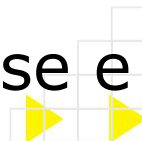




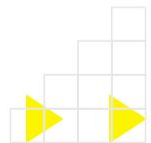
## Curso 3: Administração de Dados Complexos em Larga Escala -- Data warehouse --

Prof. Jose Fernando Rodrigues Junior

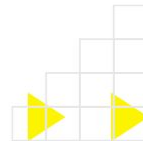
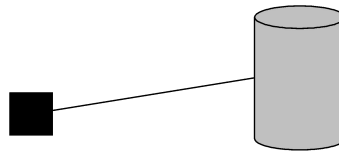
Objetivo: apresentar o conceito de data warehouse e seu papel na coleta de dados



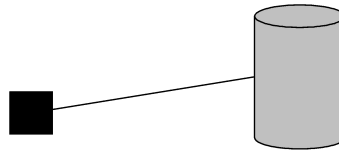
# Big Data warehousing



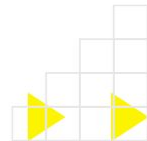
No início, uma única fonte de dados



No início, uma única fonte de dados

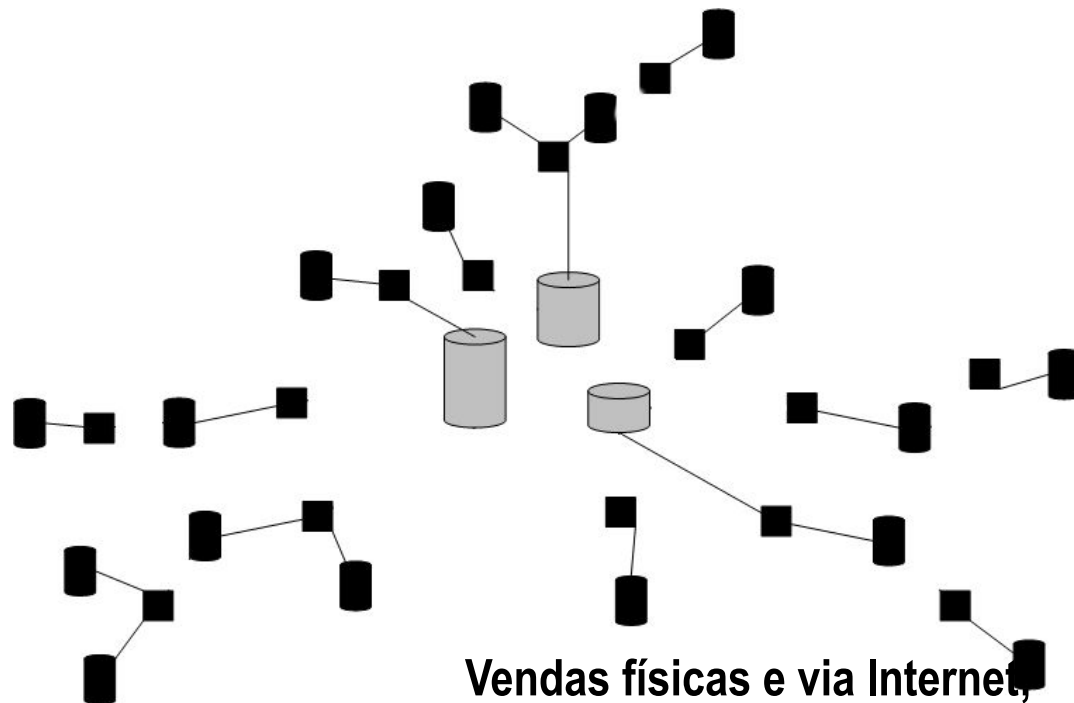


**Exemplo: movimento do caixa da loja física,  
Clientes locais, sem propaganda.**

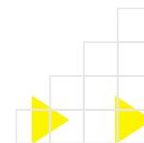


## Mais tarde

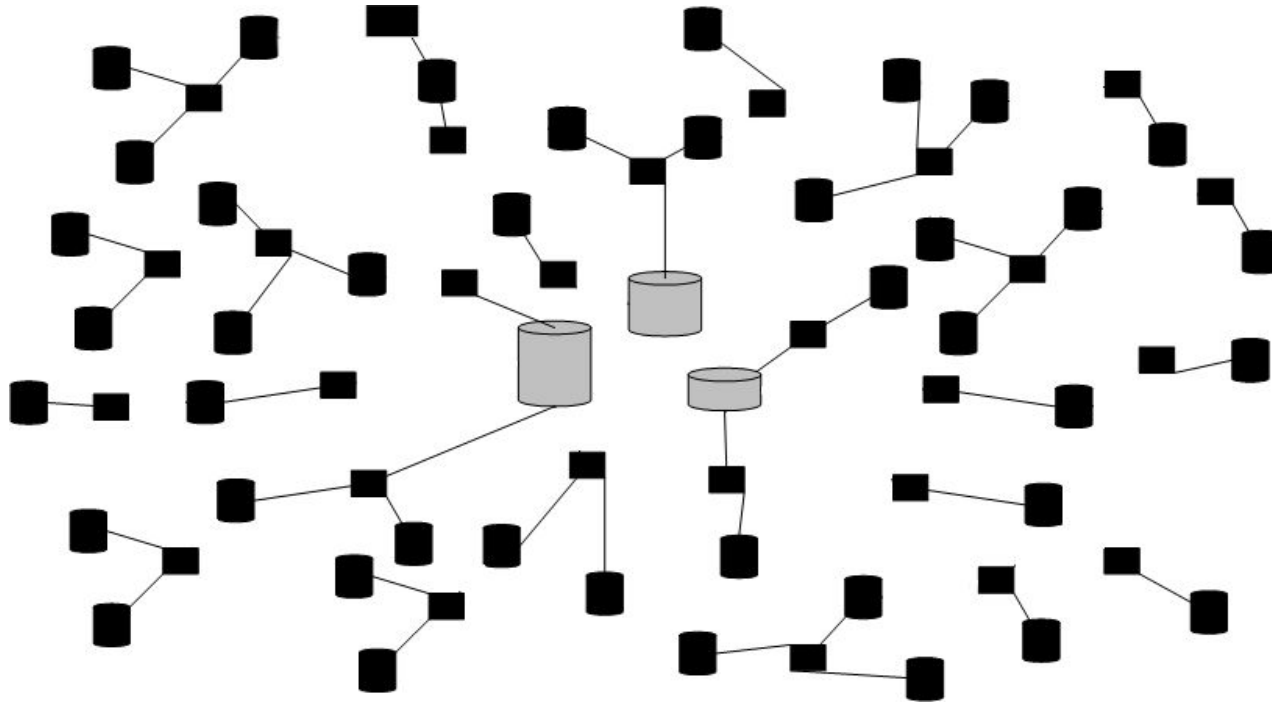
E assim sucessivamente



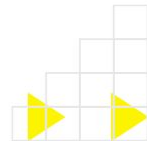
**Vendas físicas e via Internet,  
clientes em todos os locais do país, e nas redes sociais;  
propaganda eletrônica.**



# Finalmente













**Vendas internacionais; clientes em dezenas de países;  
propaganda eletrônica personalizada por atividade e localização.**

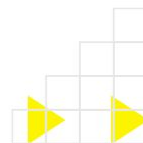


# Até que...



## **O que está acontecendo no meu negócio?**

-  O que é mais vendido?
-  Quando?
-  Por qual loja/meio?
-  Qual o melhor meio de propaganda?
-  Quanto estou faturando?
-  Qual seria uma boa oferta?
-  As vendas cresceram ou subiram?
-  Quais regiões compraram mais?
-  Como fazer propaganda?
-  ...





# Até que...



## O que está acontecendo no meu negócio?

 O que é mais vendido?

 Quando?

Solução:

Usar os dados das lojas para analisar o que está acontecendo no meu negócio.

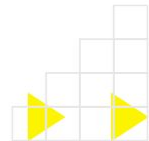
Boa idéia!!!

 As vendas cresceram ou subiram?

 Quais regiões compraram mais?

 Como fazer propaganda?

 ...



# Até que...



## O que está acontecendo no meu negócio?

 O que é mais vendido?

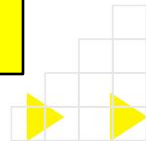
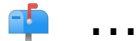
 Quando?

Solução:

Usar os dados das lojas para analisar o que está acontecendo no meu negócio.




Boa idéia!!!

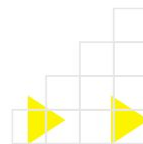
Mas,...





# Empecilhos ao uso dos dados...

-  **Diversos** sistemas de formatos/bancos de dados em uso nos diferentes setores da empresa: **PostgreSQL, Oracle, NoSQL, texto, planilhas, geolocalização, logs, ...;**
-  Cada loja, um **controle transacional diferente**, de acordo com suas necessidades regionais;
-  **Esquemas diferentes, tipos de dados diferentes, distribuição geográfica, falta de interconexão.**





# Empecilhos ao uso dos dados...

**Diversos** sistemas de formatos/bancos de dados em uso nos diferentes departamentos: Excel, Access, Oracle, PostgreSQL, texto, planilhas, etc.

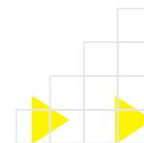
Cada loja possui seus próprios dados, com suas particularidades e formatos.

**Esquemas** diferentes, falta de padronização, distribuição geográfica, falta de interconexão.

Total falta de conformidade!!!

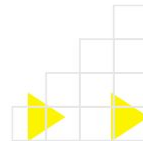
E muitos, muitos dados!!!

O que fazer?

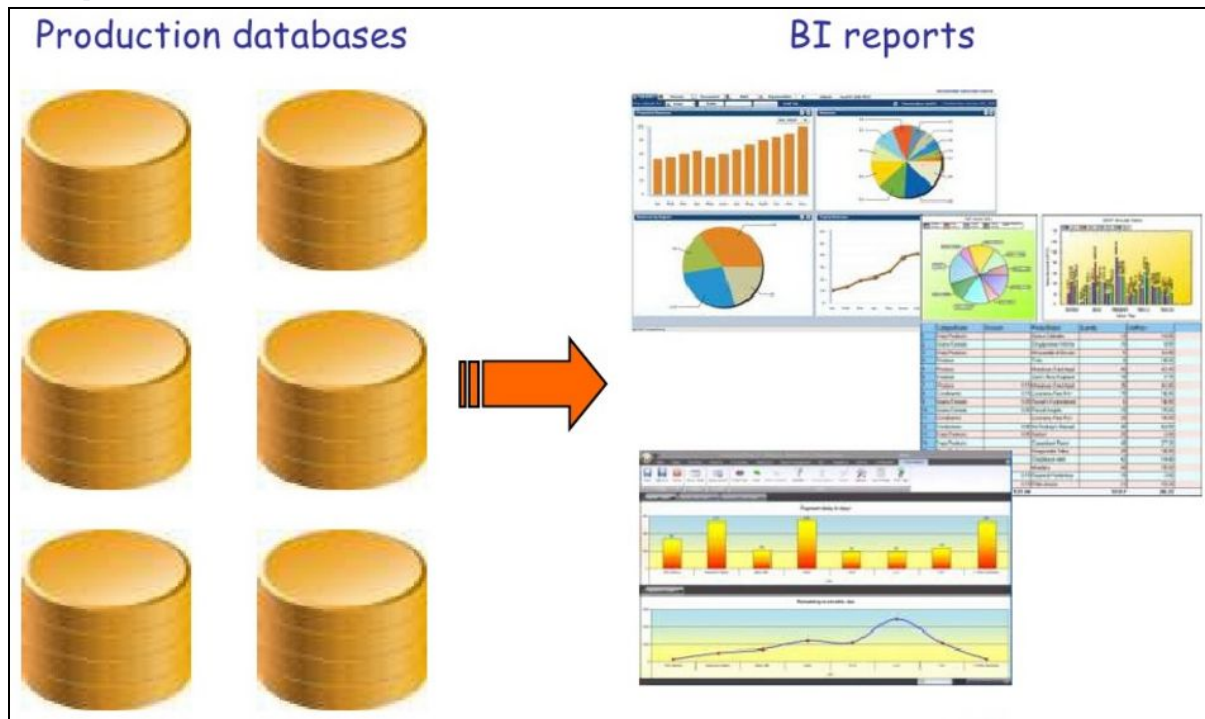




**Data warehouse** se refere a uma classe de tecnologias fundamentais para a administração de dados em larga escala, especialmente na indústria.



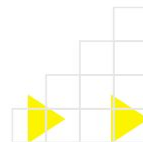
# Empecilhos ao uso dos dados...



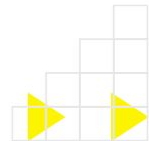
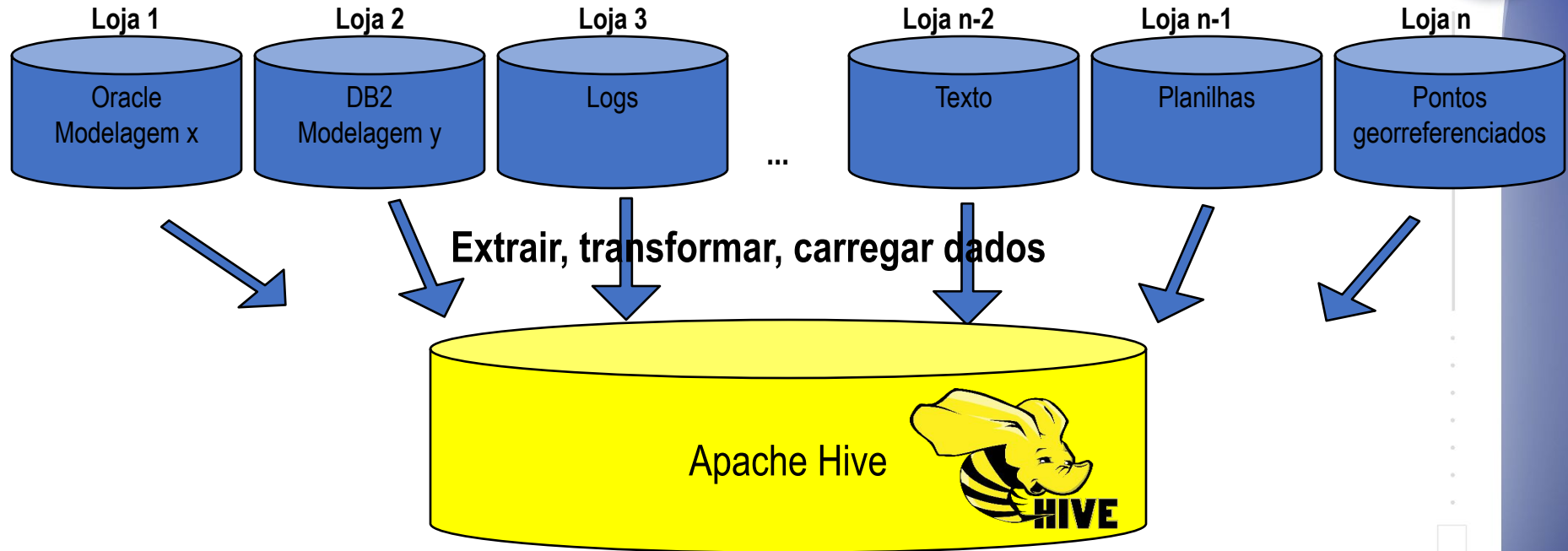


# Objetivos das ferramentas

- Descoberta de conhecimento
  - Avaliar
  - Prever
  - Identificar alternativas
    - Suporte à decisão



# Possível solução

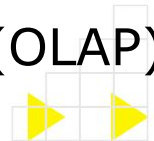




# Data warehouse

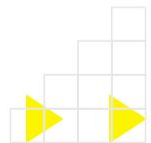


- **Big Data warehouse**
- **Consolidação** dos dados de uma empresa
- **Desempenho** na consulta aos dados
- **Separação** entre suporte à decisão e bancos de dados operacionais
- **Suporte a ferramentas:** Online Analytical Processing (OLAP), aprendizado de máquina, e visualização





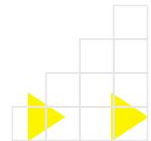
# Arquitetura de um data warehouse





## Arquitetura de um data warehouse

- Arquitetura definida pelo contexto da organização
- De maneira geral, tem as seguintes camadas:
  1. **Operacional (OLTPs):** fornecem dados
  2. **De acesso aos dados:** ETL
  3. **Acesso à informação:** ferramentas de acesso a dados, geração de relatórios, e análise (**OLAP**) □ **Business Intelligence**
  4. **Metadados:** detalhamento do conteúdo do data warehouse □ **dicionário de dados**





## Dicionário de dados

- Descrição dos dados do DW:
  - origem;
  - regras de transformação;
  - nomes e aliases;
  - formato dos dados;
  - histórico de atualizações;
  - acesso e segurança;
  - responsabilidades sobre os dados.
- Sem o dicionário, não há sistema

