Comparing MLB Run Value and Awards

Andrew Eross, Owen Wassel, Nick McConnell

Introduction

Baseball has always been called "America's Pastime." However, beginning in the end of the 20th century and the early 2000's, the evolution of baseball analytics has greatly evolved the game. In the past, baseball people had just looked at offensive stats such as batting average, home runs, RBI, and strikeouts. But with the ability to track more and more data related to baseball, new stats have been developed to analyze and evaluate teams and players.

Many attribute the "Moneyball A's" as being a source renaissance for baseball analytics, where the early 2000's Oakland Athetics focused their roster building and talent acquisition by evaluating players on less popular statistics such as OBP (on base percentage), an effective way to acquire impact players who were overlooked by other teams. However, this age has now past, and new numbers drive baseball decision making.

MLB introduced Statcast in 2016, which, from the MLB website, is described as the "state-of-the-art tracking technology that allows for the collection and analysis of a massive amount of baseball data" through means such as cameras, radars, and other tracking devices. This data is accessible on Baseball Savant (an MLB licensed website). Baseball Savant takes these observations and calculates statistics related to how well players move and how balls are hit or thrown by players based on their spin, direction, and velocity. Some well-known Baseball Savant stats are hard-hit rates (how often a player hits a ball well) and xBA (the expected batting average of a player given how well they hit the ball).

The statistic we would like to analyze is Baseball Savant's run value, described on their website: "Every pitch is assigned a run value based on its outcome (ball, strike, home run, etc.). The sum of all of a player's contributions across a season, or multiple seasons, measures his overall batting or pitching run value. A positive value represents runs created for hitters, and runs prevented for pitchers." We are just looking at offensive run value, so how many runs hitters supposedly create at the plate (not including base running). This run value calculation is all "theoretical" meaning it is not derived by summing all the runs that score as a result of a hitter's plate appearance, rather how many *should* score based on how well they hit the ball, neutralizing factors such as opponent defense and teammate base running.