

An exploration to...



European Professional Football

**Zhenyu Zhu
Yueqing Shen
Jingran Su**

COM-480 Data Visualization

Contents

1. Introduction	1
1.1. European Professional Football	1
1.2. Our consideration and motivation	1
1.3. The main implementation of this project	2
2. Challenges and Solutions	3
2.1. Dataset processing and supplement	3
2.2. Main challenges and our solutions	3
2.3. Website building and implementation	4
3. Design and Implementation	5
3.1. Ideas and design	5
3.1.1. themes of our website	5
3.1.2. main function of our website	5
3.1.3. little ingenuity	6
3.2. Further improvement on preliminary proposal	6
4. Peer assessment	8

A soccer player in a white jersey is helping a player in a blue and yellow jersey up from the ground. The player in white is leaning over, holding the hand of the player in blue and yellow, who is sitting on the grass. The background is a blurred stadium with many spectators.

1. Introduction

1.1. European Professional Football

The European Professional Football events are national football competitions run by domestic leagues. Unlike the UEFA European Football Championship or UEFA Super Cup that are "competition-oriented" and involve football games between countries, the target of European Professional Football is to run domestic competitions to help their member clubs perform at the highest sporting level in a sustainable and competitively balanced environment.

The European Professional Football events are regulated by the European Professional Football Leagues (EPFL) which was founded in 1997 and now has 29 member countries: Austria, Belgium, Bulgaria, Denmark, England, Finland, France, Germany, Greece, Israel, Italy, Kazakhstan, the Netherlands, Norway, Poland, Portugal, Russia, Scotland, Slovenia, Spain, Sweden, Switzerland, Romania, Latvia, Ukraine, and Serbia, which has a big influence on European football field.

1.2. Our consideration and motivation

Though the European Professional Football games are not inter-country, we can still obtain vital information of teams' comprehensive competitiveness on a domestic level. For instance, the list of promoted and relegated clubs each year can be reflection of the rise and fall of a team and indicator of great events or even revolution in a country's sports industry. More importantly, from the perspective of some local football lovers, these games can be an evaluation of the competitiveness of domestic clubs.

Luckily, for contemporary football fans, the data about soccer matches are at their fingertips thanks to the high-speed internet. However, despite the fact that the European Professional Football games are of great interest and concern, there are few integrated platforms for football lovers to explore and most of the data are scattered on different websites.

Thus, in this project, our purpose is to develop a clear and comprehensive platform for soccer fans to discover the historical events and information of their interest interactively and more conveniently. Here we are dedicated to creating an integrated platform by collecting and visualizing the data of European soccer events from season 2008/2009 to season 2015/2016. Hopefully, this platform can help users to explore the European Professional Football games in a more convenient and decent way.

1.3. The main implementation of this project

The whole story started from a good and clean dataset on Kaggle: *European Soccer Database*, which contains comprehensive data of players, matches and detailed match events such as goal types, possession for over 10k matches from season 2008/2009 to season 2015/2016.

The data includes approximately 26.0k entries corresponding to the detailed information of each soccer event of European Professional Football ranging from season 2008/2009 to season 2015/2016, which provides us with abundant data to show on our website. The structure of our website is proposed to contain three layers where there are mainly two maps and two statistical table showing all the information.

Moreover, in order to provide figures to better illustrate the project, we also implemented a mini "crawler" task by python to obtain the needed pictures from wikipedia to give the users a better experience. In the website design, we used d3.js to build two maps for exploration and also added some useful creative elements to facilitate the whole exploration. These will be further explained in **section 2**.



2. Challenges and Solutions

2.1. Dataset processing and supplement

As we mentioned before the dataset we used is the *European Soccer Database*. Subsets we extracted from the datasets were "match" including country, league, season, stage, date, the name and attributed ID of home team and away team, goals of home team and away team and "player" including name, birthday, weight, height, FIFA score, preferred foot and a series of physical quality data. The dataset was clean and ready for data exploration and results visualization.

However, in order to build the whole website as we proposed, we also need the json files of the 11 participating countries, the coordinates of different teams and the pictures of team logo. We also wanted to construct a block showing the star player of each team in each season, and to present the players to the users, the figures of around 1000 players should also be collected and attributed to our existing dataset.

For these supplementary data to perfect our dataset, we mainly used python to do the crawler of these missing data from wikipedia(including the coordinates value, the logo of teams and the pictures of players) and the json files were collected from an online json file generator. When collecting the pictures, another issue was the various format, we used an online converter to unify the pictures to png, which was fit to the website design.

2.2. Main challenges and our solutions

The first challenge here is how to build the map in a clearer and more decent way since some teams are in quite near proximity and may overlap on a small point while some clubs are located even in outlying islands. Hence, the nice scaling of the maps is necessary.

The second problem is the radar chart. There are many JS files about radar charts on the Internet, but they failed to correctly fit to our data. To solve this problem and construct a decent radar chart, we simply drew a radar chart ourselves with lines, circles, and text elements spliced together.

2.3. Website building and implementation

In the process of building the website, the first problem we encountered was how to achieve accurate and fast switching when displaying data from different countries, seasons and teams.

We split the website into five modules: Europe map, country map, season timeline, ranking table, team and player information. And each module is constructed by a function with parameters (for example, input the season and country to get a ranking table). Each time the user clicks, these functions are called and the corresponding information is output. And for the team and player information module, we have designed a variety of construction methods. When only the country information is input, it will display this country's league championship information in the 2008/2009 season. When input country information and season information, it displays the country's league championship information in the input season. When entering country information, season information, and team information, it displays the current season. Information about the team in the country's league. The above settings and information delivery methods enable our website to display data in a correct way.



3. Design and Implementation

3.1. Ideas and design

Since the project is designed to facilitate users' exploration to European Professional Football games, as we proposed in the milestone 1 and milestone 2, the website was constructed from three points of view: league (geographical and historical) perspective, team perspective and player perspective. From these different sights, users can really have an insightful discovery when exploring the data. For instance, the rise and fall of the teams can be indicators of interesting historical events or great revolution of a country. In this section, ideas and design are illustrated with some sketches and figures.

3.1.1. theme of our website

Different from the mainstream sports websites that are colorfull and content-intensive, our web page is designed to facilitate a visually comfortable exploration to our dataset with clearly shown figures and tables. Moreover, in order to add more fun, we used a concise and warm theme.

3.1.2. main function of our website

The visualization components are listed as follows: a "global map" of 11 participating regions, 11 individual maps of each region, a score table, a selectable season timebox, and marks on the regional map corresponding to each team. The "global" map serves as an entrance to the country which users want to explore for further detailed information. The individual maps of each region along with the marks of each team will help users to get a clear idea of the location of each club. The score table will change according to the season selected.

Design and Implementation

In addition to these fundamental information, other interesting factors will also be added into the supplementary information with respect to the team players. Considering the users may curious about the relationship between team performance and figure statistics of team players, herein we provide another layer containing figure statistics of team players which will be found in detail in this section. A radar figure is added to this layer to visualize the physical strength of the best players.

3.1.3. little ingenuity

Since we are dedicated to designing a user-friendly website, we paid speacial attention on the website. For example, considering the users may be especially interested in the score table for comparison, we fixed the score table to make sure that it would appear however users browse the whole website. These were all realized by setting the css.

What's more, in order to better guide the users to their interests, we set floating windows when the mouse is over the mark of each team. After clicking the geographical mark, the corresponding score table will show up and users can further look into information of each team.

3.2. Further improvement on preliminary proposal

Fig. 1 shows the preliminary proposal of the whole website which well illustrates the main structure of our website while in **Fig. 2**, we present our final website. The structure stays the same while some little changes were applied to the real website.

First, we added a banner to stress the theme of our project with people from different regions playing the football, which also indicates the theme of European Professional Football: the promotion of common progression and healthy competition.

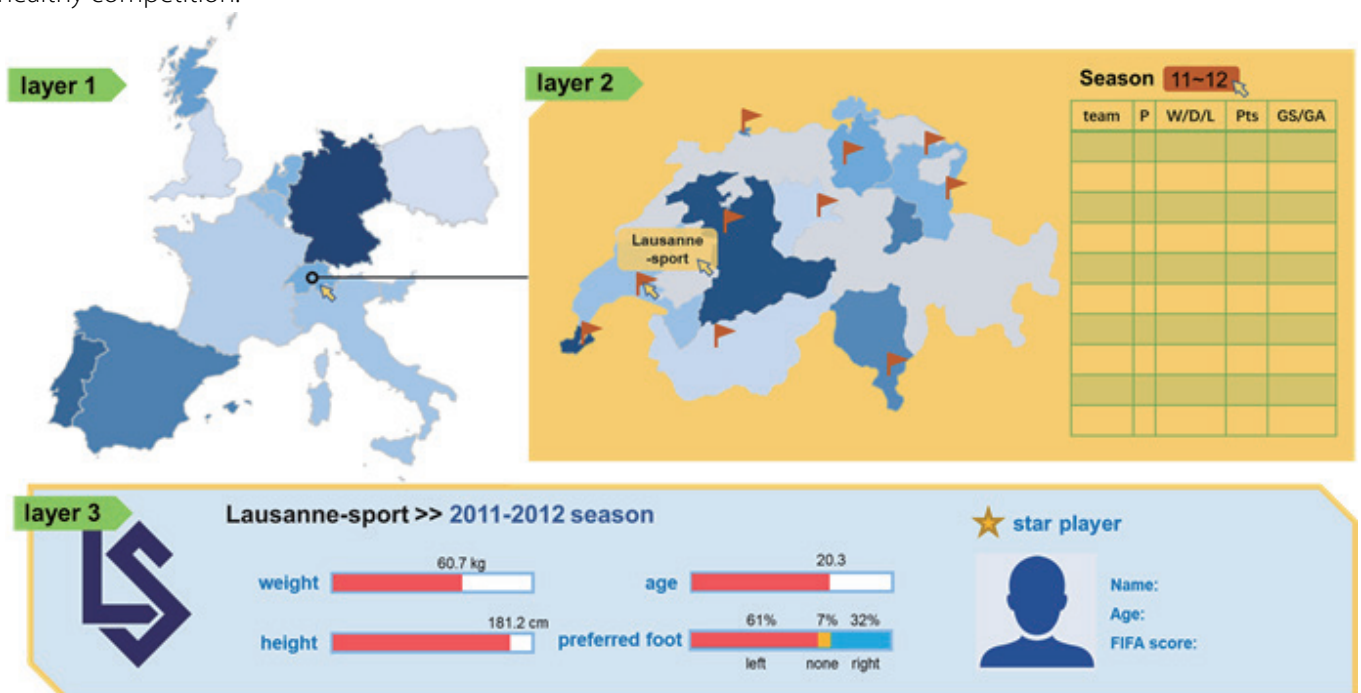


Figure 1. The preliminary sketch of the idea of our website

Design and Implementation



Figure 2a. the first and second layer of the website

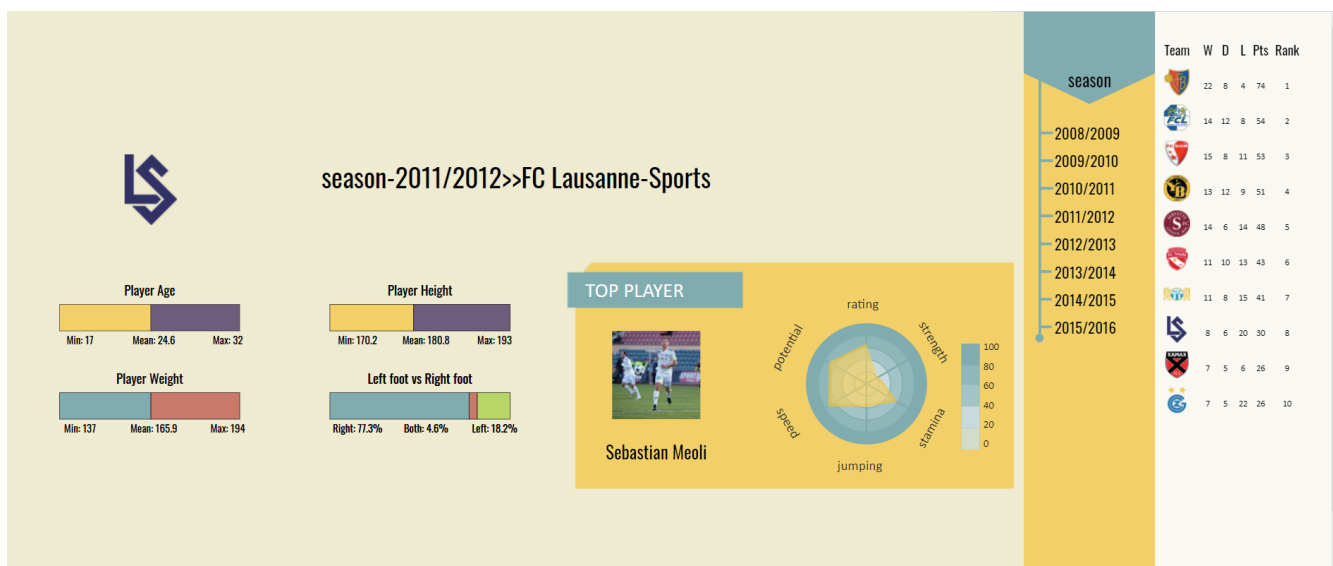


Figure 2b. the third layer of the website

Second, as we mentioned before, the theme is very precise and warm and the score table and season bar are fixed on the right side of the website to provide convenience in case users want to review the score table from time to time. The colors in the map are also removed to avoid misunderstanding. Due to the speciality of organization of the European Professional Football matches, where the games are all domestic, they are not comparable on the country level.

Second, we added more to the statistics of the top player including the rating, strength, stamina, jumping, speed, and potential scores ranging from 0-100, and visualized these data with a color-gradient radar plot based on the hope that users can intuitively view the strength and possible weakness of top players from different teams.



4. Peer Assessment

A clear peer assessment is provided as follows with detailed contributions listed:

- Jingran Su: implementation of Exploratory Data Analysis, preparation and consolidation of dataset, correction of wrong data.
- Yueqing Shen: overall design of website, construction of the website (mainly the css and html part, and js part with the first layer). the crawler of pictures and coordinates, the preparation of reports and process book.
- Zhenyu Zhu: implementation of Exploratory Data Analysis, construction of the website(mainly the js part with the second and third layer) , revision of the process book.