

SQL MODELO ENTIDADE-RELACIONAMENTO

André Feltoza de Mendonça

INFORMAÇÃO E COMUNICAÇÃO

Invoice

Invoice_id
Customer_id
Order_id
Product_id
Date_time
Status
Total
Remark

SQL MODELO ENTIDADE-RELACIONAMENTO

André Feitoza de Mendonça

INFORMAÇÃO E COMUNICAÇÃO

Invoice

Invoice_id
Customer_id
Order_id
Product_id
Date_time
Status
Total
Remark



Autor

André Feitoza de Mendonça

Bacharel em Engenharia da Computação e mestre em Ciência da Computação pela Universidade Federal de Pernambuco. Atua como gerente de projetos de sistemas voltados à automação do processo de aplicação de exames nacionais (ENADE, por exemplo), no Instituto Nacional de Estudos e Pesquisas Anísio Teixeira, INEP/MEC. Foi analista de processos com vistas à automatização do Processo Decisório do Conselho Deliberativo Nacional do SEBRAE. Possui publicações em Congressos e em revistas especializadas em Ciência da Computação.

Design Instrucional

Sarah Saraiva

Projeto Gráfico

NT Editora

Revisão

Mariana Carvalho

Ricardo Moura

Erick Guilhon

Capa

NT Editora

Ilustração

Eduardo Calazans

Editoração Eletrônica

Marcelo Moraes

Daniel Dias

NT Editora, uma empresa do Grupo NT

SCS Quadra 2 – Bl. C – 4º andar – Ed. Cedro II

CEP 70.302-914 – Brasília – DF

Fone: (61) 3421-9200

sac@grupont.com.br

www.nteditora.com.br e www.grupont.com.br

Mendonça, André Feitoza de.

SQL Modelo entidade-relacionamento / André Feitoza de Mendonça – 1. ed. reimpr. – Brasília: NT Editora, 2017.

152 p. il. ; 21,0 X 29,7 cm.

ISBN 978-85-8416-249-9

1. SQL. 2. Banco de dados.

I. Título

Copyright © 2017 por NT Editora.

Nenhuma parte desta publicação poderá ser reproduzida por qualquer modo ou meio, seja eletrônico, fotográfico, mecânico ou outros, sem autorização prévia e escrita da NT Editora.

ÍCONES

Prezado(a) aluno(a),

Ao longo dos seus estudos, você encontrará alguns ícones na coluna lateral do material didático. A presença desses ícones o(a) ajudará a compreender melhor o conteúdo abordado e a fazer os exercícios propostos. Conheça os ícones logo abaixo:



Saiba mais

Esse ícone apontará para informações complementares sobre o assunto que você está estudando. Serão curiosidades, temas afins ou exemplos do cotidiano que o ajudarão a fixar o conteúdo estudado.



Importante

O conteúdo indicado com esse ícone tem bastante importância para seus estudos. Leia com atenção e, tendo dúvida, pergunte ao seu tutor.



Dicas

Esse ícone apresenta dicas de estudo.



Exercícios

Toda vez que você vir o ícone de exercícios, responda às questões propostas.



Exercícios

Ao final das lições, você deverá responder aos exercícios no seu livro.

Bons estudos!

Sumário

1 INTRODUÇÃO À MANIPULAÇÃO DE DADOS COM SQL	9
1.1 O que são bancos de dados?	9
1.2 SGBD x bancos de dados.....	12
1.3 Projeto de banco de dados.....	15
1.4 Ciclo de vida de um dado e Linguagem SQL	18
2 MODELANDO UM BANCO DE DADOS.....	26
2.1 Conceitos básicos do modelo entidade-relacionamento	26
2.2 Entidades e atributos	28
2.3 Relacionamentos e cardinalidade	32
2.4 Modelo relacional.....	36
3 A CRIAÇÃO DE UM BANCO DE DADOS.....	47
3.1 Esquemas	47
3.2 <i>CREATE TABLE</i>	50
3.3 Criando o banco de dados da escola Caminho Feliz.....	55
3.4 <i>ALTER TABLE</i> e <i>DROP TABLE</i>	59
4 INSERÇÃO E MANUTENÇÃO DE DADOS	65
4.1 Inserção simples de registros e inserção de registros com chave estrangeira ...	65
4.2 Manutenção de dados.....	72
4.3 Gerenciamento de transações	77
5 CONSULTAS SQL.....	82
5.1 Funcionamento de consultas.....	82
5.2 Filtragem e ordenação	86
5.3 Funções.....	90
5.4 Agrupando registros	92
6 JUNÇÕES (JOINS) ENTRE TABELAS	99
6.1 Junções internas (<i>inner joins</i>)	99
6.2 Junções à esquerda (<i>left joins</i>) e junções externas completas (<i>full outer joins</i>).....	105
7 CONCEITOS AVANÇADOS.....	114
7.1 Consultas aninhadas	114
7.2 Visões (<i>views</i>) e gatilhos (<i>triggers</i>).....	119
7.3 Gerenciamento de privilégios	125

8 SISTEMA DA INFORMAÇÃO: GERANDO SCRIPTS SQL.....	131
8.1 Detalhando o sistema da informação.....	131
8.2 Cadastrar professor.....	134
8.3 Cadastrar turma	139
8.4 Relatório professor por estudante	142
GLOSSÁRIO	148
BIBLIOGRAFIA	151

Caro(a) aluno(a),

Bem-vindo aos estudos de **SQL Modelo Entidade-Relacionamento!**

Com a informatização cada vez mais presente em nossas vidas, a quantidade de dados necessária aos negócios cresceu. Empresas e órgãos públicos estão enquadrados nessa realidade. Sistemas ERPs complexos, por exemplo, guardam dados de vários setores de uma empresa, de modo a tornar sua operação mais célere. Adicionalmente, os gestores dessa empresa precisam de dados consolidados, obtidos a partir da operação para tomar decisões. No setor público, situação análoga tem-se configurado de forma cada vez mais frequente, porém, nesse caso, o foco é na qualidade do serviço público prestado.

Para que as organizações façam uso de seus dados, com o intuito de melhorar sua *performance*, é necessário que tais dados virem informação útil. Nesse sentido, os bancos de dados relacionais têm prestado grande contribuição há décadas. Grandes sistemas, aplicações, aplicativos móveis e sistemas *web* guardam seus dados em bancos de dados e os utilizam para passar as informações sob as mais diversas perspectivas a cada tipo de usuário.

A construção do banco de dados de um *software* depende inicialmente dos requisitos de *software*; uma vez que esses são conhecidos, o banco de dados pode ser modelado de forma conceitual, normalmente utilizando-se a abordagem entidade-relacionamento (ER).

Uma boa modelagem é o início de um projeto de banco de dados bem-sucedido. No entanto, uma vez modelado, é necessário, por meio de um sistema de gerenciamento de banco de dados (SGBD), construir o banco de dados que irá ser parte de cada sistema.

Para se implementar um banco de dados relacional, bem como para armazenar os dados e posteriormente recuperá-los, faz-se uso da linguagem SQL. SQL significa *Structured Query Language* ou, em português, Linguagem de Consulta Estruturada. Essa linguagem foi padronizada pela *American National Standards Institute* (ANSI) em 1986. Apesar da padronização, e de haver muito em comum no SQL adotado pelos diversos SGBDs, nem sempre um comando SQL feito para ser executado em um SGBD funcionará sem adaptações em outro.

Neste material, você terá a oportunidade de adquirir e consolidar conhecimentos sobre SQL em oito lições. Conceitos básicos serão apresentados sempre com a utilização de exemplos, e você poderá fixar o conteúdo por meio de questões ao final de cada lição.

Para garantir que você terá condições de utilizar o conteúdo apresentado no mundo real, você terá a oportunidade de realizar atividades práticas, ao fim de cada lição, que simularão o trabalho envolvido na gestão de um banco de dados, desde a sua criação até a emissão de relatórios gerenciais para os diretores de uma empresa.

Nesse contexto, este livro se divide em oito lições, cujos objetivos gerais são:

- compreender as particularidades de um banco de dados relacional para registrar as informações corporativas de empresa;
- construir bancos de dados reais, com vistas a estruturar o conhecimento de uma organização;
- aprender os comandos principais da linguagem SQL para manipulação de dados, a fim de manipular dados corporativos e emitir relatórios operacionais e estratégicos de uma corporação;
- entender conceitos como propriedades *ACID*, restrições de integridade e transações, sendo capaz de enxergar como isso se reflete no uso do SQL, com o fito de garantir segurança nas operações do SGBD selecionado;
- conhecer conceitos avançados (junções, visões e concessões de acesso) de bancos de dados relacionais e construí-los em SQL, visando à geração de conhecimento para tomada de decisão e à garantia da segurança da informação de uma organização.

Não perca tempo! Aproveite esta oportunidade para se tornar um profissional de sucesso.

Bons estudos!

André Mendonça

1 INTRODUÇÃO À MANIPULAÇÃO DE DADOS COM SQL

Pela linguagem SQL, é possível criar tabelas, relacionamentos, inserir, atualizar e deletar dados de uma tabela e recuperar esses dados diretamente ou fazendo cruzamentos e transformações. Esses são apenas alguns dos recursos que a linguagem SQL provê, independentemente do SGBD em uso.

Para aqueles que desejam ser desenvolvedores ou administradores de bancos de dados, o conhecimento da linguagem SQL é imprescindível.

Nesta lição, estudaremos os conceitos iniciais para entendimento da manipulação de dados com SQL. Assim você terá os fundamentos para compreender as lições posteriores, as quais aprofundarão o tema.

Objetivos

Ao finalizar esta lição, você deverá ser capaz de:

- conhecer a definição de banco de dados;
- compreender a diferença entre **SGBD** e banco de dados;
- analisar quais são as etapas envolvidas em um projeto de banco de dados;
- avaliar o ciclo de vida de um dado em uma empresa;
- aplicar os conceitos básicos de SQL e seus tipos de comando.

1.1 O que são bancos de dados?

Antes de definir banco de dados, é de extrema importância que os seguintes conceitos sejam apresentados apropriadamente: dado, informação e conhecimento.

Um dado é um registro de atividades ou ações ocorridas no âmbito de uma empresa. Tais registros são criados, geralmente, por sistemas informatizados. Por exemplo, imagine que uma empresa de telefonia registre todos os horários de ligações entre os seus clientes – pode-se afirmar que cada um desses registros é um dado. Em suma, dados são fatos do mundo real que são registrados para uma possível utilização posterior.

Já a informação se utiliza dos dados como matéria-prima. Realiza-se o processamento desses dados para gerar informações úteis para a tomada de decisão de uma corporação. Ainda utilizando como exemplo a empresa de



SGBD: sistema que apresenta uma série de funcionalidades que visam à definição dos dados, bem como lidam com questões de armazenamento e manipulação de registros da base de dados utilizada.

telefonia, é uma informação o fato de 90% das ligações ocorrerem após as 18 horas. Ou seja, a informação possui um significado que é obtido por meio da associação de dados.

O conhecimento, por sua vez, é uma hipótese que é construída com base nas informações disponíveis de uma empresa, podendo ser definido também como a integração de uma ou mais fontes de informação para a tomada de decisão empresarial. Dessa forma, seria um possível conhecimento, por exemplo, que os clientes da referida empresa de telefonia não são companhias (ou pessoas jurídicas), tendo em vista que a grande maioria das ligações registradas ocorre após o período comercial.



Saiba mais

Uma grande rede de supermercados dos EUA descobriu, ao analisar informações obtidas por meio de relatórios de vendas, que, durante feriados, era bastante frequente o fato de homens comprarem cerveja junto com fraldas. Dessa forma, a empresa gerou o seguinte conhecimento: existe uma alta probabilidade de quem compra fralda também leve cerveja. A partir de então, foi tomada a decisão de colocar os dois produtos disponíveis um ao lado do outro e, conseqüentemente, o resultado foi o aumento de vendas desses itens.



Uma vez que os conceitos de dado, informação e conhecimento foram estabelecidos, a definição de banco de dados pode ser realizada. Afinal de contas, o que são bancos de dados?

Primeiramente, está enganado quem pensa que os bancos de dados surgiram apenas no século passado, com a evolução dos computadores. Os bancos de dados estão presentes na nossa vida cotidiana há muito tempo. Por exemplo, uma lista telefônica de papel é considerada um banco de dados.

Neste livro, considera-se a definição de que um banco de dados é um conjunto de dados inter-relacionados sobre um determinado domínio. Partindo para uma definição mais relacionada ao cotidiano de profissionais de Tecnologia da Informação, um banco de dados é uma coleção de tabelas que podem ou não estar relacionadas.

As tabelas são estruturas de dados que possuem colunas e linhas. As colunas são as propriedades de determinado dado que se quer armazenar. Por sua vez, as linhas são os dados propriamente ditos. Note que um dado de uma tabela pode estar relacionado a outro de uma tabela distinta. Esse processo de vinculação de dados entre tabelas se dá por meio de chaves e será detalhado na lição 2.

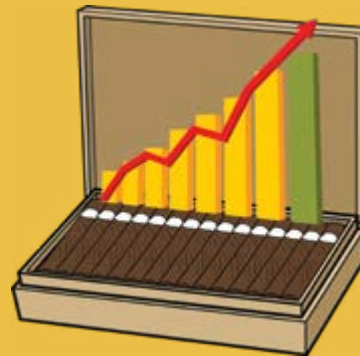
Vale destacar que as informações podem ser obtidas de um banco de dados por meio de relatórios que extraem diversos dados de uma ou mais tabelas. No decorrer deste livro, serão apresentadas diversas formas de obtenção de relatórios ricos para empresas que as auxiliam nas tomadas de decisão corporativas.

Com diversos relatórios em mãos, os gestores de uma empresa podem inferir determinados conhecimentos sobre o ambiente de negócios em que a corporação está atuando. Vale ressaltar que

está fora do escopo deste livro o detalhamento dos métodos de obtenção de conhecimentos relacionados a dados empresariais.

Dicas

Cuidado ao relacionar duas informações para a geração de um possível conhecimento. Certa vez, pesquisadores dos EUA verificaram que a média salarial dos trabalhadores de Boston (cidade americana) e os preços do charuto cubano apresentavam uma forte correlação positiva, ou seja, quando os referidos salários aumentavam, quase sempre, o preço dos charutos também aumentava. Quando esse estudo foi feito, os pesquisadores se esqueceram de considerar que o valor das moedas tende a se desvalorizar com o tempo. Dessa forma, assim como o salário e o preço dos charutos ficaram maiores, o preço de um safári na África, provavelmente, também ficou maior no intervalo de tempo considerado na pesquisa. Ou seja, nesse caso, existia um fator exterior que influenciava as informações a serem analisadas, o que distorce o potencial conhecimento obtido pelos pesquisadores.



No ambiente corporativo, um banco de dados serve para armazenar, por exemplo, dados sobre clientes, produtos e vendas realizadas. A figura 1 ilustra esse banco de dados simples. Observa-se que os dados do cliente e do produto se relacionam para formar uma venda. Por sua vez, a venda possui um preço de realização da transação comercial. Pode-se verificar que cada tabela possui seus respectivos atributos e que cada um deles possui um tipo de dado (como você poderá verificar na lição 2).

Cliente		Produto	
Código Cliente	Nome Cliente	Código Produto	Nome Produto

Venda			
Código Venda	Código Cliente	Código Produto	Preço Venda

Figura 1 – Exemplo de banco de dados

Considerando a figura 1, perceba que cada uma das tabelas apresenta uma coluna que indica um código. Na maior parte das vezes, esses códigos são atributos sequenciais empregados para distinguir um registro de outro. Dessa forma, esses códigos não representam nenhum significado relevante para a empresa.

Você também pode observar que a tabela de vendas apresenta os campos "código cliente" e "código produto". É por meio desses códigos referentes a outras tabelas que, como já mencionado, é realizada a vinculação de uma determinada venda com um cliente e com um produto.



Exercitando

Uma empresa de venda de bolos apresenta uma base de dados para poder gerenciar suas atividades. Analise as seguintes afirmações, que podem ser extraídas do banco de dados:

- o bolo de abacaxi custa R\$15.00;
- 80% das vendas ocorrem nos finais de semana;
- clientes que comprem bolo de chocolate têm grandes chances de comprar refrigerante também.

Assinale a alternativa que classifica de maneira correta cada uma das afirmações.

- Dado, conhecimento, informação.
- Dado, informação, conhecimento.
- Conhecimento, informação, dado.
- Informação, dado, conhecimento.

Comentário: a primeira afirmação se refere a um registro que é armazenado em um banco de dados. No caso em questão, trata-se de um dado sobre um determinado bolo. Já a segunda afirmação é uma informação que é gerada por meio dos dados do banco de dados da empresa. Por fim, a última afirmação indica um conhecimento que é obtido por meio de informações. Dessa forma, a alternativa correta é a "b".

1.2 SGBD x bancos de dados



Atualmente, registros de um banco de dados podem ser alterados por sistemas *online* acessados, muitas vezes, por centenas de usuários ao mesmo tempo. Imagine a quantidade de compras realizadas ao mesmo tempo no site *amazon.com*, por exemplo. Agora, imagine a complexidade de gerenciar essa quantidade de transações simultâneas? Para resolver esses problemas, bancos de dados necessitam de sistemas gerenciadores de bancos de dados (SGBDs).

O principal objetivo de se utilizar um SGBD é tornar a vida dos profissionais de tecnologia da informação mais fácil. Dessa forma, com o uso de um SGBD, você pode abstrair a complexidade das operações. Por exemplo, suponha que, ao mesmo tempo, dois clientes requisitam alterar o mesmo conjunto de dados. Nesse tipo de situação, o SGBD proporciona um controle transacional que impede que o dado seja corrompido, por exemplo.

O SGBD possui uma série de funcionalidades que visam à definição dos dados, bem como lidam com questões de armazenamento e manipulação de registros da base de dados utilizada. Esse tipo de sistema abstrai os detalhes tecnológicos dos bancos de dados utilizados pelos sistemas da informação. Ou seja, quando um sistema necessita gravar uma determinada informação na base de dados, o SGBD proporciona uma interface para que isso seja realizado. Dessa forma, detalhes de implementação são "escondidos".



Saiba mais

Suponha que exista uma tabela de estudantes e que cada estudante seja unicamente identificado pelo seu CPF. Ao tentar inserir um novo estudante nessa base de dados, o SGBD verifica previamente se já existe outro estudante com CPF igual. Caso já exista, o SGBD não permite que o estudante seja inserido novamente. Como será explicado posteriormente na lição 2, esse tipo de comportamento é denominado restrição de integridade. Esse é um exemplo de tarefa que é executada internamente pelo SGBD.



Os SGBDs também são utilizados para armazenar o dicionário de dados. Os dicionários de dados são os metadados referentes ao banco de dados gerenciado. De uma forma resumida, o metadado é o dado que descreve o banco de dados. Assim, são exemplos de metadados, por exemplo, informações sobre quem criou determinada estrutura de dados, a data de última atualização de determinado registro, nome de tabelas e respectivas colunas, etc.

A figura 2 representa uma arquitetura na qual um SGBD está inserido. Pode-se verificar que um SGBD pode operar com um ou mais bancos de dados. Por sua vez, sistemas/programas se utilizam da interface fornecida pelo SGBD para manipular dados relacionados.

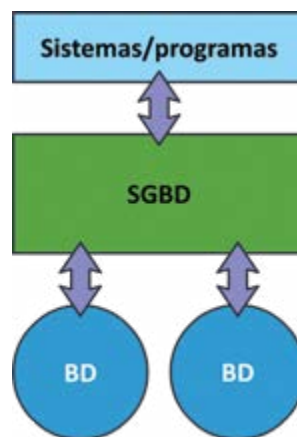


Figura 2 – Arquitetura para utilização de um SGBD

Nesse contexto, os SGBDs proporcionam inúmeras vantagens. Listamos algumas delas a seguir.

- **Múltiplas visões de dados**

Podem-se definir visões distintas para grupos de usuários diferentes. Por exemplo, o grupo de diretores de uma empresa, provavelmente, terá uma visão mais abrangente dos dados da corporação do que a visão de dados fornecida aos estagiários, por conta da experiência.

- **Redundância minimizada**

O corrompimento de dados pode ocorrer em qualquer banco de dados. Os SGBDs proporcionam tratamento de redundância da informação otimizado, o que reduz custos de armazenamento.

- **Gerenciamento de transações**

Dois programas podem, simultaneamente, atualizar o mesmo conjunto de dados. O SGBD realiza controle transacional, de modo que uma operação não afete a outra e ambas não ocorram ao mesmo tempo. Ao fim, é garantido que o banco de dados estará estável.

- **Garantia de integridade dos dados**

Quando uma informação é inserida ou atualizada, o SGBD verifica diversas restrições antes de realizar a operação. Por exemplo, o mesmo dado não pode ser inserido mais de uma vez em uma estrutura de dados.

- **Independência de dados**

A forma com que o dado é armazenado fisicamente não é relevante para os desenvolvedores de sistemas ou programas que manipulam determinada base de dados. O SGBD proporciona essa abstração por meio da linguagem SQL.

Por outro lado, a principal desvantagem de se utilizar SGBD é o custo. Além do custo de aquisição, a atualização e a manutenção desse tipo de sistema de gerenciamento são muito **onerosas**. A empresa deverá, portanto, realizar uma análise de custo-benefício para determinar se irá realizar uma aquisição de um SGBD. Vale ressaltar que, no mercado, existem versões gratuitas de SGBDs, mas que não podem ser utilizadas para fins comerciais.



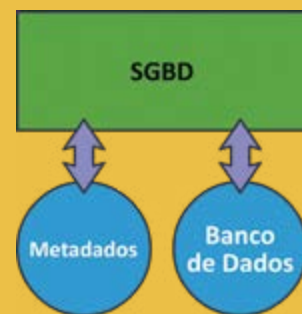
Onerosas: que impõem, envolvem ou estão sujeitas a ônus, encargo, obrigação. Que ocasionam despesas, gastos; dispendiosas.



Saiba mais

Qual é a diferença entre um banco de dados e um SGBD?

Um banco de dados é uma forma de estruturação de dados de um determinado domínio, possuindo os dados propriamente ditos. Já o SGBD está em uma camada acima do banco de dados. Ele proporciona meios para manipular os dados de maneira simplificada. Uma das formas mais utilizadas de realizar a manipulação de dados é o uso da linguagem SQL. Os SGBDs possuem componentes para interpretar e executar comandos em SQL.



Exercitando

Assinale a alternativa que traz a informação incorreta a respeito dos SGBDs.

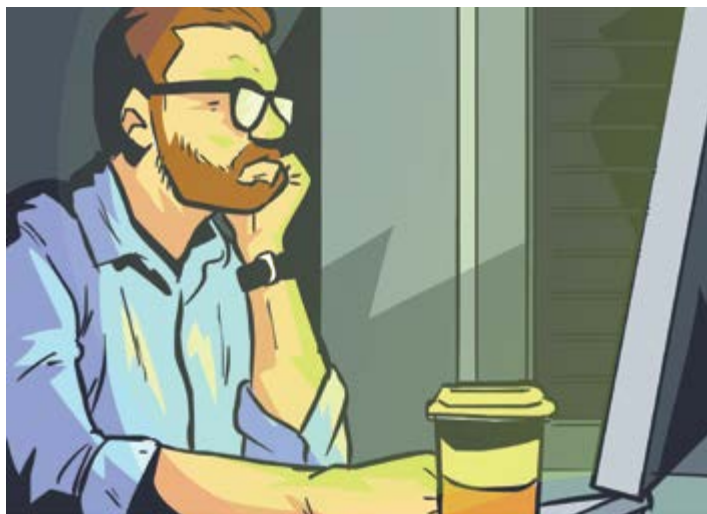
- a) Garantem a integridade dos dados por meio de restrições aos dados que são inseridos numa determinada base de dados.
- b) Reduzem a necessidade de redundância no armazenamento de dados.
- c) Possibilitam que múltiplas visões dos dados possam ser estabelecidas.
- d) Realizam o gerenciamento de transações no banco de dados, não permitindo que múltiplos usuários possam realizar operações.

Comentário: de fato, os SGBDs garantem a integridade dos dados. Existem diversos tipos de restrições de integridade que realizam esse tipo de garantia. É também uma característica dos SGBDs a sua arquitetura interna de armazenamento de dados, que otimiza o espaço ocupado. Por fim, os SGBDs podem criar visões dos dados distintas para os mais diversos usuários. A única alternativa incorreta é a "d". Os SGBDs realizam, sim, o gerenciamento de transações, mas tal mecanismo permite que múltiplos usuários operem nos bancos de dados gerenciados.

1.3 Projeto de banco de dados

Imagine que você foi contratado para criar um banco de dados de uma empresa recém-criada. Qual é o primeiro passo para o projeto de um banco de dados?

De uma forma resumida, nesse tipo de projeto, você deverá iniciar os trabalhos entendendo quais são as necessidades do negócio e, ao final, deverá implementar o banco de dados fisicamente, respeitando restrições tecnológicas e considerando questões de *performance*.



O projeto de banco de dados apresenta as etapas descritas a seguir.

- **Especificação de requisitos de dados**

Nessa etapa, deverão ser mapeados quais entidades do negócio e respectivos relacionamentos deverão ser considerados no projeto. Por exemplo, para uma loja de sapatos, são entidades relevantes os clientes, os calçados e os vendedores.

- **Projeto conceitual**

Aqui, ocorre um detalhamento das entidades especificadas. Assim, seus atributos, relacionamentos e possíveis restrições são especificados. É importante ressaltar que, nessa fase, questões tecnológicas não são consideradas. O que importa é projetar um banco de dados que esteja aderente às necessidades do negócio. Ou seja, o projeto conceitual indica quais dados devem estar presentes no banco de dados e não define como os dados devem ser armazenados fisicamente. Como resultado dessa etapa, normalmente, é produzido um modelo entidade-relacionamento. Esse modelo será detalhado posteriormente na lição 2.

- **Projeto lógico**

Tendo como base o modelo entidade-relacionamento, o projeto lógico mapeia que tabelas e respectivos relacionamentos serão necessários para atender aos requisitos do negócio. Dessa forma, ao fim dessa etapa, é criado o modelo de dados. Geralmente, utiliza-se o modelo de dados relacional (que será detalhado na lição 2). A partir desse modelo, podem ser criados os comandos ou *scripts* necessários para que a estrutura do banco de dados seja construída.

Saiba mais

Geralmente, os projetos lógicos produzem um modelo de dados relacional. Entretanto fique atento! Existem outros tipos de modelos de dados, tais como hierárquico e rede.

No modelo hierárquico, a estrutura de dados é disposta em árvore. Assim, existem registros-pais conectados a um ou mais registros-filhos. Por outro lado, um filho pode ter apenas um único pai. Esse tipo de modelo é apropriado para representar dados com uma hierarquia forte. Por exemplo, o modelo hierárquico se adaptaria bem para representar organogramas corporativos. A Figura 3 representa esse tipo de modelo de dados.



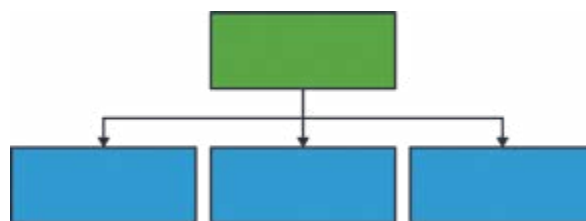


Figura 3 – Modelo de dados hierárquico

Já o modelo em rede elimina a restrição de que um filho tenha apenas um único pai. Assim, por exemplo, nesse tipo de modelagem, um funcionário que, na modelagem hierárquica, poderia estar relacionado a um único chefe, agora pode ter dois chefes. A figura 4 ilustra o modelo em rede.

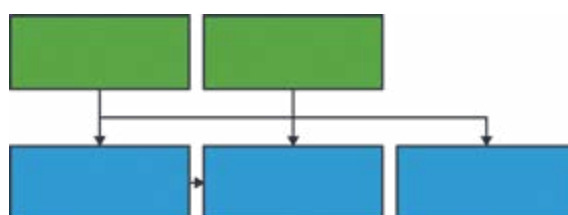


Figura 4 – Modelo de dados em rede

• Projeto físico

Por fim, no projeto físico, são realizados ajustes com vistas à melhoria da *performance* de possíveis consultas ao banco de dados. Assim, são criadas estruturas internas, tais como índices, por exemplo. Nessa etapa são também estabelecidas as permissões de acesso às tabelas criadas, bem como ocorre o dimensionamento da quantidade de dados que uma determinada estrutura de dados terá ou, pelo menos, a previsão de crescimento da mesma. Nessa etapa, pode ocorrer, também, a indicação de junção de dados disponibilizados em tabelas distintas, de forma a melhorar o desempenho de possíveis consultas.



Importante

O projeto físico de um banco de dados é realmente importante?

Essa etapa é muitas vezes negligenciada no projeto de um banco de dados, mas é de fundamental importância para o funcionamento de qualquer sistema. Você não desejará que o seu sistema funcione corretamente quando realiza testes em um ambiente de desenvolvimento (sem uma massa de dados grande)

mas, ao entrar em produção, todos os relatórios fiquem lentos. Esse tipo de situação ocorre, na maioria das vezes, por conta da falta de índices que otimizam a busca de dados em uma determinada tabela. Você também não gostaria que o espaço em disco de determinada tabela estoure e seu sistema pare de funcionar. Um projeto físico bem feito evita esses problemas.



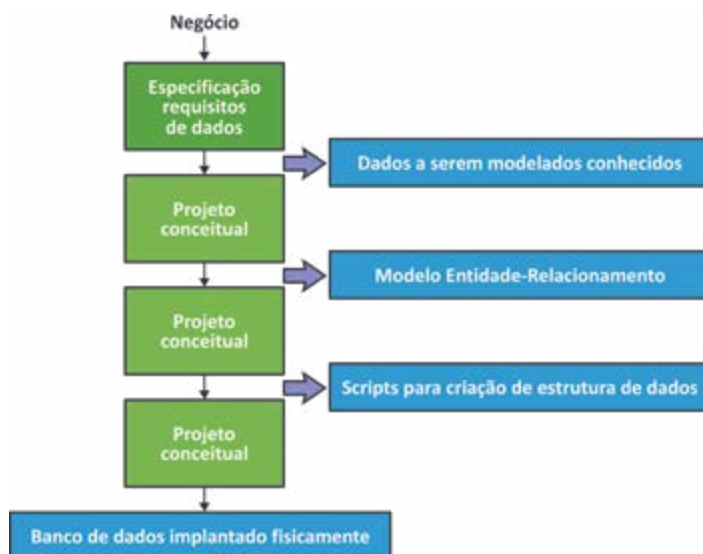


Figura 5 – Esquema resumido das etapas de um projeto de banco de dados

A figura 5 resume as etapas de um projeto de um banco de dados e apresenta os principais artefatos gerados desse conjunto de atividades. Vale destacar que o projeto de um banco de dados não é executado em cascata. Pode ocorrer, por exemplo, que seja feita uma descoberta durante o projeto lógico que demande atualização do modelo entidade-relacionamento (gerado no projeto conceitual).

Dicas

Quais perfis são necessários para realizar o gerenciamento de um banco de dados?

Para gerenciar um banco de dados, além de diversos sistemas gerenciadores, são necessários diversos tipos de profissionais. Os perfis de profissionais mais importantes estão descritos a seguir.

- **Projetista de banco de dados**

É aquele que é responsável por conduzir um projeto de banco de dados. Esse tipo de profissional analisa quais dados comerciais são relevantes para serem armazenados em um banco de dados. Enfim, ele tem de se comunicar com os clientes internos para entender suas necessidades comerciais, a fim de criar um banco de dados compatível.

- **Administrador de dados (AD)**

Esse tipo de profissional desenvolve a modelagem conceitual de um banco de dados. Além disso, ele define padrões de nomes de estruturas de dados e especifica regras de negócio que impactam a forma com que um banco de dados está disposto.

- **Administrador de banco de dados (DBA)**

Profissional responsável por aspectos de segurança e *performance* de um banco de dados. Por exemplo, é ele quem concede acesso ao banco de dados e monitora o seu uso. Além disso, ele realiza o mapeamento do modelo conceitual para o modelo lógico.

- **Desenvolvedores**

Os desenvolvedores de sistemas são usuários de um determinado banco de dados. Basicamente, eles criam *softwares* que se comunicam com as bases de dados por meio da linguagem SQL.





Exercitando

Marque a alternativa que traz a etapa de um projeto de banco de dados que elabora o modelo relacional de dados.

- a) Projeto físico.
- b) Projeto conceitual.
- c) Especificação de requisitos de dados.
- d) Projeto lógico.

Comentário: o modelo relacional de dados é criado no projeto lógico. Tal modelo toma como base o modelo entidade-relacionamento elaborado no projeto conceitual. A partir do modelo relacional de dados, as tabelas podem ser criadas, bem como o relacionamento entre elas. A alternativa correta é a letra “d”.

1.4 Ciclo de vida de um dado e Linguagem SQL



Dentro de uma empresa, existe um fluxo geral percorrido pelos dados. Após a finalização de um projeto de banco de dados, os sistemas transacionais começam a inserir dados referentes à empresa. Esse tipo de sistema, também denominado OLTP (*Online Transaction Processing*), opera cotidianamente com registro de dados operacionais.

Os dados registrados por sistemas OLTP são de extrema importância para o funcionamento de uma empresa. Neles são registrados dados que refletem o dia a dia da empresa e seus processos organizacionais. São exemplos desse tipo de sistema: ERP, CRM, etc.

Entretanto é difícil extrair informações relevantes de tais dados operacionais, seja pela não uniformização dos dados (cada sistema pode registrar dados independentemente dos demais), seja pela quantidade (muitas vezes é necessário o armazenamento de estados internos de sistemas na base de dados para o seu correto funcionamento).

Periodicamente, muitas empresas realizam um procedimento para extrair os principais dados dos sistemas transacionais da corporação, em que elas realizam as devidas transformações para a uniformização dos dados e, em seguida, fazem o carregamento desses dados nos chamados armazéns de dados ou *data warehouses* (DWs). Esse procedimento é denominado de **ETL** (*Extract, Transform, Load*).

Os *data warehouses* são bases de dados que armazenam, ao longo do tempo, dados estratégicos para a corporação. Assim, os armazéns de dados são depósitos de dados orientados por assunto, integrados, não voláteis e variáveis com o tempo.

De uma forma geral, os DWs são modelados de forma multidimensional. Assim, a base de dados possui assuntos que têm uma ou mais dimensões de pesquisa. Por exemplo, um assunto de uma



Processo ETL: procedimento de extração, transformação e carregamento de dados transacionais para sistemas de apoio à decisão.

determinada empresa poderia ser vendas realizadas. Nessa situação, as dimensões possíveis poderiam ser produto e cliente. Ou seja, esse tipo de banco de dados é modelado para responder com agilidade ao relatório de quantidade de vendas de um determinado produto para um perfil específico de cliente.

Ao fim do ciclo de vida de dados corporativos, existem os sistemas de apoio à decisão, também denominados de sistemas **OLAP** (*Online Analytical Processing*), que são capazes de gerar informações para a tomada de decisão de uma empresa. Esse tipo de sistema realiza consultas em *data warehouses*.

Para fins de consolidação, o quadro 1 indica as principais diferenças entre bancos de dados mantidos por sistemas OLTP e OLAP.

Quadro 1 – Comparativo entre bases de dados relacionadas a sistemas OLTP e OLAP

OLTP (bases de dados transacionais)	OLAP (<i>data warehouses</i>)
Dados atualizados a cada transação de sistema transacional.	Dados não são modificados. A cada vez que o processo ETL ocorre, novos registros são inseridos (sem nenhuma atualização).
Os dados são estruturados em um banco de dados relacional.	Os dados são estruturados em um banco de dados multidimensional.
Os dados são manipulados com grande agilidade, mas relatórios gerenciais são gerados com ineficiência.	Alta <i>performance</i> na geração de relatórios estratégicos.
Foco no operacional.	Foco no estratégico.



OLAP: sistemas de apoio à decisão (*Online Analytical Processing*).

A figura a seguir ilustra a forma com que os dados são utilizados em um ambiente corporativo. Pode-se observar que o projeto de um banco de dados é o começo de tudo. Sistemas transacionais inserem dados nas estruturas de dados criadas. Procedimentos ETL geram, periodicamente, bases de dados estratégicas que são consumidas por sistemas de apoio à decisão.

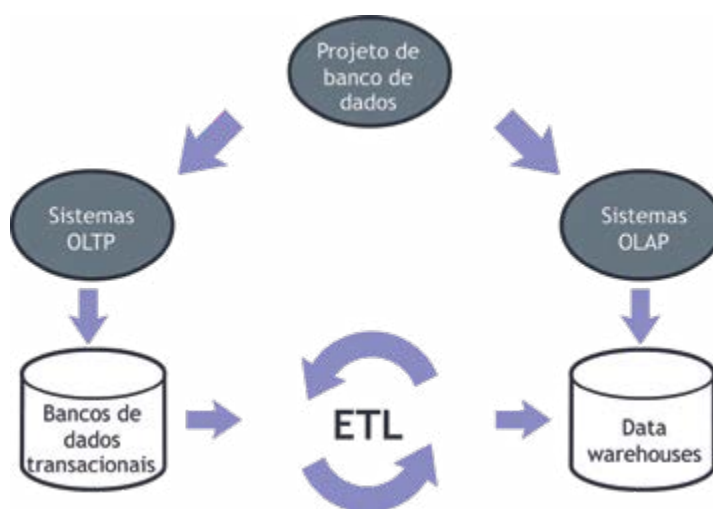


Figura 6 – Ciclo de vida de dados em uma corporação



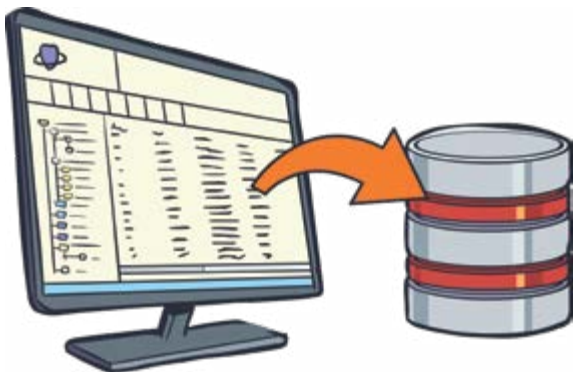
Exercitando

Marque a alternativa que traz as características corretas de um sistema do tipo OLAP.

- a) Estruturado em banco de dados relacional, operacional.
- b) Estruturado em banco de dados multidimensional, estratégico.
- c) Estruturado em banco de dados relacional, estratégico.
- d) Estruturado em banco de dados multidimensional, operacional.

Comentário: sistemas do tipo OLAP apresentam o seu banco de dados estruturado de forma multidimensional. Esse tipo de modelagem garante mais *performance* no momento de geração de relatórios. Esse tipo de sistema possui um caráter estratégico, uma vez que dados não podem ser modificados após terem sido inseridos. Dessa forma, a alternativa “b” é a correta.

Linguagem SQL



Preparado para o que interessa? Vamos aprender mais sobre SQL?

A linguagem SQL (*Structured Query Language* ou Linguagem de Consulta Estruturada) foi criada no início da década de 70 pela IBM e padronizada pela *American National Standards Institute* em 1986.

Cada SGBD possui uma interface específica para que comandos SQL possam ser executados nas suas bases de dados. Dessa forma, apesar de haver uma padronização da linguagem, algumas chamadas SQLs precisarão de adaptação se forem executadas em SGBDs distintos.

A linguagem SQL é declarativa, ou seja, foca apenas no objetivo de uma determinada tarefa a ser executada. Detalhes de implementação não são explicitados e, dessa forma, são necessários poucos comandos SQL para a execução de atividades complexas.

Comandos ou *scripts* em SQL somente podem ser utilizados em bancos de dados relacionais (ou seja, bancos de dados gerados por modelos relacionais). Esse tipo de banco de dados apresenta um conjunto de tabelas ou relações. Cada tabela possui os seguintes elementos principais:

- colunas – conjunto ordenado de atributos da entidade que está sendo armazenada;
- linhas – são registros individuais (ou tuplas) da entidade representada pela tabela em questão. Cada linha de uma tabela possui as mesmas colunas.

Saiba mais

Antes dos bancos de dados relacionais, você tem ideia de como os bancos de dados eram implementados pelas empresas?

Primeiramente, as corporações se utilizavam de fichas de papel organizadas em arquivos físicos. Nesse tipo de organização de dados, a recuperação da informação era extremamente custosa. Imagine que você tem dados de seus clientes em fichas de papel. Para localizar os dados de um determinado cliente, um funcionário da empresa teria de abrir o arquivo físico e procurar, um a um, o registro solicitado. Até mesmo a inserção de um registro novo de cliente nesse tipo de banco de dados não é trivial. O funcionário teria de preencher, numa ficha de papel, os dados do cliente e, em seguida, inserir a ficha em ordem alfabética, por exemplo. Muito arcaico, não acha?

Em seguida, em substituição aos arquivos físicos, surgiram os arquivos digitais. Dessa forma, foram criados *softwares* para gerenciar esses dados armazenados digitalmente, e, assim, as tarefas de armazenar dados e recuperá-los ficaram bem mais fáceis.

Entretanto entidades distintas, nesse tipo de abordagem, não eram relacionadas explicitamente. Esse fato dificulta muito a geração de relatórios que cruzam dados de clientes e produtos, por exemplo. A extração de informações de um banco de dados organizados por arquivos digitais não é trivial.

A linguagem SQL tem como premissa o fato de que os dados manipulados são estruturados. Dados estruturados apresentam sempre o mesmo grupo de atributos e estão organizados em tabelas ou relações.

Existem, ainda, os dados semiestruturados e os não estruturados. Esses últimos não são manipulados diretamente por comandos SQL. Dados semiestruturados apresentam uma estrutura irregular, podendo possuir apenas uma parte dos dados esperados. Como exemplo, há os arquivos XML, que representam dados semiestruturados. Já os dados não estruturados não possuem uma estrutura definida e, normalmente, são caracterizados por imagens e vídeos.

Basicamente, existem dois tipos de comandos SQL principais. Os comandos **DDL** (Linguagem de Definição de Dados) e **DML** (Linguagem de Manipulação de Dados). Os comandos DDL são utilizados para criar tabelas e seus relacionamentos. Assim, com esse subconjunto de instruções SQL, o banco de dados propriamente dito é criado.

Já os comandos DML são empregados para inserção, alterações, consultas e exclusões de dados. Vale ressaltar que, para as instruções DML serem executadas, comandos ou *scripts* DDL deverão ter sido executados previamente. Instruções DML são utilizadas para a geração de relatórios gerenciais tendo como base as diversas informações presentes no banco de dados.

Muitas vezes, desenvolvedores de sistemas não escrevem diretamente consultas SQL para manter dados referentes às suas aplicações. Eles se utilizam de *frameworks* que criam *scripts* SQL a partir de parâmetros para cada operação. Para a linguagem de programação Java, por exemplo, existe o *framework* *Hibernate*.



DDL: linguagem de definição de dados.

DML: linguagem de manipulação de dados.

Apesar de esse tipo de abordagem melhorar a produtividade dos desenvolvedores consideravelmente, uma vez que os programadores não precisam escrever *scripts* SQL enormes, em algumas situações, tais *frameworks* geram consultas ineficientes. Assim, em determinados casos, sistemas podem sofrer com a lentidão excessiva e, às vezes, não funcionarem.

Suponha você deseja criar uma tabela para armazenar dados de clientes de uma determinada empresa. A tabela deverá possuir o nome e a data de nascimento do cliente e seus contatos (telefone e e-mail). O *script* SQL DDL que poderia realizar essa tarefa é o seguinte:

```
CREATE TABLE IF NOT EXISTS cliente (  
  nome VARCHAR (20),  
  sobrenome VARCHAR (30),  
  ddd INT (2),  
  telefone VARCHAR (9),  
  data_nasc DATE,  
  email VARCHAR (30),  
);
```

O *script* DDL, primeiramente, verifica se a tabela já existe no banco de dados antes de criá-la. Em seguida, todos os atributos do cliente são definidos. Ou seja, as colunas da tabela são estabelecidas. Para cada coluna, um tipo de dado é declarado. Por exemplo, o campo "data_nasc" é do tipo *DATE* (data). Agora, suponha que você quer realizar uma consulta de todos os clientes que possuem sobrenome igual a SILVA. Nesse caso, o *script* que seria utilizado para esse fim seria:

```
Select * from cliente where sobrenome = 'SILVA'
```

Comandos SQL são de fácil entendimento. As palavras-chave da linguagem já indicam claramente qual a operação está sendo requisitada. Por exemplo, o termo *insert* indica que um registro de uma tabela está sendo inserido. Outras palavras-chave muito utilizadas são: *select*, *from*, *where*, *create* e *drop*. Mais detalhes poderão ser consultados nas próximas lições.

Existem outros subconjuntos de comandos SQL, tais como:

- DCL (Linguagem de Controle de Dados): operações de concessão de permissões para manipulação de dados e estruturas de dados. As principais palavras-chave são: *grant* e *revoke*;
- DTL (Linguagem de Transação de Dados): instruções relacionadas ao controle transacional de operações. As palavras-chave mais utilizadas são: *commit* e *rollback*.



Exercitando

Assinale a alternativa correta com relação aos subconjuntos de instruções SQL DDL, DML, DCL e DTL.

- Comandos DDL são empregados para criar estruturas de dados. Comandos DML manipulam os dados propriamente ditos.
- Comandos DDL manipulam os dados propriamente ditos. Comandos DML são empregados para criar estruturas de dados.
- Comandos DDL são empregados para criar estruturas de dados. Comandos DCL são empregados para controlar transações.
- Comandos DDL manipulam os dados propriamente ditos. Comandos DTL são empregados para conceder permissões para a manipulação de dados.

Comentário: instruções do tipo DDL definem as estruturas de dados de um banco de dados. Comandos DML são empregados para inserir, alterar e remover dados dessas estruturas de dados. Instruções DCL e DTL são utilizadas para, respectivamente, controlar transações e conceder privilégios a usuários de bancos de dados. Assim, a alternativa “a” é a correta.

Resumindo

Nesta lição, você teve a oportunidade de aprender os principais conceitos que serão empregados no decorrer deste livro. Primeiramente, foi apresentada a distinção das definições de dado, informação e conhecimento e, em seguida, o conceito de banco de dados foi detalhado.

A importância dos SGBDs foi descrita, e as suas principais vantagens foram apresentadas. Entre elas, estão a garantia de integridade dos dados e a possibilidade de múltiplas visões de dados.

Logo em seguida, a forma com que um banco de dados é construído foi exibida por meio do que é chamado de um projeto de banco de dados. Dessa forma, foram apresentadas todas as etapas previstas para esse tipo de projeto. Essas atividades são: especificação de requisitos de dados, projeto conceitual, projeto lógico e projeto físico.

Na sequência, o ciclo de vida de um dado dentro de uma empresa foi exibido. De uma forma geral, o dado nasce em sistemas transacionais (OLTP) e, periodicamente, é transferido por meio do processo de ETL para sistemas de apoio à decisão (OLAP). Por fim, os conceitos básicos sobre a linguagem SQL foram detalhados.

Veja se você se sente apto a:

- conhecer a definição de banco de dados;
- compreender a diferença entre SGBD e banco de dados;
- analisar quais são as etapas envolvidas em um projeto de banco de dados;
- avaliar o ciclo de vida de um dado em uma empresa;
- aplicar os conceitos básicos de SQL e os seus tipos de comando.

Exercícios

Questão 1 – De acordo com o que nós estudamos, um dado é um registro de atividades ou ações ocorridas no âmbito de uma empresa. Qual das alternativas se refere ao exemplo de um dado?

- a) No último ano, 60% dos alunos de um determinado curso passaram no vestibular.
- b) Histórico escolar dos estudantes do ensino médio, de 2016, da Escola União.
- c) Alunos que passam no vestibular de Engenharia, geralmente, costumam tirar notas em matemática acima da média do curso.
- d) No primeiro semestre, apenas 30% dos estudantes foram reprovados por falta.



Parabéns, você finalizou esta lição!

Agora responda às questões ao lado.

Questão 2 – Nós aprendemos que a informação se utiliza dos dados como matéria-prima. Identifique qual das alternativas se refere a um exemplo de informação.

- a) Cadastro de fornecedores.
- b) Fornecedores de uma determinada região têm grande probabilidade de entregar a matéria-prima com atraso no inverno.
- c) Durante o ano passado, 70% das entregas de fornecedores foram realizadas com atraso.
- d) Entregas realizadas pelos fornecedores no inverno.

Questão 3 – O conhecimento, por sua vez, é uma hipótese que é construída com base nas informações disponíveis de uma empresa. Identifique qual das alternativas se refere a um exemplo de conhecimento.

- a) Clientes atendidos no último final de semana apresentaram um nível de satisfação de 80%.
- b) Normalmente, clientes atendidos por funcionários que bateram suas metas apresentam um nível baixo de satisfação.
- c) 70% dos funcionários bateram suas metas de vendas no último ano.
- d) Clientes atendidos no final de semana.

Questão 4 – Assinale a alternativa que não apresenta uma característica de um banco de dado relacional.

- a) Um banco de dados é composto por tabelas que podem ou não se relacionar.
- b) As tabelas possuem colunas e linhas. As linhas são os dados propriamente ditos.
- c) Cada linha de uma determinada tabela pode apresentar uma quantidade diferente de colunas.
- d) As colunas de uma tabela se apresentam em uma determinada ordem.

Questão 5 – Identifique a alternativa que melhor define um banco de dado hierárquico.

- a) Um registro-filho pode ter mais de um registro-pai.
- b) Um registro-pai pode ter somente um único registro-filho.
- c) Um registro-pai pode ter um ou mais registros-filhos.
- d) Registros-filhos podem ter, no máximo, dois filhos.

Questão 6 – Com relação aos sistemas gerenciadores de bancos de dados, marque a alternativa correta.

- a) Esse tipo de sistema possibilita que desenvolvedores insiram dados em bancos de dados sem conhecer detalhes dessa ação. Em alguns casos, isso pode gerar problemas de *performance*.
- b) Devido à padronização da linguagem SQL, um *script* pode executar, sem adaptação, em qualquer SGBD.
- c) SGBDs podem receber comandos de um cliente por vez, a fim de evitar problemas de inconsistência de dados.
- d) Os bancos de dados enviam comandos SQL para os SGBDs para, por exemplo, criar novas estruturas de dados.

Questão 7 – Assinale a alternativa que não apresenta uma característica de um SGBD.

- a) Garantir que dados sejam inseridos de forma consistente.
- b) Possibilitar uma multiplicidade de visões para um mesmo banco de dados.
- c) Permitir a independência dos dados, o que faz com que desenvolvedores não precisem saber nem onde as informações salvas por seus sistemas são armazenadas fisicamente.
- d) Realizar o gerenciamento da concorrência, eliminando tentativas de acessos simultâneos a um banco de dados.

Questão 8 – Suponha que, na véspera do Natal, dez usuários acessam um sistema de comércio eletrônico com a intenção de realizar a compra da última unidade de determinado produto. O sistema permite que, enquanto a compra não for finalizada, qualquer cliente possa visualizar características do produto, informar dados de pagamento e solicitar que a compra seja efetivada. Por azar, os dez usuários solicitaram a efetivação da compra ao mesmo tempo. O SGBD que realiza a gestão do referido sistema seleciona uma transação (que, por sorte, foi enviada milisegundos antes das demais) e efetiva a compra. As demais são canceladas, e nenhuma cobrança é realizada para os outros nove clientes. Qual é a característica de um SGBD que está relacionada com a situação descrita?

- a) Gerenciamento de transações.
- b) Restrição de integridade.
- c) Independência de dados.
- d) Múltiplas visões dos dados.

Questão 9 – Em um sistema de cadastro de fornecedores de uma empresa, o SGBD não permitiu que fosse realizada a inserção de dois fornecedores com o mesmo CNPJ. Qual é a característica de um SGBD que está relacionada com a situação descrita?

- a) Independência de dados.
- b) Restrição de integridade.
- c) Múltiplas visões dos dados.
- d) Gerenciamento de transações.

Questão 10 – Marque a alternativa que indica, na ordem correta, a sequência de etapas de um projeto de banco de dados.

- a) Projeto conceitual, projeto lógico, especificação de requisitos de dados e projeto físico.
- b) Especificação de requisitos de dados, projeto conceitual, projeto lógico e projeto físico.
- c) Especificação de requisitos de dados, projeto lógico, projeto conceitual e projeto físico.
- d) Especificação de requisitos de dados, projeto conceitual, projeto físico e projeto lógico.