



Sistemas de informação

enterprise analytics and data
warehousing

Profº Fabiano J. Cury Marques

enterprise analytics and data warehousing

Modelagem Dimensional Parte I



Agenda

- ✕ Onde estamos?
- ✕ Modelagem Dimensional
- ✕ Fatos
- ✕ Dimensões
- ✕ Referências

1.

Em que ponto paramos?

Revisão rápida

O que vimos até agora?



- ✗ Processo decisório;
- ✗ Hierarquia da Informação;
- ✗ DW : Orientado a assunto, Integrado, Série Temporal e Não Volátil;
- ✗ Sistemas Transacionais e Dimensionais;
- ✗ Conceitos de: Data Mart e Data Warehouse;
- ✗ Tendência: Pesquisa, dados estruturados, SOA e DW em tempo real;

O que vimos até agora?



- ✖ Arquitetura de Fluxo de dados :
 - Formato de dados: Stage, DDS, NDS e ODS;
 - Acessibilidade: User-facing, Interna e Híbrida;

- ✖ Base de dados: Relacional, Normalizada, desnormalizada e dimensional;

- ✖ ETL, MetaData, Data Quality

- ✖ Arquitetura do Sistemas:
 - On premise
 - Cloud computing

O que vimos até agora?

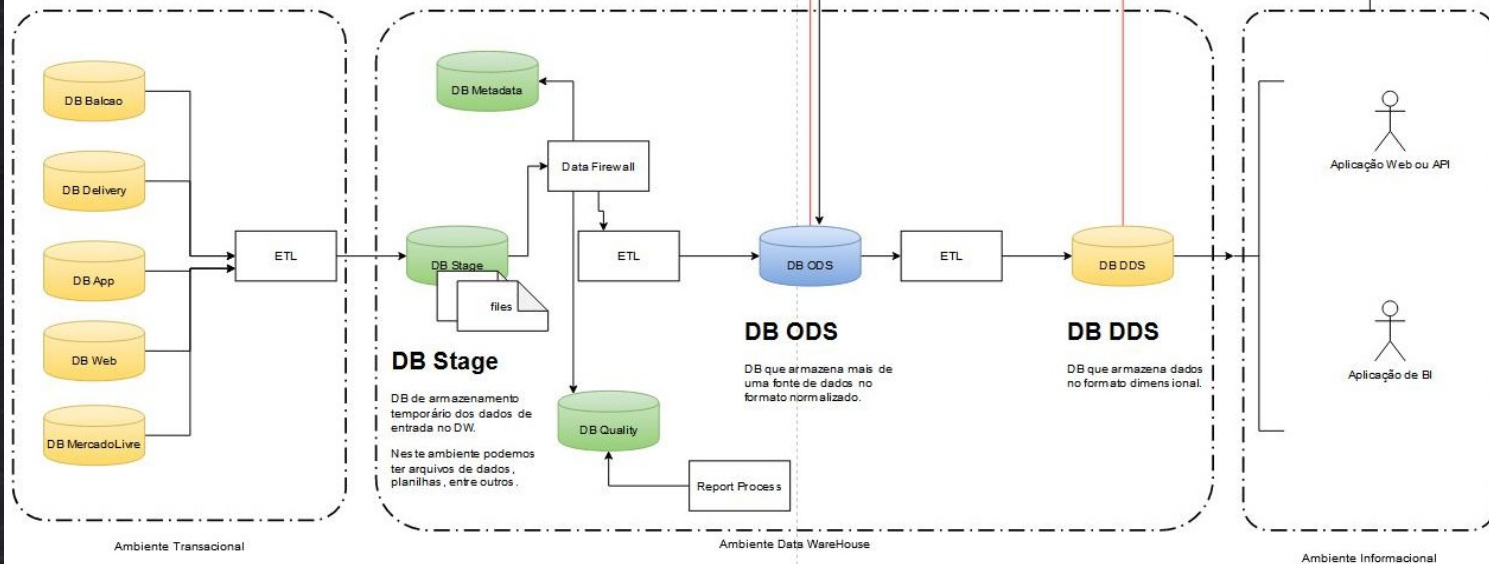


X

Arquitetura de Fluxo de Dados DW

Objetivos principais:

- Desonerar os DBs Transacionais.
- Aumentar o valor do negócio em função das análises de DW



2.

Modelagem Dimensional

Modelagem Dimensional

Modelagem Dimensional



- ✗ Um dos principais ativos da organização é a **informação**
 - Geralmente este ativo é armazenado de duas formas
 - Sistemas transacionais (onde os dados entram);
 - Data Warehouse (de onde os dados são lidos);
- ✗ Em relação a técnica de modelagem, diferenciamos cada uma das formas de armazenamento
 - Sistemas Transacionais utilizam normalmente a **modelagem 3FN**;
 - Repositórios DDS do *data warehouse* utilizam sempre uma técnica de modelagem dimensional;
- ✗ Independente da técnica de modelagem dimensional utilizada, a maior parte dos conceitos são os mesmos

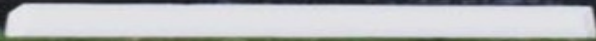
Histórico



- ✗ Apesar de ser um mito popular, não foi Ralph Kimball que inventou a terminologia e abordagem dimensional
- ✗ Pelo que se pode determinar termos como fatos e dimensões já eram utilizados na **década de 60** em pesquisas acadêmicas e corporativas
- ✗ É quase certo que não foi uma pessoa específica que **inventou a abordagem dimensional**.
- ✗ É uma força natural que acontece quando o projetista coloca o desempenho e a facilidade de entendimento como metas principais na arquitetura da base de dados



In data we
trust!



✗ Definição de fato em dicionário etimológico:



fact (n.)

1530s, "action, anything done," especially "evil deed," from Latin *factum* "an event, occurrence, deed, achievement," in Medieval Latin also "state, condition, circumstance," literally "thing done" (source also of Old French *fait*, Spanish *hecho*, Italian *fatto*), noun use of neuter of *factus*, past participle of *facere* "to do" (from PIE root ***dhe-** "to set, put"). Main modern sense of "thing known to be true" is from 1630s, from notion of "something that has actually occurred."

Compare **feat**, which is an earlier adoption of the same word via French. *Facts* "real state of things (as distinguished from a statement of belief)" is from 1630s. *In fact* "in reality" is from 1707. *Facts of life* "harsh realities" is from 1854; euphemistic sense of "human sexual functions" first recorded 1913. Alliterative pairing of *facts and figures* is from 1727.

Facts and Figures are the most stubborn Evidences; they neither yield to the most persuasive Eloquence, nor bend to the most imperious Authority. [Abel Boyer, "The Political State of Great Britain," 1727]



✗ Definição de fato no dicionário Michaelis:

fato

Cuidado com o uso da palavra **fato**. Rigorosamente falando, ela significa aquilo que ocorreu, aquilo que foi feito. Portanto, dizer "fato concreto", "fato que aconteceu" ou "fato verídico", no fundo, é uma impropriedade. Se ocorreu, é um fato e não pode ser falso; se não ocorreu, não constitui um fato. Dispense o adjetivo. Diga, portanto: Minha conclusão baseia-se em fatos (e não "em fatos concretos") / *Explique-me todos os fatos* (e não "os fatos que aconteceram").

- ✗ Usa-se o termo fato para representar uma **medição do negócio**
- ✗ Os dados de **medição numérica de desempenho** do negócio são armazenados em uma tabela tipicamente denominada FATO.

Algumas métricas de negócio



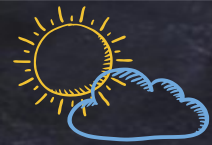
- ✗ Métricas para medir o desempenho do negócio:
- ✗ **CAC** – Custo de Aquisição do Cliente: Valor do investimento feito em Marketing e Vendas dividido pela quantidade de clientes conquistados.
- ✗ **LTV** – Valor do Ciclo de Vida do cliente: Valor médio produzido por um cliente durante seu período ativo na empresa.

Algumas métricas de negócio



- ✗ Métricas para medir o desempenho do negócio:
- ✗ **ROI** – Retorno sobre o investimento: Utilizado para medir quanto dinheiro a empresa esta ganhando/perdendo a cada investimento.
- ✗ **Churn** – Taxa de Churn: Métrica utilizada para medir a perda de clientes por um período.

Exercício



1 - A empresa Hipotética xpto pagou 5.000,00 por um box de 5 m² no evento “grandes marcas” realizado no shopping “Center Dez”. Durante sete dias, o evento recebeu aproximadamente 200.000 visitantes. A empresa gastou no total R\$ 15.000, considerando brindes e outros custos, e cadastrou 10.000 novos clientes. Sabe-se que a porcentagem de captação real é de 10% desse montante. Qual é o valor do CAC estimado?

2 - A empresa “hot dog legal” trabalha com 5 perfis diferentes de clientes. O perfil A tem um ticket médio de consumo de aproximadamente 15,00. O perfil B tem um ticket médio de consumo de aproximadamente 25,00. Sabendo que o perfil C tem período de retenção médio de aproximadamente 12 meses, ticket médio de consumo entre os perfis A e B, e aproximadamente 6 compras/mês, Qual é o valor do LTV estimado?



Para refletir!!



MARKETING, NEGÓCIOS

Burger King distribui grátis 67 mil lanches Whopper a cliente em vassoura

Grupo é conhecido pelas campanhas publicitárias agressivas

Por **Denyse Godoy, Marina Filipe**

Publicado em: 31/10/2020 às 16h40

Alterado em: 31/10/2020 às 21h38

🕒 Tempo de leitura: 2 min



Tabela Fato



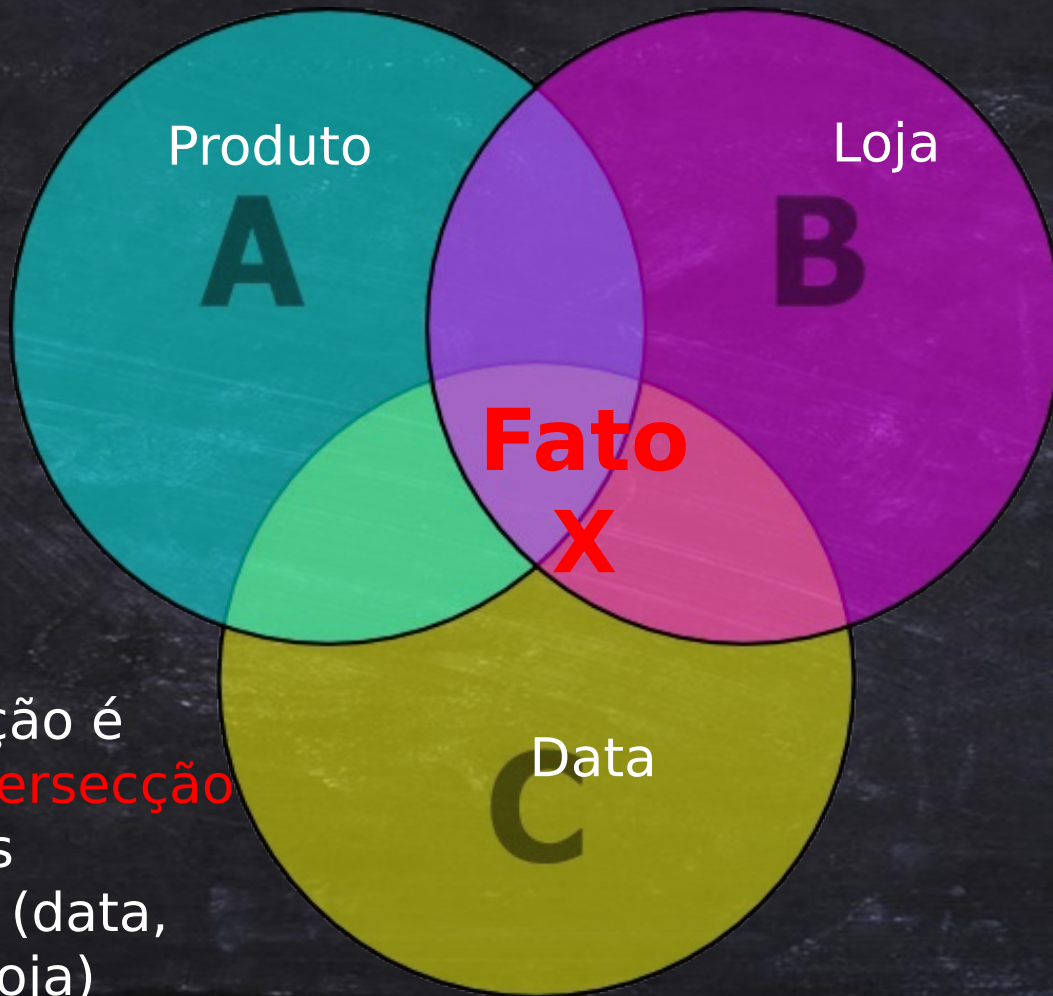
- ✗ Uma **tabela fato** é a tabela fundamental em um modelo dimensional
- ✗ É nesta tabela que os dados de **medição numérica de desempenho** do negócio são armazenados
- ✗ Por exemplo, imagine-se em um supermercado olhando **produtos que estão sendo vendidos** e anotando a **quantidade vendida** e o **total de vendas** em dólar deste produto
- ✗ Isto é um evento do negócio e ele deverá ser armazenado na tabela de fato



Exemplo de Tabela Fato

Daily Sales Fact Table
Date Key (FK)
Product Key (FK)
Store Key (FK)
Quantity Sold
Dollar Sales Amount

- ✗ Esta é uma tabela fato de Vendas Diárias
 - Observe os fatos nestas tabela ou seja as medições do negócio
 - Quantidade vendida
 - Total de vendas em dólar
- ✗ Uma medição é feita na intersecção de todas as dimensões (data, produto e loja)
- ✗ A lista de dimensões define o **grão** da tabela fato e diz qual é o **escopo da medição**



- ✗ Uma medição é feita na **intersecção** de todas as dimensões (data, produto e loja)



Exemplo de Tabela Fato

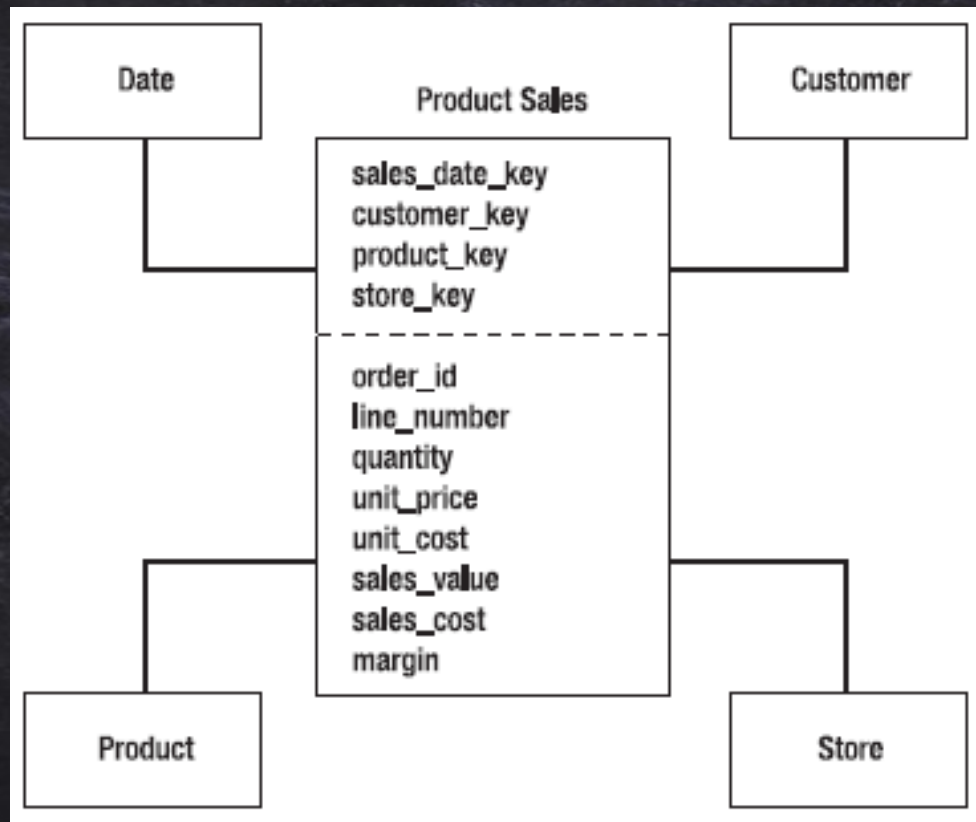


Tabela Fato



- ✗ Cada **linha da tabela** fato corresponde a uma **medição**
- ✗ Uma medição é uma linha em uma tabela fato
- ✗ Todas as medições em uma tabela fato devem estar no mesmo grão
- ✗ As medições são definidas na **fase de levantamento de requisitos** do projeto.
 - O que deseja-se medir? Faz sentido medir? O que eu vou fazer com esta medição?
- ✗ Pode-se ter **fatos calculados** na tabela fato
 - Exemplo: temos a quantidade e o valor unitário de um produto. Com a multiplicação entre os dois temos o valor vendido deste produto

Os Fatos



- ✗ Os fatos mais úteis são numéricos e aditivos
 - Por exemplo, total de vendas em dólar

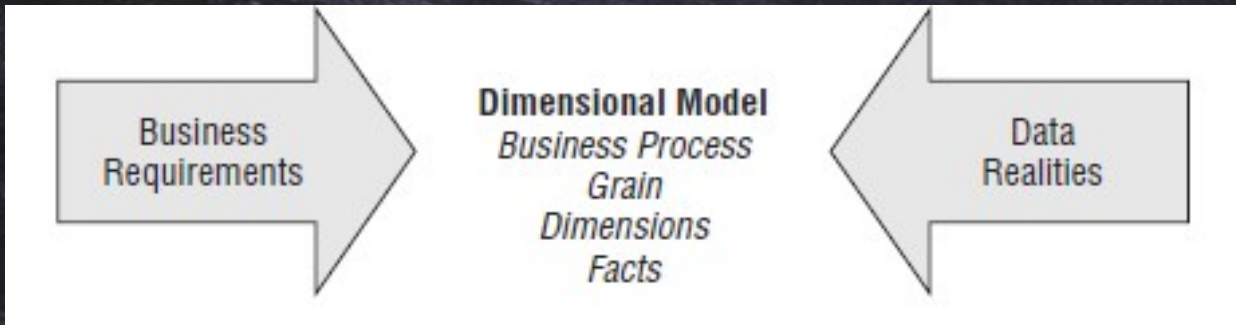
- ✗ Aditividade
 - Crucial, pois é incomum, em aplicações envolvendo data warehouses, fazer uso de apenas uma única linha da tabela fato;
 - Em vez disso, utilizam-se centenas, milhares e até milhões de linhas de uma só vez. Isso justifica a necessidade adicioná-las na tabela.
 - Perceba no nosso exemplo de tabela fato que as duas medições são aditivas
 - Existem fatos não-aditivos e semi-aditivos, também vamos estudá-los
 - Não-aditivos: não podem ser somados de maneira alguma
 - Semi-aditivos: só podem ser somados junto a algumas dimensões

- ✗ Descrevemos fatos como um continuamento valorados como um guia para ajudar a escolher o que é um **fato** e o que é um **atributo de uma dimensão**

Modelagem de Fatos



- ✗ Resista à tentação de modelar os fatos apenas olhando para os “dados de origem”
- ✗ Supere o fato de que: “Mergulhar nos dados pode ser menos intimidador do que entrevistar uma pessoa de negócios”
- ✗ Os dados não substituem os “inputs” de pessoas de negócios





Exercício em Grupo

A empresa MOVIMENTA S/A opera no mercado de locação de automóveis. As locações são realizadas para pessoas físicas por todo o território nacional e 60% do faturamento da empresa está concentrado no estado de São Paulo.

A empresa MOVIMENTA possui uma frota de diferentes tipos de veículos para atender diferentes finalidades. Em função dos diferentes tipos de veículos surgem perfis de diferentes tipos de cliente.

Uma característica interessante a ser considerada no negócio, ainda não explorada pela empresa, é a possibilidade de remanejamento de veículos entre lojas em função da demanda particular de cada uma delas.

Prepare-se para realizar uma modelagem para DW:

- Considerando o contexto apresentado acima, simule uma reunião para debater quais são os fatos envolvidos neste tipo de negócio?





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.
- ✘ Tufano, D. Michaelis Português Fácil: Tira-dúvidas de Redação, Melhoramentos, 2016
- ✘ www.etymonline.com/word/fact, ultimo acesso em 27/02/20



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).