Introduction into Big Data analytics

Assignment 3

Janusz Szwabiński

Task 1 Install Apache Spark and PySpark on your computer.

Task 2 Collect the data of all NBA players since 1980 (i.e. the year of introduction of 3-point-throws).

Task 3 By making use of PySpark:

- find best scorers sorted in ascending order,
- analyze the average number of 3 point attempts for each season in a time limit of 48 minutes (an interval corresponding to an approximate full NBA game),
- predict the average number of 3 point attempts for the next 5 years.