FIFA DATA ANALYSIS

FIFA INTRODUCTION

The full form of FIFA is **"Fédération Internationale de Football Association"**, which is French for "International Federation of Association Football."

FIFA, the **Fédération Internationale de Football Association**, is the international governing body for soccer (known as football outside North America). Founded in 1904, FIFA is responsible for overseeing the organization and promotion of international football competitions, including the FIFA World Cup, which is the most prestigious tournament in the sport.

Key Functions of FIFA:

- **Organizing Competitions:** FIFA organizes major international tournaments, such as the FIFA World Cup for men and women, and youth tournaments like the U-20 and U-17 World Cups.

- **Setting Rules:** FIFA works with the International Football Association Board (IFAB) to set and update the rules of the game, known as the Laws of the Game.

- **Promoting Football:** FIFA promotes football globally, aiming to grow the sport and make it more inclusive. It supports football development programs, especially in less developed countries.

- **Governing Football:** FIFA governs international football by enforcing regulations, organizing competitions, and ensuring fair play. It also addresses disciplinary issues and oversees the transfer system for players.

- **Membership:** FIFA is made up of 211 member associations (national football federations), making it one of the largest international sports organizations. These members are grouped into six continental confederations: Africa (CAF), Asia (AFC), Europe (UEFA), North and Central America and the Caribbean (CONCACAF), Oceania (OFC), and South America (CONMEBOL).

FIFA is headquartered in Zurich, Switzerland, and it plays a crucial role in shaping the future of football globally.

The FIFA World Cup is one of the most prestigious and widely watched sporting events in the world, held every four years. Here's an overview of how it works:

TYPES OF POSITIONS IN FOOTBALL

1. Goalkeeper (GK)

- Protects the goal and prevents the opposing team from scoring.

2. Defenders

a. Right Back (RB): Defends the right flank and supports the attack on the right side.

b. Left Back (LB): Defends the left flank and supports the attack on the left side.

c. Center Back (CB): Defends centrally and is responsible for marking the opposition's forwards and clearing the ball from the defensive area.

d. Right Center Back (RCB): Plays on the right side of the central defense, often in a 3- or 5-defender formation.

e. Left Center Back (LCB): Plays on the left side of the central defense, often in a 3- or 5-defender formation.

f. Wing Back (RWB/LWB): Operates wide on either side of the defense, with both defensive and attacking responsibilities.

3. Midfielders

a. Central Midfielder (CM): Controls the midfield area, distributing the ball and linking defense and attack.

b. Defensive Midfielder (CDM): Positioned in front of the defense, focusing on defensive duties and disrupting the opposition's attacks.

c. Attacking Midfielder (CAM): Positioned closer to the forwards, focusing on creating goal-scoring opportunities and providing assists.

d. Right Midfielder (RM): Operates on the right side of the midfield, contributing to both defense and attack.

e. Left Midfielder (LM): Operates on the left side of the midfield, contributing to both defense and attack.

f. Box-to-Box Midfielder (B2B): Moves between the defensive and attacking areas, contributing to both ends of the pitch.

4. Forwards:

a. Striker (ST): Positioned centrally, with the primary responsibility of scoring goals.

b. Center Forward (CF): Similar to a striker but may also involve holding up the ball and linking play.

c. Second Striker (SS): Positioned slightly behind the main striker, combining scoring with playmaking responsibilities.

d. Right Winger (RW): Operates on the right flank, focusing on delivering crosses and creating goal-scoring opportunities.

e. Left Winger (LW): Operates on the left flank, focusing on delivering crosses and creating goal-scoring opportunities.

5. Additional Positions

a. Attacking Wing Back (AWB): A wing back with a more attacking focus, often involved in both defensive and offensive plays.

b. Sweeper (SW): Positioned behind the center backs, clearing the ball from the defensive area and covering defensive gaps.

TYPES OF STAGES IN FIFA

1. Group stages:
- Group 1-6 same as Group A-H but former is used in older context, presently latter term is used. This is in place since 1986.

2. Round of 16:
- This is the first knockout stage of the fifa tournament. The top two teams from each of the 8 groups advance to this stage.

3. Quarter finals:
- 8 Teams compete here.

4. Semi finals:
- 4 teams play here

5. Finals:
- 2 teams play here

6. Third - place Playoff/ Playoff for third place:
- A match played between the two teams that lost in the Semifinals to determine third place.

7. Third Place:
- The final ranking awarded to the team that wins the playoff for third place.

PRELIMINARY STAGE/ QUALIFYING ROUND

1. Qualification Phase (Pre-Tournament)

Before the World Cup begins, teams must qualify through regional tournaments. The steps involved are:

a) **Regional Qualifiers:**

★ **Objective:** To determine which teams from each region will secure a spot in the World Cup.

★ **Format:** Each confederation (AFC, CAF, CONCACAF, CONMEBOL, OFC, UEFA) manages its own qualification process, which varies by region but generally follows a similar structure. The goal is to determine which teams from each region earn a spot in the World Cup.

★ **Qualification Structure(AFC):**

➢ **First round:**
   o Lower-ranked AFC teams, including those that did not automatically advance to the second round.
   o Matches are played over two legs.
   o Winning team advance to second round.

➢ **Second round:**
   o Some teams automatically advance to the second round based on their FIFA rankings and previous performance. These teams are usually the top-ranked teams from the previous World Cup or continental qualifiers.
   o In 2022 a total of 40 teams were in second round including 4 teams that qualified from round 1.
   o 40 were divided into 8 groups each having 5 teams and round robin matches were conducted and top 2 teams from all groups were qualified from third round.

➢ **Third round:**
   o Again round robin matches(Matches played over 2 legs) are played.
   o 2 teams was qualified for world cup 2022.

b) **Intercontinental Playoffs:**

★ **Setup:** Teams from different continents compete in playoff matches if needed. This is often the case for the final spots in the World Cup.

★ **Format:** Single or two-legged matches (home and away) to decide which team earns the last spot in the World Cup.

2. Tournament Preliminary Round (If Applicable)

In some World Cups, particularly before the 1998 tournament, a preliminary round within the tournament itself was used to finalize the main group stage. The steps involved are:

a) **Preliminary Round Matches:**

★ **Teams Involved:** Teams that did not automatically qualify for the group stage or were determined through earlier rounds play in this stage.
★ **Format:** Matches are played to determine which teams advance to the main group stage.
★ **Example:** In the 1998 World Cup, the preliminary round was used to determine the final teams in the group stage, though in modern tournaments, the main group stage generally includes all 32 teams directly.

b) **Determining Group Stage Line-Up:**
★ **Outcome:** The teams that succeed in the preliminary round join the 32 teams in the main group stage. The format ensures that the number of teams matches the tournament's group stage requirements.

3. Group Stage Preparation

Once the preliminary round is completed and teams for the group stage are finalized:

a) **Group Draw:**
★ **Setup:** The teams are divided into groups (e.g., 8 groups of 4 teams each in recent tournaments).
★ **Format:** Teams play a round-robin format within their group, and the top teams from each group advance to the knockout stages.

WORKFLOW

1. Data Obtained from: Unified Mentor.
2. Did some research on FIFA and football to know about it.
3. Data Cleaning: python(pandas, datetime, numpy), Excel
4. Data Visualization: Power BI.
5. Insights generated from the visualization.
6. Prepared a report and presentation based on the workflow and results.

DATA CLEANING

```python
1. import pandas as pd
   import numpy as np
   from datetime import datetime

   df = pd.read_csv("WorldCupMatches.csv")
   # Excel had 4350 rows with only data in 852 so dropped
   all the data after 852 rows
   df = df.drop(df.index[852:])
   # Dropping duplicates
   df = df.drop_duplicates()

   # Checking shape of the df
   print(df.shape)

   # To check if there is any null values
   print(df.isnull().sum())
   # found one missing value in attendance

   # To find the index of the value that is missing in
   attendance
   empty_indices = df[df['Attendance'].isnull()].index
   print(empty_indices)
```

```python
# Calculate the mean of the 'Attendance' column,
# ignoring NaN values, and round it to an integer
mean_attendance = int(round(df['Attendance'].mean()))
df['Attendance'].fillna(mean_attendance, inplace=True)

# Sliced Datetime till date
df['Datetime'] = df['Datetime'].str[:11]

# Rename the column Datetime to Date
df.rename(columns={'Datetime': 'Date'}, inplace=True)

# To sum all goals
df['Goals'] = df['Home Team Goals'] + df['Away Team
Goals']

# To replace spaces with underscores in column names
df.columns = df.columns.str.replace(' ', '_')
df.columns = df.columns.str.replace('-', '_')

# To sum half time goals
df['Half_Time_Goals'] = df['Half_time_Home_Goals'] +
df['Half_time_Away_Goals']

# To rename column names
df['Country'] = df['Home_Team_Name']
df['Win_Conditions'] = df['Win_conditions']
df['Half_Time_Home_Goals'] =
df['Half_time_Home_Goals']
df['Half_Time_Away_Goals'] =
df['Half_time_Away_Goals']

# Creating column Outcome_Of_The_Match with
# conditional function
def outcome(df):
    if df['Home_Team_Goals'] > df['Away_Team_Goals']:
```

```
        return 'Home_Team_Win'
    elif df['Home_Team_Goals'] < df['Away_Team_Goals']:
        return 'Away_Team_Win'
    else:
        return 'DRAW'

# df.apply is used to apply functions along an axis on df
that is on rows or columns axis = 1 for rows.
df['Outcome_Of_The_Match'] = df.apply(outcome,
axis=1)

# To reorder columns
new_order = ['Year','Date', 'RoundID', 'Stage', 'MatchID',
'Country', 'City', 'Stadium', 'Home_Team_Name',
'Home_Team_Initials', 'Half_Time_Home_Goals',
'Home_Team_Goals','Away_Team_Name',
'Away_Team_Initials', 'Half_Time_Away_Goals',
'Away_Team_Goals', 'Half_Time_Goals',
'Goals','Outcome_Of_The_Match', 'Win_Conditions',
'Attendance', 'Referee', 'Assistant_1', 'Assistant_2']
df = df.reindex(columns=new_order)

# df.to_csv('WC_MATCHES_BY_PY.csv')
```
2.

INSIGHTS BEFORE DATA CLEANING

1. **MatchID:** Unqiue for each match.
2. **RoundID**: Unique for each round that is (preliminary round, round of 16, quarter finals, semin-fianls, finals).
3. Event g43 implies goal scored at 43rd minute.
4. GKC: Goal keeper coach
5. **Total Squad Size:** Each team can have a squad of **23 players** for the FIFA World Cup.
6. **Match Day:** 11 starting players and up to 5 substitutes are allowed per match (with an additional substitution if the match goes into extra time).
7. Each team has one head coach and additional coaching staff.

DATA VISUALIZATION

FIFA WC MATCH ANALYSIS

Attendance	Goals	Total no of Cities	Half - Time Goals	Total no of Matches
38M	2412	151	973	851

Year — All
Home_Team_Name — All

1540 V/S 872

Away_Team_Name — All

Away_Team_Win

Attendance by Year

Sum of Goals by Year and Country

City	Stadium
Manaus	Arena Amazonia
Curitiba	Arena da Baixada
Sao Paulo	Arena de Sao Paulo
Salvador	Arena Fonte Nova
Cuiaba	Arena Pantanal
Recife	Arena Pernambuco
Vasteras	Arosvallen
Rosario	Arroyito - Estadio Dr. Lisandro de la Torre

FIFA WC Matches Top 10 Data

Top 10 Countries By Goals

Top 10 Attendance By Cities

Top 10 Attendance By Stadium

Top 10 Cities By Goals

Top 10 Stadium By Goals

FIFA WC MATCH ANALYSIS

Year		Country	
All	▼	Argentina	▼

No Of Teams Qualified

Stage	MatchIDCount
Final	2
Group 1	7
Group 2	2
Group 3	4
Group 4	3
Group A	4
Group B	8
Group C	2
Group D	3
Group F	4
Total	**52**

Frequency distribution of outcome of a match

Outcome_Of_The_Match
- Home_Team_Win
- Away_Team_Win
- DRAW

5 (9.62%)
8 (15.38%)
39 (75%)

Average Goals

2.10

Country Flag
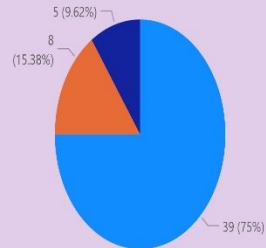
Home Team Goals

## 109

Away Team Goals

44

Half Time Goals

61

Top 10 matches with highest number of attendance

| MatchID	Sum of Attendance	Home_Team_Name	Away_Team_Name	City	Stadium	Referee
393	114600	Argentina	Germany	Mexico City	Estadio Azteca	ARPPI FILHO
392	114580	Argentina	England	Mexico City	Estadio Azteca	BENNACEUR
388	114500	Argentina	Belgium	Mexico City	Estadio Azteca	MARQUEZ RA
749	95000	Argentina	Belgium	Barcelona	Camp Nou	CHRISTOV Vo
300061502	84377	Argentina	Mexico	Johannesburg	Soccer City Stadium	ROSETTI Rob
300061458	82174	Argentina	Korea Republic	Johannesburg	Soccer City Stadium	DE BLEECKER
300186477	74738	Argentina	Bosnia and Herzegovina	Rio De Janeiro	Estadio do Maracana	AGUILAR Joel
26	73780	Argentina	Cameroon	Milan	Giuseppe Meazza	VAUTROT Mic
1088	72886	Argentina	USA	Montevideo	Estadio Centenario	LANGENUS Jc
2197	71666	Argentina	France	Buenos Aires	El Monumental - Estadio Monumental Antonio Vespuci	DUBACH Jean
Total	**898301**					

FIFA WC Matches Data By Charts
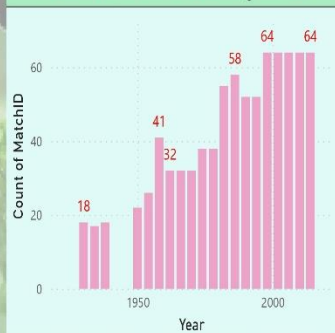
Goals Scored By Year
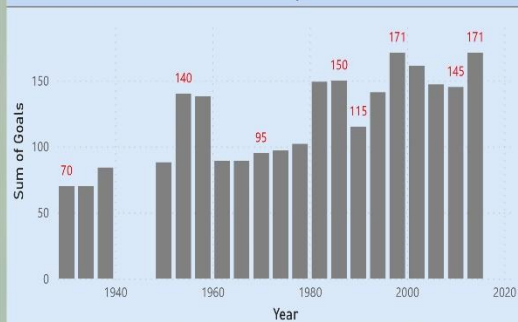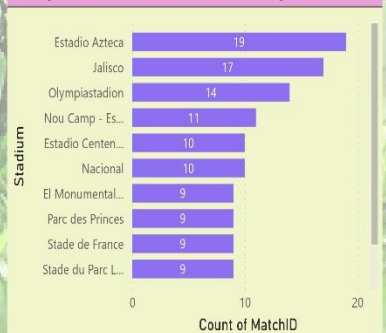

Goals Scored By Countries

Highest Number Of Matches According To Cities
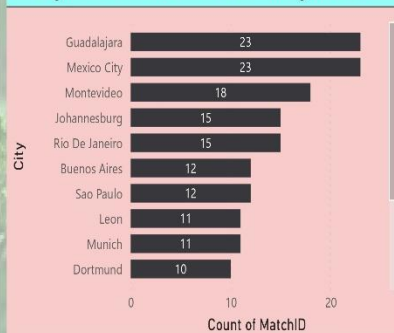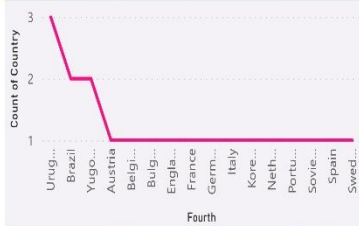

Number Of Matches According To Year
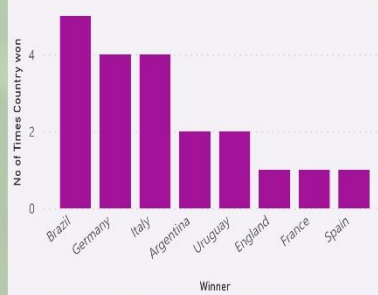

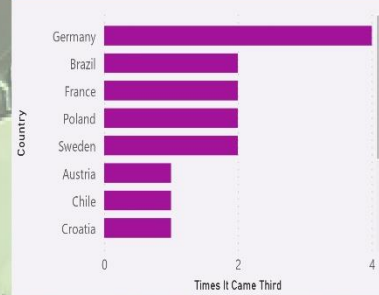### Highest number of matches according to stadium

FIFA WORLD CUP ANALYSIS

Count of Country by Fourth

Number of Times A Country Came 1st

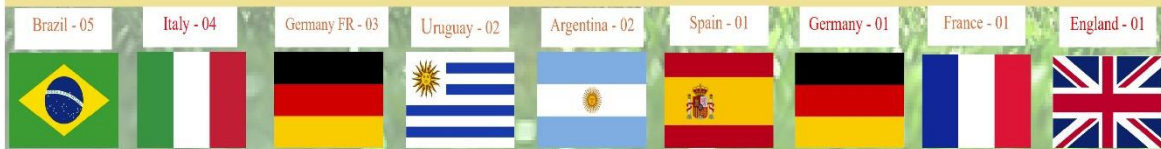## Number of Times A Country Came Third

Germany
Brazil
France
Poland
Sweden
Austria
Chile
Croatia

Count of Country by Winner

- Spain 1 (5%)
- En... 1 (5...)
- Uruguay 2 (1...)
- Argentina 2 (1...)
- Italy 4 (20%)
- Brazil 5 (25%)
- Germany 4 (20%)
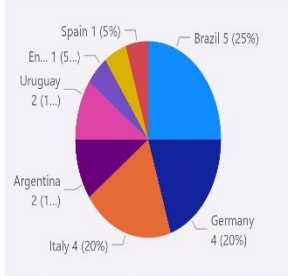
Host Countries With Respective Years & Attendance

Sweden, Spain, France, Korea, USA, Mexico, Switzerland, Brazil, South Africa, Argentina, Uruguay, Chile

NORTH AMERICA, EUROPE, ASIA, AFRICA, SOUTH AMERICA, AUSTRALIA

Microsoft Bing | © 2024 Microsoft Corporation

Count of Runners - up by Country

Germany, Argentina, Netherlands, Brazil, Czechoslov..., Hungary, Italy, France, Sweden

Countries That Won The WC

| Brazil - 05 | Italy - 04 | Germany FR - 03 | Uruguay - 02 | Argentina - 02 | Spain - 01 | Germany - 01 | France - 01 | England - 01 |

FIFA Players Analysis

Count of Players

7635

Total Goals

2412

Home_Team_Name

- Algeria
- Angola
- Argentina
- Australia
- Austria
- Belgium
- Bolivia
- Bosnia and Herzegovina
- Brazil

V/S

Away_Team_Name

- Algeria
- Angola
- Argentina
- Australia
- Austria
- Belgium
- Bolivia
- Bosnia and Herzegovina
- Brazil

Count of Players Team Wise

Team Initials	Count of Player Name
BRA	307
ITA	301
ARG	282
MEX	273
FRA	241
ESP	239
ENG	229
URU	228
BEL	211
SWE	186
SUI	178
KOR	174
CHI	172
NED	172
YUG	168
USA	166
PAR	158
FRG	152
TCH	146
HUN	145
SCO	141
GER	138
AUT	126
POL	126
BUL	123
CMR	121
ROU	119
Total	**7635**

Count of Player by Position

- 2.44K (6.46%)
- 33.64K (89.04%)

Position
- ●
- ● GK
- ● C
- ● GKC

Line Up

- N
- S

Position

-
- C
- GK
- GKC

Players Name

- ?URI?I?
- A BAUTISTA
- A COLE
- A GUARDADO
- A MEDINA
- A. AL-DOSSARY
- A. AL-GANOUBI
- A. ALMEIDA
- A. AYEW
- A. BAK

INSIGHTS FROM DATA VISUALIZATION

1. Top Countries by World Cup Appearances:
  - Brazil: 78 matches
  - Germany: 78 matches
  - Italy: 57 matches
  - Argentina: 52 matches
  - England: 35 matches

2. Winning Trends (Top Countries):
  - Brazil: 58 wins
  - Germany: 53 wins
  - Argentina: 39 wins
  - Italy: 38 wins
  - England: 22 wins

3. Average Goals per Match:
  - The average number of goals per match is 2.85.

4. Highest Attendance Matches:
  - Uruguay: 173,850 attendees
  - Brazil: 152,772 attendees
  - Brazil: 142,429 attendees
  - Brazil: 138,886 attendees
  - Argentina: 114,600 attendees

5. Top Host Cities:
  - Mexico City: 23 matches
  - Montevideo: 18 matches
  - Guadalajara: 17 matches
  - Johannesburg: 15 matches
  - Rio De Janeiro: 15 matches

6. Top Stadiums:
  - Estadio Azteca: 19 matches
  - Olympiastadion: 14 matches
  - Jalisco: 14 matches
  - Nou Camp - Estadio León: 11 matches
  - Estadio Centenario: 10 matches

7. Top Referees by Matches Officiated:
  - Ravshan Irmatov (UZB): 9 matches
  - Jorge Larrionda (URU): 8 matches
  - Joel Quiniou (FRA): 8 matches

- Benito Archundia (MEX): 8 matches
- Frank De Bleeckere (BEL): 7 matches

8. Goals Over Time:
   - Early tournaments like 1954 saw high goal averages (approx. 5.4 goals per match).
   - In recent tournaments, the average has stabilized around 2.5 goals per match.

9. Most Frequent Matchups:
   - Brazil vs. Sweden and Germany vs. Argentina are among the most common fixtures.

10. Penalty Shootouts:
   - Penalty shootouts have become more common in knockout stages. Germany has a strong record.

11. Stage Performance Analysis:
   - Teams like Brazil and Germany consistently top their groups and perform well in knockout stages.

12. Host Nation Performance:
   - Several nations (Uruguay, Italy, England, etc.) have won the World Cup as hosts.

13. Biggest Wins and Goal Differentials:
   - Hungary's 10-1 win over El Salvador (1982) is one of the largest goal margins.

14. Disciplinary Records:
   - Argentina and Germany have historically received higher numbers of cards.

15. Comeback Victories:
   - Matches like Germany vs. Sweden (2018) and Belgium vs. Japan (2018) highlight comeback victories.

16. Timing of Goals:
   - A significant number of goals are scored in the last 15 minutes of matches.

17. Youngest and Oldest Goal Scorers:
   - Pele (Brazil) was the youngest to score, while Roger Milla (Cameroon) was the oldest.

18. Regional Performance:
   - Europe has won the World Cup 12 times, while South America has won 9 times.

Insights from WC Players & WC's

1. Top Scorers (Goals):

- **RONALDO**: 13 goals
- **KLOSE**: 12 goals
- **Uwe SEELER**: 9 goals
- **MÜLLER**: 9 goals
- **Gerd MUELLER**: 9 goals

2. Player Participation by Team:

- **Brazil**: 2,403 players
- **Italy**: 1,843 players
- **Argentina**: 1,807 players
- **England**: 1,378 players
- **West Germany (FRG)**: 1,364 players

3. Coach Appearances:

- **Luiz Felipe Scolari (Brazil)**: 575 appearances
- **Helmut Schoen (FRG)**: 550 appearances
- **Carlos Alberto Parreira (Brazil)**: 514 appearances

4. Goals by Player Position:

- **C (Center)**: 144 goals

5. Position Frequency:

- **Goalkeeper (GK)**: 2,441 players
- **Center (C)**: 1,510 players

6. Goals Trend Over Time:

- Goals increased from 70 in 1930 to a peak of 171 in 1998 and 2014.

7. Top World Cup Winners:

- **Brazil**: 5 wins
- **Italy**: 4 wins
- **Germany**: 4 wins

8. Top World Cup Finalists:

- **Germany**: 4 times finalist

- **Argentina**: 3 times finalist

- **Netherlands**: 3 times finalist

9. Attendance Trends Over Time:

- Attendance rose significantly from 590,549 in 1930 to 3,587,538 in 1994, with recent tournaments drawing millions of spectators.

10. Host Country Wins:

- Host countries have won the World Cup **6 times**.

11. Host Country Top 4 Finishes:

- Host countries finished in the top 4 **15 times**.