

## Instituto Superior de Engenharia do Porto MSc in Informatics Engineering

## Armazéns de Dados

- 1. O nível de granularidade adotado numa tabela de factos é unicamente relevante segundo as perspetivas técnica e de projeto.
- 2. Abstraindo de todos os outros fatores que possam condicionar a escolha do nível de granularidade, a melhor opção consiste em adotar o nível de granularidade mais elementar.
- 3. A utilização de chaves de substituição nas dimensões, em detrimento das chaves dos sistemas operacionais, justifica-se unicamente por questões de performance.
- 4. A existência de atributos do tipo *EffectiveDate* e *ExpiredDate* ou *IsCurrent* é dispensável na implementação de um mecanismo de *Slowly Changing Dimension* (SCD) Tipo 2.
- 5. A implementação de um mecanismo de *Slowly Changing Dimension* (SCD) Tipo 2 para armazenamento do histórico das alterações que ocorrem aos atributos de uma dimensão, pode ser plenamente alcançada utilizando unicamente um atributo que indica se é o registo mais atual ou não (*isCurrent*).
- Considerando que um atributo vai sofrer inúmeras alterações ao longo do tempo, um mecanismo de controlo de alterações de Slowly Changing Dimension (SCD) – Tipo 3 é o mais adequado.
- 7. Uma dimensão *Junk*: é utilizada em armazéns de dados para armazenar as combinações possíveis de um conjunto de atributos demográficos, o que permite a sua remoção de uma dimensão.
- 8. Uma mini-dimensão é utilizada em armazéns de dados para armazenar conjuntos de atributos do tipo textual e do tipo *flag* (booleano), o que permite a sua remoção da tabela de factos.
- 9. Uma tabela de factos é sempre composta pela chave primária, formada pelo conjunto ou subconjunto de chaves estrangeiras, por um conjunto de medidas (factos) cuja análise é relevante para o negócio e, eventualmente, uma ou mais dimensões degeneradas.
- 10. Uma tabela de factos pode armazenar três tipos de medidas: aditivas; semi-aditivas; e, não aditivas.
- 11. Na extração de dados estática (a partir dos sistemas operacionais), a única estratégia disponível consiste em utilizar força bruta, ou seja, comparar os atributos dos registos do sistema operacional com os atributos dos registos no armazém de dados.

- 12. Kimball (Bus architecture) e Inmon (CIF architecture) concordam que o maior poder dos dados se encontra no nível mais atómico.
- 13. Sabe-se que a generalidade das análises/consultas de dados efetuadas num armazém de dados têm sempre o aspeto temporal (data e/ou tempo) presente. Assim, na definição da chave primária da tabela de factos e, consequentemente, do respetivo índice, há que ter este aspeto em consideração.
- 14. Numa tabela de factos existe sempre somente um único índice, i.e., o que está inerente à sua chave primária.
- 15. No processo de carregamento de dados, recorrendo ao componente *Slowly Changing Dimension (Integration Services Project)*, a *business key* corresponde necessariamente a um único atributo.
- 16. No SQL Server é possível definir-se um índice do tipo *Bitmap* para atributos de dimensões que apresentem baixa cardinalidade (i.e., reduzido n.º de valores distintos).