

AGENDA DO CURSO

1 APRESENTAÇÃO DO PROFESSOR E DA DISCIPLINA

2 MODELO DIMENSIONAL: FATOS, DIMENSÕES, MEDIDAS

3 GRANULARIDADE E AGREGAÇÃO

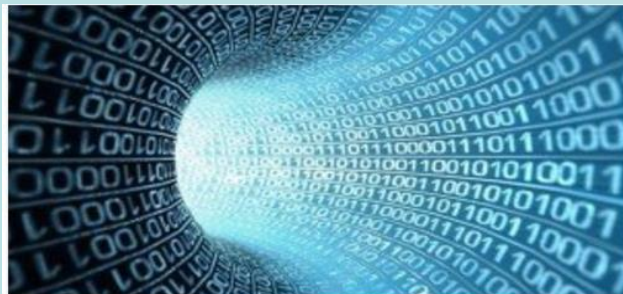
4 PRINCIPAIS TIPOS DE MODELOS

5 PRINCIPAIS OPERADORES OLAP

6 FATORES CRÍTICOS DE SUCESSO

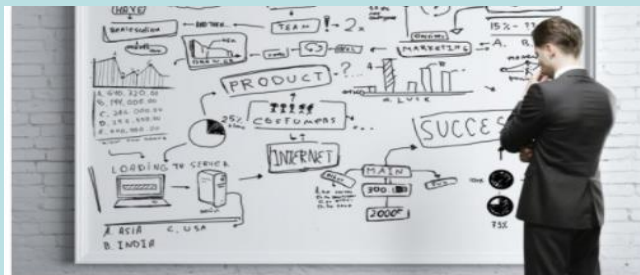
7 EXERCÍCIOS





Ferramentas de Data Discovery: o futuro das aplicações de BI?

Priscila J. Papazissis Matuck no LinkedIn
15 de agosto de 2016



Percepção sobre os perfis dos profissionais da informação no contexto de BI / Big Data

Priscila J. Papazissis Matuck no LinkedIn
21 de agosto de 2016



Tableau e Qlik Sense: como comparar?

Priscila J. Papazissis Matuck no LinkedIn
29 de agosto de 2016



A evolução do Qlikview é o Qlik Sense?

Priscila J. Papazissis Matuck no LinkedIn
1 de maio de 2017



Contando histórias por meio dos dados: ferramentas de Business Intelligence + storytelling e os benefícios para a tomada de decisão

Priscila J. Papazissis Matuck no LinkedIn
11 de junho de 2017



Empresas data-driven e o uso do big data como diferencial competitivo

Priscila Papazissis Matuck no LinkedIn
13 de junho de 2018

BIBLIOGRAFIA BÁSICA

BARBIERI, C. **Business Intelligence - Modelagem e Tecnologia**. Axcel Books, 2001.

BARBIERI, C. **BI2 - Business Intelligence - Modelagem e Qualidade**. Elsevier, 2011.

CRUZ, T. **Sistemas de Informações Gerenciais: tecnologia de informação e a empresa do século XXI**. São Paulo: Atlas, 1998.

HAAG, S., CUMMING, S. **Management Information Systems for the Information Age**, Editora McGraw Hill, 2010.

TURBAN, E. *et al.* **Decision support and business intelligence systems**. Prentice Hall, 8ª Edição, 2007.

BIBLIOGRAFIA COMPLEMENTAR

LEME, T. **Business Intelligence no Microsoft® Excel**, Axcel Books, 2004.

LOVINS, L. H., Natural Capitalism, Path to Sustainability, **The International Handbook on Environmental Technology Management**, Editado por Dora Marinova, David Annandale e John Phillimore, Massachusetts (USA): Edward Elgar Publishing, Inc, 2006;

MATUCK, P. J. P. 137p. **Dissertação**. Manifestações discursivas de comportamentos e práticas organizacionais na implementação do programa de gerenciamento da cadeia de suprimentos em uma indústria siderúrgica. Universidade FUMEC. Belo Horizonte. Dezembro, 2006.

PONNIAH, P. **Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals**. John Wiley & Sons Inc, 2001.

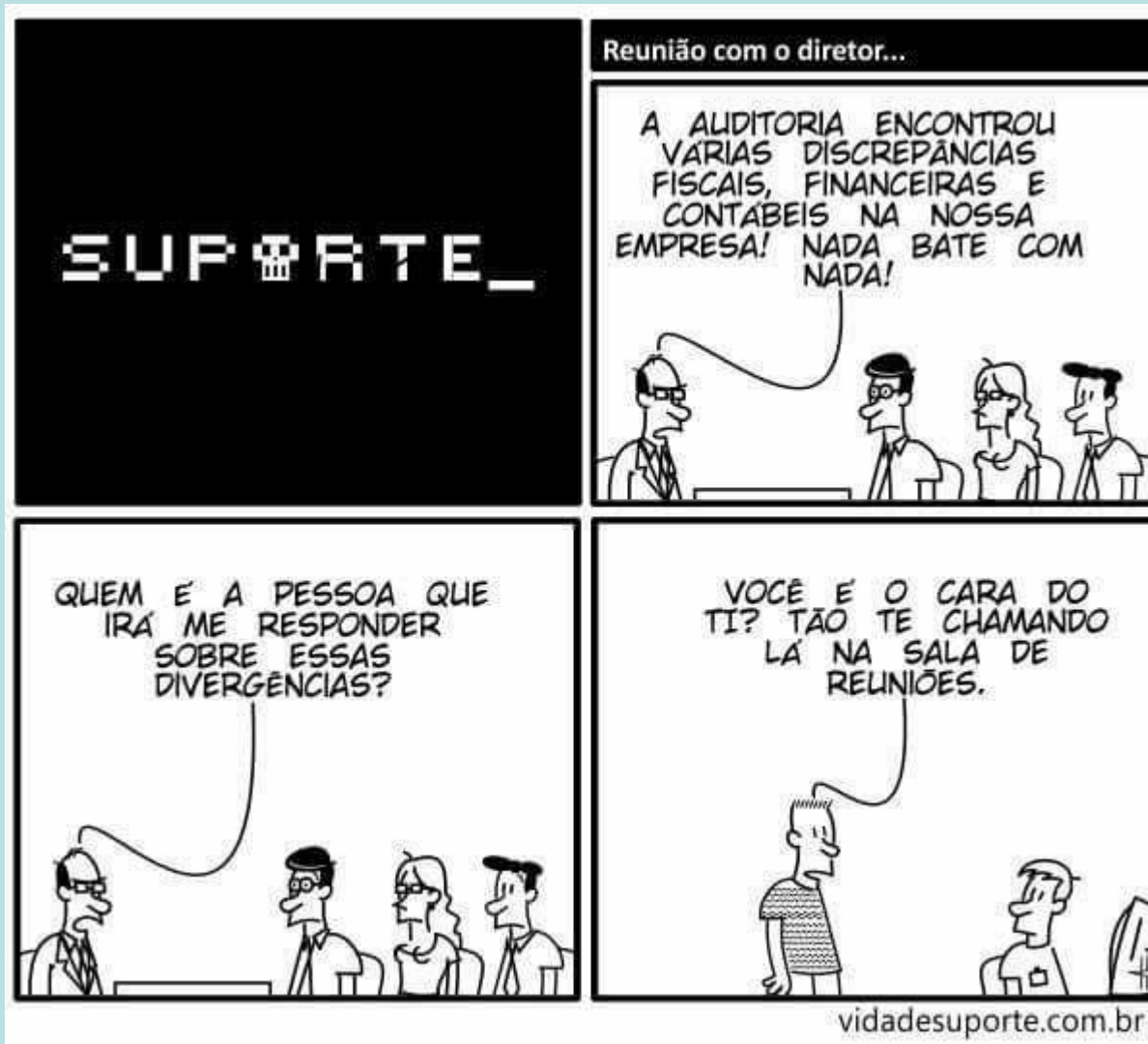
PRADO-JUNIOR, S. T. A evolução dos conceitos associados ao alinhamento entre a gestão de TI e a estratégia de negócios. In: **EnANPAD** - Encontro Nacional de programas de Pós-graduação em Administração, 2004, Curitiba - PR. Anais do 28º ENANPAD. 2004.

TURBAN, E.; RAINER, R. K.; POTTER, R. E. **Administração da Tecnologia da Informação: Teoria e Prática**. Rio de Janeiro: Campus, 2003.

2 CONTEXTUALIZAÇÃO



UMA VIDA REAL...





O ambiente empresarial tem se tornado cada dia mais **COMPLEXO**



As ações e reações que as empresas realizam para responder às **PRESSÕES**



e para se **BENEFICIAR** das oportunidades conquistadas em seus mercados



devem ser suportadas por **SISTEMAS** que tornem viável o monitoramento e controle do ambiente e proporcionem a **melhoria na tomada de decisão**. Vejamos...

2 CONTEXTUALIZAÇÃO

"Assim como estamos vendo muito mais complexidade nos dados, estamos vendo mais complexidade em como gerenciar esses dados" (GARTNER, 2016).

1958

Hans Peter Luhn, pesquisador da **IBM**, usou pela primeira vez o termo *Business Intelligence* considerando-o como um sistema capaz de mapear as inter-relações presentes em uma organização, promovendo a orientação das ações da alta administração.

1989

Howard Dresner, analista do Gartner Group, propôs o modelo de BI como um método para melhorar as decisões de negócio realizado por meio de sistemas baseados em computador, transformando dados em informações.

1996

As organizações só começaram a se interessar pelo BI, quando o conceito começou a ser espalhado nos EIS (*executive information systems*) e começaram a fornecer informações de forma simples e amigável.

2000

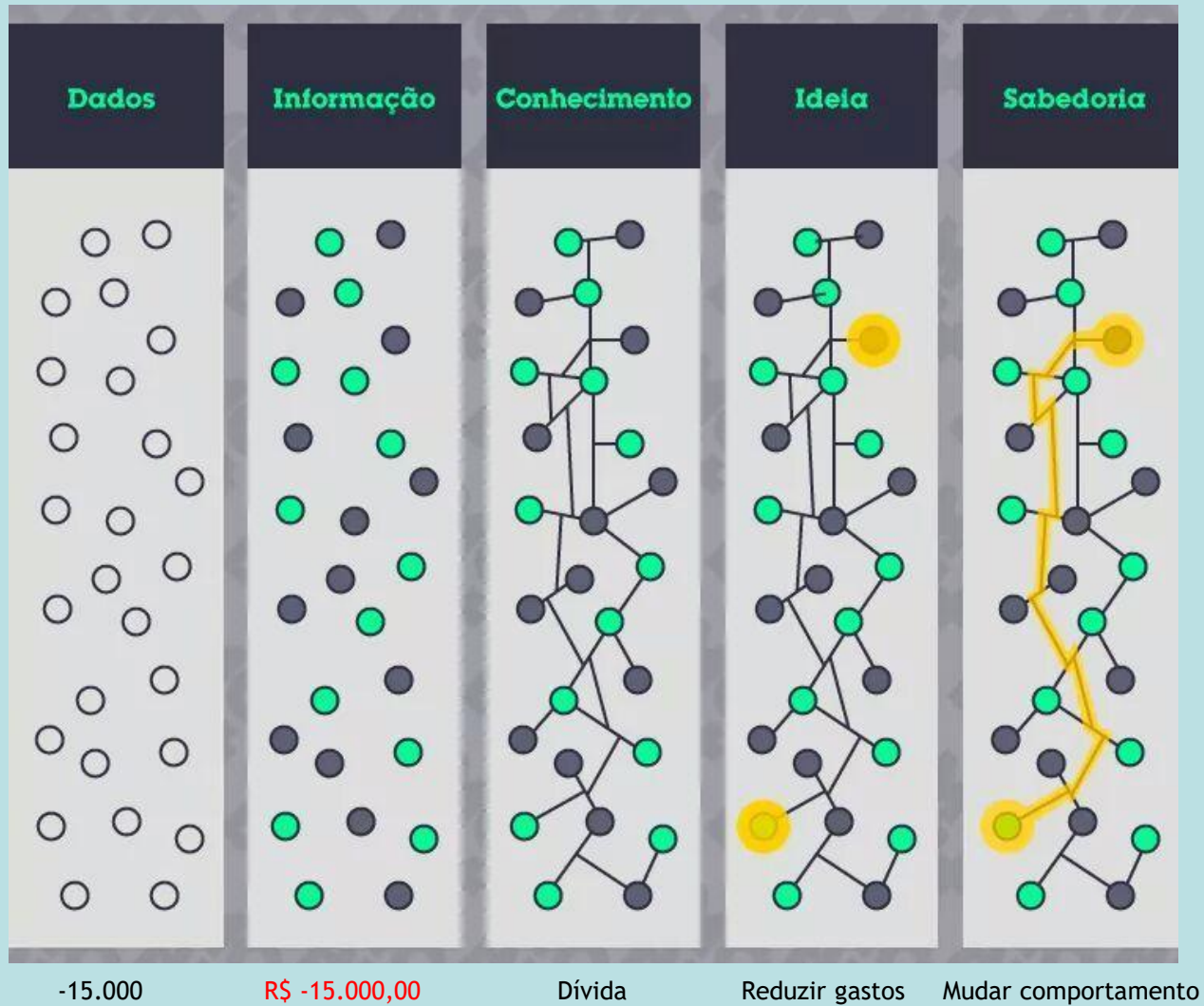
Várias ferramentas foram surgindo no mercado, como os DSS (*decision support systems*), planilhas, DW, *data marts*, *data mining*, ferramentas OLAP, entre outras...

2010

Se inicia a era de captura e tratamento de grandes volumes de dados dispersos, estruturados e não estruturados, impulsionados pelo advento da Internet.



3 ELEMENTOS CONSTITUTIVOS DO BI/BA



3 ELEMENTOS CONSTITUTIVOS DO BI/BA

O que precisamos saber antes de continuar...

O processo de **informatização** de qualquer empresa é baseado no desenvolvimento e na implantação de **sistemas transacionais** (ou operacionais). As principais funções desses sistemas são:

- ✓ **Coletar** dados existentes a partir das transações realizadas nas organizações;
- ✓ **Armazenar** os dados;
- ✓ **Ordenar** ou indexar esses dados, de modo a facilitar o acesso;
- ✓ Permitir **consultas** que retratem diferentes aspectos das operações;
- ✓ Gerar **relatórios operacionais**.

A construção de qualquer plataforma de BI requer o mínimo de informatização dos processos que deverão ser fonte de avaliação do desempenho empresarial.



3 ELEMENTOS CONSTITUTIVOS DO BI/BA

Então o que é Business Intelligence?

As oportunidades associadas a dados e análises em diferentes organizações ajudaram a gerar interesse significativo no Business Intelligence ao longo dos anos. Chen, H., Chiang, R. H. e Storely, C. (2012), afirmam que um ambiente de BI deve ter:



que analisam dados do negócio e ajudam as organizações a entenderem melhor seu direcionamento e o mercado onde estão inseridas, sendo capazes de tomar decisões sustentadas, assertivas e rápidas.



É necessário favorecer a inteligência, aprendizagem e criatividade organizacional se você pretende ter um BI na sua empresa. Sem a existência de ambientes de trabalho que permitam a colaboração, comunicação, aceitação do risco e tolerância ao erro, esses elementos podem nunca cumprir seu objetivo (SANTOS, M. Y., RAMOS, I., 2006).

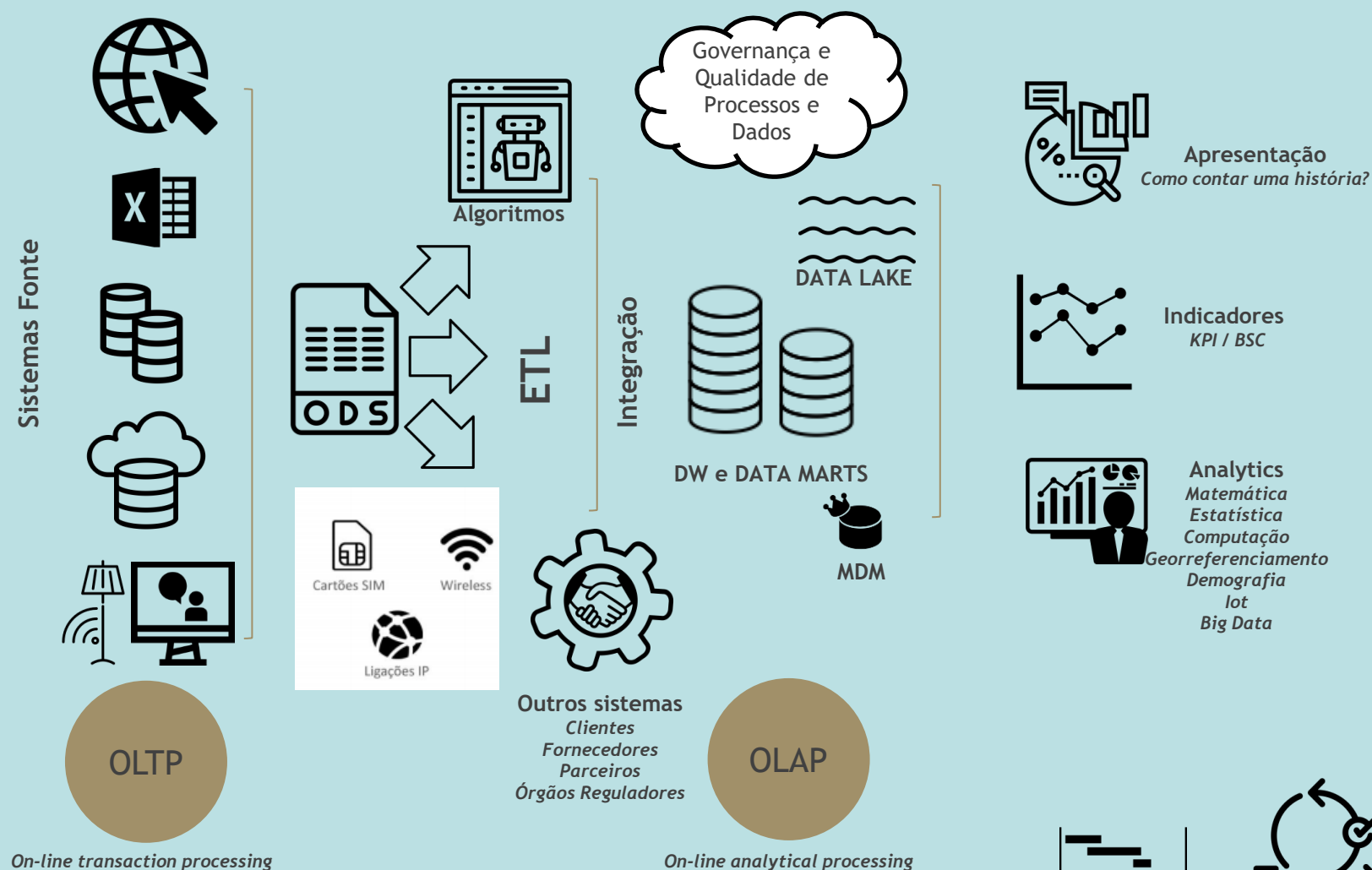
Informação operacional x voltada à decisão

Características	Dados Operacionais	Dados Informacionais
Conteúdo	Valores Correntes	Valores sumarizados, calculados, integrados de várias fontes
Organização dos Dados	Por aplicação / sistema de informação	Por assunto / negócio
Natureza dos Dados	Dinâmica	Estática até o “refreshment” dos dados
Formato das Estruturas	Relacional, própria para a computação transacional	Dimensional, simplificado, próprio para atividades analíticas
Atualização dos Dados	Atualização campo a campo	Acesso sem update
Uso	Altamente estruturado, processamento repetitivo	Desestruturado, com processamento analítico
Tempo de Resposta	Otimizado	Análises com grandes volumes de dados, complexos, com tempos de resposta maiores

(Barbieri, 2001, p. 47)

3 ELEMENTOS CONSTITUTIVOS DO BI/BA

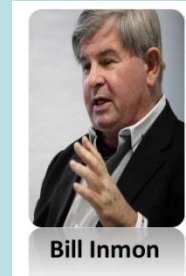
Arquitetura de Business Intelligence / Business Analytics



FILOSOFIAS DE DW e DM

Data Warehouse: Bill Inmon

- ✓ É parte da arquitetura de BI da empresa
- ✓ Normalmente a empresa tem 1 DW
- ✓ Os Data Marts buscam dele as suas informações e os clientes consomem dos DM
- ✓ Grande depósito de informações-Visão Top Down



Data Marts: Ralph Kimball

- ✓ O DW é formado logicamente por um conjunto de Data Marts, logo os DM nascem primeiro
- ✓ Os DM são desenvolvidos separadamente, observando-se dimensões conformes
- ✓ Os Data Marts estão modelados em Modelo Dimensional

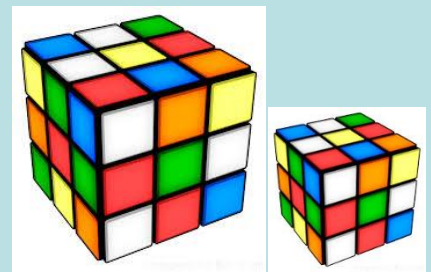


Modelo Dimensional

“Técnica que busca apresentar os dados em um modelo padronizado e intuitivo, que permita o entendimento por parte de usuários, bem como alto desempenho de acesso (Ralph Kimball).”

“O modelo dimensional disponibiliza itens necessários para modelar assuntos onde uma série de entradas (dimensões) estarão ligadas à algumas poucas métricas (fatos), criando uma notação legível e objetiva (Barbieri)”

Mas o que são fatos e dimensões?



FATO

Armazenam medidas numéricas associadas aos eventos do negócio.

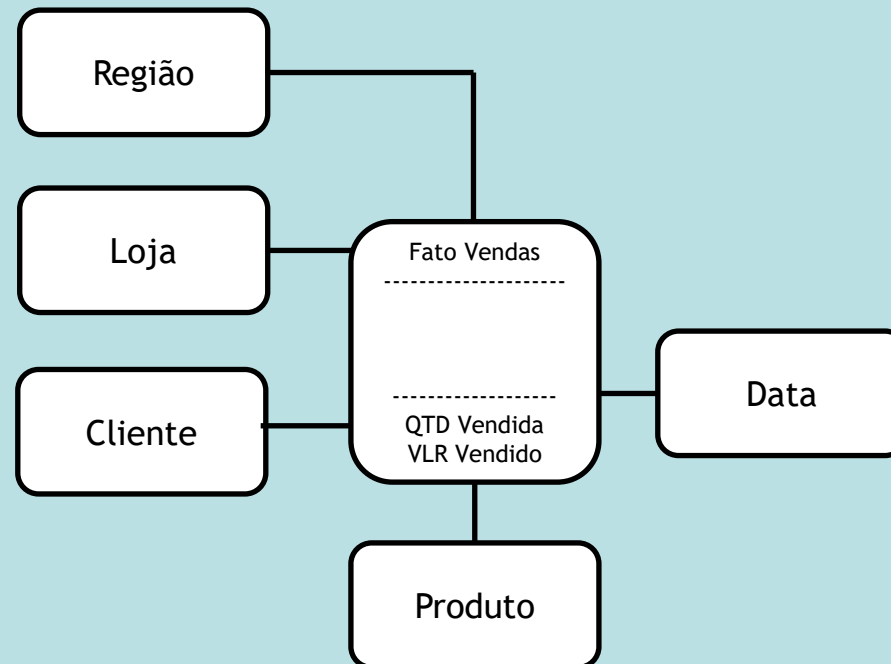
Formada pelas chaves estrangeiras das dimensões e suas métricas.



DIMENSÃO

Representam os assuntos do negócio. Perspectiva de análise do negócio. “Filtros” da informação.

Formada pelas chaves primárias das dimensões e suas características (atributos).



Modelagem Dimensional na Prática

Como modelamos dimensionalmente? Como isso acontece na prática?

As quatro decisões de negócio que devem ser tomadas durante o design de um modelo dimensional são:

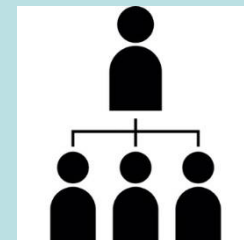
- 1) Selecione os processos de negócio
- 2) Declare a granularidade (oi??)
- 3) Identifique as dimensões
- 4) Identifique os fatos



Dessa forma, o time de desenvolvedores será capaz de determinar as tabelas e o nome das colunas, exemplos de valores de domínio e as regras de negócio mapeadas no sistema.

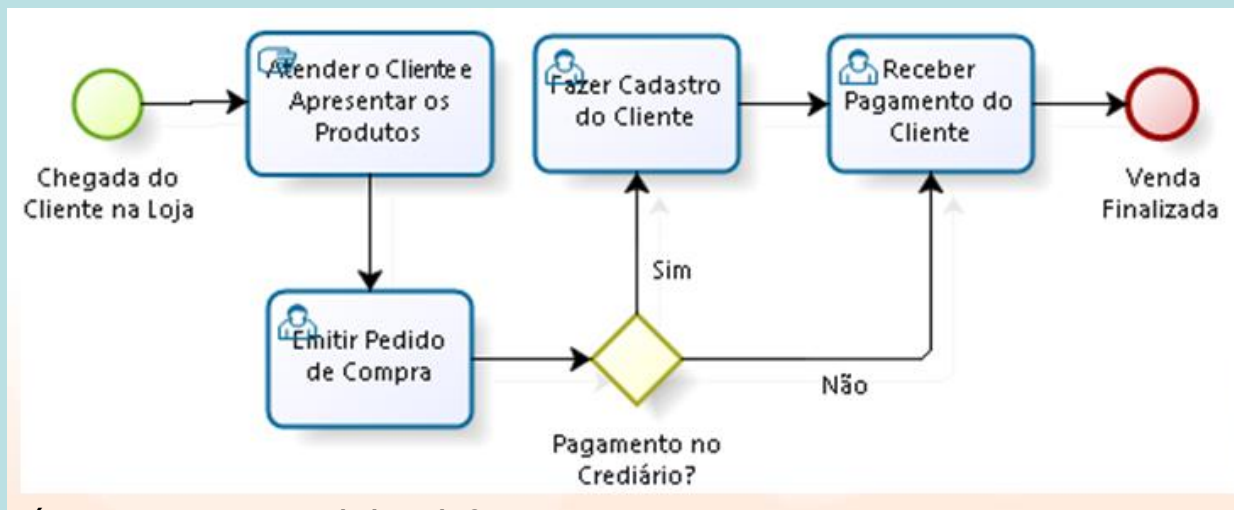
Passos da Modelagem Dimensional

1. Definição dos Processos de Negócio



Processos de negócio são atividades operacionais desempenhadas pela sua organização. Eventos de negócio capturam e geram métricas de desempenho que se traduzem em fatos de tabela fato.

Cada processo de negócio corresponde à N linhas no DW da organização.



- ✓ Que processo de negócio estamos modelando?
- ✓ O que usamos para medir este processo? Quantidades, valores, médias, somas, elementos constitutivos...

Passos da Modelagem Dimensional

2. Definição da Granularidade

Os fatores que definem a escolha da granularidade estão relacionados com o volume de dados a ser mantido na base de dados e com o processamento necessário para produzi-lo. *Perguntas: os dados serão armazenados desde quando? 2 anos? 3 anos? Os dados serão armazenados para serem recuperados até a sua granularidade mínima? Até o cliente? Até o produto?*

A declaração da granularidade deve ser feita antes da escolha das dimensões e fatos porque cada candidato à dimensão ou fato deve estar consistente à granularidade.



Diferentes granularidades não podem estar misturadas na mesma tabela fato.

Passos da Modelagem Dimensional

Pontos a avaliar:

- ✓ Dois anos de dados são suficientes para perceber os padrões de comportamento do dado?
- ✓ A granularidade temporal pode ser definida como diário, semanal, mensal, anual, etc.

Exemplo: um cliente pode fazer várias compras na mesma loja

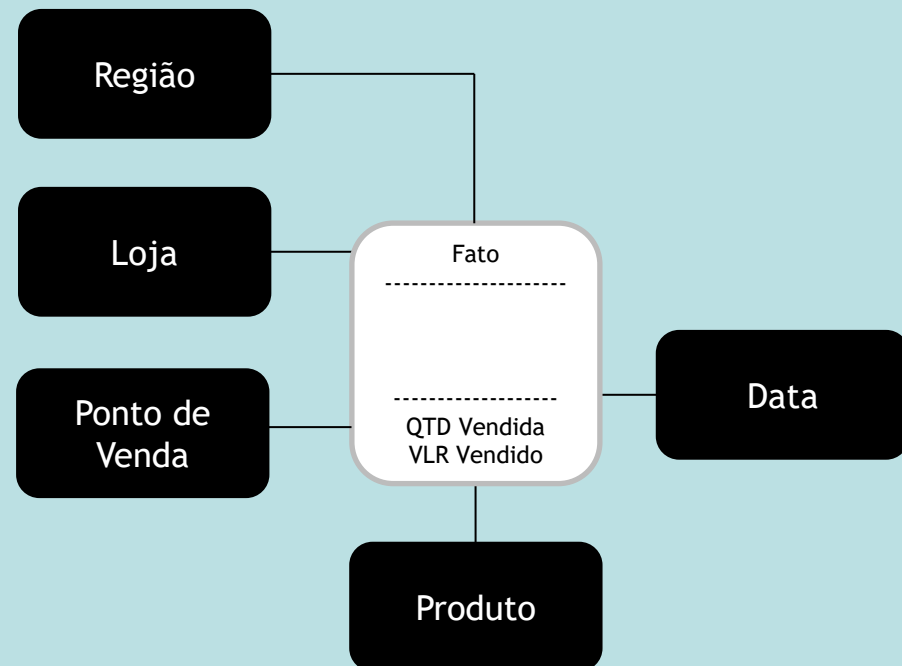
- Menor Granularidade: Armazenar no DW os dados de cada compra que o cliente realizou.
- Maior Granularidade: Armazenar no DW os dados consolidados do total de compras realizadas pelo cliente.



Passos da Modelagem Dimensional

3. Definição das Tabelas Dimensão

- ✓ Representação dos contextos relevantes para a análise de um fato, ou seja, representa uma perspectiva de análise dos usuários.
- ✓ Pontos de vista ou perspectivas do negócios sobre os quais uma organização deseja guardar registros e observar as métricas.
- ✓ Composta por chave (primária) e atributos.
- ✓ Considerados a “alma” do DW porque contém os pontos de entrada que nivelam o modelo com os processos de negócio.



Modelagem Dimensional

Dimensões

- ✓ Servem como filtro nas consultas / relatórios.

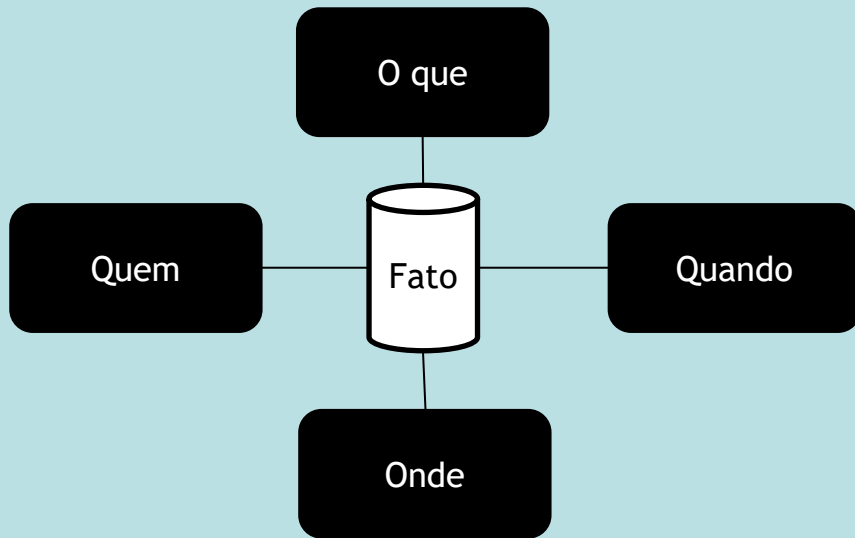
Características:

- ✓ Chaves simples (em geral, artificiais: “*surrogate keys*”).
- ✓ Descrição única para cada registro.
- ✓ Usualmente não dependente do tempo.
- ✓ Podem ser hierárquicas.



Modelagem Dimensional

4. Definição das Tabelas Fato



Fatos são medidas resultantes de um evento do processo de negócio, quase sempre numérico.

Um **FATO** (ASSUNTO) é um acontecimento de negócio e as **MÉTRICAS** (MEDIDAS) associadas a ele.

Cada fato é uma linha, que possui chaves multi-key formada pelas chaves primárias das dimensões que se relacionam com ele.

É tudo aquilo que reflete a evolução dos negócios do dia a dia de uma organização.



Modelagem Dimensional

Exemplos de Fato e suas métricas:

- ✓ O valor total de vendas, no último trimestre.
- ✓ Os índices de criminalidade no ano atual sobre os últimos dois anos.
- ✓ O número de unidades de produtos vendidas.
- ✓ A quantidade em estoque.

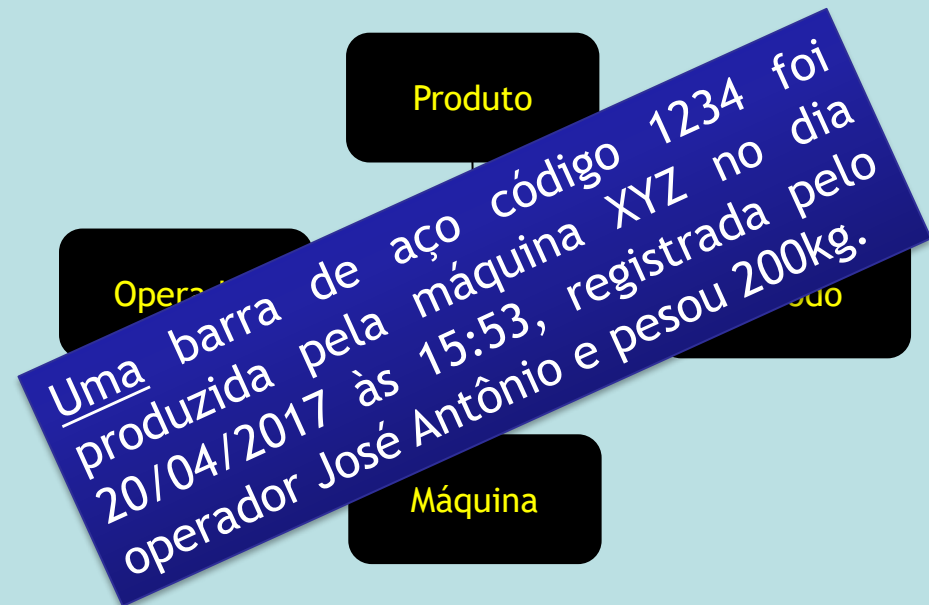
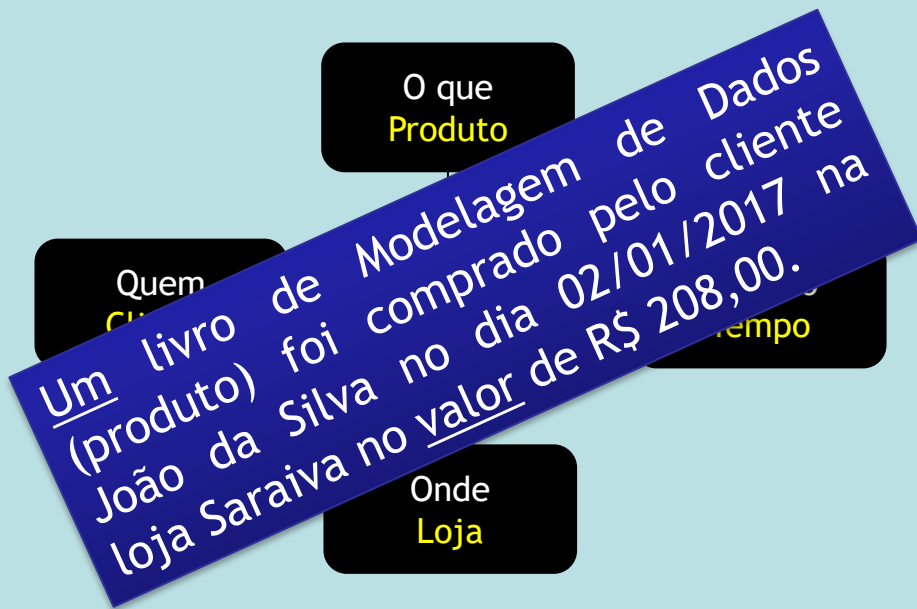
É fato ... Não é fato...



- ✓ É fato: venda (possui quantidade e valor; ocorre diariamente)
- ✓ Não é fato: fornecedor (as características não mudam com frequência)

Modelagem Dimensional

Vejamos mais exemplos



Fact tables express the many-to-many relationships between dimensions in dimensional models.

Descobrendo dimensões

Descobrendo Dimensões

- ✓ Qual é o Fato? Qual é a métrica?

Logo, a métrica é medida **POR**...?

- ✓ Geralmente as dimensões são derivadas das perguntas do usuário após entendimento das métricas.
- ✓ É necessário reescrever a pergunta do usuário utilizando a seguinte forma:

Métrica **POR** dimensão **POR** dimensão

Exemplo:

- ✓ Quais foram as vendas no período de junho a julho deste ano na região sul?
→ *Valor das vendas POR período POR região*



Uma ferramenta eficaz na identificação das dimensões: a matriz de barramento de avaliação de um DW

BUSINESS PROCESSES	COMMON DIMENSIONS							
	Date	Product	Store	Promotion	Warehouse	Vendor	Contract	Shipper
Retail Sales	X	X	X	X				
Retail Inventory	X	X	X					
Retail Deliveries	X	X	X					
Warehouse Inventory	X	X			X	X		
Warehouse Deliveries	X	X			X	X		
Purchase Orders	X	X			X	X	X	X

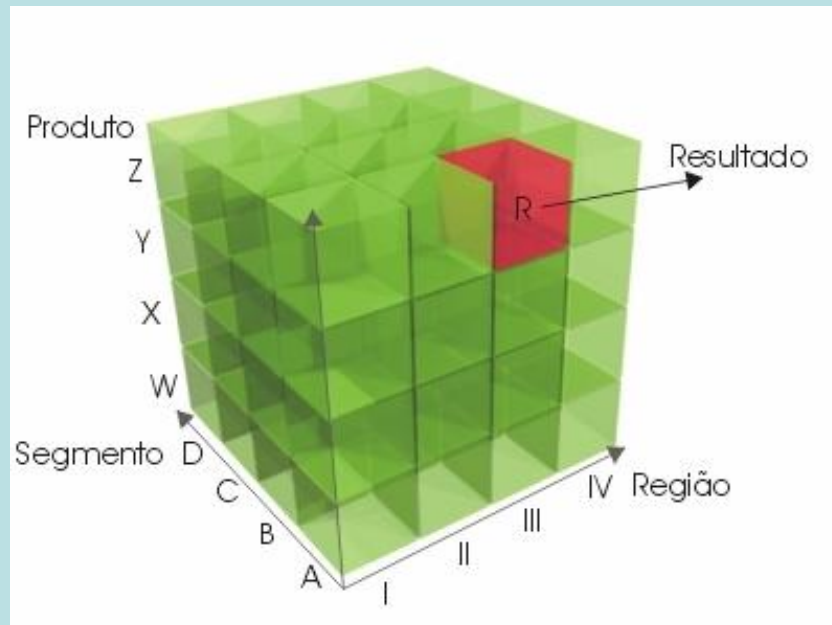
Figure 3.8 Sample data warehouse bus matrix.

As linhas da matriz de barramento correspondem aos data marts. Você deve criar matrizes se as fontes de dados forem diferentes, os processos forem diferentes, ou se a linha representa mais do que o que pode ser razoavelmente abordado em uma única implementação.

Modelagem Dimensional

O CUBO

- ✓ A metáfora do CUBO é apenas uma aproximação da forma como os dados estão organizados.
- ✓ A razão de se utilizar um cubo é para transmitir a ideia de múltiplas dimensões.



Modelagem Dimensional

Exemplo

Fato: Criminalidade.

- ✓ Informação analisada no Fato: “Os índices de criminalidade cresceram em 2004 se comparado com os anos anteriores”.

Ocorrência	2002	2003	2004
Assalto à mão armada	123	109	158
Furtos em residências	75	90	101
Furtos de veículos	243	250	332
Assassinatos	89	77	167



- ✓ Um fato pode ser evolutivo, ou seja, muda suas medidas com o tempo, de acordo com a necessidade de análise.
- ✓ Um fato é analisado a partir de um ponto de vista ou dimensão de análise.

Passos da Modelagem Dimensional

Classificação de Aditividade

- ✓ **Métricas de fluxo de valores:** São processadas por somas, média, mínimo e máximo. Há plena aditividade desta métrica, independente da dimensão que esta sendo tratada. - EX: Quantidade de Produtos vendidos no mês 1 pode ser somado a quantidade de Produtos vendidos no mês 2. Média vendida nos dois meses, ou maior ou menor venda.
- ✓ **Métricas de nível:** Relacionadas a medidas de valor cumulativo. Indica que a soma da medida não produz um valor correto, mas tem sentido tirar a média, o máximo ou mínimo. - Exemplo: Número de peças em estoque no mês 1 não pode ser somado ao número de peças em estoque no mês 2. Mas pode ser tirada a média de peças em estoque nos dois meses.
- ✓ **Métricas relativas:** Indicam valores relativizados a uma base, como por exemplo porcentagem, taxa. Também não podem ser somados, mas é possível aplicar operações de média, máximo e mínimo - Exemplo: Porcentagem de desconto. Taxa de conversão entre real e dólar.



Passos da Modelagem Dimensional

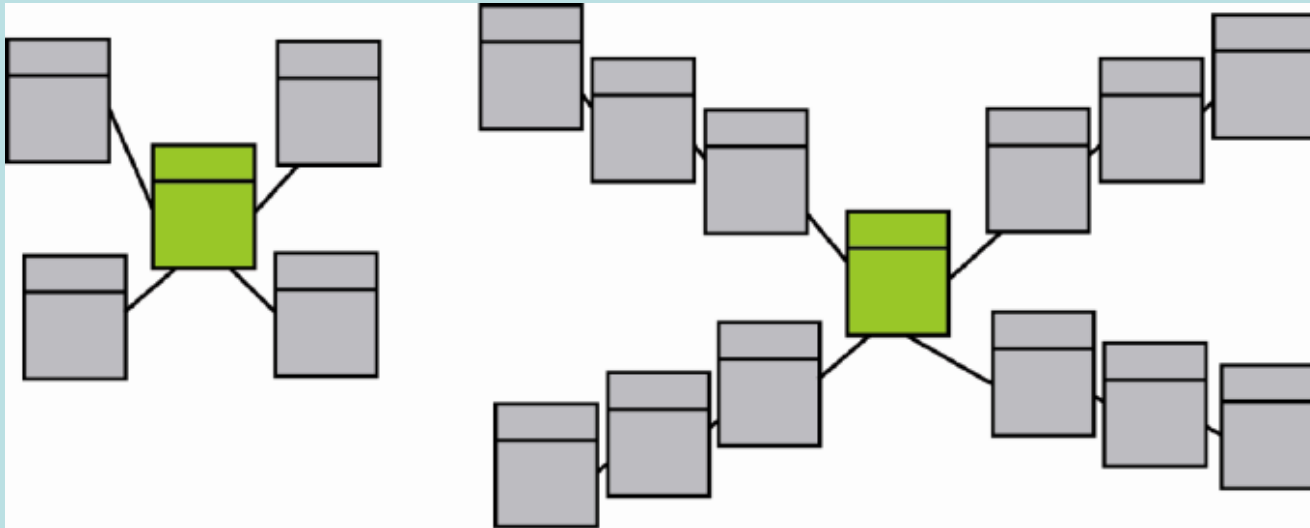
Definição do Atributos da Tabela Fato

- ✓ Com isso, podem também ser classificadas quanto à aditividade:
 - ✓ Aditivas: podem ser sumarizadas independente das dimensões utilizadas.
 - ✓ Não aditivas: não podem ser sumarizadas ao longo das dimensões - perdem o sentido (%)
 - ✓ Semi-aditivas: ela pode perder sentido para a análise caso seja agregada (saldo).

Passos da Modelagem Dimensional

Normalização das Tabelas Dimensão

Tabelas Dimensão estão mais sujeitas a normalização do que tabelas Fato.



Normalização de banco de dados é um conjunto de regras que visa a organização de um projeto de banco de dados, evitando perda ou repetição e informações.

Para normalizar o banco de dados, deve-se examinar os atributos de uma entidade, com o objetivo de se evitar anomalias observadas na inclusão, exclusão e alteração de registros.

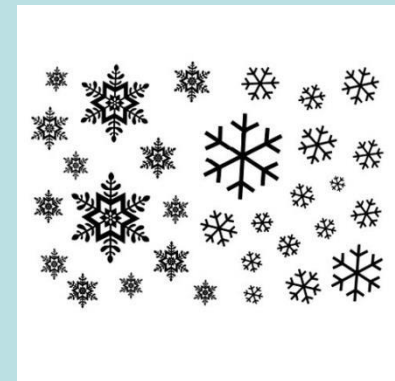
Principais Tipos de Modelos

- ✓ No modelo multidimensional, os fatos e dimensões podem ser dispostos segundo diferentes configurações:
- ✓ Esquema Estrela ou *Star Schema*.



- ✓ Esquema Constelação de Fatos.

- ✓ Esquema Flocos de Neve ou *Snow Flake*.

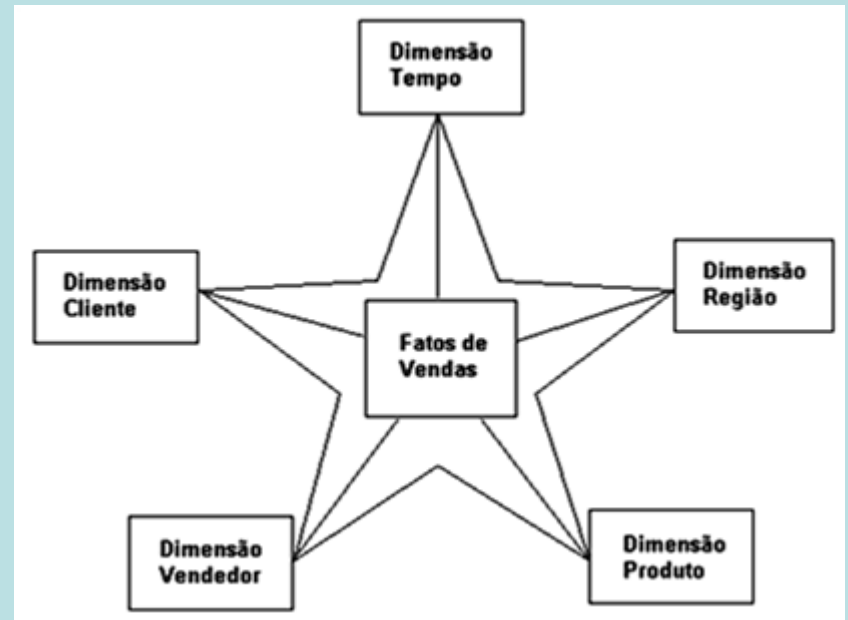


Principais Tipos de Modelos

Esquema Estrela ou Star Schema

- ✓ Proposto por Ralph Kimball.
- ✓ Dominante no projeto de DW.
- ✓ Características:

- Distingue melhor as dimensões dos fatos medidos.
- Simplifica a visualização dimensional.
- Bastante voltada para a abordagem relacional (a literatura fala sempre em tabelas).



Principais Tipos de Modelos

Esquema Estrela ou Star Schema

- ✓ Assim chamado pois é representado com uma tabela central FATO, cercada de dimensões.
- ✓ Importante é que as dimensões não são normalizadas e não existe relacionamento entre as mesmas.
- ✓ O relacionamento entre a entidade fato e cada dimensão é um relacionamento de um para muitos, no sentido da dimensão para o fato.

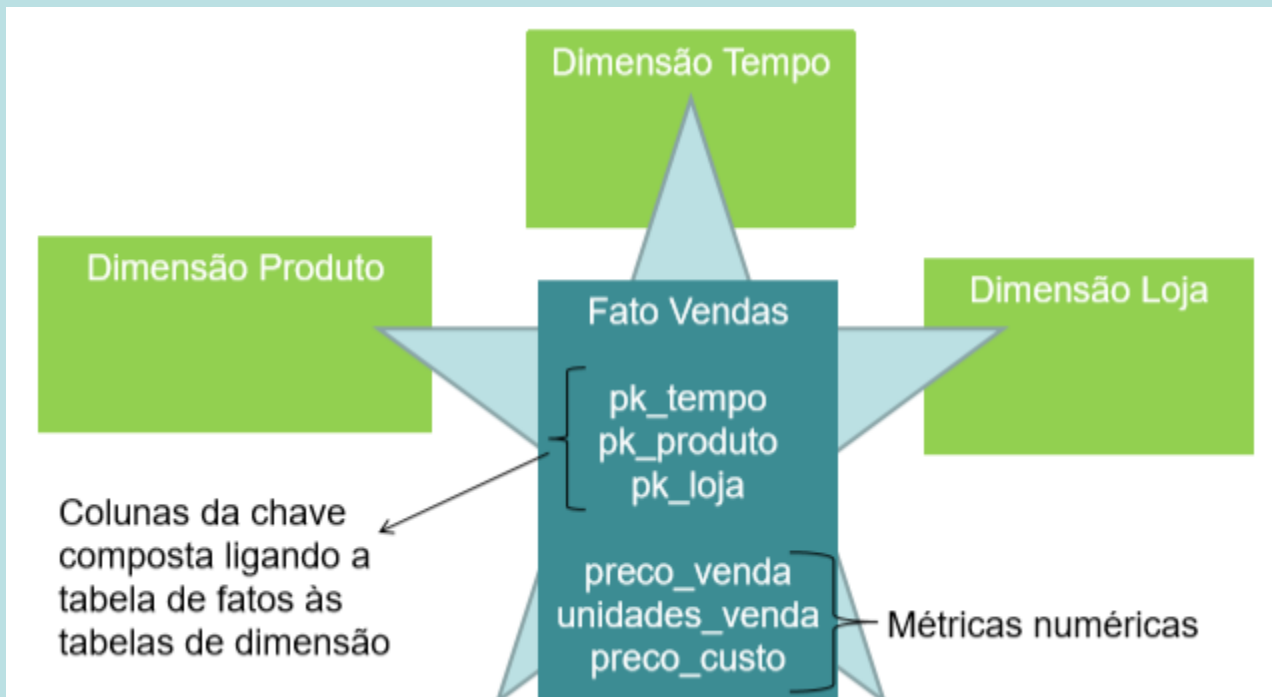


Principais Tipos de Modelos

Exemplos Esquema Estrela:

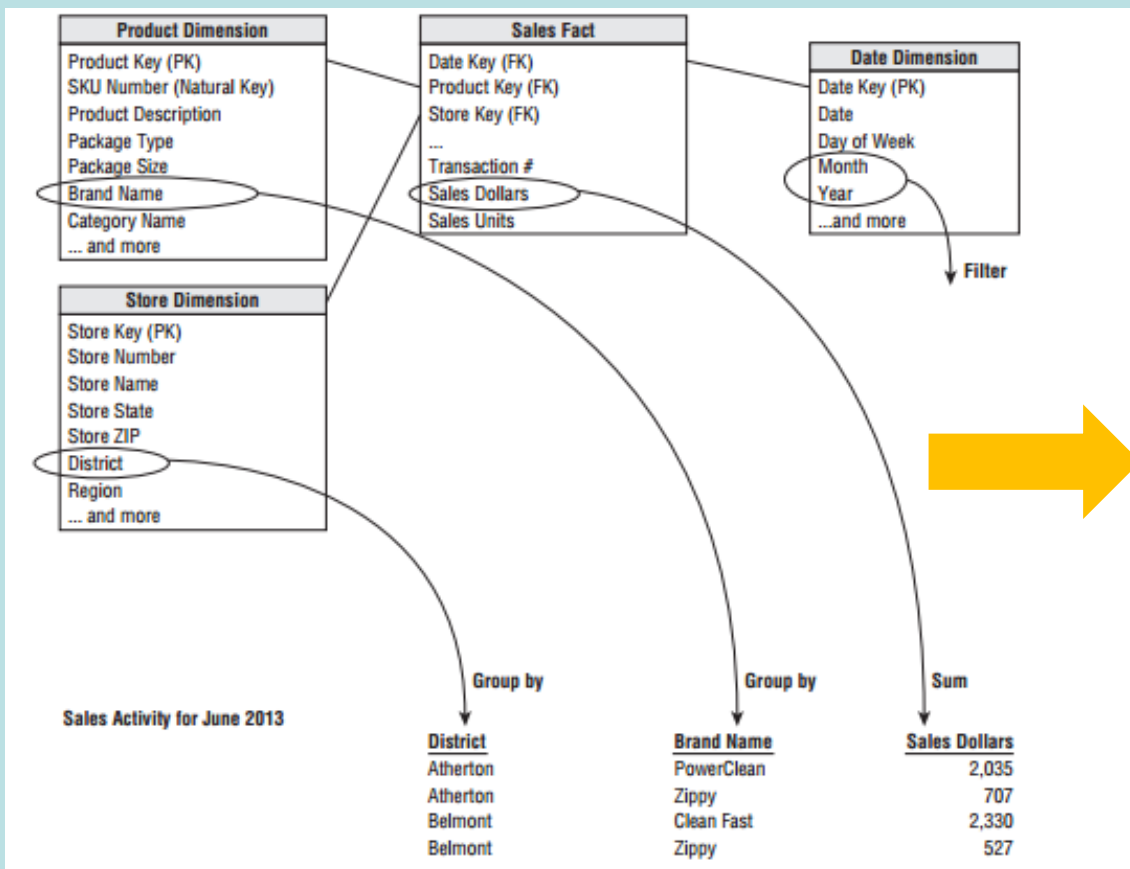
“Vendas por categoria de produto sobre os últimos seis meses.”

“Vendas por marca entre 2000 e 2016.”



Principais Tipos de Modelos

Como esse processo é montado? O Modelo Estrela x Dashboard Final x ETL de carga...

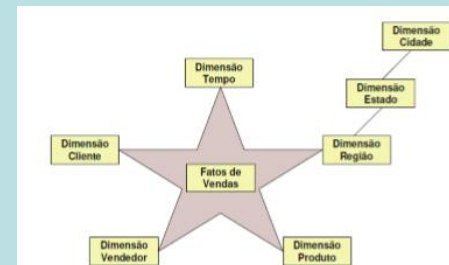


```
SELECT
    store.district_name,
    product.brand,
    sum(sales_facts.sales_dollars) AS "Sales Dollars"
FROM
    store,
    product,
    date,
    sales_facts
WHERE
    date.month_name="January" AND
    date.year=2013 AND
    store.store_key = sales_facts.store_key AND
    product.product_key = sales_facts.product_key AND
    date.date_key = sales_facts.date_key
GROUP BY
    store.district_name,
    product.brand
```

Principais Tipos de Modelos

Esquema Floco de Neve ou Snow Flake

- ✓ O esquema floco de neve é uma variação do esquema estrela no qual todas as tabelas dimensão são totalmente normalizadas (Ex.: produto, grupo produto, tipo produto).
- ✓ Reduzem a redundância mas aumentam a complexidade do esquema e consequentemente a compreensão por parte dos usuários que necessitam extrair a informação por eles mesmos.
- ✓ Dificultam a implementação em determinadas ferramentas de visualização dos dados (Qlik é um exemplo).
- ✓ Considerada uma decomposição de ou mais dimensões que possuem hierarquias.

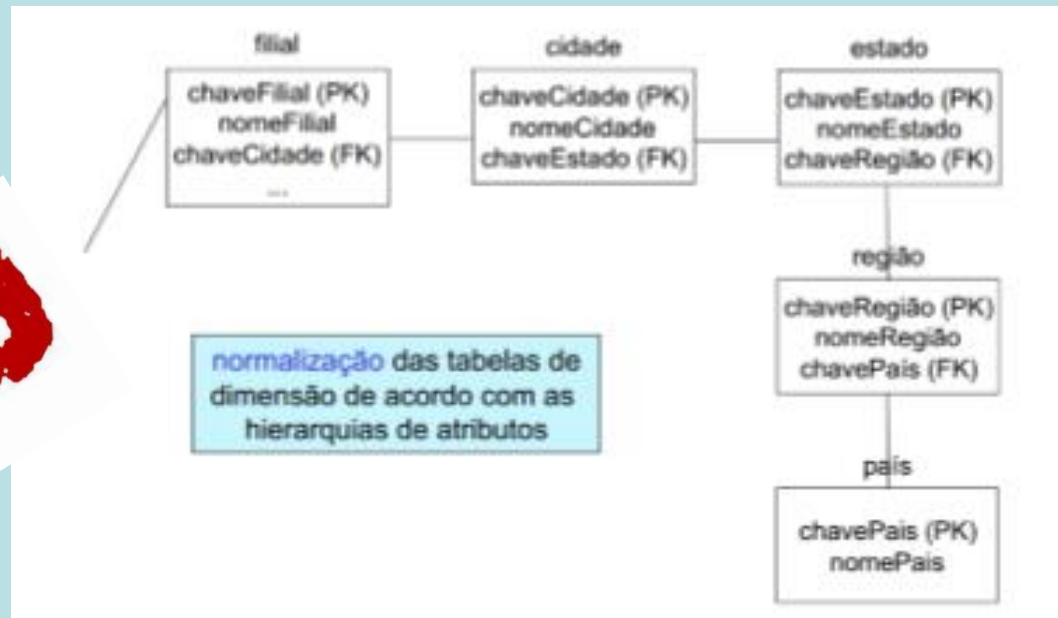


Principais Tipos de Modelos

Esquema Floco de Neve ou Snow Flake

- ✓ A dimensão principal, aquela que se relaciona com o fato, possui o menor nível da hierarquia.

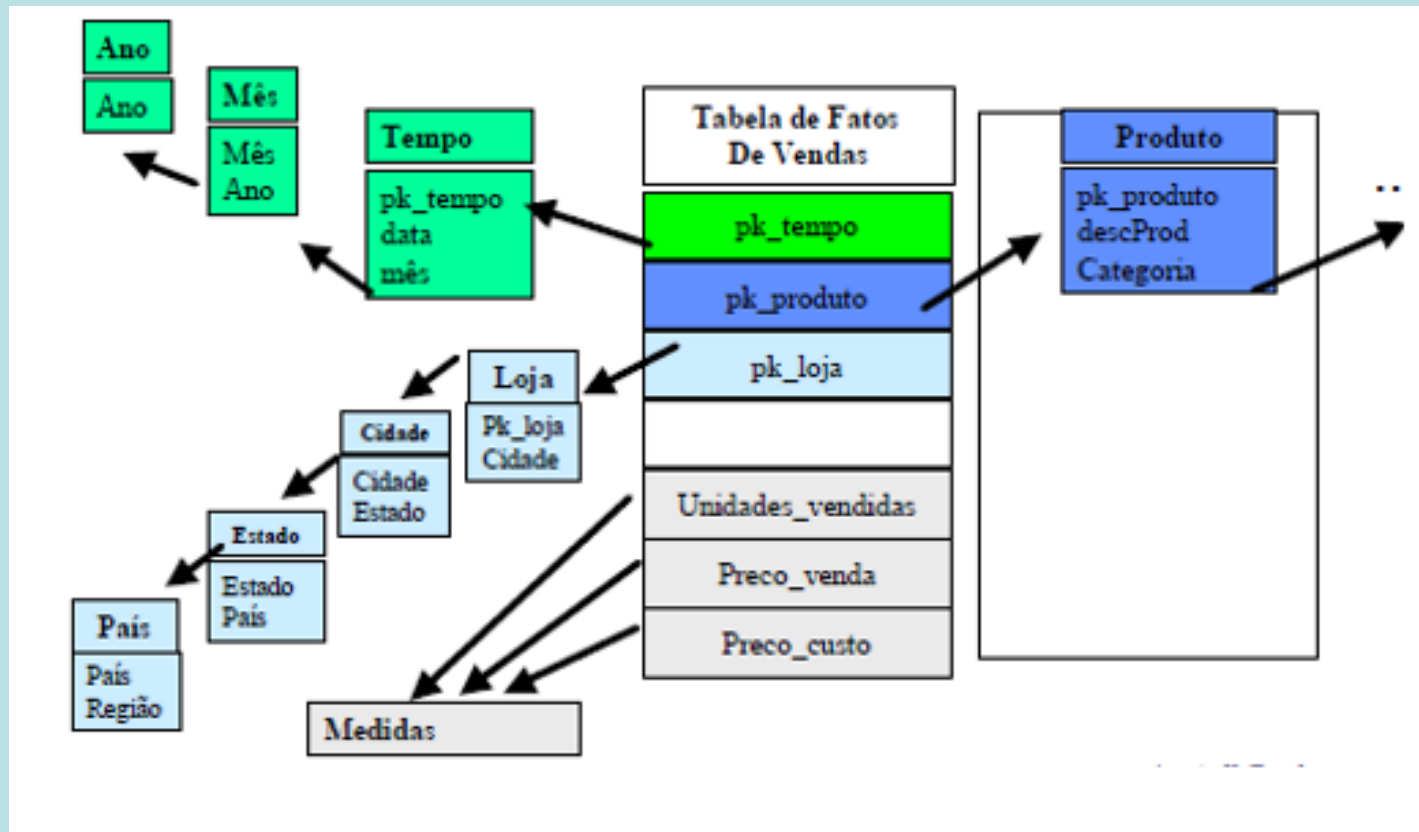
FATO



Principais Tipos de Modelos

Esquema Floco de Neve ou Snow Flake

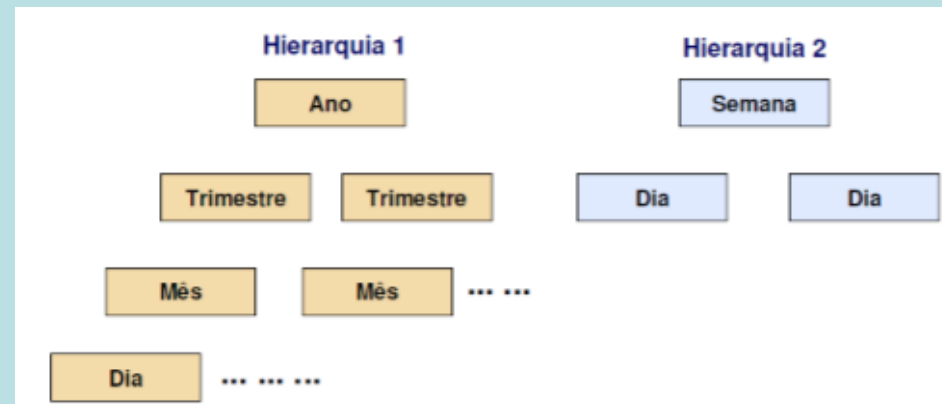
✓ Exemplo



Principais Tipos de Modelos

Hierarquias de Dimensões

- ✓ Uma dimensão pode ter múltiplas hierarquias além de outros atributos descritivos.
- ✓ 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 para uma empresa atacadista:
 - Geografia física: cep, cidade, estado, região, país
 - Geografia de vendas: território, região, zona

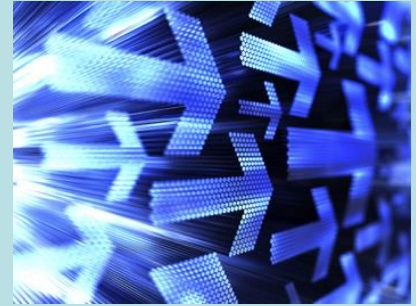


Principais Tipos de Modelos

Star Schema x Snow Flake

	Star Schema	Snow Flake
Performance	+	-
Armazenamento	-	+
Simplicidade	+	-

Principais Tipos de Modelos



Fatos - boas práticas

Existem três tipos básicos de tabela fato:

✓ **Transacional** (Transaction Grain): armazena um fato no seu nível mais granular.

Exemplo: Cada movimentação bancária (saques, depósitos), cada item vendido em cada Nota Fiscal.

✓ **Periódico** (Periodic Snapshot): armazena um acumulado num período de tempo.

Exemplo: Movimentação diária bancária. Todos os itens vendidos de um determinado produto.

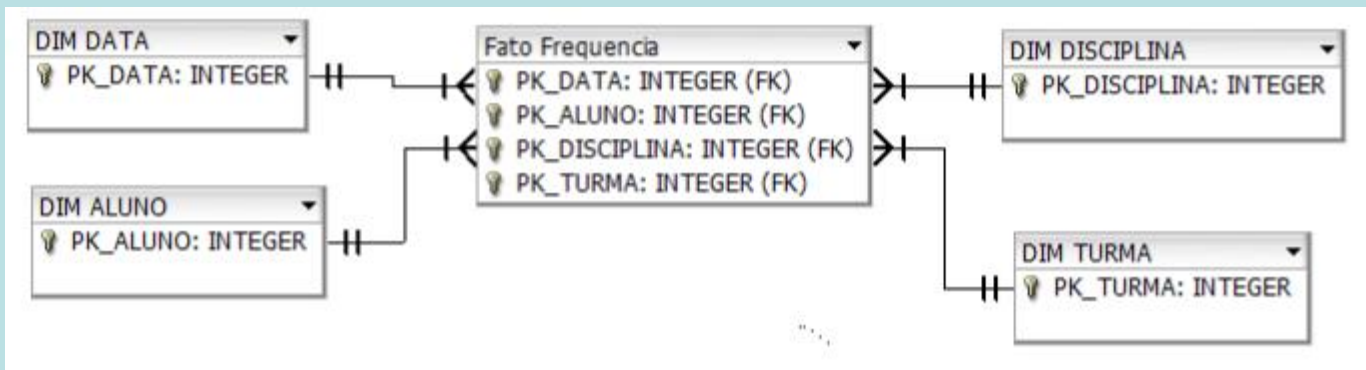
✓ **Cumulativo** (Accumulating snapshot) : armazena fato que sofre alteração num período de tempo.

Exemplo: Acompanhamento de pedidos, processo.

Principais Tipos de Modelos

Tabela Fato sem Fatos

- ✓ Quando não é necessário armazenar nenhuma métrica em uma tabela fato.
- ✓ O simples armazenamento de um relacionamento muitos-para-muitos entre chaves de dimensões é tudo que uma área de negócios necessita como métrica.

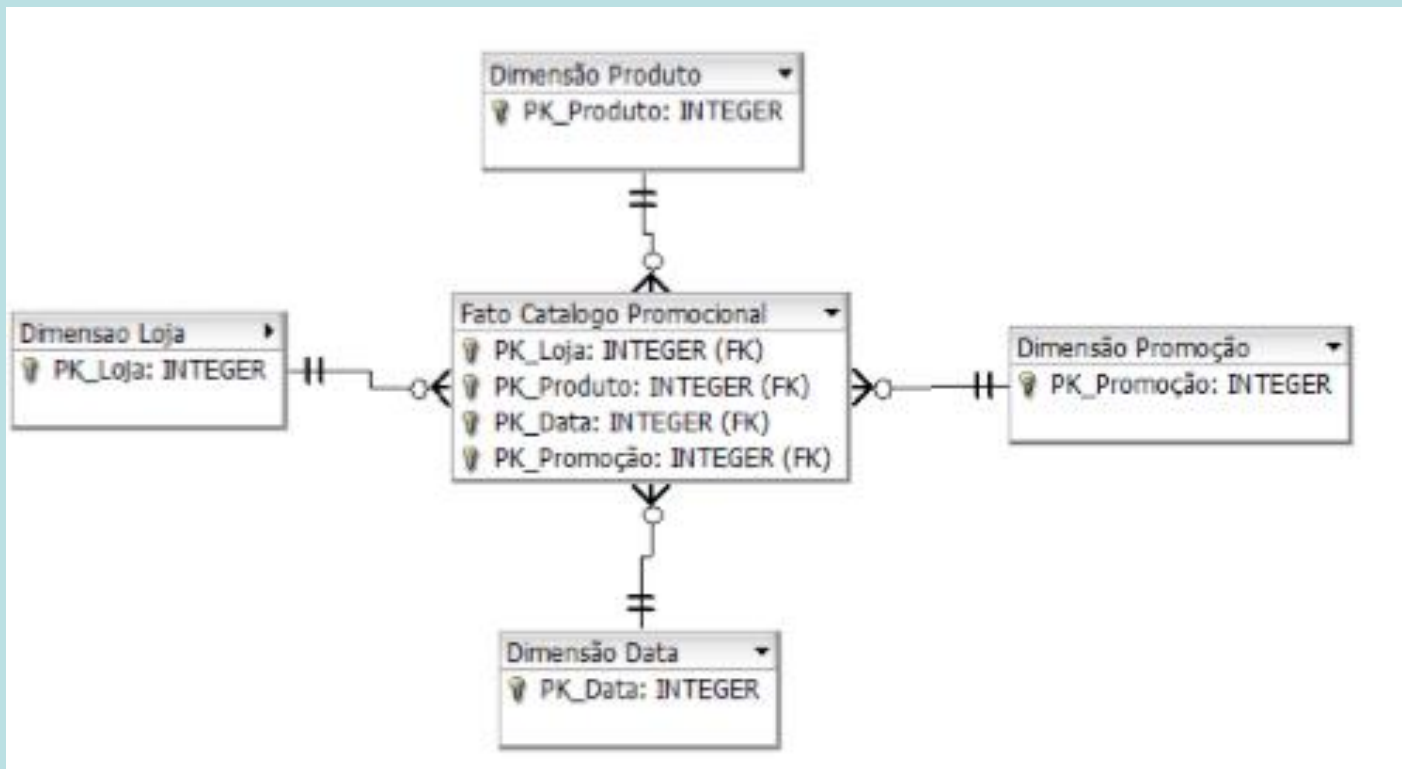


Um simples select count nessa fato determina a quantidade de registros.

Principais Tipos de Modelos

Outro exemplo

- ✓ Uma tabela de fatos, tipicamente sem fatos, que registra todos os produtos que estão em promoção numa determinada loja, independentemente de ser vendidos ou não.



Principais Tipos de Modelos

Fatos Agregadas

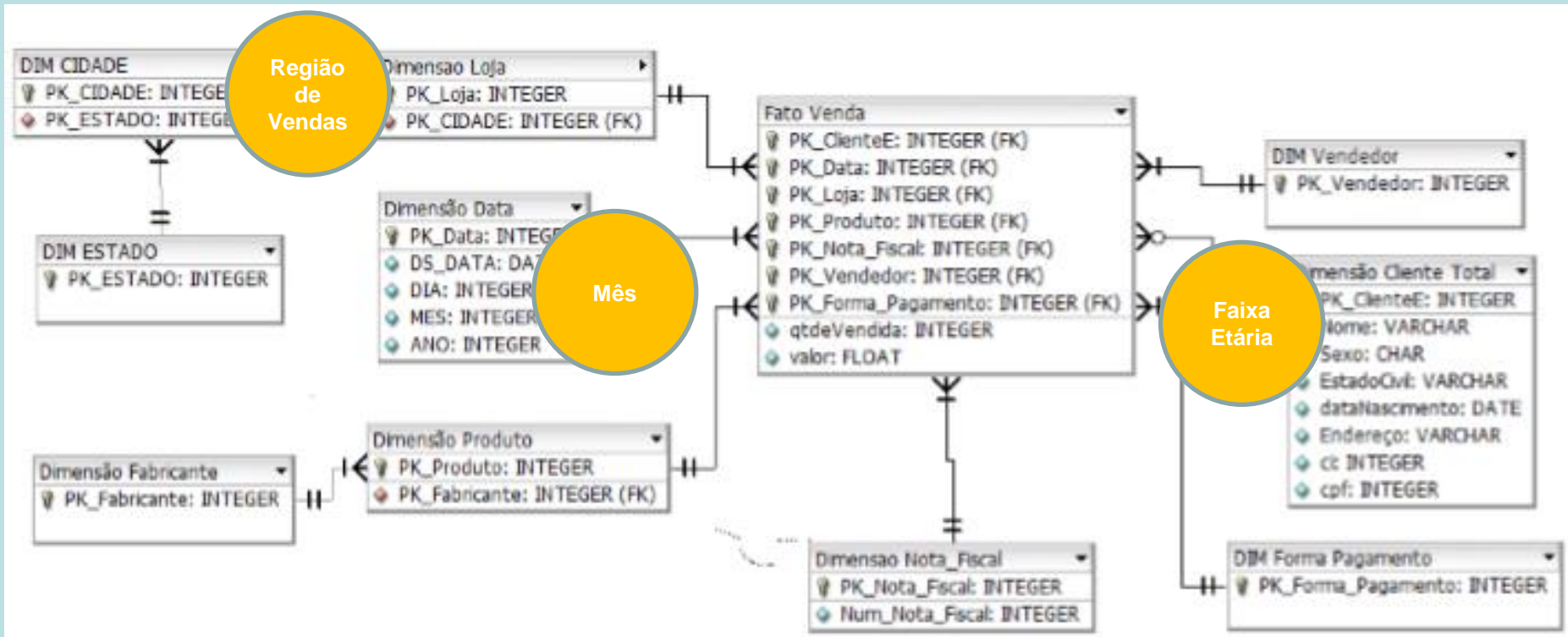
- ✓ Consiste em criar uma nova tabela fato, com os dados de outra tabela de fatos, alterando a granularidade da mesma, gerando assim tabelas bem menores, com dados sumarizados.
- ✓ A criação de tabelas fatos agregadas é uma técnica utilizada para obter ganhos de performance.

Vejamos...

Principais Tipos de Modelos

Fatos Agregados

- ✓ Quais fatos agregados podemos definir a partir do fato abaixo?



Principais Tipos de Modelos

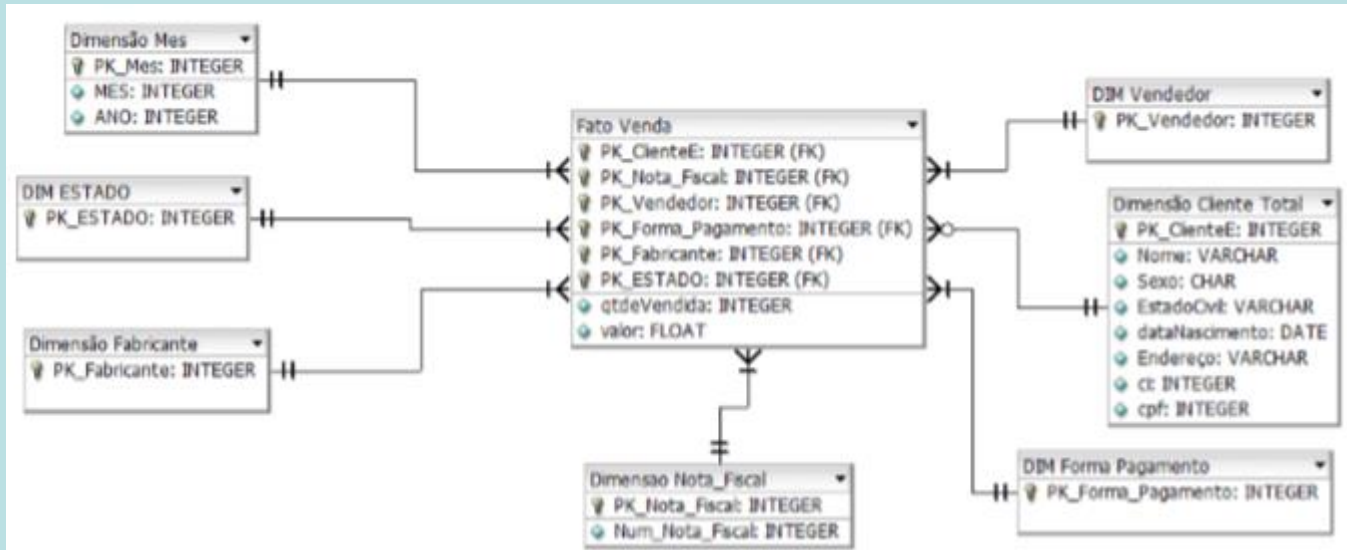
Fatos Agregadas

✓ Mais exemplos:

Dia → Mês

Produto → Fabricante

Loja → Estado



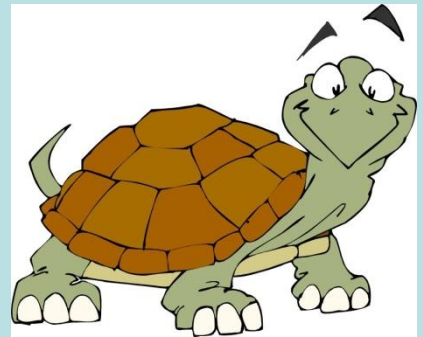
Dinâmica de Mudança das Dimensões

Estratégica de manutenção de dimensões quando ocorre atualização de dados.

1. *Slowly Change Dimension*: sofrem poucas mudanças.

Características: dimensões que tendem a ter poucas alterações.

Exemplo: Loja e Produto.



O que ocorre é que quando o processo de carga detecta que um registro foi alterado na origem, os atributos que foram afetados são atualizados na tabela de dimensão.

Existem 3 estratégias de atualização:

1. Substitui-se o valor antigo pelo novo valor, como se o valor antigo nunca tivesse existido.
2. Cria-se um novo registro do elemento alterado.
3. Mantém-se o valor original e acrescenta-se mais um campo com o novo valor.

Dinâmica de Mudança das Dimensões

Estratégica de manutenção de dimensões quando ocorre atualização de dados.

2. *Rapidly Change Dimension*: mudam constantemente.

Exemplo: Cliente

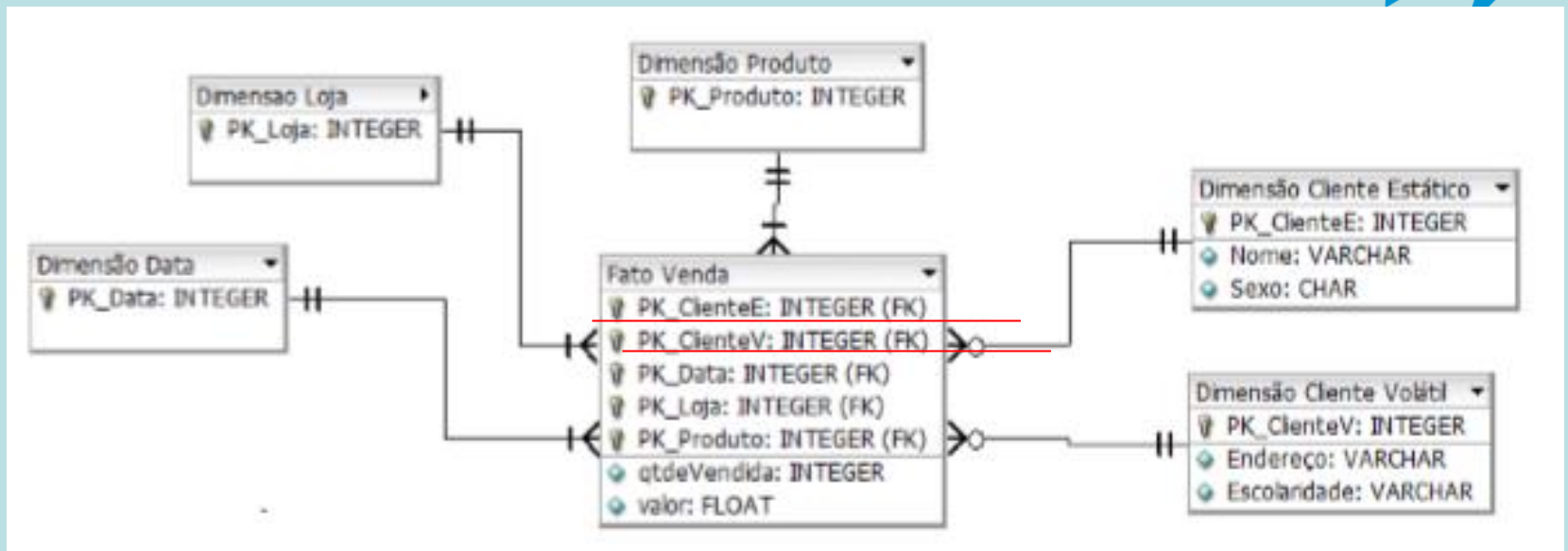
- ✓ Mudanças em atributos como: idade, endereço, escolaridade, etc.
- ✓ A estratégia de atualização recomendada é manter duas dimensões separadas, uma para os dados estáticos e outra para os dados voláteis, com um relacionamento que permita navegar entre as duas dimensões.
- ✓ Neste caso a dimensão volátil, guarda o histórico, ou seja, todas as versões do registro alterado.
- ✓ A tabela Fato mantém a chave de ambas as dimensões.



Dinâmica de Mudança das Dimensões

Estratégia de manutenção de dimensões quando ocorre atualização de dados.

Exemplo:



Dinâmica de Mudança das Dimensões

Estratégica de manutenção de dimensões quando ocorre atualização de dados.

3. *Dimensão Sucata (Junk Dimension).*

Junk Dimension (dimensão lixo) é uma coleção de códigos aleatórios transacionais, indicadores e atributos textuais com significado.

Exemplos:

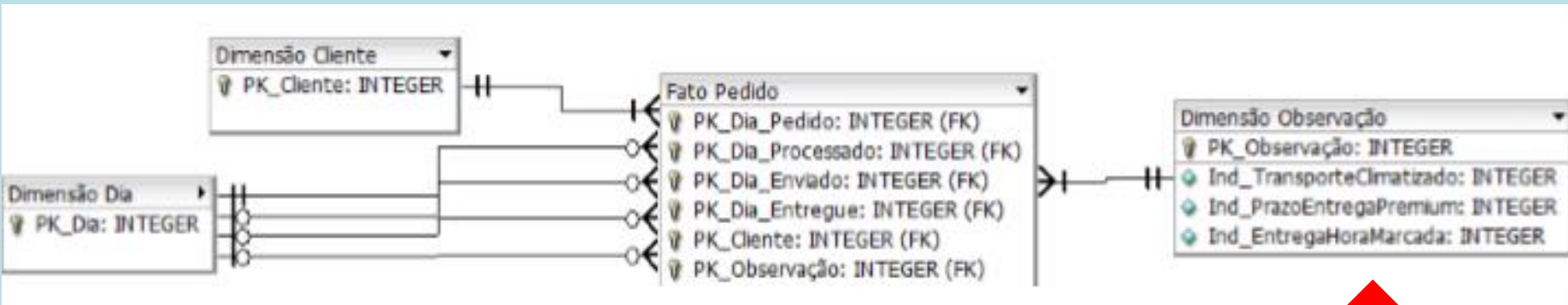
- ✓ Gênero (Feminino, Masculino);
- ✓ Estado Civil (Solteiro, Casado, etc);
- ✓ Tipo de Servidor (Federal, Estadual, Municipal);
- ✓ Indicador de Sim e Não, etc.



Dinâmica de Mudança das Dimensões

Estratégia de manutenção de dimensões quando ocorre atualização de dados

3. *Dimensão Sucata (Junk Dimension).*



Dinâmica de Mudança das Dimensões

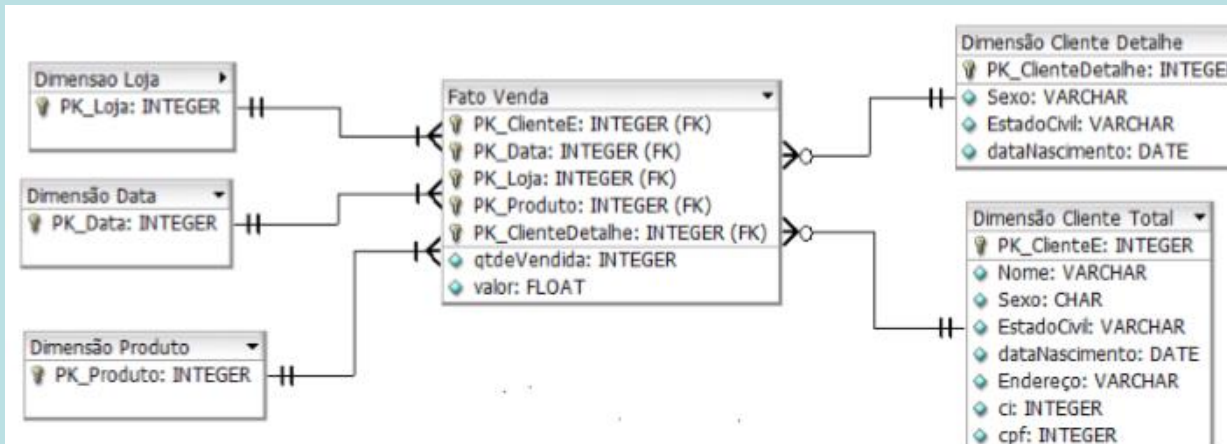
Estratégica de manutenção de dimensões quando ocorre atualização de dados.

4. Dimensão Gigante (Huge Dimension)

- ✓ Dimensões gigantes são aquelas que acumulam milhões de registros. Navegação de baixa performance.

5. Mini Dimensões

- ✓ Uma Mini Dimensão visa melhorar a performance de análise envolvendo dimensões gigantes. Possui combinações de um conjunto de atributos de uma dimensão gigante.



Dinâmica de Mudança das Dimensões

Estratégica de manutenção de dimensões quando ocorre atualização de dados.

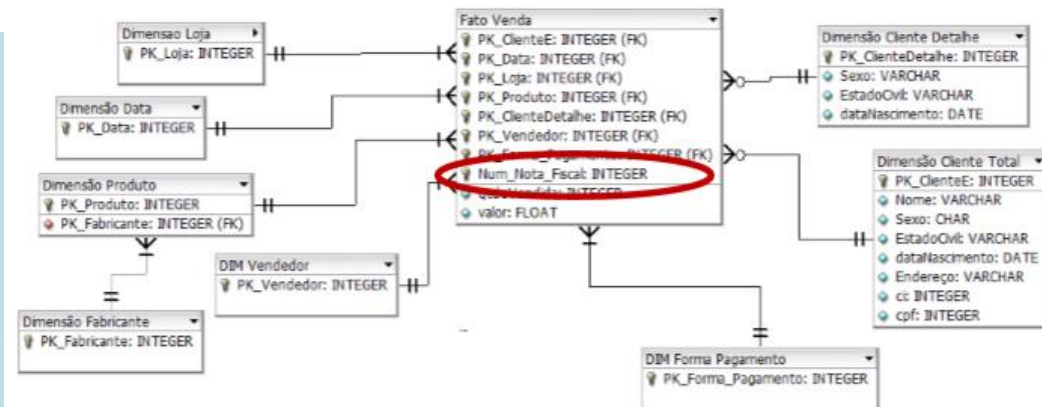
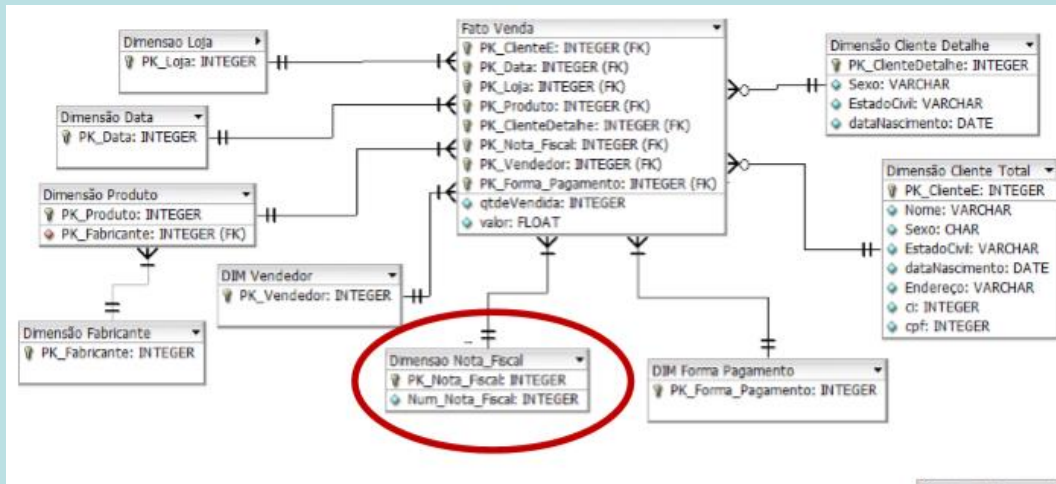
6. Dimensão Degenerada (*Degenerate Dimension*)

- ✓ Números de controle de documentos, como números de pedidos, números de fatura e números de conhecimento de embarque, normalmente são representados como dimensões descaracterizadas em tabelas de fatos em que o grão da tabela é o documento propriamente dito ou uma linha de item do documento.
- ✓ As informação relacionada está em outra dimensão.
- ✓ Pode não ser necessário criar a tabela dimensão fisicamente.

Dinâmica de Mudança das Dimensões

Estratégia de manutenção de dimensões quando ocorre atualização de dados.

6. Dimensão Degenerada (Degenerate Dimension)



Dimensões Especiais

- ✓ **TEMPO:** Presente em todos os projetos. Característica de histórico no DW.
- ✓ **ESPAÇO:** Um fato pode variar de acordo com o local. Pode ser: Loja, Órgão, Departamento, Filial, etc. Em geral, possuem uma hierarquia com localização geográfica (Cidade, estado, etc).
- ✓ **OBJETO:** Relacionado ao objetivo do negócio. Geralmente envolve os objetos que realizam o fato ou que são necessários para que ele aconteça.



Operações OLAP

Operações dimensionais de ferramentas OLAP:

✓ Drill Down

- alto nível → baixo nível: ocorre quando o usuário aumenta o nível de detalhe da informação, diminuindo o grau de granularidade.

✓ Drill Up (Roll Up)

- baixo nível → alto nível: ocorre quando o usuário diminui o nível de detalhe da informação, aumentando o grau de granularidade.

Com o Drill Down e o Roll up você pode “subir ou descer” dentro do detalhamento do dado, como, por exemplo, analisar uma informação tanto diariamente quanto anualmente, partindo da mesma base.

Operações OLAP



Operações OLAP

MES	BANCO	REGIAO	QUANTIDADE EMITIDA
jan/13	Banco do Brasil	Sul	100
jan/13	Santander	Sul	120
jan/13	Banco do Brasil	Sudeste	250
jan/13	Santander	Sudeste	160
fev/13	Bradesco	Sul	150
fev/13	Banco do Brasil	Sul	130
fev/13	Santander	Sul	120
fev/13	Bradesco	Sudeste	200
fev/13	Banco do Brasil	Sudeste	230
fev/13	Santander	Sudeste	220

Roll up

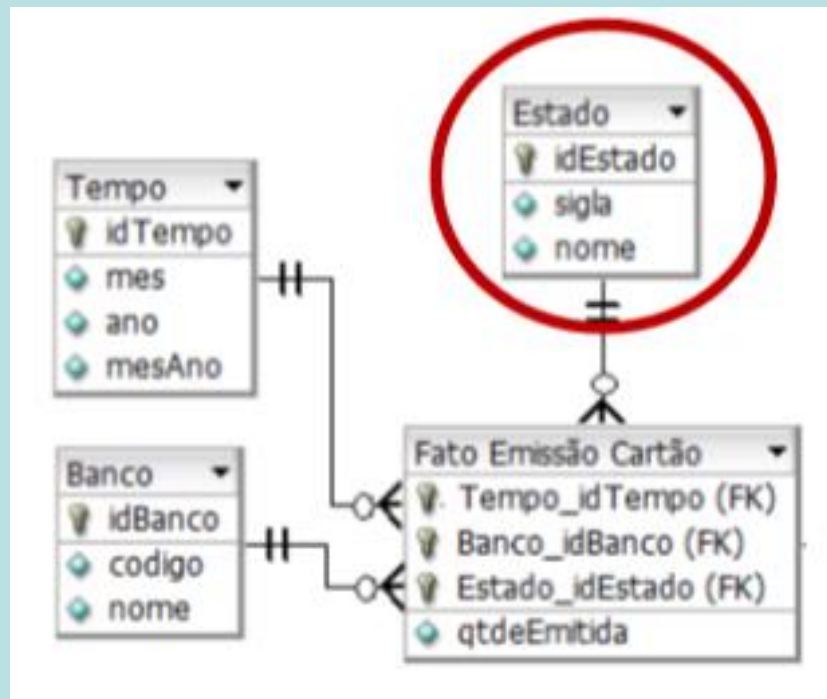
MES	BANCO	REGIAO	ESTADO	QUANTIDADE EMITIDA
jan/13	Banco do Brasil	Sul	Rio Grande do Sul	35
jan/13	Banco do Brasil	Sul	Santa Catarina	32
jan/13	Banco do Brasil	Sul	Paraná	33
jan/13	Santander	Sul	Rio Grande do Sul	41
jan/13	Santander	Sul	Santa Catarina	39
jan/13	Santander	Sul	Paraná	40
jan/13	Banco do Brasil	Sudeste	Minas Gerais	60
jan/13	Banco do Brasil	Sudeste	São Paulo	80
jan/13	Banco do Brasil	Sudeste	Rio de Janeiro	70
jan/13	Banco do Brasil	Sudeste	Espírito Santo	40
jan/13	Santander	Sudeste	Minas Gerais	40
jan/13	Santander	Sudeste	São Paulo	55
jan/13	Santander	Sudeste	Rio de Janeiro	35
jan/13	Santander	Sudeste	Espírito Santo	30
fev/13	Bradesco	Sul	Rio Grande do Sul	60
fev/13	Bradesco	Sul	Santa Catarina	40
fev/13	Bradesco	Sul	Paraná	50
fev/13	Banco do Brasil	Sul	Rio Grande do Sul	50

100

Operações OLAP

A operação de roll up não é limitada pelo grão máximo. Os dados podem ser agregados em outras dimensões mesmo após se chegar a este limite superior.

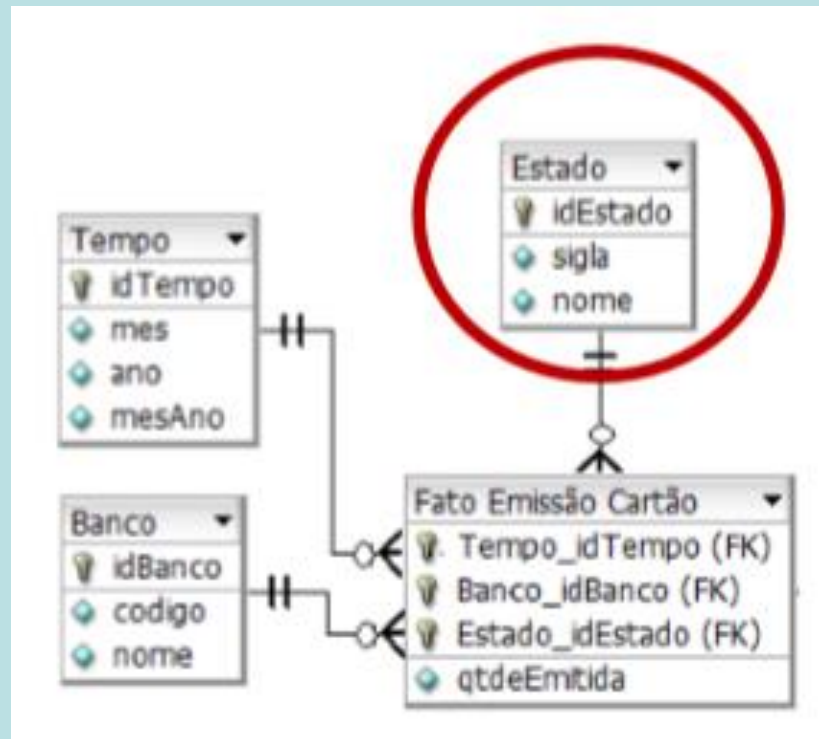
A partir de um ESTADO é possível identificar a REGIÃO que ele pertence.



Operações OLAP

A operação de drill down é limitada pelo grão mínimo.

A partir de um ESTADO não é possível identificar a CIDADE que ele pertence.



Operações OLAP

Operações dimensionais de ferramentas OLAP:

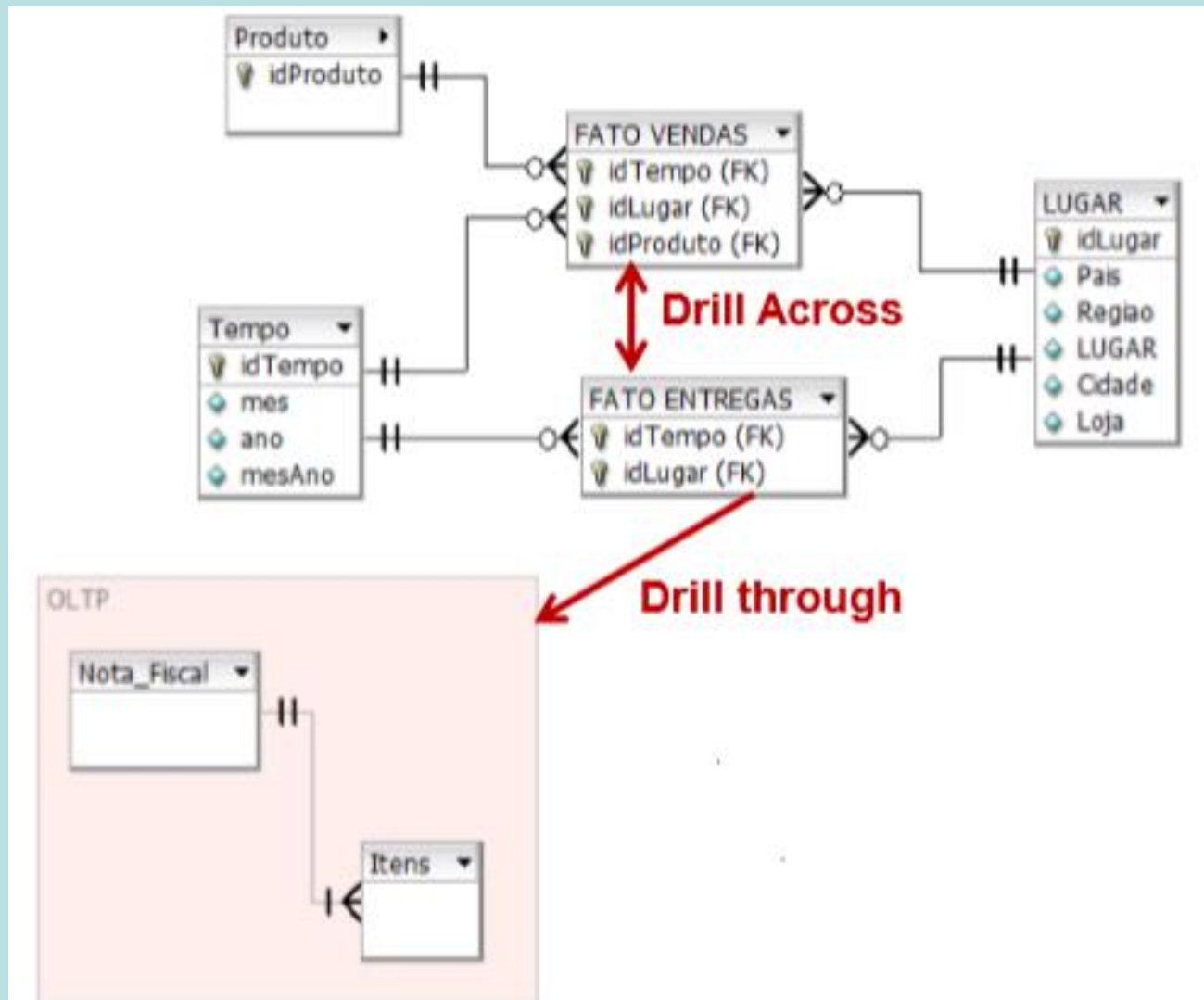
✓ Drill Across

- Envolve mais do que uma tabela de fato ou esquema.
- Devem existir dimensões compartilhadas entre estes esquemas.

✓ Drill Through

- Envolve uma pesquisa em um nível mais baixo do que o data mart permite, e que está presente no seu modelo operacional/transacional.

Operações OLAP



Operações OLAP

Operações dimensionais de ferramentas OLAP:

✓ Slice e Dice


- Fatiar e Cortar o cubo separando partes de um cubo [INMOM, 1999].
- Seleciona partes específicas de um cubo.
- Consiste na rotação do cubo, possibilitando a combinação de quaisquer dimensões.
- **Slice:** Análise de apenas uma fatia. Restringe os valores de uma dimensão, mas não diminui a cardinalidade do cubo.
- **Dice:** Reduz as dimensões ou a cardinalidade de um cubo por meio da eliminação de uma ou mais dimensões.

Operações OLAP

Operações dimensionais de ferramentas OLAP:

- ✓ Slice: corte do cubo e análise somente da fatia. (Plano)

Volume de Produção (em milhares)		Computadores	DVD
Nordeste	Rio Grande do Norte	45	20
	Paraíba	79	43



Volume de Produção (em milhares)		Computadores
Nordeste	Rio Grande do Norte	45

SELEÇÃO: WHERE estado = Rio Grande do Norte

Operações OLAP

Operações dimensionais de ferramentas OLAP:

- ✓ Dice: mudança da perspectiva (rotação do cubo).

Volume de Produção		Computadores	DVD	Celulares	Pagers
Nordeste	Rio Grande do Norte	45	20	31	34
	Paraíba	79	43	50	35

Volume de Produção	Nordeste	
	Rio Grande do Norte	Paraíba
Computadores	45	79
DVD	20	43
Celulares	31	50
Pagers	34	35



Operações OLAP - Quadro Resumo

Operação	Definição
<i>drill-down</i>	analisa os dados em níveis de agregação progressivamente mais detalhados, ou de menor granularidade
<i>roll-up</i>	analisa os dados em níveis de agregação progressivamente menos detalhados, ou de maior granularidade
<i>slice and dice</i>	restringe os dados sendo analisados a um subconjunto destes dados <i>slice</i> : corte para um valor fixo <i>dice</i> : seleção de faixas de valores
<i>drill-across</i>	compara medidas numéricas distintas que são relacionadas entre si através de pelo menos uma dimensão em comum

O Futuro da Modelagem Dimensional by Ralph Kimball

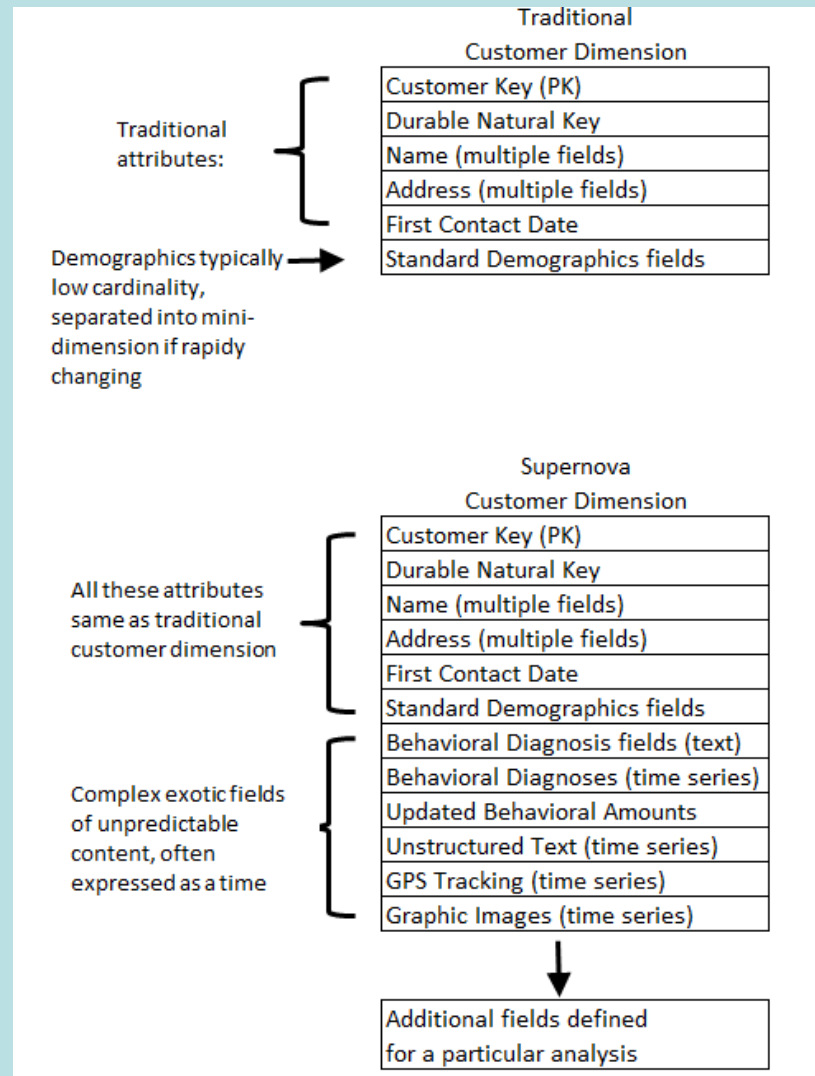
(<http://www.kimballgroup.com/2015/12/design-tip-180-the-future-is-bright/>) em 01-12-2015

Mesmo no admirável mundo de tipos de dados mais estranhos e de processamento não-relacional, a modelagem dimensional é extremamente relevante. Todos os acontecimentos têm sempre um contexto: data, hora, local, cliente / pessoa / paciente, ação, etc.

Expanda sua mente: a junção de novas dimensões a um modelo não precisa acontecer por meio de um banco de dados relacional convencional, porque a correspondência pode ser feita de outras maneiras. As dimensões, é claro, são a alma do *data warehouse*. Fatos são apenas observações que sempre existem em um contexto dimensional. Daqui para frente, podemos esperar que as dimensões se tornem mais poderosas, a fim de oferecer suporte a consultas baseadas em comportamentos mais sofisticados e análise preditiva. Já houve propostas para generalizar o esquema em estrela a um esquema novo. Nesse esquema, atributos dimensionais devem se tornar objetos complexos em vez de um texto simples. Dimensões novas também tornam-se muito mais maleáveis e extensíveis a partir de uma análise para outra. Vejamos um exemplo:

O Futuro da Modelagem Dimensional by Ralph Kimball

(<http://www.kimballgroup.com/2015/12/design-tip-180-the-future-is-bright/>) em 01-12-2015



Ferramentas

- ✓ **CA Erwin** <http://erwin.com/worldwide/portuguese-brazil>
- ✓ **IBM's Relational Rose** <http://www-142.ibm.com/software/products/br/pt/enterprise/>
- ✓ **Oracle's Designer** <http://www.oracle.com/technetwork/developer-tools/developer-suite/downloads/index.html>
- ✓ **DBDesigner** <http://www.fabforce.net/dbdesigner4/dbdesigner-4-0-5-6-1>
- ✓ **ER Studio** <https://www.idera.com/er-studio-enterprise-data-modeling-and-architecture-tools>
- ✓ **Ferramentas Free** <http://www.devmedia.com.br/business-intelligence-conhecendo-algumas-ferramentas-open-source/31963>