



# Modelo de Dados e Metadados

## Técnicas de Modelagem

Profa. Kelli de Faria Cordeiro  
kelli.faria@gmail.com



Foreign Keys, Primary Keys, and Surrogate Keys

## CHAVES ESTRANGEIRA, PRIMÁRIA E SURROGADA

30/08/2009

Business Intelligence



## Chaves

- Todas as tabelas dimensão possuem uma chave primária que identifica unicamente um registro.
- As chaves em um DW não devem ter significado.
  - As chaves originais não devem ser usadas.

30/08/2009

Business Intelligence



## Chaves Surrogadas

- Vamos assumir que uma chave surrogada é um campo do tipo inteiro com valores seqüenciais.
  - São usadas sob demanda.
- Essas chaves por si só não têm significado.
  - Consultas apenas na chave não dizem nada sobre o registro.
  - É necessário usar a chave para recuperar o registro para ver o que ele contém.

30/08/2009

Business Intelligence



## Chaves Surrogadas

- Na maior parte dos casos, um inteiro de 4 bytes é o suficiente.
  - $2^{32}$  valores
  - Mais de 2 bilhões de inteiros positivos

30/08/2009

Business Intelligence



## Chaves Surrogadas

- Identificação de uma data
  - É mandatório o uso de chaves surrogadas para identificar uma data.
  - Uma chave data, tipicamente, ocupa 8 bytes
    - 4 bytes seriam desperdiçados em cada registro da tabela fato e dimensão!
  - Exceto:
    - Consultas sobre a data executadas diretamente na tabela Fato, sem *join* com a dimensão Tempo.

30/08/2009

Business Intelligence



## Chaves Surrogadas

- Desvantagens do uso de chaves *datestamp*:
  - Impossibilita codificações do tipo:
    - “Data a ser determinada”
    - “Data não aplicável”
  - Consultas apenas pela hora necessita da criação de uma outra chave do tipo *timestamp*
    - Uma chave *datetimestamp* resolveria este problema.
- Considerar a utilização de chave surrogada para resolver essas questões.

30/08/2009

Business Intelligence



## Chaves Surrogadas

- Por que devemos evitar chaves do ambiente de produção?
  - 1 – A produção pode decidir reusar a chave.
    - Exemplo: na virada de ano, alguns registros podem ser retirados do banco de produção.
      - Um DW mantém dados históricos por anos!
  - 2 – Empresas podem se fundir.
    - Regras de construção de chaves podem ser modificadas completamente.
    - Conseqüência: chaves compostas de vários dígitos alfanuméricos.

30/08/2009

Business Intelligence



## Chaves Surrogadas

- Por que devemos evitar chaves do ambiente de produção?
  - 3 – As chaves da produção podem ser deliberadamente alteradas.
    - Impedirá a rastreabilidade

30/08/2009

Business Intelligence



## Chaves Surrogadas

- Usadas para registrar situações que não têm um código operacional:
  - “Nenhuma promoção em vigor”

30/08/2009

Business Intelligence

## Chaves Surrogadas

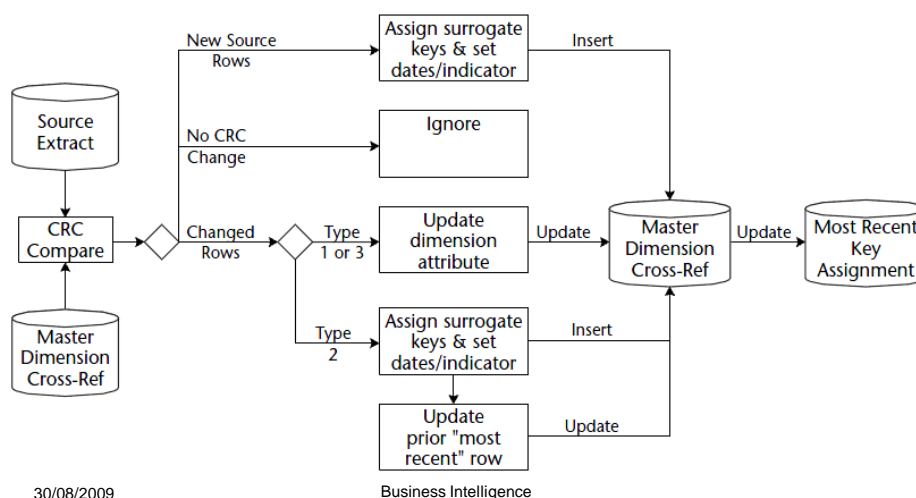
### ■ Outros nomes:

- ☐ Chaves sem significado
- ☐ Chaves de inteiros
- ☐ Chaves substitutas
- ☐ Chaves artificiais
- ☐ Chaves sintéticas
- ☐ Chaves não-naturais

30/08/2009

Business Intelligence

## Gerenciamento de Chaves Surrogadas



30/08/2009

Business Intelligence



Additive, Semiadditive, and Nonadditive Facts

## **FATOS ADITIVOS, SEMI-ADITIVOS E NÃO-ADITIVOS**

30/08/2009

Business Intelligence



### **Fatos Aditivos**

- Os fatos que mais fazem sentido para serem analisados são os perfeitamente aditivos.
  - Exemplo: valor das vendas em dólares e unidades.
- Entretanto, medidas numéricas de intensidade, não são perfeitamente aditivos.

30/08/2009

Business Intelligence



## Fatos Semi-Aditivos

- Exemplos de medidas de intensidade:
  - Níveis de inventário
  - Balanço Contábil
- Diferentemente de fatos aditivos, os fatos semi-aditivos não representam um fluxo temporal.
- Medidas de intensidade são normalmente aditivos ao longo de quase todas as dimensões, exceto a dimensão Tempo.

30/08/2009

Business Intelligence




## Fatos Semi-Aditivos

- Medida de Balanço Contábil e Níveis de Inventário devem ser consultadas ao longo do tempo através de uma média temporal.
  - Somar tudo e dividir pelo total de períodos de tempo
- Medidas de temperatura devem ser consultadas através de uma média (Função SQL AVG).

30/08/2009

Business Intelligence






## Fatos Não-Aditivos

- Fatos Textuais

30/08/2009 Business Intelligence



Factless Fact Tables

# TABELAS FATO SEM FATOS

30/08/2009 Business Intelligence



## Tabelas de Fatos Sem Fatos

- Até agora vimos tabelas Fato com as mesmas características:
  - ☐ Várias colunas chave
  - ☐ Várias medidas aditivas
- Porém, existem alguns fatos que não possuem medidas!

30/08/2009

Business Intelligence



## Tabelas Fato Sem Fatos

- Úteis para descrever eventos e cobertura.
- Podemos responder a seguinte pergunta:
  - ☐ O que não aconteceu?

30/08/2009

Business Intelligence



## Tabelas Fato Sem Fatos

- Representa um conjunto das relações muitos-para-muitos entre as tabelas dimensão.
- Registra a colisão das dimensões em um ponto no tempo ou espaço.
- A peculiaridade desses exemplos é que não temos um fato numérico, e sim um contador.

30/08/2009

Business Intelligence



## Tabelas Fato Sem Fatos

- Eventos são modelados como tabelas Fato contendo uma série de chaves, cada uma representando uma dimensão participante do evento.
- Tabelas Evento não têm nenhum fato numérico associado, por isso é chamado de tabela de fatos sem fatos.

30/08/2009

Business Intelligence

## Tabelas Fato Sem Fatos

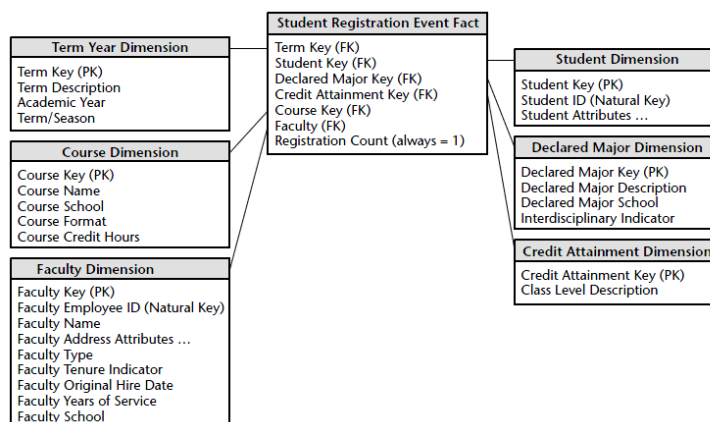
- Existem muitas situações em que os eventos precisam ser registrados.
- Exemplo:
  - Rastreamento da matrícula de estudantes por período letivo.
  - O grão da tabela fato é cada matrícula de um estudante em um período.
    - A dimensão tempo é o período letivo.

30/08/2009

Business Intelligence

## Tabelas Fato Sem Fatos

- Exemplo de Tabela Fato de Eventos



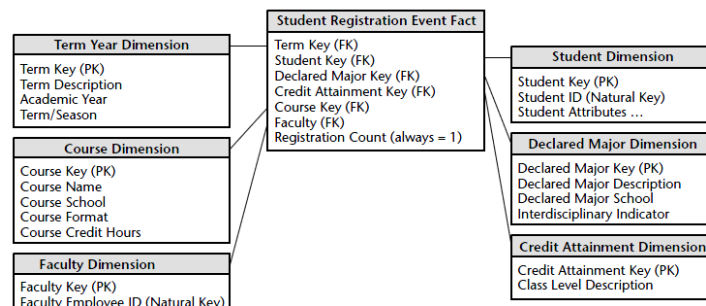
30/08/2009

Business Intelligence

## Tabelas Fato Sem Fatos

### ■ Exemplos de perguntas:

- ☐ Quais estudantes estão matriculados em que cursos?
- ☐ Quantos estudantes de matricularam em determinados cursos nos últimos 3 anos?
- ☐ Quantos estudantes de matricular em mais de um curso?

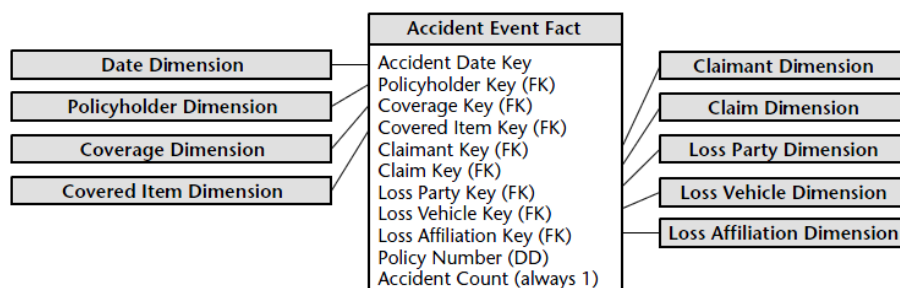


30/08/2009

## Tabelas Fato Sem Fatos

### ■ Exemplo de Tabela Fato de Acidentes de Veículos

- ☐ Quais perguntas ele pode responder?



30/08/2009

Business Intelligence



## Tabelas Fato Sem Fatos

### ■ Tabela Fato de Cobertura

- Considere um sistema que registra a utilização das instalações de uma Universidade.
- Exemplos de perguntas que podem ser respondidas com uma tabela de cobertura
  - Quais instalações são mais usadas?
  - Qual foi a utilização média das instalações em determinado período do dia?
  - O uso das instalações cai significativamente às sextas-feiras?

30/08/2009

Business Intelligence



## Tabelas Fato Sem Fatos

### ■ Tabela Fato

- Criar um registro na tabela fato para cada hora, independente da instalação estar sendo usado ou não.

### ■ Tabela Dimensão

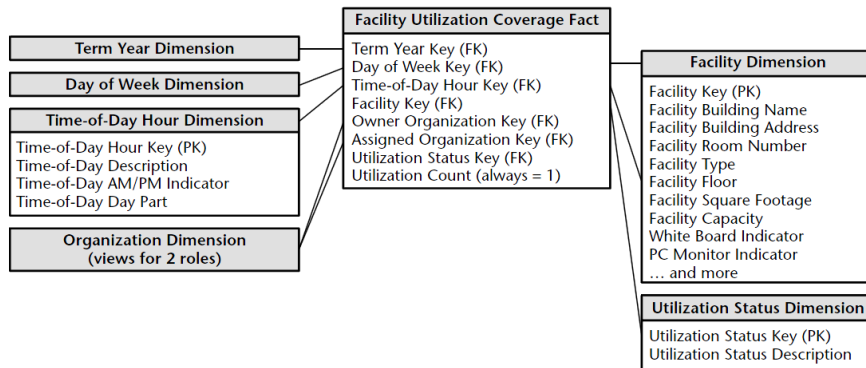
- Criar uma dimensão com a situação da instalação com um descritor com os seguintes valores:
  - "Disponível"
  - "Utilizado"

30/08/2009

Business Intelligence

## Tabelas Fato Sem Fatos

### ■ Exemplo de Tabela Fato de Cobertura da Utilização de Instalações



30/08/2009

Business Intelligence

## Tabelas Fato Sem Fatos

### ■ Exercício 5:

- Considere um sistema de registro da frequência dos alunos em cursos de uma universidade.
- Elaborar um modelo multidimensional para responder as seguintes questões analíticas:
  - Quais cursos são mais freqüentados?
  - Quais cursos sofreram a menor perda de freqüência de comparecimento durante o ano letivo?
  - Quais alunos freqüentaram quais cursos?
  - Qual membro do corpo docente teve mais alunos?

30/08/2009

Business Intelligence



## Tabelas Fato Sem Fatos

### ■ Exercício Extra:

- Elaborar um modelo multidimensional para responder a seguinte questão analítica:
  - Quais produtos estavam em promoção e não foram vendidos?