



## Parte I

---

1. Uma das características fundamentais de um data warehouse é a sua “orientação temporal”. Explique de forma sucinta em que tal consiste e justifique a importância de tal característica.
2. A definição de dimensões é uma das etapas fundamentais no desenvolvimento de um esquema de um *data warehouse*. Algumas das dimensões são definidas como sendo de variação, com manutenção de história. Apresente e caracterize um pequeno exemplo no qual se possa identificar uma dimensão do tipo referido.
3. No processo de implementação de um *data warehouse* é usual adoptar-se o método dos “4 passos”. De forma sucinta, explique em que consiste este método, quando se aplica, e que tipo de tarefas com ele são realizadas.
4. Apresente um pequeno caso pratico no qual optaria por desenhar um esquema em “flocos-de-neve” em vez de um esquema em “estrela” para um dado *data mart*. Justifique.
5. Os dados angariados nas fontes de informação de um sistema de *data warehousing* nem sempre apresentam a qualidade que gostaríamos. Identifique e caracterize dois problemas típicos de qualidade que podem ocorrer.

## Parte II

---

Considere o seguinte caso para estudo:

*Nos últimos cinco anos a ETL (Empresa Transformadora de Lixo, SA) alargou a sua área de actuação a várias regiões do país. O aumento da procura de produtos derivados da reciclagem de lixos fez com que a empresa aumentasse a sua produção e, conseqüentemente, a sua capacidade de recolha, procurando novos lugares para fazer a colecta de maiores quantidades de lixo, e a sua capacidade de transformação, construindo novas instalações para a reciclagem dos lixos. Actualmente, a ETL tem em funcionamento quatro instalações de transformação de lixos (Aires de Cima, Sopro Frio, Candeias, e Castro Franco) que são alimentadas diariamente por uma frota de 55 camiões com a capacidade para transportarem, em média, cada um deles, cerca de 50 toneladas de lixo. O processo de gestão da recolha de lixos não é muito complicado. No início de cada dia de trabalho, os condutores de cada um dos camiões levantam nas secretarias das instalações de transformação em que trabalham a sua guia de recolha, que tem um número, uma data e uma hora de emissão. Esta guia indica-lhes quais as zonas e locais em que terão de recolher lixo e que tipo de lixo é que vão recolher. Quando os camiões ficam com as suas cargas completas os seus condutores têm que, logo que possam, vir descarregá-los à sua respectiva instalação, retomando de seguida, caso ainda estejam no seu horário de trabalho, a sua ordem de trabalhos de recolha. Durante a descarga de um camião, cada posto de controlo de uma instalação (todas as instalações têm três pontos de controlo, à excepção da instalação de Sopro Frio que tem seis) supervisiona a descarga dos lixos e regista a seguinte informação: número do condutor, matrícula do camião, número de quilómetros realizados, áreas e locais de colecta visitados, peso e tipo do lixo recolhido. Toda a informação envolvida, tanto na emissão das guias de recolha como nos registos de controlo das descargas, em especial a relacionada com as instalações de transformação, os pontos de recepção, os condutores, os veículos, as áreas de recolhas e o tipo de lixo, é verificada sempre nos sistemas operacionais da empresa.*

(...)

Com base no caso apresentado, pretende-se que desenhe um possível esquema para um *data mart* especialmente orientado para as atividades de tomada de decisão relacionadas com a recepção de lixo recolhido pelos funcionários da ETL. Explique sumariamente o esquema projetado, abordando os seguintes tópicos: grão, estrutura da tabela de factos, explicação das dimensões e hierarquias envolvidas, medidas escolhidas e respectivas funções de agregação e, finalmente, a topologia geral do esquema.