



Sistemas de informação

enterprise analytics and data
warehousing

Profº Fabiano J. Cury Marques

enterprise analytics and data warehousing

Modelagem Dimensional Parte III



Agenda

- ✕ Surrogate Key
- ✕ Tabela fato sem fatos
- ✕ Referências



Surrogate Keys

Surrogate keys



Surrogate Keys

- ✗ Por vezes, projetistas relutam a desistir das **chaves naturais** pois querem navegar pela tabela fato baseado no código operacional evitando uma junção com a tabela dimensão
- ✗ Lembre-se que no modelo dimensional, **as dimensões são as portas de entrada para a fato e não o contrário**
- ✗ Um dos principais benefícios das **chaves artificiais** é que elas isolam o ambiente do DW das mudanças operacionais
- ✗ Outro benefício seria o desempenho. Chaves naturais normalmente são String. E lembre-se: em uma tabela fato com 1 bilhão de linhas, cada byte a mais em uma linha se transforma em 1 GB de armazenamento



Surrogate Keys

- ✗ Modelos dimensionais devem sempre usar **chaves artificiais** em vez de se basear em códigos operacionais
- ✗ Estas chaves são conhecidas como: **chave sem significado**, **chave de inteiros**, **chaves não-naturais**, **chaves sintéticas** etc.
- ✗ De forma simples, chaves artificiais são inteiros que são atribuídos sequencialmente quando necessário para popular uma dimensão
- ✗ Exemplo: para o primeiro registro de produto é atribuído uma SK de produto com o valor 1, para o próximo registro será atribuído uma SK com o valor 2 e assim por diante...
- ✗ Estas chaves apenas servem para fazer a **junção entre fatos e dimensões**



Surrogate Keys

- ✗ Além disso, chaves artificiais permitem trabalhar com **registros de dimensão que não têm códigos operacionais**. Exemplo:
 - Sem promoção
 - Data a ser determinada
 - Data não aplicável

- ✗ Finalmente, o uso de chaves artificiais é necessário para suportar uma das principais técnicas para lidar com **mudanças de atributos nas tabelas dimensão**

- ✗ Na verdade, esse é o principal motivo para o uso das SKs

Surrogate Keys



✗ Dilema de modelagem ...

**REGISTRO NA
DIMENSÃO DEVE
SER ALTERADO!!**

É NECESSÁRIO
TERMOS UMA
CHAVE ARTIFICIAL!!

2.

Tabela Fato sem Fatos

Tabela Fato sem Fatos

TABELA FATO SEM FATOS



- ✗ Ainda baseado no caso de vendas a varejo, existe uma pergunta que não conseguimos responder com aquele modelo
- ✗ Apesar de termos a dimensão Promoção (que indica que um produto foi vendido sob uma promoção específica) não conseguimos responder:
 - Quais produtos estavam em promoção mas não foram vendidos?
- ✗ A tabela fato só armazena SKUs realmente vendidos. Não existem linhas na tabela com fatos zero para os que não venderam pois isso tornaria nossa tabela ainda maior



RELEMBRANDO: Estudo de caso – Varejista

- ✗ Imagine-se trabalhando em uma grande cadeia varejista (supermercados). Nosso negócio tem **100 lojas** espalhadas em cinco estados.
- ✗ Cada loja tem diversos departamentos como: **doces, comidas congeladas, açougue, flores, limpeza doméstica** etc.
- ✗ Cada loja tem aproximadamente **60.000 produtos específicos (SKU)** para venda em suas prateleiras. Destes, 55.000 são de manufaturas externas e 5.000 são de departamentos como açougue, panificadora etc.
- ✗ Os dados são coletados em diversos lugares, mas o mais importante para o negócio são os dados de produtos comprados pelo cliente **coletado pelo scanner dos caixas (POS)**
- ✗ Para o lucro da empresa trabalha-se **baixando custos de aquisição e processamento** e ao mesmo tempo atraindo tantos clientes quanto possível em um ambiente de preços altamente competitivos
- ✗ Algumas das decisões gerenciais mais importantes levam em conta **preços e promoções**

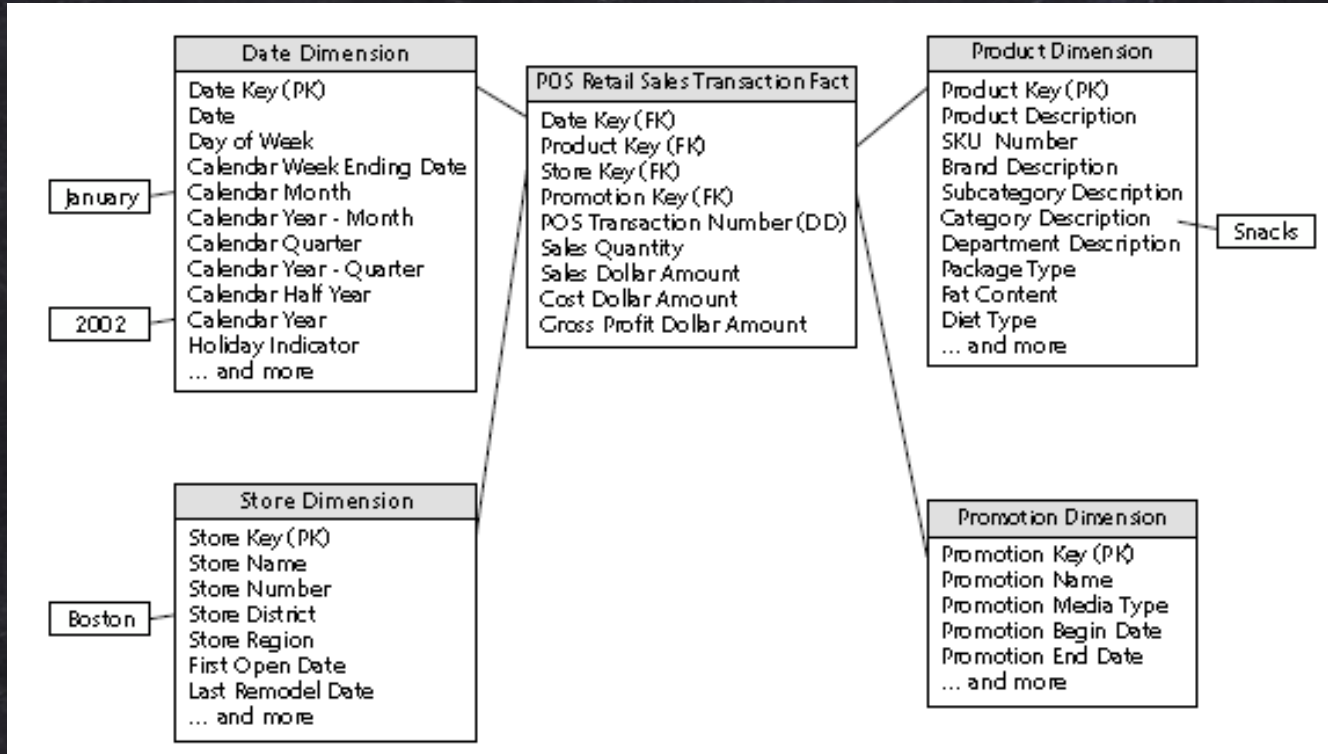


RELEMBRANDO: Estudo de caso – Varejista

1. Selecionando o **processo de negócio**:
Vendas a Varejo no POS
2. Definição do **Grão**:
item específico em uma transação do POS (uma compra no caixa)
3. Escolhendo as **Dimensões**:
data, produto, loja, promoção e Número de Transação POS
4. Identificando os **Fatos**:
quantidade vendida, total de venda, total de custo e lucro bruto



RELEMBRANDO: Estudo de caso – Varejista



Quais produtos estavam em promoção mas não foram vendidos?

TABELA FATO SEM FATOS



- ✗ Para podermos responder a questão acima, criamos uma **nova tabela fato** para a cobertura das promoções
- ✗ A tabela fato Cobertura de Promoção (*Promotion Coverage*) terá como chaves: **data**, **produto**, **loja** e **promoção**
- ✗ Nota-se que ela é muito similar com a fato de vendas entretanto o grão seria diferente (uma linha para cada produto em promoção em um dia em uma loja) não importa se o produto é vendido ou não
- ✗ A tabela fato de cobertura permite ver o relacionamento entre as chaves como definidas por uma promoção
- ✗ Como essa tabela apenas captura o relacionamento entre as chaves mas não tem nenhuma métrica de medição, chamamos ela de **Tabela Fato sem Fatos** (**Factless Table**)

TABELA FATO SEM FATOS



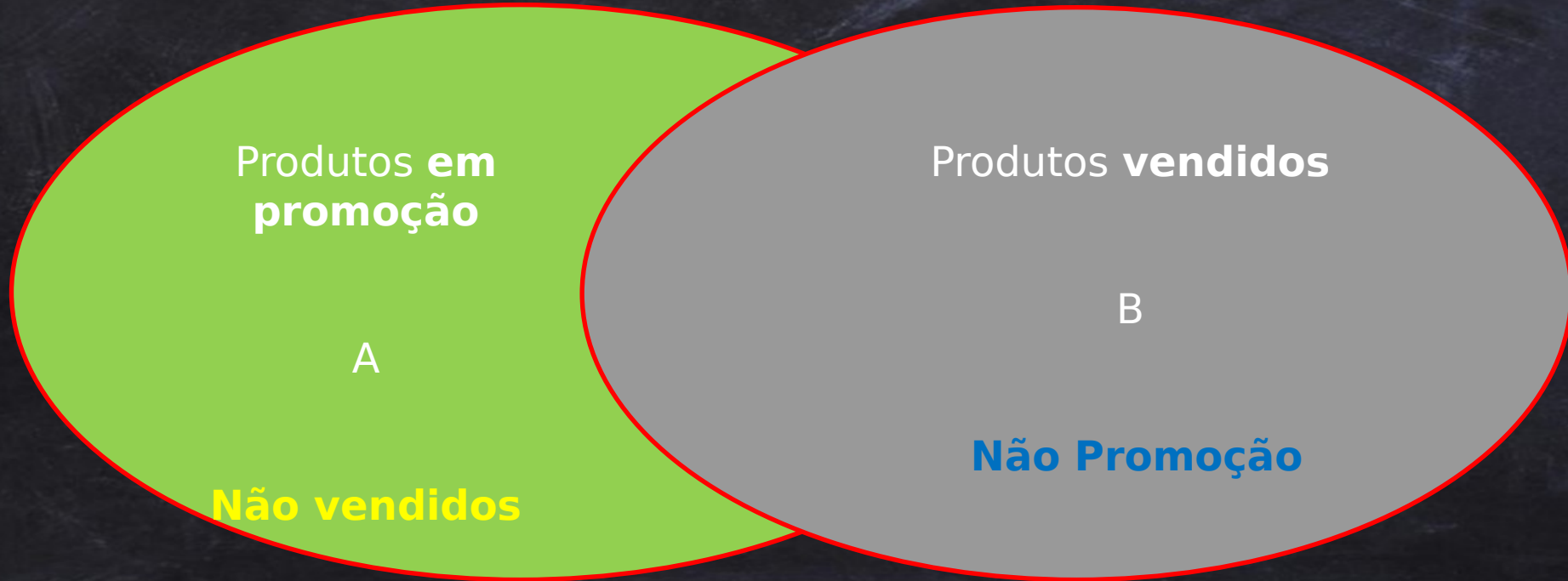
- ✗ Para responder a questão que nos levou a criar a tabela fato sem fatos devemos fazer um processo com duas etapas:
 - Primeiro fazer uma query para saber quais produtos estavam em promoção em determinado dia consultando a tabela **fato cobertura de promoção**
 - Depois, quais produtos foram vendidos a partir da tabela **fato de vendas**
- ✗ A resposta para a pergunta está na diferença entre as duas listas de produtos
- ✗ Em uma ferramenta **BI/OLAP** é muito simples obter respostas como essa





TABELA FATO SEM FATOS

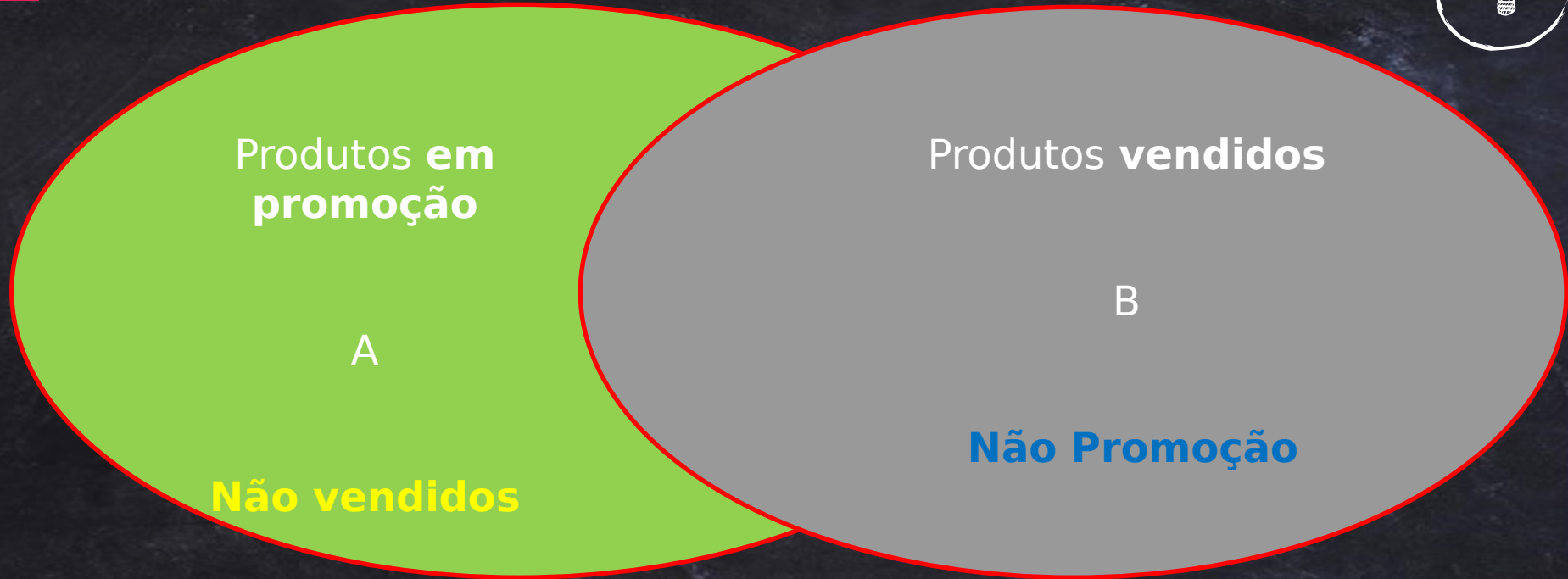
- ✗ Para responder a questão que nos levou a criar a tabela fato sem fatos devemos fazer um processo com duas etapas:



$$\text{Resposta} = A - B$$



TABELA FATO SEM FATOS

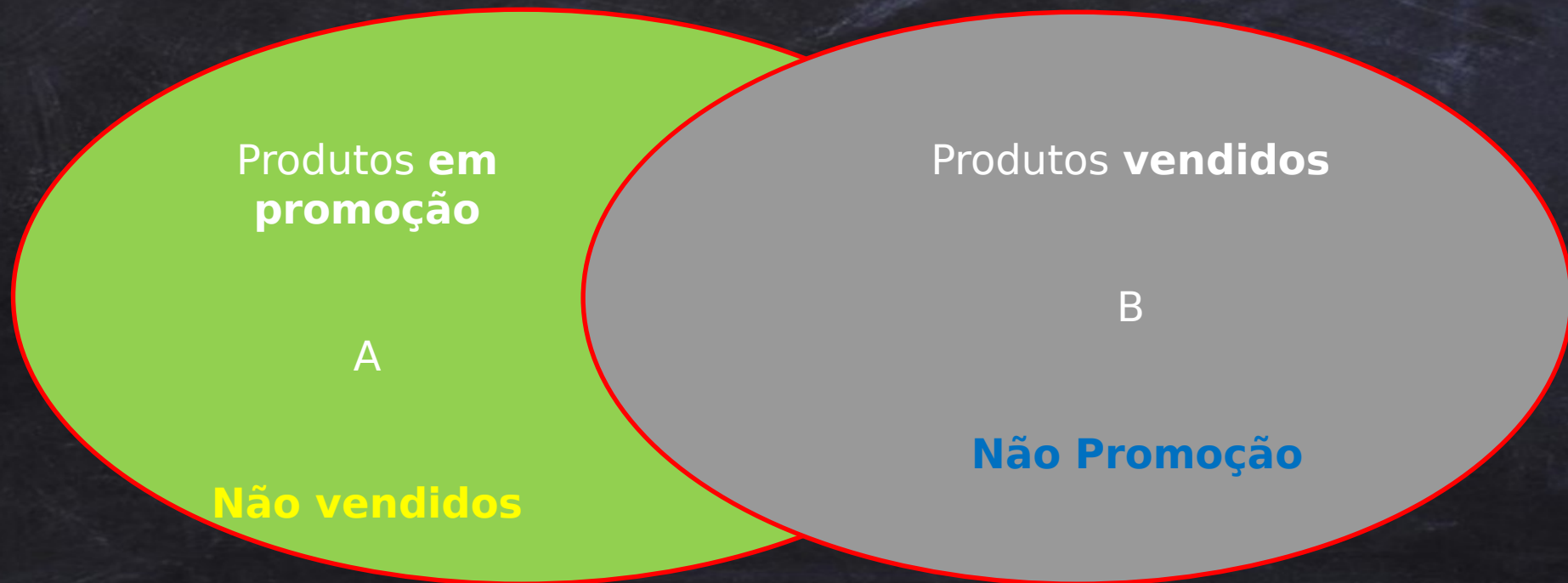


```
SELECT <select_list>  
FROM TableA A  
LEFT JOIN TableB B  
ON A.Key = B.Key  
WHERE B.Key IS NULL
```


TABELA FATO SEM FATOS



✗ Humm ...



NOT EXISTS vs NOT IN vs JOIN with NOT NULL ???



Referências

- ✕ KIMBALL, R., ROSS, M. The Data Warehouse Toolkit. 2ª ed., John Wiley Professional, 2002.
- ✕ MACHADO, F. N. R. Tecnologia e Projeto de Data Warehouse. 1ª ed., São Paulo: Ed. Érica, 2004.



Obrigado!

Copyright © 2019 Prof. MSc. Eng. Wakim B. Saba

<https://br.linkedin.com/in/wakimsaba>

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).