

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/shreemangautam/dataviscourse-undecided>

Background and Motivation

A three point field goal, sometimes called a **three-pointer** or a **three**, is a field goal in a basketball game that is made from beyond the three-point line, which is 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 shots). 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 started taking and making more three point shots and it has accelerated since 2015. As an 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, which has completely revolutionized 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 to be only for muscular and towering athletes, has become more accessible to people with different physical abilities. It has also improved the marketability of the game since the game is played more dynamically now. In this project we will visualize this revolution in the game of basketball using visualization techniques and animations.

On a personal note, Apoorv follows basketball and it was his idea to do research on the 3 point revolution since he has noticed that change, similar to how the aforementioned statistical models have noticed that change. 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 he knows enough about the sport to do research on this topic. However, Shreeman follows soccer and has a fantasy premier league team, which means that he is upto date with statistical models used to predict points for players in a given gameweek. 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.

Project Objectives

The objective of the project is to create visualizations in order to understand the change in the game of basketball 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 in basketball, and 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 5 regions, and our aim is 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 3 pointer in a basketball court is. Our heatmap, which encompasses all regions, by virtue of appropriate colors, will testify to that ratio. As an example, if region 1 has a higher ratio than region 2, region 1 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 years 2010-22, we will construct a line graph that will show 3 things for each year: % of 3 pointers, % of 2 pointers and % of a 1 pointer. Please note that these percentages add up to a 100 for each year. So, the line graph will have

3 lines and from this, we can test our hypothesis which is that 3 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 also further the cause for using data analytics in combination with data visualization to advance the game of basketball. Fascinatingly, the applications of data analytics and data science have revolutionized many industries and sports is another example.

Finally we aim to accomplish a working website that uses visualization tools to capture and represent all the information for someone who is new to basketball to get an understanding of the strategies of the modern day game in contrast to its history.

Data and Data Processing

In this article(<https://datavizardry.com/2020/01/28/nba-shot-charts-part-1/>), the author talks about wrangling shot data from the API of this website:

<https://www.nba.com/stats>. As of right now, we expect some data cleanup, if we are to follow the guidelines set by the author of the linked article.

We can use this website as well to get our data:

https://nbasavant.com/shot_search.php?hfST=&hfQ=&hfSZB=three%7C&hfSZA=&hfSZR=&ddlYear=2015&txtGameDateGT=2015-11-14&txtGameDateLT=2015-12-07&ddlGameTimeGT_min=&ddlGameTimeGT_sec=&ddlGameTimeLT_min=&ddlGameTimeLT_sec=&ddlShotClockGT=&ddlShotClockLT=&ddlDefDistanceGT=5&ddlDefDistanceLT=&ddlDribblesGT=&ddlDribblesLT=&ddlTouchTimeGT=&ddlTouchTimeLT=&ddlShotDistanceGT=&ddlShotDistanceLT=&ddlTeamShooting=&ddlTeamDefense=&hfPT=&ddlGroupBy=teamd&ddlOrderBy=field_goal_desc&hfGT=0%7C&ddlShotMade=&ddlMin=0

Must-have features

As of right now, our project has a small scope with room for increasing the scope given that this small venture succeeds. Accordingly, a must-have feature for our visualization is the number of regions in the court being five. Another must-have feature is the data gathered from years 2010-22. We will implement this as a dropdown, and for each year in the dropdown, we will show the heatmap generated from the 5 regions. Finally, having the line chart is another must-have feature since it tests our hypothesis of the 3 point revolution. Therefore, without the year dropdown, the heatmap for the 5 regions, and the linechart, we will consider our project to be a failure.

Optional features

In regards to scope, once we get our basic implementation up and running, we can feasibly increase the number of regions, which would mean that the heatmap can be more detailed. Related to this, once our 5 regions are defined, we can attach a click functionality to all the regions which will allow the user to focus on the heatmap of one region.

Also, we hope to add another dropdown, in which a player's 3 point success history is shown, regardless of year. So, for instance, we can show Stephen Curry's 3 point success history from the time he started playing until the most recent NBA season he played in. We also hope to show the 3 point history for the legends of the game(Michael Jordan, Kareem Abdul Jabbar etc).

Project Schedule

We have decided to meet on Fridays. As the semester goes on, we might meet more frequently.

Weekly deadlines:

Oct 28: Data wrangling/cleaning

Nov 4: Making sure data is in usable format

Nov 11(Project Milestone Due): Skeleton of the web interface

Nov 18: Load in data

Nov 25: Chart Visualization

Dec 2(Final Project Due): Write up

Sketches:

1. Ideas

Heatmap - 3-point shots only
- all shots (whole court)

Line graph - % of 3pt shots taken (3PA/FGA)
- all time 3pt field goals made

- Field goal attempt % comparison per type (paint, mid range, 3pt)

- Small subset of top players from different eras, plot their shooting heatmaps
- Entire league by season shooting heatmaps

2. Filter

Heatmap of 3pt shots at subregions outside 3pt line

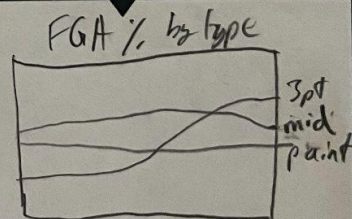
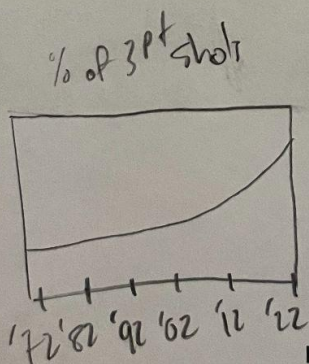
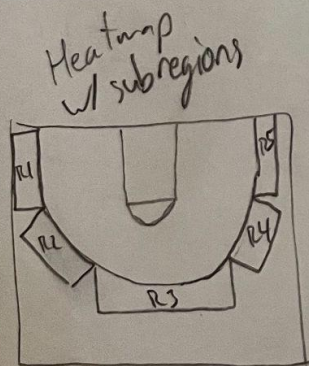
Plot both league averages per season and specific players.

3. Categorize

Heatmaps - 3pt subregion
- league average

Line graphs - % of 3pt shots per season
- FGA % compare by type

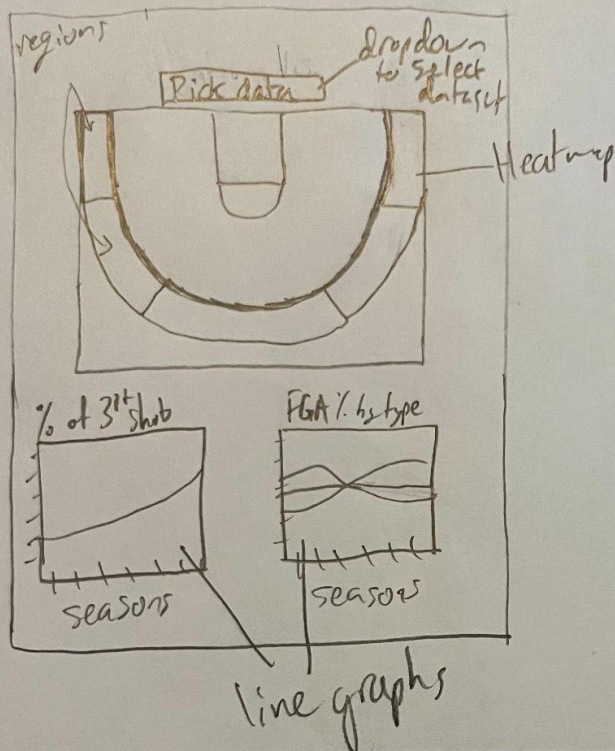
4. Combine and Refine



5. Question

most efficient position to shoot 3pt?

Layout



Title:

Author:

Date:

Sheet:

Task:

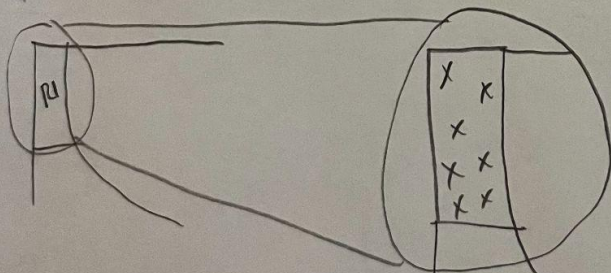
Operations

Dropdown:

- league averages by season
- player specific data (top players, legends)

Focus

Possible focus on subregions?



Show hotspots within subregion (highest % success, volume?)

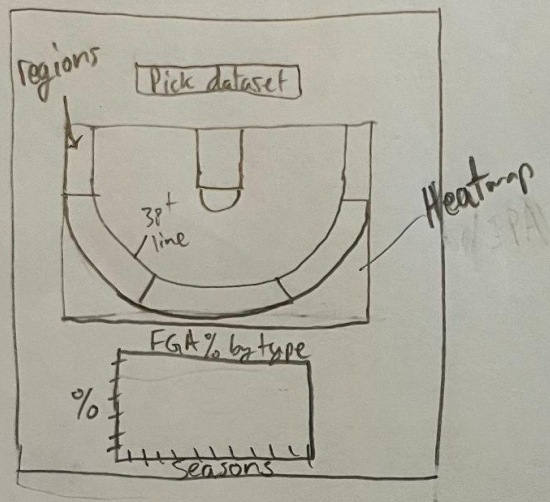
Discussion (Line graph)
metrics - $\frac{\text{FGA (3pt, mid, point)}}{\text{Total FGA}}$

Heatmap: $\frac{\text{points scored}}{\text{3pt attempt}}$

Datasets - league averages per season
- Ten players from different eras who were shooters

Layout

(Same as previous but removing
superfluous line graph showing
% 3pt attempts.)



Title:

Author:

Date:

Sheet:

Task:

Operations

- Dropdown data
Select

- Optional - focus/zoom
on region if
clicked

Focus

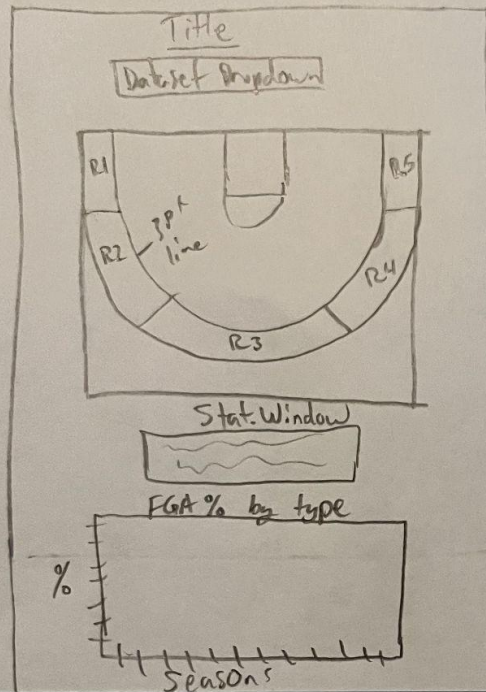
Optional feature (time permitting)

- Zoom on a region if it is
clicked
- show more detailed 3PA positions
along with statistics

Discussion

Removed 2nd graph
because it shows
essentially the
same data as
the one
remaining.

Layout



Title:

Author:

Date:

Sheet:

Task:

Operations

- Dropdown dataset selection
- Statistics window to show stats for chosen dataset
- Optional feature of region zoom-in

Focus

Optional feature:

- user clicks on a region and zoom-in on selected region
- zoom in will take place in same visualization window, over initial heatmap
- show more detailed stats for that region in stats window

Optional feature:

- Dataset dropdown includes selected "legendary" players

Detail

- User will select a dataset which will be displayed in heatmap.
- Statistics relevant to dataset will be shown in stats window below
- Time permitting, zoom feature to show more detail on selected region