






Universidade Federal Rural de Pernambuco
Departamento de Estatística e Informática - DEINFO
Bacharelado em Sistemas de Informação
Disciplina: Introdução ao Armazenamento e Análise de Dados - IAAD
Professora: Roberta Macêdo Marques Gouveia

Estudante (nome completo): Estom Paulino da Silva Junior

Exercício relacionado aos assuntos abordados nos capítulos 01 e 02 do livro “Sistemas de Banco de Dados”, 6ª Edição, de autoria de Ramez Elmasri e Shamkant Navathe.

 Pontuação: 10 pontos (cada questão: 2,0 pontos)

 Prazo: 12/07/2022 (15 dias)

 **Atenção:** está ativada a verificação de plágio do Google Classroom. As questões com respostas iguais entre estudantes ou copiadas da internet serão desconsideradas/anuladas.

Considere o estado/instância do banco de dados Startup apresentado abaixo:

Startup

id_startup	nome_startup	cidade_sede
10001	Tech4Toy	Porto Alegre
10002	Smart123	Belo Horizonte
10003	knowledgeUp	Rio de Janeiro
10004	BSI Next Level	Recife
10005	QualiHealth	São Paulo
10006	ProEdu	Florianópolis

Linguagem_Programação

id_linguagem	nome_linguagem	ano_lançamento
20001	Python	1991
20002	PHP	1995
20003	Java	1995
20004	C	1972
20005	JavaScript	1995
20006	Dart	2011

Programador

id_programador	id_startup	nome_programador	gênero	data_nascimento	email
30001	10001	João Pedro	M	23/06/1993	joaop@mail.com
30002	10002	Paula Silva	F	10/01/1986	paulas@mail.com
30003	10003	Renata Vieira	F	05/07/1991	renatav@mail.com
30004	10004	Felipe Santos	M	25/11/1976	felipes@mail.com
30005	10001	Ana Cristina	F	19/02/1968	anac@mail.com
30006	10004	Alexandre Alves	M	07/07/1988	alexandrea@mail.com
30007	10002	Laura Marques	F	04/10/1987	lauram@mail.com

Programador_Linguagem

id_programador	id_linguagem
30001	20001
30001	20002
30002	20003
30003	20004
30003	20005
30004	20005
30007	20001
30007	20002

Figura 1 - Estado de um BD que armazena informações de startups, linguagens e seus programadores.

Questão 1) Em relação ao esquema do banco de dados Startup apresentado na Figura 1:

a) Especifique mais uma tabela (escolha livre), além das 4 existentes, em seguida informe todos os relacionamentos existentes entre as 5 tabelas do banco de dados Startup.

Projetos_Startup

id_projeto	nome_projeto	id_startup	id_linguagem
50001	Apollo	10002	20003
50002	Big Husky	10002	20003
50003	Atwood	10004	20005
50004	Project X	10006	20001
50005	Agile	10004	20005

Programador -> Startup

Relacionamento um-para-um (1:1)

Startup -> Programador

Relacionamento de um pra muitos (1:n)

Programador_linguagem <-> Linguagem_programacao

Relacionamento muitos pra muitos (n:n)

b) Apresente um exemplo de consulta (de forma textual) a ser realizada neste banco de dados usando conjuntamente as tabelas Startup, Programador, Linguagem_Programação e Programador_Linguagem. (Observação: não é para especificar a consulta em SQL, já que esse assunto só será abordado nas próximas aulas).

Listar a Startup que tenha “João Pedro” que utilize o nome_linguagem Python e seu id_linguagem 20001 e imprima os nomes

c) Suponha que o atributo “id_startup” da tabela Startup tenha sido alterado, sendo incluídas 2 letras ao final dos 5 caracteres/números, cujas letras estão relacionadas às iniciais das cidades sede. Dessa forma, o atributo “id_startup” da tabela Programador também precisou ser alterado. Segue o atributo “id_startup” modificado em ambas as tabelas:

Novo “id_startup” da tabela Startup:	Novo “id_startup” da tabela Programador:
10001PA	10001PA
10002BH	10002BH
10003RJ	10003RJ
10004RC	10004RC
10005SP	10001PA
10006FL	10004RC
	10002BH

Considere a seguinte modificação no BD: a startup Tech4Toy não é mais sediada em Porto Alegre, passando a ser sediada na cidade de João Pessoa. Dessa forma, seu identificador precisará ser alterado de 10001PA para 10001JP. Quais os impactos dessa alteração da cidade sede neste banco de dados? Em sua opinião, essa nova modelagem impactou positivamente ou negativamente no banco de dados Startup? Justifique!

A modelagem de dados dessa maneira traz uma complexidade ao banco de dados e com isso causa um impacto negativo, desta maneira traz alterações nas tabelas Startup e Programador, essas alterações se fazem necessárias a alteração as tabelas que possuem relações com o atributo que está vinculado a PK

Questão 2) Explique cada uma das 5 (cinco) restrições de integridade indicadas abaixo, em seguida, exemplifique um comando (podendo ser uma inserção, atualização ou remoção) que viole cada restrição de integridade.

(Observação: os comandos a serem exemplificados devem ser apresentados de forma textual, ou seja, não é para especificá-los em SQL, já que esse assunto só será abordado nas próximas aulas).

a) Integridade de chave (singularidade) → Exemplificar considerando o atributo id_linguagem.

Restrição específica e exclusiva dos valores dos dados

cada registro de linguagem de programação deverá ter um valor exclusivo para id_linguagem

b) Integridade de domínio → Exemplificar considerando o atributo ano_lançamento.

Valores aceitáveis que a coluna pode aceitar e conter. Os valores inseridos nos atributos devem corresponder às formatações necessárias (numérico, string, cadeia de bits, booleano e etc).

caso o domínio do atributo “ano_lançamento” seja do tipo numérico a violação ocorre ao inserir valores do tipo string.

c) Integridade de vazio/null → Exemplificar considerando o atributo nome_programador.

Esta forma de integridade especifica se valores NULL são permitidos nos atributos, caso não seja permitido não é permitido um valor NULL

o atributo nome_programador é NOT NULL, dessa forma não é permitido um valor NUL, se isso acontecer será uma violação.

d) Integridade referencial → Exemplificar considerando o atributo id_startup.

Esta forma de integridade é especificada como regra de que o valor de uma FK deve ser PK de algum registro na tabela de origem

o atributo id_startup é uma FK na tabela programador, se fosse ser inserido na tabela programador um novo programador passando o id_startup inexistente, ocorrerá uma violação da integridade referencial, pois está sendo feito uma referencia a uma entidade inexistente

e) Integridade semântica (regras de negócio) → Exemplificar considerando o atributo data_nascimento.

São restrições especificadas pelo administrador para controle de valores válidos para os dados condizentes com os requisitos de aplicações, o SGBD deve controlar esta integridade com a implementação de funções básicas.

caso a tabela startup contrate programadores que nasceram antes do data 23/06/1993, o programador “João Pedro” estaria violando o atributo data_nascimento

Questão 3) Existem diversos Modelos de Dados, sendo o Modelo de Dados Relacional um dos mais utilizados na atualidade. O BD Startup ilustrado na Figura 1 utiliza o Modelo de Dados Relacional. Já os modelos Hierárquico e de Rede são exemplos de modelos mais antigos, usados em sistemas/*softwares* legados. Há também os modelos de dados contemporâneos, tais como os Modelos de Dados Não-Relacionais (também conhecidos como NoSQL), que vêm sendo bastante empregados e ganhando adesão de grandes empresas de TI.

Diante desse contexto, defina o termo “Modelo de Dados”, em seguida escolha UM dos Modelos de Dados Não-Relacionais, entre eles: Modelo orientado a coluna; Modelo orientado a documento; Modelo orientado a grafos; Modelo orientado a chave-valor; Multimodelo (ou Modelo Híbrido). Após a escolha de um desses modelos, defina-o e exemplifique como os dados sobre as startups e as linguagens em que seus programadores programam podem ser representados pelo modelo escolhido.

Modelo de dados: Definição de coleção de concetos que podem ser usados para uma estrutura do BD, oferecendo abstração de dados que ocultam detalhes de armazenamento e implementação que não são de interesses dos usuarios do banco de dados

Modelo orientado a documentos; Modelo de banco de dados em que o documento já vem definido com os dados armazenados em sua estrutura, armazenando documentos na parte de valor, tem uma estrutura que se assemelha a uma árvore, vindo dos mapas, coleções e valores escalares

```
{
  "Startup":{
    "id": "10001"
    "nome": "Tech4Toy"
    "cidade sede": "Porto Alegre"
  "Linguagem de Programção":{
    "id linguagem": "20001"
    "nome linguagem": "Python"
    "ano lançamento": "1991"
  "Programador":{
    "id programador": "30001"
    "id": "10001"
    "nome": "Joao Pedro"
    "gênero": "M"
    "data nasc": "23/06/1993"
    "email": "joaop@mail.com"
  "Programador Linguagem":{
    "id programador": "30001"
    "id linguagem": "20001"
  }
}
```

Questão 4) Antes do surgimento do Sistema Gerenciador de Banco de Dados (SGBD), os dados eram armazenados em arquivos tradicionais (diretórios) do sistema operacional. Seguem alguns exemplos de SGBDs que utilizam o modelo de dados relacional: MySQL, PostgreSQL, MariaDB etc. Seguem alguns exemplos de SGBDs que utilizam o modelo de dados NoSQL: MongoDB, Cassandra, Google Cloud Firestore (Firebase) etc. Diante desse contexto, responda:

a) Caso as informações sobre as startups, linguagens e seus programadores fossem armazenadas em arquivos tradicionais do sistema operacional, e não em SGBD, elenque as principais problemáticas e limitações desse tipo de armazenamento de dados.

Dificuldade de acessar os dados, a definição e implantação dos arquivos necessários para uma implantação do software específica é executada de modo "individual" (cada usuário) como uma parte da programação da aplicação;

Redundância não controlada dos dados, pois cada aplicação que faz uso desses dados necessitava ter uma cópia localmente

Desperdício no espaço de armazenamento para manter dados atualizados,

Acesso restrito por tipo de usuário

b) Conceitue a arquitetura de três esquemas dos SGBDs (também conhecida como arquitetura ANSI/SPARC), e explique a diferença entre a independência lógica e a independência física dos dados.

A arquitetura ANSI/SPARC é definida em três níveis: Nível externo, Nível conceitual e Nível interno.

o Nível externo é o nível mais alto de abstração baseado na visão dos usuários em pares que são de interesses, cada esquema externo é comumente implementado usando um modelo de dados representativo, baseado em um projeto de esquema externo em um modelo de dados

O Nível conceitual define a disposição das informações na base de dados, ele descreve a estrutura do banco de dados inteiro para uma comunidade de usuários, se concentra na descrição de entidades, tipo de dados, relacionamentos, de forma resumida ele descreve todos os dados armazenados e as relações existentes

O Nível interno é o nível mais baixo de abstração e define a maneira como os dados são armazenados e como vão ser acessados. Usa um modelo de dados físico e descreve detalhes do armazenamento

Independência lógica é a capacidade de alterar o esquema conceitual, sem ter de alterar os esquemas externos de uma aplicação, modifica o esquema conceitual sem a necessidade de reescrever nada.

Independência física é a capacidade de alterar o esquema interno, sem ter de alterar o esquema conceitual, os esquemas externos não precisam ser alterados, modifica o esquema físico sem a necessidade de reescrever o modelo conceitual

Questão 5) Caso você estivesse implementando um sistema *web* que armazenasse informações semelhantes ao banco de dados Startup da Figura 1, qual arquitetura de SGBD você escolheria, a saber: centralizada, cliente/servidor básica (duas camadas), cliente/servidor de três camadas ou cliente/servidor de n camadas (Conferir a seção 2.5 do Cap. 2)? Justifique a escolha (ou não) de CADA arquitetura.

Observação: idealize que o BD Startup pode se estender, criando filiais pelo Brasil e/ou exterior, incluindo informações dos projetos associados aos programadores, novos requisitos e novas regras de negócio.

Escolheria a arquitetura cliente/servidor de três camadas, já que essa arquitetura permite desempenhar um papel intermediário da execução de aplicações e armazenamento de regras de negócio. Além de melhorar a segurança do BD, conseguindo verificar as credenciais dos clientes antes de encaminhar uma solicitação.

A arquitetura centralizada não funcionaria, uma vez que a interface do usuário e o servidor de banco de dados estão em máquinas diferentes para um sistema baseado na Web.

A arquitetura cliente/servidor básica e arquitetura cliente/servidor de duas camadas: funcionaria se a lógica de negócios puder residir em um servidor diferente do Servidor de Banco de Dados.

Escolheria também cliente/servidor n camadas, já que poderia dividir a lógica de negócios em várias camadas, além de que qualquer camada pode ser executada em um processador e ser tratado independentemente. Além dos benefícios já citados acima ainda temos a questão da criptografia mais avançada tornando mais seguro o transferência de dados.