Process Book[NBA 3 point shot revolution]

Author 1: Apoorv Bhambri email: apoorv.bhambri@hsc.utah.edu UID: u1377803

Author 2: Nick de Jonge email: nick.dejonge@utah.edu UID: u0283331

Author 3: Shreeman Gautam email: u1092041@utah.edu UID: u1092041

Github repository: https://github.com/Shreeman5/dataviscourse-undecided

Overview and Motivation:

A three-point field goal, sometimes called a three-pointer, or three is a field goal in a basketball game that is made from beyond the three-point line, a designated arc surrounding the basket. A successful attempt is worth three points, in contrast to the two points awarded for field goals made within the three-point line and the one point for each made free throw.

Historically the three point shot had been used very scarcely in the NBA (National Basketball Association). It was considered to be an inefficient shot only taken in the most desperate circumstances in a game (For instance, game situations in which the shot clock is winding down and it is hard to create penetration for a two-point shot or if a team is trailing behind late in games and it is not possible to make up the difference on two-point attempts). Consequently, not many players practiced or developed the ability to take and make three-point shots. It was a rare skill in the NBA, and the coaches discouraged players since it was considered a waste of possession.

Things changed remarkably, starting in the 2010s when teams began taking and making more three-point shots, and it has accelerated since 2015. For example, the ratio of 3-point shot to mid-range shot went from 0.72 in 2010 to 3.05 in 2021. Statistical models indicated that taking the 3-point shot made the teams more offensively efficient, completely revolutionizing the way basketball is played today. As a result of this change, basketball has become less physical and more skill-demanding. This is evident from the fact that the average height of an NBA player peaked in 2010 and has been steadily declining since. The NBA, which was once considered only for muscular and towering athletes, has become more accessible to people with different physical abilities. It has also improved the game's marketability since it is now played more dynamically. In this project, we will visualize this revolution in basketball using visualization techniques and animations.

On a personal note, Apoorv follows basketball, and it was his idea to research the 3-point revolution since he noticed that change, similar to how the aforementioned statistical models have seen it. Accordingly, this project is of great interest to Apoorv. Nick also loves basketball and is highly familiar with the intricacies of basketball. Given his personal interest in the sport of basketball and the trends of modern playstyles, Nick is looking forward to visualizing the transformation of 3-point shooting. As there is a large volume of available data, Nick also thinks this is the perfect idea for a data visualization project. Shreeman does not follow basketball but knows enough about the sport to research this topic. However, Shreeman follows soccer and has a fantasy premier league team, which means he is up to date with statistical models used to predict points for players in a given game week. Shreeman also loves working with data and is

looking to pursue a career in data analysis. Therefore, this project, albeit different from soccer, is of great interest to Shreeman.

Related Work:

https://www.nba.com/news/3-point-era-nba-75

https://datavizardry.com/2020/01/28/nba-shot-charts-part-1/

Questions/Objectives:

The project's objective is to create visualizations to understand the change in the basketball game due to the frequency of the 3-point shot taken in the last 12 years(2010-22). We will also compare visualizations for modern players and past legends.

There is still a debate in the basketball community on the perfect balance between different shots. Through our work, we will visualize the efficiency of all the 3 point shots using heatmaps of the court to find the perfect winning strategy.

Firstly, for each year from 2010 to 2022, using data compiled from all teams that played in that year, we aim to divide the basketball court, beyond the 3-point line, into five regions, and we aim to find the ratio of successful shots to attempted shots for each region. This will give us an understanding of what, statistically, the best region to shoot a three-pointer in a basketball court is. Our heatmap, which encompasses all regions, by virtue of appropriate colors, will testify to that ratio. For example, if region 1 has a higher percentage than region 2, region one will have more dark green spots than region 2. In this example, a darker shade of green, compared to a lighter shade of green, represents a more successful attempt.

Secondly, using data from 2010-22, we will construct a line graph showing three things for each year: % of 3 pointers, % of 2 pointers, and % of one pointers. Please note that these percentages add up to 100 for each year. So, the line graph will have three lines, and from this, we can test our hypothesis, which is that three-pointer attempts have risen over the last 12 years.

We will also learn how the use of statistics and data analytics has completely changed the game of basketball. The objective is to further the cause by using data analytics combined with data visualization to advance the game of basketball. Fascinatingly, data analytics and data science applications have revolutionized many industries, and sports is another example.

Finally, we aim to accomplish a functional website that uses visualization tools to capture and represent all the information for someone new to basketball to understand the strategies of the modern-day game in contrast to its history.

Data:

Data extraction was relatively easy to do, just a bit tedious to compile. Python has a library called nba_api and from that library we can use the module stats.endpoints.shotchartdetail. Using shotchartdetail, we can specify the type of data we want for any given season.

- One of the parameters within shotchartdetail is 'context_measure_simple' which can take on a wide array of values. One of the values is 'FG3A', which gives us 3 pointer data, attempted vs made shots. So, 'attempted' is always 1 but 'made' can be 0 or 1. This is all the data we need for making a heatmap. Another value that 'context_measure_simple' can take is 'PTS', which gives us attempted vs made information about 3 pointers, 2 pointers and 1 pointers. That is all the data we need for making a linechart.
- Another parameter is 'season_nullable', which can take season values. For instance, the values can range from '2010-11' to '2021-22'. This, combined with the previous parameter gives us all the data from 2010-22.
- Another parameter is 'player_id'. If this value is 0, all the info about all the players from a given season is found.
- Finally, when all of these parameters are combined, we can get the data into a json format. After using python's panda module, we can get the json data into csv files.
- All of the data files(for heatmap and for linechart) have been uploaded to github as zip files.

Feedback

Peer Review Feedback - 25th October 2022

The peer review feedback for our group was provided by Caleb Johnson, Natalie Platil, and Casey Lee(they are working on the project 'Energy Sources') on 25th October 2022.

Caleb suggested using three distinct colors for the attempted Field Goal Percentage (%FGA) graph for visual effect.
The consensus was that the idea of a Heat map is promising.
Casey also suggested using a Clustering algorithm to find the best (most efficient) shooting region on the court.
Caleb proposed the idea of having team data and players and league data.
All of them agreed that data should be relatively easy to find.
Caleb also expressed that red and yellow colors are the most suitable for a heatmap.
Finally, Casey, Natalie and Caleb appreciated the project and its scope and believed that the project has the potential to add more things.

TA Feedback - 28th October 2022

TA Tripti Agarwal loved our idea. She said that if we can pull it off, it would be a really nice visualization to look at. She did express concern related to the dataset. She encouraged us to have a prototype ready in its full functionality along with the cleaned-up data by the first milestone.

EXPLORATORY DATA ANALYSIS

DESIGN EVOLUTION

IMPLEMENTATION

EVALUATION