

Linear Regression Analysis Project Writeup

Analysis on the NBA gameplay strategy

Josh Jingtian Wang

2/22/2022

Abstract:

The former general manager of the Houston Rockets, Daryl Morey, was the first GM to apply analytics to basketball. His basketball philosophy (dubbed “Moreyball”) favors three-point field goals and lay-ups over mid-range jumpers. It is evident that “Moreyball” has swept over the league in the past five years with most teams adopting the three pointer-dominant strategy, but it is inconclusive whether “Moreyball” has contributed to team wins. With basketball being one of the biggest sports in the world, it would be important to look at if “Moreyball” is actually “money”.

Design:

In this project, the Houston Rockets front office has asked us to investigate whether the MoreyBall philosophy contributes to the improve of a team’s net rating. Identifying factors that affect the net rating will improve the winning record and popularity of the team.

Data Description:

NBA Season Advanced Stats (2002-2022)

<https://basketball-reference.com/>

NBA Shot Location Data (2011-2022)

<https://basketball-reference.com/>

NBA Playtype Data (2015-2022)

<https://nba.com/>

Algorithm:

Features were first selected based on domain knowledge, then picked out by investigating multicollinearity via the correlation heatmap. Simple linear regression and Elastic Net regularization.

Tools:

Data retrieval was be done via BeautifulSoup and Selenium. Linear Regression was done vid Python's Scikilearn package. Data cleaning was be done via Python's pandas package. Data visualization was done with Python's matplotlib and seaborn packages.

Communication:

Please refer to the slides.