

Integração Access-Excel para produzir um sistema de apoio a decisão que simula um Data Warehouse e OLAP

Wilson Luiz Vinci (Faculdades IPEP) wilson@cnptia.embrapa.br

Marcelo Gonçalves Narciso (Embrapa Informática Agropec.) narciso@cnptia.embrapa.br

Resumo

No mercado existem várias ferramentas de apoio a decisão que auxiliam os gerentes ou administradores em várias situações. Dentre estas ferramentas estão Data Warehouse (DW) e OLAP (On-Line Analytical Processing). Existem situações nas quais podem ser usados outros softwares que fazem o mesmo que DW e OLAP, só que em escala menor. O software Access pode ser usado para modelar um pequeno DW (tabelas de fato e dimensões) e o software Excel pode ser usado como um software OLAP, em pequena escala, podendo fazer consultas complexas sem que o usuário possua conhecimentos em linguagem SQL, bastando apenas o conhecimento de como integrar o Access e o Excel para a montagem de tabelas dinâmicas. As consultas podem ser mostradas em forma textual, em tabelas ou graficamente. Outras características do OLAP, como Drill-Down podem ser feitas também. Um estudo de caso de uma companhia telefônica será mostrado para exemplificar o uso desta integração. Os resultados foram relatórios sobre diversos tópicos importantes que foram feitos usando-se apenas Access e Excel, sem necessitar de softwares com custo mais elevado, que seriam o DW e o OLAP.

Palavras-chave: OLAP; Data Warehouse; apoio a decisão; Excel; Access

1 . Introdução

No mundo atual, decisões são tomadas a todo momento e em curto espaço de tempo, levando os administradores a se informarem cada vez mais rápido e a processarem muitos dados em prazos cada vez menores de tempo.

Para que decisões acertadas sejam praticadas em um tempo muito pequeno, os gestores estão utilizando ferramentas de informação em uso, mas nem sempre estas

respondem dentro do tempo exigido pelas situações apresentadas e as visões que desejam e isto aumenta a margem de erros ou dificulta a localização de informações que se procura.

Essas ferramentas são os bancos de dados operacionais, contendo muitas informações, cujos dados estão sempre variando a cada operação realizada e que normalmente o mesmo tipo de informação pode apresentar variações. Porém são dados brutos para auxiliarem em decisões e mostrarem tendências de consumo.

Entretanto, sendo refinados, trabalhados de maneira correta, podem produzir informações valiosas para quem sabe lidar com elas e sendo bem aplicadas, produzirão resultados esperados, quando o ato decisório é praticado no momento oportuno.

Este trabalho procurará, dentro da proporção adequada, introduzir os conceitos e algumas abordagens sobre uma ferramenta que auxilia na tomada de decisão, que está sendo muito bem aceita no meio empresarial, seus benefícios, qualidades e que pode ser uma solução bastante útil para gerentes e executivos em seus planejamentos e estratégias nos negócios, segundo a realidade atual no segmento empresarial, os Sistema de Informação Gerencial.

Um pequeno projeto será construído para simular um ambiente de sistema de informação gerencial e o benefício desse trabalho contribuirá para os gestores da área de telecomunicações, no sentido de mostrar tendências das atividades em dados.

2. Data Warehouse

Pode-se dizer que um Data Warehouse é um banco de dados formado a partir dos dados operacionais de uma empresa, com dados tratados e agrupados por assunto, não voláteis, isto é, não se alteram em função do tempo ou a todo momento, não constituídos para transações, mas sim para consultas das mais variadas pelos usuários que utilizarem.

O DW é um ambiente, não um produto. Trata-se da construção arquitetônica de um sistema de informação que fornece a seus usuários informações históricas de apoio a decisão, as quais são de difícil acesso ou apresentação quando se utiliza de meios tradicionais de armazenamento de dados operacionais.[5]

Esse sistema tem a finalidade de facilitar o acesso a informações da empresa, como um todo, de maneira compreensível, intuitiva e óbvia para os gestores da área de negócio, permitindo a separação e combinação infinita dos dados contidos no DW[3].

Como todo sistema, um DW, possui elementos que têm uma finalidade, produzindo saídas que somadas às saídas dos outros componentes resultam em informações mais especificadas e valiosas as áreas de gestão.

É necessário compreender bem cada elemento do DW e controlá-los de forma eficaz com o intuito de que o objetivo ao qual o sistema se destina seja alcançado com o máximo de plenitude possível.

Na figura 1, que será mostrada a seguir, fica visível a arquitetura de um Data Warehouse; por ser genérica esta pode ser implementada as várias abordagens existentes para construção da ferramenta de análise decisória:

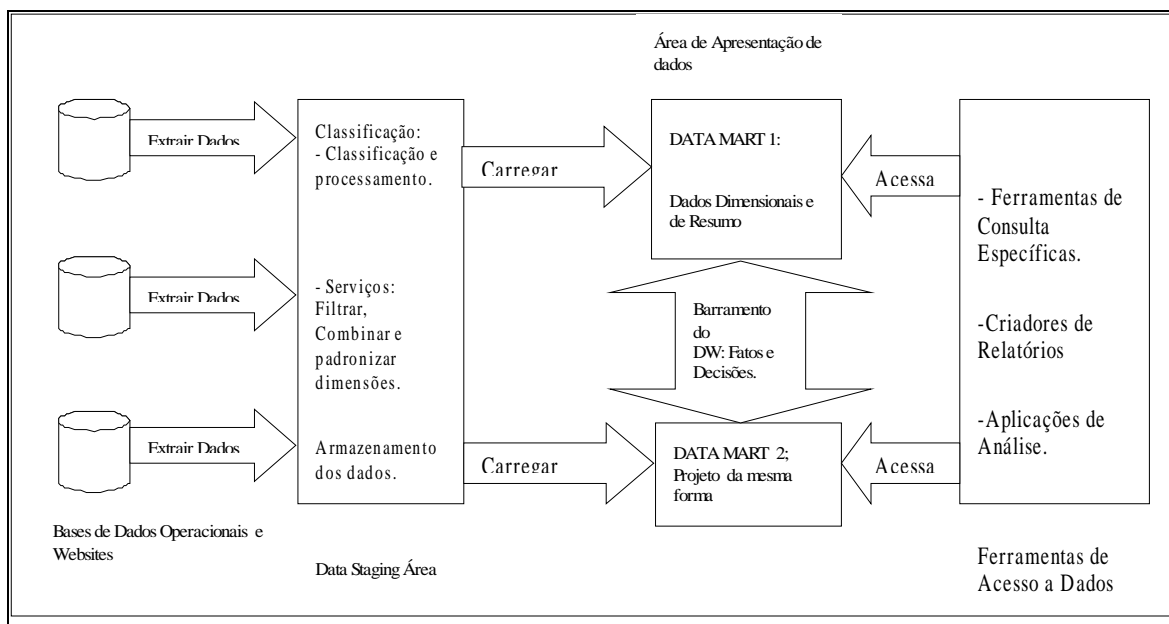


Figura 1 – Arquitetura de um Ambiente Data Warehouse

A seguir a explicação de cada item do esquema acima e suas funções dentro do sistema DW.

Base de Sistemas Operacionais e Fontes Externas. São os dados nos armazéns de informação da organização e soma-se a estes os dados obtidos e armazenados de outras fontes, como dados de websites com informações relativas ao tema incluído no armazém informativo, que serão tratados e armazenados em blocos como descrito.

É a partir dos sistemas operacionais de transação(bancos de dados normalizados) da empresa que os dados são extraídos. Tais sistemas possuem consultas limitadas, normalmente um registro apenas é consultado por vez, e que são aplicações independentes, relacionadas à área específica em que são utilizadas.

Data Mart. É uma parte do todo que compõe a área de apresentação. De uma forma bem elementar um Data Mart representa os dados de um único processo de negócio.

Tabelas de Fatos e Tabelas de Dimensão. O desempenho e as medições numéricas da empresa são armazenados em um modelo dimensional, que é a tabela de fatos, objeto principal do modelo.

Essa tabela está contida em um único Data Mart e armazena os fatos de um processo de negócio para levantamento dos resultados e análise do desempenho.

Volume vendido e quantidade em moeda(Total de Vendas) são exemplos de fatos numéricos e aditivos na tabela de fatos e estes são os dados mais úteis para medição.

Acompanham uma tabela de fatos as tabelas de dimensão. Tabelas de Dimensão contêm descritivos textuais da empresa. Compostas de muitas colunas ou atributos têm a função de restringir uma consulta numa tabela de fatos.

Mês, ano ou semana, loja, marca de produto são exemplos de atributos de uma tabela de dimensão. Através das chaves estrangeiras são relacionadas as tabelas de dimensão com as tabelas de fatos. Mais informações, ver [6].

2.1 Considerações sobre Data Warehouse

Pode-se dizer que o sistema Data Warehouse é de grande porte para servir várias áreas de gestão e administração de uma grande empresa. Suporta grandes quantidades de dados podendo, inclusive, trabalhar em abordagem distribuída ou centralizada.

O sistema é constituído de técnicas de extração, transformação e carga bastante caros, com necessidade de hardware próprio e mão-de-obra bem qualificada para desenvolvimento.

O valor do investimento é retornado, pois as informações produzidas são de suma importância para a área de gestão e os resultados para a empresa, quando bem desenvolvido, são de grande valor.

Todavia, o conjunto de técnicas, em sua totalidade, é inviável para aplicação em negócios de pequeno porte, devido ao custo do investimento e o porte do sistema para o tamanho desses negócios, além do custo da mão-de-obra para uma implementação.

No próximo item será vista outra forma de Sistema de Suporte a Decisão, as técnicas de implementação e verificar-se-á a possibilidade de implantação em negócio de dimensão menor.

3. OLAP (On-Line Analytical Processing).

Os conceitos de OLAP incluem a noção ou idéia de múltiplas dimensões hierárquicas e podem ser usados por qualquer um para que se pense mais claramente a respeito do mundo, seja o mundo material de escala atômica à escala galáctica, o mundo econômico dos micro agentes às macro economias, ou o mundo social dos relacionamentos interpessoais aos internacionais [6].

As Ferramentas OLAP são as aplicações que os usuários finais têm para extrair os dados de suas bases com os quais geram relatórios capazes de responderem as suas questões gerenciais[2].

O OLAP proporciona as condições de análise de dados ON-LINE necessárias para responder às possíveis torrentes de perguntas dos analistas, gerentes e executivos[1].

Conforme as definições acima sobre OLAP permitem a visualização multidimensional de fatos em sua menor parte (restrita a menor unidade) até sua parte maior (ampla unidade), utilizando determinadas ferramentas para a composição dessas informações, gerando meios informativos destinados à usuários da área de gerenciamento de negócios.

3.1. Como o OLAP Pode ser Necessário

A concorrência entre empresas é uma prática comum e muito acirrada, principalmente quando várias organizações trabalham no mesmo ramo do mercado consumidor.

Sair-se-ão melhor, as empresas que conhecerem bem as necessidades de seus clientes e conseguirem perceber e explorar o potencial do mercado em que atuarem.

Para alcançarem isso, a área de gestão necessita tomar decisões fundamentais a cerca de preço de produto, lançamento de novos produtos, conquista de novos mercados e para atingir tais objetivos, esses gestores, precisam de informação.

A importância da boa informação pode ser considerada como a diferença em valor entre decisões certas e decisões erradas, onde as decisões são baseadas nessa informação. A boa informação precisa ser existente, precisa (no sentido de ser verdadeira), oportuna e inteligível [10].

Isso converge para a necessidade seguinte:

“A velocidade com que executivos obtêm informações e tomam decisões determina a competitividade de uma empresa e seu sucesso de longo prazo[4].

Assim, o sistema OLAP pode auxiliar em muito à tomada de decisão, podendo ser usado em todas as áreas empresariais, produzindo informações diversas a estas utilizando o sistema multidimensional.

O sistema OLAP, no sentido lógico, precisa “dar suporte para múltiplas dimensões, hierarquias, fórmulas dimensionais e a separação de estrutura de dados e representação” [6].

Algumas características que são as técnicas que o sistema oferece são:

- Consultas Ad-Hoc – “São consultas com acesso casual único e tratamento dos dados segundo parâmetros nunca antes utilizados, geralmente executado de forma iterativa e heurística.”[3] Segundo a afirmação acima, o usuário final gera consultas para responder os questionamentos que necessita, com visões de informações ainda não vistas e com utilização de métodos que auxiliem os usuários a encontrar o que busca.

- Slice-and-Dice – Utiliza-se essa característica para fazer a análise de informações sob diversos ângulos e o limite só está na imaginação do usuário. Diversas combinações podem ser executadas de modo que outras visões sejam criadas e as informações visualizadas de outra forma.

- Drill Down-Up – Sistema que permite ao usuário subir ou descer dentro de uma hierarquia, como, por exemplo, o tempo visualizando a informação primeiro em anos, depois em meses e em seguida por semanas.

A análise multidimensional é uma das grandes utilidades da tecnologia OLAP, consistindo em ver determinados cubos de informações de diferentes ângulos, e de vários níveis de agregação. Os “cubos” são massas de dados que retornam das consultas feitas ao banco de dados, e podem ser manipulados e visualizados por inúmeros ângulos (usando a

tecnologia de slice-and-dice) e diferentes níveis de agregação (usando a tecnologia chamada “drill”) [2].

3.2 Dimensões e Cálculos da ferramenta OLAP

Os cálculos são elementos importantes na estrutura OLAP, devendo somar e tirar médias de um volume de dados bastante grande. Porém, as informações realmente importantes são extraídas após um processo de cálculos e comparação entre razões e proporções.

Quando a consulta é realizada há um processamento analítico interno para a visualização dos valores solicitados.

Para exemplificar, um gerente de vendas que deseja verificar a venda de produtos, listados de A a Z, as quantidades e o total de vendas de cada um deles, em um determinado mês, além das regiões em que foram vendidos, o sistema terá de efetuar diversos cálculos e comparações e ainda poderá dimensioná-los em valores, ordenando-os dos que venderam mais para os que venderam menos.

3.3 Comentários sobre o sistema OLAP

O sistema de dados multidimensional OLAP é um sistema bem sofisticado que exige um processo de engenharia para implementação bastante caro e mão-de-obra bastante qualificada para implementação.

Como o DW, produz informações bastante importantes e que ajudam a área de negócio em suas decisões e análises e ainda é aplicado em vários departamentos de uma grande empresa.

Esse sistema pode se utilizar de Data Mart do sistema DW para gerar informações gerenciais porque trabalha com múltiplas dimensões e grande quantidade de dados na geração de informações.

Mas para pequenas quantidades de dados os ambientes criados são grandes demais para implantação e o custo também é grande para o porte dos negócios.

Os tipos de bases de dados e ferramentas para apresentação são muito sofisticados para pequenos negócios, porém os conceitos podem ser utilizados com outras ferramentas para desenvolvimento de ambiente gerencial de pequena dimensão que se ajuste ao tamanho do negócio, explorando outras possibilidades de implementação.

4. Utilização do Microsoft Access e do Microsoft Excel para Gerar Relatórios Gerenciais. Um estudo de caso de uma empresa telefônica (Telefonica).

O Microsoft Access, com conexão ODBC no sistema de banco de dados operacional da empresa de grande porte, armazenava todos os dados referentes a

reclamações de clientes do serviço de dados banda larga ADSL e essas reclamações e instalações eram relacionadas à prestadora de serviços telefônicos.

Usando o sistema de vínculos do Access o banco de dados era atualizado sempre pela manhã com as reclamações do dia anterior. Essa atualização era uma consulta que filtrava todas as reclamações do dia ou semana ou mês finalizadas, isto é, com reparo efetuado pelo técnico de ADSL.

Da mesma maneira um outra consulta filtrava as instalações do serviço e todos esses dados (reparos e instalações) eram referentes à minha empresa. Eles eram armazenados nesse banco de dados em duas tabelas que tinham o mesmo nome reparo e instalações e colecionavam os dados de meses dessas instalações e reparos.

Com esses dados era possível filtrar a quantidade de reclamações por cidade e bairro; verificar quais clientes tinham mais de uma reclamação dentro do mês e quais técnicos ou técnico foram verificar a reclamação; a produtividade de cada técnico e outras informações.

Usando as consultas de referencia cruzada utilizava dois campos como, por exemplo, campo técnico e data do reparo, para verificar quantos reparos cada técnico havia feito por dia e isso era acumulado no mês.

Para verificar a produtividade do meu setor – de suporte – utilizava o registro do funcionário que finalizava a instalação ou reparo e a data da finalização para medir a produtividade de cada funcionário.

Utilizando a consulta encontrar duplicatas era possível verificar quais clientes reclamavam mais de uma vez dentro do mesmo mês e acrescentando o campo técnico sabia também qual ou quais técnicos tinham feito o reparo no cliente com reclamação repetida.

As instalações, também, eram contadas através do mesmo processo que se dividiam em instalações realizadas, instalações não realizadas, instalações pendentes que eram criadas por cidade e por bairro de instalação.

Acumulava os dados do setor utilizando os registros de funcionários para filtrar todas as reclamações finalizadas pela minha área para gerar o total de finalizações e comparar com o total de reclamações gerais; também comparar as finalizações do mês corrente em relação ao mês anterior para verificar se haviam diminuído ou aumentado.

E esses dados também eram armazenados em tabelas no Access para serem comparados com o mês seguinte.

Alguns cálculos, os que utilizavam proporções eram realizados nas grades da consulta do próprio Access mesmo e o restante na planilha de cálculo, como o cálculo de percentuais de reclamações de clientes acima de duas dentro de um mês com o total geral de reclamações.

Outra medida calculada no Access era a eficiência dos funcionários do meu setor que era feito somando todas as reclamações baixadas em um mês e dessas quantas reclamações se repetiam dentro do mesmo mês.

E mais uma série de comparações e filtros poderiam ser executados no banco de dados utilizando os dados filtrados, como por exemplo, do total de instalações de um mês, quantos geraram reclamações dentro do mesmo mês.

Montados esses filtros e cálculos era só gerar o relatório no Excel enviando os dados e formatando.

Os relatórios gerados diretamente das consultadas no Access pelo sistema de conexão dados, importar dados e executar uma nova consulta a banco de dados eram vinculados os dados do Access na Planilha Excel e na caixa de diálogo propriedades do intervalo de dados externos marcando a caixa para atualizar os dados assim que abrir o arquivo, fazia atualização automática dessas informações.

Eram gerados relatórios de produtividade por setor – reclamações finalizadas com uma reclamação apenas dentro do mês que era considerado uma reclamação comum. Produtividade por técnico, por atendente, por região.

Reclamações repetidas de um mesmo cliente dentro do mesmo mês, apontando o técnico que havia verificado o serviço e a cidade do cliente. Quantidade de reclamações por região, por segmento de serviço – caracteriza o nível do serviço e tipo de cliente. As instalações também tinham os mesmos relatórios.

Eficiência na baixas de reclamações repetidas por atendente e geral. Relatórios de clientes considerados críticos quanto à reclamação os quais a mão-de-obra e supervisão ficavam mais alertas e o atendimento também.

Levantamento de defeitos provocadas por modem ADSL, por problema na fiação interna do cliente – fiação de telefone, problemas de software, problemas de hardware, problemas relativos a placa de rede, vírus na máquina do cliente, problemas envolvendo os sistemas internos da área de banda larga., pela rede de telefonia – cabos aéreos e subterrâneos, etc.

Todas essas informações eram geradas com pouca intervenção nos softwares porque era carregar os dados e abrir a planilha de dados. Um esquema simples para visualizar o tipo de trabalho realizado pode ser visto na figura 2, a seguir:

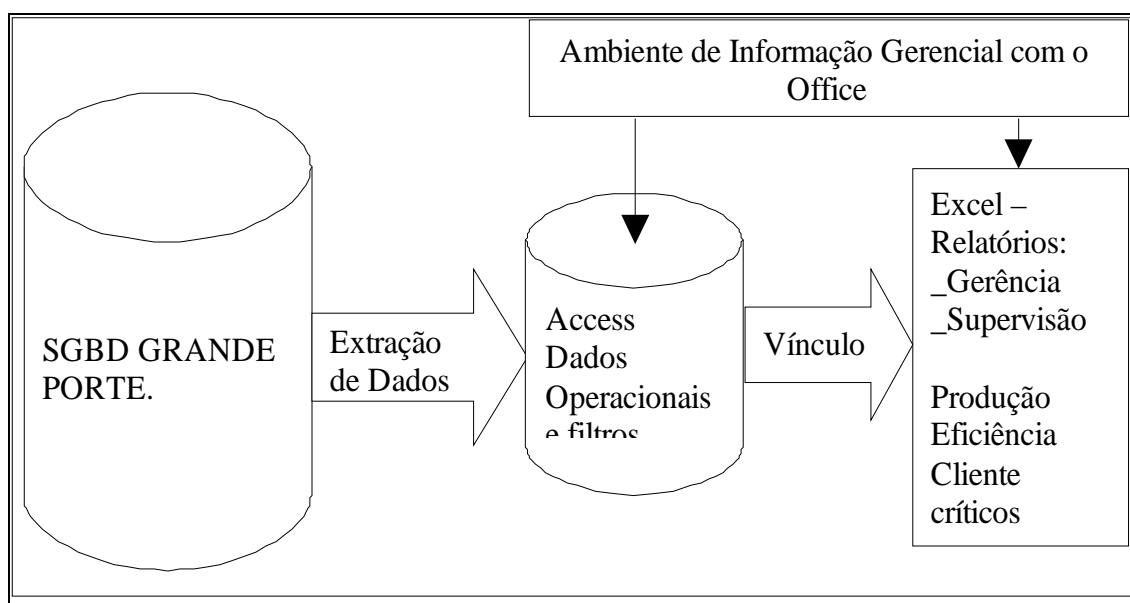


Figura 2 - Esquema de extração de dados e armazenamento para um ambiente de Informação Gerencial

De forma resumida, este processo é feito da seguinte forma:

- 1 – Fazer a modelagem (tabela de fatos e dimensões)

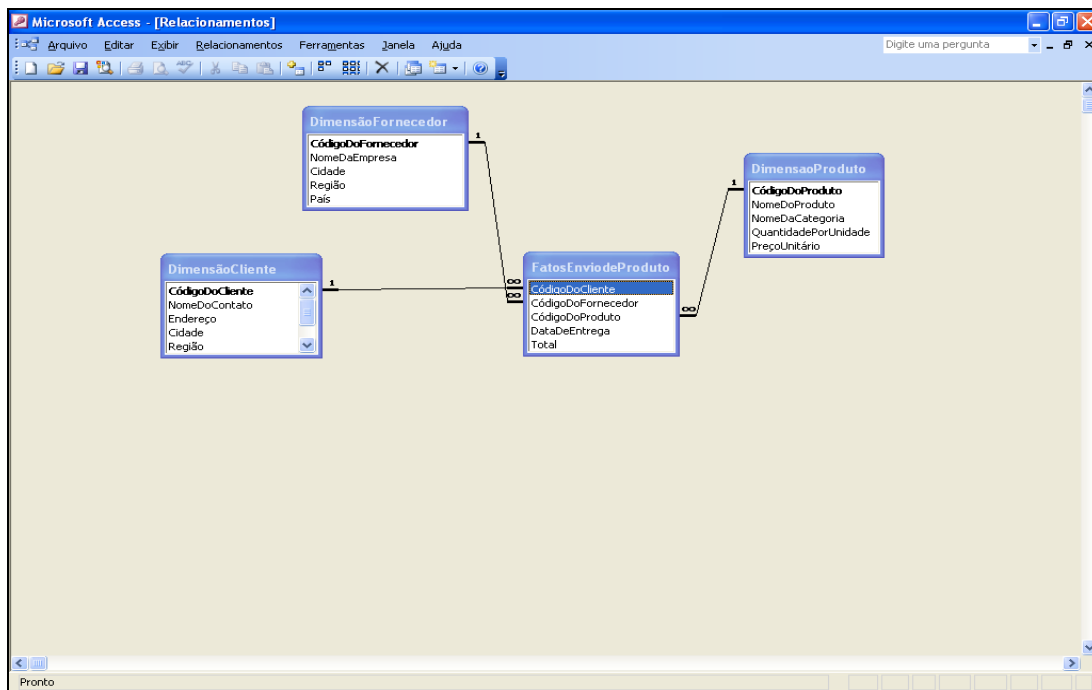


Figura 3 – Tabelas de dimensão e fato relacionadas.

2 – Em seguida, Usar o Excel para fazer a importação dos dados e construção da tabela dinâmica e visualizar cubo OLAP

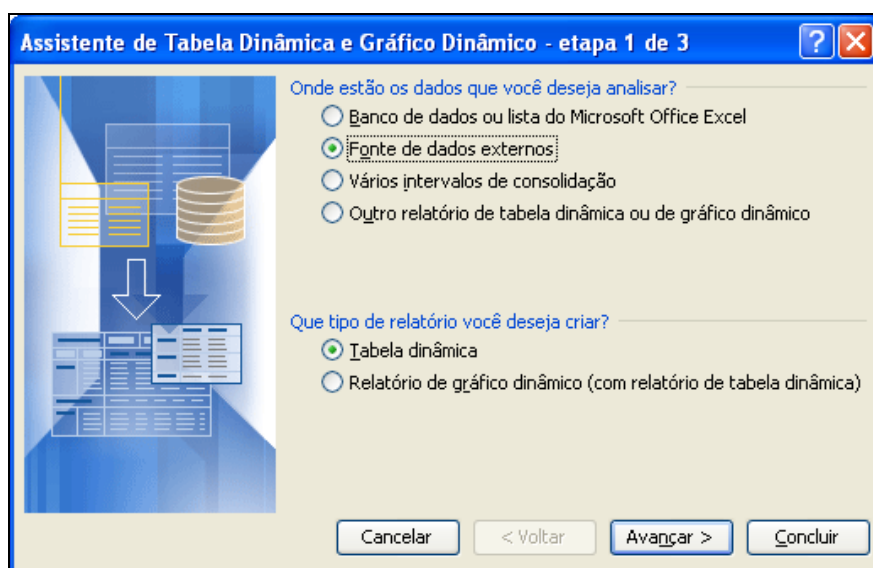


Figura 4 – Caixa de diálogo para conexão em fonte de dados externa.

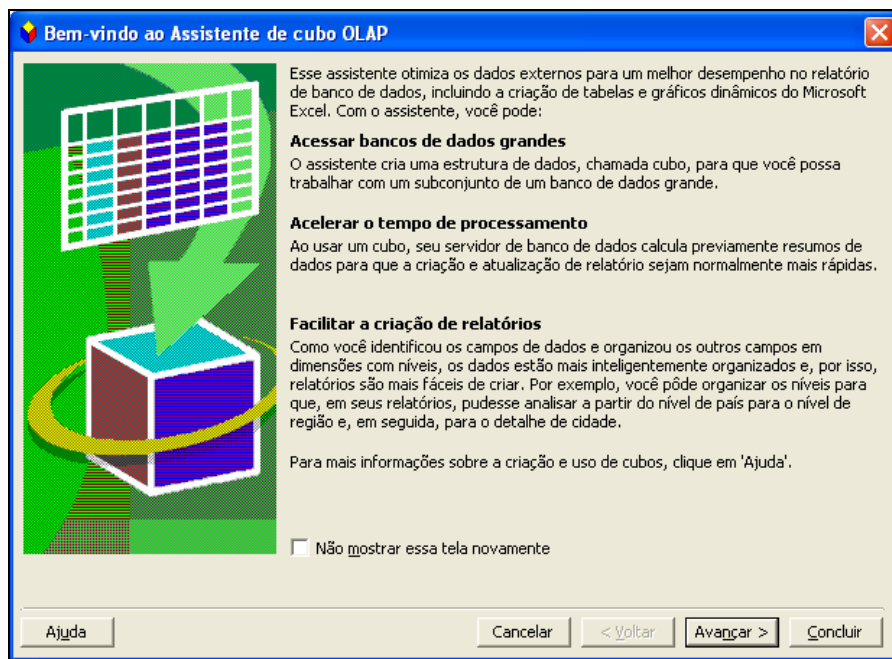


Figura 5 – apresentação do assistente de Cubo OLAP

3 – Visualização dos resultados

	A	B	C	D	E	F	G
1							
2							
3	Soma de Total	NomeDoContato	NomeDoProduto				
4		Alejandra Camino	Alejandra Camino Total	Ana Trujillo	Ana Trujillo Total	Anabela Domingues	
5	NomeDaEmpresa	Camembert Pierrot		Camembert Pierrot		Camembert Pierrot	Chartreuse verte
6	Aux joyeux ecclésiastiques						
7	Exotic Liquids						
8	Gai pâturage	51	51	51	51		51
9	New Orleans Cajun Delights						
10	Total geral	51	51	51	51		51
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							

Figura 6 – Relatório formado pelo Cubo OLAP.

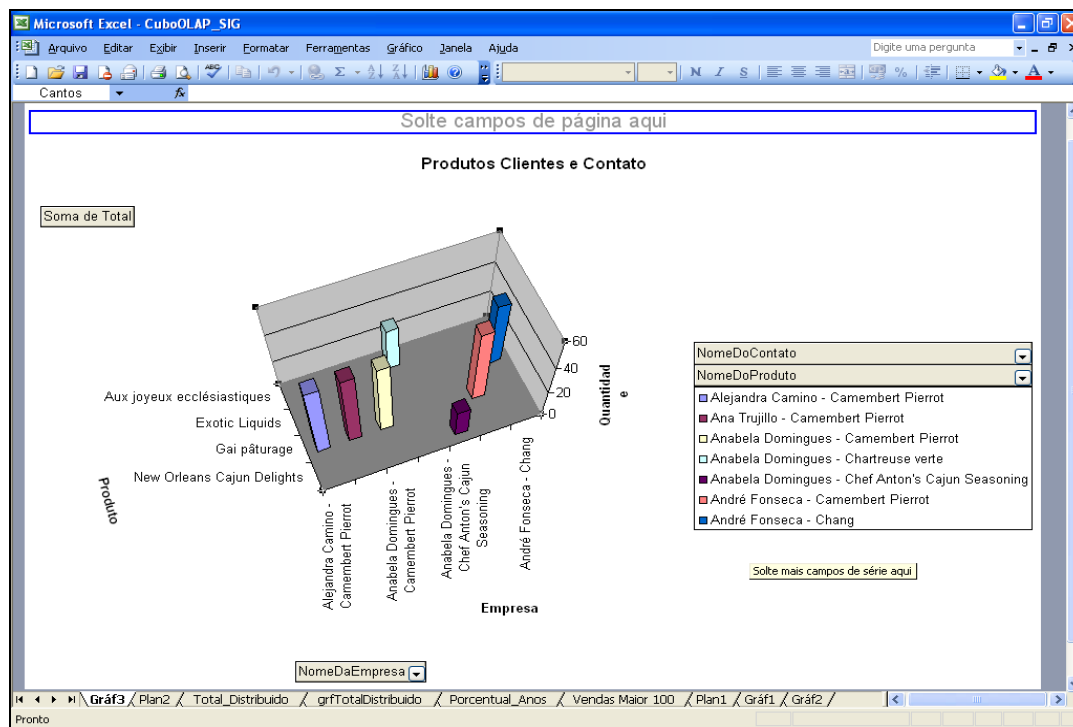


Figura 7 – gráfico em colunas 3D para dimensões e séries.

Mais informações sobre este processo de construção de um sistema de apoio à decisão usando Access e Excel podem ser vistas em [7].

5. Conclusões e Comentários

A implementação de um Sistema de Informação Gerencial com as ferramentas Ms Access e Ms Excel, do pacote Office, (ou BASE e Calc do OpenOffice, última versão) pode ser criado desde que o volume de dados seja suportado pelas ferramentas, o que significa que esse sistema não pode ser usado em grande escala, como por exemplo, em uma rede de hipermercados.

Entretanto, o sistema pode ser desenvolvido para trabalhar dados de um departamento e gerar informações melhores que os relatórios montados apenas em planilha.

Para que o ambiente seja útil, é necessário explorar bem o que o usuário do sistema de informação necessita, na análise de requisitos, pois é o usuário quem pode definir melhor a sua necessidade.

Trabalhando bem todas as etapas de desenvolvimento de software o resultado poder ser um ambiente muito produtivo e com informações melhores quando bem trabalhada.

O sistema de extração de dados deverá ser trabalhado por um especialista em informação gerencial porque qualquer modificação nesses dados pode produzir informação imprecisa e com certeza gerará inconvenientes para o gestor.

O tempo de implantação é relativamente pequeno uma vez que o porte do sistema não exige muito tempo de desenvolvimento, embora a seriedade e responsabilidade pelos dados produzidos e as informações estejam no mesmo grau que no sistema OLAP ou DW.

A base de dados com informações transformadas pode estar instalada ou em um equipamento como servidor ou mesmo no equipamento do gestor e a configuração deve ser pelo menos um equipamento com processador de 500 mhz para suportar mais de um acesso no caso de um servidor.

A qualidade das informações necessitaria de críticas de um analista de negócios para poder validar a funcionalidade do ambiente e os resultados benéficos que pode gerar, através dos relatórios gerados e seus gráficos.

Quando o usuário é bem ambientado nas ferramentas consegue produzir informações rápidas e com qualidade razoável para análise e utilização por gerentes ou supervisores.

Por outro lado, caso haja um crescimento na demanda de relatórios e de dados o sistema pode ser pequeno e o tempo de resposta poderá não ser muito agradável para o usuário quando montado na arquitetura cliente-servidor.

Se houver um crescimento na quantidade de dados armazenadas nos sistemas transacionais passando de algumas centenas de transações por dia para milhares surgirá a necessidade de migração para um sistema mais potente.

Para que um usuário possa aproveitar bem o sistema a necessidade de treinamento na ferramenta Ms Excel se torna imperativa para que possa acessar a base de dados e trabalhá-los e gerar as informações que necessita.

Referências Bibliográficas

- [1] Alenquer, Pablo Lopes, Regras de Negócio para Análise em Ambientes OLAP, UFRJ, 2002. Disponível em:
<http://genesis.nce.ufrj.br/dataware/DataWarehouse/Teses/Pablo/TesePablo.pdf>
- [2] Cielo, Ivã, Artigo OLAP, WEB SITE DW, 2000. Disponível em:
<http://www.datawarehouse.inf.br/metadados.asp>
- [3] DW Brasil, Artigos Bussiness Intelligence, WEB SITE DWBRASIL, sem data.
Disponível em: <http://www.dwbrasil.com.br/html/artigos.html>
- [4] Gama, João, DataWarehouseing, WEB SITE UNIVERSIDADE PORTOCARLENSE, 2002. Disponível em: http://www.liacc.up.pt/~jgama/Mestrado/Bases_Dados/olap.pdf
- [5] Kimball, Ralph, Ross, Margy, The Data Warehouse Toolkit: Guia Completo para Modelagem Dimensional, Editora Campus, 2002.
- [6] Thomsen, Erik, OLAP : Construindo Sistemas de Informação Multidimensionais, Editora Campus, 2002
- [7] Vinci, Wilson Luís. SISTEMA DE INFORMAÇÃO GERENCIAL DE PEQUENO PORTE. Monografia, Faculdades Integradas IPEP, Campinas, 2006.