Making 2023 NBA Draft Comparisons using NCAA and NBA Shot Charts

Nicholas Lukowsky

Department of Sport Analytics, Falk College of Sport and Human Dynamics, Syracuse University

ABSTRACT

Recent data innovation in basketball over the last ten years has introduced the idea of player-tracking data. This allows for the easy collection of a variety new data across different levels of basketball, one being shot charts. Each year the NBA Draft brings with it comparisons for its prospects with past and present NBA players. This study analyzes the current state of the 2023 NBA Draft collects shooting data from the NCAA and NBA 2022-23 seasons. A model is developed using Kernel density estimation and the Kullback-Leibler divergence to make draft comparisons based on similarities in the locations of field goal attempts in shot charts. A R Shiny application is created that allows users to see the results of this model along with producing shot chart visualizations for the NCAA draft prospects and current NBA players.

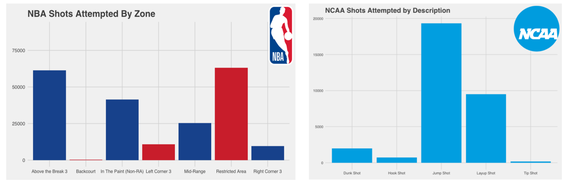
INTRODUCTION

The National Basketball Association (NBA) holds an annual draft where amateur talent is selected by teams in an organized and structured manner. NCAA players make up the largest proportion of these prospects and are the focus of this study. The order of selections is decided by a lottery system that promotes parity across the league. The draft takes place on June 22nd, 2023, and for many, the pre-draft process begins years before this date. During this time, teams get the chance to evaluate the potential future stars of their organizations. With this comes the unfair NBA player comparisons that are given to draft prospects as everyone attempts to grasp who they could become in their NBA career.

The use of data and analytics in basketball is on the rise and shot charts have become a prevalent type of analysis to evaluate players. Shot charts are useful in understanding a player's shooting ability, shot selection and overall scoring efficiency. This thesis aims to use shot chart data to assess prospects in the 2023 NBA Draft, with the goal of providing valuable insights for NBA teams and fans. This will take the bias out of the usual pre-draft predictions by forgoing traditional methods of comparing college and professional players. This is done by creating a tool that allows users to visualize the potential of their favorite amateur players through comparisons with current NBA players.

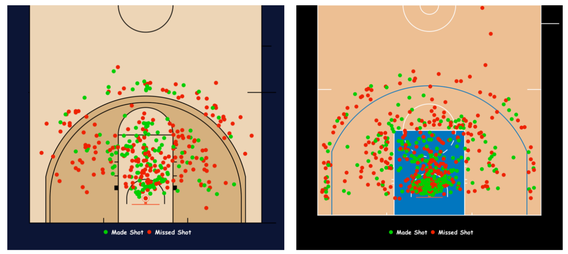
DATA & METHODOLOGY

Shooting data was collected for every NBA player to attempt 100 or more shots in the 2022-23 season using the R package, nbastatR. Conversely, the package, gamezoneR was used to collect the shot location data for around 100 of the best NBA Draft prospects from college.



The above graphs show summary visualizations for both the NBA and NCAA shots database. Both show the change in offensive philosophy in the recent years as teams focus more on jump shots and layups from the three point line and painted area, respectively.

Another R package, sportyR, was used to create the shot charts used in this study. To-scale playing surfaces were created and customized to match the look of their corresponding NBA teams. The previously collected shot data was adjusted so that it matched the dimensions of sportyR's charts.

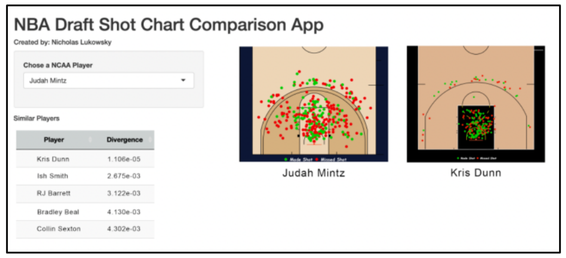


The above shot charts show an example analysis comparing Judah Mintz of Syracuse and Markelle Fultz of the Orlando Magic. Similarities and differences can be seen and overall the shot charts can be used as an aid to the upcoming model

As opposed to traditional draft comparisons this thesis used only data, specifically shot data to find equivalent NBA players to each draft prospect. A model was created using kernel density estimation and the Kullback-Leibler divergence. KDE uses smoothing to take into account nearby data and here it estimates the spatial distribution of a player's shot chart two-dimensionally. KLD is used to compare distributions and is implemented to measure the differences in the KDE of player's shot charts. It must be remembered that comparisons don't regard any defensive or non-shooting impact a player has within a game.

SHINY APP & DISCUSSION

To allow for users to make these comparisons and visualize the data, a R Shiny application was used. R Shiny is a web application framework for R that allow users a way to explore data in a simple, user-friendly way. Following the process from the Data & Methodology section, a user then begins by a selecting a NCAA player from a drop down of all tracked prospects. The model results are then generated along with the sportyR customized shot charts. The application itself includes a lot of reactive elements as the model is based on the NCAA player selected and the NBA chart are based on the model results.



The above screenshot continues with the Judah Mintz example from before as his shot data is passed through the model with a final comparison of 2022-23 Kris Dunn of the Utah Jazz. It's quite easy to see the top 5 comparisons and shot charts for a variety of draft prospects.

This study attempted to take the available data across different levels of basketball and find similarities amongst players transitioning from one league to the next. The results produced more variety in the comparisons than expected as in today's NBA a player's offensive role is more uniform than ever. Due to this, a position restriction was implemented into the shiny application for more accurate comparisons. Something that differs from non-data driven draft comparisons is more "star" NCAA prospects were compared to NBA players with smaller roles. This could be helpful, as not every draft prospect is going to live up to the expectations given to them and they may actually have secondary roles within an NBA team's offense.

Limitations of this thesis include the fact that shot charts aren't a proxy for skill and that only NCAA prospects are included in the draft player pool. Future studies could address these limitations as data becomes more accessible. Defensive shot charts could be added to account for defensive ability in the comparisons along with assist frameworks and other new-age data that better quantifies a player's overall game.

CONCLUSIONS

The analysis above tackled the challenge of projecting the career of NBA Draft prospects. By using the data of shot charts, NCAA players were examined side-by-side with every player currently in the NBA. A model containing density estimation and divergence found which NBA players had the most similar shooting dispersion through lowest divergence in shot charts. It was found that the comparisons might differ from what expected as the NBA has become more positionless, meaning every player is asked to have a more diverse shooting profile than years before. Ultimately, this analysis took a different route than a lot of other traditional thesis research, but through the creation of a R Shiny application presents an interactive tool for all to explore the idea of analytical NBA Draft comparisons themselves.

REFERENCES

1. Bresler, A. (2023). nbastatR: R's interface to NBA data. R package version 0.1.152, https://github.com/abresler/nbastatR
2. Drucker, R. (2022). sportyR: Plot Scaled 'ggplot' Representations of Sports Playing Surfaces. R package version 2.1.0, https://CRAN.R-project.org/package=sportyR
3. Lichtenstein, J. (2023). gamezoneR: A Scraping Interface for STATS LLC GameZone College Basketball Data. https://jacklich10.github.io/gamezoneR https://www.github.com/JackLich10/gamezoneR
4. Pickard, C. (2016, October 25). Nylon Calculus: Finding and quantifying similar shooters in the NBA. FanSided. Retrieved from https://fansided.com/2016/10/25/finding-quantifying-similar-shooters/
5. Reich, B. J., Hodges, J. S., Carlin, B. P., & Reich, A. M. (2006). A spatial analysis of basketball shot chart data. The American Statistician, 60(1), 3–12. https://doi.org/10.1198/000313006x90305
6. Samangy, D. (2022). 2022 NBA Draft App. Retrieved from https://dsamangy.shinyapps.io/2022_NBA_Draft_App/

ACKNOWLEDGEMENTS

I would like to acknowledge two of the authors of my sources, Chris Pickard and Dominic Samangy as main inspirations who created R Shiny applications related to the NBA themselves. Along with these authors, I would like to thank Dr. Paul for his guidance and understanding throughout the process of writing my thesis.