

Número: \_\_\_\_\_ Nome: Gonçalo dos seguros

### Grupo I - Modelação Dimensional (6,5 valores)

Uma determinada empresa de comércio eletrónico de material informático possui um sistema operacional que regista as compras de produtos feitas pelos clientes. Cada compra resulta dos produtos que foram sendo adicionados pelo cliente a um carrinho de compras, enquanto foi navegando nas várias páginas *web* que compõem o *site* da empresa. No momento em que o cliente despoleta a finalização da compra, é registado cada um dos produtos que se encontra no seu carrinho de compras, respetiva quantidade, o preço unitário (sem desconto) e, se for o caso, a percentagem de desconto a aplicar no preço unitário.

Cada cliente necessita de se registar previamente no *site* da empresa para poder realizar compras. Sempre que pretenda realizar uma compra, o cliente necessita de autenticar-se no *site* mediante o seu *login* (único para cada cliente) e *password*. Assim, para além destes dados sobre o cliente, o sistema operacional também regista: o nome do cliente; a morada; o código postal; o país; o número de identificação fiscal; e, o e-mail. A cada código postal encontra-se associada uma localidade.

Cada produto que a empresa comercializa no *site* é caracterizado por: um código (único para cada produto); um descritivo; o preço atual; a taxa de IVA atual; a data em que o produto começou a ser comercializado; e, eventualmente, a data em que o produto deixou de ser comercializado. Cada produto encontra-se também classificado numa determinada categoria (*e.g.*: tinteiros; toners; impressoras laser).

No momento da compra, é também registado o instante de tempo (data e hora) em que tal ocorreu, assim como a página *web* a partir da qual o produto foi adicionado ao carrinho de compras. Uma página *web* apresenta os vários produtos pertencentes à mesma categoria. Cada página *web* que compõe o *site* é caracterizada por: um identificador único; um descritivo; o nome da página (*e.g.*, *tinteiros.html*); o tamanho da página (em bytes); a data em que a página foi criada; e, eventualmente, a data em que a página deixou de estar acessível no *site*.

A empresa realiza frequentemente promoções pelo que os dados referente à eventual promoção associada à compra de cada produto encontra-se também registada pelo sistema operacional da empresa. Cada promoção é caracterizada por: um identificador único; uma descrição (*e.g.*: desconto de 25%); uma data de início; e, uma data de fim.

1. Seguindo a metodologia *Kimball*, desenvolva o processo de análise dimensional, a fim de definir e criar o modelo dimensional para um *data mart* que permita realizar análises multidimensionais de dados variadas às compras de produtos feitas pelos clientes, de acordo com a realidade que acabou de ser descrita. Apresente todos os factos, dimensões, granularidade e aspetos relevantes para o projeto de *data mart*.
2. Admita que se pretendem efetuar análises de dados aos totais trimestrais: das compras sem incluir o desconto; dos descontos; das compras incluindo o desconto. Explique o que acrescentaria ou alteraria a nível do esquema do modelo dimensional elaborado na alínea anterior, de modo a suportar a realização deste tipo de análises da forma mais eficiente/otimizada.

**Grupo II - Múltipla Escolha**  
(1 valor cada questão correcta/-0,5 cada questão errada)

Nas questões seguintes assinale apenas uma só alternativa correspondendo à que considera correcta.

**Grupo III – Verdadeiros ou Falsos com Justificação**  
(2 valores cada questão)

Indique se as seguintes afirmações são verdadeiras ou falsas, apresentando a respetiva justificação.

1. Aquando do carregamento de uma dimensão deve previamente proceder-se à eliminação das restrições chave primária – chave estrangeira (*primary key – foreign key constraints*), sendo estas recriadas após esse carregamento, de modo a aumentar a performance/desempenho desse processo.

Falso - Isto é feito nas tabelas de factos

---

---

---

---

---

2. Um índice do tipo *Bitmap* pode ser utilizado, no SQL Server, como chave primária de uma dimensão.

Falso - O SQL server não suporta índices do tipo Bitmap

---

---

---

---

---

3. Uma dimensão *Junk* serve para armazenar as combinações distintas dos valores de atributos demográficos que sejam discretos/enumerados ou tenham sido transformados em discretos/enumerados.

Falso - É uma mini-dimensão

#### **Grupo IV – Questão de Desenvolvimento** (2,5 valores)

Em determinadas situações há necessidade de se proceder à correção de factos já anteriormente carregados na tabela de factos. Explique as formas que podem ser utilizadas para proceder a essa correção.

3 formas -> negar o facto e recerragá-lo; atualizar o facto; Eliminar e recerregar o facto