INFO 3300, Project 2

Adam Mahar (ajm348), Kyle Brown (kmb266), Sean Viswanathan (sv287)

Introduction

The three-point line was introduced into the NBA in 1980. At its inception, hardly any teams were taking advantage of the shot, despite the fact that three-pointers have a much higher payoff at lower percentages than two-pointers; for example, shooting 28% from the three point line gives an expected return of .84 points, while a player must shoot 42% from within the three point line to achieve the same expected return. As time has gone on, players have adapted to take more three point shots, and teams have changed their strategies to focus on these higher value shots.

Since 1985, the NBA draft has included a lottery, in which the worst teams from the previous season get the first picks. In 1985, the number of draftees was seven, though it has increased in recent years. These draft prospects are thus generally brought on to the worst squads in the league, with the hope that they'll improve it in the coming years. Understanding how that draft pick will perform is important to a team's success and growth. Draft picks on average start to make significant impacts after their third season in the league.

Our project takes a closer look at these two big, innovative features in the NBA, in the context of what combination of these factors it takes to become a champion.

The Data

We scraped data extensively from basketball-reference.com, getting things including:

- 3 point percentage and attempts for championship-winning teams
- League average 3 point percentage and attempts
- Basic stats for each player on a championship-winning team
- Statistics about the distance from the basket at which teams take shots
- Wins by the teams that draft lottery picks were on

We also used nba.com to find

- Who was chosen in the draft lottery
- What year they were chosen in

We used Beautiful Soup (Python) to scrape the data. The data that we found is included as csv files in our submission. We also converted one file into JSON so that it could be more readily usable by our programs, and this is lottery_picks_stats_data.json. We performed this

conversion by means of a short script written in JavaScript. Most of the data was readily findable and usable in convenient tables on basketball-reference.

By utilizing multiple datasets from basketball-reference we were able to pull the important variables listed above to start building our story. We wanted to see the impact of the 3-point line along with attempts and percentages. This was the first important thing we needed to filter out of all of the shooting data that we pulled. Next was league averages and the shooting distances from different teams. For each championship team, we pulled their starting five players along with their statistics for that specific season and their career. Additionally, we scraped data from the top seven lottery picks and the wins for those teams that had the picks over the next ten to fifteen years.

We were able to pull all these different data sets together to make a cohesive and meaningful story to craft. By combining the teams wins with the teams lottery draft pick, we were able to demonstrate how that team is impacted by the draft pick.

Three Pointers on the Rise

Our first graph shows that three pointer percentages and number of attempts over the course of the season are both on the rise, and have been since the introduction of the three point line. It shows the championship winning team for each year, with the option to add league averages. It also allows the user to toggle between three point shooting percentage and number of three point attempts, and hovering over a team point gives year, team, three point percentage, three point attempts, and average age of each team. Hovering over an average point gives all of these except the team name, of course.

Looking at the three point shooting percentages within the past fifteen years, it's apparent that the teams that are winning championships are no longer able to do it on the power of their two point game alone, as the variance around the average line has shrunk, especially in the negative direction, while the last three years show us that exceptional three point shooting teams are beginning to consistently win, a trend that could foreseeably continue into the future.

The number of three point attempts that championship winning teams are taking does not appear to adhere to this trend-- not even close. Rather, the variance around this average line continues to be sporadic, and offers an area for future investigation. However, the data does imply that the amount of three point attempts have significantly increased for the league as a whole so it suggests that teams are moving towards small ball and outside the three point line. However, just because a team takes more attempts doesn't dictate their success towards a championship.

Player Cards

Our graphic below that graph gives an idea of the starting five players from each of those championship winning teams and their personal stats. The player cards are a comparison of the past thirty five years of teams showcasing their statistics and when they were drafted. The purpose of this visualization is to compare the elite players of the championship teams and what their playstyles are like. Additionally it is important to note players changes from season to season and how they adapt to the changing game.

As a whole, the player cards give insight to the team's typically most contributing players and how they transition to a more three point oriented game over time.

Moving Away From The Paint

Our second graphic focuses on the last fifteen years of the league as these are the years that the data is detailed enough and available for us to analyze. In this visualization we have different sized radial bands which represent how far away the team shoots from the basket. Within three feet, three to ten feet, ten to sixteen feet, sixteen feet to inside the three point line, and outside the three point line. We provide the percentage of shots of the total shots taken, for each range as well as what the team's percentage of made shots is from each section.

This goes hand in hand with the initial visualization that depicts the change of 3P% and 3PA over the past 35 years as we can see where more of the shots are being taken from for championship teams. As such, we can see teams are hitting more of their threes yet seem to be moving away from inside to outside the three line.

The Muddled Story of Draft Picks

This final graph, which works similar to how a horse race game might, takes the top seven draft picks from the selected year and by pressing the play button we can visually see how that player's team performs from year to year. Since the introduction of the draft we wanted to see how significant of an impact these draft players can make on NBA teams. The top seven is known as the "lottery" and the most valuable picks to have during the draft which is why we decided to track the top seven picks.

Very few draft picks will win championships within their first year or two of being drafted which correlates well with what the data demonstrates. However, when the player begins to reach their peak and develop a better game and abilities, we can see that their team begins to achieve more wins and has a better shot at winning the championship.

Additionally, we track when players get traded and switch teams in order to fully understand how their environment has changed and what sort of impact that has had on them. By comparing the top seven picks against each other we can see the difference between picks that are "busts" and don't contribute much to those that are significantly great such as Lebron James.

Starting the X-axis scale to be the year selected to 2014 (linear scale) allowed for enough data to show growth and change in the players and teams to understand and visually see what impact players have with their teams. Additionally, the Y axis stays constant as the number of wins for that year, where the maximum wins a team can achieve is 82. Using the teams' colors, we chart lines from year to year to show the change over time and have the "horses" (the players) move horizontally to the next tick (being the next season year) and vertically (for how many wins they have in the subsequent season).

Outside Code

We drew inspiration (and code snippets) for mouseover tooltips from http://bl.ocks.org/d3noob/a22c42db65eb00d4e369. For line animation, we used a trick from http://jaketrent.com/post/animating-d3-line/. We also made extensive use of, and are grateful for, both d3 and jQuery.

Data Citations

Most of our data came from various pages on:

http://www.basketball-reference.com

Our draft picks data came from:

http://www.nba.com/history/draft/all-time-lottery-draft-picks/index.html