



Critérios de Correção

SISTEMAS DE GESTÃO DE BASES DE DADOS | 21103 | ÉPOCA RECURSO

Período de Realização: decorre 24-07-2025 deste 10:00 com 1:30 horas de duração e tolerância de 0:15.

Data de Limite de Entrega: até 11:45 de Portugal Continental.

Temática / Tema / Conteúdos: Sistemas de gestão de bases de dados

Objetivos: Reconhecer formas de armazenamento de dados e formas de otimização de consultas; reconhecer o sistema transacional e formas de recuperação de dados; reconhecer ambientes de Data Warehouse, Data Mining e Information Retrieval.

Trabalho a desenvolver: ambiente WISEflow, tipo FLOWlock

Critérios de avaliação e cotação: A cotação deste e-fólio é de 120 pontos = 12 valores, pode encontrar as cotações parciais junto de cada pergunta. A interpretação das perguntas também faz parte da sua resolução, se encontrar alguma ambiguidade deve indicar claramente como foi resolvida. Critérios de avaliação gerais: (i) para a dificuldade de leitura (linhas cruzadas, letras com fontes desadequadas) a penalização é de 20% a 100%; (ii) para erros e omissões a penalização é de 20% a 100%.

Normas a respeitar: Deve redigir o seu E-fólio na Folha de Resolução disponibilizada na turma e preencher todos os dados do cabeçalho. Podem ser incluídas imagens e digitalizações de conteúdos produzido manualmente pelo estudante. Todas as páginas do documento devem ser numeradas. O documento A4 deve ser redigido em Times New Roman, tamanho de letra 12. O espaçamento entre linhas deve corresponder a 1,0 ou 1,5 linhas. Nomeie o ficheiro com o seu nome da plataforma. Finalmente deve gerar um PDF do documento. Deve carregar o referido ficheiro para a plataforma no dispositivo até à data e hora limite de entrega. Evite a entrega próximo da hora limite para se precaver contra eventuais problemas. O ficheiro a enviar não deve exceder 8 MB.

Votos de bom trabalho!

A informação da avaliação do estudante está contida no vetor das cotações:

Questão: 1 2 3 4 5 6

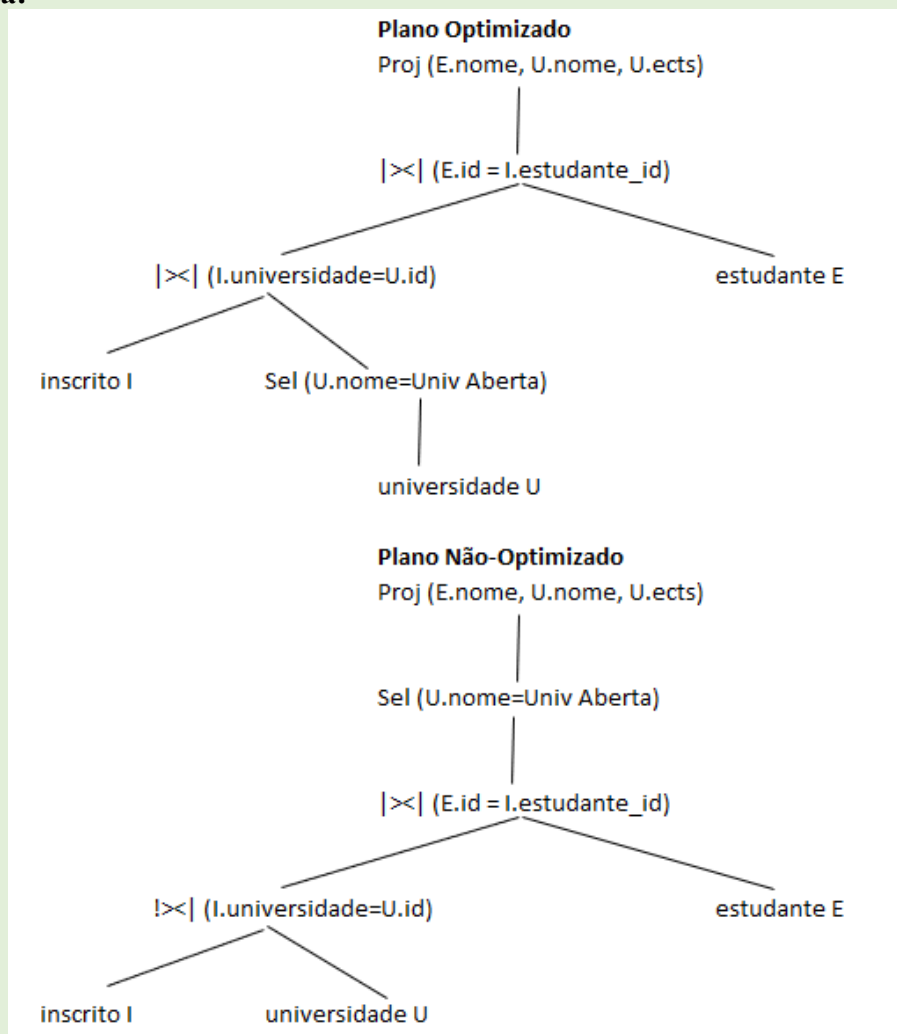
Cotação: 2 2 2 2 2 2

Tópico 1 – Consultas e Transações

1. (2 valores) Otimização de consultas de um SGDB. Para a seguinte consulta desenhe duas árvores com os planos de execução de eficiência diferente, um otimizado e outro não otimizado.

```
select E.nome, U.nome, U.ects
from estudante E, inscrito I, universidade U
where E.id = I.estudante_id
and I.universidade_id = U.id
and U.nome = 'Universidade Aberta'
```

Resposta:



Critérios de correção:

- 1,0 valor, plano otimizado com Seleção nas folhas da árvore (antes das junções)
- 1,0 valor, plano não-otimizado com Seleção depois das junções
- erros, omissões, redundâncias: -20% a -100%

2. (2 valores) Concorrência no SGBD. Considere o seguinte sequenciamento r1(A), r2(B), r2(A), r1(B), w1(A), w2(B). O referido sequenciamento é serializável pelo 2PL com ‘locks’ exclusivos e partilhados? Justifique a resposta usando a seguinte tabela.

T1	T2	justificação

Resposta:

T1	T2	T1	T2	justificação
r(A)		x-lock(A), read(A)		lock exclusivo, write(A)
	r(B)		x-lock(B), read(B)	lock exclusivo, write(B)
	r(A)		read(A)	T2 waiting
r(B)		read(B)		T1 waiting
w(A)		write(A)		deadlock
	w(B)		write(B)	

Conclusão: O sequenciamento NÃO é serializável pelo 2PL com ‘locks’ exclusivos e partilhados.

Critérios de correção:

- 1,0 valor, tabela com transações e justificações
- 1,0 valor, resposta final e justificação
- erros, omissões, redundâncias: -20% a -100%

3. (2 valores) Recuperação do SGBD. Considere o seguinte Log e preencha as colunas da tabela seguinte nas fases de Redo e Undo.

Log	Redo-Phase	Undo-Phase
Start T1		
Read T1		
Commit T1		
Checkpoint		
Start T2		
Read T2		
Start T3		
Checkpoint		
Write T2		
Commit T2		
Read T3		
Write T3		
System crash <-----		

Resposta:

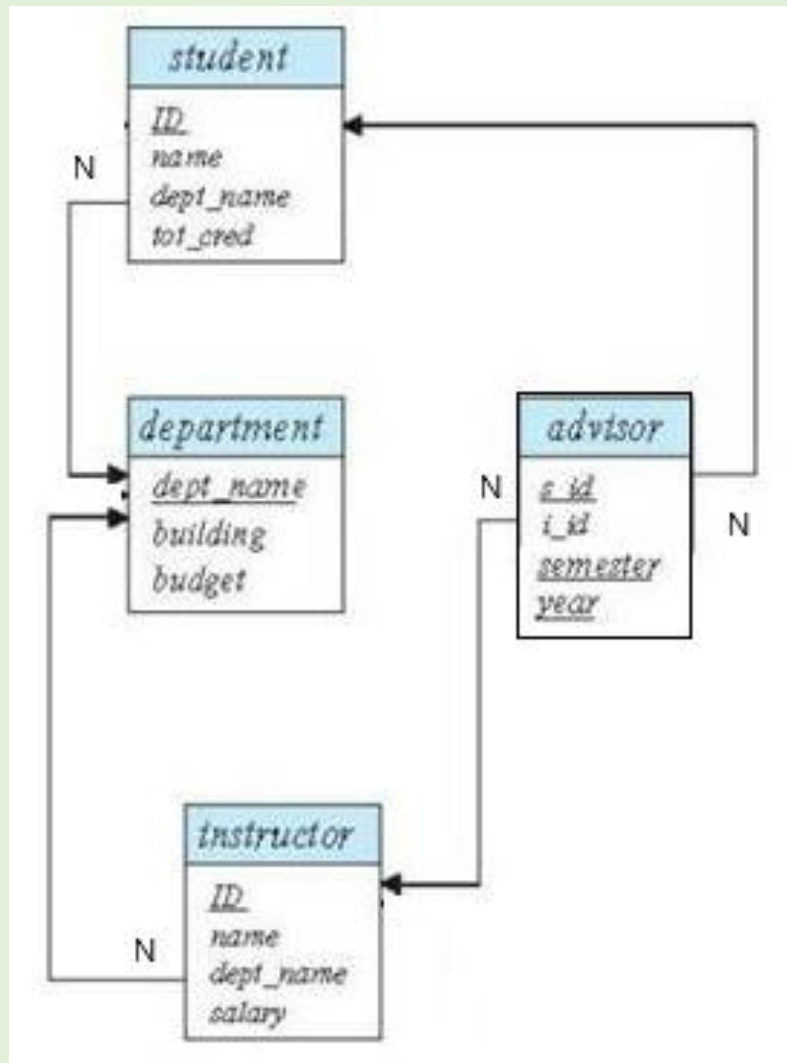
Log	Redo-Phase	Undo-Phase
Start T1	undo-list=[T1]	Undo-Phase end
Read T1		↑
Write T1	Redo T1	
Commit T1	undo-list=[]	
Checkpoint		
Start T2	undo-list=[T2]	
Read T2		undo-list=[]
Start T3	undo-list=[T2,T3]	undo T3
Checkpoint		
Write T2	Redo T2	
Commit T2	undo-list=[T3]	
Read T3		
Write T3	Redo T3	undo-list=[T3]
System crash <-----	----	Undo-Phase start
Redo T2		
Redo T3		
Undo T3		↓

Critérios de correção:

- 1,0 valor, tabela com coluna de redo-phase
- 1,0 valor, tabela com coluna de undo-phase
- erros, omissões, redundâncias: -20% a -100%

Tópico 2 – Data Warehousing, Mining e Information Retrieval

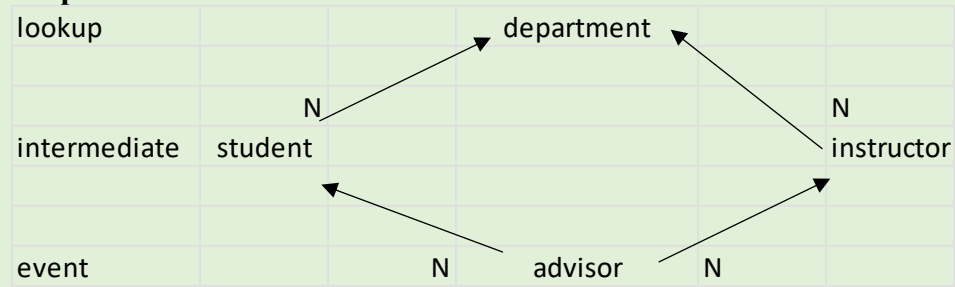
4. (2 valores) Considere a seguinte base de dados que vai servir de fonte de dados a um “Data Warehouse”.



Considere ainda, as seguintes formas desnormalizadas (FDs):

- Antes da 1FD: constituída por várias poli-árvores interligadas;
- 1FD: constituída por várias poli-árvores separadas;
- 2FD: constituída por várias árvores separadas, depois da replicação das tabelas intermédias e de lookup para evitar caminho múltiplos; esta FD é equivalente ao esquema em estrela/floco-de-neve;
- 3FD: o processo de desnormalização termina com a junção de todas as tabelas da árvore com vista a uma rápida leitura dos dados.

Em que forma de desnormalização se encontra a bases de dados? Justifique a resposta e identifique os caminhos múltiplos no caso de existirem.

Resposta:

Em que forma de desnormalização se encontra a bases de dados? 1FD dados que existe uma poli-árvore separada.

Justifique a resposta e identifique os caminhos múltiplos no caso de existirem: Existem dois caminhos múltiplos department-advisor

Critérios de correção:

- 1,0 valores, representação da 1FD
- 1,0 valores, identificação dos caminhos múltiplos
- erros, omissões, redundâncias: -20% a -100%

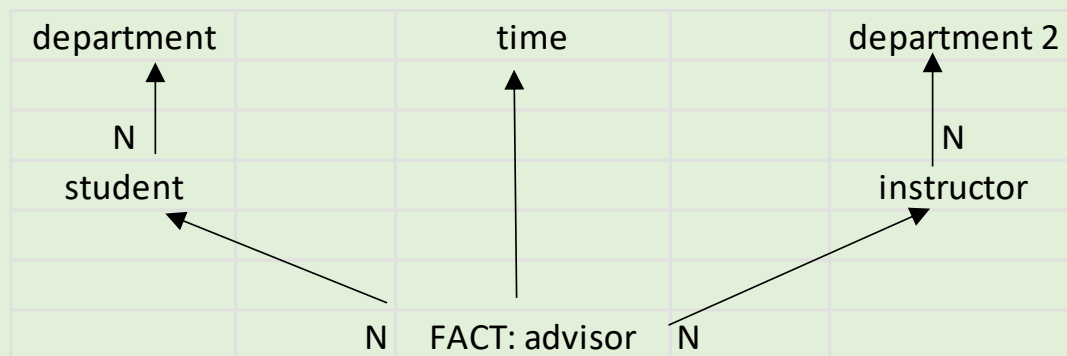
5. (2 valores) Considere a bases de dados da pergunta 4, desenhe um Data Mart, em estrela/floco-de-neve, com uma tabela de factos e com pelo menos três dimensões.

Considere a seguinte hierarquia de tipos de atributos de uma tabela:

- Aditivos: são atributos que podem ser agregados (somados) por todas as dimensões, ex: valor da venda (usar Sum()) sempre)
- Semi-aditivos: são atributos que podem ser agregados (somados) por algumas as dimensões, ex: quantidade (usar Sum() em condições particulares)
- Não-aditivos: são atributos que não podem ser agregados (somados), ex: preço unitário (usar Average() por exemplo)
- Sem factos: só existem identificadores (usar a função Count() dos identificadores).

Para a tabela de factos defina a maior hierarquia do tipo de atributos que apresenta. Justifique a resposta.

Resposta:



Conclusão: Tabela de factos sem factos, só existem identificadores (usar a função Count() dos identificadores).

Critérios de correção:

- 1,0 valor, tabela de factos com 3 dimensões, deve incluir o Tempo
- 1,0 valor, tipo de atributos: Sem Factos
- erros, omissões, redundâncias: -20% a -100%

6. (2 valor) Considere os dados da pergunta 4 e 5, e crie uma consulta OLAP (usando uma tabela de duas entradas) que evidencie o esforço (em número de estudantes) de cada professor (instructor) ao longo dos anos. Exemplifique a tabela de duas entradas e a consulta em SQL.

Resposta: Para a tabela com 10 linhas

Year	Instructor	Student
2023	Prof. Ana Martins	Lucas Oliveira
2022	Prof. Paulo Souza	Mariana Costa
2024	Prof. Ana Martins	João Pereira
2023	Prof. Bruno Mendes	Aline Rocha
2022	Prof. Paulo Souza	Fernanda Dias
2024	Prof. Bruno Mendes	Gabriel Almeida
2023	Prof. Carla Lima	Sofia Fernandes
2022	Prof. Paulo Souza	Diego Santos
2024	Prof. Paulo Souza	Camila Teixeira
2023	Prof. Ana Martins	Rafael Carvalho

Obtemos a seguinte consulta OLAP dos professores versus ano:

Year	2022	2023	2024	Total
Prof. Ana Martins		2	1	3
Prof. Bruno Mendes		1	1	2
Prof. Carla Lima		1		1
Prof. Paulo Souza	3		1	4
Total	3	4	3	10

Com o seguinte código:

```

TRANSFORM Count(Fact.[student]) AS CountOfstudent
SELECT Fact.[instructor], Count(Fact.[student]) AS [Total Of student]
FROM Fact
GROUP BY Fact.[instructor] PIVOT Table1.[year];

```

Critérios de correção:

- 1,0 valor, tabela pivot com 2 entradas
- 1,0 valor, código SQL
- erros, omissões, redundâncias: -20% a -100%

FIM