Predicting NBA Daily Fantasy Scores with Data

Drew Hibbard

Daily Fantasy Games are Dominated by Algorithms

- DraftKings and FanDuel
- Data > Anecdotes
- 91% of profits were made by 1.3% of players (McKinsey)

Scraping the Web to Obtain Stats Data

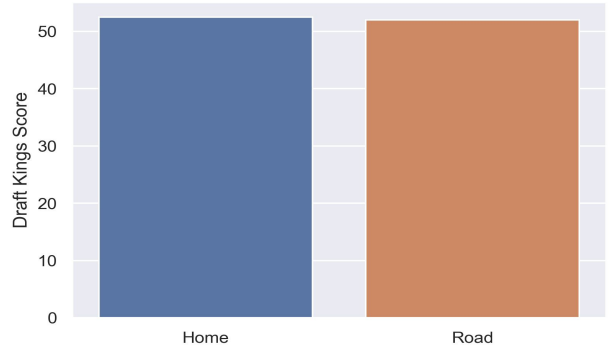
- Basketball-reference.com
- Game logs for every player 2015-2019
- Useful raw features:
  - All basic stats
  - Opposing team
  - Date of game
  - Home/Road

Feature Engineering to Obtain More Valuable Predictors

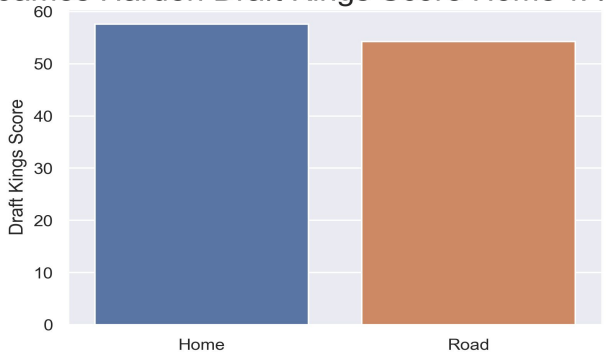
- Player average vs each opponent
- 3- and 7-game trends to capture hot/cold streaks
- Player’s overall season average
- Days rest

Testing My Assumptions - Home/Road Splits

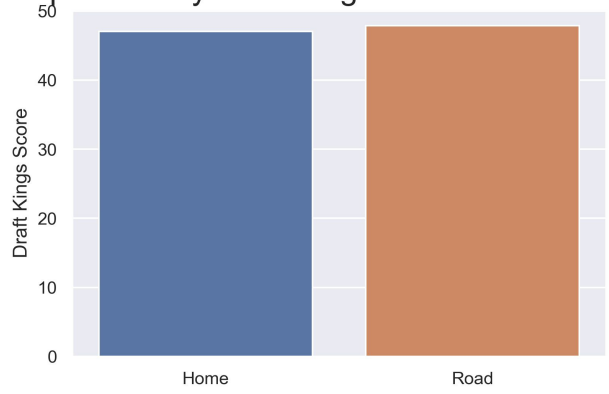
LeBron James Draft Kings Score Home v. Road



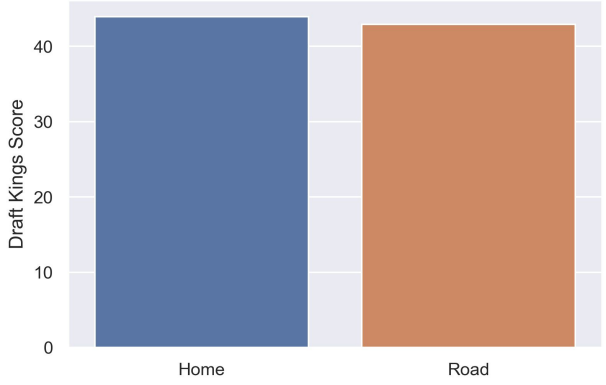
James Harden Draft Kings Score Home v. Road



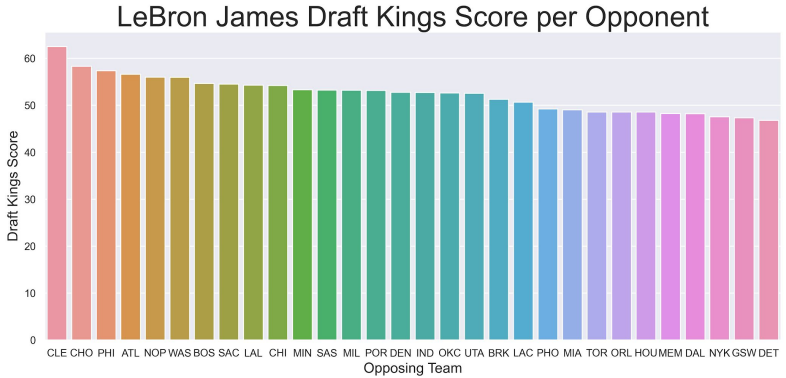
Stephen Curry Draft Kings Score Home v. Road

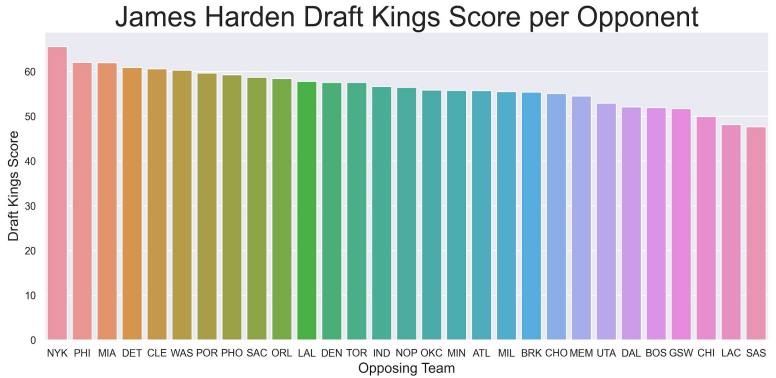


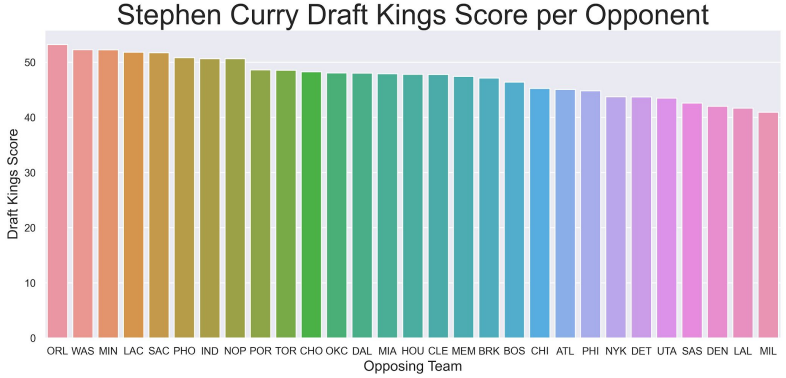
Damian Lillard Draft Kings Score Home v. Road

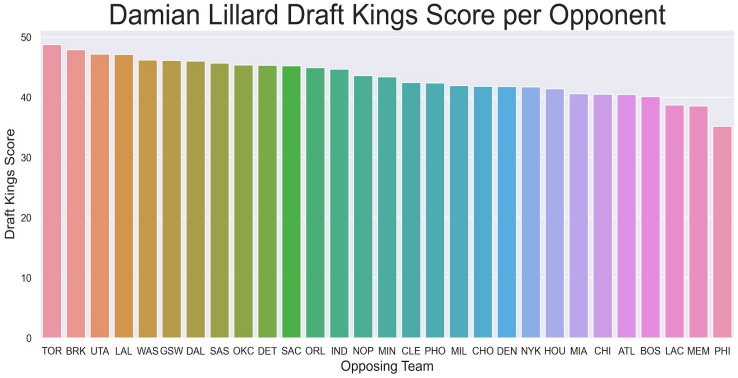


Testing My Assumptions - Opponent Splits

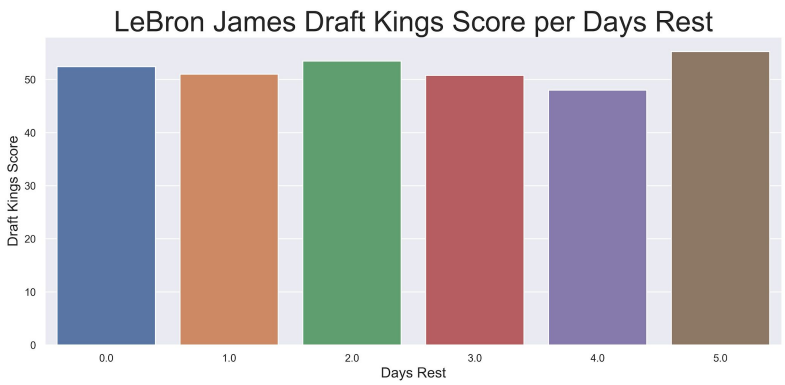


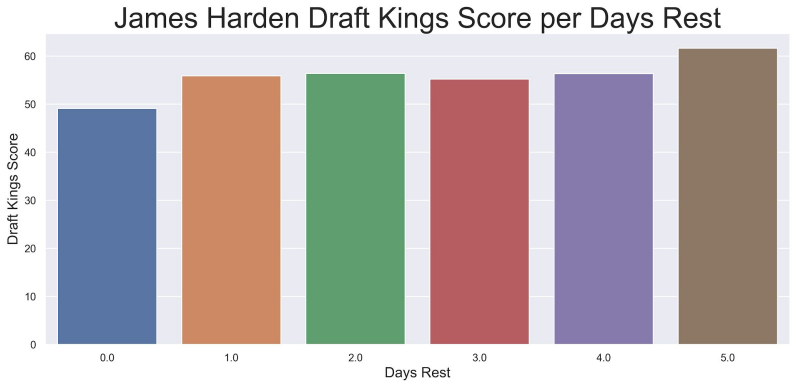


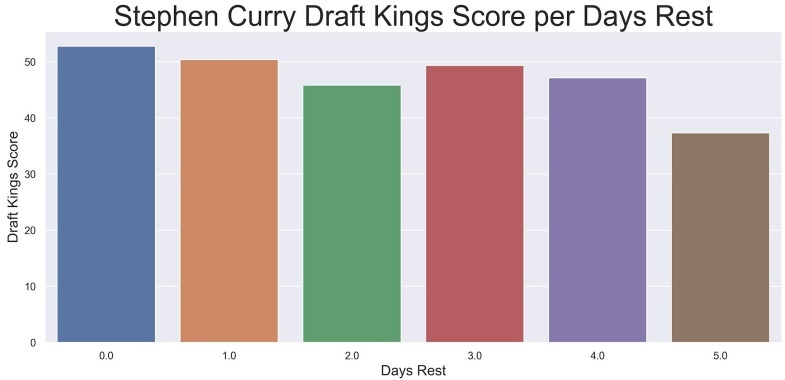


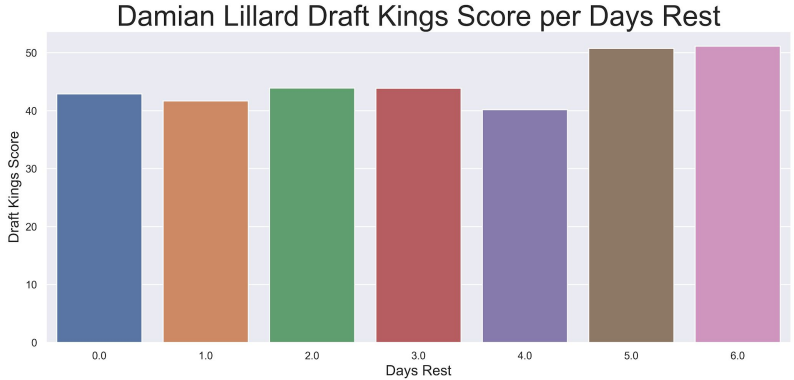


Testing My Assumptions - Days Rest

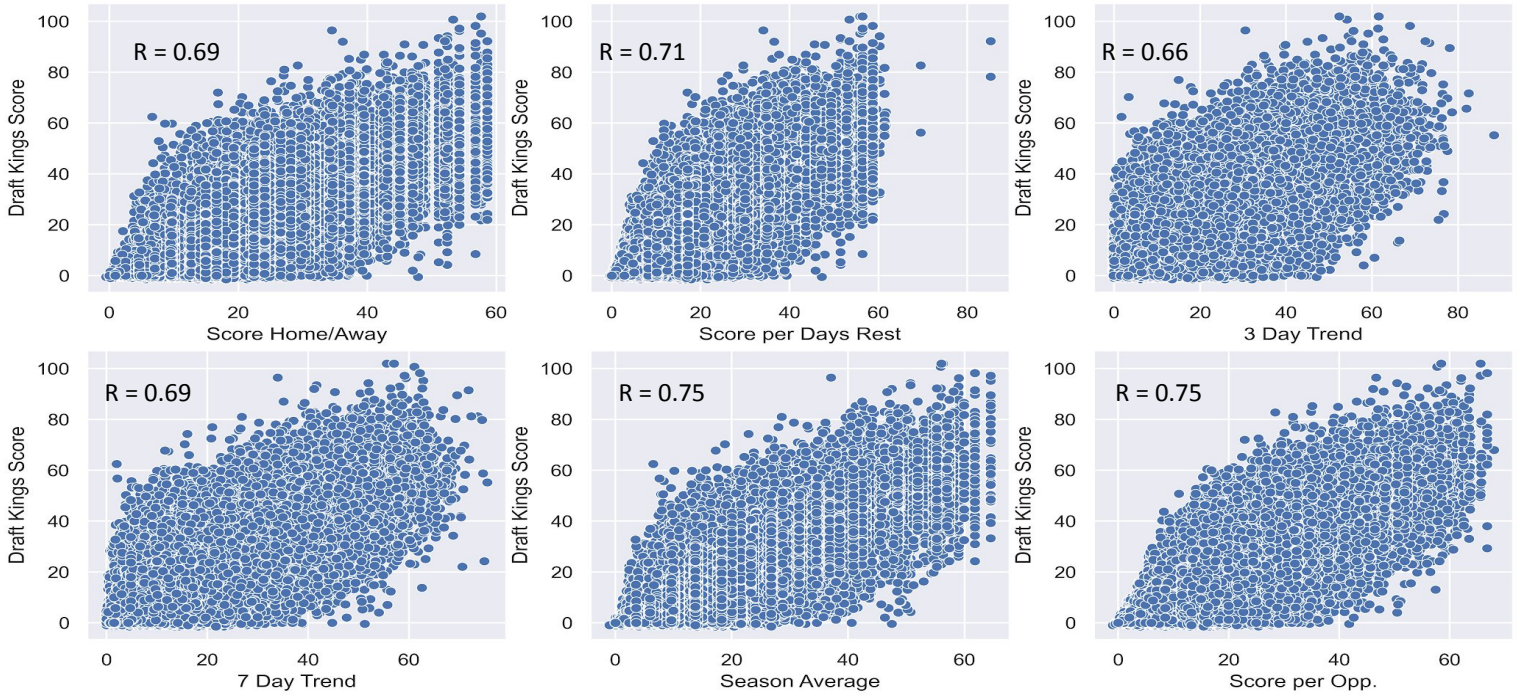






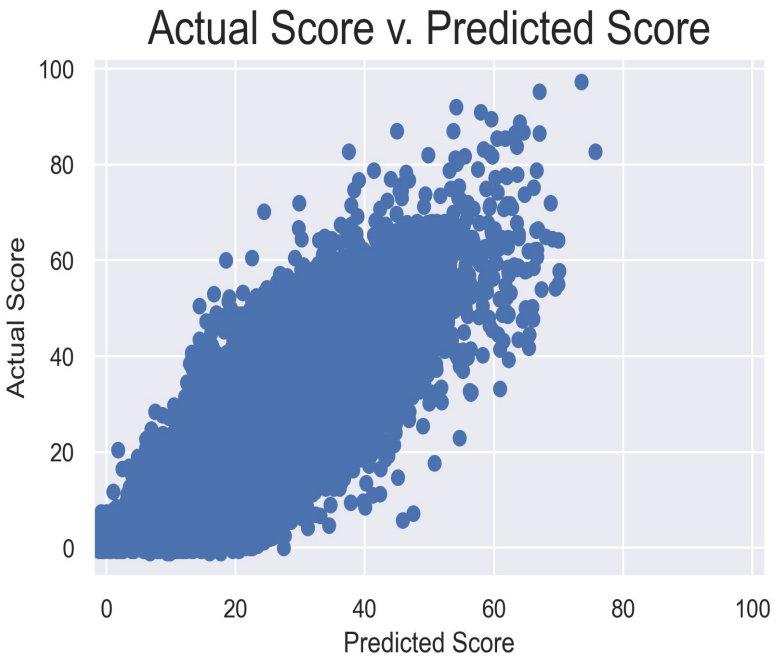


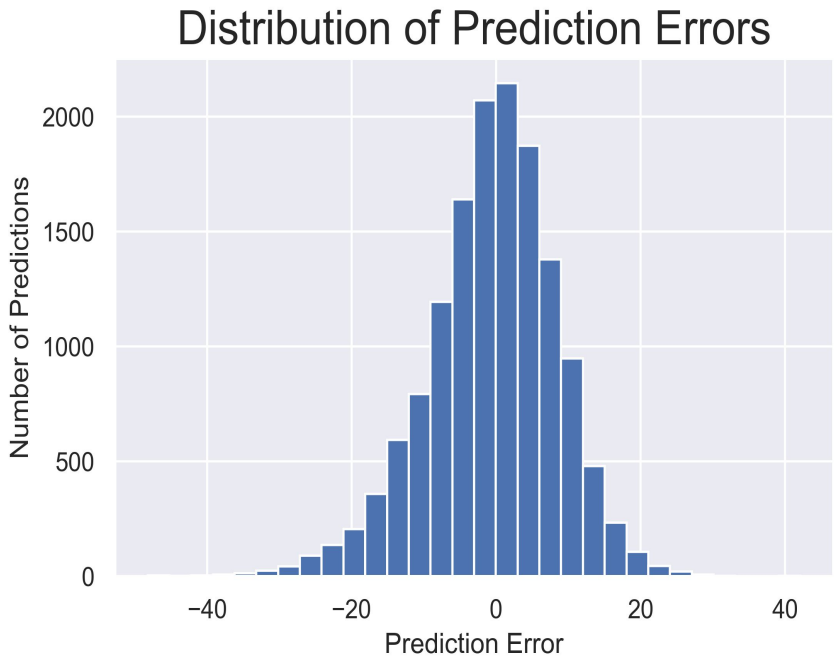
Model Development - Feature Selection



Model Results on 2019 Games

- Best: Simple Linear with all six features
- R squared: 0.66
- RMSE: 8.60 fantasy points
  - Standard deviation: 14.0


Actual Score v. Predicted Score
Actual Score
Predicted Score
0
20
40
60
80
100


Distribution of Prediction Errors
Number of Predictions
Prediction Error
0
500
1000
1500
2000
−40
−20
0
20
40

Next Steps

- Algorithm to scrape injury reports day of games
- Adjust algorithm to continually scrape new data
- Algorithm to determine optimal lineup based on player salaries
  - Undervalued players