

UM ESQUEMA DE DATA WAREHOUSE INTEGRADO COM VISUALIZAÇÕES EM DASHBOARDS

L.A. Silva^{1,*}; G.C. Bruschi¹; A.M. Mantovani¹; M.R. Brochado¹

1 Faculdade de Tecnologia de Bauru (FATEC)
R. Manoel Bento da Cruz, 3-30 - Bauru/SP, CEP.: 17015-171, Brasil.
Telefone: (14) 3223-2083

*luis.silva51@fatec.sp.gov.br

RESUMO: Atualmente, o mercado tem exigido cada vez mais que as empresas busquem diferenciais para sair à frente de seus concorrentes, demonstrando que a competitividade de uma empresa está intimamente ligada à sua capacidade de adquirir dados de diferentes fontes e informações e interpretá-las de forma rápida e eficiente, além de apresentar estas informações de maneira simples e de fácil compreensão para os gestores. Para buscar esta eficiência na recuperação de dados este artigo tem como objetivo demonstrar a importação de dados de uma base de dados relacional para um modelo multidimensional DW, e também a praticidade de apresentação de resultados em painéis de informações em forma de dashboards.

PALAVRAS-CHAVE: banco de dados; data warehouse; painéis de informação.

ABSTRACT: Currently, the market has increasingly demanded companies to seek differentials to come out ahead of its competitors, demonstrating that the competitiveness of a company is closely related to its ability to acquire data from different sources and information and interpret them quickly and efficiently, in addition to present this information in a simple and easy way to interpret for managers. To pursue this efficiency in data recovery, the concept of Data Warehouse emerged. This paper aims to demonstrate the possibility of importing data from a relational database to a multidimensional DW model and the practicality of presenting the results in dashboards.

KEYWORDS: database; data warehouse; dashboards.

1. INTRODUÇÃO

A competitividade de uma empresa é diretamente proporcional à sua capacidade de obter informações, processá-las e disponibilizá-las de forma rápida, segura e com qualidade. A aplicação de um *Data Warehouse* (DW) melhora a qualidade da apresentação das informações facilitando sua interpretação pelos gestores, podendo assim gerar um diferencial nas tomadas de decisão [1].

Fortulan e Gonçalves Filho [2] demonstraram a implantação de um sistema de Inteligência de Negócios, do inglês *Business Intelligence* (BI), com armazenagem em DW, para administrar dados gerados nas áreas de produção das fábricas, concluíram que a implantação do sistema deu um bom suporte ao gerente para a tomada de decisões, retornando informações de forma fácil e flexível. Já de acordo com Favaretto [3], as informações devem ser precisas e estarem disponíveis quando necessárias, para isto baseou-se em três itens para a melhoria da qualidade de dados: a acuracidade, a temporalidade e a integridade. Dessa forma, para essas três medidas terem sucesso na implementação do DW, é importante alinhar o projeto com a estratégia da empresa e não somente a um grupo de pessoas.

Este artigo tem como objetivo demonstrar a conversão de dados de uma base de dados relacional para um modelo multidimensional DW. O modelo multidimensional foi elaborado com base na ideia de negócio focado na venda de miniaturas de veículos clássicos, seguindo requisitos para a extração dos dados necessários para um gestor, com base no fato *Vendas* e nas dimensões: *Local*, *Produto*, *Tempo* e *Vendedor*. Devido às limitações do SGBD na forma de exibir as informações, foram utilizados *dashboards* para uma melhor visualização e compreensão.

O restante deste trabalho está organizado como segue. Na Seção 2, é abordado o referencial teórico para melhor compreensão do trabalho. Na Seção 3, estão relatados os materiais e métodos que foram utilizados nos experimentos. Na Seção 4, são demonstrados os resultados e, na Seção 5, as conclusões.

2. REFERENCIAL TEÓRICO

2.1. Dados e Informações

Em um sistema de informação, estão contidas todas as informações necessárias ao objetivo do sistema. Por definição, um computador não processa informações, mas sim, dados [4], informações são geradas a todo o momento e em situações distintas, devem ser guardadas para futuras decisões. Pode-se dizer que dados são a matéria prima da informação, ou seja a informação não tratada, enquanto informações são os dados tratados, que após passar por processamentos retornam informações com algum significado, podendo ser utilizados para tomada de decisões [5]. Normalmente esses dados são concentrados em um sistema de banco de dados que é basicamente um sistema computadorizado de manutenção de registros que permite que os usuários busquem e atualizem essas informações quando as solicitarem [6].

O *Oracle* é um Sistema Gerenciador de Banco de Dados (SGBD) que implementa o modelo relacional, apresenta recursos que facilitam a criação e manipulação de dados na forma de DWs, sendo este o banco de dados escolhido para representar o modelo DW que será implementado neste artigo.

2.2. Data Warehouse

O DW reúne dados de diferentes fontes, geralmente de bases de dados transacionais existentes ou fontes externas, organizando estes dados de forma que possibilite análises fáceis e geração de relatórios, tem como característica o direcionamento para aplicações de apoio as decisões. São otimizados para a recuperação de dados, e não para processos diários de transações [7].

Para Machado [8], algumas das principais justificativas para aplicação da tecnologia DW em uma organização, é a existência na empresa de constantes alterações nos sistemas transacionais corporativos; dificuldade acentuada na recuperação de dados históricos em períodos superiores ao ano atual de operações; existência de pacotes de sistemas de fornecedores diferentes; falta de padronização e integração dos dados existentes nos diversos sistemas; entre outros.

A diferença entre um Banco de Dados Relacional e um DW não está apenas no que é armazenado, mas também na forma como é armazenado. Os dados em um sistema transacional são armazenados a um nível de normalização mais adequado para inserções rápidas, atualizações e exclusões. No caso de um sistema para análise, é armazenado com uma forma de normalização que seja de rápida leitura. Para a modelagem de um DW, são usados diferentes modelos multidimensionais, entre eles estão o Modelo Floco de Neve, do inglês *Snowflake Schema* e o Modelo Estrela, do inglês *Star Schema*. O modelo de armazenamento multidimensional envolve dois tipos de tabelas: de dimensão e de fatos. Uma tabela de dimensão consiste em tuplas de atributos da dimensão, já uma tabela de fato pode ser pensada como possuindo tuplas, uma por fato registrado. Esse fato contém uma ou algumas variáveis medidas ou observadas e as identifica com

ponteiros para tabelas de dimensão. A tabela de fato contém os dados, e as dimensões identificam cada tupla naqueles dados. O modelo estrela é elaborado na Segunda Forma Normal (2FN) de dados, consistindo em uma tabela de fato com uma tabela para cada dimensão. Já no modelo floco de neve é aplicada a Terceira Forma Normal (3FN), sendo uma variação do modelo estrela no qual as tabelas dimensionais são organizadas em uma hierarquia por meio da normalização. O ponto fraco deste modelo é a perda de desempenho, pois acaba gerando um modelo mais ramificado, exigindo acesso em muitas tabelas [9].

O modelo multidimensional aplicado nos experimentos deste artigo foi o modelo estrela, por ter melhor desempenho em relação ao modelo floco de neve e por ser a estrutura básica de um modelo de dados multidimensional. Sua composição típica possui uma grande entidade central (fato) e um conjunto de entidades menores (dimensões), arranjas ao redor dessa entidade central, formando uma estrela.

2.3. Dashboards

Para Few [10], *dashboards* é uma exibição visual das informações mais importantes necessárias para alcançar um ou mais objetivos, consolidados e agrupados em uma única tela para que todas as informações possam ser monitoradas ao mesmo tempo. Assim como o painel de um veículo exibe informações para sua condução apropriada, um *dashboard* corporativo serve para o mesmo propósito, ou seja, tomar decisões estratégicas para conduzir as organizações. Existem diversas ferramentas para mineração e exibição de informações e dentre elas estão algumas ferramentas da Oracle, como a *Oracle BI Publisher*, que é uma solução para gerar, gerenciar e distribuir relatórios de maneira fácil e rápida, utilizando o próprio navegador da Web [11]. Em virtude da ferramenta *Oracle BI Publisher* ser facilmente integrada ao banco de dados Oracle, esta será utilizada para demonstrar as visualizações em forma de *dashboards* do DW implementado nos experimentos.

3. MATERIAIS E MÉTODOS

Para condução dos experimentos foi utilizada a base de dados relacional *classicmodels*¹, armazenada no banco de dados Oracle, utilizada somente para fins didáticos para comprovação do método. A base contém informações de um negócio de vendas de miniaturas de veículos clássicos, que possui lojas localizadas em diferentes países. Estão armazenados em suas tabelas os dados de clientes, das lojas, dos pedidos e produtos, além de informações de pedido de compra, produtos vendidos e lojas que realizaram a venda. As informações contidas no banco são dos anos de 2003, 2004 e 2005. O modelo é composto pelas tabelas descritas: i) *Offices*: contém detalhes dos escritórios de venda; ii) *Customers*: dados dos clientes; iii) *Employees*: dados dos funcionários, incluindo local de trabalho e hierarquias. iv) *Products*: modelos de carros cadastrados; v) *Products Line*: categorias das linhas de produto; vi) *Orders*: pedidos feitos pelos clientes; vii) *Orders Details*: itens dos pedidos realizados; e viii) *Payments*: pagamentos feitos pelos clientes.

Para montagem e carga do DW, foram utilizadas as informações das tabelas *Offices*, *Employees*, *Orders*, *OrderDetails* e *Products* do banco de dados relacional. Da tabela *Offices* foram retirados os dados das lojas (cidade, estado e país), da tabela *Employees* foram aproveitados os dados sobre vendedores (nome e sobrenome), a tabela *Orders* forneceu dados de quando foram feitos os pedidos (mês e ano), da tabela *OrderDetails* foram usadas as informações de quantidade e preço dos produtos e da tabela *Products* o nome e a linha do produto foram utilizados.

¹ <http://www.mysqltutorial.org/mysql-sample-database.aspx>

Após identificados os dados necessários para a criação do DW no banco de dados *Oracle*, foi elaborado o modelo multidimensional no esquema estrela, apresentado na Figura 1, possuindo uma tabela fato *Vendas* e quatro tabelas de dimensões: *Local*, *Produto*, *Tempo* e *Vendedor*. A tabela *Local* representa onde ocorrem os fatos do negócio, a tabela *Produto* representa o que foi negociado, a tabela *Tempo* representa quando ocorreram os fatos, a tabela *Vendedor* representa quem participa da ação de negócio e a tabela *Vendas* representa a junção das informações, incluindo os dados de quantidade e preço. Já para realizar a carga de dados descrita, do modelo relacional para o modelo DW, foram utilizadas funções desenvolvidas em PL/SQL (Procedural Language/Structured Query Language), diretamente no banco de dados em questão.

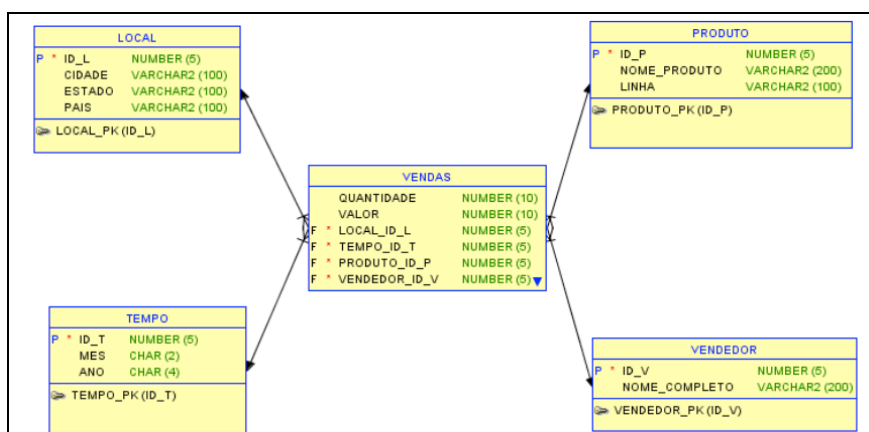


Figura 1. Modelo DW proposto.

4. RESULTADOS

Após o modelo DW ter sido alimentado pelos dados do modelo relacional, foi elaborada uma consulta baseada na Linguagem de Consulta Estruturada, do inglês Structured Query Language (SQL) apresentado os resultados de forma linear e posteriormente os mesmos dados foram retratados em forma de *dashboards* com indicadores visuais em forma de barras. A Figura 2 apresenta o SQL da consulta, que retorna o total de quantidade vendida e o lucro gerado, entre os anos 2003 e 2005, nos países em que as lojas estão instaladas. É feita uma busca nas tabelas *Local*, *Vendas* e *Tempo*, retornando os campos país, ano, quantidade e valor; onde valor é o produto do valor unitário pela quantidade vendida. A Tabela 1 exibe o resultado dessa consulta de forma linear, separando pelas colunas *PAÍS*, *ANO*, *QUANTIDADE* e *VALOR* e mostrando as informações de cada país por linha.

```

1  select l.pais, t.ano,
2  to_char(sum(v.qtd), 'FM999G999G999', 'nls_numeric_characters='
3  ',. '') as Quantidade,
4  '$'||to_char(sum(v.valor*v.qtd), 'FM999G999G999D90', 'nls_numeric_characters='
5  ',. '') as Valor
6  from local l
7  inner join vendas v,
8  on l.id_l = v.id_l
9  inner join tempo t
10 on t.id_t = v.id_t
11 group by cube (t.ano, l.pais)
12 order by t.ano, l.pais asc;

```

Figura 2. Consulta SQL.

Tabela 1. Resultado da Consulta SQL.

PAÍS	ANO	QUANTIDADE	VALOR
Australia	2003	3.529	\$343.180,00
France	2003	10.694	\$1.065.566,00
Japan	2003	2.851	\$292.509,00
UK	2003	6.017	\$616.272,00
USA	2003	13.348	\$1.359.748,00
	2003	36.439	\$3.677.275,00
...

A Figura 3 retrata a mesma consulta SQL, porém na forma de *dashboards*, apresentando no gráfico de barras os valores da venda para *France*, *USA*, *UK*, *Australia* e *Japan* nos anos de 2003, 2004 e 2005. E abaixo, as vendas são detalhadas na forma de uma tabela pela quantidade e valor exato de cada uma, sendo dividida pelos países e anos citados acima. Uma das grandes vantagens das ferramentas de *dashboards* são as facilidades de modificar as visualizações baseadas na mesma consulta realizada, o que é mais complicado nas consultas SQL, pois nestas mudanças de visualização mesmo baseadas nos mesmos dados é necessário a reconstrução da consulta SQL em questão.

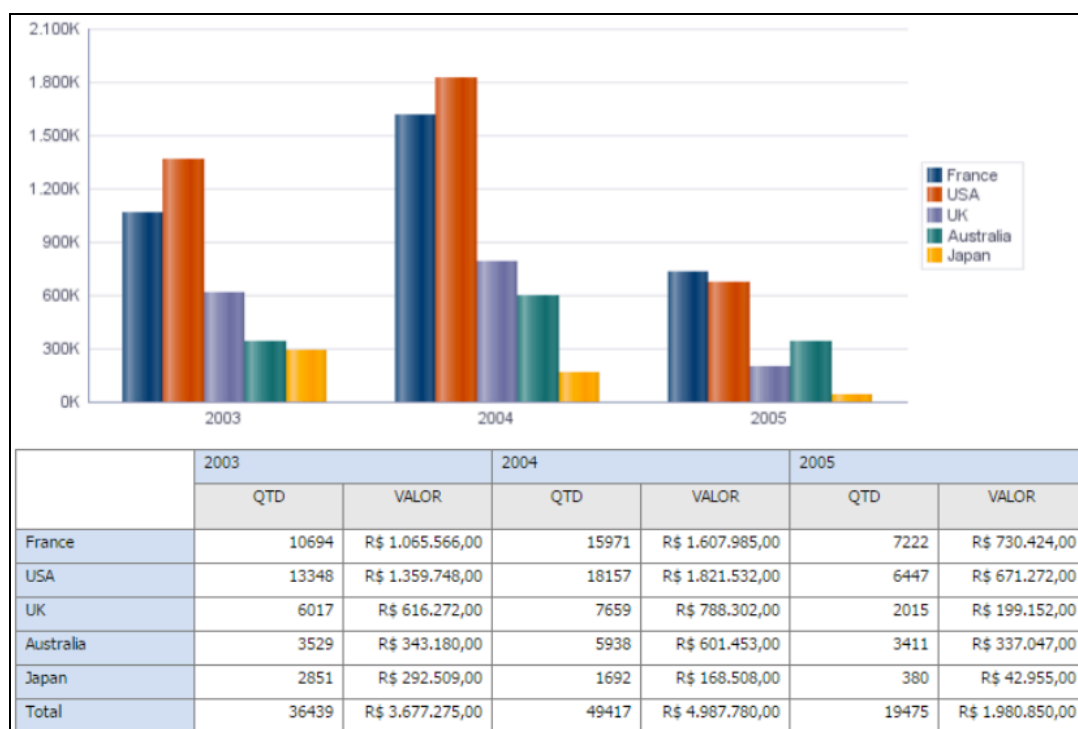


Figura 3. Dashboards da consulta SQL.

5. CONCLUSÃO

O modelo relacional é criado para armazenar todos os dados relevantes ao modelo de negócio, mas nem sempre sua estrutura favorece a extração das informações gerenciais de forma rápida, principalmente porque os dados que geram as informações estão divididos em diversas tabelas relacionadas. Um modelo DW pode ser criado, focado nas visões gerenciais desejadas, tendo em sua alimentação os dados sintetizados, o que favorece o acesso mais rápido das informações, além de possibilitar mudanças de dimensão na mesma visualização. Ferramentas que apresentam estes dados em forma de *dashboards* tem sido cada vez mais utilizada na apresentação das visões gerenciais criadas através de DWs. Dessa forma, pode-se concluir que a utilização de ferramentas para exibição das informações em forma de *dashboards* a partir de um DW facilita muito a visualização e entendimento da real situação das organizações para tomadas de decisão. Além da interatividade na filtragem de informações, o sistema permite que o resultado visual utilizando gráficos e tabelas seja mais amigável. Fica evidenciado que a aplicação de um modelo multidimensional baseado em banco de dados relacionais pode dar um ganho de velocidade na montagem das informações.

6. REFERÊNCIAS

- [1] MATTIODA, R. A.; FAVARETTO, F. *Qualidade da informação em duas empresas que utilizam Data Warehouse na perspectiva do consumidor de informação: um estudo de caso*. Gest. Prod., São Carlos, v.16, n.4, p. 645-666, Dec. 2009.
- [2] FORTULAN, M. R.; FILHO, E.V.G. *Business Intelligence, um modelo para explorar o chão-de-fábrica*. Gest. Prod., São Carlos, v.12, n.1, p. 55-56, Jan. 2005.
- [3] FAVARETTO, F. *Melhoria da qualidade da informação no controle da produção: estudo exploratório utilizando Data Warehouse*. Gest. Prod., São Paulo, v.17, n.2, p. 343-353, Ago. 2007.
- [4] MACHADO, F.; ABREU, M. *Projeto de Banco de Dados*. São Paulo: Editora Érica, 2004.
- [5] OLIVEIRA, C.H.P. *SQL Curso Pático*. São Paulo: Novatec, 2002.
- [6] DATE, C.J. *Introdução a sistemas de banco de dados*. Rio de Janeiro: Elsevier, 2003.
- [7] SILVA, L.A.; BRUSCHI, G.C.; SILVA, F. Q.; TORRES, J. R. *Proposta de Data Warehouse para ser Consumido por Gestores de Produção*. Revista FATEC Sebrae em debate: gestão, tecnologias e negócios. Brasil, v. 2, p. 97-110, 2015.
- [8] MACHADO, F.N.R. *Tecnologia e projeto de Data Warehouse*. São Paulo: Editora Érica, 2010.
- [9] SHARMA, N.; IYER, A.; BHATTACHARYA R.; MODI N.; CRIVELINI W. *Getting Started With Data Warehousing*. São Paulo: Editora Markham IBM, 2012.
- [10] FEW, S. *Information Dashboard Design*. São Paulo: Editora O'Reilly, 2006.
- [11] ORACLE. Oracle Business Intelligence Publisher. Disponível em <http://www.oracle.com/technetwork/middleware/bi-publisher/overview/index.html>. Acesso em: 10/06/2016.