

BANCO DE DADOS MULTIDIMENSIONAIS: ANÁLISE DOS CONCEITOS FUNDAMENTAIS E SUA IMPORTÂNCIA NA TOMADA DE DECISÃO NO AMBIENTE CORPORATIVO

MULTIDIMENSIONAL DATABASE: ANALYSIS OF THE FUNDAMENTAL CONCEPTS AND THEIR IMPORTANCE IN THE DECISION MAKING IN THE CORPORATE ENVIRONMENT

Marcio Maestrello Funes*¹

RESUMO

Muitas empresas e modelos de negócio têm como principal fonte de informações para tomada de decisões em diversas áreas, o seu próprio banco de dados, o modelo relacional amplamente utilizado no campo mercadológico é uma das escolhas mais requisitadas por desenvolvedores, porém não a única. Bancos de dados multidimensionais fornecem uma alternativa de modelo para banco de dados onde temos a possibilidade de busca e análise de informações em um modo multidimensional, esta particularidade permite não só criar novas possibilidades de combinação de dados, antes não conquistadas no modelo relacional, mas também fornecer novos pontos de vista a desenvolvedores que ao conceberem suas soluções, necessitam de fornecer aos seus usuários finais diversos relatórios para tomada de decisão. Este artigo fornece uma visão geral sobre os conceitos relacionados ao uso de banco de dados multidimensionais e uma análise de como é possível aplicar cada conceito em exemplos mercadológicos.

Palavras-chave: Banco de Dados. Multidimensionais. Tomada de Decisão. Ambiente Corporativo.

ABSTRACT

Many companies and business models have as their main source of information for decision making in many areas its own database, the relational model widely used in the marketing field is one of the most sought choices by developers but not the only one. Multidimensional databases provide an alternative model for database where we have the ability to search and analyze information in a multidimensional mode, this feature allows not only create new possibilities not previously achieved in the relational model but also provide new insights to developers when designing their solutions need to provide their end users and several different possibilities for decision making. This article provides an overview of the fundamental concepts necessary to multidimensional data and bank an analysis of how we can apply each concept in market examples.

Keywords: Database. Multidimensional. Decision Making. Environment Corporate.

* Universidade de São Paulo (USP) – Laboratório Intermídia. marciofunes@usp.br

Introdução

No mundo da tecnologia da informação, empresas e profissionais estão em busca de novas soluções e novas possibilidades para o uso da tecnologia e o modo com que ela pode melhorar as possibilidades já existentes. Na área de banco de dados não seria diferente, devido ao aumento de informações que são geradas dia a dia através de diversos processos e a compreensão de que diferentes formas de armazenamento podem influenciar estes processos, a busca de novos modelos de banco de dados são totalmente válidos e justificáveis.

Um sistema de banco de dados deve permitir muito além do simples armazenamento de dados, principalmente no campo mercadológico onde empresas necessitam de informações e conhecimento baseado nos dados armazenados que possuem para tomadas de decisões. Acumular dados é apenas o primeiro passo, é necessária a consulta inteligente e a busca de informações refinadas para que tomadas de decisões ocorram com a menor chance de erros (MACHADO, 2004).

Devido à importância de não apenas armazenar dados em um banco, mas sim a busca inteligente desses dados, aumentasse a procura por modelos de sistemas de banco de dados que permitam suporte a tomada de decisão e também a flexibilização de informações de acordo com as necessidades de seus usuários, neste artigo veremos um desses modelos, o modelo multidimensional.

1 Fundamentação Teórica

1.1 Sistemas de Banco de Dados Multidimensionais

Na década de 1970, muitos conceitos sobre linguagens de consulta a banco de dados surgiram, além de linguagens foram também criados vários sistemas personalizados de apoio à decisão ao redor dessas linguagens. Naquela época, o foco era o uso de geradores de relatórios mais dinâmicos e que possibilitassem a visualização de diversas informações de formas diferentes, com isso os usuários finais passaram a ter acesso direto aos depósitos de dados de computadores com a possibilidade da manipulação e visualização desses dados para o apoio a decisão, esse breve histórico mostra o início da formação do modelo para sistemas de banco de dados multidimensionais (DATE, 2004).

Sistemas de banco de dados multidimensionais estão mais focados na forma de análise e busca de dados do que especificamente na coleta e armazenamento de dados, devido a esse foco a modelagem do sistema multidimensional e a forma de sua criação afeta o modo com que as informações serão analisadas e a forma de busca e retorno de informação.

A visão multidimensional que se pode atribuir a um sistema de banco de dados, ultrapassa o conceito convencional de consultas e edições de um determinado dado ou um conjunto de dados. Modelar um banco multidimensional permite análises profundas e o estabelecimento de relações de dados que no modelo convencional são de difícil ou impossível associação.

1.2 Tabela Fato

Tabela fato é a tabela do banco multidimensional que indica um referencial central armazenando as transações que serão analisadas neste banco. Possui em seu interior as medidas (quantidades numéricas) que irão quantificar o modelo de negócio, desta forma existe um ponto inicial e que interliga as dimensões.

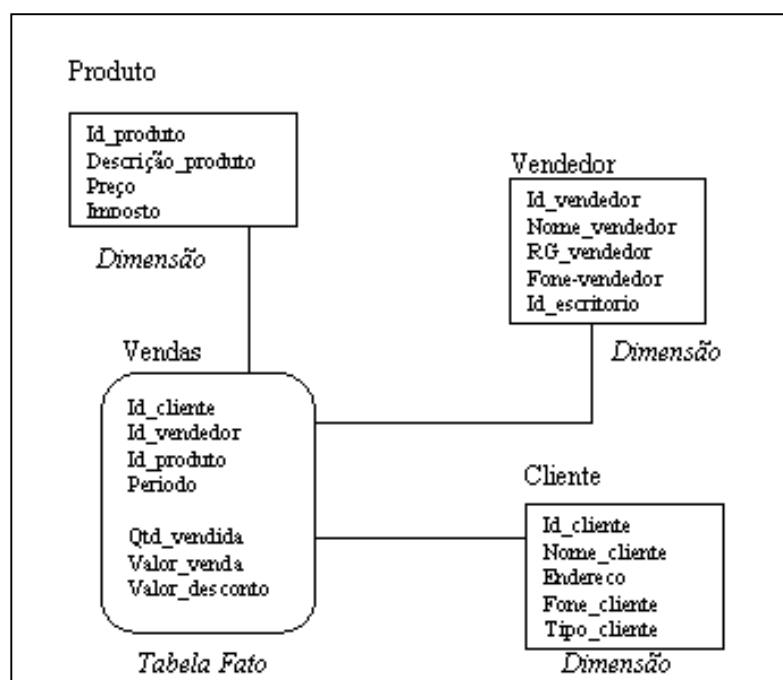


Figura 1 – Tabela Fato e Dimensões

Fonte: Revista Infotec (2010).

A tabela fato Vendas da Figura 1 possui as chaves primárias: Id_cliente, Id_vendedor, Id_produto e também possui as medidas: Qtd_vendida, Valor_venda e Valor_desconto. Nota-se que a tabela fato é construída contendo as chaves primárias das outras tabelas e também quantifica o modelo de negócio (Qtd_vendida) com a associação de valores (Valor_venda e Valor_desconto). Portanto, a Tabela Fato Vendas é a principal associação de dados desse exemplo e as dimensões Produto, Vendedor e Cliente fornecem a tabela os dados necessários para modelagem do cubo no banco de dados multidimensional.

1.3 Medidas

São os atributos sempre relacionados à numerais aos quais tem a representatividade de desempenho como indicador de negócio. Quando analisados as dimensões que se relacionam em um cubo de diversas vezes, é necessário um indicador numérico que indique meios de criação de métricas comparativas e de crescimento ou decréscimo (MACHADO, 2004).

Medidas possuem um papel especial principalmente quanto à métrica numérica relacionada a valores financeiros, quando necessário comparativos de valores de vendas, números de vendas (a partir do ponto de vista de uma ou mais dimensões) ou a quantificação de produtos em um estoque, é possível criar variáveis do tipo medidas.

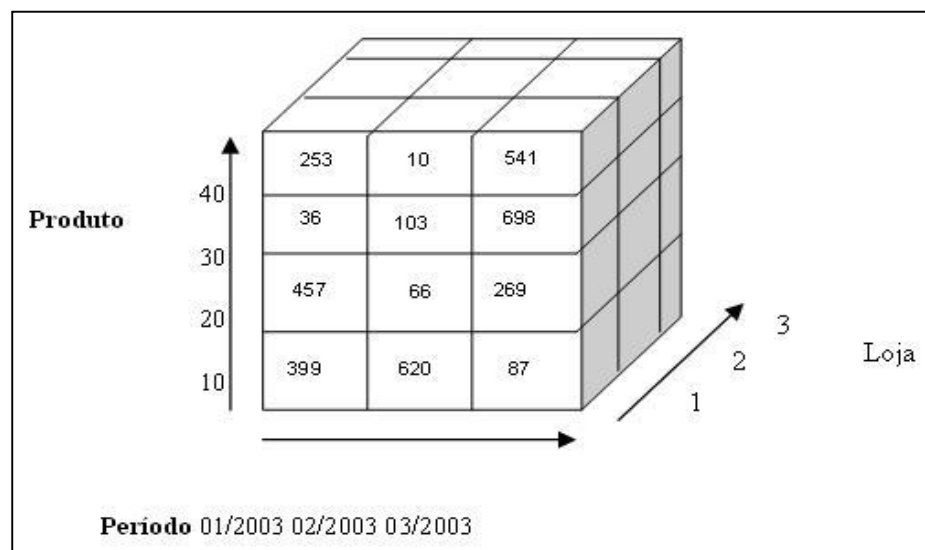


Figura 2 – Medidas dispostas em um cubo

Fonte: Campos (2005)

Na Figura 2 nota-se o uso de uma medida “66”, ela representa a quantidade de venda do produto 20 no mês de 02 de 2003 e ocorreu na Loja 1, facilmente nota-se as três dimensões que compõe o cubo e como pode-se encontrar a referência desejada e ainda apresentar múltiplas opções de consulta ao usuário final.

2 Características de multidimensionalidade

2.1 Cubos e Hiper cubos

Segundo DATE (2004), o conceito de cubo ou também conhecido como hiper cubo é um dos principais conceitos fundamentais que caracterizam os sistemas de banco de dados multidimensionais. O cubo possibilita a análise de dados ou o conjunto de dados de forma tridimensional como a figura abaixo demonstra:

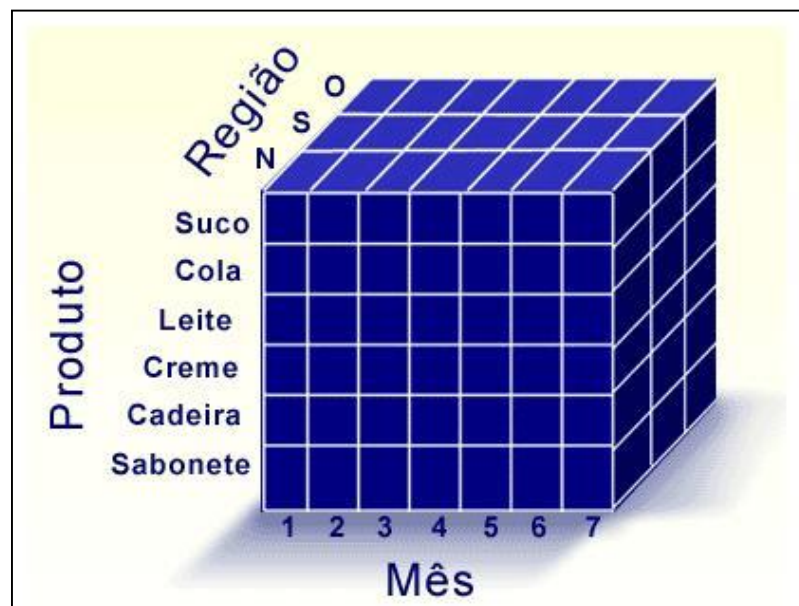


Figura 3 – Cubo multidimensional

Fonte: Nardi (2007)

As células do cubo contêm as informações que podem ser analisadas de diversos ângulos devido existir mais de um ponto de vista, essa análise se dá pelo fato dos eixos do cubo ou também conhecidos como dimensões, representarem diversos assuntos diferentes. Na Figura 3 nota-se o eixo do produto, o eixo mês e o eixo região que em conjunto possibilitam a análise de suas células na busca de informações de diversas formas possíveis,

simultaneamente pode-se analisar o produto que está ao mesmo tempo sendo referenciado há um mês e uma região.

Segundo Mannino (2010), em seu livro Projeto, Desenvolvimento de Aplicações e Administração de Banco de Dados, nota-se uma exemplificação bem clara da utilização de cubos, ele diz:

Imagine uma empresa que vende produtos eletrônicos em diferentes partes dos Estados Unidos. Essa empresa comercializa especialmente quatro diferentes impressoras (laser monocromática, jato de tinta, fotográficas e portáteis) em cinco estados distintos (Califórnia, Washington, Colorado, Utah e Arizona). Para armazenar os dados das vendas diárias para cada produto e cada local em um banco de dados relacional você precisaria dispor a informação em três colunas Produto, Local e Vendas de 20 linhas, quatro instâncias de Produtos vezes cinco instâncias de Local (MANNINO, 2010, p. 45).

Abaixo encontra-se um exemplo baseado ao que Mannino (2010) se refere:

| Produto | Local | Venda |
|---------------------|------------|-------|
| Laser monocromática | Utah | 200 |
| Jato de tinta | Califórnia | 100 |
| Fotográfica | Arizona | 50 |
| Portáteis | Colorado | 237 |

Tabela 1 – Tabela Relacional

Fonte: Mannino (2010)

Na Tabela 1, é possível visualizar que no estado de Utah foram vendidas 200 unidades de impressoras do tipo Laser monocromática. Para conseguir análises mais profundas e que suportem tomadas de decisão por meio do cruzamento dos dados, é necessário construir mais de uma tabela no modelo relacional para visualizar a informação de vários outros pontos de vista, como por exemplo, a possibilidade de visualizar a quantidade de impressoras Jato de tinta vendidas neste mesmo estado.

Outra possibilidade será analisar os produtos e a quantidade vendida juntamente ao tempo que a venda foi realizada, essa informação será útil para saber se existe uma época mais favorável de venda e assim tomar decisões referentes à estratégia de marketing e

logística, acrescentar uma nova coluna na Tabela 1 seria possível, porém seria necessário criar outras diversas tabelas para conseguir armazenar todas as variáveis possíveis utilizando as informações de Produtos, Local, Venda e agora Tempo.

A solução mais viável neste caso, será a utilização de um banco de dados multidimensional no qual pode-se compreender todas as dimensões (produtos, local, vendas e data) sem a necessidade de mudar a modelagem do sistema se acaso uma nova dimensão do modelo de negócio desejar ser acrescentada. Na Figura 4 a mesma representação da Tabela 1 incluindo a dimensão Tempo.

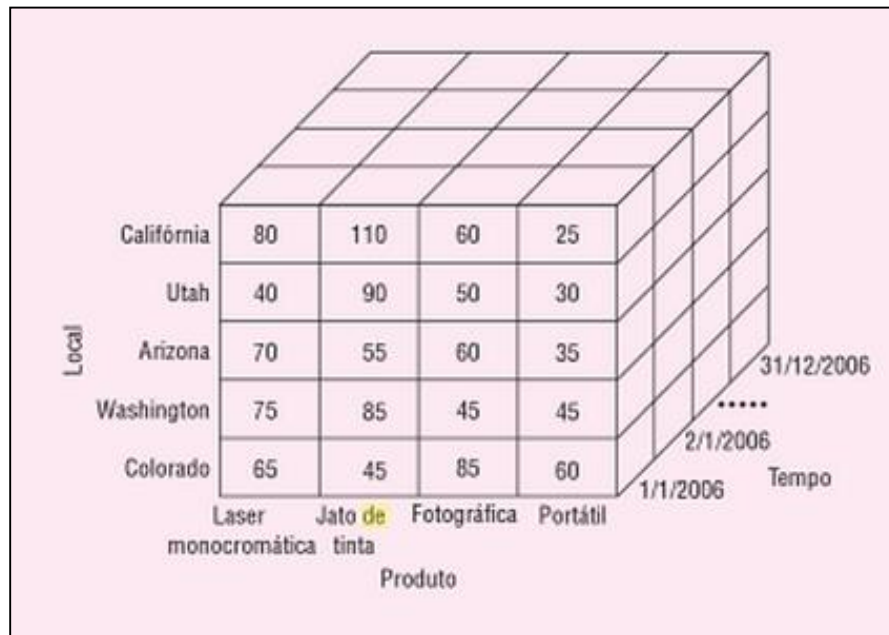


Figura 4 – Cubo Multidimensional
Fonte: Mannino (2010)

Analisando a Figura 4 pode-se, em um mesmo cubo, visualizar todas as dimensões que necessita-se para apoiar decisões no modelo de negócio da empresa do exemplo, isso demonstra claramente que em casos de muitas dimensões o modelo relacional de banco de dados pode não ser a melhor opção.

2.2 Multicubos

Também é possível representar dados na forma de multicubos, a uma estrutura de um cubo pode-se agrupar multicubos em sua composição. Esses multicubos podem ser

analisados de forma independentes na tentativa de retornar o dado de forma mais rápida e fácil. Na composição de um multicubo, existem os dados como segmento em um conjunto de pequenos cubos separados e cada cubo contém os eixos diferentes uns dos outros (MACHADO, 2004).

O aumento das possibilidades de análise de informação principalmente quando o modelo de negócio ao qual se está modelando a solução multidimensional, possui muitas variáveis como demonstra a Figura 5.

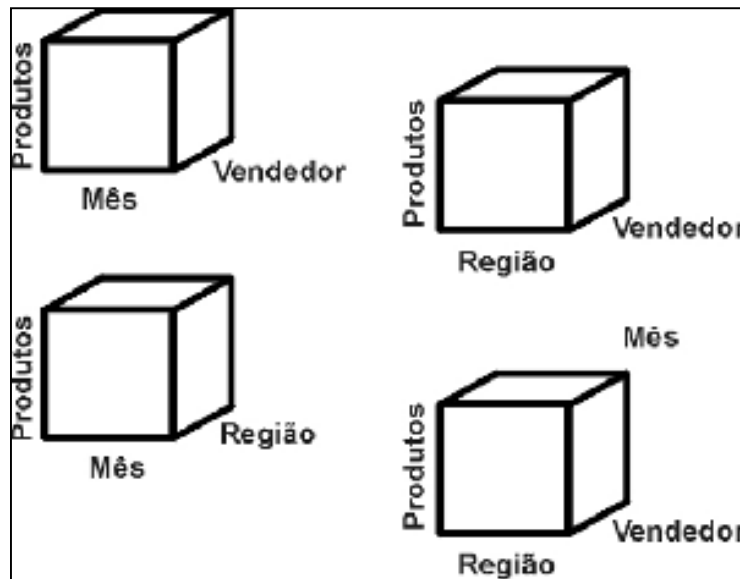


Figura 5 – Multicubos e suas dimensões
Fonte: Sodek (2001)

No modelo de negócio da Figura 5 nota-se:

- Produtos
- Mês
- Vendedor
- Região

Por existir quatro dimensões diferentes para a análise de um mesmo tipo de dado, a possibilidade de representação em multicubos é justificada. A informação está armazenada em vários cubos e se a quantidade de dimensões tende a ser maior que a quantidade de cubos, as dimensões podem ser aumentadas exponencialmente. No exemplo da Figura 5, as dimensões mês, produto e vendedor estão cruzadas em um subcubo de forma a analisar a

informação sobre o ponto de vista dessas dimensões, em outro momento deseja-se também visualizar essa mesma informação agora com o ponto de vista da dimensão região e assim cria-se mais um subcubo onde é acrescentado as dimensões Produtos e Vendedor.

Segundo Soweck (2001), a vantagem da utilização de multicubos está na menor utilização de espaços de armazenamento em disco, pois o problema de dados esparsos é diminuído, outra vantagem é o melhor desempenho em consultas em um único cubo. A desvantagem dessa utilização está na necessidade de inúmeras consultas em diversos cubos com dimensões diferentes, quando isso ocorre o processamento exige mais do hardware e software e existe queda no desempenho tornando respostas mais lentas e desempenho comprometido, conhecer as vantagens e desvantagens do uso de multicubos e o modelo de negócio que se espera da solução auxilie na escolha do sistema de banco de dados multidimensional.

Ainda segundo Soweck (2001), essas são as principais características de multicubos e hipercubos:

- Multicubos são mais eficientes e versáteis;
- Hipercubos são mais fáceis para entender;
- Usuário deve reconhecer melhor hipercubos pela sua simplicidade de uso;
- Profissionais com experiência em multidimensional preferem multicubos por causa da sua maior flexibilidade e pela facilidade de ajustes;
- Multicubos é a maneira mais eficiente de armazenar muitos dados esparsos podendo reduzir o efeito de explosão de valores pré-calculados em banco de dados;
- Produtos mais sofisticados tendem a usar multicubos;
- Aplicações pré-construídas também tendem a usar multicubos permitindo que o dado estruturado possa ser melhor ajustado para as necessidades conhecidas da aplicação;
- No hipercubo, cada dimensão pertence a um único cubo;
- Em um multicubo, a dimensão pode fazer parte de múltiplos cubos.

3 Tomada de decisão no ambiente corporativo

3.1 Dimensões

Ao analisar a Tabela 1, criada no modelo relacional, nota-se claramente que suas dimensões são formadas pelas colunas e linhas de sua tabela, nesse modelo apenas cruzando a dimensão linha e a dimensão coluna tem-se a referência do dado que é buscado.

Em sistemas multidimensionais o papel da dimensão é o mais importante, geralmente um banco de dados é construído com conceitos de multidimensionalidade quando a presença de muitas dimensões é necessária para satisfazer o modelo de negócio ao qual se está modelando o banco, portanto, as dimensões são as características fundamentais da estrutura dimensional.

Através das dimensões, pode-se agrupar e referenciar os dados ao longo de todo cubo, assim como pode-se observar em eixos de um cubo matemático, o funcionamento e a arquitetura do modelo multidimensional, que permite a inserção ou retirada de novos eixos aumentando o tamanho do cubo sem perder as referências criadas desde o início da modelagem e agregar novas possibilidades de busca de informações.

Observe na Figura 6 o uso das dimensões Local, Produto e Data em um mesmo cubo, cada lado no cubo possui referência em três dimensões diferentes e, ao mesmo tempo, dessa forma, o sistema tem a possibilidade de analisar o cubo utilizando apenas as dimensões que achar necessária para apresentar as informações ao usuário final, criando relatórios dinâmicos, por exemplo.



Figura 6 – Dimensões de um ambiente corporativo
Fonte: O Autor

3.2 Hierarquia

A hierarquia presente no modelo de sistemas multidimensionais possui o papel de organização em níveis, que permitem a disposição da informação de forma a facilitar a busca rápida de um dado procurado. Como cada dado pode estar presente em multidimensões, sua hierarquia permite o controle e gestão do banco de dados para que ele cresça de forma ordenada e balanceada (DATES, 2004).

Ao analisar uma empresa de qualquer segmento mercadológico, pode-se observar que sua estrutura se baseia em uma hierarquia onde os departamentos com diversos funcionários que respeitam uma ordem de liderança. Ao construir uma solução que atenda o cadastro dos funcionários dessa empresa, será necessário especificar na solução o nível hierárquico que cada funcionário da empresa possui.

Aplicando a mesma lógica presente no mercado de trabalho ao modelo dimensional representado na Figura 4, a dimensão Tempo poderá ser o nível primário hierárquico que permitirá a visualização de vendas por dia, por mês ou por ano. Dessa forma, cria-se um nível hierárquico que prevê todas as relações da dimensão Tempo e como hierarquias secundárias das dimensões Dia, Mês e Ano podem ainda mais criar níveis que demonstrem as vendas por hora ou por minuto.

A possibilidade de hierarquia de níveis em um modelo de banco multidimensional permite a abertura para análise do conjunto de dados de forma abrangente, pode-se gerar relatórios dinâmicos ao qual o usuário final, através de um simples filtro de busca, detalhar a visualização da informação para apoiar suas decisões e forma de aproveitamento da informação.

3.2.1 Membros de uma Hierarquia

Após a definição de uma dimensão como sendo de nível primário, será necessário definir quem serão os membros a compor os níveis desta hierarquia. No exemplo da Figura 4, a dimensão Tempo poderá ser de nível hierárquico primário e possuir como membros da hierarquia o Dia, o Mês e o Ano, criando uma relação que será útil ao modelo de negócio.

Outra possibilidade é o compartilhamento de membros no modelo multidimensional, uma vez definido que duas dimensões podem possuir relação com um mesmo membro hierárquico, o modelo permite que haja compartilhamento entre as duas dimensões sem alterar a estrutura do banco de dados.

No exemplo da Figura 4, o membro Mês que poderá pertencer à hierarquia da dimensão Tempo, pode ser compartilhado com a dimensão Local de forma a não só apresentar quantas unidades de impressora foi vendido no Local especificado em uma data selecionado, mas também, informar qual local recebeu uma quantidade maior de impressoras no geral ou de impressora específicas, essa informação poderia auxiliar uma tomada de decisão de abrir ou não uma filial no local mais visitado.

Além do compartilhamento de membros também podemos criar membros padrões, eles são utilizados como referencial entre os membros da hierarquia. Ao cadastrar funcionários de uma empresa na dimensão Funcionários de um cubo pode-se criar uma hierarquia que permita dizer ao modelo de negócio que determinado funcionário é responsável por um setor de vendas muito grande e aquela região é detalhada em várias divisões, quando queremos filtrar por determinada região podemos usar o membro padrão como referência e ainda detalha-lo quando necessário (CAMPOS, 2005).

Considerações finais

Com base nas informações contidas neste artigo nota-se que o conceito de banco de dados multidimensionais permite explorar diversas opções de busca e visualização de informação aplicadas a modelos de negócio e tomada de decisão. Em um mercado que está sempre à procura de novas possibilidades de manipulação da informação o conceito de multidimensionalidade permite que desenvolvedores criem soluções mais complexas e que possibilitam ao usuário final recursos de análise e tomada de decisão antes não exploradas.

A compreensão dos conceitos fundamentais ao qual a multidimensionalidade de um banco de dados está construída, permite a prospecção de novas ferramentas, novos aplicativos e modificações de soluções já existente que não satisfazem totalmente o usuário dessa solução.

Apesar de muitos sistemas trabalharem com a modelagem de seus bancos de dados no modelo relacional, pode-se entender a multidimensionalidade como uma evolução no modo de tecnologias voltadas a banco de dados que podem agregar novas possibilidades e novos caminhos antes não percorridos.

Referências

CAMPOS, R. A. **Qualidade de dados em Data Warehouse**. 2005. 72f. Trabalho de Conclusão de Curso (Graduação em Sistemas de Informação) - Centro de Ensino Superior de Juiz de Fora – CESJF, Juiz de Fora, 2005.

DATE, C. J. **Introdução a sistemas de bancos de dados**. São Paulo: Editora Elsevier Brasil, 2004.

INFORMATIVO TÉCNICO Nº 52. Olap - Molap – Rolap. **Revista Infotec**. Edição 4. Disponível em: <<http://www.ccuec.unicamp.br/revista/infotec/informacao/inf54.htm>>. Acesso em: 10 jun. 2014.

MACHADO, F. N. R. **Projeto de Data Warehouse: uma visão multidimensional**. São Paulo: Érica, 2004.

MANNINO, M. V. **Projeto, desenvolvimento de aplicações e administração de banco de dados**. 3. ed. São Paulo: Editora Mc Graw Hill, 2010.

NARDI, A. R. **Fundamentos e modelagem de bancos de dados multidimensionais**. MSDN – Microsoft Developer Network. Disponível em: <<https://msdn.microsoft.com/pt-br/library/cc518031.aspx#XSLTsection127121120120>>. Acesso em: 13 jun. 2014.

SOWEK, C. A. Estruturas de dados multidimensionais. **Jornal Técnico Bate Byte**, São Paulo, 2001. Disponível em: <<http://www.batebyte.pr.gov.br/modules/conteudo/conteudo.php?conteudo=1732>>. Acesso em: 15 jun. 2014.