

## **Week 2 - Undervalued Player Identification**

*By: Robert Banas, Bailey Harrison, Simran Mallik, Jai Arora*

### **Data Cleaning**

To identify 5 undervalued players, our team used three datasets. The first dataset was compiled by Premier Talent Sports and Entertainment, which had relevant variables including player names and service time (minimum 6 years). Using R, we joined this dataset with a Fangraphs dashboard dataset for hitters with a minimum of 10 plate appearances, and also joined the first dataset with a Fangraphs advanced dataset for pitchers with a minimum of 10 innings pitches. Our new hitters and pitchers datasets were created.

These datasets included the metrics that we thought would help predict successful on-field performance. For hitters, we identified metrics Base on Balls percentage (BB%), Weighted Runs Created (wRC+), EV (Exit Velocity), and Wins Above Replacement (WAR). For pitchers, we identified metrics HR/9 (Home Runs Per Nine Innings), BB% (Walk Rate), AVG (Batting Average Against), and ERA- (Earned Run Average minus).

The rest of the data cleaning was conducted collaboratively in google sheets. Our hitters and pitchers dataset contained variables including player name, position, major league service time, and 2020 salary. For each metric in each dataset, there was a metric column, a metric rank column, a weighted metric rank column, a total score column, and a final player ranking column. In the metric rank column, we found the rank of each player in terms of that metric and this value into the column. Then, we multiplied each player's metric rank by a specific weight.

We multiplied weights by each player's metric because we had metrics with average values that were very variable in magnitude. We wanted to weigh each metric differently since we believed not all metrics should hold the same value, but because of this great variability in the magnitude of the values of each metric, we multiplied weights by the metric ranks instead of by the actual scores to ensure the values were comparable. Finally, we totaled the weighted metric rank scores to create a total score and ranked the players from highest to the lowest total score. The highest score indicated the best player.

After the player ranking was finished, each person identified 5 players who they believed were undervalued. We did this by looking at the salaries that were ordered according to the ranking of the hitters and pitchers in their lists, and scanning for significant jumps in salaries for players within the upper quartile. If there was a significant negative difference in one player's salary compared to the salaries of the other players ranking immediately above or below that person, then we identified that player as a potentially undervalued player. We compiled a list of undervalued players and selected 5 from that list as our 5 undervalued players.

## Metric Selection

### Position Players

The 4 metrics we decided to evaluate position players included WAR (wins above replacement), WRC+ (weighted runs created), EV (average exit velocity), and BB% (Base on Balls percentage). WAR and WRC+ were weighted in our formula at 35% each, with BB% and EV also equally weighted at 15%. We multiplied these corresponding weights by the rank of each player in each category. Ranks were awarded based on a point system. To show this works, as an example, with 134 position players in this sample, Freddie Freeman came out as the player with the highest WRC+ in the dataset, so he was awarded 134 points multiplied by the 35% weight for the WRC+ metric here (46.9 points added). The same methodology was applied for the 3 other statistics/weights. After doing this with each player, their cumulative points were added up, with more points signifying better production. Players were finally ranked in order from highest to lowest rank score. Our team then looked for players in the upper quartile of production with low 2020 salaries relative to their production.

**Formula:**  $35\% \times \text{WRC+ rank} + 35\% \times \text{WAR rank} + 15\% \times \text{BB\% rank} + 15\% \times \text{EV rank} = \text{Total Points}$

Our intuition behind the metrics selected is that with a wide disparity in positional specialization importance in regard to defense and speed with catchers, infielders, and outfielders, we decided to focus on mostly hitting metrics only. WAR does consider the total value of a player (defense and baserunning included) which is why we weighted it the most, but other than that, we decided to focus on what a player brings to the plate. We thought WRC+ was a great statistic to value how much run production a hitter provides which is the name of the game which is why we weighted this heavily at 35% as well. We chose BB% because we thought it was an important metric that values a hitter's plate discipline. Finally, we used the Statcast statistic of exit velocity to showcase a hitter's power and see if he barrels up the ball a lot. Weak EV can signify if a player with good statistics is posed for regression or vice versa.

### Pitchers

The 4 metrics we decided to use to evaluate pitchers are ERA- (Earned Run Average minus), HR/9 (Home Runs Per Nine Innings), BB% (Walk Rate), and AVG (Batting Average Against). We didn't use WAR because of the disparity between starters and relievers. ERA- was weighted in our formula at 40% and HR/9, BB%, and Avg were each weighted at 20%. We weighted ERA- more because we thought that metric was the most important. ERA- takes a pitcher's ERA and compares it to the league's average, scoring from a center of 100. For every point below 100 in an ERA- score, that means that the pitcher's ERA is one percentage point better than the league's average. HR/9 represents the average number of home runs allowed by a pitcher on a nine-inning scale. BB% represents the frequency with which a pitcher walks hitters. Avg is determined by dividing the number of hits against a given pitcher by the number of at-bats against him. Next, we ranked all of the pitchers from 111 to 1 since there were 111 players in the sample, with 111 being the best at that metric. Then, we multiplied these corresponding weights

by the rank of each player in each category. Ranks were awarded based on a point system. Players were finally ranked in order from highest to lowest rank score. Our team then looked for players in the upper quartile of production with low 2020 salaries relative to their production.

**Formula:**  $40\% \times \text{ERA- rank} + 20\% \times \text{HR/9 rank} + 20\% \times \text{BB\%} + 20\% \times \text{AVG} = \text{Total Points}$

### **Five Undervalued Players**

The top five most undervalued players that we came to the conclusion of based on our rankings of the players were the three batters Kole Calhoun, Jose Ramirez, and Tommy La Stella, and the two pitchers Jeremy Jeffress and Greg Holland.

Jose Ramirez is ranked 7<sup>th</sup> out of 134 batters and has a \$6,850,000 salary with a \$2,714,000 adjusted salary for the shortened season of 2020. He has a WAR of 3.4, a WRC+ of 164, an EV of 88.7, and a BB% of 12%. Within our rankings, Ramirez is one rank below Bryce Harper who earns a \$27,500,000 salary, and one rank above Paul Goldschmidt who earns a \$26,000,000 salary

Kole Calhoun is ranked 15<sup>th</sup> out of 134 batters and has a \$6,000,000 salary with a \$2,200,000 adjusted salary for the shortened season of 2020. He has a WAR of 1.8, WRC+ of 125, EV of 89.4, and BB% of 12%. Within our rankings, Calhoun is making significantly less than his similarly ranked peers George Springer and Salvador Perez that make \$21,000,000 and \$14,200,000 respectively.

Tommy La Stella is ranked 33<sup>rd</sup> out of 134 batters and has a \$3,250,000 salary with a \$646,776 adjusted salary for the shortened season of 2020. He has a WAR of 1.2, WRC+ of 129, EV of 88, and BB% of 12%. Within our rankings, La Stella ranks just below Jackie Bradley Jr. who makes \$11,000,000. Additionally, La Stella ranks 2 spots above Robinson Cano, who makes \$24,000,000.

Greg Holland is ranked 3<sup>rd</sup> out of 111 pitchers and has a \$1,250,000 promised salary. He has an HR/9 of 0.32, BB% of 0.06, AVG of 0.196, and ERA- of 42. The two players ranked above Holland are Liam Hendriks and Brad Hand who have a salary of \$5,300,000 and \$7,000,000 respectively.

Jeremy Jeffress is ranked 18<sup>th</sup> out of 111 pitchers and has only a \$850,000 salary with a \$314,815 adjusted salary for the shortened season of 2020. He has an HR/9 of 0.39, BB% of 0.14, AVG of 0.135, and an ERA- of 35. Within our rankings, Jeffress ranks one below Clayton Kershaw who makes \$31,000,000. Additionally, Jeffress ranks one above Jacob deGrom, who makes \$25,000,000.

## Hitters Rankings

Rank	Name	Total Score	2020 Salary
1	Freddie Freeman	130.8	\$22,559,375.00
2	Marcell Ozuna	127.6	\$18,000,000.00
3	Mike Trout	127.45	\$37,766,667.00
4	Brandon Belt	125.1	\$17,200,000.00
5	Anthony Rendon	123.8	\$26,071,428.00
6	Bryce Harper	121.35	\$27,538,462.00
7	Jose Ramirez	120.3	\$6,850,000.00
8	Paul Goldschmidt	118.8	\$26,000,000.00
9	Mookie Betts	117.5	\$32,000,000.00
10	Manny Machado	117.1	\$32,000,000.00
11	Jose Abreu	116.55	\$12,781,666.00
12	Wil Myers	113.8	\$22,500,000.00
13	Travis d'Arnaud	111.55	\$8,000,000.00
14	George Springer	109.9	\$21,000,000.00
15	Kole Calhoun	107.6	\$6,000,000.00
16	Salvador Perez	106.35	\$14,200,000.00
17	Yasmani Grandal	106.3	\$18,250,000.00
18	Jason Heyward	105.95	\$23,500,000.00
19	Justin Turner	105.65	\$20,000,000.00
20	Josh Donaldson	105.3	\$21,000,000.00
21	Xander Bogaerts	104.85	\$20,000,000.00
22	James McCann	103.65	\$5,400,000.00
23	J.T. Realmuto	102.35	\$10,000,000.00
24	Miguel Rojas	101.4	\$4,750,000.00
25	Kyle Seager	101	\$19,500,000.00
26	Robbie Grossman	100.35	\$3,725,000.00

27	Giancarlo Stanton	100.2	\$26,000,000.00
28	Aaron Hicks	97.65	\$10,785,714.00
29	Christian Yelich	97.65	\$12,500,000.00
30	Michael Brantley	96.7	\$16,000,000.00
31	Cesar Hernandez	96.65	\$6,343,000.00
32	Jackie Jr. Bradley	96.2	\$11,000,000.00
33	Tommy La Stella	96.2	\$3,250,000.00
34	Brad Miller	95.6	\$2,000,000.00
35	Robinson Cano	95.25	\$24,000,000.00
36	Mitch Moreland	90.7	\$525,146.00
37	Jose Iglesias	89.55	\$2,500,000.00
38	Christian Vazquez	89.5	\$4,200,000.00
39	Eugenio Suarez	89.45	\$9,535,714.00
40	A.J. Pollock	88.25	\$15,000,000.00
41	Eric Hosmer	86.35	\$21,000,000.00
42	Francisco Cervelli	86.3	\$2,000,000.00
43	Jedd Gyorko	85.7	\$1,000,000.00
44	Mike Moustakas	83.1	\$12,000,000.00
45	Lorenzo Cain	83.05	\$15,000,000.00
46	C.J. Cron	82.2	\$6,100,000.00
47	Jean Segura	81.95	\$14,850,000.00
48	Anthony Rizzo	81.4	\$16,500,000.00
49	Kevin Kiermaier	80.1	\$10,166,666.00
50	Brandon Crawford	79.6	\$15,200,000.00
51	Jurickson Profar	79.6	\$5,700,000.00
52	Miguel Cabrera	79.2	\$30,000,000.00
53	Martin Maldonado	78.4	\$3,500,000.00
54	Wilmer Flores	77.35	\$3,000,000.00

55	Jonathan Schoop	76.8	\$6,100,000.00
56	David Peralta	76.05	\$7,000,000.00
57	Didi Gregorius	74.5	\$14,000,000.00
58	Carlos Santana	74.35	\$20,833,334.00
59	Jason Castro	72.9	\$2,235,219.00
60	Nick Castellanos	72.85	\$16,000,000.00
61	Marcus Semien	72.3	\$13,000,000.00
62	Tyler Flowers	72.2	\$4,000,000.00
63	Nick Ahmed	72.05	\$6,375,000.00
64	Jason Kipnis	72	\$1,000,000.00
65	Kolten Wong	70.85	\$10,250,000.00
66	Starling Marte	70.75	\$11,500,000.00
67	Randal Grichuk	69.65	\$13,000,000.00
68	Jake Marisnick	68.9	\$3,312,500.00
69	Kevin Pillar	68.45	\$4,250,000.00
70	Derek Dietrich	67.45	\$575,000.00
71	Charlie Blackmon	65	\$21,333,333.00
72	Matt Carpenter	64.05	\$18,500,000.00
73	Yan Gomes	63.35	\$4,000,000.00
74	Leury Garcia	62.35	\$3,250,000.00
75	Andrew McCutchen	60.6	\$17,000,000.00
76	Avisail Garcia	60.25	\$7,250,000.00
77	Freddy Galvis	59.2	\$5,500,000.00
78	Joe Panik	58.95	\$2,850,000.00
79	Alex Avila	58.8	\$4,250,000.00
80	Joc Pederson	58.4	\$7,750,000.00
81	Andrelton Simmons	58.3	\$15,000,000.00
82	Tucker Barnhart	57.85	\$3,937,500.00

83	Nolan Arenado	56.85	\$35,025,000.00
84	Wilson Ramos	56.05	\$10,436,000.00
85	Josh Harrison	53.85	\$1,000,000.00
86	Marwin Gonzalez	53.2	\$10,500,000.00
87	Matt Joyce	52.3	\$1,500,000.00
88	Cameron Maybin	51.15	\$1,500,000.00
89	Dexter Fowler	48.6	\$16,500,000.00
90	Todd Frazier	48.15	\$754,552.00
91	Corey Dickerson	47.2	\$8,000,000.00
92	Jake Lamb	47.15	\$5,515,000.00
93	Starlin Castro	45.65	\$6,000,000.00
94	Chris Owings	45.15	\$563,500.00
95	Jose Altuve	44.65	\$29,000,000.00
96	Logan Forsythe	44.05	\$575,000.00
97	Khris Davis	43.7	\$16,750,000.00
98	Nick Markakis	42	\$4,000,000.00
99	Howie Kendrick	41.8	\$13,200,000.00
100	Howie Kendrick	40.8	\$4,000,000.00
101	Sandy Leon	40.7	\$2,000,000.00
102	Yadier Molina	40.3	\$20,000,000.00
103	Enrique Hernandez	39.6	\$5,900,000.00
104	J.D. Martinez	39.35	\$23,750,000.00
105	Mike Zunino	39	\$4,500,000.00
106	Josh Reddick	38.55	\$13,000,000.00
107	Ehire Adrianza	37.55	\$1,600,000.00
108	Roberto Perez	36.75	\$3,625,000.00
109	Stephen Vogt	36.4	\$2,500,000.00
110	Albert Pujols	35.45	\$29,000,000.00

111	Jonathan Villar	33.95	\$1,767,825.00
112	Elvis Andrus	33.7	\$15,250,000.00
113	Gregory Polanco	32.8	\$8,600,000.00
114	Matt Wieters	30.15	\$2,000,000.00
115	Brock Holt	29.7	\$563,500.00
116	Edwin Encarnacion	29.5	\$12,000,000.00
117	Sean Rodriguez	29.5	\$1,000,000.00
118	Adam Eaton	29.05	\$9,500,000.00
119	Eduardo Escobar	28.1	\$7,166,666.00
120	Yoenis Cespedes	27.55	\$6,000,000.00
121	Adeiny Hechavarria	23.55	\$1,000,000.00
122	Eric Sogard	23.3	\$500,000.00
123	Billy Hamilton	20.6	\$171,368.00
124	Rougned Odor	20.25	\$9,333,333.00
125	Austin Romine	20.15	\$4,150,000.00
126	Ender Inciarte	17.2	\$7,700,000.00
127	Bryan Holaday	16.95	\$575,000.00
128	Dee Gordon	14.3	\$10,000,000.00
129	Robinson Chirinos	13.2	\$1,239,622.00
130	Jarrold Dyson	12.3	\$2,000,000.00
131	Daniel Murphy	10.6	\$6,000,000.00
132	Jon Jay	9.75	\$1,500,000.00
133	Drew Butera	9.3	\$1,000,000.00
134	Chris Davis	8.25	\$21,065,362.00



## Pitchers Rankings

Rank	Name	Total Score	2020 Salary
1	Liam Hendriks	104.40	\$5,300,000.00
2	Brad Hand	101.40	\$7,000,000.00
3	Greg Holland	97.40	\$1,250,000.00
4	Darren O'Day	96.40	\$2,250,000.00
5	Alex Colome	95.60	\$10,532,500.00
6	Erasmus Ramirez	95.20	\$563,500.00
7	Trevor Bauer	93.80	\$17,500,000.00
8	Yu Darvish	92.00	\$22,000,000.00
9	Zack Britton	91.00	\$13,000,000.00
10	Oliver Perez	90.00	\$3,000,000.00
11	Jake McGee	89.80	\$2,759,260.00
12	Tyler Clippard	89.40	\$2,750,000.00
13	Dallas Keuchel	88.40	\$18,000,000.00
14	Aaron Loup	88.00	\$1,000,000.00
15	Trevor Rosenthal	87.60	\$2,000,000.00
16	Jake Diekman	87.40	\$2,750,000.00
17	Clayton Kershaw	87.20	\$31,000,000.00
18	Jeremy Jeffress	85.80	\$850,000.00
19	Jacob deGrom	85.00	\$25,500,000.00
20	Drew Pomeranz	84.80	\$6,000,000.00
21	Tony Watson	83.80	\$4,481,481.00
22	Yusmeiro Petit	83.20	\$5,500,000.00
23	Andrew Miller	82.80	\$11,537,000.00
24	Hyun-Jin Ryu	81.80	\$20,000,000.00
25	Shane Greene	80.20	\$6,250,000.00
26	Zack Wheeler	80.00	\$21,500,000.00
27	Michael Pineda	78.60	\$10,000,000.00

28	Mark Melancon	78.60	\$19,000,000.00
29	Kyle Hendricks	77.80	\$12,000,000.00
30	Joe Kelly	77.40	\$8,500,000.00
31	Gerrit Cole	74.60	\$36,000,000.00
32	Drew Smyly	73.40	\$4,093,000.00
33	Joakim Soria	72.60	\$8,500,000.00
34	Adam Wainwright	72.00	\$6,111,111.00
35	Anthony Bass	71.60	\$1,500,000.00
36	Ross Detwiler	71.20	\$800,000.00
37	Taijuan Walker	71.00	\$2,370,370.00
38	Pedro Baez	69.40	\$4,000,000.00
39	Tommy Hunter	69.20	\$850,000.00
40	Rich Hill	68.60	\$6,518,519.00
41	Lance Lynn	68.00	\$11,333,333.00
42	Blake Treinen	67.60	\$10,000,000.00
43	Aroldis Chapman	67.60	\$17,200,000.00
44	Brandon Kintzler	67.20	\$3,000,000.00
45	Carlos Carrasco	67.20	\$10,250,000.00
46	Kenley Jansen	66.40	\$18,000,000.00
47	Zack Greinke	66.40	\$35,000,000.00
48	Sonny Gray	64.80	\$10,203,667.00
49	Kevin Gausman	64.80	\$9,000,000.00
50	Ryan Pressly	63.00	\$8,750,000.00
51	Trevor Cahill	63.00	\$2,000,000.00
52	Nathan Eovaldi	61.20	\$17,000,000.00
53	Brett Anderson	59.80	\$5,148,000.00
54	Justin Wilson	59.60	\$5,093,000.00
55	Sergio Romo	59.40	\$4,750,000.00

56	Jeurys Familia	57.20	\$11,666,666.00
57	Trevor May	56.80	\$2,205,000.00
58	Jose Quintana	56.20	\$10,500,000.00
59	Garrett Richards	55.60	\$9,241,000.00
60	Max Scherzer	55.40	\$35,920,616.00
61	Brad Boxberger	54.80	\$1,000,000.00
62	Josh Tomlin	54.40	\$1,000,000.00
63	Craig Stammen	54.40	\$4,185,000.00
64	Will Smith	53.60	\$13,000,000.00
65	Charlie Morton	53.60	\$15,000,000.00
66	Alex Cobb	51.40	\$14,000,000.00
67	Matt Shoemaker	49.20	\$4,200,000.00
68	Will Harris	48.20	\$8,000,000.00
69	Martin Perez	48.20	\$5,700,000.00
70	Patrick Corbin	47.60	\$19,416,666.00
71	Jon Lester	47.20	\$20,000,000.00
72	Rick Porcello	47.00	\$10,000,000.00
73	Craig Kimbrel	45.80	\$16,000,000.00
74	T.J. McFarland	44.80	\$1,800,000.00
75	Mike Minor	43.40	\$9,833,334.00
76	Nick Vincent	43.20	\$1,000,000.00
77	Danny Duffy	43.20	\$15,250,000.00
78	Jake Arrieta	42.40	\$20,000,000.00
79	Johnny Cueto	41.20	\$21,000,000.00
80	Wade Miley	39.80	\$6,000,000.00
81	David Phelps	39.00	\$1,612,000.00
82	Jared Hughes	37.20	\$700,000.00
83	Tyler Chatwood	36.80	\$13,000,000.00
84	Brad Brach	36.40	\$975,000.00

85	Adam Ottavino	35.80	\$9,000,000.00
86	Jake Odorizzi	34.60	\$17,800,000.00
87	Daniel Hudson	33.20	\$5,148,148.00
88	Kyle Gibson	32.60	\$11,555,555.00
89	Dellin Betances	32.60	\$7,797,000.00
90	Steve Cishek	32.40	\$5,250,000.00
91	Madison Bumgarner	31.80	\$6,000,000.00
92	Michael Wacha	30.40	\$3,555,555.00
93	Anibal Sanchez	30.40	\$9,555,555.00
94	Derek Holland	29.20	\$1,000,000.00
95	Gio Gonzalez	27.60	\$4,500,000.00
96	Alex Wood	26.40	\$4,000,000.00
97	Jeff Samardzija	25.60	\$19,800,000.00
98	Jordan Lyles	24.00	\$8,000,000.00
99	Wade LeBlanc	23.60	\$800,000.00
100	Chase Anderson	22.60	\$8,500,000.00
101	Robbie Ray	20.60	\$9,430,000.00
102	Anthony DeSclafani	20.40	\$5,900,000.00
103	Jesse Chavez	20.00	\$4,000,000.00
104	Brandon Workman	19.20	\$3,500,000.00
105	Ivan Nova	19.00	\$1,593,000.00
106	Ian Kennedy	18.80	\$16,500,000.00
107	Tanner Roark	15.80	\$12,000,000.00
108	Hector Rondon	14.60	\$2,500,000.00
109	Matt Harvey	13.80	\$575,000.00
110	Carlos Martinez	11.00	\$11,700,000.00
111	Julio Teheran	10.20	\$9,000,000.00

