EUROPEAN SOCCER

Database& Analysis

DATE:
03/12/2020

PREPARED BY:

Siwei Tang, Zhenyu Wang

Domain

– Soccer Club

– Time: 2010-2019

– Competitions:

– Team:

– League Rank, Match Results...

For more details, please check the design report.

Question

Country's performance as a whole

– Which country has a better soccer club performance at European level competitions in the past 10 years?
Do Europa League and the Champions League show similar or different results?

R e s u l t s

Country's performance as a whole

- Both competitions show similar results:

- Spanish teams dominate competitions (winning rate: 55% in the CL, 57% in the EL)

- England and Germany are too close to call (51%, 51% in the CL, 54%, 52% in the EL)

- Italy has better performance than France (44%, 39% in the CL, 47%, 39% in the EL)

Question

Interesting facts about top teams

- Does top-class soccer clubs' performances in their own country's league relate to its performance in the Champions League in the past 10 years?

- Top-class soccer clubs are defined as clubs that attended at least the Champions League for 7 times in the past 10 years?

Results

Interesting facts about outlier teams

- 3 teams dominate both continental and domestic competitions

- Several teams played better in their own countries than the Champions League
(Win Rate in CL: 52%, Avg League Rank: 1.0)

- Arsenal is an outlier
(Win Rate in CL: 55%, Avg League Rank: 4.4)

- There is no strong relationship between domestic and European competitions

Question

Special factors that influence the performance

- Is there a 'Natural Enemy' for clubs in one country to another country (That is, clubs in one country beat clubs in another country for most times) in past 10 years?

- Did teams play at home always have a stronger performance than away teams for different countries?

R e s u l t s

Special factors that influence the performance

- 'Natural Enemy': Spain vs. Italy
(Home win rate: 73%,
Away Unbeaten Rate: 72%)

- Home teams often have better performance
than away teams.

Challenges&Lessons

Database Design: 3 approaches to store domestic league data.
- (Year, League, Team)
- (Year, Team), (Year, Team), ...
- (Team), (Team), ...

Design Trade-off
The efficiency of importing the data
Future Extention
Readability

Consequences
When we want to use five leagues together?

Challenges&Lessons

Key of the Data
- Did not use a unique ID for each match

Design Trade-off
Less data preparation (Pandas)
Future Extention

Consequences
If we create another relation detailed
information about a match?

Important Lesson
Prepare for future extention

Questions?

THANK YOU