

Data Warehouse

Processamento Analítico de Dados
Profa. Dra. Cristina Dutra de Aguiar

Data Warehouse

- Banco de dados
 - voltado para o suporte aos processos de gerência e tomada de decisão
 - Características dos dados
 - orientados a assuntos
 - integrados
 - não voláteis
 - históricos
 - organizados multidimensionalmente
-

Características dos Dados

- Orientados a assunto
 - relativos aos temas de negócio de maior interesse da corporação
 - *exemplos*: clientes, produtos, promoções, contas e vendas
 - Integrados
 - dados obtidos dos provedores de informação corrigidos para eliminar possíveis inconsistências
-

Características dos Dados

- Não-voláteis

- o conteúdo do DW permanece estável por longos períodos de tempo

- Históricos

- relevantes a algum período de tempo

- *exemplo*: usualmente dados relativos a um grande espectro de tempo (5 a 10 anos) encontram-se disponíveis

Características dos Dados

- Organizados Multidimensionalmente
 - aspectos estáticos
 - dimensões e seus atributos
 - perspectivas de análises dos usuários de SSD
 - medidas numéricas
 - representam os assuntos de interesse
 - são determinadas pelas dimensões que as caracterizam
-

Dimensão

- Exemplos
 - **produto**: chaveProduto, nomeProduto, marca, categoria, departamento
 - **filial**: chaveFilial, nomeFilial, cidade, estado, região, país
 - **data**: chaveData, dia, mês, trimestre, semestre, ano
-

Hierarquia de Atributos

- Definição
 - permite que atributos de uma dimensão relacionem-se com outros atributos da mesma dimensão
 - especifica níveis de agregação e, portanto, granularidade dos itens de dados
 - *Exemplo*: dimensão **filial**
 - cidade → estado → região → país
 - *hierarquia de nível quatro na dimensão filial*
-

Medida Numérica Aditiva

- Pode ser **somada** considerando **todas as dimensões**
 - Exemplos
 - unidades vendidas
 - lucro
 - São agregadas usando
 - função de agregação **SOMA**
-

Medida Numérica Não Aditiva

- Não pode ser **somada**
 - Exemplos
 - **preço**, com a semântica de preço do produto
 - São agregadas usando
 - Funções de agregação **AGV**, **MAX**, **MIN**, **COUNT**
 - Outra função complexa
-

Medida Numérica Semiaditiva

- Podem ser **somadas** considerando somente **algumas dimensões**
 - Exemplo
 - **Número de clientes**, com a semântica de vendas de produtos
 - **Não aditiva**: para dois produtos P1 e P2 vendidos pela mesma filial no mesmo dia, não é possível somar o número de clientes do produto P1 com o número de clientes do produto P2, desde que o mesmo cliente pode estar sendo contabilizado duas vezes
 - **Aditiva**: número de clientes do produto P1 por dia pode ser agregado para se obter o número de clientes do produto P1 por mês
-