

**Projeto em ciência de  
dados com soluções  
para processamento  
paralelo e distribuído  
de dados**



## Construindo um *Data Warehouse* a partir de indicadores. Construindo ETL's para uso com o *Data Warehouse* criado

Bloco 1

Marcelo Tavares de Lima






## ► Objetivos

- Apresentar conceitos fundamentais de *Data Warehouse*.
- Descrever sobre a construção de um *Data Warehouse* a partir de indicadores.
- Descrever como construir ETL's para uso com *Data Warehouse*.



## ► *Data Warehouse*

- Com a globalização, “a computação se tornou fundamental para processar grandes volumes de dados”.(NOVAIS, 2012, p. 12)
  - Com o passar dos anos, as empresas passaram seus sistemas de informações da área operacional para áreas mais analíticas, com a intenção de realizar tomada de decisão baseada em dados.
- 



## ► *Data Warehouse*

- *Business Intelligence* (BI), segundo Novais (2012):  
É um conjunto de conceitos e metodologias que, fazendo uso de acontecimentos (fatos) e sistemas baseados nos mesmos, apoia a tomada de decisões em negócios. (NOVAIS, 2012, p. 13)



## ► *Data Warehouse*


- Esse conceito destaca a existência de um grande desafio para os gestores quanto à análise dos acontecimentos (fatos) relacionados ao seu dever de detectar quaisquer informações úteis para alavancar os negócios e ganhar vantagem em relação à concorrência.



## ► *Data Warehouse*

- Novais (2012) afirma que:

O propósito do *Business Intelligence* é permitir a tomada de decisões proativas, ao gerar informações necessárias ao negócio e disponibilizá-los no momento certo. (NOVAIS, 2012, p. 13)





## ► *Data Warehouse*

- Para achar o momento certo, é necessário estar muito bem aparelhado com ferramentas de sistemas de gestão de dados que possam extrair informação relevantes de grandes bases de dados, bem como realizar tal ação no momento apropriado, sem perder a qualidade e a confiabilidade.





## ► *Data Warehouse*

- *Data Warehouse* é uma arquitetura, não uma tecnologia. Representa uma grande base de dados capaz de integrar, de forma concisa e confiável, as informações de interesse para a empresa, que se encontram espalhadas pelos sistemas operacionais e em fontes externas, para posterior utilização nos sistemas de apoio à decisão (Machado 2004, p. 22 *apud* Gura e Benck, 2011, p. 21) .



## ► *Data Warehouse*

- Um dos pioneiros, no estudo de *Data Warehouse*, foi William H. Immon, que apresentou um dos primeiros conceitos a respeito: “é um conjunto de dados consolidados por assunto, não é volátil e está sempre em constante variação quanto ao tempo”. (VIEIRA, 2009, p. 13)



## ► *Data Warehouse*

- Outra definição de *Data Warehouse*, apresentada por Verzola ([s.d.]) é que:

Um *Data Warehouse* é um banco de dados contendo dados extraídos do ambiente de produção da empresa, que foram selecionados e depurados, tendo sido otimizados para processamento de consulta e não para processamento de transações. (VERZOLA, [s.d.], p. 8)



## ► *Data Warehouse*

Vieira (2009) apresenta uma listagem de características que um *Data Warehouse* deve possuir:

1. Conter um conjunto de programas para extração de dados.
2. Ter um banco de dados.
3. Ter um sistema para recuperação e visualização de dados.



## ► ***Data Warehouse***

Já em referência aos possíveis benefícios que pode trazer um *Data Warehouse*, segundo Vieira (2009), podem ser descritos:

1. Capacidade de acesso fácil e rápido a dados.
2. Capacidade de armazenamento de dados de forma consistente.
3. Ser flexível.
4. Ser detentor de ferramentas de consultas e de visualização de dados.
5. Realizar armazenamento confiável de dados.
6. Capaz de construção de dados específicos para direcionamento de negócios.



## ► *Data Warehouse*

Gura e Benck (2011) apresenta uma lista com justificativas para uma empresa implementar um *Data Warehouse* em sua rotina, que replicada a seguir:

- Existência de várias plataformas de hardware e de software.
- Frequentes alterações nos sistemas transacionais corporativos.
- Dificuldade na recuperação de dados históricos em períodos que antecedem o ano vigente de operações.
- Existência de sistemas de diferentes fornecedores.
- Existência de despadronização e de integração de dados dos diversos sistemas utilizados.
- Ausência de documentação e de segurança no armazenamento de dados.



## ► *Data Warehouse:* características

- Orientado por assunto.
- Variante no tempo.
- Não volátil.
- Granularidade.
- Integração.

## Construindo um *Data Warehouse* a partir de indicadores. Construindo ETL's para uso com o *Data Warehouse* criado

Bloco 2

Marcelo Tavares de Lima







## ► Construção de um *Data Warehouse*

- Gura e Benck (2011, p. 31) afirmam que “o processo de fazer *Data Warehouse* é denominado *Data Warehousing*”.
- O desenvolvimento de um armazém de dados tem relação direta com a necessidade dos usuários do sistema e com a realidade de informações disponíveis para sua integração e alimentação.



## ► Construção de um *Data Warehouse*

- Em diversos ambientes, dados operacionais não integrados são complexos e muito difíceis de lidar.
- Para alcançar benefícios reais de um *Data Warehouse*, porém, é necessário passar pelo exercício doloroso, complexo e demorado de integração.



## ► Construção de um *Data Warehouse*

- Aplicativos de extração/ transformação /carregamento (do inglês, *extract/ transform/ load*, ETL) podem automatizar grande parte desse processo tedioso.
- Esse processo de integração deve ser feito apenas uma vez. Entretanto, em qualquer caso, é obrigatório que os dados que migram para o *Data Warehouse* sejam integrados, não apenas lançados no *Data Warehouse* a partir do ambiente operacional, segundo Inmon (2005).



## ► Construção de um *Data Warehouse*

- O objetivo de realizar um ETL em um *Data Warehouse*, segundo Novais (2012, p. 31), é “trabalhar com toda a parte de extração de dados de fontes externas, transformação para atender às necessidades de negócios e carga de dados dentro do *Data Warehouse*”, visando a integração e a padronização dos dados que alimentarão o sistema.



## ► Construção de um *Data Warehouse*

- A extração de dados pode ser realizada com ferramentas específicas, destacando-se as ferramentas *Online Analytical Processing (OLAP)*, que são ferramentas de tratamento de dados multidimensionais, que surgiram nos anos 1960.



## ► Construção de um *Data Warehouse*

- Segundo Sarkis (2001), existe um consenso, entre os estudiosos do assunto, de que o ambiente de um *Data Warehouse* envolve três grandes componentes: *Back End*, o repositório *data warehouse* e *Front End*.

PÓS-GRADUAÇÃO

## Teoria em prática


Bloco 3

Marcelo Tavares de Lima






## ► Teoria em prática

- Imagine que você seja o gerente do departamento de compras de uma empresa de pequeno porte e, por ser uma empresa pequena, possui sistemas de dados mais simples e menos analíticos.
  - A partir de um levantamento realizado por sua equipe, você percebeu que está havendo uma tendência de crescimento no ramo em que atua a empresa.
- 






## ► Teoria em prática

- A partir desse cenário, você, como responsável por um dos principais departamentos da empresa, decide investir em sistemas de armazenamento e tratamento de dados mais avançados e robustos. Portanto, faz uma consulta ao departamento de informática para verificar a possibilidade de colocar em prática seus planos.
- 



## ► Teoria em prática

- A equipe de informática sugere que seja elaborado um sistema de *Data Warehouse* em seu departamento para, assim, permitir a você e sua equipe consultar os dados que alimentam os diversos sistemas do departamento de compras, lembrando também que há um interesse em realizar consultas para avaliar as tomadas de decisões como parte dos processos de trabalho.
- 



## ► Teoria em prática

- Pense na proposta. Como você acha que deveria ser esse sistema? Que tipo de dado seria interessante para alimentar o sistema? Muitas outras perguntas e dúvidas poderão surgir com a proposta do departamento de informática. Avalie!

## Dica do professor

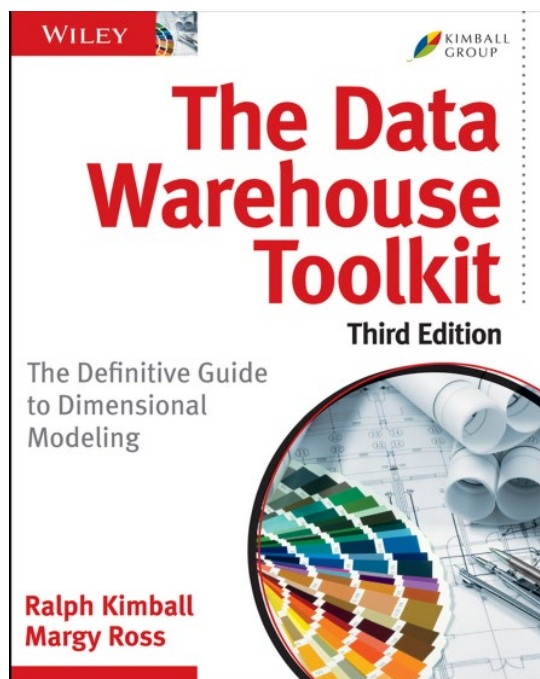
### Bloco 4

Marcelo Tavares de Lima



## ► Indicação de livro

Figura 1 - Livro



Fonte: <https://kbimages1-a.akamaihd.net/b38ba411-4f8e-4c57-b1c2-b55d252f9c85/353/569/90/False/the-data-warehouse-toolkit-1.jpg>. Acesso em: 03 fev. 2020.

**Título:** The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Third Edition.

**Autores:** Ralph Kimball e Margy Ross.

**Editora:** John Wiley & Sons, Inc.

**Ano de publicação:** 2013.



## ► Referências

GURA, E. F.; BENCK, L. L. N. **Construção de um data warehouse, aliado a uma ferramenta open source ireport na geração de informações para tomada de decisão.** Trabalho de conclusão de curso (Curso de Tecnologia em análise e desenvolvimento de sistemas). Coordenação de Informática, Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2011. Disponível em:

[http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/6449/1/P\\_G\\_COADS\\_2011\\_2\\_05.pdf](http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/6449/1/P_G_COADS_2011_2_05.pdf). Acesso em: 03 fev. 2020.

INMON, W. H. **Bulding the data warehouse.** 4. ed. Indianopolis: Wiley Publishing, Inc. 2005.



## ► Referências

NOVAIS, R. R. C. **Modelagem dimensional**. Trabalho de conclusão de curso (Curso de Tecnologia em processamento de dados). Faculdade de Tecnologia de São Paulo. São Paulo, 2012. Disponível em: <http://www.fatecsp.br/dti/tcc/tcc00071.pdf>. Acesso em: 03 fev. 2020.

SARKIS, L. C. **Data Warehouse**: o processo de migração de dados. Dissertação (Mestrado em Ciência da Computação). Universidade Federal de Santa Catarina. Florianópolis, 2001. Disponível em: <https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/80047/227423.pdf?sequence=1&isAllowed=y>. Acesso em: 03 fev. 2020.



## ► Referências

VERZOLA, I. **Data warehouse**. Barueri: Pontes Computadores e Serviços Ltda, [s.d]. Disponível em:  
<http://www.pontes.inf.br/docs/datawarehouse.pdf>. Acesso em:  
03 fev. 2020.

VIEIRA, E. **Tecnologia olap**. Trabalho de conclusão de curso (Graduação em Ciência da Computação). Instituto Municipal do Ensino Superior de Assis. Assis, 2009. Disponível em:  
<https://cepein.femanet.com.br/BDigital/arqTccs/0411150200.pdf>. Acesso em: 03 fev. 2020.



