



Universidade de São Paulo - São Carlos
Instituto de Ciências Matemáticas e de Computação
SCC-541 Laboratório de Bases de Dados
Trabalho Prático T6 – Exercícios sobre Visões (*views*)

Dr. Caetano Traina Jr., Igor Alberte R. Eleutério
1º semestre de 2023

Data para entrega: 14 de junho

Datasets

Base de dados da Fórmula 1 - FIA.

A base de dados a ser utilizada é a mesma preparada no **Trabalho 1**. Caso necessário, os arquivos originais para a carga de dados encontra-se no Google Drive, e podem ser acessados no endereço [este link](#). Para acessar os arquivos, você deve estar logado no Google Drive com sua conta USP. Cada arquivo corresponde a uma tabela, que deve ser carregada na base de dados.

Os dados originais foram ligeiramente modificados para facilitar o trabalho. Os originais podem ser obtidos nos seguintes *sites*:

- Dados da Fórmula-1: [Ergast Developer API](#)
- Países e Cidades do planeta: [GeoNames](#)
- Aeroportos: [OurAirports](#)
- O objetivo deste trabalho é explorar o uso de **índices** nos SGBDs.

Instruções

Considerando as Questões 1 e 2, devem ser definidos índices que ajudem a responder às consultas mais rapidamente. Depois, esses índices deverão ser avaliados quanto à sua eficiência. Devem ser utilizados um índice **B-tree** e um índice **Hash**, um para cada questão. A equipe deverá escolher um deles para cada questão e justificar sua escolha no relatório.

Considere os seguintes pontos:

- Avalie se é interessante utilizar as cláusulas **INCLUDE** e **WHERE** nos índices definidos. Lembre-se de justificar as escolhas no relatório.
- Para medir o tempo de execução de uma consulta, é interessante desenvolver uma função em PL/SQL para executar cada consulta 100 vezes e apresentar o tempo médio gasto. A função definida na figura 1 mostra um exemplo de como esse procedimento pode ser implementado (lembre-se que esse exemplo talvez precise ser adaptado às necessidades de cada questão).
- O procedimento de testes deve ser o seguinte:
 1. Definir uma consulta que contemple as condições de cada questão.
 2. Executar a função **Mede_tempo** antes de criar o índice, passando a consulta como parâmetro.
 3. Criar o índice.

```

CREATE OR REPLACE FUNCTION Mede.Tempo(Q TEXT)
  RETURNS TABLE (Name TEXT , Nationality TEXT) AS $$
  DECLARE
    TIni TIME;      TFin TIME;
    i DOUBLE PRECISION;
    Diff BIGINT;
  BEGIN
    -- Registra o tempo inicial
    TIni := CLOCK_TIMESTAMP ();
    FOR i IN 0..100 LOOP
      EXECUTE Q;
    END LOOP;
    -- Registra o tempo final
    TFin := CLOCK_TIMESTAMP ();
    -- Calcula a diferenca em milisegundos
    Diff := ROUND( (EXTRACT(EPOCH FROM TFin) -
                    EXTRACT(EPOCH FROM TIni) )/10);
    RAISE NOTICE '% - % := %', TFin , TIni , Diff;
    -- Retorna o resultado da consulta recebida
    RETURN QUERY EXECUTE Q;
  END;
$$ LANGUAGE plpgsql;

```

Figura 1: Função para obter o tempo médio de 100 execuções de um comando em SQL.

4. Executar novamente a função `mede.tempo` para avaliar tempo da nova execução.

Dica: o comando `EXPLAIN ANALYZE <query>` permite avaliar como um comando `<query>` é executado pelo SGBD. Por exemplo, caso apareça `Seq_Scan` é feita uma leitura sequencial. Caso apareça `Index_Scan using <nome>` é utilizado o índice que tem o nome `<nome>`.

Atividades da semana

Exercício 1) Dado o nome exato (`<Nome>=Forename|Surname`), recuperar a nacionalidade do piloto.

Exercício 2) Dado um nome completo ou o padrão do início de um nome de uma ou mais cidades brasileiras, recuperar a Latitude, Longitude e População das cidades brasileira que atendam ao nome ou padrão. Teste em uma consulta do tipo: `WHERE name LIKE 'nome%'`.

Exercício 3) (**Questão teórica**) Estruturas B-trees conseguem indexar consultas com predicados do tipo:

`<Atributo> LIKE ''%valor%''`? (Considere que o atributo seja do tipo `TEXT`). Comente.

Entrega

Cada equipe deverá entregar dois arquivos no **Escaninho** do **Tidia**, sendo que apenas o líder de cada equipe deverá colocar, no seu *Escaninho*, em uma pasta com o nome **T5**:

- Um arquivo com o *script* no formato **.zip**, contendo um arquivo **.sql** com os comandos **SQL** utilizados para cada atividade.
- Um arquivo com o relatório SUCINTO no formato **.pdf**.
- O relatório deverá apresentar capturas de tela do tempo de execução para cada caso. O procedimento seguido para executar os testes também deve ser descrito, bem como as escolhas de índices feitas e suas justificativas. Em caso de questões com perguntas, a resposta ao item deve ser escrita.

Como a atividade tem prazo de entrega máximo de uma hora antes da próxima aula, os arquivos devem ser submetidos

até às 18h00 do dia 14 de junho, *com postagem somente do líder da equipe*.

Não serão aceitos projetos feitos à mão e a organização clara das respostas também é um ponto avaliado.

Plágio será avaliado com nota zero.

Bom Trabalho!