FIFA 2019

CONTENT

INTRODUCTION

FIFA 19 is a football video game developed by Electronic Arts. As a football fan, I made an exploratory analysis on the FIFA 19 dataset using R and now converting them into visualisation using D3 to draw insights from the dataset. This project mainly focuses on creating recommendations of players through the process of visualisation.

Our intended audience can be categorised into 3 levels, namely:

1. **User**: Recommend top players to be included for playing fantasy leagues.
2. **Coaches / Team Management:** Recommend team selection/combination based on the formation.
3. **Team owners**: Recommend quality players within the age and budget limit.

As a part of our exploration, our first visualisation displays top Fifa 19 players from different countries, clubs, positional groups (i.e. Striker, Midfielders, Defenders, Goal-Keepers) and their stronger foot in an edge bundling chart. This provides a basic visualisation of currently top rated players in FIFA 19.

We then focus on finding whether the age and wages earned by the player have any correlation and what inferences can be carried out from the relation. This is done with the help of a line chart between age and the average wages earned by the player.

In our recommendation, it is very important to categorize players on the basis of their position. Thus, in our next visualisation, we try to calculate the top 5 skills sets required to play in a specific position. This is done with the help of a bar chart with different skills and the average percentage of that required skill.

Finally, the last visualisation resolves our problem statement of recommending top players for different positions in football. This narrative visualisation is carried out by

comparing top 2 players of that particular position on different skill sets using radar charts.

These unique and interactive visualisations will thus enable our intended audiences to recommend quality players along with the replacement options on the basis of different features. With these visual analysis, team management and owners can buy players within the age and budget limit.

DESIGN

The Five sheet design provides the brainstorming ideas in a greater detail. These initial ideas help in organising the thought process and methodology to be used for the final visualisations. These provide details about the alternative designs that could have been considered and justifies the final visualisations.

SHEET 1: Brainstorming

This sheet comprises all the ideas and thoughts about the problem statement, who all are impacted and how it can be resolved for the intended audience. It contains filtering and removing of duplicate or less important visual designs, categorize or organise the data and finally grouping these brainstorming ideas to collectively resolve the problem statement.

In the brainstorming sheet, the problem statement is analysed and different features of the dataset are explored. According to the problem, basic designs are thought that can help the audience to recommend quality players on the basis of different features.

I took the fifa dataset and filtered out important features like Name, Age, Position, Country, Club, Wages etc. which can be used for the recommendations. Different visual plots are considered on the basis of the dataset. These include bar charts, line charts, radar plots, choropleth maps, scatter plot, hierarchical edge bundling graph, sankey diagram, pie charts etc. In order to create interactive visualisation, different options for user interaction are noted like drop-down menu, slider bar, radio buttons or checkboxes.

SHEET 2: Initial Design - Part 1

The first step to the visualisation was to explore how the data is provided. As the dataset contained information about Countries and Clubs, this could be presented using the **Choropleth map** showing the aggregate distribution of players around the globe. Once created. it provided the useful insight that most players play from Europe, but it lacked user interaction. Secondly, a circular representation of exploring this dataset is considered using **Pie charts**. Although, it provided different sections of information but this was not elaborative enough for the users to explore this dataset.

Finally, **Hierarchical Edge-Bundling graph** is considered which contains interconnected nodes and on focus of each node, basic statistical information can be presented.

SHEET 3: Initial Design - Part 2

In the next visualisation design and check the relationship of Age and Wages of the players, initially **Scatter plot** was thought to show the distribution of players earning on their current age. This chart provided how the dataset is distributed but it lacked user interaction with the chart. Thus, in order to provide user interaction and meanwhile also check the relationship of Age and the wages earned, a **Line chart** is considered where a slider is used for filtering out the range of age.

In order to showcase the different skills set required for different positions, **Bar charts** are considered. For filtering data, instead of Radio buttons, as there were many options (like GoalKeeper-GK, Striker - ST, Striker - LW, Striker - RW, MidFielder - CAM, MidFielder - CDM, Defender - CB, Defender- LB, Defender - RB) that could have taken unnecessary space, Drop-Down menu is used containing all possible positional groups.

SHEET 4: Initial Design - Part 3

This part of the sheet provided the visualisation for my recommendations. **Radar Charts** are used for creating recommendations for different positions and also provide comparison between the recommended player and its replacement option.

Filtering of different positions can be carried out using Drop-Down Menu containing positions. For the Radar charts, different combinations of colors are used to compare different players and also showcase the percentage of skills a player has. These charts are unrealistic when more comparisons are made, thus we are restricting to only 2 comparisons i.e one Recommendation and other as a Replacement option.

SHEET 5: Realisation

The final sheet is a Realisation sheet which provides complete layout, focus and its operations for chosen visualisation. The final layout is similar to the Five design sheet, except for providing different filtering options of formations, I provided the drop-down for different positional groups and combined them to showcase my recommendation through Radar Charts.

The Initial view presents the Users to check the top rated players along with their basic statistics. User Interactions is provided to select filters for Age ranges and different positional groups and finally recommend top 2 players and see their skills statistics.

This project used different technologies like **D3, JavaScript, CSS, R, HTML, and Microsoft Excel**, created using Bracket.io. The total time to build this project including exploration and visualisation was 100 hours overall with 8 hours per day (13 days) and 16hours (2 days for testing) i.e 15 days.

This user interactive page provides useful insights and provides player recommendations.

IMPLEMENTATION

This Project is done with the use of different technologies like D3, JavaScript, CSS, HTML, R and Microsoft Excel. R language is used for data wrangling and cleaning purposes and for exploration of dataset. Microsoft Excel is used for viewing the created data frames and .csv files. HTML, CSS and Javascript are used to create a web page and D3 is Javascript library which mainly focuses on the visualisation part.

1. Edge - Bundling Chart:

Steps**:**

1. From the original dataset, top players on the basis of overall performance and their FIFA ratings, players are filtered out in R.

2. From this dataset, only relevant features like basic information in terms of Name, Overall Rating, Country, Club, Age, Position, value and wages(in euros) are recorded and stored in a .csv file.

3. A json file is created with the interconnected links in order to show relationships between other variables as each node is connected to the other node in a json array.

4. Features like Name, Countries, Club, Foot(preference), and Position are represented in a circular manner.

5. On hovering, the interconnected nodes are displayed along with the basic information of the category selected, this makes a far better user experience while exploring the dataset.

2. Line Chart:

For visualisation of the relationship of Age and the wages earned, a **Line chart** is used where a slider acts as a filtering option for ages.

Steps:

1. After exploring the dataset, players of the same Age are grouped together and the wages earned are taken as the Average Wages.

2. Grouping and filtering is done using R and .csv files are viewed in Microsoft Excel.

3. Javascript library is loaded for providing the Range Slider for Ages in D3.
4. Using D3, a line chart is created with the X axis as the Ages (in year) and Y axis as their average wages.

5. On updating the slider range, the line chart is also updated automatically which makes it more user interactive.

6. On hovering on the line charts displays the average wage earned at that Age.

3. Bar Chart:

For visualising different skills set required for different positions, Bar charts are used along with the Drop-Down menu, containing all possible positional groups.

Steps:

1. After exploring the dataset, a positional group is created namely GoalKeepers, Defenders, MidFielders and Strikers.

2. Players are filtered out on the basis of their group using R.

3. Creating .csv files for each position and providing a Drop-down menu with its options.

4. Using D3, a bar chart is created with X axis as different skill set and Y axis as their average value for that required skill set.

5. On changing the Drop-down menu, the bar chart for the position is updated automatically which makes it more user interactive.

6. On hovering on the bar charts displays the average required skill set.

4. Radar Chart:

For visualisation for recommendations, **Radar Charts** are used for different positions and also provide comparison between the recommended player and its replacement option.

Steps:

1. After exploring the dataset, a positional group is created namely GoalKeepers, Defenders, MidFielders and Strikers.

2. For each position, top 5 skills are recorded and further analysis is done.

3. Top 2 players are filtered out on the basis of their position using R.

4. This data along with their percentage of required skills are stored and providing recommendations.

5. Using a Javascript library, radar plots are created and different players are compared.

6. On changing the Drop-down menu, the radar chart for the position is also updated automatically which makes it more user interactive

7. On hovering of each point on the radar plot, highlights the area of that curve and also tells the percentage of particular skills set.

USER GUIDE

The files can be accessed using Bracket.io and the folder containing all files are used in the working directory. To run the project, run **index.html.**

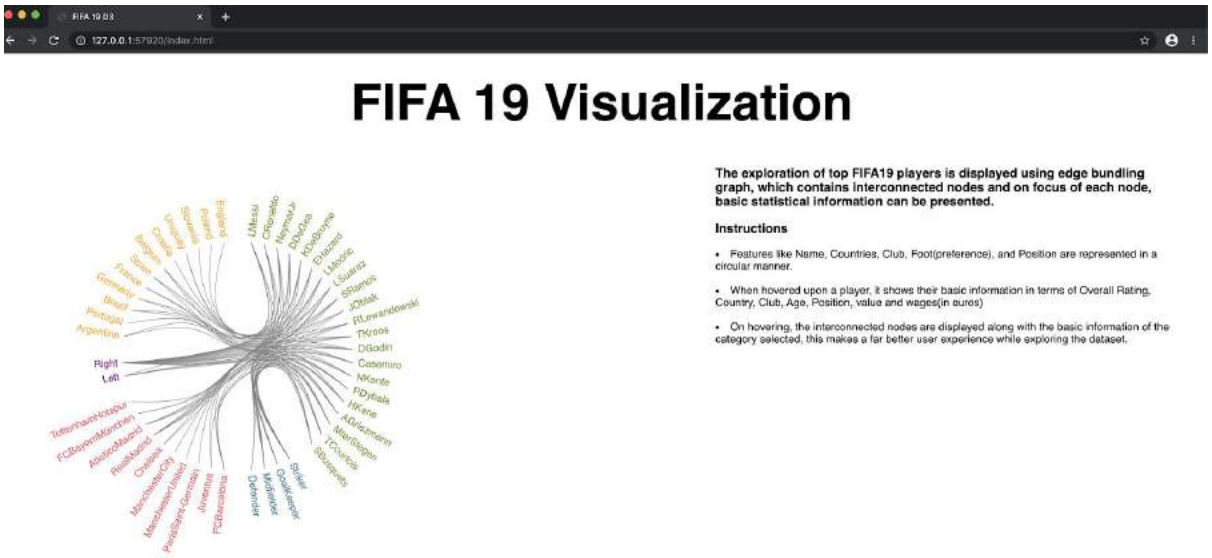


Figure: Web page

It provides a way to explore how the data is provided. **Hierarchical Edge-Bundling graph** is considered which contains interconnected nodes and basic statistical information is presented.

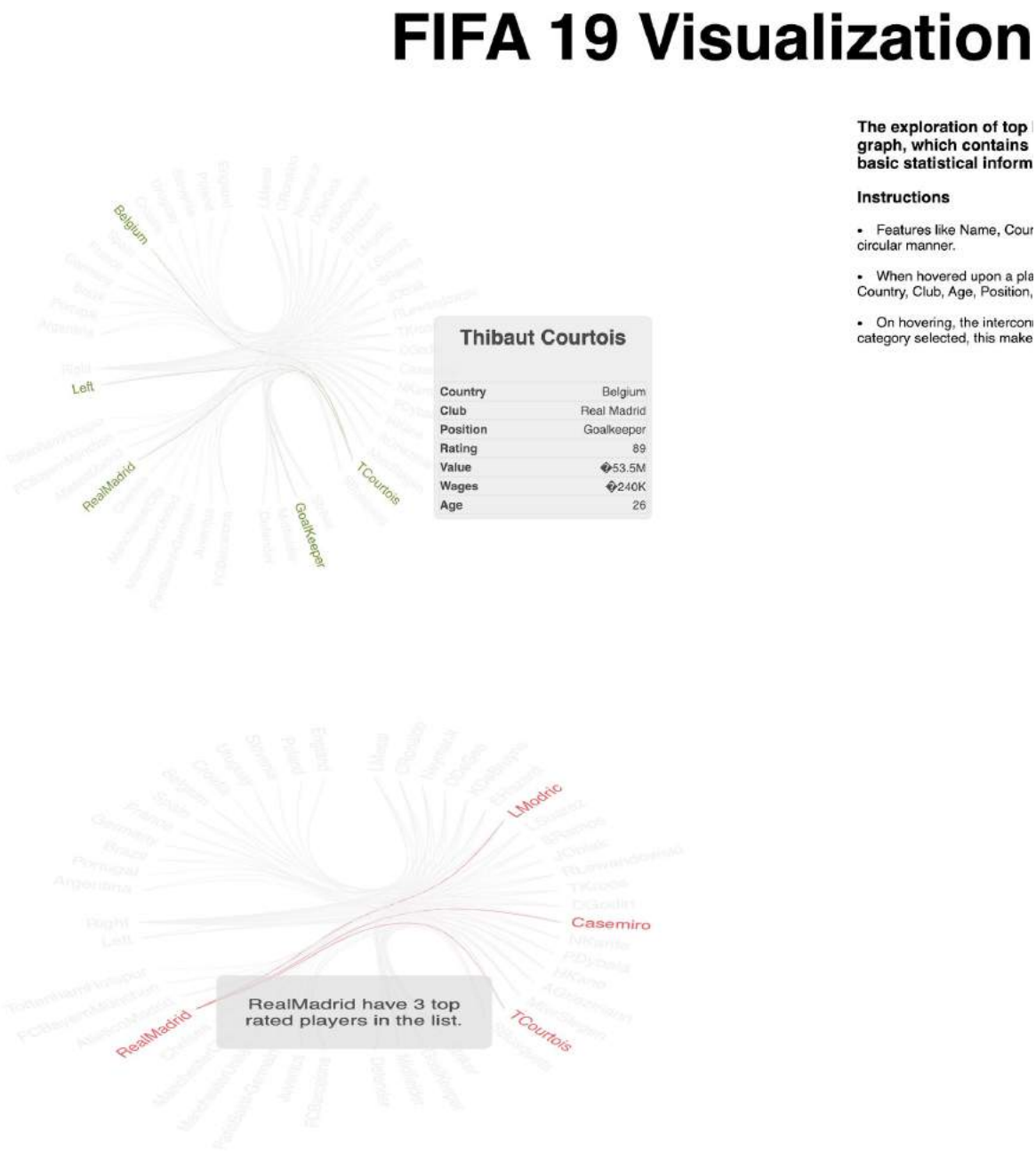


Figure: Hovering on a player provides basic information

Then the relationship between the Age and Wages earned can be visualized using line chart.

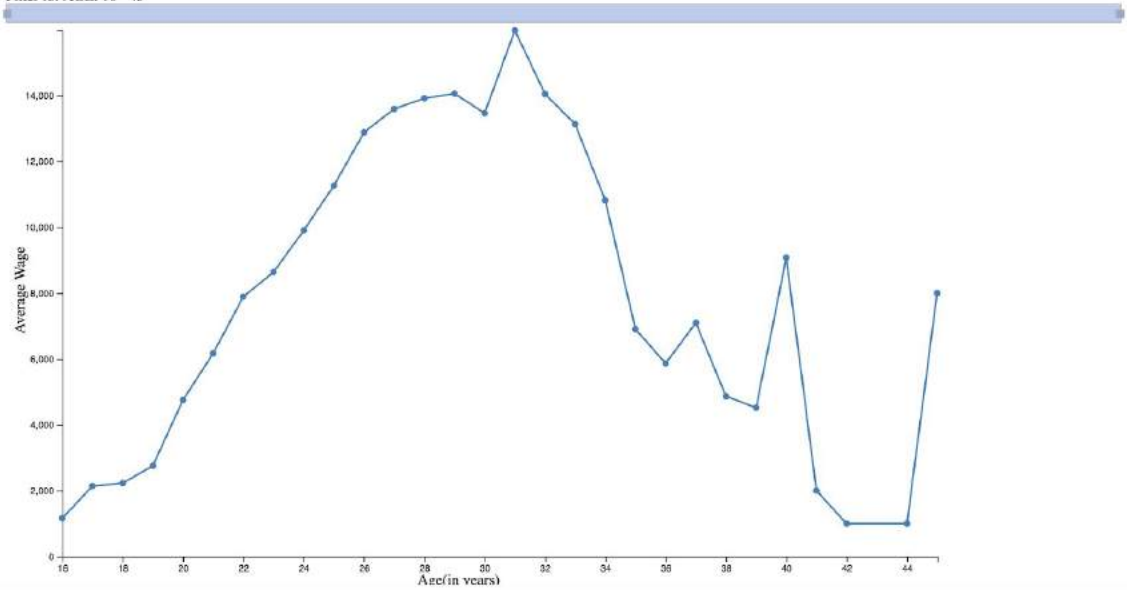
How are the wages earned and ranking of players from various positional groups (like Strikers, Mid-Field, Defence, Goal-keepers) distributed and which group has the highest value? Does the Age of the player have an impact on overall performance and their wages?

.

Instructions

- A range slider is provided, to filter out the AGE.(drag the lower or upper slider limit)
- The line chart is updated according to the filter of age selection.
- On hover of the line chart, the average wages earned by the player of that age is displayed.

Filter for AGE: 16 - 45



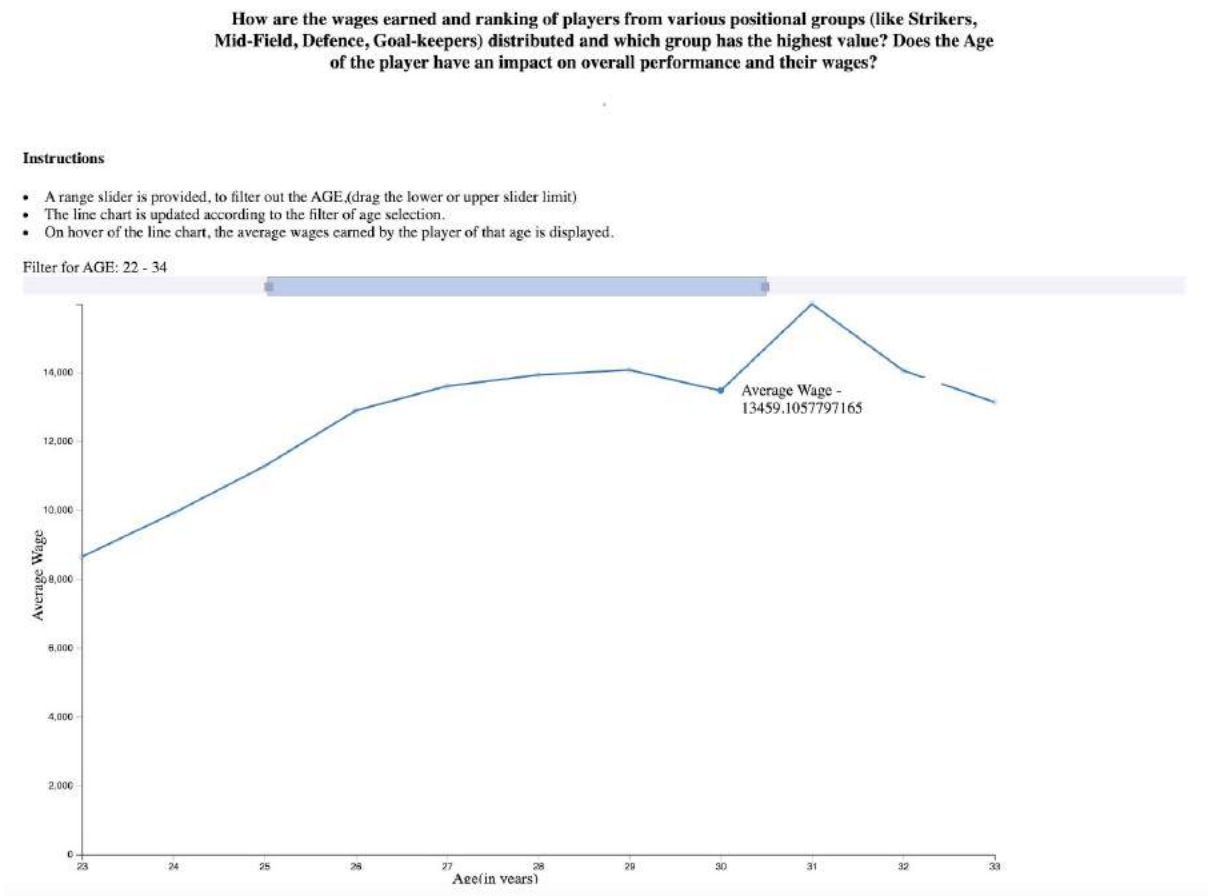


Figure displays user applied filter for AGE: 22-34, with hovering displays average wage in euros.

We then look at different positional groups and check the top 5 skills. Hovering on the bar chart displays the average required skill for that position.

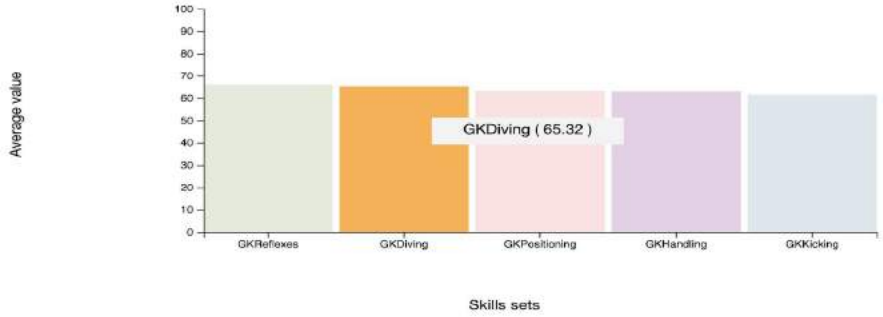
Which different skill sets are predominantly required for different positions in football?

Instructions

- According to the formations, a position is selected.
- Based on the selection, top 5 stills are displayed with average required skill set.
- Hover on the bar plot displays the average skill set

Select the position from the drop down below :

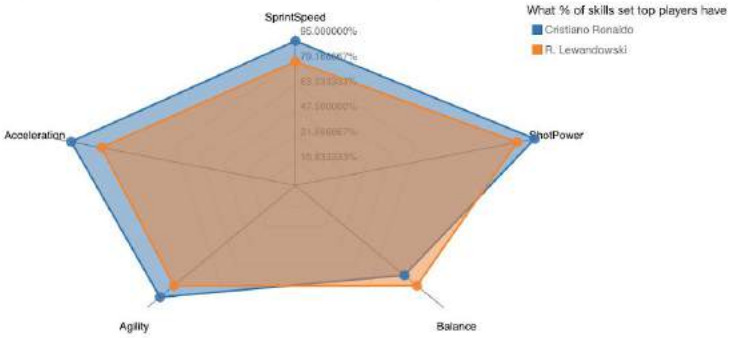
GoalKeeper - GK



After checking the top skills for a position, we can finally recommend players. This is done using RADAR charts.

What recommendations can we make to the coaches and the team management aiming to enhance their team performance by choosing correct team combinations for different formations based on different positions (as given by Q1) and also provide different options as replacement (as given by Q2)?

Upon selectionof the position, from the the above drop down menu, a recommended players are displayed along with the comparison using radar graph.



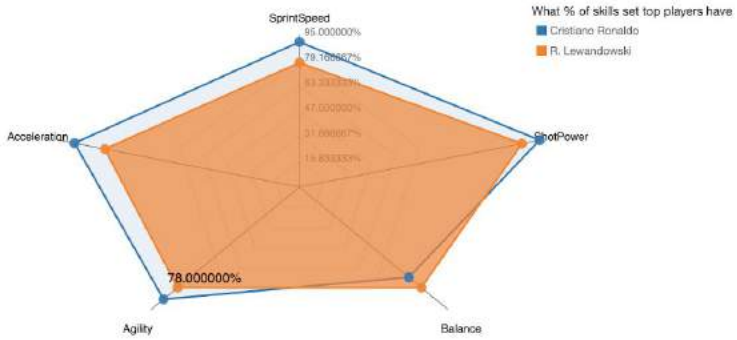
Conclusion:

For coaches or team management, recommendations can be provided to enhance their team performance. A correct team combination can be chosen for different formations based on different positions. Recommendations also provide like-for-like replacements without affecting the overall team performance.

On hovering on Radar plots, it highlights the area of that curve and also tells the percentage of particular skills set as shown below:

What recommendations can we make to the coaches and the team management aiming to enhance their team performance by choosing correct team combinations for different formations based on different positions (as given by Q1) and also provide different options as replacement (as given by Q2)?

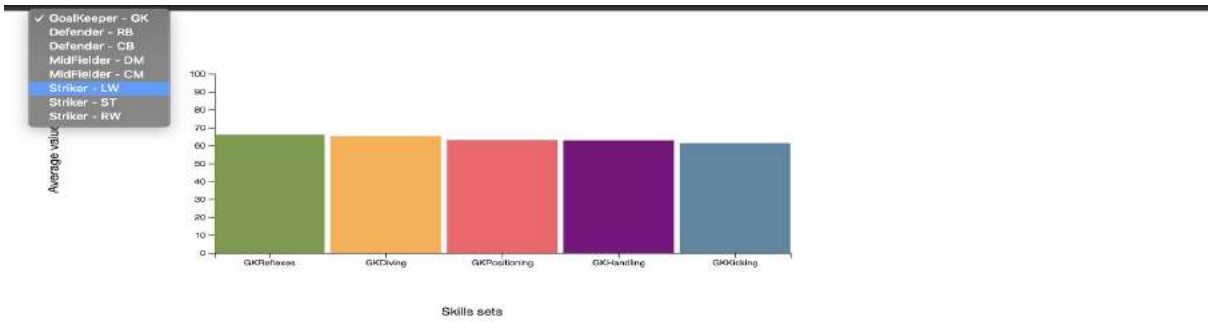
Upon selectionof the position, from the the above drop down menu, a recommended players are displayed along with the comparison using radar graph.



Conclusion:

For coaches or team management, recommendations can be provided to enhance their team performance. A correct team combination can be chosen for different formations based on different positions. Recommendations also provide like-for-like replacements without affecting the overall team performance.

A highly interactive and narrative visualisation consists of story telling of the data provided and updation of charts as per the user interaction. Thus, bar charts and Radar charts are automatically updated on choosing the positional filter as shown below:

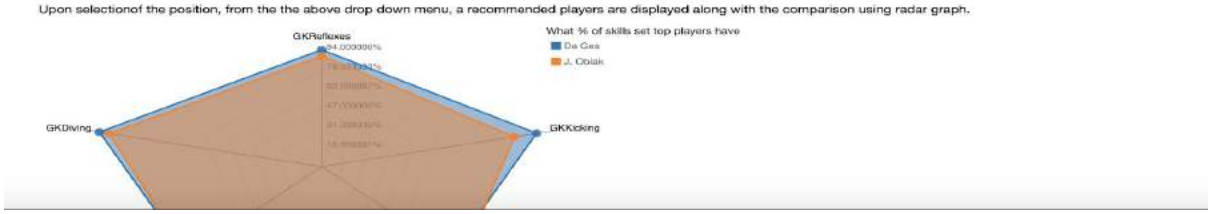


Conclusion:

In FIFA, players playing from different positions possess different skill sets. For instance, a Striker skills like SprintSpeed, Acceleration, Agility, Balance and ShotPower play a key role. In the Midfield position, players need to have strong Balance, Agility, Acceleration, SprintSpeed, and Stamina whereas for Defender, skills like Strength, Jumping, Stamina, Standing Tackle and Aggression impact their performance and ratings.

What recommendations can we make to the coaches and the team management aiming to enhance their team performance by choosing correct team combinations for different formations based on different positions (as given by Q1) and also provide different options as replacement (as given by Q2)?

Upon selectionof the position, from the the above drop down menu, a recommended players are displayed along with the comparison using radar graph.



BIBLIOGRAPHY

1. **SoFifa, powered by EA Sports, https://sofifa.com/**
2. **Kaggle dataset, retrieved from https://www.kaggle.com/karangadiya/fifa19#data.csv**
3. **Prabhu Kumar, Data Analysis of FIFA 2019, retrieved from https://medium.com/@prabhukumar1322/data-analysis-of-fifa-2019-26e052da5a64**
4. **R-Graph Gallery, retrieved from https://www.r-graph-gallery.com/**

CONCLUSION

While performing this project, I learned a lot of great visualization techniques and how storytelling can be achieved with design. I feel much more confident in exploring the dataset in R and making visualizations through D3. Following are the insights:

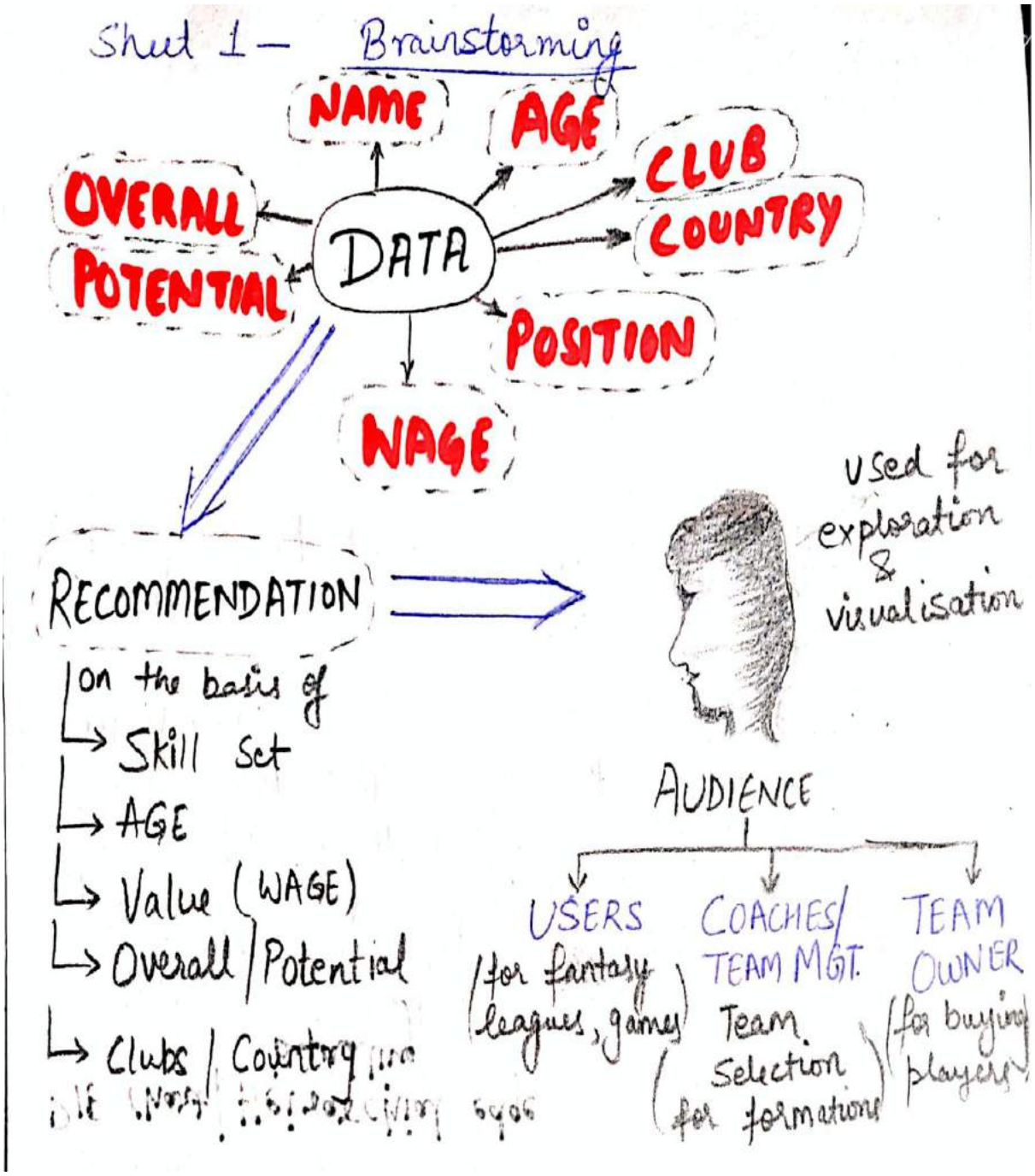
1. In Football, Age plays an effective role in the overall rating of a player and wages they earn. From the visualisation, it is observed that as Age increases the Wages earned by a player tends to increase. The overall rating and wages earned are **optimum at the age of 31** (which provides a combination of both youth and experience), and after that it tends to decrease as the player ages. The fluctionas are a result of a few experienced players which tend to earn even when their age approaches 40.

2. In FIFA, players playing from different positions possess different skill sets. For instance, a **Striker skills like SprintSpeed, Acceleration, Agility, Balance and ShotPower** play a key role. In the **Midfield** position, players need to have strong **Balance, Agility, Acceleration, SprintSpeed, and Stamina** whereas for **Defender**, skills like **Strength, Jumping, Stamina, Standing Tackle and Aggression** impact their performance and ratings.

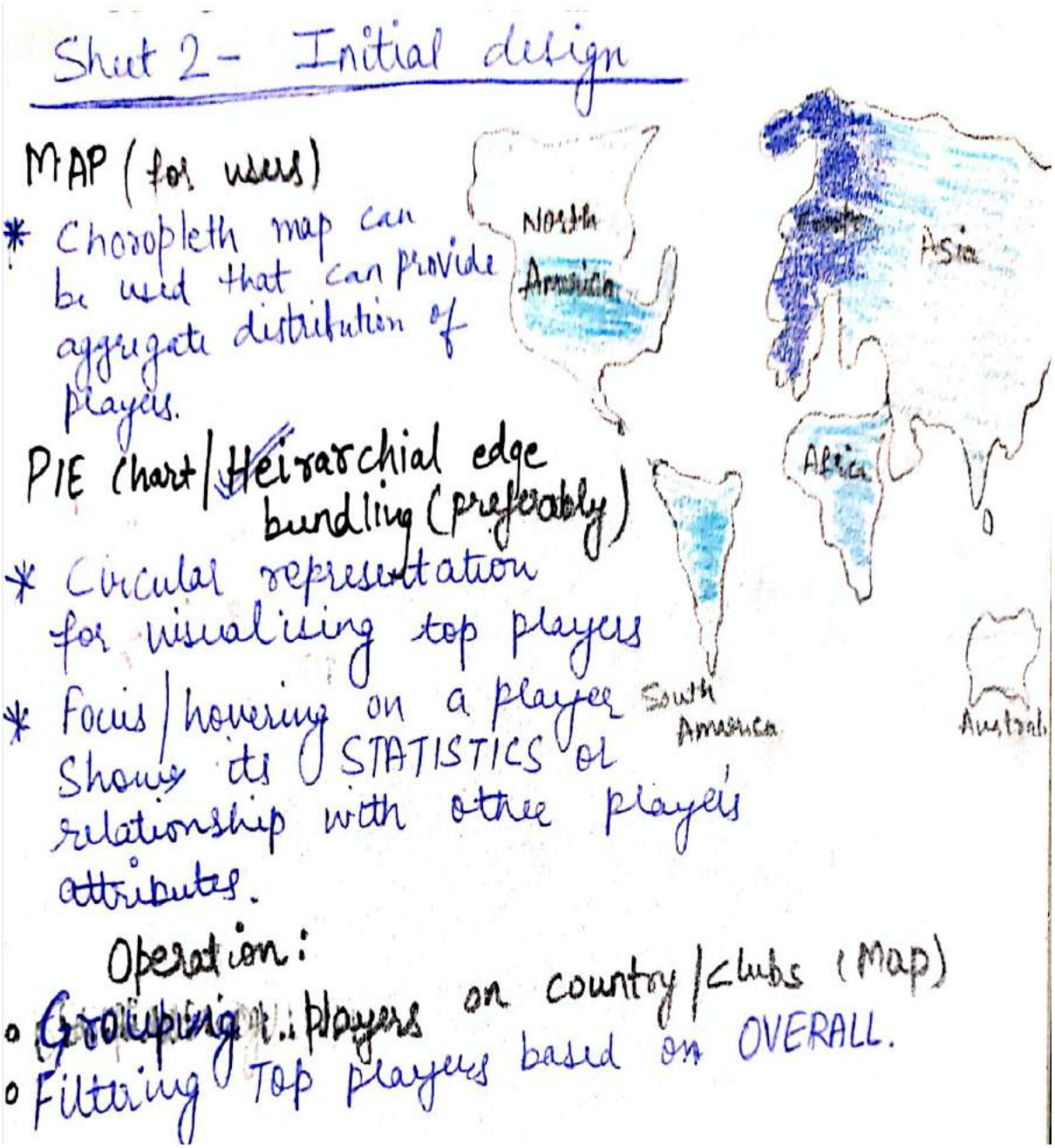
3. For coaches or team management, **recommendations** can be provided to enhance their team performance. A correct team combination can be chosen for different formations based on different positions. Recommendations also provide **like-for-like replacements** without affecting the overall team performance.

APPENDIX

SHEET 1: Brainstorming



SHEET 2: Initial Design

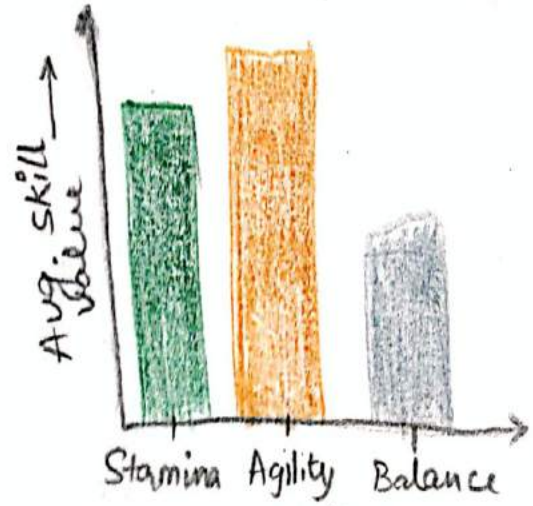


SHEET 3: Initial Design - Part 2

Sheet 3- Initial design

BAR plots :

* Players are categorised on the basis of POSITIONAL groups (Striker, Mid-fielder, Defence & Goal keepers).
* Different skill sets can be obtained for different groups.



RADIO BUTTONS :

* For selecting POSITIONAL groups
* for choosing Team Formation (4-3-3, 4-3-2-1 ...)

(PRO's)
Selection on the basis of single attribute

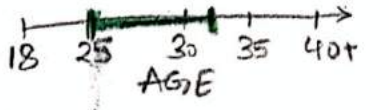
Operation :
- Grouping after creating different positional groups.
- Selecting the radio buttons changes the Bar plots.

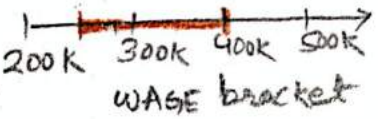
SHEET 4: Initial Design - Part 3

Sheet 4 – Initial design

* Drop-down (for COACHES / TEAM MGT.)
  - To select player from specific CLUB / COUNTRY.
  - On the basis of selection, different plots / visualisation gets updated
    HELPS to Study & Analise PLAYERS.

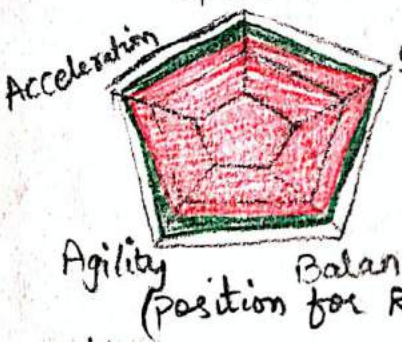
* SLIDER
  - For selecting the AGE range of players performance.
  - For creating WAGE bracket particularly useful for TEAM owner to buy players.





* RADAR plots
  - Useful for creating different positions.
  - Different FORMATIONS

RECOMMENDATIONS for are created. These are selected using drop-down and RADAR plots are dynamically updated.



- L. Messi
- L. Suárez

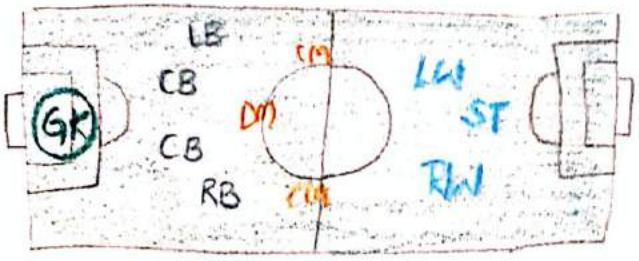
* Different colours denote different players

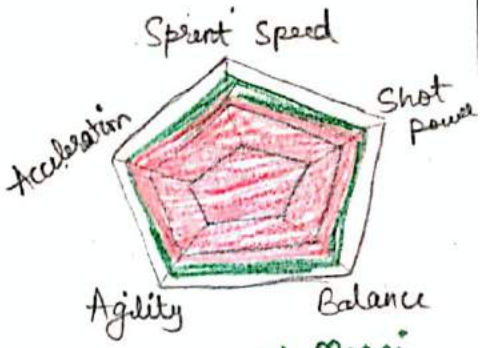
(CON) : Unreadable when more plots comparison is made (ie. >3)

SHEET 5: Realisation

Sheet 5 – Realisation

LAYOUT :–





Formation selected

(drop-down) 4-3-3

- Defenders
- Mid fielders
- Strikers
- Goal keeper

(Position for RW)
- L. Messi
- L. Suarez

OPERATION :–

- Initial view provide user to view top players & their statistics.
- The user can select filters :–
  - ↳ Country / club
  - ↳ Age
  - ↳ Wage
  - ↳ Skill set
  - ↳ Formation

  & Compare top 2 RECOMMENDED players and see their statistics.

* here user can be any kind of audience (user, coach, or owner)

DETAIL :–

- Technology → JavaScript, R, D3.
- Time to build → 100 hours for exploration & visualisation
- Estimation of COST :
  - ↳ 8 hrs per day ≃ 13 days
  - ↳ 16 hrs per day ≃ 2 days (for testing)

  ≃ 15 days.

ie. 3 weeks.