Data 608

Professor Ferrari

Baron Curtin

Final Project Writeup

The NBA is currently undergoing a major shift in the way teams view how to win the most games as possible. This shift has been catalyzed by analytics. This “Moneyball” approach proposes the idea of the 50-40-30 rule, as a baseline measure of efficiency. The 50 indicates the rate, 50%, at which players will shoot the ball or convert for points within the restricted area. The restricted area is the area closest to the basket, nicknamed “the paint.” An expected point output, at 50% while taking 10 shots, will on average yield 10 points. Similarly, 40 represents the idea that players will shoot 40% inside the 3-pt arc, but outside the points. Expected returns at that FG% for 10 shots would be 8 points. 30 represents the 30% players shoot behind the 3-pt line, whose expected yield would be 9 points. Under this logic, it makes the mid-range shot, the area outside of the restricted area and inside the 3-pt arc, unfavorable and minimizes the chance of winning due to lower expected point totals. This 50-40-30 rule is just the baseline at which analytics operates because these have historically been the conversion rates for these areas.

The idea behind the analysis and visualization is to identify players who can convert mid-range shots at a rate of 50% or better, making those shots worth shooting. Consequently, it will also identify players who should be not be shooting mid-range shots. The visualizations are relevant because it shows players who are below average. Player who are below average in those areas should avoid those shots and either look to explore other ranges, or develop their games.

Visualization Logistics:

* Data Acquisition
  + Data was acquired from <http://stats.nba.com> and also enlisted the help of a publicly available Python package: nba\_api
* Data Dictionary
  + Fields included: Player Name, Team, Age, FGM (Field Goals Made), FGA (Field Goal Attempts), FG% (Field Goal %)
    - FGM, FGA, FG% were available for six distinct distance ranges: Less then 5 ft., 5-9 ft., 10-14 ft., 15-19 ft., 20-24 ft., and 25-29 ft.
    - Only the mid-range distances were used
* Visualization
  + Dash and Plotly were the visualization tools used
  + Timeframe: 2018-2019 NBA Seasons
  + User Controls: Player Name
    - Users are able to search for a player that had a minimum of 50 attempts from each mid-range area
  + Visual Information:
    - FG% Graph
    - Points per 10 Attempts Graph
    - Each graph shows metrics, via line, for the best of that category, the average of that category, and for the player the user selects in that category
    - Data is organized into box-plots to show the spread of the rest of the NBA (minimum 50 attempts)