

21103 - Sistemas de Gestão de Bases de Dados
2015-2016
e-fólio C

PARA A RESOLUÇÃO DO E-FÓLIO, ACONSELHA-SE QUE LEIA ATENTAMENTE O SEGUINTE:

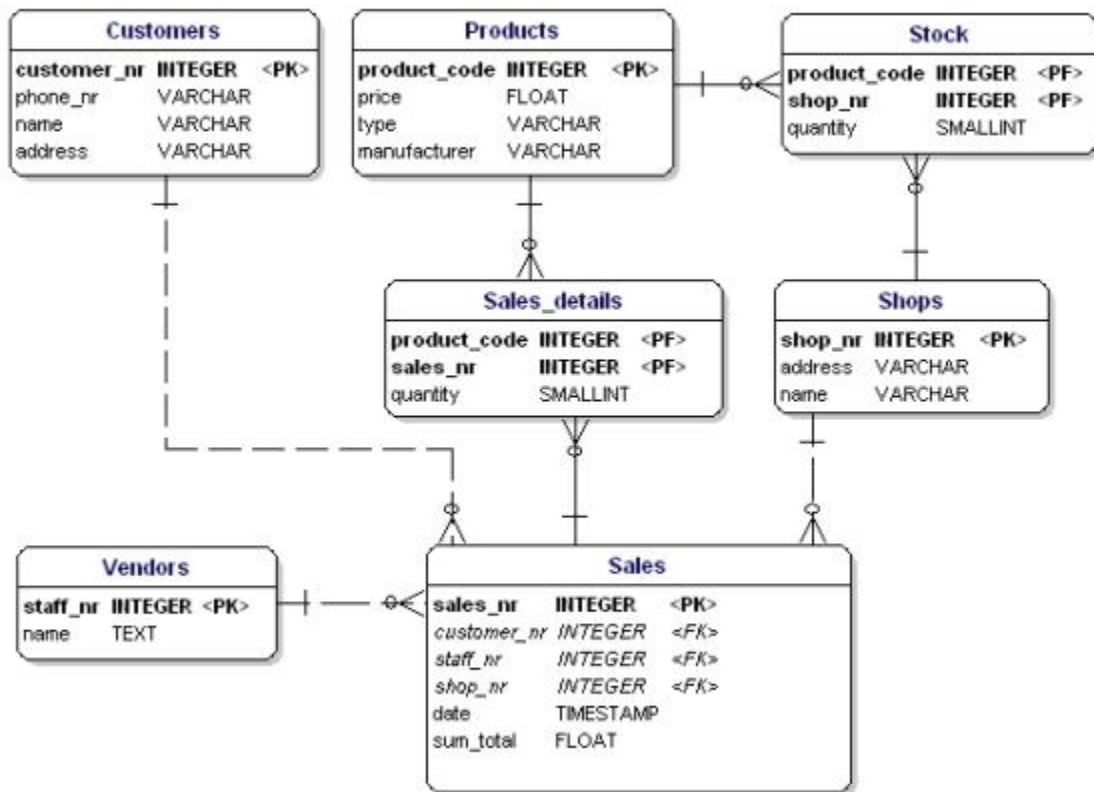
- 1) O e-fólio é constituído por 3 perguntas. A cotação global é de 3 valores.
- 2) O e-fólio deve ser entregue num único ficheiro PDF, não zipado, com fundo branco, com perguntas numeradas e sem necessidade de rodar o texto para o ler. Penalização de 1 a 3 valores.
- 3) Não são aceites e-fólios manuscritos, i.e. tem penalização de 100%.
- 4) O nome do ficheiro deve seguir a normal “eFolioC” + <nº estudante> + <nome estudante com o máximo de 3 palavras>. Penalização de 1 a 3 valores.
- 5) Na primeira página do e-fólio deve constar o nome completo do estudante bem como o seu número. Penalização de 1 a 3 valores.
- 6) Durante a realização do e-fólio, os estudantes devem concentrar-se na resolução do seu trabalho individual, não sendo permitida a colocação de perguntas ao professor ou entre colegas.
- 7) A interpretação das perguntas também faz parte da sua resolução, se encontrar alguma ambiguidade deve indicar claramente como foi resolvida.
- 8) A legibilidade, a objectividade e a clareza nas respostas serão valorizadas, pelo que, a falta destas qualidades serão penalizadas.

1) (1 valor) No processo de extração de dados de uma base de dados transacional existem 2 tipos de armadilhas no SQL (SQL traps):

- junções com múltiplos caminhos (“multiple access path problem”, “loop”)

- junções com agregações de dados de 2 tabelas (“connection trap”)

Dada a seguinte base de dados transacional de vendas de produtos



1.a) Para a base de dados da figura exemplifique uma consulta que evidencie a armadilha de junções com múltiplos caminhos, com dados e resultados.

1.b) Para a base de dados da figura exemplifique uma consulta que evidencie a junção com agregações de dados de 2 tabelas, com dados e resultados.

2) (1 valor) Para a mesma base de dados transacional de vendas de produtos:

2.a) Desenhe uma base de dados transacional equivalente, na 3ª forma normal. Faça o carregamento de dados. Na representação gráfica das ligações de 1:N, a tabela com uma única linha é desenhada em cima e a tabela com várias linhas é desenhada por baixo.

2.b) De seguida remova os caminhos múltiplos que eventualmente existam no esquema de base de dados. Na representação gráfica das ligações de 1:N, a tabela com uma única linha é desenhada em cima e a tabela com várias linhas é desenhada por baixo.

2.c) Pretendemos desenhar um “Data Warehouse” relacional em estrela ou em constelação, i.e. com duas ou mais estrelas com a granularidade dos detalhes das vendas. Defina a(s) tabela(s) de factos e mostre a tabela depois da desnormalização dos dados. Defina as dimensões com os níveis de agregação para o “Data Warehouse” relacional. Apresente a(s) tabela(s) de factos associada às dimensões. Ao juntar as tabelas transacionais tenha em consideração as eventuais armadilhas referidas na pergunta anterior.

2.d) Quais as tabelas da base de dados que foram desnormalizadas? Quais as tabelas da base de dados que não foram utilizadas no “Data Warehouse”?

2.e) Crie duas perguntas e traduza para SQL utilizando pelo menos duas dimensões (OLAP).

3) (1 valor) *Information Retrieval*

Escreva um texto, com pelo menos 500 palavras, onde descreva:

a) o que entende por PageRank

b) o que entende por SEO (Search Engine Optimization)

c) e a sua importância do PageRank no SEO