Data Visualization Project

2021/22 NBA season infographic

Tommaso Farneti (898148)
Fabio Galante (851242)
Gianni Eduard Balbin Canchanya (901609)

ACADEMIC YEAR 2022/23

The infographic is available at this **link**.

Contents

1 Introduction

Basketball is one of the most popular and widely viewed sports in the world.
Despite being so widespread, when it comes to basketball the first thought goes always to the **NBA** (National Basket Association), the professional league in North America which was formed in 1949 by the merger of two rival organizations, the National Basketball League and the Basketball Association of America.

There are several things that have helped NBA, and basketball in general, to become so popular. For example you have 48 minutes of action (if the clock is ticking the ball is in play) and it is the sport that probably requires more athleticism than any other. In addition, its games are broadcast in many other countries and this has led to a globalization of the brand. Furthermore, one of the greatest things about NBA are the iconic players the league has had, like Kobe Bryant, Lebron James, Michael Jordan obviously and many many others. Also the branding, the competition and the money behind the league are extraordinary.

1.1 NBA Teams

NBA is composed of 30 teams, 29 from the United States and only 1 from Canada, that are divided into two conferences of three divisions with five teams each:

Eastern Conference

- **Atlantic**
(Boston Celtics, Brooklyn Nets, New York Knicks, Philadelphia 76ers, Toronto Raptors)

- **Central**
(Chicago Bulls, Cleveland Cavaliers, Detroit Pistons, Indiana Pacers, Milwaukee Bucks)

- **Southeast**
(Atlanta Hawks, Charlotte Hornets, Miami Heat, Orlando Magic, Washington Wizards)

Western Conference

- **Northwest**
(Denver Nuggets, Minnesota Timberwolves, Oklahoma City Thunder, Portland Trail Blazers, Utah Jazz)

- **Pacific**
(Golden State Warriors, Los Angeles Clippers, Los Angeles Lakers, Phoenix Suns, Sacramento Kings)

- **Southwest**
(Dallas Mavericks, Houston Rockets, Memphis Grizzlies, New Orleans Pelicans, San Antonio Spurs)

1.2 NBA Season Structure

The NBA **regular season** starts in October and ends in April, with each team playing 82 games (41 each home and away). A team faces opponents in its own division four times a year (16 games). Each team plays six of the teams from the other two divisions in its conference four times (24 games), and the remaining four teams three times (12 games). Finally, each team plays all the teams in the other conference twice apiece (30 games)[1].
The eight teams (of each conference) with the best record make the playoffs, seeded from 1 to 8 based on best to worst record at the end of the regular season.
Teams finishing the regular season in positions 7 through 10 in the standings for each conference compete to determine the number 7 and 8 seeds in each conference's bracket tournament (NBA **Play-In** Tournament).

The NBA **playoffs** begin in April following the conclusion of the regular season with the top eight teams in each conference, regardless of divisional alignment, competing for the league's championship title. Seeds are awarded in strict order of regular season record (with a tiebreaker system used as needed).
Having a higher seed means a team faces a weaker team in the first round. The team in each series with the better record has home-court advantage[2], including the First Round. Therefore, the team with the best regular season record in the league is guaranteed home-court advantage in every series it plays.
The playoffs follow a *tournament* format and every round take the form of a best-of-seven series, meaning a team needs to win four games to progress to the next round.

1.3 Focus on 2021/22 season

Eastern Conference v·t·e

#	Team	W	L	PCT	GB	GP
1	**c** – Miami Heat *	53	29	.646	–	82
2	**y** – Boston Celtics *	51	31	.622	2.0	82
3	**y** – Milwaukee Bucks *	51	31	.622	2.0	82
4	**x** – Philadelphia 76ers	51	31	.622	2.0	82
5	**x** – Toronto Raptors	48	34	.585	5.0	82
6	**x** – Chicago Bulls	46	36	.561	7.0	82
7	**x** – Brooklyn Nets	44	38	.537	9.0	82
8	**pi** – Cleveland Cavaliers	44	38	.537	9.0	82
9	**x** – Atlanta Hawks	43	39	.524	10.0	82
10	**pi** – Charlotte Hornets	43	39	.524	10.0	82
11	New York Knicks	37	45	.451	16.0	82
12	Washington Wizards	35	47	.427	18.0	82
13	Indiana Pacers	25	57	.305	28.0	82
14	Detroit Pistons	23	59	.280	30.0	82
15	Orlando Magic	22	60	.268	31.0	82

Western Conference v·t·e

#	Team	W	L	PCT	GB	GP
1	**z** – Phoenix Suns *	64	18	.780	–	82
2	**y** – Memphis Grizzlies *	56	26	.683	8.0	82
3	**x** – Golden State Warriors	53	29	.646	11.0	82
4	**x** – Dallas Mavericks	52	30	.634	12.0	82
5	**y** – Utah Jazz *	49	33	.598	15.0	82
6	**x** – Denver Nuggets	48	34	.585	16.0	82
7	**x** – Minnesota Timberwolves	46	36	.561	18.0	82
8	**pi** – Los Angeles Clippers	42	40	.512	22.0	82
9	**x** – New Orleans Pelicans	36	46	.439	28.0	82
10	**pi** – San Antonio Spurs	34	48	.415	30.0	82
11	Los Angeles Lakers	33	49	.402	31.0	82
12	Sacramento Kings	30	52	.366	34.0	82
13	Portland Trail Blazers	27	55	.329	37.0	82
14	Oklahoma City Thunder	24	58	.293	40.0	82
15	Houston Rockets	20	62	.244	44.0	82

Figure 1: Regular Season by Conference - Wikipedia

[1]National Basketball Association - Regular Season (Wikipedia)

[2]The team having the home-court advantage plays the first two games of the series at home. The following two games will be one at home and then away. If another game is needed, it is played at home.

Notes

z - Clinched home court advantage for the entire playoffs

c - Clinched home court advantage for the conference playoffs

y - Clinched division title

x - Clinched playoff spot

pi - Clinched play-in tournament spot

* - Division leader

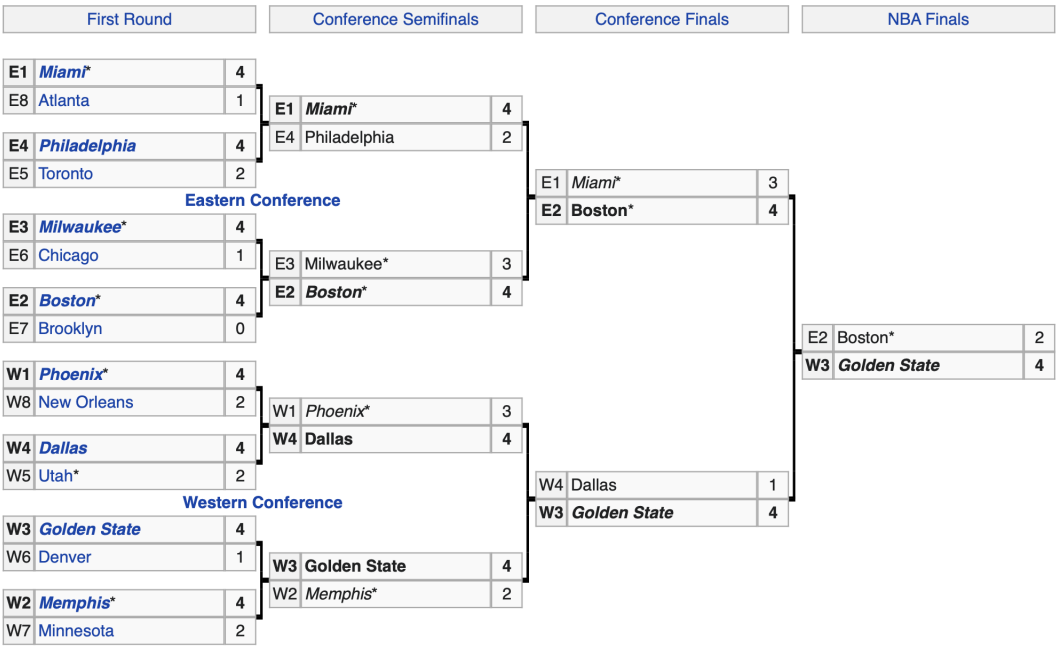


Figure 2: Playoffs Bracket - Wikipedia

Notes

* - Division Winner

Bold - Series winner

Italic - Team with home-court advantage

2 Dataset

2.1 Datasets used

We imported data about 2021/22 NBA season mostly from **Basketball Reference**. This website offers a large amount of data about players and teams, including traditional and advanced statistics for every season. We imported and merged them together obtaining the following datasets:

- Complete Player Stats 2021-22
- Playoff Player Stats
- Complete Team Stats 2021-22
- Playoff Team Stats

Another source we used is **NBA-players-data** from Kaggle; here we found information about nationality, height and weight of every player and we called this dataset:

- Player With Nationality

The last source we relied on is an R-package called **ballr**, which allows you to access data from **Basketball Reference** using simple functions. The one we used is **NBASeasonTeamByYear**, a function that imports data about every single game of 2021/22 regular season, like the scores and the results. We called the obtained dataset:

- Games

2.1.1 Variables description

We worked with a high number of variables, some more common and some a little bit more sophisticated. In order to let everybody understand our work we will explain them here:

- Rk: Rank
- G: Games
- MP: Minutes Played
- FG: Field Goals
- FGA: Field Goal Attempts
- FG%: Field Goal Percentage
- 3P: 3-Point Field Goals

- 3PA: 3-Point Field Goal Attempts
- 3P%: 3-Point Field Goal Percentage
- 2P: 2-Point Field Goals
- 2PA: 2-point Field Goal Attempts
- 2P%: 2-Point Field Goal Percentage
- FT: Free Throws
- FTA: Free Throw Attempts
- FT%: Free Throw Percentage
- ORB: Offensive Rebounds
- DRB: Defensive Rebounds
- TRB: Total Rebounds
- AST: Assists
- STL: Steals
- BLK: Blocks
- TOV: Turnovers
- PF: Personal Fouls
- PTS: Points
- Age: Player's age
- W: Wins
- L: Losses
- MOV: Margin of Victory
- SOS: Strength of Schedule; a rating of strength of schedule. The rating is denominated in points above/below average, where zero is average.
- SRS: Simple Rating System; a team rating that takes into account average point differential and strength of schedule. The rating is denominated in points above/below average, where zero is average.
- ORtg: Offensive Rating An estimate of points produced (players) or scored (teams) per 100 possessions

- DRtg: Defensive Rating An estimate of points allowed per 100 possessions

- NRtg: Net Rating; an estimate of point differential per 100 possessions.

- Pace: Pace Factor: An estimate of possessions per 48 minutes

- FTr: Free Throw Attempt Rate Number of FT Attempts Per FG Attempt

- 3PAr: 3-Point Attempt Rate Percentage of FG Attempts from 3-Point Range

- TS%: True Shooting Percentage A measure of shooting efficiency that takes into account 2-point field goals, 3-point field goals, and free throws. Offense Four Factors

- eFG%: Effective Field Goal Percentage This statistic adjusts for the fact that a 3-point field goal is worth one more point than a 2-point field goal.

- TOV%: Turnover Percentage An estimate of turnovers committed per 100 plays.

- ORB%: Offensive Rebound Percentage An estimate of the percentage of available offensive rebounds a player grabbed while they were on the floor.

- FT/FGA: Free Throws Per Field Goal Attempt Defense Four Factors

- eFG%: Opponent Effective Field Goal Percentage

- TOV%: Opponent Turnover Percentage

- DRB%: Defensive Rebound Percentage An estimate of the percentage of available defensive rebounds a player grabbed while they were on the floor.

- FT/FGA: Opponent Free Throws Per Field Goal Attempt

- PER: Player Efficiency Rating A measure of per-minute production standardized such that the league average is 15.

- TRB%: Total Rebound Percentage An estimate of the percentage of available rebounds a player grabbed while they were on the floor.

- AST%:Assist Percentage An estimate of the percentage of teammate field goals a player assisted while they were on the floor.

- STL%: Steal Percentage An estimate of the percentage of opponent possessions that end with a steal by the player while they were on the floor.

- BLK%: Block Percentage An estimate of the percentage of opponent two-point field goal attempts blocked by the player while they were on the floor.

- USG%: Usage Percentage An estimate of the percentage of team plays used by a player while they were on the floor.

- OWS: Offensive Win Shares An estimate of the number of wins contributed by a player due to offense.

- DWS: Defensive Win Shares An estimate of the number of wins contributed by a player due to defense.

- WS: Win Shares An estimate of the number of wins contributed by a player.

- WS/48: Win Shares Per 48 Minutes An estimate of the number of wins contributed by a player per 48 minutes (league average is approximately .100)

- OBPM: Offensive Box Plus/Minus A box score estimate of the offensive points per 100 possessions a player contributed above a league-average player, translated to an average team.

- DBPM: Defensive Box Plus/Minus A box score estimate of the defensive points per 100 possessions a player contributed above a league-average player, translated to an average team. BPM – Box Plus/Minus A box score estimate of the points per 100 possessions a player contributed above a league-average player, translated to an average team.

- VORP: Value over Replacement Player A box score estimate of the points per 100 TEAM possessions that a player contributed above a replacement-level (-2.0) player, translated to an average team and prorated to an 82-game season. Multiply by 2.70 to convert to wins over replacement.

- Country: Native Country

- Draft Year: Year the player got drafted in the NBA

- Squadra: Interested team playing

- Avversaria: Opposing team playing

- Punti: Point scored by the interested team

- Puntiavv: Point scored by the opposing team

- Casa: If the value is 1 the team is playing at home, if the value is 0 the team is playing a road game.

- Giornidiriposo: Rest days before the game

2.2 Elaborations done on the dataset

The datasets we created had some little problems that we had to solve:

1) The dataset *Complete Player Stats 2021-22* showed players and their statistics in a single row, but for players who changed team during the year more rows were shown, each one containing the statistics recorded by the player during the time spent on that team and one of them showing the total statistics recorded during the season. Those rows are easy to identify because the value of the attribute team is TOT. To solve this problem we created a new dataset called *Complete Player Stats 2021-22 No Rep*, where we kept all the rows including players who did not change team and rows whose team attribute value is TOT.

2) We had to change the dataset *Games* in order to make it more suitable for our analysis. First we had to create a new attribute that contained the cumulative sum of the results for every team, so that we could see how that changed, and create a ranking; we called this dataset *Games2*. Than we had to rearrange the order of the observation, since they were ordered by team instead of being ordered by data (day of the year). We changed this and called the new dataset *Games3*.

3) Last but not least, to create a certain type of chart we needed to have all statistics on the same scale, so we decided to normalize the chosen variables and created three new datasets, each one used for the corresponding chart: *Attack Player Stats Radar*,*Defense Player Stats Radar* and *Overall Player Stats Radar*.
   To normalize we applied a simple statistic rule:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

   where X is an observation.

3 Realization of the Data Visualization

3.1 Goal of the infographic

The infographic we present wants to summarize the last NBA season showing statistics of both the teams and the players. We want to show which teams and players performed better (or worse), how the standings changed over time and the main differences between each team or player.

3.1.1 Target

The target of this infographic is every person passionate about NBA or sports in general, but also all the people who, for any reason, want to know how the 2021/22 regular season and playoffs went. The infographic is built to be easily used and understood, not only by sports experts or statisticians.

3.2 Teams comparison

The *Teams Comparison* graph allow the comparison between all the teams of the league. It is a scatterplot in which is possible to manually select which variable to put on the X and on the Y axis. As we can see, on the left we have the list of teams, and there is the possibility to click on one of them to highlight it inside the graph.
If you hover your mouse over a logo you can actually see the list of the values for the chosen variables.

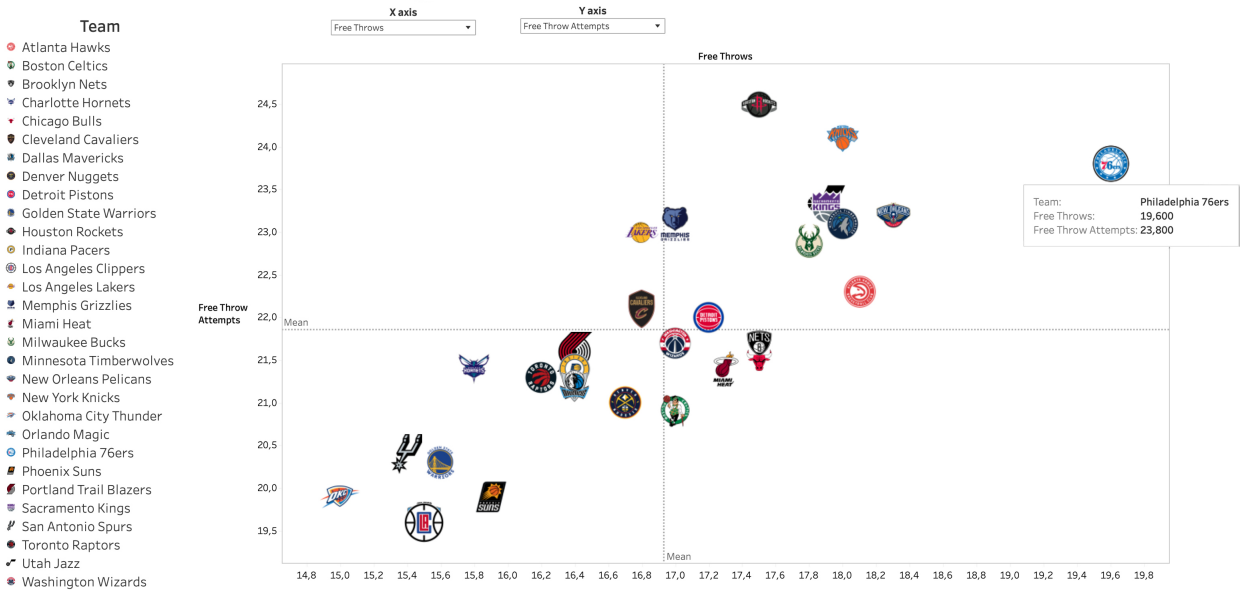


Figure 3: Teams Statistics graph - Free throws , Free throws attempts

For example we can notice how Philadelphia was the team which scored the most free throws.

3.3 Players comparison

3.3.1 Players Statistics per Position

The first graph with which we can make a comparison between players is the Players Statistics per Position, which means that we can discriminate players by their position on the field. In fact, in this case we are able to select the players that play as Center, as Power/Small Forward or Shooting/Point Guard.
Here we can select which variables to put on the X and Y axis too. Also, each player is represented with a circle colored as the team in which he plays. We can easily notice which players stand out from the others for many different statistics.

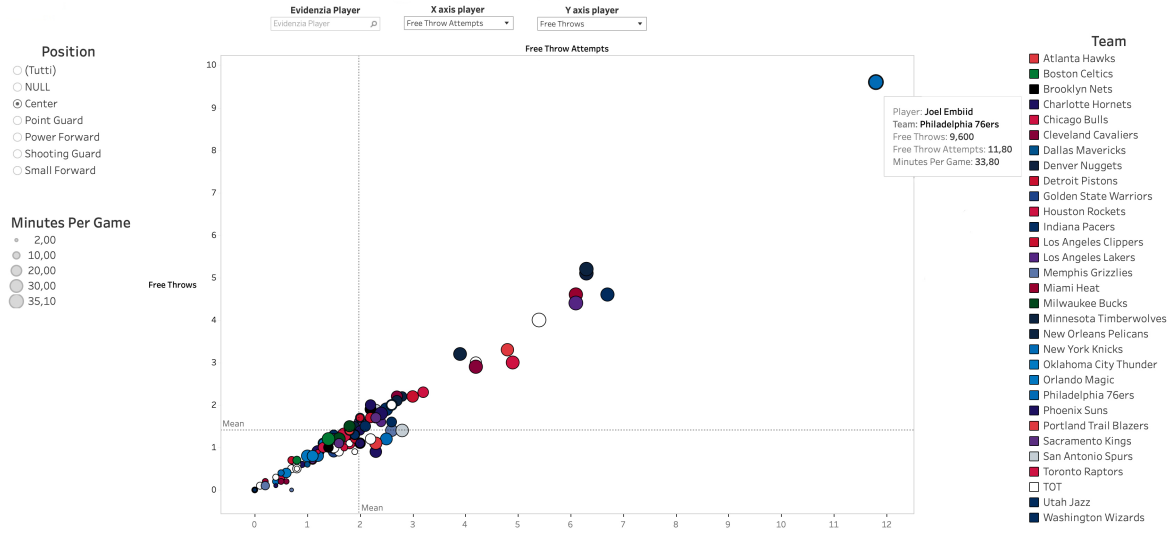


Figure 4: Players Statistics per Position

Like we did in the Comparison Team graph, we put Free Throws on the X axis and the Free Throws Attempts on the Y axis. We can notice for example that Joel Embiid (PHI) is one of the best free throw shooter with an average of 9.6 free throws scored on 11.8 attempts per game. Embiid being an outlier for free throw is one of the reasons Philadelphia has the most made free throws.

3.3.2 Radar Chart

Radar Charts are used to compare two or more items or groups on various features or characteristics. In this case we compare player based on:

- Overall statistics: Assists per game, Blocks per game, Points per game, Steals per game, Turnovers per game, Total Rebounds per game.
- Offensive statistics: Field Goal %, Free Throws per game, Points per game, Usage Percentage, 3-Point Field Goals per game, 2-Point Field Goals per game.
- Defensive statistics: Blocks per game, Defensive Rebounds per game, Defensive Rebound %, Defensive Win Shares, Personal Fouls per game, Steals per game.

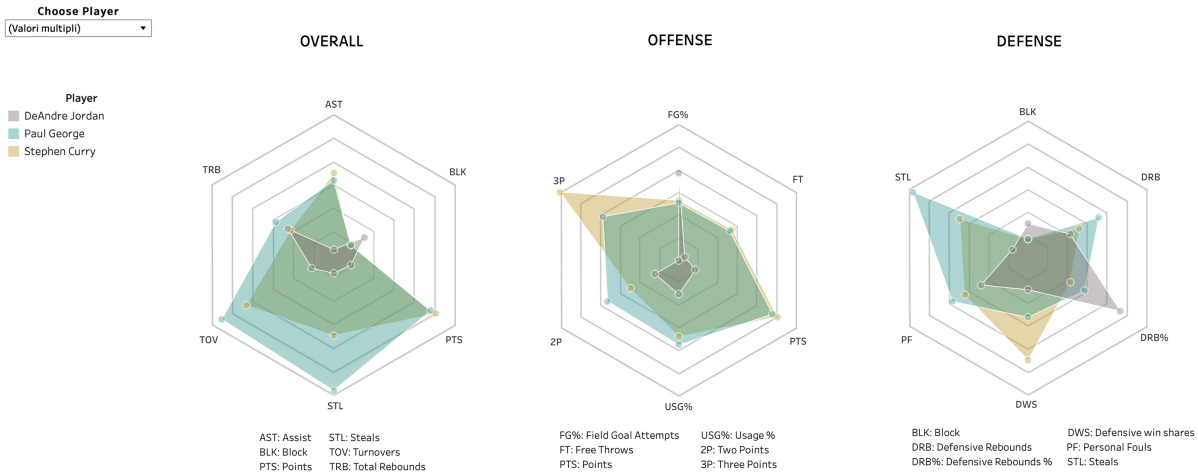


Figure 5: Radar Chart

In figure (5) we can see a comparison between Stephen Curry, Paul George and DeAndre Jordan using the Radar Charts. We can notice how Stephen Curry is the one with more Points Per Game and more 3-Point Goals Per Game, but with less valuable defensive statistics. Paul George is the one with more Steals and Turnovers per Game, while DeAndre Jordan stands out from the others for Defensive Rebounds percentage.

3.3.3 Players' Shots Statistics

Then we focused on the statistics about the players' shots during the playoffs. This graph compares all the players based on Field Goals, 3-point, 2-point and Free Throw shots. It's also possible to discriminate by minutes played per game using a dedicated slider.

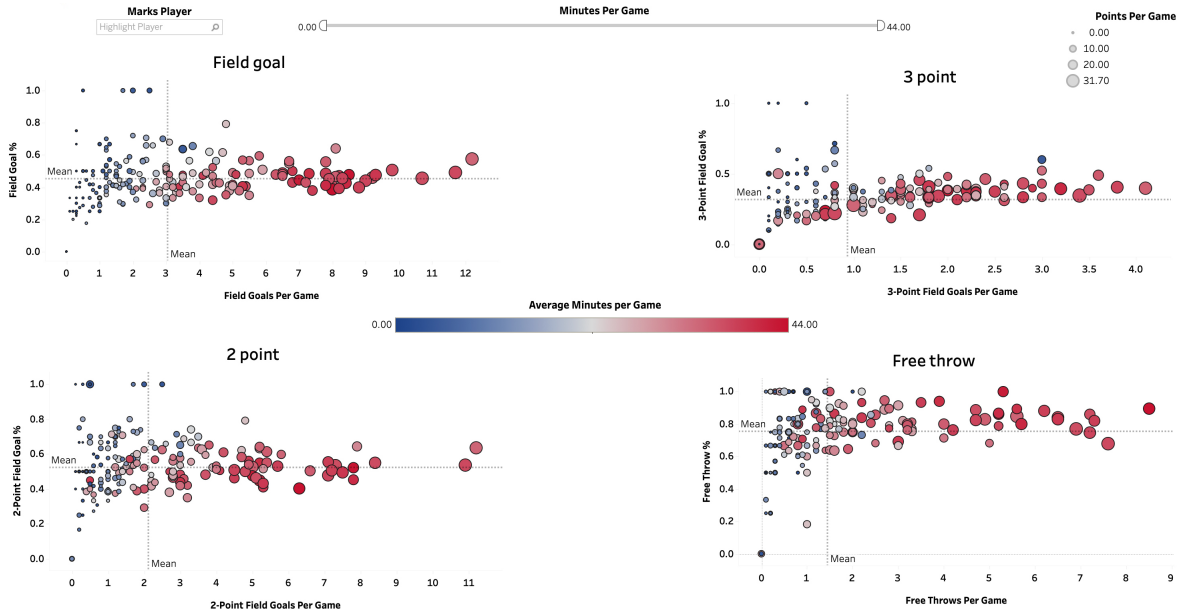
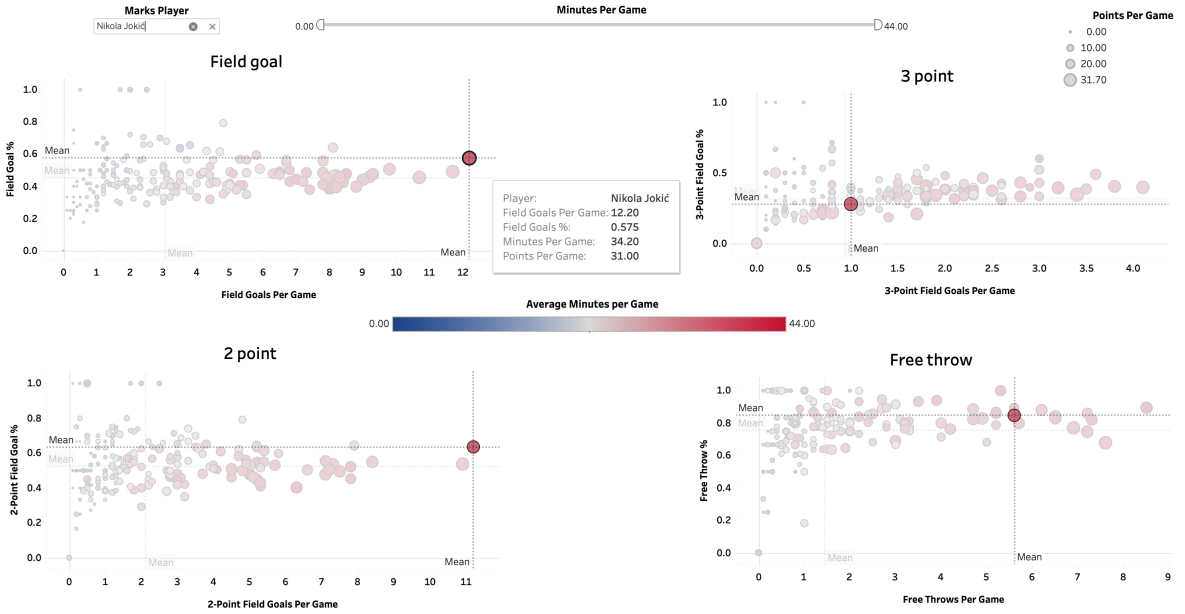


Figure 6: Players' Shots Statistics

It is possible to highlight a specific player in order to see which are his shot characterics:



We can notice how Nikola Jokic (MVP of the season) stands out from all the others for 2-point shooting. He is an above average free throw shooter, especially related to attempted free throws, and he is also an average 3-point shooter (that is still impressive given the fact that he is a center). These characteristics allow him to lead on everybody

in field goals shooting statistics, showing once again why he deserved the MVP award.

3.3.4 Players Efficiency

This graph shows which players have been the most efficient during the regular season and the playoffs, in terms of "points" on "minutes played". The redder the circle the more efficient the player is. Also, note that the number of circles are different between the two graphs, and that's because not every player that played in the Regular Season necessarily played during the Playoffs (due to an injury or due to the fact that his team did not qualify).

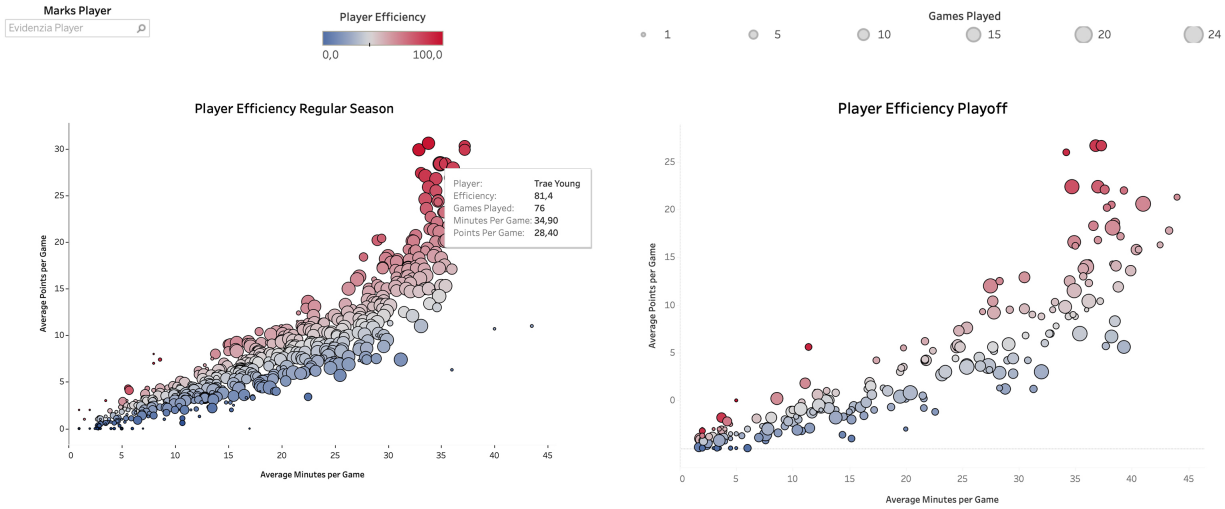


Figure 7: Players Efficiency - Regular Season & Playoffs

For example Trae Young (ATL) was one of the players with the highest efficiency during the Regular Season, with a value of 81.4 .

3.4 Ranking Regular Season

With the following chart we show dynamically the performance of the teams in the league, divided by conference, and so how the ranking changed over time. It is possible to see how the chart was on a specific day of the year or to see the bars moving over time.

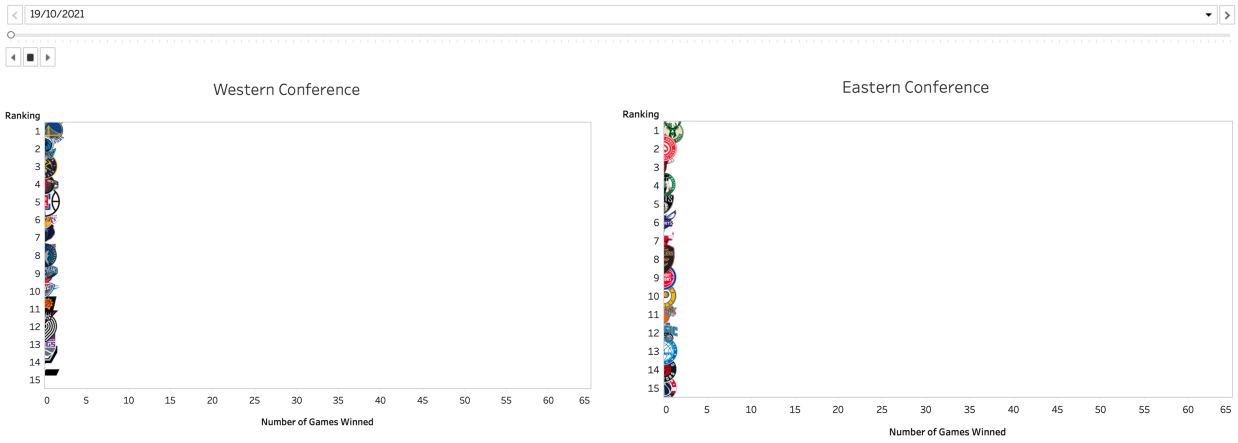


Figure 8: Ranking - 0 matches

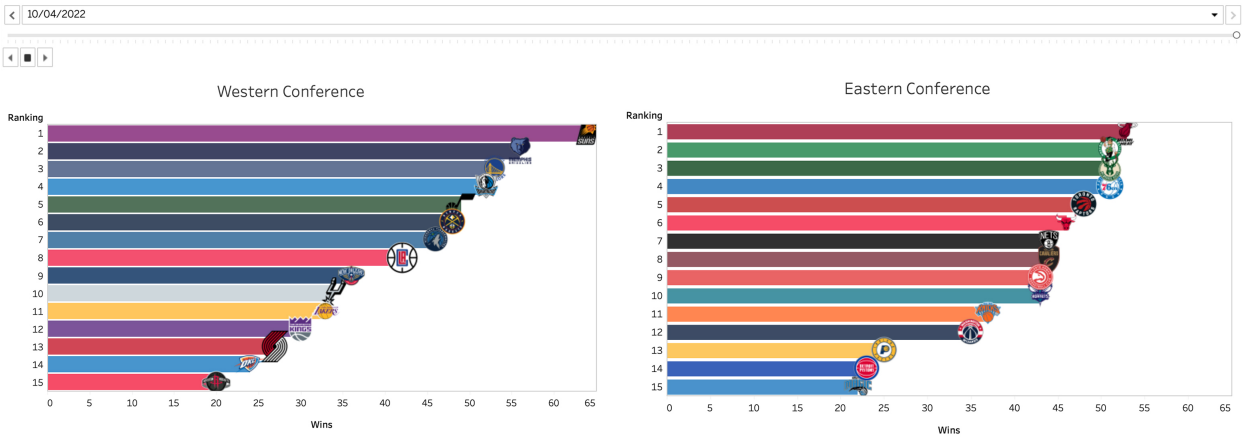


Figure 9: Ranking - 82 matches

The first team of the Western Conference during the regular season was Phoenix Suns, while in the Eastern Conference it was Miami Heat.

3.5 Teams Donut Charts

Donut charts show, for each team, how each player has contributed to a given statistic, divided by regular season and playoffs. It is possible to select both the team and the variable.

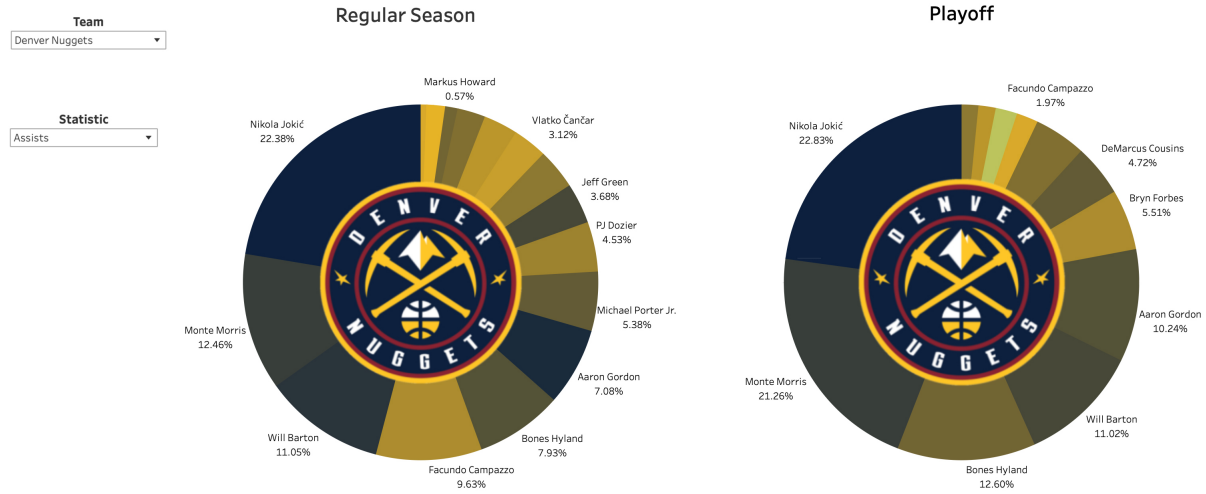


Figure 10: Denver Nuggets donut charts

For example, in the case of Denver Nuggets, during the regular season we can notice that Nikola Jokić contributed to the 22.38% of all the Assists of his team. Also, we can see that, during the playoffs, Monte Morris contributed to the 21.26% of all the Assists of Denver (more than what he did during the regular season).

3.6 Statistics Map

It is also possible to see how the various statistics are distributed geographically speaking. Here we have two different maps: the first one shows the maximum value of a given statistic by country, while the second one shows the average value of that particular statistic. It is possible to select which statistic to analyze and click on the countries to actually see the value of it.

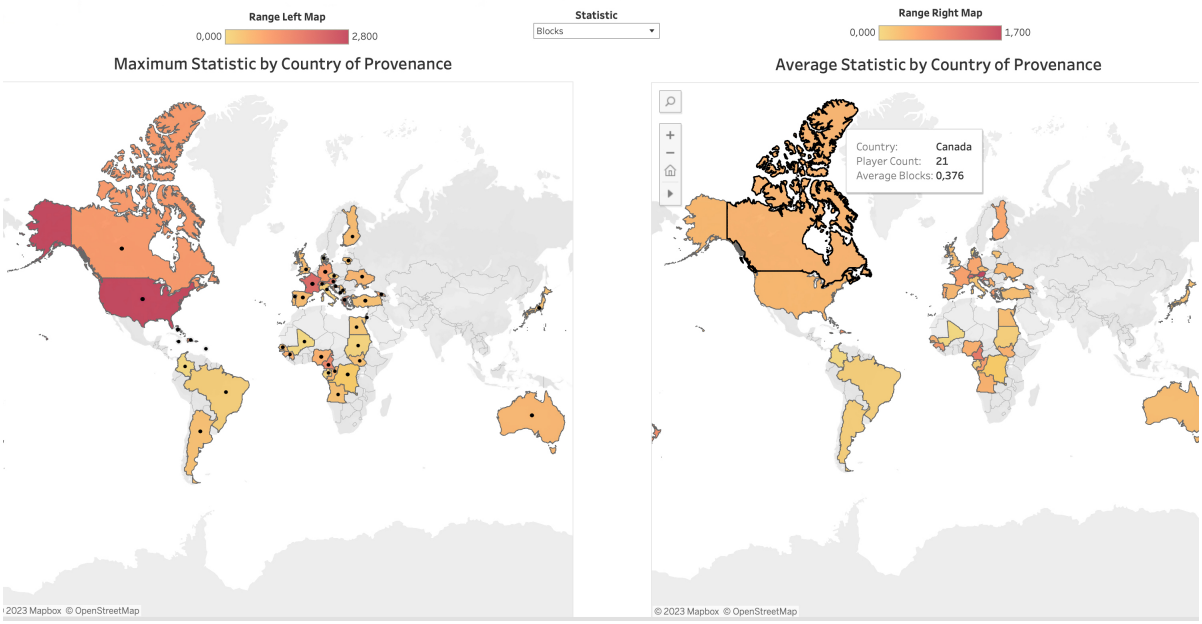


Figure 11: Map of Blocks Per Game

In this case we can notice that the country with the maximum value of Blocks per game is USA and that the Canadian players have an average number of Blocks per game of 0.376. In the Maximum Statistics Map you can also know which is the player with that maximum statistic by hover your mouse over the black dot.

3.7 End of the season

The last slide of the story recaps how the season, and so the playoffs, ended.
Playoffs began on April 16 and ended on June 16 with conclusion of the 2022 NBA Finals.
The winners of the Western Conference final, Golden State Warriors, played against the Eastern Conference winners, Boston Celtics. Golden State won 4-2 and so won the NBA 2021-22 title.
The MVP of the season was Nikola Jokic (DEN), while the MVP of the finals was Stephen Curry (GS).

Figure 12: End of the season

4 Evaluation

The quality of an infographic can be assessed on different levels:

- **Qualitative - Quantitative** → usually performed by using a heuristic evaluation, a psychometric questionnaire and a user test.
- **Absolute - Comparative** → an Absolute evaluation assesses wether a dataviz is "good" or not. A Comparative evaluation compares two or more dataviz to detect if there is any significant difference and which is the best one. Comparisons can be either *longitudinal* (between versions of the same dataviz) or *cross-sectional* (between different dataviz).
- **Formative - Summative** → the Formative evaluation is done during the development phase to early detect flaws and address them until the final version is delivered. The Summative one is done at the end of the development phase to see if the dataviz meets all the requirements that were collected at the beginning and the expectations.

4.1 Heuristic Evaluation

In a Heuristic Evaluation, designers observe their users (∼4-5) interact with their dataviz, asking them to *think aloud*, and detect usability problems (i.e. opportunities for improvement) interpreting the user's behaviour and comments in light of some good design "heuristics".
The output of a heuristic evaluation is a list of usability problems (divergences between your data viz display&behaviour and one or more heuristics) that is a list of opportunities to improve the data viz (make it more compliant with the heuristics). That's what make the heuristic evaluation so great (as a formative method).

4.1.1 Problems Identified

We asked 3 persons to interact with our infographic and to say aloud the problems they encountered in using it, in order to understand if there were problems during the usage. These are the issues we have identified related to our data visualization:

- *Teams Statistics:* "maybe there are too many available choices for the axis, I don't know which to choose".
- *Radar Charts:* "I don't understand what the abbreviations mean".
- *Players Statistics per Position:* "team logos bother the visualization and I don't understand why some players have the NBA logo. Also, I don't understand the abbreviations of the positions and I would like to have the possibility to search a specific player".

- *Ranking Regular Season:* users don't understand how to start the dynamic visualization.

- *Statistics Map:* explain better what the visualization shows.

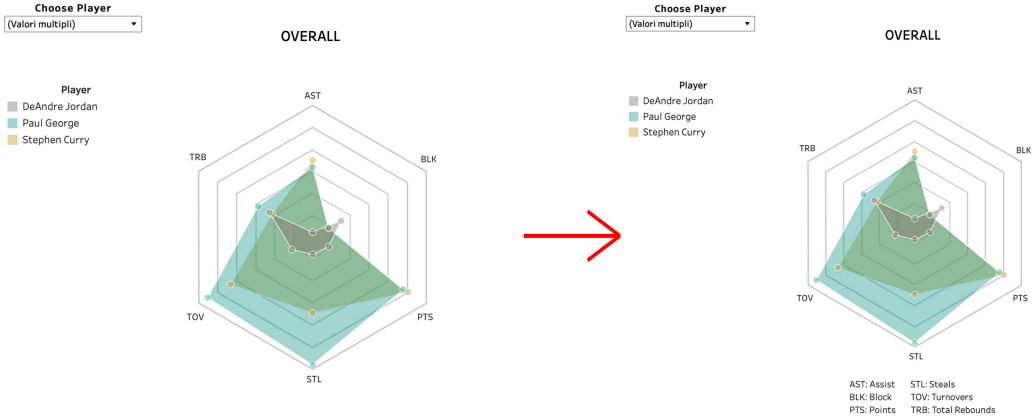
4.1.2 Changes made to the infographic

- After the evaluation we started solving all the problems we could. First of all we added some X-Y pairs hints, in order to give to all the users, even the less experts, the opportunity to create a graph which can output significant results.

Hints: some common pairs are

x = Free Throws	y = Free Throw Attempts
x = 3 Point Field Goals	y = 3 Point Field Goal Attempts
x = 2 Point Field Goals	y = 2 Point Field Goal Attempts
x = Field Goal Percentage	y = Field Goal Attempts
x = Offensive Rating	y = Defensive Rating

- Then we added a legend to each **radar chart**, so the users can easily understand what the abbreviations mean.



- The third visualization we improved is the Players Statistics per Position. First of all we substituted the teams logos with simple circles, so the graph is clearer. Then we wrote the full positions names and we added the possibility to highlight a specific player. Lastly we added inside the player information label his team, so that the colors couldn't confuse where they play.

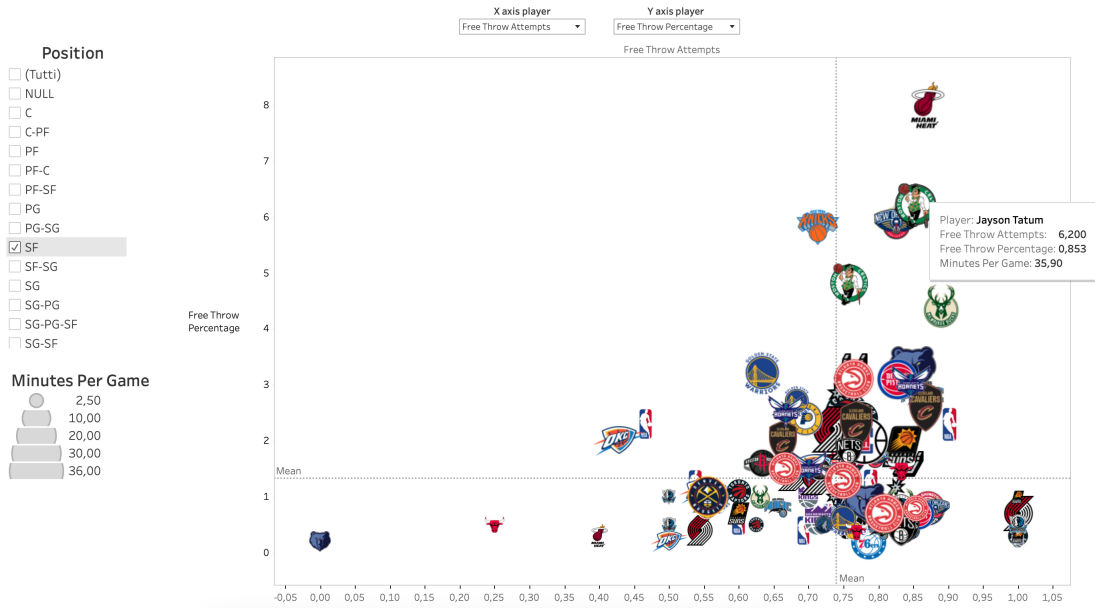


Figure 13: Players Statistics per Position before changes

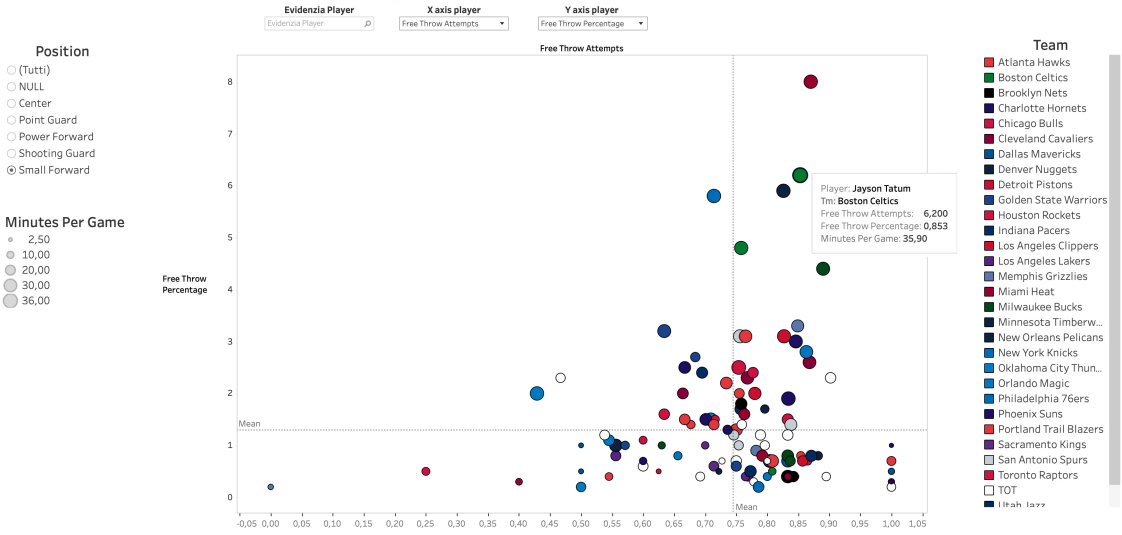
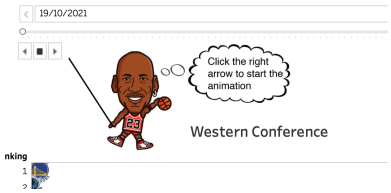


Figure 14: Players Statistics per Position after changes

- We added a tip to tell the users how they can start the dynamic visualization of the Ranking chart.



- Finally, we added a simple explanation on how the Statistics Map works.

4.2 User Test

In the user test, we observe the users (∼8) perform a task requiring to interact with our data visualization. We planned the task in advance on the basis of the objectives for which we have designed the data viz (tipically to inform, demonstrate or train the users) and focus on quantitative measures, like error rate, execution time, and the like. This assessment can be either comparative or absolute. In the former case you compare different versions of your design product. In the latter case you compare it with optimal error rate and optimal execution times or/and you compare the performances of population/sample strata. We chose the absolute one.

The tasks we conceived are:

- *Task 1:* in the *Statistics Map* chart, tell me, for the variable "*Total Rebounds*", which is the player with the highest value.
- *Task 2:* in the *Players' Shots Statistics* graph, tell me, between the players with a Minutes Per Game value over 30, which is the one with the highest percentage of two-points field goals, between those who have at least 8 two-point field goals per game.
- *Task 3:* in the *Players Efficiency* graph, tell me if Kyrie Irving had a better efficiency during the regular season or during the playoffs.
- *Task 4:* in the *Teams Statistics* graph, tell me which team had the less 3-point field goals attempts per game and which team had the most 3-point field goals per game.
- *Task 5:* in the *Teams Donut Charts* graph, tell me which Cleveland Cavalier player contributed the most for the variable "*Steals*" during the regular season.

4.2.1 Results: User test

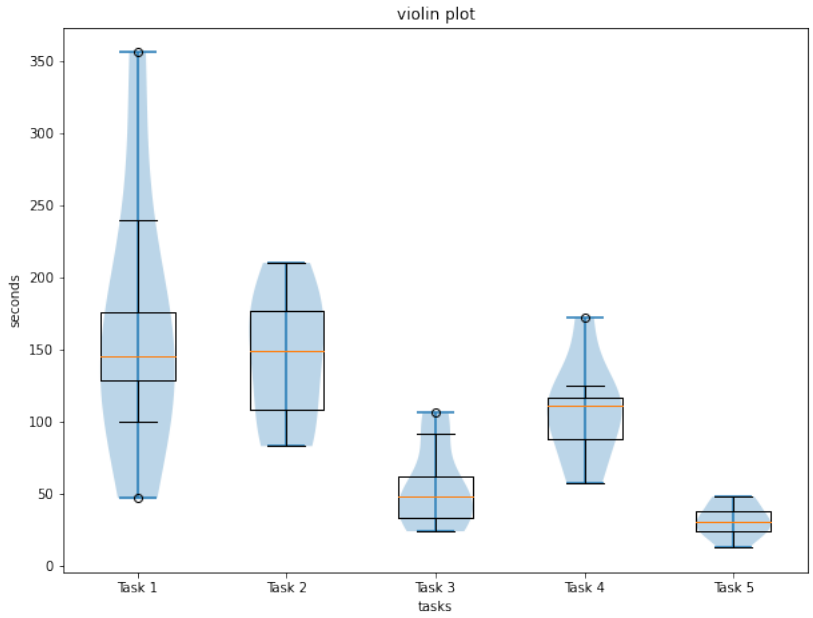
We administered the test, consisting of five tasks, to eight users. For each user, we recorded the seconds spent to complete each task. In the table below it's possible to see the results

	Task 1	Task 2	Task 3	Task 4	Task 5
U1	154"	116"	52"	125"	48"
U2	138"	174"	92"	114"	46"
U3	356"	83"	49"	71"	13"
U4	47"	160"	48"	172"	28"
U5	100"	138"	24"	57"	25"
U6	153"	186"	106"	114"	34"
U7	240"	85"	32"	108"	23"
U8	138"	210"	34"	94"	35"

Mean	166	144	55	107	32
Median	146	149	49	111	31

completed without help
completed with help
failed

The following graph shows the execution time of each task:



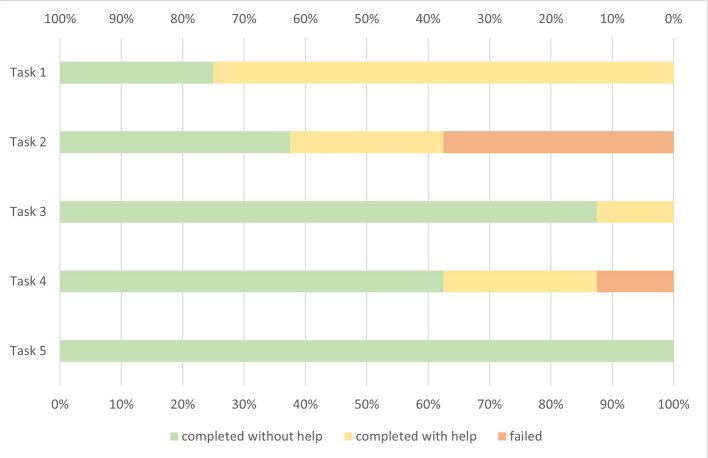
This graph is composed, for each task, of a violin plot and a box plot. Violin plots are used to visualize the distribution of numerical data. Unlike a box plot that can only show summary statistics, violin plots depict summary statistics and the density of each variable. For explanatory data viz, boxplot is easier to understand quickly. A box plot is another way of showing the relationship between a numeric variable and a categorical variable. Compared to the violin plot, the box plot leans more on summarization of the data, primarily just reporting a set of descriptive statistics for the numeric values on each categorical level.[3]

As we can see, Task 1 takes longer than the others. In fact it is the task where each one spent more time, probably because it is the first one, so users had to get the hang. Since most of the user had some problem solving the first task we gave them some hints before they started(user 6 and 8), and we noticed a little slump of solving time, so we decided to add the hints to the graphic.

Thanks to the boxplot we notice that there are outliers in our data, represented by a circle at the top of the first, third and fourth task. The other tasks have similar interval ranges. We can say that the question of the fifth task was the easier one.

[3]Violin and Box Plot, *Julien Beaulieu*

We can also show the error rates for each task:



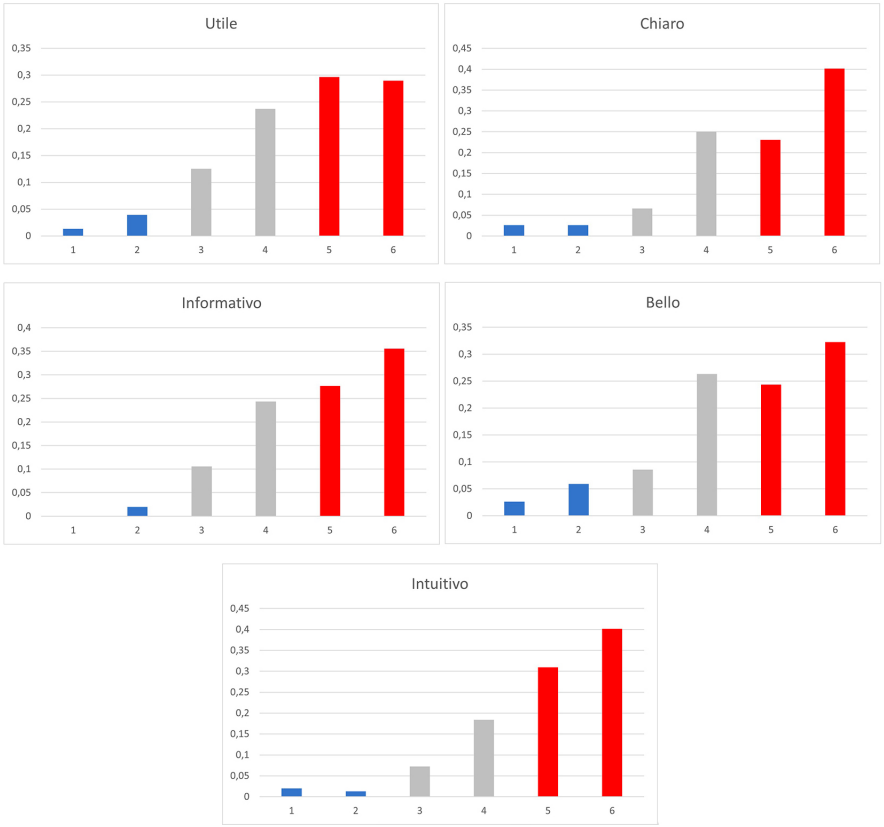
As we can see, the first task often required help to be completed, but every user managed to bring it to the end. Task 2 was the one with more fails, probably because it is the one that required more understanding of the scatterplots. Task 5, as we said, was the easier one.

4.3 Psychometric survey

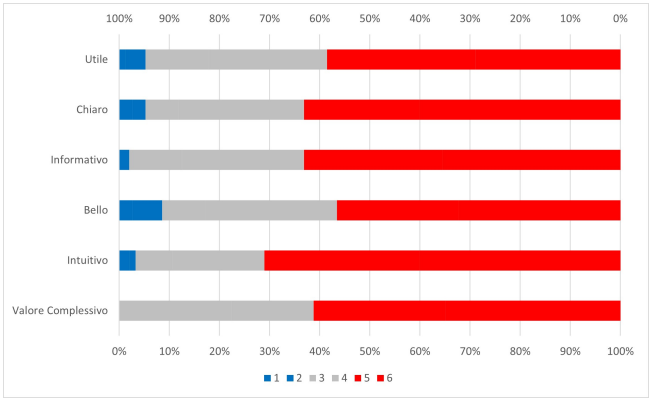
We administered a questionnaire to a sample of users (∼19) to evaluate some quality dimensions of the interaction of the users (even just a 10-second glimpse) with the data viz useful to show differences between different versions or user groups (comparison).
The survey consists of two sections. The first section is useful to assess the quality of the infographic in terms of Utility, Clarity, Informativeness, Beauty and Intuitiveness, by attributing a value from 1(very little) to 6(very much) to each attribute.
In the second section an overall evaluation is asked to the user, on a scale from 1(very low) to 6(very high).

4.3.1 Results: Psychometric survey

We administered the psychometric questionnaire, using the Cabitza-Locoro scale, to 19 users and these are the results:

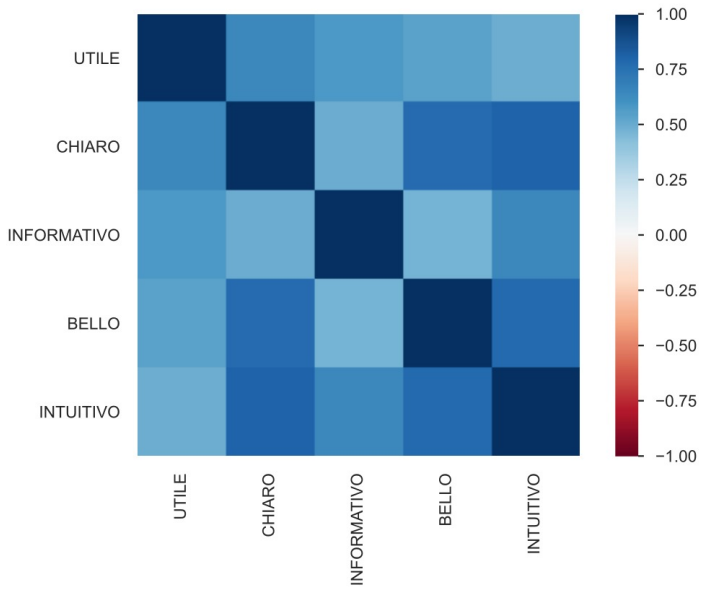


Then we built a stacked bar chart that allow a direct comparison between the various attributes.



The attributes "intuitivo" and "informativo" are the ones with the most satisfactory results. The less positive aspect is the beauty. Also the utility attribute had some negative votes. That is probably because some of the graphs we built are not so easy to understand, especially if who voted in the survey is not passionate about NBA or sports in general. However, the majority of votes are positioned in the central values, leaning towards the positive ones.

Lastly we built a correlogram in order to highlight correlations between different attributes:



	UTILE	CHIARO	INFORMATIVO	BELLO	INTUITIVO
UTILE	1.000	0.647	0.571	0.539	0.487
CHIARO	0.647	1.000	0.494	0.767	0.799
INFORMATIVO	0.571	0.494	1.000	0.468	0.645
BELLO	0.539	0.767	0.468	1.000	0.774
INTUITIVO	0.487	0.799	0.645	0.774	1.000

As we can see the highest correlations are between "chiaro" and "bello" , 0.767 , and between "chiaro" and "intuitivo", 0.799.

5 Conclusion

After all, our data visualization has resulted interesting and valuable by the users. Our purpose was to create an infographic that allow people to compare teams and players of the 2021-22 NBA season and to understand how the latter went.
Sports statistics field is full of different ideas, maybe even too many, and therefore it is not easy to decide what to choose to show to the public that has a value and that is also enjoyable visually. However, we believe that our work can be a good basis for the "story-telling" of a whole season through data.
The reason why we chose NBA data is because we are sports fans; moreover basketball, especially in the USA, is one of the pioneers of intensive statistical analysis applied to sports to improve teams and players' performances.
This project taught us not only how to use specific technical programs such as Tableau and the theory behind data visualization, but made us confront, debate, communicate and find together a way to solve our problems amd complete our tasks, as a team. Ultimately we are all satisfied with the infographic we produced and the skills we gained during the process.