

Número: _____ Nome: Toni

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

Uma determinada empresa detentora da concessão de autoestradas possui um sistema operacional que calcula em tempo real o valor de portagem a pagar, com base na passagem do veículo, nos pórticos de entrada e saída da autoestrada, à custa de um dispositivo eletrónico existente no interior do veículo. Os pórticos encontram-se nos ramais de entrada/saída das autoestradas, sendo caracterizados por: um identificador; a autoestrada em que se encontram; o respetivo Km; e, pela indicação se é um pórtico de entrada ou saída da autoestrada. O identificador do pórtico é único, não existindo em qualquer outra autoestrada um pórtico com igual identificador.

A empresa é concessionária de várias autoestradas. Cada autoestrada é caracterizada por: uma designação única (e.g., A1; A2; A3); localidade de início; localidade de fim; distância em Kms do início ao fim; n.º de locais de acesso (i.e., entradas/saídas) existentes; data de início da construção; e, data de fim da construção (do último troço), uma vez que uma autoestrada vai sendo progressivamente construída ao longo do tempo.

Para poder circular numa autoestrada, cada veículo tem que possuir obrigatoriamente um dispositivo eletrónico. Cada dispositivo é caracterizado por: um número de série (único), a data de fabrico, a marca e o modelo (do dispositivo); a matrícula do veículo em que se encontra instalado/colocado; a respetiva classe de pagamento do veículo (i.e., motociclos pertencem à classe 1; ligeiros pertencem à classe 2; pesados pertencem à classe 3); e, o código de cliente/utilizador a quem o dispositivo pertence.

Os dispositivos são adquiridos à empresa pelos clientes/utilizadores da(s) autoestrada(s). Cada cliente é caracterizado por: código de cliente (identificador único); nome; morada; código postal; e, número de identificação fiscal. A cada código postal do cliente corresponde a respetiva localidade. Cada cliente pode possuir um ou mais identificadores, em função do número de veículos que possui a circular na(s) autoestrada(s) da concessionária.

Sempre que um veículo/dispositivo entra ou sai de uma autoestrada é registado o instante de tempo em que tal ocorreu, bem como o respetivo pórtico. Nem sempre a autoestrada de entrada é a mesma de saída (e.g., um veículo/dispositivo pode entrar na A1 e sair na A2). Por outro lado, a data de saída pode não ser a mesma que a data de entrada, isto é, o(s) ocupante(s) do veículo pode(m) ter pernoitado num hotel numa das áreas de serviço existentes nas autoestradas. No momento da saída de uma autoestrada, o sistema operacional calcula o valor de portagem a pagar, assim como o n.º de Kms percorridos desde a entrada na autoestrada.

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 portagens dos veículos nas deslocações nas autoestradas, de acordo com a realidade que acabou de ser descrita. Apresente todos os factos, dimensões, granularidade e todos os aspetos relevantes para o projeto de *data mart*.
2. Admita que se pretendem efetuar análises de dados ao valor total de portagens a pagar mensalmente por cliente, assim como ao respetivo n.º de Kms percorridos por este. 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)

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. A performance na realização das análises/consultas de dados é o único motivo pelo qual se adota o modelo dimensional em armazém de dados.

Falso - são 3 ao todo, performance, simplicidade para com users e flexibilidade

2. Abstraindo de todos os outros fatores que possam condicionar a escolha do nível de granularidade de um armazém de dados, a melhor opção consiste em adotar o nível de granularidade mais elementar.

Verdadeiro - Abstraindo de condições de negócio, porque possibilita todas as consultas necessárias pelo cliente

3. O frete existente ao nível da fatura pode ser armazenado, sem qualquer alteração, como medida/facto numa tabela de factos cuja granularidade seja ao nível da linha da fatura.

Falso - O frete tem de ser dividida pelas linhas respetivas

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

Na gestão do histórico das alterações que podem ocorrer aos atributos de uma dimensão, existem três técnicas disponíveis. Caracterize cada uma dessas técnicas e exemplifique o seu funcionamento num cenário de atualização ao valor de um atributo de uma dimensão.

Tipo 2 - Adiciona um registo novo, effectiveDate/ExpiredDate

Tipo 4 - Cria uma history table

Tipo 6 - Adiciona um novo registo (tipo2) + uma nova coluna (tipo3) + Reescreve (type 1)