



INTERNATIONAL FOOTBALL PROJECT

This project is Part 1 of our analysis and demonstrates how we analyzed international football results using SQL and Power BI.



by Ahmed Youssef

INTRODUCTION

This project explores international football results data to uncover trends in hosting cities, team performance, and match outcomes using SQL and Power BI.

■ Main objectives

- 1- Which cities have hosted the largest number of international matches?**
- 2-Do home teams generally have a higher chance of winning?**
- 3-Which tournaments have the highest percentage of matches that end in a draw?**
- 4-Which teams score the most goals on average and which teams concede the most goals on average ?**
- 5-What are the matches with the highest total number of goals**



PART OF SQL



2. DATASET OVERVIEW

■ Data Source

This project explores international football results data to uncover trends in hosting cities, team performance, and match outcomes using SQL and Power BI.

■ Dataset Size (Rows & Columns)

The dataset contains over 40,000 match records across 9 main columns.

The image shows two side-by-side SQL query windows. The left window displays a single row of data with the column name 'NUMOFROWS' and the value '48673'. The right window shows a similar query with the same result. A red arrow points from the 'NUMOFROWS' value in the left window down to the column headers at the bottom of the image, which are enclosed in a red box. The column headers are: date, home_team, away_team, home_score, away_score, tournament, city, country, and neutral.

```
SELECT TOP 0 *
FROM results;
```

date	home_team	away_team	home_score	away_score	tournament	city	country	neutral
------	-----------	-----------	------------	------------	------------	------	---------	---------

```
--Display count of data
SELECT count(*) AS NUMOFROWS
FROM results
```

NUMOFROWS
48673

2. DATASET OVERVIEW

■ Dataset Features

- **date** - The date when the match was played
- **home_team** - The team playing at home
- **away_team** - The team playing away
- **home_score** - Number of goals scored by the home team
- **away_score** - Number of goals scored by the away team
- **tournament** - The name of the competition (e.g., World Cup qualification, Friendly, regional cups)
- **city** - The city where the match took place
- **country** - The country hosting the match
- **neutral** - Indicates whether the match was played on neutral ground (0 = not neutral, 1 = neutral)



2. DATASET OVERVIEW

■ Data Cleaning

-- checking for missing value

```
SELECT *  
FROM results  
WHERE date IS NULL  
    OR home_team IS NULL  
    OR away_team IS NULL  
    OR home_score IS NULL  
    OR away_score IS NULL  
    OR tournament IS NULL  
    OR city IS NULL  
    OR country IS NULL  
    OR neutral IS NULL;
```

The screenshot shows a SQL query being run in a database interface. The query is designed to check for missing values in various columns of the 'results' table. A red arrow points from the text 'Checking for Missing Values' to the WHERE clause of the query. Another red arrow points from the text 'No values are missing' to the empty result set in the 'Messages' tab.

Checking for Missing Values

No values are missing

Results Messages

date	home_team	away_team	home_score	away_score	tournament	city	country	neutral



2. DATASET OVERVIEW

■ Data Cleaning

```
--- checking for duplicates
SELECT date, home_team, away_team, home_score, away_score, tournament, city, country, neutral,
       COUNT(*) AS duplicates_count
FROM results
GROUP BY date, home_team, away_team, home_score, away_score, tournament, city, country, neutral
HAVING COUNT(*) > 1;
```

The screenshot shows a SQL query being run in a database environment. The query is designed to check for duplicate entries in the 'results' table based on specific columns. A red arrow points from the text 'Check for duplicates' to the WHERE clause of the query. Another red arrow points from the text 'No values are duplicates' to the results table header.

date	home_team	away_team	home_score	away_score	tournament	city	country	neutral	duplicates_count

Check for duplicates

No values are duplicates



GOALS

1- Which cities have hosted the largest number of international matches, and what might explain their popularity?

```
--1- Which cities have hosted the largest number of international matches?  
SELECT TOP 5 city ,  
       COUNT(city) AS MOSTCITY  
FROM results  
GROUP BY city  
ORDER BY MOSTCITY DESC ;
```

results Messages

city	MOSTCITY
Kuala Lumpur	737
Bangkok	579
Doha	552
London	441
Budapest	433



city	MOSTCITY
Kuala Lumpur	737
Bangkok	579
Doha	552
London	441
Budapest	433

Kuala Lumpur is the most active city in hosting international football matches because it is home to the Asian Football Confederation (AFC) and has major stadiums like Bukit Jalil that are used for many Asian tournaments.

GOALS

2- Do home teams generally have a higher chance of winning? What is the win percentage for home teams overall?

```
--2-Do home teams generally have a higher chance of winning?  
SELECT  
    (SUM(CASE WHEN home_score > away_score THEN 1 ELSE 0 END) * 100.0  
     / COUNT(*)) AS HomeTeamWinPercentage,  
  
    (SUM(CASE WHEN home_score = away_score THEN 1 ELSE 0 END) * 100.0  
     / COUNT(*)) AS DrawPercentage,  
    (SUM (CASE WHEN home_score< away_score THEN 1 ELSE 0 END ) * 100.0  
     /COUNT(*)) AS AwayTeamWinPercentage  
FROM results;
```

The home team's winning percentage =49.03 % The Draw percentage =22.72 % The away team's winning percentage =28.23 %

HomeTeamWinPercentage	DrawPercentage	AwayTeamWinPercentage
49.039508557105	22.725124812524	28.235366630370

A red arrow points from the SQL query area to the first row of the results table, highlighting the HomeTeamWinPercentage value.

Yes, home teams generally have a higher chance of winning. The overall home win percentage is 49.03%.

GOALS

3- Which tournaments have the highest percentage of matches that end in a draw? What does this say about the competitiveness of these tournaments?

```
--3- Which tournaments have the highest  
--percentage of matches that end in a draw?  
  
SELECT TOP 1 tournament , COUNT (*) AS CountdrawperTournament ,  
CONCAT(  
    (COUNT(*) * 100 /  
     (SELECT COUNT(*) FROM results WHERE home_score = away_score))  
,  
    '%'  
) AS Percentage  
FROM results  
WHERE home_score = away_score  
GROUP BY tournament  
ORDER BY CountdrawperTournament DESC;
```

tournament	CountdrawperTournament	Percentage
Friendly	4559	41%

The tournament with the highest percentage of matches ending in a draw is Friendly matches, with 41% of games drawn. This suggests that Friendly matches are generally less competitive, as teams may experiment with lineups and strategies rather than playing to win at all costs.

GOALS

4-Which teams score the most goals on average?

```
--4-Which teams score the most goals on average?  
SELECT TOP 5 team,  
       AVG(goals_scored) AS avg_goals_scored  
FROM (  
    SELECT home_team AS team, home_score AS goals_scored  
    FROM results  
    UNION ALL  
    SELECT away_team AS team, away_score AS goals_scored  
    FROM results  
) AS all_goals  
GROUP BY team  
HAVING AVG(goals_scored) > (SELECT AVG(away_score + home_score ) FROM results)  
ORDER BY avg_goals_scored DESC;
```

Results

team	avg_goals_scored
Elba Island	5
Surrey	3
Cascadia	3
Isle of Man	3
Occitania	3

team avg_goals_scored

team	avg_goals_scored
1 Elba Island	5
2 Surrey	3
3 Cascadia	3
4 Isle of Man	3
5 Occitania	3



Elba Island has the highest average goals scored per match (5 goals per game). Surrey, Cascadia, Isle of Man, and Occitania follow with an average of 3 goals per match, showing that these teams are the strongest offensively.

GOALS

5-which teams concede the most goals on average?

```
-- which teams concede the most goals on average?  
SELECT TOP 5 team,  
       AVG(goals_conceded) AS avg_conceded  
FROM (  
    SELECT home_team AS team, away_score AS goals_conceded  
      FROM results  
  
    UNION ALL  
  
    SELECT away_team AS team, home_score AS goals_conceded  
      FROM results  
) AS all_conceded  
GROUP BY team  
HAVING AVG(goals_conceded) > (SELECT AVG(away_score+home_score) FROM results)  
ORDER BY avg_conceded DESC;
```



team	avg_conceded
Sark	17
Niue	16
Darfur	14
Kiribati	11
Saint Pierre and Miquelon	11

	team	avg_conceded
1	Sark	17
2	Niue	16
3	Darfur	14
4	Kiribati	11
5	Saint Pierre and Miquelon	11

Sark concedes the most goals on average (17 goals per match). Niue and Darfur also concede very high averages, followed by Kiribati and Saint Pierre and Miquelon, indicating that these teams have the weakest defensive performance.

GOALS

5-which teams concede the most goals on average?

```
-- What are the matches with the highest total number of goals,  
-- and what patterns can be observed from them?  
  
SELECT TOP 5 home_team ,  
       away_team ,  
       tournament ,  
       city ,  
       country ,  
       (away_score + home_score)AS Match_score  
FROM results  
ORDER BY Match_score DESC ;
```

	home_team	away_team	tournament	city	country	Match_score
1	Australia	American Samoa	FIFA World Cup qualification	Coffs Harbour	Australia	31
2	Tahiti	Cook Islands	South Pacific Games	Papeete	Tahiti	30
3	Fiji	Kiribati	South Pacific Games	Nausori	Fiji	24
4	Australia	Tonga	FIFA World Cup qualification	Coffs Harbour	Australia	22
5	Sápmi	Monaco	Viva World Cup	Hyères	France	22

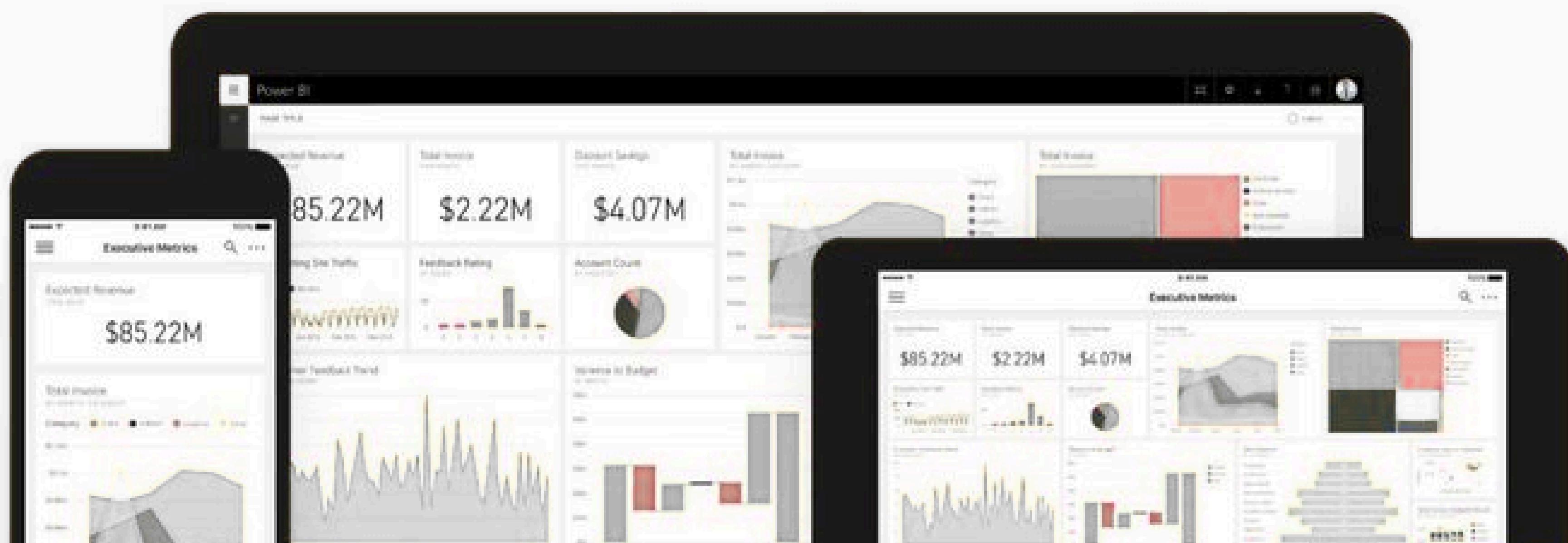
Results						
home_team	away_team	tournament	city	country	Match_score	
Australia	American Samoa	FIFA World Cup qualification	Coffs Harbour	Australia	31	
Tahiti	Cook Islands	South Pacific Games	Papeete	Tahiti	30	
Fiji	Kiribati	South Pacific Games	Nausori	Fiji	24	
Australia	Tonga	FIFA World Cup qualification	Coffs Harbour	Australia	22	
Sápmi	Monaco	Viva World Cup	Hyères	France	22	



The matches with the highest total goals all involve a large mismatch in team strength, leading to very high scores. Most of these games come from regional or less competitive tournaments like the South Pacific Games and Viva World Cup. Teams from Oceania appear frequently, where strength gaps are big. Also, several matches were played on the home ground of the stronger team, which may have contributed to the large score differences.

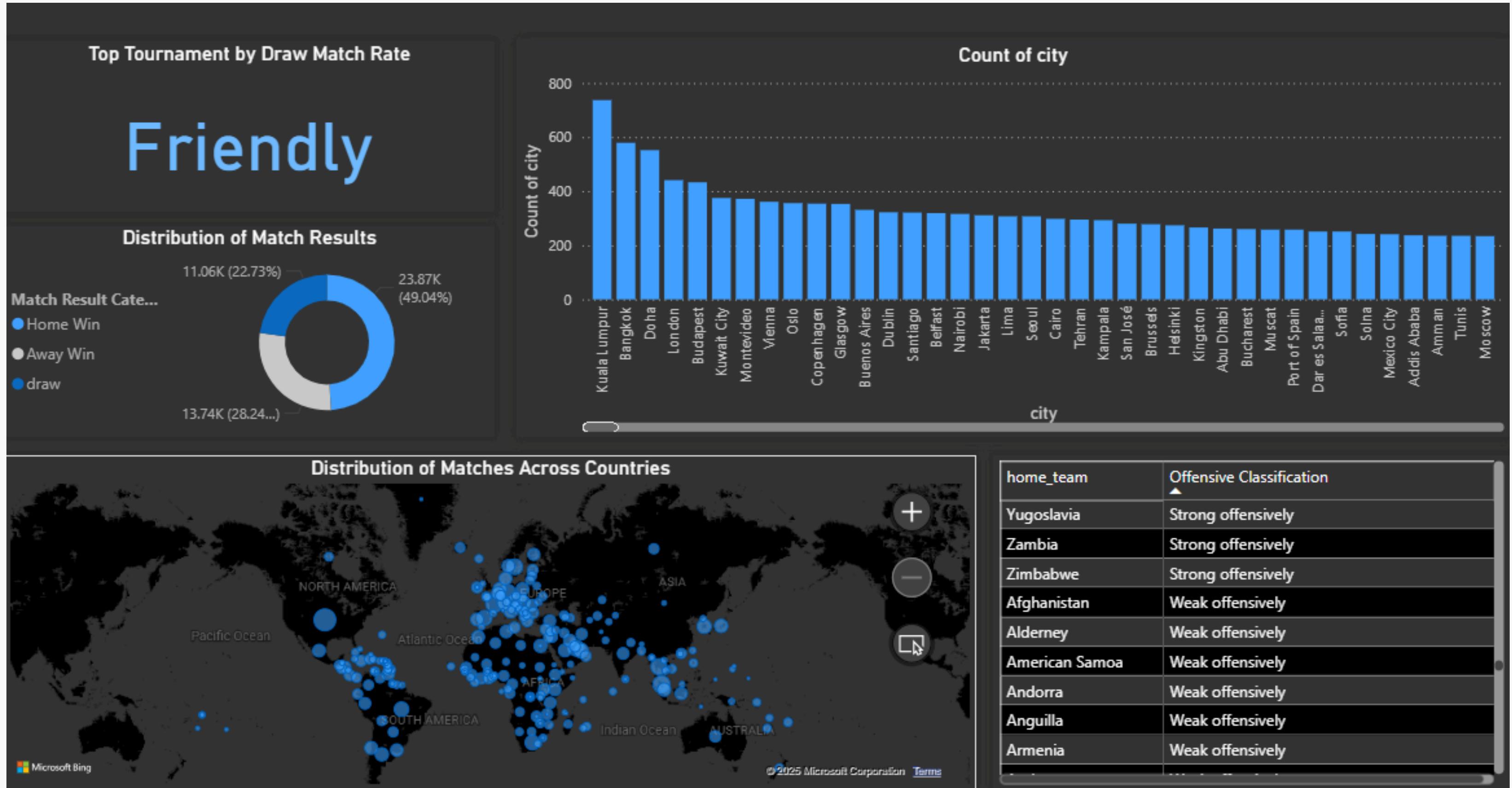


PART OF POWER BI

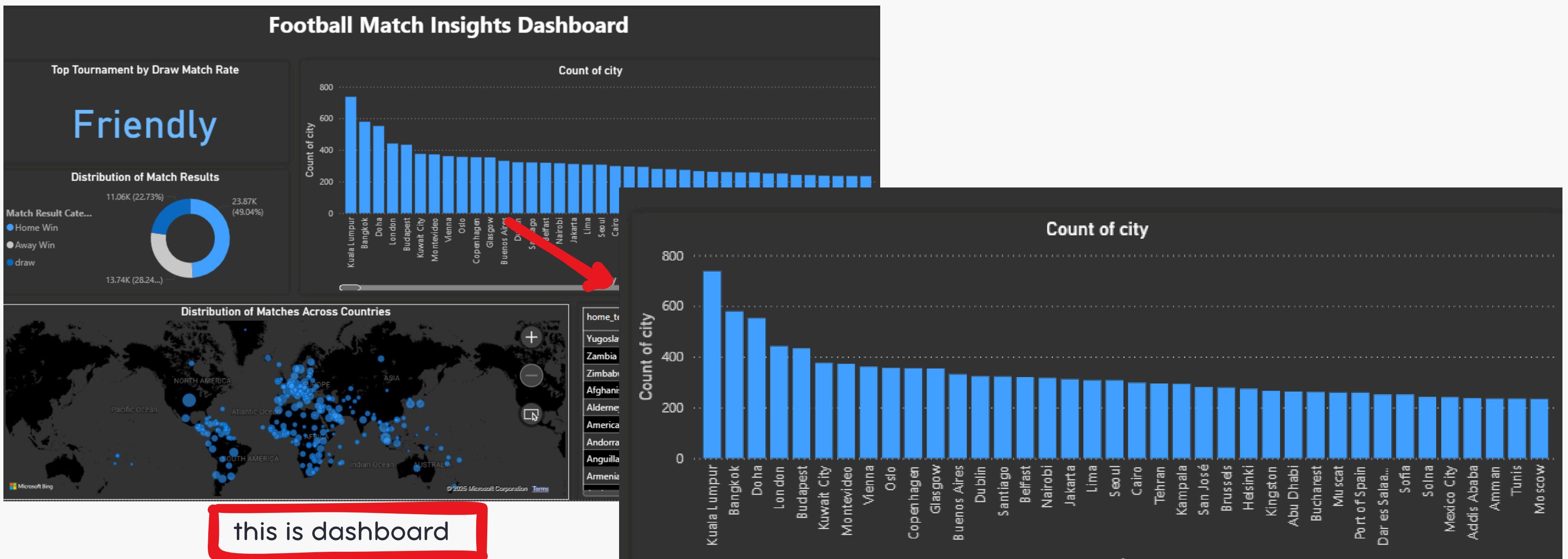


Football Analytics Dashboard

(Built with Power BI)

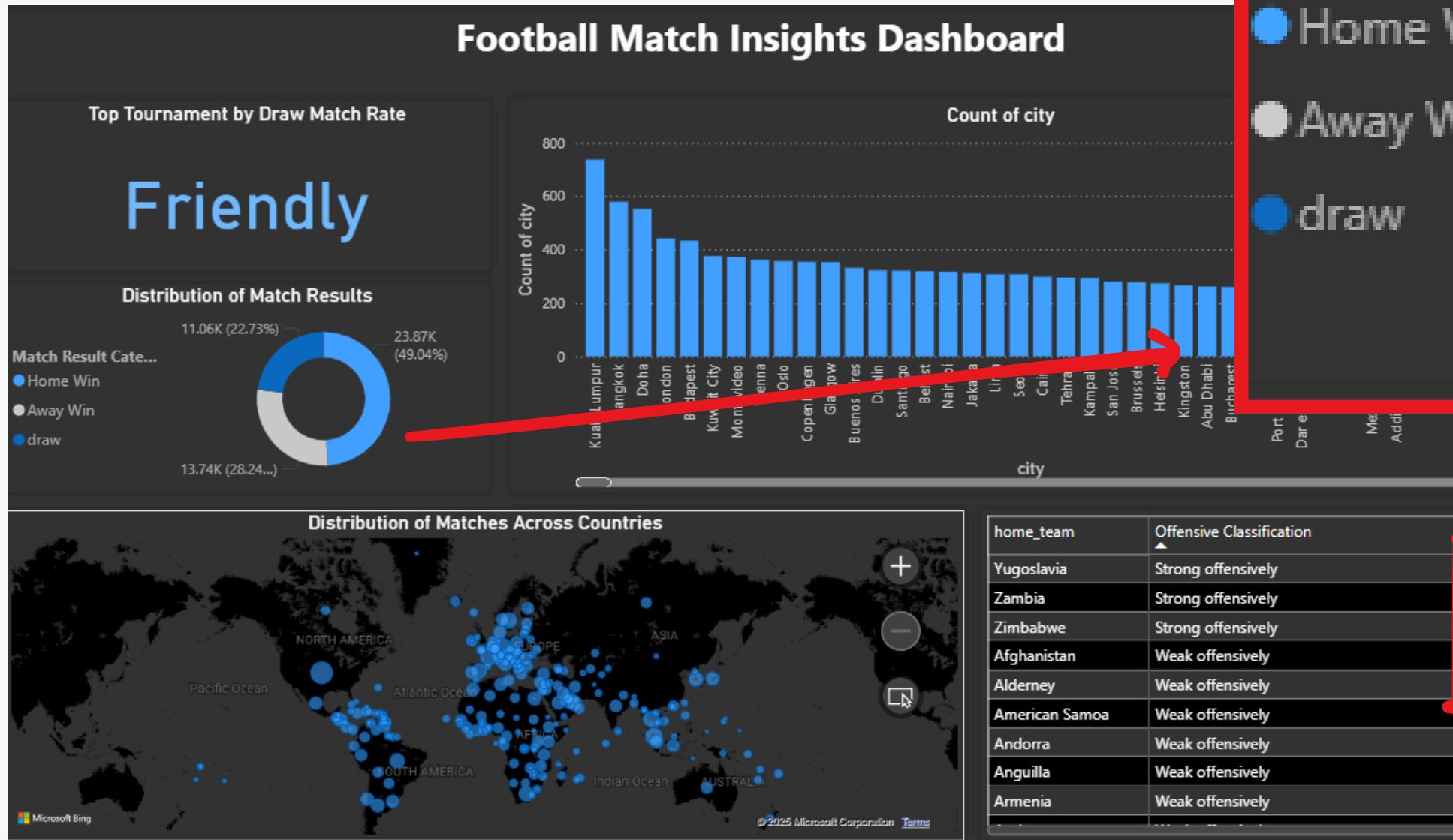


1-Which cities have hosted the largest number of international matches ?

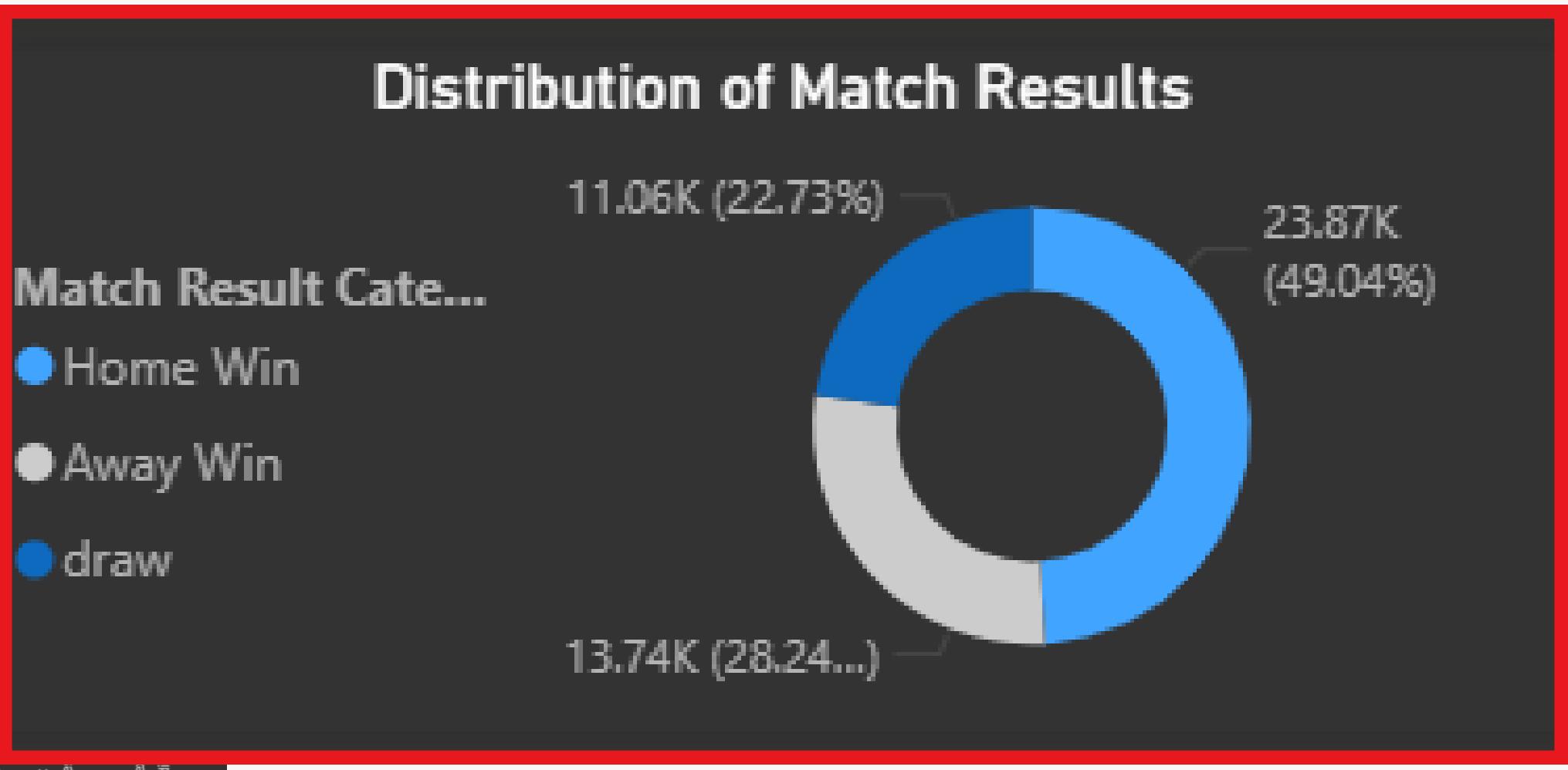


This section of the dashboard shows the cities that have hosted the tournament the most.

2-Do home teams generally have a higher chance of winning?

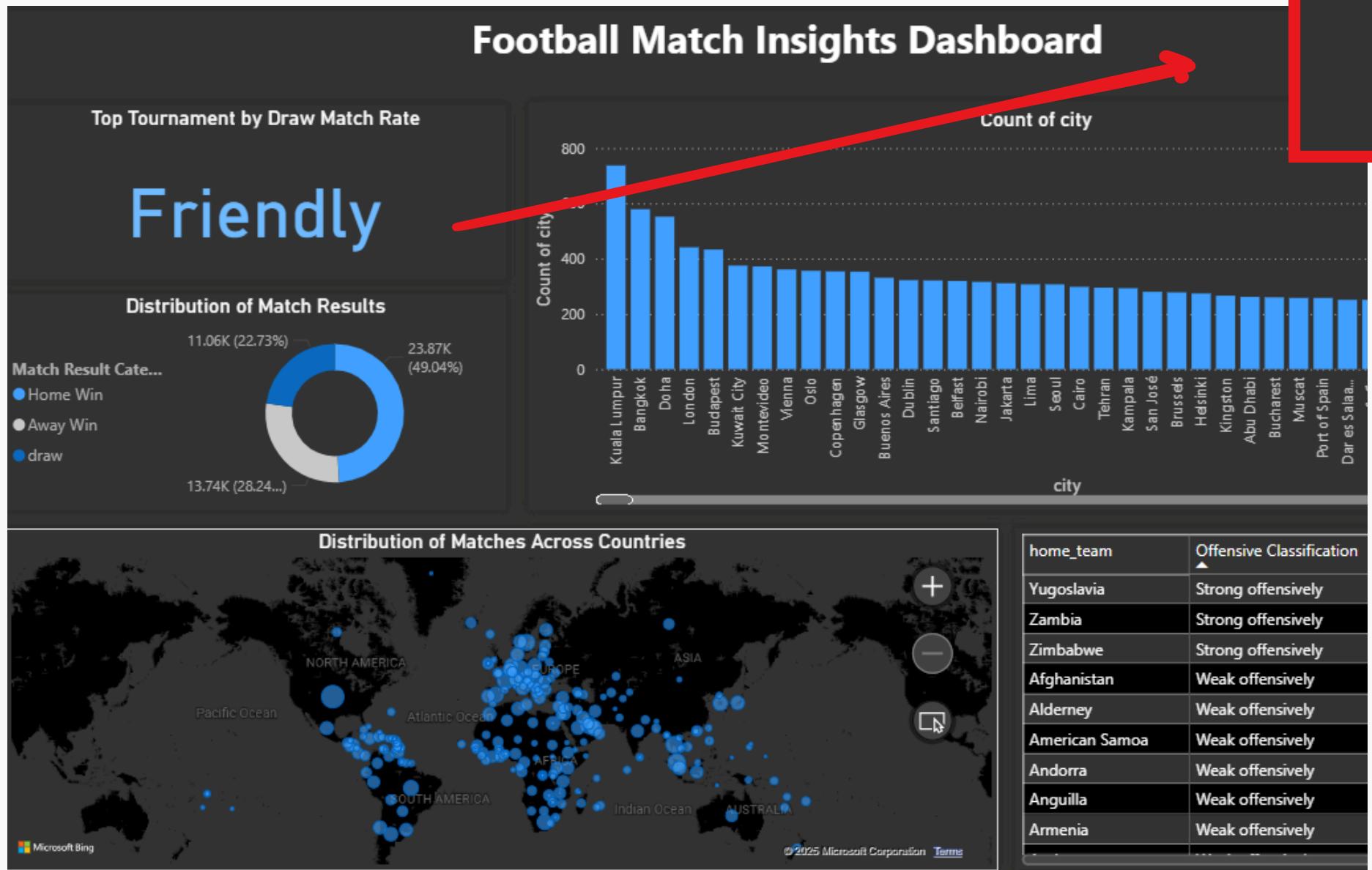


This is the dashboard



This section of the dashboard shows the distribution of match results.

3-Which tournaments have the highest percentage of matches that end in a draw?



this is dashboard

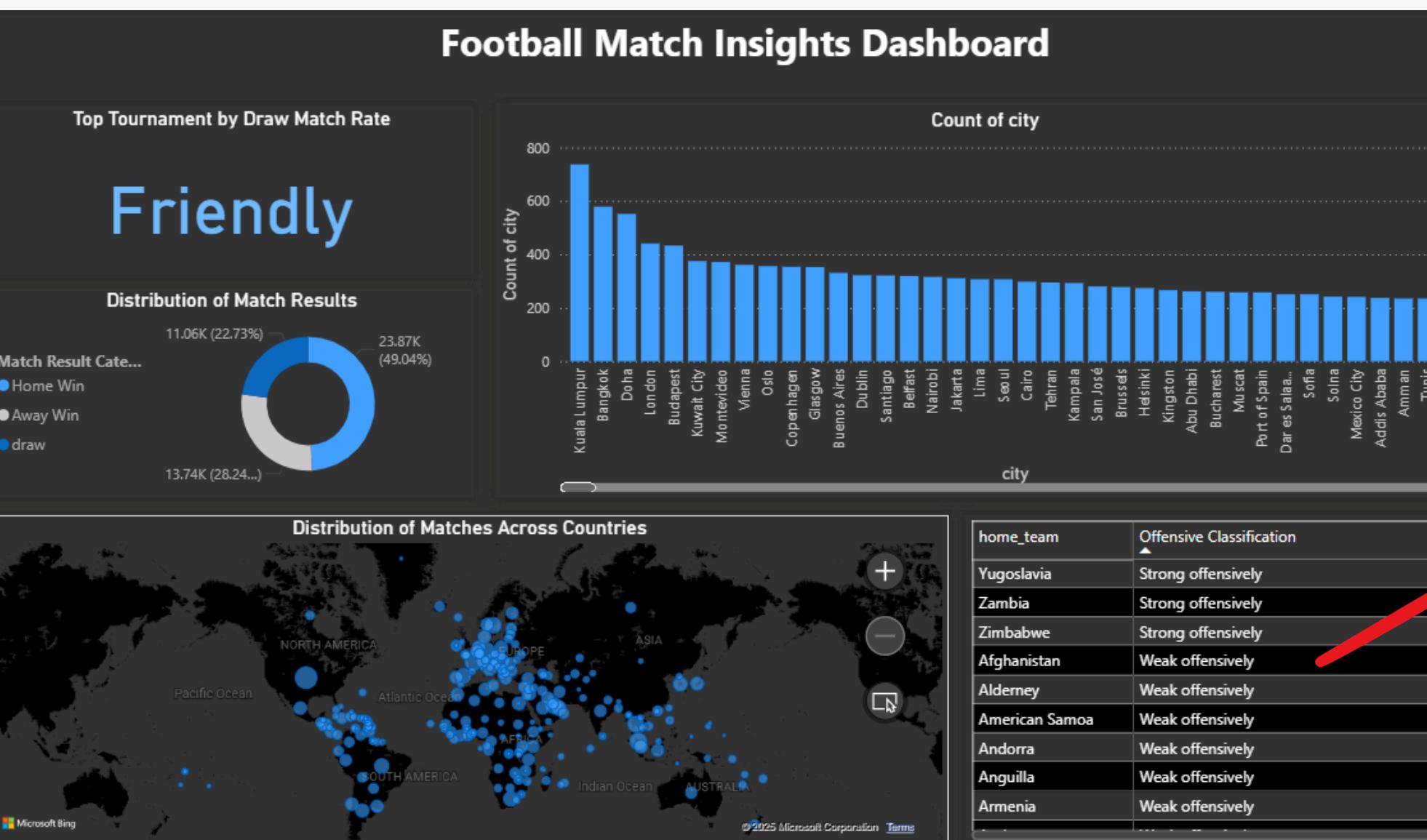
Top Tournament by Draw Match Rate

Friendly

This section of the dashboard shows the Top Tournaments by Draw Match Rate.

home_team	Offensive Classification
Yugoslavia	Strong offensively
Zambia	Strong offensively
Zimbabwe	Strong offensively
Afghanistan	Weak offensively
Alderney	Weak offensively
American Samoa	Weak offensively
Andorra	Weak offensively
Anguilla	Weak offensively
Armenia	Weak offensively

3-Which teams score the most goals on average ?



home_team	Offensive Classification
Yugoslavia	Strong offensively
Zambia	Strong offensively
Zimbabwe	Strong offensively
Afghanistan	Weak offensively
Alderney	Weak offensively
American Samoa	Weak offensively
Andorra	Weak offensively
Anguilla	Weak offensively
Armenia	Weak offensively

This section of the dashboard shows the teams compared to the overall average goals. Teams with an average above the overall average are stronger offensively, while teams below the overall average are weaker.

this is dashboard

THANK YOU