HOMEWORK 1

In this homework you will learn to setup Hadoop on your labtop, and experiment with Sqoop, Hive or Impala.

1. Download, install and configure Hadoop on your Labtop
2. Consider the MySQL database movielens, derived from the MovieLens project. (https://grouplens.org/datasets/movielens/). The database consists of several related tables, but we will import only two of these: movie, which contains about 3,900 movies; and movierating, which has about 1,000,000 ratingof those movies.Importing a table (your arbitrary data) from MySQL to HDFS.
3. Using Sqoop to import directly into Hive or Impala
4. Create and manipulate the Movilens tables with Hive or Impala
5. Bonus task: try to evaluate and compare the processing performance between MySQL and Hive/Impale on HDFS.