

Quiz 1:

You are given access to a database that contains the test scores of the TOEFL (Test of English as a Foreign Language) for a very large population. The records in the database might have various formats, but you can always extract the legal name, the city of birth, the number of an official identity document such as a passport, and the birthdate of the test taker as well as their score and the date of the test. The records are kept in different files according to the test center and of course, a test taker can take the test in different locations so that the records with a

The test is taken by students to be admitted to studies in the USA. Students can take the test multiple times and submit their best score to the university to which they are applying. Unfortunately, fraud is a real problem and it is known that people will take the test for others. In order to find these frauds, you are asked to find instances where the same person has shown a dramatic increase in their test scores over a short time, as this is an indicator of fraud.

Organize a map reduce job in order to list test taker information with all the TOEFL test scores of the test taker.