**Football Transfers Data Analyze**

**Team members:**

* **Roy Liberman –** [**roy.liberman@** **mail.huji.ac.il.com**](mailto:roy.liberman@%20mail.huji.ac.il.com)**, roy.liberman**
* **Eden Yerushalmy –** [**eden.yerushalmy@mail.huji. ac.il**](mailto:eden.yerushalmy@mail.huji.%20ac.il) **, eden107**
* **Moran Polack –** [**moran.polack@mail.huji.ac.il**](mailto:moran.polack@mail.huji.ac.il) **, polack1989**

**European football transfer market introduction:**

The football industry in Europe is valuated in hundreds billions of dollars. One of the major business aspects of football is the Transfer Market. Every football season there are two period of times which called the transfer window. In those period of times, each team can buy and sell players according to the regulation of FIFA (International Federation of Association Football) and UEFA (Union of European Football Associations). There are three ways for a player to switch team –

1) Transfer – the destination team pays money to the origin team.

2) Loan – the player moves to the destination team for a limited time and returns to his team.

3) Free – the contract of the player in the origin team has ended, the player has the right to join any club.

FIFA and UEFA regulation supposed to prevent situation where teams has transfers agreements. The reason this regulation exists is to give more power to the players and keep the market open (without monopoles).

**Problem description:**

Analyze teams transfer history in order to study existence of financial connection between European football clubs in terms of the amount of player and money that was transferred between them in the last decade.

Our assumption was that exceptional connections as described above exist. We wanted to figure out if indeed the phenomenon occur and in which extent.

**Data:**

Our data includes **all** **transfers** in Europe between the years 2007 - 2017.

Record is a tuple of the following values: *Player Name, Transfer Year, Original Team, Original Country, Destination Team, Destination Country, Transfer Type and Price.*

Where Country refers to a team (etc. Barcelona – Spain), Price scale is Millions of Euro.

The data contains 5,932 transfer records, 4,018 Players, 723 Teams and 79 Countries.

The size of the data is 1 MB.

We used a crawler to collect all the information. Site address - <http://www.soccernews.com/soccer-transfers>

Note – FIFA and UEFA **do not** publish this information, mainly because of pressure from the biggest teams in Europe, who believe that this kind of information will increase the price of players and regulation.

**Solution:**

* **Data collection and organization –** we built a crawler using python package “Scrapy” to collect the date from the website. The website data was not fully associative. We addressed the following issues:
  + - Player names, Teams names – the names were not consistently (etc. case sensitive, sometimes written in Latin letters, spaces and special signs). Moreover, teams were written in several names (etc. Barcelona = “Barca”, “FC Barcelona”, “Barcelona”).

We changed all string to upper case, deleted spaces, and created a team mapping using dictionary to avoid multiple team names.

* + - Price – prices were written in several formats (etc. “15M”, “15Millions”, “15,000,000”, “15”, “500,000”, “NA”, “Loan”, “Free” and more).

We parsed it into two attributes – Transfer Type and Price. Transfer type is either Transfer, Loan or Free (see above introduction for more details). Price is number (Millions of Euros) in case it is transfer, null otherwise.

* + - Countries – the website data does not include the country of each team. We added the Countries manually while creating the team mapping dictionary.

We used python package named “XslxWritter” to create several excel tables of the data in a file named “**project\_data**.xls”.

* **Third – algorithm for finding cliques and communities – maybe add explanation about the excel graph**

**Future Work:**

Investigate the correlation between the amount of money paid for a player and the time period that player stayed in the team. Moreover, check if there is a correlation between the amount of money team invests and the achievements of the team and player specifically.

To do so, more data about the players (etc. personal achievements, amount of goal scored and more) and about the teams (achievements and awards)

Another interesting aspect is to add to our data information about the team’s budget and understand the influence of transfer market on it. Team budget constructed by few major components:

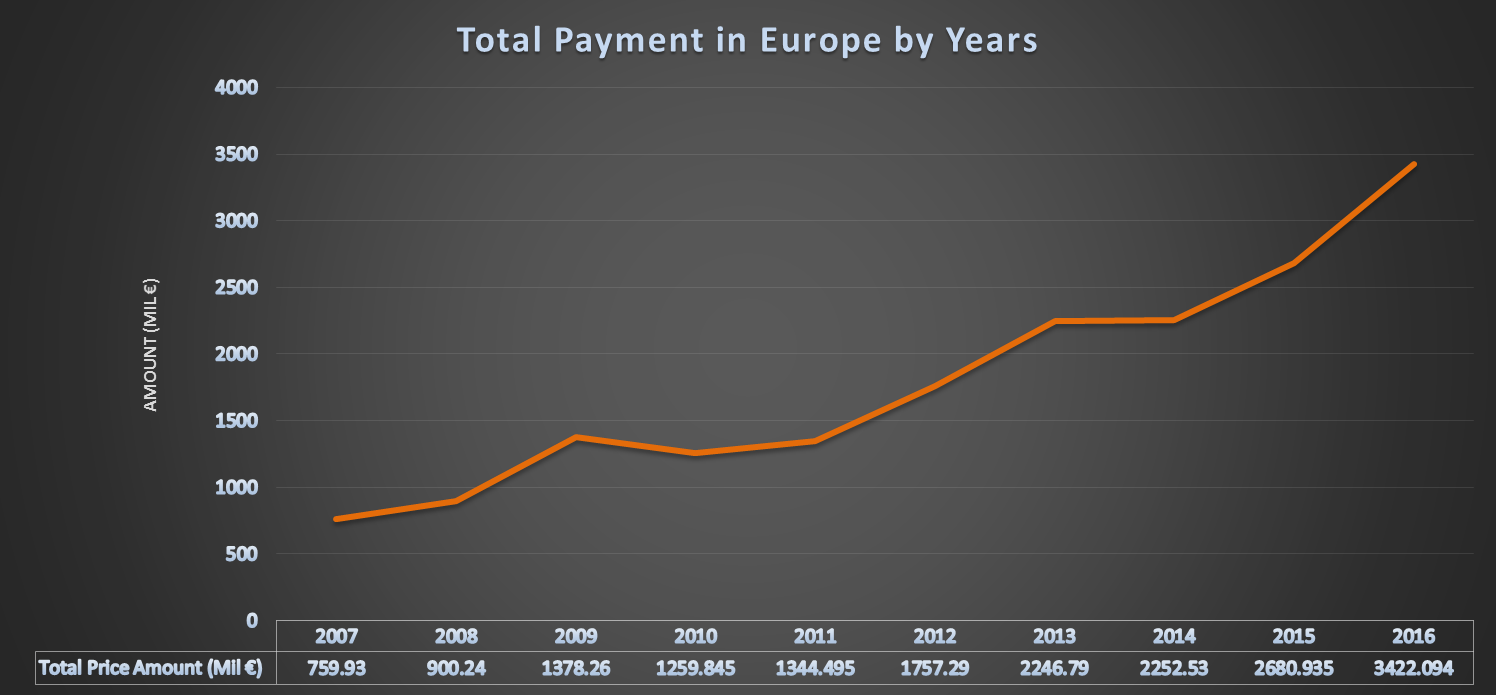
* Transfer Market
* Broadcasting Rights and competition money (winning awards)
* Brands and Sponsors

It will be interesting to discover and try to understand if there a relation between the amounts of money team invest in transfer market and the bank balance of the teams.

FIFA and UEFA do not publish information about budgets of the teams.

By the way – it seems like the amount of money does not related to the bank balance of a team, but to the amount of money owner have...).

**Brief conclusion:**

****

|  |
| --- |
| The amounts clubs spending on buying players increased significantly over the years.  There are 3 main reasons :  1) Sponsors and Media increased the amount of money that has been given to the clubs (mainly because of popularity). 2) New owners of a big number of teams – the new owners invested big amount of money in their teams after they bought it.  3) Reasons 1-2 increased the competition between clubs about each and each player. |
|
|
|

