

DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO, UFMG
DCC011: INTRODUÇÃO A BANCO DE DADOS - TRABALHO PRÁTICO 2, 2018/2

INTRODUÇÃO

O objetivo deste trabalho é projetar e implementar um banco de dados relacional para realizar consultas interessantes. O projeto do banco de dados deverá seguir o processo apresentado em aula. Em particular, deverão ser apresentados o diagrama entidade-relacionamento e o esquema relacional referente ao banco de dados projetado. O banco de dados deverá ser instanciado com dados reais (de preferência) ou gerados sinteticamente. Consultas deverão ser realizadas sobre o banco e discutidas quanto à sua eficiência.

Características básicas do banco de dados (até 20%). Para definição do tema, cada grupo deverá escolher entre: (i) um conjunto de dados do portal <http://dados.gov.br>; (ii) um conjunto de dados real que possa ser coletado da Web (e.g, twitter, imdb, github, dblp, etc); ou (iii) uma aplicação atual ou inovadora para a qual dados sintéticos possam ser gerados. Para o banco de dados escolhido, deverão ser apresentados o diagrama ER e o esquema relacional correspondentes, atendendo aos seguintes requisitos:

- Pelo menos 4 tipos de entidade, cada tipo com ao menos 2 atributos (além de um atributo identificador);
- Pelo menos 3 tipos de relacionamento, ao menos um com cardinalidade M:N;
- **Pode ser necessário alterar os dados originais para atender a esses requisitos.**

Consultas (até 30%). Deverão ser especificadas e executadas um total de 10 consultas em SQL, sendo:

- 2 consultas envolvendo as operações de seleção e projeção;
- 3 consultas envolvendo a junção de duas relações;
- 3 consultas envolvendo a junção de três ou mais relações;
- 2 consultas envolvendo funções de agregação sobre o resultado da junção de pelo menos duas relações.

Características avançadas (até 50%). A segunda metade da avaliação deste trabalho considera três dessas quatro opções: (i) avaliação de eficiência das consultas (cada uma das 10 consultas especificadas é formulada de pelo menos 2 formas diferentes considerando, por exemplo, utilização ou não do operador JOIN, utilização ou não de consultas aninhadas, criação de índices, entre outras); (ii) para deixar os dados e as consultas ainda mais interessantes, o banco pode armazenar também dados de tipo não-convencional (texto, imagens, som, vídeo, etc); (iii) além das 10 consultas obrigatórias, pode-se especificar mais três consultas do tipo relatório (por exemplo, um relatório seria composto por indicadores agrupados por região geográfica ou por grupos demográficos); (iv) interface gráfica (e.g., um website) para visualização *interativa* dos dados modelados por meio das consultas produzidas.

NOTA: pontos extras se o grupo optar por realizar as quatro opções.

OBSERVAÇÕES

- O trabalho deverá ser feito em **grupo de 2 a 5 pessoas, onde uma é líder** (ficará responsável por fazer avaliações individuais de cada um dos integrantes do grupo, os quais também a avaliarão). Dado o tempo limitado para realização do trabalho, é importante que os membros do grupo dividam as atividades entre si para evitar desperdício de tempo. A avaliação é *individual*, e o trabalho é proporcional ao grupo;
- Em hipótese alguma se pode utilizar de aplicações cujos esquemas conceituais e/ou relacionais tenham sido vistos ou discutidos total ou parcialmente em aula ou nas provas e exercícios;
- Deve-se usar o SGBD MySQL ou PostgreSQL para implementação do banco de dados e suas consultas;
- Se for o caso, ao colocar o resultado das consultas no relatório, deve-se evitar fazer *printscreen* da tela do DOS com fundo preto e letra branca :-/;
- A origem dos dados deve ser informada no relatório final;
- A realização das características avançadas deve estar clara no relatório, e para desenvolvimento de interface, o link para seu website ou um conjunto de *printscreens* deve ser adicionado ao relatório final;
- Dificuldades encontradas e problemas com suas soluções devem ser detalhadas no relatório final;
- A avaliação compreenderá o relatório final entregue e a apresentação em aula. Além da funcionalidade básica prevista, a avaliação irá considerar a criatividade e a qualidade do banco de dados desenvolvido.

CALENDÁRIO

31/10: Proposta: grupo e tema (máx. 1 página; via moodle)

A proposta deverá descrever o conjunto de dados escolhido. Em particular, deverão ser descritas as entidades e relacionamentos presentes, e possíveis consultas a serem formuladas sobre esses dados.

26/11: Relatório Final (pelo líder via moodle)

O relatório final deve conter: descrição textual do banco de dados, diagramas ER e Relacional, consultas realizadas com explicação textual e respectivo comando SQL, e a descrição das atividades realizadas individualmente por participante do grupo.

26/11 a 03/12: Apresentação

Cada grupo terá em torno de 10 min (a confirmar, dependendo do número de grupos) para apresentar o tema do trabalho, a modelagem desenvolvida (conceitual e relacional), uma seleção das várias consultas formuladas e de seus resultados, e demonstração da interface (se tiver sido desenvolvida).

05/12: Última chance para entregar o relatório (com possíveis correções e complementações, e.g., interface).