

	Universidade Federal Fluminense
	Disciplina: Projeto de Banco de Dados para SI
	Professor: Daniel de Oliveira
	Semestre: 2023.1
	Curso: Sistemas de Informação

Avaliação Individual

Leia as Instruções:

- Não há pegadinhas na prova.
- Você tem um prazo razoável para realizar essa prova. Faça com calma!
- É essencial ter comportamento **ÉTICO**. Não copie respostas de colegas!

Considere o esquema relacional a seguir de uma livraria para responder as questões 1 e 2. As chaves primárias se encontram sublinhadas.

```

CLIENTE(cid, cnome, uf) % Identificador, nome e UF (estado) do cliente
LIVRO(lid, titulo, categoria, ISBN, anopublicacao) % Identificador, título, categoria, ISBN e ano de publicação
NOTA_FISCAL(nfid, cid, datacompra, valortotal) % Identificador da nota fiscal, identificador do cliente, data da compra e o preço total da nota
      cid referencia CLIENTE(cid)
ITEM_NF(nfid,lid,quantidade)
      nfid referencia NOTA_FISCAL(nfid)
      lid referencia LIVRO(lid)

```

Considere também as consultas a seguir como as mais significativas para resolver as questões 1 e 2.

- Q1. $\pi_{\text{titulo}}(\sigma_{\text{categoria}='Terror' \wedge \text{anopublicacao} > 2021}(\text{Livro}) \bowtie (\pi_{\text{lid}}(\text{Livro}) - \pi_{\text{lid}}(\text{Item_NF})))$
- Q2. $\pi_{\text{titulo}, \text{ISBN}}((\text{Livro} \bowtie \text{Item_NF}) \bowtie (\sigma_{\text{uf}='RJ'}(\text{Cliente}) \bowtie \text{Nota_Fiscal}))$
- Q3. $\pi_{\text{cnome}}(\sigma_{\text{uf}='AM'}(\text{Cliente}))$
- Q4. $\pi_{\text{cnome}}(\sigma_{\text{categoria}='Romance' \wedge \text{anopublicacao} < 1999}(\text{Livro}) \bowtie \text{Item_NF}) \bowtie (\sigma_{\text{cnome} \neq 'Daniel'}(\text{Cliente}) \bowtie \text{Nota_Fiscal})$
- Q5. $\pi_{\text{cnome}, \text{cid}}(\sigma_{\text{uf}='CE'}(\text{Cliente}))$

- Utilizando a base de dados apresentada anteriormente, faça o projeto de distribuição dessa aplicação considerando os algoritmos de Fragmentação Horizontal Primária e Derivada apresentada em sala de aula. Mostre todos os passos do seu projeto de distribuição, desde a geração do grafo dono x membro (*owner x member*) até a definição dos fragmentos para cada relação. Apresente ao final de seu projeto de fragmentação uma consulta (em SQL ou álgebra relacional) que não se beneficie da fragmentação realizada. **(2,5 pontos)**
- Apresente o projeto de fragmentação da relação LIVRO usando o algoritmo de Fragmentação Vertical apresentado no livro e visto em sala de aula. Mostre todos os passos do projeto de fragmentação. A sua resposta deve conter todas as matrizes (*i.e.*, Uso, AA e CA) intermediárias e cálculos de contribuição e energia de ligação. Considere que A1=lid, A2=titulo, A3=categoria, A4=ISBN e A5=anopublicacao e as seguintes frequências para cada consulta $\text{freq}(Q1) = 100$, $\text{freq}(Q2) = 80$, $\text{freq}(Q3) = 30$, $\text{freq}(Q4) = 50$, $\text{freq}(Q5) = 10$. Apresente ao final de seu projeto de fragmentação uma consulta (em SQL ou álgebra relacional) que não se beneficie da fragmentação realizada. **(2,5 pontos)**

3. A empresa D&D Engenharia executa projetos. Em seu sistema de controle de projetos encontramos duas tabelas com a seguinte estrutura (PKs se encontram sublinhadas) **(2,0 pontos)** :

```
Projeto(IdProjeto:int, Orçamento:double, Cidade:varchar(50))
Despesa(TipoDespesa:char(1), NrDocumento:int, IdProjeto:int, Valor:double)
        IdProjeto referencia Projeto(IdProjeto)
```

Considere três soluções de indexação ou ordenação das tabelas de **Despesa** e **Projeto**:

- (a) índice na forma de árvore B+;
- (b) índice hash;
- (c) não há índice, mas há um *script* Python de ordenação externa para este campo implementada pelo usuário.

Para os seguintes cenários alternativos apresentados a seguir recomende uma ou mais das soluções (a), (b) ou (c) e justifique sua resposta. Indique em que campo ou campos serão aplicados os índices. Apresente o *script* de criação dos índices (quando for o caso). Comente aspectos como consumo de espaço em disco e desempenho de operações de INSERT, UPDATE e DELETE.

- O engenheiro-chefe comumente submete consultas que retornam as despesas com valor igual a R\$100.000,00 **(0,5 pontos)**
 - Existe uma rotina de controle de projeto que precisa do total de despesas de um dado projeto a partir do seu valor de IdProjeto. **(0,5 pontos)**
 - Ao final de cada mês é gerado um relatório apresentando o total de despesas por projeto da cidade do Rio de Janeiro **(0,5 pontos)**
 - Os dados da tabela **Despesa** são usados diariamente por rotinas analíticas que fazem cruzamentos de dados e respondem perguntas como: Há despesas de material de consumo pagas pela contratante (TipoDespesa = 'Contratante')? **(0,5 pontos)**
4. Considere a relação R(A:int,B:varchar(20),C:varchar(10),D:int) e a consulta SELECT B,C FROM R WHERE D > 10. Um colega seu criou o seguinte índice para acelerar essa consulta: CREATE INDEX idx_D_R ON R USING HASH (D). Esse índice é o mais indicado para essa consulta? Justifique sua resposta. **(1,0 pontos)**
5. Considere que você iniciou em uma nova empresa como DBA (*Database Administrator*) e está muito empolgado. Entretanto, não há documentação e você não conhece nada do banco de dados da empresa. Dessa forma, começou a explorar os diversos *scripts* disponíveis nos diretórios da rede. Um dos *scripts* encontrados é apresentado a seguir e foi feito por um colega que não trabalha mais na empresa. Esse *script* cria duas tabelas chamadas de **Pessoas** e **Saldos** e um índice associado à tabela **Saldos**.

```
CREATE TABLE Pessoas (ID INT, Nome VARCHAR(50))
CREATE TABLE Saldos (DataSaldo DATE, IDPessoa INT, Valor SMALLMONEY)
ALTER TABLE Pessoas ADD PRIMARY KEY (ID)
ALTER TABLE Saldos ADD PRIMARY KEY (DataSaldo, IDPessoa)
ALTER TABLE Saldos ADD FOREIGN KEY (IDPessoa) REFERENCES Pessoas (ID)
INSERT INTO Pessoas VALUES (1,'Daniel'),(2,'Wagner'),(3,'Marcos'),(4,'Maria')
INSERT INTO Saldos VALUES ('2012-08-15',4,5000)
CREATE INDEX IDX_Pessoa ON Saldos (IDPessoa)
```

Uma vez que as tabelas estão criadas, a seguinte consulta foi submetida por você a pedido de seu chefe para retornar o nome e os registros de saldo de um cliente específico:

```
SELECT P.Nome, S.* FROM Pessoas As P INNER JOIN Saldos As S ON P.ID = S.IDPessoa WHERE P.ID = 4
```

O índice **IDX_Pessoa** que foi criado por seu colega de trabalho seria útil para essa consulta? Justifique sua resposta. **(2,0 pontos)**