



# Applying MFAS to rank football teams based on UCL performance

Applying MFAS to rank football teams is an idea I have had in mind since the moment I have read about MFAS.

Currently, most European leagues and even the group stages in the champions league follow a round-robin based scoring method where each team plays the other twice : once home and then away. The winner of a match gets three points, in case of a tie both teams get one point. In the standard procedure, the final ranking of the teams in the championship is made up by adding these points.

In the UCL, the number of matches played by the teams depend on the round which they are able to progress to. Hence, applying a point-based system to draw conclusions regarding a team's performance in the UCL over a span of 10 years does not make sense.

Also, not all teams are able to participate in every edition of the tournament. Since the format isn't round robin among all the participating teams in a particular edition, it is difficult to compare 2 teams who haven't played each other.

Also, most of the times, two best teams A and B meet earlier than the final and in such a situation, saying a team C (which reached the final) is better than the losing team among A and B is simply not fair.

**Also, ideally, winning against a top team should count more than beating an average team while judging the performance of a team.**

MFAS focusses on producing an ordering so as to minimize the sum of back-edges in the graph (whose structure I have defined below).

## Input format and representing the data

The dataset considers Champions League matches (excluding the preliminary qualifying matches). The dataset has been extracted from here.

openfootball/europe-champions-league

A free open public domain football database & schema for use in any (programming) language e.g. uses datasets in (structured) text using the football.txt format. More football.db Project Site " Free open public domain football datasets for Europe.

 <https://github.com/openfootball/europe-champions-league>



There are  $n = 102$  teams which have played in the Champions league from **2005-2006 season to the 2015-2016 season**. However, while some teams are regulars in the tournament, others don't make it every year. The input data is parsed so as to create a  $(n * n)$  matrix  $H$  where the  $H_{ij}$  entry is the sum of all margins (goal difference) with which team  $i$  has won across all matches against team  $j$ . **Since a team plays atmost**

**(3+1+1+1+1) 7 opponents in a single season**, it is quite natural that most entries in this matrix are zero even though the data is being considered over a 10-year span.

I treated the  $(n * n)$  matrix as a graph with 'n' vertices and placed an edge between two vertices  $i,j$  if both of them played atleast one match where the  $i^{th}$  team has won (the edge weight being the sum of goal differences). So, the aim is to obtain a linear ordering of all vertices so as to minimize the sum of the back edges.

## Preliminary results and corrections based on them

After including all 102 teams in the graph and trying to find best possible ordering after applying **Chanas and kobiyancki heuristic** some unexpected results were obtained.

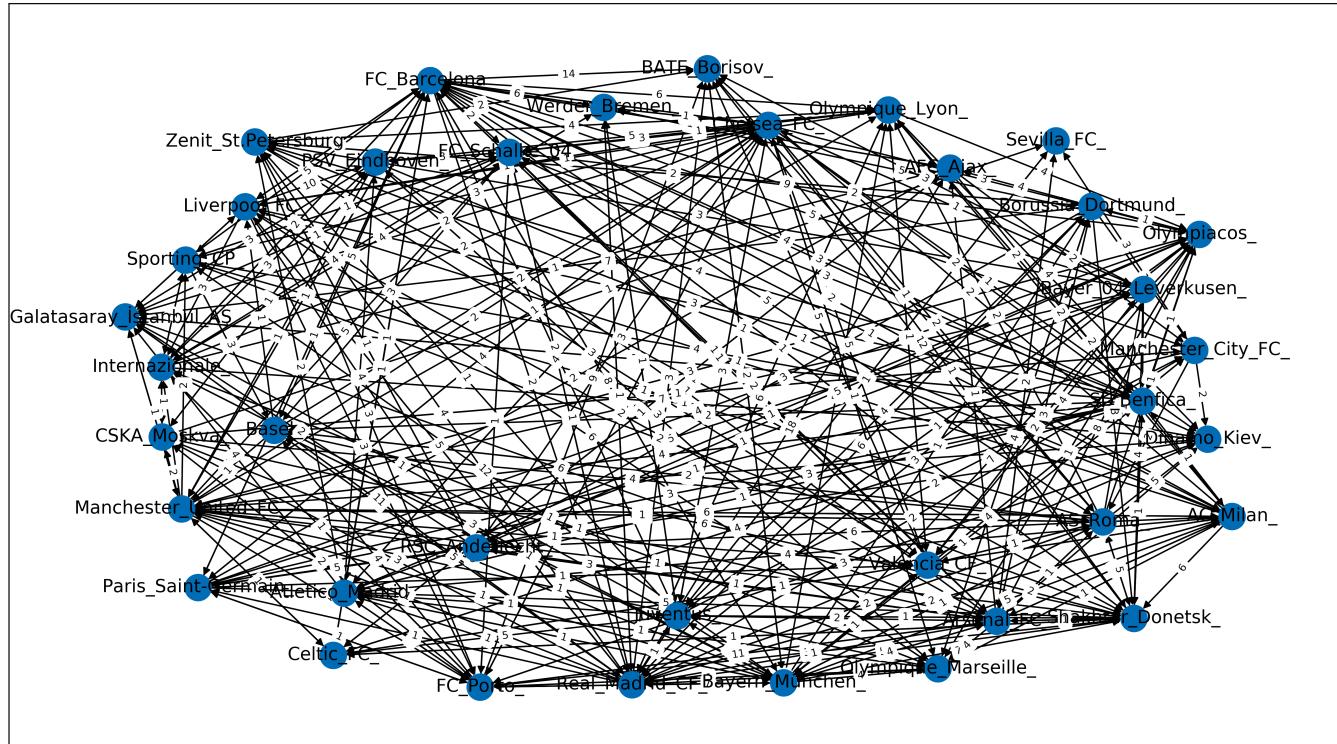
The preliminary rankings obtained can be found here :

[https://s3-us-west-2.amazonaws.com/secure.notion-static.com/a3337ddc-82e9-4419-abd4-379eb2376024/ranking\\_all.txt](https://s3-us-west-2.amazonaws.com/secure.notion-static.com/a3337ddc-82e9-4419-abd4-379eb2376024/ranking_all.txt)

Some anomalies were teams like **Rosenborg\_BK\_ and AS\_Roma\_being placed higher than Internazionale\_, the UCL winner of the 2009-2010 season**. Throughout the ordering, several such instances were found. However, there is a simple explanation for this. Imagine a team which has played in the UCL in just one of the above 10 seasons. Considering that the team went out in the group stages, it has played just 6 matches. In a span of 6 matches, it is possible that the due to draws/smaller win-lose margins, the weights of the edges connecting this team is small in magnitude and hence, placing this team almost anywhere in the ordering does not impact the optimisation function to a great extent.

## Final results

In order to take care of such anomalies, I filtered the dataset to consider matches only between those teams which have played in atleast 20 matches across the season. This ensured that most teams have played enough matches to be able to draw a conclusive verdict by considering the in and out edges.



I found that just 37 teams fulfill this criteria and the ordering obtained after considering just these teams was as follows:

**Bayern\_München\_**  
**FC\_Barcelona\_**  
**Juventus\_**

**Borussia\_Dortmund\_**  
**Real\_Madrid\_CF\_**  
Paris\_Saint-Germain\_  
Manchester\_United\_FC\_  
Chelsea\_FC\_  
Bayer\_04\_Leverkusen\_  
Zenit\_St.Petersburg  
Sporting\_CP\_  
FC\_Schalke\_04\_  
Olympique\_Lyon\_  
Valencia\_CF\_  
SL\_Benfica\_  
Liverpool\_FC\_  
Arsenal\_FC\_  
FC\_Porto\_  
Internazionale\_  
AFC\_Ajax\_  
Atletico\_Madrid\_  
PSV\_Eindhoven\_  
AC\_Milan\_  
Shakhtar\_Donetsk\_  
BATE\_Borisov\_  
Basel\_  
Manchester\_City\_FC\_  
AS\_Roma\_  
CSKA\_Moskva\_  
Sevilla\_FC\_  
Olympique\_Marseille\_  
Olympiacos\_  
RSC\_Anderlecht\_  
Galatasaray\_İstanbul\_AŞ\_  
Celtic\_FC\_  
Dinamo\_Kiev\_  
Werder\_Bremen\_

The ordering looks satisfactory at first glance. The top 5 teams have all either won the final or appeared in the final and have reached the latter stages of the tournament consistently. Since goal difference plays a major role in the edge weights, it would explain via Bayern Munich consistently ranks above Barcelona (😢) in all possible randomisations I did (the 7-0 aggregate loss really cost us our place at the top).

**PS:** The datasets consists of matches from the **2005-2006 season to the 2015-2016** season, So, any significant changes in football royalty hasn't yet reflected here.