

1. Defina detalhadamente *Data Warehousing*, apresentado e explicando um esquema da infraestrutura. (2 valores)
2. Considere um talão de uma venda de supermercado (imagem 1), associado a uma qualquer cadeia de retalho, com diversas lojas numa cidade ou conjunto de cidades. Os dados das vendas são recolhidos nas caixas, com a indicação dos diversos produtos adquiridos por cada cliente. Em termos de estratégia de negócio para a atração de novos clientes, ou retenção dos já existentes, podem ser desencadeados diversas ações de marketing que podem passar, por exemplo, por promoções de produtos. Para poder vender, a cadeia de retalho, necessita comprar, encomendar, entre outros processos de suporte às atividades do dia a dia desta organização.

Supermercado APAD 2017/2018, LDA	
Av. Cor. José Maria Vale de Andrade	
Cidade de Viseu	
Nif: 680033548	
Fatura Simplificada Nº 354769282	17/05/2018 16:45
Loja: L001	
Caixa: F0234(Ana Silva)	
P000123456 Papel A4 80gr.	22,50€
5*4,5€	
P000345678 Capa Arquivo A4 Larga	6,00€
2*3€	
P000456789 Capa Arquivo A4 Fina	7,50€
3*2,5€	
Total	36,00€
Valor Pago (Multibanco)	36,00€
Obrigado pela sua visita	

Figura 1 - Talão de venda

No caso do talão de vendas apresentado, e tentando restringir o exemplo, espera-se perceber o perfil de aquisições por parte dos clientes, de forma a disponibilizar indicadores para a gestão do negócio.

- 2.1) Utilize o modelo Kimball, contemplando todas as etapas, para desenvolver um modelo dimensional em estrela para o DW (5 valores). Justifique as suas opções, tabelas e campos.
- 2.2) Transforme o modelo em estrela da resposta anterior, num modelo em floco de neve. Justifique as suas opções. (1 valor)
3. Efetue uma comparação, enumerando vantagens e desvantagens entre uma base de dados relacional e uma base de dados multidimensional (2 valores)
4. Explique detalhadamente o que é o *OLAP* apresentando exemplos da sua utilização. (2 valores)

**5 - Responda, com uma cruz na célula adequada do quadro correto no final desta página indicando a letra com a resposta pretendida para as perguntas 1 a 8. A resposta pretendida deve ser sempre aquela que for considerada correta(s) do restante grupo de respostas, isto é, nalguns casos isso pode corresponder identificar uma resposta correta entre três erradas, mas noutros casos isso pode corresponder ao inverso. (A cotação destas respostas é: correta= 1 valor; Não respondido= 0 valores; Errada= -100% do peso da questão, até ao mínimo de 0 valores).**

5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8

**5.1. O Nível de estagio de dados (Data Staging) para uma arquitetura genérica de uma Data Warehouse representa:**

- A – Todas as fontes de dados originais que vão alimentar o DW, como BD's, ficheiros de texto, excel, etc;
- B – Primeiro armazenamento dos dados e conjunto de processos que vão iniciar o tratamento dos dados;
- C – Informação que é apresentada, de várias formas, aos utilizadores. Neste nível, são geralmente usadas ferramentas Online Analytical Processing (OLAP) e de geração de relatórios;
- D – Área onde os dados são armazenados após a execução de todos os processos de filtragem e transformação.

**5.2. Nos sistemas que incorporam base de dados operacionais:**

- A – o foco da atenção são as transações diárias características da operação normal do negócio;
- B – Para conseguir uma visão de um assunto, ter-se-ia de aceder a muitas tabelas e de várias aplicações (se usados dados operacionais);
- C – É focado nas entidades de alto nível do negócio;
- D – Nenhuma das respostas anteriores está correta.

**5.3. Um Data Warehouse com ODS (Operational Data Store),**

- A – Permite efetuar Faturação, Controlo de encomendas, Contabilidade geral, Gestão de Stocks, Processamento de Salários;
- B – Não é mais que uma base de dados desenhada para integrar dados de múltiplas fontes em que estes são tratados antes de serem enviados ao DW para armazenamento;
- C – Inclui um subconjunto de dados de DW referentes a determinado assunto ou departamento;
- D – Nenhuma das respostas anteriores está correta.

**5.4. O processo responsável por capturar os dados das bases de dados de origem e transformar estes dados de acordo com as necessidades de utilização pelo negócio é chamado de:**

- A – Data Mart.
- B – Modelagem Dimensional
- C – ETL
- D – BI
- E – OLAP

**5.5. A arquitectura de Chaudhuri é composta por 5 fluxos. O quarto fluxo:**

- A – sumariza e agrupa os dados dentro de “vistas” mais adequadas aos utilizadores finais e às aplicações de frontend (tabelas sumarizadas, gráficos, páginas HTML, base de dados pessoais, entre outros formatos).
- B – envolve a entrada dos dados no sistema (inflow), que contempla extrair, filtrar, transformar, integrar e carregar os dados de várias fontes no DW;
- C – em tempos pré determinados, de dois a cinco anos dependendo da empresa, os dados armazenados no DW passam para o estado de dados antigos.
- D – Nenhuma das respostas anteriores está correta

**5.6. OLAP, ou On-line Analytical Processing é a capacidade de manipular e analisar um grande volume de dados sob múltiplas perspectivas. Dentro do OLAP, existe a estrutura ROLAP. Qual das afirmações melhor se adequa ao ROLAP:**

- A – Estamos em presença duma aproximação, em que é utilizada uma base de dados desenhada especificamente para suportar dados multidimensionais e que pré - calcula todos os valores derivados
- B – O ROLAP tem tempos de resposta muito baixos e pouca necessidade de espaço de armazenamento

- C – Estes servidores suportam vistas multidimensionais sobre os dados, recorrendo a mecanismos de armazenamento também multidimensionais.

D – Nenhuma das respostas anteriores está correta

**5.7. A operação típica OLAP - Drill-Down (roll-down):**

- A – permite navegar de dados generalizados para dados mais detalhados
- B – consiste na operação do movimento de nível de consolidação mais baixo para mais alto (sumarizar dados)
- C – permite selecionar um subconjunto de dados do cubo, sendo uma das dimensões do cubo limitada aos dados que obedecem a uma determinada condição
- D – permite definir um subcubo para o qual podem ser especificados critérios de seleção para duas ou mais dimensões.

**5.8. Toda modelagem dimensional possui dois elementos imprescindíveis: as tabelas Fatos e as tabelas Dimensões. São características das tabelas de dimensão:**

- A – Existem sempre menos dimensões quantas as vertentes pelas quais se pretende analisar os factos
- B – Serem a fonte principal das cláusulas das consultas, agrupamentos e títulos de relatórios
- C – Apresentarem-se sempre normalizadas, integrando na maioria dos casos uma grande quantidade de atributos;
- D – Contêm mais registos quando comparadas com a tabela de factos
- E – As respostas B e A estão corretas
- F – As respostas C e D estão corretas
- G – As respostas A e D estão corretas
- H – As respostas B e D estão corretas