FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO
Mestrado Integrado em Engenharia Informática e Computação

Database Technology

Assignment nr. 3

Version 1 – Elections

NoSQL Assignment

Goals

Build a small DB using a NoSQL approach and compare with the results obtained from the relational approach.

Work groups

The assignment should be executed by a group of one or two elements.

Deadline

To submit in Moodle by 2017-05-31.

Subject

The context for this assignment is the Parliament elections of October, 10th 1999.

The country is divided in circumscriptions, called district (distrito), though on the islands the circumscription coincides with the region. The parties (partido) present lists (lista) for the circumscriptions where they stand up for election. It is at the district level that the MP mandates are obtained. MP elected by the emigrants are not considered, so the total number of mandates is 226.

However, voting results in parties are recorded with a finer-grain, at the parish (freguesia) level. Parishes belong to municipalities (concelho), which belong to districts (distrito).

The total number of votes in parties, plus the white (brancos) and null (nulos) votes are the number of voting citizens (votantes). Adding up the abstentions, one gets the enrolled (inscritos). The information on enrolled, abstentions, white and null votes is only known at the district level.

Parties may submit lists just to some districts.

This database has 7 tables with actual data. (Oracle database, user GTD7)

- The tables **freguesias** (parishes), **concelhos** (municipalities) and **distritos** (districts) are relative to electoral circumscriptions and administrative areas. The region is ('C' - Continent, 'A' - Açores, 'M' - Madeira).
- The table **listas** (candidate lists) has information about the number of mandates obtained by each party (sigla) in the districts (codigo) where it has participated. The value 0 means that the party has presented a list but it got no MP mandate.
- The table **votacoes** (voting) keeps the number of votes received by the party in the parishes.
- The table **participacoes** (participations) contains, for each district, the number of enrolled citizens (inscritos), voting (votantes), abstentions (abstenções), null (nulos) and white (brancos) votes. Note that:

  inscritos = votantes + abstenções

  votantes = sum of votes in each party in the distrito + brancos + nulos.

- The table **partidos** (parties) records the acronym (sigla) and the designation of the 12 existing parties.

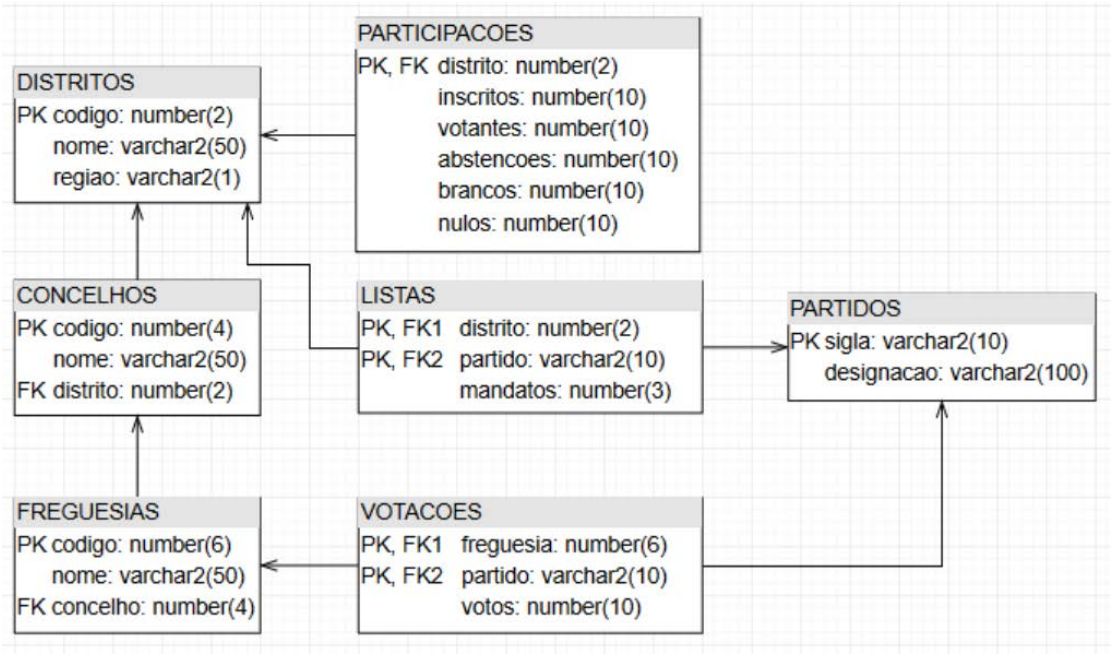

Figure 1: Relational schema.

The tables are in Oracle user GTD7.

TASKS

The following tasks should be performed, described in a report and presented on the deadline:

1) Design a Mongo document model for the elections example, explaining the decisions made. The information available under this model should be equivalent to the information in the relational database.
2) Migrate the data from the Oracle user GTD7 into the NoSQL database using one or more times the following method: write a PL/SQL package able to extract the data using an appropriate SQL query, deliver it in XML format, and process the XML result using a dynamic XSL stylesheet to produce in a browser a set of Mongo method calls to populate the NoSQL database designed in 1). Include in the report the SQL query, a sample of the XML result, the XSL stylesheet, and a sample of the method calls.
3) Prepare Mongo queries for the following questions:
    a. Calculate the total number of Members of the Parliament that each party has got.
    b. Find the number of votes in each parish of the municipality ‘Lisboa’ that the party ‘Partido Popular’ has got.
    c. Find the names of the districts and the designations of the parties for the cases where the party got an absolute majority in the district (meaning more votes than the sum of votes of the remaining lists).
    d. Ask the database a query you think is interesting.
4) Compare the Mongo and the Oracle implementations from the viewpoints of data size, processing time, and query easiness.