

Project 1

# FIFA WORLD CUP ANALYSIS

UM2024917 | Sumeetkumar Tripathi



# DATA OVERVIEW

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	Year	852 non-null	float64
1	Datetime	852 non-null	object
2	Stage	852 non-null	object
3	Stadium	852 non-null	object
4	City	852 non-null	object
5	Home Team Name	852 non-null	object
6	Home Team Goals	852 non-null	float64
7	Away Team Goals	852 non-null	float64
8	Away Team Name	852 non-null	object
9	Win conditions	852 non-null	object
10	Attendance	850 non-null	float64
11	Half-time Home Goals	852 non-null	float64
12	Half-time Away Goals	852 non-null	float64
13	Referee	852 non-null	object
14	Assistant 1	852 non-null	object
15	Assistant 2	852 non-null	object
16	RoundID	852 non-null	float64
17	MatchID	852 non-null	float64
18	Home Team Initials	852 non-null	object
19	Away Team Initials	852 non-null	object

Matches Dataset

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	RoundID	37784 non-null	int64
1	MatchID	37784 non-null	int64
2	Team Initials	37784 non-null	object
3	Coach Name	37784 non-null	object
4	Line-up	37784 non-null	object
5	Shirt Number	37784 non-null	int64
6	Player Name	37784 non-null	object
7	Position	4143 non-null	object
8	Event	9069 non-null	object

Players Dataset

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	Year	20 non-null	int64
1	Country	20 non-null	object
2	Winner	20 non-null	object
3	Runners-Up	20 non-null	object
4	Third	20 non-null	object
5	Fourth	20 non-null	object
6	GoalsScored	20 non-null	int64
7	QualifiedTeams	20 non-null	int64
8	MatchesPlayed	20 non-null	int64
9	Attendance	20 non-null	object

Cups Dataset

# DATA CLEANING DROPPING IRRELEVANT COLUMNS

#	Column
1	Datetime
3	Stadium
4	City
9	Win conditions
10	Attendance
13	Referee
14	Assistant
15	Assistant

**Matches Dataset**  
8 Unwanted Columns  
Dropped

#	Column
5	Shirt Number
8	Event

**Players Dataset**  
2 Unwanted Columns  
Dropped

#	Column
1	Country
9	Attendance

**Cups Dataset**  
2 Unwanted Columns  
Dropped

# DATA CLEANING DROPPING NULL COLUMNS

RoundID	0.000000
MatchID	0.000000
Team Initials	0.000000
Coach Name	0.000000
Line-up	0.000000
Player Name	0.000000
Position	89.035041

**Matches Dataset**  
**Position column is**  
**89.035% null**

# DATA CLEANING STANDARDIZING DATASETS

**New column matchWinner based on goals scored in each match.**

**Renaming all columns in camelCase.**

**Assigning the correct datatypes to all the columns.**

- C🇳te d'Ivoire
- rn">Trinidad and Tobago
- rn">Serbia and Montenegro
- rn">Bosnia and Herzegovina
- rn">United Arab Emirates
- rn">Republic of Ireland

**Correction of Character Encoding**

**New column containing Full names of team**

- Germany FR -> Germany
- IR Iran -> Iran

**Team names referring same country**

- C🇳te d'Ivoire
- rn">Trinidad and Tobago
- rn">Serbia and Montenegro
- rn">Bosnia and Herzegovina
- rn">United Arab Emirates
- rn">Republic of Ireland

**Correction of Character Encoding**

# DATA CLEANING   FINAL RESULT

#	Column	Non-Null Count	Dtype
0	year	852 non-null	int64
1	datetime	852 non-null	object
2	stage	852 non-null	object
3	homeTeam	852 non-null	int64
4	homeGoals	852 non-null	int64
5	awayGoals	852 non-null	float64
6	awayTeam	852 non-null	object
7	homeHalfTimeGoals	852 non-null	int64
8	awayHalfTimeGoals	852 non-null	int64
9	roundId	852 non-null	int64
10	matchId	852 non-null	int64
11	matchWinner	852 non-null	object

Matches Dataset

#	Column	Non-Null Count	Dtype
0	roundId	37784 non-null	int64
1	matchId	37784 non-null	int64
2	coachName	37784 non-null	object
3	lineUp	37784 non-null	object
4	playerName	37784 non-null	object
5	teamName	37784 non-null	object

Players Dataset

#	Column	Non-Null Count	Dtype
0	year	20 non-null	int64
1	winner	20 non-null	object
2	runnersUp	20 non-null	object
3	third	20 non-null	object
4	fourth	20 non-null	object
5	goalsScored	20 non-null	int64
6	qualifiedTeams	20 non-null	int64
7	matchesPlayed	20 non-null	int64

Cups Dataset

# DATA VISUALIZATION TOP WINNING TEAMS

**BRAZIL**

**GERMANY**

**ITALY**

**URUGUAY**

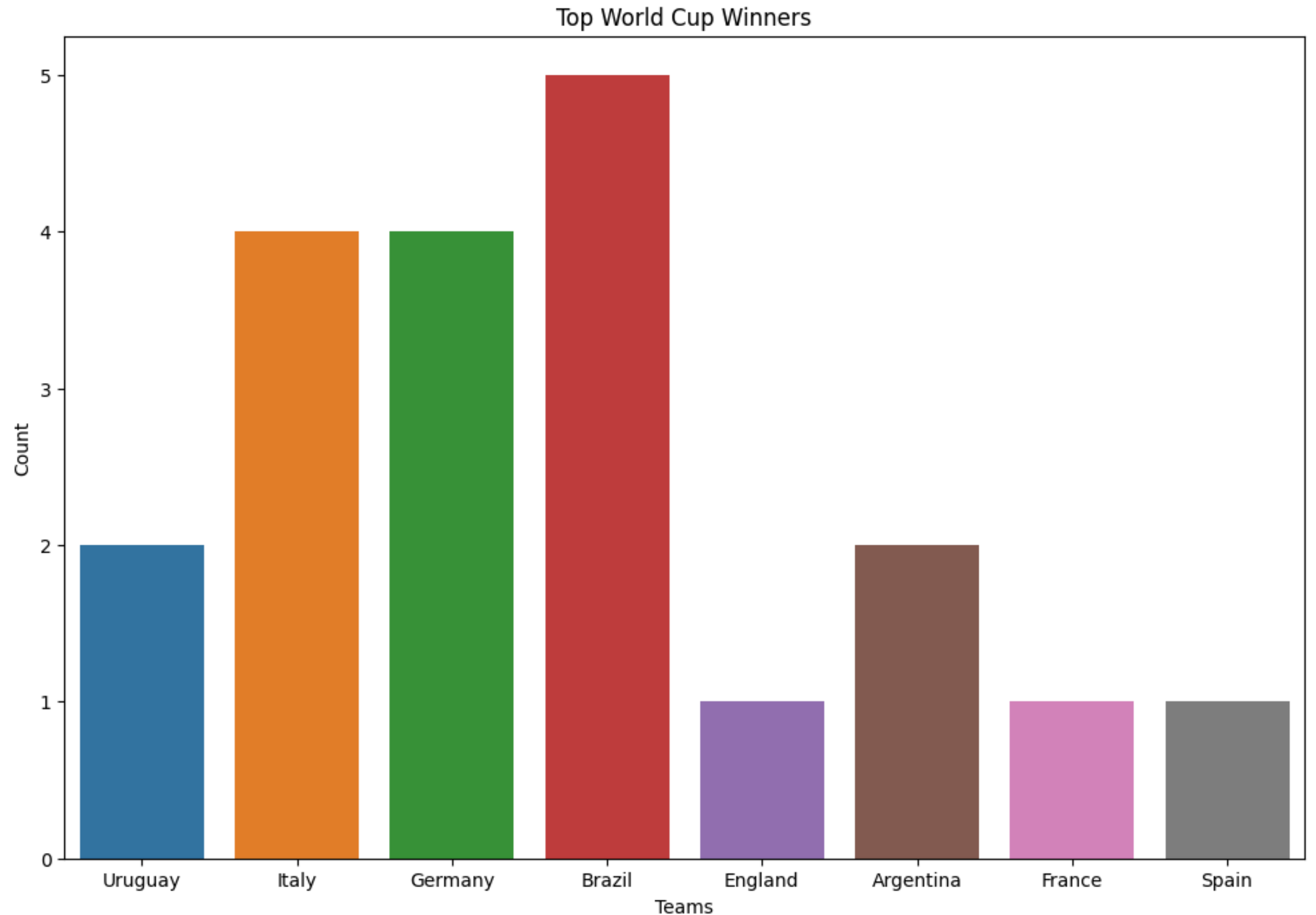
**ARGENTINA**

**ENGLAND**

**FRANCE**

**SPAIN**

Are the top world  
cup winnning  
teams.



# DATA VISUALIZATION TEAM PERFORMANCES

In the winning years,

**BRAZIL**

**GERMANY**

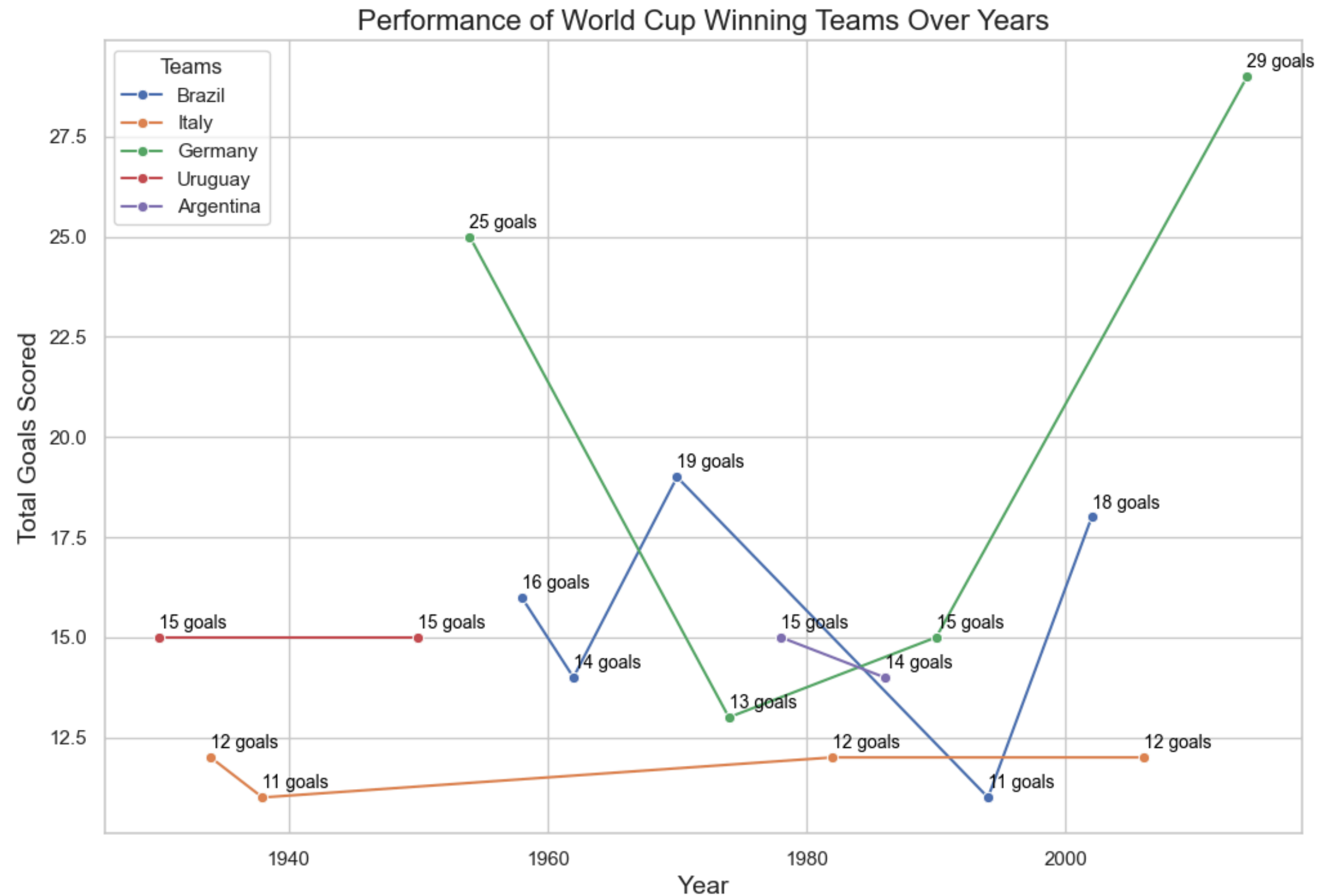
play aggressively while,

**ITALY**

**URUGUAY**

**ARGENTINA**

focussed on a  
defensive approach.





# DATA VISUALIZATION TEAM PERFORMANCES

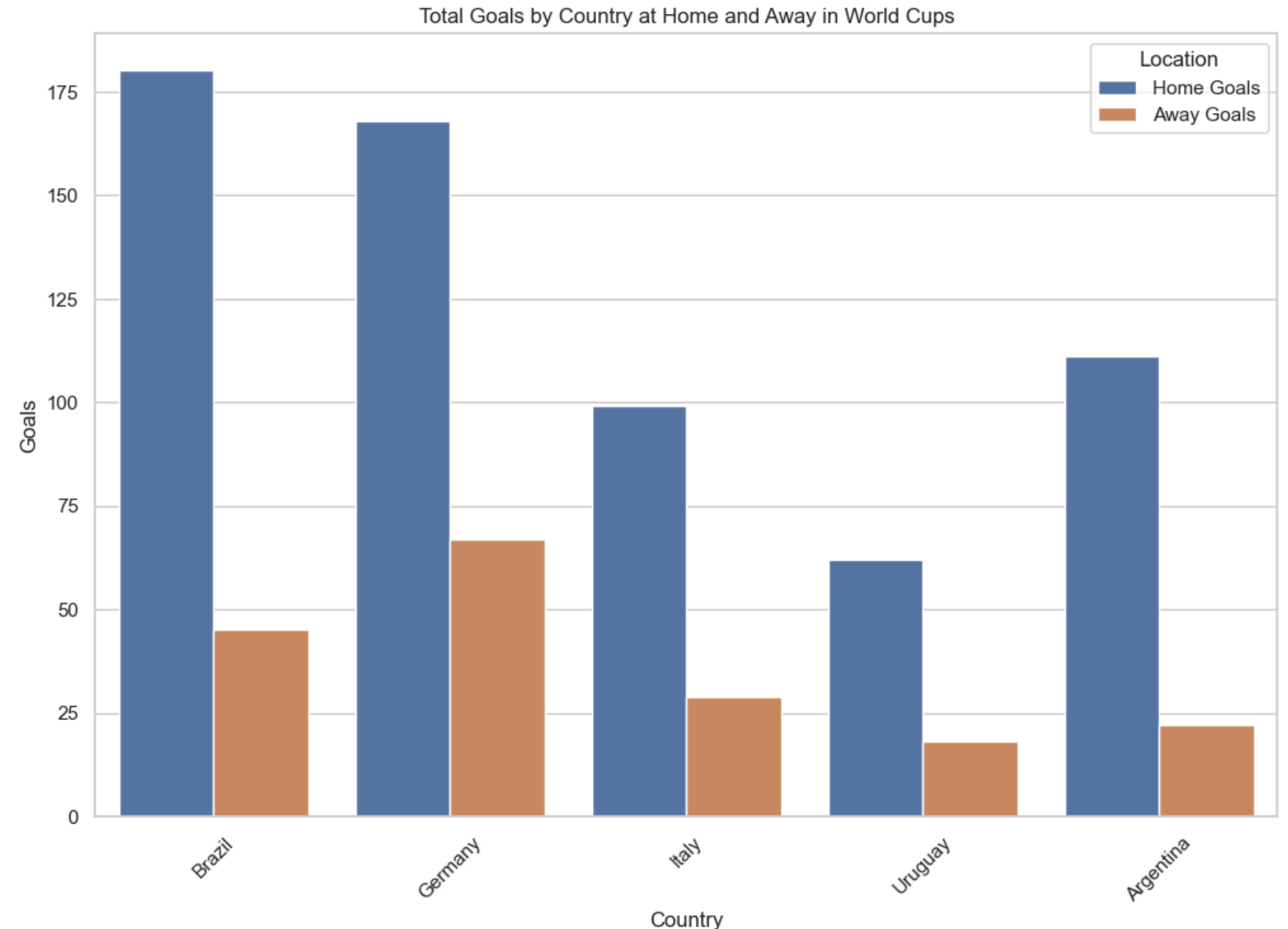
Same results for the half time goals in the winning seasons.



# DATA VISUALIZATION HOME/AWAY

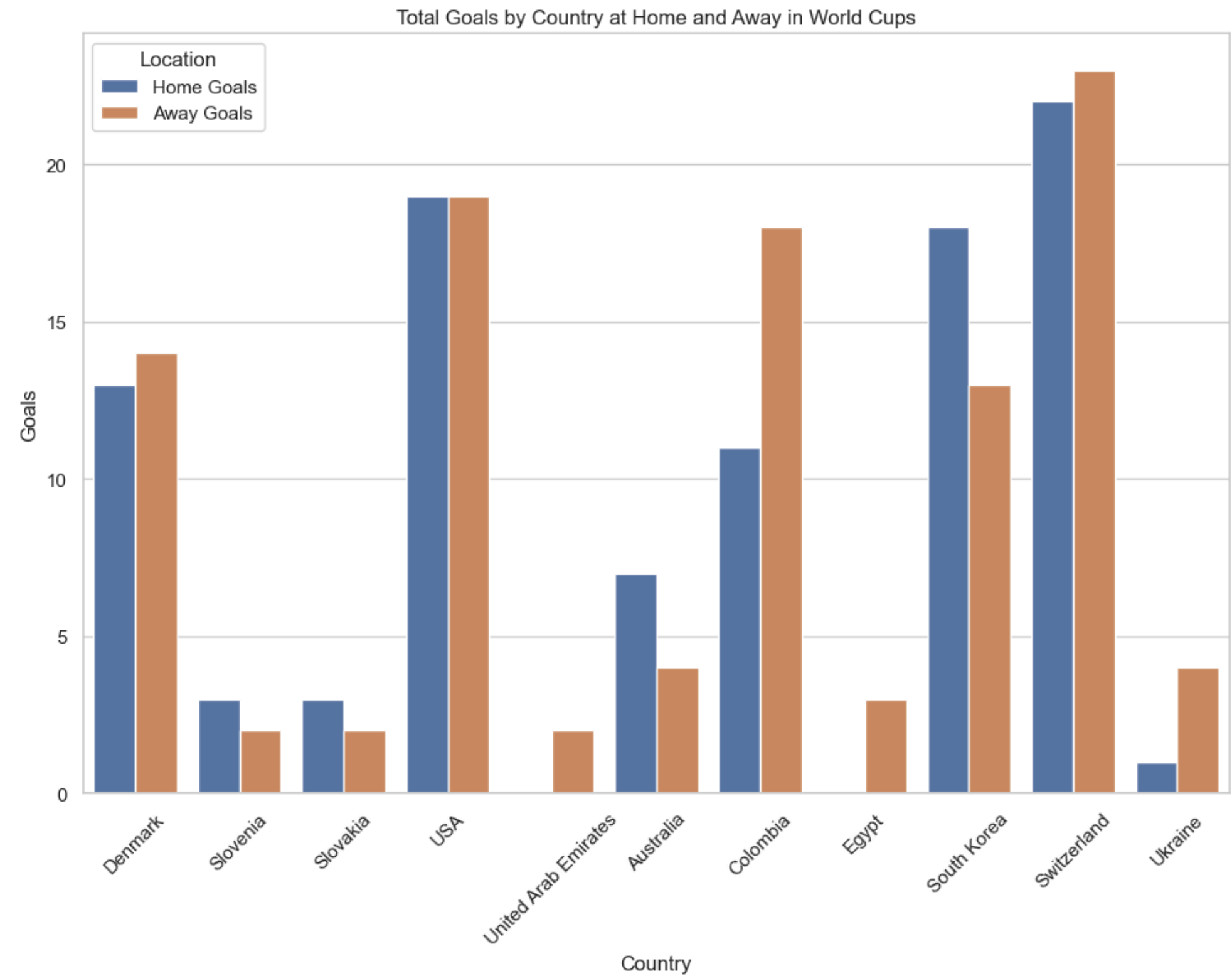
The world cup winning teams performed the best while playing in the **home** grounds

But the performance was moderate while playing in the **away** grounds.



# DATA VISUALIZATION HOME/AWAY

For the other teams, no notable differences were observed in performances between home and away grounds.



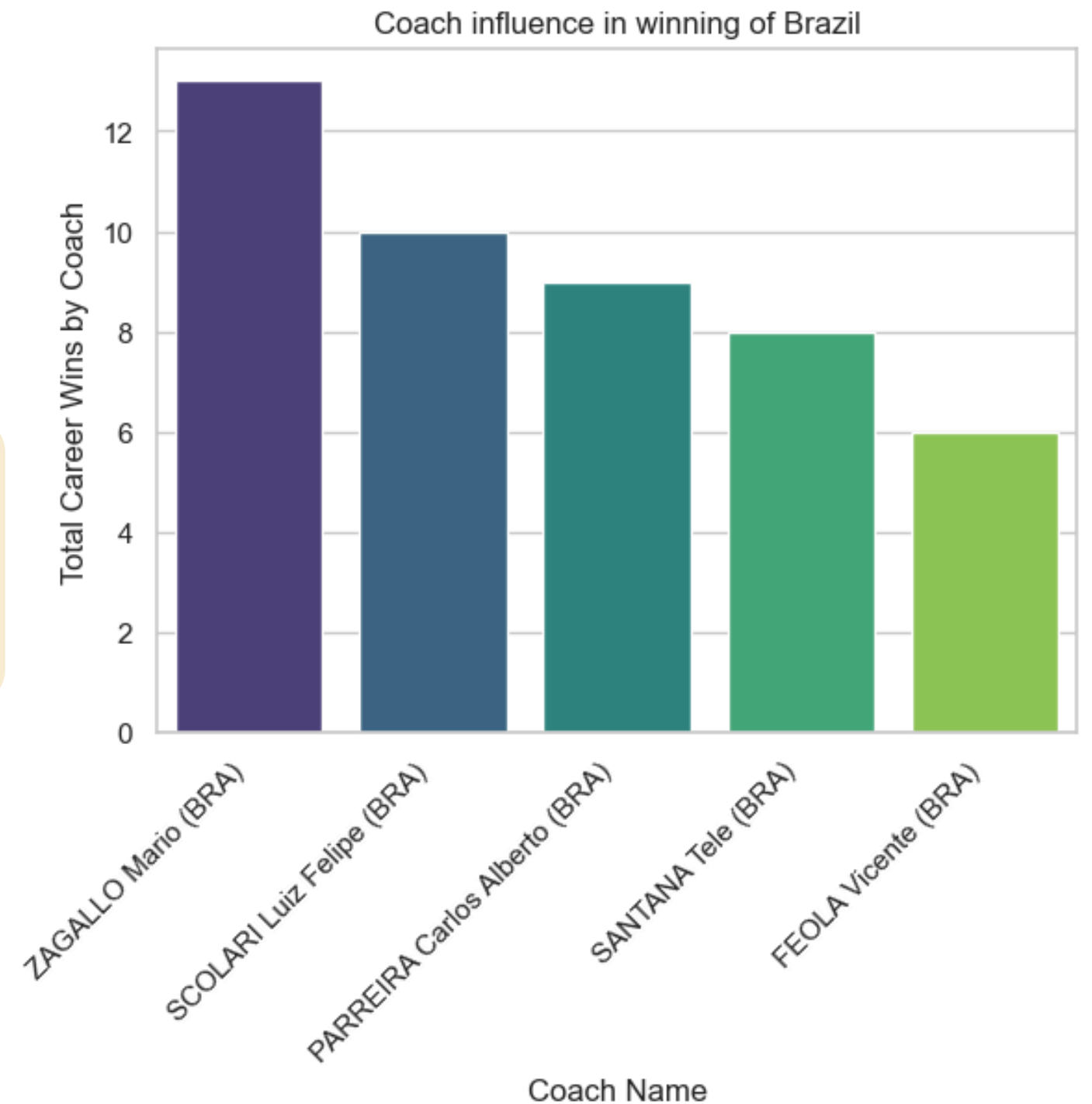
# DATA VISUALIZATION COACH INFLUENCE

## BRAZIL

Coaches who have made their teams win the **most number of times** in their respective careers are also winning the world cups for their teams.

### WorldCup Winning Coaches

0. ZAGALLO Mario (BRA)
1. SCOLARI Luiz Felipe
2. MOREIRA Aymore (BRA)
3. FEOLA Vicente (BRA)



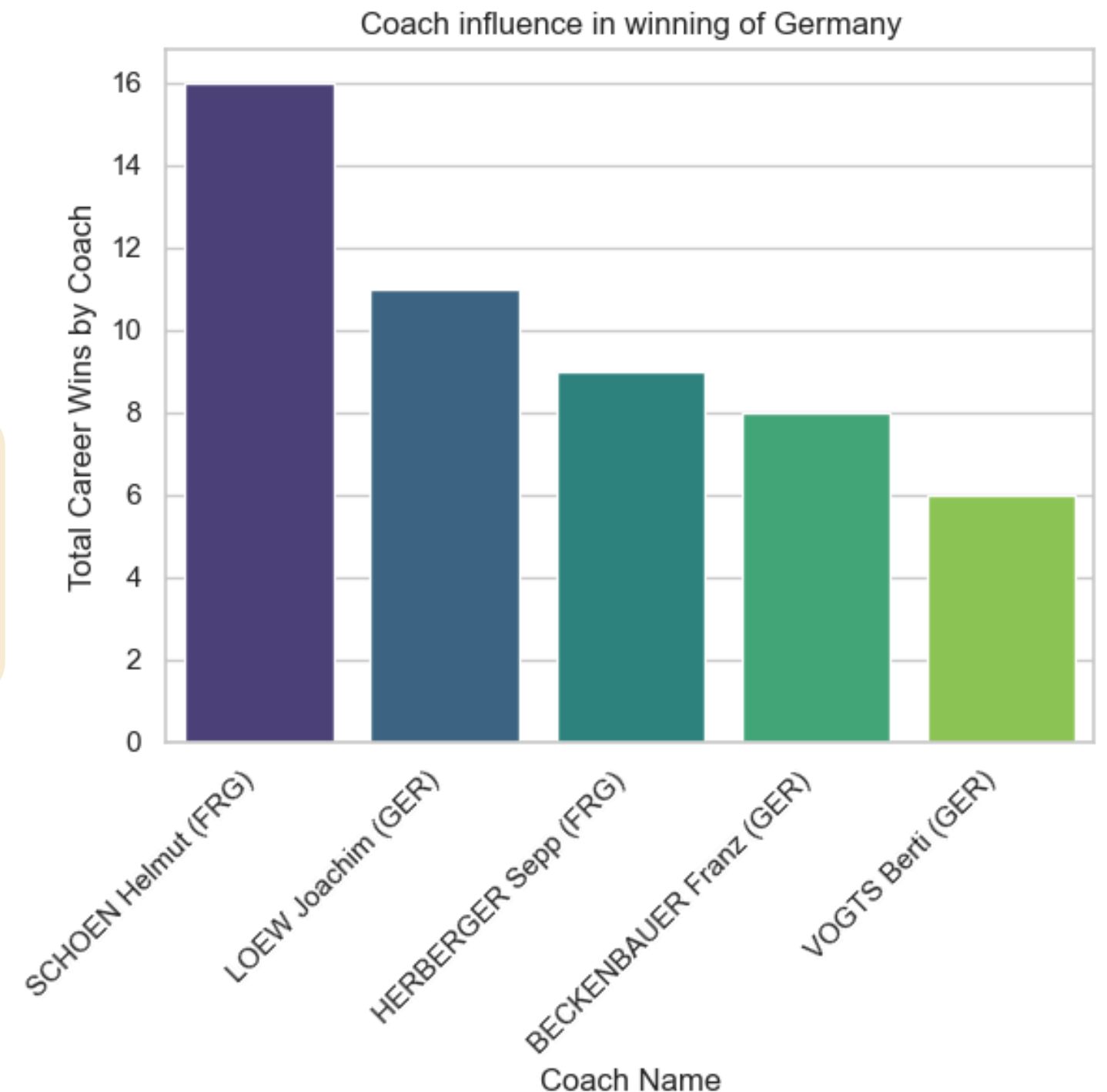
# DATA VISUALIZATION COACH INFLUENCE

## GERMANY

Coaches who have made their teams win the **most number of times** in their respective careers are also winning the world cups for their teams.

### WorldCup Winning Coaches

0. SCHOEN Helmut (FRG)
1. HERBERGER Sepp (FRG)
2. LOEW Joachim (GER)
3. BECKENBAUER Franz (GER)

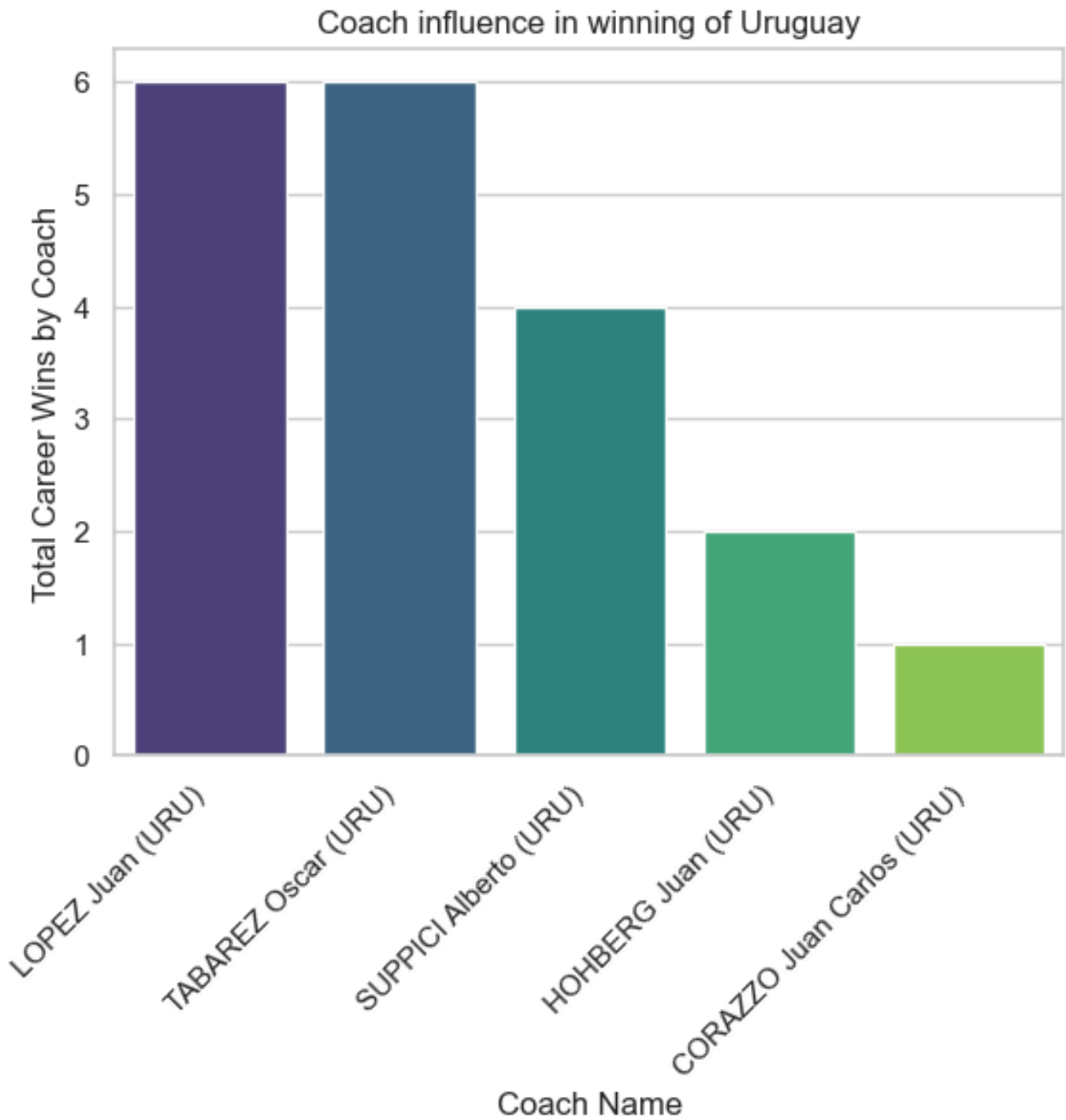


# DATA VISUALIZATION COACH INFLUENCE

## URUGUAY

Coaches who have made their teams win the **most number of times** in their respective careers are also winning the world cups for their teams.

**WorldCup Winning Coaches**  
0. SUPPICI Alberto (URU)

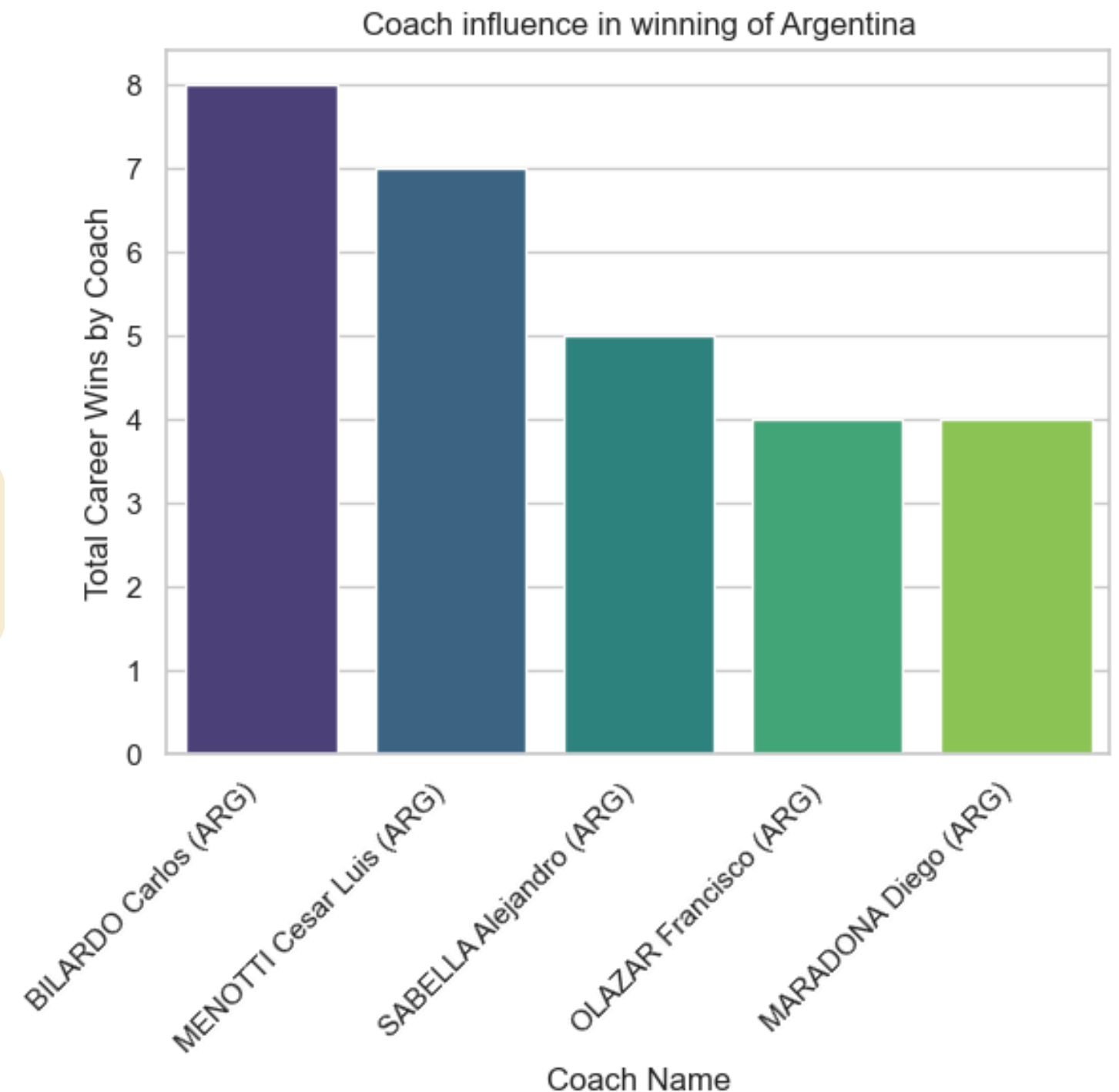


# DATA VISUALIZATION COACH INFLUENCE

## ARGENTINA

Coaches who have made their teams win the **most number of times** in their respective careers are also winning the world cups for their teams.

- WorldCup Winning Coaches**
0. BILARDO Carlos (ARG)
  1. MENOTTI Cesar Luis (ARG)

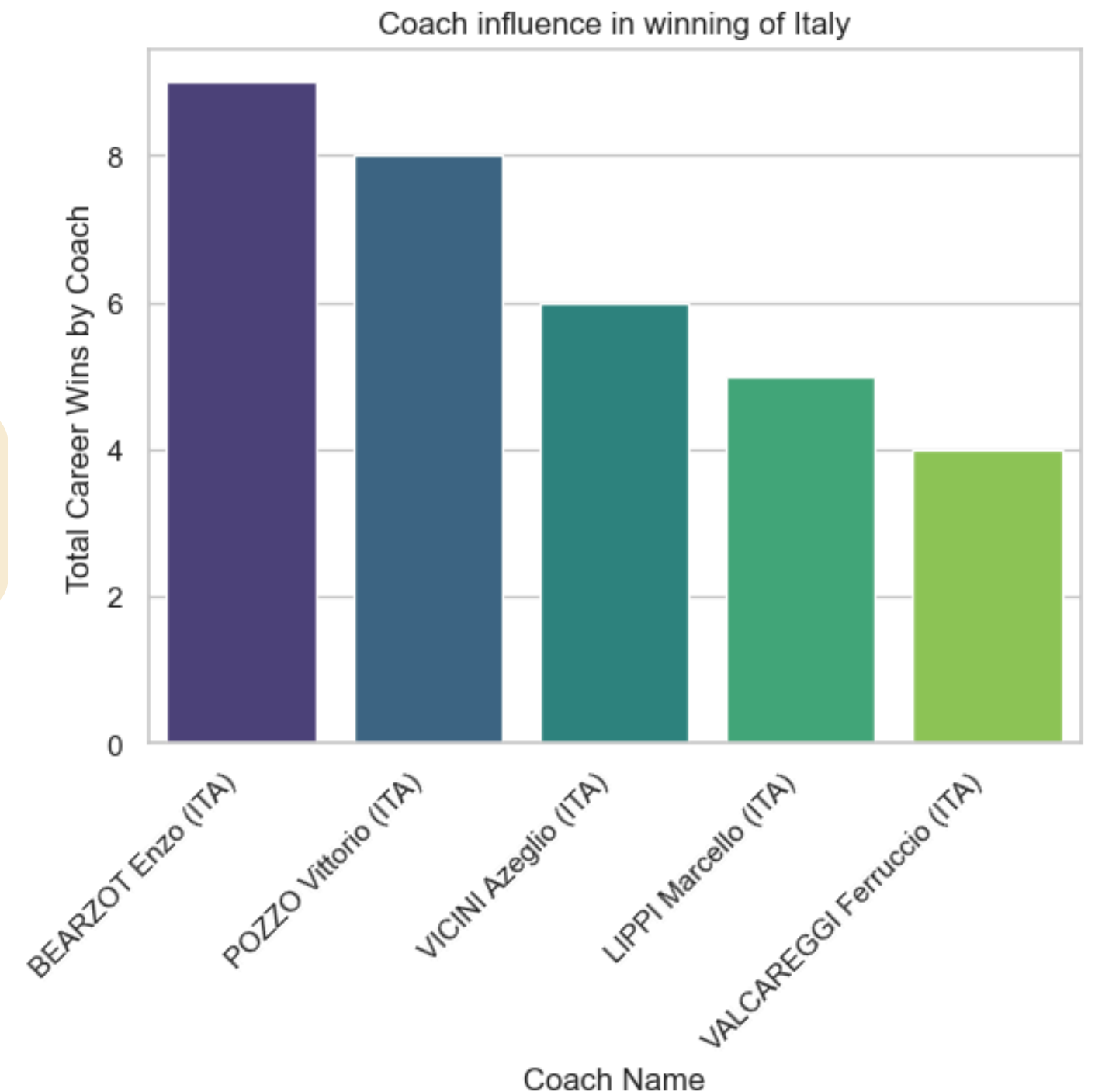


# DATA VISUALIZATION COACH INFLUENCE

## ITALY

Coaches who have made their teams win the **most number of times** in their respective careers are also winning the world cups for their teams.

- WorldCup Winning Coaches**
0. BEARZOT Enzo (ITA)
  1. POZZO Vittorio (ITA)

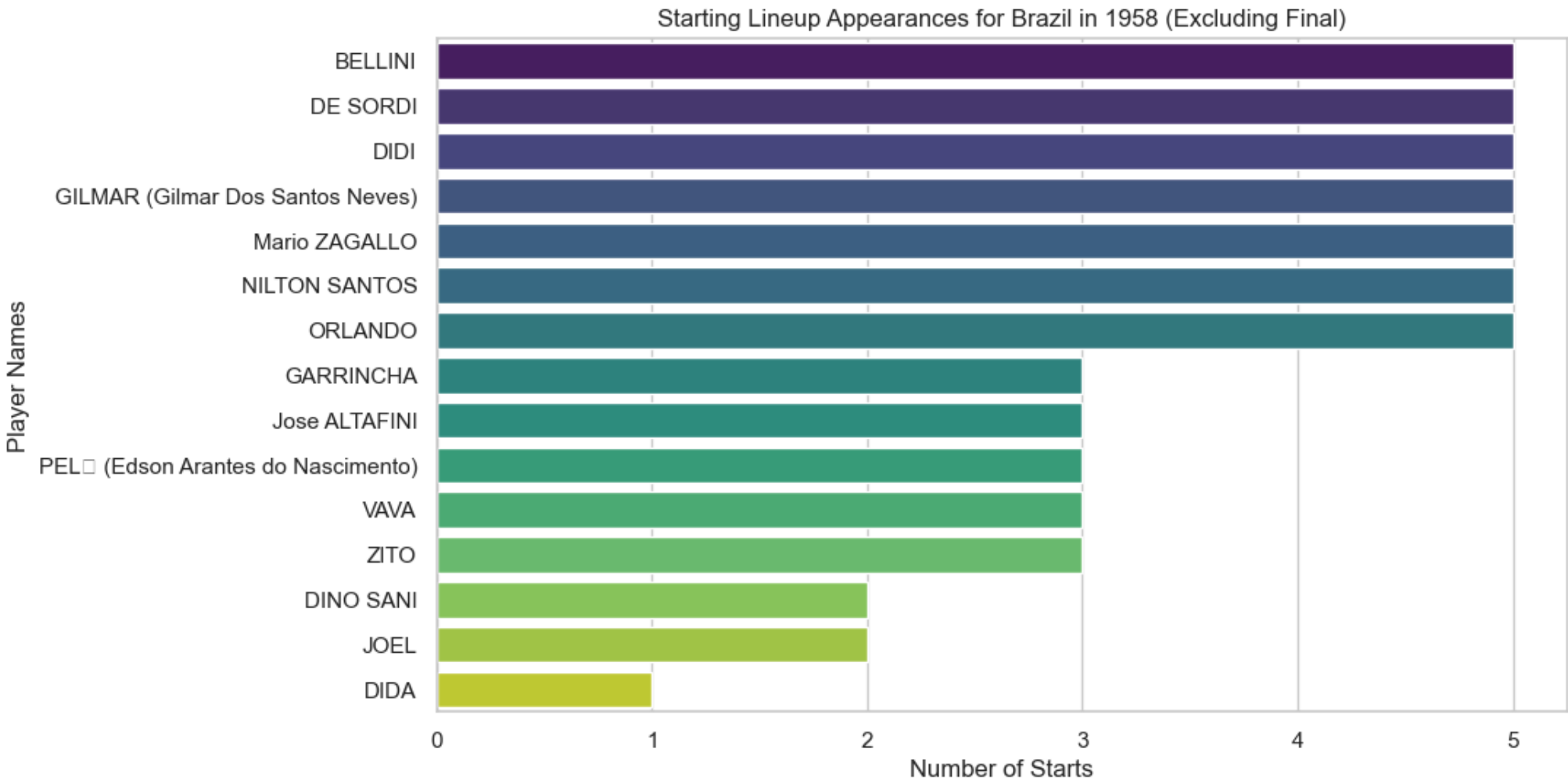




# DATA VISUALIZATION LINEUP ANALYSIS

BRAZIL

Players who have played the **maximum starting matches** in the world cup season have also played the final match and won the world cup.



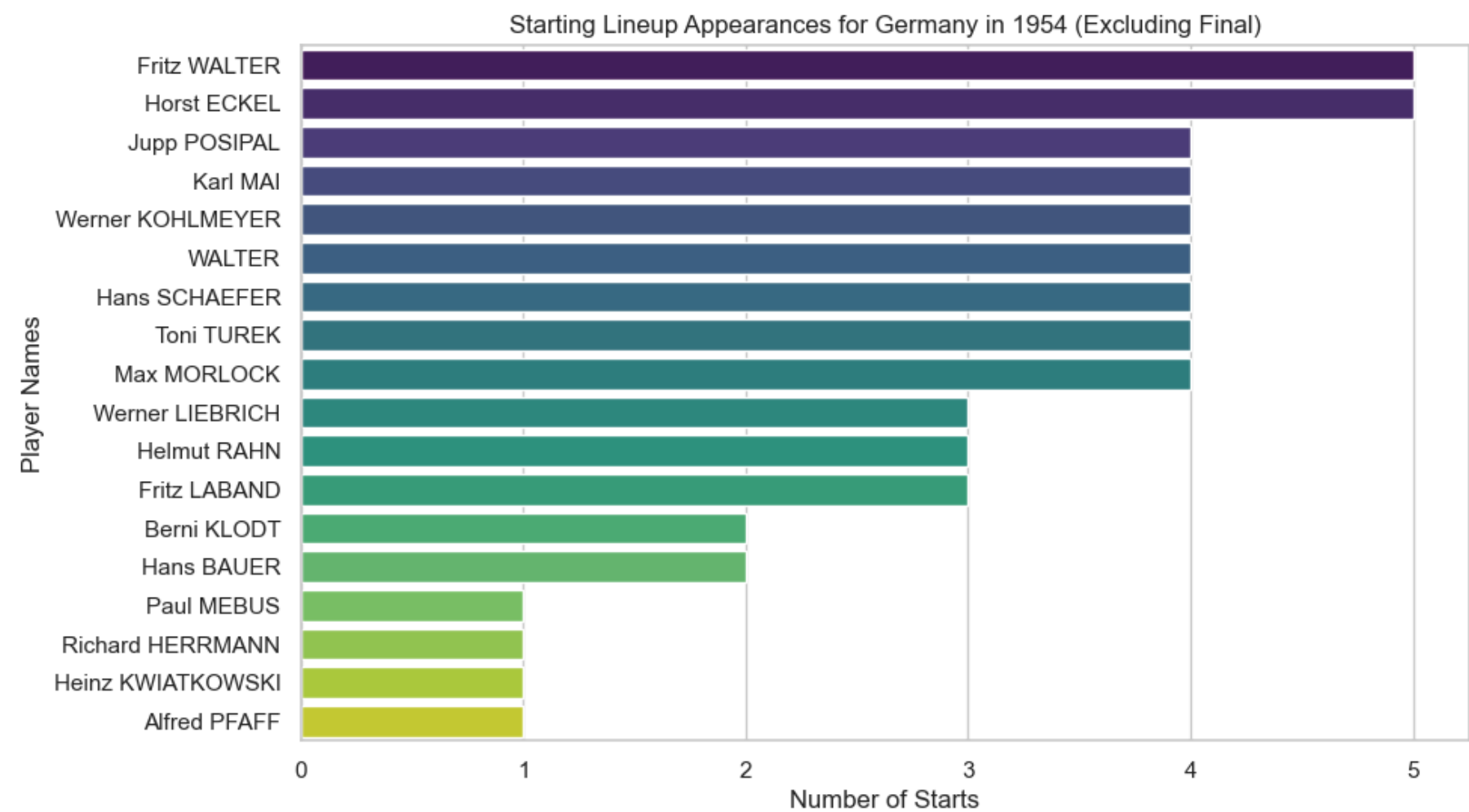
- Players in the Final Match Lineup:**
- 1: GARRINCHA
  - 2: BELLINI
  - 3: Mario ZAGALLO
  - 4: DIDI
  - 5: PELÉ (Edson Arantes do Nascimento)
  - 6: NILTON SANTOS
  - 7: DJALMA SANTOS
  - 8: ORLANDO
  - 9: ZITO
  - 10: VAVA
  - 11: GILMAR (Gilmar Dos Santos Neves)

1958

# DATA VISUALIZATION

GERMANY

Players who have played the **maximum starting matches** in the world cup season have also played the final match and won the world cup.



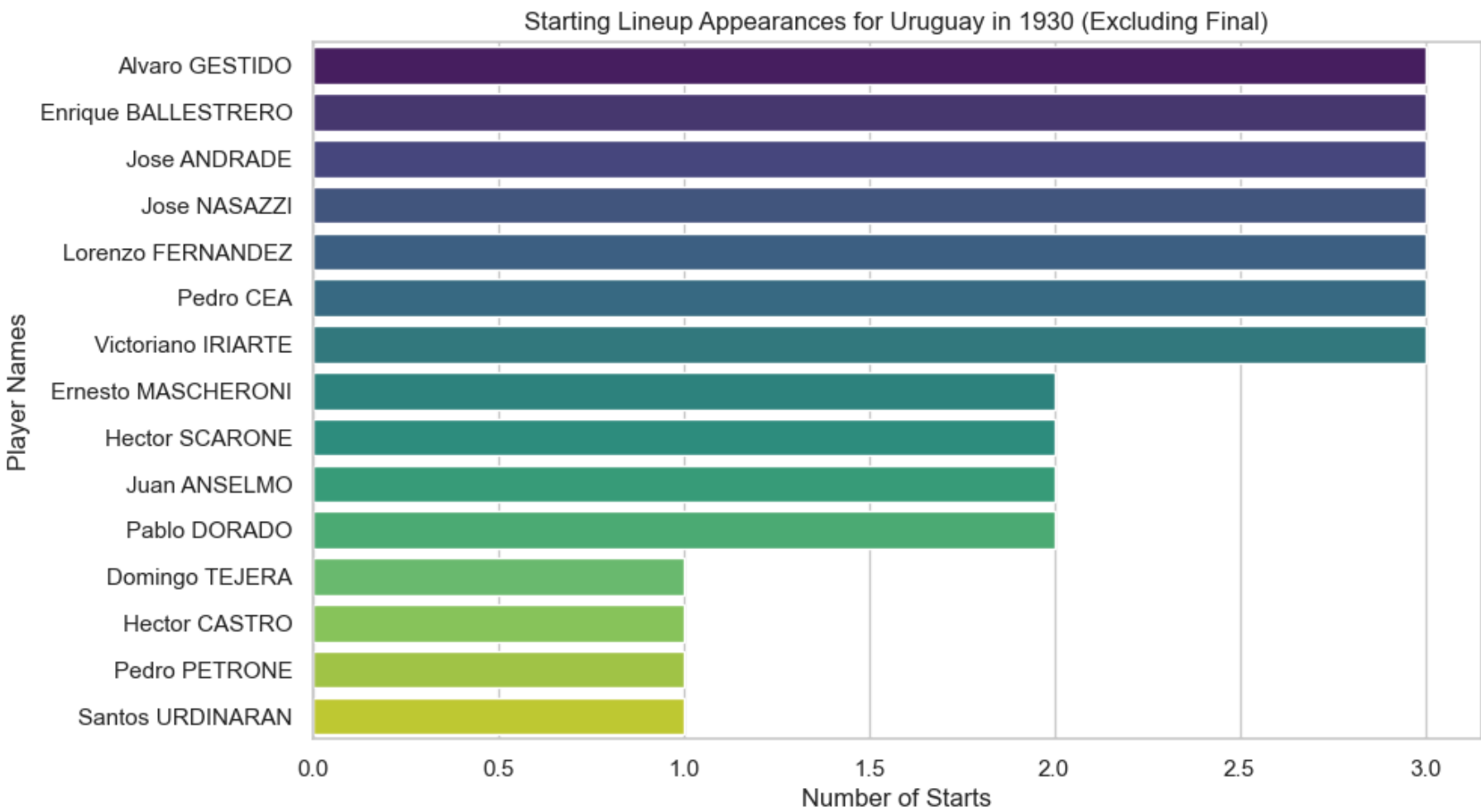
- Players in the Final Match Lineup:**
- 1: Hans SCHAEFER
  - 2: Toni TUREK
  - 3: Jupp POSIPAL
  - 4: Max MORLOCK
  - 5: Horst ECKEL
  - 6: Helmut RAHN
  - 7: Fritz WALTER
  - 8: WALTER
  - 9: Karl MAI
  - 10: Werner KOHLMEYER
  - 11: Werner LIEBRICH

1954

# DATA VISUALIZATION

## URUGUAY

Players who have played the **maximum starting matches** in the world cup season have also played the final match and won the world cup.



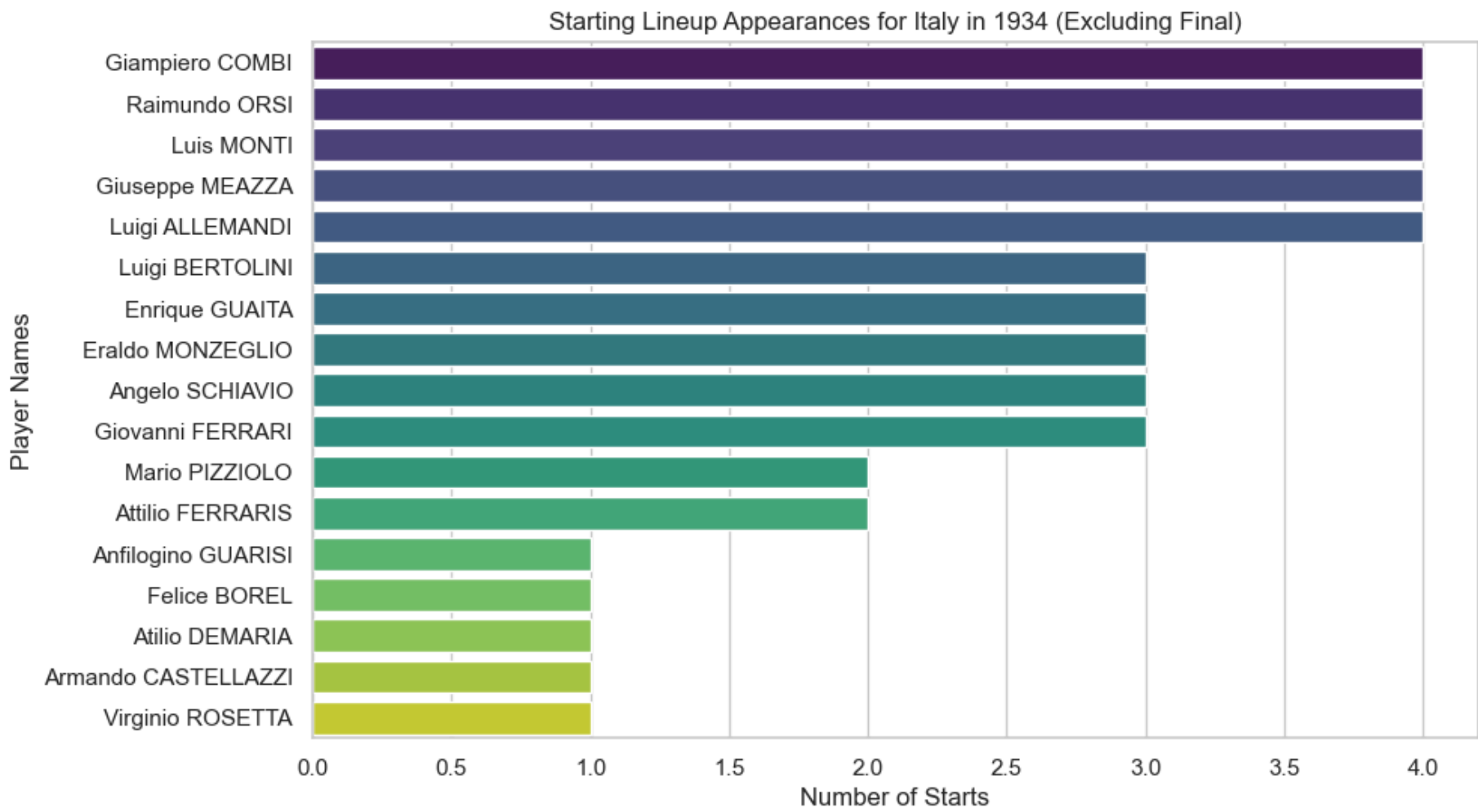
**Players in the Final Match Lineup:**

- 1: Enrique BALLESTRERO
- 2: Pablo DORADO
- 3: Victoriano IRIARTE
- 4: Hector SCARONE
- 5: Hector CASTRO
- 6: Lorenzo FERNANDEZ
- 7: Jose ANDRADE
- 8: Jose NASAZZI
- 9: Ernesto MASCHERONI
- 10: Alvaro GESTIDO
- 11: Pedro CEA

# DATA VISUALIZATION

ITALY

Players who have played the **maximum starting matches** in the world cup season have also played the final match and won the world cup.



**Players in the Final Match Lineup**

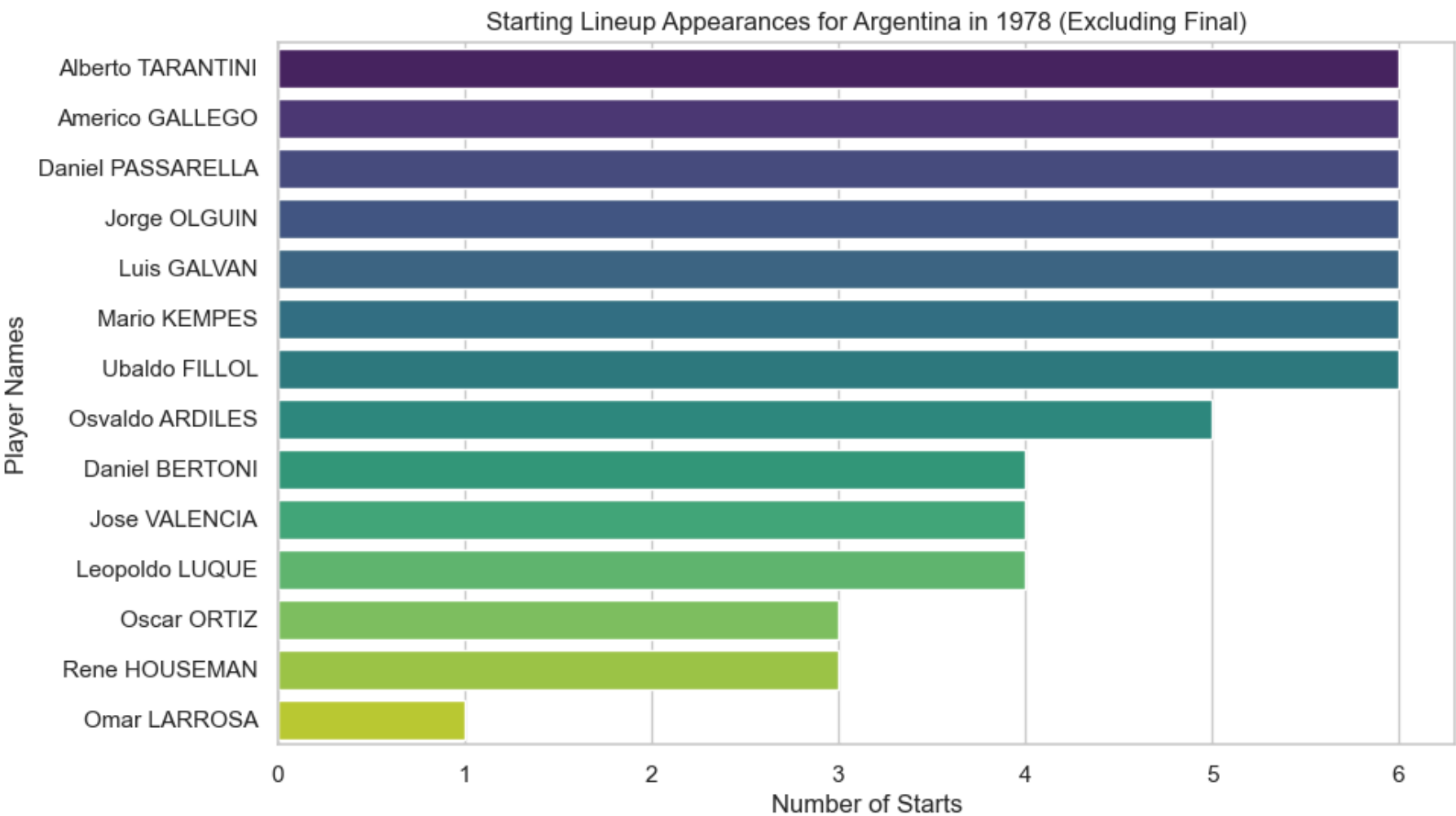
- 1: Angelo SCHIAVIO
- 2: Luis MONTI
- 3: Giovanni FERRARI
- 4: Giampiero COMBI
- 5: Attilio FERRARIS
- 6: Enrique GUAITA
- 7: Eraldo MONZEGLIO
- 8: Giuseppe MEAZZA
- 9: Raimundo ORSI
- 10: Luigi BERTOLINI
- 11: Luigi ALLEMANDI

1934

# DATA VISUALIZATION

## ARGENTINA

Players who have played the **maximum starting matches** in the world cup season have also played the final match and won the world cup.



**Players in the Final Match Lineup:**

- 1: Jorge OLGUIN
- 2: Osvaldo ARDILES
- 3: Alberto TARANTINI
- 4: Oscar ORTIZ
- 5: Americo GALLEGO
- 6: Daniel PASSARELLA
- 7: Ubaldo FILLLOL
- 8: Leopoldo LUQUE
- 9: Luis GALVAN
- 10: Daniel BERTONI
- 11: Mario KEMPES

# CONCLUSION

- Both Brazil and Germany have consistently adopted an aggressive playing style during the seasons in which they secured World Cup victories.
- In contrast, Argentina, France, Uruguay, and Italy have demonstrated a more defensive approach in their winning campaigns.
- The top five teams, each with multiple World Cup titles, tend to perform exceptionally well when playing on home grounds.
- For these teams, the coaches who led their squads to the most victories have also secured World Cup titles.
- Additionally, players who were part of the starting lineup for the majority of matches were also selected for the final match and contributed to their team's ultimate victory.