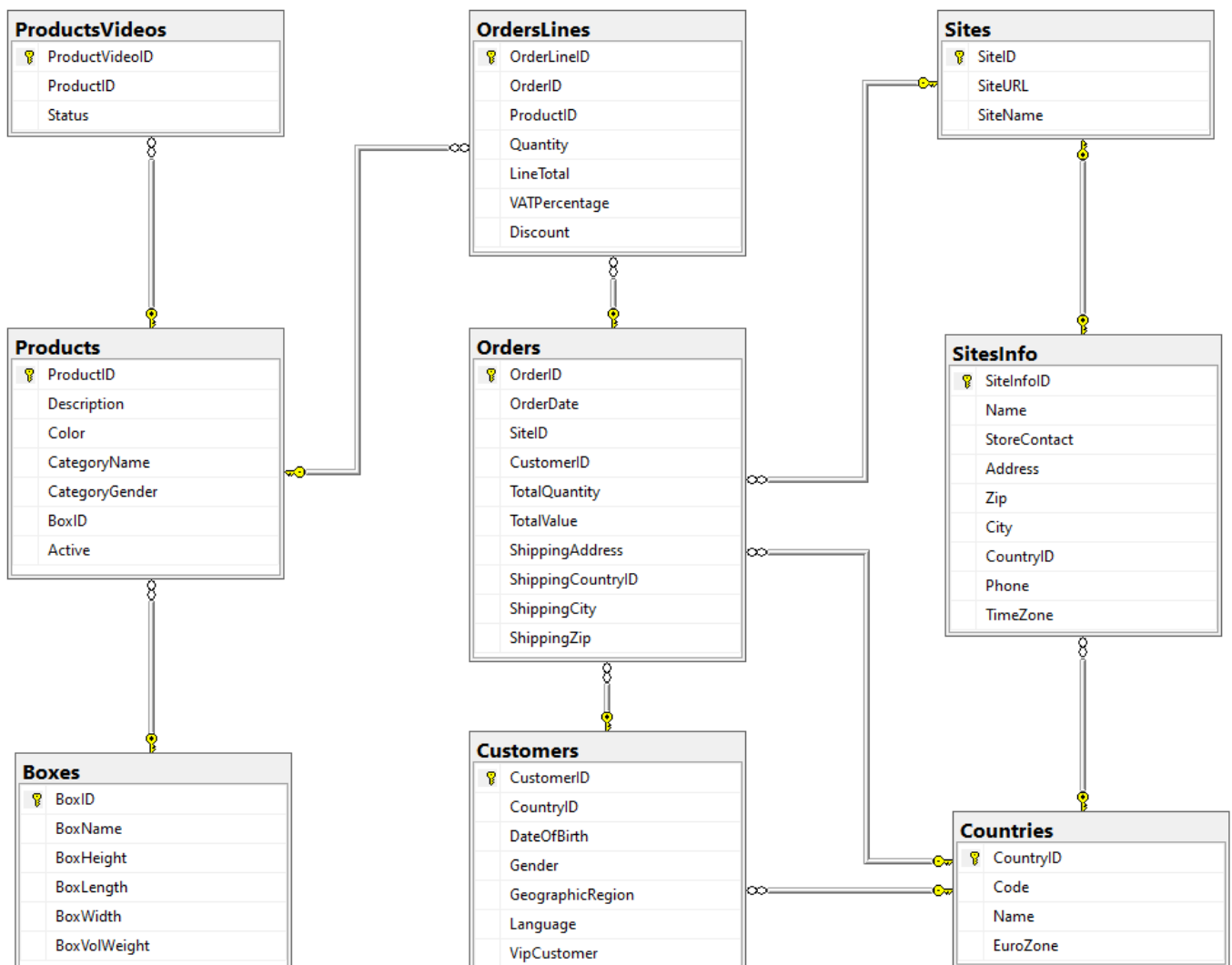


Número: _____ Nome: Rui Costa

Grupo I - Modelação Dimensional (6,5 valores)

Uma determinada empresa de comércio eletrónico dedica-se à venda de relógios de luxo. A empresa possui sites na Internet, nos quais os clientes podem obter informações/detalhes sobre os produtos (inclusivamente visualizar vídeos) e realizar as respetivas encomendas. Dado tratar-se de comércio eletrónico, cada produto encontra-se associado a uma caixa que é usada no seu embalamento. O sistema operacional da empresa centraliza todas as encomendas realizadas nos diversos sites. A empresa apenas atua na área territorial da união europeia pelo que a moeda utilizada é somente o Euro. Na figura seguinte apresenta-se o excerto do modelo de dados relativo à gestão de encomendas desse sistema operacional.



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 dados sobre as encomendas de produtos dos clientes, de acordo com a realidade descrita e o modelo de dados apresentado. Apresente todos os factos, dimensões, granularidade e aspetos relevantes para o projeto de *data mart*.

Grupo II - Múltipla Escolha

(1 valor cada resposta correta/-0,5 cada resposta errada)

Grupo III – Verdadeiros ou Falsos com Justificação

(2 valores cada resposta correta)

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

1. No processo de carregamento de uma dimensão com elevado volume de dados, recorrendo ao componente *Slowly Changing Dimension (Integration Services Project)*, deve definir-se um índice único (i.e., não admite valores duplicados) para o atributo que constitui a *business key* (chave primária no sistema operacional).

Verdadeiro. No processo de carregamento de uma dimensão com um elevado volume de dados usando o componente Slowly Changing Dimension (SCD) em um projeto do Integration Services (SSIS), é comum definir um índice único para o atributo que constitui a chave de negócios (business key), que normalmente corresponde à chave primária no sistema operacional. a businessKey pode representar várias chaves, mas só haverá uma businessKey

2. 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 dividido pelas linhas respetivas

Verdadeiro - Caso contrário, não podem ser aplicadas funções estatísticas

Uma estratégia de otimização que pode ser adotada em armazéns de dados consiste na criação de particionamentos/partições (normalmente, na tabela de factos por ser aquela que contém um elevado volume de dados). Explique o conceito de particionamento, assim como os diferentes tipos de particionamentos existentes e apresente as vantagens e desvantagens da criação de particionamentos.

Partição horizontal -> Divide a tabela ao colocar linhas diferentes em vários ficheiros físicos.

vantagens-> otimização/performance

Desvantagens -> aumento de complexidade