

Universidade Estadual de Maringá
Centro de Tecnologia
Departamento de Informática
Curso de Engenharia de Produção

**A utilização do Microsoft Excel como ferramenta de
Business Intelligence**

Valmir Pereira Alvas

TCC-EP-78-2008

Universidade Estadual de Maringá
Centro de Tecnologia
Departamento de Informática
Curso de Engenharia de Produção

A Utilização do Microsoft Excel como ferramenta de Business Intelligence

Valmir Pereira Alvas

TCC-EP-78-2008

Trabalho de Conclusão de Curso apresentado Curso de Engenharia de Produção, do Centro de Tecnologia, da Universidade Estadual de Maringá.
Orientador: Prof.^o Lafaiete Henrique Rosa Leme

**Maringá - Paraná
2008**

Valmir Pereira Alvas

A Utilização do Microsoft Excel como Ferramenta de Business Intelligence

Este exemplar corresponde à redação final do Trabalho de Conclusão de Curso aprovado como requisito parcial para obtenção do grau de Bacharel em Engenharia de Produção da Universidade Estadual de Maringá, pela comissão formada pelos professores:

Orientador: Prof^o Lafaiete Henrique Rosa Leme
Departamento de Informática, CTC

Prof^a. Márcia Marcondes Altimari Samed
Departamento de Informática, CTC

Maringá, setembro de 2008

DEDICATÓRIA

Dedico esse trabalho primeiramente a minha Família, mas principalmente ao meu grande amigo e primo Bill, falecido tragicamente neste ano, e que vai fazer uma falta danada para mim e para toda a nossa família.

AGRADECIMENTOS

Agradeço primeiramente a minha Família e colegas de trabalho, que me ajudaram em momentos complicados. Agradeço também ao colega Heber, e finalmente ao Professor Lafaiete.

RESUMO

Com a crescente utilização de sistema de ERP pelas empresas, aumentou-se consideravelmente a quantidade de dados gerados por tais sistemas nas empresas. Tal utilização, porém, não refletiu positivamente na tomada de decisões pelos diretores das empresas, visto que essa grande massa de dados gerados não possuía atributos que auxiliassem na tomada de decisões. Para resolver esse problema, surgiram os sistemas de *Business Intelligence*, que através de várias técnicas conseguem extrair informações desses dados de maneira a ajudar os tomadores de decisão na sua tarefa.

Palavras-chave: *Business Intelligence*, SAD, Sistemas de Apoio à Decisão, ERP, OLAP.

SUMÁRIO

LISTA DE FIGURAS.....	viii
LISTA DE TABELAS.....	iv
LISTA DE ABREVIATURAS E SIGLAS.....	x
1 INTRODUÇÃO.....	1
1.1 JUSTIFICATIVA.....	1
1.2 DEFINIÇÃO E DELIMITAÇÃO DO PROBLEMA.....	1
1.3 OBJETIVOS.....	2
1.3.1 <i>Objetivo geral</i>	2
1.3.2 <i>Objetivos específicos</i>	2
2 REVISÃO DA LITERATURA.....	3
2.1 INTRODUÇÃO AO <i>BUSINESS INTELLIGENCE</i>	4
2.1.1 <i>Histórico</i>	4
2.1.2 <i>Evolução</i>	4
2.2 FERRAMENTAS DE <i>BUSINESS INTELLIGENCE</i>	5
2.2.1 <i>BPM</i>	6
2.2.2 <i>BAM</i>	6
2.2.3 <i>Data Warehouse</i>	7
2.2.4 <i>Metadados</i>	11
2.2.5 <i>Data Mart</i>	11
2.2.6 <i>Data Mining</i>	11
2.2.7 <i>EIS</i>	12
2.2.8 <i>On-Line Analytical Processing</i>	13
2.2.9 <i>DSS</i>	16
2.2.10 <i>Árvore de Decisão</i>	16
2.2.11 <i>Redes Neurais</i>	17
2.3 DICAS NA IMPLEMENTAÇÃO DO <i>BUSINESS INTELLIGENCE</i>	17
3 DESENVOLVIMENTO.....	19
3.1 UTILIZANDO O BROFFICE CALC.....	19
3.2 UTILIZANDO O MICROSOFT EXCEL.....	25
CONCLUSÕES.....	35
SUGESTÃO DE TRABALHOS FUTUROS.....	36

LISTA DE FIGURAS

FIGURA 1 – SELEÇÃO DOS DADOS PARA ANÁLISE NO BROFFICE CALC.....	19
FIGURA 2 – SELEÇÃO DA OPÇÃO DE ASSISTENTE DE DADOS NO BROFFICE CALC.....	20
FIGURA 3 – INDICAÇÃO DA ORIGEM DOS DADOS DO BROFFICE CALC.....	20
FIGURA 4 – ASSISTENTE DE DADOS DO BROFFICE CALC.....	21
FIGURA 5 – EXEMPLO DO ASSISTENTE DE DADOS DO BROFFICE COM CAMPOS ARRASTADOS.....	22
FIGURA 6 – RESULTADO DO ESTUDO DE CASO NO BROFFICE CALC.....	23
FIGURA 7 – SEGUNDO EXEMPLO DO ASSISTENTE DE DADOS DO BROFFICE.....	23
FIGURA 8 – RESULTADO DO SEGUNDO EXEMPLO DO BROFFICE.....	24
FIGURA 9 – ELIMINAÇÃO DO FILTRO NO SEGUNDO EXEMPLO DO BROFFICE.....	24
FIGURA 10 – DADOS CARREGADOS NO MICROSOFT EXCEL.....	26
FIGURA 11 – SELEÇÃO DA OPÇÃO DE TABELAS DINÂMICAS NO EXCEL.....	26
FIGURA 12 – SELEÇÃO DA ORIGEM DOS DADOS NO EXCEL.....	27
FIGURA 13 – SELEÇÃO DAS CÉLULAS DOS DADOS NO EXCEL.....	27
FIGURA 14 – SELEÇÃO DO DESTINO DOS DADOS NO EXCEL.....	28
FIGURA 15 – SELEÇÃO DOS DADOS PARA ANÁLISE NO MICROSOFT EXCEL.....	29
FIGURA 16 – TELA DE ANÁLISE DE DADOS NO MICROSOFT EXCEL.....	30
FIGURA 17 – SELEÇÃO DE DADOS EXTERNOS NO EXCEL.....	31
FIGURA 18 – SELEÇÃO DO BANCO DE DADOS NO EXCEL.....	31
FIGURA 19 – SELEÇÃO DOS CAMPOS NO EXCEL.....	32
FIGURA 20 – MENSAGEM DE RELACIONAMENTO MANUAL DASTABELAS NO EXCEL.....	32
FIGURA 21 – TELA DO MICROSOFT QUERY DO EXCEL.....	33

LISTA DE TABELAS

TABELA 1 – COMPARAÇÃO ENTRE UM BANCO DE DADOS OPERACIONAL E UM DATA WAREHOUSE	8
---	---

LISTA DE ABREVIATURAS E SIGLAS

BI	<i>Business Intelligence</i>
SAD	Sistema de Apoio à Decisão
ERP	<i>Enterprise Resource Planning</i>
DASD	<i>Direct Access Storage Device</i>
SGDB	Sistema Gerenciador de Banco de Dados
EIS	<i>Executive Information System</i>
DW	<i>Data Warehouse</i>
DM	<i>Data Mart</i>
OSD	<i>Operational Data Store</i>
OLAP	<i>On-line Analytical Processing</i>

1 INTRODUÇÃO

Atualmente, com a proliferação dos sistemas de ERP, que permitem uma maior integração dos dados de uma empresa, obtém-se uma grande quantidade de informações do dia a dia da mesma. Porém, tal quantidade de informação não significa necessariamente que de posse delas será facilitada à tomada de decisões.

É nesse contexto que se tornam necessários os sistemas de *Business Intelligence* (BI), também conhecidos com Sistemas de Apoio à Decisão (SAD). Tais sistemas possibilitam a análise dos dados gerados por outros sistemas de maneira a obterem dados que facilitem a tomada de decisões.

Pretende-se com este trabalho demonstrar a utilização do software Microsoft Excel como ferramenta de Business Intelligence nas empresas, visando demonstrar que a utilização dessa ferramenta, relativamente barata, ajudaria sobremaneira na tomada de decisões, principalmente nas empresas de menor porte.

1.1 Justificativa

A escolha do tema deu-se primeiramente ao fato deste autor já ter trabalhado no desenvolvimento de sistemas ERP, principalmente no tocante à parte relacionada ao banco de dados. Tal envolvimento propiciará um melhor aproveitamento da experiência adquirida neste período, além de permitir a aplicação de conceitos adquiridos nas áreas de banco de dados e engenharia de software.

1.2 Definição e delimitação do problema

Pretende-se demonstrar a utilização do Microsoft Excel como ferramenta para obtenção de dados pertinentes à tomada de decisões em uma empresa de vendas de representação comercial, que vende produtos variados, pois atendem empresas de diferentes segmentos de mercado, tais como vassouras e caixas de fósforos.

1.3 Objetivos

1.3.1 Objetivo geral

Demonstrar que é perfeitamente possível a utilização do Microsoft Excel como ferramenta de *Business Intelligence*.

1.3.2 Objetivos específicos

À empresa foi solicitada uma cópia do seu banco de dados, que foi utilizado para o desenvolvimento deste trabalho. Posteriormente, será realizado um treinamento com o proprietário da empresa sobre o uso do Microsoft Excel na tomada de decisão.

2 REVISÃO DA LITERATURA

Segundo HEINRICHS E LIMI (2003), para ter competitividade, é necessário que as empresas detenham muito mais conhecimento do que há pouco tempo atrás e, ainda, para obter sucesso, é necessário que elas saibam mais sobre seus clientes, mercados, tecnologias e processos, e precisam ter essas informações antes que seus concorrentes.

SHIM et al (2002) diz que os SAD são soluções computacionais desenvolvidas para apoiar a tomada de decisões complexas durante a resolução de problemas. Ferramentas clássicas de SAD compreendem componentes para gerenciamento de sofisticados bancos de dados, poderosas funções de modelagem e poderosos, embora simples, projetos de interface com o usuário, que permitem trabalhar interativamente com questões, relatórios e funções gráficas.

CARLSSON e TURBAN (2002) explicam que o termo SAD, propriamente dito, tem sido visto cada vez menos, tanto em revistas quanto em *Websites* de vendas, e no seu lugar tem sido cada vez mais freqüente o uso de termos como *Business Intelligence* e *OLAP*. Por extensão está crescendo o reconhecimento de que *BI* está se tornando um componente necessário na chamada segunda geração dos sistemas *ERP*, que claramente reconhece a necessidade de dar suporte não apenas ao processamento de transações operacionais, mas também ao processamento de análises.

Complementa em GRIGORI et al (2004) que com dados limpos e agregados sobre um determinado processo, armazenados em um *Data Warehouse*, é possível realizar análises utilizando-se tecnologias de *BI* e extrair conhecimento sobre as circunstâncias que levaram a determinado resultado no passado, tenha o resultado sido satisfatório ou não. Assim, é possível utilizar essas informações para explicar por que tais circunstâncias ocorreram e para prever potenciais problemas nos processos em andamento.

2.1 Introdução ao *Business Intelligence*

2.1.1 Histórico

Para PRIMAK (2008), a utilização dos conceitos de *Business Intelligence*, termo cuja paternidade é atribuída à empresa *Gartner Group*, ou simplesmente BI, não é tão recente como se pode imaginar. Seus conceitos práticos já eram utilizados por povos antigos, quando cruzavam informações colhidas na natureza em benefício de suas aldeias.

Embora o mundo em que vivemos tenha mudado muito desde então, o conceito de BI permanece praticamente inalterado, ou seja, atualmente ainda é extremamente importante o cruzamento de informações que auxiliem a tomada de decisões eficazes na gestão de uma empresa (PRIMAK, 2008).

2.1.2 Evolução

PRIMAK (2008) sugeriu que entre trinta e quarenta anos atrás, quando os computadores começaram a diminuir de tamanho e deixaram de ocupar salas enormes, as empresas começaram a perceber que os dados eram importantes fontes geradoras de informações que poderiam auxiliar a tomada de decisões. Nessa época, porém, não existiam recursos de hardware e software eficientes, que possibilitassem a obtenção e análise desses dados de maneira a utilizá-los na tomada de decisões. Nos anos 70, porém, isso começou a mudar com a evolução da forma de armazenamento de dados. O surgimento do DASD (*Direct Access Storage Device* – Dispositivo de Armazenamento de Acesso Direto) e do SGBD (Sistema Gerenciador de Banco de Dados) possibilitaram o estabelecimento de uma única fonte de dados para todo o processamento.

No início dos anos 90 surgiu o *Data Warehouse* (DW) que se trata de uma grande base de dados únicos, já consolidados e organizados considerado por especialistas como fundamental para a prática do *Business Intelligence*. Ainda na década de 90, mais precisamente em 1996, o BI passou a ser considerado uma evolução do EIS (*Executive Information System*), sistema criado no final da década de 70 a partir de trabalhos de pesquisadores do MIT (*Massachusetts Institute of Technology*), nos EUA. Esses sistemas possuem uma tecnologia cujo objetivo

principal é fornecer informações empresariais a partir de uma base de dados (PRIMAK, 2008).

Também em PRIMAK (2008) conclui-se que com a evolução tecnológica dos últimos anos, o BI ganhou muita abrangência e importância com a inclusão de uma série de ferramentas, incluindo a própria EIS além dos DSS (*Decision Support System* – sistemas de suporte à decisão), Planilhas Eletrônicas, Geradores de Consultas e Relatórios, *Data Smarts*, *Data Mining*, Ferramentas OLAP (*On-line Analytical Processing*) entre outras. O BI também está diretamente ligado ao ERP (*Enterprise Resource Planning*), que representa os sistemas empresariais integrados de gestão cuja função é facilitar os processos operacionais nas empresas, tais como controle de estoques, emissão de notas fiscais, pagamento de contas, etc.

2.2 Ferramentas de *Business Intelligence*

Para PRIMAK (2008), os sistemas integrados de gestão que constituem a base dos processos das empresas são estruturados de forma transacional, o que dificulta o seu tratamento informacional. É aí que entram os Sistemas de BI, que permitem o desenvolvimento de estruturas especiais de armazenamento de informações como o *Data Warehouse* (DW), *Data Mart* (DM) e o *Operational Data Store* (OSD). Os dois primeiros referem-se a estruturas de dados remodelados para análises diferenciais e o último trata da análise consolidada, porém, sem as características dimensionais.

O BI também possibilita o desenvolvimento de aplicações especiais de tratamento de dados como a *On-line Analytical Processing* (OLAP) e o *Data Mining*. O OLAP possibilita o tratamento dos dados de forma dimensional, possibilitando uma forma múltipla e combinada de análise. Já o *Data Mining* está mais relacionado com a análise de inferência baseada em algoritmos que objetivam o reconhecimento de padrões escondidos nos dados (PRIMAK, 2008).

A seguir uma análise das principais ferramentas do *Business Intelligence*.

2.2.1 BPM

O BPM (*Business Performance Management*) permite atuar no negócio de maneira ativa em direção ao desejado, em vez de apenas criticar o ocorrido após ter ocorrido. Tem como objetivo melhorar a capacidade de entendimento dos negócios e gerenciar seu desempenho em todos os níveis, incluindo acionistas, gerentes, funcionários, clientes e fornecedores dentro de um ambiente de desenvolvimento integrado. (PRIMAK, 2008).

Existem basicamente dois tipos de BPM, um relacionado a medidas, ou *Business Performance Measurement* e outro ligado ao gerenciamento, ou *Business Performance Management*.

Em PRIMAK (2008), temos a definição dos tipos de BMP. O de medidas é uma aplicação que trabalha com KPI (*Key Performances Indicators* – indicadores chaves de desempenho) que permitem medir o desempenho ou ainda utilizam informações comparativas do mercado (comparações com os concorrentes). Já o BMP de gerenciamento se apóia na modelagem ou exploração de cenários que permitem ao usuário considerar as implicações de cursos alternativos ao invés de apenas explorar os fatos ocorridos e os seus motivos.

2.2.2 BAM

O BAM (*Business Activity Monitoring* – monitoramento da atividade de negócio) é uma evolução do BMP. A idéia é ter uma interface parecida a de um velocímetro de um automóvel e nela coloque dados críticos de desempenho em tempo real, tais como vendas por hora, por exemplo (PRIMAK, 2008).

A comparação com o velocímetro do carro é mantido quando se tem em mente que seu objetivo final é que o executivo, ao avaliar as medidas do BAM, deve tomar providências a respeito, fato que também ocorreria se estivesse dirigindo um automóvel e necessitasse de adaptar a velocidade às condições da estrada, por exemplo.

PRIMAK (2008) explica os três tipos de BAMs existentes:

- a) Métricas de Processo: Tecnologias de integração de processos que mostram a integração em tempo real como parte do mecanismo de integração. Permite criar metaprocessos acima dos processos existentes na empresa.
- b) BAM Passivo: Servidores de integração que permitem mostrar as informações em tempo real de forma simples para os usuários finais. Os quais podem observar o estado de seus negócios e podem fazer as mudanças que julgar necessárias, embora tais mudanças não sejam implementadas pelo BAM.
- c) BAM Ativo: Possibilita, além de monitorar e calcular usando os dados oferecidos em tempo real, também tomar decisões usando lógicas pré-programadas.

2.2.3 *Data Warehouse*

Ao longo do tempo, os bancos de dados foram desenvolvidos para fins de processamentos de dados operacionais e analíticos, havendo maior ênfase no primeiro caso, ainda que ambos tivessem usuários com diferentes necessidades. Uma vez compreendida essa diferença, foram criados bancos de dados separados para fins analíticos, chamados de *Data Warehouse*, explicam GRAY e WATSON (1999).

Para LEME (2004), o *Data Warehouse* nada mais é do que um grande armazém de dados. Ele é, talvez, o mais importante componente que possibilita a criação de um bom projeto de *Business Intelligence*.

Já para PRIMAK (2008), o perfeito entendimento de *Data Warehouse* passa pelo entendimento do conceito tradicional de banco de dados, para que se possa fazer uma comparação entre ambos.

Conforme HARDING e YU (1999), o *Data Warehouse* extrai, de uma maneira eficiente, informações valiosas dos bancos de dados operacionais que estão espalhados por toda a empresa e que, certamente, contêm dados duplicados ou parcialmente duplicados. Por isso, antes de povoar o *Data Warehouse* com os dados operacionais, estes dados precisam ser preparados para estarem aptos a serem questionados, pesquisados, analisados e apresentarem conclusões. Nas aplicações de *BI*, o foco está mais voltado para as tendências e agregações do

que para cada transação individual; portanto, os dados armazenados devem oferecer níveis de agregação apropriados.

“Um banco de dados é uma coleção de dados operacionais armazenados e utilizados pelo sistema de aplicações de uma empresa específica. Esses dados são chamados de operacionais ou primitivos”, define BATINI (1992).

A partir desta definição sobre os dados operacionais, pode-se dizer que o *Data Warehouse* (DW) é uma coleção de dados derivados dos dados operacionais, também definidos como gerenciais, informacionais ou ainda analíticos. Em resumo, os sistemas transacionais das empresas são onde os dados são colocados enquanto que o DW é onde os dados são enriquecidos e pesquisados, conclui LEME (2004).

Abaixo uma tabela, extraída de PRIMAK (2008), contendo as principais diferenças entre um banco de dados operacional e o *Data Warehouse*.

Tabela 1 – Comparação entre um banco de dados operacional e um *Data Warehouse*

Características	Banco de Dados Operacional	Data Warehouse
Objetivo	Operações diárias do negócio	Análise do negócio
Uso	Operacional	Informativo
Tipo de Processamento	OLTP	OLAP
Unidade de Trabalho	Inclusão, Alteração, Exclusão	Carga e consulta
Número de Usuários	Milhares	Centenas
Tipo de Usuário	Operadores	Comunidade Gerencial
Interação do Usuário	Somente pré-definida	Pré-definida e <i>ad-hoc</i>
Volume	Megabytes – Gigabytes	Gigabytes – Terabytes
Histórico	60 a 90 dias	5 a 10 anos
Granularidade	Detalhados	Detalhados e resumidos
Redundância	Não ocorre	Ocorre
Estrutura	Estática	Variável
Manutenção desejada	Mínima	Constante
Acesso a registros	Dezenas	Milhares
Atualização	Contínua (tempo real)	Periódica (em lote)
Integridade	Transação	A cada atualização
Número de Índices	Poucos / Simples	Muitos / complexos
Intenção dos Índices	Localizar um registro	Aperfeiçoar consultas

Fonte: Primak (2008)

A partir desses conceitos percebe-se que os dados contidos no DW são resumos provenientes de múltiplos sistemas utilizados normalmente há muito tempo e que ainda estão em operação.

Os DW são construídos de tal forma que os dados sejam armazenados e consultados de forma que não estejam limitados a tabelas e linhas estritamente relacionais,

Os dados do DW devem estar separados fisicamente dos dados operacionais e antes de serem deslocados dos sistemas operacionais normalmente devem sofrer alterações através de técnicas de filtragem, agrupamento e/ou refinamento dos dados, explica PRIMAK (2008).

PRIMAK (2008) também explica que o DW tem por objetivo organizar os dados corporativos de maneira integrada, com uma única versão da verdade, ou seja, possuindo uma única cópia, histórico, variável com o tempo e gerando uma única fonte de dados, usada para abastecer os *Data Marts* (DM).

Para essa organização, novos métodos de armazenamento, estruturação e novas tecnologias são necessárias para que permitam a geração e recuperação de tais informações. Tais tecnologias diferem dos padrões operacionais de sistemas de banco de dados de três maneiras, conclui PRIMAK (2008):

- a) Permitem visualizações informativas, pesquisando, reportando e modelando capacidades que vão além dos padrões de sistemas operacionais freqüentemente oferecidos;
- b) Armazenam dados freqüentemente em formato de cubo (OLAP) multidimensional, que permite rápida agregação de dados e detalhamento de análises;
- c) Disponibilizam habilidades para extrair, tratar e agregar dados de múltiplos sistemas operacionais em *Data Marts* ou *Data Warehouses* separados.

Para LEME (2004), o DW possui algumas características importantes, vejamos algumas dessas características:

- a) Não volátil: Os dados nunca são alterados ou apagados, pois, num DW os dados são apenas incluídos e posteriormente acessados.

- b) **Estrutura por Assunto:** Diferentemente dos sistemas operacionais, onde os dados são estruturados por aplicação, onde temos, por exemplo, estoque, faturamento, financeiro, etc. No DW os dados são estruturados por assunto, mais especificamente pelos assuntos do negócio da empresa, o que vem a ser essencial para a tomada de decisões.
- c) **Integração:** Talvez a característica mais importante do DW. A integração dos dados e de conceitos. Essa integração é obtida graças à consolidação e tratamento que os dados dos sistemas operacionais recebem antes de serem incluídos no DW. Isso permite que esses dados tenham um único significado para todos que os consultarem.
- d) **Variável com o Tempo:** Os dados num DW são precisos em algum instante do tempo. Tal característica possibilita aos administradores analisar a evolução do negócio ao longo de um período, possibilitando analisar sazonalidades, eventos externos que influenciam diretamente os rumos da empresa.

PRIMAK (2008) define que várias são as maneiras de se recuperar uma informação no DW, As mais comuns são: relatórios, consultas, os EIS, ferramentas que utilizam características OLAP e também as ferramentas de *Data Mining*. A tendência dessas soluções está na integração com a Web, o que permitirá maior agilidade em consultas.

O acesso aos dados pode ser realizado de duas maneiras: direta ou indiretamente. Para o acesso direto é necessário que a aplicação do ambiente operacional faça uma solicitação diretamente aos dados localizados no DW. A solicitação é passada então ao ambiente do DW que localiza os dados e os repassa ao ambiente operacional. Embora pareça simples, esse acesso ocorre de maneira rara, visto que uma solicitação poderá sofrer uma série de limitações, dentre elas o tempo de solicitação, que não possibilita ao processamento operacional ter qualquer exigência de tempo. A forma de acesso indireto é mais eficiente e rápida. Isso é possível graças a um programa que é rodado periodicamente nos dados que analisa as características e critérios previamente definidos. Tal análise cria um ambiente operacional que possibilita o acesso rápido aos dados do DW. Este programa, que normalmente possui características de inteligência artificial, e também é executado em segundo plano, permitindo que o DW possa ser acessado por outras tarefas, comenta PRIMAK (2008).

2.2.4 Metadados

PRIMAK (2008) definiu Metadados como sendo os “dados dos dados”. Num DW, os metadados são fundamentais na elaboração do mesmo. Isso ocorre devido à grande documentação gerada em um DW, diferentemente de um banco de dados tradicional.

O metadado de um DW é gerado a partir de um levantamento dos relatórios a serem gerados, de onde vêm os dados que alimentam o DW, os processos de extração, o tratamento de rotinas de cargas de dados. Além disso, também são adicionadas as regras de negócios da empresa, as mudanças ao longo do tempo e a frequência de acesso aos dados, concluiu também PRIMAK (2008).

2.2.5 Data Mart

Para PRIMAK (2008), o *Data Mart* é na verdade um *Data Warehouse* reduzido, utilizado apenas por pequenos grupos dentro da organização. São normalmente preferidos em relação a um DW devido ao custo reduzido e ao menor tempo de implementação. Por fornecer os Sistemas de Apoio a Decisão para pequenos grupos, o *Data Mart* pode ser utilizado como um teste para a implementação de um *Data Warehouse* para os outros setores da empresa.

Deve-se ter em mente que tanto o *Data Mart* como o *Data Warehouse* possuem basicamente as mesmas definições de problemas e os requisitos de dados. O que muda é apenas o tamanho e o escopo do problema. Enquanto o *Data Mart* trata o problema departamental ou local, o DW envolve o esforço de toda a companhia para que o suporte a decisão atue em todos os níveis da organização.

2.2.6 Data Mining

Conforme BISPO e CAZARINI (1998), o objetivo desta ferramenta é extrair informações potencialmente úteis a partir de dados “crus” armazenados em um *Data Warehouse* ou em bancos de dados dos sistemas operacionais, porém, preferencialmente nos primeiros, por já estarem consolidados. Para tal, utilizam-se técnicas sofisticadas de procura, como algoritmos de Inteligência Artificial, Redes Neurais, Árvores de Decisões, Regras de Indução, ou ainda,

combinações entre elas. O resultado obtido pela aplicação do *Data Mining* deve ser compacto, legível (apresentado de alguma forma simbólica), interpretável e deve representar fielmente os dados que lhe deram origem. Isso é necessário porque as pessoas normalmente estão mais interessadas em obter o conhecimento propriamente dito do que os modelos matemáticos, afirma BÄCK (2002).

Segundo PRIMAK (2008), o *Data Mining* é uma das respostas para um dos maiores problemas relacionados à análise e transformação dos dados em informações. Como fazê-lo em um menor tempo possível e de forma automática. Para o *Data Mining* a solução é utilizar uma combinação de estatística convencional com técnicas de inteligência artificial.

Os processos de *Data Mining* são extremamente complexos, tanto que dificilmente apenas um especialista será capaz de resolvê-los, sendo sempre necessário mais de um profissional especializado. Também é obrigatória a participação de um profissional com o conhecimento do negócio, que deverá possuir domínio total do assunto, evitando assim que a modelagem acabe não tendo utilidade no apoio a decisão (PRIMAK, 2008).

PRIMAK (2008) comenta que várias técnicas de *Data Mining* podem ser usadas para a tomada de decisão. Dentre elas podemos citar: Análise de Conglomerados, Métodos Preditivos com Séries Temporais, Árvores de Decisão, Redes Neurais e Algoritmos Genéticos.

2.2.7 EIS

Para PRIMAK (2008), o *Executive Information System* (EIS) é um sistema voltado a atender as necessidades dos altos executivos da empresa. Ele fornece informações gerenciais de forma rápida e simples. Deve possuir uma interface amigável, já que seus usuários estão sempre muito ocupados. Os dados mostrados em sua interface sempre devem mostrar as informações de maneira resumida, já que as decisões neste nível administrativo não se atêm aos detalhes.

Embora possa ser construído tomando como base os dados de sistemas transacionais, o ideal é que sejam buscados dados do *Data Warehouse* o que facilitaria a busca visto que no DW os dados já estão agrupados em uma única base de dados (PRIMAK, 2008).

Para PRIMAK (2008) o EIS deve possuir recursos de customização, para que possa se adaptar ao usuário que irá operá-lo. Também deve apresentar recursos gráficos que facilitem a visualização e interpretação das informações. Assim como o usuário não tem muito tempo, o seu uso não deve demandar muito treinamento para que possam ser utilizados.

Embora os dados mostrados sejam sempre resumidos, o EIS deve permitir ao usuário aumentar o nível de detalhamento dos dados, utilizando uma ferramenta OLAP. Isso possibilita, numa consulta onde é mostrado o total de vendas do ano, por exemplo, que seja possível consultar os valores mês a mês (PRIMAK, 2008).

PRIMAK (2008), conclui sugerindo que embora os conceitos de EIS e de SAD possam gerar certa confusão, por estarem relacionados, tratam de problemas diferenciados, o que faz com que atendam a um público-alvo diferente. Enquanto o EIS é voltado para os executivos, o SAD tem como usuários os níveis gerenciais intermediários.

2.2.8 On-Line Analytical Processing

BISPO e CAZARINI (1998) apresentam o *OLAP* como uma ferramenta capaz de efetuar análises de dados com visão multidimensional do negócio, comparando-os por diversos ângulos.

Os cubos são os principais objetos de um *OLAP*. Construídos com tecnologia que permite rápido acesso aos dados, normalmente eles são construídos a partir de subconjuntos de um *Data Warehouse* e são organizados e sumarizados dentro de estruturas multidimensionais definidas por dimensões e medidas. Os bancos de dados multidimensionais guardam as informações numa espécie de cubo de dimensão “n”, o que resulta em diversas matrizes esparsas que permitem trabalhar simultaneamente com diversos cenários definidos por combinações de dados, como produtos, região, vendas, tempo, etc. A vantagem dos bancos de dados multidimensionais sobre os bancos de dados relacionais é que os primeiros estão otimizados para obter velocidade e facilidade de resposta, complementam GRAY e WATSON (1999).

Segundo PRIMAK (2008), também é possível analisar as informações sob diferentes prismas, através da técnica do *slice-and-dice*. Já o *Drill Down/Up* permite fazer uma exploração em diferentes níveis de detalhes das informações. Isso possibilita ao usuário “subir e descer” dentro do detalhamento dos dados, sendo possível, por exemplo, partir de uma consulta de valores anual para uma mensal e finalmente para uma consulta diária.

O surgimento dessa ferramenta veio facilitar a vida das empresas, que antes, quando necessitassem, por exemplo, de um relatório que listassem quais foram os clientes que mais compraram num determinado semestre, perdiam muito tempo entre a localização dos dados e construção propriamente dita do relatório. Mesmo após o surgimento dos SGBD e a evolução das linguagens de programação, que facilitou a vida dos analistas de sistemas e conseqüentemente facilitou o acesso aos dados, o panorama não melhorou muito para quem necessitasse de um relatório, visto que ainda precisaria do auxílio de um técnico da área de informática (PRIMAK, 2008).

PRIMAK (2008) conclui explicando que com o OLAP, o acesso aos dados pelos usuários finais tornou-se mais fácil, visto que o próprio usuário pode elaborar a consulta que desejar, graças à interface amigável do OLAP, sem necessitar do auxílio de um técnico, como ocorria anteriormente.

Abaixo as 12 regras que caracterizam uma ferramenta OLAP, que foram citadas pela primeira vez em 1992 em um artigo escrito por E.F.Codd:

- a) Visão conceitual multidimensional;
- b) Transparência;
- c) Acessibilidade;
- d) Informações de performance consistente;
- e) Arquitetura cliente-servidor;
- f) Dimensionalidade genética;
- g) Manipulação de matrizes dinamicamente;
- h) Suporte a multiusuários;
- i) Operações ilimitadas em referências cruzadas;
- j) Manipulação de dados intuitivamente;
- k) Consultas flexíveis;

1) Níveis de dimensões a agregações ilimitadas.

Entre as funções mais tradicionais de uma ferramenta OLAP estão a criação de tabelas cruzadas, a explosão de informações e a criação de dimensões. Ao trabalhar de modo interativo, a ferramenta cria outros questionamentos a partir de uma resposta do usuário, permitindo ao usuário analisar os porquês dos resultados (PRIMAK, 2008).

Um exemplo desses questionamentos seria: “Quais foram os 10 produtos que obtiveram o maior lucro no primeiro trimestre de 2008”. A resposta a este e a outros questionamentos é baseada em fatos históricos que vão mostrar uma tendência de comportamento das variáveis selecionadas (PRIMAK, 2008).

LEME (2004), explica as variações de modelos possíveis dentro da ferramenta OLAP:

ROLAP – Relational On-line Analytical Processing – modalidade onde os dados utilizados serão localizados através do próprio SGBD, com as tabelas implementadas da maneira relacional clássica. Embora isso traga as vantagens da utilização de um SGBD, tem como desvantagem a necessidade de um gerenciamento do desempenho, que pode ser comprometido pelo excesso de tabelas.

MOLAP – Multi Dimensional On-line Analytical Processing – também localiza os dados através de um SGBD, porém, esse SGBD possui características especiais de armazenamento além de ferramentas para o tratamento dimensional de dados. A desvantagem é que neste caso não se ter todos os recursos do SGBD, além de necessitar de migração dos dados da estrutura relacional para a multidimensional e também de atualizações constantes.

HOLAP – Hybrid On-line Analytical Processing – misto entre o ROLAP e o MOLAP. Estruturas relacionais são utilizadas para armazenamento de dados com maior granularidade e estruturas dimensionais são utilizadas no armazenamento de dados consolidados.

DOLAP – Desktop On-line Analytical Processing – as estruturas dimensionais ou multidimensionais são transferidas para a estação de trabalho do cliente, objetivando facilitar o desempenho, minimizando o tráfego de informações na rede da empresa.

2.2.9 DSS

Surgidos a partir dos sistemas transacionais das empresas, os *Decision Support Systems*, ou Sistemas de Suporte à Decisão, auxiliam no processo decisório utilizando modelos para resolver problemas não estruturados. Para sua utilização é necessário primeiramente definir quais dados serão gerados a partir dos sistemas transacionais, esses dados devem então ser armazenados no *Data Warehouse*. Em seguida deve-se partir para a modelagem a partir da qual será gerado o modelo físico (PRIMAK, 2008).

Com o modelo físico pronto, os dados devem ser extraídos do DW através de rotinas de extração que podem ser desenvolvidas por um programador em qualquer linguagem de programação. Finalmente, deve ser feita uma checagem da consistência dos dados, pois qualquer erro nos mesmos poderá causar problemas em uma análise futura (PRIMAK, 2008).

PRIMAK (2008) finaliza comentando que embora pareçam aparentemente simples, os DSS exigem um profundo conhecimento técnico e de negócios para que sejam utilizados em êxito.

2.2.10 Árvore de Decisão

PRIMAK (2008) define a Árvore de Decisão como uma técnica utilizada para criar e organizar regras de classificação e decisão em formato de diagrama de árvores, a partir de uma grande massa de dados, como um *Data Mart* ou *Data Warehouse*. As regras de classificação criadas possibilitarão prever resultados futuros a partir dos dados atuais.

A árvore inicia-se com um único grupo que reúne todos os casos a serem estudados. A base então é dividida em nódulos que representam categorias das variáveis analisadas, à medida que a árvore vai se expandindo. Cada galho formado então se abre em subgrupos que são mutuamente exclusivos (PRIMAK, 2008).

2.2.11 Redes Neurais

Cada vez mais utilizadas em *Data Mining*, as Redes Neurais têm como vantagem a aprendizagem a partir das experiências, não ficando restrita a uma ordem pré-estabelecida. São compostas de algoritmos e procedimentos computacionais que imitam a capacidade de aprendizagem do cérebro humano. A técnica consiste na formação de nódulos cujo processamento se assemelha ao dos neurônios, o que deu nome á técnica (PRIMAK, 2008).

Ainda segundo PRIMAK (2008), os nódulos devem ser conectados como em uma rede, onde funcionam paralelamente. A primeira fase de nódulos é composta pelos nódulos de entrada, que recebem a entrada das variáveis fornecidas pelo banco de dados. Essa entrada é então transformada de acordo com a função de ativação, produzindo uma saída que será enviada à entrada de um nódulo da fase de nódulos seguinte.

A fase de nódulos é chamada de nódulos ocultos, que podem formar diversas camadas, dependendo da complexidade da rede em questão. Finalmente, tem-se os nódulos que saída, que recebem uma entrada, processam, porém não a enviam a outro nódulo, já que este é o nódulo que dará a resposta final da rede (PRIMAK, 2008).

2.3 Dicas na implementação do *Business Intelligence*

Para Boris Evelson, analista da *Forrest Research*, é necessário estabelecer a visão estratégica antes de iniciar um processo de implementação de *Business Intelligence* (BI) numa organização. Abaixo algumas dicas que devem ser seguidas:

- a) A implementação de BI não deve, em absoluto, ser responsabilidade de alguém de TI. Ao invés, o BI deveria estar nas mãos de um executivo da linha de frente. Alguém que tenha uma boa idéia dos objetivos, estratégias e metas da companhia e que saiba traduzir a missão em indicadores de desempenho que a suportarão. Essa pessoa deve governar a implementação com documentação adequada do *case* de negócio e ser responsável por mudanças;
- b) Crie definições comuns. Sem definições comuns, a implementação de BI não pode ser bem sucedida. A falta de acordo é um problema comum das empresas hoje. Por

exemplo, as áreas de finanças e vendas podem definir “faturamento” de forma diferente;

- c) Tenha acesso à atual situação. Você deve analisar em que ponto está o BI (processos, estrutura organizacional). Tanto a área de tecnologia da informação como de negócio devem estar envolvidas;
- d) Crie um plano para armazenamento de dados. Muitas organizações começam com apenas um centro comercial, já que é mais barato e rápido, mas considere que essa tática significa uma nova construção assim que as informações precisarem de mais espaço, o que pode sair de controle em alguns anos;
- e) Entenda o que os usuários precisam. As três partes da estratégia de BI são: estratégica, tática e operacional. Usuários estratégicos tomam poucas decisões, mas cada uma delas pode ter um efeito profundo – por exemplo, deveríamos fechar a filial em uma cidade e abri-la em outra. Usuários táticos tomam várias decisões por semana, e precisam de atualizações diárias. Usuários operacionais são os funcionários de linha de frente, como os atendentes de *call center*. Precisam de dados com ajustes próprios para realizar uma quantidade imensa de transações. Entender quem usará o BI e para que propósito, pode mostrar o tipo de informação necessária e com que frequência, e assim guiar as decisões de BI.
- f) Escolha algo simples para começar. Recomenda-se escolher componentes de alto valor, mas simples para começar. Por exemplo, dados de análise de vendas podem trazer alto valor e têm uma série de modelos no mercado com as melhores práticas.

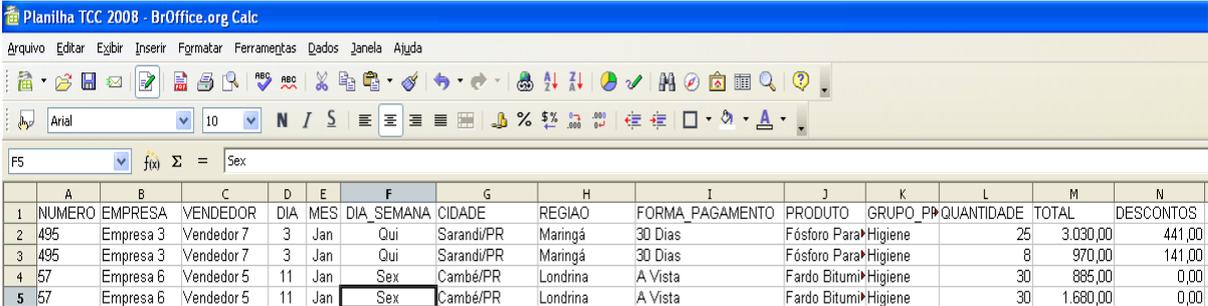
3 DESENVOLVIMENTO

Para o desenvolvimento do estudo de caso que pretende mostrar como utilizar o Excel como ferramenta de BI, foram utilizados dados de uma empresa de Representação Comercial, que representa vários tipos de produtos para empresas de seguimentos de mercados diferentes. Essa base de dados possui cerca de 5000 tuplas que estão armazenadas num banco de dados Firebird.

3.1 Utilizando o BrOffice Calc

Embora o objetivo seja mostrar a utilização do Microsoft Excel como ferramenta de BI, iremos mostrar também a utilização do programa BrOffice Calc com o mesmo objetivo. Isso foi feito para possibilitar a utilização de um programa gratuito, que pode ser necessário para o caso do Microsoft Excel não estar disponível, por se tratar de um programa que para ser utilizado necessita do pagamento de uma licença de uso.

Primeiramente, será necessário que os dados que serão analisados estejam carregados no programa, ou mais precisamente, numa planilha. Cada informação deverá estar em uma coluna e a primeira linha deve conter a identificação de cada campo, conforme a Figura 1.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	NUMERO	EMPRESA	VENDEDOR	DIA	MES	DIA_SEMANA	CIDADE	REGIAO	FORMA PAGAMENTO	PRODUTO	GRUPO_PROD	QUANTIDADE	TOTAL	DESCONTOS
2	495	Empresa 3	Vendedor 7	3	Jan	Qui	Sarandi/PR	Maringá	30 Dias	Fósforo Para	Higiene	25	3.030,00	441,00
3	495	Empresa 3	Vendedor 7	3	Jan	Qui	Sarandi/PR	Maringá	30 Dias	Fósforo Para	Higiene	8	970,00	141,00
4	57	Empresa 6	Vendedor 5	11	Jan	Sex	Cambé/PR	Londrina	A Vista	Fardo Bitumi	Higiene	30	885,00	0,00
5	57	Empresa 6	Vendedor 5	11	Jan	Sex	Cambé/PR	Londrina	A Vista	Fardo Bitumi	Higiene	30	1.680,00	0,00

Figura 1 – Seleção dos Dados para Análise no BrOffice Calc

Carregados os dados, deve-se posicionar o cursor na primeira linha e coluna e em seguida o menu “Dados” deve ser selecionada e a opção “Assistente de Dados / Iniciar” ativada, conforme mostra a Figura 2.

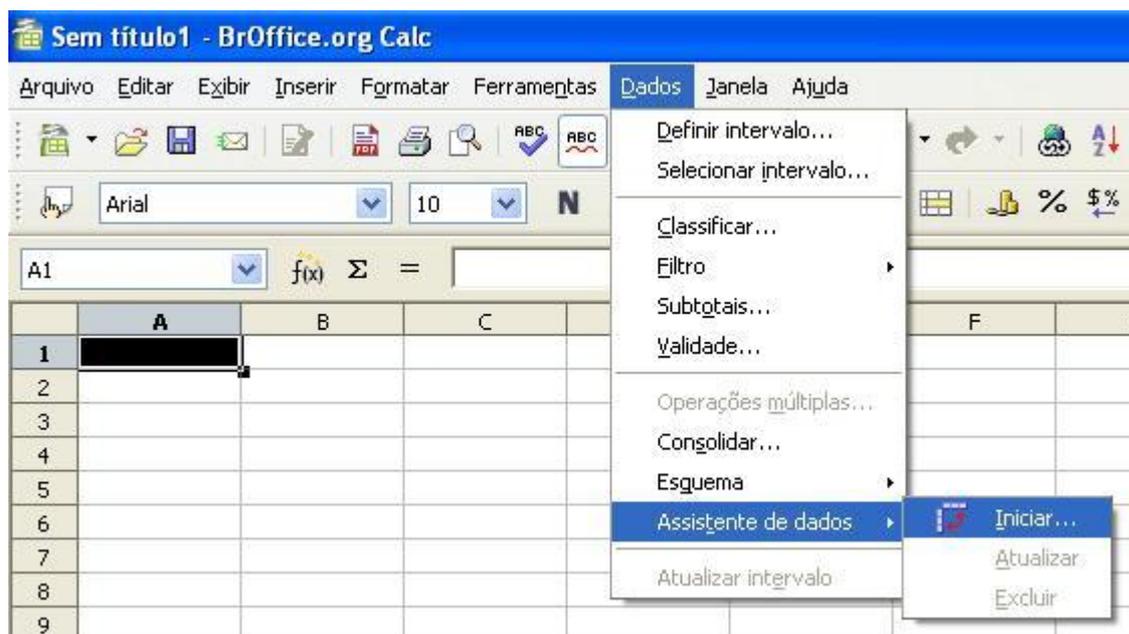


Figura 2 – Seleção da opção de Assistente de Dados no BrOffice Calc.

Após selecionar a opção, uma pequena janela, mostrada na Figura 3, será mostrada.

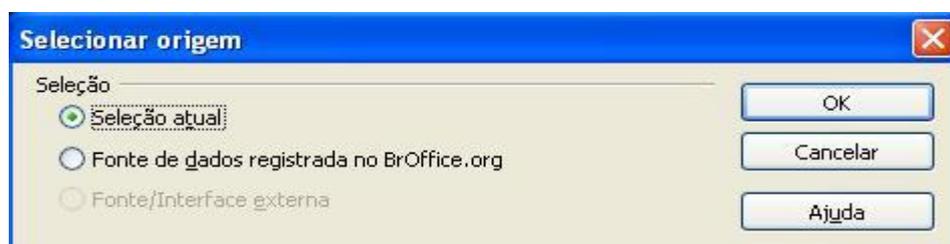


Figura 3 – Indicação da Origem dos dados do BrOffice Calc.

Deve-se então deixar a opção “Seleção Atual” marcada e clicar em “OK”. Será então mostrada a janela na Figura 4.

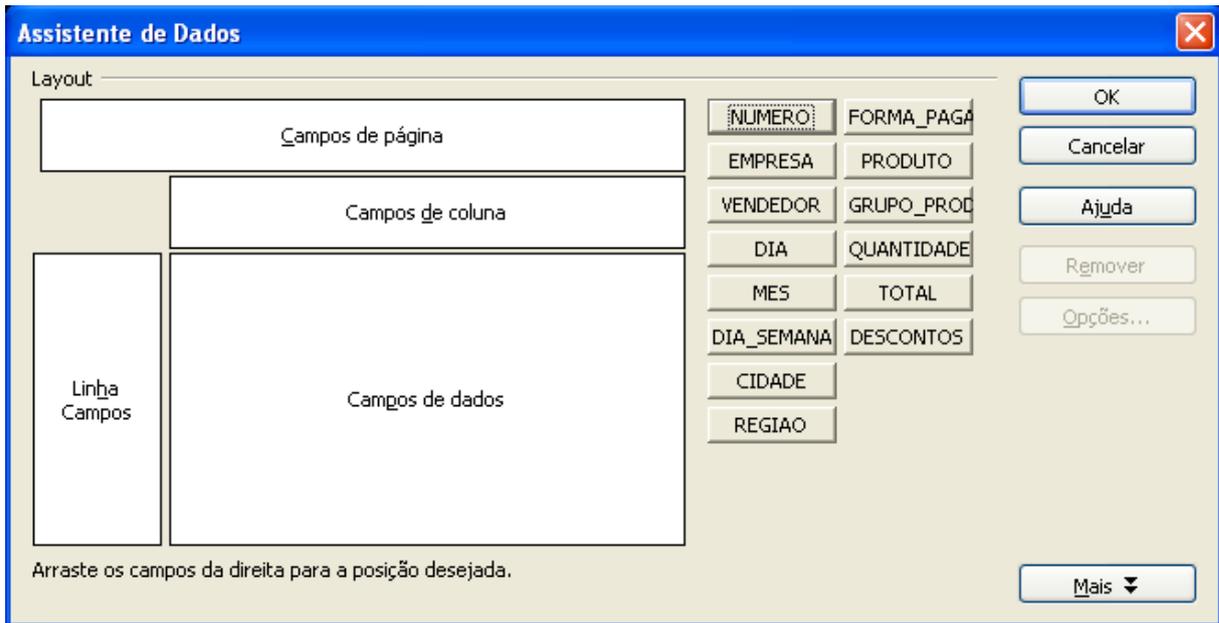


Figura 4 – Assistente de Dados do BrOffice Calc.

Nesta tela, os campos devem ser arrastados para os espaços correspondentes, que tratarão a informação, dependendo do seu tipo:

Os “Campos de página” servem para definição de filtros, ou seja, ao colocar um ou mais campos nesta área, será possível selecionar um dado desse campo de maneira que os dados consultados que serão cruzados refiram-se apenas ao selecionado. Pode-se, por exemplo, colocar nessa área os campos “EMPRESA” e “VENDEDOR”. Dessa forma, será possível depurar a consulta dos dados selecionando uma única empresa e/ou produto.

Os “Campos de Coluna” e “Linha Campos” servem para indicar os dados que serão cruzados, ou seja, nessa área devem ser colocados os campos cujos dados serão diretamente comparados. Arrastando-se para essa área o campo “REGIÃO” e para a área “Campos de coluna” o campo “DIA_SEMANA”, será possível fazer uma avaliação do comportamento das vendas em cada região conforme o dia da semana.

Para “Campos de Dados” devem ser arrastados os campos que serão analisados no cruzamento de dados. Ao arrastar para essa área o campo “QUANTIDADE” será possível comparar as vendas dos produtos baseados na quantidade das vendas.

Essa comparação pode ser feita em mais de uma informação ao mesmo tempo, bastando para isso arrastar outros campos para a área desejada. O campo “TOTAL” também pode ser arrastado para a área “Campos de Dados”, o que permitirá uma comparação dos dados tanto em termos de quantidade quanto em termos de valores.

Note que os campos colocados na área “Campo de dados” foram antecedidos pelo termo “Soma”. Isso significa que os valores desses campos serão somados, ou seja, a avaliação do dado será feita com a soma de todas as quantidades por exemplo. O termo “soma” é atribuído automaticamente a todos os campos adicionados. Esse termo pode ser substituído conforme o tipo de avaliação desejada. Pode-se, por exemplo, substituí-lo por “Máximo”, que mostrará o maior valor encontrado para o campo arrastado. Para trocar esse termo, basta selecionar o campo já dentro da área e teclar no botão “Opções”

Na figura 5 temos a tela de assistente de dados com os campos arrastados conforme os exemplos sugeridos anteriormente.

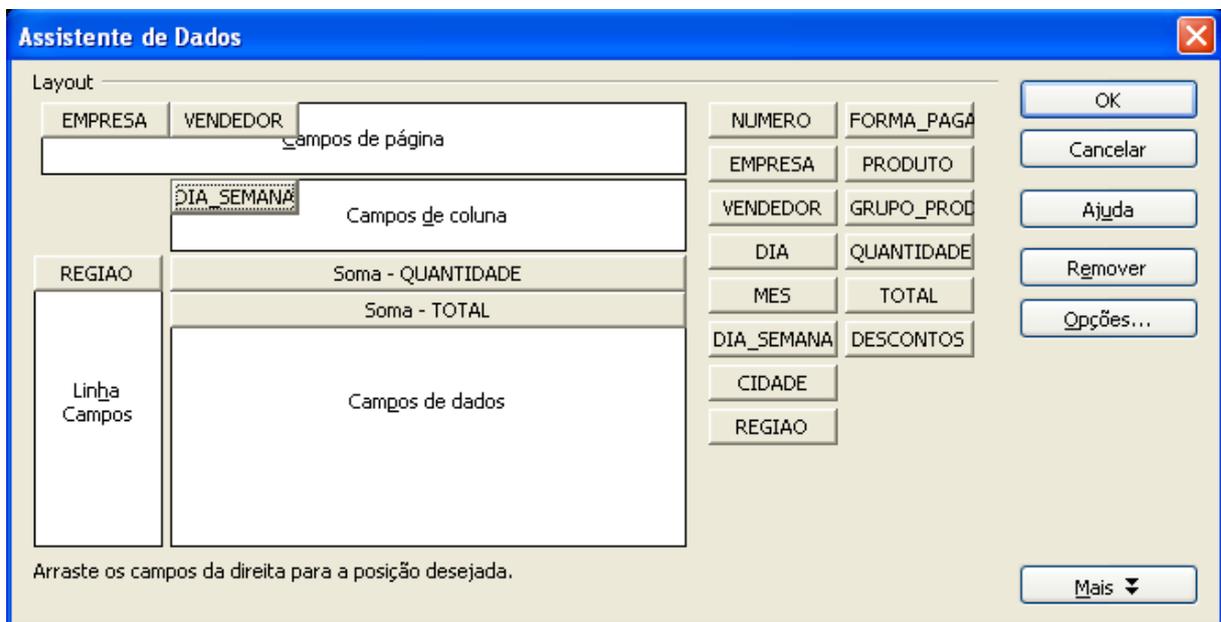


Figura 5 – Exemplo do Assistente de Dados do BrOffice com campos arrastados

Após incluir os campos nos espaços desejados, deve-se clicar no botão “OK”. Feito isso, a consulta dos dados será montada e transferida para a planilha existente na janela anterior, conforme é mostrado na Figura 6. Os dados mostrados na figura serão adicionados à planilha logo abaixo dos dados que foram selecionados.

Filtro									
EMPRESA	- todas -								
VENDEDOR	- todas -								
		DIA_SEMANA							
REGIAO	Dados	Qua	Qui	Sab	Seg	Sex	Ter	Total Resultado	
Bahia	Soma - QUANTIDADE	2369	1716		4400	2750	1200	12435	
	Soma - TOTAL	83.135,00	34.565,00		177.562,00	105.778,00	45.477,00	446.517,00	
Brasilia	Soma - QUANTIDADE	2850			5200	2300	1250	11600	
	Soma - TOTAL	98.934,00			198.258,00	91.076,00	45.233,00	433.501,00	
Cascavel	Soma - QUANTIDADE	10731	8911	2808	19216	13577	1972	57215	
	Soma - TOTAL	230.356,00	153.535,00	54.058,00	422.737,00	321.432,00	108.133,00	1.290.251,00	
Curitiba	Soma - QUANTIDADE		610		1336		917	2863	
	Soma - TOTAL		14.400,00		31.518,00		19.887,00	65.805,00	
Londrina	Soma - QUANTIDADE	11190	9065	67	4576	9687	7882	42467	
	Soma - TOTAL	210.546,00	216.650,00	1.577,00	93.886,00	195.972,00	156.472,00	875.103,00	
Maringá	Soma - QUANTIDADE	21054	13127	747	57003	11811	28410	132152	
	Soma - TOTAL	351.927,00	478.413,00	21.851,00	665.062,00	463.671,00	412.953,00	2.393.877,00	
Pato Branco	Soma - QUANTIDADE	335	202	100	572	243	444	1896	
	Soma - TOTAL	6.600,00	4.449,00	2.250,00	11.811,00	5.137,00	10.601,00	40.848,00	
Ponta Grossa	Soma - QUANTIDADE	108	241	52	496	466	408	1771	
	Soma - TOTAL	1.844,00	4.423,00	1.275,00	9.896,00	11.100,00	6.886,00	35.424,00	
Total Soma - QUANTIDADE		48637	33872	3774	92799	40834	42483	262399	
Total Soma - TOTAL		983.342,00	906.435,00	81.011,00	1.610.730,00	1.194.166,00	805.642,00	5.581.326,00	

Figura 6 – Resultado do Estudo de Caso no BrOffice Calc.

A partir dos dados mostrados na figura 6, os dados podem ser avaliados e os filtros podem ser utilizados para facilitar a avaliação dos dados. Para isso basta clicar na pequena “seta” que aparece à frente dos campos “EMPRESA” e “VENDEDOR”.

Caso seja necessário alterar os campos na consulta, deve-se selecionar novamente a opção apresentada na Figura 2. Daí a janela mostrada na Figura 4 será novamente mostrada, retornando-se daí explicações subseqüentes. A figura 7 mostra uma nova seleção de dados para análise.

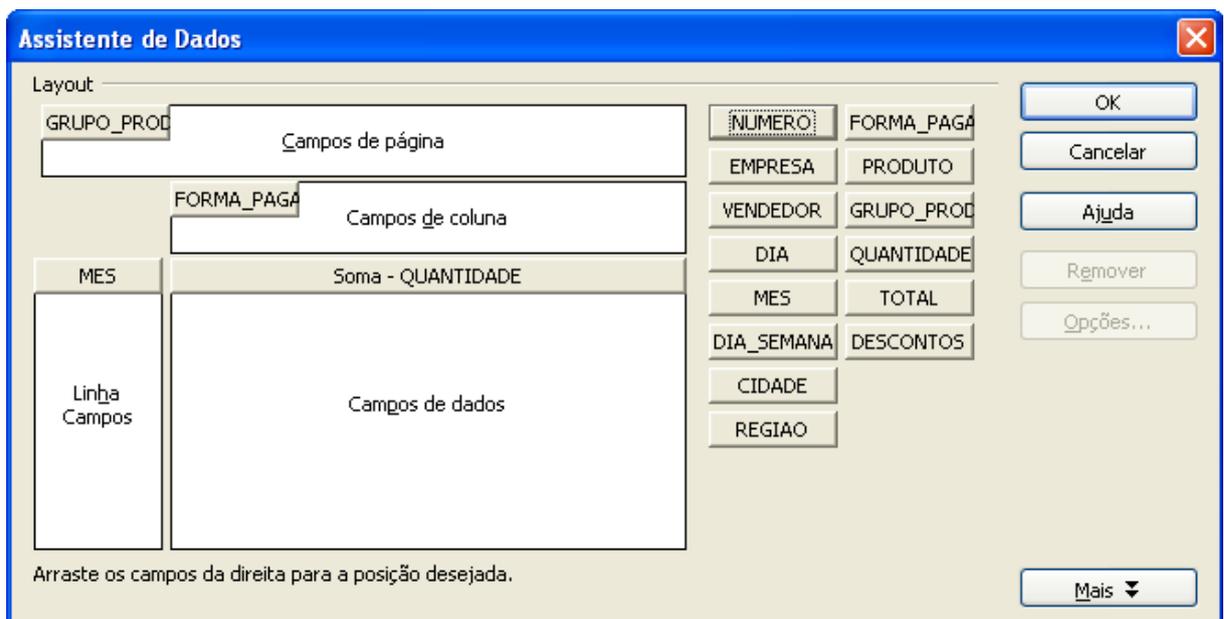


Figura 7 – Segundo Exemplo do Assistente de Dados do BrOffice

A partir dos campos mostrados na figura 7, será gerada a consulta mostrada na figura 8.

Filtro								
GRUPO_PRODUTO	Limpeza							
Soma - QUANTIDADE	FORMA_PAGAMENTO							
MES	10 Dias	20 Dias	30 Dias	45 Dias	60 Dias	A Vista	Total Resultado	
Abr	2822	82	1148	4118			48	8218
Ago	395	800	1403	1250	370			4218
Fev	2029	90	1079	2486	1560			7244
Jan	1229	70	1551	3167			10	6027
Jul	1828	110	1895	1728	1320		42	6923
Jun	1849		1039	1583				4471
Mai	1614		1209	1791				4614
Mar	1319		1675	1748	290		8	5040
Set	1354		1301	1072	700		141	4568
Total Resultado	14439	1152	12300	18943	4240		249	51323

Figura 8 – Resultado do Segundo Exemplo do BrOffice.

Nesse novo exemplo, apenas o campo “QUANTIDADE” esta sendo avaliado num cruzamento de dados que compara a quantidade vendida no mês em relação à forma de pagamento. Nota-se também que os dados estão filtrados conforme o “GRUPO_PRODUTO”, no caso o de “Limpeza”. Dessa forma, todas as quantidades somadas e separadas conforme o “MÊS” e a “FORMA_PAGAMENTO” referem-se apenas as vendas desse tipo de produto.

A figura 9 mostra também o resultado do segundo exemplo, porém, sem o filtro do “GRUPO_PRODUTO” ativado. Nota-se que sem o filtro, os valores são automaticamente ajustados.

Filtro							
GRUPO_PRODUTO	todas -						
Soma - QUANTIDADE	FORMA_PAGAMENTO						
MES	10 Dias	20 Dias	30 Dias	45 Dias	60 Dias	A Vista	Total Resultado
Abr	7102	85	12770	8824	23	2603	31407
Ago	7699	2155	4963	7057	370	2450	24694
Fev	15404	190	3711	7644	1560	2897	31406
Jan	2817	73	3993	4562	10000	2520	23965
Jul	5862	113	14272	4745	1584	2501	29077
Jun	8198	640	12883	7189		2353	31263
Mai	5835		5013	5475		2583	18906
Mar	3189		5185	4010	290	2458	15132
Set	27865	204	16595	4275	4904	2706	56549
Total Resultado	83971	3460	79385	53781	18731	23071	262399

Figura 9 – Eliminação do Filtro no Segundo Exemplo do BrOffice.

Durante o desenvolvimento do estudo de caso no tocante à utilização do BrOffice Calc, constatou-se que o mesmo pode ser utilizado no auxílio à tomada de decisões, porem, uma restrição importante ao seu uso reside no fato do mesmo não possibilitar a conexão direta com o banco de dados, o que torna necessário a exportação de dados sempre que uma análise for necessária.

Outra recurso não oferecido pelo BrOffice Calc e presente no Microsoft Excel refere-se à possibilidade de reduzir o nível da avaliação. Isso permitira, por exemplo, avaliar um a um quais foram os pedidos que geraram a venda dos 640 itens vendidos com prazo de pagamento de 20 dias no mês de junho, mostrados na figura 9.

3.2 Utilizando o Microsoft Excel.

Abaixo será mostrado passo a passo como utilizar o Microsoft Excel como ferramenta uma ferramenta de BI. Na verdade, o Microsoft Excel é uma ferramenta do tipo OLAP, já abordada anteriormente.

Esse roteiro foi feito baseado na versão 2003 do referido software, porém, sua utilização em versões anteriores não apresentará diferença substancial. Não foi avaliada a versão 2007, que é a mais recente do Microsoft Excel.

O Microsoft Excel oferece duas maneiras de se obter os dados que serão avaliados. A primeira é feita através da exportação dos dados que serão analisados para uma planilha do Excel ou até para um arquivo do tipo texto, com os campos separados por vírgulas. Esse método tem a vantagem de ser mais simples, pois não requer um grande conhecimento no tocante ao funcionamento dos bancos de dados, porém, tem o inconveniente de ser necessária uma nova exportação toda vez que a ferramenta for utilizada.

O segundo método permite conectar o Microsoft Excel diretamente ao Banco de Dados onde estão as informações que serão avaliadas. Para isso é utilizada uma outra ferramenta contida no Microsoft Office denominada Microsoft Query. Esse método tem como principal vantagem a possibilidade de utilização da ferramenta a qualquer momento, visto que os dados já estão acessíveis, porém, para sua utilização torna-se necessário que o usuário possua um bom conhecimento no tocante ao funcionamento do banco de dados utilizado pela empresa.

Primeiramente será demonstrada a utilização do método onde os dados são exportados. Nesta caso, a primeira tarefa será carregar no excel os dados que foram exportados. Uma janela com os dados já carregados é mostrada na figura 10.

1	NUMERO	EMPRESA	VENDEDOR	DIA	MES	DIA_SEMANA	CIDADE	REGIAO	FORMA	PRODUTO	GRUPO_PRO	QUANTIDADE	TOTAL	DESCONTOS
2	495	Empresa 3	Vendedor 7	3	Jan	Qui	Sarandi/PR	Maringá	30 Dias	Fósforo Paraná	Higiene	25	3.030,00	441,00
3	495	Empresa 3	Vendedor 7	3	Jan	Qui	Sarandi/PR	Maringá	30 Dias	Fósforo Paraná	Higiene	8	970,00	141,00
4	57	Empresa 6	Vendedor 5	11	Jan	Sex	Cambé/PR	Londrina	A Vista	Fardo Bituminir	Higiene	30	885,00	0,00
5	57	Empresa 6	Vendedor 5	11	Jan	Sex	Cambé/PR	Londrina	A Vista	Fardo Bituminir	Higiene	30	1.680,00	0,00
6	707	Empresa 1	Vendedor 7	14	Jan	San	Umuarama/PR	Maringá	30 Dias	Futroto da Tami	Melhor	60	1.236,00	0,00

Figura 10 – Dados carregados no Microsoft Excel

Em seguida o curso deve ser posicionado na primeira linha e coluna dos dados, conforme mostrado na figura 10. O menu “Dados”, que fica no menu principal do Microsoft Excel deve ser acionado e nele a opção “Relatório de tabela e gráfico dinâmicos” deve ser selecionada. Isso é mostrado na figura 11.

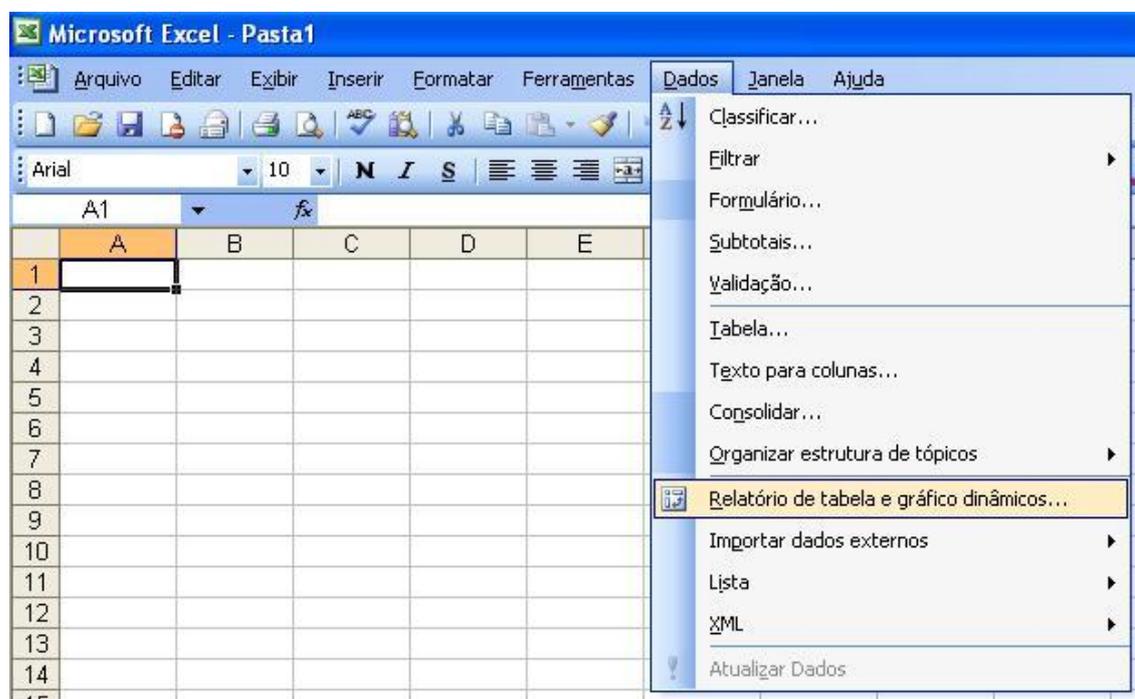


Figura 11 – Seleção da Opção de Tabelas Dinâmicas no Excel

Selecionada a opção, um novo menu será aberto, conforme mostra a Figura 11.



Figura 12 – Seleção da Origem dos Dados no Excel

Nesta janela a opção “Banco de dados ou lista do Microsoft Office Excel” deverá ser marcada. A tecla “Avançar” deverá ser pressionada. Surgirá então a janela mostrada na figura 13.



Figura 13 – Seleção das células dos dados no Excel

O campo “Intervalo” será preenchido automaticamente com o intervalo onde estão os dados da planilha. O botão “Avançar” deverá ser pressionado. A janela da figura 14 será mostrada.

No caso dos dados estarem em outra planilha, basta clicar no botão “Procurar” e informar onde a referida planilha esta armazenada.



Figura 14 – Seleção do Destino dos Dados no Excel

Nesta janela o usuário poderá selecionar se deseja gerar os dados para consulta na mesma planilha onde estão os dados ou então uma nova planilha. Deve-se então marcar a opção “Na nova planilha”, visto que se os dados forem adicionados na mesma planilha onde estão os dados, isso poderá gerar uma certa confusão. Feita a seleção, basta ao usuário clicar no botão “Concluir”.

A figura 15 mostra a janela que será mostrada. É nela onde o usuário deverá definir quais campos deseja avaliar.

Para a área “Solte campos de página aqui” devem ser arrastados os campos sobre os quais se deseja realizar alguma espécie de filtragem, ou seja, quando o usuário desejar fazer uma análise de um dado específico, de maneira que os dados consultados que serão cruzados refiram-se apenas ao selecionado. Pode-se, por exemplo, colocar nessa área os campos “EMPRESA” e “VENDEDOR”. Dessa forma, será possível depurar a consulta dos dados selecionando uma única empresa e/ou produto.

Nas áreas “Solte campos de coluna aqui” e “Solte campos de linha aqui” devem ser adicionados os campos cujos dados que serão diretamente comparados. Arrastando-se para uma dessas áreas o campo “REGIÃO” e para a outra o campo “DIA_SEMANA”, será possível fazer uma avaliação do comportamento das vendas em cada região conforme o dia da semana.

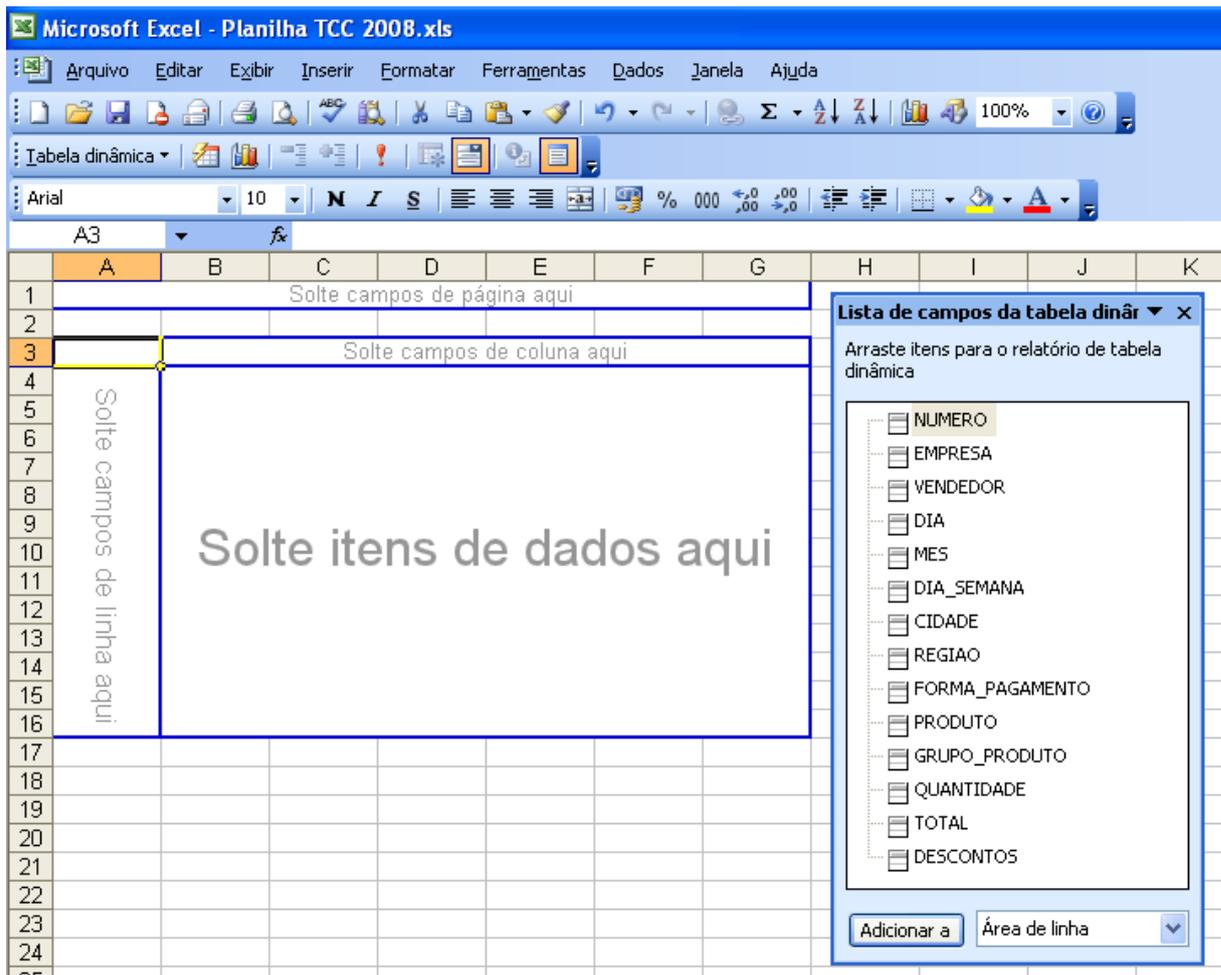


Figura 15 – Seleção dos Dados para análise no Microsoft Excel

Para a área “Solte itens de dados aqui” devem ser arrastados os campos que serão analisados no cruzamento de dados. Ao arrastar para essa área o campo “QUANTIDADE” será possível comparar as vendas dos produtos baseados na quantidade das vendas.

Essa comparação pode ser feita em mais de uma informação ao mesmo tempo, bastando para isso arrastar outros campos para a área desejada. O campo “TOTAL” também pode ser arrastado para a área “Solte itens de dados aqui”, o que permitirá uma comparação dos dados tanto em termos de quantidade quanto em termos de valores.

Note que a medida que os campos são arrastados para a área desejada, a ferramenta automaticamente altera o cenário para essa nova informação. Veja na figura 16 a janela com os campos citados já nas áreas correspondentes.

The screenshot shows a Microsoft Excel spreadsheet with a data table and a dynamic list of fields. The table has columns for 'REGIAO', 'DIA_SEMANA', and 'Total geral'. The dynamic list on the right includes fields like EMPRESA, DIA_SEMANA, and TOTAL.

REGIAO	Dados	Seg	Ter	Qua	Qui	Sex	Sab	Total geral
Bahia	Soma de QUANTIDADE	4400	1200	2369	1716	2750		12435
	Soma de TOTAL	177562	45477	83135	34565	105778		446517
Brasilia	Soma de QUANTIDADE	5200	1250	2850		2300		11600
	Soma de TOTAL	198258	45233	98934		91076		433501
Cascavel	Soma de QUANTIDADE	19216	1972	10731	8911	13577	2608	57215
	Soma de TOTAL	422737	108133	230356	153535	321432	54058	1290251
Curitiba	Soma de QUANTIDADE	1336	917		610			2663
	Soma de TOTAL	31518	19887		14400			65805
Londrina	Soma de QUANTIDADE	4576	7882	11190	9065	9687	67	42467
	Soma de TOTAL	93886	156472	210546	216650	195972	1577	875103
Maringá	Soma de QUANTIDADE	57003	28410	21054	13127	11811	747	132152
	Soma de TOTAL	665062	412953	351927	478413	463671	21851	2393877
Pato Branco	Soma de QUANTIDADE	572	444	335	202	243	100	1896
	Soma de TOTAL	11811	10601	6600	4449	5137	2250	40648
Ponta Grossa	Soma de QUANTIDADE	496	408	108	241	466	52	1771
	Soma de TOTAL	9896	6886	1844	4423	11100	1275	35424
Total Soma de QUANTIDADE		92799	42483	48637	33872	40834	3774	262399
Total Soma de TOTAL		1610730	805642	983342	906435	1194166	81011	5581326

Figura 16 – Tela de Análise de dados no Microsoft Excel

Após arrastar todos os campos e montar o cenário desejado, ainda é possível utilizar vários recursos que a ferramenta oferece. Pode-se, por exemplo, selecionar uma única informação dentro de um campo que foi arrastado, de maneira que a mesma possa ser analisada individualmente. Para fazer isso, basta clicar na “seta” existente à frente dos campos. Ao clicar nessa será possível selecionar uma das informações contidas nesse campo.

Clicando-se duas vezes sobre um valor na área “Itens de Dados” é possível obter a origem desse valor, por exemplo, caso seja necessário avaliar quais pedidos geraram a quantidade de 572 itens vendidos as segundas-feiras na região de Pato Branco, basta clicar duas vezes nesse valor, que a ferramenta mostrará quais foram os pedidos que geraram esse valor.

Também é possível alterar a maneira como os campos arrastados para a área de “Itens de Dados” valores são calculados. Veja que os campos “QUANTIDADE” e “TOTAL” que foram arrastados para esta área tiveram seu cálculos definidos como “Soma”. Isso ocorre porque a ferramenta identifica automaticamente que como esses valores são numéricos, ou seja, valores, a tendência é que esses dados serão somados, porém, é possível alterar essa metodologia para, por exemplo, obter qual foi o menor valor vendido. Para isso, basta clicar com o botão direito do mouse sobre o campo e no menu que surge selecionar a opção “Configurações de Campo”. Na janela que se abre, deve-se selecionar a opção “Mínimo” no campo “Resumir por”. Nessa campo pode-se perceber os vários tipos de análises que podem ser obtidas.

Ao arrastar um campo que não seja numérico para essa área, será atribuído o cálculo como “Contar”, ou seja, os valores não serão somados e sim contados. Isso pode ser útil caso seja necessário obter a quantidade de pedidos além da quantidade vendida, por exemplo.

Passemos agora o segundo método de obtenção de dados. Primeiramente, voltemos para a janela mostrada na figura 12. Ao marcar a opção “Fonte de Dados Externos”, possibilitará que os dados que será avaliados seja conectados diretamente pelo Excel. Essa conexão depende dos drives de “Fontes de Dados ODBC”, instalados nas “Ferramentas Administrativas” do “Painel de Controle” do computador.

Selecionando essa opção, a janela mostrada na Figura 17 será aberta.

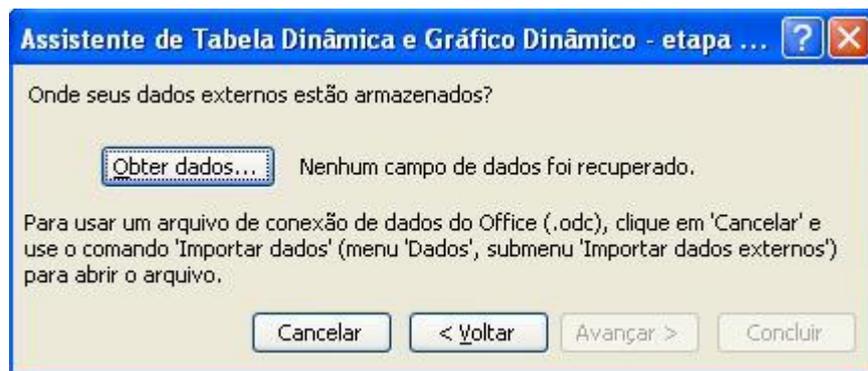


Figura 17 – Seleção de dados externos no Excel

Deve-se então clicar no botão “Obter Dados”, que causará a abertura de outra janela, mostrada na Figura 18.

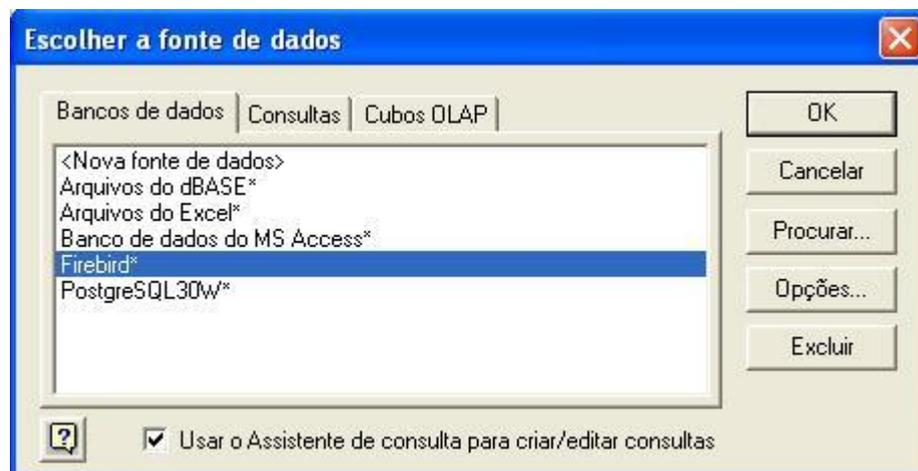


Figura 18 – Seleção do Banco de Dados no Excel

Nessa janela deverá ser selecionado o banco de dados ao qual o Excel deverá se conectar para extrair os dados para a análise. Selecionaremos o banco de dados Firebird, que é o banco de dados onde estão os dados utilizados no estudo do caso. Selecionado o banco, deve-se clicar no botão “OK”.

Uma nova janela será aberta, onde será mostrada, onde os campos de cada tabela do banco de dados deverão ser selecionados. Note que deverão ser selecionados apenas campos que terão alguma relevância na análise. A Figura 19 mostra a janela onde os campos deverão ser selecionados.

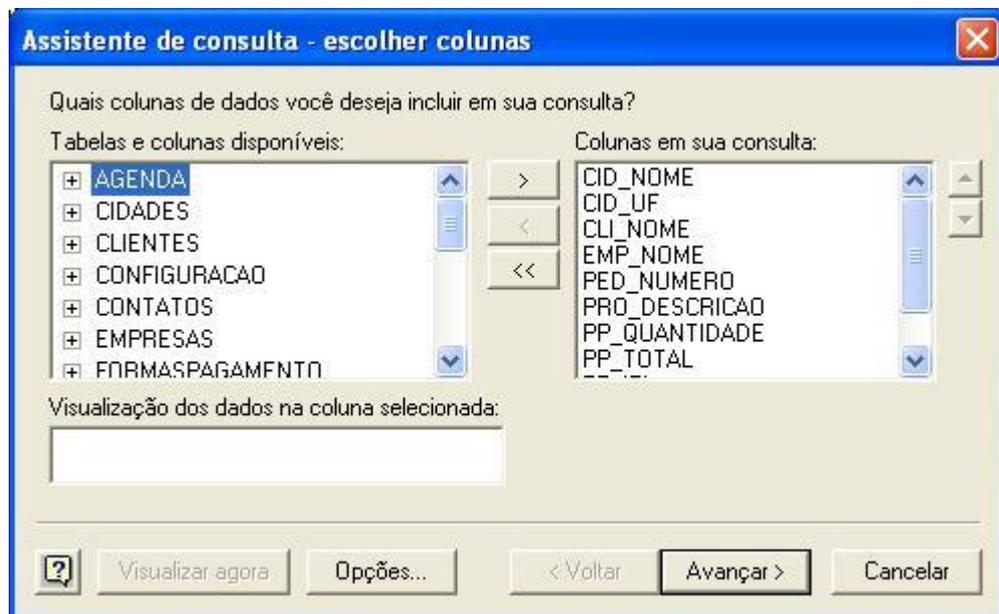


Figura 19 – Seleção dos Campos no Excel

Terminada a seleção dos campos, o botão “Avançar” deve ser acionado. Dependendo do banco de dados selecionado, será mostrada uma mensagem, vista na Figura 20, informando que o relacionamento entre as tabelas não poderá ser realizado, fazendo com que isso tenha que ser feito manualmente. Essa mensagem ocorrerá quando o banco de dados não possuir um vínculo entre as tabelas, também conhecido como *Foreign Keys*.



Figura 20 – Mensagem de Relacionamento Manual das tabelas no Excel

Na próxima etapa (Figura 21) será aberto um programa acessório externo da Microsoft denominado Microsoft Query, que possibilitará o vínculo entre as tabelas. Na figura os relacionamentos entre as tabelas já foi feito. Trata-se das linhas ligando uma tabela a outra. Note que tais linhas conectam um campo de uma tabela a um campo na outra. Veja, por exemplo, que as tabelas “Clientes” e “Cidades” estão vinculadas pelos campos “CLI_CIDADE” e “CID_CODIGO” respectivamente. Isso permitirá ao Excel vincular cada cliente a uma cidade, o mesmo ocorrendo com todas as tabelas.

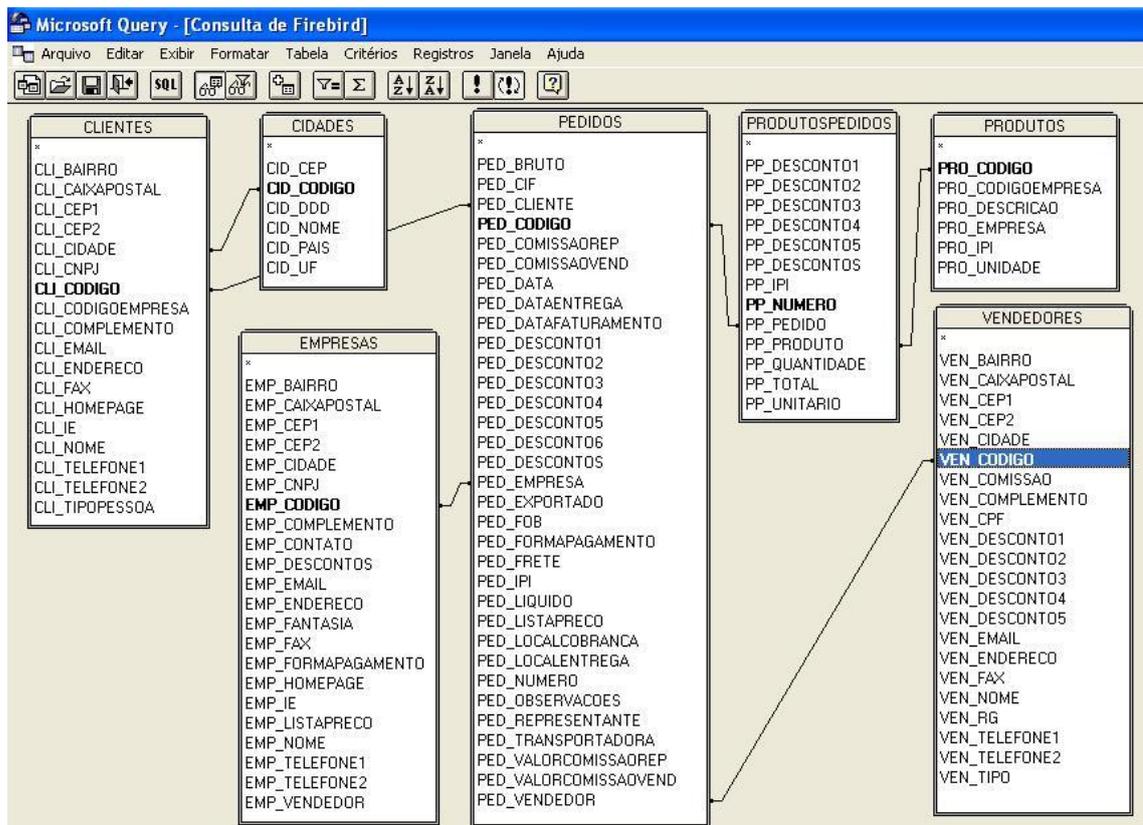


Figura 21 – Tela do Microsoft Query do Excel

Após todos os vínculos terem sido feitos, deve-se retornar os dados ao Excel, clicando no botão “Retornar Dados” (quarto botão da esquerda para a direita). Pode-se ainda selecionar esta opção através do menu “Arquivo”..

Retornaremos então à janela mostrada na Figura 17, onde então se deve teclar no botão “Avançar”. Será aberta a janela mostrada na Figura 14. Nessa janela deverá ser selecionado onde os dados trazidos do Microsoft Query deverão ser armazenados. Selecione o destino desejado e tecler “Concluir”.

A partir desse ponto, o usuário chegará à situação mostrada a partir da figura 15 e as explicações feitas a partir daí são as mesmas para esse segundo método.

Observa-se que os campos contidos na demonstração do primeiro e segundo método são diferentes. Isso ocorre porque para o primeiro método foi feita uma exportação dos dados. Para facilitar a utilização da ferramenta, são exportados apenas os dados necessários para a análise, além de também serem modificados os nomes dos campos, de maneira que fiquem mais inteligíveis ao usuário.

CONCLUSÕES

Nos dias atuais é extremamente importante que as empresas tenham mecanismos que facilitem a tomada de decisões, principalmente devido à acirrada concorrência existente no mercado. Embora exista uma enorme gama de softwares que auxiliam as empresas nessa tarefa, dependendo do tamanho da empresa, isso poderá ter um custo muito elevado, que inviabilizaria tal empreitada.

A adoção de uma ferramenta que a grande maioria das empresas já dispõe como é o caso do Microsoft Excel, ou ainda do BrOffice Calc, vem permitir a essas empresas a utilização de uma ferramenta desconhecida da grande maioria dos seus usuários, que é a criação de tabelas dinâmicas.

Com tais tabelas, é possível ao tomador de decisão, avaliar vários cenários e baseados neles, tomar decisões acertadas a cerca do futuro da empresa.

Trabalhando já há algum tempo com o desenvolvimento de sistemas comerciais, principalmente os mais simples, voltados para as pequenas empresas, tais como lojas de roupas e pequenos mercados “de bairros”, sempre notou-se que após a instalação de um sistema, passado o período de adaptação, normal nesse tipo de implantação, os clientes sempre solicitavam relatórios dos mais diversos, normalmente complicados e que na maioria das vezes eram utilizados apenas uma vez, pois após sua utilização, outras variáveis eram acrescentadas ao problema que gerou a sua solicitação.

Com a utilização de ferramentas como as mostradas neste trabalho, tais relatórios deixariam de ser necessários, visto que após um treinamento na utilização da ferramenta, seria possível ao próprio usuário elaborar os cenários necessários para a análise dos dados.

Concluí-se, portanto, que a utilização do Microsoft Excel como Ferramenta de BI vem facilitar, e muito, a vida dos tomadores de decisão, principalmente das pequenas empresas.

SUGESTÕES DE TRABALHOS FUTUROS

- Implementação do Estudo de caso na versão 2007 do Microsoft Excel.
- Utilização real do estudo em uma empresa;
- Estudo do desempenho em uma base de dados com mais de 100.000 tuplas;
- Utilização da ferramenta através da opção de Cubos OLAP (mostrada na Figura 18);
- Verificar a possibilidade da ferramenta se conectar a dois bancos de dados distintos para a análise dos dados.

REFERÊNCIAS

- BÄCK, T. **Adaptative business intelligence based on evolution strategies: some application examples of self-adaptative software.** Information Sciences, 2002.
- BATINI, C. E LENZERINI, M. **Comparative Analysis Of Methodologies For Database Schema Integration,** ACM Computing Surveys, 1986.
- BISPO, C. A. F.; CAZARINI, E. W. **A nova geração de sistemas de apoio à decisão.** In: **ENEGEP, 18, 1998,** Niterói, Rio de Janeiro, Brasil. Anais... Niterói: ABEPRO, 1998.
- CARLSSON, C.; TURBAN, E. **DSS: directions for the next decade.** Decision Support Systems, 2002.
- GRAY, P.; WATSON, H. J. **The new DSS: data warehouses, OLAP, MDD and KDD.** 1999. Disponível em: <<http://hsb.baylor.edu/ramsover/ais.ac.96/papers/graywats.htm>>.
- GRIGORI, D., CASATI, F.; CASTELLANOS, M.; DAYAL, U.; SAYAL, M.; SHAN, M. C. **Business Process Intelligence.** Computers in Industry, 2004.
- HARDING, J. A.; YU, B. **Information-centred enterprise design supported by a factory data model and data warehousing.** Computers in Industry, 1999.
- HEINRICHS, J. H.; LIM, J. S. **Integrating web-based datamining tolls with business models for knowledge management.** Decision Support Systems, 2003.
- LEME, Trajano Filho. **Business Intelligence no Microsoft Excel,** 2004, Ed. Excel Books, Rio de Janeiro, Brasil
- PRIMAK, Fábio Vinícius. **Decisões com B.I. (Business Ingelligence),** 2008, Ed. Ciência Moderna Rio de Janeiro, Brasil.
- SANTOS, Antonio Raimundo dos, **Gestão do Conhecimento – Uma Experiência para o Sucesso Empresarial** (SERPRO / PUCPR / ESAF (Escola de administração Financeira), 2001.
- SERRA, Laércio. **A Essência do Business Intelligence,** 2002, Ed. Berkeley, São Paulo, Brasil.
- SHIM, J. P.; WARKENTIN, M.; COURTNEY, J.; POWER, D. J.; SHARDA, R.; CARLSSON, C. **Past, present, and future of decision support technology.** Decision Support System, 2002.

Universidade Estadual de Maringá
Departamento de Informática
Curso de Engenharia de Produção
Av. Colombo 5790, Maringá-PR
CEP 87020-900
Tel: (044) 3261-4196 / Fax: (044) 3261-5874