Forecasting NBA Scores

Final project I

Lightning presentation

David Goldman

Background

The NBA betting markets are known to be *highly efficient* in producing accurate betting lines (forecasts), yet there is surprisingly very little public information as how these forecasts are actually created

This project aims to gain a better understanding of the mechanics involved in forecasting NBA scores

Current Research Question

- Can NBA betting lines be *reasonably accurately* recreated through *simple* Data Science techniques?

Future Research Questions

- Can NBA betting lines then be matched further through more *advanced* Data Science techniques (i.e. machine learning, referee analytics, advanced statistical modeling techniques, etc.)

- Can NBA betting lines then actually be *beaten* (surpassed in accuracy) through betting into the cases where the advanced model *differs* from the betting line (and thus earn a positive ROI over the long-term)?

The Hypothesis

- NBA basketball scores can roughly forecasted using ***simple*** Data Science techniques

- Given that NBA betting lines are ***highly efficient***, a successful first attempt should be just to *roughly replicate* the market betting lines (within a reasonable margin)

The Data

- **Basketball-reference.com** – Best for the overall variety of data
- **Bigdataball.com** – Best for exportable box scores
- **NBA.com** – Best for team-level advanced analytics
- **Pinnacle.com** – Industry-leading sportsbook for current betting lines
- **Rotoword.com** – Best for current player news

Additional Reference

Within the widely respected book, "**Basketball on Paper**" the author identifies the "*four factors of basketball success*" and their associated weights of importance:

- **Shooting (40%)**
- **Turnovers (25%)**
- **Rebounds (20%)**
- **Free Throws (15%)**

… which could potentially serve as a starting point