The FIFA Story

Final Report

Daksh (2022101068) **Jay** (2021101035) **Kritika** (2021101125)

Introduction

We have tried to work out some of the visuals for the world's most prestigious Football tounrnament - FIFA.

Proceeded with visuals that help us understand information regarding the previous years datasets as well as making different analysis of strategies & that of some of the top selected players too based on their past records and physicalities.

We intend and aim to provide sort of a story-line flow for FIFA data that builds from basics and is useful for the absolute beginners as well as some of the enthusiastic analysts too.

Content Links

Website LINK : https://daxvshah11.github.io/the_FIFA_Story
Video LINK : https://youtu.be/hxgGMX5X9fg
GitHub Repo LINK : https://github.com/Daxvshah11/the_FIFA_Story

Dataset

The dataset chosen has 3 different subsets in it. There are 3 csv files within, namely:

- **World Cup Matches :** *contains columns for WC matches with datapoints like year, date-time, stage, stadium, city, home team, away team, goals for both teams, win conditions if any, half time scores, referees, round IDs and initials for both teams*
- **World Cup Players :** *it is about data related to all the players who have ever participated in the WC with columns like round ID, match ID, team they belong to, their coach, line-up, shirt number, player name, position and event*
- **World Cups :** *this is about the world cups that have been conducted containing info like the year, host country, WC winner of that year, runners-up, third & fourth place positions, goals scored, qualified teams, matches played, attendance*

Other than these 3 files, we have used some of the custom datasets from various resources which includes :

- data for physical details of top 5 chosen players for comparision (eg. height, weight, age etc.) with other data points like their number of International FIFA competitive matches played and number of goals scored in them. Wikipedia, custom data
- png image of the Official FIFA trpophy as a data-point for masking used in creating a well masked word cloud over it
- Python code generated png images for heatmaps as well as field (shot) maps that are chosen according to the interaction from the user for selected country against France

Phase 1 Brief

In the first phase, we intended to do a little research on the topic oursevles and try to understand the actual world usecases of the topic as well as why it should be done in the first place. We also tried to analyse its necessity along the same lines.

Further, once we were learned about the same, we moved onto knowing the humongous amounts of data generated per match in the sport of Football. This also led us to choose specifically FIFA as that is a tournament which is valued by each and every human soul supporting the sport over the whole globe.

In addition to the same, we then finally moved onto the part of choosing a suitable and implementable data as well as also selecting some graphs and charts that we actually felt like important and worth enough to be visualised.
Once those were finalised, we concluded our phase 1 with a lot of learnings about the topic in general and in specifics with the course too.

Phase 2 Brief

In the second phase, after positive comments from the TAs as well as the professor on the concept of the project, we actually started implementing the charts and were also successful in doing so for first 2 selected charts as well as a third one partially.

While implementing the charts in reality, we encountered some other innovative thoughts and ideas about better visualisations also. Plus, there was also a valuable feedback from the TAs on the same so we decided to remove some simpler ones from the expected charts in the final project and added some other in place of those to make the final visualisation even better than before.

We also planned to expand our dataset compared to the initially found data and thus, also found and stored the data set that we were planning to work on.
Thus, we implemented the first chart which literally introduces people and viewers to the *"WORLD of FIFA"* and which country has participated how many times as well as won the WC for what number of times.

After that, we implemented a bar race chart talking about how different countries have progressed in the main interesting aspect of the sport, *ie. Goal Scoring*. The race chart follows the time stamps of each 4-yearly occurring FIFA and how accumulatted total of the goals scored by each country progressed.

This indeed concluded our phase 2 on a good note. Below are some images for the same.

Figure 1: Choropleth Vis

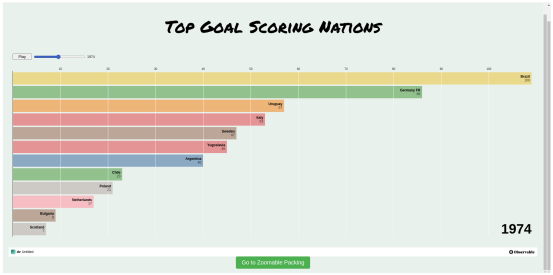

Figure 2: Race Chart Vis

1 Visualisation : Choropleth

- **Location Focus :** Choropleth maps excel at displaying variations across geographical regions, making them ideal for visualizing trends in FIFA data tied to specific countries. This could include metrics like viewership figures or national team's performance.
- **Accessibility :** Choropleth maps are known for their clarity and ease of interpretation. This is particularly beneficial for communicating complex FIFA data to a broad audience, ensuring everyone can grasp the key insights from the visualization.

2 Visualisation : Word-Cloud

- **Keyword Emphasis :** Word clouds effectively highlight frequently used terms within your FIFA data text source. This can reveal prominent themes, player names, popular teams, better performing team etc. at a glance
- **Engagement Factor :** The visual nature of word clouds makes them inherently engaging. This can be particularly useful as we have used the shape of the Official FIFA trophy itself to mask the word cloud of most winning teams.

3 Visualisation : Bar Race Chart

- **Dynamic Storytelling :** Bar race charts present data as an animation where bars representing different variables compete over time. This creates a dynamic and engaging way to showcase how metrics in your FIFA data evolve and shift positions throughout periods.
- **Comparative Analysis :** By visualizing multiple variables simultaneously, bar race charts facilitate easy comparison between different aspects of your FIFA data. This allows viewers to quickly identify trends, track the rise and fall of teams in rankings, or compare goal scorers over time.
- **Focus on Change :** The animated nature of bar race charts naturally draws attention to how data points change over time. This makes them ideal for highlighting significant shifts in FIFA data, such as unexpected team performance surges or historical ranking changes. We have used these same reasons to back our choice of using a bar race chart for showing number of goals scored over the years by each team in FIFA

4 Visualisation : Zoomable Circle Packing

- **Hierarchical Exploration :** Zoomable circle packing charts excel at displaying hierarchical data structures. This allows viewers to explore FIFA data at different levels of detail. For instance, you could visualize player positions within a team, historical performance trends within a specific region, or even zoom in to see individual players and their statistics.
- **Relationship Visualization :** By highlighting connected circles, these charts can reveal relationships between different elements in your FIFA data. This can be helpful for showcasing team rivalries, player transfers between clubs, or dependencies between various game statistics.
- **Interactive Data Discovery :** The zooming functionality empowers viewers to delve deeper into specific sections of the chart. This allows for a more interactive exploration of your FIFA data, enabling viewers to focus on areas of particular interest, like a specific team's performance over time or player connections within a league.

5 Visualisation : Heat Maps

Chosen country for this was ***"FRANCE"***. After their 2018 WC win, their playing technique as well as their team was being called as the ***"most Balanced"*** team by most Football analysts as well as critics. Hence the choice!

- **Spatial Analysis :** Heatmaps excel at visualizing data with a spatial component, making them ideal for analyzing trends related to player positioning on the field. This could be helpful for showcasing shot locations, player movement patterns during a match, or even analyzing team formations across different leagues.
- **Identifying Trends :** By using color intensity to represent data values, heatmaps effectively reveal trends and patterns within your FIFA data. This could be useful for visualizing player performance metrics across different positions, analyzing goal scoring hotspots on the field, or comparing fan engagement on social media for various teams.

6 Visualisation : Field Maps

- **Tactical Insights :** Field maps provide a clear representation of the football pitch, making them ideal for visualizing tactical aspects of FIFA data. This could be helpful for analyzing player positioning during set pieces, tracking player movement patterns throughout a match, or even showcasing heatmaps of specific zones on the field (e.g., penalty area activity). This was the same reason we used it in our story telling to analyse performance of Team FRANCE with some other countries.
- **Action Visualization :** By overlaying data points onto the field map, you can effectively visualize in-game actions. This could be useful for showcasing pass networks between players, analyzing shot locations and trajectories, or even tracking player movement during defensive formations.

7 Visualisation : Multiple Parallel Coordinates

- **Comparative Analysis of Players :** Multiple parallel coordinate plots excel at visually comparing players across various attributes. Each player is represented by a distinct line, with their values for different attributes plotted along the axes. This allows viewers to identify similarities and differences between players' various aspects and records as we have used in our visualisations.
- **Identifying Trends and Outliers :** By analyzing the lines in a parallel coordinate plot, you can identify trends and potential outliers among the players. Players with lines that run close together might exhibit similar skillsets, while those with lines deviating significantly from the pack could be outliers with unique strengths or weaknesses.

Utilities & Use-cases

There are innumerable use-cases to this visualisation and some of them are as mentioned below :

- Identifying trends in participation by demographic to plan future tournaments and modifying marketing strategies
- Vivid and modifiably-interactive comparision between some of the best players to ever play the sport
- Analyze the popularity of the World Cup in different regions to attract sponsorships effectively
- Gain insights into historical participation patterns and interesting facts about team performances as well as strategies
- Explore past champions and relive iconic moments in World Cup history
- Create data-driven stories about the evolution of the World Cup along with the Sport
- Learn interesting facts about the World Cup in a visually appealing way

These are some of the use-cases of this visualisation based on our scope of the project.

Users of the Visualisation

According to the scope of the project and based on the use-cases mentioned above, here are some of the typical users for this visualisation :

- **FIFA Organisers**
- **FIFA Sponsors**
- **Football fans**
- **Journalists & Media**
- **Casual viewers**

Improvement Prospects

- **Scalable :** We have tried to incorporate as much as possible in each of the visualisation at this scale but the way we have chosen the visualisations, they are easily scalable to as many number as wanted. For eg. multiple parallel coordinates chart can be easily scaled to compare any number of players and even the domain for choosing the players can be increased to full capacity.
- **Limited Aspects :** We have considered only aspects like country-wise analysis as well as player based or field based analysis. But, there are many other important aspects for potential great visuals for eg. about viewership from each country and its growth over the years for marketing purpose, about better players and their relations with physical attributes or their history in clubs etc. and the list would keep going.

Contributions

Daksh Shah

- Choropleth vis
- Multiple Parallel Coordinates vis
- Word Cloud Generation vis
- Heatmaps vis
- Website CSS
- All documentation & reports

Jay Muppidi

- Bar Race Chart vis
- Zoomable Circle Packing vis
- Field Shot Maps vis
- Website CSS
- PPT & Video

7.1 Kritika Gautam

- Partly HTML handled
- Partly CSS handled