

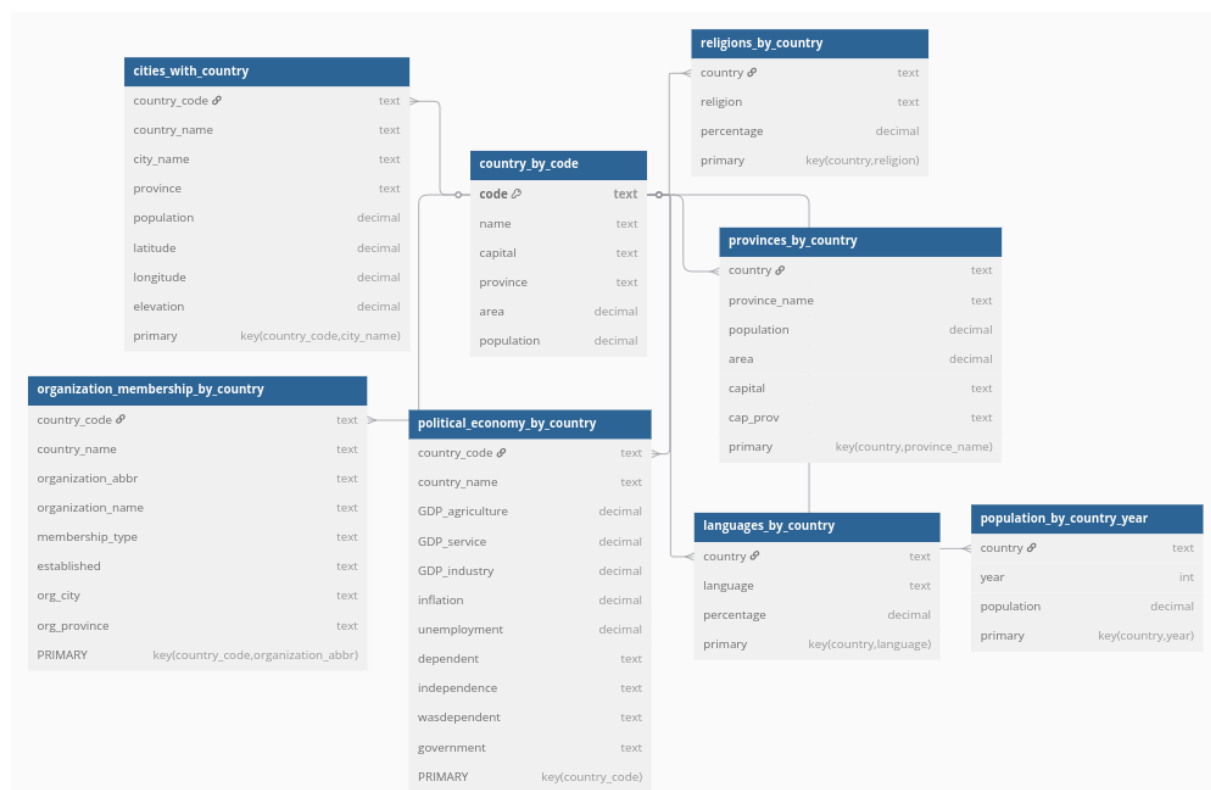
DBMS Practical - Project 1: MONDIAL

Roman Christof

DBMS Design Choices

Cassandra

Schema



(not denormalized tables from the original Postgres data schema only partly shown)

Reasoning

In Cassandra, schema design follows a query-driven approach rather than traditional normalization principles. Since Cassandra does not support joins, we structured the tables in a way that directly supports the queries they are intended to answer. As a result, related tables from the relational model are merged into denormalized tables if their data is often queried together. For instance, the City and Country tables are merged into a single cities_with_country table. This enables queries such as retrieving the name of the country for each city and counting the number of cities within a country. These operations would normally require a join between city and country data, which is not possible in Cassandra. By combining them into a single table, all relevant information is stored together, allowing

efficient querying. Similarly, the tables Organization, isMember, and Country are merged into organization_membership_by_country. This combined table supports queries such as determining the number of organizations a country belongs to and retrieving detailed information about those organizations. Again, merging is necessary here because the original relational model relies on join operations between multiple tables, which Cassandra cannot perform. The political_economy_by_country table is designed following the same principle. It merges attributes from the relational Economy and Politics tables, along with the country name. This structure supports queries such as retrieving a country's economic indicators (GDP by sector, inflation, unemployment) alongside its political attributes (form of government, independence, dependencies). These attributes are often analyzed together in geopolitical or socioeconomic analyses. However, not all tables are merged. Some, such as Province, geo_river and River, are intentionally left in a normalized structure to illustrate Cassandra's limitations when it comes to flexible queries.