

Few facts about World Cups

1930 - 2022

Contents

1. Questions and hypothesis
2. Approach and analysis
3. Q1. inverse proportionality?
4. Q2. the influence of "his land" and fans
5. Q3. Europe and South America are the hegemonic of football?
6. Q4. Deffensive style of play
7. Results

Questions and hypothesis

1. How many rounds were played in each World Cup tournament and how has the average number of goals per match changed over the years?

Hypothesis: From tournament to tournament there were more and more teams, while the average number of goals per game decreased.

2. Does the fact that a given country is the host of the event and has fans on its side affect the result in a given championship?

Hypothesis: Yes, it does.

3. Which continents dominate when it comes to won matches and the entire championship?

Hypothesis: Europe and South America, because the culture of football is most developed there.

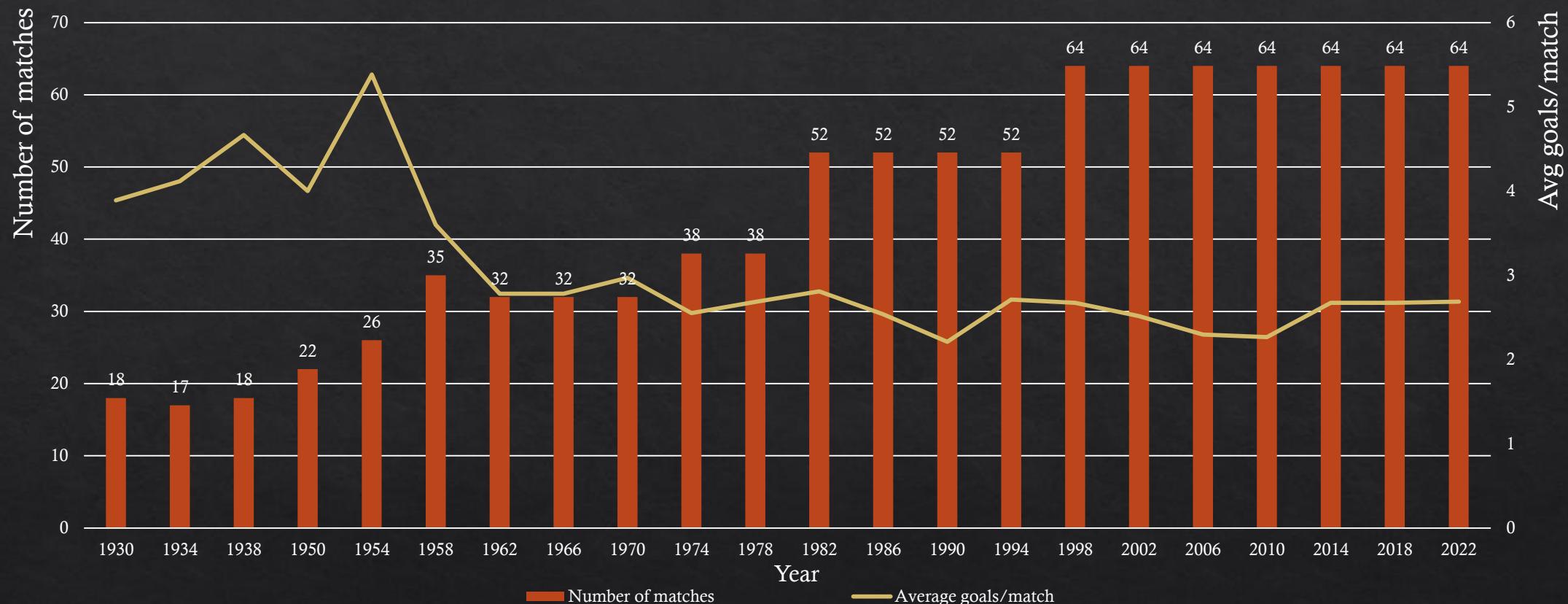
4. Which countries play the most defensive football and had the most draws in the group stage?

Hypothesis: For the past decade, the Italian team has been one that has played very defensive football.

Approach and analysis

1. I downloaded 2 datasets:
 - ◊ World cup from: <https://data.world/sportsvizsunday/sports-viz-sundays-2018/workspace/file?filename=World+Cup+Results.xlsx>
 - ◊ Continents from: https://github.com/dbouquin/IS_608/blob/master/NanosatDB_munging/Countries-Continents.csv
2. Formatted and cleaned datasets. For example: to the "continents" dataset, I added countries that no longer exist (e.g. Yugoslavia) and have participated in the World Cup in their history.
3. I performed the queries in SQL and extracted the information needed to answer the questions posed earlier.
4. In order to present the results, I made their graphical visualization

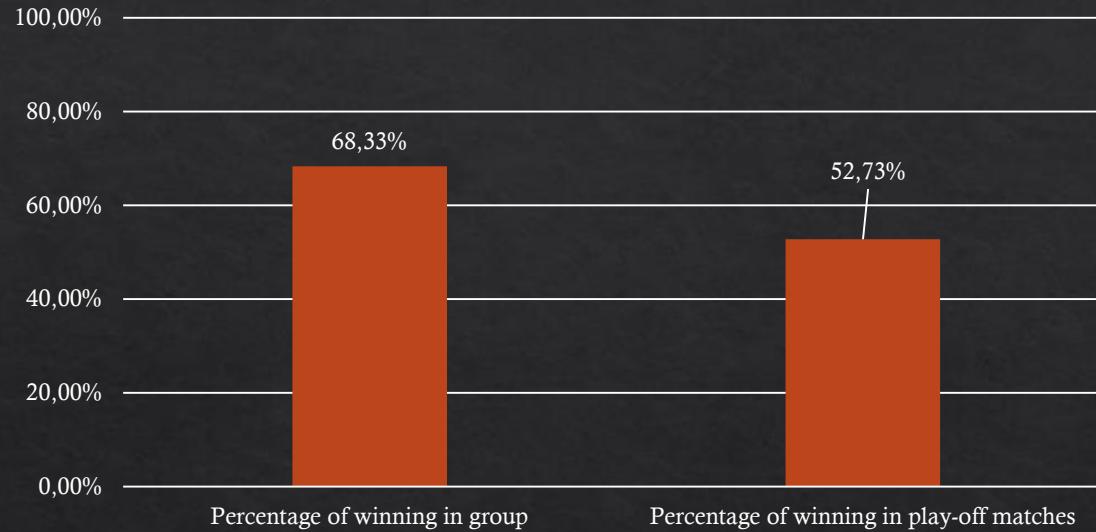
Q1. Inverse proportionality?



As the number of World Cup matches increases, the average number of goals per match decreases.

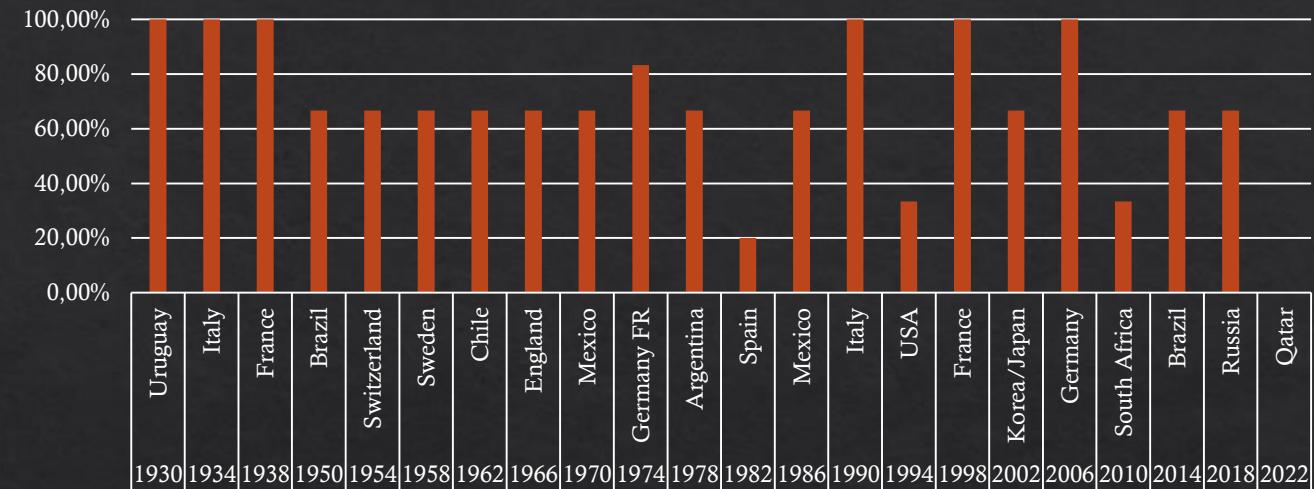
Q2. The influence of "his land" and fans

Teams as hosts on worldcup 1930-2022

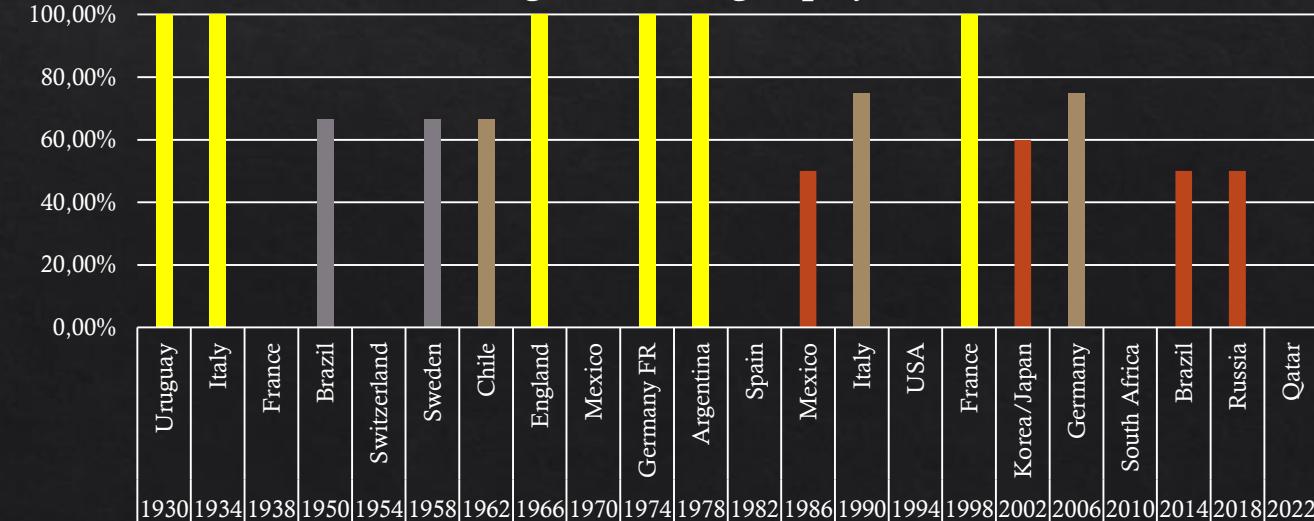


The results of the visualization show how much influence being a host has on the results. 6 out of 21 hosts won the gold medal in front of their audience.

Percentage of winning in group

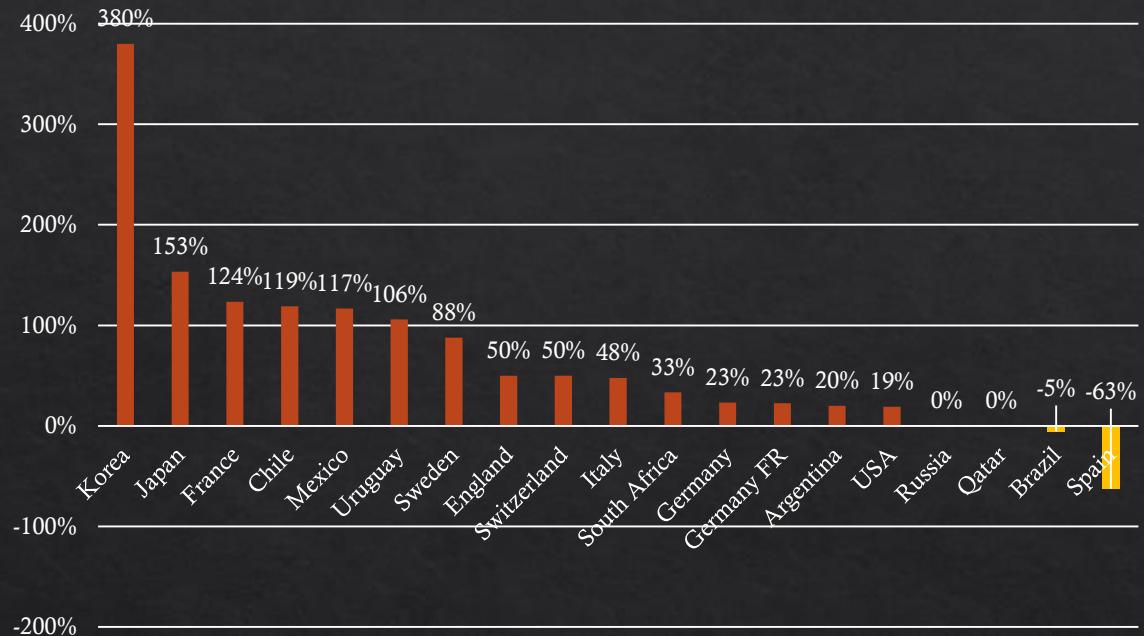


Percentage of winning in play-off matches

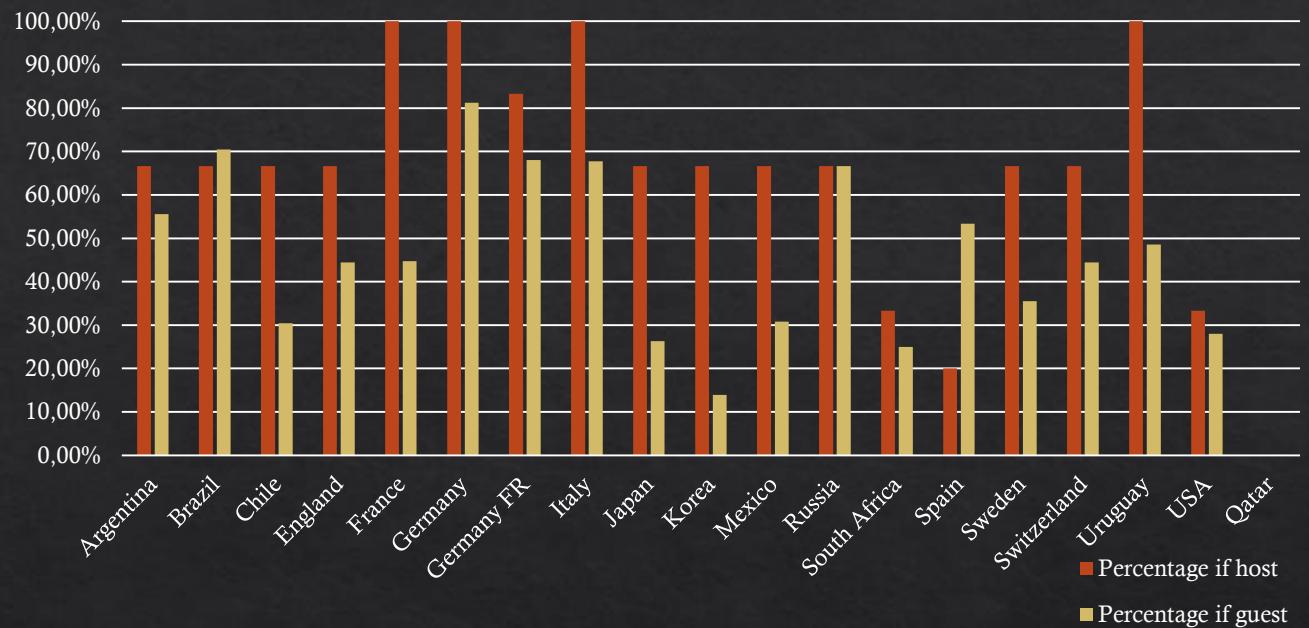


Q2. The influence of "his land" and fans

How many percent more group matches were won as host than as guest



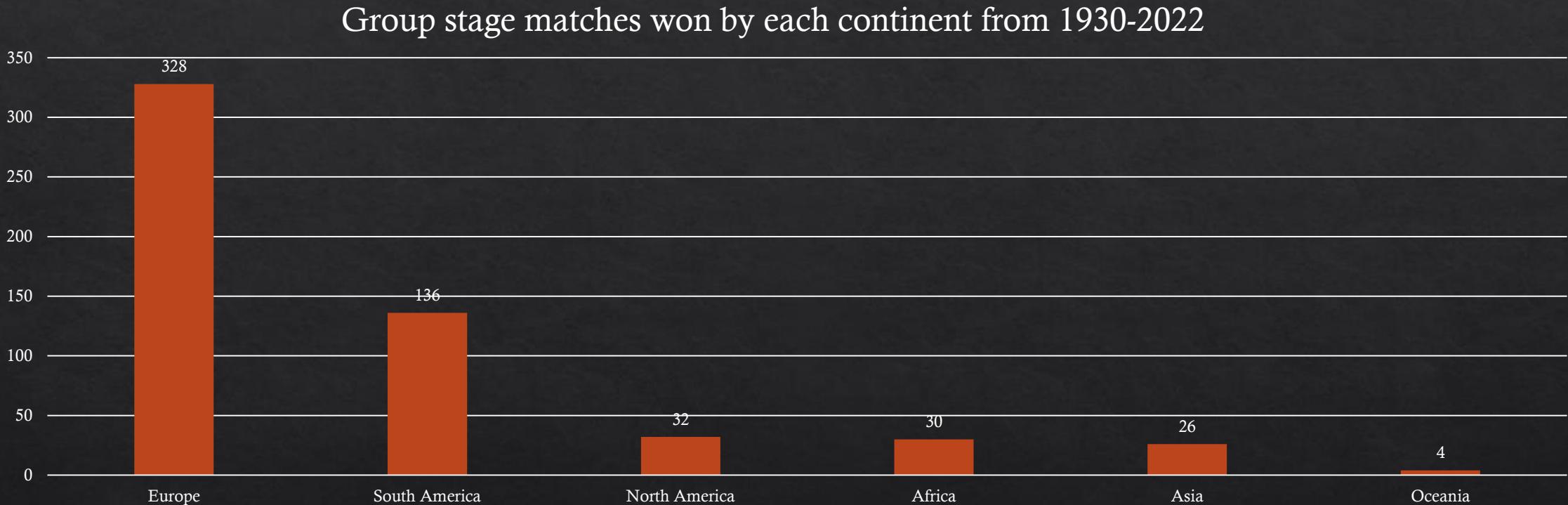
Percentage winning matches in group stage



The graph showing the percentage difference between group matches as home and as away shows how much impact playing at home has. The difference in Korea can be alarming - it is interesting to read the article about it:

<https://www.vice.com/en/article/ywgx4y/how-the-2002-world-cup-became-the-most-controversial-tournament-in-recent-memory>

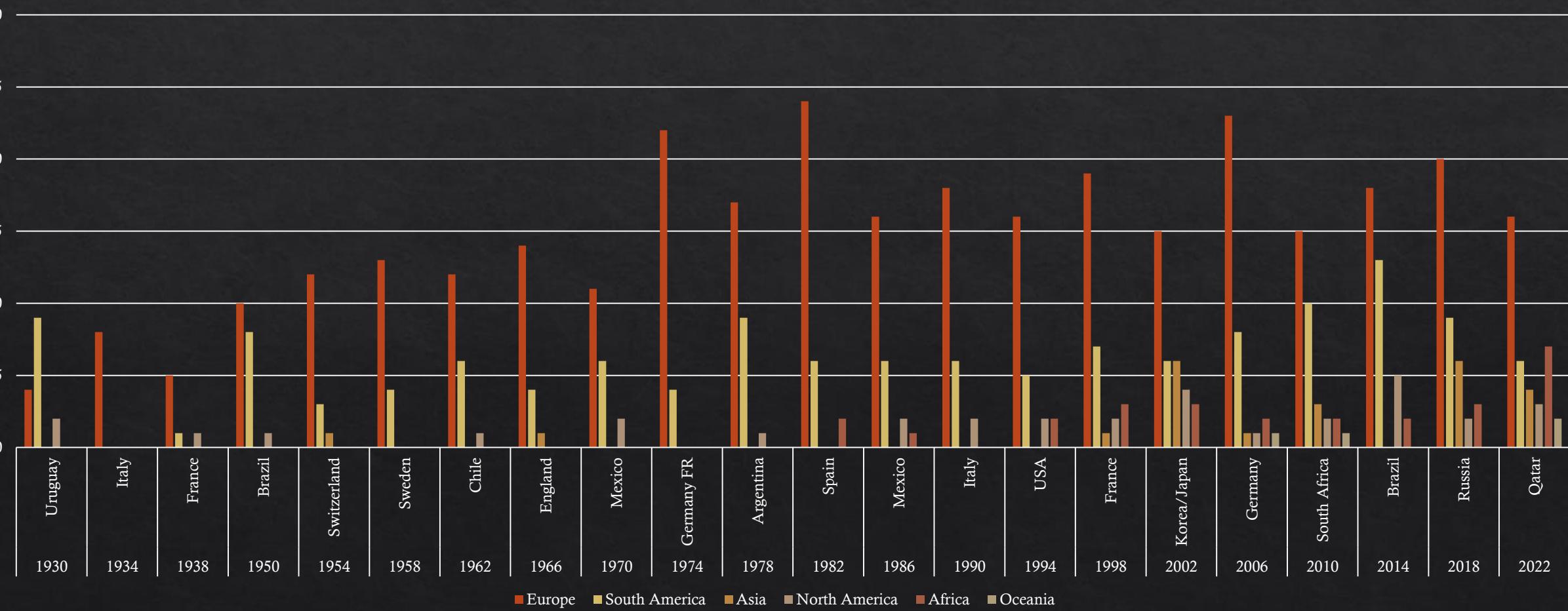
Q3. Europe and South America are the hegemonic of football?



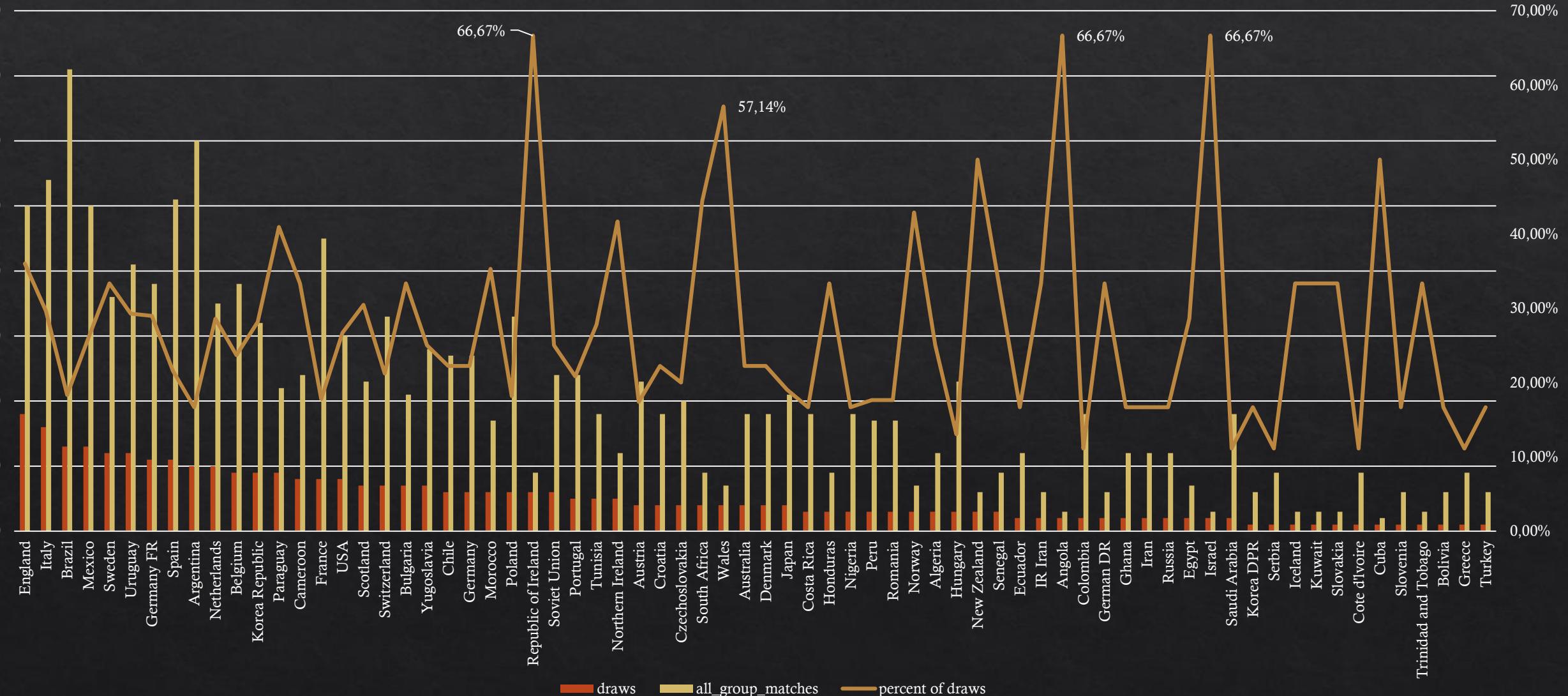
Europe and South America are unquestionably the two continents that have dominated football for less than a century. From 1934, these two continents led the group stage until 2022, when Africa had more games won than South America.

Q3. Europe and South America are the hegemonic of football?

Group stage matches won for each worldcup



Q4. Deffensive style of play



Results

1. **How many rounds were played in each World Cup tournament and how has the average number of goals per match changed over the years?**

Answer: The hypothesis was correct - the more matches in the tournament, the lower the average number of goals per match. Three things that I described in the report may be significant: "greater intensity of matches, which forces players to spread their strength better throughout the tournament; the level of refereeing and the rules that have changed over the years; sophistication of players - football has become more popular, so its level is constantly increasing, the entry threshold is higher and the level (between positions and teams) is more even.,,

2. **Does the fact that a given country is the host of the event and has fans on its side affect the result in a given championship?**

Answer: It turned out to be very significant. Only two countries (Brazil and Spain) from all the hosts of the World Cup played better on "not their territory".

3. **Which continents dominate when it comes to won matches and the entire championship?**

Answer: Here, too, the assumptions were correct. In one year alone - at the first World Cup in Uruguay in 1930 - Europe won fewer group matches than any other continent (South America). In the rest, Europe was the undisputed favourite. The last championship is also a change - because after 92 years of holding the first or second place in this classification, South America had to give way to second place to Africa, which made it third itself.

4. **Which countries play the most defensive football and had the most draws in the group stage?**

Answer: The answer to this hypothesis was completely different from reality. The team that has played in at least 3 World Cups and has the highest percentage of games drawn to games won is Ireland.