NBA+: An Information Visualization Application about current season of NBA

Pedro Tavares – 93103

MSc Robotics and Intelligent Systems

Ricardo Cruz – 93118

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

1. 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.

2. 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.

2.1. 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 are familiar with the NBA but don't have time to watch it. 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.

2.2. 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.

1. 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?
2. 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?

3. What is the "rebounding" stat for player Lebron James?
4. Compare the values of the stat "outsideScoring" between the players Lebron James, Kevin Durant and Andre Drummond. Who has the higher score?
5. What is the average 3 Points Misses and 3 Points Goals for the players Lebron James, Andre Drummond e Kevin Durant?
6. What are the teams that have an average age between 23 and 25?
7. 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?
8. Who is the player from the team Los Angeles Lakers that has the higher average of blocks per game?

3. Dataset

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

The data was obtained through the Kaggle online platform [1].

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

Since we also wanted to do visualizations about the teams, we had to create methods that did the necessary calculations and grouping to get data from the teams.

4. 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

4.1. 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 anticipating there would not be enough time to develop in the functional prototype, however, we presented it in the tests to get the opinions of the users. The results will be discussed later in this section).
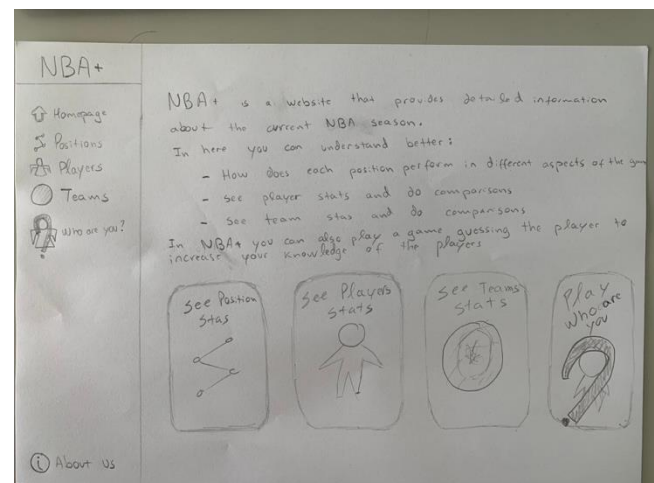

Figure 1. Homepage

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.
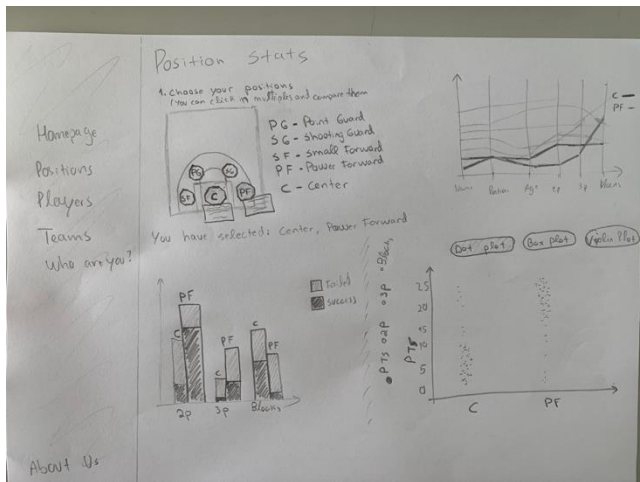

Figure 2. Position Stats Page

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 thought 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.
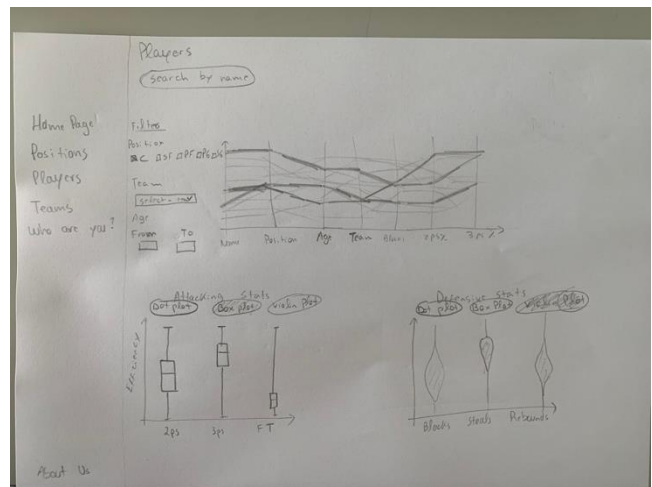

Figure 3. Players page when search input is not focused

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).
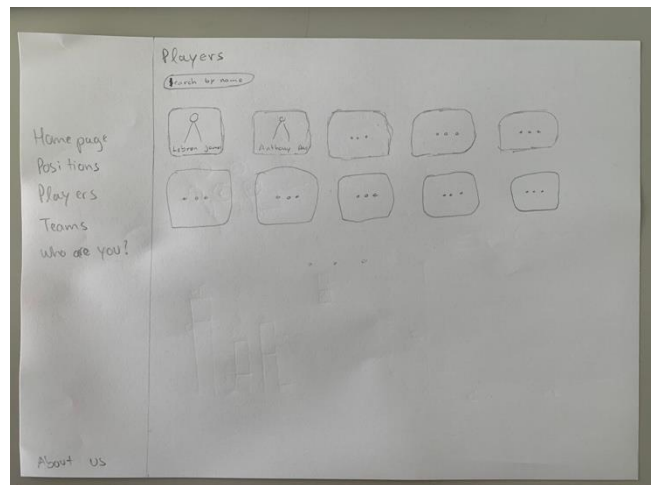

Figure 4. Players page when search input is focused

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.
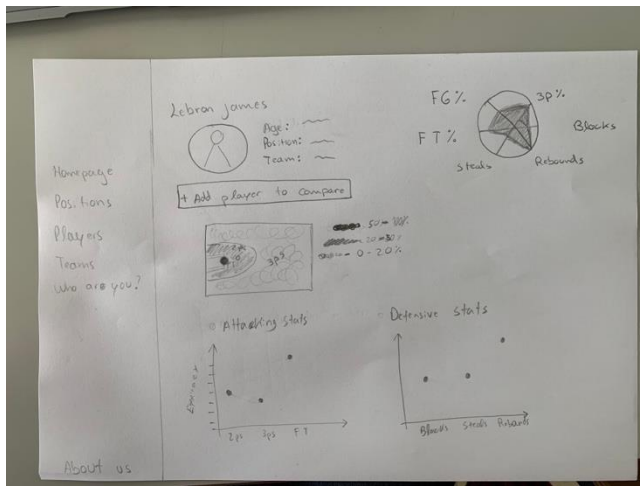

Figure 5. Player page with one player

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.
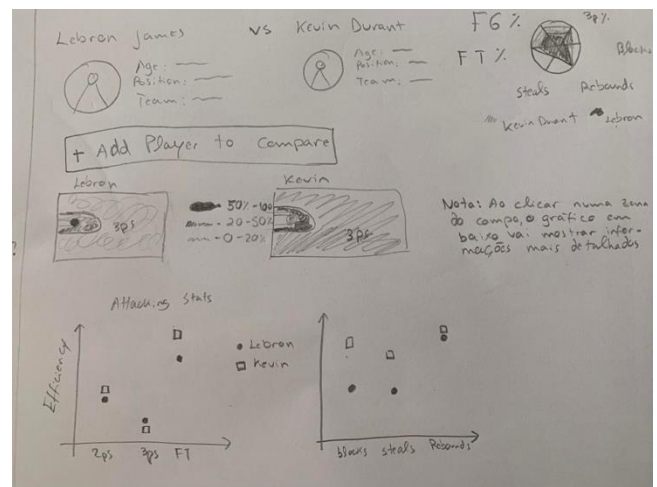

Figure 6. Player page in comparison mode with two players

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

- The page with all the teams, 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.

Figure 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.
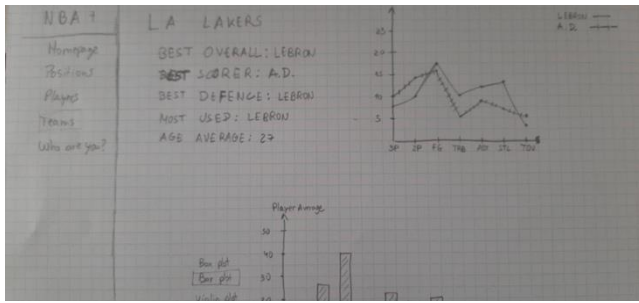
Figure 8. Team page

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 (section 2.2, generic questions), 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.

4.2. Functional prototype

Based on the prototype presented in the previous section, and considering the feedback obtained in the evaluation phase, we decided to make some changes:

- Removed the "Who are You" game because we were not going to have time to develop it.
- Encapsulated the positions page in the players page through filters.
- In the visual representations that included dot plot or box plot or violin plot, we chose to represent only the box plot, because we considered it as the most complete of the three. Taking into account the time for development, and the fact that the other were redundant, we made this decision.
- Limited the comparison between players to 3, so that the page wouldn't be too confusing.
- Removed the half-courts in the player comparison page, because in the data we had no way to make a heatmap, so it didn't make sense.
- Also on the player comparison page, we added a bar plot that quantitatively analyzes a certain attribute for the selected players, and this attribute can be changed through the filter select box or the radar plot labels.

In all pages, we present a sidebar that allows the user to navigate easily between pages.

For the homepage, we present a brief contextualization of what is possible to do on NBA+ as well as the quantification of the total number of

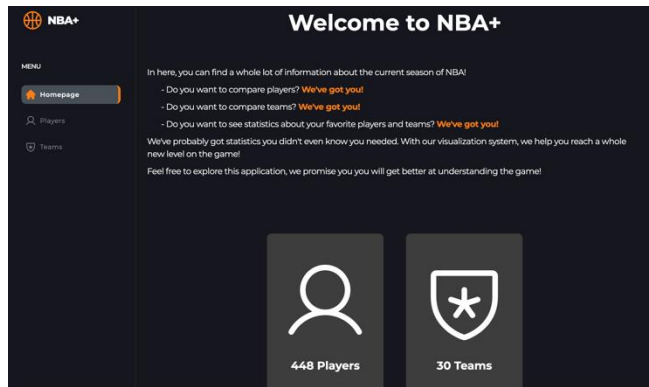players and teams. User can navigate to those pages by clicking in the cards.

Figure 9. Functional prototype Homepage

When clicking in players via the card or via the sidebar, the user gets in the page presented in Figure 10. In here, he is presented with a parallel plot grouped by positions and two box plot graphics to check the distribution of attacking and defensive stats.

Additionally, since there are 448 players, we added filters so the user can see the visualizations more clearly.

Figure 10. Functional prototype all players page. Answers Q1, Q2

In the case user overlays a certain line in the parallel plot, a tooltip appears showing the concrete stats for that player and highlights all the players of that position. Additionally, in the filters, the user can choose to see only a certain position to see more clearly the influence of the position on the other variables.
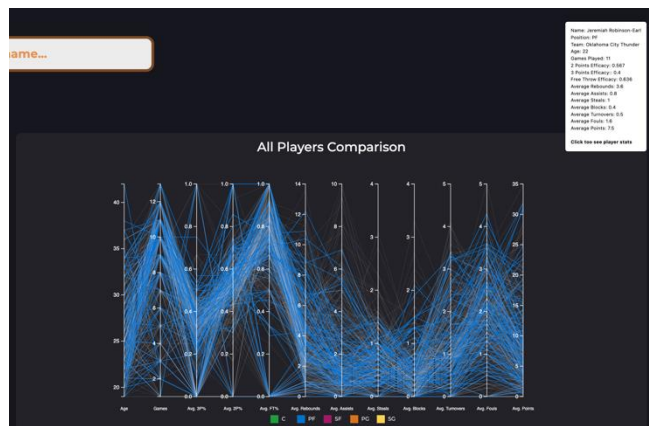

Figure 11. Parallel Plot for all players, filtered by PF. Answers Q2

When the user starts typing, we then add cards with the players name that start with the search user did.
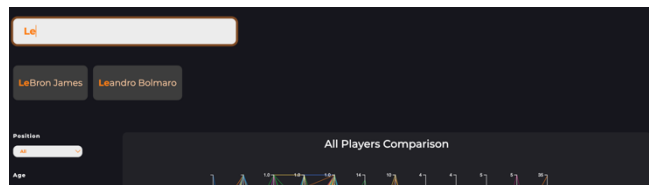

Figure 12. Functional prototype all players page when start typing player name

When the user enters in a certain player page, in this case the player Lebron James, he is presented with a radar plot to check the attributes of the player in several dimensions and a bar plot that specifies the quantified data used to normalize the value on the radar plot. The bar plot value can be changed either by the select box or by clicking in the radar plot attribute.
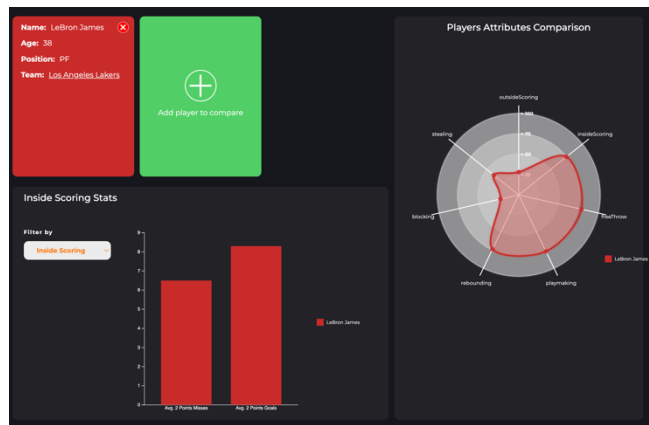
Figure 13. Functional prototype player comparison page with 1 player. Answers Q3

If the user wants to compare the current player with other players, he can add up to 3 by clicking in a new card. When clicked, a modal pops up, and presents a similar search bar with Figure 12. When there is more than one player, the visualizations remain the same, however they have data for the other players as well.

Figure 14. Functional prototype player comparison page with 3 players. Answers Q3, Q4, Q5

When user mouse gets over a certain point on the radar plot, a tooltip shows indicating the player selected and the attribute value.
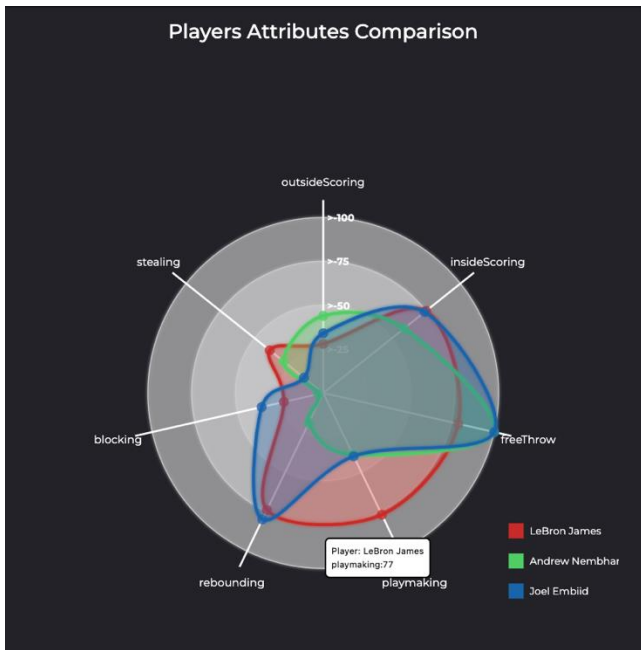

Figure 15. Radar Plot for players comparison. Answers Q3, Q4

By clicking in the label of a certain attribute, the bar plot of Figure 16 changes to present the data for that attribute.

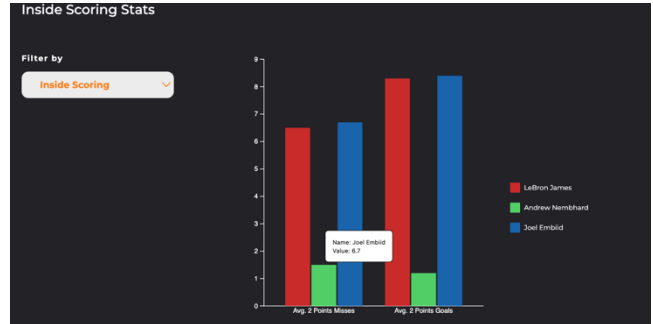When mouse is over a certain bar, it is indicated via a tooltip the exact value and the player.

Figure 16. Bar Plot for players comparison. Answers Q4, Q5

In the teams' sections, the user is presented with a similar interface when compared to players, presenting some filters and a parallel plot with all teams as well as two box plot graphics to see the distribution of the teams across the different attributes.

When user is hovering a certain team, a tooltip is shown indicating the exact values of the team for the different attributes shown.
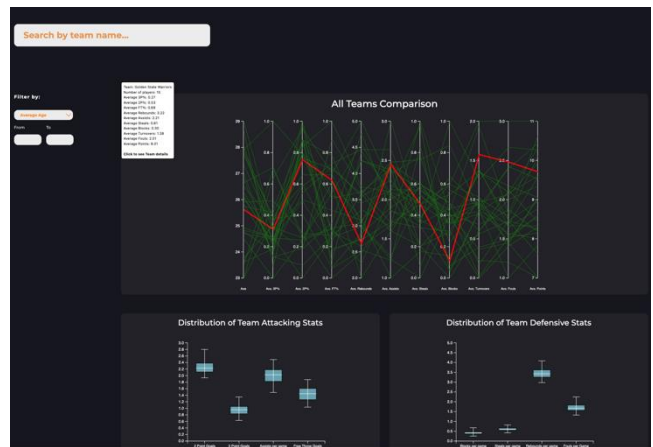

Figure 17. Functional Prototype all teams page. Answers Q6, Q7

Similarly with the search bar for players, the search bar for teams shows the team with an image when the users start typing.
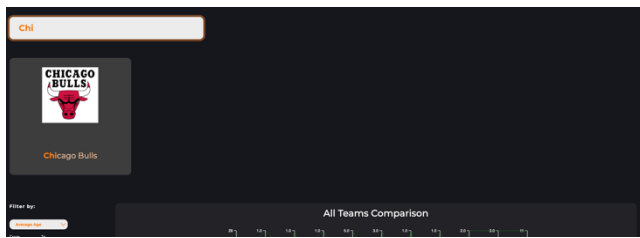
Figure 18. Functional prototype all teams page when start typing team name

Finally, in the team page, we present a similar interface with all teams page, but in this case the visual representations are for the players of that specific team. When user is over a certain line, the player is highlighted and a tooltip appears, so the user can check the exact attributes for that player.
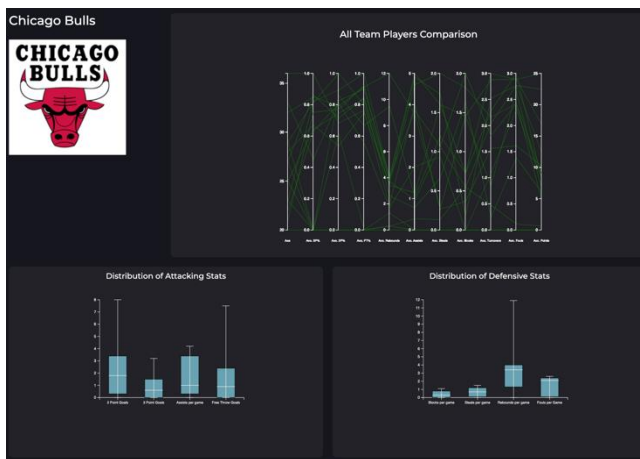

Figure 19. Functional Prototype team page. Answers Q8

4.3. Implementation challenges

The main challenge in developing the functional prototype was our ambition to combine Information Visualization with the platform's usability. Finding the balance between the most important part of the curricular unit, the visualization of information, but at the same time trying to develop a responsive platform that would adjust to the type of device used and that would provide a simple interface so that the users could enjoy the platform to the maximum was what caused the most difficulties, more specifically due to the development time.

Another challenge, related to the topic above, was the cost of implementation time caused due to the design and operation logic being done by us.

For the development, we ended up using the following technologies:

- Semantic HTML

- CSS

- Javascript

- D3.js

Regarding the graphical representations, as the d3 technology was new to both elements of the group, we ended up having some initial difficulties in creating the ideal graphics. All the representations were developed by us, however, we made use of resources that provided code examples that were able to inspire us in achieving better results. As resources we used the official website of the d3.js library [2] and the D3 Graph Gallery [3]

4.4. Evaluation and changes in the prototype

Regarding the evaluation of our functional prototype, we chose to perform usability tests, days before the presentation, in order to determine possible changes to the platform. We developed a participant form and an observer form and applied them to our target audience (people who like the NBA but don't have much time to watch the games and want to keep up with the statistics), and we were able to perform 5 tests with friends that fit the target audience.

During the test, the participants were asked to answer the questions presented in section 2.2, namely those identified as more specific. During the tests, data such as difficulty of execution, need for help, and the participant's ability to complete the task were noted.

At the end of the test, we presented a questionnaire [4] to the participants with 10 questions, in order to quantify the system usability, using the System Usability Scale (SUS) [5].

Regarding the results obtained during the tests, we were able to find that the participants did

most of the tasks effectively. However, we found that in tasks which required the identification of a certain attribute through the x axis, there were some difficulties, and there was even one participant who needed help with a task. In the participants' form we were also able to corroborate this fact, as most identified this as a possible improvement. In addition, we also had the suggestion of adding a tooltip in the parallel plots to see the player/team specific data that the mouse overlays.

Although in the tests we identified one problem that made it difficult to perform some tasks, we got a SUS score of 90.5, which translates to excellent usability. However, the fact that the participants were our friends and had a good technological literacy might have influenced the score.

5. Conclusion and Future Work

Regarding future work, since our second phase of usability testing was before the presentation, we would like to reiterate that it would be important to perform usability tests again, understanding if the problems identified from this phase, namely the x-axis of the graphs was solved.

Furthermore, we would also like, in a future phase, to conduct interviews with experts in the area of information visualization in order to be able to identify improvements and perhaps new graphs that would answer more efficiently the questions.

Finally, we would also like to look for an extra dataset that has more data about the teams, so that we can also expand the visualizations in that section.

In conclusion, we were very pleased with the final result. Although we had problems with the development time, we were able to develop graphical representations that answer the questions we thought about, thus responding to the challenge of the curricular unit, improving our skills with the d3.js tool, and especially applying the theory of information visualization, something very important nowadays with the exponential increase of data.

Furthermore, we were also able to improve our skills in developing a platform with good usability and corroborate these results by conducting usability tests.

6. References

[1] Kaggle, 2022-2023 NBA Player Stats, https://www.kaggle.com/datasets/vivovinco/20222023-nba-player-stats-regular

[2] D3.js, Data Driven Documents, https://d3js.org/

[3] The D3.js Graph Gallery, https://d3-graph-gallery.com/

[4] NBA+ Usability Questionnaire, https://forms.gle/UkGijM4iuyu74E4H7

[5] Beyond the NPS: Measuring Perceived Usability with the SUS, NASA-TLX, and the Single Ease Question After Tasks and Usability Tests, https://www.nngroup.com/articles/measuring-perceived-usability/