Daily Fantasy Baseball Lineup Optimization

Group Members:
Brian Broeking 9070985453
Victor Calimao 9073620347
Alec Scheele 9066283624
Kairat Satbekov

1. Brief description of the problem. If you'll be using real data, where will you find it and how much will you need?

Draftkings is one of the major providers of daily fantasy sports in the United States. Everyday, Draftkings runs online competitions where between 2 to 50,000+ people compete to build the best fantasy baseball lineup for that day. The best lineups are formed by correctly picking players who have good performances in live Major League Baseball games during that day.

There are two major classes of competition. Head to head games involve the players trying to build lineups that will place in the top half of the tournament to receive their reward. Rewards for having the highest scoring lineup are equal to the rewards for having a lineup in the 50th percentile. The other format will be the one we are optimizing for in this project. GPP's or guaranteed prize pool tournaments are ones where the top 2-5% of players receive an overwhelmingly high reward compared to the rest of the entrants.

In the project, we hope to discover methods of creating optimal lineups based on data that we have collected over the past year of Draftkings events. Some of the data sources we have to work with are listed here. We will not be using all of these, but all these options are available for exploration and we have a start on collecting the data from most of these sources.

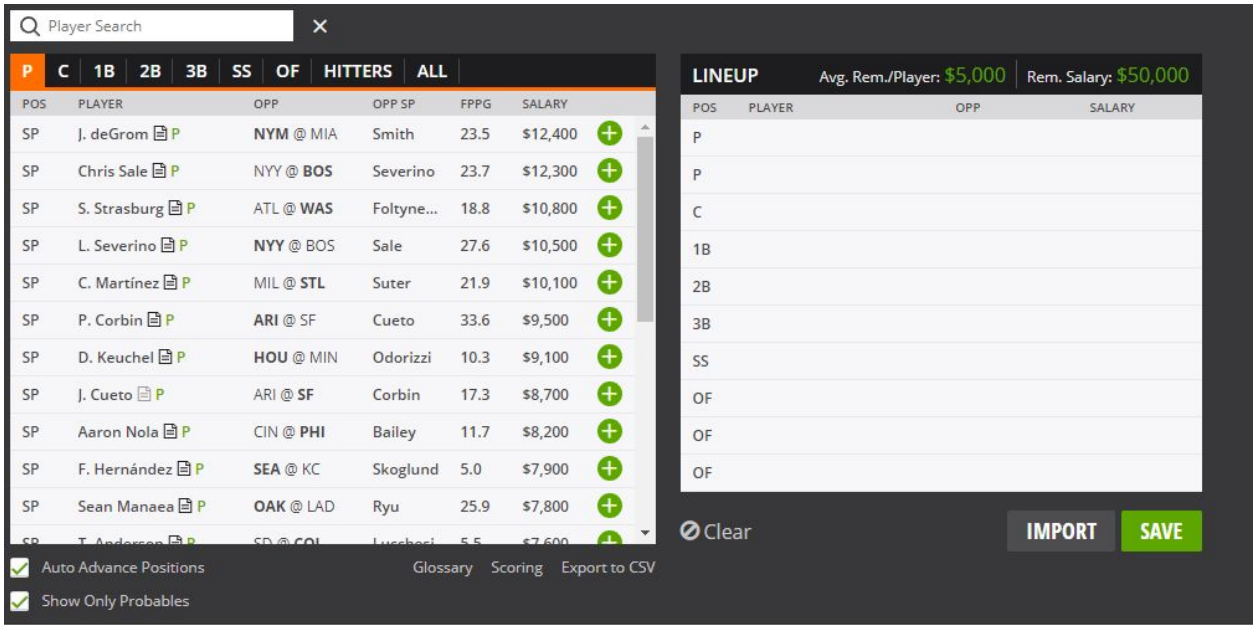
Sources.
Baseballreference.com or Fangraphs.com - Both of these websites contain statistical data for Major League Baseball players dating back to last century.

BaseballSavant.com - Major League Baseball partnered with Amazon to install sensors in all ballparks over the past few season. Because of their efforts, we have physics data for every play of the past two MLB seasons. This data can be used to further optimize the lineup selection.

Draftkings.com - Draftkings provides a CSV everyday of each players salary. We will use this to keep our lineups under the salary cap.

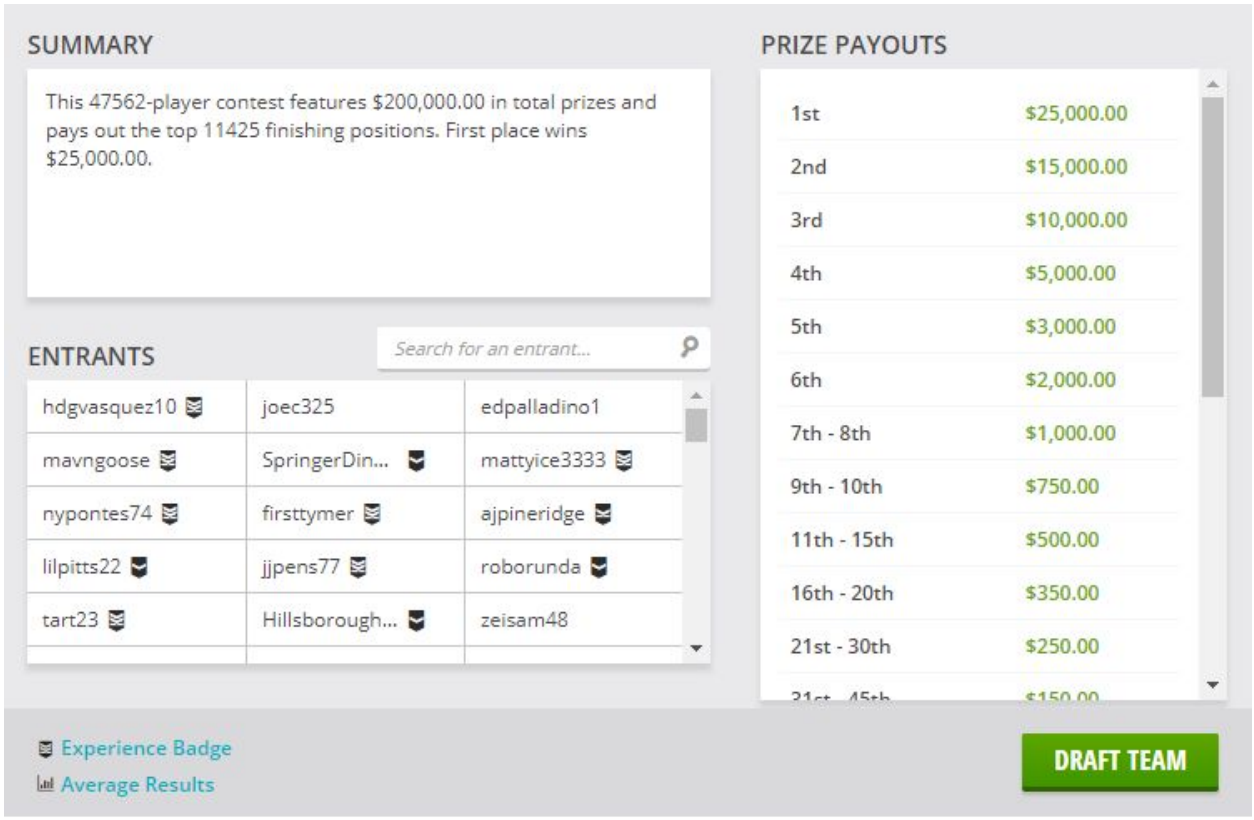
Past data gathered - We also have past contest data from last season and earlier this year. This data can be used to backtest the model.

Process of making a lineup.



This is the GUI interface for making Draftkings lineups. Each player is assigned a salary and users must decide how to optimally spend their $ 50,000 salary cap. This cap cannot be exceeded and each lineup must include 2 Pitchers, 1 Catcher, 1 First Baseman, 1 Second Baseman, 1 Third Baseman, 1 Shortstop, and 3 Outfielders.

Payouts.



The goal will to be to optimize the lineups such that they have the greatest likelihood of placing in the top of the competition. As you can see above, the prizing is heavily weighted towards the top of players. These competitions allow for multiple entries, so we hope to optimize many lineups to be entered in these competitions.

2. Type of model (LP, QP, MIP, etc.) and an approximate count of the number of variables and constraints in the model:

Variables: Players with
Constraints: Salary Cap ($50,000), Number of Players allowed in each position (2 Pitchers, 1 Catcher, 1 First Baseman, 1 Second Baseman, 1 Third Baseman, 1 Shortstop, and 3 Outfielders), Maximum number of players chosen from each team

Model: This will be a Linear Program with a binary layout where players in the lineup are marked with a 1 and those without are marked 0. Each player will be given a daily projection, and the objective of the model will be to optimize the lineup with respect to the daily projections. We are looking to find the lineups that maximize the projected points based on other variables. This is where the interesting parts of the project appear. Lineups that simply optimize based on daily projected points are usually not successful. Straightforward optimizers lack that ability to understand the correlation between batters of the same team. Due to the nature of the sport, the success of some hitters on a team will allow for other to have more opportunity. This creates scenarios where you want to “stack” players from the same team. One example of a basic

constraint we create from this idea is that we want all of our lineups to contain at least 4 players from a team. These are the lineups that have succeeded historically. Other possibilities include removing some players from the player pool based on the pitcher they are facing or based on the weather at the ballpark that day.