Analyzing MLB Expected Batting Average: Insights from 2023 and Predictions for 2024

Harry Svendsgaard

March 26, 2024

Introduction

As the 2024 MLB season's Opening Day draws near, my first goal is to assess the accuracy of my model, which relies on 2022 data, by comparing it with the statistics from the 2023 season. Additionally, I aim to implement significant changes to the code and conduct a similar analysis using the latest data from the conclusion of the 2023 season.

Previously, my focus wasn't on predicting upcoming season statistics. Instead, I centered my work on analyzing the numbers from the 2022 season to identify players who performed better or worse than expected at the plate. Towards the end of my research, I did examine the first two months of data from the 2023 season to evaluate how accurate my model was – essentially, were players who outperformed in the previous season returning to average, and vice versa. This helped gauge the effectiveness of the model. Based on the initial data from the start of the 2023 season, it seemed like the model was performing admirably. Players with projected higher batting averages than their actual averages were showing notable improvements in their performance at the beginning of the 2023 season. However, I recognized the need to reassess the model after the season concluded to ensure its reliability.

This time around, I'm thrilled to reveal the adjustments I've implemented to enhance the model's accuracy. I'll also be introducing predictions based on performance and age trends in the MLB. By reading this article, you'll easily identify which players exceeded expectations and which ones fell short last season, along with those who could have a standout performance in the upcoming 2024 season.

A Look at Last Year's Data

Last year, I waited until late May to publish the article. This gave me the chance to examine the data from the beginning of the 2023 season and compare it with the model projections based on the previous year's data. While this did offer some insights into the model's accuracy, it was limited by the fact that only a quarter of the season had passed. Now, with a full season's worth of data available, we can expect to draw more conclusive findings regarding the model's accuracy.

If you've read last season's article, you might recognize the tables below. These are the same ones I used in the final section, where I analyzed how the model's predictions stacked up against the actual statistics from players in the early part of the 2023 season. I've now updated the last two columns of these tables to reflect the finalized stats from the entire 2023 season.

Last year, I decided to dig deeper than just the overall league stats. Specifically, I wanted to focus on players under the age of 25 since they're likely still developing and haven't reached their peak performance yet. It feels more crucial to predict the future success of a 23-year-old compared to that of a 33-year-old. The first table below examines players aged 25 or younger who had the largest difference between their expected

batting average (xBA) and their actual batting average (True BA). To keep it simple, we can think of this as the list of underperformers because these players were projected to have significantly more hits than they actually achieved in the 2022 season. Let's see if these trends continued into the 2023 season.

Player	2022 PA's	2022 Batting Avg	2022 xBA	2023 PA's	2023 Batting Avg
Josh Lowe	198	0.221	0.298	501	0.292
Cristian Pache	260	0.166	0.236	95	0.238
Taylor Walls	466	0.172	0.241	349	0.201
Trent Grisham	524	0.184	0.233	555	0.198
Geraldo Perdomo	500	0.195	0.242	495	0.246
Spencer Torkelson	404	0.203	0.247	684	0.233
Abraham Toro	352	0.185	0.228	414(AAA)	0.291(AAA)
Jose Barrero	174	0.152	0.194	149	0.218
Diego Castillo	283	0.206	0.247	556 (AAA)	0.313 (AAA)
Nolan Gorman	313	0.226	0.262	464	0.236

In this instance, the model proved to be reliable, accurately predicting which players would improve their performance in the following season. For example, in 2022, Josh Lowe had a batting average of just 0.221 over 198 plate appearances. However, according to the model I developed, his expected batting average was 0.298. In 2023, Lowe significantly boosted his performance, batting 0.292 over 501 plate appearances. Similarly, all the other players on this list also showed improvements in their batting compared to the 2022 season. Despite Abraham Toro and Diego Castillo spending most of the season in the Minor Leagues, I felt it was important to include their stats from the season as well.

Next up are the players under the age of 25 who ended up with notably more hits than expected. We'll refer to this list as the overperformers list. Based on the stats, these players achieved a higher batting average than what was predicted for them.

Player	2022 PA's	2022 Batting Avg	2022 xBA	2023 PA's	2023 Batting Avg
Mickey Moniak	112	0.170	0.042	323	0.280
Taylor Trammell	117	0.196	0.087	56	0.130
Richie Palacios	123	0.232	0.138	102	0.258
Bubba Thompson	181	0.265	0.193	60	0.170
Spencer Steer	108	0.211	0.141	665	0.271
Peyton Burdick	102	0.207	0.143	37	0.182
Michael Toglia	120	0.216	0.154	152	0.163
Vinnie Pasquantino	298	0.295	0.242	260	0.247
Luis Arraez	603	0.316	0.265	617	0.354
Dermis Garcia	125	0.207	0.156	198 (AAA)	0.263 (AAA)

Examining these results made me realize that there's a need for some adjustments if we aim to predict a player's future statistics accurately. While there are a few players who experienced significant drops in their batting average, like Vinnie Pasquantino, the accuracy doesn't seem as high as in the top ten table mentioned earlier. Pasquantino's batting average was almost identical to his expected batting average from the previous season. Ultimately, this is the main objective, especially for young players in the league: trying to forecast their future performance.

Young players generally improve from season to season, which is why I think there are some inconsistencies in the table above. The expected batting average was solely calculated based on the previous season's data and didn't consider player progression. I'll address this factor later on.

Following up on that, I aimed to assess how well the model performed with players over the age of 25. If my theory regarding progression holds true, the upcoming table should provide more reliable outcomes. In other

words, the 2023 statistics of the players listed should closely match their expected batting average from the previous season.

Player	2022 PA's	2022 Batting Avg	2022 xBA	2023 PA's	2023 Batting Avg
Alex Call	131	0.237	0.124	439	0.200
Matt Carpenter	154	0.305	0.193	237	0.176
Eli White	117	0.200	0.095	17	0.071
Kris Bryant	181	0.306	0.204	335	0.233
Garrett Stubbs	121	0.264	0.171	125	0.204
Stuart Fairchild	110	0.247	0.159	255	0.228
Oscar Mercado	128	0.207	0.132	32	0.290

We're definitely noticing much more consistent outcomes in this table. The only player who actually improved their batting average from the previous season was Oscar Mercado, and he only had 32 plate appearances that year. This shows that older players were more likely to follow the pattern of the model, as opposed to the younger players.

New Features

Model

To improve the accuracy of the model, I made a few minor adjustments. Previously, I relied on stepwise regression, allowing the computer to determine the best model for the data. However, this time, I experimented with incorporating additional factors. Last year, hits were estimated based on plate appearances, average launch angle, and the percentage of balls hit in the sweet spot. This year, I've included average exit velocity, strikeout percentage, and stolen bases per plate appearance. These three new factors aim to give a boost to players who hit the ball harder, strike out less, and utilize speed to their advantage in getting on base.

A player like Luis Arraez, who appeared on the overperformers list from the 2022 season, but hit for an incredible average (0.354) in 2023 could benefit from this change.

Rk.	Player	Year	Age	AB	PA	H	K%
1	Brantley Jr., Michael	2023	36	54	57	15	3.5
2	Arraez, Luis	2023	26	574	617	203	5.5
3	Vargas, Ildemaro	2023	31	262	286	66	7.0
4	Clement, Ernie	2023	27	50	52	19	7.7
5	Madrigal, Nick	2023	26	270	294	71	8.2
6	Fletcher, David	2023	29	89	97	22	9.3
7	Kemp, Tony	2023	31	359	419	75	9.5
8	McNeil, Jeff	2023	31	585	648	158	10.0
9	Ruiz, Keibert	2023	24	523	562	136	10.3
10	Kwan, Steven	2023	25	638	718	171	10.4

Figure 1: Lowest Strikeout Percentage 2023

Arraez had the lowest strikeout percentage in the entire league for players that qualified for the model. Michael Brantley did have a lower strikeout percentage, but this was only accomplished in 57 total plate appearances for the season, as compared to Arraez's 617.

Players who strike out less often tend to have higher batting averages. When you make contact with the ball, there's a chance for anything to happen. Luis Arraez showcased this last season by achieving the highest batting average in a full season since Josh Hamilton in 2010.

Data

A more minor change that was made included a change in the parameters for a player to be qualified for the model. Previously, I set 100 plate appearances as the minimum, but I have increased the minumum value to 200 for this past season's data.

With the minimum requirement raised, it's less likely to see any unusual cases. I observed this particularly with the young players. Among the names listed in the previous section are Mickey Moniak, Spencer Steer, and Luis Arraez. Steer and Moniak significantly increased their plate appearances from the 2022 season to the 2023 season, resulting in substantially improved batting averages for both. Luis Arraez was one of two players on that list to meet the 200 plate appearance minimum in the 2022 season, the other being Vinnie Pasquantino. Interestingly, Pasquantino almost matched his expected batting average from the previous year in the 2023 season.

Implementing this approach will not only decrease errors but also showcase players who are more developed than others. Expanding on that, players who are more developed are likely to align better with the model's predictions, as evidenced by Vinnie Pasquantino. With that said, forecasting a player's future success will become much simpler, with reduced margin for error.

2023 Season Results

In this section, I'll present the top ten and bottom ten players in the league based on (xHits - True Hits) divided by At Bats (AB). It's more meaningful to examine this value per At Bat, rather than just looking at the difference between expected and true hits. This way, we can identify players whose batting average deviated the most from what was expected versus what actually happened.

Under-Performers

The following table is the top ten players, age 25 or younger, in $(xHits - TrueHits)/AB$. Simply, it is the players that were expected a significantly higher batting average than their true average from the 2023 season.

Top 10 Players Under 25 in (Expected Hits - Hits) per AB

Last	First	Age	AB	PA	H	BA	xHits	xHits - H	(xHits - H)/AB	xBA	BAdiff
Matos	Luis	21	228	253	57	0.250	68.53492	11.534924	0.0505918	0.3005918	0.0505918
Volpe	Anthony	22	541	601	113	0.209	136.98577	23.985771	0.0443360	0.2532084	0.0442084
Burleson	Alec	24	315	347	77	0.244	90.89309	13.893085	0.0441050	0.2885495	0.0445495
Pasquantino	Vinnie	25	231	260	57	0.247	66.16372	9.163722	0.0396698	0.2864230	0.0394230
Turang	Brice	23	404	448	88	0.218	102.01502	14.015019	0.0346906	0.2525124	0.0345124
Carlson	Dylan	24	219	255	48	0.219	55.22283	7.222833	0.0329810	0.2521591	0.0331591
Baty	Brett	23	353	389	75	0.212	86.61002	11.610021	0.0328896	0.2453542	0.0333542
Arias	Gabriel	23	315	345	66	0.210	74.86843	8.868428	0.0281537	0.2376775	0.0276775
Wiemer Jr.	Joey	24	367	410	75	0.204	84.94364	9.943635	0.0270944	0.2314540	0.0274540
Massey	Michael	25	428	461	98	0.229	109.24853	11.248534	0.0262816	0.2552536	0.0262536

Figure 2: 2023 Top 10 Players in $xBA - TrueBA$ (25 or Younger)

This list features several names that are still regarded as promising prospects in the league. It's worth noting that Luis Matos, who topped the list, is the youngest player here. With already over 250 plate appearances as a 21-year-old, we can expect his numbers to keep climbing. Another standout from this list is Anthony Volpe. In the 2023 season, he had over 600 plate appearances, significantly more than anyone else on the list. As he enters his 23-year-old season, Volpe should continue to receive ample opportunities. Although his true batting average from last season wasn't eye-opening at 0.209, he was projected to hit at a 0.253 clip, which is notably higher than the average 22-year-old.

Instead of shifting focus to players under 25 who overperformed, let's continue examining the underperformers for now and explore the league-wide leaders in this category.

Top 10 Players in (Expected Hits - Hits) per AB

Last	First	Age	AB	PA	H	BA	xHits	xHits - H	(xHits - H)/AB	xBA	BAdiff
Smith	Pavin	27	191	228	36	0.188	49.25191	13.25191	0.0693817	0.2578634	0.0698634
Kemp	Tony	31	359	419	75	0.209	97.78865	22.78865	0.0634781	0.2723918	0.0633918
Wendle	Joey	33	297	318	63	0.212	78.71524	15.71524	0.0529133	0.2650345	0.0530345
Schwarber	Kyle	30	585	720	115	0.197	145.61377	30.61377	0.0523312	0.2489124	0.0519124
Matos	Luis	21	228	253	57	0.250	68.53492	11.53492	0.0505918	0.3005918	0.0505918
Refsnyder	Rob	32	202	243	50	0.248	59.88397	9.88397	0.0489305	0.2964553	0.0484553
d'Arnaud	Travis	34	267	292	60	0.225	72.79133	12.79133	0.0479076	0.2726267	0.0476267
Grisham	Trent	26	469	555	93	0.198	114.34020	21.34020	0.0455015	0.2437957	0.0457957
Lopez	Nicky	28	225	262	52	0.231	62.20579	10.20579	0.0453590	0.2764702	0.0454702
Volpe	Anthony	22	541	601	113	0.209	136.98577	23.98577	0.0443360	0.2532084	0.0442084

Figure 3: 2023 Top 10 Players in $xBA - TrueBA$ (All)

While Pavin Smith tops this list, I'd like to highlight Tony Kemp first because he had nearly double the plate

appearances during the season. Kemp finished with a batting average of 0.209, but he was expected to hit around 0.270. This marked his lowest batting average in his career so far. Kemp has now joined the Orioles, who boast a strong lineup heading into the season. Considering Kemp's age and the Orioles' talented young players, I anticipate his plate appearances will decrease significantly. However, he could still prove to be a valuable asset to the team if one of their players happens to go down with an injury.

One standout name on this list is Kyle Schwarber, who ranked second in the league in home runs. Like Kemp, Schwarber posted his lowest batting average of his career, excluding the shortened COVID season. Despite not being known for his speed and having a relatively high strikeout rate, Schwarber still made it onto this list. This suggests that Schwarber excels in at least a few areas such as average exit velocity, average launch angle, and sweet spot percentage, making him one of the top performers in the league in those aspects. This can be confirmed using the data from Baseball Savant, which has Schwarber 11th in the league in exit velocity and 10th in launch angle.

Over-Performers

Moving on from the players that under-performed in the 2023 season, let's take a look at the players that over-performed. The names on the following lists hit for a higher true batting average than their expected batting average in the 2023 season.

As we discovered earlier, the model appears to produce more precise outcomes with older players who outperformed expectations. With that in mind, the table below showcases the bottom ten players in the league based on the difference between expected batting average (xBA) and true batting average (True BA).

Bottom 10 Players in (Expected Hits - Hits) per AB

Last	First	Age	AB	PA	H	BA	xHits	xHits - H	(xHits - H)/AB	xBA	BAdiff
Triolo	Jared	25	181	209	54	0.298	41.08640	-12.91360	-0.0713458	0.2269967	-0.0710033
Altuve	Jose	33	360	410	112	0.311	87.48398	-24.51602	-0.0681000	0.2430111	-0.0679889
McGuire	Reese	28	187	206	50	0.267	37.39629	-12.60371	-0.0673995	0.1999802	-0.0670198
Lewis	Royce	24	217	239	67	0.309	52.95812	-14.04188	-0.0647091	0.2440466	-0.0649534
Hicks	Aaron	33	269	312	68	0.253	52.20291	-15.79709	-0.0587252	0.1940629	-0.0589371
Moniak	Mickey	25	311	323	87	0.280	69.14063	-17.85937	-0.0574256	0.2223171	-0.0576829
Ford	Mike	30	219	251	50	0.228	37.66849	-12.33151	-0.0563083	0.1720022	-0.0559978
Ramirez	Harold	28	400	434	125	0.313	103.23295	-21.76705	-0.0544176	0.2580824	-0.0549176
Arraez	Luis	26	574	617	203	0.354	171.92223	-31.07777	-0.0541425	0.2995161	-0.0544839
Duvall	Adam	34	320	353	79	0.247	62.09756	-16.90244	-0.0528201	0.1940549	-0.0529451

Figure 4: 2023 Bottom 10 Players in $xBA - TrueBA$ (All)

One notable name on this list is Jose Altuve. He's become one of the most recognized figures in baseball, thanks to both his individual achievements and his team's success over the years. Although Altuve managed to maintain a batting average of 0.311 last season, it seems that his advanced metrics indicated a significantly lower expected batting average. We'll be monitoring Altuve closely as we head into the 2024 season to see if his actual performance sees any decline.

Luis Arraez managed to find his way onto this list again. I believe that this mostly stems from him having such a standout batting average last season. Hitting 0.354 is not a common occurrence, and I am inclined

to believe that the model is attempting to overcorrect for his case. I look at Arraez's case as a slight outlier, simply because of the incredible season he had last year.

Finally, we will look at the players under the age of 25 that were at the bottom of the league, in respect to expected batting average vs. true batting average.

Bottom 10 Players Under 25 in (Expected Hits - Hits) per AB

Last	First	Age	AB	PA	H	BA	xHits	xHits - H	(xHits - H)/AB	xBA	BAdiff
Triolo	Jared	25	181	209	54	0.298	41.08640	-12.913596	-0.0713458	0.2269967	-0.0710033
Lewis	Royce	24	217	239	67	0.309	52.95812	-14.041879	-0.0647091	0.2440466	-0.0649534
Moniak	Mickey	25	311	323	87	0.280	69.14063	-17.859375	-0.0574256	0.2223171	-0.0576829
Wallner	Matt	25	213	254	53	0.249	42.34047	-10.659529	-0.0500447	0.1987816	-0.0502184
Montero	Elehuris	24	284	307	69	0.243	55.26790	-13.732102	-0.0483525	0.1946053	-0.0483947
Jones	Nolan	25	367	424	109	0.297	94.10823	-14.891768	-0.0405770	0.2564257	-0.0405743
Frelick	Sal	23	191	223	47	0.246	39.48087	-7.519131	-0.0393672	0.2067061	-0.0392939
Peguero	Liover	22	198	213	47	0.237	39.70766	-7.292345	-0.0368300	0.2005437	-0.0364563
Sabol	Blake	25	310	344	73	0.235	62.36566	-10.634343	-0.0343043	0.2011795	-0.0338205
Encarnacion-Strand	Christian	23	222	241	60	0.270	52.55420	-7.445799	-0.0335396	0.2367306	-0.0332694

Figure 5: 2023 Bottom 10 Players in $xBA - TrueBA$ (25 or Younger)

Based on our earlier findings, it's tougher to rely on expected batting average from the previous season as a guide for young players' true batting average in the next season. Once more, I believe the increase in minimum plate appearances will help us get results that better match up for the following season.
Nolan Jones saw the most plate appearances on this list, at 424 for the season. He batted nearly 0.300 on the year, but was expected just shy of 0.260. With a healthy amount of plate appearances on the year, and Jones being at the top age last season for players qualified on this list (25) I expect he will see a dip in his average this coming season.

Sal Frelick and Liover Peguero are among the youngest players on this list, and they had just enough plate appearances to qualify for the model. Considering the typical development trajectory for young players and the fact that they'll likely get more chances to play, it wouldn't be surprising if these players exceed their stats from last season, even though they overperformed then.

Predicting Future Success

Expanding on the end of the previous section, our goal is to make reasonably accurate predictions for upcoming seasons, particularly for young players. Teams want to give their promising young talents plenty of chances to shine and develop. In this section, I've crafted a simple age curve to apply to players and devised a formula that combines true batting average and expected batting average from the previous season. The goal is to produce predictions that closely match how players actually perform in the 2024 season.

Formulas

There are a few factors that need to be addressed when predicting the future success of players. The first of which being the age curve. Young players are likely to progress year after year, to a certain point. On the other end, older players are likely to start seeing a decline in their numbers at a certain point in their career.

Though there are some oddball cases of players peaking very early or late in their career, it is typically true that a player's "prime" is in their mid to late 20's.

To create a variable that takes the age curve into account, I first tallied the amount of hits and at-bats for each specific age, using data from both the 2022 and 2023 season. This allowed me to get a batting average for each age. I used a log function to normalize the total hits for each age and multiplied the value by that age's batting average. This created a weight that I could use for each age. The results of this computation are shown in the table below. The values from this table will be utilized in the formulas to come.

Age	Total AB	Total Hits	BA	$normHits$	$normHits * BA$
42	450	118	0.262	4.77	1.25
41	448	105	0.234	4.65	1.09
40	334	86	0.257	4.45	1.15
39	1265	287	0.227	5.67	1.28
38	1926	442	0.229	6.09	1.40
37	1750	394	0.225	5.98	1.35
36	3497	804	0.230	6.69	1.54
35	6543	1642	0.251	7.40	1.86
34	9237	2298	0.249	7.74	1.93
33	13746	3338	0.243	8.11	1.97
32	20086	4879	0.243	8.49	2.06
31	23855	5823	0.244	8.67	2.12
30	27935	6909	0.247	8.84	2.19
29	28358	7022	0.248	8.86	2.19
28	35726	8713	0.244	9.07	2.21
27	34246	8316	0.243	9.03	2.19
26	29394	7357	0.250	8.90	2.23
25	32246	8108	0.251	9.00	2.26
24	24154	5946	0.246	8.69	2.14
23	18010	4311	0.239	8.37	2.00
22	9259	2275	0.246	7.73	1.90
21	4871	1242	0.255	7.12	1.82
20	169	40	0.237	3.69	0.87
19	34	8	0.235	2.08	0.49

The $normHits * BA$ value will serve as a weight for each specific age. This weight will be important in generating a formula for next season's projected batting average.

The formula used to solve for next year's projected batting average is shown below.

$$K * \frac{AW_{2024}}{AW_{2023}} * AA_{2023} = xBA_{2024}$$

Where K is:

$$K = \frac{xBA_{2023} + TrueBA_{2023}}{2 * AA_{2023}}$$

In these functions, AW is the age weight for the age of the player for the specific season which is subscripted. AA is the average batting average for the specific age of the player in the subscripted season. This value is pulled from the BA column in the table above.

Now that we've settled on a formula for projecting batting average in the 2024 season, the next step is to put this formula into action. In the following section, we'll apply this formula to the players and make our best estimate for their production in the 2024 season.

Results: 2024 Season Predictions

Let's start with the players that were highlighted in the 2023 season results section. The following table takes a look back at the top ten players under 25 from the previous season. As a reminder, these players were expected to boast a much higher batting average than their true value.

Player	2023 Age	2023 BA	2023 xBA	2024 Age	2024 xBA
Luis Matos	21	0.250	0.300	22	0.277
Anthony Volpe	22	0.209	0.253	23	0.237
Alec Burleson	24	0.244	0.289	25	0.288
Vinnie Pasquantino	25	0.247	0.286	26	0.261
Brice Turang	23	0.218	0.253	24	0.258
Dylan Carlson	24	0.219	0.251	25	0.255
Brett Baty	23	0.212	0.245	24	0.251
Gabriel Arias	23	0.210	0.238	24	0.246
Joey Wiemer Jr.	24	0.204	0.231	25	0.235
Michael Massey	25	0.229	0.255	26	0.237

Looking at this list, we notice several players who are anticipated to make significant progress from the previous season. Alec Burleson and Brice Turang stand out as having the biggest jumps from their actual batting average in 2023 to their expected batting average for the 2024 season. Luis Matos also shows a considerable improvement and is definitely someone to keep an eye on, especially since he's the youngest player on this list. As he enters his second season in his budding career, with only 76 games under his belt from his rookie season, Matos should have plenty of opportunities to shine, and I expect him to have an excellent season.

We could continue going on about the other players on that list, but I believe it's more appropriate to shift our focus to the young players who outperformed expectations last season. Once again, these players exceeded their anticipated batting average, according to the model. As we've mentioned before, there are inconsistencies for these players due to their overall development and increased opportunities as they advance in their careers. With the formula we discussed earlier, our goal is to gain a clearer understanding of what to anticipate from these players in the 2024 season.

Player	2023 Age	2023 BA	2023 xBA	2024 Age	2024 xBA
Jared Triolo	25	0.298	0.227	26	0.257
Royce Lewis	24	0.309	0.244	25	0.299
Mickey Moniak	25	0.280	0.222	26	0.246
Matt Wallner	25	0.249	0.199	26	0.219
Elehuris Montero	24	0.243	0.195	25	0.236
Nolan Jones	25	0.297	0.256	26	0.271
Sal Frelick	23	0.246	0.207	24	0.249
Liover Peguero	22	0.237	0.201	23	0.225
Blake Sabol	25	0.235	0.201	26	0.214
Christian Encarnacion-Strand	23	0.270	0.237	24	0.278

Despite having a lower expected batting average than their actual batting average in 2023, Sal Frelick and Christian Encarnacion-Strand are projected to hit even higher than their actual average from the previous season. This is probably due to a few factors. Firstly, they're among the younger players in this group. Based on the age weights we discovered earlier, the difference between age 23 and 24 is more significant than for most other ages, and it appears to mark the beginning of a player's "prime." Secondly, these players batted at a higher average than their peers who were the same age in 2023. This certainly played a significant role in their positive outlook for the next season.

Though Royce Lewis hit an incredible 0.309 last season, he was only expected to hit 0.244. His projected numbers for the following season do not take much of a dip, as he is projected to hit 0.299. Despite the lower expected batting average from last season, we will still anticipate another solid season for Royce Lewis.

Finally, I wanted to look at the players that are expected to lead the league in batting in the 2024 season. Below are the ten players with the highest expected batting averages for the upcoming season.

Top 10 Players in xBA for the 2024 Season

Last	First	2023 Age	2023 BA	2023 xBA	2024 Age	2024 xBA
Acuna Jr.	Ronald	25	0.337	0.3056469	26	0.3149454
Arraez	Luis	26	0.354	0.2995161	27	0.3118081
Freeman	Freddie	33	0.331	0.2907025	34	0.3112519
Garcia	Maikel	23	0.272	0.2873988	24	0.3071909
Seager	Corey	29	0.327	0.2889104	30	0.3066577
Moreno	Gabriel	23	0.284	0.2716163	24	0.3051138
Witt Jr.	Bobby	23	0.276	0.2761360	24	0.3032026
McLain	Matt	23	0.290	0.2605303	24	0.3023208
Diaz	Yainer	24	0.282	0.2718588	25	0.2992275
Lewis	Royce	24	0.309	0.2440466	25	0.2987887

Figure 6: 2024 Top 10 Players in xBA

This list gives us a fantastic mix of players. There are some who've already established themselves as among the best in the world, alongside a bunch of young talents with promising futures ahead of them.

We see that Ronald Acuna is projected to be the league leader in batting average this season. That should not come as a major surprise, seeing as he is still on the early side of his prime, and is coming off an incredible year. Luis Arraez finally sees the respect he deserves, as he is projected to be just behind Acuna in batting this upcoming season.

Many baseball fans know about the promising talent of Bobby Witt Jr. He's likely to have another impressive year, building on what he displayed in the 2023 season. As for younger players who haven't quite made a big splash in the league yet, we can anticipate Maikel Garcia and Yainer Diaz to have standout seasons.

Conclusion

I trust that this article has been engaging, at the very least. The beauty of baseball lies in its unpredictability. Nevertheless, data analysts strive to forecast the future as best as they can, aiming to give teams an edge. Whether it's analyzing daily pitching matchups and game situations or scouting prospects and predicting player performance down the road, analysts work hard to contribute to teams' success. Exploring expected batting average in the MLB sheds light on the intricate dynamics of player performance. By utilizing statistical models and data analysis, we can gain valuable insights into players' abilities and potential. From seasoned veterans to emerging talents, understanding xBA offers a glimpse into the future of the game.

Looking ahead, I'm excited to share that I'm currently developing a user-friendly tool for everyone to explore this data. Rather than simply seeing the select few names on the top ten lists, you would be able to search

for a player and see what their projections look like for the 2024 season, and potentially more seasons to come.

Furthermore, there are still some aspects of the code that I plan to refine moving forward. This became evident in the final section when examining the top ten players for the 2024 season in terms of expected batting average. It's apparent that the model might be overly influenced by statistical averages. A young and exceptionally skilled player like Ronald Acuna shouldn't anticipate a significant decrease in expected batting average for the upcoming season. Additionally, I aim to incorporate more variables into the formula, such as considering a player returning from a major injury.

There's still plenty to discover and refine, and I'm excited to continue making progress. But for now, I'm truly pleased with the results we've discussed in this article. With the 2024 season underway, I'm eager to see how the players we've talked about here fare. It's definitely a goal to conduct a similar analysis next offseason and share the behind-the-scenes adjustments that were made.