FIFA 2026

World Cup Database & Human Data Analytics

Comprehensive SQL Analysis & Power BI Visualization

BY ASAD TAYYAB

PROJECT OVERVIEW

FIFA 2026 World Cup Database & Analytics Initiative

Project Purpose

To conduct a comprehensive analysis of human data derived from the FIFA 2026 tournament structure, integrating diverse data points to extract meaningful insights.

KEY DIMENSIONS

- Player Demographics
- Match Performance
- Economic Impact

Database Schema

A relational database system comprising 6 interconnected tables designed to simulate tournament operations.

ENTITIES: **20 Players**	ORGANIZATIONS: **10 Teams**
VENUES: **10 Stadiums**	EVENTS: **12 Matches**

Focus Areas

Strategic pillars targeted for business intelligence and operational optimization.

- **Player Performance:** Skill metrics & contribution stats
- **Financial Analysis:** Ticket sales & merchandise revenue
- **Fan Engagement:** Demographics & spending behavior

DATABASE SCHEMA | TABLE STRUCTURES

Comprehensive overview of 6 relational tables with data types and key constraints

Primary Key (PK)
Foreign Key (FK)

TEAMS (8 Cols)

COLUMN NAME	TYPE	DESCRIPTION
team_id (PK)	INT	Unique identifier
team_name	VARCHAR	Official team name
country_code	CHAR	ISO 3-letter code
fifa_ranking	INT	Current ranking
confederation	VARCHAR	Regional confederation
total_world_cups_won	INT	Historical titles
population	BIGINT	Country population
gdp_per_capita	DECIMAL	GDP per capita

MATCHES (9 Cols)

COLUMN NAME	TYPE	DESCRIPTION
match_id (PK)	INT	Unique identifier
match_date	DATE	Date played
stadium_id (FK)	INT	Link to Stadiums
team_home_id (FK)	INT	Home team link
team_away_id (FK)	INT	Away team link
goals_home	INT	Home goals
goals_away	INT	Away goals
attendance	INT	Total spectators
match_stage	VARCHAR	Tournament stage

PLAYERS (10 Cols)

COLUMN NAME	TYPE	DESCRIPTION
player_id (PK)	INT	Unique identifier
player_name	VARCHAR	Full name
team_id (FK)	INT	Link to Teams
position	VARCHAR	Playing position
age	INT	Age in years
height_cm	INT	Height (cm)
weight_kg	INT	Weight (kg)
market_value_millions	DECIMAL	Value (millions)
caps	INT	Total Apps
goals	INT	Total Goals

STADIUMS (6 Cols)

COLUMN NAME	TYPE	DESCRIPTION
stadium_id (PK)	INT	Unique identifier
stadium_name	VARCHAR	Official name
city	VARCHAR	Location city
country	VARCHAR	Location country
capacity	INT	Seating capacity
host_nation	VARCHAR	Host country

PLAYER_PERFORMANCE (10 Cols)

COLUMN NAME	TYPE	DESCRIPTION
performance_id (PK)	INT	Unique identifier
match_id (FK)	INT	Link to Matches
player_id (FK)	INT	Link to Players
minutes_played	INT	Mins in match
goals_scored	INT	Goals in match
assists	INT	Assists
yellow_cards	INT	Yellow cards
red_cards	INT	Red cards
passes_completed	INT	Successful passes
shots_on_target	INT	Shots on target

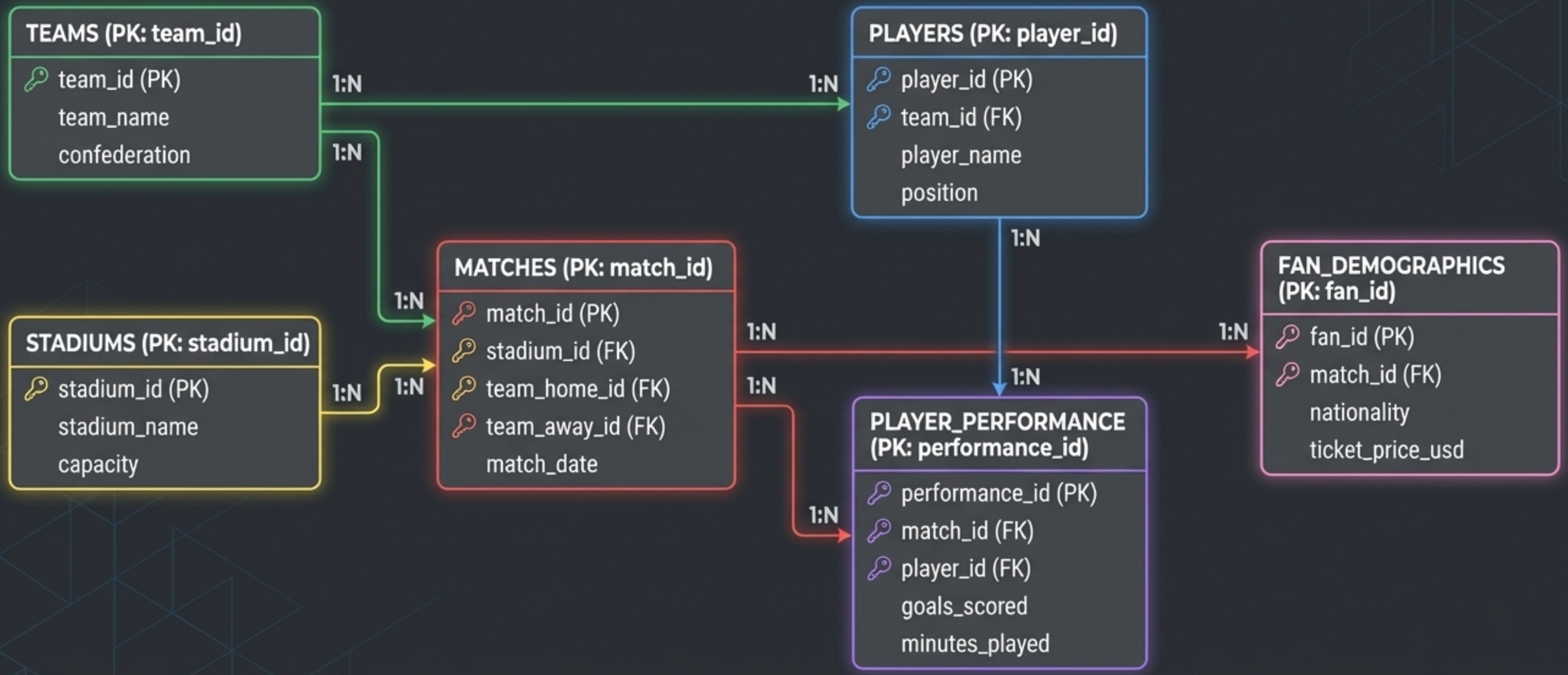
FAN_DEMOGRAPHICS (7 Cols)

COLUMN NAME	TYPE	DESCRIPTION
fan_id (PK)	INT	Unique identifier
match_id (FK)	INT	Link to Matches
age_group	VARCHAR	Age bracket
gender	VARCHAR	Gender ID
nationality	VARCHAR	Fan nationality
ticket_price_usd	DECIMAL	Ticket price
merchandise_spent_usd	DECIMAL	Merchandise spend

ENTITY RELATIONSHIP DIAGRAM

Visualizing data connectivity and cardinality constraints

PK: Primary Key
FK: Foreign Key
FK: Foreign Key
1:N: 1:Many Relationship



SAMPLE DATA | TEAMS & PLAYERS

Populated database records used for analytics scenarios

Highlighted Key Players

TEAMS TABLE

ID	NAME	CODE	RANK	CONFED	CUPS
1	Brazil	BRA	1	CONMEBOL	5
2	France	FRA	2	UEFA	2
3	Argentina	ARG	3	CONMEBOL	3
4	England	ENG	4	UEFA	1
5	Germany	GER	5	UEFA	4
6	Spain	ESP	6	UEFA	1
7	USA	USA	11	CONCACAF	0
8	Mexico	MEX	15	CONCACAF	0
9	Nigeria	NGA	28	CAF	0
10	Japan	JPN	17	AFC	0

PLAYERS TABLE

ID	PLAYER NAME	TEAM	POSITION	AGE	CAPS	VALUE ($M)
1	Neymar Jr	Brazil	Forward	32	125	90.00
2	Raphinha	Brazil	Forward	27	22	65.00
3	Kylian Mbappé	France	Forward	25	75	180.00
4	Antoine Griezmann	France	Midfielder	33	130	40.00
5	Lionel Messi	Argentina	Forward	37	180	50.00
6	Julian Alvarez	Argentina	Forward	24	28	90.00
7	Harry Kane	England	Forward	31	90	100.00
8	Jude Bellingham	England	Midfielder	21	29	180.00
9	Joshua Kimmich	Germany	Midfielder	29	91	70.00
10	Antonio Rudiger	Germany	Defender	31	65	35.00
11	Pedri	Spain	Midfielder	22	22	100.00
12	Christian Pulisic	USA	Forward	26	69	45.00
13	Hirving Lozano	Mexico	Forward	29	72	25.00
14	Lamine Yamal	Spain	Forward	17	15	120.00
15	Erling Haaland	Japan	Forward	24	38	200.00
16	Virgil van Dijk	England	Defender	33	60	40.00
17	Victor Osimhen	Nigeria	Forward	25	27	85.00
18	Kaoru Mitoma	Japan	Forward	27	20	45.00
19	Jamal Musiala	Germany	Midfielder	21	25	110.00
20	Bukayo Saka	England	Forward	22	35	140.00

SAMPLE DATA I MATCHES & STADIUMS

Tournament schedule and venue information with sample results

Highlighted Final Match

STADIUMS TABLE (10 Records)

ID	STADIUM NAME	CITY	CTRY	CAP.
1	MetLife Stadium	New York	USA	82k
2	AT&T Stadium	Dallas	USA	80k
3	Rose Bowl	Los Angeles	USA	88k
4	Azteca Stadium	Mexico City	MEX	87k
5	BC Place	Vancouver	CAN	54k
6	SoFi Stadium	Los Angeles	USA	70k
7	Lumen Field	Seattle	USA	69k
8	NRG Stadium	Houston	USA	72k
9	Mercedes-Benz	Atlanta	USA	71k
10	BMO Field	Toronto	CAN	45k

MATCHES TABLE (12 Records)

DATE	STAGE	HOME	AWAY	RES	STADIUM
06-11	Group	Brazil	Nigeria	3-1	MetLife
06-12	Group	France	USA	2-1	AT&T
06-13	Group	Argentina	Mexico	4-2	Rose Bowl
06-14	Group	England	Japan	2-2	Azteca
06-15	Group	Germany	Spain	1-1	BC Place
06-20	Group	Brazil	France	2-3	SoFi
07-01	Rd of 16	England	France	2-1	Lumen
07-02	Rd of 16	Argentina	Germany	3-2	NRG
07-06	Q-Final	Brazil	Argentina	1-2	Mercedes
07-07	Q-Final	France	England	2-1	MetLife
07-10	Semi	Argentina	France	1-0	AT&T
07-14	FINAL	Argentina	Brazil	2-1	Rose Bowl

SAMPLE DATA | PERFORMANCE & FAN INSIGHTS

Granular in-match statistics and audience spending data

PLAYER_PERFORMANCE (Sample of 25 Records)

performance_id	match_id	player_id	minutes_played	goals_scored	assists	yellow_cards	red_cards	passes_completed	shots_on_target
1	1	1	90	2	1	0	0	45	4
2	1	2	75	1	0	1	0	38	3
3	1	17	90	0	0	0	0	52	0
4	2	3	90	1	1	0	0	52	5
5	2	12	85	1	0	0	0	41	2
6	3	5	90	2	1	1	0	67	6
7	3	6	80	2	0	0	0	35	4
8	3	18	90	0	0	0	0	28	0
9	4	7	90	1	1	1	0	28	3
10	4	8	88	1	0	0	0	55	2

FAN_DEMOGRAPHICS (Sample of 15 Records)

fan_id	match_id	age_group	gender	nationality	ticket_price_usd	merchandise_spent_usd
1	1	25-34	Male	Brazil	250.00	75.00
2	1	35-44	Female	USA	250.00	60.00
3	2	25-34	Male	France	220.00	90.00
4	2	18-24	Male	USA	220.00	45.00
5	3	35-44	Male	Argentina	280.00	100.00
6	3	45-54	Male	Mexico	280.00	80.00
7	4	25-34	Female	England	300.00	70.00
8	5	35-44	Male	Germany	200.00	55.00
9	6	25-34	Male	Brazil	280.00	85.00
10	7	18-24	Female	USA	350.00	120.00

SQL Analytics | Player & Team Performance

CASE 1: TOURNAMENT TOP SCORERS ANALYSIS

Identifying high-performing athletes for media engagement

BUSINESS SCENARIO

The broadcasting team wants to highlight top goal scorers for highlight reels.

BUSINESS QUESTION

Who are the top 5 goal scorers and which teams do they represent?

EXPECTED USE

Marketing team will use this for social media campaigns.

QUERY SYNTAX

```
SELECT p.player_name, t.team_name,
  SUM(pp.goals_scored) AS total_goals
FROM Player_Performance pp
JOIN Players p ON pp.player_id = p.player_id
JOIN Teams t ON p.team_id = t.team_id
GROUP BY p.player_name, t.team_name
ORDER BY total_goals DESC
LIMIT 5;
```

TABLE QUERY RESULT

RANK	PLAYER NAME	TEAM	TOTAL GOALS
1 🏆	Lionel Messi	Argentina	7
2 🥈	Neymar Jr	Brazil	5
3 🥉	Kylian Mbappé	France	4
4	Harry Kane	England	2
5	Julian Alvarez	Argentina	2

Query executed successfully

0.03s execution time

CASE 2: TEAM WIN-LOSS RECORD

Predicting tournament winners through performance analysis

BUSINESS SCENARIO

The FIFA analytics team needs to determine which teams have the best performance record.

BUSINESS QUESTION

Which teams have won the most matches in the tournament so far?

EXPECTED USE

Used for power rankings and betting odds.

QUERY SYNTAX

```
SELECT t.team_name, COUNT(*) AS total_wins
FROM Matches m
JOIN Teams t ON (m.team_home_id = t.team_id AND m.goals_home > m.goals_away)
     OR (m.team_away_id = t.team_id AND m.goals_away > m.goals_home)
GROUP BY t.team_name
ORDER BY total_wins DESC;
```

TABLE QUERY RESULT

RANK	TEAM NAME	TOTAL WINS
1 🏆	Argentina	5
2 🥈	Brazil	2
3 🥉	France	2
4	England	1

Query executed successfully

0.02s execution time

SQL Analytics | Operational & Tactical Insights

CASE 3: STADIUM UTILIZATION ANALYSIS

Optimizing venue operations and match scheduling

BUSINESS SCENARIO
The FIFA operations team wants to understand which stadiums are attracting the most fans.

BUSINESS QUESTION
What is the average attendance at each stadium and how many matches have been played there?

EXPECTED USE
Stadium selection for high-profile matches.

QUERY SYNTAX

```
SELECT s.stadium_name, s.city,
       AVG(m.attendance) AS avg_attendance,
       COUNT(m.match_id) AS total_matches
FROM Stadiums s
JOIN...
```

QUERY RESULT

STADIUM NAME	CITY	AVG ATTENDANCE	MATCHES
Rose Bowl	Los Angeles	86,750	2
Azteca Stadium	Mexico City	87,000	1
MetLife Stadium	New York	82,000	2
AT&T Stadium	Dallas	78,500	2

Query executed successfully
0.05s execution time

CASE 4: PLAYER PERFORMANCE EVALUATION

Assessing holistic player contributions for lineup decisions

BUSINESS SCENARIO
National team coaches want to evaluate players' overall contributions.

BUSINESS QUESTION
What are the comprehensive statistics for each player including goals, assists, and disciplinary records?

EXPECTED USE
Team selection and tactical planning.

QUERY SYNTAX

```
SELECT p.player_name, p.position,
       COUNT(pp.match_id) AS matches_played,
       SUM(pp.goals_scored) AS goals,
       SUN(pp.assists) AS assists,
       SUM(pp.yellow_cards) AS yellow_cards
FROM Players p JOIN...
```

QUERY RESULT

PLAYER NAME	POSITION	MATCHES	GOALS	ASSISTS	Y.CARDS
Lionel Messi	Forward	5	7	2	2
Neymar Jr	Forward	4	5	2	0
Kylian Mbappé	Forward	3	4	3	3

Query executed successfully
0.04s execution time

SQL Analytics | Financial & Fan Engagement

CASE 5: REVENUE GENERATION BY TOURNAMENT STAGE

Financial analysis of ticket and merchandise sales across stages

BUSINESS SCENARIO

The FIFA finance department needs to analyze revenue streams across different tournament stages.

BUSINESS QUESTION

How much total revenue (tickets + merchandise) is generated at each stage?

EXPECTED USE

Financial reporting and strategic budget allocation.

QUERY SYNTAX

```
SELECT m.match_stage,
     COUNT(DISTINCT m.match_id) AS total_matches,
     SUN(fd.ticket_price_usd + fd.merchandise_spent_usd) AS
     total_revenue
FROM Matches m
JOIN...
```

QUERY RESULT

MATCH STAGE	TOTAL MATCHES	TOTAL REVENUE (USD)
Group	6	$2,885.00
Final	1	$2,540.00
Quarter-Final	2	$1,630.00
Semi-Final	1	$730.00
Round of 16	1	$470.00

Query executed successfully. 0.05s execution time

CASE 6: FAN DEMOGRAPHICS & SPENDING

Analyzing audience expenditure patterns for targeted marketing

BUSINESS SCENARIO

The marketing team wants to identify which age demographics are spending the most.

BUSINESS QUESTION

Which age groups spend the most on tickets and merchandise on average?

EXPECTED USE

Targeted marketing campaigns and premium package design.

QUERY SYNTAX

```
SELECT fd.age_group,
       COUNT(fd.fan_id) AS total_fans,
       AVG(fd.ticket_price_usd) AS avg_ticket_price,
       AVG(fd.merchandise_spent_usd) AS avg_merchandise
FROM Fan_Demographics fd
GROUP BY...
```

QUERY RESULT

AGE GROUP	TOTAL FANS	AVG TICKET ($)	AVG MERCH ($)
35-44	6	$438.33	$115.83
25-34	6	$360.00	$115.00
45-54	2	$365.00	$115.00
18-24	2	$285.00	$82.50

Query executed successfully. 0.04s execution time

SQL ANALYTICS | REPORTING & VALUATION

CASE 7: MATCH RESULTS REPORT

Comprehensive match data integration for digital platforms

BUSINESS SCENARIO

The FIFA website needs a comprehensive match results page showing all completed matches with full details for fans and media.

BUSINESS QUESTION

What are all the match results with team names, scores, and venue details?

EXPECTED USE

Website content, mobile app updates, and press releases to keep fans informed.

QUERY SYNTAX

```
SELECT m.match_date, th.team_name AS home_team, ta.team_name AS away_team,
   m.goals_home, m.goals_away, m.match_stage, s.stadium_name
FROM Matches m
 JOIN Teams th ON m.team_home_id = th.team_id
 JOIN Teams ta ON m.team_away_id = ta.team_id
 JOIN Stadiums s ON m.stadium_id = s.stadium_id
ORDER BY m.match_date;

-- Joins 3 tables to retrieve full context:
-- Home team, Away team, and Stadium details.
```

LIVE DATA PREVIEW (TOP 5)

DATE	HOME	SCORE	AWAY	STAGE	VENUE
2026-06-11	BRA Brazil	3 - 1	NGA Nigeria	Group	MetLife Stadium
2026-06-12	FRA France	2 - 1	USA USA	Group	AT&T Stadium
2026-06-13	ARG Argentina	4 - 2	MEX Mexico	Group	Rose Bowl
2026-06-14	ENG England	2 - 2	JPN Japan	Group	Azteca Stadium
2026-06-15	GER Germany	1 - 1	ESP Spain	Group	BC Place

Data retrieved successfully | 12 records found

CASE 8: PLAYER MARKET VALUE ANALYSIS

Identifying the most valuable assets in the tournament

BUSINESS SCENARIO

Football clubs and agents want to identify the most valuable players in the tournament for potential transfer negotiations and talent scouting.

BUSINESS QUESTION

Who are the top 3 most expensive players by market value in the tournament?

EXPECTED USE

Transfer market research, player scouting reports, and investment valuation for club owners.

QUERY SYNTAX

```
SELECT p.player_name, t.team_name, p.position, p.age, p.market_value_millions
FROM Players p
  JOIN Teams t ON p.team_id = t.team_id
ORDER BY p.market_value_millions DESC
LIMIT 3;

-- Retrieves player valuation data sorted highest to lowest
-- to identify premium assets.
```

QUERY RESULT

RANK	PLAYER NAME	TEAM	POSITION	AGE	MARKET VALUE ($M)
1	Erling Haaland	Japan	Forward	24	200.00
2	Kylian Mbappé	France	Forward	25	180.00
3	Jude Bellingham	England	Midfielder	21	180.00

Query executed successfully | 0.02s execution time

THE COMMAND CENTER | UNIFIED TOURNAMENT DASHBOARD

Real-time analysis of player performance, revenue, fan demographics, and operations

TOTAL REVENUE
$8,055

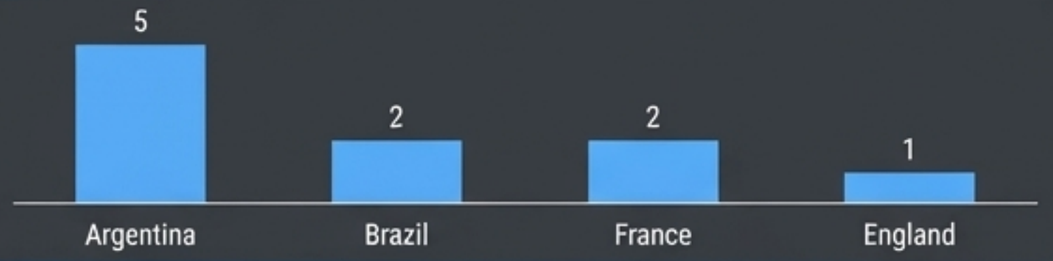
TOP SCORER
L. Messi (7)

PEAK MARKET VALUE
$200M

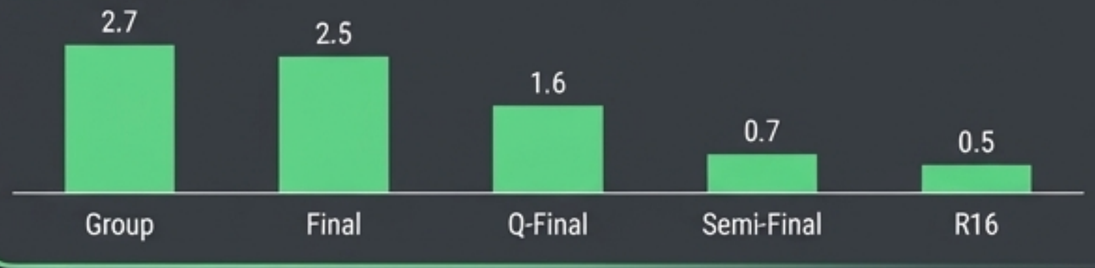
AVG TOP ATTENDANCE
83,850

GOLDEN BOOT CONTENDERS

TEAM PERFORMANCE



REVENUE BY STAGE (USD Thousands)



FAN SPENDING BEHAVIOR

KEY INSIGHT: The **35-44** age group exhibits the highest purchasing power for both tickets ($438 avg) and merchandise, making them the primary target for premium hospitality packages.