

UNIVERSIDADE SÃO FRANCISCO

Engenharia de Computação

BRUNO CAMPANELLA CRUZ

BRUNO GABRIEL CORREA MIRANDA

FELLIPE BARRETTO TURCHETTE

**CONCEITOS DE BUSINESS INTELLIGENCE POR MEIO DE
ESTUDOS DE CASO: FERRAMENTAS PENTAHO E
QLIKVIEW**

Itatiba

2014

BRUNO CAMPANELLA CRUZ - R.A. 002201000286
BRUNO GABRIEL CORREA MIRANDA - R.A. 002201001110
FELLIPE BARRETTO TURCHETTE - R.A. 002201000196

**CONCEITOS DE BUSINESS INTELLIGENCE POR MEIO DE
ESTUDOS DE CASO: FERRAMENTAS PENTAH0 E
QLIKVIEW**

Monografia apresentada à disciplina
Trabalho de Conclusão de Curso, do
curso de Engenharia de Computação da
Universidade São Francisco como
exigência para conclusão do curso de
graduação.

Orientador: Prof. Glauco Kiss Leme

Itatiba
2014

AGRADECIMENTOS

Agradecemos a todos os familiares que sempre nos apoiaram incondicionalmente, aos amigos e colegas de sala que fizeram parte da nossa formação.

A Universidade São Francisco pela oportunidade de fazer o curso.

Ao nosso orientador Glauco Kiss Leme, pelo suporte, correções e incentivo.

A todos aqueles que direta ou indiretamente fizeram parte da nossa formação e conclusão desse projeto, o nosso muito obrigado.

RESUMO

Atualmente, o tempo necessário para a tomada de decisão por parte das empresas está cada vez menor. Estas decisões determinam, muitas vezes, o sucesso ou fracasso das empresas em seus respectivos projetos. Com base neste cenário é fundamental que um gestor ou executivo tenha ao seu alcance todas as informações necessárias para tomar a decisão da melhor maneira possível. E o mais importante, baseado em dados confiáveis. O conceito de Business Intelligence tem o intuito de auxiliar na tomada de decisão, abstraindo informações provenientes de diversas origens de fonte de dados e transformando esses dados em conhecimento através de relatórios analíticos e sintéticos. Além de indicadores para melhor legibilidade destas informações. Tendo em mãos a informação necessária, o gestor pode criar estratégias para direcionar e concluir seus objetivos, dentro da empresa, com uma maior probabilidade de acerto. Neste trabalho foi explicado o conceito de Business Intelligence e também foram estudadas as etapas padrões de Business Intelligence como ETL, DataWareHouse, Datamarts e OLAP. Os conceitos apresentados foram aplicados na prática, através de ferramentas voltadas ao Business Intelligence, como é o caso das ferramentas Pentaho e QlikView.

ABSTRACT

Nowadays, the time needed for making decisions by companies is smaller and smaller. These decisions determine often the success or failure of the companies in their projects. Based on this scenario it is essential that a manager or executive has all the information necessary to make the decision in the best possible way. And most importantly, based on reliable data. The concept of Business Intelligence is intended to assist in decision making, leaving aside, information from different data source and transforming that data into knowledge through analytical and synthetic reports, and indicators for readability of this information. Having in hand the necessary information, the manager can develop strategies to direct and complete the goals within the company with a higher probability of success. This work explains the concept of Business Intelligence and also considers the standard steps of Business Intelligence and ETL, DataWareHouse, Datamarts and OLAP. The concepts presented were applied in practice through tools aimed at business intelligence, as is the case with Pentaho and QlikView tools.

LISTA DE ILUSTRAÇÕES

Figura 1 - Processos para criação ferramenta BI.....	18
Figura 2 - Composição do BI	20
Figura 3 - Ilustração do ciclo de vida do ETL em todo seu processo	24
Figura 4 - Ilustração da grande quantidade de fontes de dados para extração do ETL	25
Figura 5 - Passos para a criação de um Data Warehouse	29
Figura 6 - Modelo dimensional.....	30
Figura 7: Modelo Estrela.....	31
Figura 8: Modelo floco de neve.....	32
Figura 9 - Tipos de sistemas analíticos para soluções informatizadas em BI.....	34
Figura 10 - Modelo Cubo OLAP.....	34
Figura 11 - Exemplo cubo que contém duas medidas, pacotes, e último, e três dimensões relacionadas, Rota, Origem e Horário	36
Figura 12 - Arquitetura de um sistema OLAP	36
Figura 13: Ciclo de Data Mining.....	38
Figura 14 - Fases do QlikView.....	40
Figura 15 - Diagrama em Star Schema	41
Figura 16 - Diagrama em Snow Flake	42
Figura 17 - Janela para conexão da base de dados através do ODBC.....	43
Figura 18 - Janela de Scripts que demonstra os scripts padrões e o script de conexão ODBC	44
Figura 19 - Importação de dados interface gráfica e seu código SQL gerado abaixo .	45
Figura 20 - Script de seleção modificado para obter Ano do pedido e Mês do pedido	46
Figura 21 - Botão Recarregar que carrega as informações em memória.....	46
Figura 22 - Objetos que podem ser usados para apresentar e filtrar dados no QlikView	47

Figura 23 - Listas para filtrar por data	48
Figura 24 - Gráfico em Pizza mostrando vendas por país	49
Figura 25 - Objeto Seleções Atuais em funcionamento	50
Figura 26 - Projeto QlikView criado a partir do modelo Star Schema.....	50
Figura 27 - Projeto QlikView criado a partir do modelo Snow Flake.....	51
Figura 28 - Banco de dados Postgres.....	53
Figura 29 - Pentaho Data Integration.....	54
Figura 30 - Pentaho Schema Workbench	55
Figura 31 - Console Pentaho.....	55
Figura 32 - Visualização de cubo OLAP	56
Figura 33 - Cubos de análise de dados	57
Figura 34 - Saiku Analytics	58

LISTAS DE ABREVIATURAS E SIGLAS

BI – *Business Intelligence*

CSO – Diretor de Estratégia

DM – *Data Mining*

DW – *Data Warehouse*

EIS – *Executive Information Systems*

ETL – *Extract Tranform Load*

MIT – *Massachusets Institute of Tecnology*

OLAP – *On-Line Analytics Processing*

OLTP - Processamento de transações on-line

RAM - *Random Access Memory*

SAD – Sistema de apoio à decisão

PCP – Processo e controle de produção

SI - Sistema Integrado

CRISP-DM - Cross-Industry Standard Process of Data Mining

SUMÁRIO

1 INTRODUÇÃO.....	11
1.1 Apresentação	11
1.2 Resultados Esperados	12
2 DESENVOLVIMENTO	13
2.1 Fundamentação Teórica.....	13
2.1.1 Definições e conceitos	13
2.1.2 História	14
2.1.3 Evolução	15
2.1.4 O que é Business Intelligence?	17
2.1.5 Benefícios do BI	19
2.1.6 Ferramentas de BI	21
2.1.7 Como se inicia um projeto de BI?	22
2.2 Metodologia	23
2.3 Etapas de BI	23
2.3.1 ETL	23
2.3.2 Datawarehouse	26
2.3.2.1 Modelo dimensional para data warehouse	29
2.3.2.2 Tipos de modelos dimensionais	30
2.3.2.3 Considerações sobre ambos os modelos	32
2.3.3 Data marts	33
2.3.4 OLAP (on-line analytics processing)	33
2.3.5 Data mining	36
2.4 Ferramentas	39
2.4.1 Qlik View	39
2.4.1.1 Fases do Qlikview	40
2.4.1.2 Requisitos	40
2.4.1.3 ETL	43
2.4.1.4 Construção da interface	47
2.4.2 Pentaho BI platform	52
2.4.2.1 Banco de dados	52
2.4.2.2 Pentaho data integration (PDI)	53
2.4.2.3 Pentaho schema workbench (PSW)	54

2.4.2.4 Console Pentaho.....	55
2.4.2.5 Visualização pivot analysis	56
2.4.2.6 Analysis pivot – cubos de análise de dados	56
2.4.2.7 Saiku analytics.....	57
2.5 Resultados Obtidos	58
3 CONCLUSÃO	59
REFERÊNCIAS	60

1 INTRODUÇÃO

1.1 Apresentação

O conceito de *Business Intelligence* (BI) não é recente, é um conceito formado desde os primórdios englobando os persas, egípcios, no cenário das guerras mundiais. Utilizando os princípios básicos do BI, estes cruzavam informações obtidas junto à natureza em seu próprio benefício, tomando decisões estratégicas analisando os comportamentos de marés, os períodos de chuva ou seca entre outras formas de obter informações que seriam, por sua vez, usadas para tomar as melhores decisões com o cenário do momento e do futuro. Em estudos sobre economia inteligente, Yves-Michel Marti, cientista, professor e fundador da Egideria cita exemplos como no final do século XVI, quando a Rainha Elizabeth I, visando ocupar os territórios conquistados, determinou que a base da força inglesa fosse “informação e comércio” e solicitou ao filósofo Francis Bacon que inventasse um sistema dinâmico de informação, o qual foi amplamente aplicado pelos ingleses na época.

Desde então o mundo mudou, mas permanece o mesmo conceito. O desenvolvimento tecnológico ocorrido a partir da década de 70 possibilitou a criação de ferramentas, a fim de facilitar todo o processo de captação, extração, armazenamento, filtragem, disponibilidade e personalização dos dados. Porém, o grande problema era que esses produtos exigiam intensa e exaustiva programação e não disponibilizavam informação em tempo hábil nem de forma flexível, e além de tudo tinham alto custo de implantação. Em meados da década de 90 o ambiente empresarial foi fortemente afetado com essas mudanças, porém, após mais uma década, uma nova corrente começa a se fortalecer, sugerindo que o verdadeiro patrimônio das companhias não são seus colaboradores, mas sim suas informações (LEME FILHO, 2004, p.21).

Com o surgimento dos bancos de dados relacionais, dos computadores pessoais e dos sistemas operacionais com interfaces gráficas como o Windows, aliados ao aumento da complexidade dos negócios, começaram a surgir os primeiros produtos realmente direcionados aos analistas de negócios que possibilitavam rapidez e uma maior flexibilidade de análise.

Neste novo cenário, manifestou-se a necessidade de adquirir informações em tempo ágil e eficaz, promovendo aos executivos alicerces em suas tomadas de decisões, implementadas através dessas ferramentas de software, cuja função é proporcionar ganhos nos processos decisórios gerenciais e da alta administração nas organizações, assim sendo, a falta de informação em tempo hábil pode causar um problema à empresa, uma vez

que as tomadas de decisão serão empíricas ao invés de embasadas em dados reais. No intuito de evitar esses problemas foi implementado o conceito BI, de modo que a forma de gerir informação em uma empresa faz com que a mesma obtenha ou não sucesso.

Atualmente o interesse pelo BI vem se expandindo, pois a necessidade de cruzar dados diversos a fim de criar uma série de análises e projeções, de forma a agilizar os processos relacionados às tomadas de decisões vem sendo cada vez mais procurada. Logo, pode se concluir que *Business Intelligence* não é apenas um conceito de processo, mas também uma ferramenta de tecnologia.

Através de uma ferramenta implementada utilizando SAD (Sistema de Apoio à Decisão) e tecnologia BI o executivo pode manipular dados, gerando relatórios específicos às necessidades vigentes da empresa. Estes relatórios podem proporcionar oportunidades de negócios, aliado ao nível de conhecimento desse gestor, convertendo ameaças em boas chances de lucro e aumentar a capacidade de resposta à mudança, de acordo as características de seu nicho de mercado.

O objetivo deste trabalho é explicar e aplicar, na prática, os conceitos estudados sobre *Business Intelligence* nas ferramentas Pentaho e QlikView. Serão mostradas cada uma das etapas de uma implantação padrão de *Business Intelligence* como ETL, *Data Warehouse*, *Datamarts* e OLAP. Este trabalho também tem o objetivo de destacar os pontos fortes e pontos fracos das plataformas analisadas, mostrando as características distintas entre Pentaho e QlikView.

1.2 Resultados Esperados

Os resultados esperados deste trabalho são estudar o conceito de *Business Intelligence* e suas etapas, utilizando na prática as ferramentas QlikView e Pentaho. Esse estudo das ferramentas deve mostrar como resultado os lados positivos e negativos de cada ferramenta e se essas ferramentas utilizam do conceito de *Business Intelligence* estudado. Também espera-se que esta aplicação das ferramentas na prática mostre se elas podem ser aplicadas em empresas de pequeno e médio porte.

2 DESENVOLVIMENTO

2.1 Fundamentação Teórica

Segundo Pavani (2006), “para atender a crescente demanda por projetos de BI dos últimos anos, muitas consultorias de informática vêm se especializando na implantação de projetos classificados como informacionais”.

Atualmente, com a grande pressão que as empresas recebem do mercado financeiro, uma tomada de decisão pode ser determinante em um bom negócio, além do mais, na maioria das vezes, há pouco tempo para pensar e agir. Assim sendo, nesta perspectiva, uma decisão errada poderá ser fatal para o destino de um empreendimento. Assim, optar por ferramentas corretas para a gestão e organização tecnológica do *Business Intelligence*, irá fornecer melhora nas decisões da empresa.

A filosofia de trabalho com foco voltado para o cliente foi suprimida pelas soluções tecnológicas. Atualmente o BI, assim como aconteceu com o CRM na década passada, está sendo difundido por empresas desenvolvedoras e consultorias que vendem e prestam serviços de implantação. Empresas contratantes passam a se interessar somente pelas diferenças arquiteturais e possibilidades de integração tecnológica com outras plataformas de mercado. Profissionais de BI buscam incansavelmente atualizar seus conhecimentos técnicos para manterem-se empregáveis e desfrutar das altas quantias pagas para tais funções específicas.

Muitas empresas ainda possuem uma visão muito simplista, ou seja, meramente tecnológica dos benefícios que podem ser adquiridos com a utilização dessas soluções. Fornecedoras de serviços de implantação não buscam alinhar a grande capacidade das ferramentas de BI ao planejamento estratégico dos contratantes, até mesmo porque a maioria de seus consultores migrou do desenvolvimento de sistemas operacionais para o mundo informacional nos últimos anos.

2.1.1 Definições e conceitos

Alguns conceitos serão utilizados nas próximas seções e são definidos abaixo:

- **Dado:** Nos dizeres de Alter (1996, apud PLETSCHE, 2003, p.17) são “fatos, imagens ou sons que podem ou não ser pertinentes ou úteis para uma tarefa particular”. É importante salientar que um dado sozinho geralmente não traz relevante utilidade.

- **Informação:** Para Romney e Steinbart (2000, p.13), são “dados que foram organizados e processados de forma que sejam significativos”, ou seja, conjunto de dados coletados, de forma pré determinada transformando assim dados brutos em informação útil e aplicável a uma ou mais situações.
- **Conhecimento:** Nos dizeres de Alter (1996, apud LEITE, 2002, p.27) “é uma combinação de instintos, ideias, regras e procedimentos que guiam ações e decisões”. Com isso, o conhecimento auxilia a tomada de decisão.
- **Decisão:** Para Revista TECAP (nº 3, ano 3, volume 3 – 2009 anual) é o processo pelo qual são escolhidas algumas ou apenas uma entre muitas alternativas para as ações a serem realizadas, esta escolha deve estar embasada pelo maior número possível de informações e conhecimento para que a decisão escolhida seja a melhor dentre as disponíveis.

2.1.2 História

Uma referência a *Business Intelligence* aparece no livro *Art of War* (Sun Tzu 2006),. Sun Tzu dizia que para conseguir sucesso na Guerra, deve-se ter conhecimento total de suas próprias virtudes e fraquezas, assim como conhecimento total das principais virtudes e fraquezas de seus inimigos. A falta de um destes resulta na derrota.

Antes do começo da era da informação no século 20, os negócios algumas vezes enfrentavam problemas com o esforço para obter informações de fontes não automáticas. Faltavam aos negócios recursos computacionais para analisar corretamente os dados, e quase sempre, tomaram decisões comerciais primárias baseadas apenas na intuição.

Já hoje em dia nos negócios modernos, os padrões, a automação, e as tecnologias crescentes conduziram para que cada vez mais dados fossem disponibilizados. Tecnologias de *Data Warehouse* possuem um conjunto de repositórios para armazenar esses dados. ETL melhoradas e recentemente *Enterprise Application Integration* tem aumentado a velocidade de coleta de dados. OLAP mostra tecnologias que têm permitido gerar novos relatórios mais rapidamente que analisam os dados.

Com todas essas tecnologias surgiu o termo *Business Intelligence* que, resumidamente é um conjunto de conceitos e metodologias que, fazendo uso de dados extraídos de uma organização, apoiam a tomada de decisões. De forma mais detalhada, atualmente existem muitas definições para o termo, como a de Angeloni e Reis (2006, p.3),

que descreve *Business Intelligence* como: “O conceito de *Business Intelligence* com o entendimento de que é inteligência de negócios ou inteligência empresarial que compõe-se de um conjunto de metodologias de gestão, implementadas através de ferramentas de software, cuja função é proporcionar ganhos nos processos decisórios gerenciais e da alta administração. Ferramentas que integram em um só lugar todas as informações necessárias ao processo decisório”. Reforça-se que o objetivo do *Business Intelligence* é transformar dados em conhecimento, que suporta o processo decisório com o objetivo de gerar vantagens competitivas.

Já segundo Batista (2004, apud ANGELONI e REIS,2006, p.5): “Podem fornecer uma visão sistêmica do negócio e ajudar na distribuição uniforme dos dados entre os usuários, sendo seu objetivo principal transformar grandes quantidades de dados em informações de qualidade para a tomada de decisões. Através delas, é possível cruzar dados, visualizar informações em várias dimensões e analisar os principais indicadores de desempenho empresarial”.

Barbieri (2001, apud ANGELONI e REIS,2006, p.5), por sua vez, descreve como: “Um guarda-chuva conceitual, visto que se dedica à captura de dados, informações e conhecimentos que permitam às empresas competirem com maior eficiência em uma abordagem evolutiva de modelagem de dados, capazes de promover a estruturação de informações em depósitos retrospectivos e históricos, permitindo sua modelagem por ferramentas analíticas. Seu conceito é abrangente e envolve todos os recursos necessários para o processamento e disponibilização da informação ao usuário.”.

2.1.3 Evolução

O termo *Business Intelligence* surgiu apenas na década de 80, pela empresa Gartner Group, segundo Primak (2008). O mesmo autor, ainda definiu o *Business Intelligence* como “o processo inteligente de coleta, organização, análise, compartilhamento e monitoração de dados contidos em *Data Warehouse* e/ou *Data Mart*, gerando informações para o suporte à tomada de decisões no ambiente de negócios”.

Para Primak (2008) o interesse expressivo das soluções de BI pelo setor corporativo ocorreu “no final de 1996, quando o conceito começou a ser espalhado com um processo de evolução do EIS (*Executive Information Systems*) – um sistema criado no final da década 70” a partir dos estudos e trabalhos desenvolvidos pelo MIT (*Massachusetts Institute of Technology* – EUA).

Para Howard Dresner, ex-analista da Gartner o emprego do BI por parte das grandes empresas possibilita realizar uma série de análises e projeções, de forma a agilizar os processos relacionados às tomadas de decisões.

Com o desenvolvimento da informática e a crescente busca pela padronização de dados os sistemas ficaram mais poderosos e as informações mais disponíveis, inclusive com uma crescente tendência de busca pela comunicação entre empresas diferentes, o processo de recuperação de dados ficou facilitado, o que impulsionou bastante a necessidade de uma melhor análise de tais dados, e é exatamente aí que o *Business Intelligence* entre em ação.

Atualmente as aplicações e tecnologias de BI podem ajudar as empresas a analisarem as tendências de transformação do mercado, alterações no comportamento de clientes, os padrões de consumo, preferências de clientes, recursos das empresas e condições de mercado.

No competitivo setor de atendimento ao cliente, as empresas têm que dispor de informações precisas e atualizadas sobre as preferências dos clientes, para que as empresas possam adaptar-se rapidamente à demanda em constante mutação. Os sistemas de BI permitem que as empresas colem informações sobre as tendências do mercado e ofereçam produtos e serviços inovadores, antecipando as transformações das exigências dos clientes.

Com o advento da internet e a globalização é cada vez mais acirrada a disputa entre as empresas por espaço no mercado, há uma procura cada vez maior por formas de diferenciação da concorrência. Para tomar decisões importantes e vencer a concorrência é necessário que os diretores das empresas consigam, de alguma forma, prever o próximo passo dos concorrentes e agir de forma antecipada.

Dados passam a não ser tão úteis se não estiverem ligados a um contexto, sendo usados como informação necessária a um conhecimento da própria empresa e para a melhor decisão para seus clientes, parceiros e concorrentes.

Até poucas décadas atrás a recuperação de tais dados não era fácil. Os dados de cada empresa e de cada sistema seguiam diferentes padrões.

Os sistemas de BI transformaram-se na arte de examinar grandes quantidades de dados, extraindo as informações pertinentes e transformando as informações em conhecimento, com base no qual, as decisões podem ser tomadas.

Em 1989, Howard Dresner, um membro de pesquisa do Gartner Group popularizou “BI” com um termo genérico, usado para descrever um conjunto de conceitos e métodos para aperfeiçoar a tomada de decisões de negócios utilizando sistemas de suporte baseados em fatos. Dresner deixou o Gartner em 2005 e entrou para a Hyperion Solutions como seu CSO (*Chief Strategic Officer* – Diretor de Estratégia).

2.1.4 O que é Business Intelligence?

Com a atual ascensão da tecnologia aparente para o mundo, as empresas vêm cada vez mais a necessidade de avançar e atualizar seus processos e métodos, a fim de acompanhar a demanda cada vez maior que o mercado mundial exige e alcançar sucesso através de resultados ágeis, proporcionados por métodos eficientes.

Sendo um desses métodos, um conjunto de aplicações para apoio à tomada de decisões da empresa reunindo informações armazenadas em banco de dados, e outras fontes internas e externas, organizando-as, analisando-as e trazendo um retorno para a empresa de forma que seja possível que esta estude o melhor planejamento possível e tome as decisões certas.

O *Business Intelligence* (BI), que em português trás o significado “Inteligência de Negócios”, surgiu na década de 80 sendo difundido pelo Gartner Group e é uma tecnologia que permite a transformação de dados em informações quantitativas e importantes para a tomada de decisão da empresa. Essa definição é apresentada por Noronha (1997, apud Tyson, 2013) onde cita *Business Intelligence* como sendo “um processo que envolve a coleta, análise e validação de informações sobre concorrentes, clientes, fornecedores, candidatos potenciais à aquisição, candidatos à *joint-venture* e alianças estratégicas. Incluem também eventos econômicos, reguladores e políticos que tenham impacto sobre os negócios da empresa. O processo de BI analisa e valida todas essas informações e as transforma em conhecimento estratégico”

O *Business Intelligence* também oferece auxílio gerencial nas áreas de vendas, compras, estoque, faturamento, PCP, custos, finanças, contabilidade, telemarketing, recursos humanos e indicadores de gestão. Além disso, em caso de erros decisórios, o sistema pode emitir alertas e explorar as informações até encontrar a raiz do problema, informando ao responsável.

O termo BI, inteligência de negócios, refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte a gestão

de negócio. É o conjunto de teorias, metodologias, processos, estruturas e tecnologias que transformam uma grande quantidade de dados brutos em informação útil para tomadas de decisões estratégicas. Tudo começa com coleção de dados, *Data Warehousing*, a integração de dados de uma ou mais fontes e assim, cria um repositório central de dados, um *Data Warehouse* – os armazéns de dados. Com essa imensidão de dados, *Data Mining*, aplica-se à mineração desses dados, o processo de explorar grandes quantidades de dados à procura de padrões consistentes para detectar relacionamentos e novos subconjuntos de dados a serem mapeados e extraírem informações privilegiadas. Análises, *Analytics*, de minerações geral relatórios, *Reporting*, detalhados para fortalecer o esclarecimento do cenário. Reengenharia de processo de negócios (BPR) trata se de uma estratégia de gestão de negócios para a análise e desenho dos fluxos de trabalho e dos processos de negócios visando a reestruturação organizacional, com o foco no design de baixo para cima de processos de negócios dentro de uma organização. E *Benchmarking*, a busca das melhores práticas com o propósito de maximizar o desempenho.



Fonte: <http://www.oficinadanet.com.br/post/13153-o-que-e-business-intelligence>

Figura 1 - Processos para criação ferramenta BI

Segundo (Noronha, 2013) *“a grande vantagem do conceito de Business Intelligence é justamente a capacidade que o sistema possui para ‘traduzir’ os dados armazenados em uma linguagem de fácil assimilação pelo corpo gerencial das empresas. Esse ambiente gerencial geralmente é caracterizado por gráficos que permitem a rápida interpretação de uma situação. Por exemplo, se perguntamos ao sistema quais são os nossos maiores clientes, a ferramenta de BI imediatamente monta um ranking entre os maiores, aplicando critérios como faturamento, por exemplo.”*

Para concluir, pode-se citar um exemplo de aplicação do BI, também fornecido por Noronha (2013):

“Wall-Mart, uma das maiores cadeias de hipermercados do mundo, consegue reunir informações em tempo real das vendas de 4,5 mil lojas distribuídas pelo mundo, cerca de duas mil delas nos Estados Unidos e, esse tremendo data warehouse já se tornou parâmetro até mesmo para o presidente da república e do banco central dos Estados Unidos. O Bush e o Alan Greenspan ligam toda semana para o CEO do Wall-Mart para saber das vendas. Assim eles acompanham o andamento da economia mundial e local. Além disto o sistema permite identificar as correlações entre produtos que em muito ajudam no cross merchandising. Assim os gerentes das lojas experimentam novas posições para os produtos que comprovam fazer as vendas crescer em função de uma relação até então desconhecida. No Wall-Mart, por exemplo, os salgadinhos no caminho da cerveja ou o molho perto dos legumes são fórmulas comprovadas.”

2.1.5 Benefícios do BI

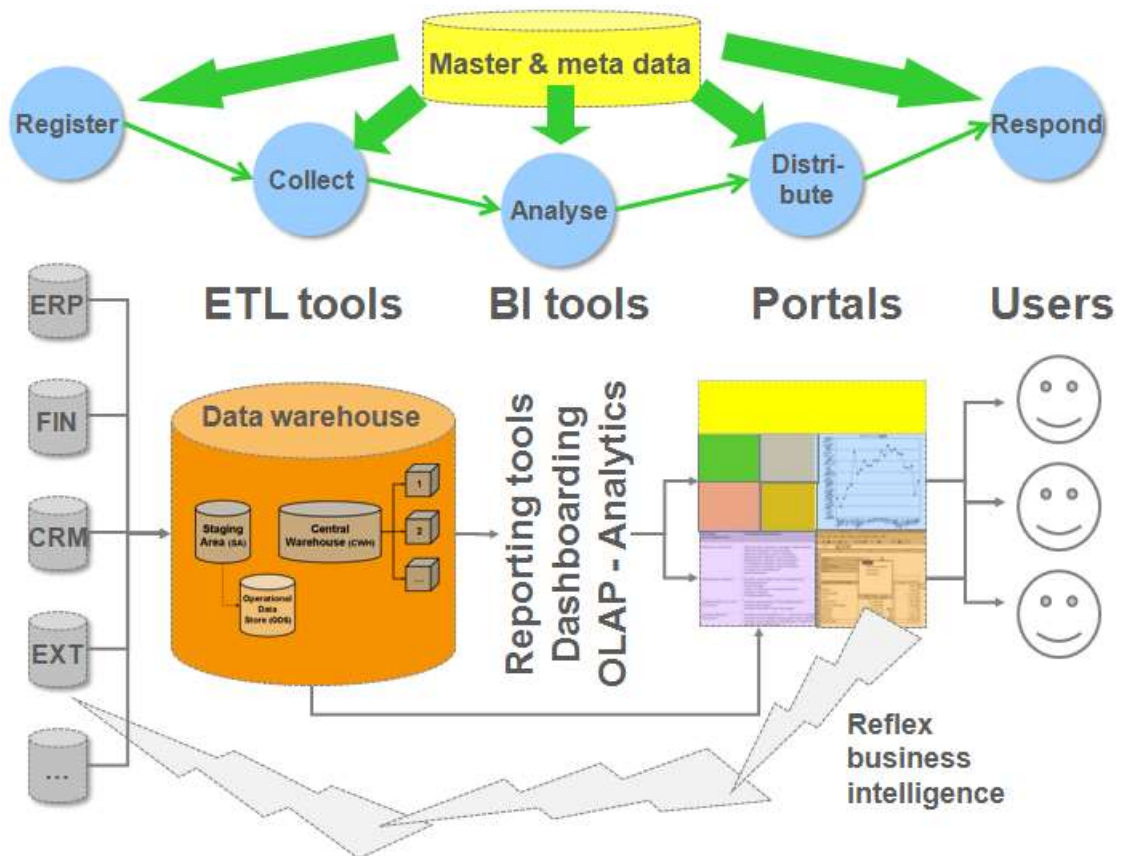
Segundo Primak (2008), a estratégia constante que é encontrada em BI precisa estar sempre alinhada aos compromissos e interesses de cada empresa, direcionada às metas previamente estabelecidas.

Fatores imprescindíveis na estrutura tecnológica de BI são: o desenvolvimento de um repositório de dados como, por exemplo, um *Data Warehouse* ou um *Data Mart*, bem como, ferramentas a serem usadas tais como ferramentas ETL, OLAP, entre outras.

Porém, de nada adianta a empresa possuir todas as ferramentas disponíveis no mercado para a extração das informações e não saber ao certo interpretá-las. É necessário que a empresa saiba nortear o seu capital intelectual para que dessa forma as informações propiciadas pelo BI atendam às expectativas esperadas.

Os gestores poderão ter acesso às informações “rapidamente” e poderão abreviar o tempo de resposta melhorando assim os processos decisórios. Dessa forma, a informação será o verdadeiro capital integralizado da empresa trazendo conhecimento para as decisões imediatas e para aquelas que virão no futuro.

Contudo, trabalhar o conhecimento usando o BI é uma linha tênue e precisa estar sempre bem “presa” às definições dos processos evolutivos da empresa em conjunto com novas práticas comerciais, em melhores maneiras de relacionamento com os clientes e em novas formas de sobrevivência visando sempre usar a inteligência na tomada de decisão precisa e coerente.



Fonte: <http://www.passionned.com/business-intelligence/what-is-business-intelligence-bi/>

Figura 2 - Composição do BI

A seguir são listados alguns benefícios do uso e aplicação do BI nas instituições:

- Incorporar os projetos de tecnologia com as metas estabelecidas pelas empresas na busca do máximo retorno do investimento;

- Compreender as tendências dos negócios, melhorando a consistência no momento de decisão de estratégias e ações a serem tomadas;
- Facilitar a identificação de riscos;
- Planejamento corporativo mais amplo;
- Facilitar o acesso e distribuir informação de modo mais amplo para obter envolvimento de todos dentro da empresa;
- Oferecer dados estratégicos para análise com um mínimo de atraso em relação a uma transação ou evento dentro da empresa.

O BI possui diversas vantagens podendo realizar os procedimentos abaixo:

- Automatização da informação de mapas de indicadores, evitando procedimentos manuais e rotineiros, obtendo diminuição dos custos operacionais;
- Visualização dinâmica cruzada: apresentação disponibilizada em várias perspectivas;
- Importação direta de dados de outras aplicações (ex.: Excel, Sistemas Legados, Sistemas Satélites e outras bases de dados), para tratamento dessas informações;
- Passagem dos dados estáticos para informações dinâmicas;
- Permite maior flexibilidade das informações das diversas fontes;
- Utilização da aplicação por seleção das variáveis de análise (arrastamento e/ou filtragem por um valor determinado dessa variável);
- Informação sempre atualizada, mediante definição e parametrização prévia.

2.1.6 Ferramentas de BI

Segundo ainda Primak (2008), as primeiras ferramentas de BI começaram a surgir no início dos anos 70 e tinham como características o uso intenso de programação linear, o que elevava os custos de análise e desenvolvimento.

Com o passar do tempo e o consequente surgimento dos bancos de dados relacionais, dos computadores pessoais, das interfaces gráficas, e da consolidação do modelo cliente-servidor, os desenvolvedores de soluções começaram a colocar no mercado produtos direcionados para os analistas de decisão, bem mais “amigáveis”.

Hoje, o conjunto de soluções para BI multiplicou-se e a diversidade de produtos é muito grande e continua em constante evolução e crescimento tecnológico.

É possível encontrar desde pacotes pré-configuráveis, até ferramentas “engessadas” e inclusive soluções que permitem às empresas se aventurarem no desenvolvimento de um sistema totalmente caseiro.

Estas ferramentas têm em comum a característica de facilitar a transformação dos “amontoados de dados” em informações de forma a auxiliar os diversos níveis de uma empresa na tomada segura de decisões.

A seguir são enumerados algumas ferramentas que auxiliam a implantar o conceito de BI:

- Planilhas eletrônicas;
- Geradores de consultas baseadas em SQL;
- Sistemas de apoio à decisão (DSS);
- EIS;
- Ferramentas OLAP;
- Ferramentas de BAM;
- Ferramentas ETLs;
- Ferramentas metadados;
- Ferramentas BPM;
- Ferramentas *Data Mining*;

No tópico de ferramentas encontram-se explicações mais detalhadas para ferramentas, que foram escolhidas para serem estudadas neste trabalho, utilizando-as em cada uma das etapas de BI e criando um projeto desde o início.

2.1.7 Como se inicia um projeto de BI?

Primeiramente deve identificar as reais necessidades da empresa, especialmente as das áreas de vendas e marketing e, posteriormente, das finanças. Ou, no caso da geração de indicadores de desempenho, todas as principais áreas da empresa. Também deve ficar claro que apesar desses projetos envolverem o uso das ferramentas e soluções de tecnologia da informação, é importante entender que BI é um projeto de negócios e por isso deve estar alinhado à estratégia global da corporação.

Nesse sentido, esse tipo de iniciativa exige o envolvimento dos profissionais das áreas de negócios, que serão seus principais usuários, e essa participação deve ocorrer

desde a fase de análise e desenvolvimento até a implementação efetiva das soluções. Caberá à área de TI verificar a viabilidade de aquisição de ferramentas que se mostram mais adequadas ao projeto e às possibilidades de integração com os sistemas já instalados na empresa.

A integração de todos os sistemas utilizados, antes de se iniciar a implementação de um projeto de BI é de extrema importância, considerando os dados vindos de diversas fontes como: mainframes, bancos de dados relacionais, servidores, estações cliente, dispositivos móveis entre outros. Mas o pré-requisito fundamental para o sucesso da iniciativa é entender a cadeia de valor da informação para a corporação como um todo.

2.2 Metodologia

Para a realização deste trabalho serão aplicados os conceitos estudados sobre BI com foco em duas ferramentas: Pentaho BI Platform, desenvolvida pela empresa Pentaho Corporation e a ferramenta Qlikview, desenvolvida pela empresa QlikTech. Maiores informações sobre as características de cada ferramenta utilizada poderão ser consultadas no tópico de ferramentas deste projeto.

O trabalho foi construído para que, primeiramente, seja entregue ao leitor uma introdução ao BI e suas definições assim como explicar o seu conceito. Uma explicação em cada uma das principais etapas do BI, posteriormente uma apresentação sobre as ferramentas escolhidas a serem estudadas e por fim uma conclusão destes casos de uso.

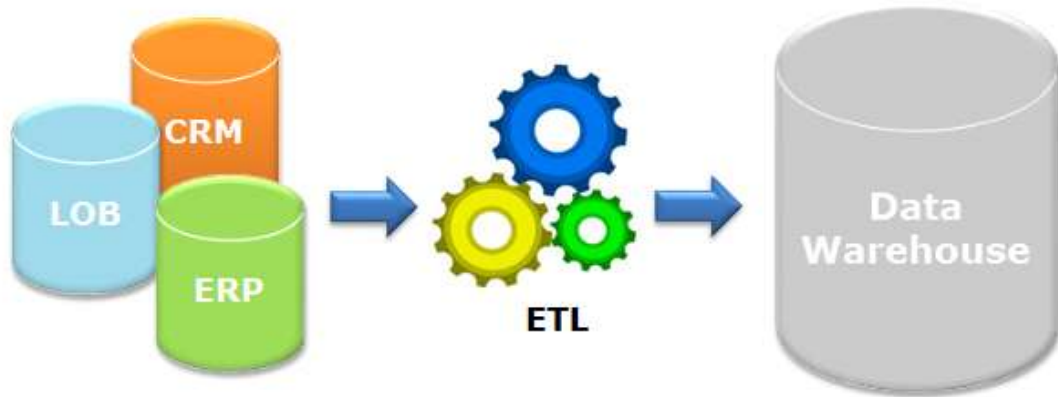
2.3 Etapas de BI

2.3.1 ETL

Essa fase é responsável pela preparação dos dados que farão parte do DW. Ela se dá basicamente em três passos: extração, transformação e carga de dados.

No passo da extração o ETL utiliza de fontes de dados (*Data Sources*) que fornecerão toda a informação utilizada no processo de transformação. Uma vez extraída a informação o ETL entra no segundo passo, este sendo o de transformação (também chamado de Limpeza, Ajustes e Consolidação), nesta etapa os ajustes na informação são

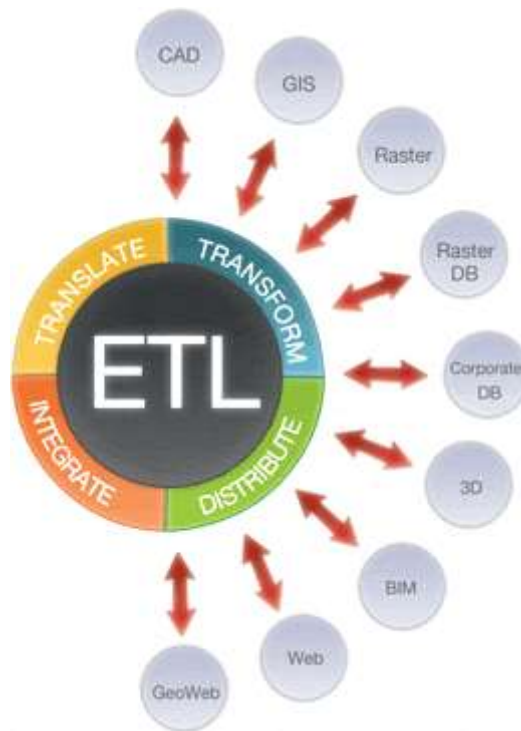
efetuados melhorando sua qualidade para o terceiro e último passo: passo de carga (ou entrega) de dados o ETL fisicamente estrutura e carrega os dados dentro da camada de apresentação fazendo a entrega ao *Data Warehouse*.



Fonte: <http://www.arcplan.com/en/blog/tag/etl/>

Figura 3 - Ilustração do ciclo de vida do ETL em todo seu processo

O ETL é responsável por toda a parte de extração de informação de fontes externas. Mesmo a maioria dessas fontes sendo banco de dados o ETL não se limita a eles, podendo também extrair dados de planilhas do Excel, documentos do ACCESS, documentos de texto, programas de ERP, programas de CRM e diversas outras fontes.



Fonte: <http://mundogeo.com/blog/2010/01/29/spatial-etl/>

Figura 4 - Ilustração da grande quantidade de fontes de dados para extração do ETL

Quando se obtém dados de uma fonte, os mesmos possuem muitas informações incoerentes. Por exemplo, quando um vendedor executa uma venda, a preocupação dele é somente vender, não se importando com a qualidade dos dados. Então se o cliente não estiver com o número de CPF em mãos, ele cadastra um número qualquer, desde que o sistema aceite. Caso o diretor dessa empresa consulte o seu DW para ver quais são os seus maiores clientes corre o risco de em primeiro lugar aparecer o cliente que tem o CPF genérico inserido pelo vendedor. Por isso nessa fase é verificada a compatibilidade entre os dados.

No momento que os dados são levados para o DW eles devem ser padronizados, ou seja, quando o usuário foi consultar os dados ele não verá informações iguais em formatos diferentes. Alguns fatores que devem ser analisados antes de começar a fase de extração dos dados são:

- A extração de dados do ambiente operacional para o DW demanda uma mudança na tecnologia. Os dados são transferidos de bancos de dados hierárquicos para uma nova estrutura de SGBD relacional para *Data Warehouse*.

- A seleção de dados do ambiente operacional pode ser muito complexa, pois muitas vezes é necessário selecionar vários campos de um sistema transacional para compor um único campo no *Data Warehouse*.
- Outro fator que deve ser levado em conta é que dificilmente há o modelo de dados dos sistemas antigos, e, se existem, não são documentados. Os dados são reformatados. Por exemplo: um campo do sistema operacional do tipo DD/MM/AAAA pode ser passado para o outro sistema do tipo ano e mês como AAAA/MM.
- Quando há vários arquivos de entrada, a escolha das chaves deverá ser feita antes que os arquivos sejam intercalados. Isto significa que, se diferentes estruturas de chaves são usadas nos diferentes arquivos de entrada, então, opta-se por apenas uma dessas estruturas.
- Os arquivos devem ser gerados obedecendo à mesma ordem das colunas estipuladas no ambiente de *Data Warehouse*.
- Pode haver vários resultados. Dados podem ser produzidos em diferentes níveis de resumo pelo mesmo programa de geração de cargas.
- Valores *Default* devem ser fornecidos. Às vezes pode existir um campo no *Data Warehouse* que não possui fonte de dados, então a solução é definir um valor padrão para estes campos.

Apesar de existirem ferramentas de ETL ainda tem-se a necessidade de criar rotinas de cargas para atender determinadas situações que poderão ocorrer. O que se pode afirmar sobre os benefícios do uso das ferramentas de ETL é que os benefícios serão bastante vistosos e a produtividade aumentará consideravelmente. (PRIMAK, 2008).

2.3.2 Datawarehouse

Hoje em dia, *Business Intelligence* (BI) é um dos temas discutidos com frequência na maior parte dos mercados de trabalho em todo o mundo. A maioria das empresas estão estabelecendo ou planejando estabelecer um sistema de *Business Intelligence* e um *data warehouse* (DW). Conhecimentos relacionados ao BI e *data warehouse* estão em grande demanda no mercado de trabalho.

O *data warehouse* é o núcleo do sistema de BI. Um *data warehouse* é um banco de dados construído com a finalidade de análise de dados e elaboração de relatórios. Este propósito muda o design deste banco de dados também. Os bancos de dados operacionais são construídos sobre padrões de normalização, que são eficientes para sistemas

transacionais, como por exemplo, para reduzir a redundância. Um banco de dados concebido para um sistema de vendas contém várias tabelas relacionadas entre si. Assim, por exemplo, um relatório sobre as informações de vendas pode consumir mais de 10 *joined conditions*, o que diminui o tempo de resposta de consulta e da criação de um relatório. Um *data warehouse* vem com um novo design que reduz o tempo de resposta e aumenta o desempenho de consultas para relatórios e análises.

O Ciclo de vida do *Data Warehouse* é composto por uma série de etapas. Inicia-se pelo planejamento do Programa ou Projeto, passa pelo levantamento e definição dos requisitos de negócios e então dividido em três caminhos, um seguindo a parte de Arquitetura e Design Técnico, outro trabalhando a definição da Modelagem Dimensional, o Desenho físico e a parte de ETL, e o ultimo e terceiro caminho que trata do planejamento e desenvolvimento da aplicação de BI, o *front end* propriamente dito.

Algumas das vantagens de se utilizar um *Data Warehouse* são citadas abaixo:

Ao fornecer dados de várias fontes, gerentes e executivos não precisam tomar decisões de negócios com base em dados limitados. Além disso, *data warehouses* e BI relacionados podem ser aplicados diretamente aos processos de negócios, incluindo segmentação de marketing, gestão de inventário, gestão financeira e de vendas.

Como os usuários de negócios podem acessar rapidamente dados críticos de uma série de fontes, tudo em um só lugar, podem rapidamente tomar decisões informadas sobre as principais iniciativas. Não perderão tempo recuperando dados preciosos à partir de múltiplas fontes.

Não apenas isso, mas, os executivos, podem consultar os dados com pouca ou nenhuma ajuda de TI, ou seja, poupar mais tempo e mais dinheiro. Isso significa que os usuários de negócios não terão que esperar até que TI gere os relatórios, e as pessoas que trabalham duro em TI poderão fazer o que eles fazem de melhor: manter o negócio funcionando.

A implementação de um *Data Warehouse* inclui a conversão de dados de inúmeros sistemas de origem em um formato comum. Uma vez que cada dado de diversos departamentos é padronizado, cada departamento produzirá resultados que estão de acordo com todos os outros departamentos. Assim, o usuário pode ter mais confiança na precisão de seus dados, dados precisos são a base para as decisões de negócios fortes.

Um *data warehouse* armazena grandes quantidades de dados históricos para que o usuário possa analisar diferentes períodos de tempo e as tendências, a fim de fazer previsões futuras. Esses dados normalmente não podem ser armazenados em um banco de dados transacional ou usados para gerar relatórios a partir de um sistema transacional.

Para a construção de um *Data Warehouse* são necessárias sete etapas:

- Levantamento das necessidades: deve-se antes de tudo fazer o levantamento de todas as informações desejadas pelo usuário. Nesse primeiro momento é realizado o cruzamento de Dimensões e Fatos necessários para alcançar os anseios dos gestores. Nesse primeiro momento deve ser levado em consideração quais os dados que serão armazenados no *Data Warehouse*, logo, não deve se preocupar com a existência efetiva dos dados e sim com os desejos requisitados.
- Mapeamento dos dados: nessa etapa é realizado o mapeamento dos dados, identificando a fonte e como chegar até eles. Aqui vê-se a viabilidade dos desejos elencados na primeira etapa, verificando a existência ou não dos dados para o alcance das necessidades solicitadas.
- Construção da *Staging Area*: após o mapeamento, constrói-se a estrutura chamada *Staging Area*, que é a área de transição dos dados do DW. Nessa área os dados são copiados e desacoplados dos sistemas de operação (OLTP) e recebem o devido tratamento para as futuras cargas nas tabelas de Fatos e Dimensões.
- Construção das Dimensões: constrói-se nessa etapa a estrutura das Dimensões que farão parte do DW. São definidos também a historicidade que os dados irão possuir nas Dimensões.
- Construção dos Fatos: constrói-se nessa etapa as estruturas dos Fatos. Aqui é avaliado e definido a granularidade da informação que será armazenada em cada Fato. São avaliados também a expectativa de crescimento e de armazenamento que serão utilizados.
- Definição do processo geral de carga: após a conclusão das etapas anteriores, é necessária a criação do motor para que tudo seja carregado, atualizado, orquestrado e processado de forma automática e ordenada. Por isso, a necessidade do processo geral de carga que é o “cérebro” do DW.
- Criação dos metadados: por fim, precisa ser desenvolvida toda a documentação dos metadados, que incluem o processo de construção e o dicionário de dados. Os metadados fornecem apoio importante para a gestão do conhecimento.

Assim, pode-se organizar um fluxo de processo da seguinte maneira:

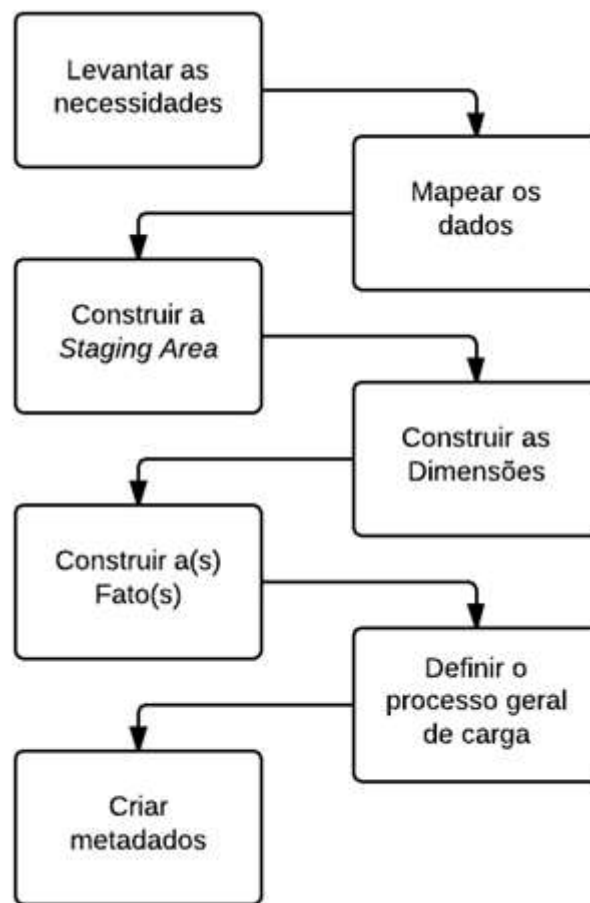


Figura 5 - Passos para a criação de um *Data Warehouse*

2.3.2.1 Modelo dimensional para data warehouse

Segundo Moreira Eduardo o modelo dimensional para construção de banco de dados para *Data Warehouse* é uma forma de modelagem onde as informações se relacionam de forma que podem ser representadas como um cubo. Sendo assim pode-se fatiar este cubo e aprofundar em cada dimensão ou eixo para extrair mais detalhes sobre os processos internos que ocorrem na empresa.

O modelo dimensional permite visualizar dados abstratos de forma simples e relacionar informações de diferentes setores da empresa de forma muito eficaz.

O que torna o *Data Warehouse* mais poderoso é que informações que se situam em vários sistemas, planilhas e arquivos espalhados por todos os setores da empresa, são reunidos em um banco de dados de forma dimensional, tendo informações unificadas e padronizadas em um mesmo local.

Um modelo de dados dimensional é extremamente simples, isto facilita para os usuários deste banco de dados identificarem onde estão localizadas as informações e permite que os softwares naveguem por estes bancos de dados com eficiência. Outro fator importante para a modelagem dimensional é a velocidade de acesso à uma informação, com modelos simples sem muitas tabelas para relacionar, é muito rápido para extrair as informações necessárias.

Um modelo dimensional conta basicamente com uma tabela de fatos central e tabelas dimensionais ligadas diretamente a elas.

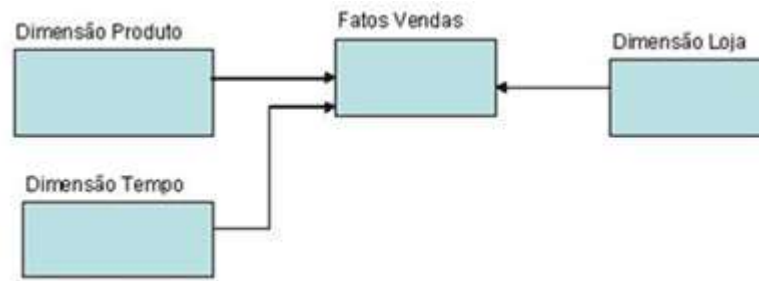


Figura 6 - Modelo dimensional

Fonte (<http://imasters.com.br/artigo/3836/gerencia-de-ti/modelo-dimENSIONAL-para-data-warehouse/>)

Os Fatos e Dimensões são tabelas do banco de dados, só que no modelo dimensional adquirem nomes de Fatos e Dimensões de acordo com a função da tabela.

2.3.2.2 Tipos de modelos dimensionais

- **Modelo Estrela (*Star Schema*)**

No modelo estrela todas as tabelas relacionam-se diretamente com a tabela de fatos, sendo assim as tabelas dimensionais devem conter todas as descrições que são necessárias para definir uma classe como Produto, Tempo ou Loja nela mesma, ou seja, as tabelas de dimensões não são normalizadas no modelo estrela, então campos como Categoria, Departamento, Marca contém suas descrições repetidas em cada registro, assim

aumentando o tamanho das tabelas de dimensão por repetirem estas descrições de forma textual em todos os registros.

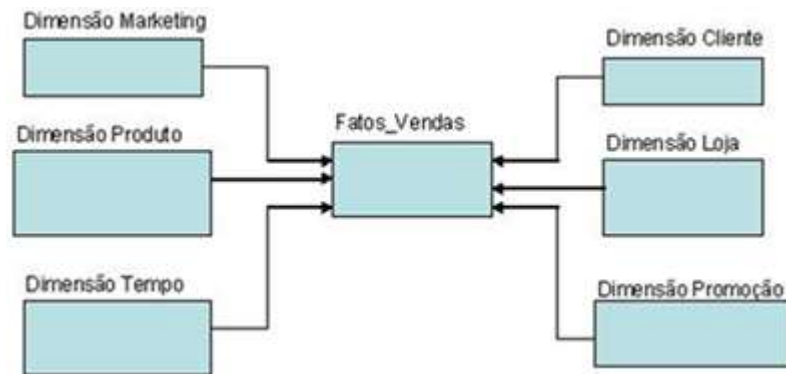


Figura 7: Modelo Estrela

Fonte (<http://imasters.com.br/artigo/3836/gerencia-de-ti/modelo-dimENSIONAL-para-data-warehouse/>)

Este modelo é chamado de estrela porque a tabela de fatos fica ao centro cercada das tabelas dimensionais assemelhando-se a uma estrela. Mas o ponto forte a fixar é que as dimensões não são normalizadas.

- **Modelo Floco de Neve (*Snow Flake*)**

No modelo Floco as tabelas dimensionais relacionam-se com a tabela de fatos, mas algumas dimensões relacionam-se apenas entre elas, isto ocorre para fins de normalização das tabelas dimensionais, visando diminuir o espaço ocupado por estas tabelas, então informações como Categoria, Departamento e Marca tornaram-se tabelas de dimensões auxiliares.

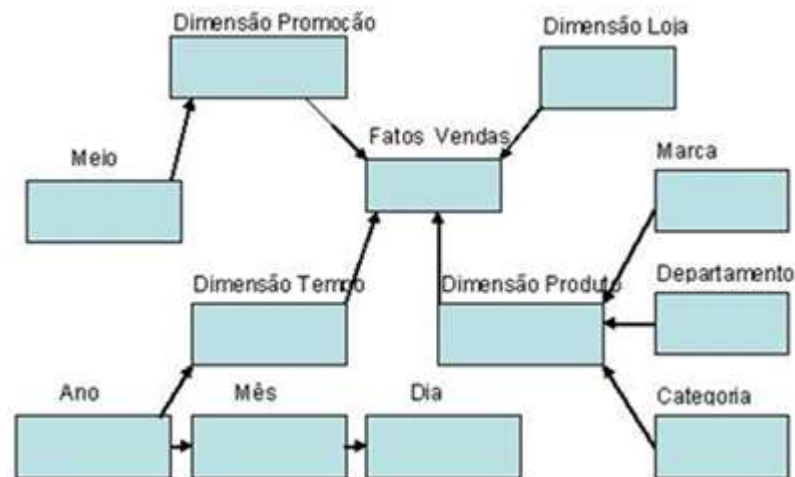


Figura 8: Modelo floco de neve

No modelo Floco existem tabelas de dimensões auxiliares que normalizam as tabelas de dimensões principais. Na figura anterior estas tabelas são (Ano, Mês e Dia) que normalizam a Dimensão Tempo, (Categoria, Departamento e Marca) que normalizam a Dimensão Produto e a tabela Meio que normaliza a Dimensão Promoção.

Construindo a base de dados desta forma, passou-se a utilizar mais tabelas para representar as mesmas dimensões, mas ocupando um espaço em disco menor do que o modelo estrela. Este modelo chama-se floco de neve, pois cada dimensão se divide em várias outras tabelas, onde organizadas de certa forma lembra um floco de neve.

2.3.2.3 Considerações sobre ambos os modelos

O Modelo Floco (Snow Flake) reduz o espaço de armazenamento dos dados dimensionais mas acrescenta várias tabelas ao modelo, deixando-o mais complexo, tornando mais difícil a navegação pelos softwares que utilizarão o banco de dados. Um outro fator é que mais tabelas serão utilizadas para executar uma consulta, então mais JOINS de instrução SQL serão feitos, tornando o acesso aos dados mais lento do que no modelo estrela.

O Modelo Estrela (Star Schema) é mais simples e mais fácil de navegação pelos software, porém desperdiça espaço repetindo as mesmas descrições ao longo de toda a tabela, porém análises feitas mostram que o ganho de espaço normalizando este esquema

resulta em um ganho menor que 1% do espaço total no banco de dados, sendo assim existem outros fatores mais importantes para serem avaliados para redução do espaço em disco como a adição de agregados e alteração na granularidade dos dados, estes temas serão abordados em tópicos posteriormente.

2.3.3 Data marts

Os *Data Marts* (DM) são DWs de escopo limitados, pequenos armazéns de dados específicos para cada área de negócios ou departamentos das empresas. Tanto os DWs como DMs formam a base a partir da qual as empresas podem utilizar as ferramentas de BI para a extração de informações gerenciais (BARBIERI, 2001).

Segundo Serra (2002, p.12) um DM é “uma parte do *Data Warehouse* restrita a um único processo de negócio, ou a um grupo de processos de negócio relacionados entre si e voltados para um grupo de negócio particular.” O DM trata de problema local ou departamental, e o DW envolve a empresa como um todo para que o apoio a decisões atue em todos os níveis da empresa.

2.3.4 OLAP (on-line analytics processing)

Segundo Ribeiro (2011) OLAP é um software cuja tecnologia de construção permite aos analistas de negócios, gerentes e executivos analisarem e visualizarem os dados corporativos de forma rápida, consistente e principalmente interativa.

A funcionalidade OLAP é inicialmente caracterizada pela análise dinâmica e multidimensional dos dados consolidados de uma organização permitindo que as atividades do usuário final sejam tanto analíticas quanto navegacionais.

As ferramentas OLAP (do inglês, *Online Analytical Processing*) são geralmente desenvolvidas para trabalhar com banco de dados desnormalizados. Essas ferramentas são capazes de navegar pelos dados de um *Data Warehouse*, possuindo uma estrutura adequada tanto para a realização de pesquisas como para a apresentação de informações.

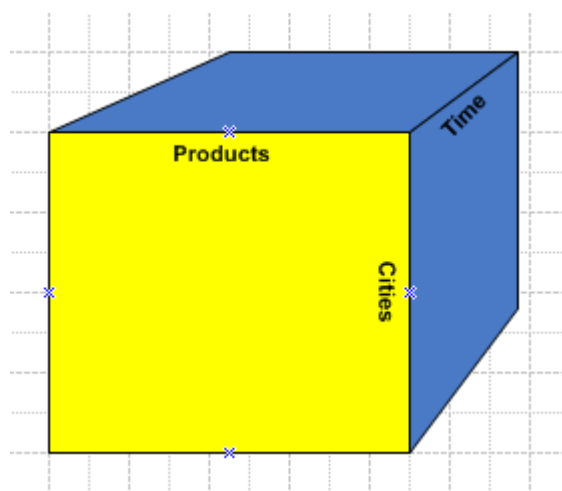
Soluções informatizadas de BI geralmente contém sistemas analíticos, que podem ser de diversos tipos, dependendo do objetivo das análises e do perfil do usuário, conforme ilustrado na Figura:



Fonte: <http://vivianeribeiro1.wordpress.com/2011/07/12/o-que-e-olap/>

Figura 9 - Tipos de sistemas analíticos para soluções informatizadas em BI

Segundo Teban (2012), em paralelo a isso existe o termo "cubo OLAP" que dita pelo Codd Edgar como sendo um banco de dados multidimensional, em que o armazenamento físico dos dados é executado em um vetor multidimensional. Os cubos OLAP podem ser considerados como uma extensão das duas dimensões de uma folha de cálculo. Muitas vezes pensa-se que tudo o que os usuários podem querer de um sistema de informação poderia ser feito em um banco de dados relacional. No entanto Codd foi um dos precursores dos bancos de dados relacionais, de modo que seus pontos de vista foram e são respeitados.



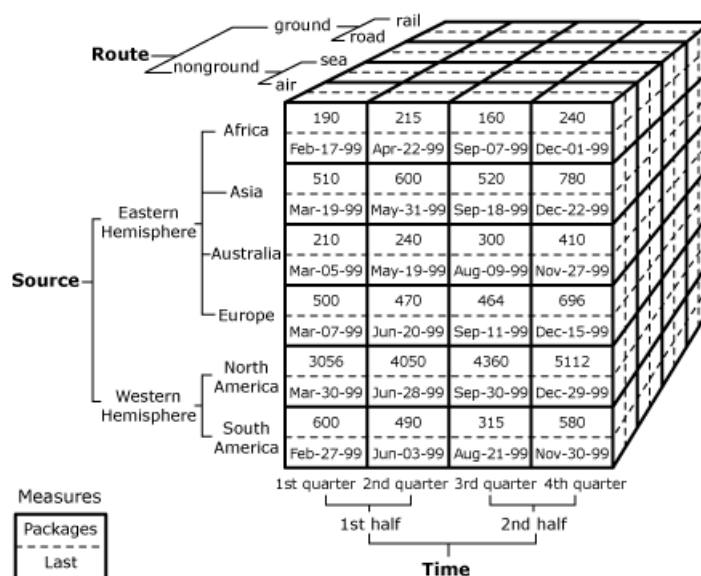
Fonte: http://en.wikipedia.org/wiki/OLAP_cube

Figura 10 - Modelo Cubo OLAP

A proposta Codd era fazer uma disposição dos vetores de dados para permitir a análise rápida. Esses vetores são chamados de cubos. Organizar os dados em cubos evita uma limitação de bancos de dados relacionais, que não são adequados para a análise imediata de grandes quantidades de dados. Os bancos de dados relacionais são mais adequados para gravação de dados de transações (conhecidos como OLTP ou processamento transações on-line). Embora existam muitas ferramentas de relatórios para bancos de dados relacionais, eles são lentos, quando deveria ser explorado em todo o banco de dados. Por exemplo, uma empresa pode analisar alguns dados financeiros por produto, por tempo, por cidade, por tipo de receitas e despesas, e comparando os dados reais com um orçamento. Esses parâmetros, dependendo de quais análises de dados são conhecidas como dimensões.

Para acessar os dados só precisa posicioná-los a partir dos valores das dimensões ou eixos. Armazenar fisicamente os dados desta forma tem seus prós e contras. Por exemplo, nestes bancos de dados de consultas seleção são muito rápidos (na verdade, quase instantâneo), mas um dos maiores problemas desta forma de armazenamento é que uma vez preenchido o banco de dados não pode receber mudanças em sua estrutura, exigindo redesenhar o cubo.

OLAP num sistema pode ter mais de três dimensões, de modo que os cubos de OLAP também são chamados *hypercubes*.



Fonte: [http://technet.microsoft.com/pt-br/library/ms174587\(v=sql.105\).aspx](http://technet.microsoft.com/pt-br/library/ms174587(v=sql.105).aspx)

Figura 11 - Exemplo cubo que contém duas medidas, pacotes, e último, e três dimensões relacionadas, Rota, Origem e Horário

A principal característica que permite OLAP, é que ele é mais rápido quando a execução de instruções SQL tipo SELEÇÃO, ao contrário de OLTP é a melhor opção para operações do tipo INSERIR, ATUALIZAR e EXCLUIR.

Em resumo o OLAP parte da ideia de analisar as informações e capacitar atribuindo aos SI que permitem aos gestores examinarem e manipularem, interativamente grandes quantidades de dados detalhados e consolidados a partir de análises de diversas perspectivas, de forma on-line.



Fonte: (<http://technet.microsoft.com/pt-br/library/cc668463.aspx>)

Figura 12 - Arquitetura de um sistema OLAP

2.3.5 Data mining

Outro componente importante em um sistema de BI: a mineração de dados. A mineração de dados, ou *data mining*, pode ser realizada de duas maneiras; análise dos dados e previsões atuais, com base em um modelo obtido a partir de um conjunto de dados

existente. O conceito de análise de um conjunto de dados existente é chamado o modelo descritivo.

A mineração de dados é o processo de encontrar o problema, pensando na solução, testando modelos de mineração de dados em um conjunto de dados de teste e implantação de um ou mais modelos de mineração em um ambiente vivo. O processo de mineração de dados não conclui nenhuma dessas etapas; é um processo circular que continua a melhorar a análise de dados com o tempo.

Diversos processos definem as atividades e as fases de *Data Mining*, apesar das diferenças, em geral, todos possuem uma estrutura similar. Neste trabalho, foi utilizado o CRISP como modelo. O processo consiste de maneira cíclica, apesar de ser composto por fases o fluxo pode ir e voltar entre as fases:

- **Entendimento dos negócios:** Compreender os objetivos do projeto e os requisitos de uma perspectiva de negócios, e depois converter esse conhecimento em uma definição do problema de DM e um plano preliminar projetado para atingir os objetivos.
- **Entendimento dos dados:** Começa por recolher dados, então, se familiarizar com os dados, para identificar problemas de qualidade de dados, para descobrir primeiros insights sobre os dados, ou para detectar subconjuntos interessantes para formar hipóteses sobre a informação escondida.
- **Preparação de dados:** Inclui todas as atividades necessárias para a construção do conjunto de dados final (dados que serão alimentados na ferramenta de modelagem) a partir dos dados brutos iniciais. As tarefas incluem a tabela, caso, e seleção de atributos, bem como a transformação e limpeza de dados para ferramentas de modelagem.
- **Modelagem:** Selecionar e aplicar uma variedade de técnicas de modelagem e calibrar parâmetros da ferramenta com os valores ideais. Tipicamente, existem várias técnicas para o mesmo tipo de problema de mineração de dados. Algumas técnicas têm requisitos específicos sobre a forma de dados. Portanto, recuar para a fase de preparação de dados é muitas vezes necessário.
- **Avaliação:** Avaliar minuciosamente o modelo e rever os passos executados para a construção do modelo, para ter certeza que atinge adequadamente os objetivos de negócios. Determinar se há alguma questão de negócios importante que não foi suficientemente analisado. No final desta fase, ser tomada uma decisão quanto ao uso dos resultados de mineração de dados.

- **Distribuição:** Organizar e apresentar os resultados de mineração de dados. A implantação pode ser tão simples como gerar um relatório ou tão complexo como a implementação de um processo de mineração de dados repetível.

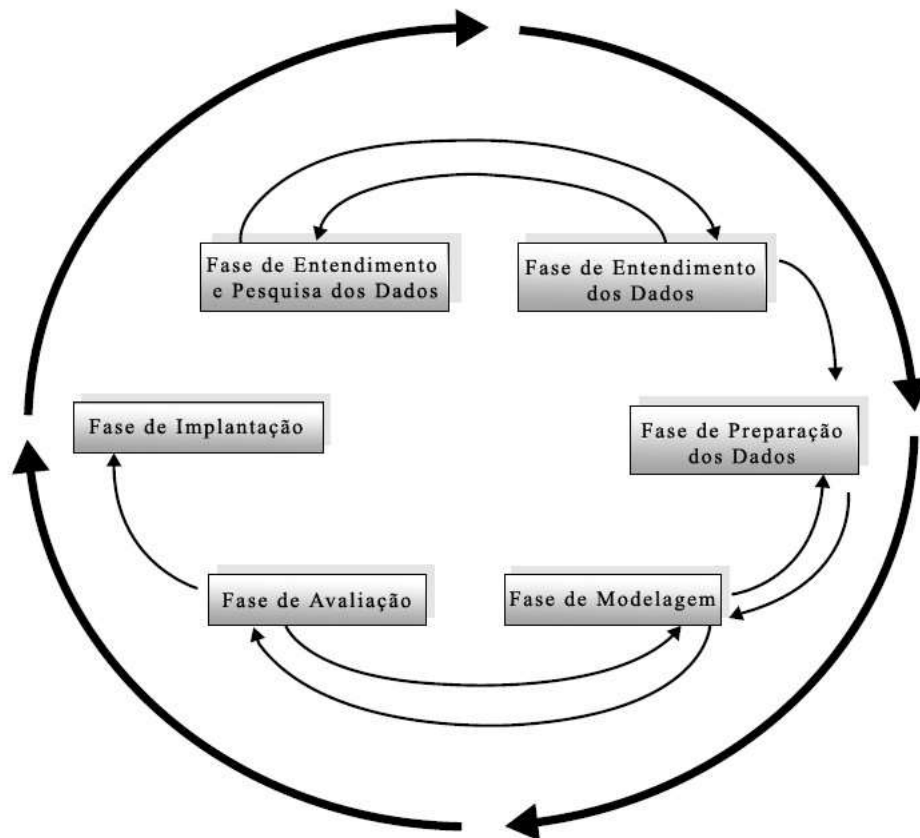


Figura 13: Ciclo de Data Mining

De acordo com Barbieri (2001) o *Data Mining* pode desempenhar ainda um número limitado de tarefas, tais como:

- **Classificação:** exame dos aspectos de um objeto e atribuição a um conjunto de classes predefinidas, sendo esses objetos representados por registros em um banco de dados e sua classificação, consistindo em um preenchimento de um campo com um código de classe de algum tipo;
- **Estimativa:** depois do fornecimento de alguns dados, a estimativa é usada para estipular um valor para alguma variável contínua e desconhecida como, por exemplo, receita ou saldo de cartão de crédito;

- **Previsão:** tendo a mesma conotação da classificação e da estimativa, a previsão se difere pelo fato de que os registros são classificados de acordo com alguma atitude futura prevista ou valor futuro estimado;
- **Agrupamento por Afinidade ou Análise de Seleção Estatística:** esta atividade se resume pela determinação das coisas possíveis de serem agrupadas, podendo ser usado também, para identificar oportunidades de vendas casadas projetando pacotes atraentes de produtos e serviços;
- **Segmentação:** corresponde ao processo de agrupamento de uma população heterogênea em vários subgrupos ou clusters mais homogêneos. Sua principal função está relacionada à mineração de dados, ela realiza a extração para verificar padrões nos dados e realiza a interpretação dos mesmos. Esta tarefa tem como função converter os dados coletados em oportunidades comerciais, e resultados acionáveis para o estágio de ação, satisfazendo assim, as necessidades da empresa.

2.4 Ferramentas

No mercado existem diversos software que utilizam do conceito de *Business Intelligence* fazendo com que as empresas escolham aquele que acharem melhor para trabalhar com suas informações.

Como são várias ferramentas de soluções em BI, sendo cada uma de uma empresa diferente, seria impossível descrever cada uma delas nesse trabalho por isso as descrições apresentadas nesse trabalho foram feitas em cima de duas ferramentas, sendo estas: QlikView e Pentaho BI Platform.

2.4.1 Qlik View

Desenvolvido pela QlikTech, o QlikView é uma plataforma de BI que combina a interação do usuário com análises de informações gráficas de fácil compreensão. É possível utilizá-lo como uma ferramenta de relatórios de tabelas ou a criação de elaborados gráficos em seu painel de controle (*dashboard*).

O QlikView é a primeira plataforma a utilizar o conceito de “*in-memory*”, onde todos os dados das tabelas são carregados na memória RAM do computador tendo como

resultado uma velocidade incrível toda vez que há a necessidade de disponibilizar estes dados na tela.

Outro ponto que deve ser citado é que o QlikView é fácil de aprender e de configurar podendo ser utilizado também por gestores de empresas e pessoas que não possuem um conhecimento avançado na área de tecnologia.

2.4.1.1 Fases do Qlikview

Segundo o treinamento disponibilizado gratuitamente no site do QlikView é necessário seguir quatro passos para utilizar a plataforma com sucesso. Esses passos são: Requisitos (*Requirements*), *Extract Transform Load* (ETL), Construção da Interface (Build Interface) e por último Distribuição (*Distribute*). Este trabalho terá como foco os três primeiros passos devido ao último passo “*Distribute*” necessitar de um software pago (*QlikView Server*).



Figura 14 - Fases do QlikView

Fonte (vídeo - http://qvtraining.qlikview.com/DOWNLOADS_FREE-ELEARNING/Build_Your_First_QlikView/Mod_1_Process_Overview/Process_Overview.htm)

2.4.1.2 Requisitos

Os requisitos dentro de um projeto QlikView não são diferentes dos projetos de programação ou de outras plataformas de BI. Deve-se primeiramente saber o que será analisado, quais as tabelas dentro de sua base de dados que serão utilizadas e dentro de quais as perspectivas serão mostradas essas informações.

Para este trabalho foram desenvolvidos dois projetos, um deles dentro do formato *Star Schema* e o outro no formato *Snow Flake*.

O projeto em *Star Schema* possui os seguintes requisitos:

- Análise do número de vendas.
- Tabelas utilizadas: *Orders*, *Employees* e *Customers* (Vendas, Funcionários e Clientes).
- As Perspectivas são: Data da venda, mês da venda, ano da venda, país do cliente, venda por funcionários, cargos de funcionários e venda por cliente.

Com estes requisitos bem definidos pode-se montar um diagrama da base em *Star Schema* para prosseguir com os demais passos.

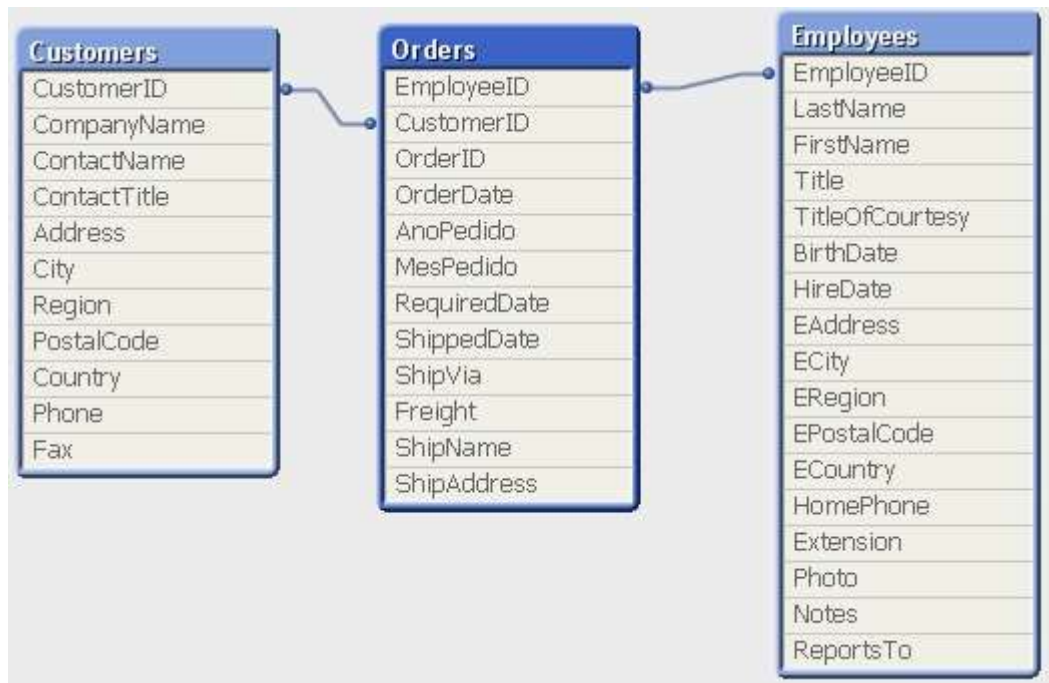


Figura 15 - Diagrama em *Star Schema*

O projeto em *Snow Flake* possui os seguintes requisitos:

- Análise do número de vendas e do número de itens vendidos.
- Tabelas utilizadas: *Orders*, *Employees*, *Customers*, *Order Details* e *Products* (Vendas, Funcionários, Clientes, Detalhes da Venda e Produtos).
- As Perspectivas são: data da venda, mês da venda, ano da venda, país do cliente, venda por funcionários, cargos de funcionários, venda por cliente, número de vendas de itens por produtos.

Com estes requisitos bem definidos pode-se montar um diagrama da base em *Snow Flake* para prosseguir com os demais passos.

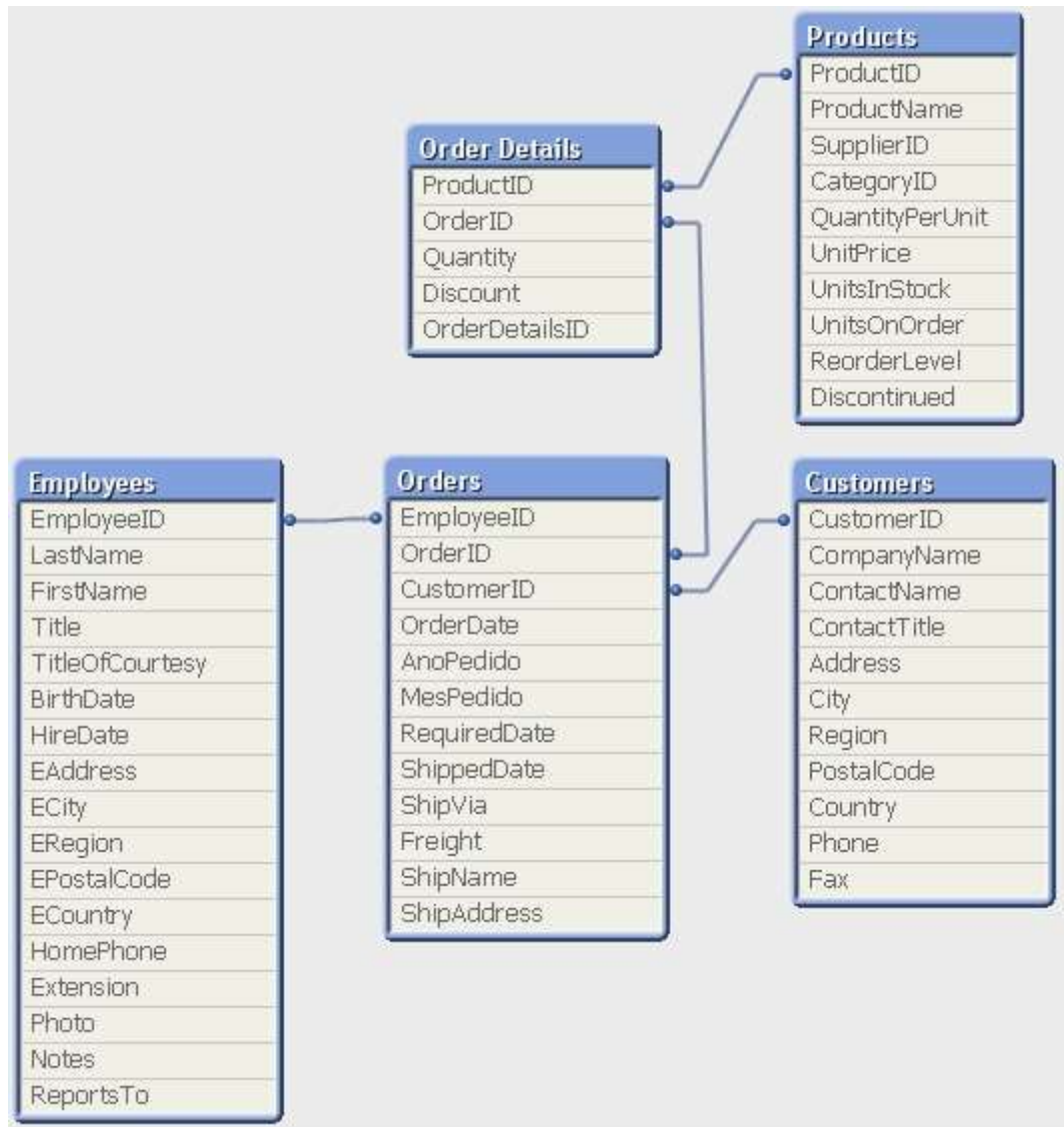


Figura 16 - Diagrama em *Snow Flake*

O *QlikView* cria as próprias ligações entre tabelas sem necessitar de Chaves Estrangeiras (*Foreign Keys*). Ele cria essas ligações através do nome das colunas, portanto se uma base possuir em duas tabelas nomes de colunas semelhantes deve-se, antes de iniciar o processo de ETL, criar seu *Data Warehouse* com ao menos uma dessas tabelas tendo suas colunas renomeadas, evitando assim uma possível ligação indesejada e evitando problemas. No exemplo nesse trabalho algumas colunas em *Employees* (Funcionários) tiveram de ser alteradas como demonstrado nas figuras a cima.

2.4.1.3 ETL

A plataforma QlikView conta com suporte para extrair informação de qualquer tipo de base de dados desde os SGBD com grandes bases de dados até mesmo planilhas em MS Excel e bases em MS Access. Além destes tipos citados também é possível importar informações de arquivos de texto (txt), arquivos XML, aplicativos corporativos como SAP e qualquer tipo de dado que possa ser importado através do driver ODBC.

Para este trabalho foram testados arquivos MS Excel, MS Access, bases de dados em SQL Server e em PostgreSQL, todos funcionaram com sucesso utilizando o driver ODBC sendo o servidor SQL Server escolhido para ser utilizado nos projetos.

Ao configurar o driver ODBC e apontá-lo para a base de dados, o QlikView importa a conexão dentro de uma janela de Scripts.

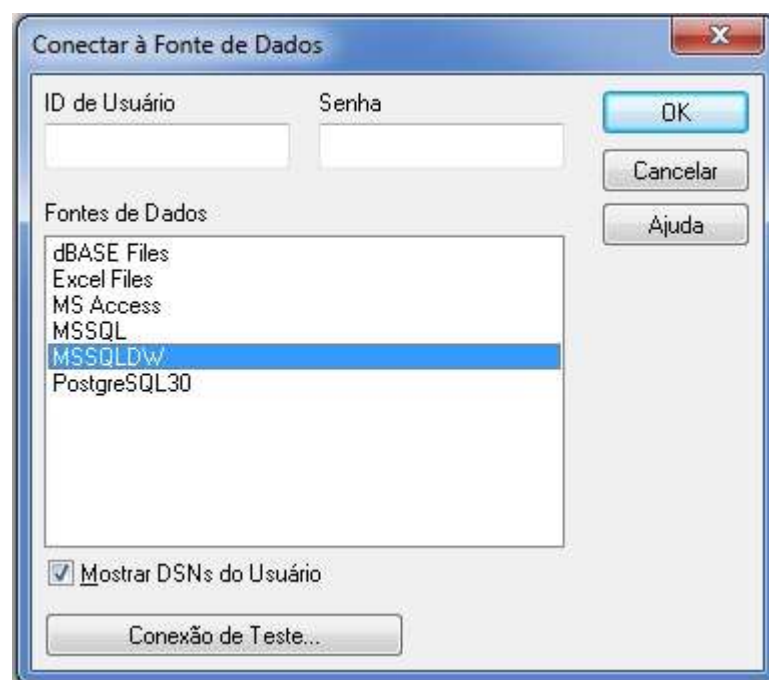


Figura 17 - Janela para conexão da base de dados através do ODBC

A figura 18 abaixo exibe os scripts padrão do QlikView e logo abaixo o script gerado para conexão pelo ODBC que foi feita na figura 17.

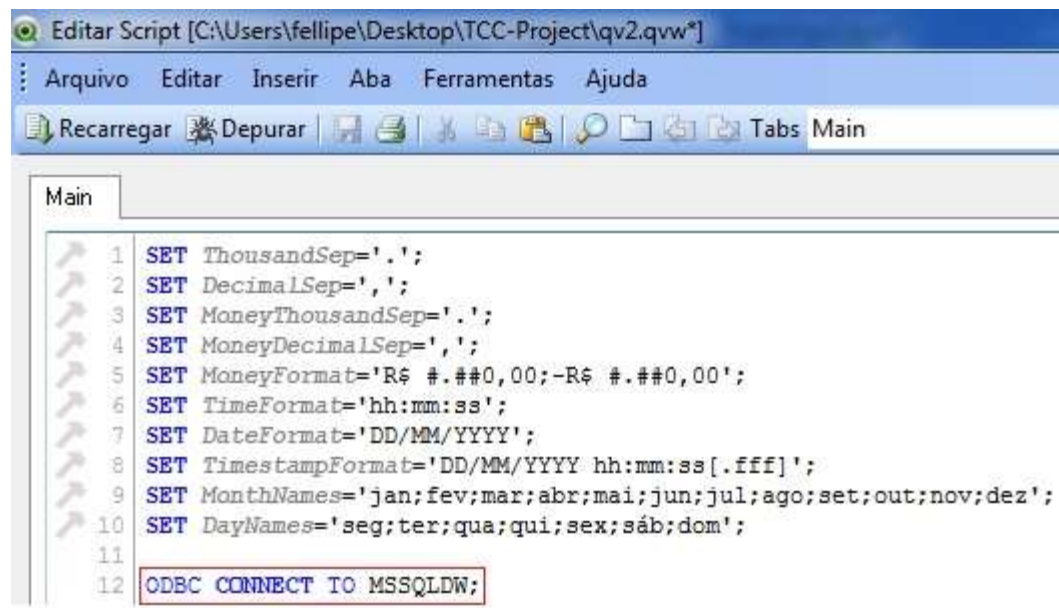


Figura 18 - Janela de Scripts que demonstra os scripts padrões e o script de conexão ODBC

A plataforma QlikView utiliza de scripts em SQL para que seja feito a extração de dados (ETL) e carregar estes mesmos dados em memória, pode-se fazer isso através da interface gráfica escolhendo quais tabelas ou colunas das tabelas deseja carregar ou simplesmente digitar o comando em SQL. Segue Figura abaixo para melhor compreensão.

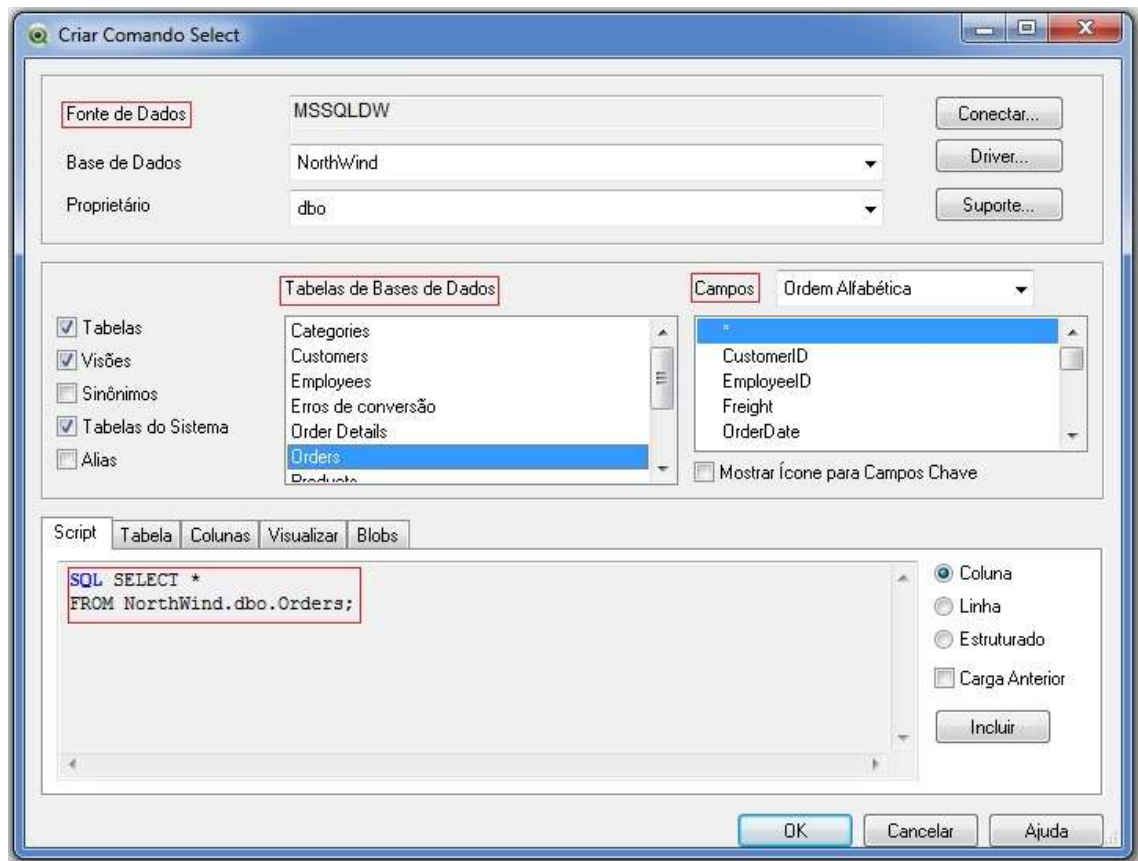


Figura 19 - Importação de dados interface gráfica e seu código SQL gerado abaixo

Para satisfazer os requisitos dentro de ambos os modelos (*Star Schema* e *Snow Flake*) deve-se ter acesso aos campos Mês do Pedido e Ano do Pedido de forma separada da data. Na tabela *Orders* tem-se a data do pedido em formatação única colocando o dia, mês e ano em apenas uma coluna chamada de *OrderDate*.

Nesse exemplo então foi necessário modificar manualmente a seleção de dados que foi gerada automaticamente para criar dois novos filtros que são;

- AnoPedido: Ano em formato de número de forma a filtrar os dados na perspectiva desses anos
- MesPedido: Mês em formato de nome de forma a filtrar os dados na perspectiva desses meses.

Segue abaixo o script pronto de acordo com a necessidade descrita no exemplo de requisito.

```

12 ODBC CONNECT TO MSSQLDW;
13 //----- Start Multiple Select Statements -----
14 SQL SELECT *
15 FROM NorthWind.dbo.Customers;
16
17
18 SQL SELECT OrderID,
19             CustomerID,
20             EmployeeID,
21             OrderDate,
22             Year(OrderDate) as AnoPedido,
23             CASE Month(OrderDate)
24                 WHEN '1' THEN '01 - Jan'
25                 WHEN '2' THEN '02 - Feb'
26                 WHEN '3' THEN '03 - Mar'
27                 WHEN '4' THEN '04 - Abr'
28                 WHEN '5' THEN '05 - Mai'
29                 WHEN '6' THEN '06 - Jun'
30                 WHEN '7' THEN '07 - Jul'
31                 WHEN '8' THEN '08 - Ago'
32                 WHEN '9' THEN '09 - Set'
33                 WHEN '10' THEN '10 - Out'
34                 WHEN '11' THEN '11 - Nov'
35                 ELSE '12 - Dez'
36             END as MesPedido,
37             RequiredDate,
38             ShippedDate,
39             ShipVia,
40             Freight,
41             ShipName,
42             ShipAddress
43 FROM NorthWind.dbo.Orders;
44

```

Figura 20 - Script de seleção modificado para obter Ano do pedido e Mês do pedido

Depois de ter escolhido as demais tabelas e colunas a serem utilizadas pode-se clicar em Recarregar para que estas informações sejam carregadas em memória.



Figura 21 - Botão Recarregar que carrega as informações em memória

2.4.1.4 Construção da interface

Esta é a fase em que o QlikView tem seu maior foco, desenvolver uma interface que seja de fácil entendimento ao usuário e que este mesmo usuário possa interagir com listas, gráficos e demais parâmetros a fim de conseguir extrair a informação que precisa para dirigir seu negócio.

A ferramenta QlikView conta com diversos objetos como listas, gráficos, tabelas, caixas de estatísticas, botões, objetos de pesquisas e demais objetos citados na figura abaixo

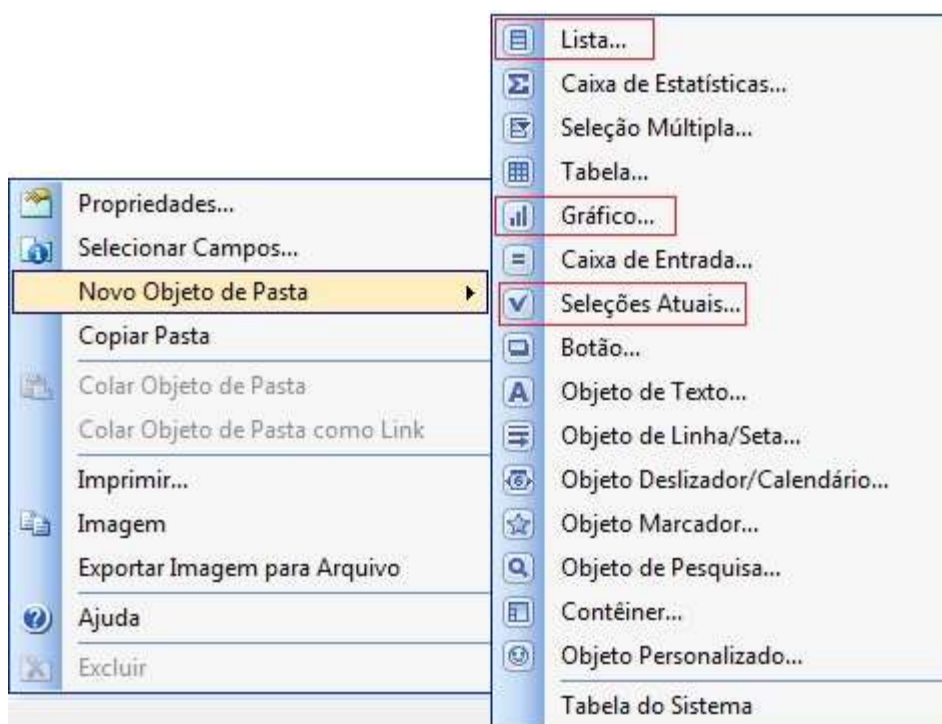


Figura 22 - Objetos que podem ser usados para apresentar e filtrar dados no QlikView

Os Objetos que foram utilizados neste trabalho estão em evidência na figura 22.

Neste trabalho, listas foram utilizadas para filtrar a consulta pelas datas, ano do pedido e mês do pedido. Na figura abaixo pode-se ver que o “AnoPedido” selecionado é “1995” e o “MesPedido” selecionado é o mês “Jun” que é o mês 06 em nosso calendário, logo pode-se ver que em OrderDate as datas de vendas que correspondem a este ano e

mês são filtradas sendo ordenadas a cima das datas que não correspondem com o filtro (datas de 08/06/1995 até 26/06/1995 correspondem enquanto as demais em negrito não).

AnoPedido	
1994	1995
1996	
OrderDate	MesPedido
08/06/1995	Ago
12/06/1995	Set
14/06/1995	Out
15/06/1995	Nov
19/06/1995	Dez
21/06/1995	Jan
22/06/1995	Feb
23/06/1995	Mar
26/06/1995	Abr
04/08/1994	Mai
05/08/1994	Jun
08/08/1994	Jul

Figura 23 - Listas para filtrar por data

Gráficos em formato de pizza e em barras proporcionam uma visualização melhor dos dados e também é possível utiliza-los para filtrar informações apenas clicando no mesmo. Na figura 24 apresenta um exemplo de gráfico em pizza, ao lado estão os países (Country) dos clientes (*Customers*), pode-se ver os números de venda ao lado de cada nome de país e no centro está o gráfico para melhor visualização destes números. É interessante optar pelo gráfico em pizza, pois na maioria das vezes são tantas informações que um gráfico em barra ficaria muito maior do que o monitor do usuário, já o gráfico em pizza se divide em várias “fatias” e não ocupa tanto espaço em tela.

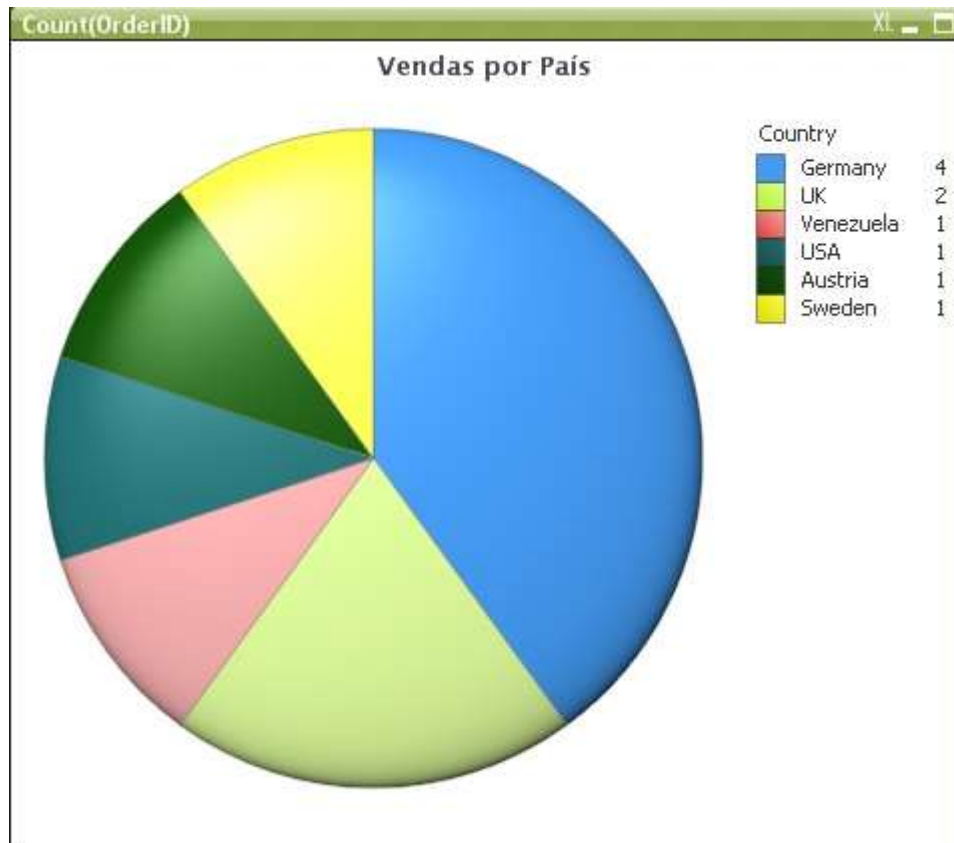


Figura 24 - Gráfico em Pizza mostrando vendas por país

O objeto “Seleções Atuais” é de grande importância uma vez que pode ser visto todos os filtros aplicados no relatório e através dessa ferramenta pode-se também retirar alguma seleção indesejada.

Na figura 25 o objeto “Seleções Atuais” mostra que os filtros selecionados são para “Ano Pedido” o ano 1995, MesPedido o mês “Jun” e os nomes dos funcionários (*Employees*) “Anne, Janet, Laura”. Isso demonstra que é possível selecionar mais de um parâmetro para um mesmo filtro, podendo neste caso fazer uma comparação entre as vendas dos três funcionários selecionados neste ano e neste mês.



Figura 25 - Objeto Seleções Atuais em funcionamento

Neste trabalho foram criados dois modelos de interface de usuário, um para o modelo *Star Schema* e o outro para o modelo *Snow Flake*, ambos já comentados anteriormente.

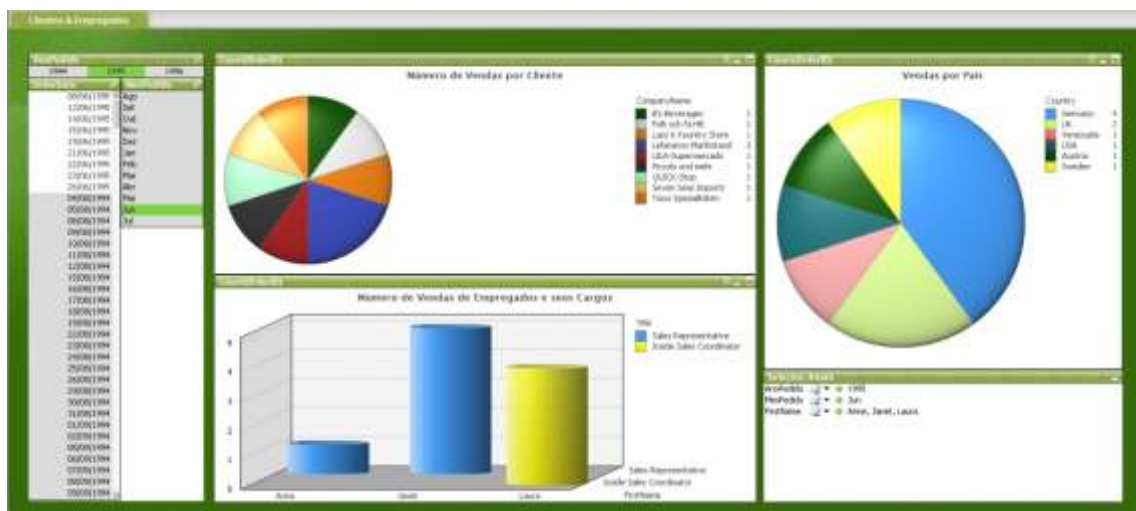


Figura 26 - Projeto QlikView criado a partir do modelo *Star Schema*

A figura 26 mostra o projeto criado no modelo *Star Schema* tendo como fato as Vendas (*Orders*) e demais requisitos citados anteriormente. No canto esquerdo estão as listas para filtrar por data, ano do pedido e mês do pedido; No centro tem-se dois gráficos, sendo o gráfico em pizza referente ao número de vendas por cliente (*Customers*) e no gráfico em barras tem-se as vendas por funcionários (*Employees*) que também divide os funcionários por cargos (cor da barra); No canto direito tem-se um gráfico de pizza de vendas por país e o objeto de seleções atuais.

Os filtros podem ser feitos a partir das listas e também dos gráficos, clicando no desenho do gráfico ou selecionando um ou mais nomes listados a direita.

Quando o usuário escolhe um filtro automaticamente todos os gráficos atualizam filtrando segundo o novo critério de filtro, fazendo com que o usuário possa combinar diversos filtros podendo encontrar a informação que procura mais fácil e dinamicamente.

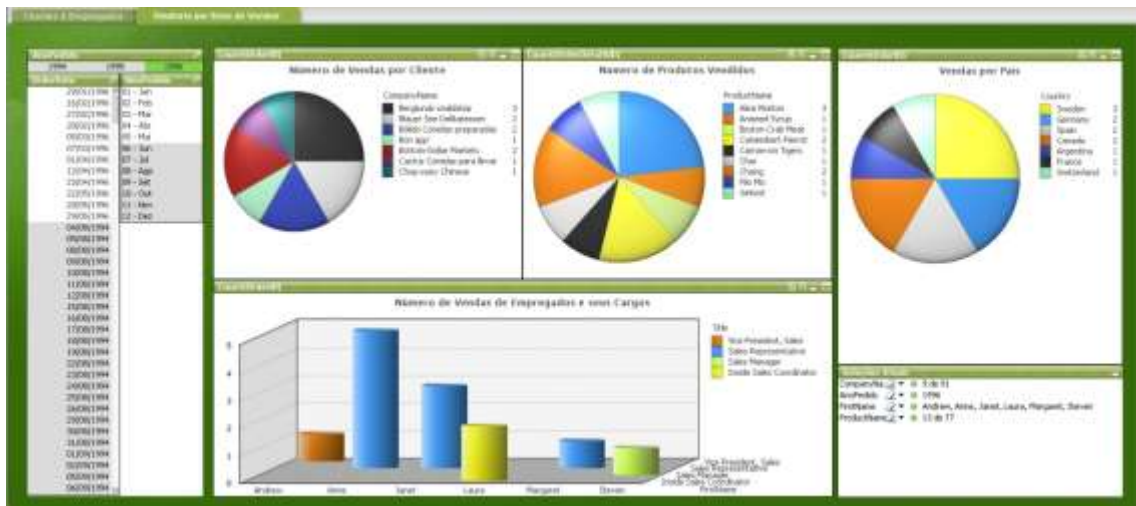


Figura 27 - Projeto QlikView criado a partir do modelo *Snow Flake*

Pode-se ver na figura 27 o projeto criado a partir do modelo *Snow Flake* tendo os mesmos requisitos já citados para esse modelo em Requisitos.

Este projeto segue com a mesma ideia do projeto anterior, porém ao adicionarmos a tabela produtos (*Products*) e os detalhes da venda (*Orders Details*) que é onde está a contagem de itens vendidos, tendo essa informação em memória pode-se criar um gráfico em pizza para contagem de itens vendidos e mostrar quais os produtos vendidos, podendo também filtrar por produto ou saber quais produtos foram vendidos em diferentes configurações de filtros.

2.4.2 Pentaho BI platform

O Pentaho é um software desenvolvido em Java utilizado na área de *Business Intelligence* (BI) que oferece, principalmente, soluções nas áreas de integração de dados (ETL - *Extract Transform Load*), relatórios, análises online (OLAP – *Online Analytical Processing*) e mineração de dados. Ele se tornou o preferido na organização do *Big Data* por ser *open source* (código aberto) e por se integrar facilmente a qualquer infraestrutura de TI. Ainda, os relatórios, vistas de análises e *dashboards* gerados pelo Pentaho são customizáveis e intuitivos.

Além de trazer, organizar e apresentar os dados de acordo com as necessidades da empresa, o Pentaho é focado nos processos de trabalho da organização e traz soluções para a empresa. Ele implementa esses processos na resolução de problemas detectados, além de apresentar uma visão geral do funcionamento da empresa.

Vantagens do Pentaho

- *Open Source*: O Pentaho pode ser modificado de acordo com as necessidades do cliente. Apesar de ainda ser necessário o investimento em implantação, treinamento e suporte, o código aberto é garantia de baixo custo acessível;
- Qualidade, segurança e independência de fornecedores;
- Definição de indicadores adequados às necessidades da empresa;
- Fácil integração a outras infraestruturas de TI;
- Relatórios e *dashboards* intuitivos.

A seguir são apresentados os fatos realizados na configuração, implementação e testes da plataforma Pentaho com o intuito de disponibilizar o conceito BI.

2.4.2.1 Banco de dados

Para criar o banco de dados OLTP e o *Datawarehouse* foi utilizado SGBD (Sistema de Gerenciamento de Banco de Dados) *Postgres*, na imagem 28 é possível visualizar os dois bancos criados e as dimensões do *Datawarehouse*.

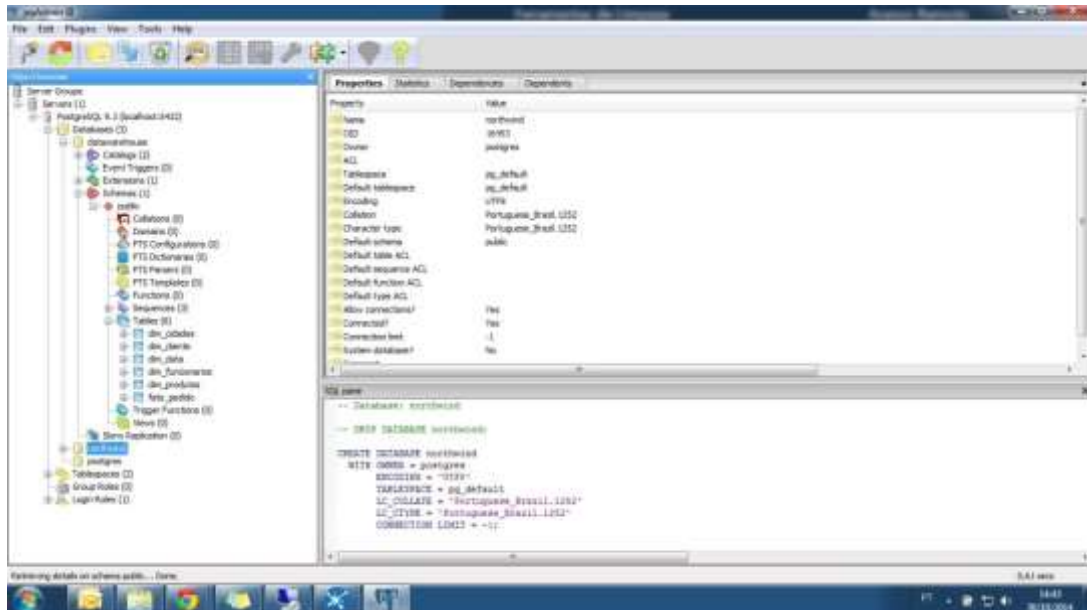


Figura 28 - Banco de dados *Postgres*

2.4.2.2 Pentaho data integration (PDI)

Para que seja iniciada a criação do ambiente Pentaho é necessário realizar a configuração com o PDI onde será feita a extração, transformação e carga (ETL) de dados.

Como é possível visualizar na figura 29 a ferramenta é utilizada para realizar as leituras dos dados do banco nortwind (OLTP), a criação das dimensões do cubo OLAP e posteriormente os *selects* para a criação da tabela fato.

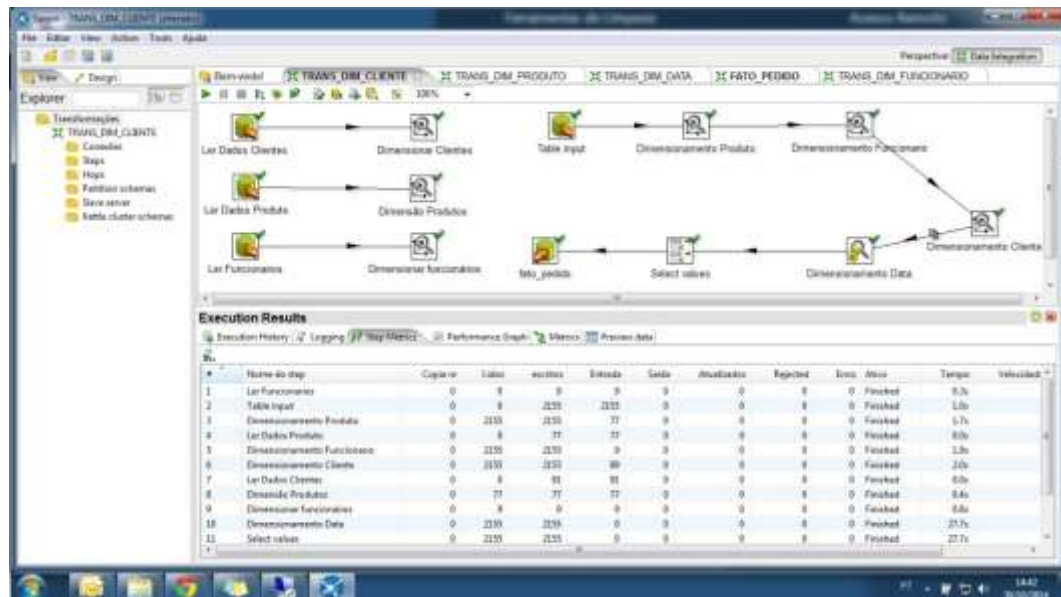


Figura 29 - Pentaho Data Integration

2.4.2.3 Pentaho schema workbench (PSW)

O Schema Workbench Mondrian é uma interface do designer, que permite criar e testar visualmente o esquema do cubo OLAP. O motor de Mondrian processa os pedidos MDX com o esquema ROLAP (*Relational OLAP*). Estes arquivos de esquema são modelos de metadados XML e são criados em uma estrutura específica utilizada pelo motor Mondrian.

Figura 30 ilustra o esquema de criação do cubo OLAP inserindo suas dimensões e métricas.

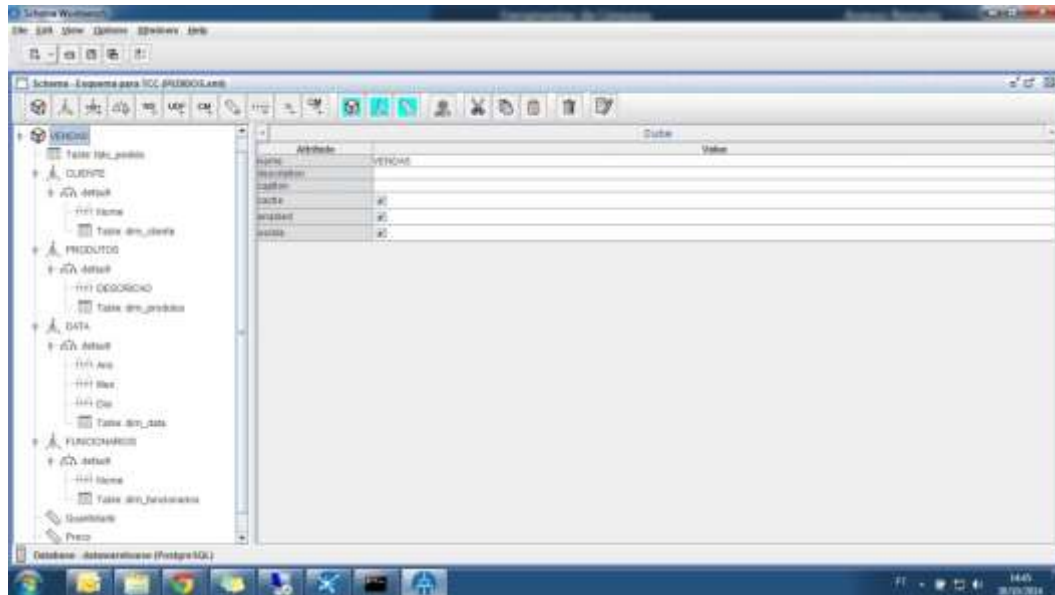


Figura 30 - Pentaho Schema Workbench

2.4.2.4 Console Pentaho

Figura 31 mostra o console via web para acessar a interface do Pentaho



Figura 31 - Console Pentaho

2.4.2.5 Visualização pivot analysis

Ao acessar o console Pentaho e o JPivot deverá ser dada a opção de visualização do cubo já montado podendo assim inserir as métricas e as dimensões a serem vistas pelo usuário final, visualizando o resultado final

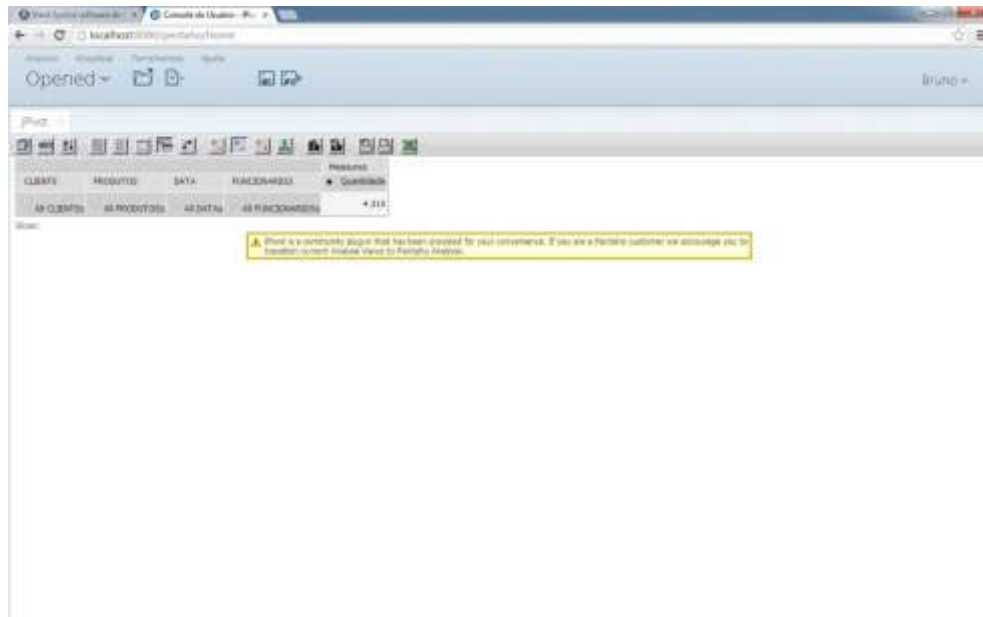


Figura 32 - Visualização de cubo OLAP

2.4.2.6 Analysis pivot – cubos de análise de dados

Os cubos podem ser manipulados pelo usuário de acordo com a necessidade, conforme pode ser visto na figura 32 sendo possível realizar a análise do cubo OLAP, além da visualização do resultado do cubo, já contemplando as métricas e dimensões inseridas pelo usuário que irá visualizar os dados e gráficos.

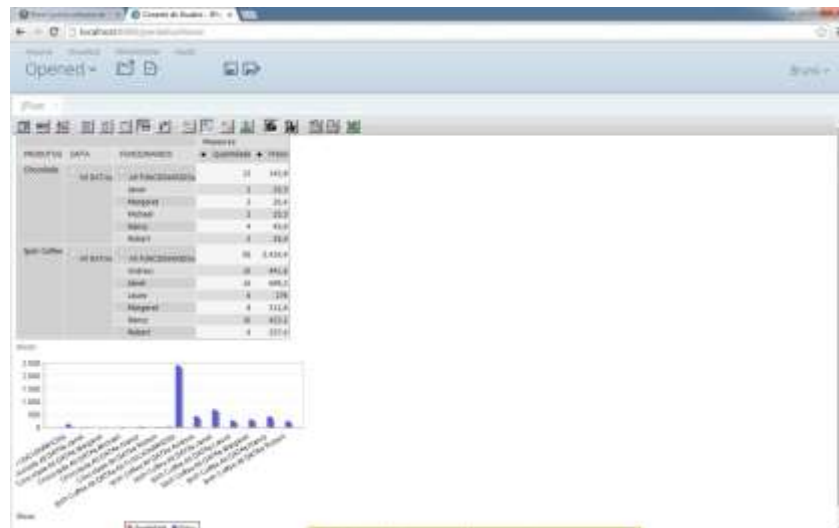


Figura 33 - Cubos de análise de dados

2.4.2.7 Saiku analytics

O Saiku Analytics é um cliente web disponível como *plug-in* para o Pentaho BI Server. Ele usa o Pentaho Analysis (Mondrian) para proporcionar uma forma fácil de usar o recurso de Cubos OLAP com uma experiência simples para usuário final.

A fim de validar e gerar um relatório de fácil visualização e criação podem ser realizado testes com o cubo Vendas, gerando relatório gráfico com as mesmas dimensões e métricas utilizadas no console principal do Pentaho.

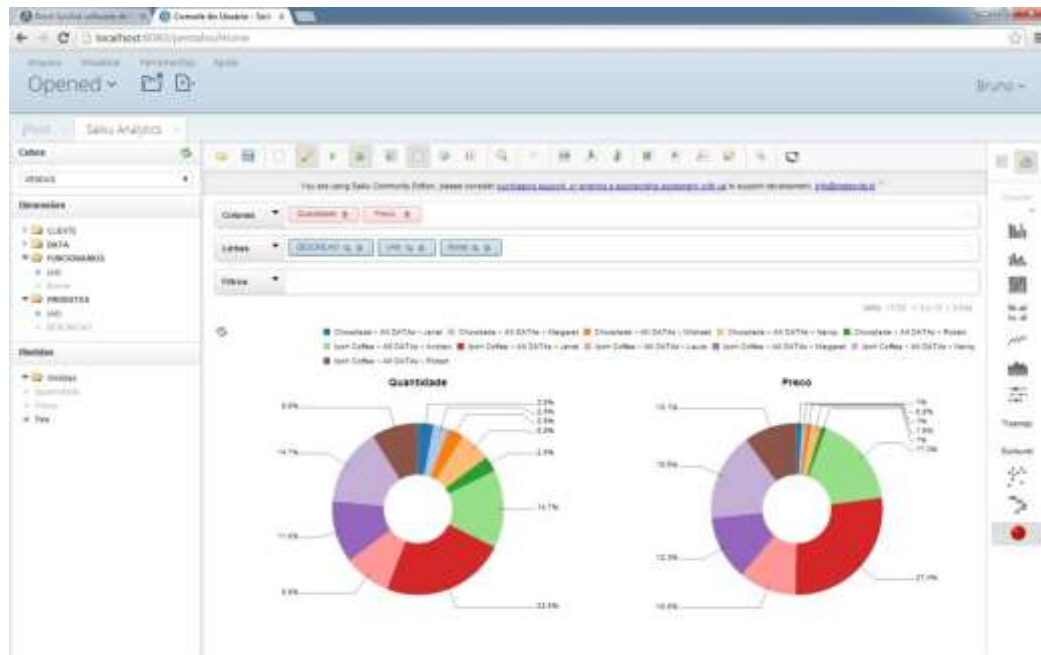


Figura 34 - Saiku Analytics

2.5 Resultados Obtidos

Os resultados obtidos nos estudos das ferramentas Pentaho e QlikView foram demonstrados de forma prática os conceitos de *Business Intelligence* e de suas etapas estudadas no trabalho provando que as duas ferramentas seguem o conceito de BI. O QlikView por sua vez não possui uma ferramenta própria para a criação do *Data Warehouse* sendo necessário ao usuário criá-lo direto no banco de dados, mesmo assim ainda é possível seguir as demais etapas uma vez que feito isso.

Enquanto o Pentaho pode ser utilizado gratuitamente e posteriormente adquirir bibliotecas pagas, para utilizar o QlikView é necessário comprar o servidor QlikView Server, para que este possa ser acessado pelo navegador. Em contrapartida o QlikView é mais simples e amigável ao usuário podendo criar relatórios de forma intuitiva.

As duas ferramentas podem ser utilizadas por empresas de pequeno e médio porte e poderão trazer todos os benefícios de BI citados neste trabalho.

3 CONCLUSÃO

Ao final deste trabalho, foi possível entender melhor as características de cada solução desenvolvida com as duas ferramentas.

O Pentaho, apesar de ser uma plataforma *open-source*, possui diversas funcionalidades a serem compradas, podendo ser utilizado com ou sem as bibliotecas com pagas, sendo mais atrativo para pequenas e médias empresas podendo optar pela opção sem custo. Ainda analisando a ferramenta Pentaho pode-se notar que ela possui diversas funcionalidades e módulos que permitem uma customização desde a mineração dos dados, a extração, transformação e carregamento das informações de uma forma interativa para pessoas com um entendimento na área da computação, uma montagem dinâmica dos cubos obtendo assim diversas dimensões disponibilizando melhor análise para o gerente de negócios podendo visualizar na própria interface do programa, *dashboards* e também exportando para outros tipos de extensões como planilhas do Excel.

Pode-se concluir quanto a ferramenta de BI QlikView é uma ótima ferramenta para pequenas e médias empresas. Sua funcionalidade em in-memory faz com que a informação seja carregada rapidamente e a interação que o usuário tem com os relatórios impressiona qualquer gerente ou executivo. O QlikView também possui pontos negativos, um deles é que apesar de toda a plataforma ser gratuita, o servidor para armazenar o projeto e disponibilizar o mesmo para acesso ao usuário final é pago, também pode-se citar que como toda a informação é carregada em memória, se a quantidade de informação for de muitas tabelas com muitas linhas cada, o servidor no qual o QlikView estiver instalado deverá ter muita memória RAM ou obterá um erro de estouro de memória (*Out of memory*) limitando o QlikView à necessidade de um computador com bom hardware para o acesso à um grande volume de dados.

A aplicação das métricas de vendas possibilitou um estudo objetivo das duas plataformas de BI escolhidas, resultando em uma aplicação do conceito adquirido no trabalho às plataformas. Conclui-se que independente da plataforma escolhida, sendo utilizados os conceitos de e as etapas de BI apresentadas, foi possível criar soluções para empresas de pequeno e médio porte.

REFERÊNCIAS

Alcântara, S. O. (07 de Junho de 2010). *Fatec Zona Leste*. Acesso em 25 de Abril de 2014, disponível em Site da Fatec Zona Leste: <http://fateczl.edu.br/TCC/2010-1/TCC-012.pdf>

<http://revistas.utfpr.edu.br/pb/index.php/CAP/article/viewFile/933/544>

ALTER, Steven. *Information systems: a management perspective*. 3ª edition EUA: Addison-Wesley Educational Publishers Inc, 1996.

ANGELONI, Maria T.; REIS, Eduardo S. *Business Intelligence como Tecnologia de Suporte à Definição de estratégias para melhoria da qualidade do ensino*. In: Encontro da ANPAD, 2006, Salvador. XXX Encontro Nacional de Pós-Graduação em Administração, 2006, 2006. v.1.p. 16 paginas

BARBIERI, Carlos. *BI- Business Intelligence: modelagem & tecnologia*. Rio de Janeiro: Axel Books do Brasil, 2001.

BATISTA E. O. *Sistemas de informação*. São Paulo: Saraiva, 2004.

BATTISTI, Julio. O modelo relacional de dados. Disponível em: <http://www.juliobattisti.com.br/artigos/office/modelorelacional_p1.asp> Acesso em: 21 de Maio de 2014.

Noronha, Vanessa (2013) - *Business Intelligence* - Conjunto de softwares que ajudam em decisões estratégicas

GATES, Bill. *A empresa na velocidade do pensamento: com um sistema nervoso digital*. Tradução Pedro Maia Soares. São Paulo: Companhia de Letras, 1999.

It4biz, (2007). It4biz IT Solutions. Disponível em Site da It4biz IT Solutions: <http://www.it4biz.com.br/novosite/comunidade/open-source-business-intelligence/o-que-e-business-intelligence/> Acesso em 25 de Abril de 2014.

LEME FILHO, Trajano. *Business intelligence no microsoft Excel*. Rio de Janeiro: Axel

Books do Brasil, 2004.

PAVANI, Renato (2006). Business Intelligence: é estratégia ou tecnologia? Disponível em: <http://webinsider.com.br/2006/01/23/business-intelligence-e-estrategia-ou-tecnologia/>
Acesso em: 18 de Maio de 2014.

Primak, (2008). Decisões com BI. Disponível no e-book em: http://books.google.com.br/books/about/Decis%C3%B5es_com_BI_Business_Intelligence.html?id=cGBneHPeLvkC

Revista TECAP – Número 03 – Ano 3 – Volume 3 – 2009. Disponível em revistas.utfpr.edu.br/pb/index.php/CAP/article/download/933/544

ROMNEY, Marshall B; STEINBART, Paul John. Accounting Information Systems. 8ª ed. New Jersey. Prentice-Hall, Inc. 2000.

SERRA, Laércio, (2002). Livro: A essência do Business Intelligence. São Paulo: Berkeley, 2002

Sun Tzu, 2006. Livro: *Art of War* (Arte da Guerra).

Teban, 2012. Analise de OLAP Multidimensional (*Analisis Multidimensional OLAP*) Disponível em (<http://analisismultidimensional6.blogspot.com.br/2012/03/que-son-los-cubos-olap.html>)

<http://www.jornallivre.com.br/19102/o-que-e-business-intelligence.html>

<http://www.microsoft.com/en-us/server-cloud/solutions/business-intelligence/overview.aspx#fbid=5f8XeVF5pX8>

<http://www.oracle.com/br/solutions/midsized/oracle-products/business-intelligence/index.html>

<http://www.pentaho.com/product/data-integration>

<http://vivianeribeiro1.wordpress.com/2011/07/12/o-que-e-olap/>

<http://technet.microsoft.com/pt-br/library/cc668463.aspx>

[http://technet.microsoft.com/pt-br/library/ms174587\(v=sql.105\).aspx](http://technet.microsoft.com/pt-br/library/ms174587(v=sql.105).aspx)

<http://analisismultidimensional6.blogspot.com.br/2012/03/que-son-los-cubos-olap.html>

http://www.oficinadanet.com.br/artigo/business_intelligence/business_intelligence_e_estrategia_ou_tecnologia

<http://www.oficinadanet.com.br/post/13153-o-que-e-business-intelligence>

<http://imasters.com.br/artigo/3836/gerencia-de-ti/modelo-dimENSIONAL-para-data-warehouse/>

<http://www.knowsolution.com.br/index.php/pentaho/o-que-e-pentaho#ixzz3lBiurs1W>

<http://www.ambientelivre.com.br/pentaho.html>