Lab Training Database Systems 2020

Task 3

For this task, you can use your database along with your test data from Task 1 in Postgres. The goal is to prepare the database for a possible future data warehouse, if the use case from Task 1 gets larger.

Tasks

a) You already have tables with content from the first task. Transform the schema accordingly to a **star schema** and also a **snowflake schema** (both are commonly found in data warehouse systems). For both schemas, create a new database each in Postgres, with corresponding tables (*fact* and *dimension* tables). For your tuples from Task 1, perform the transformations (attribute split) accordingly!

   => **For the PDF**: Add both schemas (table names, attributes, keys) and the tuples inside! Any decisions for splitting tables, e.g., for the snowflake schema should be describe, too. Simply assume that your database would hold 100 million tuples - how would you distribute them according to star and snowflake?

b) Run your five queries from Task 1 again on both of the schemas. You probably have to adapt the SQL syntax, though.

   => **For the PDF**: Add the queries in SQL for both schemas (output is not necessary).

Further comments on the third task:

I. If you are not familiar with star and snowflake schema - there are a lot of good descriptions around in the literature as well as the web.

II. You have to do it on your own, **do not copy&paste** commands and results from your fellow students!

III. When you have finished your tasks, upload the PDF on Moodle **within the given time**.