Data generator for data sources

1. Goal

The goal of the task is to implement schemas of data sources and implement a data generator for these sources.

2. Subtasks for implementation
   a. Implementing two schemas (the one for relational database is obligatory and the second one depends on the type of data source),
   b. Data generator development (the data are needed! It's up to you how do you achieve them, everything is allowed – the generation way is not evaluated, the generated data are evaluated).
3. Products:
   a. Schemas.rar
   b. Generator.rar
4. Requirements for data
   a. Number of tuples corresponding to fact rows in data warehouses must be scalable (up to 1 milion rows),
   b. Two data snaphosts (for both data sources) must be generated (in two time moments: T1 and later T2). Data in T2 are the data from T1 expanded with new tuples (corresponding to fact rows) and with new tuples (corresponding to changes in dimension elements – at least one such change must be done),
   c. Data from both data sources must enable integration (the same business keys).
5. Maximum score – 5 points.

Attention!
There are three ways of loading data into the database:
- directly from the application,
- from files with INSERT statements,
- BULK loading.
Due to the huge number of rows to be generated BULK loading is the best option!