CIS-568 Data Visualization Project Report

"International Football Match Analysis"

Authors

Nikhil Premachandra rao (02105149)
Prajwal Sreeram Vasanth kumar (02133549)
Rakshith Bangalore Siddalingaiah (02117995)

Section 1:
What data is being visualized?

This dataset visualizes international football results from 1872 to 2017, providing insights into the outcomes of football matches on a global scale.

Explore the extensive historical data spanning team performances, match details, and tournament outcomes, providing a full view of the growth of international football across time.

What are the tasks (Action/Target) of the visualizations?

The visualizations aim to uncover patterns, trends, and statistical insights within international football data, allowing for a more in-depth understanding of team dynamics, match outcomes, and historical developments to inform strategic decisions, analyze performance metrics, and draw meaningful conclusions for football fans and analysts alike.

HOW: The final visualization will contain:

1. GeoGeo-Political Football: Mapping Global Match Outcomes

How: Use D3.js to create a world map visualization. Plot match locations as points on the map. Upon selection, trigger a red highlight and display match specifics at the top. Use geographical coordinates to position match locations on the map accurately and have a drop-down to select the year.

Show data variation: Used red color to highlight match outcomes and keep the design simple to focus on match details.
Maximise data-ink ratio: Avoid unnecessary clutter; emphasize match data and minimize distractions in the map visualization on selecting a year from drop down.

2. Analysing Home and Away Game Distributions Across Continents

How: Created dropdown menu options for continents. Based on selection, generate histograms separately for home and away game statistics for the chosen continent. Utilise D3.js to plot histograms showcasing game distributions.

Data density: Ensure the charts show clear home and away game statistics distributions for easy comparison.

3. Deciphering Match Diversity Through Bar Graph Visualisation

How: Use D3.js to generate a bar chart showing the top 10 different match types on the x-axis and their frequency on the y-axis. Each bar's height represents the count of each match type.

Provide comparisons: Arrange bars for quick comparison and facilitate understanding of the frequency of each match type.

4. Exploring League Dynamics: Top 10 Teams Across Diverse Leagues

How: Create a bar graph using D3.js to display the top 10 winning teams in a tournament. Used different colors for each team.
Clear, detailed, and accurate labeling: Label teams clearly on the graph.
Avoid chartjunk: Keep the focus on the top 10 teams winning distribution without unnecessary elements.

5. Chronological Match Analysis Through Heat Map Visualization

How: Use D3.js to create a heat map representing match frequency over time. Uses shades of green to indicate match density across different years.
Clear and detailed scale: Ensure the heat map's legend or scale is clear to convey match density accurately.
Effective use of space: Use the available space to represent match frequency across the timeline without overcrowding

Section 2: Design

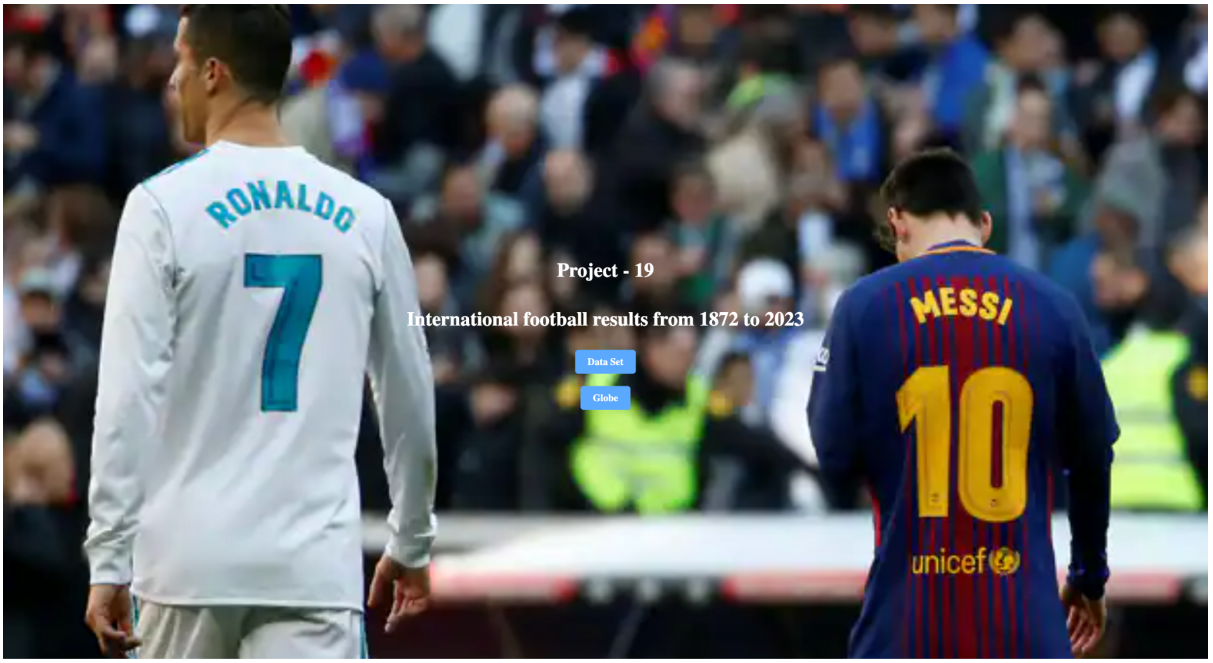

Fig 1: Homepage

Goalscorers

date	home_team	away_team	home_score	away_score	tournament	city	country	neutral
2019-01-14	United Arab Emirates	Thailand	1	1	AFC Asian Cup	Al Ain	United Arab Emirates	false
2010-06-25	Chile	Spain	1	2	FIFA World Cup	Pretoria	South Africa	true
1954-12-05	Italy	Argentina	2	0	Friendly	Rome	Italy	false
2011-01-16	Qatar	Kuwait	3	0	AFC Asian Cup	Doha	Qatar	false
2013-08-20	Afghanistan	Pakistan	3	0	Friendly	Kabul	Afghanistan	false
2017-10-10	South Korea	Morocco	1	3	Friendly	Biel	Switzerland	true
1958-10-19	Réunion	Madagascar	1	3	Friendly	Saint-Pierre	Réunion	false
1992-11-22	Jamaica	El Salvador	0	2	FIFA World Cup qualification	Kingston	Jamaica	false
1983-07-25	Gambia	Guinea-Bissau	2	2	Amílcar Cabral Cup	Nouakchott	Mauritania	true
1991-07-17	Argentina	Brazil	3	2	Copa América	Santiago	Chile	true
2019-07-27	Burundi	South Sudan	2	0	African Nations Championship qualification	Bujumbura	Burundi	false

Results

date	home_team	away_team	home_score	away_score	tournament	city	country	neutral
1984-10-31	Czechoslovakia	Malta	4	0	FIFA World Cup qualification	Prague	Czechoslovakia	false
1988-04-27	Hungary	England	0	0	Friendly	Budapest	Hungary	false
1976-10-12	Peru	Uruguay	0	0	Friendly	Lima	Peru	false
2015-10-17	Uganda	Sudan	2	0	African Nations Championship	Kampala	Uganda	false
2012-04-12	Egypt	Nigeria	3	2	Friendly	Dubai	United Arab Emirates	true
1968-11-08	Sudan	Zambia	4	2	FIFA World Cup qualification	Khartoum	Sudan	false
2011-09-03	Kenya	Guinea-Bissau	2	1	African Cup of Nations qualification	Nairobi	Kenya	false
1977-04-09	South Africa	Zimbabwe	7	0	Friendly	Johannesburg	South Africa	false
2023-06-19	Hong Kong	Thailand	0	1	Friendly	Hong Kong	Hong Kong	false
2015-10-13	Uruguay	Colombia	3	0	FIFA World Cup qualification	Montevideo	Uruguay	false
2017-12-05	Kenya	Libya	0	0	CECAFA Cup	Machakos	Kenya	false

Shootouts

date	home_team	away_team	winner	first_shooter
2010-07-02	Uruguay	Ghana	Uruguay	Uruguay
1993-06-26	Colombia	Uruguay	Colombia	Colombia
2022-02-03	Cameroon	Egypt	Egypt	
1984-12-19	Central African Republic	Gabon	Gabon	
2015-07-19	Trinidad and Tobago	Panama	Panama	
2006-07-01	England	Portugal	Portugal	Portugal
1985-10-11	Kenya	Uganda	Kenya	
2005-07-28	Kuwait	United Arab Emirates	United Arab Emirates	
2016-06-03	Bosnia and Herzegovina	Denmark	Bosnia and Herzegovina	
2000-10-28	Kenya	Uganda	Uganda	
2008-06-22	Spain	Italy	Spain	Spain

Next

Fig 2: Dataset page

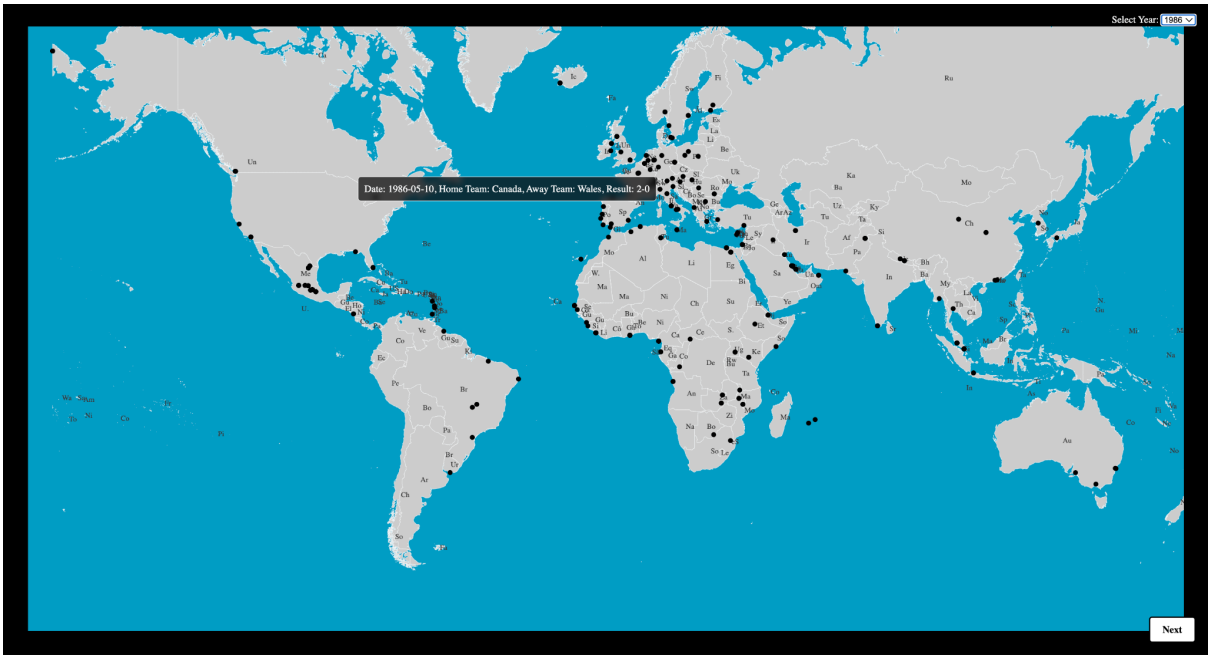

Fig 3: Political Map of the World containing matches held at different locations.

Geo-Political Football: Mapping Global Match Outcomes

Axis: The map serves as the axis, displaying the geographical locations.

Marks: Points representing match locations on the map.

Position Encoding: Latitude and longitude coordinates for placing match points accurately on the map.

Color Encoding: Red highlight for the selected match point.

Tooltip Encoding: Detailed match information upon selection.

Dropdown Encoding: Upon selecting the year the points are filtered accordingly.

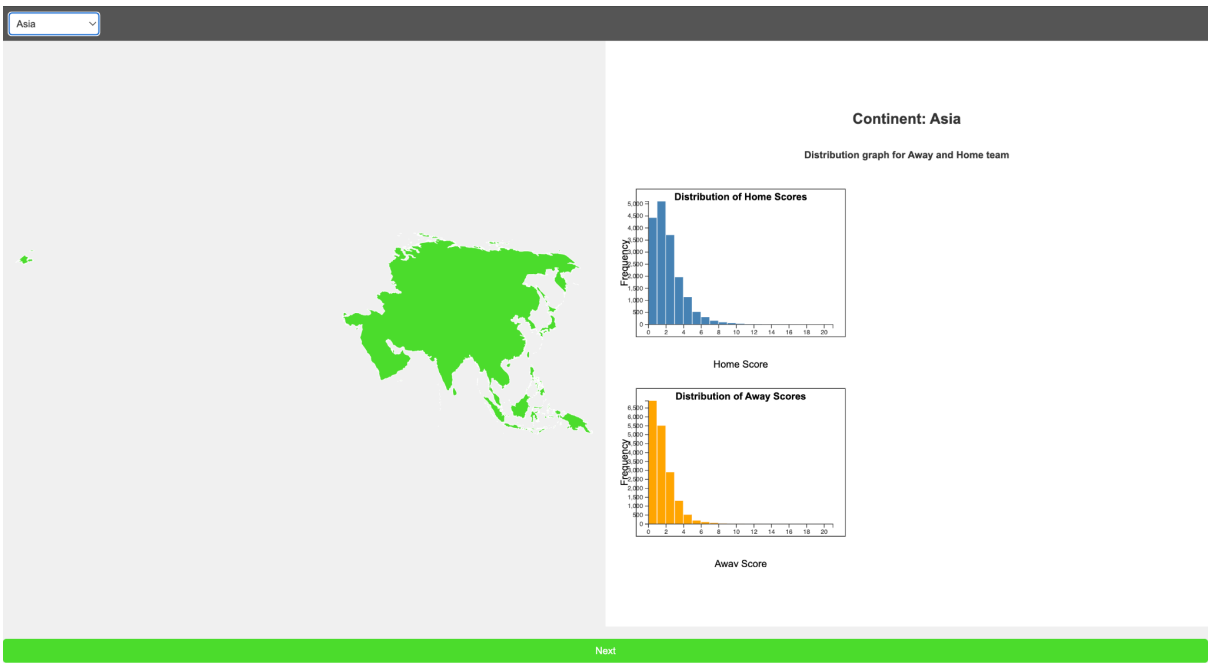

Fig 4: Analysing home and away matches using charts.

Analysing Home and Away Game Distributions Across Continents

Axis: Horizontal axis for game score statistics (home and away), vertical axis for frequency/distribution.

Marks: Histogram bars showing frequency or count.

Position Encoding: Placement of bars based on game statistics and their frequency.

Color Encoding: Distinct color for home and away game histograms.

Dropdown Encoding: Selection of continents to display specific histogram data.

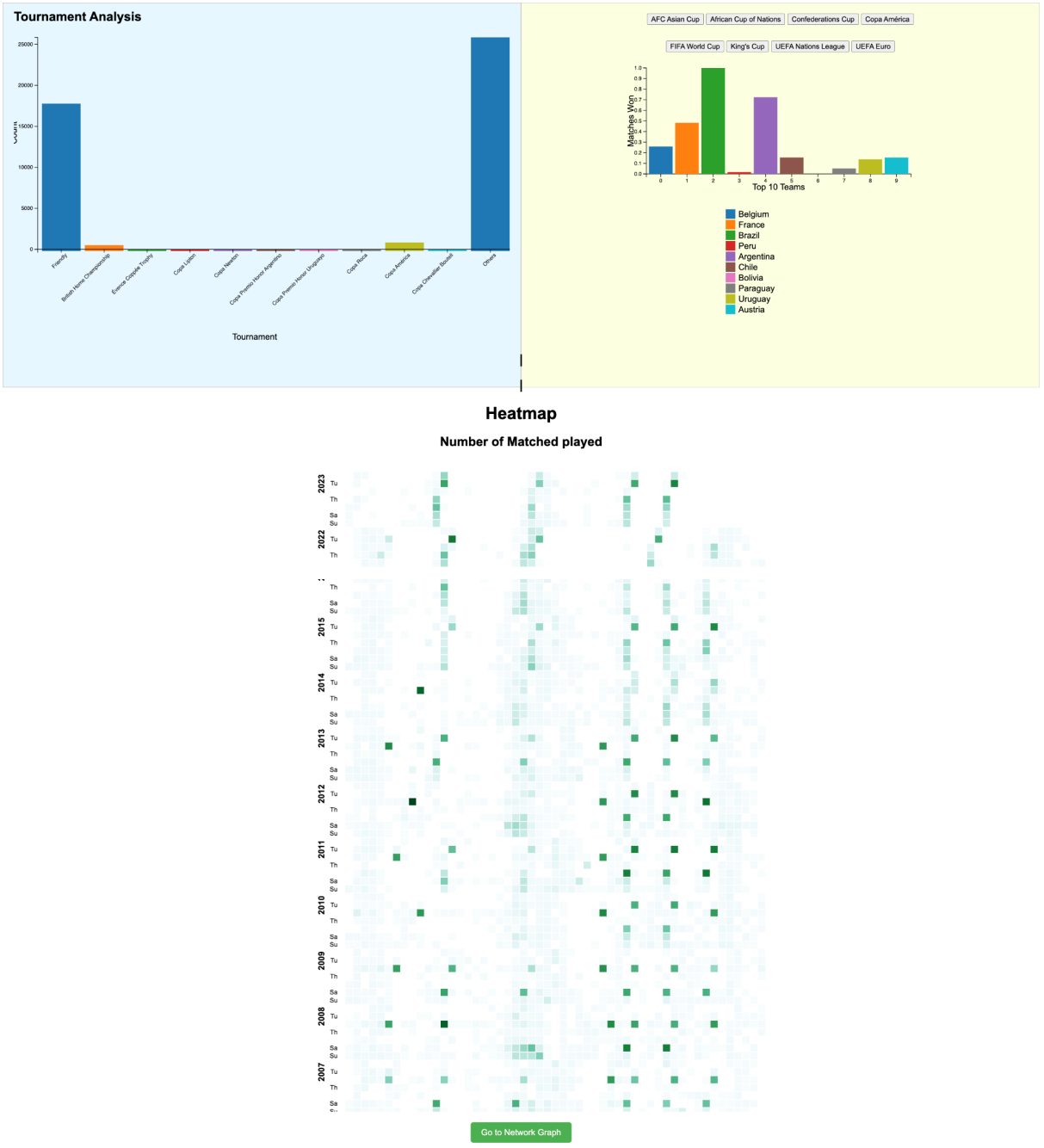

Fig 5: Visualising a Chronological Match Analysis using Heat Maps and other Bar graphs

Deciphering Match Diversity Through Bar Graph Visualisation

Axis: The horizontal axis has the Top 10 teams for different match types, and the vertical axis for frequency/count.

Marks: Bars representing the count of each match type.

Position Encoding: Placement of bars based on match type and frequency.

Height Encoding: Height of bars representing the count of each match type.

Network Graph

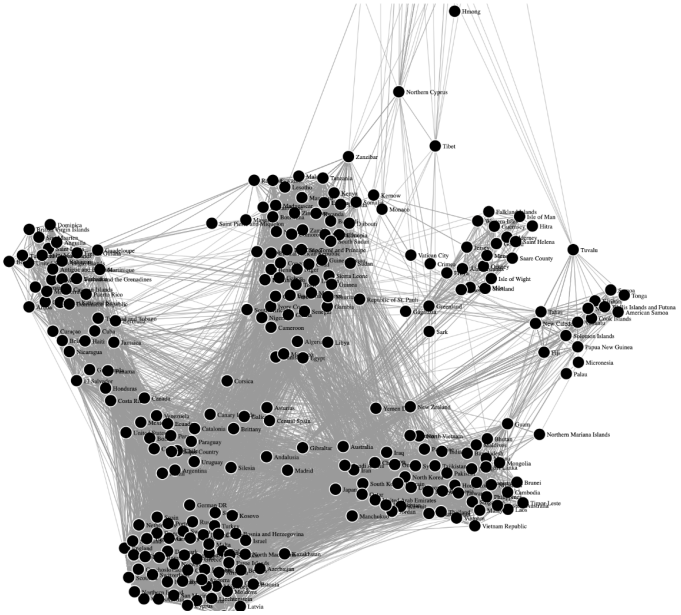

Fig 5: Visualising all Match played and the home and away team Analysis using Network graphs Analyzing Football Team Relationships Through Network Graph Visualization

Nodes and Links: Nodes represent football teams, interconnected by links indicating relationships.

Layout: The graph layout is based on force-directed simulation, ensuring an aesthetically pleasing and interactive display.

Node Positioning: Nodes are strategically positioned to ensure clarity and minimize overlap.

Interactivity: Users can drag nodes for better exploration and understanding of the network.

Node Labels: Displaying team names as labels to identify nodes easily.

Axis: The horizontal and vertical axes represent the Top 10 teams for different match types and the frequency/count, respectively.

Marks: The bars on the graph represent the count of each match type.

Position Encoding: Bar placement on the horizontal axis is based on match type, while the vertical axis indicates the frequency/count.

Height Encoding: The height of each bar represents the count of each match type.

Code Modularity: The code is organized into functions for better readability and maintenance.

Section 3: Discussion

Goals achieved:

Phase 1 :

Political Map: Initially we had a political world map with points plotted on the map of matches played using latitude and longitude and it had the match points of all the years from 1872 to 2023.

Phase 2 :

Political Map: Points on a political map when selected shows a red circle with details of the match and we added a drop-down of the year column when selected it filters out the specific matches played in that year.

Home and Away Distribution page: This page has a drop-down button for continents upon selecting a particular continent home and away statistics are plotted using histograms.

Statistics Page: This page has three different Visualised charts:

- Tournament analysis of the top 10 matches played is visualized using a bar graph.
- The top 10 teams were visualized using a bar graph of different tournaments upon selection from the button on the top.
- Heat Map which contains the density of matches played over the period.

Network graph: It connects all the matches played to the home and away teams and plots the network graph.

Obstacles:

- While plotting the graphs we were facing issues in printing the axes.
- In the Data we didn't have a latitude and longitude column, we had to pre-process the data points to map them from state and city.
- In plotting a bar graph in the Statistics page it was clustered. We had to filter the top 10 teams.
- The Heat Map contained an extensive amount of data, and due to its size, we were unable to handle it horizontally; instead, we had to approach it vertically.

Limitations:

- The bar graph shows the top 10 teams of a particular tournament; if it has less than 10 teams it will show all the teams from that tournament.
- If we choose a year on the political map page with no recorded matches played, it won't be displayed on the map.

Code:

GitHub link for Code
The web page can be accessed at this Link