



## **DATA VISUALISATION**

**ISM 6419**

### **FINAL PROJECT REPORT**

**World Cup Results and Socio-Economic Impacts, 1962-2014**

**Professor**

**Johannes Reichgelt, PHD**

By

Priyanka Prabakarrao (U52145257)

## TABLE OF CONTENTS

1. Introduction .....	3
2. Methodology .....	4
Data Sources.....	4
Data Processing.....	4
Analysis Tools.....	5
3. Analysis .....	5
Dashboard .....	5
Visualization 1 .....	6
Visualization 2 .....	7
Visualization 3 .....	7
Visualization 4 .....	8
Visualization 5 .....	8
Visualization 6 .....	9
Visualization 7 .....	10
Visualization 8 .....	11
Visualization 9 .....	12
4. Conclusion .....	12
Key Findings .....	12
Implications and Future Research .....	13

# 1. Introduction

The FIFA World Cup, held every four years since 1930, represents the zenith of international football competition. It transcends mere athletic contest, embodying a cultural phenomenon that captivates billions across the globe. Beyond the spectacle of sport, the World Cup also serves as a mirror reflecting the socio-economic conditions and capacities of its host nations. This analysis is predicated on the premise that the organization and outcomes of these tournaments are intricately linked to the economic and demographic profiles of the countries that host them.

This project delves into the World Cup results from 1962 to 2014, employing a dataset that includes detailed records of matches, teams, and final standings, enriched with socio-economic indicators such as GDP and population statistics. By integrating these diverse data sources, the study aims to uncover patterns and derive insights into how economic and demographic factors influence key aspects of the World Cup, including attendance figures, the number of matches hosted, and overall tournament success.

The research questions guiding this analysis are:

1. How does the population size of the host countries affect the attendance of the World Cup tournaments?
2. What is the relationship between the economic power (measured in GDP value) of a country and its history of participation in the World Cup, particularly in terms of total matches played?
3. Which income groups do the host countries fall under, and how might these classifications relate to their capabilities and effectiveness in hosting the tournament?

Additional questions such as Countries that won the world cup, Teams by total finals and semifinals, Countries that hosted the world cup, Countries qualified to the world cup, Competitions stats by year, Audience by country and year to provide a comprehensive overview of different facets of the World Cup's history and its global impact.

## 2. Methodology

### Data Sources

World Cup Results (1962-2014):

- Source: Kaggle.
- Dataset Link: <https://www.kaggle.com/datasets/abecklas/fifa-world-cup/data>
- Details: This dataset contains detailed records of match results, including scores, locations, and dates, along with broader tournament statistics such as standings and progression through stages. It provides a comprehensive overview of the games played during each World Cup from 1962 to 2014.

GDP and Population Data (1960-2022):

- Source: The World Bank.
- GDP (1960-2022) - <https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG>
- Population (1960-2022) - <https://data.worldbank.org/indicator/SP.POP.TOTL>
- Details: This data set offers annual GDP and population figures for all countries. For this project, special attention was given to the World Cup host nations during the years they hosted the tournaments, which allows for an examination of the relationship between a country's economic and demographic context and its hosting of the World Cup.

### Data Processing

- Tool Used: Python.
- Procedure:
  - Extraction: Data relevant to the specific years when the World Cup was held was extracted, focusing on host nations during those particular years.
  - Filtering: The datasets were filtered to isolate only the necessary years and data points pertinent to the World Cup host years.
  - Transformation: The data was transformed to align the formats and ensure consistency across the different datasets, making it suitable for comparative analysis.
  - Integration: The World Cup results data was integrated with the socio-economic data (GDP and Population) using the country and year as keys. This integration facilitated a multi-dimensional analysis of how socio-economic factors might influence World Cup outcomes.

## Analysis Tools

- Tableau: The prepared data was then imported into Tableau, where it was used to create various visualizations and dashboards.

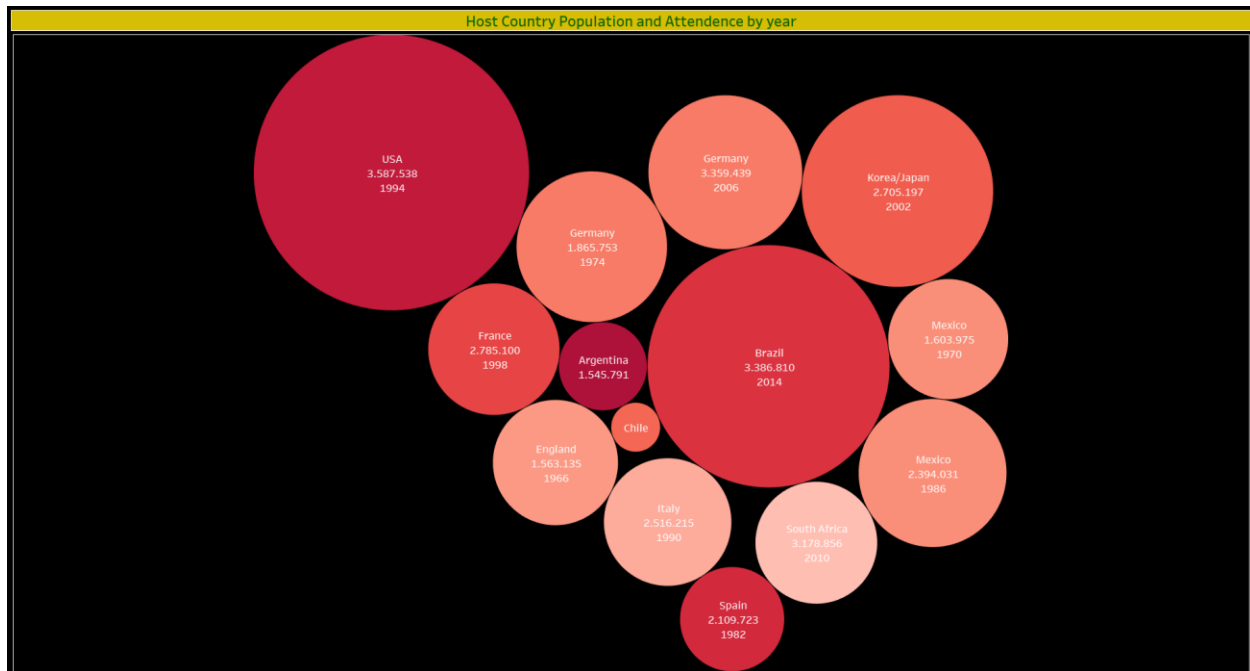
## 3. Analysis

### Dashboard



The dashboard consolidates all the visualizations as depicted, with each individual visualization described in detail below.

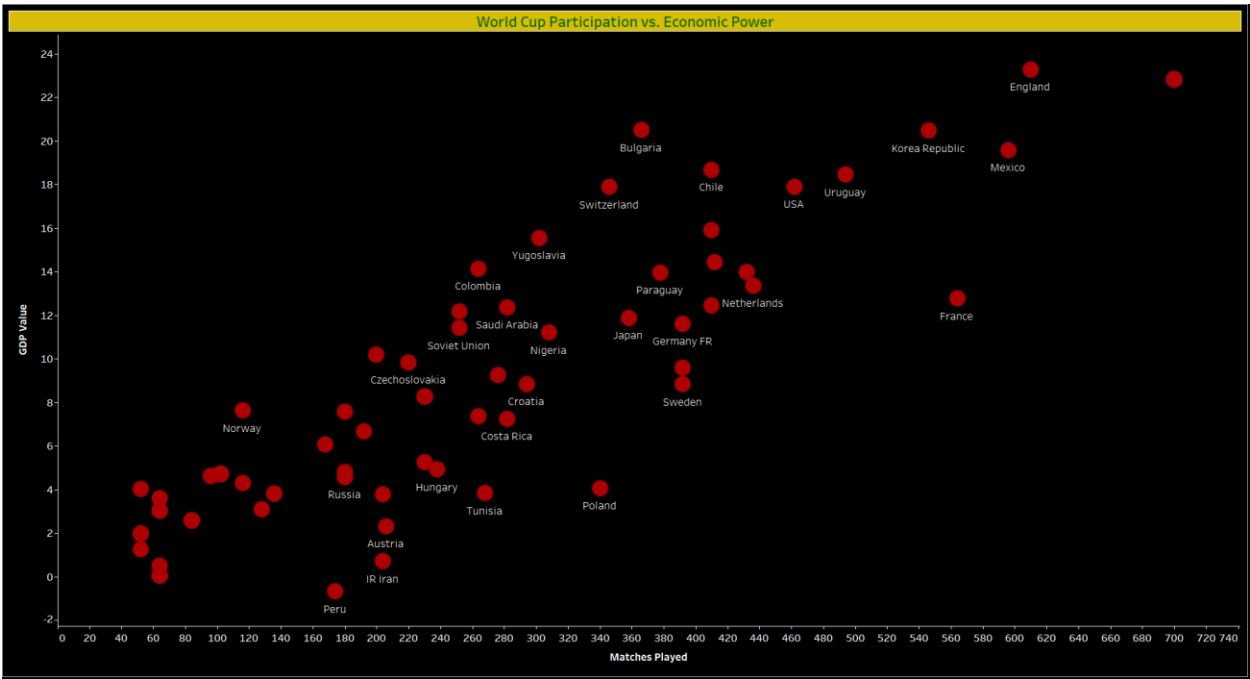
## Visualization 1



The visualization created to explore the impact of host country population size on World Cup tournament attendance employs a packed bubble chart, effectively illustrating the relationship between these two variables across different tournament years.

The observation drawn from this visualization indicates a clear trend: countries with larger populations tend to experience higher attendance figures at World Cup events. For instance, the United States, which hosted the World Cup in 1994, had the highest recorded population among host countries at approximately 263,126,000 people, correlating with the highest attendance recorded at 3,587,538. Conversely, Chile, which hosted the tournament in 1962 with the smallest host population of 8,486,120, saw the lowest attendance, totaling only 893,172. This trend suggests that more populous nations are better able to attract larger crowds.

## Visualization 2



The relationship between a country's economic power, measured by GDP value, and its World Cup participation history, specifically the total number of matches played, was analyzed through a scatter plot visualization.

The results clearly indicate a positive correlation between GDP and World Cup activity, showing that countries with higher GDPs have participated in more matches over the tournament's history. For instance, economically robust nations such as Argentina, England, Mexico, and Korea Republic, each with a GDP exceeding 19.5, have all played more than 500 matches. In contrast, countries with lower economic power like China PR, Kuwait, Egypt, Serbia, Bolivia, and Korea DPR, all of whom have GDP values below 5, have participated in fewer than 100 matches. This pattern underscores the influence of economic resources on a nation's ability to support and sustain competitive international football programs, suggesting that higher GDP provides the infrastructure and training opportunities necessary for frequent and successful World Cup participation.

## Visualization 3

Host countries and their Income Group	
Argentina	Upper middle income
Brazil	Upper middle income
Chile	High income
England	High income
France	High income
Germany	High income
Italy	High income
Korea/Japan	High income
Mexico	Upper middle income
South Africa	Upper middle income
Spain	High income
USA	High income

The research question aimed to identify the income groups of World Cup host countries and explore how these economic classifications might relate to their capabilities and effectiveness in hosting the tournament.

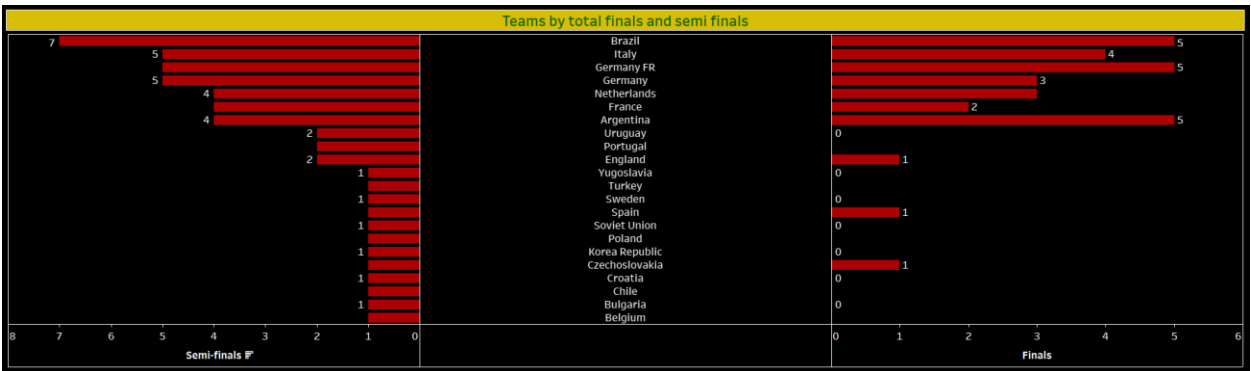
To analyze this, a text table format was employed, which revealed a notable trend: all the host countries fell within the ‘upper-middle income’ and ‘high income’ brackets. This observation suggests a possible correlation between economic prosperity and the ability to host large-scale international events like the FIFA World Cup. Countries within these higher income groups are likely to have more developed infrastructure, better facilities, and greater resources, all of which are critical for efficiently managing an event of such magnitude and complexity. This trend highlights the significant role that economic status plays not only in a country's ability to bid for and secure hosting rights but also in successfully delivering a global event that meets international standards.

Visualization 4

Countries that won the World Cup							
Brazil	Italy	Germany FR	Argentina	Spain	Germany	France	England
4	2	2	2	1	1	1	1

The analysis of World Cup victories from 1962 to 2014 highlights the dominance of certain countries in international football. Brazil emerges as the most successful nation during this period, capturing the World Cup trophy four times, which underscores its longstanding prowess in the sport. Italy, Germany FR (Federal Republic of Germany before reunification in 1990), and Argentina each secured the title twice, indicating their strong football traditions and consistent performance at the highest level. Meanwhile, Spain, Germany (after reunification), France, and England each won the tournament once.

Visualization 5



The visualization employing a butterfly chart provides a clear and comparative overview of the total number of finals and semi-finals appearances made by various teams from 1962 to 2014. Notably, Brazil and Germany FR lead with five finals appearances each, underscoring their

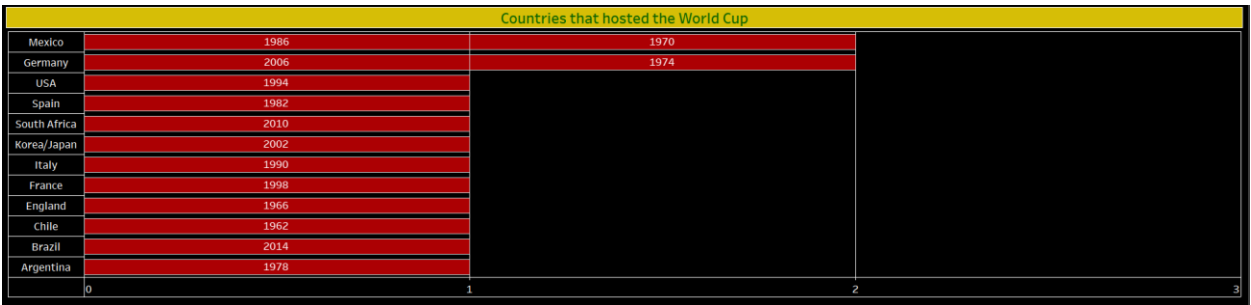


historical dominance in the World Cup arena. Brazil further stands out with a total of seven semi-finals appearances, the highest among all teams. Italy follows closely with impressive records of four finals and five semi-finals appearances, reflecting their consistent high-level performance.

Other significant contenders include Germany with three finals and five semi-finals appearances and the Netherlands with three finals and four semi-finals appearances, indicating their strong competitive presence throughout multiple tournaments. France and Argentina also show notable success, with France making it to the finals twice and semi-finals four times, and Argentina reaching the finals five times and semi-finals four times.

This distribution provides insight into the elite footballing nations that have consistently excelled on the world stage, managing to reach the later stages of the tournament multiple times. Teams like Belgium, Bulgaria, Chile, and others, despite fewer appearances in the semi-finals or finals, reflect the broader competitive nature of the tournament.

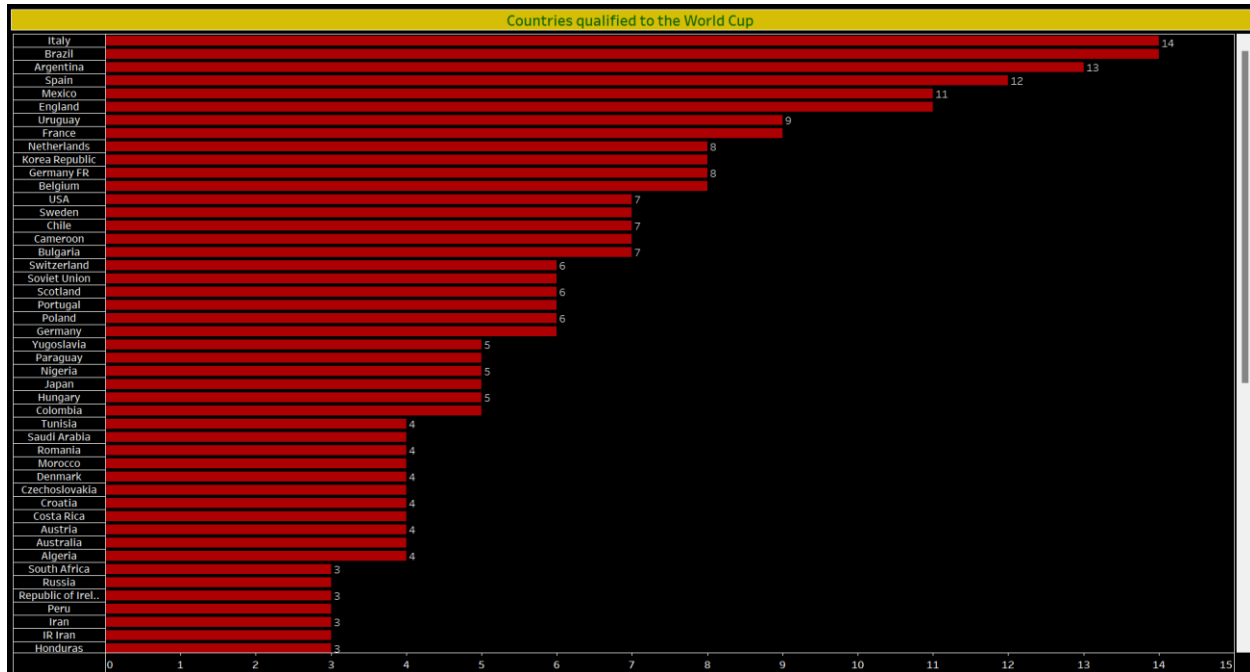
Visualization 6



The analysis of World Cup hosting nations from 1962 to 2014 is illustrated through a horizontal bar chart, which effectively outlines the frequency and specific years in which different countries hosted the tournament. Notably, Germany and Mexico stand out as each has hosted the World Cup twice—Germany in 1974 and again in 2006, and Mexico in 1970 and 1986. This indicates their significant roles in World Cup history, not only in terms of competitive play but also in their capabilities to organize such a large-scale international event.

Other countries such as Argentina, Brazil, Chile, England, France, Italy, Korea/Japan (joint hosts), South Africa, Spain, and the USA have each hosted the tournament once.

## Visualization 7

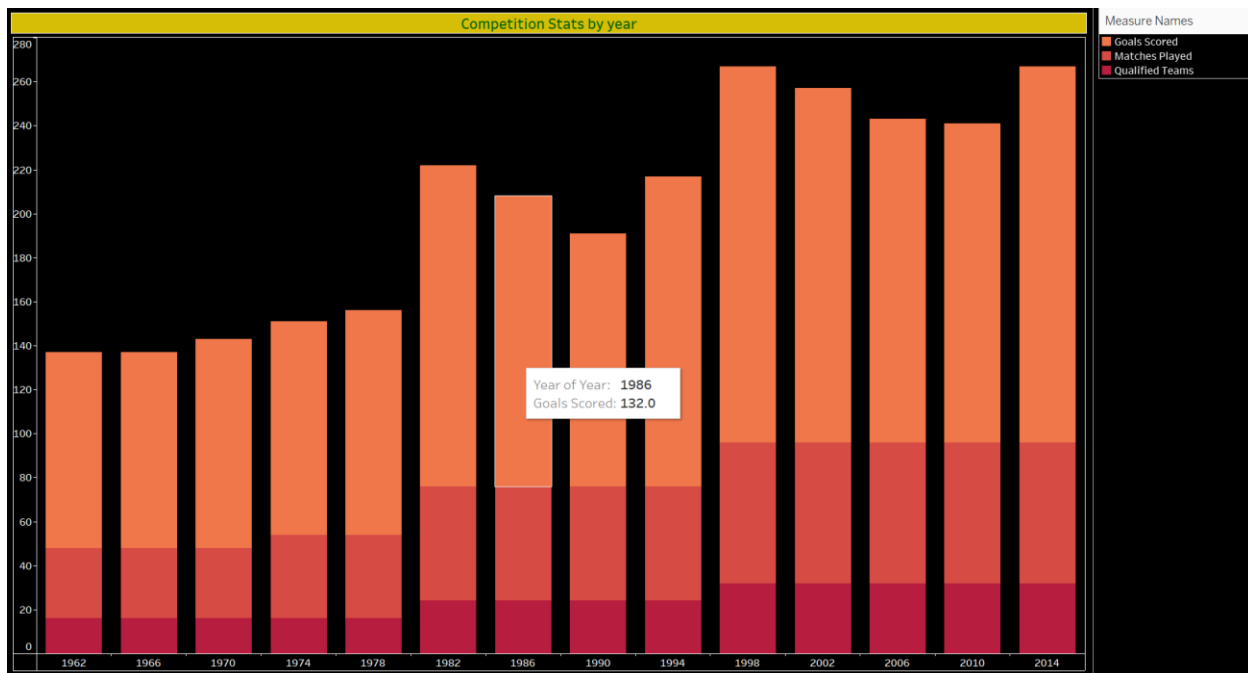


The analysis of World Cup qualifications, presented through a horizontal bar chart, offers a detailed look at the number of times various countries have qualified for the tournament from 1962 to 2014. The visualization highlights the broad spectrum of global football participation, with a significant disparity in the frequency of qualifications among nations.

Countries such as Brazil and Italy top the chart with 14 qualifications each, underscoring their dominant presence and consistent performance in international football. Following closely are Argentina and Spain, with 13 and 12 qualifications respectively, which further emphasizes the strong football traditions and competitive continuity within these nations.

On the other end of the spectrum, numerous countries such as Angola, Bolivia, Bosnia and Herzegovina, and several others have qualified only once, illustrating the challenges and competitive hurdles faced by smaller or less historically prominent footballing nations.

## Visualization 8

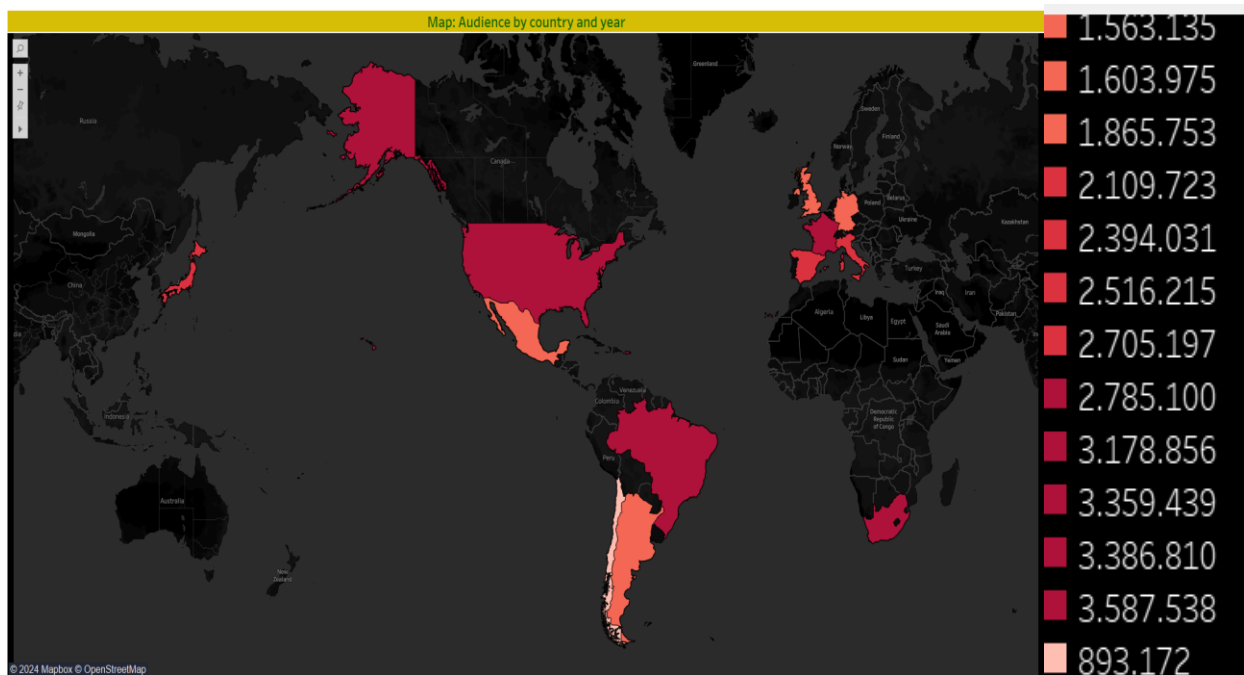


The visualization of competition statistics by year, presented through a stacked bar chart, effectively illustrates the evolution of the FIFA World Cup from 1962 to 2014. Key metrics such as goals scored, matches played, and the number of qualified teams are tracked across different tournaments to highlight trends and changes in the competition's structure.

The data shows a steady increase in the number of goals scored, matches played, and teams qualified over the years. For instance, from 1962 to 1978, the number of goals gradually increased from 89 to 102, and the matches played were consistent at 32 for each tournament until 1974, when they rose to 38. A significant change occurred in 1982, when the matches jumped to 52 and the teams increased from 16 to 24, reflecting an expansion in the tournament's format. This trend continued, with the number of goals reaching a peak of 171 in 1998 and again in 2014, coinciding with an increase in matches played to 64 and teams qualified to 32 starting in 1998.

This progression not only reflects the growing popularity and scale of the World Cup but also indicates the increasing competitiveness and global reach of the event, allowing more nations to participate and a greater number of matches to be played, ultimately leading to more goals being scored.

## Visualization 9



The analysis of World Cup audience sizes by country and year, visualized using a geographical map, effectively highlights the variations in attendance across different host nations and tournaments from 1962 to 2014. The map visualization provides a clear representation of how audience sizes have increased significantly over the years, reflecting both the growing global interest in football and the expanding capacity of host nations to accommodate larger numbers of spectators.

For instance, Chile hosted the World Cup in 1962 with the lowest recorded attendance of 893,172. Over the decades, there has been a substantial increase in attendance figures, culminating in the United States hosting in 1994 with a record high of 3,587,538 spectators. This trend is indicative of the increasing popularity of the World Cup and the improvements in infrastructure and facilities that have enabled host countries to manage larger crowds. Notably, countries like Germany, Brazil, and South Africa also show significant audience sizes in their respective hosting years, highlighting their effectiveness in organizing such a massive international event.

## 4. Conclusion

### Key Findings

- Population Impact on Attendance: The analysis confirmed a positive correlation between the population size of the host country and the attendance of the World Cup tournaments.

Countries with larger populations, such as the USA in 1994, demonstrated significantly higher attendance figures. This suggests that more populous nations have a greater capacity to draw larger crowds, likely due to a more extensive domestic fan base and the ability to support larger visitor inflows.

- **Economic Influence on Participation:** The relationship between a country's economic power (GDP) and its history of participation in the World Cup revealed that wealthier nations tend to participate more frequently and successfully. Economically stronger countries like Argentina and Germany have consistently had the resources to develop robust football programs, contributing to their frequent appearances and successes in the tournament.
- **Socio-Economic Status of Hosts:** All World Cup host countries fell into the upper-middle income or high-income brackets. This indicates that certain economic thresholds are likely necessary to meet the infrastructural and organizational demands of hosting such a large-scale international event. The economic capacity of these countries likely provides the necessary resources for successful event hosting, including stadium maintenance, security, and hospitality services.

## **Implications and Future Research**

The findings suggest that socio-economic factors play a significant role in the organization and impact of hosting the World Cup. Future research could explore:

1. How does the economic growth rate in the years leading up to the World Cup affect the infrastructural readiness of the host nations?
2. What are the long-term impacts on tourism and international perception of host countries post the World Cup?
3. **Impact of Digital Engagement:** Future studies could explore how digital engagement and media coverage influence attendance and viewership, especially in less populous countries.

In conclusion, the World Cup is not only a global sporting event but also a mirror reflecting the socio-economic capabilities of its hosts and participants. The patterns observed in this study underscore the significant role of economic and demographic factors in shaping the operational and competitive landscapes of the World Cup.