

U.C. 21103

Sistemas de Gestão de Bases de Dados

13 de julho de 2017

-- INSTRUÇÕES --

- O tempo de duração da prova de p-fólio é de 90 minutos.
- O estudante deverá responder à prova na folha de ponto e preencher o cabeçalho e todos os espaços reservados à sua identificação, com letra legível.
- Verifique no momento da entrega da(s) folha(s) de ponto se todas as páginas estão rubricadas pelo vigilante. Caso necessite de mais do que uma folha de ponto, deverá numerá-las no canto superior direito.
- Em hipótese alguma serão aceites folhas de ponto dobradas ou danificadas.
- Exclui-se, para efeitos de classificação, toda e qualquer resposta apresentada em folhas de rascunho.
- Os telemóveis deverão ser desligados durante toda a prova e os objetos pessoais deixados em local próprio da sala de exame.
- O enunciado da prova é constituído por 2 páginas e termina com a palavra FIM. Verifique o seu exemplar do enunciado e, caso encontre alguma anomalia, dirija-se ao professor vigilante nos primeiros 15 minutos da mesma, pois qualquer reclamação sobre defeitos de formatação e/ou de impressão que dificultem a leitura não será aceite depois deste período.
- Utilize unicamente tinta azul ou preta.
- O p-fólio é sem consulta. A interpretação das perguntas também faz parte da sua resolução, se encontrar alguma ambiguidade deve indicar claramente como foi resolvida

Grupo A – Sistemas de Bases de Dados

1. (2,5 valores) Considere a seguinte consulta de SQL (os campos ID são chave das respetivas tabelas):

```
SELECT delito.nome_delito, regiao.nome_regiao, Count(*) As casos
FROM delito, regiao
WHERE delito.regiao_id=regiao.regiao_id
GROUP BY delito.nome_delito;
HAVING casos>4;
```

Escreva uma expressão eficiente em álgebra relacional, equivalente à consulta. Justifique.

2. (2,5 valores) Suponha-se que existe um sistema de base de dados que nunca falha. É necessário um gestor de recuperação para este sistema?
3. (2,5 valores) Explique o propósito do mecanismo de *checkpoints* nos *logs* das bases de dados. Quantas vezes deve ser executado? Como é que a frequência de *checkpoints* afeta:
- O desempenho do sistema quando nenhuma falha ocorre?
 - O tempo que leva para se recuperar de uma falha do sistema?
 - O tempo que leva para se recuperar de uma falha de disco?
4. (2,5 valores) Proponha uma forma de implementar a técnica iterativa para *PageRank*, mesmo que a matriz T (tendo a representação como lista de adjacências) não caiba na memória de do computador.

Grupo B – Prática em “Data Warehousing”

5. (2 valores) Um departamento do Estado pretende registar numa base de dados informação relativa a áreas florestais. Sobre cada terreno pretende guardar a sua localização (freguesia, concelho e distrito), o nome da entidade responsável e a sua área. Cada freguesia, concelho e distrito possuem um código identificativo e um nome, sobre cada concelho deve ser registado o distrito a que pretende e sobre cada freguesia deve ser registado o concelho a que pertence. Existem áreas florestais de grande dimensão que ocupam mais de uma freguesia. Para cada entidade responsável deve estar registado com nome e contactos.

Pretendemos desenhar um “*Data Warehouse*” deste sistema. Defina as tabelas de factos em primeiro lugar e as tabelas de dimensões, bem como as respetivas associações.

FIM