Capstone Project 1: Prediction when MLB players will see a decline in production (WAR)

Problem:

MLB teams are constantly seeking to improve their rosters through free agency signings, trades and the development of minor league players in farm system. For these teams, player salaries represent a significant financial cost and it is important to be able to effectively evaluate a player's value.

WAR (Wins Above Replacement) is commonly used as the metric to determine player value; it is the sum of a player's total offensive and defensive contributions and attempts to measure the total number of additional wins a team can expect with the player over an average replacement level player[1]. A simplified version of the equation is:

$$WAR = (Batting\ Runs + Base\ Running\ Runs + Fielding\ Runs + Positional\ Adjustment + League\ Adjustment + Replacement\ Runs) / (Runs\ Per\ Win)$$

For this project, I will attempt to identify trends and indicators that may signal that a position player (non-pitcher) will see a significant decline in WAR. A common problem that occurs in the MLB is that teams will acquire a star player at the height of their careers for a large contract, only to discover that they've anchored themselves and a significant portion of payroll to a player whose production is quickly diminishing. With this information, MLB teams may be better able to forecast when a player's production is on the verge of decline, which in turn will improve their ability to make informed player acquisition decisions.

Data Sets:

Lahman's Baseball Database: http://www.seanlahman.com/baseball-archive/statistics/
Baseball-reference.com: https://www.baseball-reference.com/
MLB.com: http://mlb.mlb.com/stats/

Approach:

The first step in the analysis will be to look back at past MLB seasons and identify the players that experienced the biggest declines in WAR between consecutive seasons. After identifying these players, I plan on looking at the player's statistics the season prior to the drop off to see if there are common trends/leading indicators that these players share (ex. a dropping off of OBP%).

Deliverables:

- Python code posted on GitHub
- Jupyter Notebook report with results
- Slide deck presentation

[1] Replacement level player: A player of common skills available for minimum cost to a major league baseball team. A team of replacement-level players would be expected to win a baseline minimum number of games, typically 40-50, per 162 game season.