

Project 2 (Luther) Proposal

Predicting Salaries of Impending (Summer 2019) NBA Free Agents

Young Jeong

Overview

In a typical American professional sport, the concept of Free Agency is an exciting proposition for players and teams alike. Teams are always aligning themselves with certain Free Agents depending on their needs. The major component of the conversations that happen between Free Agents and different teams is the salaries. My goal, as part of the fictional NBA consulting company, is to give a prediction on the upcoming 2019 NBA Free Agents based on performances of past NBA free agents. This will give teams or a specific team an insight into who they might target as their free agent signing this summer.

Data:

I found my data on various websites; basketball-reference.com is a sports statistics website that has data on individual players and their performance metrics over the past NBA (or even pre-NBA) seasons. Spotrac.com is another sports website that lists player contract information (and Free Agency Year list). Spotrac has Free Agent lists from 2011 on but I will search for previous dates on Google.

Details about the data

- Player stats:
 - Numerical (in Average per game): Points, Rebounds (Offensive, Defensive, Total), Assists, Turnovers, Fouls, Blocks, Steals, Free Throw Percentage, Field Goal Percentage, 2PT Percentage, 3PT Percentage, Free Throw Attempts, 3PT attempts, 2PT attempts, 2PT made, 3PT made, FT made, Field Goals Made, Field Goals Attempted and Minutes Played
 - Numerical (Cumulative for the season): Games Played and Games Started.
 - Categorical: Team, Position and Age, although age could be numerical.
- Salary data: Salary are given in average per year in dollars

As I progress through the data, I may find that not all of these statistics are necessary (some can be eliminated because they are redundant). Regardless, one thing I do want to do at this point is to aggregate Free Agent player data for the previous 3 seasons; the approach behind this is the fact that the longest injury for a basketball player is usually around a year (mostly either an Achilles Tendon or a Knee Ligament tear). 3 years of data should better combat against any outlier due to a player injury affecting their stats (1 year recovery + 1 year re-adjustment to NBA speed).

After the MVP:

Obviously contract negotiation in anything, let alone the NBA, comprises of different aspects; particular player-team relations, competitions between teams to sign a particular player, player's affinity to the location of the prospective team and its current players, and other data outside of on-court performances can change how much a player can command.

For my next iteration of the design, permitting time, I want to include player information such as immediate family size (# of children, marital status) at the time of signing (or current situation when making prediction), cost of living in their pre-Free Agent teams (and current teams for the predictions), and twitter mentions for a player's popularity during the season.

While I don't expect linear regression solely on the on-court metrics to be the most accurate results, it will be interesting to see if it can actually correctly predict (or get close to) the NBA Free Agents' next contract dollar amount, and what improvements can be made if adding off-court data.