Fantasy Baseball Draft Tool

Documentation and Methodology

Claude + Kevin

February 2026

Outline

What Does This Tool Do?

Goal: Rank players by their marginal contribution to winning H2H categories.

Key Components:

1. `create_league_stats.py` – Generates hitter CSVs with z-scores
2. `create_pitching_stats.py` – Generates pitcher CSVs
3. `normalize_pa.py` – Aligns PA between projection systems
4. `draft_tool.html` – Interactive draft UI with real-time valuations

League Format: 14-category H2H weekly

- Hitting: R, HR, RBI, SB, SO, TB, OBP
- Pitching: W, SV, K, HLD, ERA, WHIP, QS

The Core Idea: Marginal Win Probability

In H2H, you win a category if your weekly total beats your opponent's.

Assuming team totals are normally distributed:

$$P(\text{win category}) = \Phi\left(\frac{\mu_{\text{my team}} - \mu_{\text{opponent}}}{\sigma \cdot \sqrt{2}}\right)$$

where:

- Φ = standard normal CDF
- σ = weekly standard deviation for that category (from 2024 data)
- The $\sqrt{2}$ comes from the variance of the difference of two normals

A player's value = how much they move $P(\text{win})$ vs. replacement.

Weekly Standard Deviations (2024 League Data)

Hitting Categories:

Category	SD	Avg	Notes
R	6.03	28.96	
HR	2.93	8.02	
RBI	6.72	27.86	
SB	**2.57**	4.74	Tight – high leverage
SO	7.45	50.11	Lower is better
TB	15.94	88.87	
OBP	0.04	0.320	

Key insight: Categories with tight SDs (SB, SV) have outsized impact. A steal is worth $\frac{1}{2.57} = 0.39$ SDs. A strikeout is only $\frac{1}{7.45} = 0.13$ SDs.

Pitching Standard Deviations

Category	SD	Avg	Notes
L	1.83	3.08	Lower is better
SV	**1.54**	2.27	Very tight – closers valuable
K	11.79	50.90	Wide – less leverage
HLD	1.64	2.30	Tight
ERA	1.31	3.79	Lower is better
WHIP	0.21	1.20	Lower is better
QS	1.40	3.25	

Why Edwin Díaz matters: +1.35 saves/week ÷ 1.54 SD = **0.88 standard deviations**.
Compare: +0.94 K/week ÷ 11.79 SD = 0.08 SDs. Saves are 11× more impactful.

How Z-Scores Are Calculated in CSVs

For **counting stats** (R, HR, RBI, TB, SB):

$$z = \frac{\text{Season Total}/25}{\text{Weekly SD}}$$

For **strikeouts** (lower is better):

$$z_{SO} = -\frac{\text{Season SO}/25}{\text{SD}_{SO}}$$

[?] For OBP, the formula is:

$$z_{OBP} = \frac{(\text{OBP} - 0.320)}{9 \times \text{SD}_{OBP}}$$

Why divide by 9? I believe this is meant to scale OBP's contribution to match counting stats, but the factor of 9 is not obviously derived. This needs verification.

Total Z-Score

$$z_{\text{total}} = z_R + z_{HR} + z_{RBI} + z_{SB} + z_{TB} + z_{SO} + z_{OBP}$$

This is a simplified ranking metric used to sort players in the CSV.

Important distinction:

- **CSV z-scores**: Quick approximation for ranking
- **Draft tool marginal value**: Exact calculation considering current roster

Both should *mostly* agree on rankings, but the draft tool is more precise because it accounts for your specific team composition.

What Is Replacement Level?

Definition: The production available from freely available players (waiver wire).

Methodology:

1. Rank all hitters by z_{total} using Depth Charts projections
2. Take players ranked 155–175 (just beyond 16 teams $\times$ 9 hitters $=$ 144)
3. Average their per-PA rates

The cohort (21 players):
Spencer Steer, Miguel Andujar, Ezequiel Tovar, Jonathan Aranda, Addison Barger, Nathan Lukes, Kyle Manzardo, Colt Keith, Josh Lowe, Romy Gonzalez, Samuel Basallo, Francisco Alvarez, Lars Nootbaar, Joey Ortiz, Tyler O'Neill, Ryan O'Hearn, Victor Robles, Jake Fraley, JJ Bleday, Munetaka Murakami, Chase Meidroth

Replacement Level Per-PA Rates

Stat	Per-PA Rate	Full Season (600 PA)
R	0.1212	73
HR	0.0330	20
RBI	0.1208	72
SO	0.2226	134
TB	0.3725	224
SB	0.0152	9
OBP	**0.320**	(capped at league avg)

[?] OBP was hardcoded to 0.320 because the cohort's actual OBP (0.324) was *above* league average, which would make replacement players OBP-positive. This seemed wrong.

PA Supplementation

Problem: Low-PA players look bad in counting stats.

Solution: Supplement everyone to 600 PA with replacement-level production.

For a player with 450 PA:

$$\text{Gap PA} = 600 - 450 = 150$$

$$\text{Runs} = \text{Projected Runs} + 150 \times 0.1212$$

$$\text{OBP} = \frac{450 \times \text{proj OBP} + 150 \times 0.320}{600}$$

Interpretation: "What would this player produce if they played full-time, with a replacement-level player filling in the rest?"

RP Replacement Level (Per Slot, Weekly)

Stat	Replacement RP
IP/week	2.48
L/week	0.118
SV/week	0.121
HLD/week	0.848
K/week	2.69
ER/week	0.96
WH/week	2.85

Key observation: Replacement RPs get **holds, not saves**.
Elite closers get saves but sacrifice holds. This creates an explicit tradeoff that the model captures.

Case Study: Edwin Díaz vs Replacement

Stat	Díaz	Replacement	Diff
SV/week	1.472	0.121	**+1.35**
HLD/week	0.099	0.848	**-0.75**
K/week	3.63	2.69	+0.94
ER/week	0.90	0.96	-0.06

Win probability impact:

- Saves: **+20.9%**
- Holds: **-12.8%**
- Strikeouts: +1.9%
- ERA/WHIP: $<$1% (diluted across 40 IP/week)

Net marginal value: $\approx$ **0.11** (Díaz adds 0.11 expected category wins/week)

Why ERA/WHIP Don't Matter Much for RPs

ERA/WHIP are innings-weighted ratio stats.

Typical team: 40 IP/week total

- 5 SP $\times$ 6.5 IP $=$ 32.5 IP
- 3 RP $\times$ 2.5 IP $=$ 7.5 IP

Díaz contributes **2.66 IP / 40 IP = 6.7%** of team innings.

Even if Díaz has much better ERA than replacement (3.06 vs 3.50), the team ERA only improves by $\approx$0.02 points.
With ERA SD $=$ 1.31, that's $0.02/1.31 = 0.015$ SDs $\rightarrow$ **+0.4% win probability**.

Conclusion: RPs earn value through saves/holds/K, not ERA/WHIP.

The Problem: Projection Systems Disagree on PT

Different projection systems predict different playing time:

Player	The Bat PA	Depth Charts PA
Player A	550	620
Player B	480	510

Problem: We want to compare *skill*, not playing time estimates.

Solution: Normalize all systems to use Depth Charts PA.

PA Normalization Process

`normalize_pa.py`:

1. Read Depth Charts PA for each player
2. For each player in The Bat/BatX:
   - Scale = DC_PA / TheBat_PA
   - Multiply all counting stats by Scale
   - Keep rate stats (OBP, K%) unchanged
3. Output normalized CSV

[?] Keeping rate stats unchanged assumes that the rate estimates are independent of playing time. This is approximately true but not perfect (e.g., platoon players might have different rates in limited samples).

Three Projection Systems Available

1. **Depth Charts** (default) – Composite of multiple systems
2. **The Bat** – Tom Tango's projection system
3. **The BatX** – Extended/experimental version

All three use:

- Same PA (normalized to Depth Charts)
- Same replacement level rates
- Same weekly SDs

Only difference: Rate stat projections (HR/PA, SB/PA, K%, OBP, etc.)

Things I'm Uncertain About

1. **[?] OBP z-score formula divides by 9** – Why 9? This seems to arbitrarily scale down OBP's contribution. Should verify this is intentional.
2. **[?] Replacement OBP was hardcoded** – The cohort (ranks 155-175) had OBP $= 0.324$, but I capped it at 0.320 to be OBP-neutral. Is this right?
3. **[?] Choice of ranks 155-175** – Why not 145-165 or 160-180? The exact cutoff affects replacement level.
4. **[?] Rate stats in PA normalization** – Should K%, OBP stay fixed when scaling PA? Small sample effects might distort rates.

Things I'm Uncertain About (continued)

5. **[?] SP replacement level** – I didn't examine this closely. Need to verify the methodology matches hitter replacement.
6. **[?] Keepers integration** – Keepers were added but I don't recall the full implementation details.
7. **[?] Weekly SD derivation** – These come from 2024 league data. Are they calculated correctly? Do they need smoothing?
8. **[?] The $\sqrt{2}$ factor** – The formula uses $\sigma\sqrt{2}$ for the difference of two normals, which is correct IF both teams have the same variance. Does this hold?

Feature Timeline

1. **Base draft tool** – UI, player data, basic rankings
2. **Marginal win probability** – Exact $\Phi()$ calculation
3. **PA supplementation to 600 PA floor** – Fill with replacement
4. **Projection toggle** – The Bat vs Depth Charts
5. **PA normalization** – All systems use DC playing time
6. **300 PA minimum filter** – Exclude low-PA players
7. **Replacement recalibration** – Ranks 155-175 cohort
8. **OBP cap at league average** – Prevent OBP-positive replacement
9. **The BatX projection system** – Third toggle option
10. **Keepers support** – Pre-load owned players

Key Code Files

Python scripts:

- `create_league_stats.py` – Hitter z-scores, replacement supplementation
- `create_pitching_stats.py` – Pitcher processing
- `normalize_pa.py` – Align PA across systems

Main application:

- `draft_tool.html` – All-in-one HTML/CSS/JS app
  - Lines 482-506: SD and replacement constants
  - Lines 580-599: `normalCDF` and `winProbability`
  - Lines 604+: Roster projection calculations

Data files:

- `fantasy_hitters_dc_2026.csv`, `fantasy_hitters_thebat_2026.csv`, `fantasy_hitters_batx_2026.csv`

Summary

What the tool does well:

- Captures category leverage via weekly SDs
- Models saves vs holds tradeoff for relievers
- Accounts for PA differences via supplementation
- Allows projection system comparison

What needs review:

- OBP z-score formula (the $\div 9$)
- Replacement level cohort selection
- SP replacement methodology
- Edge cases in keepers