Robert Harakaly

The Python script `facts.py` show us different facts about the dataset of Soccer.

To launch this code, use the command `python3 facts.py`.

This dataset shows us all match between 1993 and 2022.

We can see the date, home team and the away team, where they are from, who win, the score and the score of players.

This is what we get when we launch the code:

```
(base) MacBook-Pro-de-Robert:src robert$ python3 facts.py

Team that scored the most goals at Home is: USA with 589 goals
Team that scored the most goals Away is: Brazil with 396 goals
Team that scored the most goals Overall is: Brazil with 954 goals

Team that have the highest goals per match at Home is: New Caledonia with 2.6785714285714284 goals
Team that have the highest goals per match Away is: Vanuatu with 2.314814814814815 goals
Team that have the highest goals per match overall is: New Caledonia with 2.2582046332046333 goals

The team who have to most clean sheet at Home is: USA with 152 clean sheets
This is 48.40764331210191 % clean sheets
The team who have to most clean sheet Away is: Zambia with 96 clean sheets
This is 39.50617283950617 % clean sheets
The team who have to most clean sheet overall is: Brazil with 229 clean sheets
This is 52.88683602771363 % clean sheets

The team at home have: 49.166004765687056 % chances to win
The away team have: 28.305672839764224 % chances to win
There is 22.528322394548724 % chances to have a draw

The team who won the most of match is: USA with 194 victory at Home which is 61.78343949044586 % victory
The team who won the most of match is: Brazil with 120 victory Away which is 60.0 % victory
The team who won the most of match is: Brazil with 301 victory Overall which is 69.5150115473441 % victory
```

1 Who scored the more goals?

In the first part, I wanted to know which team scored the more goals at Home, Away and Overall.

The USA scored 589 goals at Home, away it's Brazil with 396 goals and overall, the team who score the more goals is Brazil with 954 goals.

To find these results, I used Pandas with `groupby` to group all home and away teams and then I use `.sum()` to take all goals. Then I used the function `max()` to know who scored the more goals.

2 Who have the more goals per game?

In this part I wanted to know if the teams who scored the more goals scored because they score many goals pre matches. As we can see, it's not the case. It's New Caledonia who scores the more goals pre match at Home and Overall, with approximately 2.25 goals per match. This means that having the most goals do not imply that we have the most goals per matches.

For this fact, I used the same code as before for the part one but instead of `sum()` I used `mean()`.

3 Who have the Cleaner Sheet

A clean Sheet means that one team kept the opposing team from scoring a goal so 0 goal. I wanted to know if a good team who won games it's not only about scoring goals but also having a good defense. At Home the USA have more than 48% clean sheets for all their games. Away it's Zambia and overall, it's Brazil with more that 52% clean sheets.

For this part I filtered data to say that I only want data with teams who scored zero goal.

4 Who have the more chance to win between a team at Home or away?

For this part I was thinking that it can be interesting to know if the concept of game at Home or Away really can be important. As we can see, a team at home have approximately 1/2 to win a game! This is a huge advantage for the home team, the away team have 1/4 chances to lose or to have at least a draw.

So yes, playing at Home or playing Away can strongly impact the final score.

For this part I used the filtration that I want only to see the count of game that was won by a home team of not.

5 Who is the Best?

The part that everyone wants to know, who is the Best? Who won the more games? At Home it's USA who won more that 61% of their games but then Brazil is too strong Away and Overall. Thanks to the part 1 and 3, we can see something interesting that USA is the best team to score at Home and the best defense at Home, this explains easily why they win many games, and we can see the same things with Brazil, best team to score and best defense so this explains why they won many matches.

Here also I filtered data to have only the count of how much a team Won.