CS 6630 - Visualisation for Data Science

PROCESS BOOK
FIFA Viz Play

Project Name : FIFA VizPlay

Aditya Gattu - u1418754@umail.utah.edu [UID : u1418754]

Pranav Madhav Bagepalli - u1471494@umail.utah.edu [UID : u1471494]

Nitish Sai Kumar Kadiyam - u1466435@umail.utah.edu [UID : u1466435]

Project repository : https://github.com/dataviscourse2023/final-project-fifa-vizplay

Background and Motivation. Discuss your motivations and reasons for choosing this project, especially any background or research interests that may have influenced your decision.

As we all know, Football is one of the most famous sports which is often referred to as the "beautiful game", stands as a universal language, uniting individuals across different corners of the globe. As the sport enjoys an immense following, there exists a substantial group of enthusiasts who are keen to dive deep into the intricacies of the game, exploring beyond just the goals scored and matches won. This project proposal seeks to satisfy this curiosity, offering not just a statistical dive into the world of football but an interactive and engaging visualization of data, enhancing the user's understanding and appreciation of the game.

So our primary motivation behind this project is our genuine interest in the game of football, a sport that brings together people from diverse backgrounds and fosters a sense of community. We as a team thought that we can go beyond traditional metrics and venture into more detailed analyses, thus providing more depth to the narratives surrounding the sport. Inorder to do this we are planning to utilize the vast FIFA player dataset and our aim is to offer users the opportunity to explore and analyze a plethora of variables ranging from player performances, match outcomes, to intricate details like player's physic and skills metrics. The project is also inspired by our intent to show the sports analytics, offering innovative, insightful, and visually appealing interpretations of data that can cater to both avid football enthusiasts and researchers alike.

Project Objectives. Provide the primary questions you are trying to answer with your visualization. What would you like to learn and accomplish? List the benefits.

The primary objective of our project is to delve deep into the FIFA player dataset to carve out interactive and insightful visualizations that can facilitate a more enriched understanding of the game of football. Our main aim is to address critical questions such as identifying the top players based on various metrics like attack, defense, and shooting scores, and understanding the trends in goals scored across different countries over the years. This project intends to create a dynamic platform where users can explore data interactively, providing a concrete base to football discussions which are often driven by personal opinions.

Furthermore, this visualization venture stands as a beneficial tool for football enthusiasts, researchers, and analysts, offering a data-backed approach to analyzing and appreciating the multifaceted nature of football. By facilitating detailed comparisons and analyses through interactive graphs, it opens up new avenues in the field of sports analytics. Essentially, this project seeks to enhance the community's comprehension of the sport, paving the way for informed, data-driven discourse and fostering a new wave of engagement in the football community.

Data. From where and how are you collecting your data? If appropriate, provide a link to your data sources.

For our project, we are utilizing comprehensive datasets pertaining to the FIFA player database to formulate a rich and detailed analysis. The first dataset harbors an array of information focused on match specifics, encompassing various identifiers such as match_id, team_id, player_id, among others, providing us with an expansive view into the various tournaments and player dynamics within each match. The second dataset, on the other hand, offers a deep dive into individual player statistics and traits, including data columns that outline a player's skills, career trajectory, and market value, facilitating a nuanced analysis of player competencies and standings in the world of football. The combination of these datasets will be instrumental in developing an interactive visualization tool that fosters a deeper understanding and engagement with the sport, by allowing users to explore and analyze player performances and trends in a more immersive manner.

Dataset - https://www.kaggle.com/code/sivsankar/fifa22-recommender-system/input

Data Processing. Do you expect to do substantial data cleanup? What quantities do you plan to derive from your data? How will data processing be implemented?

Before we delve into the analysis, we anticipate undertaking a significant data cleanup process to ensure the integrity and reliability of our visualization project. This will involve identifying and handling missing or inconsistent data entries, as well as merging the two datasets harmoniously by linking relevant columns such as player IDs and team IDs. Moreover, we aim to derive meaningful insights from the data by calculating metrics such as

player performance scores based on various attributes, and aggregating goal statistics on a yearly and country-wise basis. The data processing will be implemented using robust data manipulation libraries and tools, allowing us to efficiently filter, transform, and prepare the data for a seamless visualization experience.

Visualization Design :

Potential quantities and insights that we aim to derive from the dataset:

Player Performance Metrics:
- Overall and potential player ratings.
- Player age, height, weight, and nationality distribution.
- Player preferred foot, weak foot, and skill moves.

Player Comparison:
- Comparisons of players based on various attributes (e.g., pace, shooting, passing).
- Clustering or grouping of players with similar skill sets.

National Team Analysis:
- Analysis of players representing their national teams.
- Distribution of players across different national teams.
- Performance metrics of national teams.

Trends and Patterns:
- Changes in player attributes over time.
- Correlations between player attributes and performance.

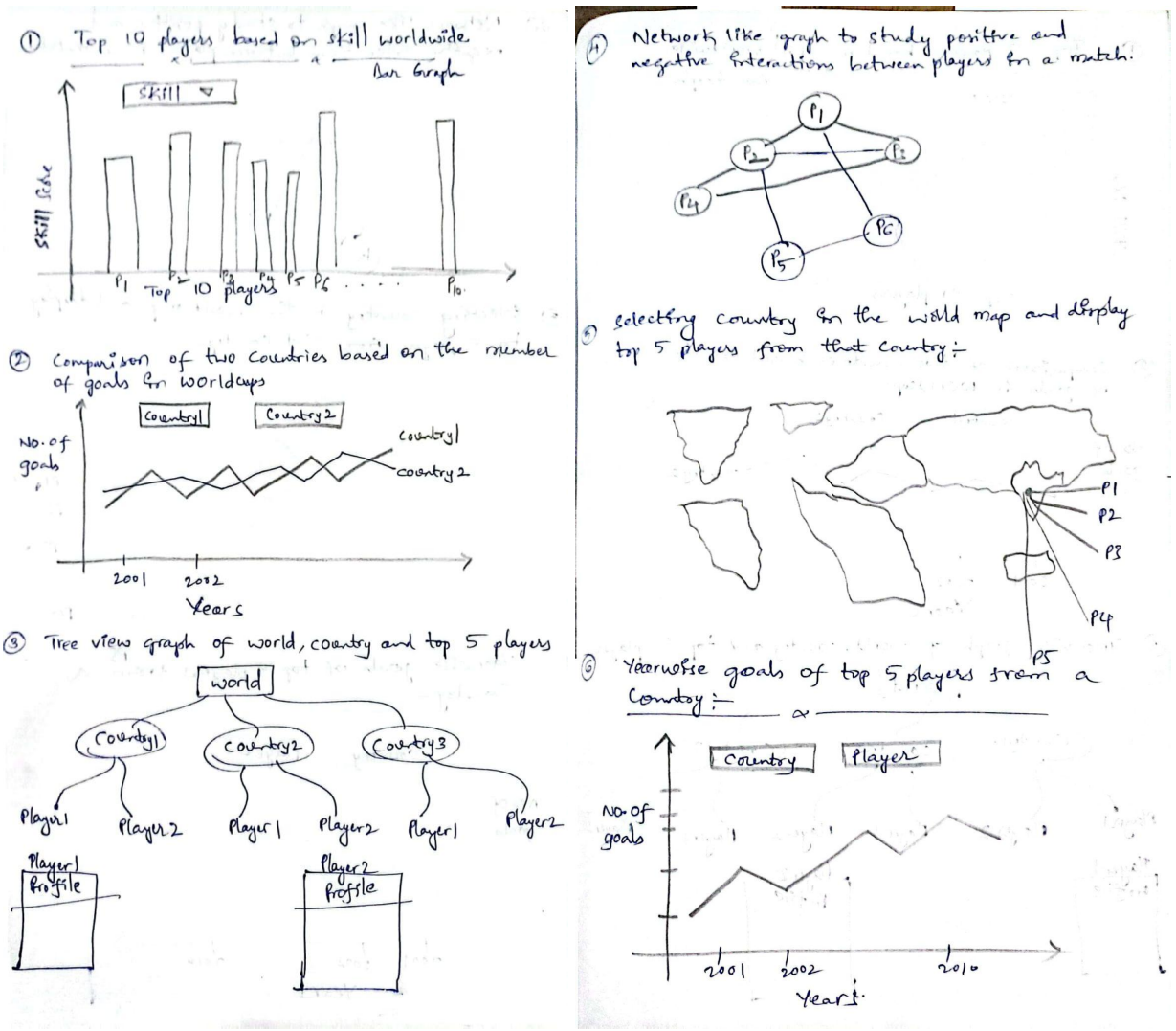
Visualizations:
- Creating charts, graphs, and visualizations to help convey insights more effectively.

Player and Club Profiles:
- Creating profiles for individual players or clubs that summarize their key statistics and attributes.

Club-Level Insights:
- Club-wise statistics, such as average player ratings.
- Club budgets based on player values and wages.
- Analysis of club positions in the league.
- Distribution of players across different club positions.

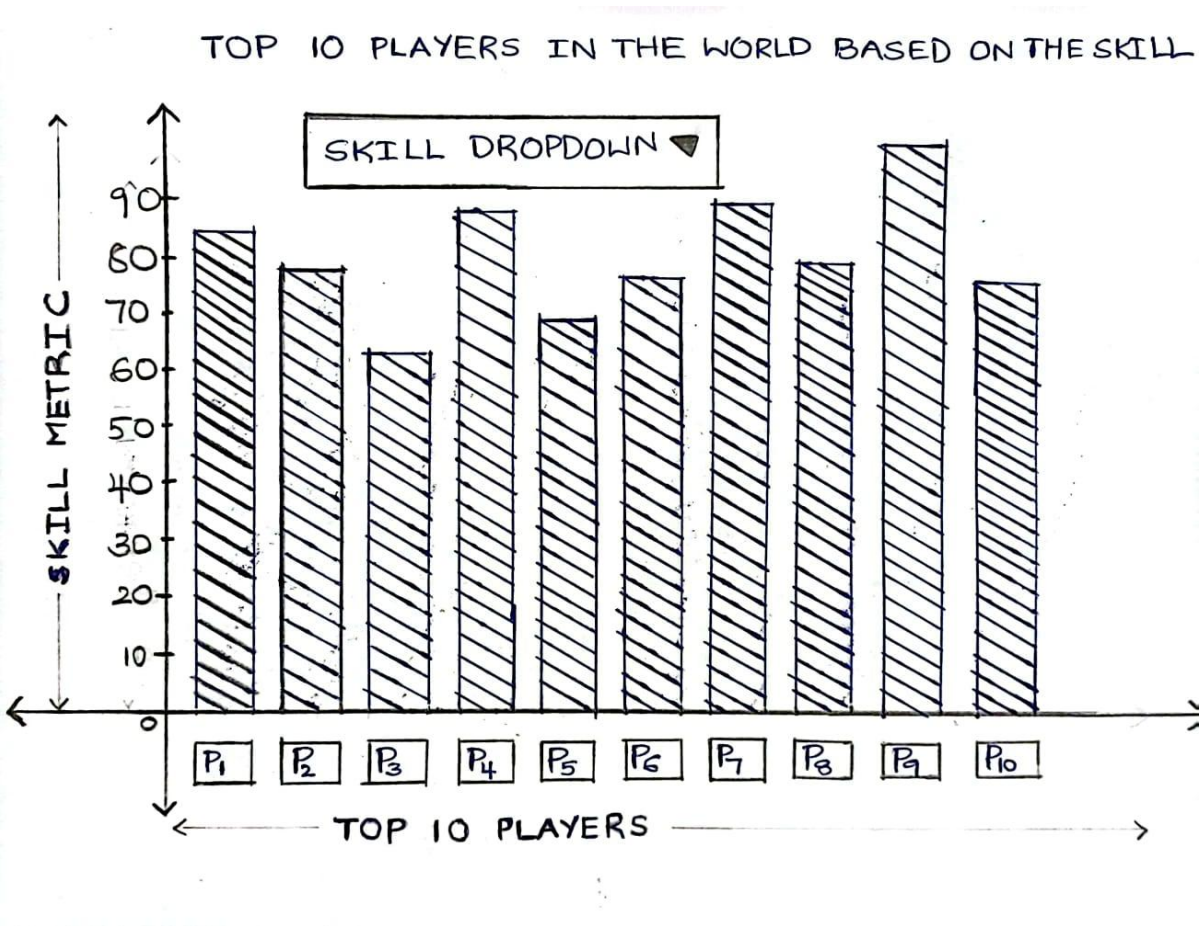
League-Level Insights:
- Analysis of leagues based on the level (e.g., top-tier, second-tier).
- Statistical summaries of leagues, including average player ratings, ages, etc.

We came up with the initial designs as follows :

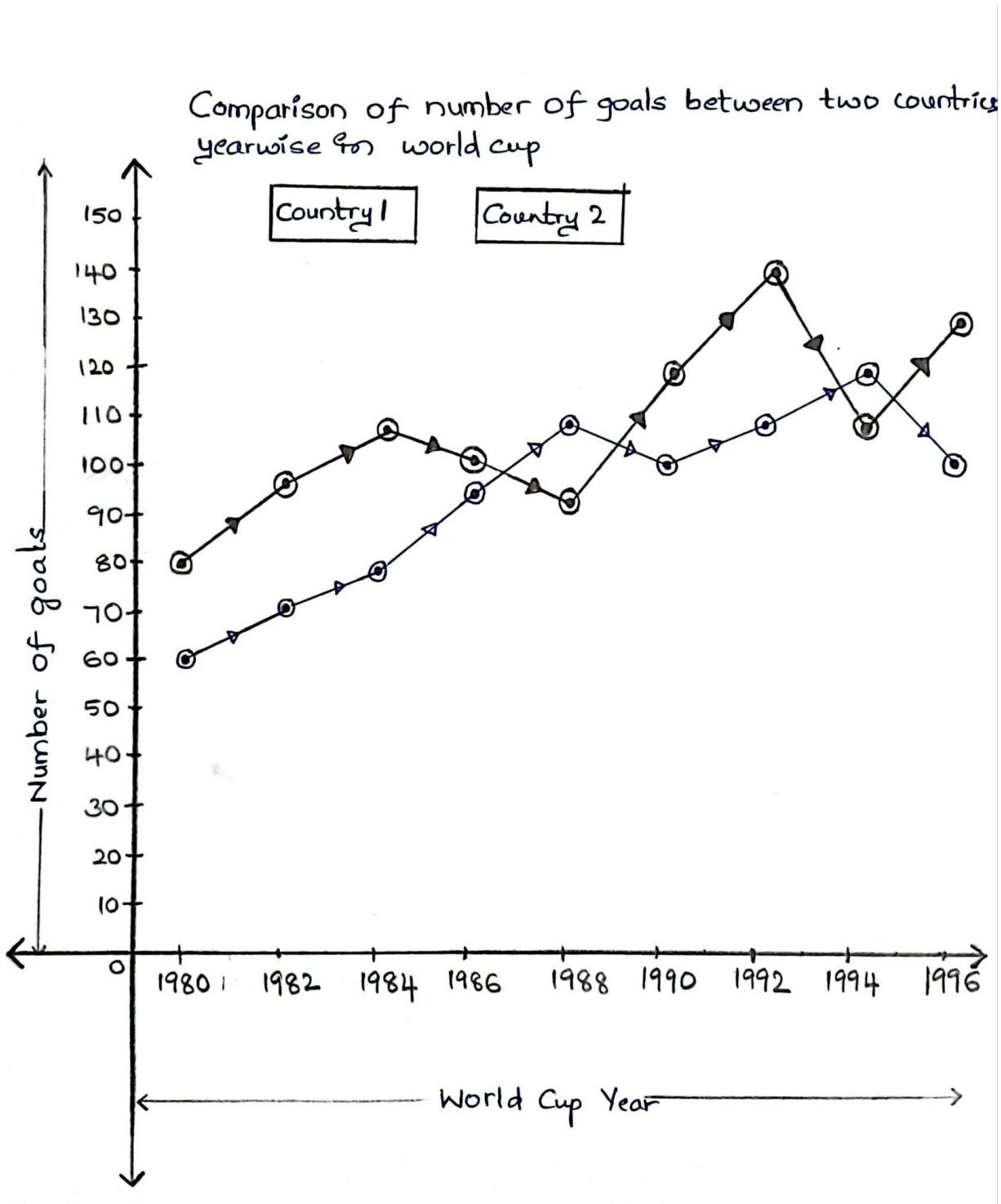


These are our Final Design sketches :

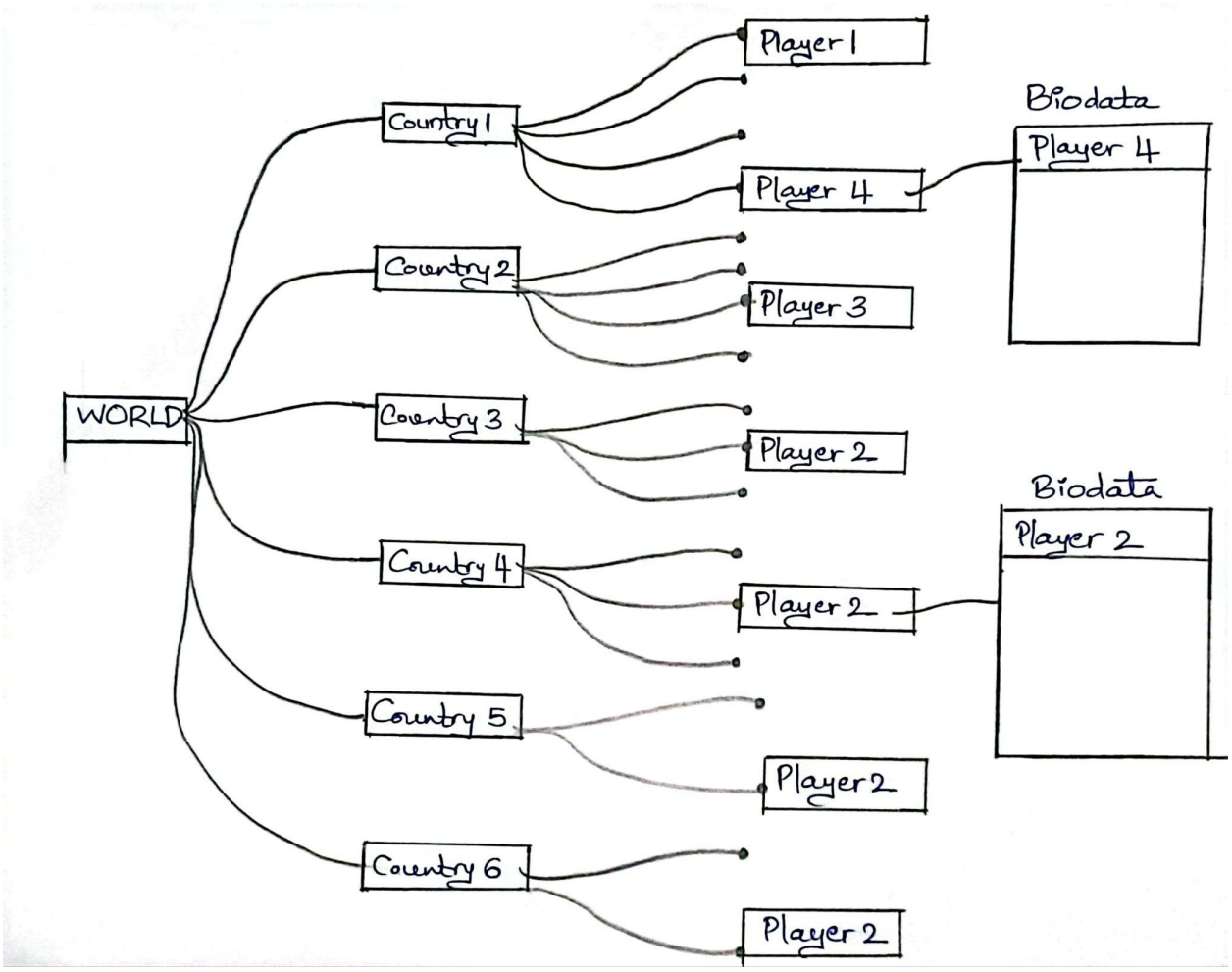
Sketch 1:


TOP 10 PLAYERS IN THE WORLD BASED ON THE SKILL
SKILL DROPDOWN
SKILL METRIC
90
80
70
60
50
40
30
20
10
0
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
TOP 10 PLAYERS

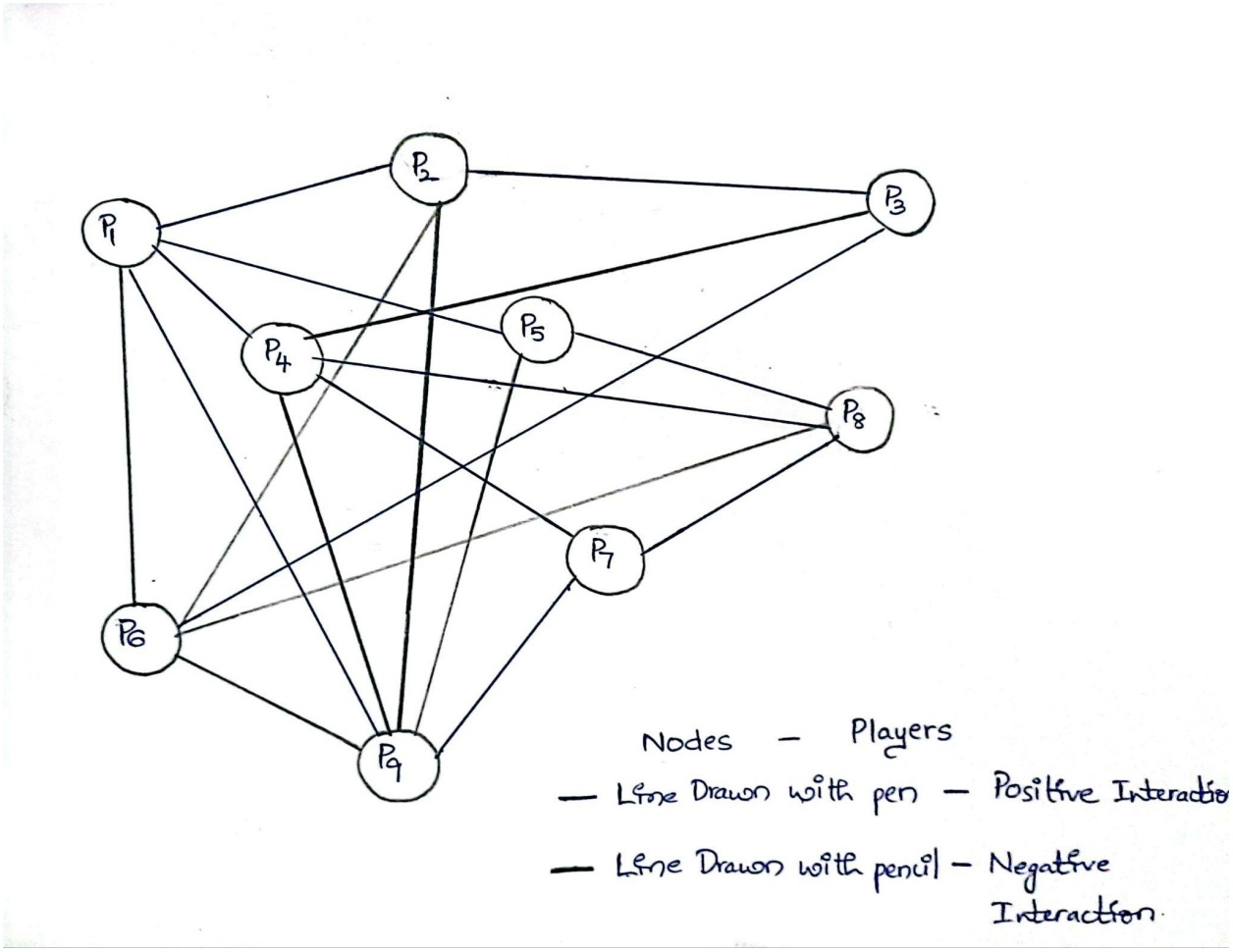
Sketch 2:


Comparison of number of goals between two countries yearwise for world cup
Country 1
Country 2
Number of goals
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
1980
1982
1984
1986
1988
1990
1992
1994
1996
World Cup Year

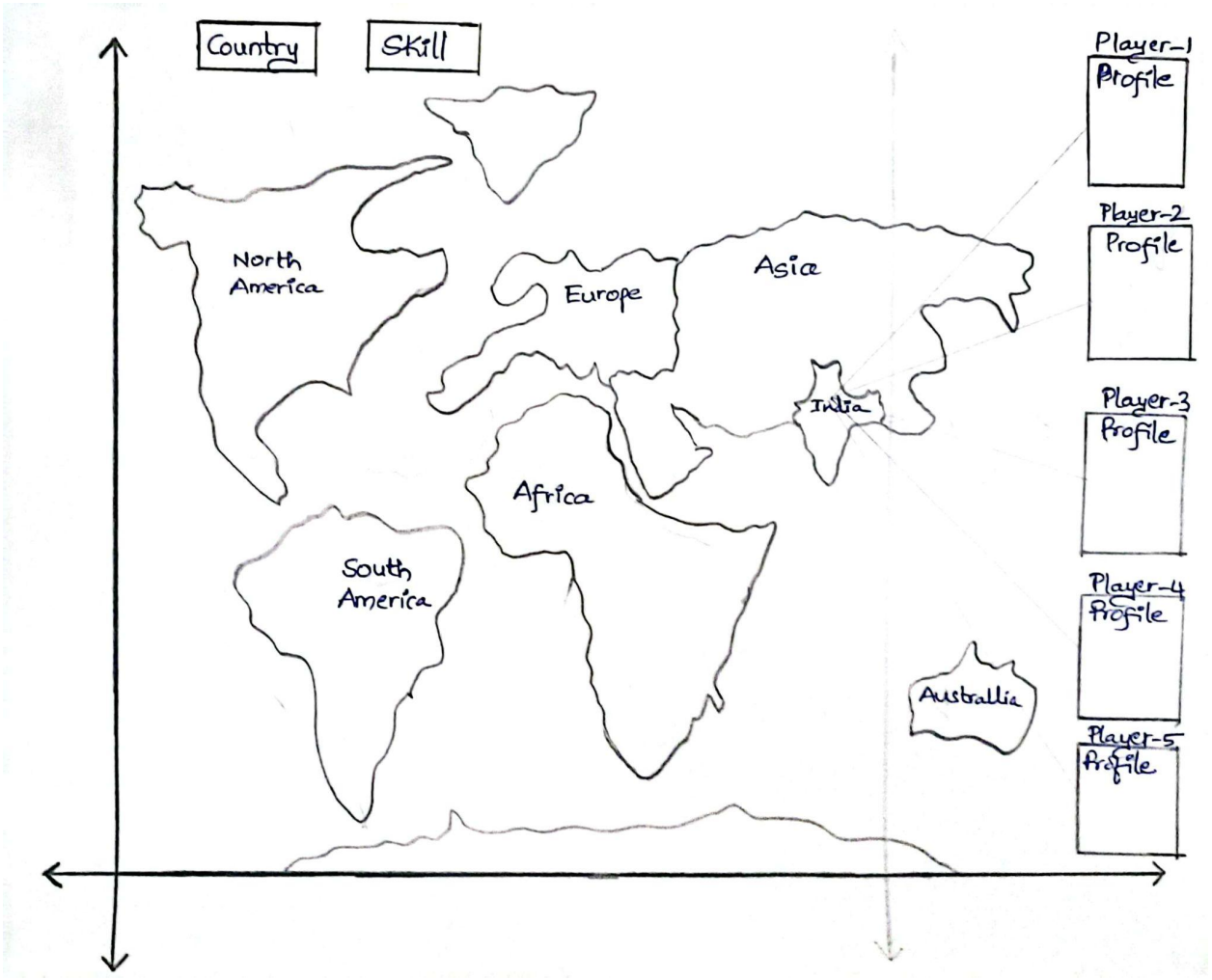
Sketch 3:


WORLD
Country 1
Country 2
Country 3
Country 4
Country 5
Country 6
Player 1
Player 4
Player 3
Player 2
Player 2
Player 2
Player 2
Biodata
Player 4
Biodata
Player 2

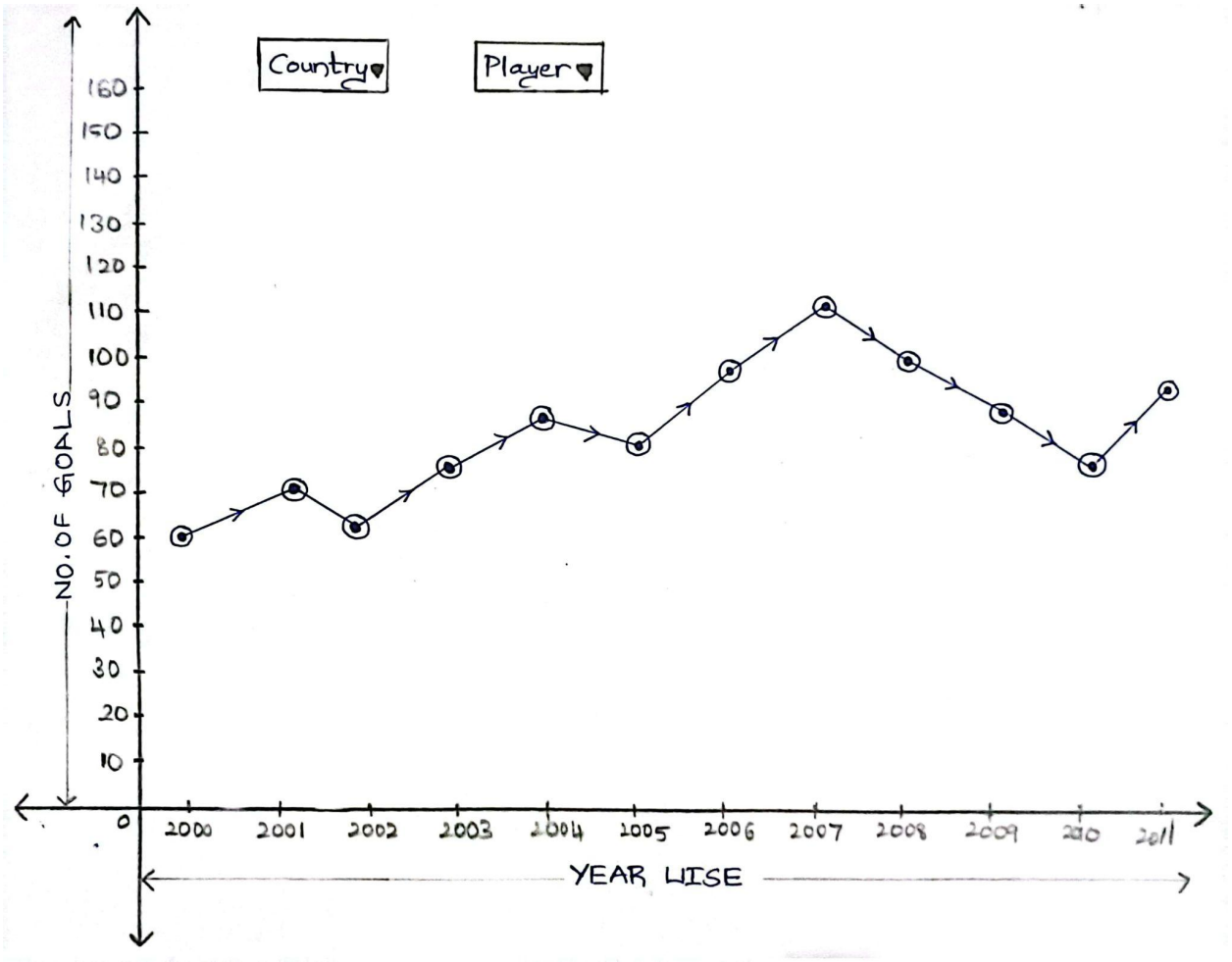
Sketch 4:


P1
P2
P3
P4
P5
P6
P7
P8
P9
Nodes – Players
— Line Drawn with pen – Positive Interactio
— Line Drawn with pencil – Negative Interaction.

Sketch 5:


Country
Skill
North America
Europe
Asia
India
Africa
South America
Australia
Player-1 Profile
Player-2 Profile
Player-3 Profile
Player-4 Profile
Player-5 Profile

Sketch 6:



Steps to first Project Milestone :

Scope of Prototype:

The prototype "FIFA VizPlay" seeks to provide an interactive visualization of the intricate details within the world of football, going beyond the traditional metrics. The scope includes providing insights into player performances, match outcomes, player's physic and skills metrics, and fostering data-driven discussions in the football community.

Data Collected:

Data is sourced from the FIFA player database and comprises two main datasets.
The first dataset focuses on match specifics like match_id, team_id, player_id, etc.
The second dataset delves into individual player statistics, skills, career trajectory, and market value.

Dataset source: https://www.kaggle.com/code/sivsankar/fifa22-recommender-system/input

Data Still Needed:

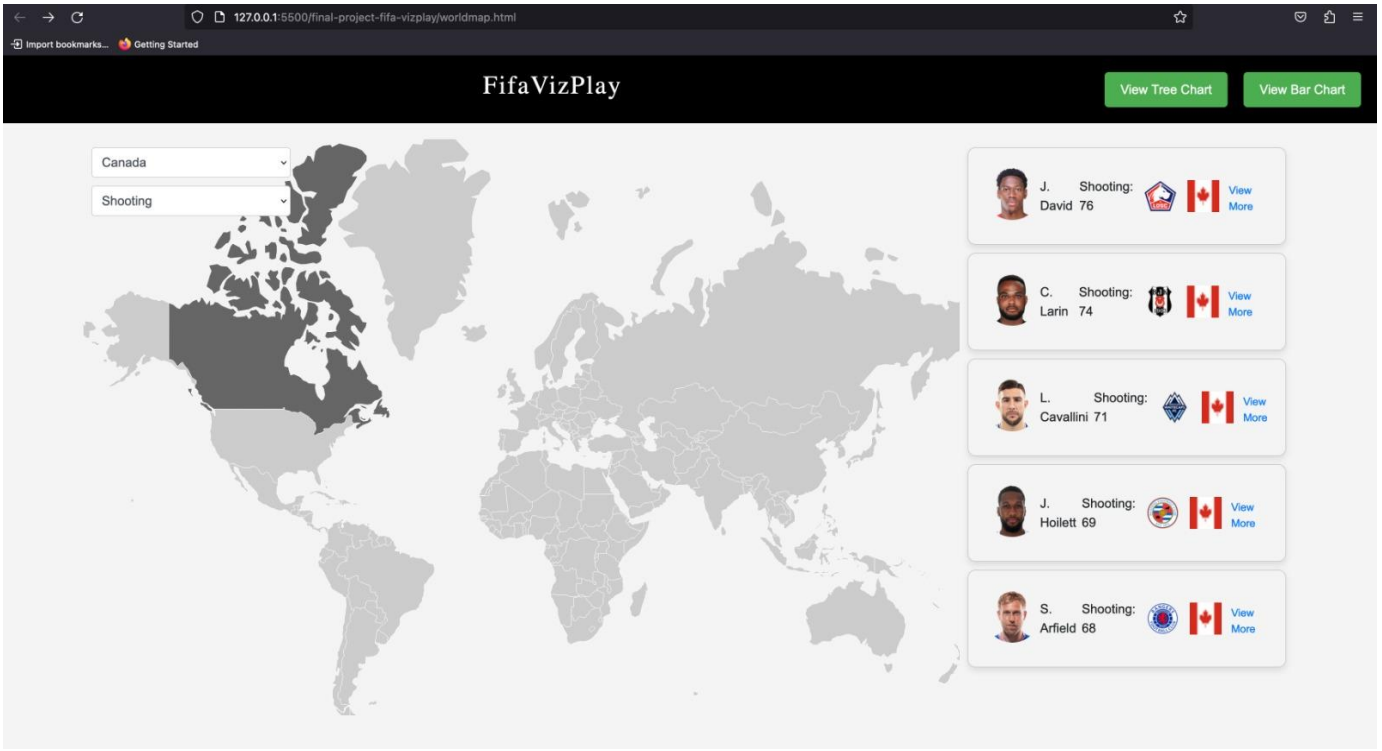
As of now, we have a comprehensive collection of player and match-specific data. However, future iterations may require data on team and player interactions etc., to offer a more holistic view.

Important Data Structures:

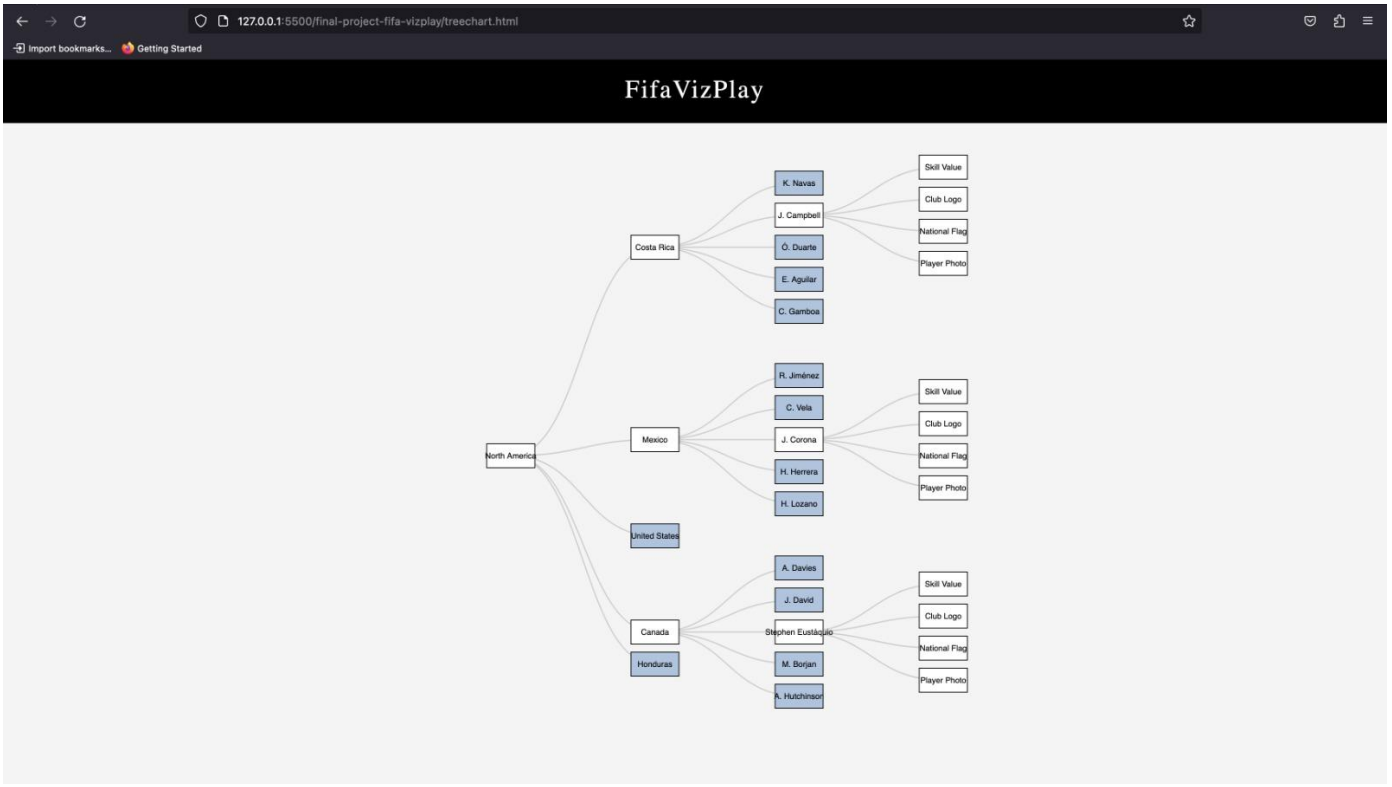
Match specific identifiers such as match_id, team_id, player_id.
Individual player statistics, encompassing player skills, career trajectory, market value, player attributes, and performance metrics.
Aggregated metrics like goal statistics on a yearly and country-wise basis.

Implementation Code Fragments:

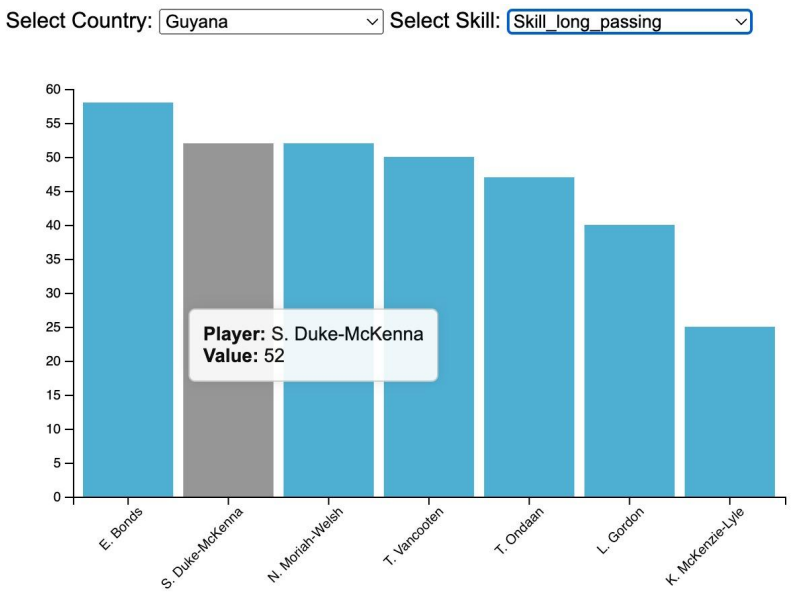
- **Data Cleanup:** Identify and handle missing data, merge datasets by linking relevant columns like player IDs and team IDs, player IDs and skills.
- **Metric Derivation**: Calculate player performance scores based on attributes, aggregate goal statistics.
- **Visualization:** Implement charts, graphs, and visualizations using D3.js and other visualization tools.
- **Geographic and Hierarchical Data -** World and Tree Map Prototype: Preliminary and refined designs illustrate the world map view of top football countries and a tree map breakdown of their premier players. This interactive interface offers a comprehensive insight into the global football landscape, emphasizing both country prominence and individual player excellence.



The Global Football Atlas: Click on any country to unveil its top 5 football stars according to the skill selected from the drop down as shown in the above image, also each with a detailed player card on the right. Dive deep into this dynamic map to witness the spread of world-class talent and get a closer look at each nation's football icons.

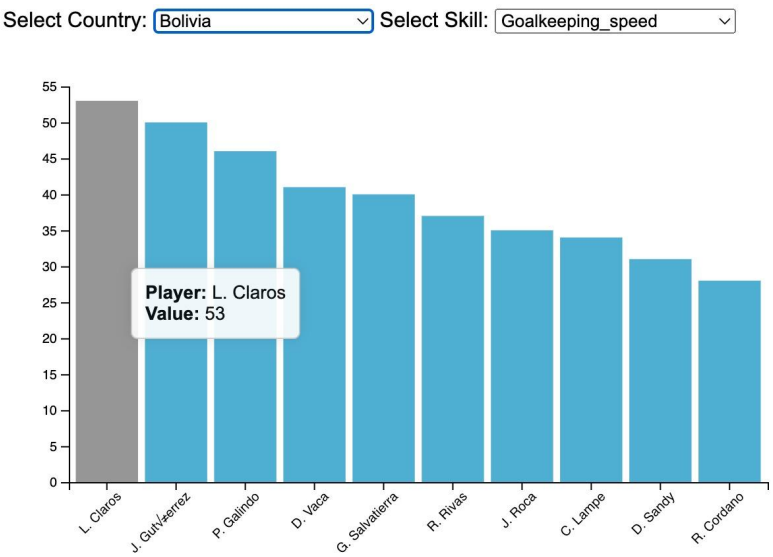


Football Elite Hierarchy: Explore the world's top 5 football nations and delve deeper into their standout players. This chart provides a layered view, first showcasing leading countries in football, and then spotlighting their top 5 players. A visual journey into the heart of global football excellence.



Guyana graph (long passing):

- "Comparison of long passing skills among Guyanese footballers. S. Duke-McKenna stands out with a value of 52."
- "Showcasing Guyana's talents: Long passing abilities of prominent players."



Bolivia graph (goalkeeping speed):

- "Bolivian goalkeepers ranked by their speed. L. Claros leads the pack with a score of 53."
- "A glance at Bolivia's finest: Evaluating goalkeeping speed of top players."

FifaVizPlay Overview and Use Cases:

- **Overview:** "FIFA VizPlay" offers a dynamic platform for users to delve deep into football data, fostering informed, data-driven discourse.

Use Cases:

- **Football Enthusiasts:** Dive deep into player statistics, compare players based on attributes, and analyze performance metrics of national teams.
- **Researchers:** Understand trends and patterns, analyze correlations between player attributes and performance.
- **Football Analysts:** Get insights on club-level statistics, understand league-level data and dynamics, and create profiles for individual players or clubs summarizing their key attributes.
- The visual sketches can be included where they fit best, especially under the “Structural Data" and FifaVizPlay Overview and Use Cases sections to give a visual representation of the interface and data structures.