# Football Data Challenge

prepared by Dimitri Lajou, Pijus Simonaitis
ENS Lyon, DBDM M1

April 25, 2016

## 1 Choosing the data set

Football data challenges data set from kaggle in Class attracted our attention because of our interest in football and curiosity about how predictable the outcomes of football matches are. In addition to that the given data set happend to require little of preprocessing.

## 2 Description of the data set

Our data set consists of:

1. Unique ID associated to every match

2. Date of the match

3. Home and away teams

4. Sets of odds proposed by six betting agencies: Bet365, Bet&Win, Interwetten, Ladbrokes, VC Bet and William Hill

 A set of odds consists of scores for:

1. Home win

2. Away win

3. Draws