

# Los Angeles Dodgers Team & Player Performance Data Analysis: 2018-2020

Acusio Bivona



# My Why



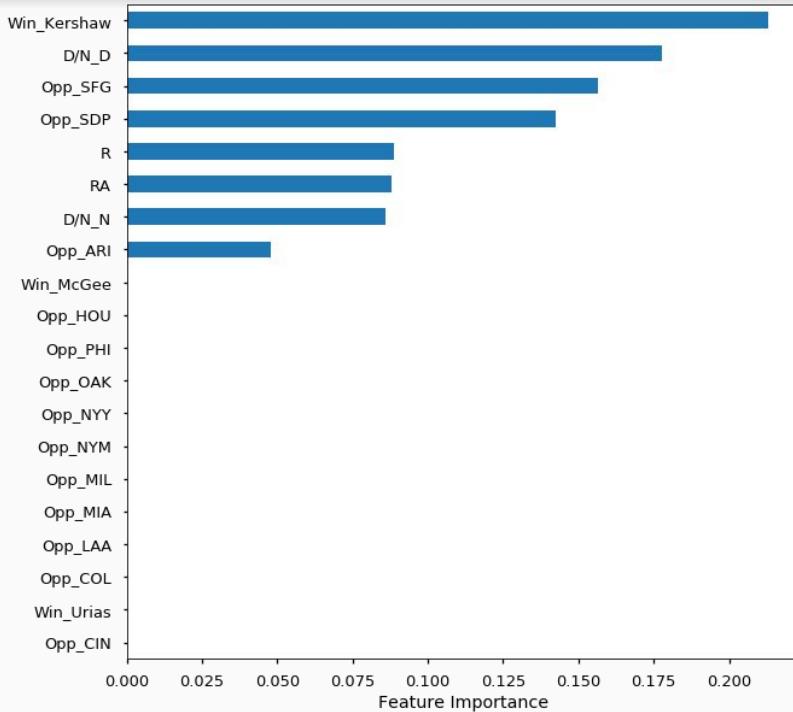
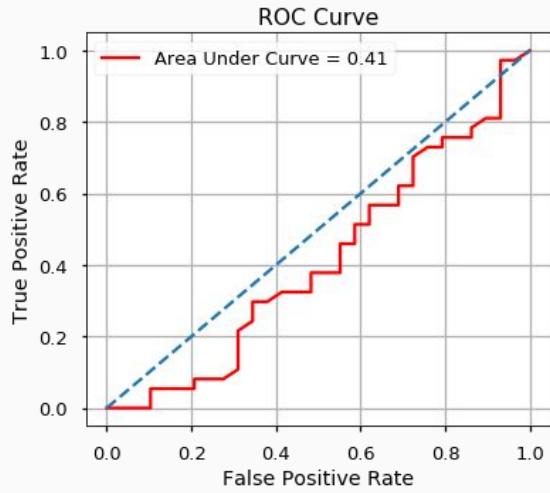
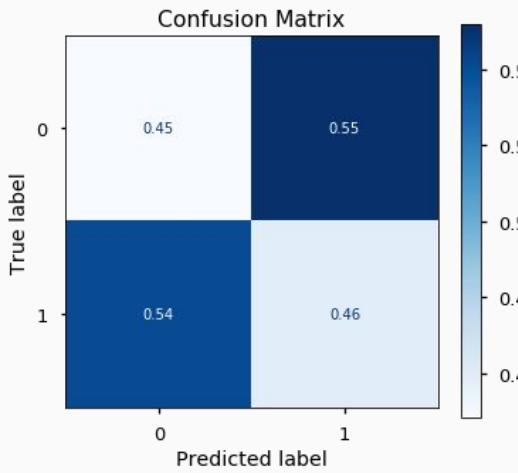
- Dodgers fan for 12 seasons
- Interesting season due to COVID-19
- What if last two seasons were only 60 games?
- Allow other Dodger & baseball fans to explore the data
  - Look for trends, compare performance between seasons, hot/cold streaks, etc.

# Objectives

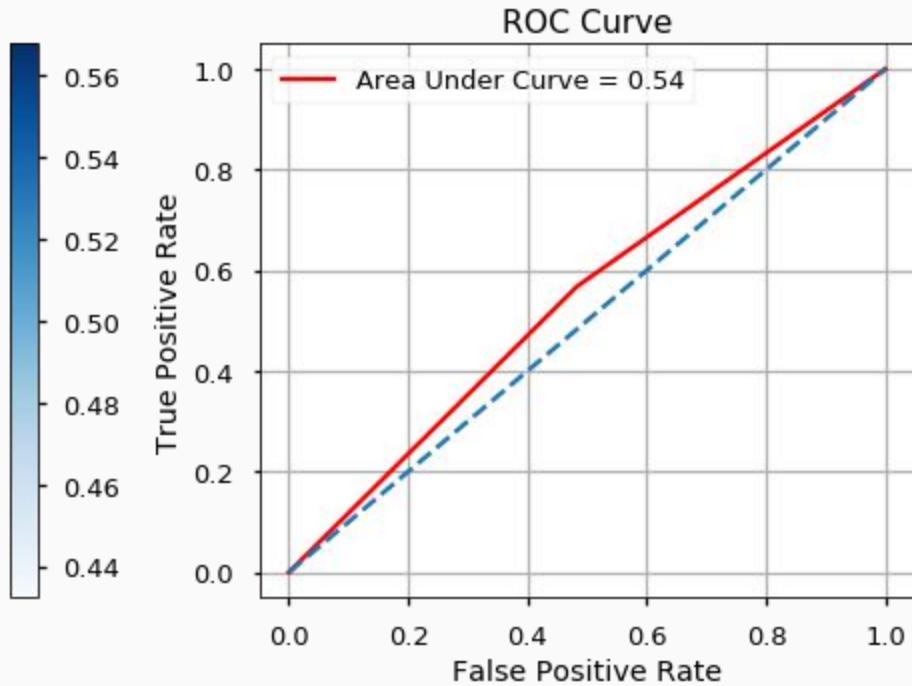
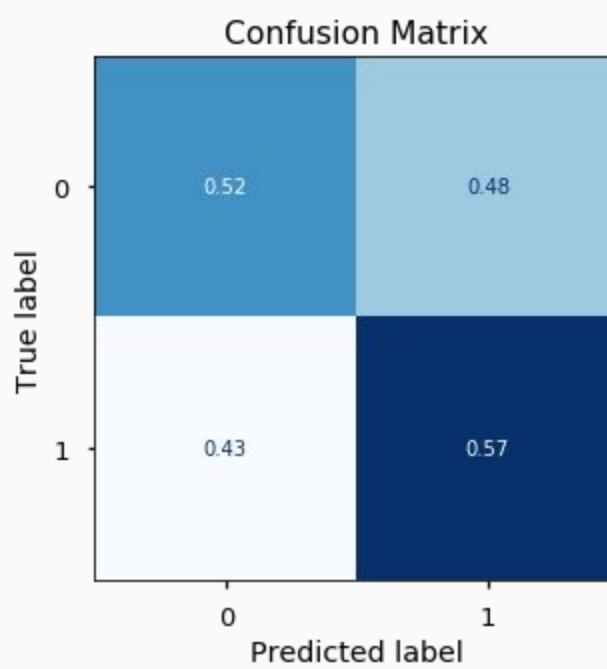


- Collect team & player data for 2018-2020 seasons
  - Done through using API(s) and webscraping
  - Box scores, pitching data, hitting data, game logs
  - Sources: mlb.com, baseballsavant.com, & baseball-reference.com
- Create machine learning model to predict future wins and losses
  - 4 models: Random Forest, XGBoost, Naive Bayes, K-Nearest Neighbors
- Create a variety of EDA plots on players' performance
  - Used to visually compare player performance this season with 2018 & 2019 seasons
- Based on first 60 games of each season
- Create data-driven recommendations

# W/L Predictions - XGBoost



# W/L Predictions - KNN



# Performance Visuals

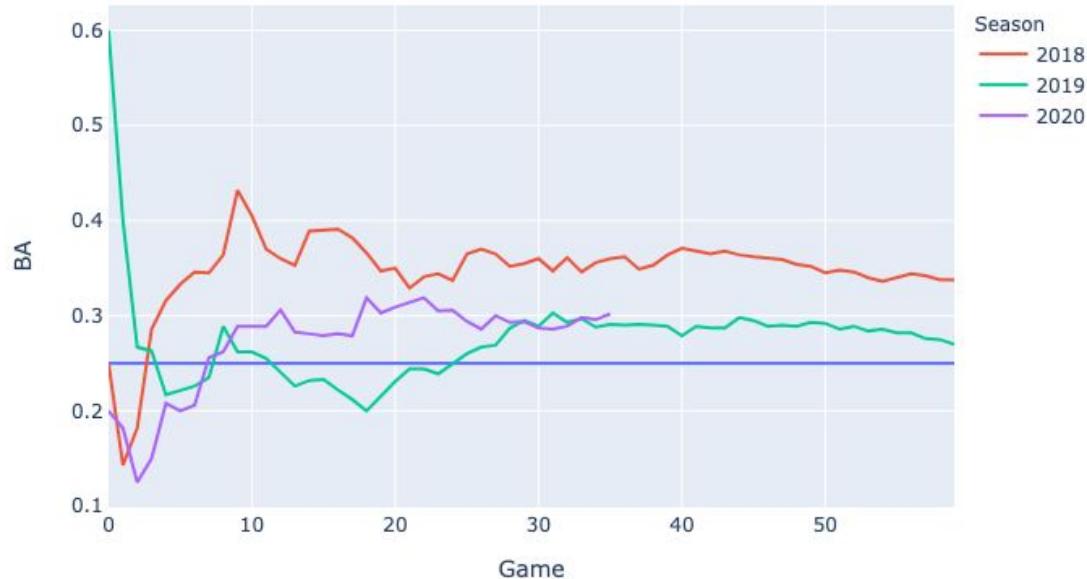


- Pitchers & Position Players placed into their respective groups
  - Pitchers = Pitching Performance Data
  - Position Players = Hitting Performance Data
- Dashboards wanted, but still in progress
  - Will allow users to interact and view whichever data they want
- Examples: Mookie Betts for Position Players, Clayton Kershaw for Pitchers

# Mookie Betts - BA



Batting Average - Betts, Mookie



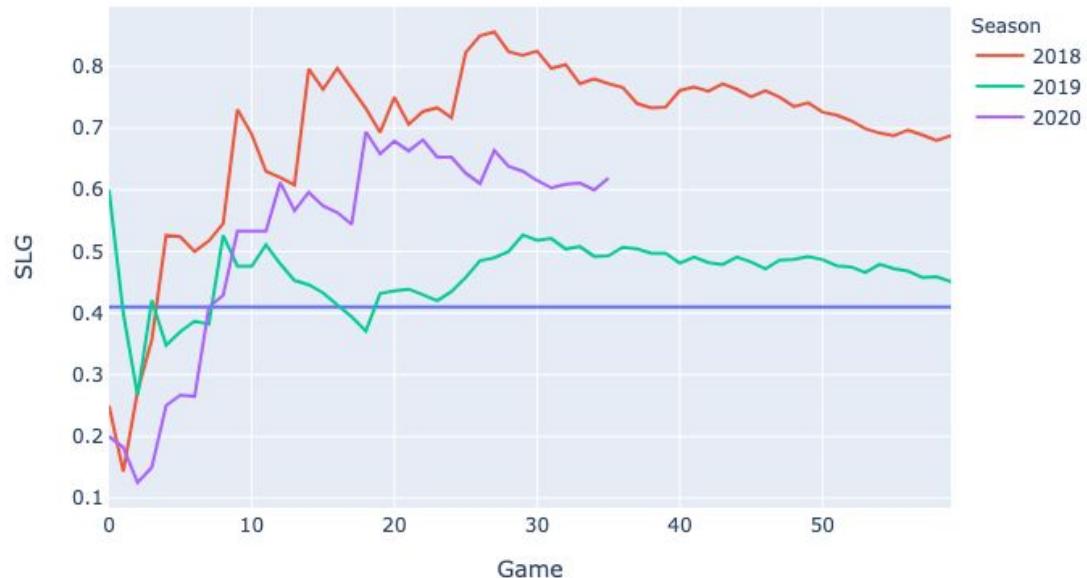
# Mookie Betts - OBP



# Mookie Betts - SLG%



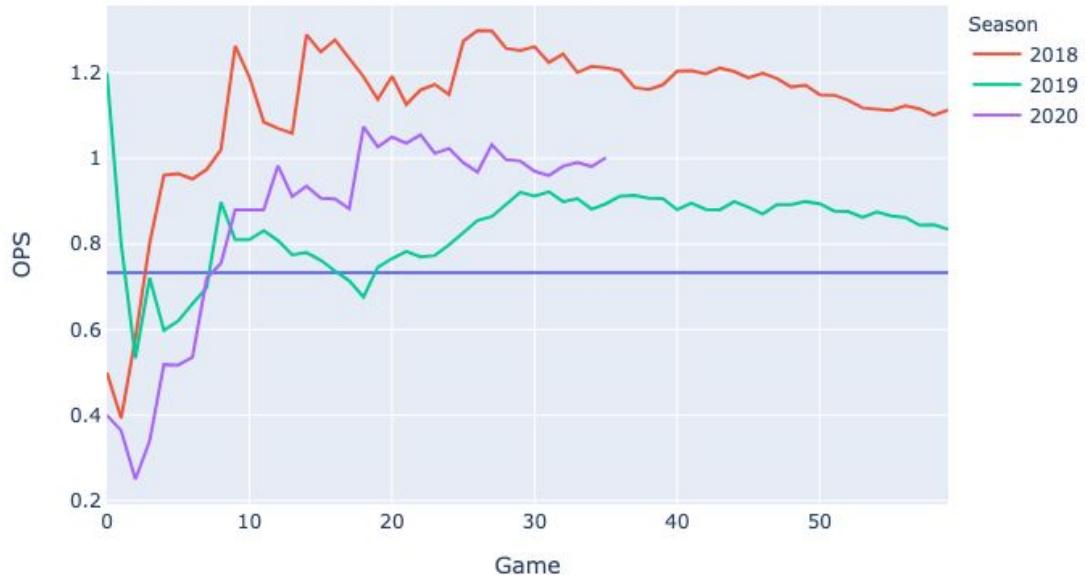
Slugging Percentage - Betts, Mookie



# Mookie Betts - OPS



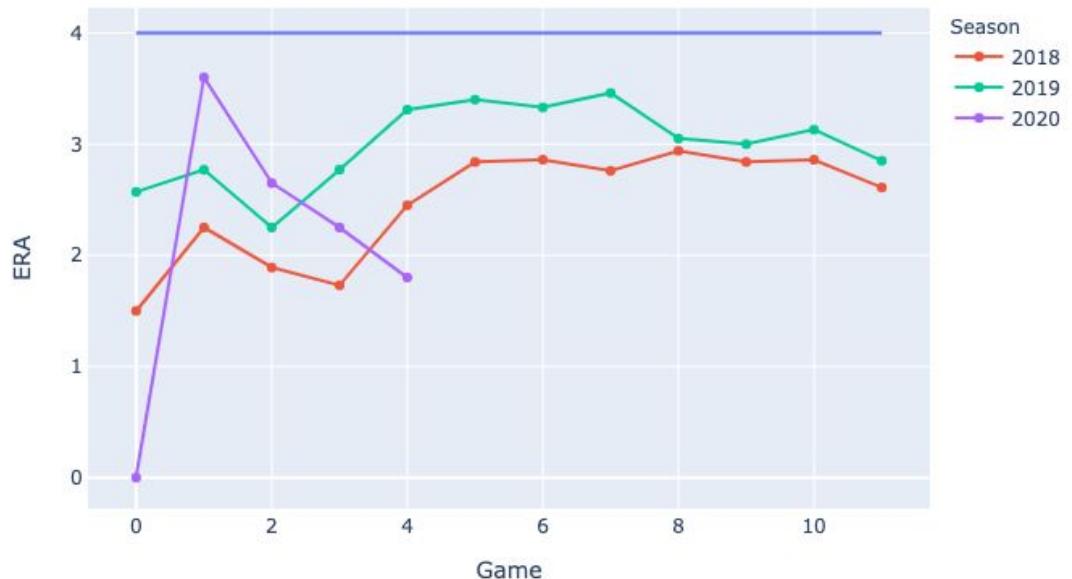
On Base % + Slugging % - Betts, Mookie



# Clayton Kershaw - ERA



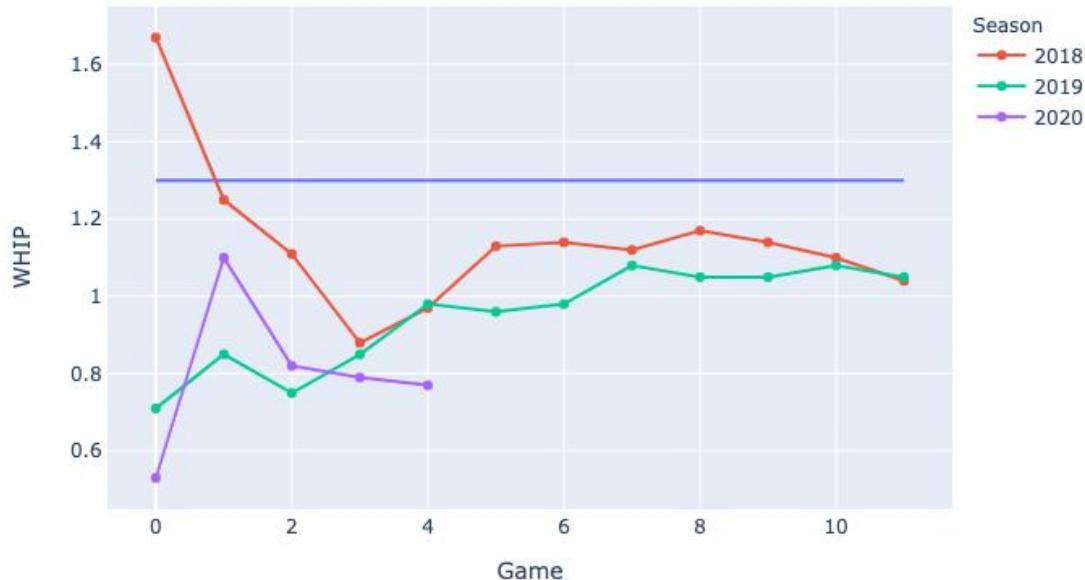
Earned Run Average - Kershaw, Clayton



# Clayton Kershaw - WHIP



Walks & Hits Per Inning Pitched - Kershaw, Clayton



# Recommendations



- Position Players:
  - Determine Batting Orders - Examples:
    - High AVG, Low SLG% = 1st or 2nd
    - Low AVG, High SLG% = 5th or 6th
    - Balance of both = 3rd or 4th
- Pitchers:
  - Starters:
    - Better performance = more trust = longer starts
  - Relievers:
    - Better performance = pitch in higher-leverage situations

# Summary



- Created prediction model for W/L
  - XG 46% W accuracy, AUC 0.41; KNN 57% W accuracy, AUC 0.54
    - More data & different tuning should yield better results
- Performance plots
  - User-friendly data analysis
  - Provides performance idea for 60-game seasons
- Daily updates = game-time decisions on the fly
  - Changing batting orders, deciding who will relieve & when, etc.

# Future Works



- Possibilities are seemingly endless
  - Include more seasons
  - Include more (or all) MLB teams
  - Include defensive performance
  - Go into advanced statistics
    - Hitters - XBA, XSLG%, Hard Hit %, K%
    - Types of pitches thrown, average velocities, pitching quadrants
  - More complex prediction models
    - Season simulations, predict player performance before season

# Thank you!



- Questions are welcomed
- Images credit to [Getty Images](#)