

Data Modelling and Databases II - Assignment 1

Alisa Martyanova BS18-05

- Python script moves database from Postgres to MongoDB using export from Postgres to JSON file
COPY (SELECT ROW_TO_JSON(t) FROM (SELECT * FROM table_name) t) TO name.json
and import from JSON file to MongoDB
table.append(json.loads(line))
client['dvdrental'][table_name].insert_many(table).

Used libraries: json, pymongo, psycopg2

It is assumed that you already have database **dvdrental** in Postgres

```
16 db_name = 'dvdrental'
17 user_name = 'postgres'
18 password = 'Alisa'
19 host = 'localhost'
20 path_to_files = '/home/postgres/json_files/' #
21
```

You need to change in the code **user_name** (17) and **password** (18) from your Postgres client and **path_to_files** (20) - path where intermediate .json files will be stored (it can be current directory).

- All the queries are written in Python using pymongo. In order to make some of them simpler I merged two tables in Postgres, and then export them in MongoDB with the rest tables from Postgres (merged_q1, merged_q3).
- Results of queries 1-3 and 5 are written in the tables in .csv files, output for query 4 is written in .txt file. For query 4 recommendation is based on the most popular films (top 5 films that were mostly rented) of the same category that customer rented in particular date.
- The average execution time for script itself and for queries 3,4 is 6-7 seconds. The reason is that about 20 000 records needs to be processed. Queries 2, 5 works much better - for about 2 sec and query 1 - for 0.1 sec.
- Component diagram you can see in same .zip archive in **Alisa_Martyanova_component_diagram.vpd.svg** file.
- Link to github:
https://github.com/AlisaMartyanova/Data_Modelling_and_Databases_II/tree/master/Assignment_1