

Introdução ao Business Intelligence

Data Warehouse

Business Intelligence

- Um “guarda chuva” que inclui arquiteturas, ferramentas, bancos de dados e metodologias (TURBAN et al., 2009)

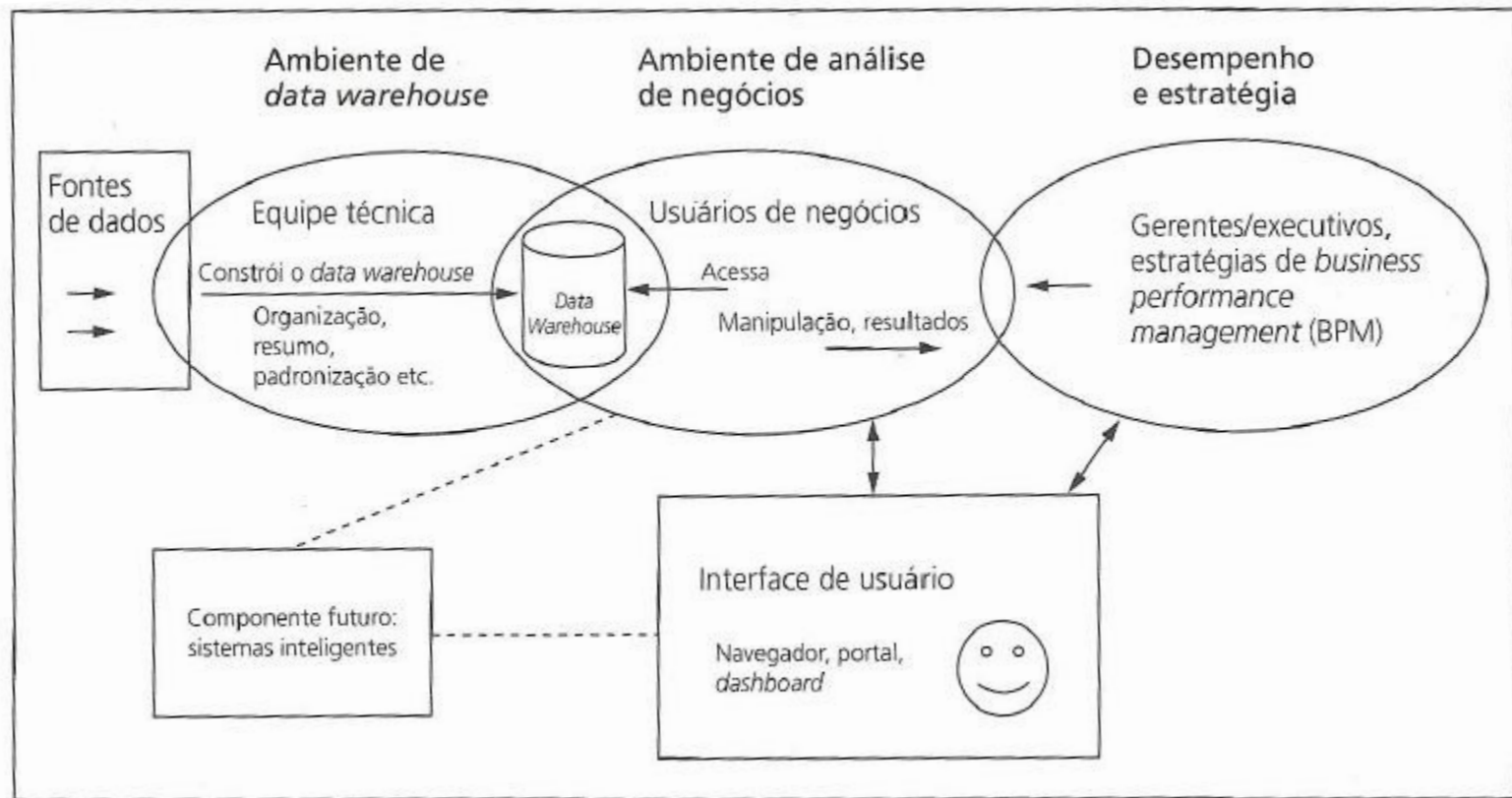
Business Intelligence

- Os principais objetivos do BI
 - permitir o acesso interativo aos dados (às vezes, em tempo real),
 - proporcionar a manipulação desses dados
 - fornecer aos gerentes e analistas de negócios a capacidade de realizar a análise adequada.
- Resumindo:
 - Facilitar o processo de tomada de decisão

Business Intelligence

- Principais componentes
 - Um repositório (Data Warehouse)
 - Ferramentas para visualização
 - Ferramentas para manipular dados (Data mining, Text Mining...)

Business Intelligence



Fonte: TURBAN et al., 2009

Problema...

- Produzir relatórios que acessem um número considerável de dados pode comprometer o andamento dos trabalhos
 - Exemplo: Compras de clientes sendo registradas em um sistema, relatório com resumo dos valores pagos nos últimos 5 anos sendo gerado. O processamento ficará mais lento, o atendimento aos clientes...mais lento

OLTP

- *OnLine Transaction Processing* – Processamento de Transações *OnLine*
- O tempo de resposta é crítico.
 - Ex. Fila em caixa de banco...
- OLTP ⇔ Sistemas de Processamento de Transações
 - Inclui
 - ERP - *Enterprise Resource Planning* (Sistemas Integrados de Gestão Empresarial)
 - SCM - Supply Chain Management (Gestão da cadeia de suprimentos)
 - CRM - Customer Relationship Management (Gerenciamento de Relacionamento com o Cliente)

Dados: Primitivos X Derivados

- Primitivos:
 - baseados em aplicações
 - detalhados
 - operações cotidianas
 - são atualizados
- Derivados:
 - baseados em assuntos ou negócios
 - resumidos e/ou refinados
 - atendem atividade gerencial
 - não são atualizados
- Difícil, alguns julgam impossível, mantê-los no mesmo banco

Data Warehouse - DW

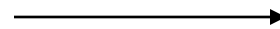
- “É um conjunto de dados baseado em assuntos, integrado, não volátil, e variável em relação ao tempo, de apoio às decisões gerenciais” W.H. Inmon

DW – Organizado por assuntos

- Os Sistemas de Processamento de Transações são organizados em torno das aplicações
 - Contabilidade, Contas a pagar, Contas a receber
- Um DW é organizado a partir de assuntos, das principais áreas de interesse
 - Ex: Vendas

DW – Integrado

- Aspecto mais importante
- Várias aplicações com informações similares (até mesmo redundantes)
- Problema nas aplicações
 - nome dos atributos
 - unidades de medida...
- Exemplo: codificação atributo estado civil
 - Aplicação A: c,s
 - Aplicação B: 1,2



DW: c,s

DW – Não-Volátil

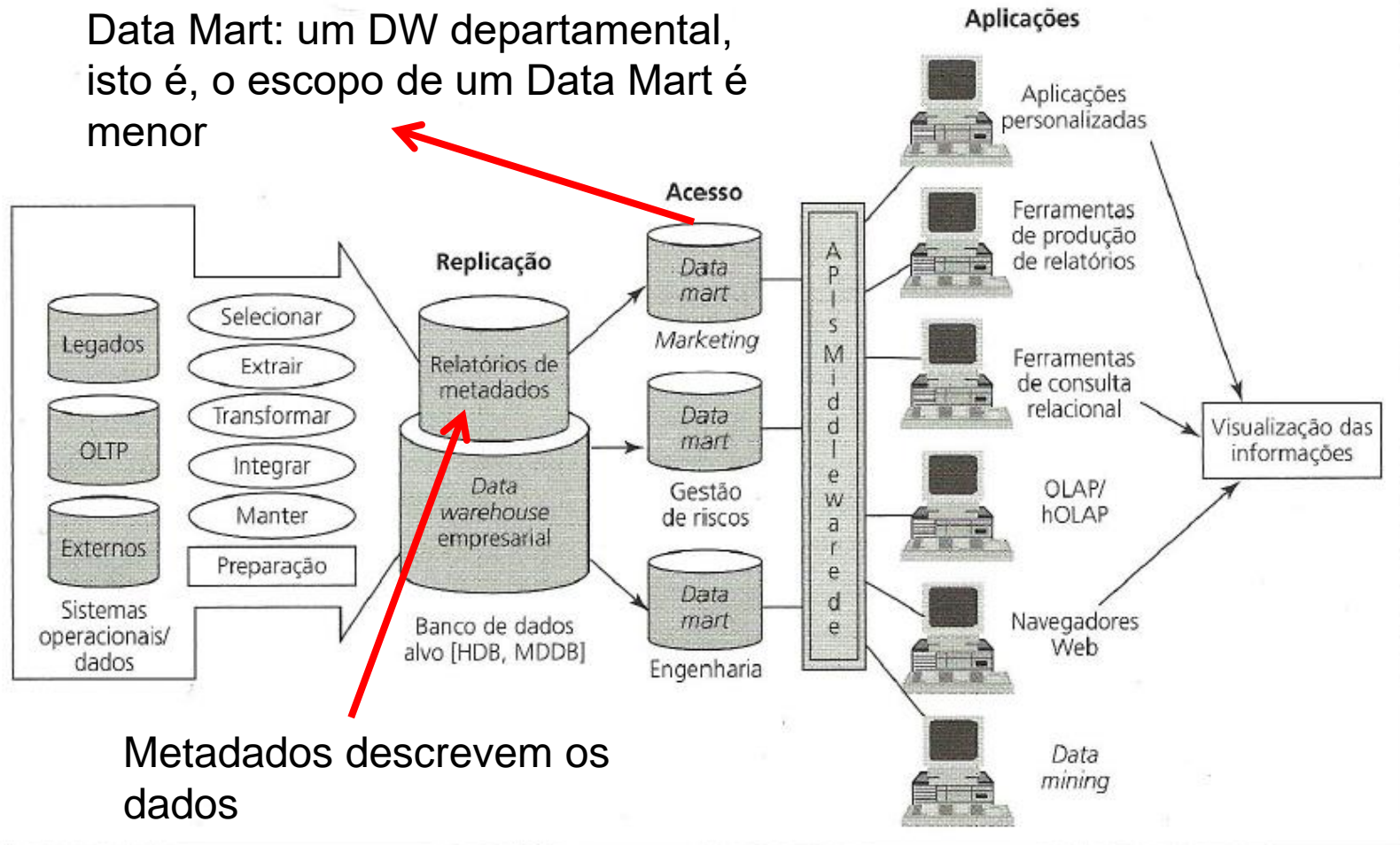
- No ambiente operacional os dados sofrem atualizações
- No DW ocorrem cargas periódicas e consultas
- Atualizações normalmente não ocorrem no DW

DW – Variável em Relação ao Tempo

- São mantidos dados históricos
- Série temporal
- O tempo é uma dimensão importante:
 - Variações ao longo do tempo
 - Diárias, Semanais, Mensais, Anuais...

DW - Ambiente

Data Mart: um DW departamental, isto é, o escopo de um Data Mart é menor



Fonte: TURBAN et al., 2009

ETL - Extract Transform Load ***(Extração Transformação Carga)***

- Extrair os dados de diversas fontes
- Alguns fabricantes fornecem softwares específicos
 - Capacidade de conectar a diferentes tipos de fontes
 - Transformação: formulação de regras
 - Carga: colocação dos dados no Data Warehouse
 - Mantém histórico do processo
- Desenvolvimento de programas específicos

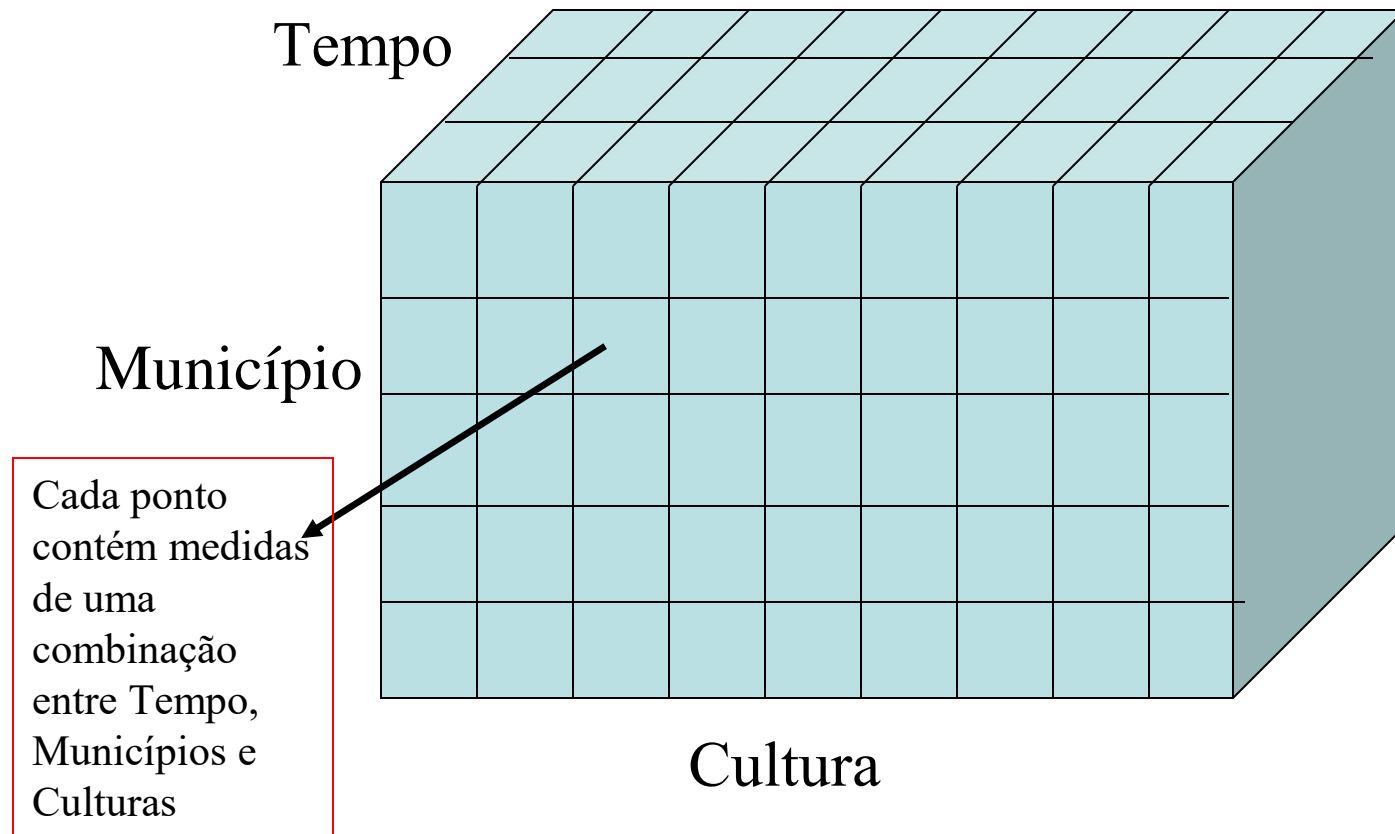
Projeto de DW

- Diferentes abordagens
- Ralph Kimball
 - Enfatiza desenvolvimento *Bottom-up*
 - Modelagem dimensional

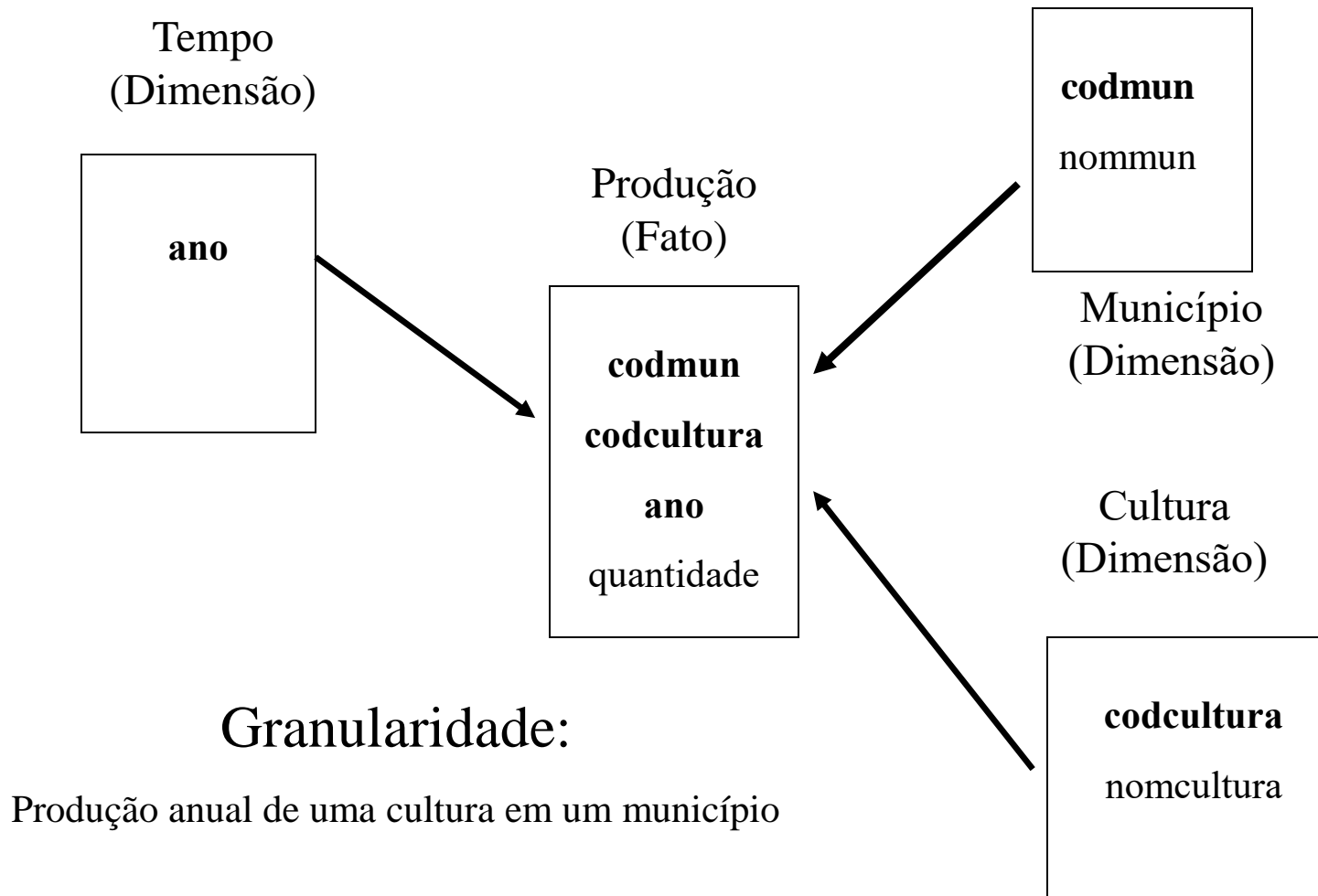
Projeto de DW

Modelo Dimensional

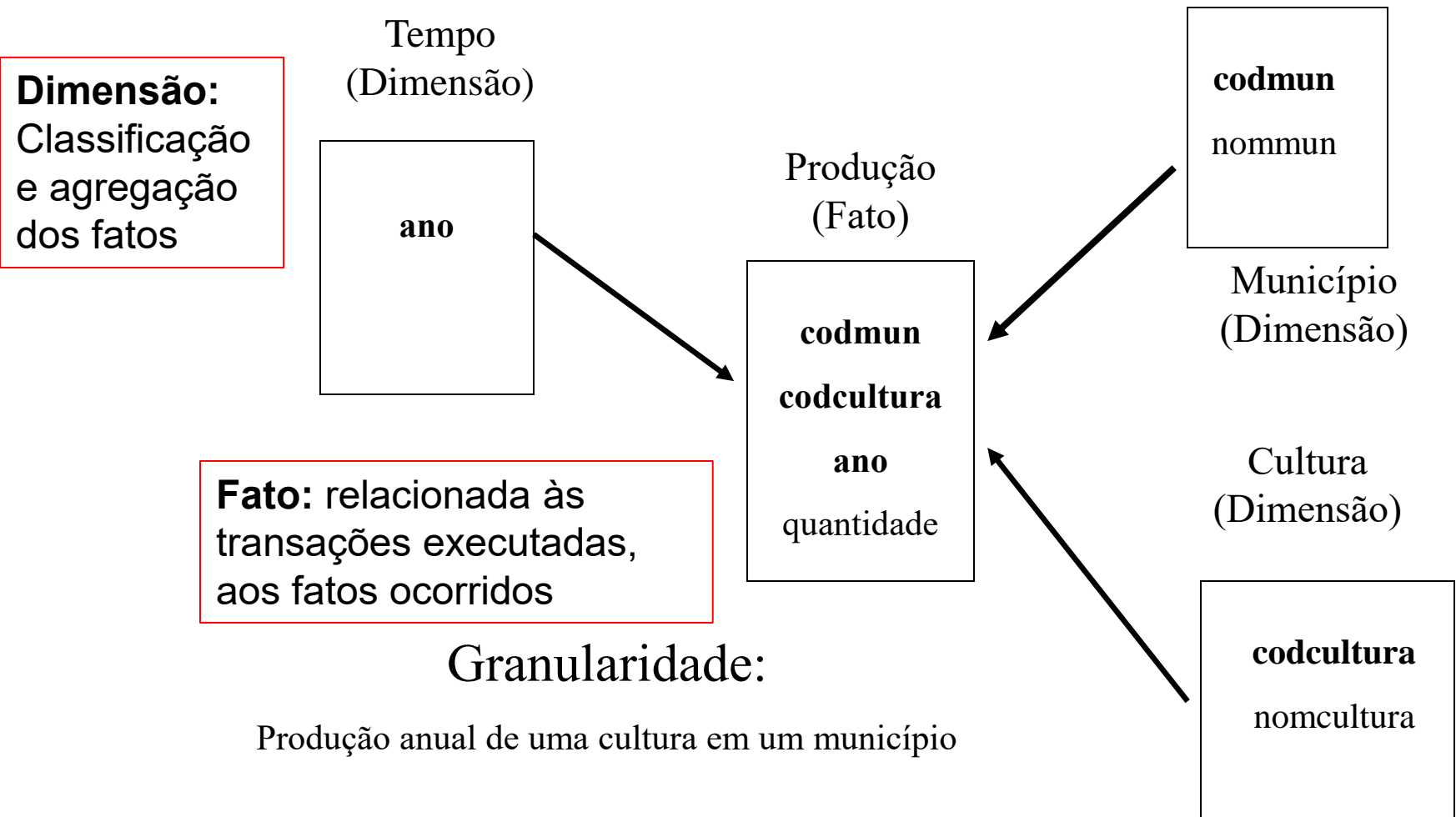
- “Pensar no negócio como um cubo de dados”



Modelo Dimensional – Star Schema



Modelo Dimensional – Star Schema



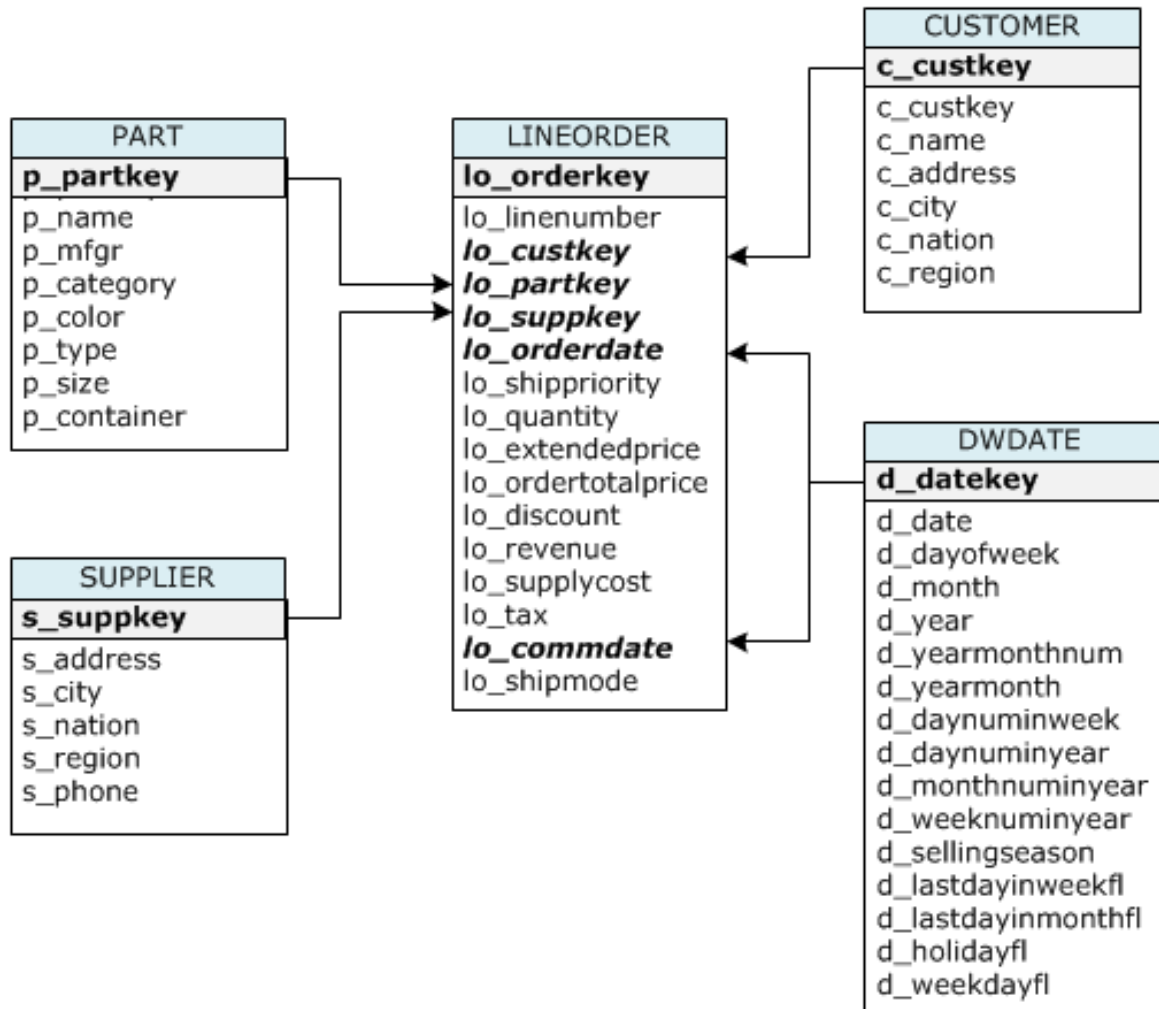
Dimensão Tempo

- Sempre presente!!
- Não é simplesmente uma data
 - Pode ter outros atributos

Dimensão Tempo

- Os dados da dimensão tempo não vem dos sistemas OLTP!!!
 - Pegar todas datas
 - Considerar calendário (e.g. feriados)
 - Considerar eventos (e.g. chuva, tempestades, etc.)
- Snowflake (subdividir o tempo)
 - Dia/Semana/Mês/Trimestre/Ano
 - Não recomendado por alguns autores
(KIMBALL et al. The Data Warehouse LifeCycle Toolkit. 2ed. , 2008 ver
pág. 265)

Dimensão Tempo - Exemplo



Star Schema Benchmark
(SSB) schema

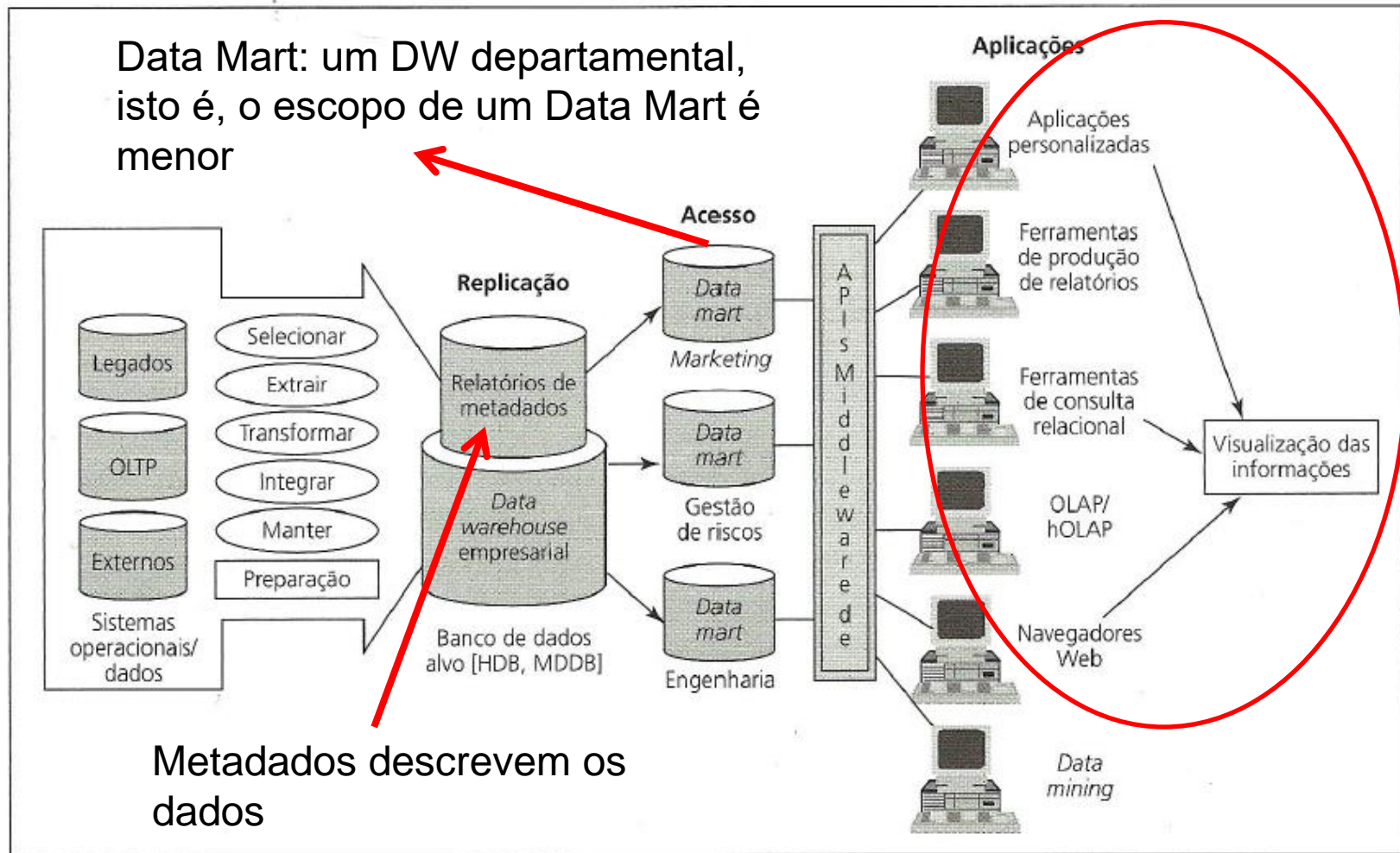
Fonte:

<https://docs.aws.amazon.com/redshift/latest/dg/tutorial-tuning-tables-create-test-data.html>

Implementação do Data Warehouse

- É possível usar um Sistema Gerenciador de Banco de Dados (SGBD) “convencional”
- MAS existem SGBDs específicos para Data Warehouse
 - Volume de dados maior
 - Operações lidam com grande volume de dados
 - Atualizações periódicas!!!
- <https://www.trustradius.com/data-warehouse#products>
- <https://www.softwaretestinghelp.com/data-warehouse-tools/>
- <https://docs.aws.amazon.com/redshift/latest/dg/welcome.html>

DW - Ambiente



Fonte: TURBAN et al., 2009

Ferramentas de Visualização

- Power BI (Microsoft)
 - <https://powerbi.microsoft.com/pt-br/>
- Data Studio (Google)
 - <https://datastudio.google.com/>

Data Studio

- **Primeiro acesso Data Studio**

<https://youtu.be/9RzLL-LsVcU>

- **Segundo acesso ao Data Studio**

<https://youtu.be/dzwDIK6x8bl>

Bibliografia usada na aula

- AUDY, L.N.J.; BRODEBECK, A. F. Sistema de Informação: Planejamento e Alinhamento Estratégico nas Organizações Porto Alegre: Bookman, 2003.
- LAUDON, K. C.; LAUDON, J. P. Sistemas de informações gerenciais São Paulo: Pearson Prentice Hall, 2007.
- TURBAN, E; RAINER, R. K. POTTER, R. E. Administração de Tecnologia da Informação . 3. ed. Rio de Janeiro : Elsevier, 2005.
- INMON, W. H. Como construir o Data Warehouse. Rio de Janeiro: Campus, 1997.
- KIMBALL, R. Data Warehouse Toolkit. São Paulo: Makron Books, 1998
- TURBAN, E. et al. Business Intelligence: Um Enfoque Gerencial para a Inteligência do Negócio. Porto Alegre: Bookman, 2009.