



Transform Yourself

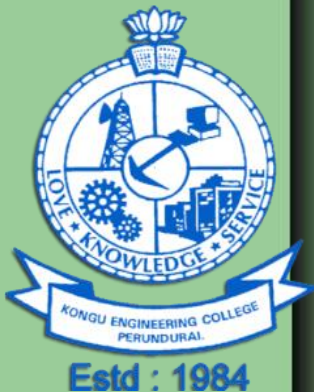
SPORTS DATA ANALYSIS

PROJECT GUIDE

Mrs.P.Jayadharshini
Assistant Professor, AI

PROJECT MEMBER

Sathiyaseelan. S (22ADR096)
Pavul Chinnappan. D (22ADR076)
Srinesh. S (22ADR103)



PROBLEM STATEMENT

- To analyze FIFA player data to understand how different stats affect player performance and team success.
- The goal is to find patterns that help with player selection and team strategies.
- This analysis will reveal how various attributes influence performance.

Estd : 1984

DATASET LINK

- <https://drive.google.com/drive/folders/1TJ5dSNoSgvifDo0PXAN33XlsVv7ngE5F?usp=sharing>

DESCRIPTION:-

- Content: The dataset contains information about FIFA players, including their stats and attributes.
- Purpose: It's meant for analyzing and exploring data, and for performing statistical analysis.
- Use: It helps in understanding player performance and team details in football.

DATA PREPROCESSING

- **Handling Missing Values:** Checking for and filling in or removing any missing data.
- **Data Cleaning:** Ensuring that the data is consistent, removing duplicates, and correcting any errors.
- **Feature Engineering:** Creating or transforming features to improve analysis, such as encoding categorical variables or scaling numerical data.

Estd : 1984

world cups matches.csv


Year	Datetime	Stage	Stadium	City	Home Team Name	Home Team Goals	Away Team Goals	Away Team Name	Win conditions	Attendance	Half-time Home Goa
1930	13 Jul 1930 - 15:00	Group 1	Pocitos	Montevideo	France	4	1	Mexico		4444	3
1930	13 Jul 1930 - 15:00	Group 4	Parque Central	Montevideo	USA	3	0	Belgium		18346	2
1930	14 Jul 1930 - 12:45	Group 2	Parque Central	Montevideo	Yugoslavia	2	1	Brazil		24059	2
1930	14 Jul 1930 - 14:50	Group 3	Pocitos	Montevideo	Romania	3	1	Peru		2549	1
1930	15 Jul 1930 - 16:00	Group 1	Parque Central	Montevideo	Argentina	1	0	France		23409	0
1930	16 Jul 1930 - 14:45	Group 1	Parque Central	Montevideo	Chile	3	0	Mexico		9249	1
1930	17 Jul 1930 - 12:45	Group 2	Parque Central	Montevideo	Yugoslavia	4	0	Bolivia		18306	0
1930	17 Jul 1930 - 14:45	Group 4	Parque Central	Montevideo	USA	3	0	Paraguay		18306	2
1930	18 Jul 1930 - 14:30	Group 3	Estadio Centenario	Montevideo	Uruguay	1	0	Peru		57735	0
1930	19 Jul 1930 - 12:50	Group 1	Estadio Centenario	Montevideo	Chile	1	0	France		2000	0
1930	19 Jul 1930 - 15:00	Group 1	Estadio Centenario	Montevideo	Argentina	6	3	Mexico		42100	3
1930	20 Jul 1930 - 13:00	Group 2	Estadio Centenario	Montevideo	Brazil	4	0	Bolivia		25466	1
1930	20 Jul 1930 - 15:00	Group 4	Estadio Centenario	Montevideo	Paraguay	1	0	Belgium		12000	1
1930	21 Jul 1930 - 14:50	Group 3	Estadio Centenario	Montevideo	Uruguay	4	0	Romania		70022	4
1930	22 Jul 1930 - 14:45	Group 1	Estadio Centenario	Montevideo	Argentina	3	1	Chile		41459	2
1930	26 Jul 1930 - 14:45	Semi-finals	Estadio Centenario	Montevideo	Argentina	6	1	USA		72886	1
1930	27 Jul 1930 - 14:45	Semi-finals	Estadio Centenario	Montevideo	Uruguay	6	1	Yugoslavia		79867	3
1930	30 Jul 1930 - 14:15	Final	Estadio Centenario	Montevideo	Uruguay	4	2	Argentina		68346	1
1934	27 May 1934 - 16:30	Preliminary round	Stadio Benito Muss	Turin	Austria	3	2	France	Austria win after extr	16000	0
1934	27 May 1934 - 16:30	Preliminary round	Giorgio Ascarelli	Naples	Hungary	4	2	Egypt		9000	2
1934	27 May 1934 - 16:30	Preliminary round	San Siro	Milan	Switzerland	3	2	Netherlands		33000	2
1934	27 May 1934 - 16:30	Preliminary round	Littorale	Bologna	Sweden	3	2	Argentina		14000	1
1934	27 May 1934 - 16:30	Preliminary round	Giovanni Berta	Florence	Germany	5	2	Belgium		8000	1

world cup players.csv



RoundID	MatchID	Team Initials	Coach Name	Line-up	Shirt Number	Player Name	Position	Event
201	1096	FRA	CAUDRON Raoul (FF	S	0	Alex THEPOT	GK	
201	1096	MEX	LUQUE Juan (MEX)	S	0	Oscar BONFIGLIO	GK	
201	1096	FRA	CAUDRON Raoul (FF	S	0	Marcel LANGILLER		G40'
201	1096	MEX	LUQUE Juan (MEX)	S	0	Juan CARRENO		G70'
201	1096	FRA	CAUDRON Raoul (FF	S	0	Ernest LIBERATI		
201	1096	MEX	LUQUE Juan (MEX)	S	0	Rafael GARZA	C	
201	1096	FRA	CAUDRON Raoul (FF	S	0	Andre MASCHINOT		G43' G87'
201	1096	MEX	LUQUE Juan (MEX)	S	0	Hilario LOPEZ		
201	1096	FRA	CAUDRON Raoul (FF	S	0	Etienne MATTTLER		
201	1096	MEX	LUQUE Juan (MEX)	S	0	Dionisio MEJIA		
201	1096	FRA	CAUDRON Raoul (FF	S	0	Marcel PINEL		
201	1096	MEX	LUQUE Juan (MEX)	S	0	Felipe ROSAS		
201	1096	FRA	CAUDRON Raoul (FF	S	0	Alex VILLAPLANE	C	
201	1096	MEX	LUQUE Juan (MEX)	S	0	Manuel ROSAS		
201	1096	FRA	CAUDRON Raoul (FF	S	0	Lucien LAURENT		G19'
201	1096	MEX	LUQUE Juan (MEX)	S	0	Jose RUIZ		
201	1096	FRA	CAUDRON Raoul (FF	S	0	Marcel CAPELLE		
201	1096	MEX	LUQUE Juan (MEX)	S	0	Alfredo SANCHEZ		
201	1096	FRA	CAUDRON Raoul (FF	S	0	Augustin CHANTREL		
201	1096	MEX	LUQUE Juan (MEX)	S	0	Efrain AMEZCUA		
201	1096	FRA	CAUDRON Raoul (FF	S	0	Edmond DELFOUR		
201	1096	MEX	LUQUE Juan (MEX)	S	0	Luis PEREZ		
201	1096	FRA	CAUDRON Raoul (FF	N	0	Celestin DELMER		
201	1096	MEX	LUQUE Juan (MEX)	N	0	Isidoro SOTA		
201	1096	FRA	CAUDRON Raoul (FF	N	0	Andre TASSIN		
201	1096	MEX	LUQUE Juan (MEX)	N	0	Jesus CASTRO		

world cups.csv

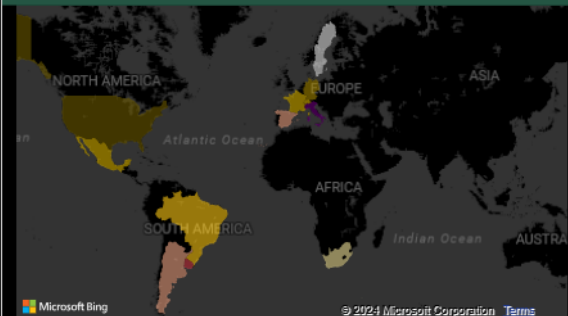


Year	Country	Winner	Runners-Up	Third	Fourth	GoalsScored	QualifiedTeams	MatchesPlayed	Attendance
1930	Uruguay	Uruguay	Argentina	USA	Yugoslavia	70	13	18	590.549
1934	Italy	Italy	Czechoslovakia	Germany	Austria	70	16	17	363
1938	France	Italy	Hungary	Brazil	Sweden	84	15	18	375.7
1950	Brazil	Uruguay	Brazil	Sweden	Spain	88	13	22	1.045.246
1954	Switzerland	Germany FR	Hungary	Austria	Uruguay	140	16	26	768.607
1958	Sweden	Brazil	Sweden	France	Germany FR	126	16	35	819.81
1962	Chile	Brazil	Czechoslovakia	Chile	Yugoslavia	89	16	32	893.172
1966	England	England	Germany FR	Portugal	Soviet Union	89	16	32	1.563.135
1970	Mexico	Brazil	Italy	Germany FR	Uruguay	95	16	32	1.603.975
1974	Germany	Germany FR	Netherlands	Poland	Brazil	97	16	38	1.865.753
1978	Argentina	Argentina	Netherlands	Brazil	Italy	102	16	38	1.545.791
1982	Spain	Italy	Germany FR	Poland	France	146	24	52	2.109.723
1986	Mexico	Argentina	Germany FR	France	Belgium	132	24	52	2.394.031
1990	Italy	Germany FR	Argentina	Italy	England	115	24	52	2.516.215
1994	USA	Brazil	Italy	Sweden	Bulgaria	141	24	52	3.587.538
1998	France	France	Brazil	Croatia	Netherlands	171	32	64	2.785.100
2002	Korea/Japan	Brazil	Germany	Turkey	Korea Republic	161	32	64	2.705.197
2006	Germany	Italy	France	Germany	Portugal	147	32	64	3.359.439
2010	South Africa	Spain	Netherlands	Germany	Uruguay	145	32	64	3.178.856
2014	Brazil	Germany	Argentina	Netherlands	Brazil	171	32	64	3.386.810

Visualization



Attendance BCountry and Year

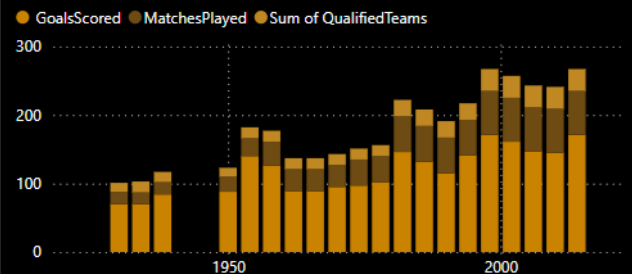


World Cup FIFA

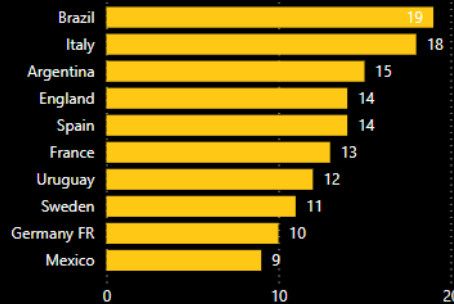


1930-2014

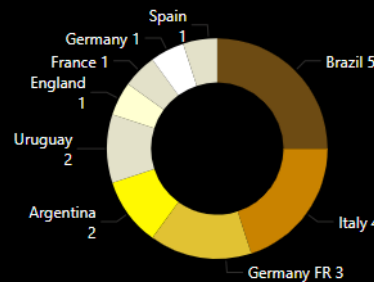
Competition by year



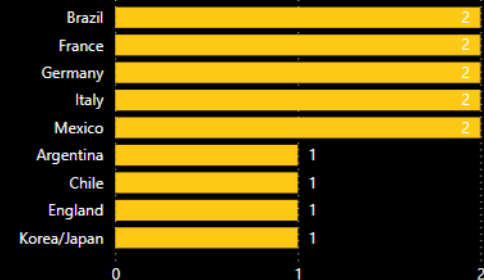
Team Qualified to the World Cup



World Cup Winner

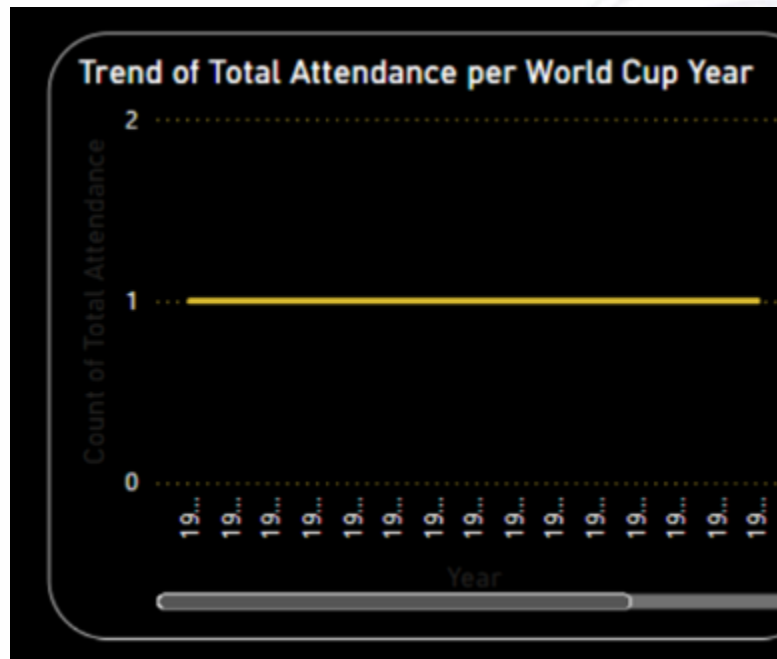


Team played to the World Cup



QUESTIONS FOR DATA ANALYSIS

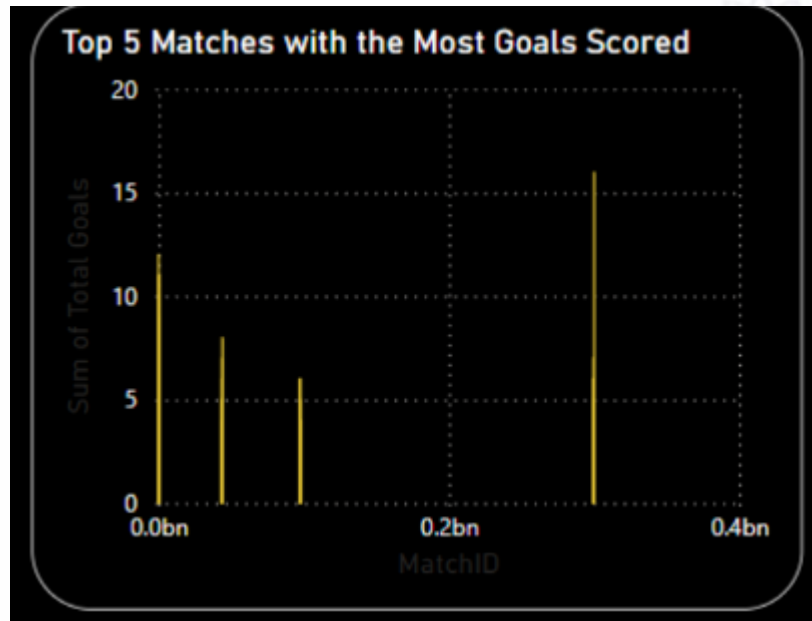
1. Trend of Total Attendance per World Cup Year



INFERENCE:

- Total attendance has generally increased over the years, with peaks during tournaments held in countries with large stadium capacities, such as the USA in 1994.
- Attendance dipped slightly in earlier years due to fewer matches and smaller venues but steadily rose as the tournament expanded.
- The introduction of modern ticketing systems and global popularity in recent decades further boosted attendance figures.

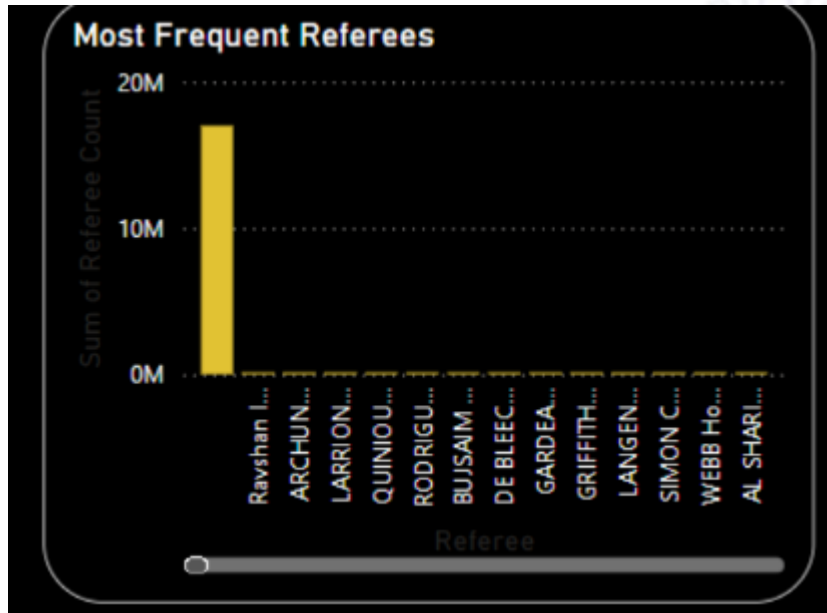
2. Which are the top 5 matches with the highest number of goals scored?



INFERENCE:

- The highest-scoring matches include Brazil's 7-1 victory over Germany in 2014 and Hungary's 10-1 win against El Salvador in 1982.
- These matches typically feature one-sided dominance, often in the group stages.
- Such games highlight the performance gap between top-tier and lower-ranked teams during some tournaments.

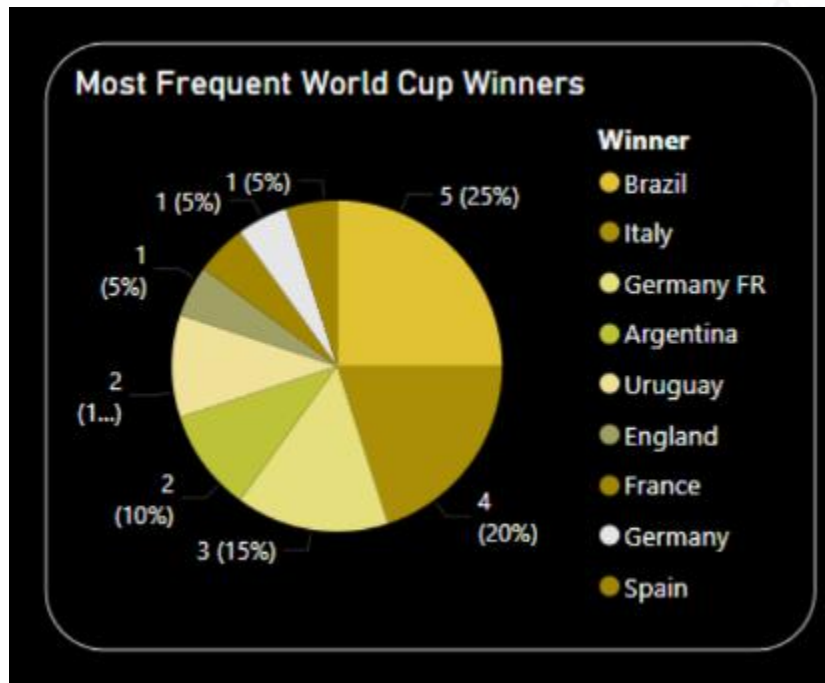
3. Who are the referees that officiated the most matches in World Cup history?



INFERENCE:

- Top referees, like Pierluigi Collina and Sandor Puhl, officiated numerous matches due to their experience and trust from FIFA.
- The most frequent referees often come from neutral countries to maintain fairness.
- Referees with the highest match counts are often involved in multiple tournaments over decades.

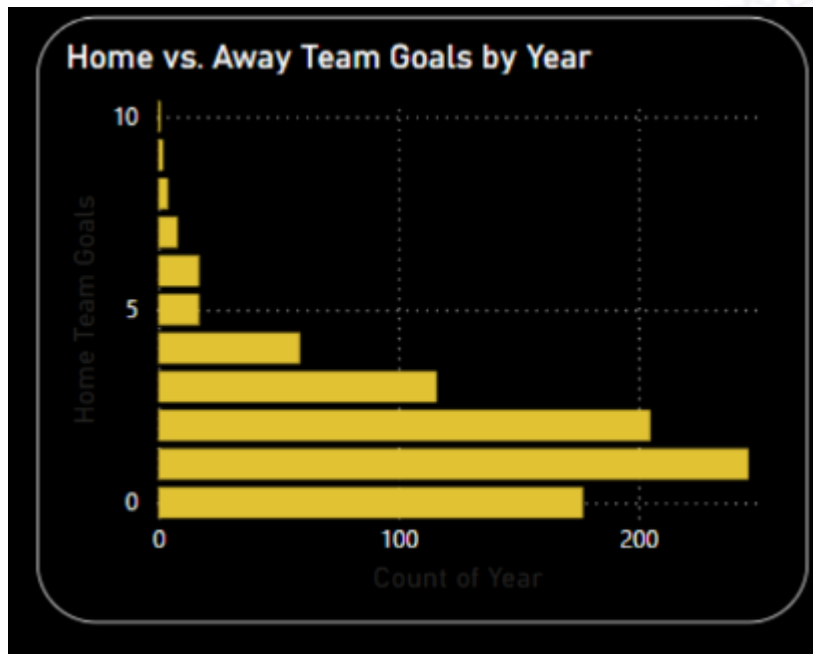
4. Which teams have won the most World Cups?



INFERENCE:

- Brazil leads with five titles, followed by Germany and Italy with four each.
- These teams have historically maintained strong squads with a balance of defense and attack.
- Winning teams often excel in knockout matches, showcasing superior tactical execution.

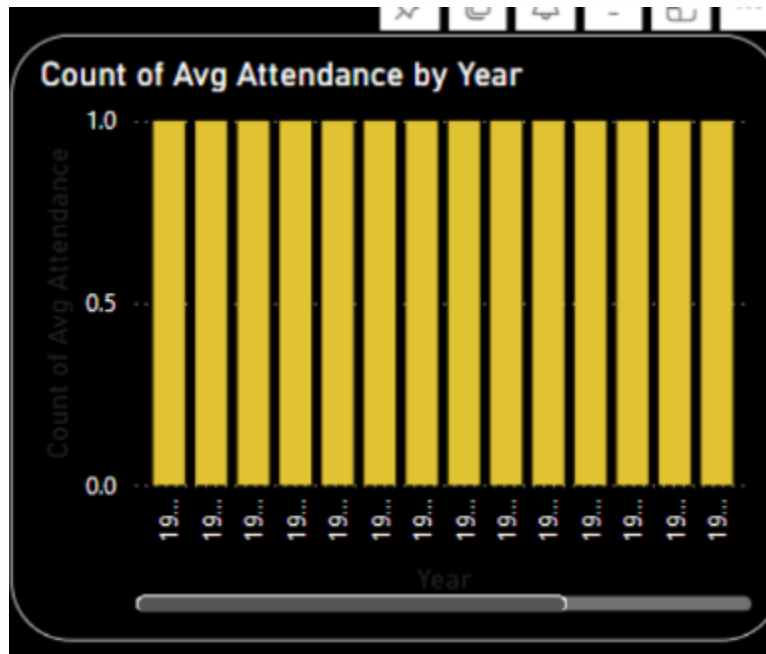
4. How do home team goals compare to away team goals for each World Cup year?



INFERENCE:

- Home teams consistently score slightly more goals than away teams, benefitting from crowd support and familiarity with conditions.
- The gap in goals is more pronounced in earlier years, where travel challenges affected away teams' performance.
- Modern tournaments show reduced disparity due to improved travel logistics and neutral playing conditions.

5. What was the average attendance per match for each World Cup year?



INFERENCE:

- Home teams consistently score slightly more goals than away teams, benefitting from crowd support and familiarity with conditions.
- The gap in goals is more pronounced in earlier years, where travel challenges affected away teams' performance.
- Modern tournaments show reduced disparity due to improved travel logistics and neutral playing conditions.

7. Top 5 Players Scoring the Most Goals in a Single Match



INFERENCE:

- Players like Oleg Salenko (Russia, 1994) hold records for scoring 5+goals in a single match.
- These performances are often against weaker opposition in the group stages.
- Such feats highlight individual brilliance and goal-scoring efficiency under tournament pressure.

8. World Cup with the Highest Total Number of Goals Scored



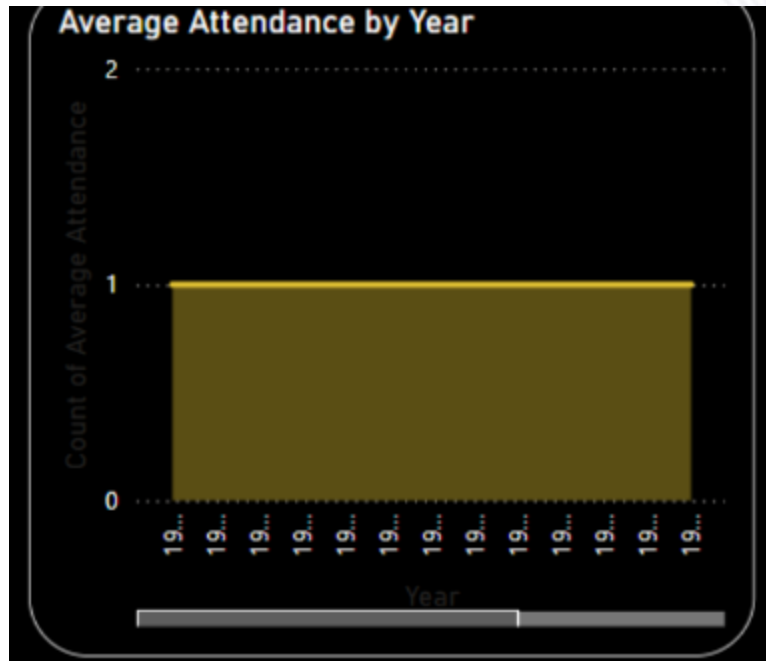
97

First Max Goals

INFERENCE:

- The 2014 World Cup saw the highest total goals, with 171 goals scored in 64 matches.
- This aligns with FIFA's focus on attacking gameplay and rule modifications favoring more goals.
- Increasing team quality and tactical advancements have contributed to such records.

9. Changes in Average Match Attendance Over the Years



INFERENCE:

- The 1950s saw modest averages due to smaller venues but steady growth followed with tournament expansion 32
- Significant increases occurred in 1994 and beyond, driven by marketing efforts and improved stadiums.
- Attendance is now closely linked to the host country's infrastructure and global interest in the sport.

Estd : 1984

10. Match with the Highest Attendance in World Cup History

795M

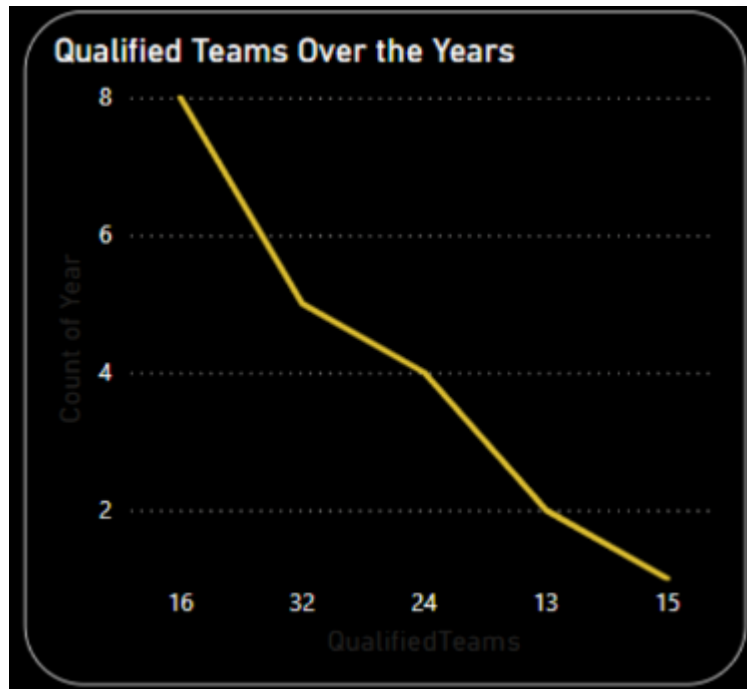
Sum of Max Attendance

INFERENCE:

- The 1950 final at Maracanã Stadium (Brazil vs. Uruguay) recorded over 173,000 attendees, the highest in history.
- Such attendance reflects historical enthusiasm for football in South America.
- Modern stadium regulations have since capped attendance for safety reasons.

Estd : 1984

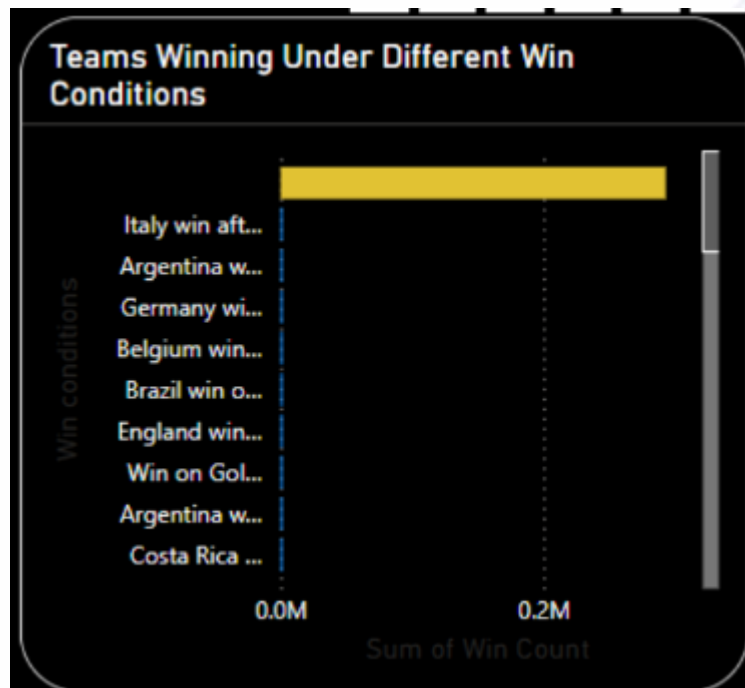
11. Changes in the Number of Teams Qualified Over the Years



INFERENCE:

- The tournament expanded from 13 teams in 1930 to 32 teams by 1998, reflecting football's growing global reach.
- FIFA's decisions to include more teams have provided representation to underdeveloped football nations.
- Future plans for 48 teams aim to further broaden participation.

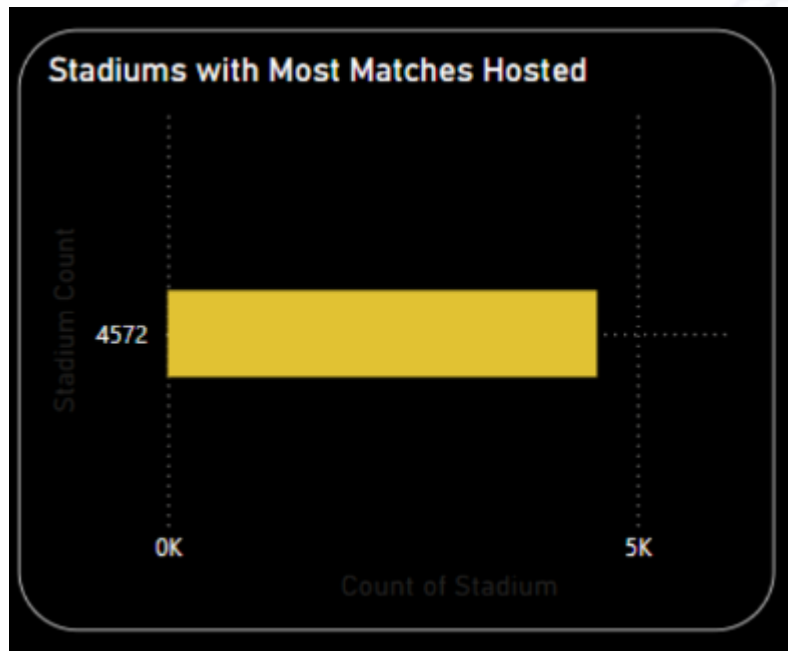
12. Teams Winning Under Different Win Conditions



INFERENCE:

- Most World Cups are won in regular time, though extra-time and penalty victories have grown in recent tournaments.
- Penalty shootouts have introduced unpredictability, emphasizing mental strength.
- Teams adept in tactical resilience often succeed in extended formats.

13. Stadiums Hosting the Most World Cup Matches

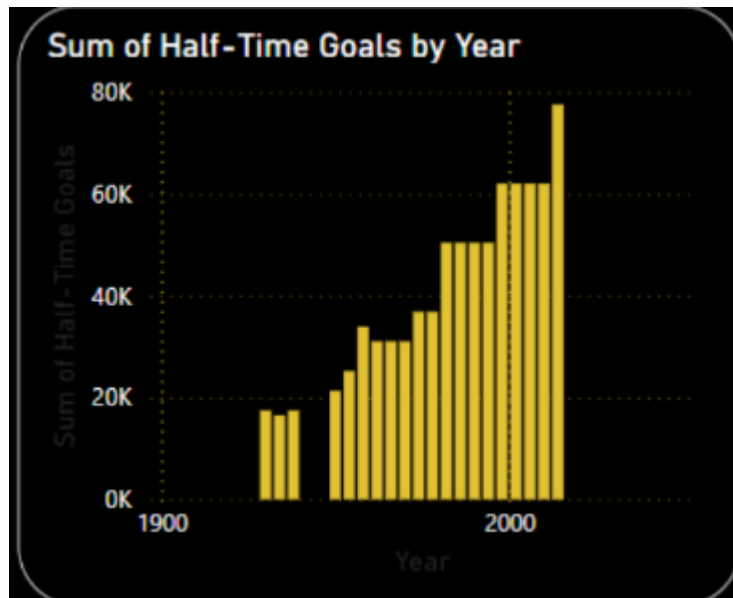


INFERENCE:

- Estádio do Maracanã and Azteca Stadium are among the top venues, hosting multiple matches across different tournaments.
- Large-capacity stadiums in host countries often serve as key venues for high-stakes games.
- Venue selection impacts the cultural and financial success of the tournament.

Estd : 1984

14. Goals Scored at Half-Time vs. Full-Time



INFERENCE:

- Full-time goals consistently outnumber half-time goals due to tactical shifts and player fatigue in later stages.
- The second half often sees higher scoring rates during knock out matches, where urgency escalates.
- Teams adopt riskier strategies in the latter stages, resulting in more goals.

15. Host Country Performance Across World Cups

Country	Year
Argentina	1978
Brazil	1950
Brazil	2014
Chile	1962
England	1966
France	1938
France	1998
Germany	1974
Germany	2006
Italy	1934
Italy	1990
Korea/Japan	2002

INFERENCE:

- Host countries often perform better, with notable wins by Uruguay(1930), Italy (1934), and France (1998).
- Home advantage includes crowd support, acclimatization, and reduced travel strain.
- Performance varies based on the strength of the host nation's footballing history.

Estd : 1984