**Team Group: #6 – Kevin, Vincent, Bhavani (Team MLB) 04-09-2018**

**Project #1 – Baseball statistics**

**Topic Investigation – Parameters of interest**

Team salary vs. Winning Percentage (1)

* Bar plot, Scatter (X-axis: $), Overlays

Batter salary vs. Average Stats (Batting Average, On-Base%, Slugging%) (2)

Batter salary vs. Counting Stats (Homeruns, RBI, R) (3)

Pitcher’s salary vs. ERA (4)

Pitcher’s salary vs. Counting Stats (# of Strikeout, Wins) (5)

Pitcher’s salary vs Average Stats (batting average against, on base% against, slugging% against) (6)

Team Salary vs WAR (Wins Above Replacement) (7)

Player salary vs WAR (Wins Above Replacement) (8)

**Scope:** Data - 10 years

Raw data: <https://www.baseball-reference.com/>

API: <https://developer.sportradar.com/files/indexBaseball.html#daily-boxscore>

**Schedule:**

1. Data gathering completed – 4/11 (Github) Wed
   1. All data is complete
2. 1st team meeting for division of work – 4/11 Wed
   1. Divide by which key plots
3. 2nd team meeting for results review – 4/14 Sat
   1. Team members share their investigation
   2. Discussion results and discuss next set of work by team member – presentation flow (google-doc)
4. 3rd team meeting – come back for integration of results – 4/16 Mon
   1. Finalize the remaining work and coordination the story for presentation

**Data Analysis:**

Answer the questions:

* Do teams that pay more win more? Is it a direct relationship?
* For players (pitchers and batters), what are key performance metrics correlated to player’s salary
* Explain any trends or correlations