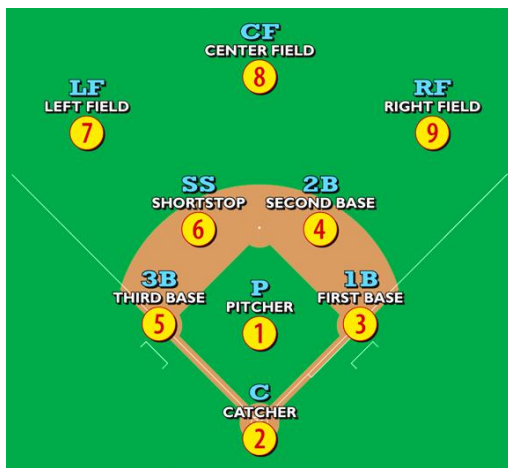# Group 2: Simple Input - Sophisticated Suggestions for Major League Baseball

Use cases: This application is designed to address two use cases

1. Buy a new player: When a manager wants to buy a new player for his team, he should be able to get a list of payers for a position, type of player and budget that he will be able to input.

2. Replace an existing player: When the manager wants to replace an existing player for a particular position, he should be able to get a list of players from other teams who are similar to this player in terms of performance and compare salaries.

Baseball field

Position Category	Positions in the adjacent picture
Pitcher	1
Catcher	2
Infielders	3, 4, 5, 6
Outfielders	7, 8, 9
Designated Hitter	Not pictured

Dataset description

Baseball data from http://seanlahman.com/baseball-archive/statistics is used for this project. This database contains pitching, hitting, and fielding statistics for Major League Baseball from 1871 through 2013. This database is also available in the form of an R package and the details can be found in https://cran.r-project.org/web/packages/Lahman/Lahman.pdf

Data dictionary: Scan the adjacent QR code for data dictionary.

Population of interest:

1. Players active in 2014 (defined as players with 2014 salary information) and their performance statistics from 2009-2013.
2. Batting: Players with AB (At-Bats) >= 30
3. Fielding: Players with number of games >= 5
4. Pitching: Players with IPouts >= 90

Performance metrics:

The performance statistics used for identifying similar players in use case 1 and the performance statistics displayed on the application are calculated as below:

- For Batting - OPS (On-base Percentage plus Slugging) = [(Hits + Walks + Hit by Pitch) / (At Bats + Walks + Hit by Pitch + Sacrifice Flies)] + (Total Bases / At-Bats)
- For Pitching - ERA (Earned Run Average) = (Earned runs * 9)/ (IPouts)

Performance statistics used for identifying similar players (or clusters) in use case 2 included all available outcomes. This includes PutOuts, Assists, Errors, and Double plays for fielding; StrikeOuts, and HomeRuns for pitching and HitByPitch, SacrificeFlies, and SacrificeHits for batting (see Data Dictionary for exhaustive list).

Example screen shot