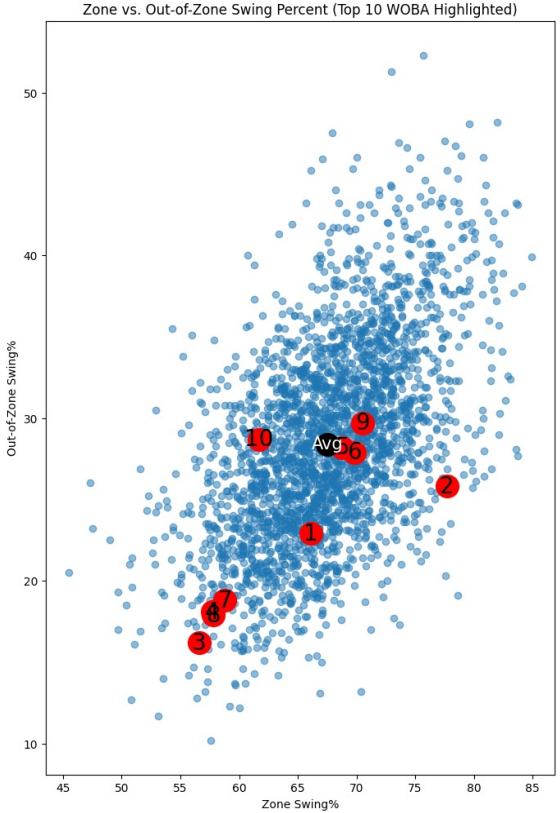
Data Exploration

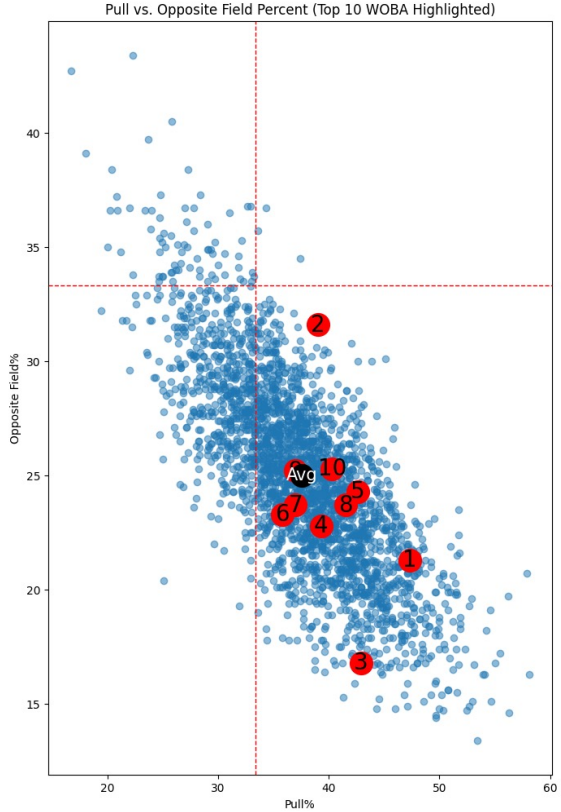


- Dataset comes from Baseball Savant, which compiles traditional statistics (HR, AVG, OPS, etc.) with newer and more advanced swing tracking statistics (Exit Velocity, Launch Angle, Zone Swing%, etc.)
- 4264 rows and 28 columns
- All batting seasons from 2016-2023 w/ more than 50 PAs are included
- Plot shows top 10 batting seasons by WOBA
- For calculations where rate statistics are used, data is filtered to 200 or more PAs

Data Exploration



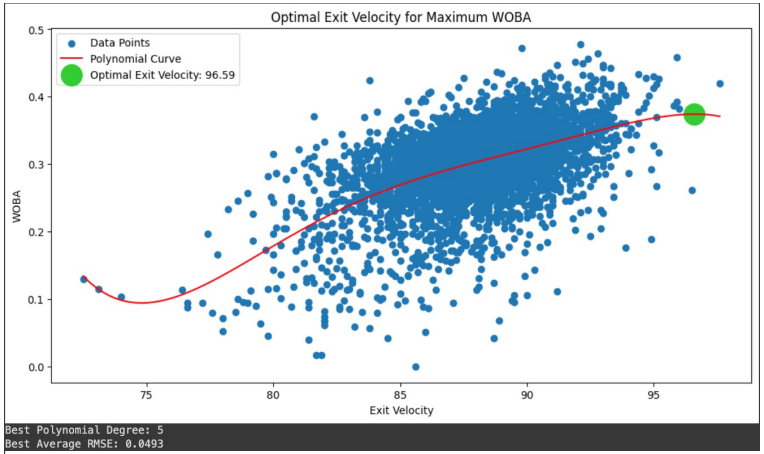


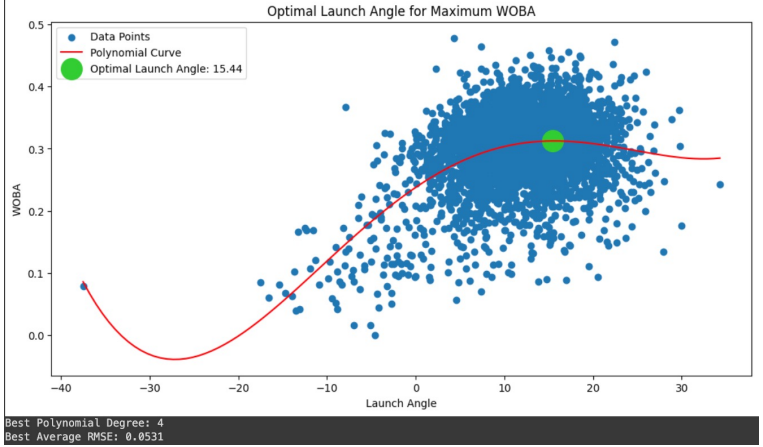


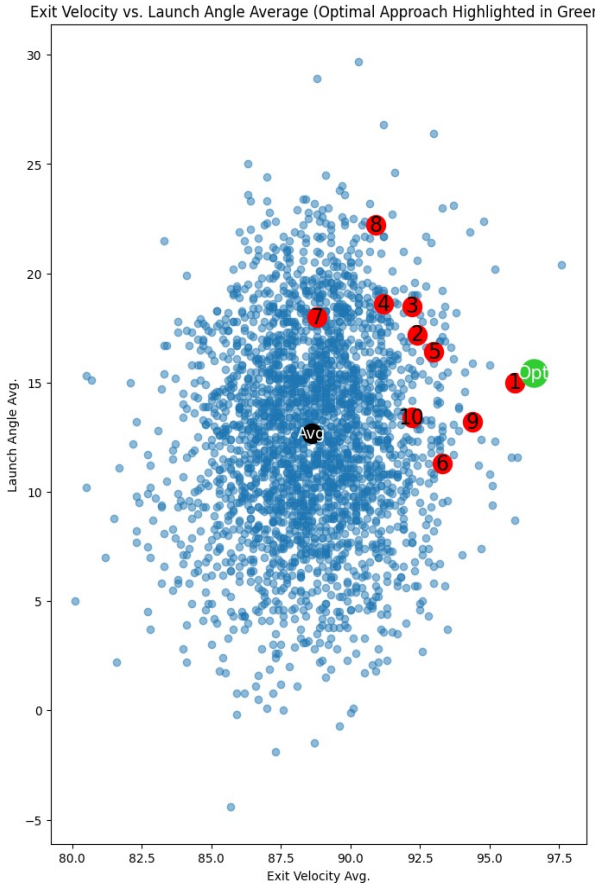
Methodology

- Using the top 10 batting seasons by WOBA in the dataset you will be able to see where elite hitters rank in each metric
- To find optimal metric values a polynomial regression is fit to each metric and WOBA, and "peak" or metric value where WOBA is optimized is found
- Using k-fold cross validation different degrees are tested and the degree with the minimum RMSE value is used
- For statistic projections a similar method is used to that of popular MLB projection system Marcel, which was created by Tom Tango
- "Uses three years of MLB data, with the most recent data weighted heavier. It regresses toward the mean. And it has an age factor"
- A function will be created that when a player name is inputted, statistic projections will be made for the next season that would occur

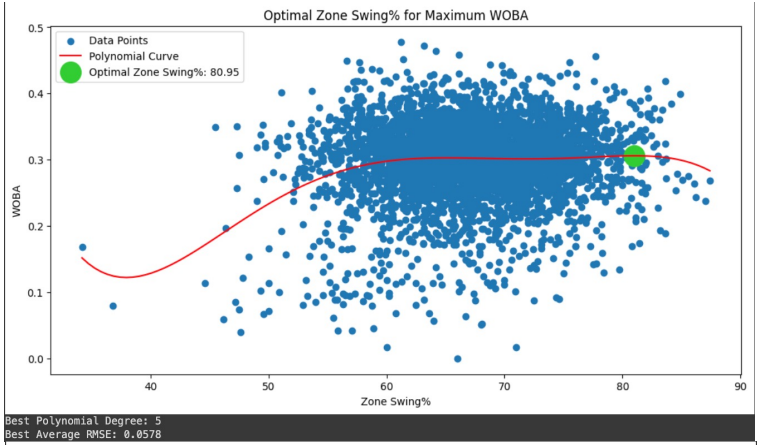
Exit Velocity & Launch Angle

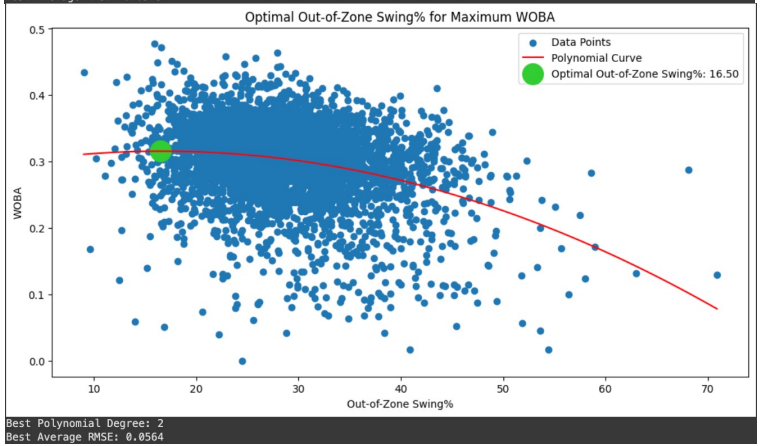


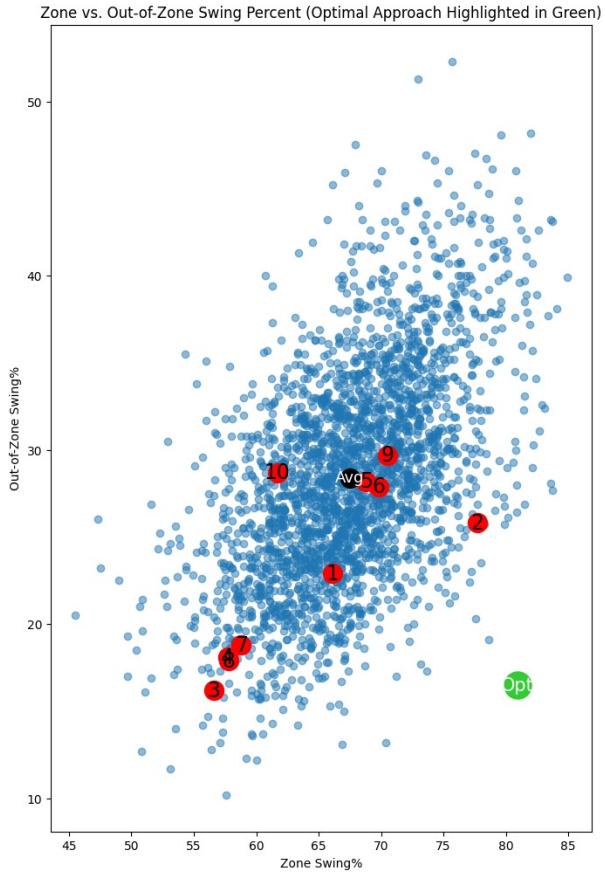




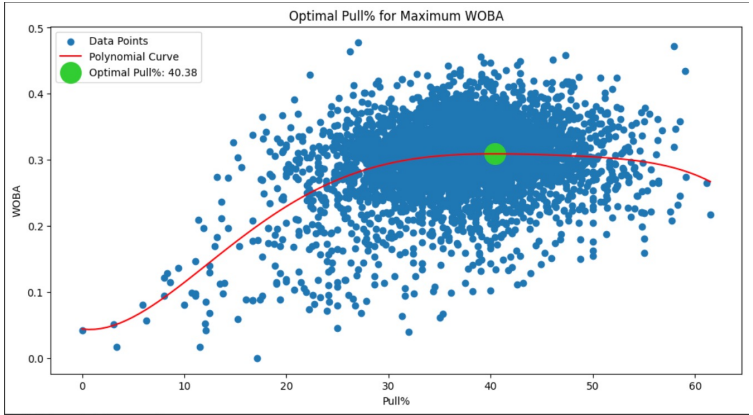
Swing Metrics

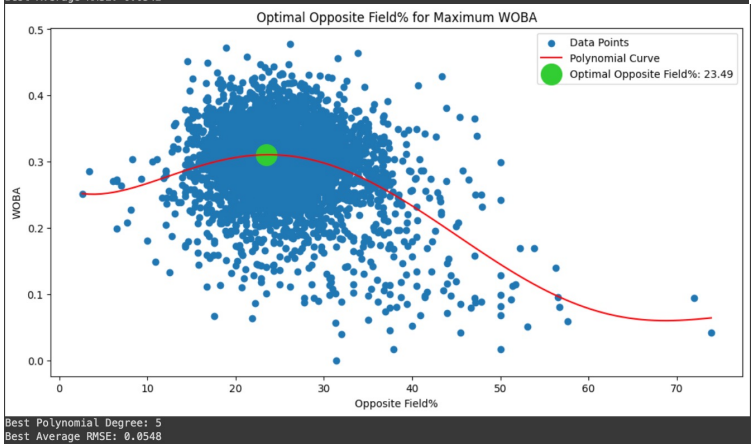


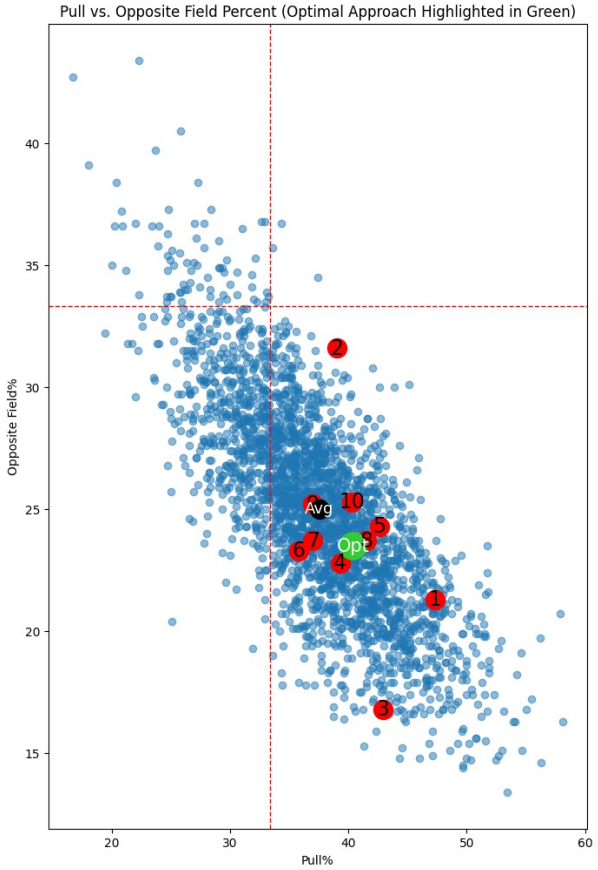




Hit Location

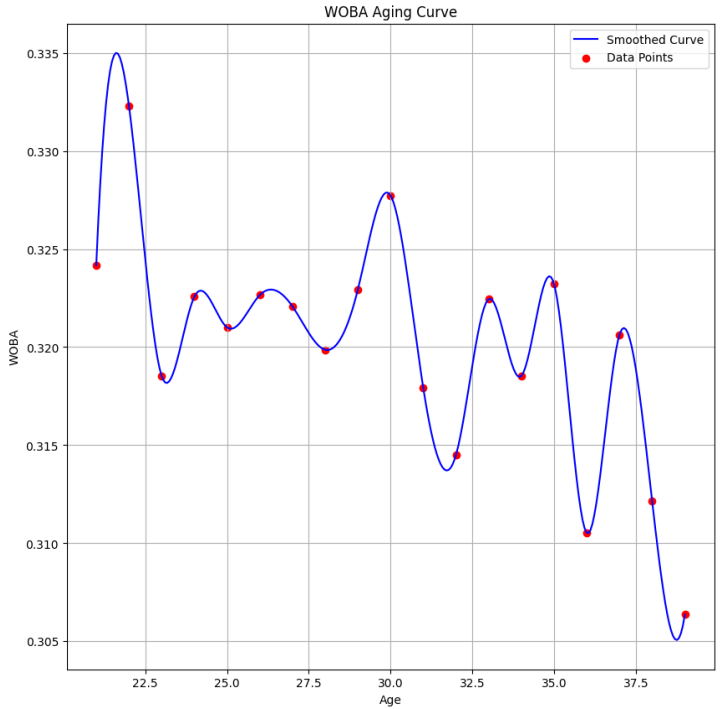






Projections: Aging Curves

	player_age	avg_woba_aging_curve	percent_change_woba
0	21	0.324167	NaN
1	22	0.332300	2.508997
2	23	0.318528	-4.144448
3	24	0.322566	1.267599
4	25	0.321004	-0.484090
5	26	0.322650	0.512606
6	27	0.322075	-0.178099
7	28	0.319864	-0.686439
8	29	0.322932	0.959150
9	30	0.327730	1.485678
10	31	0.317921	-2.992828
11	32	0.314497	-1.077152
12	33	0.322447	2.527905
13	34	0.318529	-1.215213
14	35	0.323211	1.469817
15	36	0.310529	-3.923484
16	37	0.320625	3.251089
17	38	0.312143	-2.645503
18	39	0.306364	-1.851467



2024 Projections Example

- **Formula to create each statistic projection:**
- **Weighted Mean = ((3rd most recent season x * 0.1) + (2nd most recent season x * 0.3) + (Most recent season x * 0.6))**
- **Projection = Weighted Mean + (Weighted Mean * (Aging curve percent change x/100))**

name	year	age	pa	hit	home_run	avg	ops	woba	ev	la	z_swing%	oz_swing%	pull%	opp_field%
Ohtani, Shohei	2021	26	639	138	46	0.257	0.964	0.393	93.6	16.6	69.7	27.3	46.6	22.9
Ohtani, Shohei	2022	27	666	160	34	0.273	0.875	0.370	92.9	12.1	72.1	28.4	36.0	27.8
Ohtani, Shohei	2023	28	599	151	44	0.304	1.066	0.433	94.4	13.2	70.5	29.7	37.0	25.2

name	year	age	pa	hit	home_run	avg	ops	woba	ev	la	z_swing%	oz_swing%	pull%	opp_field%
Ohtani, Shohei	2024	29	645	159	45	0.293	1.01	0.414	94.0	12.8	70.4	29.0	37.5	25.7

Projections (Scaled) to Real-Life 2024 Statistics:

name	year	age	pa	hit	home_run	avg	ops	woba	ev	la	z_swing%	oz_swing%	pull%	opp_field%
Ohtani, Shohei	2024	29	110	27	8	0.293	1.01	0.414	94.0	12.8	70.4	29.0	37.5	25.7

name	year	age	pa	hit	home_run	avg	ops	woba	ev	la	z_swing%	oz_swing%	pull%	opp_field%
Ohtani, Shohei	2024	29	110	35	5	0.368	1.094	0.467	95	13.8	71.4	27.1	46.8	23.4

Summary

Optimal batting approach metrics are as follows:

- Exit Velocity: 96.59 mph
- Launch Angle: 15.44 degrees
- Zone Swing%: 80.95%
- Out-of-Zone Swing%: 16.50%
- Pull%: 40.38%
- Opposite Field%: 23.49%

Aging curves were found for all relevant statistics and metrics and used in projections

Using a similar method to real-life projection systems a function that produces projections for a given player was created

Findings can be used by MLB batters to adapt their game using new advanced tracking metrics and methods

Challenges & Next Steps

- Context of approach trends (Exit Velo/Contact tradeoff)
- Positional aging curve differences
- Some component of luck adjustment added into projections
- Juiced ball occurrences & covid year
- Projecting young players
- Ability to put optimal approach suggestions into action through appropriate training

Resources

- https://baseballsavant.mlb.com/leaderboard/custom?year=2023%2C2022%2C2021%2C2020%2C2019%2C2018%2C2017%2C2016&type=batter&filter=&min=50&selections=player_age%2Cpa%2Chit%2Chome_run%2Ck_percent%2Cbb_percent%2Cbatting_avg%2Con_base_plus_slg%2Cbabip%2Cwoba%2Cxwoba%2Cexit_velocity_avg%2Claunch_angle_avg%2Csweet_spot_percent%2Cbarrel_batted_rate%2Chard_hit_percent%2Cavg_best_speed%2Cavg_hyper_speed%2Cz_swing_percent%2Coz_swing_percent%2Cwhiff_percent%2Cswing_percent%2Cpull_percent%2Cstraightaway_percent%2Copposite_percent&chart=false&x=player_age&y=player_age&r=no&chartType=beeswarm&sort=xwoba&sortDir=desc

- https://aamir07.medium.com/polynomial-regression-with-k-fold-cross-validation-bc5275137546

- https://blogs.fangraphs.com/checking-in-on-the-aging-curve/

- https://www.mlb.com/glossary/projection-systems/marcel-the-monkey-forecasting-system

Thanks!

Questions?