



CrITÉrios de Correção

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

Período de Realização: decorre 29-01-2026 deste 15:00 com 1:30 horas de duração e tolerância de 0:15.

Data de Limite de Entrega: até 16:45 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) Relativo ao tema das transações, considere o seguinte sequenciamento $r1(X), r1(Y), r2(X), r2(Y), w1(X), w2(X), w2(Y), w1(Y)$. O referido sequenciamento é serializável a conflitos? Justifique a resposta.

Resposta:

Serialização a conflitos: um escalonamento é seriável por conflitos se o seu grafo de precedência não contiver ciclos. Na construção do grafo de precedências as arestas são montadas a partir das observações das transações que participa, da escala sendo duas transações T_i e T_j haverá uma aresta: $T_i \rightarrow T_j$ se forem observadas as seguintes condições:

- T_i executa *write*(q) antes de T_j executar *read*(q)
- T_i executa *read*(q) antes de T_j executar *write*(q)
- T_i executa *write*(q) antes de T_j executar *write*(q)

Para o sequenciamento: $r1(X), r1(Y), r2(X), r2(Y), w1(X), w2(X), w2(Y), w1(Y)$

T1	T2				
r(X)					
r(Y)					
	r(X)				
	r(Y)				
w(X)					
	w(X)				
	w(Y)				
w(Y)					

		RW por recurso	Precedências
item X	r1, r2, w1, w2	r1->w2, r2->w1, w1->w2	
item Y	r1, r2, w2, w1	r1->w2, r2->w1, w2->w1	

Conclusão: Existe um ciclo, $T1 \rightarrow T2 \rightarrow T1$, pelo que o escalonamento não é seriável a conflitos.

Critérios de correção:

- 1,0 valor, divisão por recurso/item e precedências
- 1,0 valor, conclusão e justificação
- erros, omissões, redundâncias: -20% a -100%

2. (2 valores) Considere o sequenciamento da pergunta 1. O referido sequenciamento é serializável a vistas? Justifique a resposta.

Resposta:

Serialização por vistas: Diz-se que “ T_i lê x de T_j ”, se $W_j(x)$ for a última operação de escrita antes de $R_i(x)$. Dois escalamentos são equivalentes a vistas, se a relação “lê de” em ambas for a mesma.

Dois escalonamentos S1 e S2 são equivalentes se:

- Todos os elementos A têm o mesmo valor inicial
Se S1: $r_i(A)$ lê o valor inicial, então S2: $r_i(A)$ também lê o mesmo valor inicial
- Um elemento A tem a mesma vista nos dois escalonamentos
Se S1: $w_j(A) \Rightarrow r_i(A)$ então em S2: $w_j(A) \Rightarrow r_i(A)$; têm a mesma vista
- Todos os elementos A têm a mesma vista final
Se S1: T_i termina $w_i(A)$, então S2: T_i também termina com $w_i(A)$; os valores finais são os mesmos.

S1			S2		
T1	T2		T1	T2	
r(X)			r(X)		
r(Y)			r(Y)		
	r(X)		w(X)		
	r(Y)		w(Y)		
w(X)				r(X)	
	w(X)			r(Y)	
	w(Y)			w(X)	
w(Y)				w(Y)	
S1	X	Y	S2	X	Y
inicial R	T1	T1	inicial R	T1	T1
update RW			update RW		
final W	T2	T1	final W	T2	T2

Conclusão: S1 e S2 não são equivalentes, pelo que a sequencia não é seriável a vistas

Critérios de correção:

- 1,0 valor, tabela inicial R e final W
- 1,0 valor, conclusão e justificação
- erros, omissões, redundâncias: -20% a -100%

3. (2 valores) Considere o sequenciamento da pergunta 1 e 2. O referido sequenciamento é serializável pelo 2PL com 'locks' exclusivos e partilhados? Justifique a resposta.

Resposta:

O protocolo 2-PL (*Two-Phase Locking Protocol*), requer que cada transação recorra a bloqueios, garantindo assim a seriabilidade, para aceder aos recursos em duas fases:

1- Fase de Crescimento: Uma transação pode obter bloqueios, mas não pode libertar nenhum;

2- Fase de Encolhimento: Uma transação pode libertar bloqueios, mas não consegue obter nenhum bloqueio novo.

Este protocolo, apesar de garantir a serialização não garante a recuperação a falhas nem ao 'deadlocks'.

Em 2-PL usamos a seguinte matriz de compatibilidade de 'locks':

State of the lock	Lock request type	
	Shared	Exclusive
Shared	Yes	No
Exclusive	No	No

T1	T2	comentários
x-lock (X), r(X)		
x-lock(Y), r(Y)		
	r(X)	T2 aguarda lock
	r(Y)	T2 aguarda
w(X)		T2 aguarda
	w(X)	T2 aguarda
	w(Y)	T2 aguarda
w(Y)		T2 aguarda
unlock(X)		T2 aguarda
unlock(Y)		
	x-lock (X), r(X)	T2 avança
	x-lock(Y), r(Y)	
	w(X)	
	w(Y)	
	unlock(X)	
	unlock(Y)	

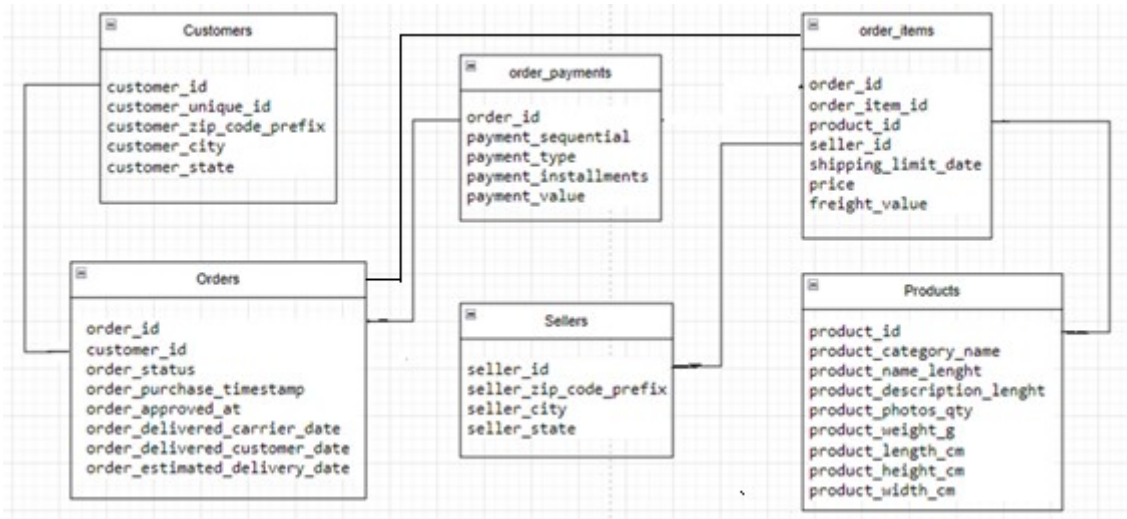
Conclusão: a sequência é serializável pelo 2PL

Critérios de correção:

- 1,0 valor, tabelas com transações e justificações
- 1,0 valor, resposta final e justificação
- 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):

- 1FD – constituída por árvores e poli-árvores separadas;
- 2FD – constituída por várias árvores separadas, com a replicação das tabelas intermédias e de lookup que forem necessárias para evitar caminho múltiplos; esta FD é equivalente ao esquema em estrela ou ao esquema 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 respostas usando a seguinte tabela e referindo as ligações 1-N e os caminhos múltiplos, caso existam.

Antes 1 FD	
1FD	
2FD	
3FD	

Resposta:



Antes 1 FD	resposta existem duas árvores não separadas
1FD	NA
2FD	NA
3FD	NA

Critérios de correção:

- 1,0 valores, resposta que existem duas árvores não separadas
- 1,0 valores, justificação
- erros, omissões, redundâncias: -20% a -100%

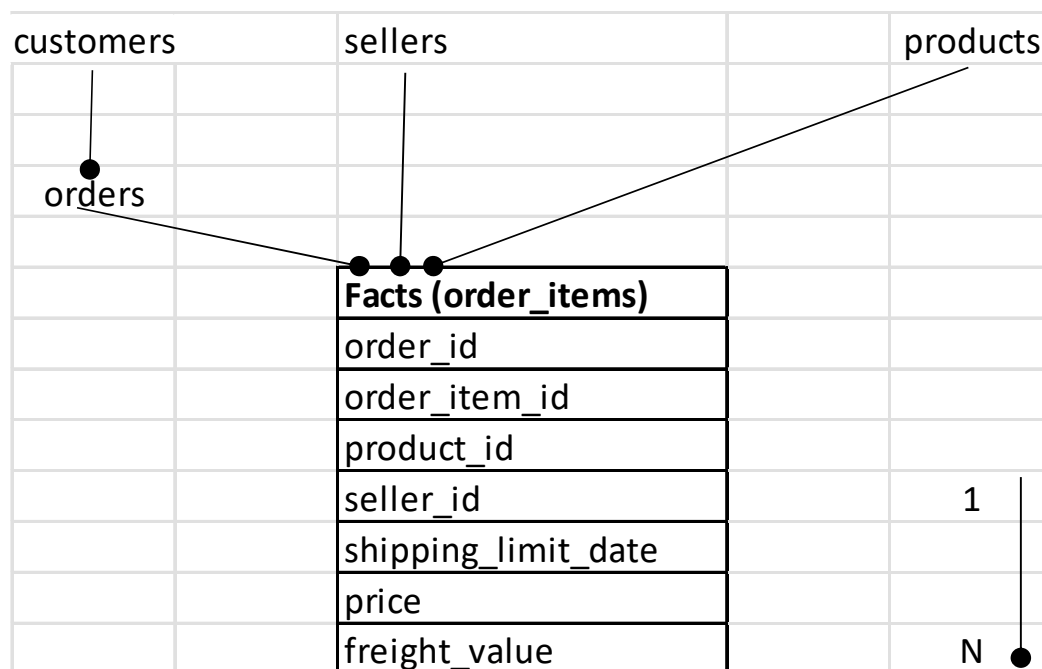
5. (2 valores) Considere a bases de dados da pergunta 4, desenhe um Data Warehouse, com as tabelas de factos com pelo menos três dimensões.

Considere os seguintes tipos de atributos da tabela de factos, do mais elevado para o menos elevado:

- 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 as tabelas de factos defina o tipo de atributos mais elevado que elas apresentam. Justifique a resposta.

Resposta:



Os atributos da tabela de Factos são Aditivos, dado que podemos encontrar um valor para cada Order item para os atributos Price e Freight value.

Critérios de correção:

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

6. (2 valor) Data Mining

Para a base dados da pergunta 4, escreva em português corrente perguntas para serem respondidas usando:

6.1) um algoritmo de associação

6.2) um algoritmo de agrupamento

6.3) um algoritmo de classificação

Nas respostas detalhe as estruturas de dados que devem ser utilizadas. Justifique a resposta.

Resposta:

6.1	associação	quais os N produtos mais comprados em conjunto algoritmo: Apriori
6.2	agrupamento	quais os K grupos de clientes, ex: k=3 algoritmo K-means
6.3	classificação	quais as variáveis que melhor explicam o preço algoritmo regressão múltipla

Critérios de correção:

- 1,0 valor, pergunta em português corrente
- 1,0 valor, estrutura de dados ou algoritmo a utilizar
- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%