



Prova 2 – ECN-010, ECN-098
Ciência dos Dados – 2º/2021

Nome: Maria Rafaela Soares da Silva

1. Resposta:

Vantagem 1: Capacidade de uso de número de linhas ilimitadas nos softwares de SQL, ao contrário do suportado no Excel, sendo a capacidade o único gargalo no manuseio dos dados no caso do SQL;

Vantagem 2: SQL oferece maior velocidade no carregamento de grandes volumes de dados, ganhando em performance.

2. Resposta:

Tablespace: corresponde a uma área lógica para armazenamento. Dentro de uma tablespace são armazenados objetos do banco de dados além dos próprios dados, inclusive os dados sobre o banco de dados (dicionário de dados).

Tabela: é a unidade básica de armazenamento de dados sendo composta por linhas e colunas.

Chave primária: refere-se aos conjuntos de um ou mais campos, cujos valores, considerando a combinação de valores em caso de mais de uma chave primária, nunca se repetem na mesma tabela e, desta forma, podem ser usadas como um índice de referência para criar relacionamentos com as demais tabelas do banco de dados (daí vem o nome banco de dados relacional).

Chave estrangeira: é um campo que aponta para a chave primária de outra tabela ou da mesma tabela. Ou seja, passa a existir uma relação entre duplas de duas tabelas ou de uma única tabela. A finalidade da chave estrangeira é garantir

a integridade dos dados referenciais, pois apenas serão permitidos valores que supostamente vão aparecer na base de dados.

3. Resposta:

Visando uma busca otimizada, o ideal seria a construção de diferentes tabelas para as informações que se repetem, visto que é necessário uma estrutura robusta para a utilização de chaves primárias.

Para as informações de identificação de cada paciente, é construída uma tabela de identificação, chamada a fins de exemplificação de “pacientes”, contendo campos como o CPF, nome, data de nascimento e endereço de residência. Para esse caso, a chave primária será o CPF de cada paciente.

A segunda tabela, por sua vez, chamada de “consultas”, conterá as informações relacionadas à consulta do paciente. Dentre essas informações, teremos a data da consulta, um ID único relacionado a consulta do paciente cuja restrição é a não repetição e o CPF do paciente. Nesse caso, a chave primária será o ID exclusivo da consulta.

Para o caso de relacionamento das informações entre a tabela “pacientes” e “consultas”, a chave estrangeira será o CPF dos pacientes.

Uma terceira tabela seria criada com as informações detalhadas da consulta, chamada “detalhes_consulta”. Nela, será inserido o código da(s) doença(s) diagnosticada(s) durante o atendimento (para esse campo, se o paciente for diagnosticado com mais de um CID, estes estarão em linhas diferentes), o ID exclusivo da consulta, sendo esta coluna utilizada como chave estrangeira para se relacionar com a tabela “consultas”, uma coluna com valor único para a tabela cujo valor será o ID da consulta contido na coluna de ID, acrescentado o CID do diagnóstico do paciente (isto, é, a coluna conterá um único valor, cujo formato

será uma junção do ID da consulta com o CID diagnosticado. Tal formato será importante pois se um paciente obtiver diagnóstico de múltiplos CIDs, a junção tornará a chave como exclusiva, visto que o ID da consulta pode se repetir com a discriminação entre os CIDs, mas a junção de ambos não - exemplo, se o ID da consulta é XXX8 e o CID é YYYY8, então o valor da coluna será XXX8YYYY9) - este campo será a chave primária da tabela, o modo, ou seja, se é convênio pelo SUS ou se é um atendimento particular com a clínica (aqui assumindo se tratar de uma clínica médica privada com convênio com o SUS) - e o CPF do médico que realizou o atendimento.

Outra tabela a ser criada é uma com as informações dos médicos, chamada "profissionais", contendo CPF do médico, nome, data de nascimento e telefone de contato.

Por fim, tem-se a tabela "pagamentos", na qual terá as colunas de ID da consulta, chave primária da tabela e estrangeira no caso de relacionamento com algumas das demais criadas, e outra coluna com o valor fechado da consulta, mais uma contendo a forma de pagamento - se em convênio, dinheiro, débito ou crédito que, a fins de simplificação, serão as únicas aceitas pela clínica.

4. Resposta:

123 ano_avaliacao T	123 nota T	123 frequencia T
2.019	0	141
2.019	1	37
2.019	2	73
2.019	3	165
2.019	4	235
2.019	5	96
2.020	0	144
2.020	1	53
2.020	2	92
2.020	3	183
2.020	4	250
2.020	5	102

5. *Resposta:*

O gênero de filme com lançamento entre os anos de 2011 e 2012 com menor nota é DRAMA. Sua posição no ranking de nota média por gêneros de filmes é 3º lugar. (Script anexado detalha resultados)

ABC genero T↑↓	123 rank_maior_nota_media T↑↓
Romance	1
Acao	2
Drama	3
Suspense	4
Comedia	5

6. *Resposta:*

123 nota_media T↑↓
2,916666667