Ryan Hance
4-15-18
CSCI 4831
Final Project, Part 1

For my final project, I would like to examine how hitters perform with two outs. I think that if a team can score with two outs, they have a much higher chance of winning because they are adding more runs before giving up the last out of the inning. Therefore, if I can determine which players perform the best when there are two outs, I can find who is the best at adding extra runs for their team. The new statistic will be call Two-Out Performance Percentage, or TOP%.

One critical element of TOP% is how often the batter can score a runner in scoring position with two outs. The calculation for this is the number of times the batter gets a hit, walk, or hit by pitch to score the runner in scoring position with two outs divided by the number of opportunities he has. Each runner will count as a point, so scoring runners from second and third counts as two points for one attempt. In addition, if the batter gets a hit, walk, or hit by pitch and the runner does not score (for example, if the runner simply moved from second to third or the runner got thrown out at home), the batter will be given a half point. If the batter scores a runner from first or he scores himself, he will get one and a half points per score since these players are much harder to score from one at-bat. This will be combined into a new statistic called Adjusted Two-Out RBI Percentage (ATOR%).

Another critical element of TOP% is the batter's success at getting on base with two outs, which could mean he got an RBI or he put himself in position for a teammate to get an RBI with two outs. This calculation is simply on-base percentage (OBP) for the batter when he comes to bat with two outs. Then, the ATOR% and OBP will be summed together, much like OBP and SLG are summed for OPS, and the result will be the player's TOP% score.

The two most similar current statistics to TOP% are RBI and OBP. The first half of TOP%, ATOR%, is a weighted measure of how many runners a batter is able to drive in with two outs, whereas RBI is a simple counting statistic of the number of runners a batter drives in. ATOR% can be thought of as a weighted measure of RBI based on how difficult it is to score the runners he drove in and with a two-out constraint. The second half of TOP% is simply OBP with two outs. The calculation for OBP is performed with the same two-out constraint as ATOR% and the result is summed with ATOR% to get the complete TOP% statistic.

One of the main differences between TOP% and other statistics is that TOP% is focused specifically on two-out scoring, while most other stats are focused on overall player performance without considering out count. In addition, TOP% measures the percentage of times a batter is successful at scoring runners when given opportunities instead of how many runners he was able to drive in.

The data to be used for this statistic comes from the Retrosheet database used in previous class assignments. It consists of finding the times when a batter comes up to hit with two outs and the number of times a hit, walk, or HBP is achieved as well as if the batter drove in any runs. I have made a webpage to display the results of this statistic.

The top 15 players in terms of TOP% from the 2010-13 data available via Retrosheet are displayed below in Table 1. The data was filtered out to remove players who had less than 100 opportunities to hit with runners in scoring position and two-outs. This removed several players who were very successful at scoring runners but with a very small sample size that may not hold as the player gets more at-bats at the MLB level. As can be seen, many of the top hitters of the

early 2010s were among the top two-out producers, as should be expected. Established players like Miguel Cabrera, Albert Pujols, and Prince Fielder score well in TOP%, while emerging stars like Joey Votto, Giancarlo Stanton, and Paul Goldschmidt begin to show that they are capable of producing at a top level. It is interesting to note that this list includes 9 players who spent a significant amount of time as first basemen during these years.

Table 1: Top 15 players from 2010-13 in TOP% statistic.

First Name	Last Name	ATOR%	OBP	TOP%
Miguel	Cabrera	0.828	0.438	1.266
Jose	Bautista	0.744	0.409	1.153
Evan	Longoria	0.783	0.369	1.151
Joey	Votto	0.680	0.448	1.128
Edwin	Encarnacion	0.771	0.347	1.117
Giancarlo	Stanton	0.747	0.347	1.094
Ryan	Braun	0.658	0.380	1.039
Carlos	Gonzalez	0.645	0.385	1.030
Albert	Pujols	0.652	0.377	1.029
Ryan	Zimmerman	0.647	0.379	1.026
Paul	Goldschmidt	0.639	0.384	1.023
Freddie	Freeman	0.643	0.379	1.022
Derrek	Lee	0.696	0.324	1.020
Allen	Craig	0.653	0.359	1.012
Prince	Fielder	0.605	0.406	1.011

Tables 2 and 3 show the top 15 players in RBI and OBP, respectively, from 2010-13. As can be seen, there are 6 players on both the TOP% and RBI lists and 6 players on both the TOP% and OBP lists. Four players appear on all three lists. This shows a loose correlation between TOP% and high RBI or OBP numbers.

Table 2: Top 15 players from 2010-13 in RBI.

First Name	Last Name	RBI
Miguel	Cabrera	507
Robinson	Cano	428
Adrian	Gonzalez	426
Prince	Fielder	417
Josh	Hamilton	401
Adrian	Beltre	401
Hunter	Pence	391
Albert	Pujols	386
Alfonso	Soriano	376
Jay	Bruce	375
Matt	Holliday	374
Jose	Bautista	365
Ryan	Braun	364

Carlos	Gonzalez	364
Billy	Butler	362

Table 3: Top 15 players from 2010-13 in OBP.

First Name	Last Name	OBP
Joey	Votto	0.437
Miguel	Cabrera	0.429
Mike	Trout	0.408
Joe	Mauer	0.403
Prince	Fielder	0.400
David	Ortiz	0.395
Jose	Bautista	0.393
Shin-Soo	Choo	0.392
Matt	Holliday	0.388
Andrew	McCutchen	0.386
Ryan	Braun	0.385
Buster	Posey	0.384
Lance	Berkman	0.382
Troy	Tulowitzki	0.382
Paul	Goldschmidt	0.380