

BI – Business Intelligence

BI – Business Intelligence

O conceito de BI de forma mais ampla, pode ser entendido como a utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios das organizações, estas, possuem uma enormidade de dados, mas enfrentam dificuldades na extração de informações a partir dela. O que dificulta o processo de tomada de decisão, na medida em que os gestores se sentem impotentes no processo de sua busca e recuperação. [Barbieri 2001]

BI – *Business Intelligence* - Inteligência Empresarial

- Traduzido para o português como **inteligência empresarial**.
- Conceito abrangente.
- Trata-se de um conjunto de ferramentas, arquiteturas, aplicativos e tecnologias, que transforma grandes quantidades de dados em informações de qualidade para a tomada de decisões.
- As ferramentas de BI permitem cruzar dados, visualizar informações em vários cenários e analisar os principais indicadores de desempenho empresarial.
- As quatro principais ferramentas do conjunto de inteligência empresarial, segundo Batista (2004), são o *data warehouse*, as ferramentas *OLAP*, o *data mining* e as ferramentas de modelagem analítica e de previsões.
- Composto por processos de coleta, transformação, análise e transformação dos dados.

Onde usar?

- Avaliar o ambiente empresarial
- Análises competitivas
- Permitem ancorar decisões
- Indicar necessidades de decisões que nem se tem idéia
- ...

Fatores Críticos de Sucesso do BI

- Foco bem definido – alinhamento de metas.
- Conseguir patrocinador forte.
- Apenas dados necessários—métricas identificadas e selecionadas.
- Evitar overdose de informação.
- Boa arquitetura tecnológica.
- Mensuração metodológica – melhor maneira de medir.
- Medir com um propósito.
- Fase de especificação (levantamento e análise).
- Fator humano, profissionais mais analíticos. Fator humano, profissionais mais analíticos. Como ocorrem os processos decisórios? Pilar do BI é humano. Experiência Humana é relevante.

Fatores Críticos de Sucesso:

- Evitar hipnose dashboard: Evitar a supervalorização dos dashboards.
- Disponibilidade e facilidade de acesso.
- Acompanhamento – ajustes necessários para assegurar que os objetivos estão sendo alcançados.

Questões culturais

- A maioria dos projetos de BI carece da disseminação das melhores práticas
- O que determina a eficácia das soluções de BI são os resultados que essas soluções apoiam.
- O que realmente as pessoas percebem ao se depararem com as informações é influenciado pelas suas experiências, modelos mentais, níveis de motivação, contexto, educação, valores culturais, cultura organizacional, ...
- O aperfeiçoamento corporativo resulta da soma de informações, conhecimentos, experiências, contexto e prática.

O quanto a empresa é analítica?

- Como ocorreremos processos decisórios?
- Onde as decisões custam mais caro?
- É necessário ensinar a empresa a pensar
- O que de fato leva a empresa a tomar as melhores decisões?

Fatores Críticos de Sucesso:

- Petrini, Pozzebon e Freitas (2004) pesquisaram a utilização de *BI* em grandes empresas brasileiras, através de uma técnica de *survey*. O resultado do estudo demonstrou que a utilização dessa ferramenta está ocorrendo há pouco mais de três anos em 73% das empresas pesquisadas. Os autores concluíram, também, que geralmente esse tipo de tecnologia é criada a partir de objetivos tecnológicos, sem levar em conta as necessidades informacionais: “há uma falta de foco na determinação de quais informações são mais relevantes para o negócio, ou até mesmo alinhar indicadores, que seriam incluídos no sistema com objetivos estratégicos” (PETRINI; POZZEBON; FREITAS, 2004, p. 12). Os autores identificaram que as empresas que utilizam algum tipo de ferramenta de *BI*, voltam sua atenção à metodologia de avaliação de desempenho e não à geração e flexibilização da informação.

Data warehouse/Data Mart (*armazém de dados*) *consolida, organiza e armazena dados que foram extraídos dos vários bancos de dados organizacionais, sejam eles operacionais, históricos ou externos, de forma que possam ser mais bem utilizados para análise.*

Ferramentas OLAP (processamento analítico on-line) permitem o exame e a manipulação interativa de grandes volumes de dados a partir de diversas perspectivas.

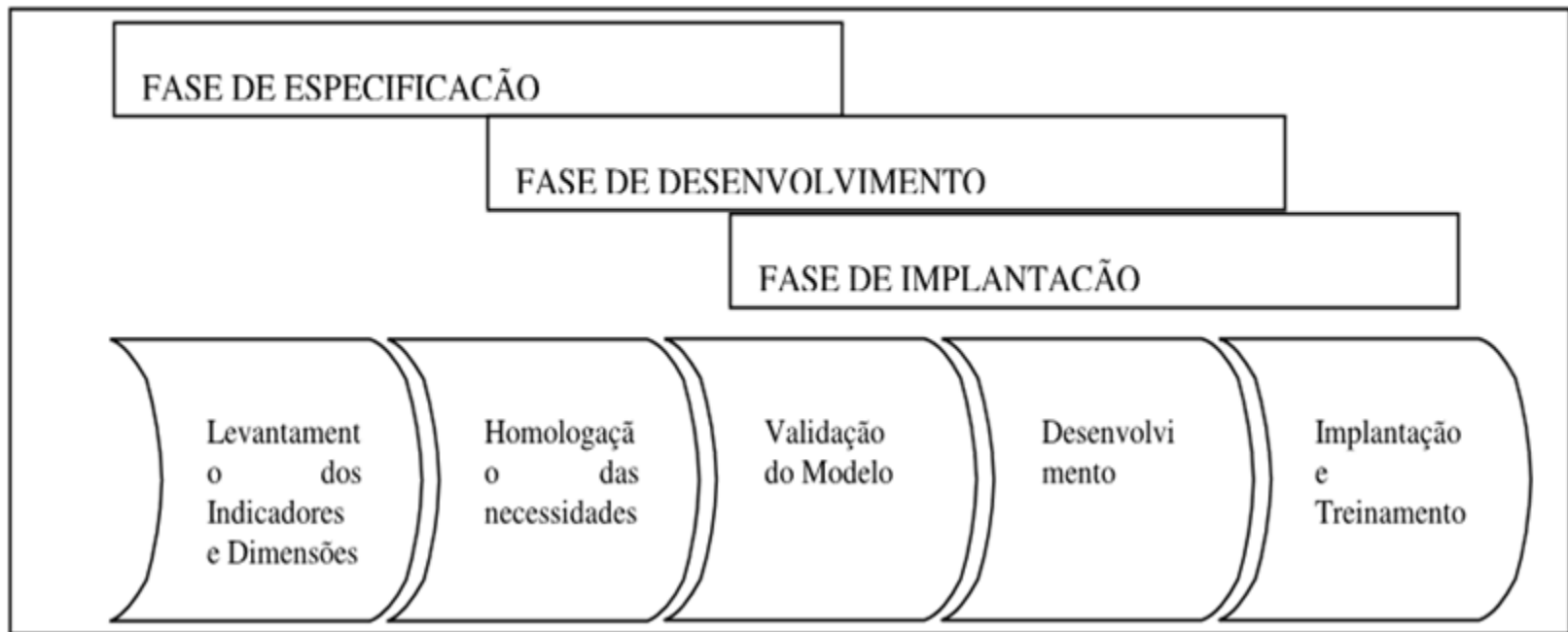
Data mining (mineração de dados) procura descobrir padrões, tendências e correlações ocultas nos dados, que possam propiciar uma vantagem competitiva a uma empresa.

Data Warehouse (DW)	Data mart (DM)	OLAP	Data Mining
<ul style="list-style-type: none"> • Conjunto de dados organizado por assunto e integrado por <i>data</i>; • Ferramenta capaz de gerenciar grandes quantidades de dados, modelando-os para suprir as necessidades dos executivos por informações mais rápidas sobre o desempenho da empresa. 	<ul style="list-style-type: none"> • Subconjunto lógico e físico do <i>DW</i>, suscetível às consultas inesperadas dos usuários; • Estruturas moldadas com dados encontrados no <i>DW</i>, pertencentes a áreas específicas na empresa, como finanças, contabilidade, vendas etc. 	<ul style="list-style-type: none"> • Facilita o acesso do usuário à base <i>DW</i> em que são realizadas consultas possibilitando melhor análise das informações; • Capacidade atribuída aos sistemas que permite aos gestores examinar e manipular interativamente grandes quantidades de dados detalhados e consolidados a partir de diversas perspectivas. 	<ul style="list-style-type: none"> • Utiliza modelos sofisticados para gerar modelos de previsões; • Exploração e análise, por meios automáticos e semi-automáticos, de grandes quantidades de dados para descobrir padrões e regras significativos; • Atende a fluxo de trabalho imprevisível, e propicia a análise em dados atuais e históricos para determinar futuras ações.

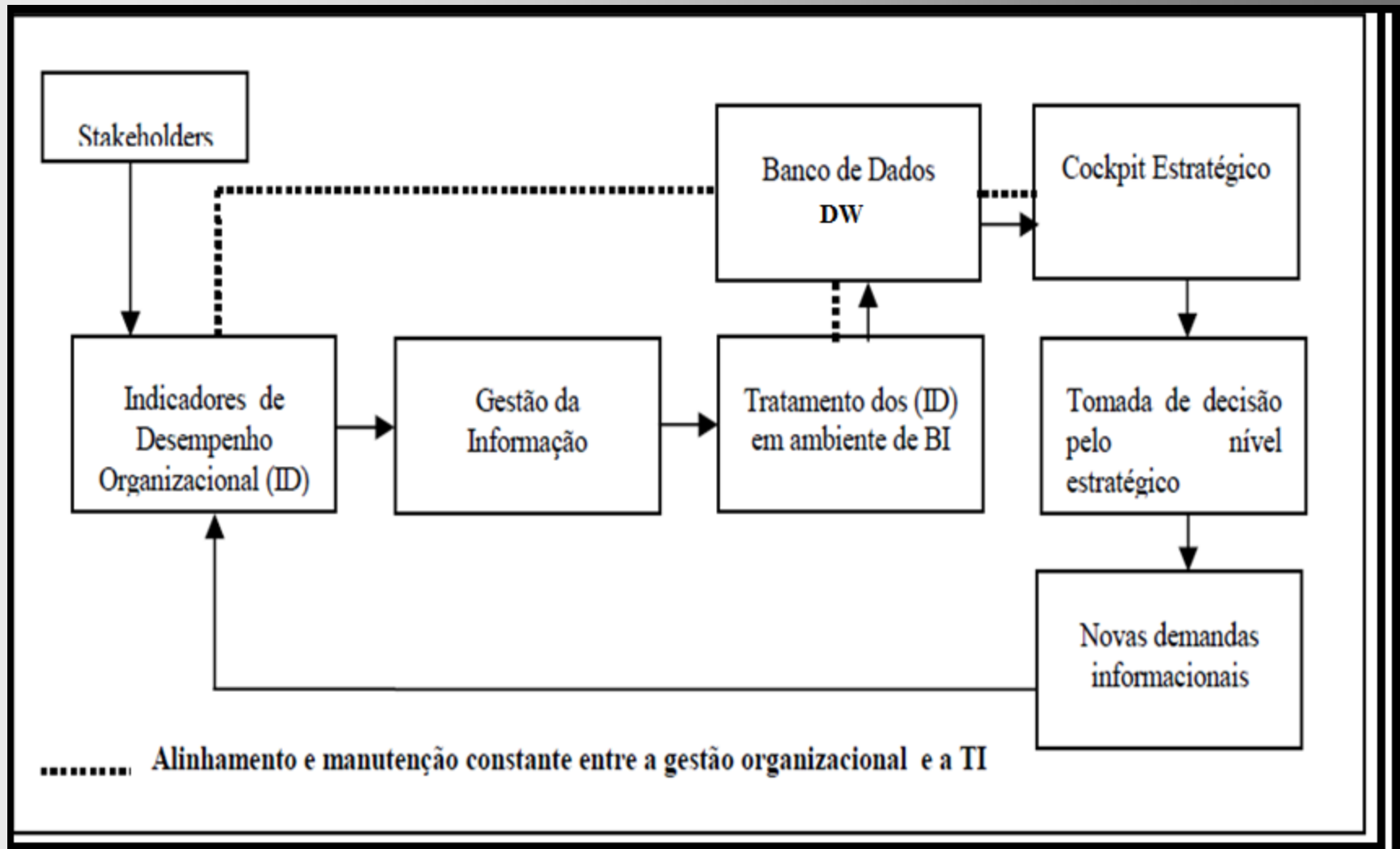
Fonte: Elaborado a partir de Barbieri (2001); Inmon (1997); Harrison (1998)

Metodologia Geral de Desenvolvimento e Implantação de BI

Adaptado de [Ecs 2004]



Exemplo de utilização:



Exemplo de utilização:

- stakeholders – é fundamental entender e conhecer as necessidades dos stakeholder.
- identifica-se os indicadores de desempenho organizacionais que devem ser mapeados.
- gestão da informação – processo de seleção, arquivamento, descarte e outras atividades no processo de gestão da informação, por exemplo, padronizar regras de atualização dos dados.
- cockpit estratégico – onde o tomador da decisão estratégica buscará, interpretará, cruzará informação, visando a administração por resultados da sua organização.

Exemplo de utilização:

- quando houver necessidades de novas informações pelo nível estratégico, deve-se rever os indicadores, bem como seu tratamento no ambiente de BI – compatibilizando assim, o ambiente de gestão com o ambiente tecnológico – ambos, insumos para a gestão estratégica organizacional.
- o objetivo da linha tracejada ligando os indicadores com banco de dados –DW e o tratamento dos Indicadores em ambiente de BI está justamente no alinhamento constante entre a construção e reconstrução de estratégias e indicadores estratégicos com o ambiente tecnológico.

Evolução das teorias estratégicas

Significativas mudanças no mundo dos negócios sucedem-se ininterruptamente... Fronteiras econômicas são redefinidas... Competições, no contexto nacional e internacional, são intensificadas... Constantes avanços da tecnologia emergem, a todo momento, criando cenários de negócios substancialmente diferentes e transitórios.

Os gestores – verdadeiros responsáveis pela performance das empresas – devem estar preparados para essas mudanças, aprimorando, cada vez mais, suas habilidades para atuarem em um ambiente de negócios diferente e complexo que, em muitos aspectos, ainda não existe.

Nesse contexto, espera-se que o executivo deste novo milênio desenvolva uma percepção sistêmica do processo de planejamento estratégico e desenhe a visão estratégica do negócio, compreendendo tanto as questões fundamentais das diversas áreas funcionais de negócios quanto a inter-relação entre essas áreas.

Sob essa ótica, é necessário analisar os processos e modelos a serem trabalhados pela organização, de modo a otimizar a administração das mudanças e a obtenção de objetivos, como também elaborar planos de ação como instrumentos para potencializar os pontos fortes e as oportunidades da empresa, neutralizando suas ameaças e seus pontos fracos.

Estratégia

Para um dos principais pensadores da Escola da Administração Estratégica – Igor Ansoff –, estratégia consiste, basicamente, em um conjunto de regras de decisão para orientar o comportamento de uma organização.

Estratégia pode ser entendida como o conjunto de opções, diretrizes e valores, em todas as suas dimensões – mentalidade, pensamento, visão, gestão, planejamento, administração, desenvolvimento e até controle – que um grupo de líderes determina para que uma organização atinja desenvolvimento sustentável de longo prazo.

Planejamento Financeiro

Características principais:

- ✓ Orçamento anual.
- ✓ Controle Financeiro.
- ✓ Administração por objetivo.

Valor: cumprir o orçamento.

Fonte: FGV

Planejamento Financeiro

- Planejamento financeiro- Processo que envolve a transformação dos dados financeiros em uma forma, de modo a orientar a posição financeira da empresa, avaliar a necessidade de aumento da capacidade produtiva e determinar que tipo de financiamento adicional deve ser feito.
- Orçamento anual - Processo de planejamento, alocação e controle das variáveis do orçamento anual da empresa, tais como receitas, investimentos, despesas.
- Utilizava o enfoque top-down – de cima para baixo –, ou seja, contava com um único estrategista, que era o executivo do topo da pirâmide organizacional.
- A Escola do Planejamento Financeiro era altamente formal – quase mecanicamente programada pelo orçamento anual – na busca da eficiência dos processos. A empresa estimava seus vários gastos com base na previsão das receitas.

Planejamento a Longo Prazo

Características principais:

- ✓ Projeção de tendências.
- ✓ Análise de lacunas.
- ✓ Curva de experiência.
- ✓ Cenários

Valor: projeção do futuro.

Fonte: FGV

Planejamento a Longo Prazo

- Elaboração de poucos cenários - técnica de formulação de várias descrições hipotéticas, em termos de variáveis qualitativas e quantitativas de eventos futuros. Uso de regras conhecidas de causa e efeito.
- Técnica do preenchimento de lacunas existentes - estimativa dos recursos necessários para uma situação futura, obtida a partir do estudo e da projeção do comportamento de indicadores atuais.
- Projeção de futuro. O futuro pode ser estimado a partir da projeção de indicadores passados e atuais.
- Valores dão sentido comum além dos objetivos a atingir.
- Curva de experiência - técnica de análise estratégica aonde a empresa podia prever a diminuição progressiva dos custos de um produto à medida que aumentava sua produção, considerada com produtos ou mercados em crescimento, ou seja, não impactados por descontinuidade. (ex: de uso: desencorajar os concorrentes, consolidando uma vantagem de custos, e podendo usufruir de margens significativas e estáveis).

Planejamento Estratégico

Características principais:

- ✓ Pensamento estratégico.
- ✓ Análise de mudanças do ambiente.
- ✓ Análise de recursos internos e competências.

Valor: definição de estratégia.

Fonte: FGV

Planejamento Estratégico

- Pensamento estratégico: sistema de levantamento de situações com objetivo de definir estratégia.
- Análise de mudanças do ambiente – levantamento dos principais fatores ambientais que afetam a empresa. É uma das etapas iniciais do planejamento estratégico. Identifica ameaças e oportunidades no ambiente externo em que a empresa está inserida, bem como os pontos fortes e fracos de seu ambiente interno.
- Análise de recursos e competências – esforço de ampliação de conhecimento dos recursos da organização com objetivo de alocação eficiente.

Triângulo de Robert Anthony divide a organização em 3 subsistemas.



Administração Estratégica

Características principais:

- ✓ Análise de competitividade
- ✓ Análise da estrutura da indústria.
- ✓ Cadeia de valor.
- ✓ Foco na Análise e implementação
- ✓ Pesquisa e informações com base analítica
- ✓ Estratégia genéricas.

Valor: determinar a atratividade da indústria.

Fonte: FGV

Administração Estratégica

- Importância da implementação das estratégias, e não apenas do processo pelo qual elas foram formuladas. Diagnóstico seguido de prescrição e depois ação.
- Estratégia genérica de competição - Voltado para o longo prazo, para as decisões não estruturadas, para a eficácia, para as informações e para a evolução da empresa.
- Análise da competitividade – lida com a decisão de como vai obter êxito em cada área de negócio que a empresa vai atuar.
- Cadeia de valor- segmenta a organiza as atividades em atividades primárias (logística de entrada, operações, logística de saída, marketing, vendas, serviços pós-venda) e de suporte (rh, suprimentos, infraestrutura da organização).
- Criação e aperfeiçoamento de ferramentas analíticas.
- Análise da estrutura da indústria - a essência da formulação de uma estratégia é a relação da empresa com seu meio ambiente. Porter propôs a análise das cinco forças competitivas: Rivalidade entre concorrentes, Entrada de novos concorrentes, Ameaça de produtos substitutos, Poder de negociação dos fornecedores, Poder de negociação dos clientes. A escola reforçou dois tipos básicos de competitividade: baixo custo e diferenciação.

Gestão Estratégica

Características principais:

- ✓ Pensamento Sistêmico.
- ✓ Integração planejamento e controle.
- ✓ Organização estratégica.
- ✓ Direção estratégica.
- ✓ Coordenação estratégica.

Valor: sintonia com ambiente interno e externo.

Fonte: FGV

- ✓ **Organização estratégica** – estabelecimento da estrutura formal de autoridades. (grupo par fazer o plano estratégico, consultor, alta administração..)
- ✓ **Coordenação estratégica** – articula e harmoniza os esforços. Elabora e revisa o plano estratégico, treina em educação estratégica, coleta informações, articula e harmoniza os processos de gestão estratégica e outras áreas da empresa.
- ✓ **Direção estratégica** – fase de implementação do plano. Orientação das operações a serem executadas. Basicamente, comunicação, estímulo e liderança.
- ✓ **Planejamento estratégico** – aponta o que deve ser feito para que a empresa esteja preparada para as incertezas. Estabelece um meio sistemático para a tomada de decisões. Aponta o que deve ser feito, para que a empresa atinja seus objetivos atuais e futuros. Encara o planejamento como um meio e não um fim.
- ✓ **Controle estratégico** – estabelecimento de padrões de medidas. Medição e Comparação para identificar e corrigir problemas e evitar reincidência. Verifica a utilização dos recursos.

Exemplo:

RHODIA ANALISA VENDAS COM USO DE BI

A fabricante de produtos químicos, Rhodia, adotou um sistema de *business intelligence* para monitorar e gerenciar o desempenho de diversas áreas de atuação da empresa. O projeto resultou na construção de uma base de dados que reúne informações de vendas, estoques e rentabilidade em âmbito mundial.

Com o sistema, a Rhodia passou a acompanhar o ciclo diário do faturamento, incluindo a comercialização e a cadeia de suprimentos (*supply chain*), além de fazer a medição da entrega aos clientes, monitoração dos estoques e análise da rentabilidade sobre as vendas.

Aproximadamente 400 funcionários da empresa em todo o mundo têm acesso aos dados utilizados na rotina de tomada de decisões estratégicas. As informações são consolidadas na matriz, na França, e cada unidade possui versão local. No Brasil estão os dados do país e da Venezuela.

O sistema brasileiro, conhecido como DW Brasil, abrange seis das 13 unidades de negócios da Rhodia e está dividido em duas áreas. Uma é de análise das operações de vendas e *supply chain*, que conta com 50 usuários e deve chegar a 150 em um ano, e a outra é de gestão de pessoal, para controle do bônus, que integra dados sobre cargos, salários, custos, histórico dos pagamentos e estatísticas e conta com 12 usuários.

Os benefícios do sistema, de acordo com a Rhodia, são aumento de produtividade na análise de dados, com a melhoria de processos e identificação de oportunidades. A empresa diz que o tempo gasto para coletar e organizar informações foi reduzido e, em alguns casos, o que era feito em dois ou três dias passou a ser realizado em alguns minutos.

O DW Brasil está sendo expandido horizontal e verticalmente, com a adição de dados de novas unidades de negócios e áreas internas. Neste momento, a Rhodia planeja a implementação global da metodologia *balanced scorecard* para 2005. A adoção do BSC deve acompanhar o projeto global para integração única do sistema de gestão da SAP, que deverá ser concluído em 2007.

FONTE: Computerworld, 8/4/2005. Disponível em:< www.computerworld.com.br>

DATA WAREHOUSE

Conceito:

“O DW é uma coleção de dados, orientado a assuntos, integrados, não volátil, com históricos, organizados sob o prisma da decisão. Bill Inmom - 1992”

“Conjunto de técnicas que aplicadas geram um sistema de dados que proporcionam informações para tomada de decisões”

Objetivo:

- Transformar um Sistema de Produção (OLTP – on line transaction processing) em um Sistema de Decisão (OLAP – on line analytical processing).
- Transformar dados de produção em informações estratégicas.
Exemplo: Evolução do mercado x produtos.
 Prospecção de clientes, segmentação de clientes.
 Perfil dos funcionários mais performáticos.

Função:

- Recuperar, integrar e analisar dados de diversas fontes.
- Oferecer uma visão transversal dos vários bancos de dados de uma empresa.
- Os dados devem ser: extraídos, transformados, “limpos”, relacionados e agregados.
- Armazenar dados históricos.

Acesso:

- Para quem decide.
- Através de ferramentas que possibilitem: agrupar, analisar, integrar e minerar. Os próprios usuários criam suas consultas.
- Acesso à informação: rápida e intuitiva.
Exemplo: visualização multidimensional dos dados.
- Permitir acessos por análise, queries ad hoc, sem necessidade de profissional especializado e sem desenvolvimento de novos sistemas de informação.



Sistema de Produção (OLTP – on line transaction processing) x Sistema de Decisão (OLAP – on line analytical processing

OLTP – On Line Transaction Processing

- Processamentos que executam operações do dia a dia da organização.
- Ênfase ao suporte do negócio, através de um processamento rápido, acurado e eficiente de dados.

Exemplo: movimento bancário.
crédito.
débito.

OLTP – On Line Transaction Processing

Exemplo:

- Entrada de pedidos de vendas.
- Emissão de notas fiscais.
- Requisição de material.

**Tanto os dados de entrada quanto as formas de processamento
são conhecidos.**

OLTP – On Line Transaction Processing

Venda - Gerenciamento de Vendas, Estoque, Clientes e Financeiro

Entrada Rápida de Pedidos de Venda [Incluir Novo Pedido](#)

Informações do Cliente:

Nome do Cliente [Incluir](#) [Alterar](#) Data Entrega: 08/09/2005

Endereço Entrega: Número: Complemento:

Cidade: Estado: C.E.P.:

Observações do Pedido:

Forma de Pagamento

Total Pedido (Bruto): 20,00 Forma de Pagamento: [Incluir](#) [Alterar](#) Boletim Unibanco

Desconto (%): Juros (%): Despesas Cobrança: 0,00 0,00 3,40 ☐ Cobra despesas?

Forma de Entrega

Forma de Entrega: [Incluir](#) [Alterar](#) Região de Entrega: Território Nacional

Volume Total: 0,000 Kg(s) Despesas Para Entrega: 0,00 Quantidade Total: 1 ☐ Cobra despesas Entrega?

Vendedor

[Incluir](#) [Alterar](#) Edsonia Brazil (Proprio)

Total Pedido

20,00

[Voltar](#) [Cancela Pedido](#) [Confirma Pedido](#) [Gera Evento](#)

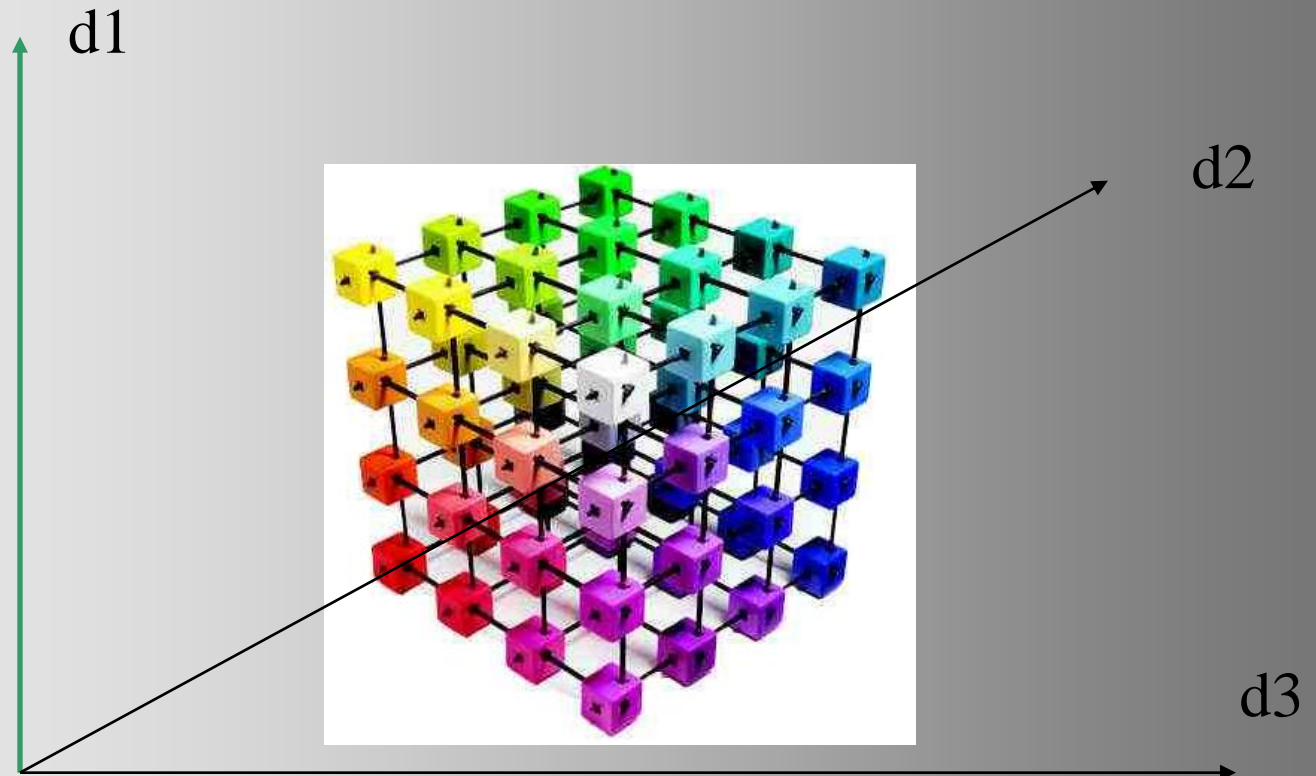
OLTP: dados em formato plano, normalizado

Id Cliente	Nome Cliente	Endereço	Estado	Cidade
01	José Silva	Rua X	SP	São Paulo
02	Marcos Souza	Rua Y	RS	Gravataí
03	Rosa Chaves	Rua Z	SP	Campinas
04	Alexandre Martins	Rua H	RS	Porto Alegre

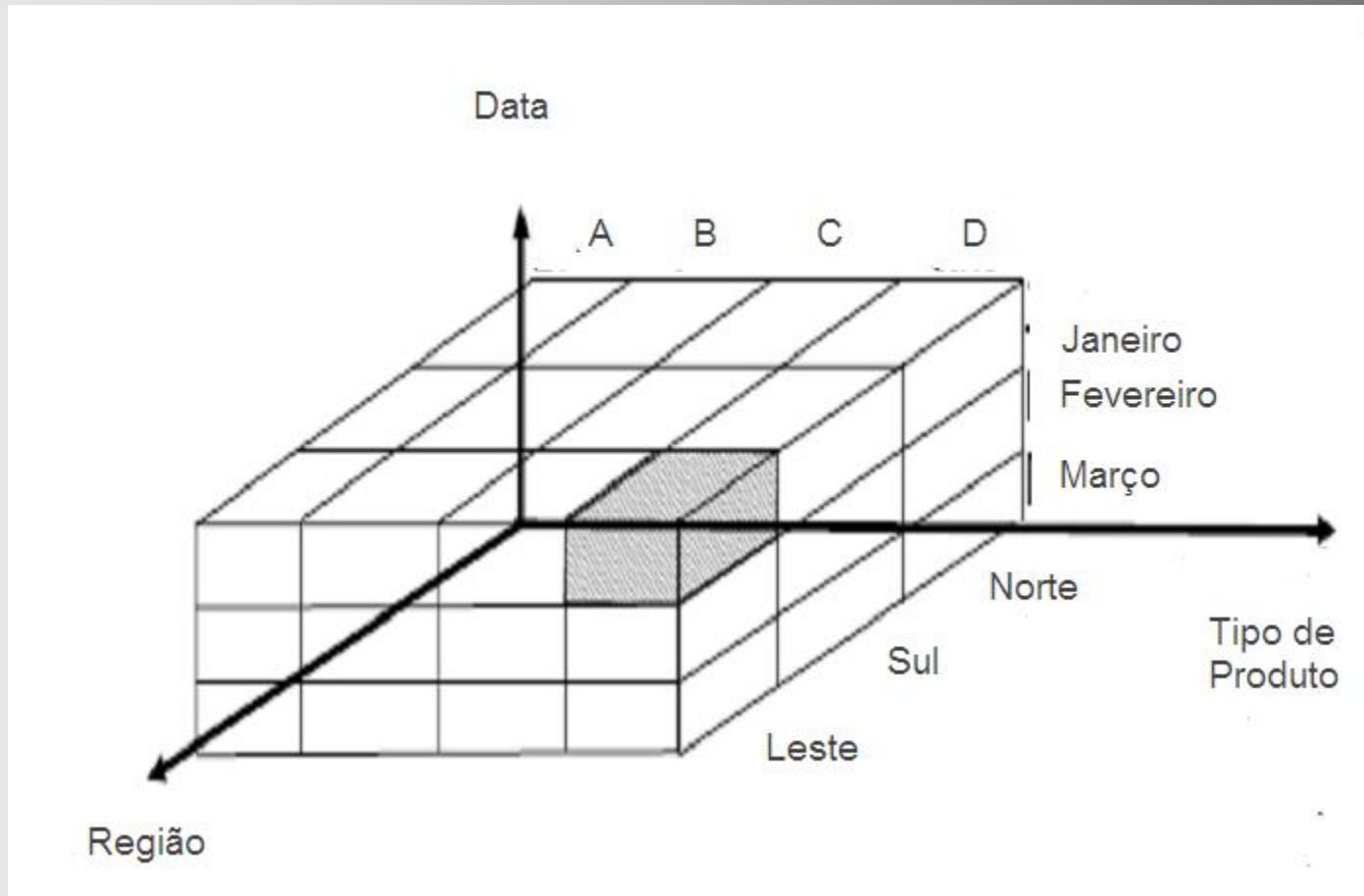
OLAP – On Line Analytical Processing

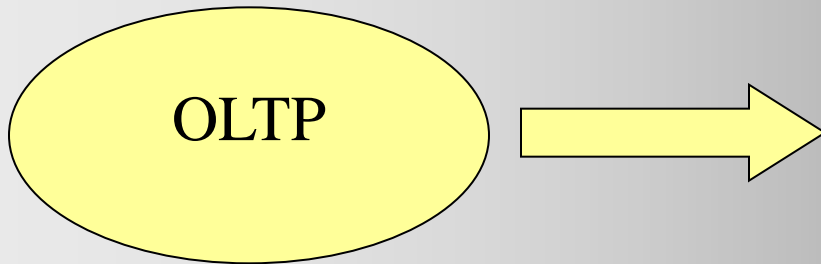
- Processamentos para suportar a tomada de decisões.
- Permitem analisar tendências e padrões em grande quantidade de dados.

OLAP: dados em formato multidimensional

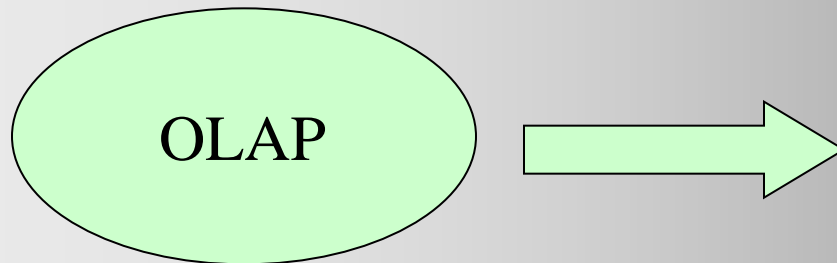


OLAP: formato multidimensional





Solicitações simples:
Quem, a quem..?
Exemplo: vendas de X, qual
o preço?
de qual fornecedor ?



Possibilita agregar dados
em várias dimensões:
Número de vendas por
fornecedor, por mês, por ano,
por região, por produto...

Porque OLTP e OLAP?

BD OLTP não são compatíveis para suportar visões OLAP

- Não possuem os mesmos objetivos.
- Formatos do dado: as tabelas apresentam uma visão plana.
- Tipo de requisição.
- Uma estrutura para o OLAP deve suportar:
 - Representação em múltiplas dimensões.
 - Fácil manipulação.

Porque OLTP e OLAP?

Porque separar:

- Desempenho: requisições OLAP podem degradar seriamente o ambiente de produção.
- Tipo do dado: histórico, agregado, dados “limpos, qualificados”, de diversas fontes.

Porque OLTP e OLAP?

Porque separar:

Características	OLTP	OLAP
Aplicação	produção	suporte à Decisão
Usuário	profissional TI, Clientes, operacional da empresa	quem decide
Acesso	simples, numeroso, não complexo, regulares, repetitivo	agrupado, não numeroso, imprevisível, irregulares
Número de dados	dezenas	milhões
Tamnahó	GB	TB
Idade dos dados	recentes, do dia a dia	histórico

“O DW é uma coleção de dados, orientado a assuntos, integrados, não volátil, com históricos, organizados sob o prisma da decisão. Bill Inmom - 1992”

Bill Inmon...

Orientado a assuntos

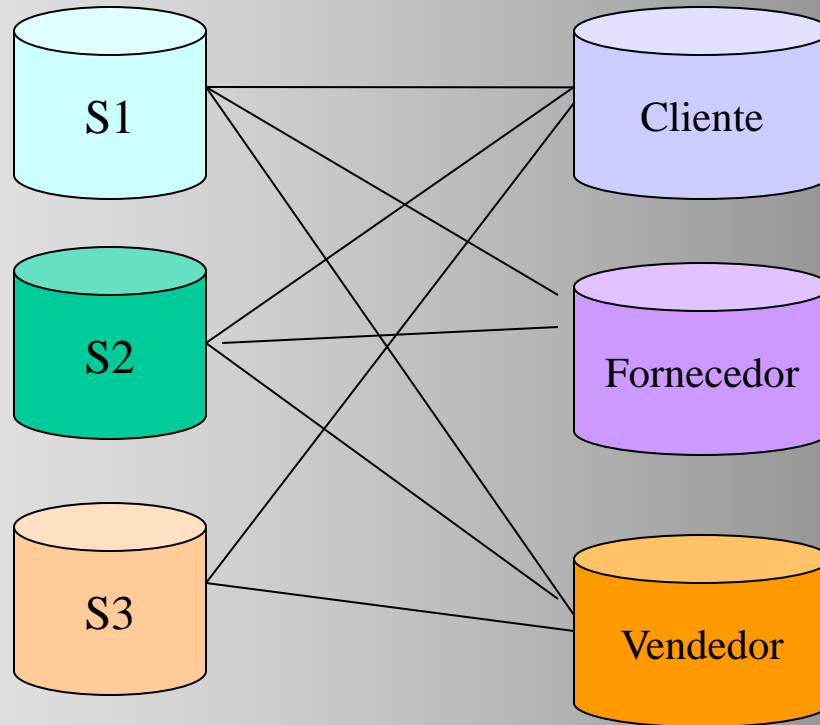
- Armazena informações sobre temas específicos importantes para o negócio da empresa. Requer dados de muitas fontes para compor uma visão corporativa do assunto.

Exemplo: produto, clientes, fornecedores.

- Ambiente operacional (OLTP): organizado por aplicações funcionais.

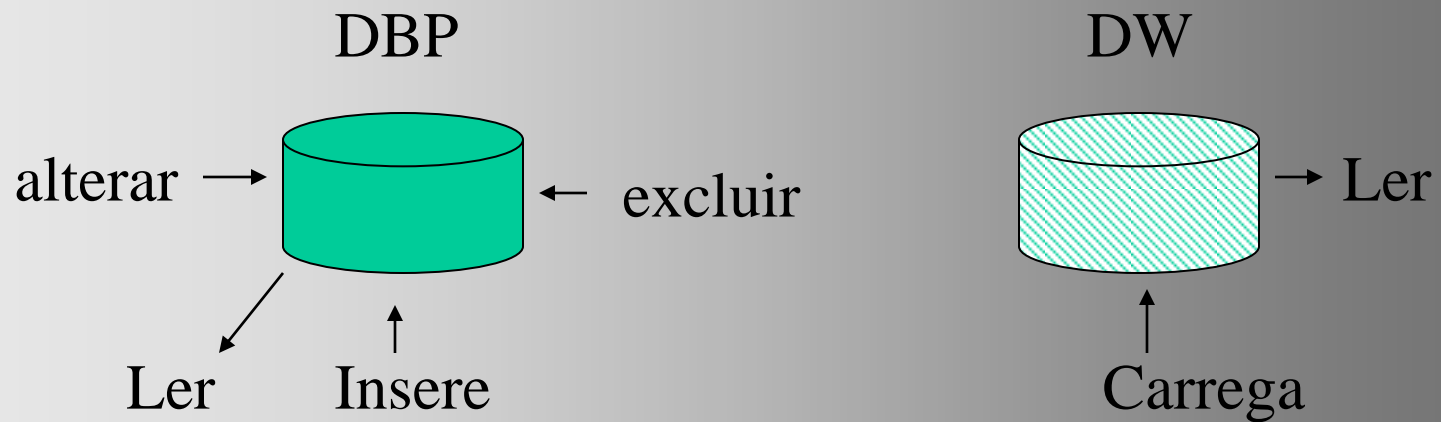
Exemplo: sistema bancário: empréstimo, investimento, seguros.

Orientado a assuntos



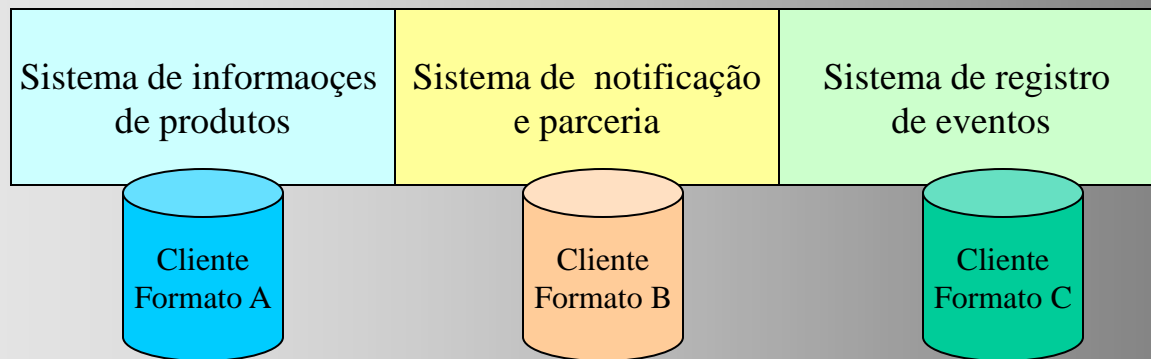
Não volátil

- O usuário apenas lê a informação. Os dados do DW não são excluídos.



Integrado

- Integra várias fontes de dados.



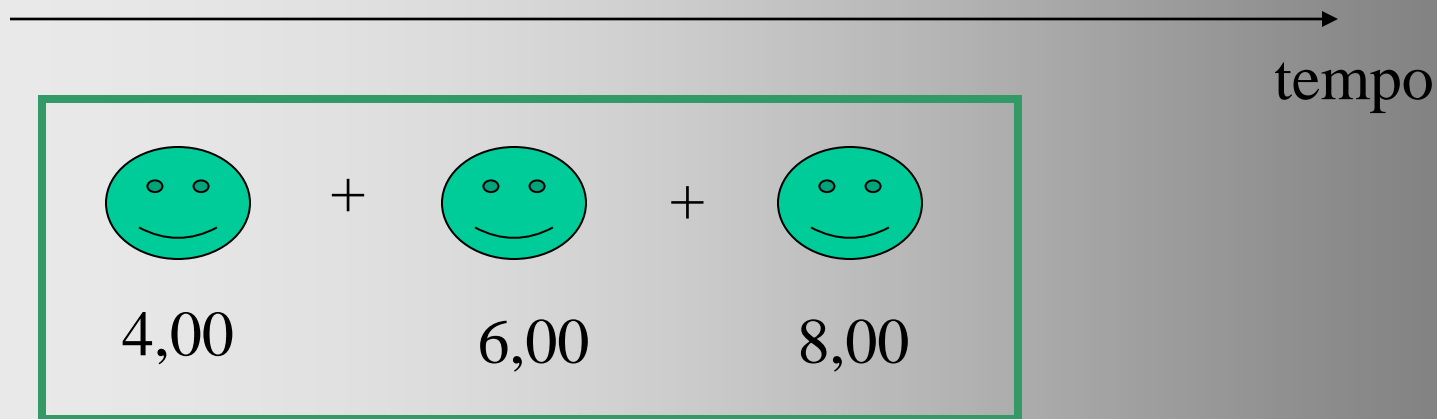
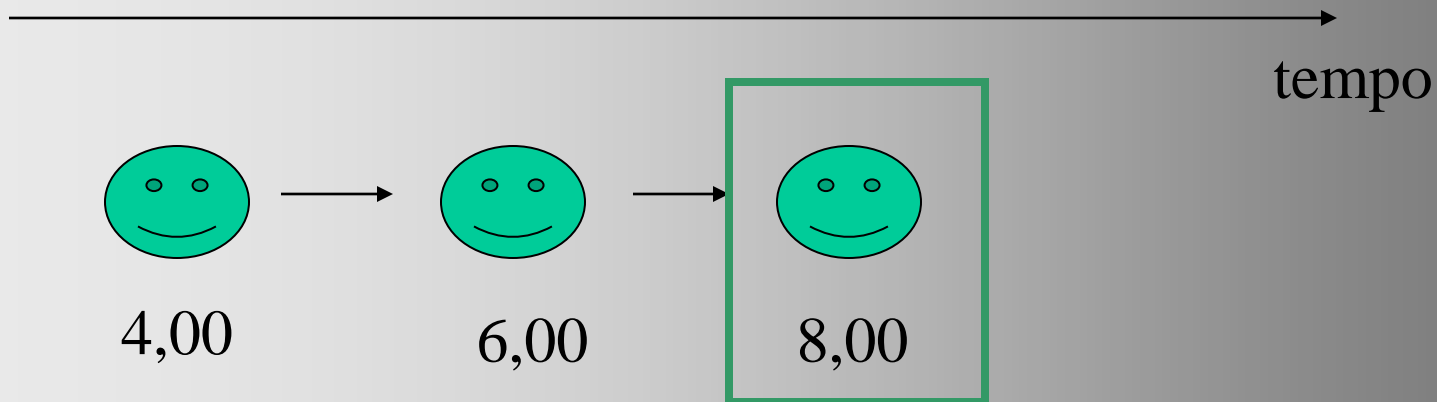
Visão unificada do cliente

Variante no tempo

- Não é atualizado. Está relacionado a um momento específico.
Exemplo: valor de um produto.
- No ambiente OLTP, o dado é atualizado de acordo com mudanças de estado do objeto, refletindo em geral o estado do objeto no momento do acesso.

Histórico

- Armazena a evolução dos indicadores.

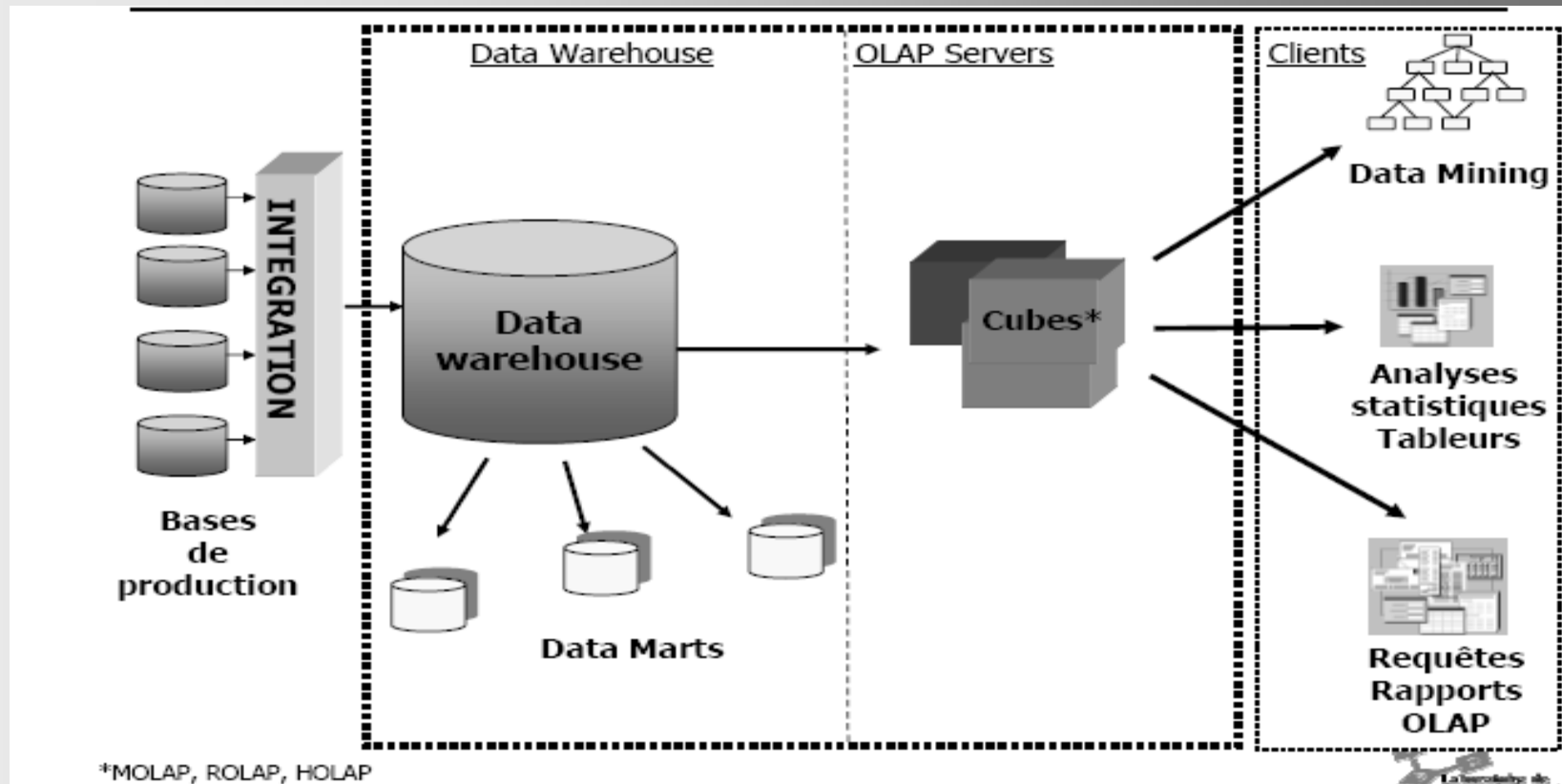


Os dados

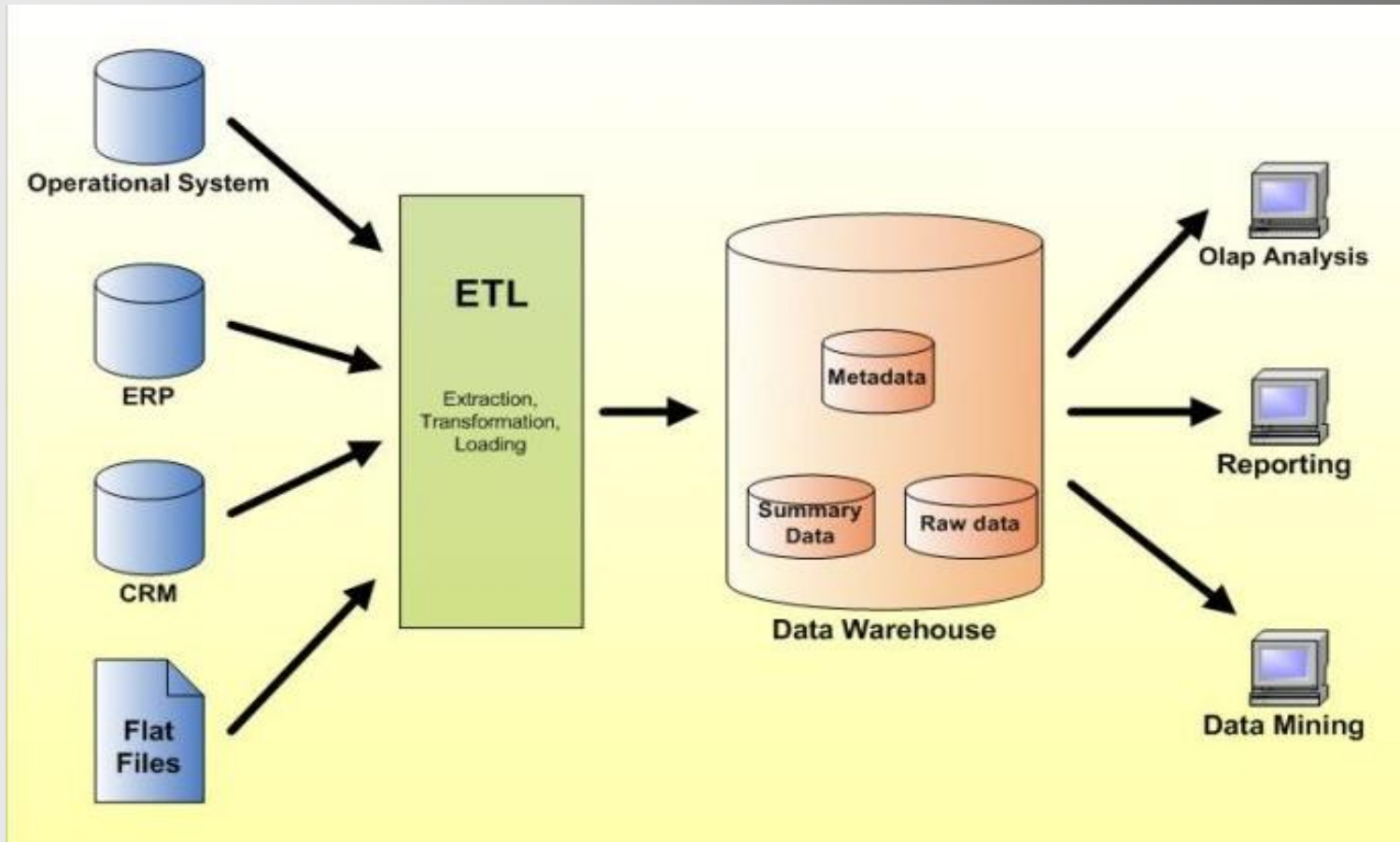
- O volume é muito importante.
- Origem: dados dispersos, às vezes de difícil acesso. Mal integrados. Na origem , não estão estruturados para uma aplicação de decisão.
- Sumarizados, agregados.

Por exemplo, um sistema de controle de estoques armazena todas as transações realizadas. Uma possível agregação para estes dados é a movimentação total para um determinado produto, em uma semana ou em um mês. Esta operação também é chamada de sumarização e tem o objetivo de agregar grandes volumes de dados em informações significativas.

Arquitetura Básica



Arquitetura Básica



Arquitetura

Visão em camadas

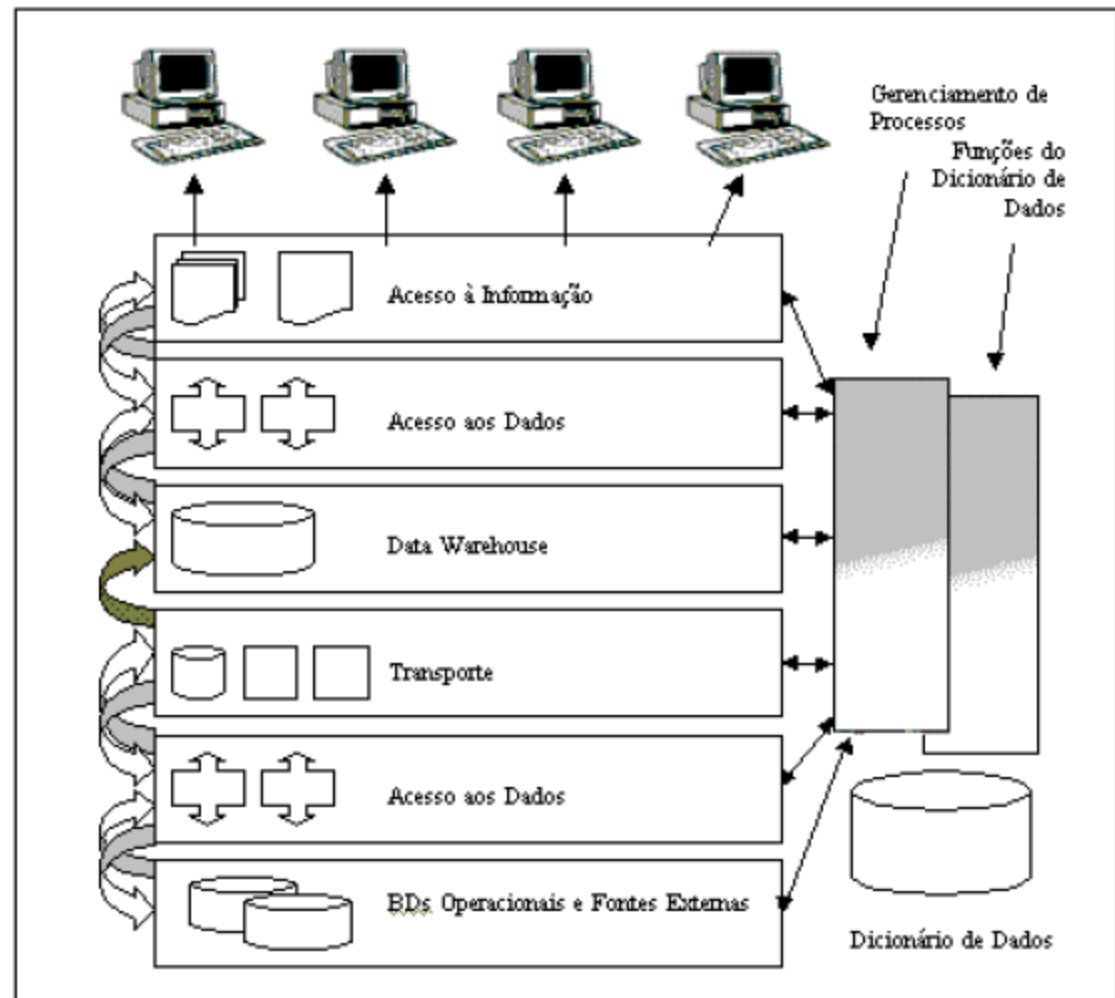


Figura 2 Arquitetura em camadas.



Camada de Acesso aos Dados: compõe o elo de ligação entre as ferramentas de acesso à informação e os bancos de dados operacionais, comunicando-se com diversos Sistemas de Gerenciamento de Banco de Dados (SGBDs) e sistemas de arquivos, sendo que a este conjunto de características dá-se o nome de acesso universal de dados;



Camada de Transporte ou *Middleware*: tem a função de gerenciar a transmissão das informações pelo ambiente de rede que serve de suporte para o sistema como um todo, separando as aplicações operacionais do formato real dos dados, realiza ainda a coleta de mensagens e transações e se encarrega de entregá-las nos locais e nos tempos determinados;

✚ Camada de Metadados (Dicionário de Dados): os metadados descrevem os dados e a organização do sistema, podem ser ainda fórmulas utilizadas para cálculo, descrições das tabelas disponíveis aos usuários, descrições dos campos das tabelas, permissões de acesso, informações sobre os administradores do sistema, entre outras;

✚ Camada de Gerenciamento de Processos: faz o controle destas tarefas que mantêm o sistema atualizado e consistente, gerenciando as diversas tarefas que são realizadas durante a construção e a manutenção dos componentes de um sistema de DW;

✚ Camada de Gerenciamento de Replicação: serve para selecionar, editar, resumir, combinar e carregar no DW as informações a partir das bases operacionais e das fontes externas, envolvendo programação bastante complexa, sendo que existem ferramentas poderosas que permitem que estes processos sejam gerenciados de forma mais amigável, além do controle da qualidade dos dados que serão carregados.

Arquitetura - Bill e Inmon

Componentes da Arquitetura

Bill Inmon

- Camada de Transformação/Integração
- Data warehouse corporativo
- ODS corrente
- Data Mart
- Exploration DW
- Metadado

Ralph Kimball

- Data staging area
- Coleção de DM's = DW
- ODS histórico
- Data Mart – papel diferente
- Metadado

Kimball vs Inmon in data warehouse architecture

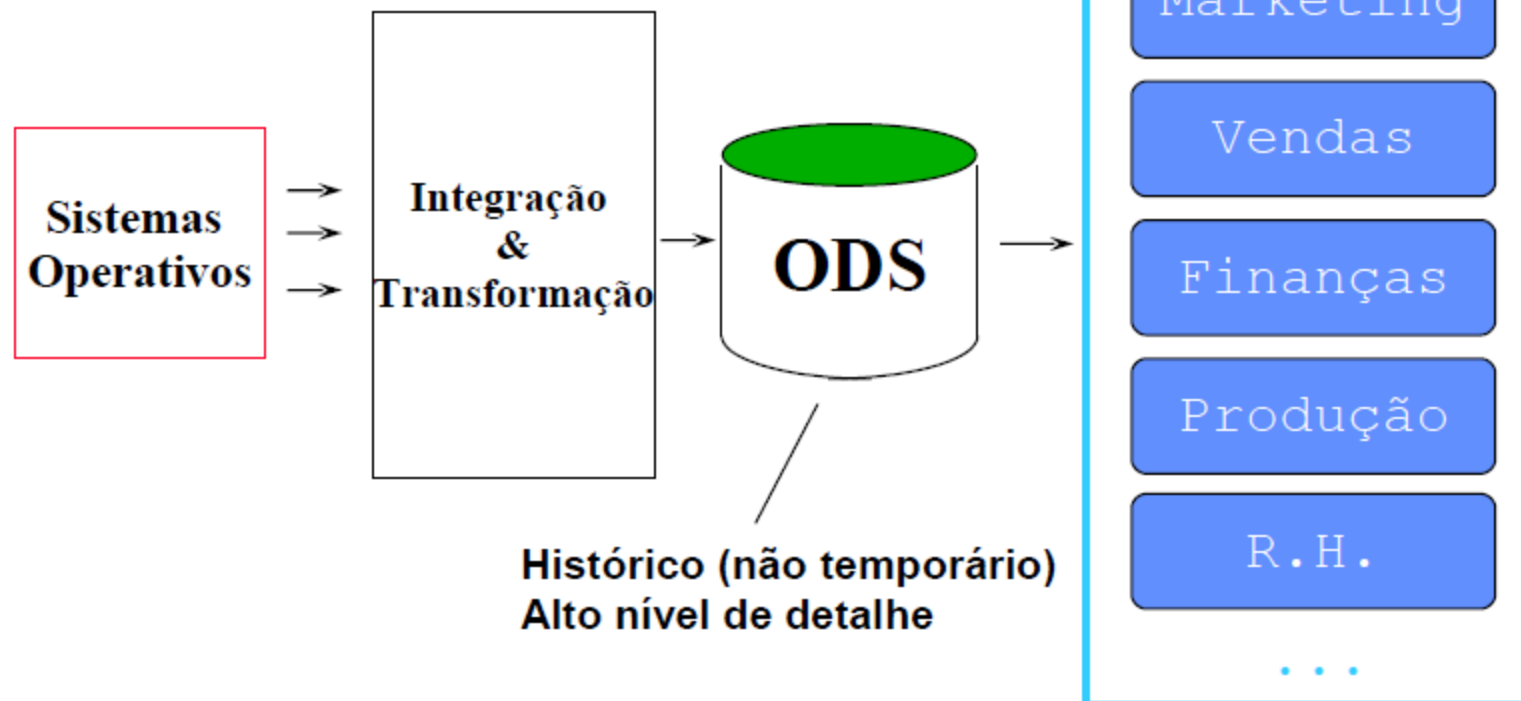
Both Kimball and Inmon's architectures share a same common feature that each has a single integrated repository of atomic data. In Inmon's architecture, it is called *enterprise data warehouse*. And in Kimball's architecture, it is known as *dimensional data warehouse*. Both architectures have an enterprise focus that supports information analysis across the organization. This approach enables to address the business requirements not only within a subject area but also across subject areas.

However there are some differences in the data warehouse architectures of both experts:

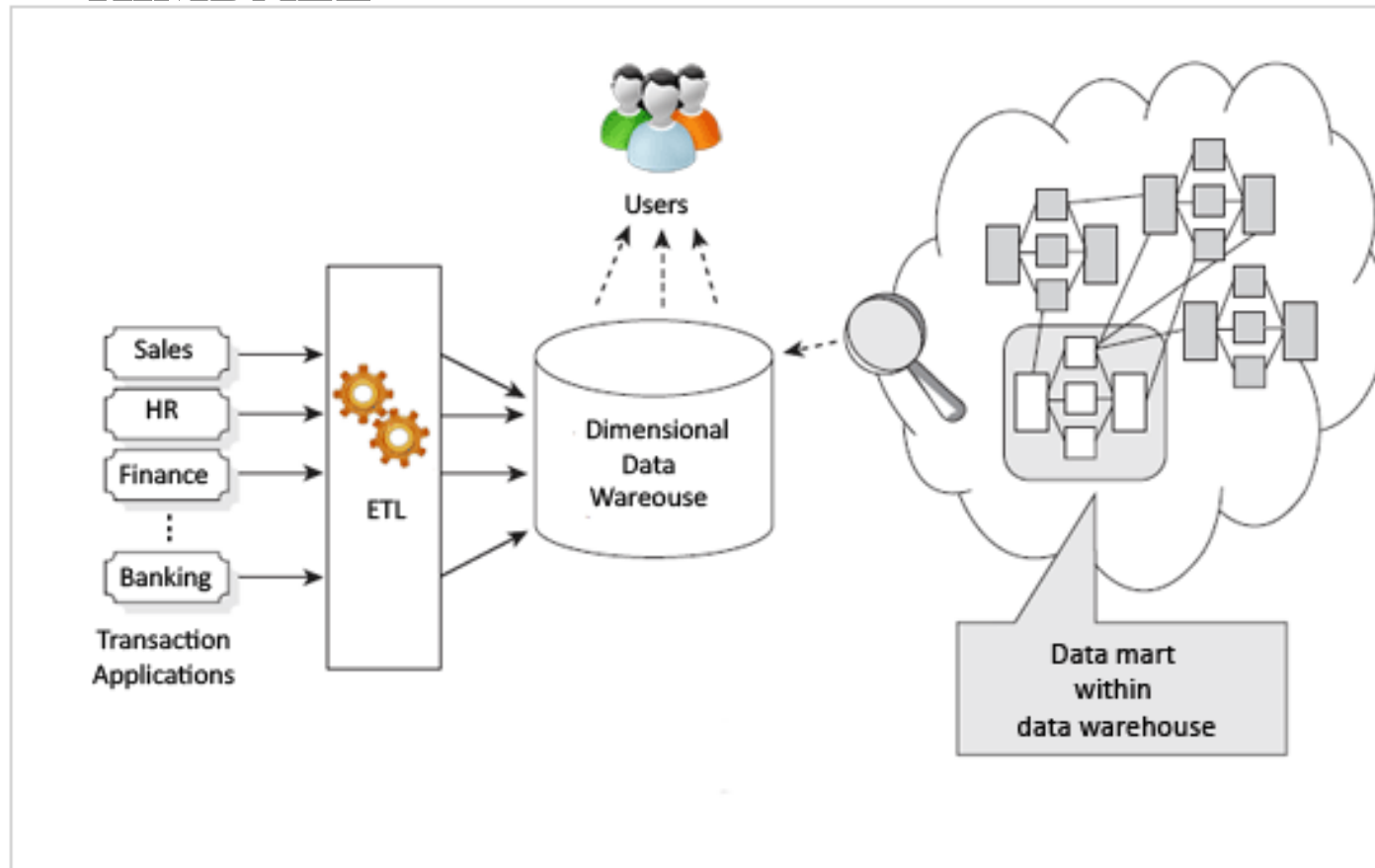
- ▶ Kimball uses dimensional model such as star schemas or snowflakes to organize the data in *dimensional data warehouse* while Inmon uses ER model in enterprise data warehouse. Inmon only uses dimensional model for data marts only while Kimball uses it for all data
- ▶ Inmon uses data marts as physical separation from enterprise data warehouse and they are built for departmental uses. While in Kimball's architecture, it is unnecessary to separate the data marts from the dimensional data warehouse.
- ▶ In dimensional data warehouse of Kimball, analytic systems can access data directly. While in Inmon's architecture, analytic systems can only access data in enterprise data warehouse via data marts.

Arquitetura de Data Warehouse de Kimball

DW = Operational Data Store + Data Marts Integrados



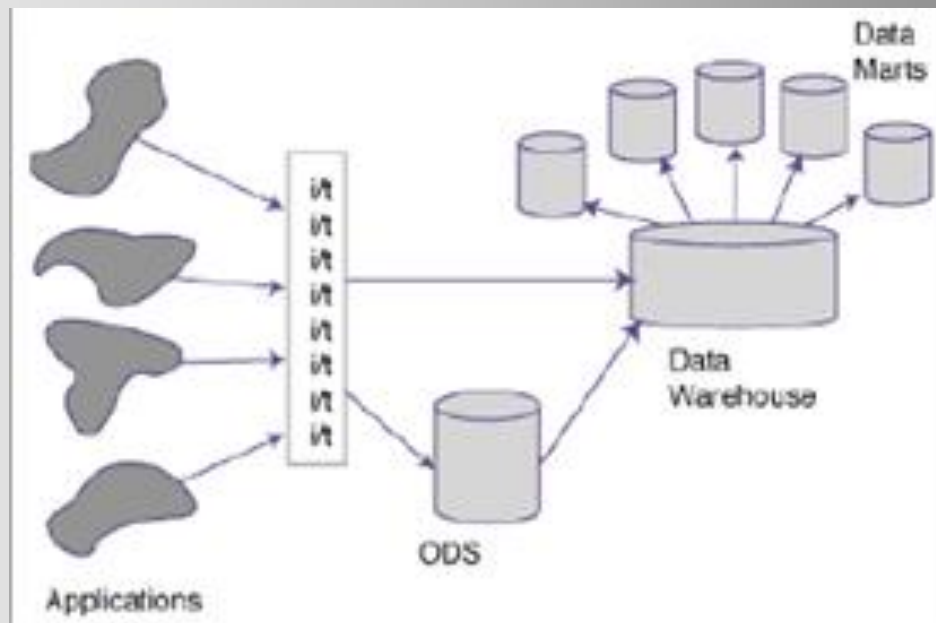
KIMBALL



Ralph Kimball Data Warehouse Architecture

- Aplicações transacionais (sistemas fontes) – sistemas operacionais transacionais. Estruturados em SGBD's relacionais ou flat files.
- ETL – processos que extraem, transformam e carregam os dados dos sistemas transacionais no Data Warehouse.
- Datamart - é apenas uma distinção lógica dentro do Data Warehouse dimensional. Apresenta do dado relacionado a um processo e negócio que permeiam áreas funcionais.
- Data Warehouse Dimensional – contém dados corporativos em formato granular. As ferramentas analíticas acessam esse banco. **Modelo dimensional granular.**

INMON



- Aplicações corporativas - sistemas de produção, transacionais que apoiam o negocio.
- ETL - processos para extrair, transformar e carregar dados.
- Enterprise Data Warehouse - elemento central da arquitetura de Inmom. Dados atômicos integrados. Modelo de dados relacional em 3FN (terceira forma normal).
- Data Mart - visões departamentais, orientado a assunto. Modelo Dimensional.
- Operational Data Store - estrutura híbrida com características de Data Warehouse e transacional. Integra sistemas fontes e fornece uma base de dados tática para sistemas de suporte à decisão. É atualizado. orientado a assunto, integrado (várias aplicações), volátil. Dados corporativos Deve ter alto desempenho e suporta atualizações on-line.
- Ferramentas analíticas extraem dados dos datamarts.

Data Marts de Inmon

Data Mart

- OLAP
- Multidimensional



Estruturas
Star ou Snowflake

Para Inmon, Data Marts são depósitos secundários derivados do Data Warehouse.

Onde Inmon e Kimball concordam

- **É necessária uma arquitetura**
- **Dimensões compartilhadas e medidas definidas através de todas as áreas da empresa**
- **O esquema estrela é útil para apresentar informações aos usuários**
- **Construir o DW iterativamente (mesmo que projetado globalmente)**
- **Metadado é fundamental**

Papel do Data Mart

Bill Inmon afirma:

- Populado pelo DW e dados externos
- Dados agregados ou atômicos
- Pode estar em plataforma diferente do DW

Ralph Kimball afirma:

- Formam o DW
- Pode ser atômico
- Dados de um único processo de negócio

Modelo ER x Modelo Dimensional

Modelo ER/Relacional: complexo, 3FN, dificuldade na junção, mais especialização para manipular.

Modelo dimensional: mais intuitivo. Tabelas fatos e de dimensão. Junção entre fatos e dimensão. Podem estar não normalizadas.

Modelo ER x Modelo Dimensional

Bill Inmon afirma:

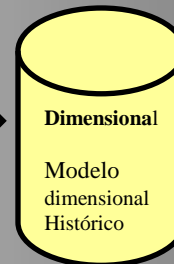
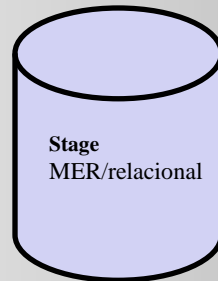
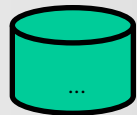
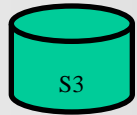
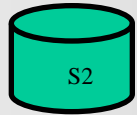
- Modelo dimensional é razoável para os DMs
- Modelo Dimensional não é flexível para o DW
- Dado normalizado é ideal para o DW e ODS. MER é consistente e flexível

Ralph Kimball afirma:

- Modelo dimensional - DM e DW
- Dados granulares
- Todo ER tem um DM equivalente
- Qualquer negócio pode ser representado como um cubo
- MER é complicado para usuário final
- MER viável para ODS e ambiente transacional.

Em geral:

**Ambiente de Produção,
transacional, operacional**



Usuários

Consultas com ferramentas
olap, data mining ...



Abordagens de Implantação do DW

Implementação Bottom Up:

Teoricamente, por ser mais fácil e menos custosa de ser implementada do que a Top Down, é muito usada, até pelo retorno rápido de investimento. Ela permite que o planejamento e a construção dos Data Mart's possam ser realizados antes mesmo que a infraestrutura do DW (Data Warehouse) seja definida. O propósito desta implementação é a construção de DW incremental a partir do desenvolvimento de DM independentes.

O processo para construção começa com o ETL para um ou mais DM's, que são modelados com base em um modelo dimensional. Um ponto negativo para tal implementação é a falta de um gerenciador que garanta padrões únicos de metadados, mesmo com a independência dos Data Marts, além de se tornar crítico na interferência junto aos sistemas transacionais, pelo processo de ETL.

Vantagens:

Implementação rápida: a construção dos DM é altamente direcionada, permitindo um rápido desenvolvimento. Normalmente, pode ser colocado em produção em um período de 6 a 9 meses.

Retorno rápido: a arquitetura baseada em DM com incremento demonstra rapidamente seu valor, permitindo uma base para investimentos adicionais com um nível mais elevado de confiança.

Manutenção do enfoque da equipe: a elaboração de DM's permite que os principais negócios sejam enfocados inicialmente, sem que haja gastos no desenvolvimento de áreas que não são essenciais ao problema.

Herança incremental: a estratégia de DM's incrementais obriga a entrega de recursos de informação passo a passo, o que permite a equipe crescer e aprender, reduzindo os riscos etc.

Desvantagens:

Perigo de legamarts: um dos maiores perigos da Administração de DW é a criação de DM's independentes,. O advento das ferramentas de drag and drop facilitou o desenvolvimento de soluções individuais que podem não considerar a arquitetura de forma global. Desse modo, os DM's independentes transformam-se em DM legados ou legamarts.

Desafio de possuir a visão de empreendimento: durante a construção dos DM's incrementais é necessário que se mantenha um rígido controle do negócio, como um todo. Esse controle requer um maior trabalho ao extrair e combinar as fontes individuais do que utilizar um DW.

Administrar e coordenar múltiplas equipes e iniciativas: normalmente esse tipo de arquitetura emprega o desenvolvimento de DM's em paralelo, no entanto, isso pode conduzir a uma rígida administração, tentando coordenar os esforços e recursos das múltiplas equipes, especialmente nas áreas de regras e semântica empresariais.

A maldição de sucesso: a arquitetura com DM's incrementais carrega a “maldição de sucesso”. Nesses casos, os usuários finais dos DM's encontram-se felizes querendo mais informação para seus DM's e ao mesmo tempo outros usuários de outros DM's aguardam o incremento de seus DM's. Isso conduz a equipe de DM a vencer desafios de recurso e de administração.

Implementação Top Down: contrário do Bottom Up

Vantagens:

Herança de arquitetura: todos os DM originados de um DW utilizam a arquitetura e os dados desse DW, permitindo uma fácil implementação.

Visão de empreendimento: o DW concentra todos os negócios da empresa, sendo possível extrair dele níveis menores de informações.

Repositório de metadados centralizado e simples: o DW provê de um repositório de metadados central para o sistema. Essa centralização permite manutenções mais simples do que aquelas realizadas em múltiplos repositórios.

Controle e centralização de regras: a arquitetura top down garante a existência de um único conjunto de aplicações para extração, limpeza e integração dos dados, além de processos centralizados de manutenção e monitoração.

Desvantagens:

Implementação muito longa: os DW são, normalmente, desenvolvidos de modo iterativo, por áreas de assuntos, como por exemplo, vendas, finanças e recursos humanos. Mesmo assim, são necessários, em média 15 ou mais meses para que área de assunto entre em produção, dificultando a garantia de apoio político e orçamentário.

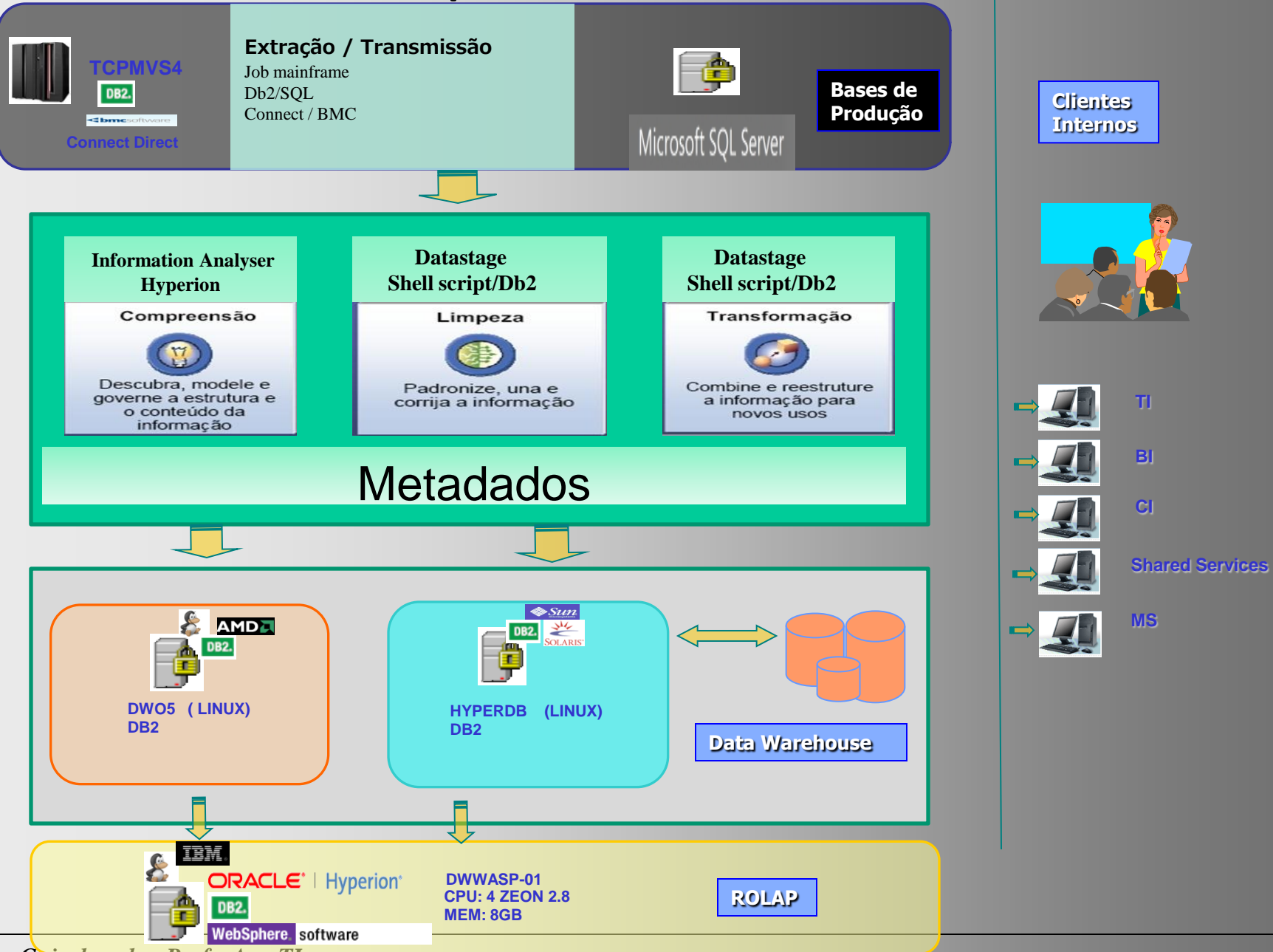
Alta taxa de risco: não existem garantias para o investimento nesse tipo de ambiente.

Heranças de cruzamentos funcionais: é necessária uma equipe de desenvolvedores e usuários finais altamente capacitados para avaliar as informações e consultas que garantam à empresa habilidade para sobreviver e prosperar na arena de mudanças de competições políticas, geográficas e organizacionais.

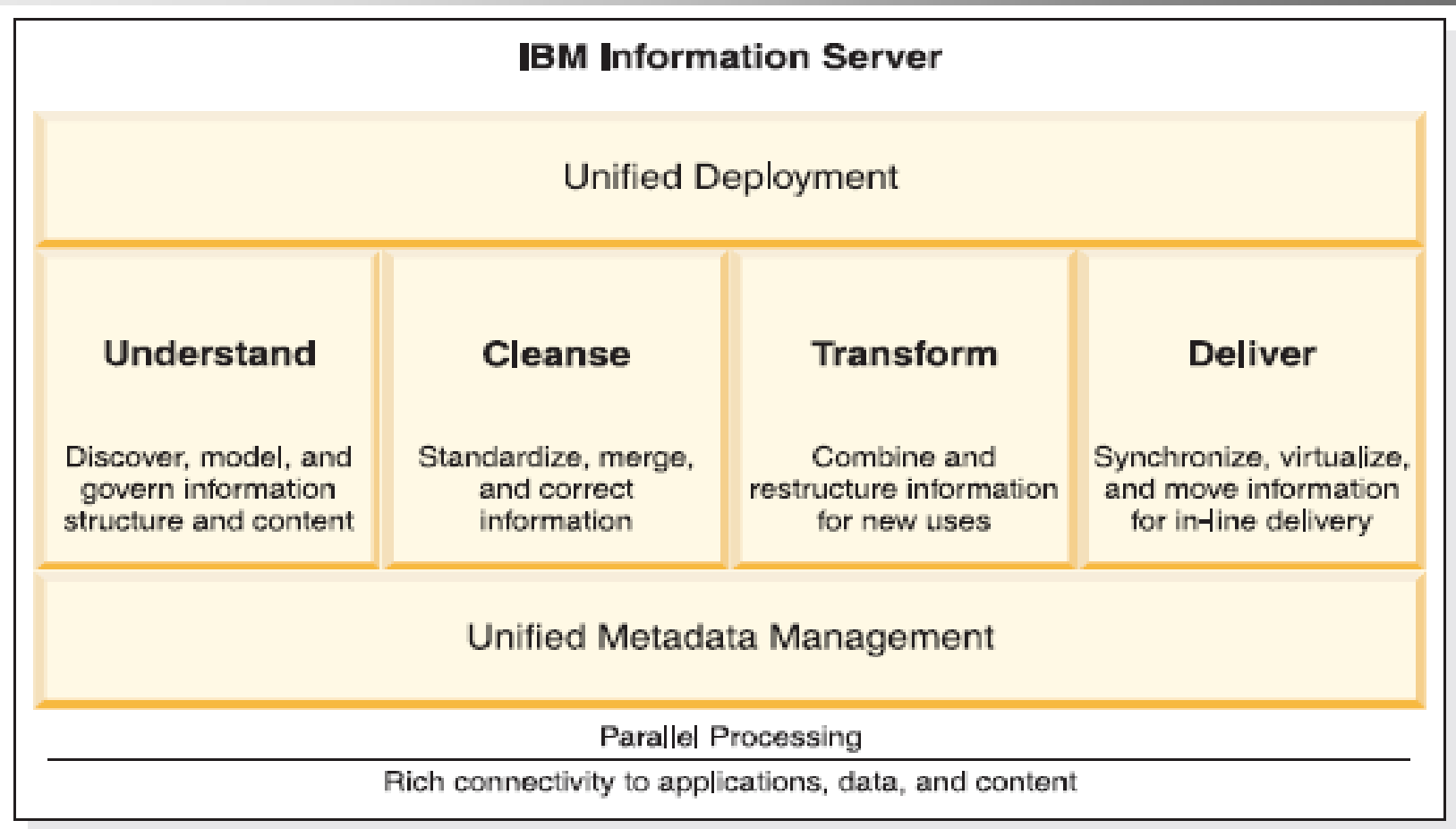
Expectativas relacionadas ao ambiente: a demora do projeto e a falta de retorno podem induzir expectativas nos usuários.

Exemplo

Exemplo: ETL e Camada de acesso informacional



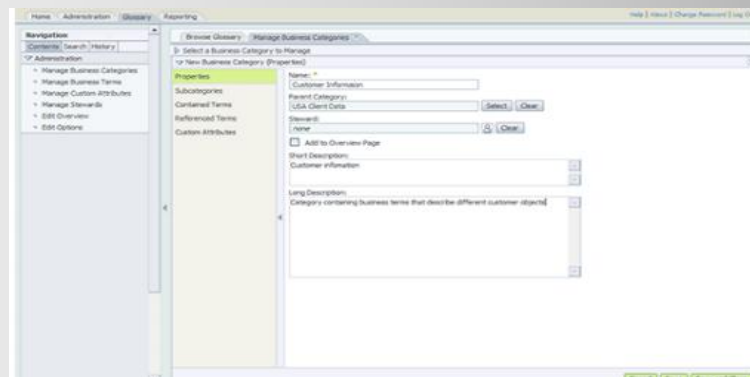
Exemplo:



Metadata services

WebSphere Business Glossary

WebSphere Business Glossary is a Web-based application that provides a business-oriented view into the data integration environment. By using WebSphere Business Glossary, you can view and update business descriptions and access technical metadata.



Dicionário de dados corporativo
Vocabulário organizacional
Regras corporativas
Alinhamento de TI e negócio

Information Analyzer

Profiling

Análise de grande volumes de dados

Data monitoring qualitativa - compara análises

Análise de estruturas

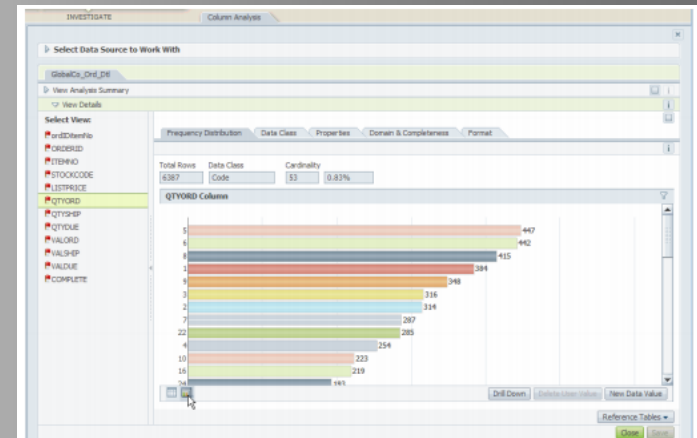


Figure 31. Column analysis example graphical view

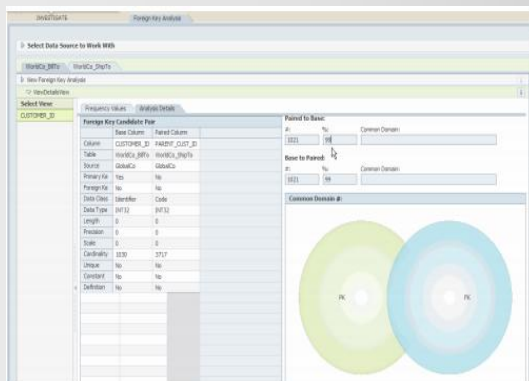


Figure 33. Foreign key analysis

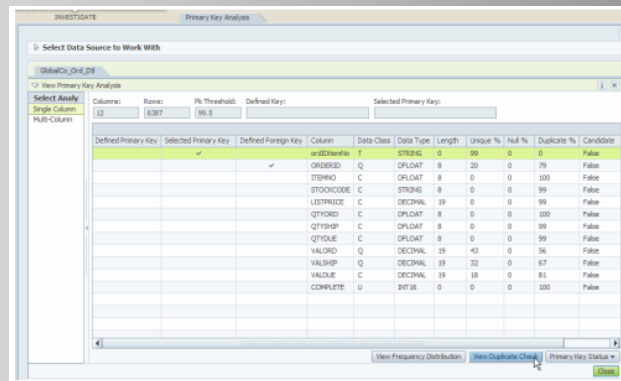


Figure 32. Primary key analysis

Exemplo:

DataStage

O IBM InfoSphere DataStage - módulo de produto principal da plataforma de integração de informações do InfoSphere Information Server.

Permite coletar, integrar e transformar dados de uma ampla variedade de origens nos data warehouses e em outros aplicativos para informações confiáveis e na hora certa.

Processa volumes de dados massivos em lotes ou em tempo real, conectando-se com um número ilimitado virtualmente de origens de dados heterogêneos em uma única tarefa.

DataStage

