

A Visual Data Analysis

# FIFA World Cup Analysis

Presented by Aditya Singh

# Introduction

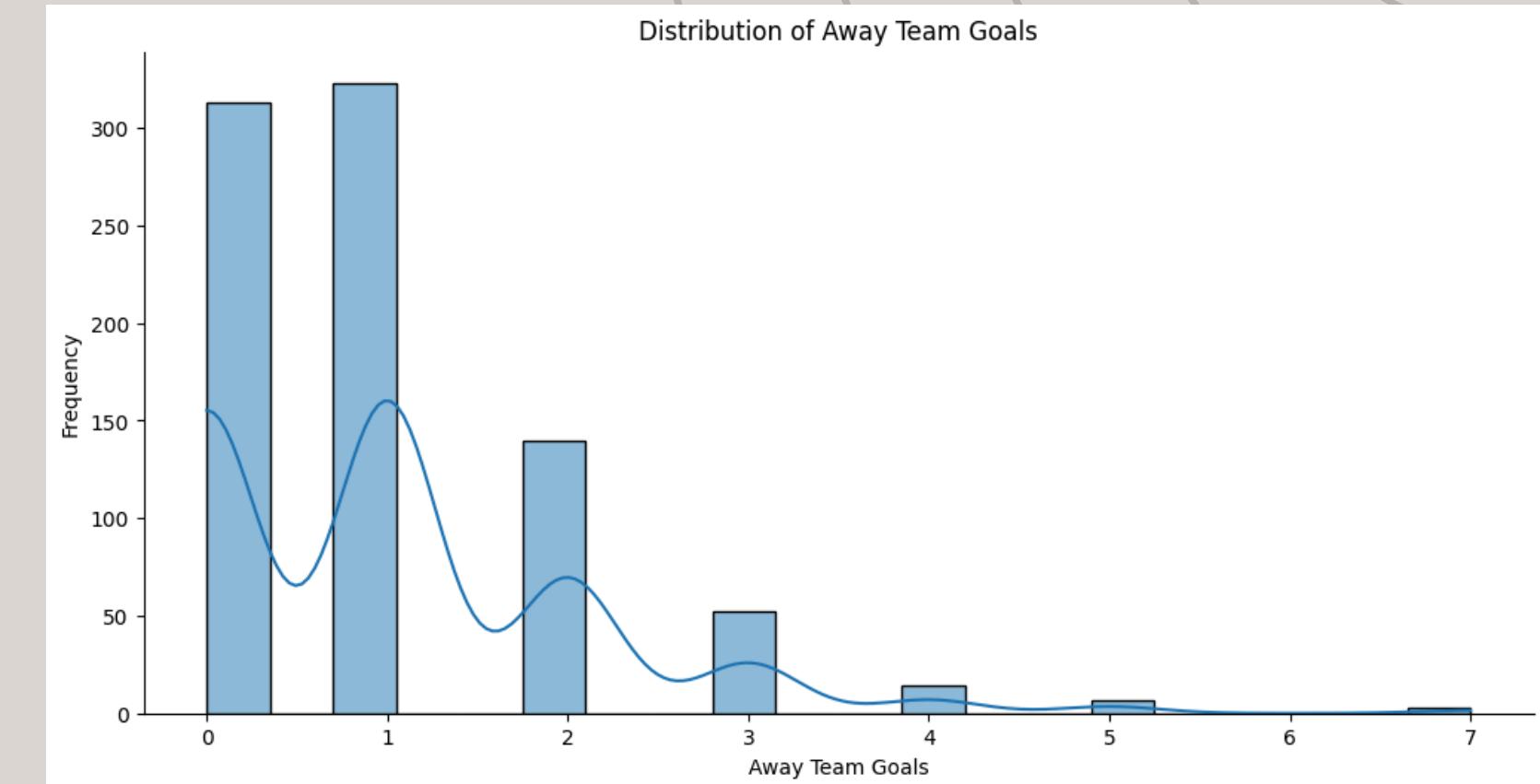
Welcome to our presentation on the analysis of historical football match data, encompassing 852 matches with detailed information across 20 columns. This dataset includes precise temporal details (Year and Datetime), contextual information (Stage, Stadium, and City), team performance metrics (Home and Away Team Goals), win conditions, and attendance figures. Additionally, it provides half-time scores, names of match officials, unique match identifiers, and team initials for quick reference. Through this rich dataset, we aim to uncover patterns and insights that illuminate the dynamics of football matches across different years and competition stages. Let's dive into the world of football data analytics to explore these intriguing aspects.

# Important KPIs

1. Total number of matches
2. Average goals per match (home and away)
3. Highest and lowest attendance
4. Average attendance per match
5. Number of matches with zero goals
6. Distribution of goals scored (both home and away)
7. Year-over-year change in average goals
8. Year-over-year change in average attendance
9. Win rate for home teams
10. Frequency of matches decided by win conditions
11. Goals scored in the first half vs. second half
12. Distribution of match stages (e.g., group stage, knockout stage)
13. Most frequent referees and assistant referees
14. Number of matches per stadium and city
15. Average goals per team (home and away teams separately)

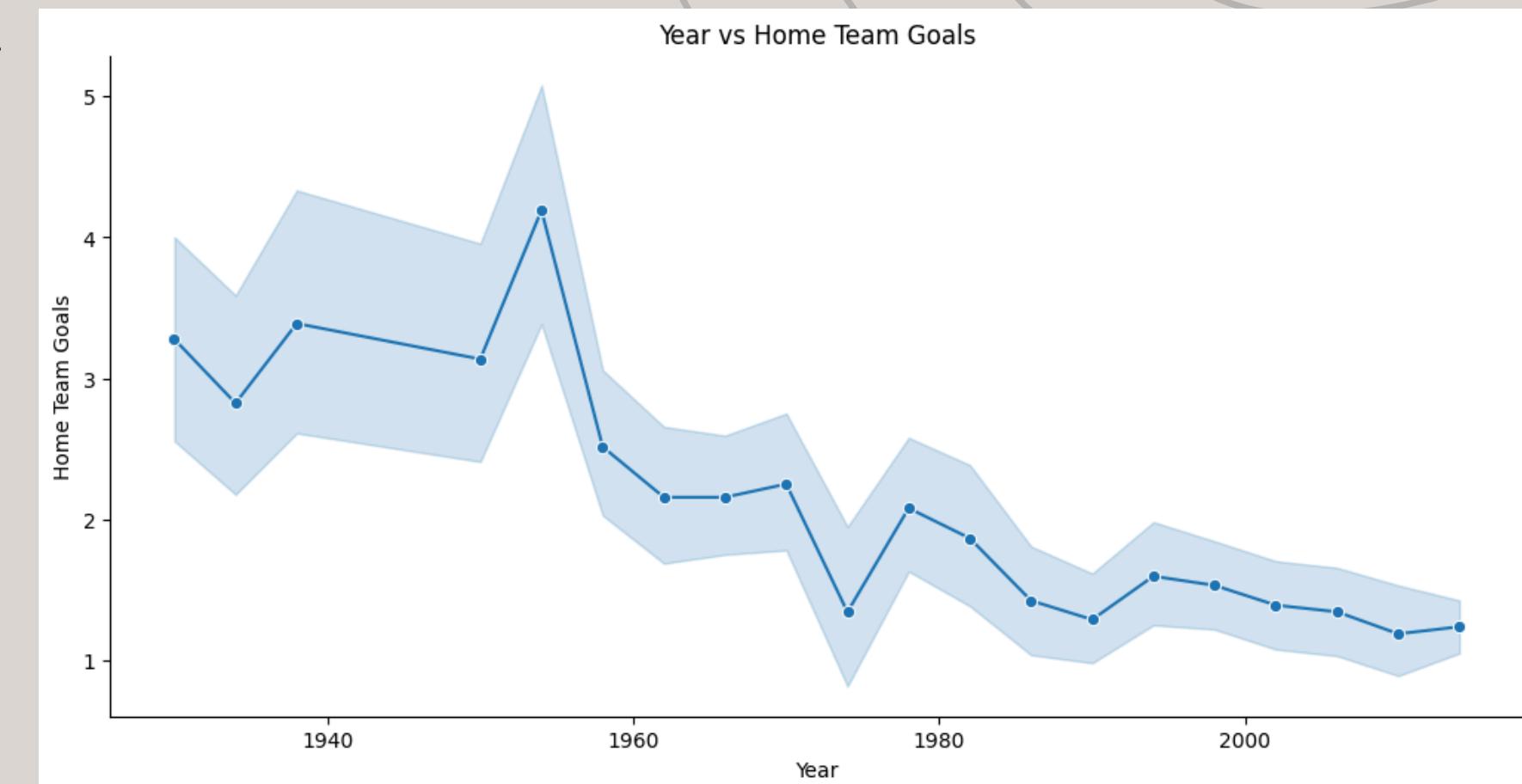
# Frequency of Away Team Goals

The graph displays the distribution of goals scored by away teams in 852 football matches. The x-axis represents the number of goals, ranging from 0 to 7, while the y-axis shows the frequency of matches with those goal counts. Most away teams scored either 0 or 1 goal, with over 300 occurrences each. The frequency significantly drops as the number of goals increases, indicating that higher goal counts are less common. The overlaid line adds a density estimate, highlighting the distribution's peaks and valleys.



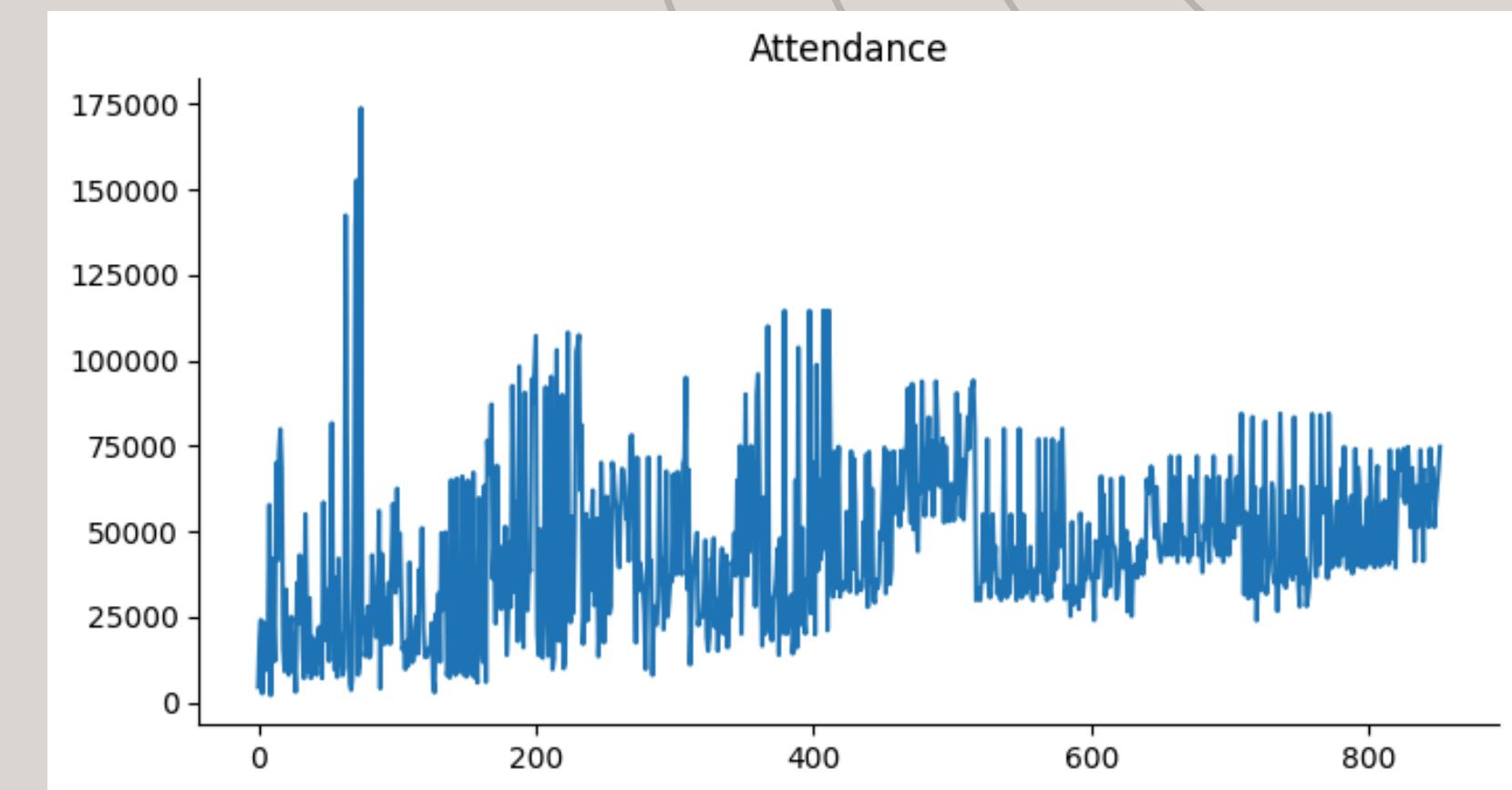
# Year vs Home Team Goals

The graph illustrates the trend of home team goals scored across different years. The x-axis represents the years, spanning from the 1930s to the 2010s, while the y-axis shows the average number of goals scored by home teams. Initially, there is some fluctuation around an average of 3 to 4 goals. A significant peak is observed around 1960, followed by a sharp decline. Post-1960, there's a general downward trend, with minor fluctuations, indicating a decrease in the average goals scored by home teams over time. The shaded area represents the variability in the data.

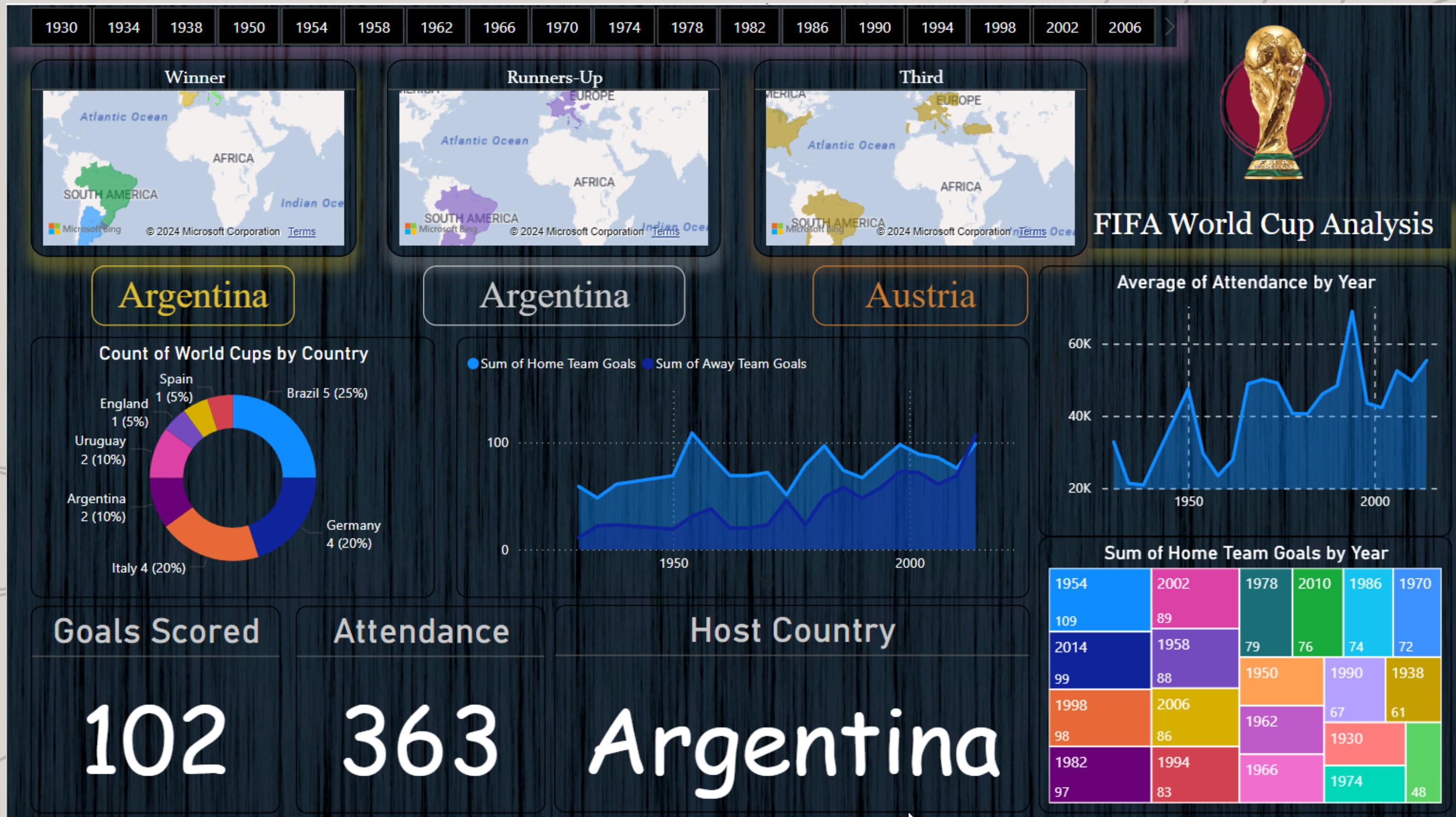


# Attendance Count

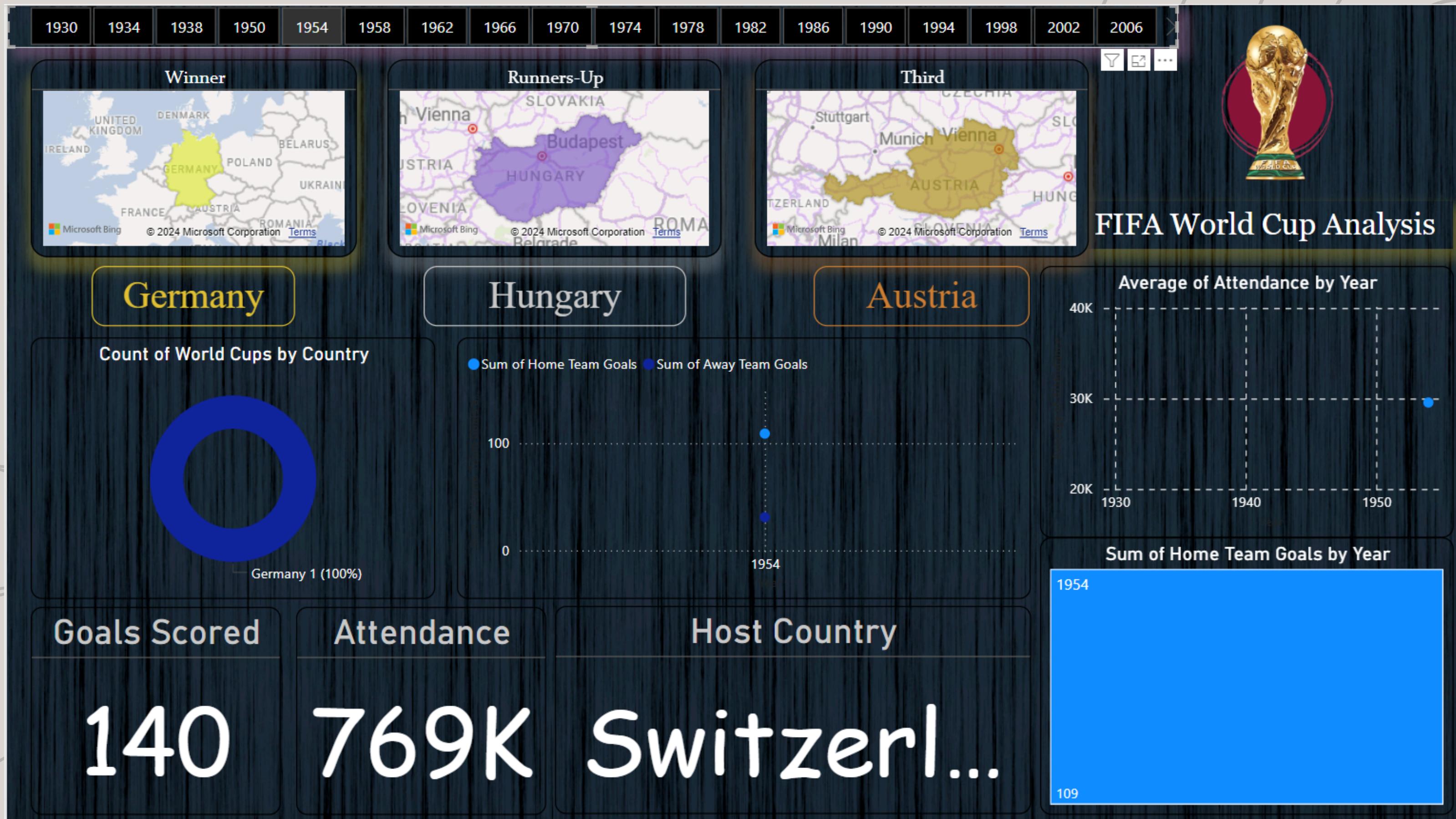
The graph displays the attendance figures for 852 football matches, with the x-axis representing the match index and the y-axis showing the number of attendees. There is a notable peak early on, with attendance surpassing 175,000, followed by high variability. The data shows several clusters of high attendance, particularly around the 200th and 400th matches, where figures frequently exceed 100,000. Over time, there is a slight decline in maximum attendance, stabilizing around 50,000 to 75,000 attendees towards the later matches.



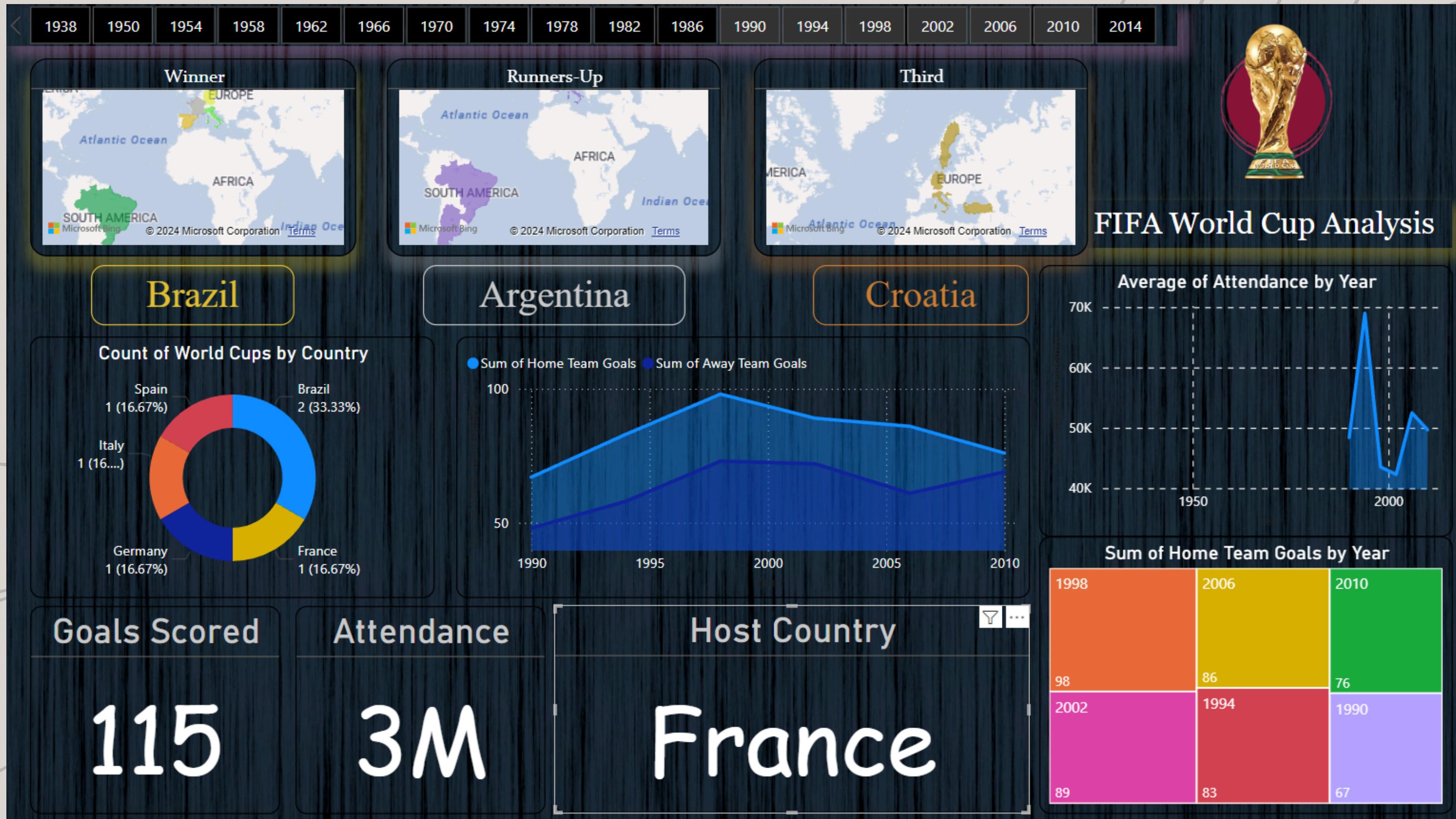
# Visual Dashboard (1)



## Visual Dashboard (2)



# Visual Dashboard (3)



# Thank You

Aditya Singh  
[csaditya44@gmail.com](mailto:csaditya44@gmail.com)  
(+91) 7379356275