NBA+: An Information Visualization about current season of NBA

Pedro Tavares – 93103 Ricardo Cruz – 93118

Msc Robotics MSc Communication and Web Technologies

Information Visualization, University of Aveiro, 2022

Abstract

Data is precious and valuable, and nowadays, there is more and more data being generated, which introduces many problems. Raw data is nothing more than words or numbers, which by themselves do not have much meaning. The field of Information Visualization helps fixing this problem by giving meaning to the data through the best visualization techniques, so that our data becomes meaningful.

In this report we present the description of the work developed in the Information Visualization curricular unit, where we tried to solve a problem regarding meaningless data, more specifically about the current NBA season.

To solve this issue, we developed a web application, NBA+, that applies different data visualization techniques to give meaning to the raw data.

Through our tests, we were able to find that we achieved good results and that our application, which is for now a functional prototype, could with some more work integrate the market to help those who are curious about the NBA

# Motivation and objectives

The main motivation for choosing this theme is the fact that we are NBA fans, and we don't have time to watch the games because they are shown at dawn. Beyond that, and important for our decision, we did not find any data visualization application in our benchmarking research, that could provide efficient and detailed information about the current season of NBA.

That said, our main goal is to address the lack of an online platform that allows the visualization of this type of data in a detailed and efficient way, allowing us to draw conclusions and support decisions based on the analysis of the data visualizations developed.

Additionally, we have considered a list of goals for our platform, so that it can really help in the points that we consider crucial to understand the NBA season.

* View player and team statistics of the current NBA season.
* Compare players and teams.
* Relate variables to player and team statistics (e.g., How position affects shooting effectiveness).
* Understand which players are performing best within their team.

# Users and the Questions

When planning a project, it is very important to define well the target audience of the platform. That was our first step, to understand which audience our application should focus on. Afterwards, and based on the goals of our platform, we made a list of requirements that were mapped to questions that our app should be capable to answer through visualizations.

## Characterization of the users and their context

In this study, which later served as a support for the implementation, we found that the group of potential users of our platform would be something in the context of people who like the NBA and are familiar with the sport, however people that does not have time to watch it because the games play at dawn. Since they like to discuss it with their friends and they are not aware of the performance of players and teams, an application of this type is needed.

## Questions to Answer

Once we had our target audience defined, we then started thinking about the questions we wanted to answer with our platform. For these questions, we first started to think in a more general way and formulated a few generic questions.

* What are the intrinsic attributes of the player?
* Which player, x or y, has been performing better, offensively and defensively?
* Does the position and age affect the performance of the player?
* Which teams are doing better in the current season in terms of statistics?
* Who are the best players on the team, offensively and defensively?

After that, and since we wanted to do usability tests, we created tasks that are also examples of more specific questions that the platform can answer.

* Compare the average of assists of the players of position C (Center) and position PG (Point Guard). In which are the values of the distribution higher?
* Compare the average of Steals and Blocks per games between players in the range of 30-35 years and players in the range 20-25. In which are the values of the distribution higher?
* What is the “rebounding” stat for player Lebron James?
* Compare the values of the stat “outsideScoring” between the players Lebron James, Kevin Durant and Andre Drummond. Who has the higher score?
* What is the average 3 Points Misses and 3 Points Goals for the players Lebron James, Andre Drummond e Kevin Durant?
* What are the teams that have an average age between 23 and 25?
* Compare the distribution of fouls per game between the teams that have an average age between 23 and 25 and 27 and 29. In which are the values of distribution higher?
* Who is the player from the team Los Angeles Lakers that has the higher average of blocks per game?

# Dataset

The dataset used for the development of the platform contains information about the statistics of the players on the current season of NBA 2022-2023.

The data was obtained through the Kaggle online platform.

It contains the following data:

* Rk: Rank
* Player: Player's name
* Pos: Position
* Age: Player's age
* Tm: Team
* G: Games played
* FG: Field goals per game
* FGA: Field goal attempts per game
* FG%: Field goal percentage
* 3P: 3-point field goals per game
* 3PA: 3-point field goal attempts per game
* 3P%: 3-point field goal percentage
* 2P: 2-point field goals per game
* 2PA: 2-point field goal attempts per game
* 2P%: 2-point field goal percentage
* FT: Free throws per game
* FTA: Free throw attempts per game
* FT%: Free throw percentage
* TRB: Total rebounds per game
* AST: Assists per game
* STL: Steals per game
* BLK: Blocks per game
* TOV: Turnovers per game
* PF: Personal fouls per game
* PTS: Points per game

# Visualization Solution

To answer the problem and the questions mentioned in the previous section, we started by developing a low fidelity prototype, in paper, in order to perform usability tests and make any changes at a lower cost stage.

Regarding usability testing, in the first phase of evaluation we chose not to follow the traditional method, i.e., we did not create participant and observer forms, since we wanted the tests to be more open-ended to collect more qualitative data and, from there, draw insights of new functionalities for the platform. We did an approach of a semi-structured interview, with predefined questions, but also giving freedom for the user to speak and provide insights. At this stage of the project, we chose to follow this methodology because we considered the project could improve with this type of feedback.

In a second phase of evaluation, we chose to follow the traditional method, which will be mentioned in section 4.4

## Low fidelity prototype and user feedback

As said before, in order to perform usability tests, although not in the traditional way, we proceeded to the creation of a low fidelity prototype that is represented next.

In the homepage (Figure 1), we present a small contextualization to the user so he can situate himself in what is possible to do on the platform. Here, we also present cards so that the user can navigate to different sections of the platform, such as Positions Stats page, Players Stats page, Teams Stats page and to the game “Who are you?” (This idea was not developed in the low fidelity prototype since we were antecipating there would not be enough time to develop in the functional prototype, however, we presented it in the tests to hear the opinions of the users. The results will be discussed later in this section).

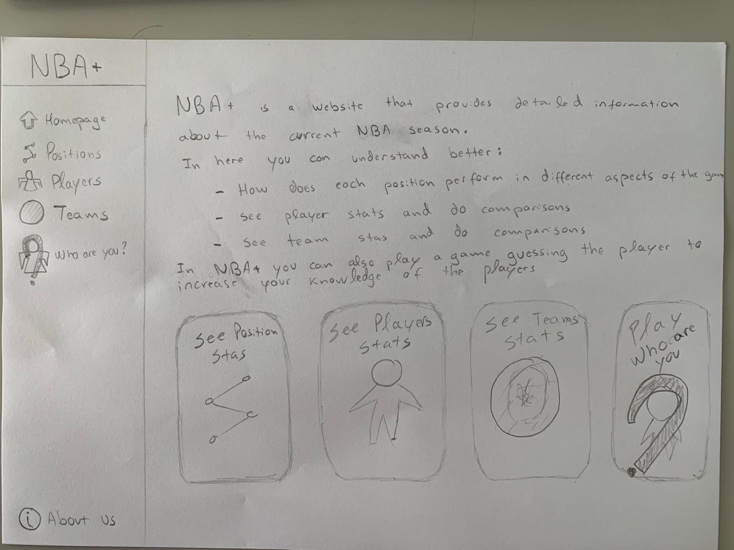


Figure 1. Homepage

This page (Figure 2) presents data regarding the different positions of the players, and how different positions influence the attacking and defensive attributes of the players.

Here, we had thought to have a creative way to choose the positions to compare, in this case an NBA half-court, and then have the data from the selected positions appear in the visualizations. For the visualizations we had thought of:

* Parallel plot to look for relations between the position and the other dimensions, which represent the players attributes.
* Stacked Bar Plot to quantitatively see the ratio between misses and hits in the target.
* DotPlot/BoxPlot/ViolinPlot to see the distribution of a certain dimension, which could be filtered.

Uma imagem com texto, quadro branco

Descrição gerada automaticamente

Figure 2. Position Stats Page

For the players page, we separated it into two phases.

The first phase would be when the user enters the page, and it is presented visualizations with the players statistics in general (Figure 3). For the visualizations of this phase, we considered adding a parallel plot to view all players and look for patterns in their attributes. Besides that, we also though of putting a dot plot/box plot/violin plot to check the distribution of the players in general, both at the attacking and defensive stats, where the x axis would have the attacking and defensive variables, respectively.

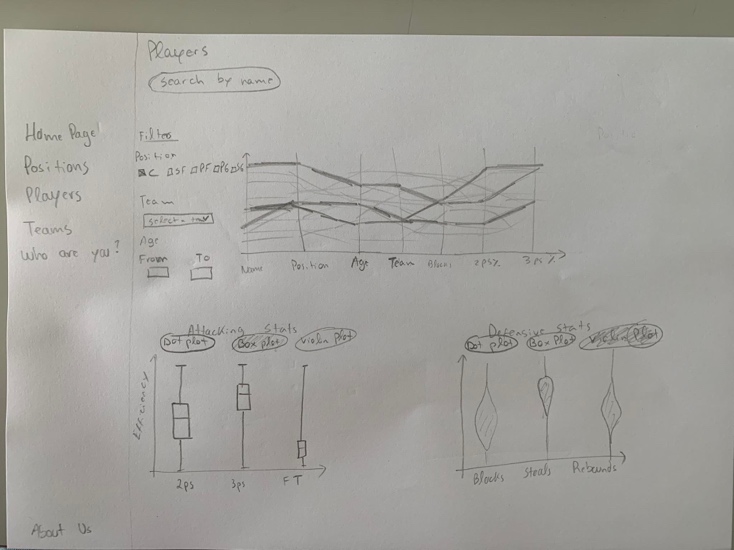


Figure 3. Players page when search input is not focused

The second phase would be when the user starts searching for a certain player. In this case, by accessing the search bar, the visualizations would be hidden and only the players whose search by name was verified would be displayed (Figure 4).

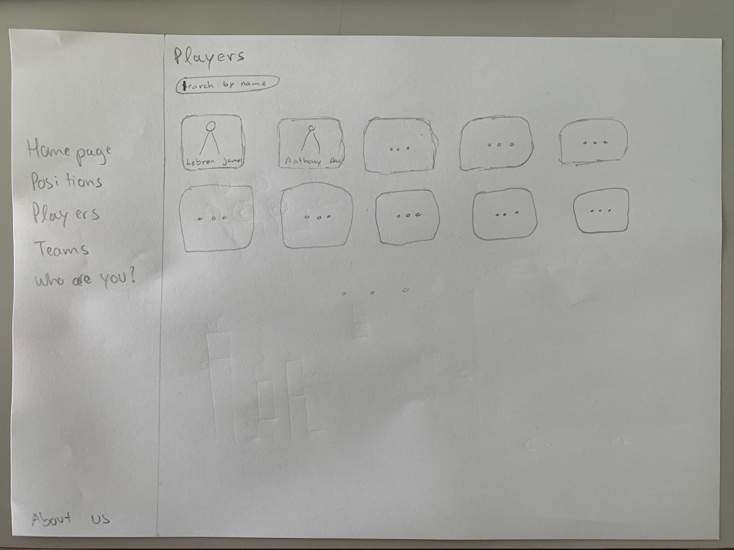


Figure 4. Players page when search input is focused

After the second phase mentioned above (Figure 4), if the user chooses the player he wants to analyze, he comes to the page of a specific player (Figure 5) which corresponds to a detailed analysis of the attributes of a specific player.

On this page we present several visualizations that help analyze a player in detail, such as:

* Radar plot, to see the player's attributes in a normalized way, where we can quickly observe how the player is doing on different levels.
* NBA half-court to represent the throwing effectiveness of the different zones of the field, however, we could only get data from three more general zones (3Points, 2Points and Free Throw).
* Points chart to see the exact data for attacking and defensive attributes.

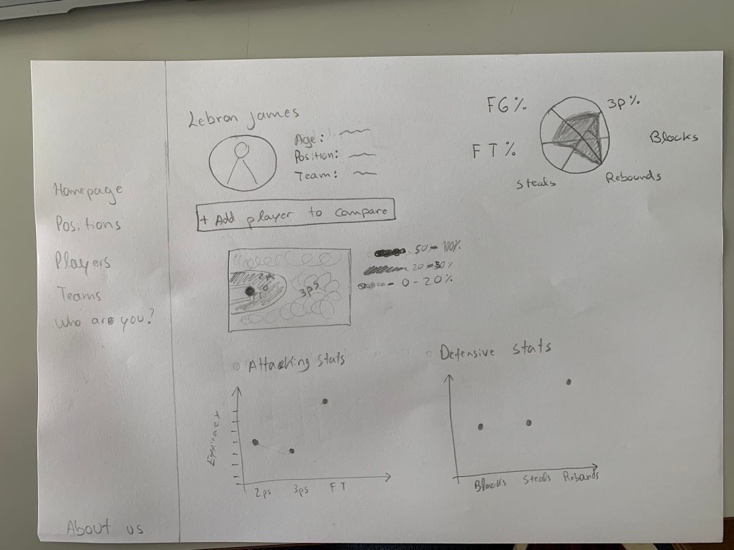


Figure 5. Player page with one player

On the previous page (Figure 5), it was also possible to add a new player to make the comparison. Figure 6 shows a sketch of what the page would look like for two players. The visualizations would remain the same, however, the entries of the new player would be added in order to make the comparison between the two.

Uma imagem com texto, quadro branco

Descrição gerada automaticamente

Figure 6. Player page in comparison mode with two players

Finally, for the team’s section, similarly to the players we have two pages:

* The page with all the players, which also has the two different phases, i.e. when you are searching for a team or when you are just looking at the visualizations.
* The page of a specific team.

For the page that contains all teams (Figure 7), we have thought of presenting a parallel plot that shows all the teams attributes and, a dot plot/box plot/violin plot to check the distribution of the attacking stats and defensive stats of the teams.

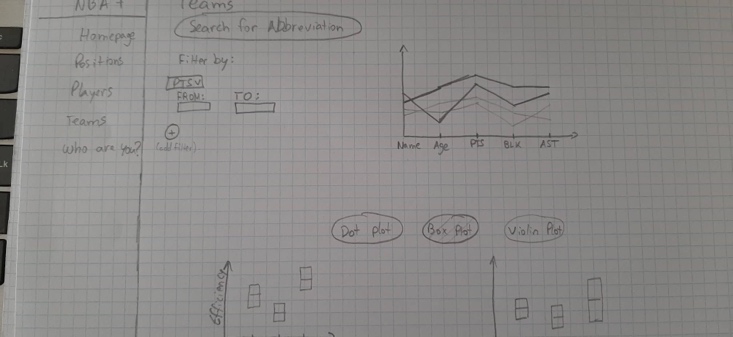


Figura 7. Teams page

Finally, for the team page (Figure 8), we added a parallel plot to check the players of that specific team and, a bar plot to check the attributes of the team quantitatively.

Uma imagem com texto, quadro branco

Descrição gerada automaticamente

Figura 8. Team page

After the low fidelity prototype was developed, we then moved on to the part of performing usability tests to validate the solution we adopted and how we could improve it.

As mentioned before, we did not follow the traditional methodology of creating a participant form and an observer form. Instead, and since we wanted to get more qualitative data, we mixed usability testing with semi-structured interviews, i.e. we prepared a set of open-ended questions and tasks to be performed, however, we also gave the user the freedom to talk about other topics and explore the prototype.

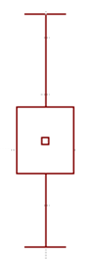
The feedback we got was as follows:

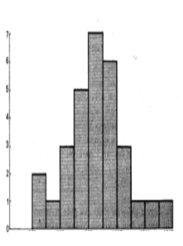
* Position stats page (Figure 2) can be encapsulated in all players page (Figure 3) because positions is heavily linked with the players. When asked to see relations between positions and the other variables, users tended to check the players page first.
* Overall, where we give the user the freedom to choose between dot plot, box plot and violin plot, remove the dot plot since the other ones are more complete.
* In the player page (Figure 5 & Figure 6), limit the comparison to three players so the radar plot does not get confusing.
* In all players page (Figure 3 & Figure 4) and all teams page (Figure 7), don’t show all players and teams when the search bar is focused. Instead, when user starts typing, show the ones that start as same as the user search.

With this feedback, we then started developing the functional prototype, presented in the next section.

## Functional prototype

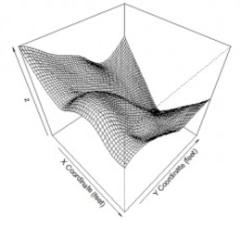
Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exercit­ation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulpu-tate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blan-dit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.





1. Visualizations to answer question Q1.

## 



1. Visualizations to answer the question Q2.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exercit­ation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulpu-tate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blan-dit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

## Implementation challenges

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat (<https://d3js.org/>). Ut wisi enim ad minim veniam, quis nostrud exercit­ation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulpu-tate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blan-dit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

## Evaluation and changes in the prototype

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exercit­ation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulpu-tate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blan-dit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

# Conclusion and Future Work

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exercit­ation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulpu-tate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blan-dit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exercit­ation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulpu-tate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blan-dit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

References

1. T. Minzner, *Visualization Analysis and Design*, A K Peters/CRC Press, 2014
2. A. Kirk, *Data Visualization : a successful design process*, Packt Publishing, 2012
3. J. Heer, M. Bostock, V. Ogievetsky, A tour through the visualiza-tion zoo. *Communications of the ACM*, vol. 8, n.1, pp. 59-67, 2010