

Análise de Dados com Base em Processamento Massivo em Paralelo

Introdução: Lista de Exercícios

Profa. Dra. Cristina Dutra de Aguiar Ciferri

André Perez

Guilherme Muzzi da Rocha

Jadson José Monteiro Oliveira

João Pedro de Carvalho Castro

Leonardo Mauro Pereira Moraes

Piero Lima Capelo

Observação:

Recomenda-se fortemente que a lista de exercícios seja respondida antes de se consultar as respostas dos exercícios.

1. Qual a diferença entre OLTP e OLAP?
2. Qual a diferença entre *data warehouse* e *data warehousing*? Qual a relação existente entre esses conceitos?
3. Considere a seguinte “chuva de expressões”:

“inserção/remoção/atualização”, “OLAP”, “grande número de usuários concorrentes”, “interação dinâmica”, “OLTP”, “interação estática/predefinida”, “poucos usuários concorrentes”, “leitura (consulta)”, “muitos registros acessados por vez”, “poucos registros acessados por vez”

Preencha a tabela a seguir utilizando as expressões supracitadas.

Ambiente Operacional	Ambiente Informacional

4. Considere as seguintes descrições:

- (a) Dados brutos sem significado semântico.
- (b) Informação interpretada, analisada, processada.
- (c) Dados organizados, estruturados, contextualizados.

Associe as descrições supracitadas com os conceitos a seguir.

- (a) Dado.
- (b) Informação.
- (c) Conhecimento.

5. Considere uma empresa de supermercados que possui várias filiais. Cada filial possui um sistema diferente para contabilizar os produtos vendidos e as promoções realizadas. Um executivo dessa empresa deseja fazer uma análise para descobrir filiais que precisam ser fechadas ou remodeladas. Por que não seria ideal que esse executivo realizasse essa análise sobre os sistemas existentes da empresa?

6. Considere a seguinte situação:

Uma empresa brasileira especializada em análise de dados busca estudar e entender as maiores consequências que a pandemia gerada pelo novo coronavírus ocasionou nas cidades brasileiras. O principal objetivo é determinar quais diferentes medidas podem ser tomadas, levando em consideração as diversas características socioeconômicas que cada cidade ou região do país pode ter. Para dar início ao estudo, a empresa catalogou conjuntos de dados que contêm índices socioeconômicos e dados referentes aos índices de contaminação, quantidade de testes, recuperação e óbitos para cada cidade.

Com base nessa descrição, cite quais dados podem ser extraídos, quais informações podem ser obtidas e quais conhecimentos podem ser construídos.

7. Pesquise uma situação real na qual seria interessante aplicar consultas analíticas. Descreva o problema e elabore três perguntas para exemplificar possíveis análises. A partir dessas perguntas, descreva três exemplos de conhecimento que podem ser gerados por meio dessas análises.

