

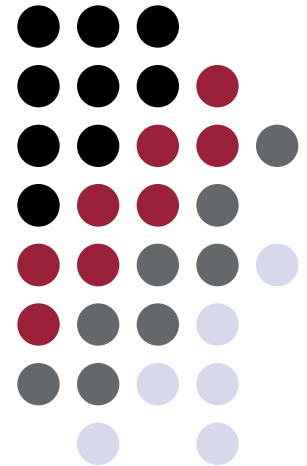
# Modelagem de Data Warehouse



Universidade Federal  
de Ouro Preto

## CEA462 – Sistemas de Apoio à Decisão

Prof. MSc. George H. G. Fonseca  
Universidade Federal de Ouro Preto





- Características do Modelo Entidade-Relacionamento
  - Foco em aplicações transacionais
  - Foco no armazenamento momentâneo (não-histórico) da informação
  - Tende a um grande número de tabelas
  - Eficiente apenas para consultas simples e diretas



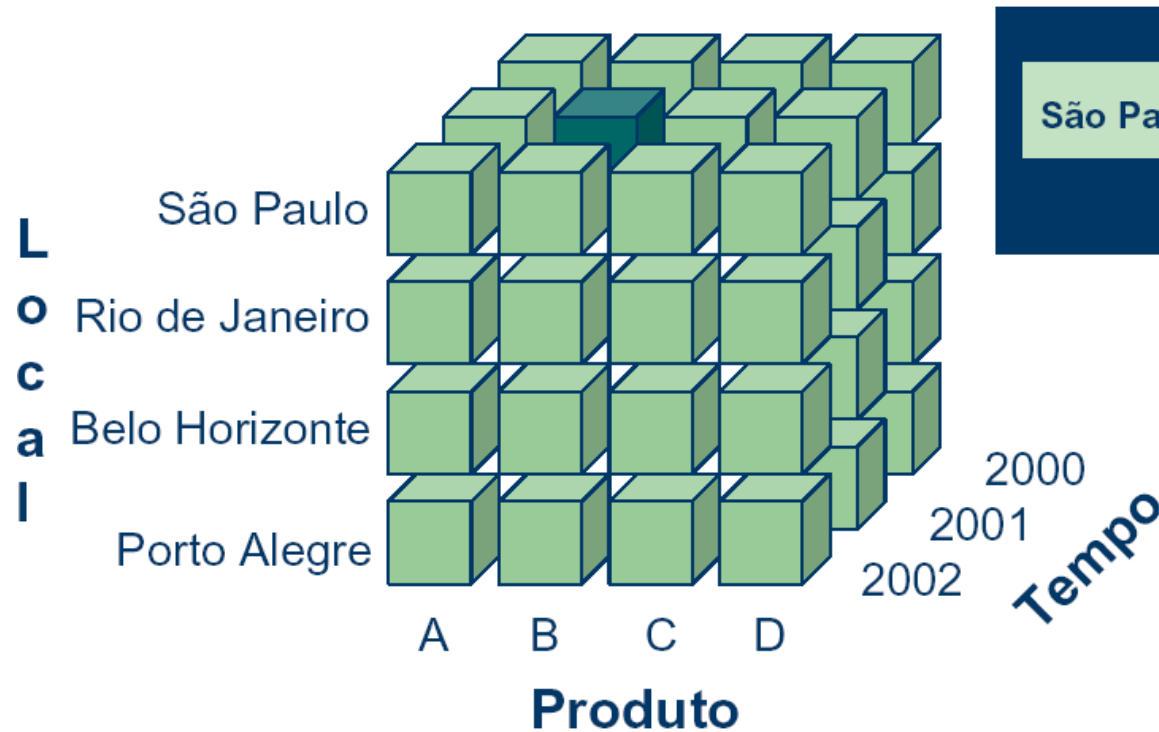
- Necessidades em um ambiente de Data Warehouse
  - Foco em aplicações gerenciais
  - Análise histórica das informações
  - Visão ampla das informações (sumarizações, cruzamentos)
  - Visualizar os dados sob diferentes perspectivas (consultas complexas)
- Para implementar um Data Warehouse necessitamos de um novo modelo, diferente do ER tradicional...

# Modelagem Multidimensional



- Representar os tipos de dados por uma estrutura chamada cubo de dados
  - Células contêm valores
  - Lados definem as dimensões de análise
- Normalmente também refere-se a cubo de dados mesmo quando há mais de 3 dimensões
  - No entanto, o termo técnico para tal estrutura é **Hipercubo**

# Modelagem Multidimensional



# Visão Relacional X Visão Multidimensional



- Visão relacional
  - Volume de vendas de uma loja de instrumentos musicais por instrumento e estado

<b>Instrumento</b>	<b>Estado</b>	<b>Qtde. Vendas</b>
Violão	MG	140
Violão	RJ	100
Violão	SP	150
Guitarra	MG	140
Guitarra	RJ	120
Guitarra	SP	80
Bateria	MG	30
Bateria	RJ	20
Bateria	SP	50

# Visão Relacional X Visão Multidimensional



- Visão multidimensional
  - Volume de vendas de uma loja de instrumentos musicais por instrumento e estado

**Volume de Vendas**

<b>Instrumento</b>	<b>Violão</b>	140	100	150
	<b>Guitarra</b>	140	120	80
	<b>Bateria</b>	30	20	50
		<b>MG</b>	<b>RJ</b>	<b>SP</b>

**Estado**

# Visão Relacional X Visão Multidimensional



- Visão relacional

- Volume de vendas de uma loja de instrumentos musicais por instrumento e estado

Instrumento	Estado	Qtde. Vendas
Violão	MG	30
Violão	MG	60
Violão	MG	50
Violão	RJ	10
Violão	RJ	30
Violão	RJ	60
Violão	SP	20
Violão	SP	40
Violão	SP	55
Guitarra	MG	15
Guitarra	MG	30
Guitarra	MG	50
Guitarra	RJ	70
Guitarra	RJ	60
Guitarra	RJ	45
Guitarra	SP	30
Guitarra	SP	20
Guitarra ...	SP	50

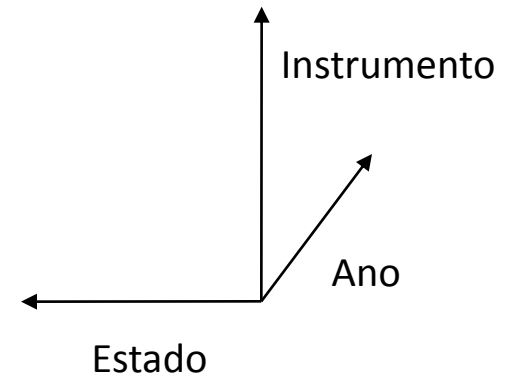
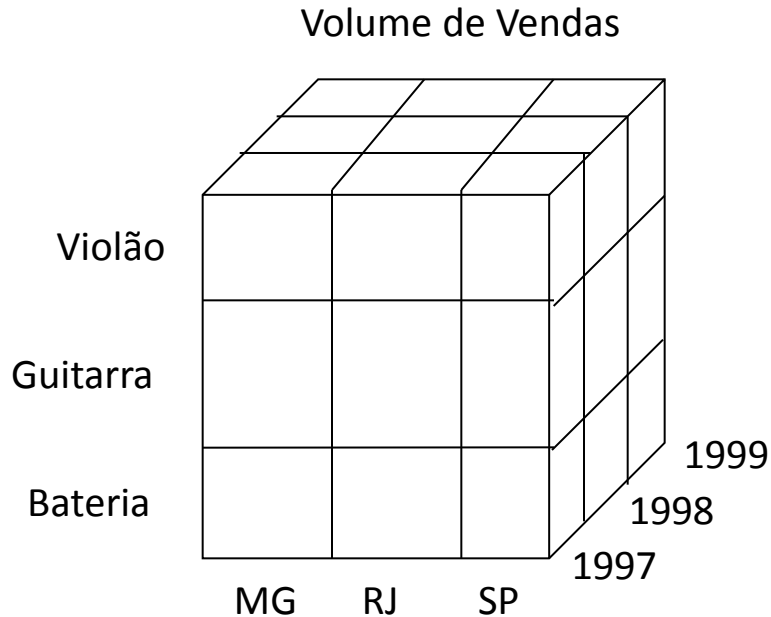


# Visão Relacional X Visão Multidimensional



- **Visão multidimensional**

- Volume de vendas de uma loja de instrumentos musicais por instrumento e estado

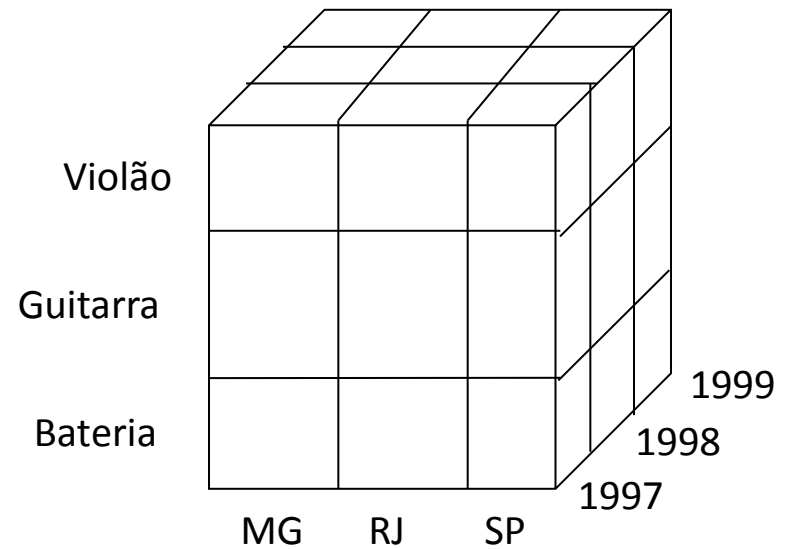
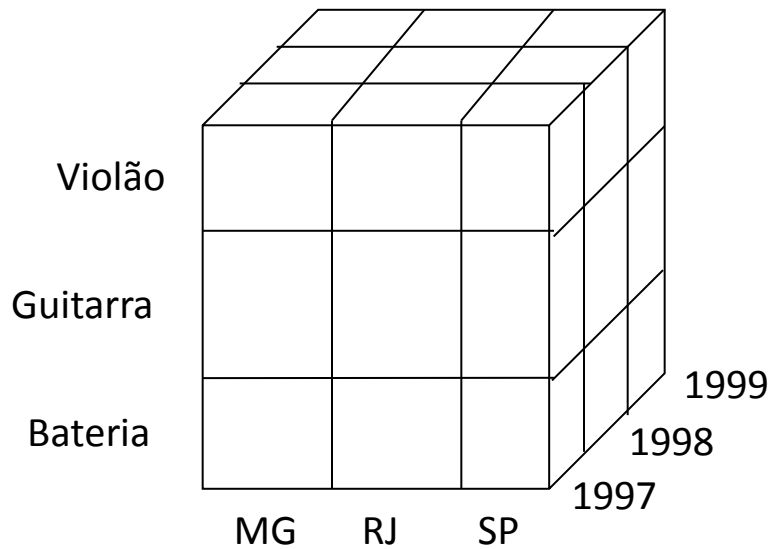


# Visão Relacional X Visão Multidimensional



- Visão multidimensional
  - Hiper cubo

Volume de Vendas



Importado

Nacional

# Modelagem Multidimensional



- Elementos básicos

- Fatos – aquilo que pode ser representado por valores numéricos. Esse conjunto de valores é também chamado métricas ou medidas
  - Ex.: Vendas
- Dimensões – determinam o contexto no qual os fatos são analisados
  - Ex.: Local, Ano e Produto
- Variáveis – atributos numéricos que representam os fatos
  - Ex.: Valor (R\$) das vendas, Unidades vendidas

# Modelagem Multidimensional

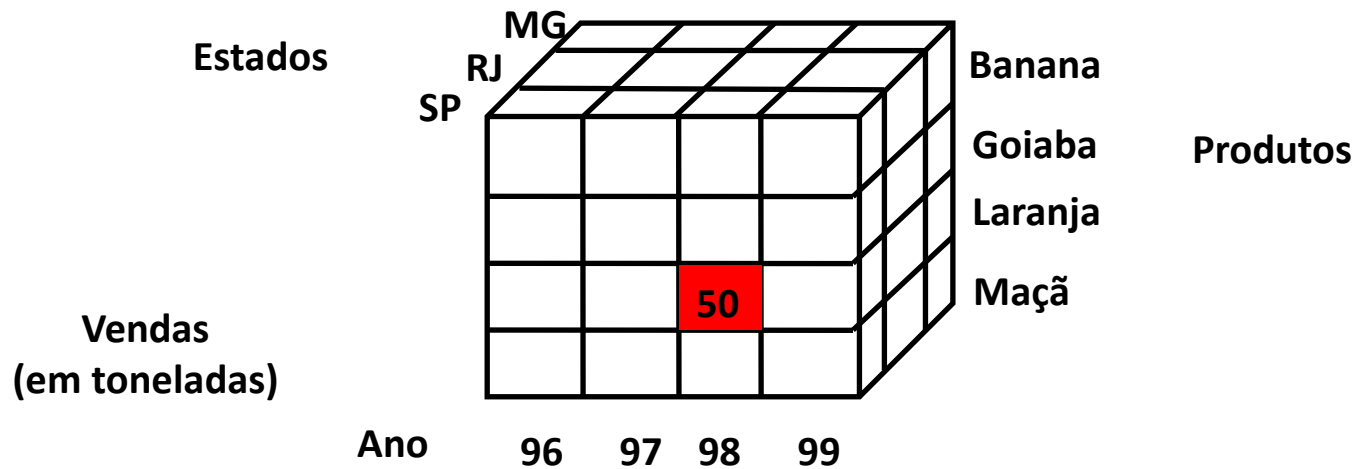


- Star Schema (Esquema Estrela)
  - Forma de **dispor as tabelas do banco** para simular um banco de dados multidimensional
  - Composto por uma tabela dominante, chamada **tabela de fatos**, rodeada de tabelas auxiliares, chamadas **tabelas de dimensão**;
  - A tabela de fatos conecta-se às demais por **múltiplas junções** e as tabelas de dimensões se conectam com **apenas uma junção** à tabela de fatos.

# Modelagem Multidimensional



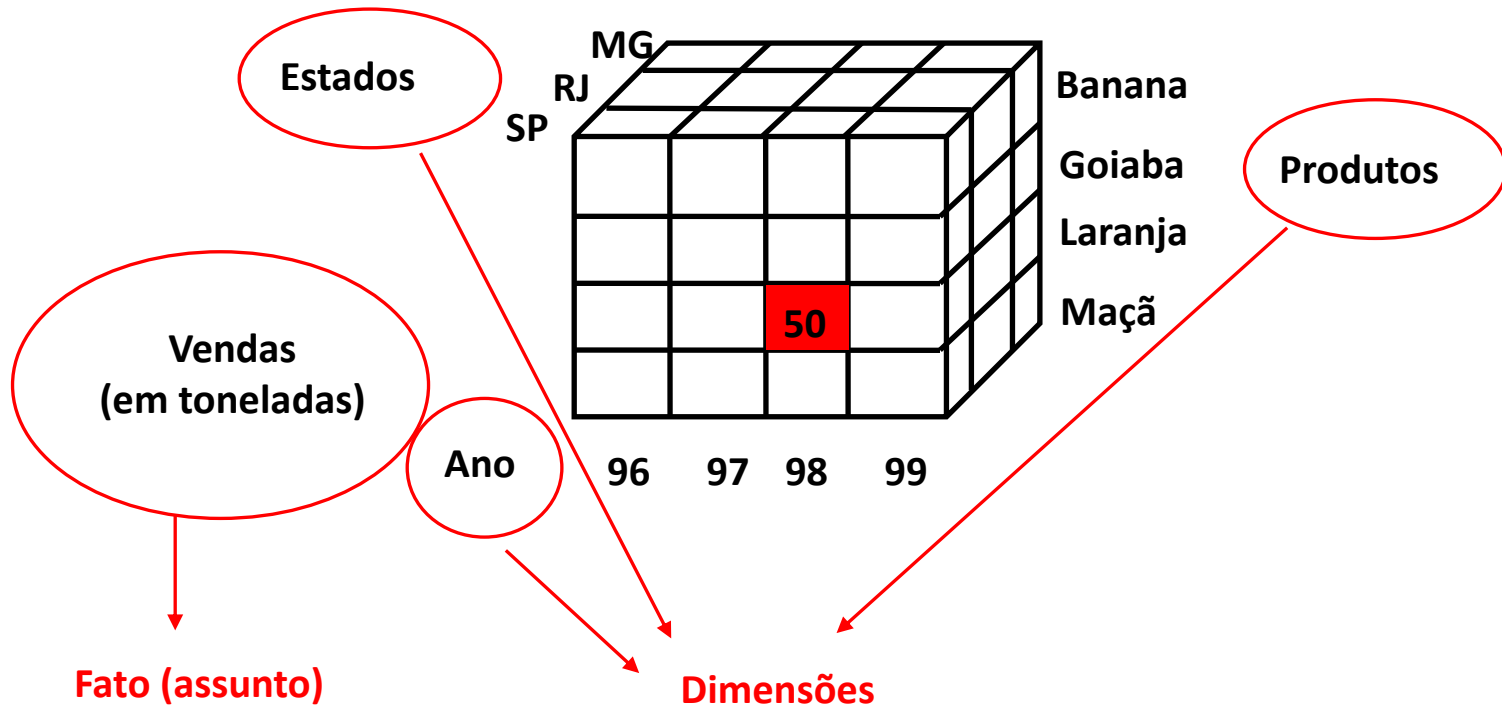
- Star Schema (Esquema Estrela)



# Modelagem Multidimensional



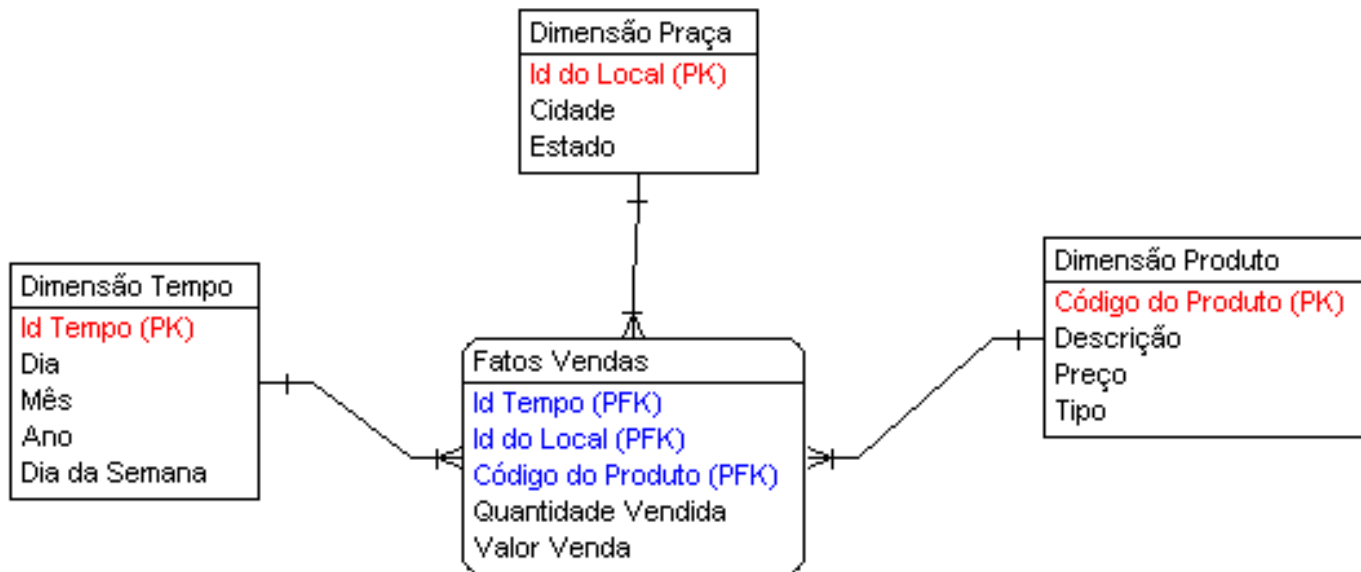
- Star Schema (Esquema Estrela)



# Modelagem Multidimensional



- Star Schema (Esquema Estrela)



# Modelagem Multidimensional



- Star Schema (Esquema Estrela)
  - Exemplo de tabela de dimensão resultante
- Produto

Id do Produto	Descrição	Preço	Tipo
101	Espaguete	10	Massa
102	Hamburguer	5	Carne
103	Talharim	15	Massa
104	Peito de Frango	20	Carne



# Modelagem Multidimensional



- Star Schema (Esquema Estrela)
  - Exemplo de tabela de fatos
    - Vendas

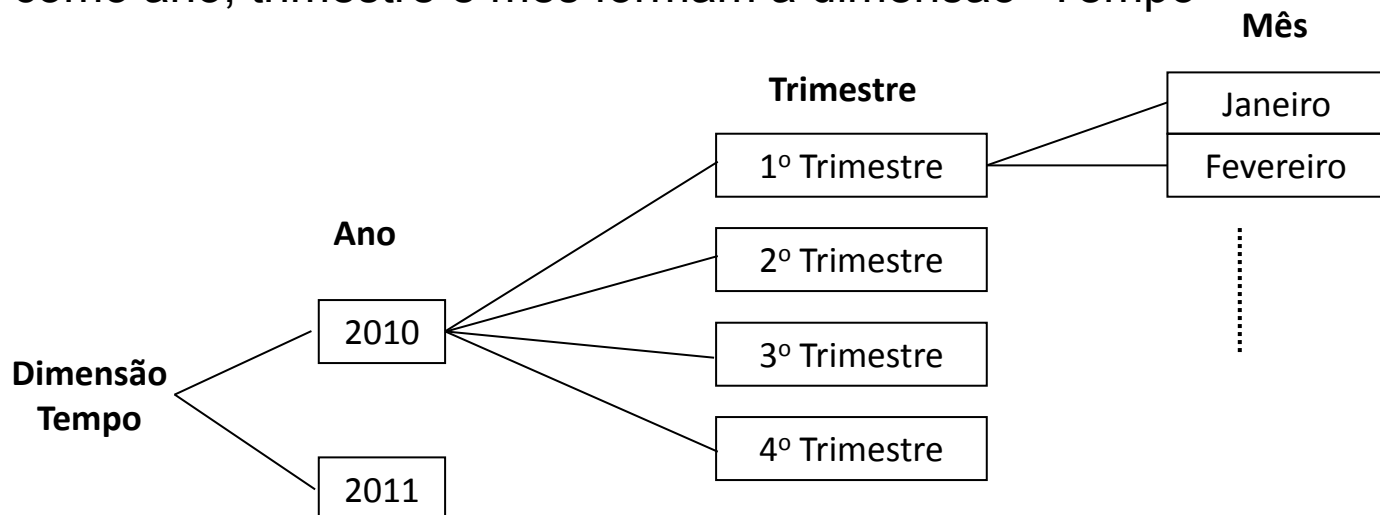
Id do Tempo	Id do Produto	Id do Funcionário	Unidades Vendidas	Valor de Venda
031011	101	200	10	500
041011	101	200	13	650
051011	101	200	15	700
061011	101	200	20	1000

# Modelagem Multidimensional



- **Membros de uma dimensão**
  - São os elementos das dimensões
  - Hierarquia de dimensão

- Ex.: Cidade, estados e regiões formam a dimensão “Local”, assim como ano, trimestre e mês formam a dimensão “Tempo”



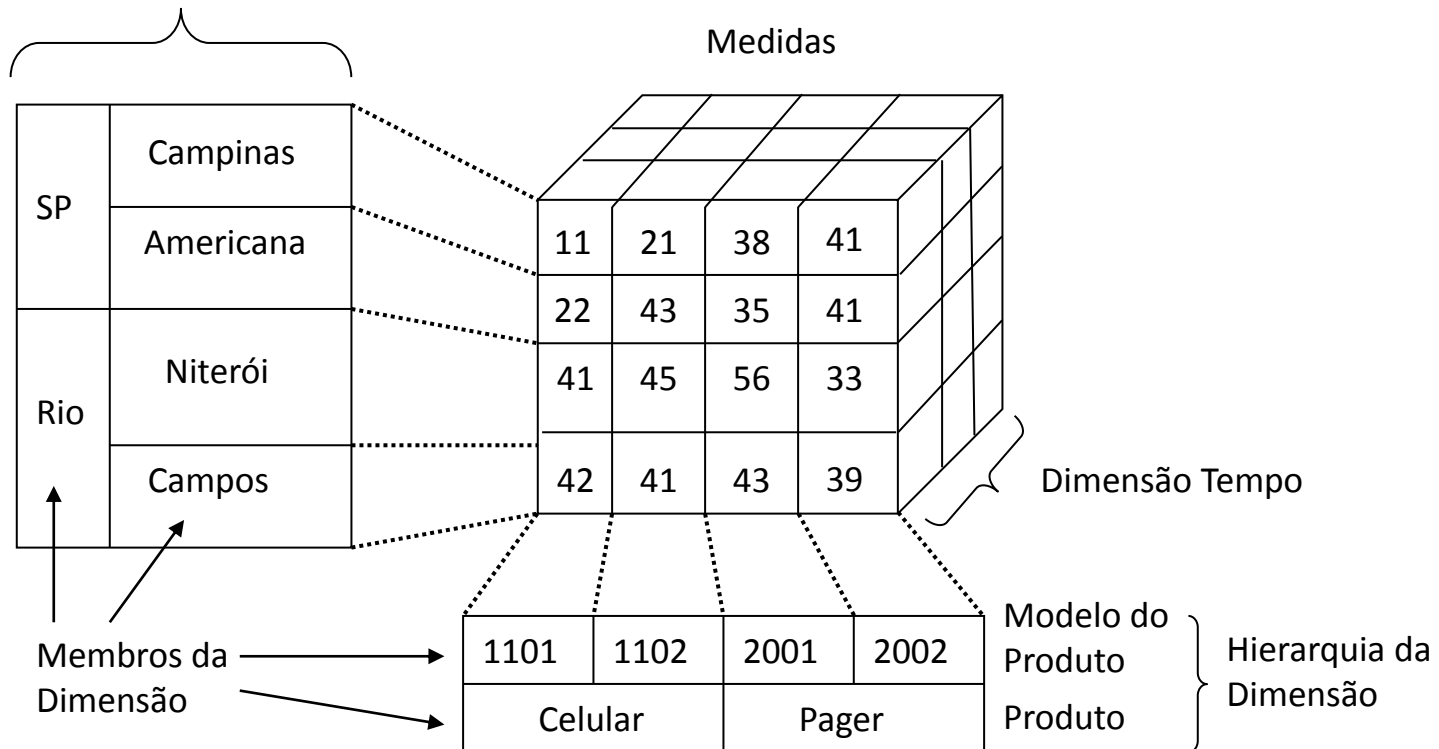
# Modelagem Multidimensional



- **Membros de uma dimensão**

Dimensão Localização  
Hierarquia da Dimensão

## Visão Multidimensional



# Modelagem Multidimensional



- As dimensões representam entidades que evoluem com o tempo..
  - Por exemplo, um cliente pode deixar de ser solteiro e casar-se
  - Para tratar essas atualizações, pode-se tratar as dimensões de três formas diferentes
    - De acordo com a importância de se ter informações históricas!

# Modelagem Multidimensional



- Dimensão Tipo 1
  - O histórico não é relevante!
  - As alterações podem ser feitas diretamente no registro em questão sem salvar o valor anterior
  - Ex.: Godofredo tinha seu estado civil solteiro até 02/07/2013  
Godofredo casou-se dia 02/07/2013
  - Godofredo teria seu estado civil atualizado para casado



# Modelagem Multidimensional



- Dimensão Tipo 1

Id do Cliente	Nome	Estado_Civil
101	Godofredo	Solteiro

Id do Cliente	Nome	Estado_Civil
101	Godofredo	Casado

# Modelagem Multidimensional



- Dimensão Tipo 2
  - O histórico é relevante!
  - Inserção de um novo registro na mesma entidade dimensional refletindo a mudança
  - Ex.: Existirão dois registros do Godofredo, o 1º referente a seu estado civil até 02/07/13 e o outro após essa data como casado
  - Na tabela de fatos vendas, o primeiro registro de Godo está vinculado às vendas anteriores a 02/07/13 e o outro às vendas posteriores

# Modelagem Multidimensional



- Dimensão Tipo 2

Id do Cliente	Nome	Estado_Civil	Status
101	Godofredo	Solteiro	Antigo
101	Godofredo	Casado	Atual



# Modelagem Multidimensional



- Dimensão Tipo 3

- O histórico é relevante e deseja-se analisar dados usando os status original e atual
- São necessários campos para armazenar
  - Status original do atributo
  - Status atual do atributo
  - Data efetiva da última alteração do campo
- Apenas dois status podem ser rastreados: o atual e o original!

# Modelagem Multidimensional



- Dimensão Tipo 3

Id do Cliente	Nome	Estado_Civil_Original	Estado_Civil_Atual	Data_Efetiva
101	Godofredo	Solteiro	Casado	02/07/2013

- E se a esposa de Godo largá-lo???

Id do Cliente	Nome	Estado_Civil_Original	Estado_Civil_Atual	Data_Efetiva
101	Godofredo	Solteiro	Divorciado	04/07/2013

- Não terá como fazer análise sobre informações de quando Godo era casado

# Modelagem Multidimensional

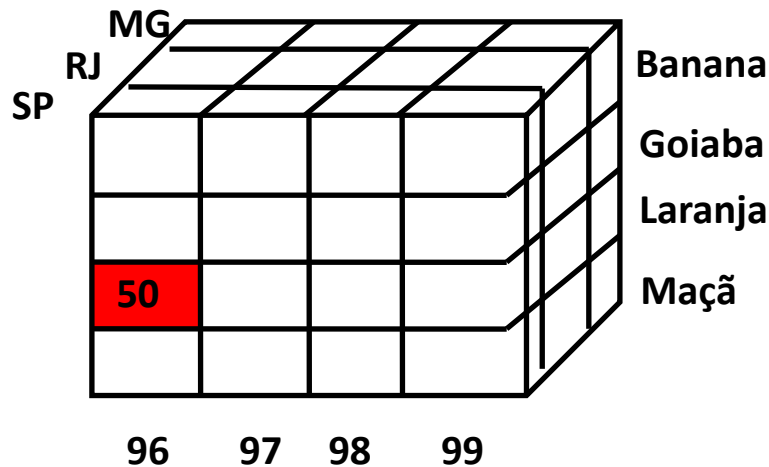
## *Conceitos Importantes*



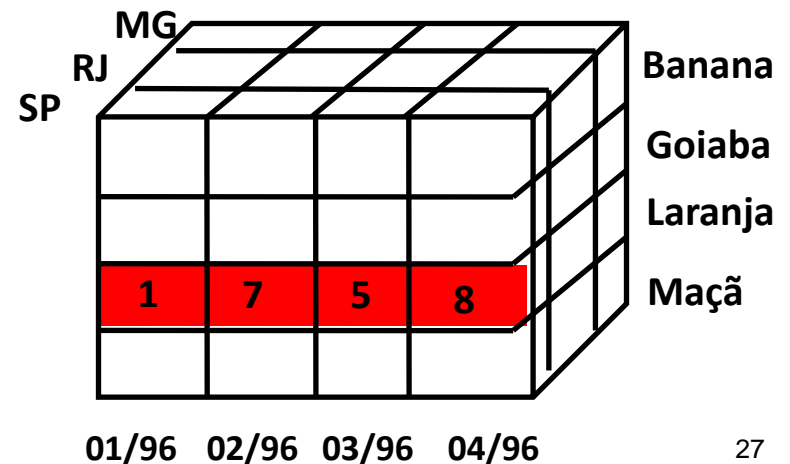
- Granularidade

- É o nível de detalhe das tabelas
- Quanto menor o “grão” maior a granularidade

MENOR Granularidade



MAIOR Granularidade

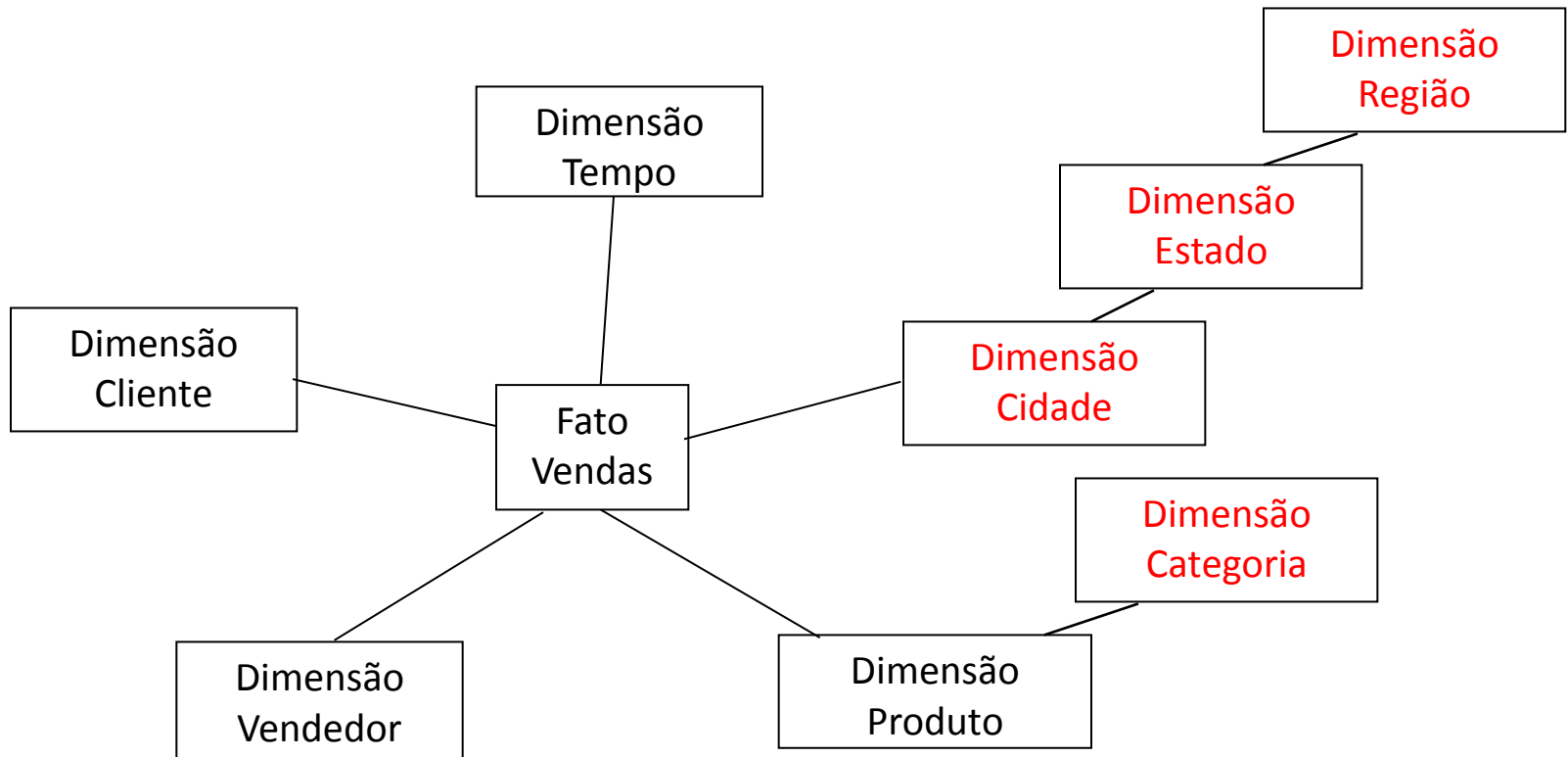


# Modelagem Multidimensional

## *Conceitos Importantes*



- Esquema Snowflake (Bloco de neve)



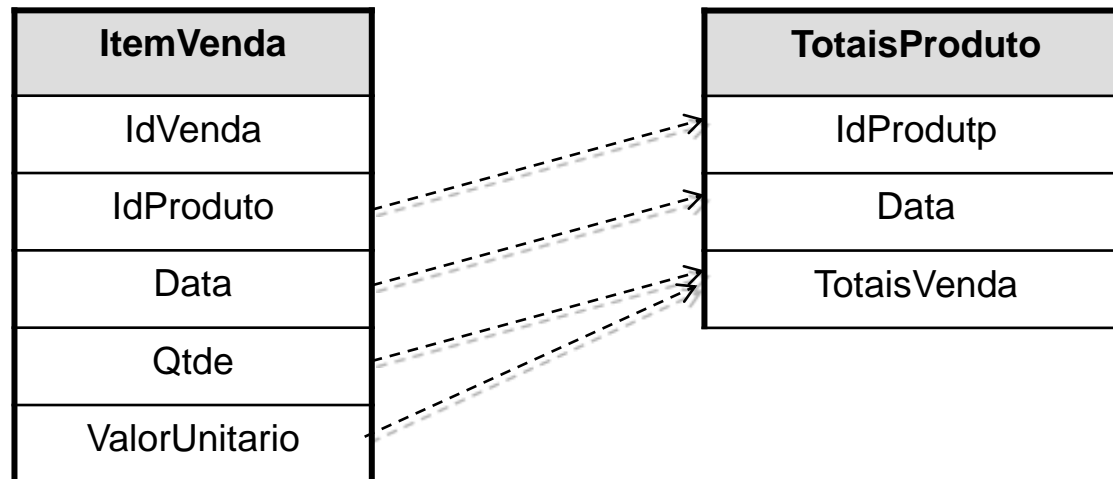
# Modelagem Multidimensional

## *Conceitos Importantes*



- **Agregação**

- Através da agregação cria-se novas entidades contendo dados sumarizados



# Modelagem Multidimensional

## *Procedimento Geral*



- Quatro passos básicos
  - 1º Definir o FATO de negócio
  - 2º Definir a GRANULARIDADE utilizada
  - 3º Definir as DIMENSÕES do fato
  - 4º Definir as MEDIDAS do fato

# Modelagem Multidimensional

## *Procedimento Geral*



- Exemplo

- Uma rede de restaurantes tem 50 filiais localizadas em vários estados da federação. Cada filial oferece mais de 1000 **produtos** diferentes nas categorias bebidas e pratos.
- A diretoria da empresa deseja analisar as **vendas**, os **custos** e os **lucros** obtidos bem como os **funcionários** mais ativos.
- **Promoções** e festivais são utilizados para atrair clientes e potencializar as vendas.

# Modelagem Multidimensional

## *Procedimento Geral*



- Exemplo

- A diretoria da empresa determinou que é estratégico para a tomada de decisões analisar o **movimento diário de cada produto**, para que possa direcionar as promoções ou festivais de acordo com os resultados das análises realizadas
- Avaliar o movimento diário de cada produto consiste em analisar as **vendas** de produtos, levando em conta os **preços praticados** e as **filiais** que realizaram tais vendas



# Modelagem Multidimensional

## *Procedimento Geral*



- Exemplo
  - 1º Definir o FATO
  - Qual elemento central a empresa deseja analisar???
  - R.: **Vendas**

# Modelagem Multidimensional

## *Procedimento Geral*



- Exemplo
  - 2º Definir a GRANULARIDADE
  - Em que nível de detalhe a empresa deseja analisar???
  - “é estratégico para a tomada de decisões analisar o movimento diário de cada produto”
- R.: **Diário**
  - (com respeito a outras dimensões não foi especificado)

# Modelagem Multidimensional

## *Procedimento Geral*



- Exemplo

- 3º Definir as DIMENSÕES

- Quais aspectos são relevantes para se realizar as análises que a empresa solicita do fato???

- De forma geral, alguns fatores a se observar são

- O quê → Produto
    - Quem → Funcionário
    - Quando → Tempo
    - Onde → Local (filial)

- Pode-se ainda levar em conta outros objetivos especificados para a análise dos fatos, como Promoções

# Modelagem Multidimensional

## *Procedimento Geral*



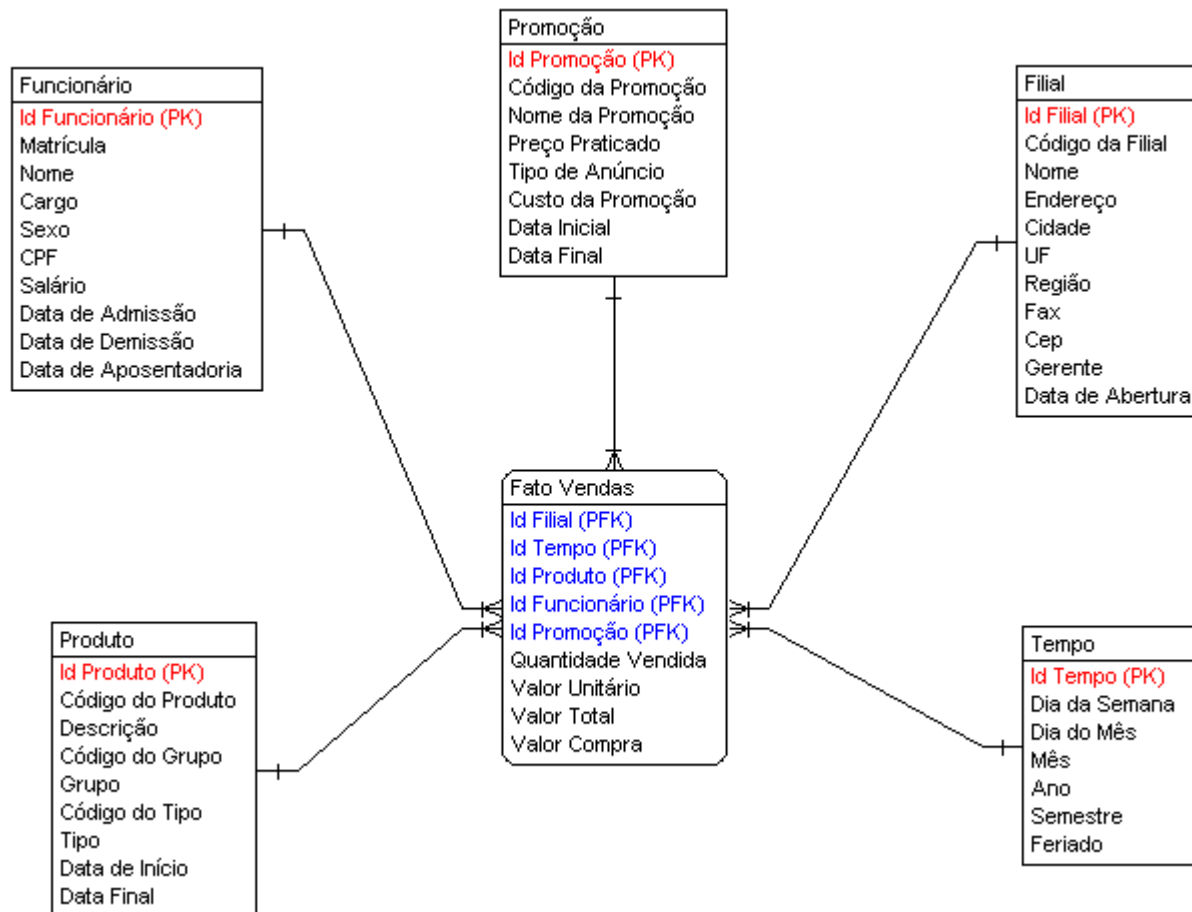
- Exemplo
  - 4º Definir as MEDIDAS
  - Como o desempenho de vendas pode ser medido???
  - R.: Quantidade vendida, valor unitário e total da venda, valor da compra

# Modelagem Multidimensional

## *Procedimento Geral*



- Exemplo – Modelo estrela correspondente





- *Tecnologia e Projeto de Data Warehouse.* Machado, F.N.R.. São Paulo, Erica, 2010.

