

Número: _____ Nome: Alexandre

Grupo I - Modelação Dimensional (6,5 valores)

Uma Instituição de Ensino Superior de grande dimensão possui um sistema de informação que, entre outros, armazena dados relativos à realização de exames. Um exame pode ser referente à época normal, época de recurso/melhoria, ou época especial, de um dado ano letivo. Cada ano letivo é caracterizado por: um código numérico (identificador único), o descritivo (*e.g.*, 2022/23), as datas de início e de fim do 1º e 2º semestre, as datas de início e de fim das épocas de exame do 1º e 2º semestre, e o número de dias efetivos de aulas anual (varia em função dos feriados).

Um exame ocorre num determinado dia e hora (de início), possui uma duração e, eventualmente, uma tolerância adicional. Um exame diz respeito a uma determinada disciplina, de um dado curso, sendo a disciplina da responsabilidade de um determinado docente (regente). Cada disciplina é caracterizada por: um código numérico (identificador único), a sigla, e o nome. A mesma disciplina pode ser comum a cursos diferentes (*e.g.*, licenciatura e mestrado), em anos curriculares diferentes (*e.g.*, 3º ano da licenciatura e 1º ano de mestrado), e possuir um número de créditos diferente. Cada curso é caracterizado por: um código numérico (identificador único), o descritivo do curso (*e.g.*, mestrado em engenharia informática), a data em que foi aprovado, a data da última remodelação, o ano letivo de arranque, e o código do departamento a que pertence. Cada docente é caracterizado por: um código numérico (identificador único), a sigla, o nome, a área de formação (*e.g.*, engenharia informática), o grau académico (*e.g.*, doutoramento), e o código do departamento a que pertence. A cada código do departamento (*e.g.*, DEI) encontra-se sempre associado um descritivo (*e.g.*, Departamento de Engenharia Informática).

Em cada exame de uma disciplina, em cada curso, o sistema de informação armazena dados sobre o número de alunos inscritos, o número de alunos que efetivamente estiveram presentes e, entre estes, o número de alunos que obtiveram aprovação. No caso das disciplinas comuns a cursos diferentes, normalmente o exame é realizado no mesmo dia e hora, mas nem sempre isto acontece.

Seguindo a metodologia *Kimball*, desenvolva o processo de análise dimensional, a fim de definir e criar o esquema conceptual para um *data mart* que permita realizar análises multidimensionais de dados variadas aos exames realizados, de acordo com a realidade descrita. Apresente todos os factos, dimensões, granularidade, e todos os aspetos relevantes para o projecto 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. Quando num armazém de dados se integram dados com proveniência de diferentes fontes de dados, o armazenamento da origem/fonte de cada registo é opcional.

Falso - É imperativo associar os dados às suas origens, por motivos de transparência aos utilizadores e saber a origem

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

3. Na tabela de factos apenas podem existir atributos cujo tipo de dados seja numérico.

Verdadeiro - Caso contrário não se podem aplicar funções estatísticas

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

Um aspeto fundamental num armazém de dados é a manutenção da integridade referencial entre as tabelas de factos e as dimensões. Apresente e explique as diferentes formas como a integridade referencial pode ser verificada/assegurada.

Check before-> antes de adicionar factos e antes de remover registos das dimensoes

check while -> engine da DB faz por default

check after -> realizar periodicamente verificações