Project Narrative

* Modern baseball advanced analytics (sabermetrics) came to prominence with the success of the 2002 Oakland Athletics, along with the subsequent book about the team by financial journalist Michael Lewis, *Moneyball*
* Since then, they have become more prevalent in Major League front offices, as more and more analysts have been hired
* Initially, sabermetric principles were used to acquire discounted talent
  + Players possessing skillsets typically undervalued by conventional baseball thinking
  + Used by teams which did not have the resources to pay top dollar for the perceived best players
* Theoretically, as these concepts permeated baseball front offices, adjustments to which players were awarded which contracts should have followed suit
  + More teams wanting a similar skill set should lead to bidding wars for the same players, as it did for traditionally valued players
* So, does value correlate with player salary?
* One metric used to measure the overall contribution which a player provides to his team is Wins Above Replacement
  + As WAR is a product of the sabermetric revolution, it should better correlate with salary now than it did in, say, the 90s
  + (Show charts revealing that the correlation actually dropped)
* Why is this the case?
  + WAR is a complex metric with many different components, so maybe the actual correlation is simply lost in the intricacy?
  + We ran tests on salary and other advanced metrics which comprise WAR, specifically OPS+ and ERA+
    - There was essentially no correlation
* This made us question how prevalent these concepts actually are in baseball today
  + We created linear regression models to determine the relationship between WAR and both plate attempts (for batters) and innings pitched (for pitchers)
    - WAR was a very significant factor for both, but there was not much change between the 90s and the past 5 seasons
  + We then tried doing the same for PA/OPS+ and IP/ERA+
    - Results were not very significant for ERA+
    - There was a rather notable increase in the relationship between PA and OPS+
  + We thought that one way to see how player valuations have changed was to examine the relationship between both PA and IP and more traditional stats
    - Home runs, runs batted in, strikeouts, earned run average, and pitching wins
    - Linear regression models showed a clear shift away from these stats
  + This supports the idea that advanced metrics have altered the way players are valued
  + The question, then, is what replaced those factors in determining salary?
* We created a correlation matrix for our dataset to see what correlated with salary
  + The only variable of note was age
* However, the relationship of these two variables decreased considerably between the two time periods for batters specifically
  + Furthermore, age accounts for 55% of what determines salary from 1995-1999; this number drops to 36% between 2015-2019
* One explanation of this could be a drop in average salary awarded to players in free agency
  + Such has been a concern sited by many commentators and industry experts
  + This is difficult idea to prove due to the rise in overall player salary due to inflation and increases in league/team revenue
  + Still, there was a clear decrease in the relationship between salary and players with 4+ WAR, signaling a decline in top talent receiving historically standard levels of compensation
  + Also, there were fewer annual salaries above league average in 2015-2019, despite having 1,000 more players with major league service time than 1995-1999
  + This is also supported by 2019 experiencing the first decline in average salary in 10 seasons
* Also, there was a clear shift in contribution as measure by WAR
  + In the 90s, players in their late 20s/earlier 30s comprised lion’s share
  + In the 2010s, players in their early 20s dominated
* This is notable because teams will inherently have to pay younger players less
  + Teams hold exclusive rights to players for their first 6 seasons of Major League service time
  + While after three years players can be awarded higher pay through arbitration, they will never make as much as they historically have on the open market
* **Conclusion**: It appears that, instead of utilizing sabermetrics to pay deserving players more, teams have used these principles to save money by maximizing the use of players before they reach free agency