

Processo de Criação de um Esquema Estrela

Processamento Analítico de Dados
Profa. Dra. Cristina Dutra de Aguiar Ciferri
Prof. Dr. Ricardo Rodrigues Ciferri

Etapas

1. Selecione o **processo do negócio**

2. Identifique o **grão** do processo do negócio

- exatamente *o que* cada tupla da tabela de fatos significa

3. Escolha as **dimensões** que aplicam a cada linha da tabela fato

4. Identifique os **fatos** que preencherão cada tupla da tabela de fatos

Exemplo: Vendas a Varejo

- Empresa do ramo de alimentos
 - 100 supermercados em 5 estados
- Cada filial
 - possui diversos departamentos
 - contém \approx 60.000 produtos individuais, identificados por SKU
 - 55.000 produtos de fornecedores externos, identificados por códigos de barra (UPC)
 - 5.000 produtos produzidos internamente, identificados por SKU

Exemplo: Vendas a Varejo

- Coleta de dados
 - ponto de venda do supermercado
 - ex: caixa
 - entrada de mercadorias
- Logística
 - compra, armazenamento e venda de produtos, visando maximizar os lucros
 - decisões mais importantes relacionadas a preços e promoções

Projeto: Vendas a Varejo

- Processo do negócio
 - ponto de venda do supermercado
 - quais produtos são vendidos em quais filiais em quais dias sob quais promoções
- Grão
 - item individual de venda
 - um DW quase sempre requer dados em seu nível de granularidade mais detalhado

Projeto: Vendas a Varejo

dimensões



data,
produto,
filial,
promoção,
número da transação
de venda

fatos



quantidade vendida,
preço unitário,
preço total,
lucro total sem deduções

• preço total =
quantidade vendida *
preço unitário

Esquema Estrela

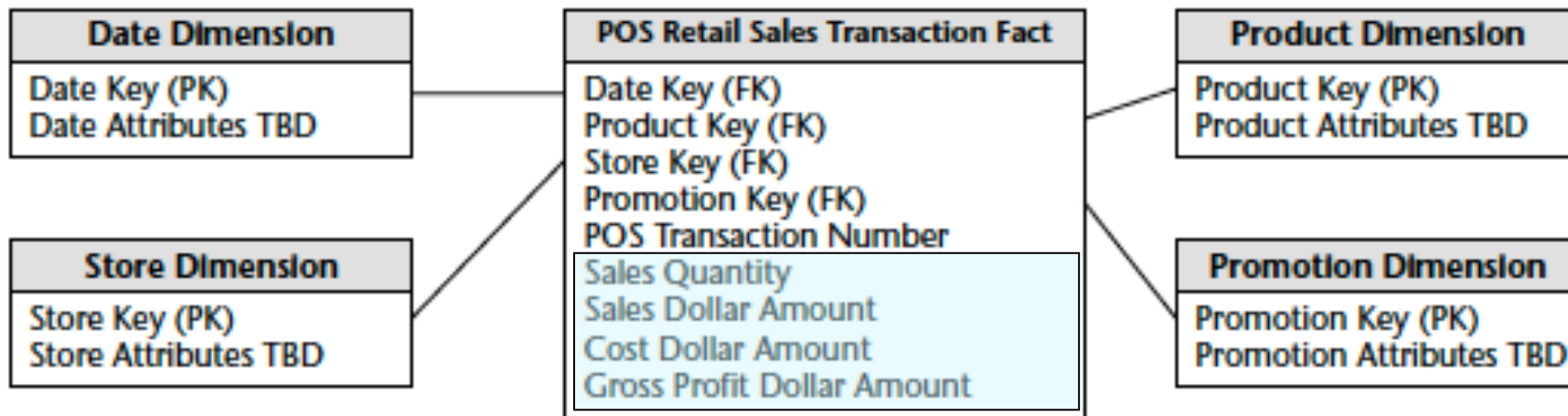


Tabela de Dimensão Data

- Característica
 - representa o aspecto temporal
 - pode ser povoada antes das demais tabelas
- Granularidade
 - diária

Date Dimension
Date Key (PK)
Date
Full Date Description
Day of Week
Day Number in Epoch
Week Number in Epoch
Month Number in Epoch
Day Number in Calendar Month
Day Number in Calendar Year
Day Number in Fiscal Month
Day Number in Fiscal Year
Last Day in Week Indicator
Last Day in Month Indicator
Calendar Week Ending Date
Calendar Week Number in Year
Calendar Month Name
Calendar Month Number in Year
Calendar Year-Month (YYYY-MM)
Calendar Quarter
Calendar Year-Quarter
Calendar Half Year
Calendar Year
Fiscal Week
Fiscal Week Number in Year
Fiscal Month
Fiscal Month Number in Year
Fiscal Year-Month
Fiscal Quarter
Fiscal Year-Quarter
Fiscal Half Year
Fiscal Year
Holiday Indicator
Weekday Indicator
Selling Season
Major Event
SQL Date Stamp
... and more

Tabela de Dimensão Produto

- Característica
 - povoada praticamente com dados oriundos do ambiente operacional
- Granularidade
 - cada SKY

Product Dimension
Product Key (PK)
Product Description
SKU Number (Natural Key)
Brand Description
Category Description
Department Description
Package Type Description
Package Size
Fat Content
Diet Type
Weight
Weight Units of Measure
Storage Type
Shelf Life Type
Shelf Width
Shelf Height
Shelf Depth
... and more

Tabela de Dimensão

Filial

- Característica
 - similar a uma dimensão geográfica
- Granularidade
 - cada filial

Store Dimension
Store Key (PK)
Store Name
Store Number (Natural Key)
Store Street Address
Store City
Store County
Store State
Store Zip Code
Store Manager
Store District
Store Region
Floor Plan Type
Photo Processing Type
Financial Service Type
Selling Square Footage
Total Square Footage
First Open Date
Last Remodel Date
... and more

Tabela de Dimensão

Promoção

- Granularidade
 - cada combinação de condições de promoção
- Discussão
 - armazenamento em uma ou mais tabelas
 - tupla adicional: *sem promoção*

Promotion Dimension
Promotion Key (PK)
Promotion Name
Price Reduction Type
Promotion Media Type
Ad Type
Display Type
Coupon Type
Ad Media Name
Display Provider
Promotion Cost
Promotion Begin Date
Promotion End Date
... and more

Dimensão Degenerada

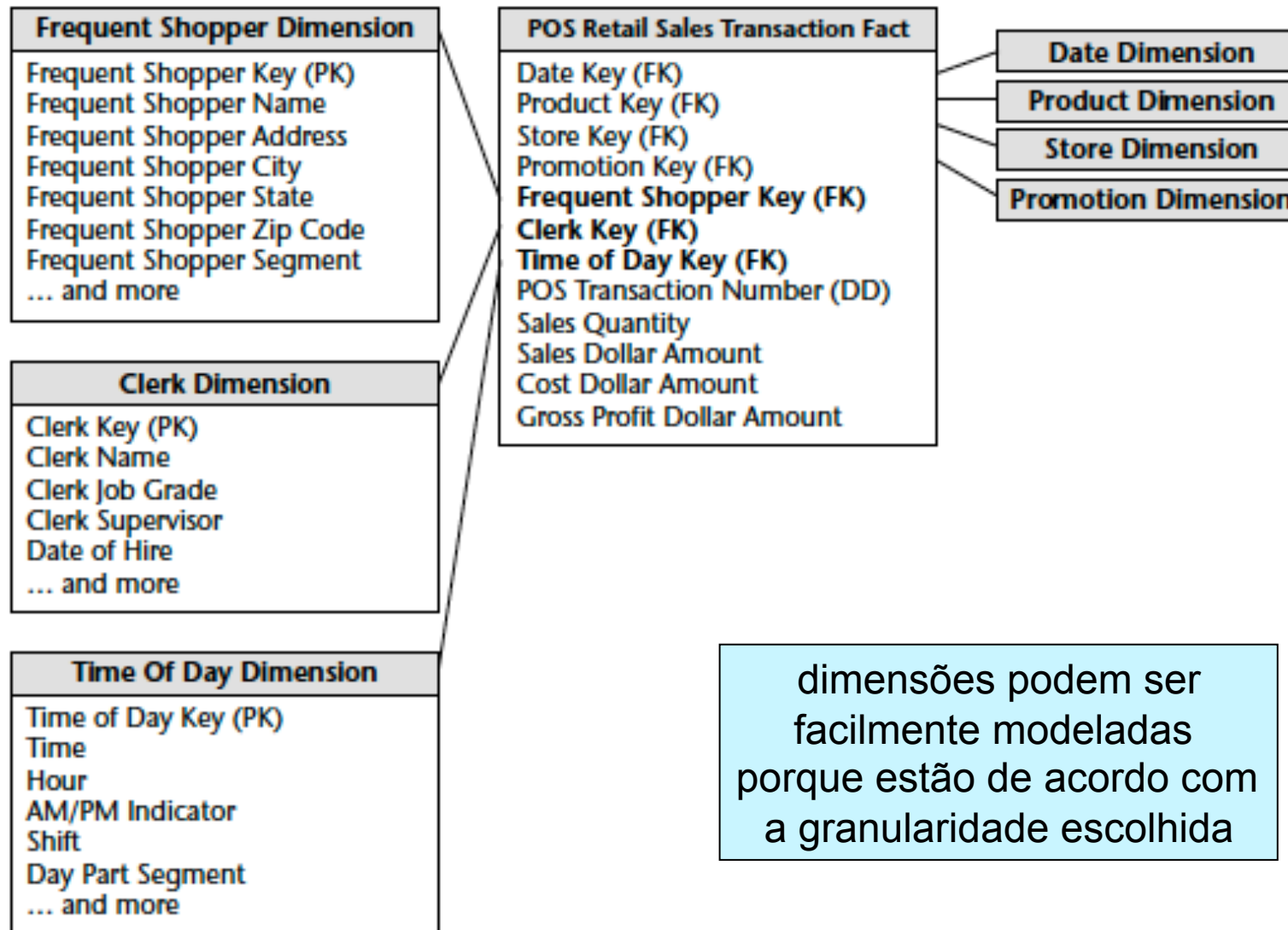
Número da Transação

- Utilidade
 - permite agrupar todos os produtos vendidos em uma transação
- Característica
 - não possui atributos adicionais
 - muito comum quando a granularidade da tabela de fatos representa cada item da transação

Projeto Estendido: Varejo

- Requisitos adicionais
 - cliente que efetuou a compra
 - período do dia no qual a compra foi feita
 - vendedor que efetuou a venda
- Para cada requisito adicional
 - nova tabela de dimensão
 - nova chave estrangeira na tabela de fatos
 - tupla artificial (quando necessário) para identificar compras realizadas anteriormente

Esquema Estrela Estendido



dimensões podem ser facilmente modeladas porque estão de acordo com a granularidade escolhida

Projeto Estendido

- Novos atributos de dimensão
 - adicionar novas colunas à tabela de dimensão
 - incluir uma tupla artificial “não disponível” para representar situações nas quais os novos atributos são válidos somente a partir de um dado tempo

Projeto Estendido

- Novas dimensões
 - adicionar novas tabelas de fatos
 - incluir chaves estrangeiras apropriadas
 - incluir tuplas artificiais
- Novos fatos
 - adicionar novas colunas à tabela de fatos
 - preencher os campos com valores apropriados
 - **analisar a granularidade!**