



## **SISTEMAS DE GESTÃO DE BASES DE DADOS | 21103**

### **Resolução e Critérios de Correção**

#### **Data e hora de realização**

21 de fevereiro de 2024, às 15h00 de Portugal Continental

#### **Duração da prova**

2h00 + 30m de tolerância

#### **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**

Resolução de um conjunto de exercícios.

#### **Critérios de avaliação e cotação**

A cotação deste e-fólio global 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 número de estudante, seguido da identificação do e-fólio, segundo o exemplo apresentado:  
000000EfólioGlobal.

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!

Paulo Pombinho

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

Questão: 1	2	3	4	5	
Cotação: 2	3	2	3	2	valores

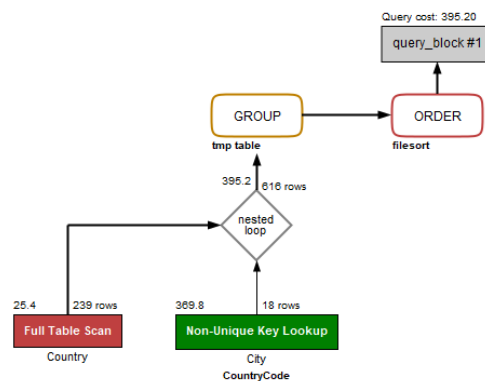
1. (2 valores) Considere a base de dados World que contem as tabelas City, Country e CountryLanguage:

City (id -> name, countrycode, district, population, ...)

Country (code -> name, continent, region, surfaceArea, indepYear, ...)

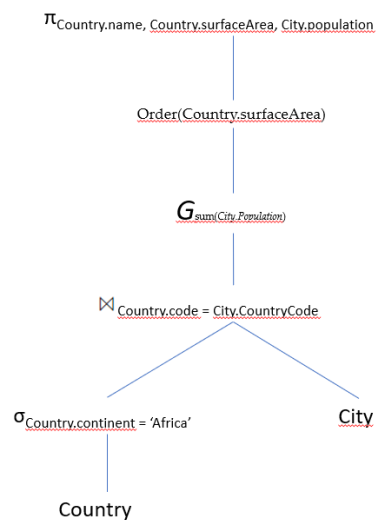
Countrylanguage (countrycode, language -> isOfficial, perctange, ...)

Considere que foi feita uma consulta em SQL que lista todos os países do continente africano, incluindo o nome do país, a sua área de superfície, e a população total das suas cidades, com os resultados ordenados por área de superfície em ordem decrescente.



Explique como faria para obter o plano de execução dessa consulta, apresentado acima, descreva o plano de execução e a análise do resultado e desenhe a árvore do plano de execução.

### Resposta Parcial:



Critérios de correção:

- Explicar como usar o Explain 0.3
- Descrição do plano e análise do resultado 0.7
- Árvore do plano de execução 1
- erros, omissões, redundâncias ou indentações desadequadas: -20% a -100%

## 2. Protocolo 2-PL:

2.1. (1 valor) Defina o protocolo 2-PL. Quais as vantagens e desvantagens? Justifique a resposta e dê exemplos.

### Resposta Parcial:

O protocolo 2PL (2-phase locking) é um algoritmo de bloqueio utilizado para o controlo de concorrência entre transações. Utiliza dois tipos de bloqueios (locks):

Bloqueio partilhado (S-lock) (utilizado nas operações de leitura (Read)): o item de dados pode ser partilhado por várias transações.

Bloqueio exclusivo (X-lock) (utilizado nas operações de escrita (Write)): o item de dados não pode ser partilhado por várias transações.

O protocolo 2PL exige que todas as transações solicitem todos os bloqueios que necessitem antes de libertar qualquer um dos bloqueios que detenha. Desta forma a gestão de bloqueios é realizada em duas fases:

Fase de crescimento ou expansão: a transação apenas pode adquirir bloqueios.

Fase de encolhimento ou contenção: a transação apenas pode libertar bloqueios.

O ponto de mudança de fases é designado de ponto de bloqueio ('lock point').

CrITÉrios de correção:

- definição: 0.5

- vantagens / desvantagens: 0.5

- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%

2.2. (2 valores) Considere o protocolo 2-PL e explique detalhadamente a execução das seguintes transações, usando os operadores X-lock(\_), S-lock(\_) e Unlock(\_). Como classifica a concorrência destas duas transações? Justifique a resposta.

2.3.

### Resposta Parcial:

T1	T2	T3	Comentários
S-Lock(A), R(A)			
		S-Lock(C), R(C)	
X-Lock(A), W(A)			
Unlock (A)			
Commit			
	S-lock(B), R(B)		
		X-lock(C), W(C)	
	Enqueue (C)		Pretende-se R(C) Não é possível T2 fazer X-lock(C) Deve aguardar
		Enqueue (B)	Pretende-se R(B) e W(B) Não é possível T3 fazer X-lock(B) Deve aguardar
			T2 e T3 aguardam a libertação de recursos, uma da outra. Estamos na presença de 'Deadlock'

CrITÉrios de correção:

- erros, omissões, redundâncias ou apresentação desadequada: -20% a -100%

3. (2 valores) Explique o conceito de 'Snapshot Isolation'. Como é que este difere do isolamento de transação tradicional? Explique quais as vantagens e desvantagens associadas e dê exemplos destas.

**Resposta Parcial:**

O 'Snapshot Isolation' é um nível de isolamento em bases de dados que permite que uma transação trabalhe com uma versão dos dados conforme existiam no início da transação, independente das mudanças feitas por outras transações paralelas.

Oferece uma visão mais estável dos dados e evitando o problema de "non-repeatable read", tem a desvantagem de não garantir serialização das transações.

Critérios de correção:

- Definição 1.0; Vantagens/Desvantagens 1.0
- erros, omissões, redundâncias: -20% a -100%

4. (3 valores) Pretendemos criar um 'Data Warehouse' para um sistema de gestão de uma rede de bibliotecas.

A base de dados deve armazenar informações sobre os livros, as bibliotecas, os empréstimos realizados, os leitores e as reservas de livros.

Livros: devem ter um identificador, o título, o autor, o género e o ano de publicação.

Bibliotecas: devem registar o identificador, a capacidade e a localização.

Empréstimos: devem registar o identificador, a data de empréstimo, a data de devolução, o identificador do livro e do leitor.

Leitores: devem ter um identificador, nome, e data de nascimento.

Reservas: devem registar o identificador, a data da reserva, o livro reservado e a biblioteca onde será recolhido.

Defina a tabela de factos em primeiro lugar. De seguida, defina dimensões para o 'Data Warehouse' e apresente a tabela de factos associada.

**Resposta Parcial:**

dimensao\_livros (id\_livro PK, titulo, autor, género, ano)  
dimensao\_leitores (id\_leitor PK, nome, data\_nascimento)  
dimensao\_bibliotecas (id\_biblioteca PK, capacidade, local)  
dimensao\_reservas (id\_reserva PK, data, id\_livro FK, id\_biblioteca FK)

factos\_emprestimo (id\_emprestimo PK, id\_livro FK, id\_leitor FK, id\_biblioteca FK)

Critérios de correção:

- erros, omissões, redundâncias: -20% a -100%

5. (2 valores) Como se diferenciam tabelas de dimensão e tabelas de factos em termos de estrutura e propósito? Exemplifique utilizando as tabelas da pergunta anterior.

**Resposta Parcial:**

As tabelas de dimensão contêm atributos descritivos ou qualitativos que oferecem contexto para os dados e que mudam com menos frequência. O propósito principal é fornecer um meio de categorizar e organizar dados dentro do 'Data Warehouse', permitindo realizar consultas e análises detalhadas.

As tabelas de factos registam eventos ou transações que ocorrem ao longo do tempo e estão sujeitas a mudanças frequentes. O propósito é quantificar e armazenar dados transacionais ou eventos, permitindo a realização de análises de tendências, padrões e desempenho ao longo do tempo.

Exemplo: Ao analisar a frequência de empréstimos (tabela de factos 'Empréstimos'), podemos usar a tabela de dimensão 'Livros' para ver quais os títulos mais populares, ou a tabela 'Leitores' para identificar os perfis dos utilizadores mais ativos.

Critérios de correção:

- Erros, omissões ou redundâncias: -20% a -100%

**FIM**