

**TSW017**  
**GRANDES SISTEMAS DE**  
**DADOS**

Terças-Feiras - 19h30 às 22h30  
Prof. Dr. Jorge Rady de Almeida Jr.  
Escola Politécnica da USP

1



*Grandes Sistemas de Dados - Aspectos Iniciais*

Motivação

Sistemas de Informação

BD Relacionais OO e Pós-Relac.

Data Warehouse e Data Mining

3

*Grandes Sistemas de Dados - Aspectos Iniciais*

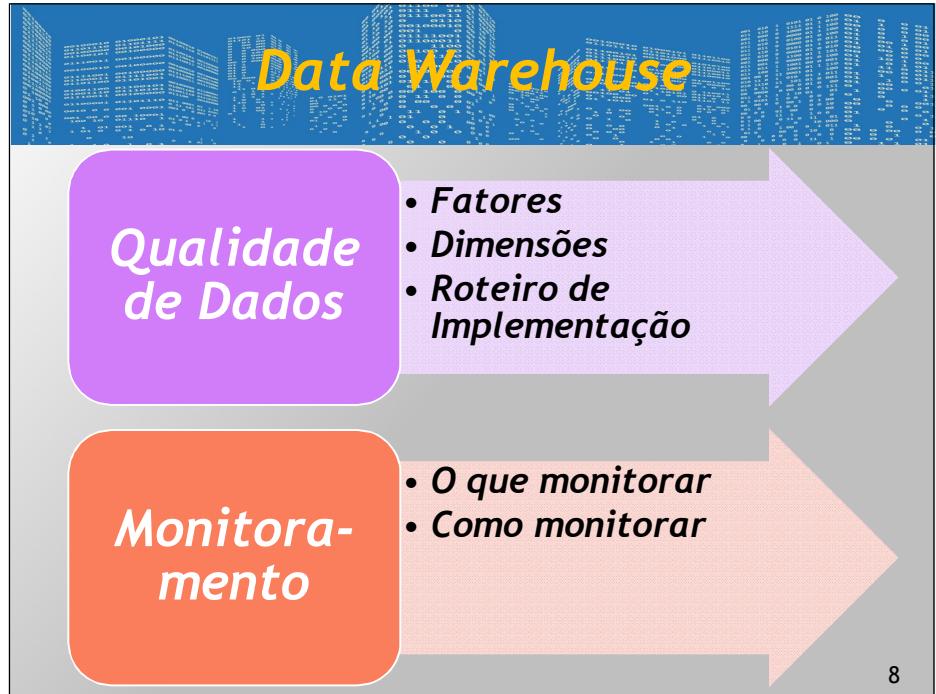
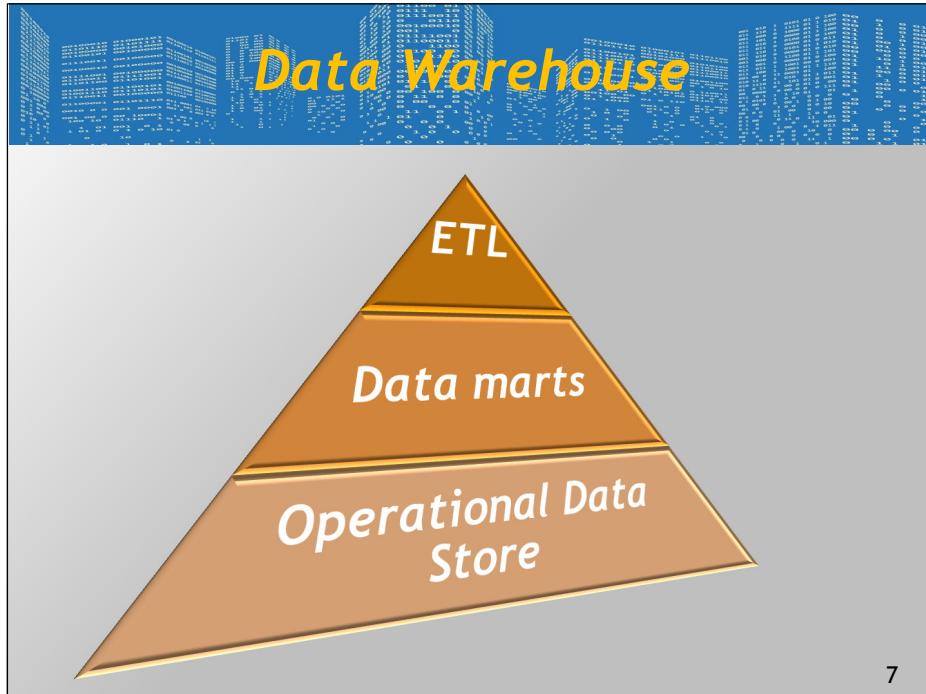
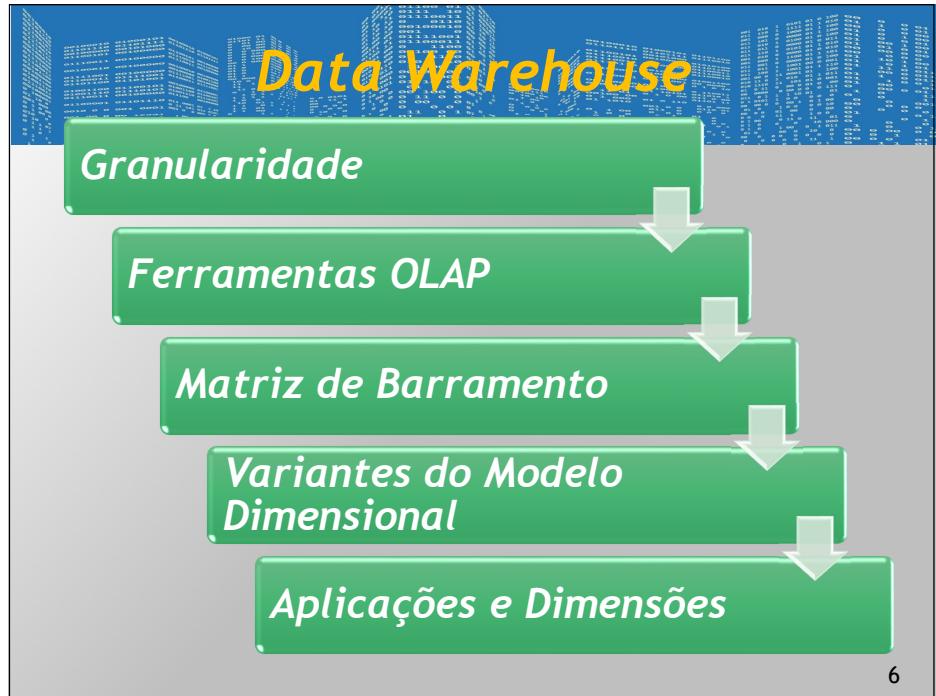
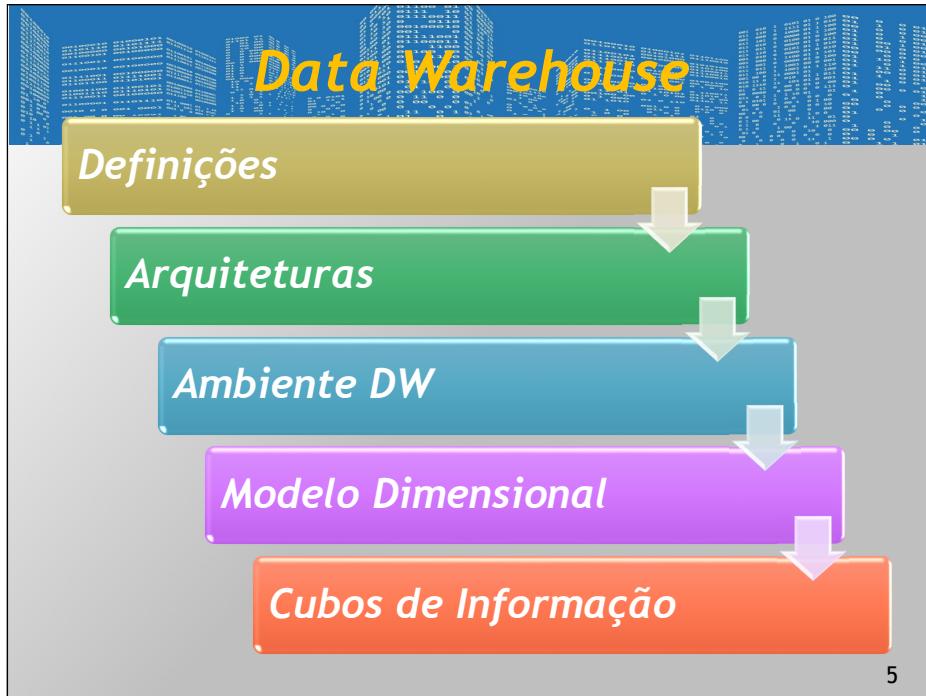
Big Data e Data Lake

Áreas de Aplicação

Qualidade e Segurança de Dados

Governança de Dados

4



# Data Warehouse

Kimball ou Inmon ?

9

# Big Data

Motivação

Os Vs do Big Data

Evolução

10

# Big Data

Definições

Tipos de Análise

Estrutura

Hadoop e Map Reduce

11

# Data Lake

Definições

Estrutura

Tipos de Dados

Armazenar x Organizar

12

# Data Mining

Conceitos

Visualização

Estatística e IA

Algoritmos

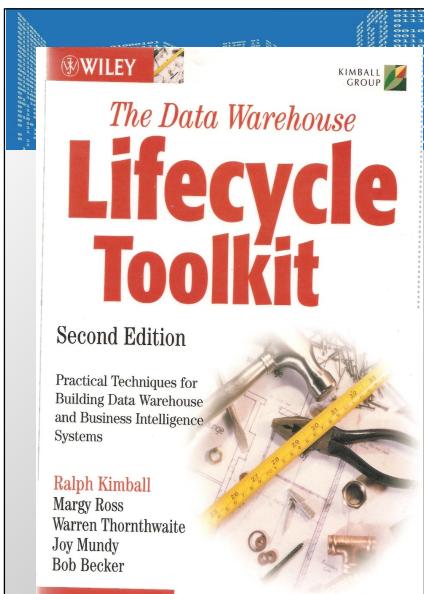
13

# Grandes Sistemas de Dados

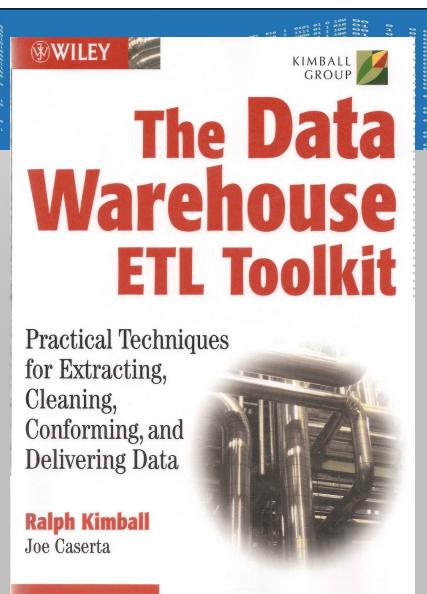
Checklist de Projeto

Equipe de projeto

14

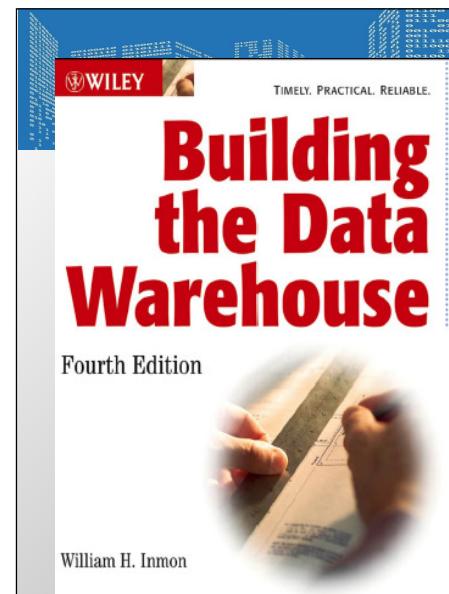


2008

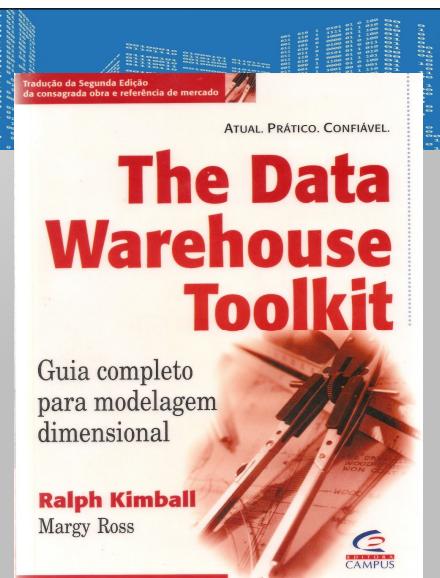


2004

15

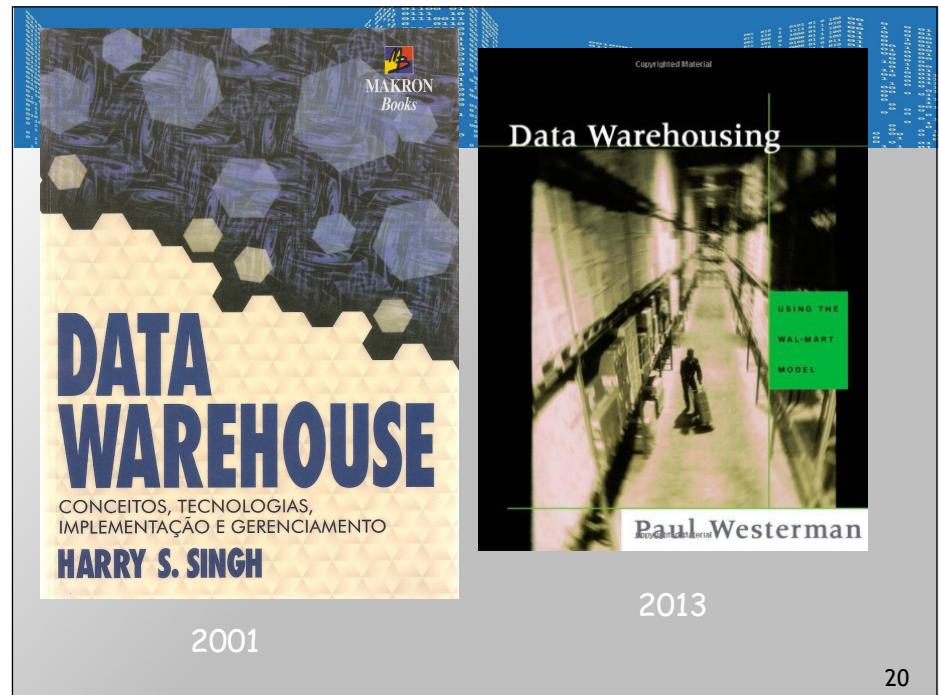
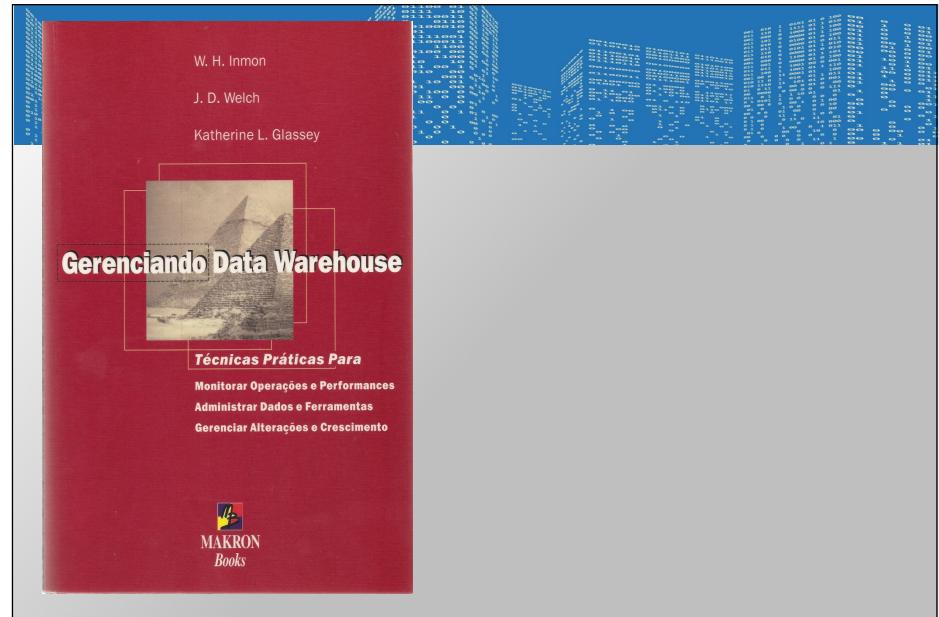
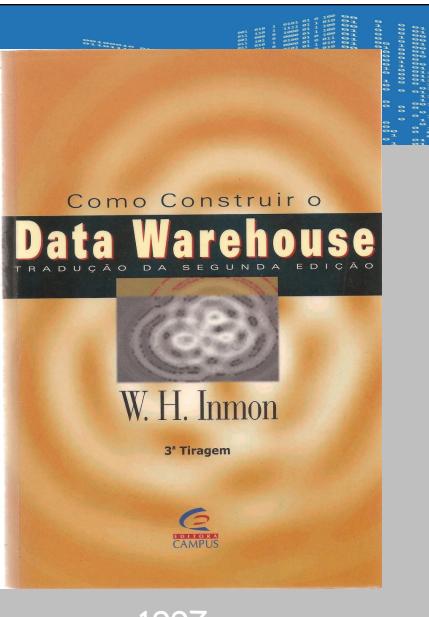
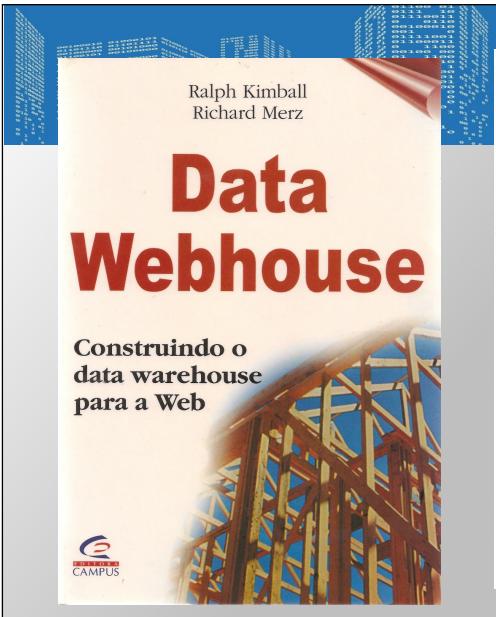


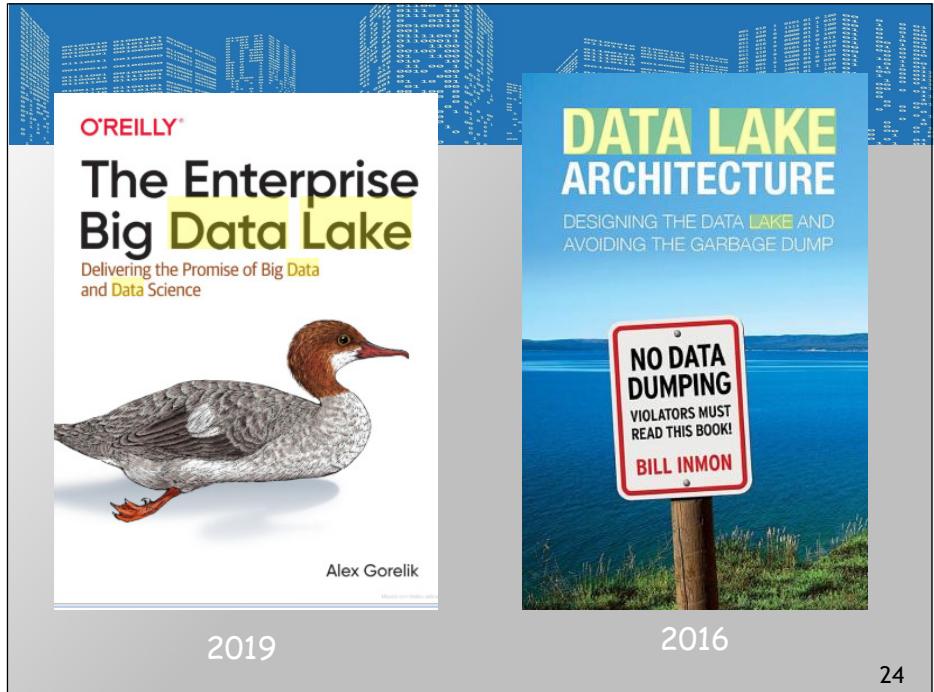
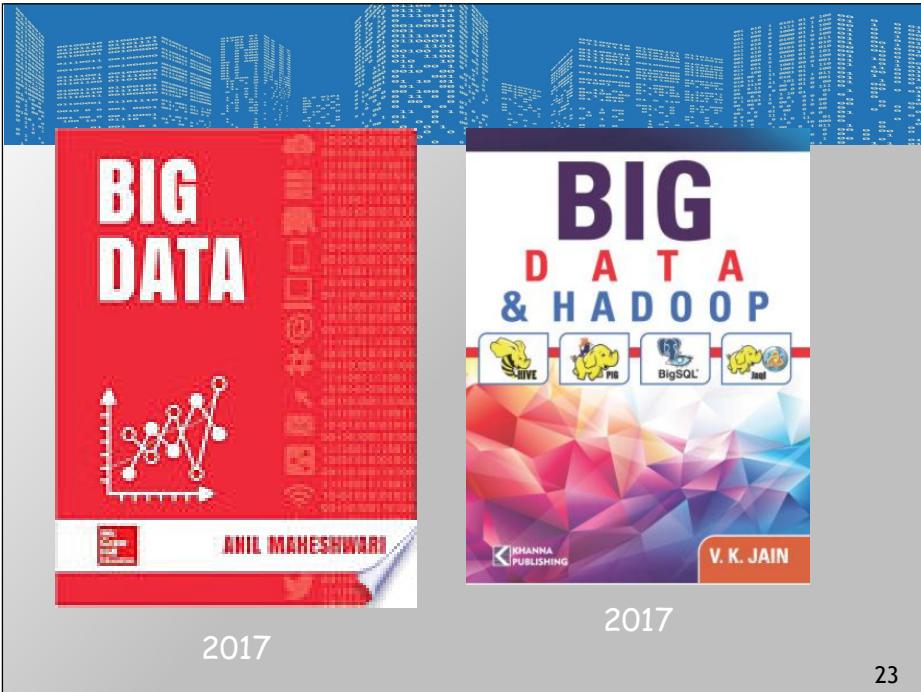
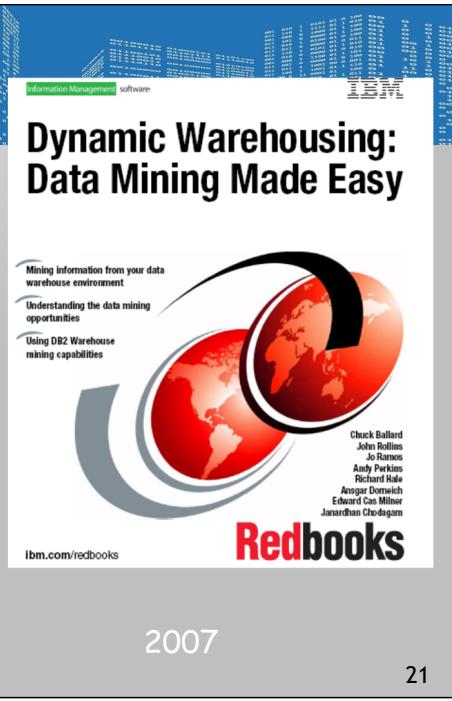
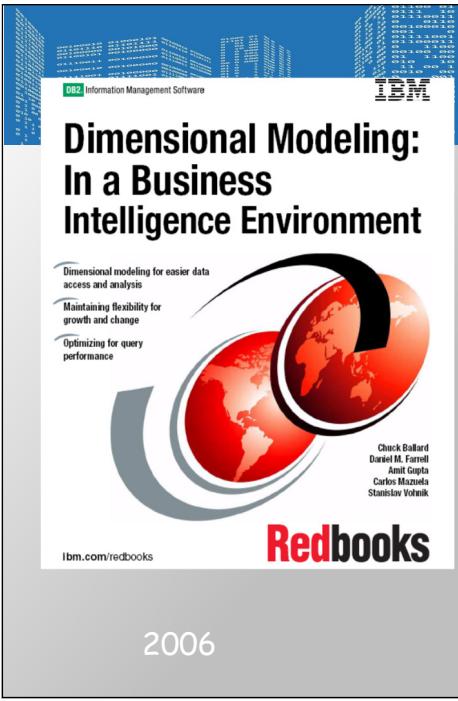
2005

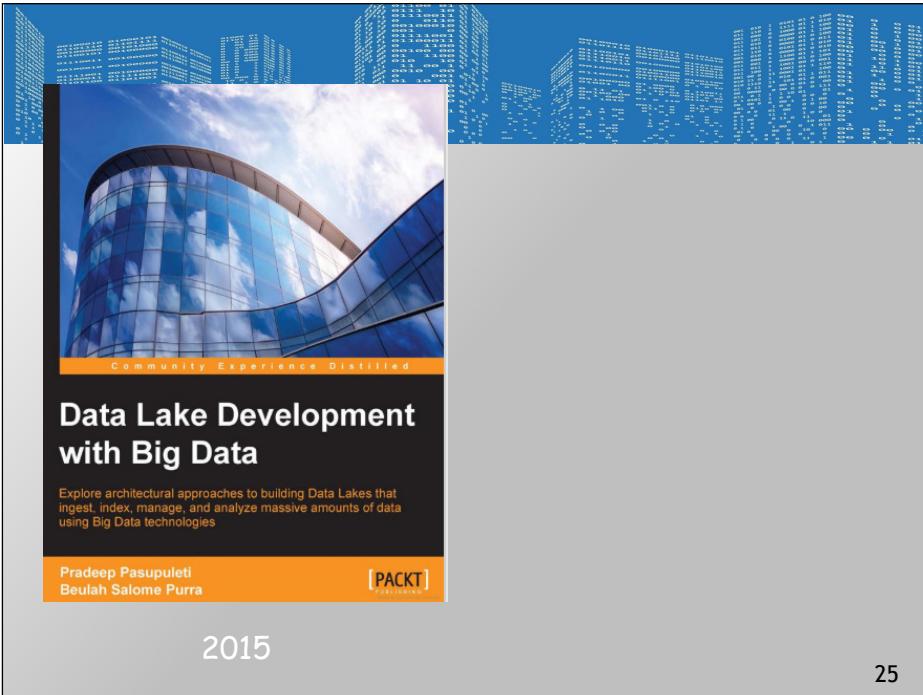


2002

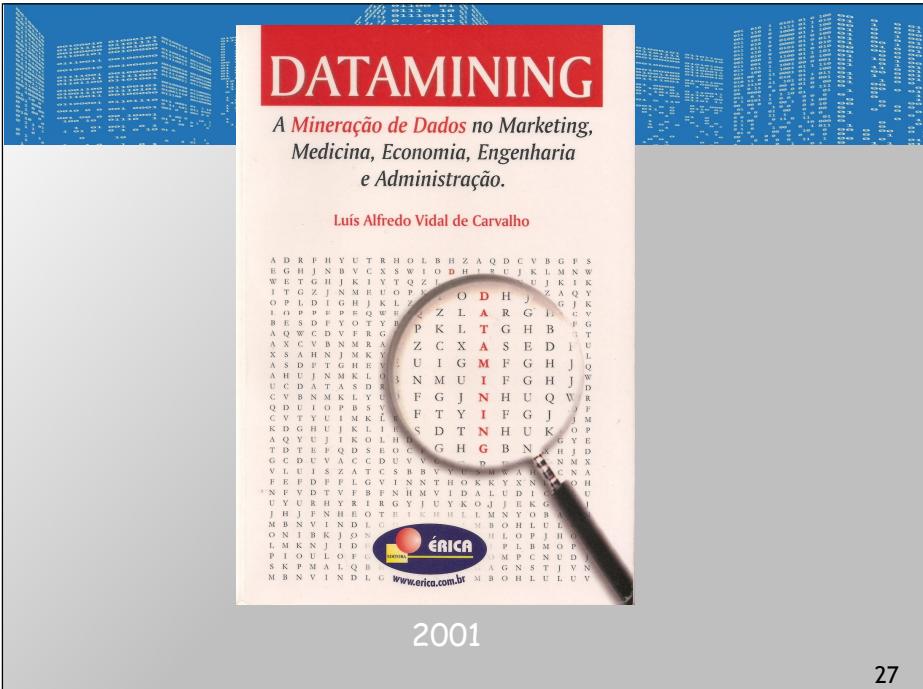
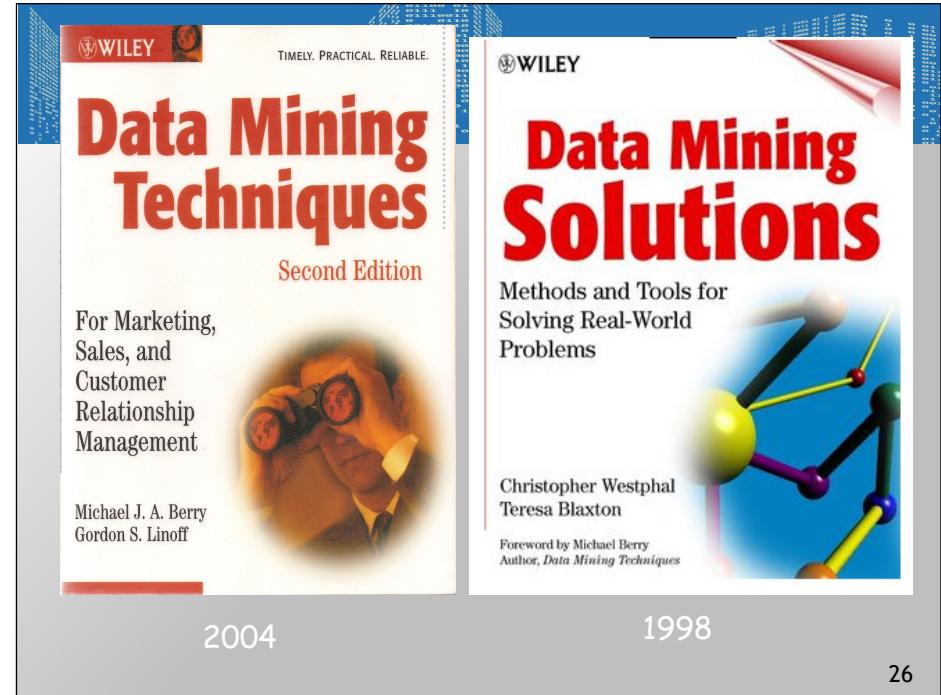
16



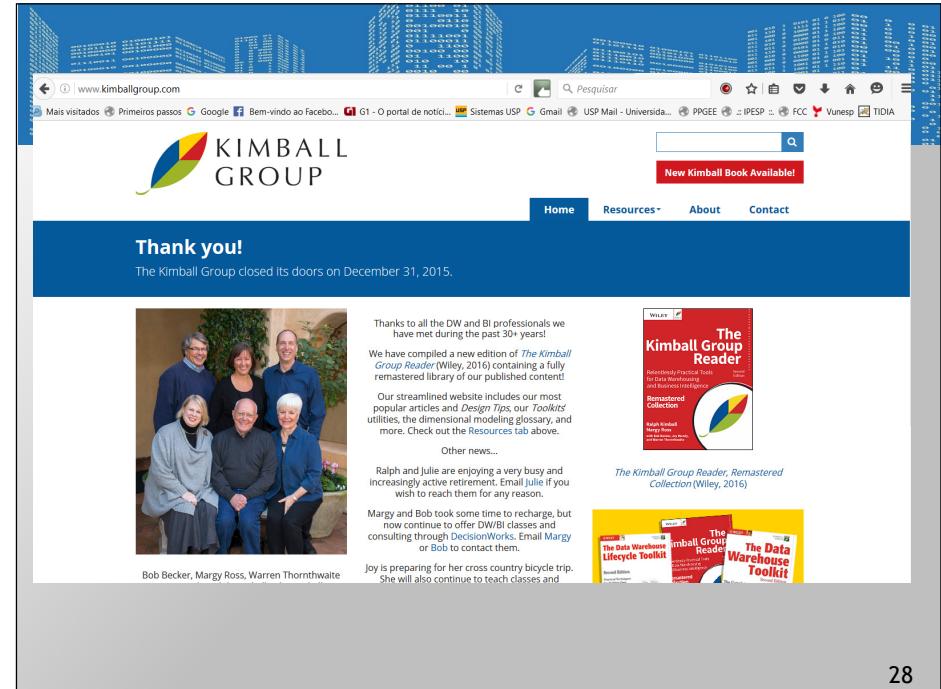




25



27



The screenshot shows the homepage of the Corporate Information Factory website. At the top, there's a navigation bar with links to various websites like Google, G1, and Gmail. Below the bar, the title "Corporate Information Factory" is displayed. A large portrait of Bill Inmon is on the left. To his right, there's a section titled "Products & Services" with links to Consulting, Books, Seminars & Speaking Engagements, Web Events, Training & Certifications, and a "Consulting" link under "CONSULTING". There's also a "CIF Resources Portal" link. Below these sections, there's a "ABOUT THIS WEBSITE" section with a detailed description of Bill Inmon's work in data warehousing. At the bottom, there's a "NEW DOWNLOAD NEW WHITE PAPERS ON:" button.

29

The slide has a blue header with the word "Tarefas" in yellow. The main content area contains a bulleted list of tasks:

- Testes ao final de algumas aulas:  $M_{Testes}$
- Preparação de resumos (solicitados em algumas aulas):  $M_{Resumos}$
- Prova Final:  $N_{Prova}$

Below the list is a mathematical formula: 
$$M_{Final} = (M_{Testes} + M_{Resumos} + 2 * N_{Prova}) / 4$$

30