**PROJETO CHINOOK DATABASE**

Antes de ir ao projeto, quero que respondam as 7 seguintes perguntas.

No futuro, não é necessário upar essas perguntas ao GitHub.

1. Apesar de existirem muitos DBMS (PostgreSQL, SQLite, Hive, Oracle...), eles costumam ser bastantes similares entre si. O que mais costuma diferir entre os DBMS são funções para tratar um tipo de dado. Que tipo de dado é esse?
2. Os bancos de dados relacionais não têm esse nome a toa. Eles possuem tabelas que se relacionam para evitar a redundância de dados. Quais são os 2 elementos(chaves) que permitem esse relacionamento? Qual a diferença entre eles?
3. Existem 4 principais tipos de JOIN, qual a diferença entre eles?
4. O que são funções de agregação?
5. Em um SELECT statement em que trazemos 5 colunas (col1, col2, ..., col5), quando colocamos col1 e col2 no GROUP BY, como as colunas restantes devem estar no SELECT?
6. Qual a diferença entre os resultados dos SELECTs abaixo?

Dica para entender a partir da visualização dos resultados:

* Criar um arquivo csv contendo 2 colunas: cliente e valor\_da\_compra.
* "Popular" esse csv com linhas referentes a várias compras de 3 clientes diferentes. É interessante ter pelo menos 1 cliente que tenha feito mais de uma compra. Salvar o csv. (Popular alguma coisa em tecnologia significa preencher com dados)
* Criar um banco de dados no DBBrowser, selecionar Importar -> Tabela a partir de csv. O nome da tabela deverá ser compras.
* Importar o .csv criado, rodar a query e analisar os resultados

**SELECT #1**

**SELECT**

cliente,

**SUM**(valor\_da\_compra) OVER (PARTITION BY cliente) AS total\_compras

**FROM** compras

**ORDER BY** cliente;

**SELECT #2**

**SELECT**

cliente,

**SUM**(valor\_da\_compra) AS total\_compras

**FROM** compras

**GROUP BY** cliente

**ORDER BY** cliente;

1. Caso eu use a função ROW\_NUMBER() como função janela, como funcionarão os parâmetros PARTITION BY e ORDER BY?

Exemplo de ROW\_NUMBER() como função janela:

**SELECT**

cliente,

valor\_da\_compra,

**ROW\_NUMBER**() OVER (PARTITION BY cliente ORDER BY valor\_da\_compra DESC) AS row\_num

**FROM** compras.

**PROJETO SQL**

O projeto consiste em:

* Levantar perguntas sobre os dados da Chinook database
* Fazer queries para responder a essas perguntas
* Plotar os dados retornados das queries
* Montar uma apresentação

Regras:

* Devem ser levantadas 10 perguntas.
* Nas queries, devem ser usadas pelo menos 4 GROUP BYs
* Devem existir pelo menos 3 queries com pelo menos 1 JOIN.
* Deve existir pelo menos uma query com 4 ou mais JOINs no mesmo SELECT.
* Deve ser utilizada pelo menos uma SUBQUERY (SELECT dentro de SELECT, com template WITH ou não).
* Deve ser utilizado ao menos 2 funções como funções-janela.
* Os códigos SQL devem estar em um único arquivo .txt quando forem me mandar. Não esqueçam de colocar a pergunta a qual cada código SQL está associado.
* A apresentação deve conter 1 slide que responda a cada pergunta. No caso, um total de 10 slides.