

Tecnologia de Bases de Dados
PROJ1 – Data Limite de entrega: a 23 março 2025

OTIMIZAÇÃO DE INTERROGAÇÕES

OBJETIVOS DO TRABALHO

Analisar os planos de execução de diferentes interrogações SQL a uma BD de teste. Avaliar o impacto da existência de índices e de estatísticas e do recurso a diferentes estratégias de estruturação das perguntas. O relatório a produzir deve, para cada situação, indicar a formulação da pergunta em SQL, a resposta obtida, as condições de execução do teste incluindo duração se relevante, o plano de execução e as conclusões a que chegou.

GRUPOS DE TRABALHO

O trabalho deve ser realizado por grupos de um ou dois elementos.

TEMA

Pretende-se registar a informação relativa às eleições para a Assembleia da República, de 10 de Outubro de 1999.

O país está dividido em círculos eleitorais, aqui designados por distritos, embora nas ilhas o círculo eleitoral coincida com a região. Os partidos apresentam listas nos círculos eleitorais a que concorrem, sendo a esse nível que são obtidos os mandatos de deputado. Não são considerados os círculos dos emigrantes, pelo que o total de mandatos é apenas de 226.

No entanto, as votações em partidos são registadas com granularidade mais fina, ao nível de freguesia. As freguesias pertencem a concelhos, os quais por sua vez pertencem a distritos. Todas as divisões administrativas têm um código e um nome. Os distritos estão em regiões ('C'- Continente, 'A'- Açores, 'M'- Madeira).

O total de votos em partidos, mais os votos em branco e os nulos, constituem os votantes. Somando votantes com abstenções obtém-se os inscritos. A informação sobre inscritos, abstenções, brancos e nulos só está presente ao nível de distrito.

Os partidos podem apresentar lista só nalguns distritos.

MODELO DE DADOS

DISTRITOS(codigo, nome, regioao)

CONCELHOS(codigo, nome, distrito → DISTRITO)

FREGUESIAS(codigo, nome, concelho → CONCELHO)

PARTIDOS(sigla, designacao)

VOTACOES(partido → PARTIDO, freguesia → FREGUESIA, votos)

LISTAS(distrito → DISTRITO, partido → PARTIDO, mandatos)

PARTICIPACOES(distrito → DISTRITO, inscritos, votantes, abstencoes, brancos, nulos)

Esta base de dados tem 7 tabelas povoadas com dados reais.

- As tabelas **freguesias**, **concelhos** e **distritos** dizem respeito às divisões administrativas e círculos eleitorais. Para cada freguesia é possível saber o seu código, o seu nome e o código do concelho a que pertence. Para cada concelho tem-se informação do seu código, nome e código do distrito (círculo eleitoral) a que pertence. Para cada distrito tem-se informação sobre o seu código, nome e região a que pertence ('C' - Continente, 'A' - Açores, 'M' - Madeira).
- A tabela **listas** possui informação sobre o nº de mandatos obtido por cada partido (sigla) nos distritos (código) em que concorreu. O valor 0 significa que concorreu mas não obteve nenhum mandato.
- A tabela **votacoes** guarda o resultado obtido pelos diferentes partidos nas freguesias. Um valor nulo em votos significa que o partido concorreu mas não houve contagem de votos (boicote).
- A tabela **participacoes** tem, para cada código de distrito, a informação relativa aos nº de inscritos, votantes, abstenções, votos nulos e votos em branco. Note que:

inscritos = votantes + abstenções

votantes = soma de votos no distrito + brancos + nulos.

- A tabela **partidos** guarda a sigla e designação dos 12 partidos existentes.

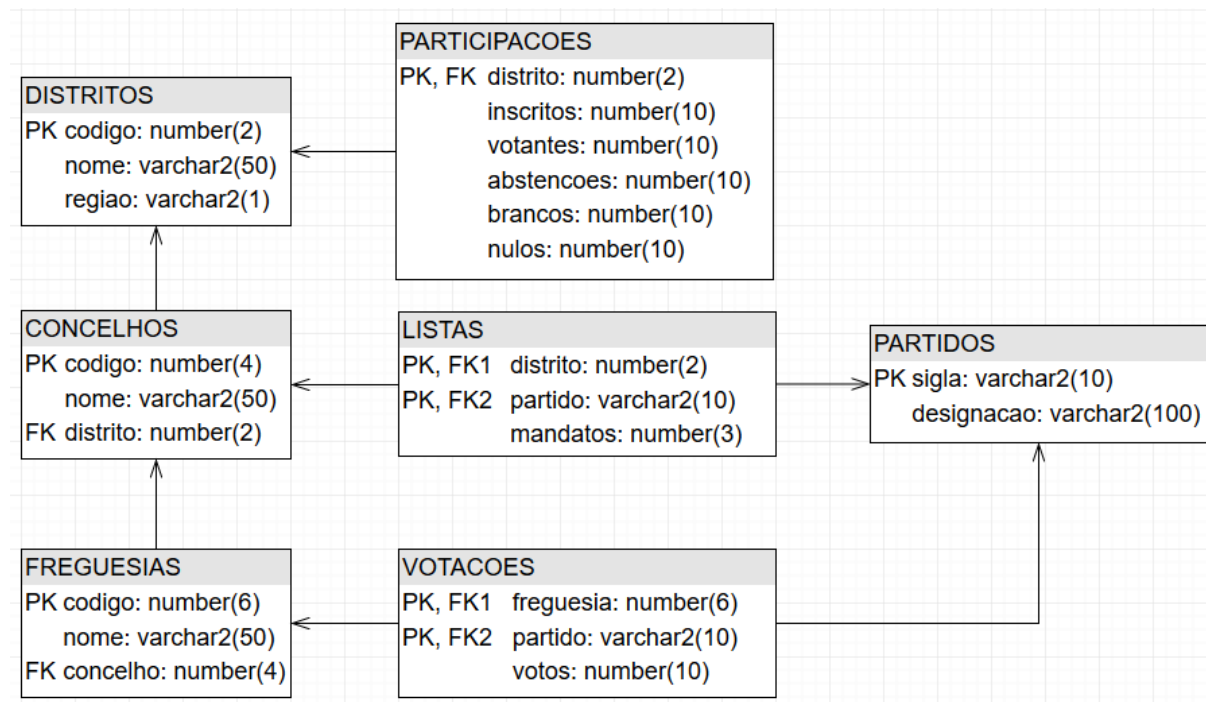


Figura 1: Esquema relacional.

DADOS

As tabelas encontram-se no utilizador GTD7, de onde devem ser copiadas, a triplicar, para as tabelas xlistas, xvotacoes, ..., ylistas, yvotacoes, ..., zlistas, zvotacoes, O objetivo é criar três ambientes de experimentação:

X - Sem índices nem restrições de integridade.

Y - Com as restrições de integridade standard (chave primária e chave externa).

Z - Com as restrições de integridade standard e os índices extra que achar convenientes, justificando a razão de inclusão de cada um deles.

MÉTODO

Pretende-se, relativamente às perguntas abaixo, quais as formulações SQL e as respetivas respostas, uma análise dos planos de execução dessas interrogações e qual o esforço estimado. As perguntas devem ser executadas nos três contextos.

QUESTÕES A ANALISAR

1. Seleção e junção.
 - a. Quais os códigos e nomes de freguesias do concelho 1103? E do concelho “Azambuja”?
 - b. Indique as siglas e designações dos partidos e o respetivo número de mandatos obtidos no distrito de Lisboa.
 - c. Indique o número de votos obtido pelo BE nas freguesias do distrito de Lisboa.
2. Agregação.
 - a. Quantos votos teve o ‘PS’ a nível nacional?
 - b. Quantos votos teve cada partido, em cada distrito?
 - c. Qual o partido que, ao nível de freguesia, registou o maior número de votos? Indique a sigla do partido, o nome da freguesia e os votos correspondentes.

-
- d. Para cada distrito indique qual o seu nome e a designação e número de votos do partido que nele teve melhor votação.
 3. Negação. Analise de igual forma a questão “Quais os partidos que não concorreram no distrito de Lisboa.”
 4. A pergunta “Houve algum partido a vencer em todos as freguesias de um concelho do distrito do Porto? Indique código do concelho e sigla do partido.” é de natureza universal. Compare do ponto de vista temporal e de plano de execução as estratégias da dupla negação e da contagem em três situações diferentes (só no contexto Z):
 - a. Sem vista para calcular o vencedor em cada freguesia.
 - b. Com vista.
 - c. Com vista materializada (eventualmente com índices).
 5. Compare os planos de execução da pesquisa “Quantos votos tiveram o PS e o PSD nos distritos 11, 15 e 17”, considerando no contexto Z
 - a. Com índices árvore-B em `zconcelhos.distrito` e `zvotacoes.partido`.
 - b. Com índices bitmap.