Lab: Load and manipulate Data on both Postgres and Mariadb and stream the results to Kafka

1. Spin up 2 Dockers: postgres and mariadb **(2)**
* postgres should listen on 5533
* mariadb should listen on 3366

2. Load the file "**big_alien_witness_data.csv**" into **mariadb (2)**
* c)
* load the file into the table "raw_input"

3. Think of a "normalized" design for the "raw_input" file and implement it on **POSTGRES(3) minimum 2 tables.**
* create the necessary normalized tables in postgres

4. Sample data question **(3):**
1. Which city had the most alien witnesses?
2. Which Region had the least alien witnesses?
3. Which time of day (day/night) had the most alien witnesses?
4. How many distinct locations were involved?

5. Create a script that reads the data from mariadb "raw_input" table, and writes the data to postgres new tables **(3)**
* it can be several scripts
* you can do it with file as intermediate step

** bonus – do it without an intermediate file **(3)**
feel free to use bash or any language you wish

6. Spin off a local kafka cluster (can be standalone) **(2)**

- Create topics based on postgres tables

7. Create a script that reads data from **postgres** and streams it to appropriate topic in kafka (5)

* script should be called stream.sh
* stream.sh TABLE_NAME

selects from TABLE_NAME and produces to topic TABLE_NAME

8. Create A script that reads topics from #6 and writes the data to the same tables in **mariadb (5)**

Results:

Scripts for create table (DDL) should be named create_table_TABLENAME.sql and placed inside ./sql/ folder in your repo

Same for creating kafka topics, place them into ./kfk/ dir.

Create separated directory in -repo with name “db-kafka_domainName”, for example “db-kafka _ipysmennyi”. All code should be created in feature branch with name “feature/domainName”, for example “feature/ipysmennyi”. After merge to “master” branch.