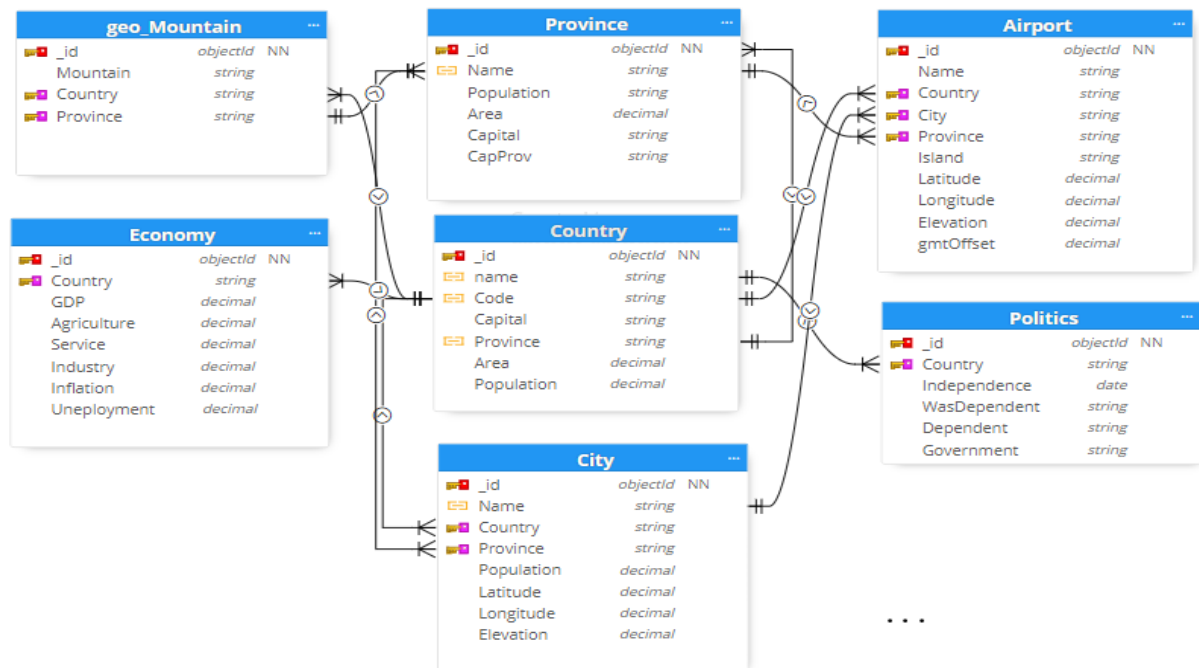


DBMS Practical - Project 1: MONDIAL

Mahmoud Alhag Ali



(Collections of mongoDB from the original Postgres data schema only partly shown)

Reasoning

MongoDB is a document-oriented database that utilizes JSON-like documents with flexible, optional schemas. In this context, relational tables from PostgreSQL are reinterpreted as collections, where each collection consists of documents formatted in a JSON-like structure. Within these documents, fields store their corresponding values.

MongoDB allows us to store unstructured data, thereby saving storage space. For this project, the MONDIAL data was retrieved from PostgreSQL using a custom Python script (located in utils folder), processed, and then saved in CSV format. To provision the data, we adapted the Docker Compose configuration from the Enron project and modified the corresponding `init_mongo.sh` setup. Finally, using `mongoimport`, we loaded the data and created the collections.