Summary

Task: Submit a short summary of the selected dataset for the group project. Describe the dataset as much as you can and try to find at least 5 interesting facts from the dataset and write a page of summary with the findings.

The five fun facts that I found from the dataset:

1. The home team that has the most victories is the USA

```
USA                          194
Mexico                       190
Brazil                       181
Saudi Arabia                 164
Egypt                        163
                             ...
Cook Islands                   2
Turks and Caicos Islands       2
US Virgin Islands              2
Anguilla                       1
San Marino                     1
Name: home_team, Length: 211, dtype: int64
```
   a.

2. The away team that has the most victories is Brazil

```
Brazil                       120
Spain                        101
Germany                      101
Zambia                        88
Argentina                     86
                             ...
Timor-Leste                    1
Anguilla                       1
American Samoa                 1
São Tomé and Príncipe          1
Andorra                        1
Name: away_team, Length: 209, dtype: int64
```
   a.

3. The continent that won the most games as the home team is Europe

```
Europe           3517
Africa           2942
Asia             2585
North America    1369
South America    1062
Oceania           286
```
   a.

4. The continent that won the most games as an away team is also Europe

```
  Europe          2435
  Africa          1513
  Asia            1286
  North America    700
  South America    657
  Oceania          180
  Name: away_team_continent, dtype: int64
```
a.

5. The city that has the most games is Doha

	index	counts
0	Doha	397
1	Bangkok	215
2	Muscat	212
3	Kuwait City	202
4	Abu Dhabi	191
...	...	...
1571	Dunedin	1
1572	Manisa	1
1573	Mondercange	1
1574	Veracruz	1
1575	Vitsebsk	1

1576 rows × 2 columns

a.

For now, the dataset has all the data from previous matches related to the FIFA world cup. It has its date, location, both team's name, continent, ranks, FIFA points presented, but I notice that in the last few columns, there are a lot of empty spaces for the matches before the year 2004, and I think we would have to consider those as null data and replace them with 0.