
A Sports Analysis of the Boston Red Sox

By: Owen Fitzhugh



Who/What are the Red Sox?

- The Red Sox are a Major League Baseball (MLB) team
 - The MLB is divided into two sections the National League (NL) and the American League (AL).
- They are located right outside the heart of Boston, Massachusetts
- Founded in 1901 as one of the American League's eight charter franchises, the Red Sox' home ballpark has been Fenway Park since 1912



Who was the best and most efficient hitter in the 2021 Boston Red Sox lineup?

FENWAY PARK

HOME OF THE

BOSTON

RED SOX



All the Red Sox Hitters in 2021

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Name	Games Played	Games Started	Batting Average	At Bat's	Runs	Hits	Doubles	Triples	Homeruns	Runs Batted In	Stolen Bases	Caught Stealing	Walks (BB)	Strikeouts	On-Base Percentage	Slugging Percentage	On-Base Slugging Percentage
2	Garrett Whitlock RP	46	0	0.5	2	0	1	0	0	0	0	0	0	0	1	0.5	0.5	1
3	Jose Iglesias 2B	23	17	0.356	59	8	21	4	1	1	7	0	0	3	9	0.406	0.508	0.914
4	Taylor Motter 2B	3	1	0.333	6	3	2	1	1	0	1	0	0	1	2	0.429	0.833	1.262
5	Connor Wong C	6	3	0.308	13	3	4	1	1	0	1	0	0	1	7	0.357	0.538	0.895
6	Xander Bogaerts SS	144	143	0.295	529	90	156	34	1	23	79	5	1	62	113	0.37	0.493	0.863
7	Kyle Schwarber LF	41	37	0.291	134	34	39	10	0	7	18	0	0	33	39	0.435	0.522	0.957
8	Alex Verdugo LF	146	137	0.289	544	88	157	32	2	13	63	6	2	51	96	0.351	0.426	0.777
9	Kevin Plawecki C	64	45	0.287	157	15	45	7	0	3	15	0	0	12	26	0.349	0.389	0.738
10	J.D. Martinez DH	148	148	0.286	570	92	163	42	3	28	99	0	0	55	150	0.349	0.518	0.867
11	Rafael Devers 3B	156	155	0.279	591	101	165	37	1	38	113	5	5	62	143	0.352	0.538	0.89
12	Christian Arroyo 2B	57	44	0.262	164	22	43	12	0	6	25	1	0	8	44	0.324	0.445	0.769
13	Hunter Renfroe RF	144	140	0.259	521	89	135	33	0	31	96	1	2	44	130	0.315	0.501	0.816
14	Christian Vazquez C	138	122	0.258	458	51	118	23	1	6	49	8	4	33	84	0.308	0.352	0.66
15	Enrique Hernandez CF	134	130	0.25	508	84	127	35	3	20	60	1	0	61	110	0.337	0.449	0.786
16	Bobby Dalbec 1B	133	115	0.24	417	50	100	21	5	25	78	2	0	28	156	0.298	0.494	0.792
17	Travis Shaw 1B	28	7	0.238	42	6	10	3	0	3	11	0	0	5	17	0.319	0.524	0.843
18	Jaren Duran CF	33	26	0.215	107	17	23	3	2	2	10	2	1	4	40	0.241	0.336	0.577
19	Marwin Gonzalez 3B	77	67	0.202	242	25	49	14	0	2	20	3	2	19	70	0.281	0.285	0.566
20	Garrett Richards SP	40	22	0.2	5	0	1	1	0	0	1	0	0	0	2	0.2	0.4	0.6
21	Michael Chavis 2B	31	19	0.19	79	12	15	4	1	2	6	1	1	1	32	0.207	0.342	0.549
22	Franchy Cordero LF	48	37	0.189	127	12	24	6	0	1	9	1	1	8	51	0.237	0.26	0.497
23	Jonathan Arauz SS	28	19	0.185	65	9	12	3	0	3	8	0	0	8	15	0.274	0.369	0.643
24	Danny Santana 1B	38	29	0.181	116	15	21	2	1	5	14	4	2	10	30	0.252	0.345	0.597
25	Jack Lopez 2B	7	5	0.154	13	2	2	2	0	0	0	0	0	1	6	0.214	0.308	0.522
26	Yairo Munoz 2B	5	2	0.091	11	0	1	0	0	0	0	0	0	0	2	0.091	0.091	0.182
27	Tanner Houck SP	18	13	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0
28	Nathan Eovaldi SP	32	32	0	2	1	0	0	0	0	0	0	0	1	1	0.333	0	0.333
29	Martin Perez SP	36	22	0	3	0	0	0	0	0	0	0	0	0	3	0	0	0
30	Nick Pivetta SP	31	30	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0
31	Eduardo Rodriguez SP	32	31	0	6	0	0	0	0	0	0	0	0	0	3	0	0	0
32	Chris Sale SP	9	9	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0
33	Austin Davis RP	19	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0

Calculations

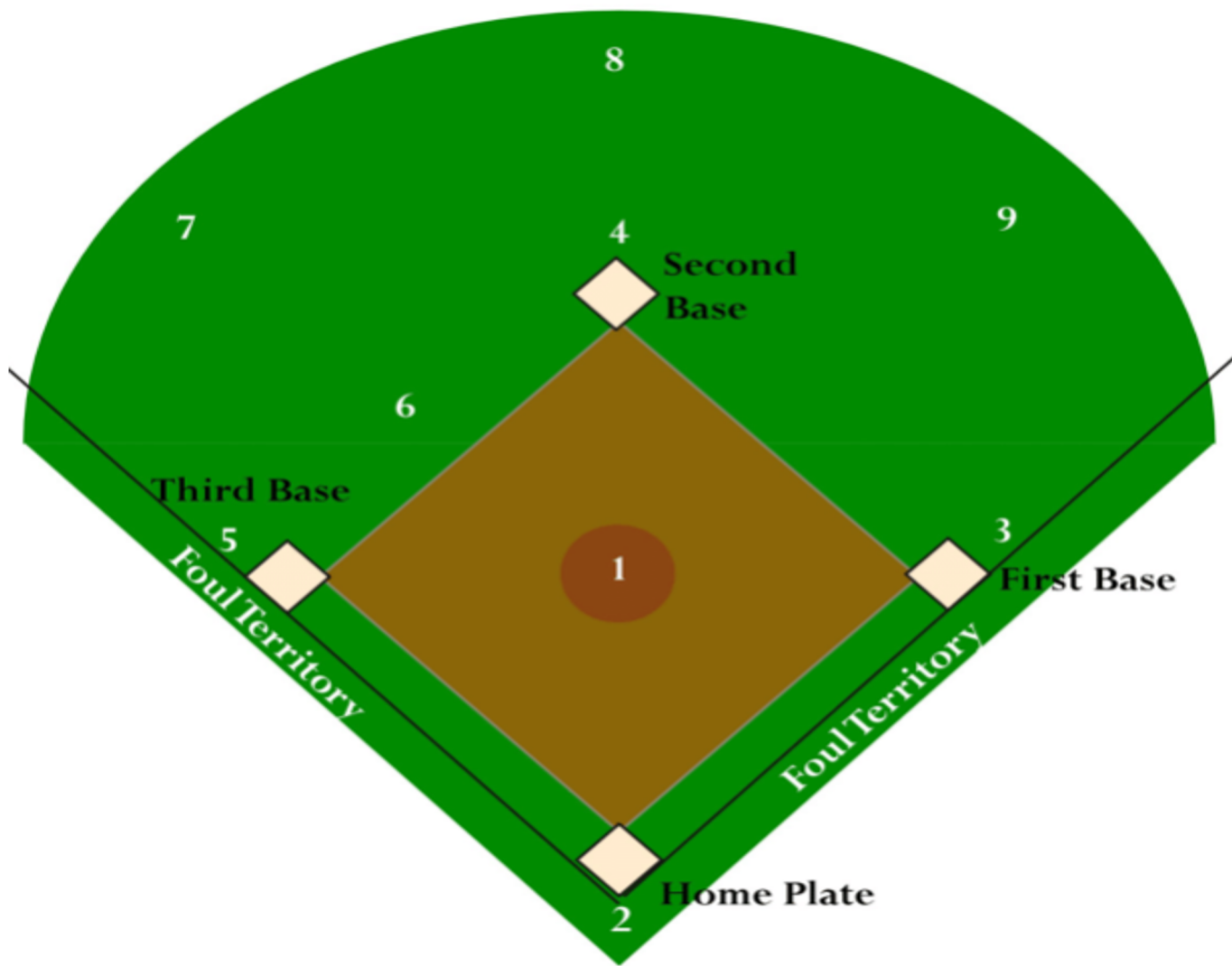
In order to determine which hitters are the most proficient these categories below will be determined.

- Runs batted in (RBI's)
 - Batting Average = $(\text{Hits}) / (\text{At-Bats})$
 - Slugging % = $(\text{Singles} * (1) + \text{Doubles} * (2) + \text{Triples} * (3) + \text{Homeruns} * (4)) / (\text{At-Bats})$
 - On-Base % = $(\text{Hits} + \text{Walks}) / (\text{At-Bats} + \text{Walks})$
 - On-Base + Slugging % = $(\text{On-Base \%}) + (\text{Slugging \%})$
-

```
1 import numpy as np
2 import csv
3
4 fhandle = open('red.csv')
5 data = list(csv.reader(fhandle))
6
7 data_arr = np.array(data)
8 transposed_arr = np.transpose(data_arr)
9 #print(transposed_arr)
10 names = transposed_arr[0].astype(str)
11 games_played = transposed_arr[1].astype(int)
12 games_started = transposed_arr[2].astype(float)
13 batting_av = transposed_arr[3].astype(float)
14 at_bats = transposed_arr[4].astype(int)
15 runs = transposed_arr[5].astype(int)
16 hits = transposed_arr[6].astype(int)
17 doubles = transposed_arr[7].astype(int)
18 triples = transposed_arr[8].astype(int)
19 hrs = transposed_arr[9].astype(int)
20 runs_batted_in = transposed_arr[10].astype(int)
21 stolen_bases = transposed_arr[11].astype(int)
22 caught_stealing = transposed_arr[12].astype(int)
23 walks = transposed_arr[13].astype(int)
24 strikeouts = transposed_arr[14].astype(int)
25 on_base_percent = transposed_arr[15].astype(float)
26 slugging_percent = transposed_arr[16].astype(float)
27 on_base_slug = transposed_arr[17].astype(float)
28
```

Category One (RBI's)

- My goal was to figure out which Red Sox hitter had the most Runs Batted In (RBI's)
 - This is a crucial aspect of a hitters game and can be affected by other people's inability to get on base, as well as where a player hits in a lineup
 - However it is still an important stat because it keeps track of the total amount of batters a batter scores
-



Category One (RBI's)

- My goal was to figure out which Red Sox hitter had the most Runs Batted In (RBI's)
 - This is a crucial aspect of a hitters game and can be affected by other people's inability to get on base, as well as where a player hits in a lineup
 - However it is still an important stat because it keeps track of the total amount of batters a batter scores

```
#####Task 1#####  
#print(f'The most RBI/s: {np.max(runs_batted_in)}')  
print(f'The players with the most RBI/s: {names[np.where(np.max(runs_batted_in) == runs_batted_in)]}')  
rbi_king = np.where(np.max(runs_batted_in) == runs_batted_in)
```

```
File Edit View Search Terminal Help  
[ofitzhug@csr04 ~]$ cd '/students/home/ofitzhug/Documents/PresentationCode'  
[ofitzhug@csr04 PresentationCode]$ python3 red_sox.py  
The players with the most RBI/s: ['Rafael Devers_3B']
```

Category 2 (Batting Average)

My goal was to find the most efficient hitters per plate appearance.

- Since batting average is $(\text{Hits})/(\text{At-Bats})$ it is tracking the percentage of how many times a batter will get a hit every at bat
- This is a useful stat because it is only affected by the individual and tracks the likelihood of a batter getting a hit

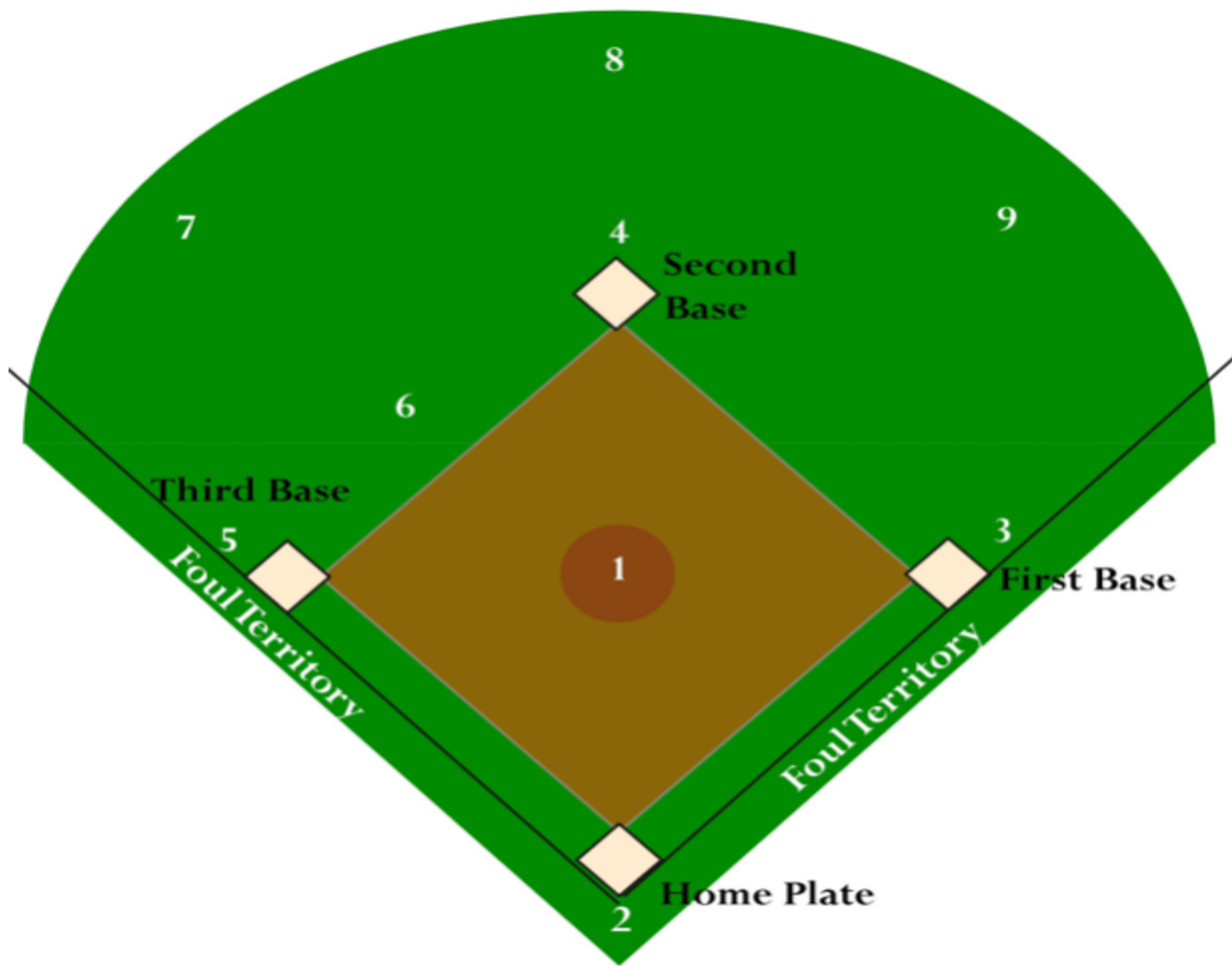
```
33 #####Task 2#####
34 print(f'The Players with a Batting Average of at least .250 (min 75 at bats) :
35 {np.intersect1d(names[np.where(at_bats >= 75)], names[np.where(batting_av>= .250)])}')
36 batting_av250 = np.intersect1d(np.where(at_bats >= 75), np.where(batting_av>= .250))
```

```
The Players with a Batting Average of at least .250 (min 75 at bats) : ['Alex Verdugo_LF' 'Christian Arroyo_2B' 'Christian Vazquez_C'
'Enrique Hernandez_CF' 'Hunter Renfroe_RF' 'J.D. Martinez_DH'
'Kevin Plawecki_C' 'Kyle Schwarber_LF' 'Rafael Devers_3B'
'Xander Bogaerts_SS']
```

Category 3 (Slugging Percentage)

$$\text{Slugging \%} = (1B)*(1) + (2B)*(2) + (3B)*(3) + (HR)*(4) / (ABs)$$

- Slugging percentage is just the average number of bases a player gets every at bat.
 - This is also similar to Batting Average but instead assigns a number to every base earned and thus is
-



Category 3 (Slugging Percentage)

Slugging % = $(1B) \cdot (1) + (2B) \cdot (2) + (3B) \cdot (3) + (HR) \cdot (4) / (ABs)$

- Slugging percentage is just the average number of bases a player gets every at bat.
- This is also similar to Batting Average but instead assigns a number to every base earned and thus is equally important

```
37 #####Task 3#####
38 print(f'The Players with a Slugging Percent of at least .5 (min 75 at bats) :
39 {np.intersect1d(names[np.where(at_bats >= 75)], names[np.where(slugging_percent >= .5)])}')
40 sluggers = np.intersect1d(np.where(at_bats >= 75), np.where(slugging_percent >= .5))
```

```
The Players with a Slugging Percent of at least .5 (min 75 at bats) : ['Hunter Renfroe_RF' 'J.D. Martinez_DH' 'Kyle Schwarber_LF'
'Rafael Devers_3B']
```

Category 4 (On-Base Percentage)

On-Base % = (Hits + Walks)/(At-Bats + Walks)

- Tracks how often a batter gets on base whether it is a hit or if it is a walk
- Also an individual statistic and isn't dependent on any one else
- Important because it matters how often a hitter can get on base, but is the only statistic here that takes into account walks

```
41 #####Task 4#####
42 print(f'The Players with a On-Base Percent of at least .350 (min 75 at bats) :
43 {np.intersect1d(names[np.where(at_bats >= 75)], names[np.where(on_base_percent>= .350)])}')
44 onbasers = np.intersect1d(np.where(at_bats >= 75), np.where(on_base_percent>= .350))
```

```
The Players with a On-Base Percent of at least .350 (min 75 at bats) : ['Alex Verdugo_LF' 'Kyle Schwarber_LF' 'Rafael Devers_3B'
'Xander Bogaerts_SS']
```

Category 5 (On-Base Slugging Percent)

On-Base + Slugging % = (On-Base %)
+ (Slugging %)

- Takes into account both slugging and on base percent and is just the addition of both of the statistics, which just simplifies two categories
- Shows another category for efficiency of a hitter

```
45 #####Task 5#####
46
47 print(f'The Players with a On-Base Plus Slugging Percent of at least .850 (min 75 at bats) :
48 {np.intersect1d(names[np.where(at_bats >= 75)], names[np.where(on_base_slug >= .850)])}')
49 obp_sluggers = np.intersect1d(np.where(at_bats >= 75), np.where(on_base_slug >= .850))
```

```
The Players with a On-Base Plus Slugging Percent of at least .850 (min 75 at bats) : ['J.D. Martinez_DH' 'Kyle Schwarber_LF' 'Rafael Devers_3B'
'Xander Bogaerts_SS']
```

Final Findings

```
50 #####Task 6#####
51 first = (np.intersect1d(obp_sluggers,onbasers))
52 second = (np.intersect1d(first,sluggers))
53 third = (np.intersect1d(second,batting_av250))
54 fourth = (np.intersect1d(third,rbi_king))
55 print(names[fourth])
```

By these findings the only common hitter to satisfy all of the categories is third baseman Rafael Devers



Sources

<https://www.baseballamerica.com/teams/1003/boston-red-sox/stats/>

<https://bosoxinjection.com/2021/05/07/red-sox-rafael-devers-crushing-baseball/>

<https://www.bostonherald.com/2021/04/28/red-sox-notebook-rafael-devers-provides-another-reminder-of-his-clutchness/>

<https://theathletic.com/2689218/2021/07/04/rafael-devers-vastly-improved-defense-putting-him-in-mvp-conversation/>
