

HUGO MORAIS MENDES

**USO DE FERRAMENTAS OPEN SOURCE PARA
SOLUÇÕES DE BUSINESS INTELLIGENCE: ESTUDO
DE CASO COM A SUÍTE PENTAHO**

BACHARELADO
EM
CIÊNCIA DA COMPUTAÇÃO

FIC – MINAS GERAIS

2014

HUGO MORAIS MENDES

**USO DE FERRAMENTAS OPEN SOURCE PARA
SOLUÇÕES DE BUSINESS INTELLIGENCE: ESTUDO
DE CASO COM A SUÍTE PENTAHO**

Monografia apresentada à banca examinadora da Faculdade de Ciência da Computação das Faculdades Integradas de Caratinga como exigência parcial para obtenção do grau de bacharel em Ciência da Computação, sob orientação da professora Msc. Fabrícia Pires Souza Tiola.

FIC – CARATINGA

2014

HUGO MORAIS MENDES

**USO DE FERRAMENTAS OPEN SOURCE PARA
SOLUÇÕES DE BUSINESS INTELLIGENCE: ESTUDO
DE CASO COM A SUÍTE PENTAHO**

Monografia submetida à Comissão examinadora designada pelo Curso de Graduação em Ciência da Computação como requisito para obtenção do grau de Bacharel.

Prof. Msc. Fabrícia Pires Souza Tiola
Faculdades Integradas de Caratinga

Prof. Msc. Glauber Luis Costa
Faculdades Integradas de Caratinga

Prof. Maicon Vinicius Ribeiro
Faculdades Integradas de Caratinga

Caratinga, 12/2014

AGRADECIMENTOS

Primeiramente agradeço a Deus por tudo o que tem me dado. Agradeço a todos da empresa DPC Distribuidor Atacadista S/A que me ajudaram direta ou indiretamente para que eu chegasse até aqui.

A minha namorada Stefane pelo companheirismo, aos meus pais que fizeram tudo que estive ao alcance deles, aos meus amigos que não me abandonaram mesmo quando não pude estar presente, aos meus professores que me deram valiosos ensinamentos.

Em especial a Fabrícia Pires, Paulo Eustáquio, Glauber Costa, Maicon Ribeiro, Joabe Machado e Edson Ferreira pelas suas contribuições para que esse trabalho se realizasse.

Obrigado a todos.

“Confie no Senhor de todo o coração e não se apóie em seu próprio entendimento.”

Provérbios 3.5

RESUMO

Business Intelligence é uma ferramenta que utiliza um conjunto de tecnologias, que são utilizadas para levar informações de alta qualidade aos tomadores de decisões nas empresas, para que eles possam transformar essas informações em decisões importantes. Com essas informações os tomadores de decisões podem gerar relatórios para fazer comparativos, criar estratégias de mercado ou encontrar setores oportunos, levando assim a uma maior lucratividade, controle de gastos e clientes mais satisfeitos. O processo de levar informações aos tomadores de decisões deve-se iniciar identificando quais informações são necessárias, seguindo pela extração, tratamento e por fim armazenando em outra base de dados conhecida como *Data Warehouse (DW)*. O DW é projetado para armazenar dados analíticos de forma consolidada, tendo em seu conteúdo informações históricas da empresa, devendo responder a solicitação do usuário mais rápido possível. Compreendendo que a informação é importante para as empresas, este trabalho tem por finalidade apresentar uma solução *Business Intelligence* para a empresa DPC Distribuidor Atacadista S/A usando a suíte *Pentaho*. Para avaliar a ferramenta implantada foi feito um estudo de caso através de um questionário avaliando a usabilidade da ferramenta com o intuito de descobrir se ela foi bem aceita na empresa. Os resultados obtidos pelo questionário foram satisfatórios, de forma geral os participantes aceitaram a suíte *Pentaho*, passando a ser usado pela maioria todos os dias nos processos decisórios.

Palavras-chave: *Business Intelligence, Data Warehouse, Pentaho.*

ABSTRACT

Business Intelligence is a tool that uses a set of technologies, which are used to bring high quality information to decision makers in companies, so they can transform that information into important decisions. With this information, decision makers can generate reports to create comparative market strategies or find appropriate sectors, leading to greater profitability, cost control and more satisfied customers. The process of bringing information to the decision makers should begin by identifying what information is required, following the extraction, treatment and ultimately storing in another database known as Data Warehouse (DW). The DW is designed to store analysis data in a consolidated manner, having in its content the company's historical information, and respond to user's request as soon as possible. Understanding that the information is important for companies, this work aims to present a Business Intelligence solution for the company's DPC Distribuidor Atacadista using the Pentaho suite. To evaluate the tool deployed was made a case study through a questionnaire evaluating the usability of the tool in order to find out if she was well accepted in the company. The results obtained by the questionnaire were satisfactory, generally the participants accepted the Pentaho suite to be used by most every day in the decision-making processes.

Keywords: *Business Intelligence, Data Warehouse, Pentaho.*

LISTA DE ILUSTRAÇÕES

Figura 1 – Funcionalidade da ferramenta OLAP sobre o cubo.	21
Figura 2 – Operação <i>Drill Down</i> na dimensão localização geográfica.	22
Figura 3 – Técnica de <i>Drill Up</i> na dimensão de tempo.	23
Figura 4 – Integração de dados no <i>Data Warehouse</i>	25
Figura 5 – Principais componentes de um <i>Data Warehouse</i>	28
Figura 6 – Representação de um <i>Data Mart</i>	30
Figura 7 – Funcionamento de uma <i>Staging Area</i>	32
Figura 8 – Cubo de dados e suas dimensões.	33
Figura 9 – Modelo de dados <i>Star Scheme</i>	35
Figura 10 – Modelagem <i>Snowflake Scheme</i>	36
Figura 11 – Modelagem de dados do <i>Data Warehouse</i>	48
Figura 12 – Modelagem de dados do sistema de origem.	50
Figura 13 – Modelagem de dados das tabelas de destino.	51
Figura 14 – Transformação de dados no PDI.	53
Figura 15 - Cubo de informação no PSW.	54
Figura 16 - Relatório sendo criado pelo PRD.	55
Figura 17 - <i>Dashboard</i> de comparativo de vendas.	56
Gráfico 1 – Questão 1.	58
Gráfico 2 – Questão 2.	59
Gráfico 3 – Questão 3.	60
Gráfico 4 - Questão 4.	61
Gráfico 5 – Questão 5.	62
Gráfico 6 – Questão 6.	63
Gráfico 7 – Questão 7.	64
Gráfico 8 – Questão 8.	65
Gráfico 9 – Questão 9.	66
Gráfico 10 – Questão 10.	67
Gráfico 11 – Questão 11.	68
Gráfico 12 – Questão 12.	69
Gráfico 13 – Questão 13.	70

Gráfico 14 – Questão 14.	71
Gráfico 15 – Questão 15.	72
Gráfico 16 – Questão 16.	73
Gráfico 17 – Questão 17.	74
Gráfico 18 – Questão 18.	75
Gráfico 19 – Questão 19.	76
Gráfico 20 – Questão 20.	77
Gráfico 21 – Questão 21.	78
Gráfico 22 – Questão 22.	79
Gráfico 23 – Questão 23.	80
Gráfico 24 – Questão 24.	81
Gráfico 25 – Questão 25.	82
Gráfico 26 – Questão 26.	83
Gráfico 27 – Questão 27.	84
Gráfico 28 – Questão 28.	85

LISTA DE TABELAS

Tabela 1 – Respostas da Questão 1.....	57
Tabela 2 – Resultado da Questão 2.....	58
Tabela 3 – Resultado da Questão 3.....	59
Tabela 4 – Resultado da Questão 4.....	60
Tabela 5 – Resultado da Questão 5.....	61
Tabela 6 – Resultado da Questão 6.....	62
Tabela 7 – Resultado da Questão 7.....	63
Tabela 8 – Resultado da Questão 8.....	64
Tabela 9 – Resultado da questão 9.....	65
Tabela 10 – Resultado da questão 10.....	67
Tabela 11– Resultado da Questão 11.....	67
Tabela 12 – Respostas da Questão 12.....	68
Tabela 13 – Respostas da Questão 13.....	69
Tabela 14 – Respostas da Questão 14.....	70
Tabela 15 – Respostas da Questão 15.....	71
Tabela 16 – Respostas da Questão 16.....	72
Tabela 17 – Respostas da Questão 17.....	73
Tabela 18 – Respostas da Questão 18.....	74
Tabela 19 – Respostas da Questão 19.....	75
Tabela 20 – Respostas da Questão 20.....	76
Tabela 21 – Respostas da Questão 21.....	77
Tabela 22 – Respostas da Questão 22.....	78
Tabela 23 – Respostas da Questão 23.....	79
Tabela 24 – Respostas da Questão 24.....	80
Tabela 25 – Respostas da Questão 25.....	81
Tabela 26 – Respostas da Questão 26.....	82
Tabela 27 – Respostas da Questão 27.....	83
Tabela 28 – Respostas da Questão 28.....	84

LISTA DE SIGLAS

BI – *Business Intelligence*

DM – *Data Mart*

DOLAP – *Desktop On Line Analytical Processing*

DW – *Data Warehouse*

EIS – *Executive Information System*

ER - Entidade Relacionamento

ES – *Engenharia de Software*

ETL – *Extract, Transform, Load*

HOLAP – *Hybrid On Line Analytical Processing*

JDBC – *Java Database Connectivity*

JVM – *Java Virtual Machine*

KETTLE – *Kettle, Extraction, Transport, Transformation and Loading Environment*

MDX – *Multi-Dimensional Expressions*

MOLAP – *Multidimensional On Line Analytical Processing*

ODBC – *Open Database Connectivity*

OLAP – *On-line Analytical Processing*

PDI – *Pentaho Data Integration*

PRD – *Pentaho Report Designer*

PSW – *Pentaho Schema Workbench*

ROLAP – *Relational On Line Analytical Processing*

SQL – *Structured Query Language*

TI – *Tecnologia da Informação*

SUMÁRIO

INTRODUÇÃO.....	14
1 REFERENCIAL TEÓRICO.....	16
1.1 BUSINESS INTELLIGENCE	16
1.1.1 CONCEITO	16
1.1.2 HISTÓRICO	18
1.1.3 ARQUITETURA E COMPONENTES.....	19
1.1.4 PROCESSAMENTO ANALÍTICO ONLINE (OLAP).....	20
1.1.5 ETL - EXTRACT TRANSFORM LOAD	23
1.1.6 BENEFÍCIOS DO BI PARA AS EMPRESAS	25
1.1.5 PRESENTE E FUTURO	26
1.1.6 SOLUÇÕES DISPONÍVEIS	27
1.2 DATA WAREHOUSE	27
1.2.1 DEFINIÇÕES E CONCEITOS DE DW	27
1.2.2 CARACTERÍSTICAS DA TECNOLOGIA DE <i>DATA WAREHOUSING</i>	29
1.2.3 <i>DATA MART</i> - DM.....	30
1.2.4 <i>OPERACIONAL DATA STORES</i> - ODS.....	31
1.2.5 GRANULARIDADE	32
1.2.6 CUBOS.....	33
1.2.7 MODELAGEM MULTIDIMENSIONAL.....	34
1.3 FERRAMENTA DE BUSINESS INTELLIGENCE	37
1.3.1 PENTAHO.....	37
1.3.2 COMPONENTES	38
1.4 ENGENHARIA DE <i>SOFTWARE</i>	39
1.4.1 QUALIDADE <i>SOFTWARE</i>	40
1.4.2 USABILIDADE.....	41
1.4.3 TESTE DE USABILIDADE	43
2 METODOLOGIA.....	44
2.1 PÚBLICO ALVO DO QUESTIONÁRIO.....	44
2.2 ELABORAÇÃO DO QUESTIONÁRIO.....	44
2.3 COLETA DOS DADOS	45
2.4 TRATAMENTO DOS DADOS.....	45
2.5 AMBIENTE DE ESTUDO.....	45

2.6 IMPLANTAÇÃO.....	46
2.6.1 TER APOIO DA GESTÃO.....	47
2.6.2 LEVANTAMENTO DAS NECESSIDADES.....	48
2.6.3 MAPEAMENTO DOS SISTEMAS E DOS DADOS.....	49
2.6.4 DEFINIR UTILIZADORES.....	51
2.6.5 GARANTIR A INTEGRAÇÃO E QUALIDADE DOS DADOS.....	52
2.6.6 IMPLANTAÇÃO.....	52
3 ANÁLISE DOS RESULTADOS.....	57
3.1 ANÁLISES DAS QUESTÕES DA PESQUISA.....	57
3.1.1 PERGUNTAS PARA CONHECIMENTO DOS PARTICIPANTES.....	57
3.1.2 QUESTÕES DE ACESSO AO <i>PENTAHO</i>	60
3.1.3 QUESTÕES DE AVALIAÇÃO DA INTERFACE DO <i>PENTAHO</i>	62
3.1.4 QUESTÕES DE AVALIAÇÃO DOS RELATÓRIOS DO <i>PENTAHO</i>	66
3.1.5 QUESTÕES DE AVALIAÇÃO DOS <i>DASHBOARDS</i> DO <i>PENTAHO</i>	72
3.1.6 QUESTÕES DE AVALIAÇÃO DA ANÁLISE DE INFORMAÇÕES DO <i>PENTAHO</i>	74
3.1.7 QUESTÕES DE AVALIAÇÃO GERAL DO <i>PENTAHO</i>	78
4 CONCLUSÃO.....	86
5 TRABALHOS FUTUROS.....	87
6 REFERÊNCIAS.....	88
7 ANEXOS.....	91
7.1 ANEXO I: QUESTIONÁRIO.....	91
7.2 ANEXO II: TERMO DE AUTORIZAÇÃO.....	94
7.3 ANEXO III: <i>VIEWS</i>	95

INTRODUÇÃO

A informação tem um papel importante nas empresas, a produção dela tem aumentado gradativamente nos últimos anos, acarretando em dificuldades para a manipulação da mesma com planilhas inteligentes. As empresas trabalham com as informações visando o ambiente empresarial, a fim de melhorar os processos internos e descobrir novos mercados. Junto com o aumento das informações, os sistemas de informação foram evoluindo e passaram a fornecer informações empresariais a partir de uma base de dados. Naturalmente a tecnologia continuou a evoluir surgindo então o termo inteligência de negócio ou *Business Intelligence* (BI) que são sistemas de informação que auxiliam nas mais diversas situações do processo de tomada de decisão.

Com as mudanças repentinas do mercado, as empresas têm de se adaptar as mudanças, quanto mais tempo elas demoram em tomar as decisões relacionadas ao mercado, há uma chance maior de levar a empresa a fracassar. Todo o processo do BI baseia-se em transformar os dados em informações, para que essas informações tornem-se conhecimento, e por fim tornam-se ações, podendo aumentar o sucesso da empresa no ambiente de negócios. Quanto mais rápida estas informações forem obtidas, mais rapidamente os tomadores de decisão poderão trabalhar os processos e tomar decisões com base em dados estatísticos de forma interativa.

A maioria dos sistemas de BI trabalha com banco de dados que utiliza os conceitos de *Data Warehouse* e armazena dados analíticos, desenvolvido especialmente para atender as necessidades de informação da alta gerência no processo de tomada de decisão.

Os tomadores de decisão na empresa DPC são os diretores, gerentes, coordenadores de venda da área comercial da empresa.

O respectivo trabalho tem como objeto de pesquisa a área de *Business Intelligence* sendo que a suíte *Pentaho* e o conceito de *Data Warehouse* são necessários para concretizar o trabalho. O objetivo geral desse trabalho tem por finalidade analisar se a suíte *Pentaho* foi bem aceita pelos tomadores de decisões que trabalham na área comercial da empresa DPC Distribuidor Atacadista, visando que as informações tornem-se conhecimentos e por fim virem ações. Para atingir o objetivo geral, os seguintes objetivos específicos foram atingidos:

- Estudar sobre as melhores práticas de *Business Intelligence*.
- Conhecer sobre a suíte *Pentaho*.

- Coletar e modelar as informações que serão importantes para a área comercial da empresa DPC.
- Implantar a suíte *Pentaho*.
- Analisar a grau de satisfação dos tomadores de decisão da área comercial da empresa DPC.

Quando o BI é implantado adequadamente, traz muitos benefícios sendo os principais: economia de tempo, economia de custo, melhores estratégias, melhores planos, relatórios rápidos e precisos. Logo a empresa tende a possuir melhor desempenho tornando-se mais competitiva.

Para analisar o grau de satisfação do usuário com a ferramenta de *Business Intelligence* implantada, foi aplicado um teste de usabilidade nos tomadores de decisões para comprovar a aceitação deles com o *Pentaho*.

O presente trabalho está organizado em três capítulos. O primeiro capítulo faz uma introdução ao tema do trabalho com uma fundamentação teórica sobre os assuntos *Business Intelligence*, *Data Warehouse* e Suíte *Pentaho*. O segundo capítulo apresenta o processo de desenvolvimento do trabalho, abordando o ambiente empresarial. O terceiro capítulo descreve os resultados do questionário utilizado para avaliar o grau de satisfação do usuário.

1 REFERENCIAL TEÓRICO

1.1 BUSINESS INTELLIGENCE

Nesta sessão será apresentado como surgiu o termo *Business Intelligence*, como ele é aplicado na empresa, o conceito e os principais componentes que compõem a arquitetura.

1.1.1 Conceito

Nas organizações existem pessoas que tomam decisões, as quais podem afetar toda a organização, alguns departamentos ou apenas algumas pessoas. Para as organizações essas decisões fazem parte do cotidiano (SCHEPS, 2008).

Porém, ainda existem pessoas que tomam decisões do modo antigo, com base em algumas informações atuais, algumas das melhores do passado e associa com opiniões de outras pessoas além do seu próprio intuito, e então avaliam qual caminho seguir para conseguir o melhor resultado possível (SCHEPS, 2008).

Um novo conceito de organização de informação é formado com o *Business Intelligence*, através de suas tecnologias ele gera um maior valor ao negócio (BARBIERI, 2001).

Em momentos críticos, tomar boas decisões pode tornar a organização bem mais sucedida a curto e a longo prazo, pois, podem trazer operações mais eficientes, deixar clientes satisfeitos ou deixar a organização mais rentável. Embora as organizações sejam limitadas de alguns recursos para realizarem suas missões, suas escolhas podem ser ilimitadas (SCHEPS, 2008).

A tradução do termo *Business Intelligence* significa Inteligência de Negócio. Atualmente existem muitas definições para este termo, a cada nova pesquisa pode se achar uma nova definição, pois cada pessoa a define tendenciando para seu interesse em particular, mas independentemente de qual seja a sua definição, os elementos que o constituem sempre são os mesmos, que são tecnologias e ferramentas para apoiar a tomada de decisão (SALLES, 2013).

Para Barbieri (2001), os processos de *Business Intelligence* trabalham com tecnologias para coletar, armazenar, analisar e disponibilizar os dados, afim de que os gestores tomem decisões transformando os dados em informações mais precisas e em conhecimentos.

O *Business Intelligence* (BI) auxilia as organizações nas tomadas de decisões, as quais podem ser importantes ou modestas, pois, o BI é um recurso flexível que pode trabalhar em vários níveis da organização. As decisões tomadas utilizando este recurso podem fazer toda a diferença na organização (SCHEPS, 2008).

O autor Scheeps (2008, p.11) define BI como: “Em essência, o BI é qualquer atividade, ferramenta ou processo usado para obter a melhor informação para apoiar o processo de tomada de decisões”.

Segundo Barbieri (2001, p.34) “O conceito de BI – Business Intelligence, de forma mais ampla, pode ser entendido como a utilização de variadas formas de informação para se definir estratégias de competitividade nos negócios da empresa”.

Para entender melhor a definição de BI, pode ser imaginado um cenário com um gerente prestes a tomar decisões, com ele há muitos consultores, ou muitos computadores processando dados da empresa, se esses dados derem uma visão ao gerente sobre a atual situação da empresa e prover percepções de como agir no futuro, então é BI (SCHEPS, 2008).

O BI é definido como um conjunto de conceitos, métodos e processos que visam melhorar a tomada de decisão nas empresas e também tomar como base a realização de estratégias, por isso o BI pode ser considerado a chave para o gerenciamento de uma organização (OLSZAK, 2003).

O BI é composto por diversos componentes tecnológicos. Por essa diversidade de componentes ele tem significados diferentes para cada pessoa, o seu principal objetivo é trabalhar com os dados, proporcionando-os aos tomadores de decisões após a sua análise, conseguem-se os *Insights* que podem servir como base nas decisões (TURBAN, 2009).

Os *Insights* são visões ou esclarecimentos que as pessoas têm sobre negócios, podendo ser algum fato ou hipótese sobre algum aspecto na organização que antes estava oculto ou era desconhecido e de repente tornou-se visível. Os *Insights* podem ser precisos, valiosos, oportunos e acionáveis (SCHEPS, 2008).

- Preciso: Devem refletir a realidade objetiva da organização aderindo aos rígidos padrões de veracidade, executivos se basearam nas informações emitidas pelos relatórios (SCHEPS, 2008);
- Valioso: Quando os dados analisados são cruzados, podem produzir informações que causem um grande impacto podendo gerar grandes retornos (SCHEPS, 2008);
- Oportuno: Surgem nos momentos em que a informação é mais necessária e não depois que o fato tenha ocorrido (SCHEPS, 2008);

- Acionável: Deve-se tirar conclusões sobre fatos viáveis, que podem levar a ações (SCHEPS, 2008).

Em BI os *Insights* muitas vezes serão surpreendentes e intuitivos, portanto devem apresentar as informações mais próximo possíveis da verdade, não só para produzir bons resultados, mas para evitar uma má reputação com os céticos. Nas organizações existem pessoas inteligentes trabalhando para que possam ligar os pontos óbvios, mas os *Insights* nem sempre são óbvios, mas seu impacto pode ser enorme (SCHEPS, 2008).

Para Turban (2009, p. 27) “O processo do BI, baseia-se na transformação do dado em informação, depois em decisões e finalmente em ações”.

O processo de BI tem como base os dados, que após a sua coleta deve organizá-los de forma que o torne fácil de analisar, e depois desta análise, esses dados se transformam em informações, para que os tomadores de decisões tomem ações importantes nas organizações (SOUZA, 2010).

1.1.2 Histórico

Na década de 1990 surgiu o termo BI, criado pela empresa Gartner Group. Entretanto no início de 1970 já havia uma definição para o termo, que surgiu através dos sistemas de geração de relatórios SIG – Sistemas de Informação Gerenciais. Os relatórios obtidos pelo sistema eram estáticos, bidimensionais, e não tinham suporte à análise (TURBAN, 2009).

Para o autor Primak (2008), o conceito prático de BI surgiu com o povo antigo, eles obtinham informações analisando o comportamento da maré, os períodos de chuva e seca entre outros. Depois cruzavam estas informações, usando os princípios básicos de BI, para trazer benefícios para as suas aldeias.

Os sistemas na década de 1970 atendiam apenas o nível operacional das organizações, contudo o avanço da Tecnologia da Informação trouxe no início dos anos 1980, o conceito de sistemas de informações executivas EIS, esses sistemas começaram a atender o nível tático e o nível estratégico. Com a evolução da tecnologia os sistemas de geração de relatórios começaram a ser dinâmicos e multidimensionais, prover previsões, análises de tendências, possibilidade de maior detalhamento, acesso a status e fatores críticos de sucesso (TURBAN, 2009).

Nos anos de 1980, surgiu à administração dos dados, devido a eles atingirem uma espécie de estrelato. Também surgiu nesta década a modelagem de dados modelo relacional e

a engenharia da informação, trocando a estrutura hierárquica pelas relações entre tabelas (BARBIERI, 2001).

Na década de 1990 muitos sistemas já tinham esses recursos, que atendiam os níveis hierárquicos das organizações, então logo surgiu o conceito de BI com poucos recursos a mais, atualmente é reconhecido que os executivos necessitam das informações que podem estar baseadas em sistema de BI. Dessa forma a definição de sistema de informação executiva tornou-se BI (TURBAN, 2009).

1.1.3 Arquitetura e Componentes

O BI é composto por quatro componentes, o primeiro é o *Data Warehouse* (DW) com seus dados que servem para fazer a análise de negócios, um conjunto de ferramentas que proporciona ao usuário manipulação e análise dos dados contidos no *Data Warehouse*; o *Business Performance Management* (BPM) para monitorar e analisar o desempenho da organização e uma interface de usuário (TURBAN, 2009).

1.1.3.1 *Data Warehouse*

O *Data Warehouse* pode ser definido como um banco ou repositório de dados estruturado, que deve estar sempre disponível. Nele há uma coleção de dados relevantes para a organização com o objetivo de dar suporte a aplicações de tomada de decisão. Sendo as suas características fundamentais: separados por assuntos, integração, variáveis com o tempo e não sofrem alterações. Este conjunto de dados produzidos são utilizados para a tomada de decisão gerencial. As aplicações de BI que usam o *Data Warehouse* nas organizações variam de pequenos relatórios a complexas otimizações. O DW é construído com as metodologias, metadados e o processo de carga de dados ETL (TURBAN, 2009; MACHADO, 2007).

1.1.3.2 Análise de Negócio

A análise de negócio é composta por diversas ferramentas de *software* e metodologias que consistem em trabalhar e fornecer acesso aos dados, com o objetivo de ajudar os usuários. Essas ferramentas são chamadas de processamento analítico *online* – OLAP, pois, permitem ao usuário fazer análises de dados do negócio com uma visão dimensional através de relatórios, com diversos ângulos de dados, consultas sob demanda e análise de dados,

permitindo assim comparar, identificar com rapidez e facilidade as tendências de desempenho da empresa (TURBAN, 2009; FORTULAN, 2005).

1.1.3.3 *Business Performance Management*

O *Business Performance Management* também conhecido pela sigla de BPM, é uma forma de direcionar a empresa para onde é desejado de forma rápida e precisa. Este componente baseia-se em definir, desenvolver e gerenciar a estratégia de negócio de uma organização, para isso ele usa a análise, a geração de relatórios e as consultas de BI, cujo principais objetivos são: aperfeiçoar o desempenho geral da organização e melhorar a capacidade de entendimento dos negócios. O BPM não crítica os fatos depois de terem ocorrido, ele exhibe antes de forma passivamente (TURBAN, 2009; PRIMAK, 2008).

1.1.3.4 Interface de usuário

São ferramentas para visualização de informação, que exibem as informações de forma clara e abrangente aos usuários. As ferramentas podem ser os *dashboards* que são painéis de informações que fornecem uma visão ampla do desempenho, tendências e exceções de desempenho, ferramentas de análise de dados, relatórios que podem ser criados pelo usuário e consultas *ad hoc* (TURBAN, 2009).

1.1.4 Processamento Analítico Online (OLAP)

O termo processamento analítico online se refere a um conjunto de ferramentas que possibilita aos usuários explorar os dados contidos em um DW através dos sistemas online. O OLAP e DW trabalham em conjunto uma completando a outra, o DW fica responsável por armazenar as informações de forma eficiente, e o OLAP deve acessar essas informações com rapidez (TURBAN, 2009; PRIMAK, 2008; MACHADO, 2010).

As ferramentas OLAP proporcionam aos sistemas de *Business Intelligence* visualizar os dados de uma nova forma (SCHEPS, 2008). As ferramentas do tipo OLAP têm as suas funcionalidades caracterizadas pela análise multidimensional dinâmica dos dados. Essas funcionalidades normalmente permitem uma série de visões, tais como geração e respostas de consultas, solicitação de relatórios e gráficos *ad hoc*, realização de análises estatísticas tradicionais ou modernas entre outros. Os relatórios *ad hoc* são consultas construídas

dinamicamente pelo usuário de acordo com as suas necessidades através de uma interface, utilizando parâmetros nunca antes utilizados, que o levem a encontrar o que procura, apoiando o usuário final nas suas atividades. Tudo isso ocorre de maneira simples, amigável e transparente, sendo necessário ter um conhecimento mínimo de informática (TURBAN, 2009; PRIMAK, 2008; MACHADO, 2010).

O OLAP está em uma categoria de software que permitem aos tomadores de decisões manipularem e obterem respostas dentro dos dados de forma multidimensional, onde os dados estão agrupados em várias dimensões tais como: produto, cliente, fornecedor, vendedor entre outros. As ferramentas OLAP permitem que os usuários consigam descobrir o motivo do resultado obtido através da ferramenta (PRIMAK, 2008).

Para Fortulan (2005) o OLAP possui algumas particularidades que o define, sendo eles: fazem análises de tendências e previsões; executa cálculos complexos; concede uma visão multidimensional dos dados; simulam cenários fictícios a partir de valores que são supostos.

Ainda segundo Anzanello (2008), o OLAP é utilizado para visualizar os dados agregados que estão no *Data Warehouse*, dessa forma ele atende as necessidades do *Business Intelligence*, pois a finalidade dessa ferramenta é apoiar a tomada de decisão.

O OLAP possibilita analisar as informações sob diferentes pontos de vista sempre que preciso. Essa técnica é chamada de *slice-and-dice*, ela permite ao analista navegar nas dimensões do cubo de informação, modificar a posição de uma informação, alterar linhas por colunas de forma a facilitar a compreensão pelo usuário (MACHADO, 2010).

A Figura 1 mostra a aplicação da técnica *slice-and-dice* da ferramenta OLAP em um cubo de informações, onde é reduzida a quantidade de informação exibida de acordo com o utilizador para facilitar a análise dos dados.

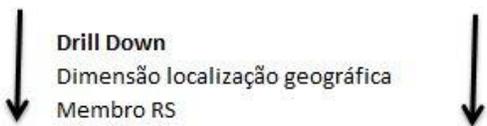


Figura 1 – Funcionalidade da ferramenta OLAP sobre o cubo.

O *Drill Down/Up* são técnicas utilizados nos sistemas OLAP para possibilitar a navegação nos dados, modificando o nível de granularidade da consulta, ou seja, fazer uma exploração dos dados em diferentes níveis de detalhes. Essa técnica permite ao usuário aumentar ou diminuir os detalhes dos dados. O *drill down* ocorre quando há um aumento no nível de detalhe da informação, diminuindo o nível de granularidade, assim o usuário pode ir do mais alto nível de informação até o detalhamento. O *drill up* ou *roll up* é o contrário do *drill down*, ele ocorre quando há um aumento no nível de granularidade, levando a uma diminuição de detalhe da informação, assim o usuário pode ir da informação mais detalhada até o nível mais alto dela (MACHADO, 2010).

Na Figura 2 é possível visualizar a operação de *drill down* em um conjunto de dados, onde no primeiro quadro da Figura 2 tem informações resumidas de venda por estado na Região Sul. Ao aplicar a técnica de *drill down* no estado Rio Grande do Sul – RS, os valores de venda são detalhados exibindo por cidade como mostra no segundo quadro da Figura 2.

Volume de Venda		Smartphone		Tablet	
		S1	S2	T1	T2
Região Sul	RS	33	12	8	12
	RC	45	34	20	23



Drill Down
 Dimensão localização geográfica
 Membro RS

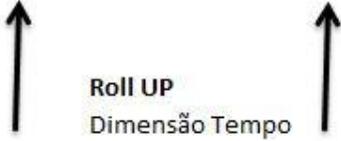
Volume de Venda		Smartphone		Tablet	
		S1	S2	T1	T2
RS	Canoas	13	4	2	5
	Porto Alegre	20	8	6	7

Figura 2 – Operação *Drill Down* na dimensão localização geográfica.

Fonte: (MACHADO, 2010, p. 87).

A Figura 3 é possível visualizar a operação de *Drill Up* na dimensão tempo, onde as informações estavam detalhadas por mês como mostra o segundo quadro da Figura 3, ao aplicar a técnica de *Roll Up* na dimensão tempo os valores que estavam detalhados por mês ficaram resumidos por trimestre como mostra o primeiro quadro da Figura 3.

Volume de Venda		2010			
		Trim. 1	Trim. 2	Trim. 3	Trim. 4
Região Sul	RS	78	67	22	56
	RC	90	67	88	99



Roll UP
 Dimensão Tempo

Volume de Venda		Trimestre 1		
		Janeiro	Fevereiro	Março
Região Sul	RS	30	26	22
	RC	28	30	32

Figura 3 – Técnica de *Drill Up* na dimensão de tempo.

Fonte: (MACHADO, 2010, p. 86).

Para Primak (2008) e Han (2006) os sistemas OLAP não armazenam dados apenas os acessam, mas sistemas OLAP são nomeados segundo a estrutura de dados que estão armazenados no banco de dados, esses sistemas podem ser assim classificados:

- MOLAP – Quando o OLAP é desenvolvido sobre um banco de dados multidimensional, o usuário trabalha com os dados diretamente no servidor (PRIMAK, 2008; HAN 2006).
- ROLAP – Quando o OLAP é desenvolvido através de um banco de dados relacional, a consulta do cliente é enviada ao servidor e processada no mesmo, mantendo o cubo de informação no mesmo (PRIMAK, 2008; HAN 2006).
- DOLAP – A ferramenta dispara uma consulta SQL ao banco de dados de um cliente, e recebe de volta um micro-cubo de informações para ser analisado no computador (PRIMAK, 2008).
- HOLAP – É uma mistura das tecnologias MOLAP e ROLAP, com essa mistura se extrai o que é melhor de cada uma das tecnologias (PRIMAK, 2008; HAN 2006).

1.1.5 ETL - Extract Transform Load

O processo de ETL também conhecido como KETTLE é fundamental para a integração de dados, ele consiste em coletar as informações do local de origem, prepará-los

e armazenar no DW. Este processo é considerado uma das etapas mais críticas de um projeto de DW. Por mais que a ferramenta ETL seja sofisticada todo processo é complexo, detalhado e trabalhoso (PRIMAK, 2008; SCHEPS, 2008).

O objetivo do *Data Warehouse* é integrar os dados de múltiplas fontes. O processo de integração de dados é composto por três processos principais sendo eles, a extração, transformação e carregamento dos dados (FOUCHÉ, 2011).

Segundo Primak (2008) e Scheps (2008) todo o processo de ETL é composto por cinco etapas. A primeira é o *Extract*, este processo consiste em extrair os dados, que podem estar em banco de dados internos e externos, planilhas, arquivos de texto entre outros. A extração dos dados deve ocorrer de forma precavida devido aos dados serem importantes para a empresa.

Segundo Fujiwara (2006), há ferramentas de ETL que conectam em diversas bases de dados e extraem os dados no formato que é necessitado. Logo utilizar estas ferramentas de ETL diminui a carga de processos dos bancos de dados transacionais.

A segunda etapa consiste em pegar os dados extraídos no processo anterior e separar. O processo de separar os dados tem por finalidade retirar os dados que não são uteis ou não agregam valor, corrigir algumas imperfeições nos dados extraídos, a fim de manter os dados concisos e com qualidade (PRIMAK, 2008; SCHEPS, 2008).

A terceira etapa é o processo de transformação dos dados, como os eles podem vir de várias fontes, a mesma informação pode ter vários formatos, por exemplo, o sexo de uma pessoa, em um sistema pode estar H para homem e M para mulher, mas em outro sistema pode estar M para masculino e F para feminino, logo o processo transforma esses dados para um único formato uniforme (PRIMAK, 2008; SCHEPS, 2008).

Para Fujiwara (2006), a etapa de transformação dos dados verifica os padrões nos dados não padronizados, criando novos campos através dos existentes.

A Figura 4 exemplifica o processo de transformação de dados, onde o valor que representa o sexo de uma pessoa é tratado de forma diferente em cada base de dados, o processo de transformação de dados verifica os dados e define qual será o padrão para aquela informação.

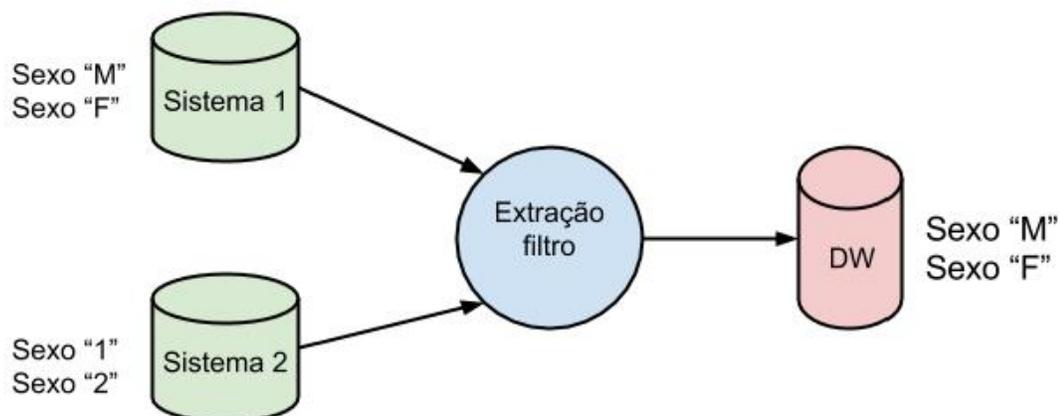


Figura 4 – Integração de dados no *Data Warehouse*.

Fonte: (MACHADO, 2010, p. 30).

A quarta etapa é o processo de carga dos dados da etapa anterior, ela tem por objetivo fazer a carga dos dados no DW, e normalmente este processo é feito a partir do banco de dados temporário conhecido como *staging area* (PRIMAK, 2008; TURBAN, 2009).

A quinta etapa é responsável por fazer as atualizações dos dados no DW, contudo os dados podem sofrer alterações nos sistemas de onde foram originados.

Quando o processo de ETL tem que extrair dados oriundos de diversas fontes e que pertencem ao mesmo assunto, estes dados nem sempre estarão prontos no mesmo horário, então o processo de ETL grava estes dados na *staging area*, para quando os dados do mesmo assunto estiverem prontos sejam feito novamente o processo de ETL para levar os dados da *staging area* para o DW (PRIMAK, 2008; SCHEPS, 2008).

O administrador do DW ao analisar os dados que serão carregados para o DW sem terem passado pela *staging area*, e ao analisar ele notar que é um volume grande de dados e que é um horário de produção deve armazená-los na *staging area* temporariamente para depois carregar no DW fora deste horário, evitando assim lentidões na base e valores alterados no dia para o analista de BI (PRIMAK, 2008).

1.1.6 Benefícios do BI para as empresas

O BI é capaz de exibir informações precisas da organização quando necessário, junto de uma visão em tempo real do desempenho corporativo de modo geral e individual, sendo

este o seu principal benefício. Tais informações são necessárias para as tomadas de decisões, planejamentos e até mesmo para se manter ativo no mercado (TURBAN, 2009).

Para Eckerson (2003) uma pesquisa feita com usuários de sistemas de BI, tendo o propósito de listar os benefícios da ferramenta através da visão deles, os benefícios são: economia de tempo, versão única da verdade, melhores estratégias e planos, melhores decisões táticas, processos mais eficientes e economia de custos.

Segundo Primak (2008), quando a implantação do BI ocorre com sucesso alguns benefícios podem ser alcançados, sendo alguns deles:

- Vantagem competitiva;
- Redução de custos com softwares;
- Alinhamento de informação estratégica e operacional;
- Rapidez na informação para a tomada de decisão estratégica;
- Informação consistente em vários locais dispersos.

Para a empresa obter vantagem competitiva o fator crucial é saber direcionar seu capital intelectual para que o projeto de BI atenda as expectativas.

Para Turban (2009) muitos dos benefícios do BI não podem ser medidos, devido a complexidade das metodologias que são utilizadas para avaliá-lo, e ao fato do *Data Warehouse* servir para outras aplicações na organização.

1.1.5 Presente e Futuro

As organizações que estão atualmente no mercado, estão precisando de qualidade e agilidade nas trocas de informações, mas estes requisitos não são para lucrarem mais, são para conseguirem sobreviver ou evitar falência. Consequentemente a organização deve levar em consideração os benefícios do BI. A previsão é que nos próximos anos, mais pessoas utilizaram ferramentas de BI para análise todos os dias (TURBAN, 2009).

As organizações que beneficiam-se com o uso de BI e ainda estendem as informações importantes aos colaboradores estão aumentando o uso dos ativos de dados existentes. Entretanto mais ferramentas de análises estarão disponíveis no mercado, para facilitar a tomada de decisão, essas ferramentas atenderam todos os níveis hierárquicos da organização, podendo também ser usados por produtores, varejistas, governos e órgãos estaduais (TURBAN, 2009).

1.1.6 Soluções Disponíveis

Atualmente existem muitas ferramentas de BI que estão no mercado, tornando mais difícil a escolha de uma para ser implantada na organização. Os fornecedores mais conhecidos são: SAP, IBM, Oracle, Microsoft, SAS, MicroStrategy, Qlinkview, Pentaho e Jasper, das ferramentas disponíveis tem as proprietárias e as da comunidade (SOUZA, 2010).

1.2 DATA WAREHOUSE

O *Data Warehouse* (DW) é um conjunto de dados com informações da organização, esses dados têm origens em diversas fontes, que constantemente são fragmentados por outros sistemas operacionais. Com o Data Warehouse não se tem este problema de dados parciais, pois, ele acessa, integra e organiza os principais dados, assim os gerentes podem usar sistemas de suporte a decisão e outras ferramentas de BI para conduzi-los a tomarem melhores decisões baseando nos dados do DW (TURBAN, 2009).

O termo *Data Warehouse* é utilizado para definir um conjunto de informações que foram coletadas de diversas fontes de dados, que passaram por um processo de tratamento de dados e armazenados em um único local. Os dados armazenados no *Data Warehouse* segue um determinada esquema de organização de forma a facilitar o acesso e o entendimento da estrutura (SILBERSCHATZ, 2011).

1.2.1 Definições e Conceitos de DW

Os bancos de dados são muito importantes para as empresas, e analisar os dados neles contidos é difícil, pois normalmente as empresas tem uma grande quantidade de dados neles. Os bancos de dados operacionais armazenam uma coleção de dados das operações diárias das empresas por um determinado sistema de aplicações. Esses dados mantidos são chamados de operacionais (PRIMAK, 2008).

O *Data Warehouse* é um banco de dados, no qual os seus dados são derivados dos bancos de dados operacionais, armazenando dados analíticos de muitos anos, que são destinados ao processo de tomada de decisão. Ele é construído para armazenar os dados retirados de vários sistemas que podem ser distintos, tornando se um meio de análise das informações para as empresas, tais dados não são limitados por tabelas e linhas estritamente relacionais (PRIMAK, 2008).

A Figura 5 exibe os principais componentes que compõe um *Data Warehouse*, na mesma figura os dados que são carregados no *Data Warehouse* podem vir de qualquer fonte de dados, tais como uma base dados operacional ou dados externos que podem ser arquivos de texto, planilhas, documentos entre outros, após os dados serem carregados no *Data Warehouse* pelo processo de ETL, eles ficam disponíveis para acesso através de inúmeras ferramentas.

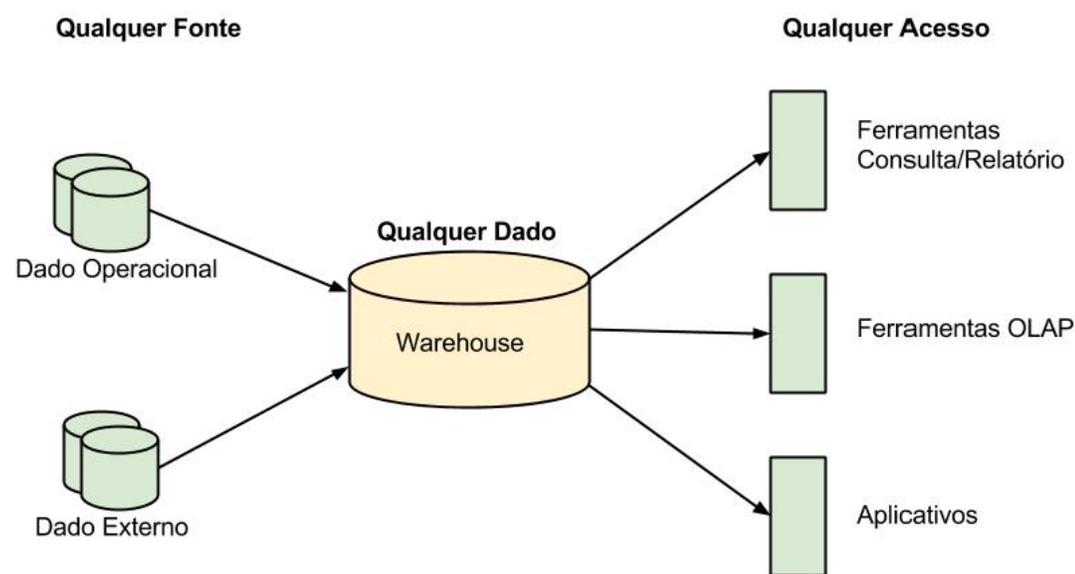


Figura 5 – Principais componentes de um *Data Warehouse*.
Fonte: (PRIMAK, 2008, p. 40).

Para Inmon (2005) os motivos para haver a distinção das bases de dados podem ser vários, sendo alguns deles: os dados dos sistemas transacionais são diferentes dos dados analíticos; os usuários dos sistemas transacionais são diferentes dos usuários dos sistemas analíticos; a quantidade de usuários que estão usando o sistema transacional é superior ao sistema analítico; na base transacional os dados sofrem constantes mudanças já na base analítica os dados são usados somente para consultas entre outros.

Com o histórico de dados é possível tornar a empresa mais competitiva estrategicamente. Os dados armazenados no DW encontram-se detalhados e resumidos e sempre estarão separados de qualquer outro sistema transacional. Por armazenar dados analíticos as consultas ao DW podem tornar-se complexas por acessar um grande número de dados, logo é importante haver muitos índices criados para que tenha performance no DW (PRIMAK, 2008; TURBAN, 2009).

Os tomadores de decisões precisam de informações precisas das operações internas embasados em fatos concretos, tendências e mudanças, logo o DW organiza os dados corporativos de maneira integrada através do cruzamento de informações de diversas fontes com dados históricos e gerando uma única fonte de dados, que será usada para carregar os *Data Marts*, consequentemente agilizar o processo de tomada de decisão diminuindo o índice de erros, e aumentando os lucros (PRIMAK, 2008; TURBAN, 2009).

1.2.2 Características da Tecnologia de *Data Warehousing*

Para Inmon (2005) e Machado (2010) o *Data Warehouse* é definido como um conjunto de dados para apoiar o processo de tomada de decisão tendo as seguintes características principais: orientação por assunto, variação com o tempo, não volatilidade e integração.

Um *Data Warehouse* armazena os dados por assuntos de interesse da organização de modo a facilitar a análise de dados, relevante a tomada de decisões, tais como, vendas, produtos ou clientes, em oposição aos bancos de dados tradicionais que são orientados por aplicações. Os dados armazenados por assuntos são referentes aos principais processos, permitindo aos usuários uma visão melhor e motivos do desempenho da organização (TURBAN, 2009; MACHADO, 2010; INMON, 2005).

Um *Data Warehouse* é variável com o tempo devido ele ter como dado chave uma data, através desta data é possível navegar nas hierarquias de tempo, esse tipo de campo leva a tomada de decisões, por serem precisos e podem ser usado para detectar tendências, variações, previsões e fazer comparações. Os dados do *Data Warehouse* não são atualizáveis, logo são relativos a um determinado instante de tempo (TURBAN, 2009; MACHADO, 2010; INMON, 2005).

Um *Data Warehouse* não é volátil devido ele ter duas operações básicas, inserção e a consulta dos dados, ou seja, após os dados serem integrados, transformados e incluídos não podem ser alterados. Quando há a necessidade de atualização os dados antigos são descartados e os novos são inseridos, tornando o *Data Warehouse* exclusivamente para acesso aos dados (TURBAN, 2009; MACHADO, 2010; INMON, 2005).

A integração de dados no *Data Warehouse* é fundamental, ela ocorre quando os dados passam das bases de dados transacionais para o a base de dados analítica, pois os dados originados em fontes diferentes podem ter formatos diferentes para a mesma finalidade, logo

os dados devem ser tratados, evitando conflitos de nomenclaturas e valores entre unidades medidas, tornando-os consistentes (TURBAN, 2009; MACHADO, 2010; INMON, 2005).

1.2.3 *Data Mart* - DM

Um *Data Mart* é um pequeno armazém de dados que normalmente se concentra em um assunto ou departamento específico, ele representa uma parte do conteúdo do *Data Warehouse*, normalmente ele é modelado seguindo o esquema estrela. Os seus dados são direcionados a uma área específica da empresa. O DM armazena informações por área temática tendo o escopo mais limitado e são mais identificados com grupos de necessidades dos usuários, enquanto o DW trata de problemas de toda a empresa para que o suporte a tomada de decisão atenda todos os níveis organizacionais. Um DM pode ser criado a partir de um DW, ou o DM pode ser usado para popular o DW (MACHADO, 2010; HAMMERGREN, 2009).

A Figura 6 apresenta melhor o que é um *Data Mart*, um *Data Warehouse* é composto por um conjunto de *Data Marts*.

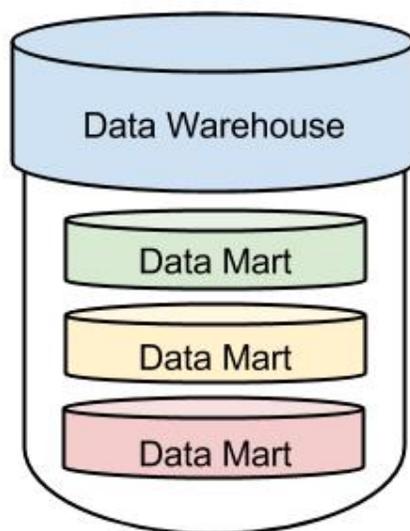


Figura 6 – Representação de um *Data Mart*.
Fonte: (MACHADO, 2010, p. 45).

As diferenças entre o DM e o DW são apenas em relação ao tamanho e ao escopo a ser resolvido. Uma das vantagens de usar o DM, é que ele possibilita um retorno rápido das

informações, garantindo um rápido processo de tomada de decisão. A ideia do DM é começar do pequeno pensando no todo, dessa forma as empresas estão optando pelo DW que inicia a partir de um departamento ou um pequeno grupo de pessoas já visando no crescimento (PRIMAK, 2008; MACHADO, 2010).

Os *Data Marts* para o autor Primak (2008, p.49) “podem ser prototipados muito mais rápido, com alguns lotes sendo construídos entre 30 e 120 dias e sistemas completos sendo construídos entre três e seis meses.”

A criação de um DW requer tempo e dinheiro, e esses critérios variam dependendo do tamanho da base de dados a ser trabalhada. Estudos ocorridos nas indústrias em geral que tem DW em produção chegaram a conclusão de que aproximadamente setenta a oitenta por cento delas começaram o DW através do DM. (PRIMAK, 2008; TURBAN, 2009; MACHADO, 2010).

1.2.4 Operacional Data Stores - ODS

O ODS também conhecido como *staging area* é um banco de dados usado como armazenamento intermediário dos dados antes deles serem passados para o DW, os seus dados são atualizados durante as operações comerciais ocorridas no banco de dados transacional. O ODS é usado também para decisões de curto prazo, pois ele só armazena informações muito recentes. O processo de ETL de um ODS é parecido ao de um *Data Warehouse* (TURBAN, 2009).

A Figura 7 mostra o funcionamento da *staging area*, onde os dados são coletados do sistema de vendas entre uma hora e duas horas e são armazenados no ODS, depois os dados são coletados do sistema financeiro das quatro horas e cinco horas e armazenados no ODS, após o processo de coleta de dados eles são tratados e sincronizados para o *Data Warehouse*.

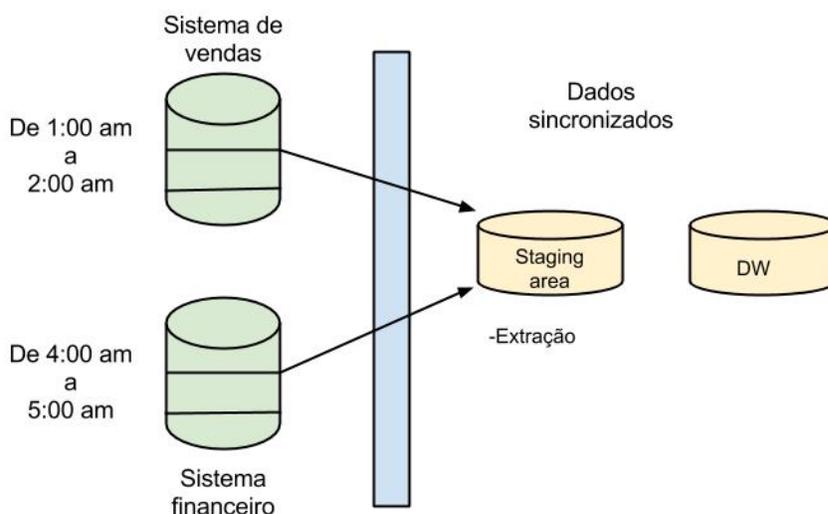


Figura 7 – Funcionamento de uma *Staging Area*.
 Fonte: (MACHADO, 2010, p. 40).

Uma *staging area* facilita a integração dos dados no DW, pois os dados são coletados dos sistemas legados pela *staging area* logo que estão disponíveis, ao fim de todo o processo o administrador do DW pode integrar os dados fora do horário de pico de operações, pode ser usado como base de análises do ambiente operacional, pois os dados não são sumarizados e o nível de detalhamento é o mesmo do ambiente retirado (MACHADO, 2010).

A *staging area* proporciona um melhor desempenho no processo de transmissão de dados para o DW, caso ocorra algum erro no processo o administrador do DW pode verificar onde ocorreram os erros. (TURBAN, 2009; MACHADO, 2010).

1.2.5 Granularidade

A granularidade dos dados refere-se ao nível de detalhe das informações. Quanto maior for o nível de detalhe menor será a granularidade, quanto mais detalhes houver, maior será a granularidade (MACHADO, 2010).

Segundo Inmon (2005) é o aspecto mais importante do *Data Warehouse*, pois ele influencia o volume de dados contido no mesmo e os tipos de consultas que podem ser realizadas.

A granularidade no DW está relacionada com o volume de dados contido nele. Quanto maior o nível de granularidade, menor é o espaço ocupado em disco e o número de índices

criados. Quanto menor for o nível de granularidade, temos economia de espaço usado em disco e o número de índices criados. (MACHADO, 2010; INMON, 2005).

Para Cramer (2006), o nível de granularidade em um *Data Warehouse* influencia diretamente no desempenho das consultas de dados realizadas no DW. Logo quanto maior for a granularidade melhor desempenho terá o a consulta. Porém uma granularidade baixa permite uma melhor análise dos dados.

1.2.6 Cubos

Os cubos são formas de visualização de dados de um banco de dados multidimensional, onde cada lado do cubo representa um assunto que se deseja analisar. Dessa forma é possível visualizar de várias formas os dados que estão sendo apresentados (ROSINI, 2003).

Os cubos de dados normalmente são multidimensionais, nele é permitido armazenar, dados de interesse do usuário. Cada lado do cubo representa um assunto, assim é possível analisar de forma rápida várias informações que está sendo apresentada pelo cubo (MACHADO, 2010).

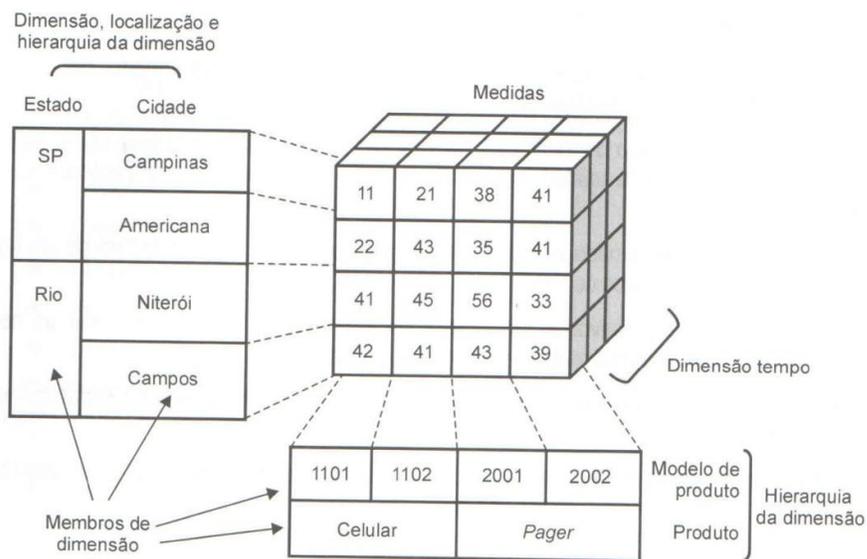


Figura 8 – Cubo de dados e suas dimensões.

Fonte: (MACHADO, 2010, p. 82).

A Figura 8 é a representação de um cubo de informações as dimensões deste cubo é localização, tempo e produto. A dimensão localização e produto possuem dois níveis de

hierarquia. A dimensão localização tem o estado e a cidade, já a dimensão produto tem a descrição do produto e o modelo e a dimensão tempo é por ano de venda. A tabela Fato do cubo é relacionada a vendas, as dimensões são produto, localização e tempo a medida deste cubo é volume de venda (MACHADO, 2010).

1.2.7 Modelagem Multidimensional

A modelagem multidimensional é uma técnica utilizada na construção de um *Data Warehouse* e envolve concepção e visão de um modelo de dados. Esta modelagem é mais nova relativamente, e não possui detalhes de técnicas que nem o modelo ER (Entidade Relacionamento), mas ela é mais simples de entender que a modelagem ER (MACHADO, 2010). Segundo Moreira (2006) esta modelagem tem por objetivo facilitar como as informações são visualização pelo usuário, as informações são representadas como um cubo e nele é possível visualizar por diversos ângulos, permitindo fatiar o cubo e trabalhar com partes dele.

A visualização de um modelo multidimensional é melhor representada por um cubo de informações, que pode ser observado de vários ângulos, apesar de ser representado por um cubo a modelagem multidimensional é mais fácil de entender. O centro do cubo é chamado de fato, pois é composta de dados de medidas, e cada lado do cubo é uma dimensão, pois permite ver a mesma informação de vários ângulos. Com essa forma de visualização é possível fatiar os dados e as dimensões para análise (MACHADO, 2010).

Para Kimball (2013) a modelagem dimensional consiste em uma técnica de projeto que visa buscar e apresentar os dados de uma estrutura padrão de forma intuitiva, tudo isso com um acesso muito rápido.

O modelo multidimensional usa dois tipos de esquemas de dados, sendo eles o *Star schema* e o *Snowflake schema* (MACHADO, 2010).

Para o autor Machado (2010), o *Star schema* ou também chamado de modelo estrela, ele é a estrutura básica de um modelo de dados multidimensional, e em sua composição possui uma tabela central chamada de fato e um conjunto de tabelas chamadas de dimensões conectadas a tabela Fato como mostra a Figura 9.

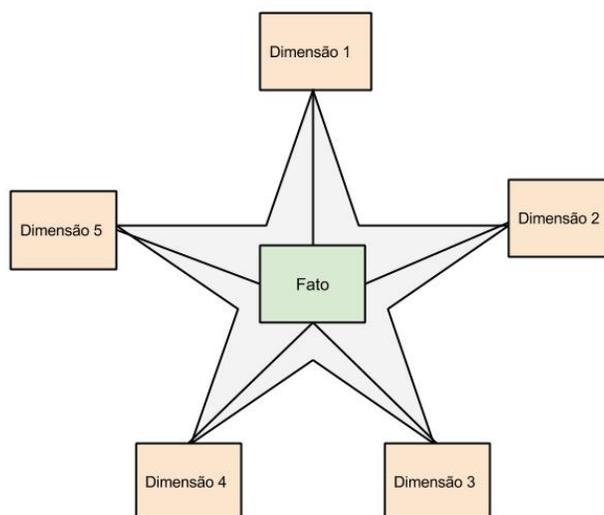


Figura 9 – Modelo de dados *Star Scheme*.

Os relacionamentos entre as tabelas fato e dimensão são ligações simples em que o relacionamento é de um para muitos. Ou seja, uma linha na dimensão se repetirá várias vezes na tabela Fato (MACHADO 2010).

Para Ballard (2006), o esquema estrela resulta em uma menor quantidade de tabelas no *Data Warehouse*, tendo como consequência um fator característico desta modelagem uma repetição de dados nas tabelas.

Segundo Machado (2010), o *Snowflake schema* conhecido como modelo floco de neve, este modelo é parecido com o modelo estrela, seguindo a composição de tabela dimensão ligada à tabela Fato, mas neste modelo há também ligações de tabela dimensão em outra tabela dimensão.

Para Ballard (2006), no esquema floco de neve as dimensões estão padronizadas, eliminando assim a repetição dos dados diferenciando do esquema estrela. Neste esquema os dados das tabelas dimensões estão distribuídos em outras tabelas.

A Figura 10 mostra a modelagem de dados do esquema *Snowflake*.

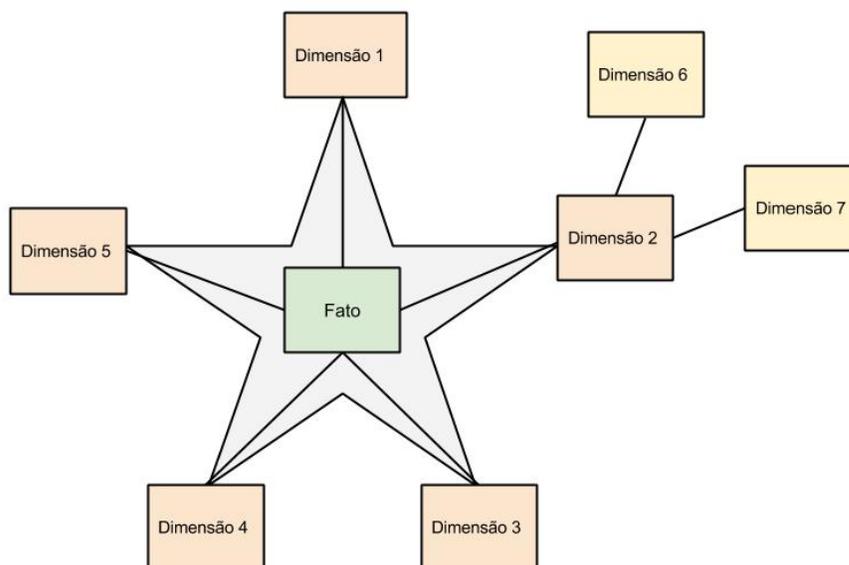


Figura 10 – Modelagem *Snowflake Scheme*.

O modelo floco de neve é o desmembramento de uma dimensão em outras dimensões que possuem hierarquia de dados. O relacionamento entre estas tabelas é de muitos para um de acordo com que vai descendo a hierarquia até chegar à tabela Fato. Este modelo evita a redundância de dados de dados em tabelas (MACHADO, 2010).

Nas tabelas dimensões suas estruturas são separadas por níveis, no qual cada nível da tabela dimensão deve corresponder a uma outra coluna da tabela dimensão. Assim os dados são classificados por grau de detalhe de forma hierárquica (ITALIANO, 2004).

A modelagem multidimensional é composta por três tipos de estruturas, sendo eles fatos, dimensões e medidas (MACHADO 2010; REIS, 2009).

A tabela Fato é a tabela central nela há uma coleção de dados que representa um evento de negócio ocorrido na empresa e é usado para analisar o processo de negócio. A tabela Fato tem como característica básica valores numéricos e datas. Nas empresas tudo que é importante para a tomada de decisão é caracterizado como um fato, por uma tabela Fato conter datas ela é evolutiva, ou seja, muda com o tempo. As características da tabela Fato são que os valores variam ao longo do tempo, os valores são numéricos e o histórico pode ser mantido e cresce com o passar do tempo. A tabela Fato possui chaves que se relacionam com as tabelas dimensões (MACHADO 2010; REIS, 2009).

A tabela dimensão são elementos que participam de algum fato, assunto de negócio e são uma das possíveis formas de visualizar os dados que estão na tabela Fato. Normalmente as dimensões possuem atributos descritivos dos elementos que estão na tabela Fato e descrevem com detalhe os elementos que fazem parte da análise dessa dimensão (MACHADO, 2010). Para Reis (2009) os dados da dimensão representam o que está sendo analisado e por isso oferece ao usuário várias possibilidades de combinações e interseções na hora de analisar os dados.

As dimensões são caracterizadas como perguntas realizadas para as tabelas fatos como, por exemplo, a tabela Fato tem informações de venda e quando se quer saber qual produto foi vendido ou onde realizou a venda usa se as dimensões para responder a pergunta (MACHADO, 2010).

As medidas são atributos do tipo numéricos armazenadas na tabela Fato e caracterizam o desempenho de uma medida em relação as dimensões que participam do fato. A medida, por exemplo, pode ser valor de compra ou de venda, o numero de unidades vendidas ou devolvidas, quantidade em estoque entre outros (MACHADO 2010; REIS, 2009).

1.3 FERRAMENTA DE BUSINESS INTELLIGENCE

Existem sistemas de BI no mercado que dão suporte a empresas, a fim de que os seus problemas de tomadas de decisão sejam solucionados, baseado na qualidade da informação. Muitos desses sistemas possuem preços elevados e para uma empresa investir inicialmente em um produto, a mesma deve ter retornos relevantes.

Dentre muitas ferramentas disponíveis BI no mercado, foram encontrados alguns pontos relevantes que ajudaram a decidir qual ferramenta escolher. Para os autores BOBSIN (2010, p.30) e SOUZA (2012, p.30) das ferramentas disponíveis a que fornece um maior número de recursos e é *Open Source* é a suíte Pentaho, além de ser a mais usada em todo o mundo, logo ela foi escolhida para ser a ferramenta de BI.

Sabendo que a maioria das pessoas que utilizaram o sistema não possui conhecimento de como funciona a ferramenta, então utilizar inicialmente software *Open Source* é a melhor maneira para implantá-lo em uma empresa.

1.3.1 Pentaho

A suíte Pentaho é composta por um conjunto de software livre desenvolvido na linguagem Java, e rodam em qualquer plataforma que tem uma JVM, os *softwares* servem

para gerar soluções de BI, os *softwares* execução processo de ETL (*Extract Transform Load*) que fazem a carga do *Data Warehouse*, criam relatórios pré-formatados, painéis de visualização, garimpagem de dados (*Data Mining*) entre outros, quando combinados todos estes recursos se tem soluções mais aprimoradas (SALLES, 2013).

O *Pentaho* possui duas versões sendo eles: *Pentaho Community Edition* (CE) onde há uma comunidade que é responsável por todo o suporte da ferramenta; *Pentaho Enterprise Edition* (EE), onde a empresa *Pentaho* que fornece o suporte para a ferramenta.

Os fundadores da empresa queriam desenvolver uma ferramenta de BI, pois eles achavam que as ferramentas disponíveis no mercado não eram flexíveis e poderosas o suficiente para atender qualquer necessidade, assim eles acreditavam que deveria haver uma maneira melhor de montar soluções de BI, contudo surgiu a empresa (SALLES, 2013).

Após desenvolver a primeira versão do *Pentaho* e uma demonstração de como usá-lo, chamado de PCI (*Pre-Configured Installation*), o PCI fez tanto sucesso que a *Pentaho* acabou adotando e melhorando o PCI, este novo componente foi levado ao *Pentaho BI Server*, que tem o console de usuário e o console de administração (SALLES, 2013).

1.3.2 Componentes

A suíte do *Pentaho BI* é composta por dois componentes, sendo eles o *Pentaho BI Platform*, implementada na forma de um servidor web chamado de *BI Server*, o cliente de desenvolvimento, que são as ferramentas que criam conteúdo para a plataforma *BI Server*. (SALLES, 2013).

O *BI Server* e seus componentes são responsáveis pela execução e controle das soluções, a sua base consiste em uma máquina de *workflow*, que chama o componente desejado para trazer o resultado para o usuário. Ele possui níveis de acesso ao ambiente por usuário e senha, controle de acesso aos objetos baseado em usuário e grupos, controle de acesso aos dados, controle de acesso a opções do sistema e *logs* (SALLES, 2013; PENTAHO, 2014).

No *BI Server* usuários podem criar relatórios *Ad Hoc*, navegador nos cubos de informações com a ferramenta de OLAP disponível, criar relatórios pré-configurados e personalizados, agendar a execução automática de relatórios e recebê-lo no e-mail, executar relatórios em segundo plano, visualizar relatórios e *Dashboards* criados para eles (SALLES, 2013; PENTAHO, 2014).

A suíte Pentaho possui ferramentas para que ajudem no processo de construção de um Data Warehouse, e ferramentas que utilizam as informações contidas no DW para criar os conteúdos para os usuários. Dentre essas ferramentas temos:

- *Pentaho Report Designer*: Que é uma ferramenta para geração de relatórios com muitos recursos, tais como gráficos, imagens, sub-relatórios dentre outros (SALLES, 2013);
- *Pentaho Metadata Editor*: É uma ferramenta para mapear as tabelas físicas de um banco de dados e dar nomes a elas, para que o usuário consiga gerar relatórios *Ad Hoc*. (SALLES, 2013);
- *Pentaho Schema Workbench*: Esta ferramenta possibilita a criação de cubos OLAP, com métricas, dimensões e hierarquias (SALLES, 2013);
- *Pentaho Design Studio*: É um ambiente de desenvolvimento para Eclipse (*Plugin*) que fornece ao desenvolvedor acesso aos recursos do Pentaho (SALLES, 2013);
- *Pentaho Weka*: A *Weka* é um ambiente para o usuário criar e testar hipóteses na base de dados (*Data Mining*) (SALLES, 2013);
- *Pentaho Data Integration*: É uma plataforma de integração de dados completa, ela realiza a extração, transformação e a integração de dados para alimentar o *Data Warehouse*, possui uma interface que facilita em todo o processo, para usar os componentes é preciso arrastar e soltar, dessa forma ela elimina a codificação e complexidade e consegue uma maior produtividade (SALLES, 2013; PENTAHO CORPORATION, 2014);

Para Salles (2013) a suíte Pentaho trabalha com todos os bancos de dados que exista um *driver* JDBC, entretanto na plataforma *Windows* ele usa o *driver* ODBC, a performance do *driver* ODBC é menor quando comparado ao JDBC, pois possui uma camada extra no processo de buscar os dados.

1.4 ENGENHARIA DE SOFTWARE

No presente trabalho, a Engenharia de *Software* apoia a criação de um questionário para avaliar a usabilidade de um software de *Business Intelligence* verificando se o mesmo proporciona facilidade em ser operado. Para que tal objetivo seja alcançado, algumas áreas de conhecimento da Engenharia de *Software* devem ser consideradas. Eles seguem relacionados nas subseções seguintes.

1.4.1 Qualidade *software*

Pouco a pouco os softwares se tornaram parte essencial da vida da maioria dos profissionais, independente de sua área de atuação, com isso cada vez mais se faz necessário *softwares* que apresentem os resultados que se propõem.

Segundo PRESSMAN (2011), códigos mal feitos de software chegou a ser responsável por até 45% do tempo inatividade dos sistemas computacionais, causando um prejuízo às empresas americanas estimado em US\$ 100 bilhões em termos de manutenção e redução da produtividade, assim, cada vez mais se pede softwares de qualidade, isto é, softwares que apresentem os resultados propostos em suas especificações, por isso Sommerville (2010), nos últimos anos a qualidade de software tem se aprimorado significadamente.

1.4.1.1 Conceito

Para Sommerville (2010), a qualidade no contexto dos softwares é um conceito complexo, diferente de qualidade em outros contextos, como da manufatura, contudo, Pressman (2011), explica que qualidade de software pode ser entendido como um software que, mediante um projeto de qualidade, apresenta para o usuário final as funções e os recursos que o usuário deseja, fornecendo confiabilidade e isenção de erros, agregando valor, isto é, trazendo benefícios para tanto para o fabricante como para o usuário.

Para que a qualidade do software seja validada, diversas abordagens podem ser utilizadas, sendo que cada uma delas descrevem o que deve ser observado para um software seja considerado um software de qualidade. Dentre essas abordagens temos os fatores de qualidade de McCall (1977), que descreve que um software para ser considerado de qualidade deve atender aos seguintes fatores:

- Correção: O quanto o programa atende aos objetivos do cliente;
- Confiabilidade: O quanto o programa realiza o que se propõe a realizar;
- Eficiência: O quanto o programa custa para o hardware;
- Integridade: O quanto o programa pode ser usado por não autorizados;
- Usabilidade: Quanto esforço é necessário para aprender a operar o programa;
- Facilidade de manutenção: Quanto esforço é necessário para localizar um erro e corrigi-lo;
- Flexibilidade: Quando esforço é necessário para modificar o programa em execução;

- Testabilidade: Quanto esforço é necessário para testar o programa a fim de garantir que ele desempenhe a função destinada;
- Portabilidade: O esforço necessário para transferir o programa para outro ambiente;
- Reusabilidade: O quanto o programa, ou partes dele, pode ser utilizado por outras aplicações.

1.4.2 Usabilidade

Usabilidade é um conceito da Engenharia de *Software* (ES) que se refere à facilidade de aprender, operar, dar entradas e interpretar as saídas de algum software (MCCALL, 1977), abordando elementos como componentes do sistema, operadores e o ambiente de operação (SOMMERVILLE, 2010).

No contexto da ES, este termo é abordado na etapa de requerimentos Não-funcionais de um software, que é a parte da ES que aborda requerimentos que não estão diretamente ligados a serviços específicos entregues ao usuário, abordando assuntos como a segurança do software e também o comportamento que o mesmo deve apresentar para o usuário final (SOMMERVILLE, 2010).

Para que haja usabilidade, (SOMMERVILLE, 2010) diz que o software precisa ser usável, além de apresentar interfaces e documentações adequadas para seu uso.

Segundo Mccall (1977), usabilidade é um dos fatores que determinam o nível de qualidade de um software, sendo que para verificar a usabilidade de um software é necessário analisar os conceitos que estão relacionados a ele, como:

- Operabilidade: Atributos do software que determinam os operações e procedimentos concernentes à operação do software;
- Treino: Atributos do software que tratam da familiarização do usuário com o software;
- Comunicatividade: Atributos do software que proveem entradas úteis e saídas que podem ser interpretadas pelo usuário.

No ano de 1994 Jakob Nielsen desenvolveu dez heurísticas de usabilidade com a intenção de auxiliar os desenvolvedores, sendo elas:

- **Visibilidade do status do sistema:** O sistema deve manter os usuários bem informados sobre o que está acontecendo, todas as ações precisam de *feedback* instantâneo para orientá-lo.

- **Correspondência entre o sistema e o mundo real:** A linguagem usado pelo sistema para comunicar com o usuário precisa ser facilmente entendidas pelo mesmo;
- **Controle do usuário e liberdade:** Usuários frequentemente escolhem algumas funções do sistema por engano e vão precisar sempre de uma “saída de emergência” claramente marcada para sair daquele estado indesejado sem ter que passar por um extenso “diálogo”. Apoio ao desfazer e refazer;
- **Consistência e padrões:** Os usuários não precisam adivinhar que diferentes palavras, situações ou ações significam a mesma coisa. Ou seja, um mesmo ícone ou palavra deve significar a mesma coisa em todos os lugares que aparece;
- **Prevenção de erros:** Informar ao usuário um caminho para voltar de um erro é importante, mas evita-lo é fundamental. Eliminando as condições passíveis de erros ou verificá-las, apresentado aos usuários uma opção de confirmar antes de se comprometerem com uma determinada ação;
- **Reconhecimento em vez de recordação:** O sistema deve evitar que o usuário precise pensar em como executar uma tarefa todas as vezes que precisar dela. Logo minimizar a carga de memória do usuário tornando objetos, ações e opções visíveis. Instruções de uso do sistema devem estar visíveis e serem facilmente recuperáveis quando necessário;
- **Flexibilidade e eficiência de utilização Aceleradores:** Usuários mais avançados geralmente usam atalhos do teclado, logo a interface deve oferecer estas opções, que acabam tornando a navegação mais flexível;
- **Estética e design minimalista:** Mostrar para o usuário o que ele apenas precisa saber. Procurando tornar as coisas simples, diretas e naturais;
- **Ajude os usuários a reconhecer, diagnosticar e resolver erros:** Mensagens de erro devem estar preparadas para dar suporte ao usuário caso ocorra. Indicar com precisão o problema e construtivamente sugerir uma solução;
- **Ajuda e documentação:** Dúvidas podem surgir durante o uso do sistema, logo qualquer informação deve ser fácil de pesquisar, a fim de oferecer ao usuário um local para ele resolver as dúvidas.

1.4.3 Teste de Usabilidade

Segundo Ferreira (2002), o teste de usabilidade é utilizado para verificar o nível de usabilidade de algum *software*, podendo servir para diferentes propósitos, como identificar problemas de usabilidade ou até mesmo comparar dois ou mais produtos.

Renata (2012), diz que a usabilidade está diretamente relacionada à interface, ao usuário e ao sistema computacional que formam os três principais pilares da Interação Homem-Computador (IHC), e verifica se o *software* é:

- Fácil de aprender;
- Eficiente para usar;
- Fácil de lembrar;
- Pouco sujeito a erros;
- Agradável de usar.

Para que este teste seja eficiente, pelo menos quatro sub-testes são realizados (FERREIRA, 2002), sendo eles:

- Exploração: Usado para avaliar a efetividade do desenho preliminar e conhecer a concepção do usuário ou modelo mental do produto;
- Avaliação: Verifica como um usuário consegue desenvolver tarefas reais, identificando deficiências específicas de usabilidade;
- Validação: Seu objetivo é verificar como o produto se enquadra em relação a padrões de usabilidade, padrões de desempenho e padrões históricos;
- Comparação: O teste de comparação pode ser usado como uma junção dos demais testes sendo que seu objetivo é realização de comparações em todos os níveis.

2 METODOLOGIA

O presente trabalho teve por objetivo analisar a usabilidade de uma ferramenta de *Business Intelligence*, sendo ela a suíte *Pentaho*. Para alcançar este objetivo foram realizadas pesquisas, a fim de compreender como funciona o processo de análise de usabilidade de um *software*. Procure-se conhecer como a análise é realizada por meio de entrevistas e quais os modelos existentes.

Após a realização do estudo de análise de usabilidade, através de pesquisa bibliográfica foi possível identificar o quanto é importante um sistema de *Business Intelligence* ser eficiente para uma empresa.

Foi realizada por meio de um questionário (Anexo I) uma pesquisa com os funcionários da empresa DPC Distribuidor Atacadista S/A a fim de conhecer o quanto a suíte *Pentaho* é agradável de usar para a tomada de decisão. Mais detalhes do questionário e do público serão descritos na subseção seguinte.

2.1 PÚBLICO ALVO DO QUESTIONÁRIO

Como público respondente do questionário foi escolhido os colaboradores que trabalham na área comercial da empresa DPC Distribuidor Atacadista S/A, segundo o papel deles dentro da empresa onde trabalham são: Gerentes e coordenadores de vendas, colaboradores e prestadores de serviço para a empresa totalizando em 9 (nove) participantes.

Para os respondentes, o questionário foi apresentado por meio de um e-mail, contendo o endereço para acessá-lo.

2.2 ELABORAÇÃO DO QUESTIONÁRIO

As perguntas do questionário foram elaboradas fundamentando nos conceitos de usabilidade expostos no tópico 1.4 (Engenharia de *Software*) que se refere a facilidade de operar o *software*. O questionário é composto por 28 (vinte e oito) questões, sendo todas de múltipla escolha.

Com o objetivo de facilitar o entendimento do questionário por parte dos funcionários, foi realizada uma breve introdução do motivo do questionário. Dessa maneira os entrevistados

responderam as questões de acordo com o acesso delas ao sistema, visando tornar o processo de respostas menos cansativo. Nem todos os entrevistados responderam a todas as perguntas.

2.3 COLETA DOS DADOS

Com o objetivo de alcançar os resultados esperados, foi utilizado o *Google Docs* para criar um formulário, este meio é seguro e eficiente. Por meio desta ferramenta foi possível compartilhar o questionário, além de possibilitar o envio das respostas, tudo isso através da internet.

O questionário foi divulgado através de um endereço *web* que redirecionava o entrevistado para a página de perguntas. O endereço *web* foi enviado no e-mail dos funcionários onde ele foi aplicado pessoalmente em cada participante, marcando o horário com o funcionário e enviando para o e-mail.

A aplicação do questionário ocorreu entre os dias na primeira semana de novembro do ano de 2014, sendo entre os dias 3 a 7 do mesmo mês informado.

2.4 TRATAMENTO DOS DADOS

Foram obtidos ao todo 9 respostas dos participantes, no entanto havia questões que somente alguns participantes podiam responder, devido as funcionalidades que foram liberadas para cada um. Uma análise preliminar das respostas possibilitou perceber que os objetivos de analisar a ferramenta de *Business Intelligence* foram alcançados de maneira geral.

Os resultados foram exibidos através de gráficos e tabelas com a intenção de facilitar a compreensão.

2.5 AMBIENTE DE ESTUDO

O presente estudo foi realizado na área comercial da empresa DPC (que autorizou este trabalho (Anexo II)), que é responsável por, analisar e avaliar os resultados obtidos da empresa informando o atual desempenho aos responsáveis e caso necessário corrigir decisões tomadas, definir áreas de venda ou identificar novas, elaborar e negociar propostas de venda, analisar o mercado e condições da empresa para propor melhorias nas vendas, prospectar clientes, analisar a carteira de trabalho da equipe de vendas, informar a equipe de vendas

sobre mudanças ocorridas na entrega, cobrança entre outras, identificar gargalos e propor novas medidas, assegurar a satisfação do cliente com o serviço fornecido pela empresa, propor soluções de problemas para clientes especiais, entre outras responsabilidades.

A empresa DPC Distribuidor Atacadista S/A atua no mercado de distribuição de cosméticos, perfumaria e higiene pessoal a mais de dez anos, e está presente em muitos estados brasileiros, com a sua matriz localizada na cidade de Caratinga no estado de Minas Gerais, e os demais estados possuem filiais. Agrega-se ao empreendimento, uma das empresas do grupo, a DPC Logística e Transportes, que é responsável por entregar as vendas, contando com sofisticado sistema de rastreamento via satélite, o qual monitora seus veículos 24 horas por dia (DPC, 2013).

O principal objetivo do BI para a empresa DPC é permitir que os diretores, gerentes, coordenadores de vendas tenham acesso interativo aos dados, proporcionando para eles uma análise adequada, servindo como base para melhores decisões.

A principal base de dados da empresa DPC tinha o tamanho aproximadamente de um *terabyte* de dados em maio de 2014, sendo considerado um tamanho grande para um banco de dados. Extrair informações ricas em detalhes de um banco de dados transacional é um grande problema, pois, deve-se analisar e estudar estes dados que geralmente ficam desnormalizados, para depois transformá-los em informações e guardá-los em um repositório central, chamado de *Data Warehouse*.

O sistema de gestão corporativa que a empresa usa foi desenvolvido para atender a diversos ramos de atividade e por isso relatórios e informações finais que são gerados demoram e em alguns casos é necessário incorporar a outros relatórios para atenderem as necessidades para a tomada de decisão.

2.6 IMPLANTAÇÃO

Buscando a solução para o problema descrito, foram definidos alguns caminhos a serem seguidos, para concretizar o trabalho. Para maior controle de todo o projeto, a implantação da solução de BI foi dividida em algumas fases:

- Ter apoio da gestão
- Levantamento das necessidades
- Mapeamento dos sistemas e dos dados
- Definir utilizadores

- Garantir a integração e qualidade dos dados
- Implantação

Na fase de ter apoio da gestão, o projeto de BI deve ter apoio dos diretores e gerentes da área comercial de venda para que o projeto não fique mal visto ou fracasse.

A fase de levantamento das necessidades procura entender quais as informações são precisas pelo setor comercial, sem se preocupar com a existência dela. Nesta fase foram realizadas entrevistas para obter o conhecimento das necessidades.

Na fase de mapeamento dos sistemas e dos dados é analisada a integridade do banco de dados conhecendo as informações armazenadas. Nesta fase foram realizadas entrevistas com o gerente do departamento de TI da empresa para obter conhecimento da estrutura do banco de dados.

Na fase de definir utilizadores são definidas quais pessoas terão acesso ao sistema e quais informações poderão ser visualizadas por cada utilizador.

A fase de garantir a integração e qualidade dos dados, para que o projeto de BI seja um fator decisivo de sucesso na empresa as informações exibidas devem estar o mais próximo possível da realidade.

A fase de implantação é a última e a mais importante, nela a ferramenta escolhida foi implantada, as informações serão disponibilizadas no sistema, e os colaboradores usufruíram da ferramenta.

2.6.1 Ter apoio da gestão

A fase de obter apoio gestão iniciou-se com uma reunião marcada com o gerente de projetos onde o mesmo é responsável pelo departamento de BI da empresa. Na reunião foi feito a proposta de implantar o *Pentaho* que é ferramenta de BI, informando para ele quais benefícios que a ferramenta traria, sendo alguns deles, melhoria como as informações serão visualizadas pelos colaboradores da empresa, também ganho na produtividade dos funcionários ao obter as informações desejadas dentre outras. Inicialmente a solução de BI será aplicada somente a área comercial da empresa, mas deixando as outras áreas como projeto futuro. O gerente de projetos aprovou a solução de BI e se dispôs a ajudar na implantação do projeto e no levantamento das necessidades com a área comercial da empresa. Sem a participação dos *stakeholders* o projeto tende a falhar.

2.6.2 Levantamento das necessidades

Esta fase foi marcada por duas reuniões com o gerente de projetos. Na primeira reunião ele mostrou um sistema de BI que a empresa possui e que foi desenvolvido pelo mesmo. O gerente de projetos sugeriu replicar as informações do sistema de BI existentes na nova solução que foi implantada, sugerindo utilizar o *Data Warehouse* existente para aproveitamento das informações. A Figura 11 mostra a modelagem do *Data Warehouse* existente.

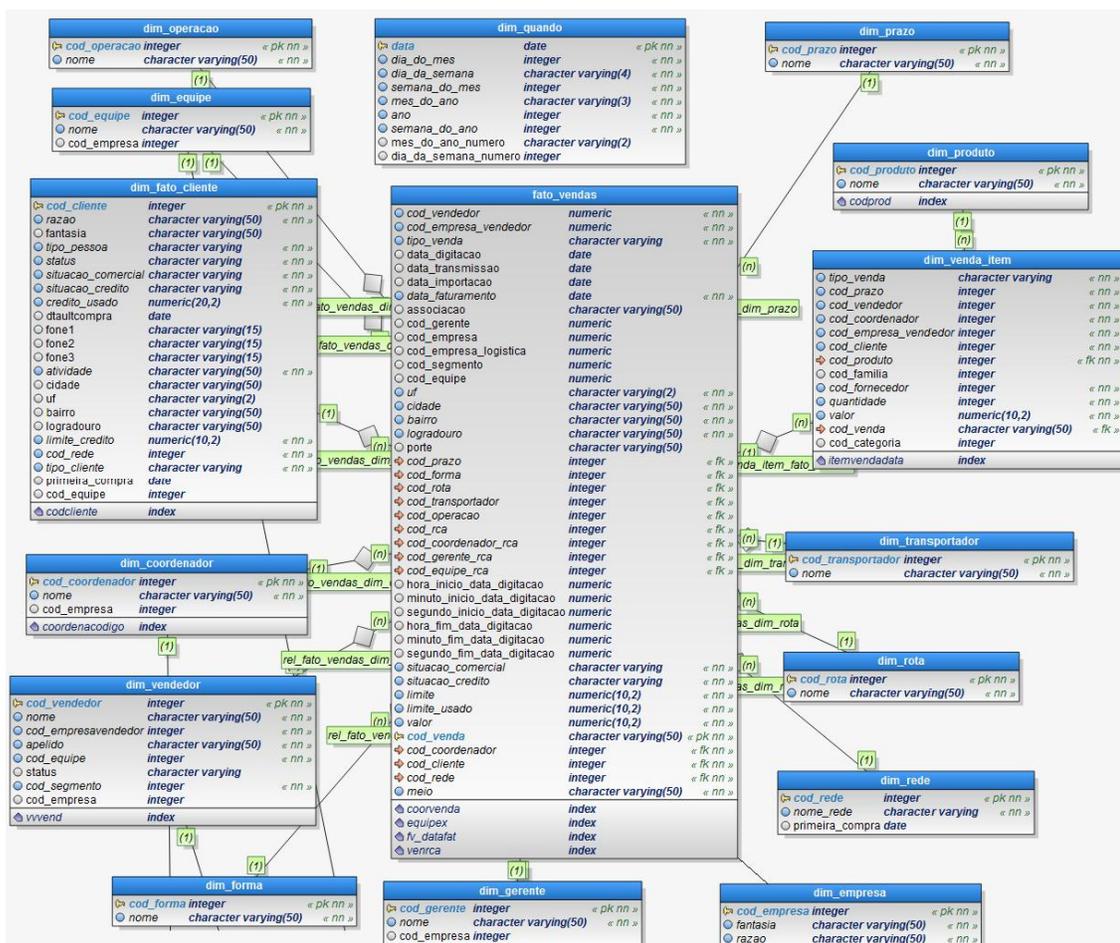


Figura 11 – Modelagem de dados do *Data Warehouse*.

A modelagem de dados que é representada na Figura 11 segue o modelo floco de neve, tendo como tabela Fato a fato_vendas onde estão todos os dados históricos de venda da empresa desde o ano de 2011. E as dimensões dessa modelagem trazem os detalhes da venda, sendo eles:

- Dim_quando – dimensão de tempo com suas hierarquias;

- Dim_prazo – dimensão de prazo de pagamento do pedido;
- Dim_produto – dimensão de informações detalhadas dos produtos;
- Dim_venda_item – dimensão de produtos que foi vendido;
- Dim_transportador – dimensão de transportador do pedido;
- Dim_rota – dimensão de rota de entrega;
- Dim_rede – dimensão de rede de clientes;
- Dim_empresa – dimensão de empresas da DPC;
- Dim_gerente – dimensão de gerente de vendas;
- Dim_forma – dimensão de forma de pagamento;
- Dim_vendedor – dimensão de vendedores;
- Dim_coordenador – dimensão de coordenadores de venda;
- Dim_fato_cliente – dimensão de dados dos clientes;
- Dim_equipe – dimensão de equipe de vendas;
- Dim_operacao – dimensão de operações de vendas;

Na segunda reunião com o gerente de projeto foi discutida quais informações a área comercial da empresa precisa, e que não estão disponíveis, e quais demoram para serem geradas pelo sistema, quando precisam estar disponíveis e em qual formato deve ser disponibilizado.

2.6.3 Mapeamento dos sistemas e dos dados

A fase de mapeamento dos sistemas e dos dados teve início após ter registrado o levantamento das necessidades, ela deu início com uma reunião contando com a presença do gerente de projetos e do gerente do departamento de TI, nessa reunião foi discutida a viabilidade das solicitações levantadas na fase anterior e foi feito o mapeamento da fonte de origem dos dados até o destino com os respectivos sistemas que efetuam a inserção dos dados.

A Figura 12 mostra a modelagem das tabelas de origem, os dados são oriundos de um sistema de informação chamado Maracanã. Nesta modelagem são tratados assuntos de meta de venda que serve para mostrar ao vendedor aonde ele precisa chegar e qual o resultado a empresa espera dele, as metas de venda nem sempre são as mesmas para cada pessoa. A seguir a uma breve descrição de cada tabela, sendo elas:

- Dpc_com_meta - são gravadas informações gerais da meta;

A Figura 13 mostra a modelagem de dados das tabelas de destino que são representadas na Figura 12, o número de tabelas representada na Figura 13 diminuiu em comparação com o número de tabelas de origem, devido os dados serem organizados de forma resumida no *Data Warehouse*. As informações de meta cadastradas continuaram sendo as mesmas do sistema de origem, no entanto os dados ficaram resumidos e repetitivos.

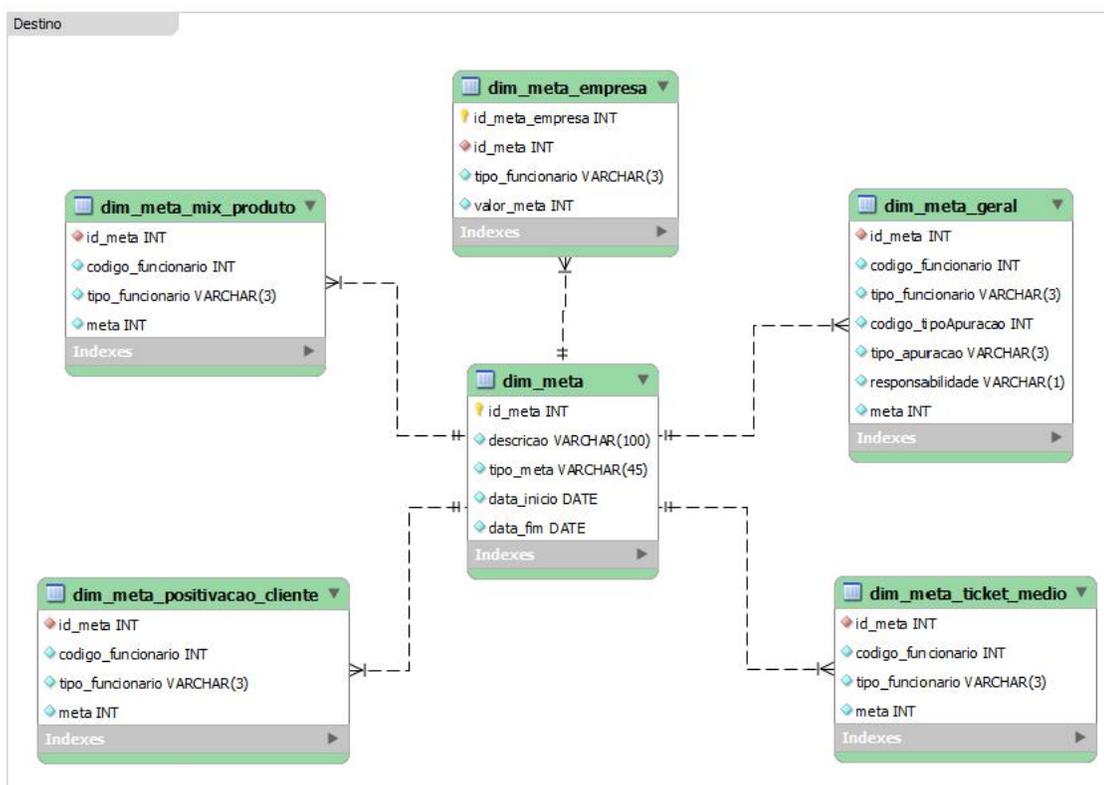


Figura 13 – Modelagem de dados das tabelas de destino.

Analisando a Figura 13 nota-se que a modelagem de dados usada é a modelagem estrela, as dimensões trabalham em conjunto com a tabela Fato_Vendas para permitir a análise das metas de venda.

2.6.4 Definir Utilizadores

Nesta fase de definir utilizadores, foram determinados quais colaboradores inicialmente poderiam usar o sistema, qual grupo cada utilizador pertence se é utilizador geral

de relatórios, analistas de dados ou gestores que decidem os objetivos. Dessa forma é feita uma revisão assegurando que os requisitos essenciais sejam incluídos desde o início.

2.6.5 Garantir a integração e qualidade dos dados

A fase de garantir a integração e qualidade dos dados é muito importante para conseguir o sucesso do projeto de BI, pois é nestas informações que os colaboradores se basearão para tomar suas decisões, e as informações devem estar mais próximas dos valores reais. As informações de venda extraídas da base de origem e que foram carregadas no *Data Warehouse* são verificadas diariamente, afim de garantir a qualidade e integração dos dados para evitar que seja questionado sobre possíveis erros.

2.6.6 Implantação

Para que os tomadores de decisões tenham acesso às informações que foram obtidas foi necessário implantar a ferramenta *Pentaho*.

O primeiro passo da implantação foi configurar o ambiente de instalação, o servidor estava com o sistema operacional *Debian Wheezy*. No servidor foi atualizada a versão do Java para a versão mais recente e foram configuradas também as variáveis de ambiente do Linux. Após a configuração o *BI Server* foi iniciado.

O segundo passo da implantação do *Pentaho*, foi utilizado a ferramenta *Pentaho Data Integration* (PDI) para fazer a integração dos dados, nesta ferramenta foi configurado a conexão com o banco de dados transacional e com o *Data Warehouse*. Com o PDI foi desenvolvido as transformações de dados que já existiam para substituir o sistema antigo, também desenvolveu as novas transformações.

A Figura 14 mostra a montagem do processo de povoamento das novas informações no *Data Warehouse*, utilizando a ferramenta *Pentaho Data Integration* (PDI), foram criadas *Views* (Anexo III) para facilitar este processo de povoamento, sendo os nomes delas: *dpcv_dw_meta*; *dpcv_dw_meta_geral*; *dpcv_dw_meta_mix_produto*; *dpcv_dw_meta_ticket_medio*; *dpcv_dw_meta_positivacao_cli*; *dpcv_dw_meta_empresa*.

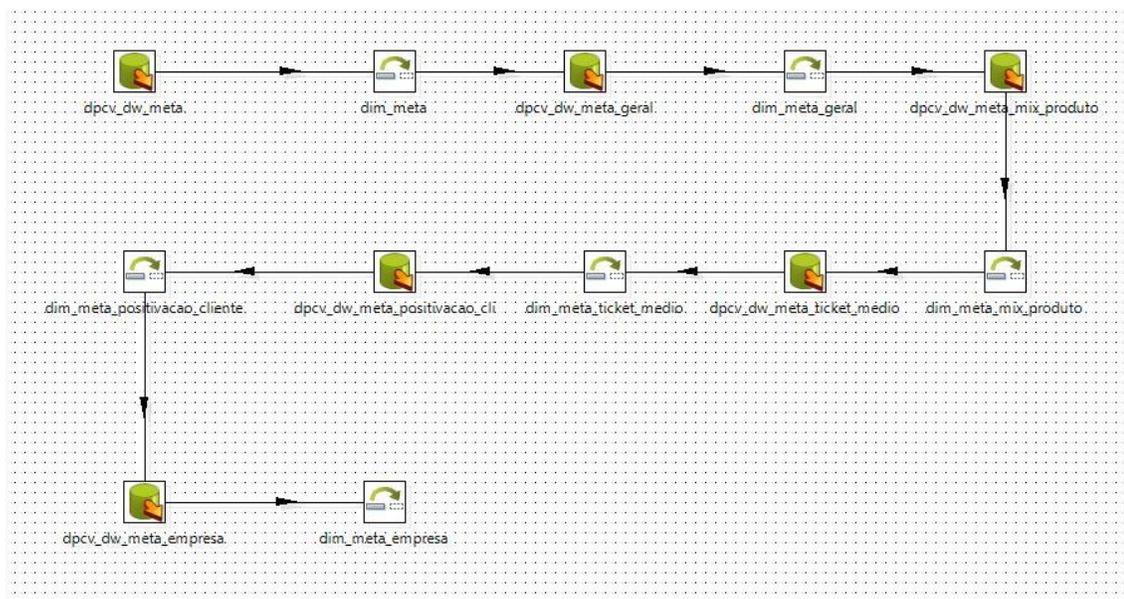


Figura 14 – Transformação de dados no PDI.

Os dados são oriundos de *Views* criadas no banco de dados transacional. Após ter criado as transformações foram criados os *Jobs* no PDI, que são processos que executam as transformações criadas e pode ser configurado para enviar e-mail após o término da execução.

O terceiro passo teve por objetivo criar as informações para os utilizadores. Para criar os cubos de informações foi usada a ferramenta *Pentaho Schema Workbench* (PSW) e pela mesma ferramenta os dados foram publicados no *BI Server*.

A Figura 15 mostra o cubo de informações sendo criado com a ferramenta *Pentaho Schema Workbench*, neste cubo a tabela Fato é a *fato_vendas* e data, cliente, segmento, forma, prazo, gerente, coordenador, equipe, vendedor, empresa, produto, família, fornecedor, UF são as dimensões do cubo e a medida deste cubo é o valor de venda que está na tabela *Fato*.

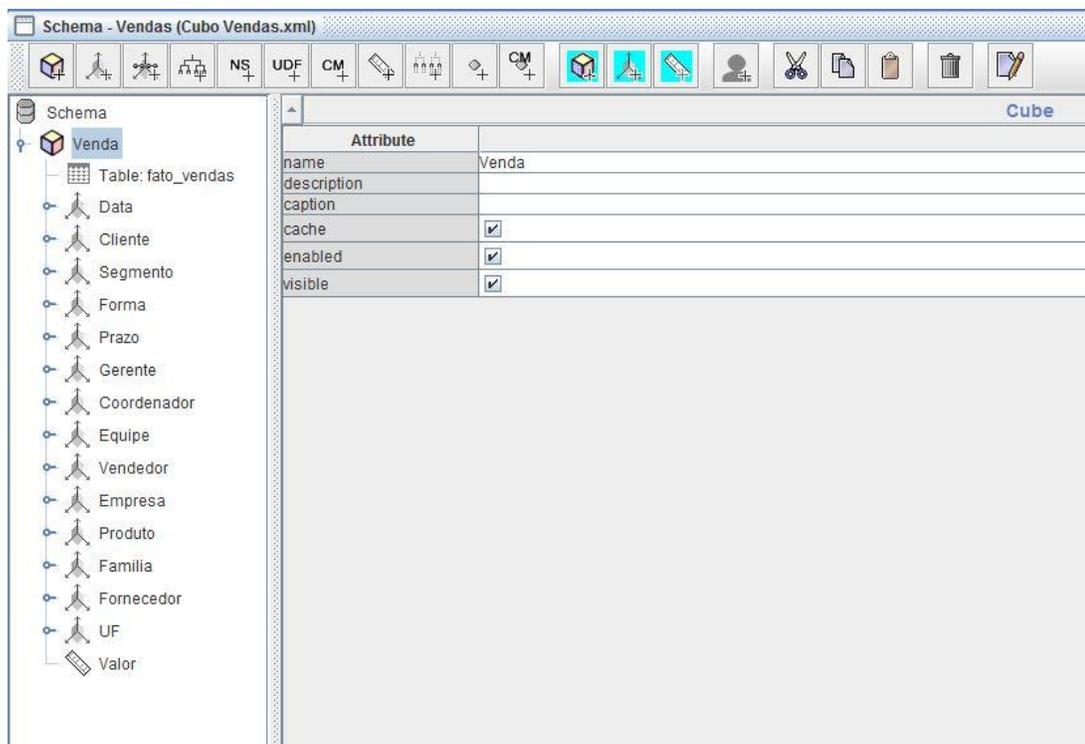


Figura 15 - Cubo de informação no PSW.

A ferramenta *Pentaho Report Designer* (PRD) foi utilizada para criar os relatórios de acordo com as necessidades dos utilizados, ao fim de cada relatório criado o mesmo é publicado no BI Server.

Na Figura 16 mostra um dos relatórios de vendas de um funcionário dos últimos seis meses em edição que foi criado pela ferramenta PRD, no relatório há parâmetros para filtrar pelo funcionário que o utilizador quiser visualizar, ele exibe também um gráfico de vendas que cuja visualização do gráfico é controlada por parâmetro. Neste relatório os valores exibidos são fictícios.

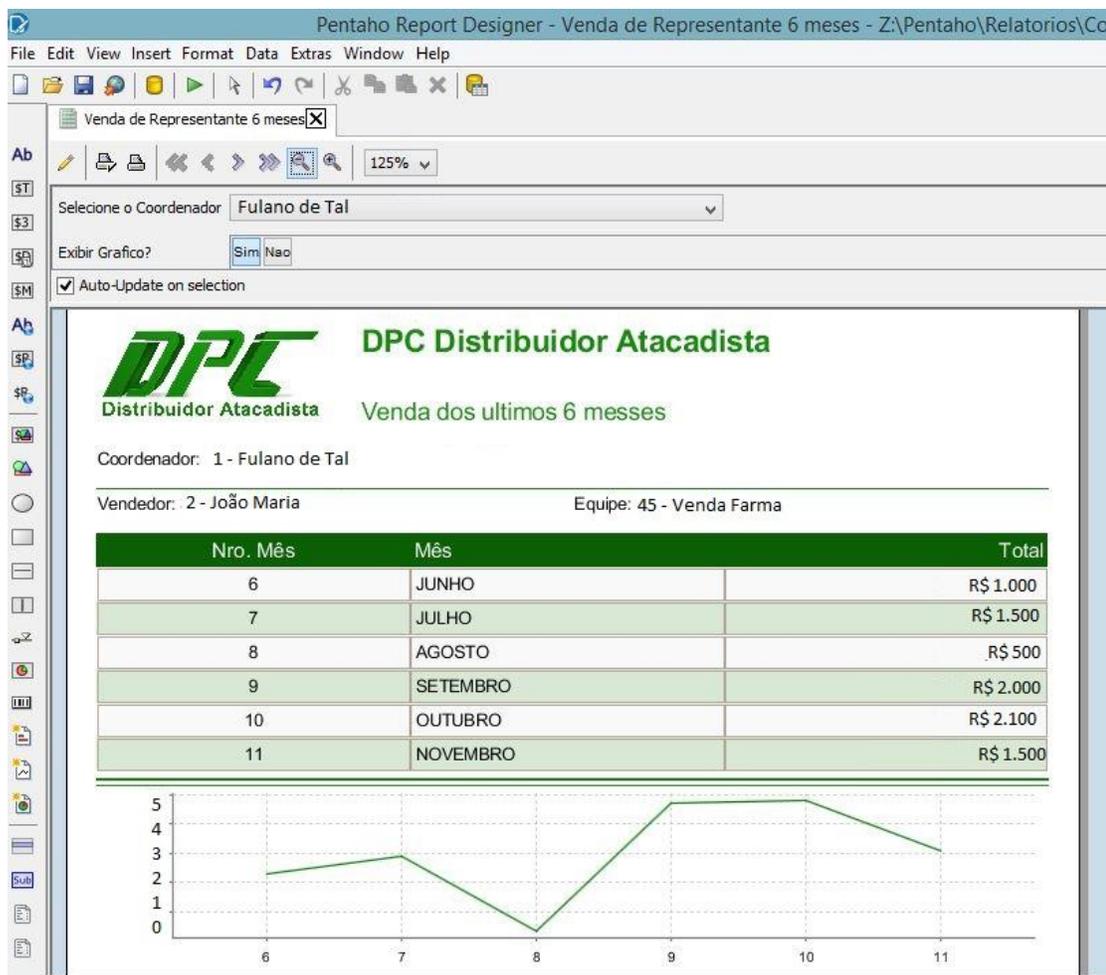


Figura 16 - Relatório sendo criado pelo PRD.

Os *dashboards* foram criados através de consultas MDX geradas pela ferramenta de análise de dados disponibilizada no BI Server, essa ferramenta é chamada de *JPivot View* e já vem no BI Server. Esta ferramenta lê os dados do cubo de informação criado pela ferramenta PSW e exibe as dimensões e medidas criadas, nela o utilizador consegue criar a visualização de dados da forma que quiser. Quando criado a visualização de dados tem uma opção na ferramenta para exibir a consulta MDX resultante da operação. A consulta é copiada e inserida em outra ferramenta disponível no BI Server para a criação de *Dashboards*.

A Figura 17 exibe um *dashboard* de comparativo de vendas da empresa mês a mês dos anos de 2013 e 2014, onde o ano 2013 está na cor verde e o ano 2014 está na cor vermelha, esse *dashboard* foi disponibilizado para alguns utilizadores e os valores exibidos são fictícios.

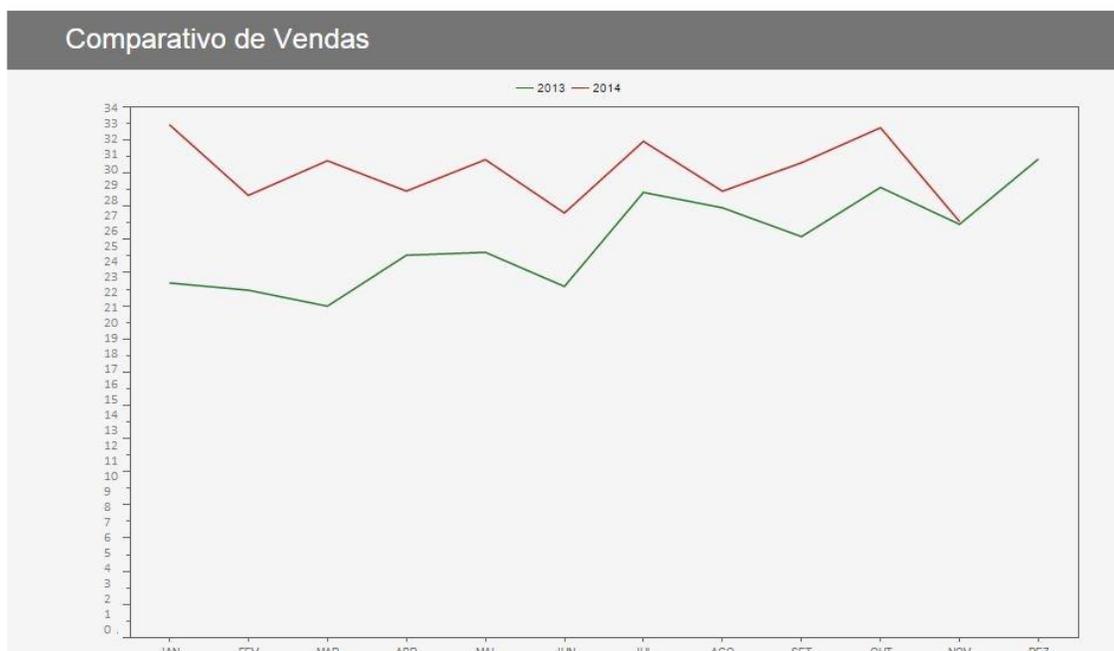


Figura 17 - *Dashboard* de comparativo de vendas.

Com todos os conteúdos criados para os utilizadores foram publicados no *BI Server*, foi preciso criar os usuário e grupos, definir quais usuários pertence a qual grupo e aplicar as permissões nos conteúdos. Os usuários e senhas foram extraídos do ambiente transacional.

Após ter criado os usuário foi feito um treinamento com os utilizadores e a equipe de suporte ensinando-os a acessar e utilizar todo o conteúdo disponibilizado no *Pentaho*, que são os relatórios, *dashboards*, ferramentas de análises de informação e ferramentas de criação de relatórios personalizados. Quando o treinamento terminou foi considerado que a ferramenta está implantada. Depois de ter implantado o *Pentaho*, foi realizado um acompanhamento com a equipe de suporte para quaisquer problemas ou dúvidas que pudessem ocorrer.

3 ANÁLISE DOS RESULTADOS

Nesta sessão serão apresentados os resultados obtidos por meio das repostas do questionário. A pesquisa contou com a participação de 9 (nove) colaboradores da empresa, sendo que todos trabalham ou participam do processo de tomada de decisões da área comercial. A análise de cada questão está separada, abordando os seus resultados e seus respectivos dados.

Com o objetivo de apresentar melhor os resultados, foram montados ilustrações a partir de gráficos e tabelas especificando detalhadamente os resultados obtidos. Abaixo de cada ilustração foi feito comentários para que não restem dúvidas com relação aos dados coletados.

3.1 ANÁLISES DAS QUESTÕES DA PESQUISA

3.1.1 Perguntas para conhecimento dos participantes

Neste grupo de questões todos os participantes responderam as questões, com a finalidade de conhecer o público.

Questão 1 – Qual a sua idade?

A primeira pergunta visou conhecer a idade dos participantes. Por meio dela, será possível verificar se a idade do participante pode influenciar nos testes de usabilidade do sistema. O Gráfico 1 e a Tabela 1 trazem os resultados dessa questão.

Tabela 1 – Respostas da Questão 1.

Respostas	Quantidade	Porcentagem
18 a 25	3	33%
26 a 35	4	44%
36 a 45	2	22%
Acima 45	0	0%

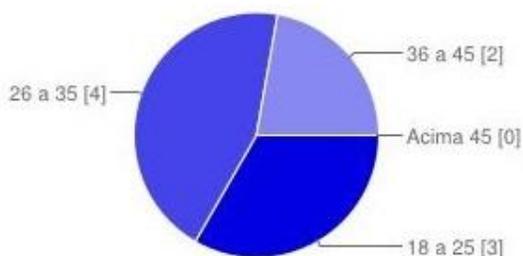


Gráfico 1 – Questão 1.

Pode-se perceber pela Tabela 1, 44% (quarenta e quatro por cento) dos participantes tinham idade entre 26 (vinte e seis) e 35 (trinta e cinco) anos, 33% (trinta e três por cento) tinham idade entre 18 (dezoito) e 25 (vinte e cinco) anos, 22% (vinte e dois por cento) tinham idade entre 36 (trinta e seis) e 45 (quarenta e cinco) anos e nenhum dos participantes era maior de 45 (quarenta e cinco) anos de idade.

Fundamentando nas respostas das demais questões do questionário, de modo geral, a idade dos participantes não é um fator que influencia no teste de usabilidade, pois as respostas foram bem diversificadas mediante a idade.

Questão 2 – Qual cargo você ocupa?

Essa questão especificamente tem o papel de relacionar o cargo que o participante ocupa com o padrão de respostas, o Gráfico 2 e a Tabela 2 mostram os resultados dessa questão.

Tabela 2 – Resultado da Questão 2.

Respostas	Quantidade	Porcentagem
Diretor	0	0%
Gerente	2	22%
Coordenador	1	11%
Colaborador	3	33%
Outro	3	33%

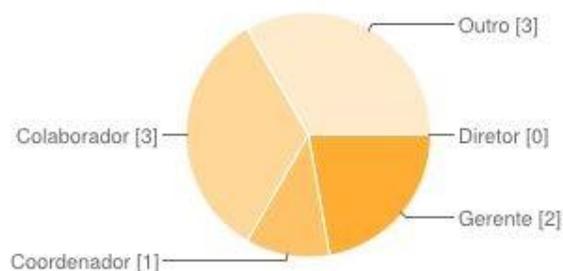


Gráfico 2 – Questão 2.

Com um empate de 33% (trinta e três por cento) dos cargos ocupados, está o colaborador e o cargo outro, no qual o cargo outro, refere-se a pessoas que prestam serviços para a empresa. 22% (vinte e dois por cento) são gerentes e 11% (onze por cento) são coordenadores que usaram a ferramenta.

Com a análise das respostas que estão na Tabela 2, é seguro afirmar que o questionário alcançou o público desejado na empresa. Pode-se concluir mediante as respostas das demais questões analisando cada cargo ocupado separadamente que, os entrevistados que ocupavam os mesmos cargos avaliaram as questões de forma parecida. No entanto não foi possível analisar as respostas do cargo de coordenador, pois apenas um participante ocupava este cargo.

Questão 3 – Com que frequência você usa o Pentaho?

Esta questão foi abordada no questionário com o intuito de descobrir com que frequência o participante usou o *Pentaho* para analisar se a frequência de uso poderia influenciar na usabilidade da ferramenta, os resultados da questão podem ser visualizados na Tabela 3 e no Gráfico 3.

Tabela 3 – Resultado da Questão 3.

Respostas	Quantidade	Porcentagem
Diariamente	7	78%
Semanalmente	0	0%
Mensalmente	2	22%
Outro	0	0%

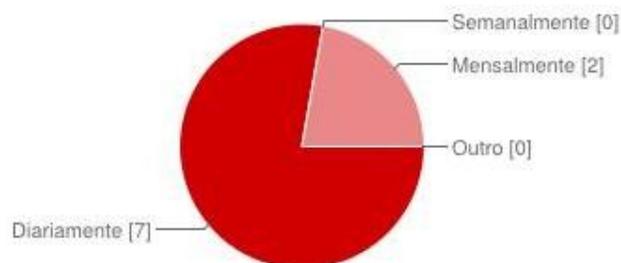


Gráfico 3 – Questão 3.

Com base nas repostas dos participantes como mostra o Gráfico 3 e a Tabela 3. A maioria dos participantes com 78% (setenta e oito por cento) usaram o *Pentaho* diariamente, 22% (vinte e dois por cento) dos participantes responderam que usaram mensalmente.

De acordo com a Tabela 3 e o Gráfico 3, de modo geral é possível inferir que o teste de usabilidade pôde ser aplicado com sucesso, pois a frequência de uso não impediu do participante avaliar a ferramenta. No entanto um dos participantes que respondeu que usou a ferramenta mensalmente não respondeu uma questão que abordava uma funcionalidade alegando não conhecê-la.

3.1.2 Questões de acesso ao *Pentaho*

Neste grupo de questões todos os participantes responderam as questões, com o objetivo de avaliar a usabilidade do *Pentaho*.

Questão 4 – Quão fácil foi acessar o Pentaho?

A quarta questão teve por finalidade avaliar se os participantes acharam fácil acessar o *Pentaho* pelo navegador. Na questão o participante tinha que escolher uma nota entre 1 (um) e 5 (cinco), no qual a nota 1 (um) ele achou muito difícil e a nota 5 (cinco) ele achou muito fácil acessar o *Pentaho*. O Gráfico 4 e a Tabela 4 exibem os valores da questão.

Tabela 4 – Resultado da Questão 4.

	Respostas	Quantidade	Porcentagem
Muito difícil	1	0	0%
	2	0	0%
	3	3	33%
	4	4	44%
Muito fácil	5	2	22%

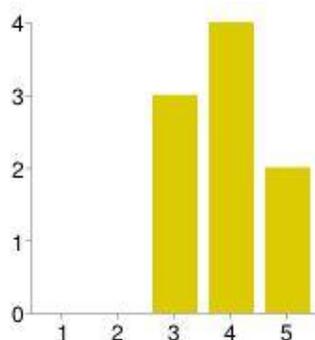


Gráfico 4 - Questão 4.

Como pode-se verificar na Tabela 4, 44% (quarenta e quatro por cento) dos participantes deram a nota 4 (quatro) para a facilidade de acesso ao *Pentaho*, 33% (trinta e três por cento) deram a nota 3 (três) e 22% (vinte e dois por cento) deram a nota 5 (cinco).

Segundo Mccall (1977) a operabilidade é uma medida da facilidade de operação de um programa. Dessa forma é possível afirmar através da maioria das respostas, que o acesso ao *Pentaho* foi fácil sem causar nenhum transtorno a eles atendendo a este conceito.

Questão 5 – O *Pentaho* ao seu ver é seguro?

A quinta pergunta tem como objetivo avaliar a segurança do *Pentaho*, para saber se ele é seguro ao ver do participante. Na pergunta o participante tinha que dar uma entre 1 (um) e 5 (cinco), no qual a nota 1 (um) ele achou pouco seguro e a nota 5 (cinco) achou muito seguro acessar o *Pentaho*. O Gráfico 5 e a Tabela 5 mostram o resultado da questão.

Tabela 5 – Resultado da Questão 5.

	Respostas	Quantidade	Porcentagem
Pouco seguro	1	0	0%
	2	0	0%
	3	0	0%
	4	6	67%
Muito seguro	5	3	33%

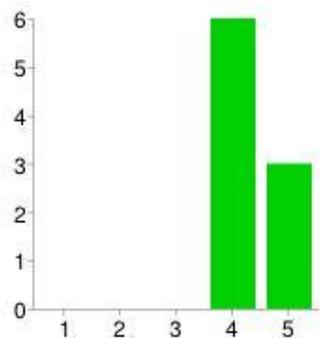


Gráfico 5 – Questão 5.

Percebe-se que a maioria dos participantes deram a nota 4 (quatro) para a segurança do *Pentaho* ficando com 67% (sessenta e sete por cento), 33% (trinta e três por cento) deram a nota 5 (cinco).

De acordo com o Gráfico 5, a maioria dos participantes acharam o *Pentaho* seguro. Contudo, não é possível inferir nenhuma conclusão devido à segurança de um *Software* não ser uma métrica de usabilidade.

3.1.3 Questões de avaliação da interface do *Pentaho*

Neste grupo de questões todos os participantes puderam responder as questões, com o intuito de avaliar se a interface é agradável e intuitiva de usar.

Questão 6 – Como você avaliaria a interface do *Pentaho*?

A pergunta de número 6 (seis) tem como finalidade mensurar a interface do *Pentaho*, para saber se ela é agradável e intuitiva, nela o participante tinha que escolher uma nota entre 1 (um) e 5 (cinco) para avaliar, sendo que a nota 1 (um) a interface é muito confusa e a nota 5 (cinco) refere-se a uma interface é intuitiva e agradável. O Gráfico 6 e a Tabela 6 exibem o resultado da questão.

Tabela 6 – Resultado da Questão 6.

	Respostas	Quantidade	Porcentagem
Muito confusa	1	0	0%
	2	0	0%
	3	3	33%

	4	5	56%
Intuitiva e agradável	5	1	11%

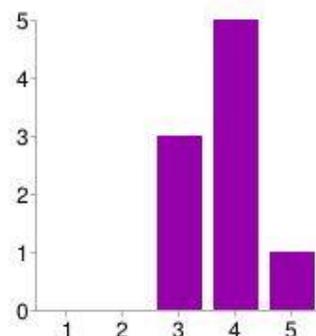


Gráfico 6 – Questão 6.

A Tabela 6 mostra que 56% (cinquenta e seis por cento) dos participantes deram a nota 4 (quatro) para a interface do *Pentaho*, 33% (trinta e três por cento) deram a nota 3, 11% (onze por cento) deram a nota 5 (cinco).

Segundo Mccall (1977), o treino de um *Software* se refere ao quanto um sistema auxilia o usuário na familiarização de um sistema. A maioria dos entrevistados com 56% (cinquenta e seis por cento) deram a nota 4, dessa forma é possível inferir que a interface do *Pentaho* intuitiva e agradável de utilizar ao ver dos participantes atendendo ao critério de treinamento.

Questão 7 – Quão fácil foi encontrar a informação que você deseja?

A sétima pergunta, tem por finalidade descobrir se o participante achou facilmente a informação que procurava, nela o participante dava uma nota entre 1 (um) e 5 (cinco) para avaliar a questão, sendo que a nota 1 (um) refere-se a pouco fácil encontrar a informação e a nota 5 (cinco) refere-se a muito fácil. O Gráfico 7 junto com a Tabela 7 exibem o resultado da questão.

Tabela 7 – Resultado da Questão 7.

	Respostas	Quantidade	Porcentagem
Pouco fácil	1	0	0%
	2	0	0%
	3	1	11%
	4	5	56%
Muito fácil	5	3	33%

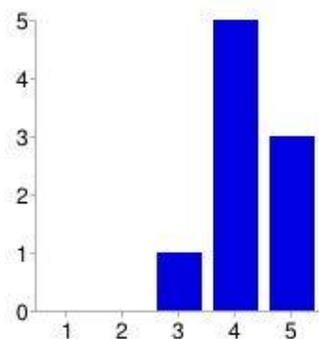


Gráfico 7 – Questão 7.

De acordo com a Tabela 7, 56% (cinquenta e seis por cento) dos participantes deram a nota 4 (quatro) para a interface do *Pentaho*, 33% (trinta e três por cento) deram a nota 5, 11% (onze por cento) deram a nota 3 (cinco).

Para Nielsen (1994) a métrica de reconhecimento em vez de recordação aborda que o sistema deve evitar que o usuário precise pensar em como executar uma tarefa. Pode-se inferir pelas respostas dos participantes, que a maioria achou facilmente as informações que procuravam no *Pentaho*, concluindo que o *Pentaho* atende ao presente critério heurístico.

Questão 8 – Há muitas maneiras de encontrar a informação que você deseja?

A pergunta de número 8 (oito) tem por finalidade descobrir se havia muitas maneiras do participante encontrar a informação que precisava. O participante escolhia uma nota entre 1 (um) e 5 (cinco) para avaliar a questão, a nota 1 (um) refere-se havia poucas maneiras, a nota 5 (cinco) refere-se que havia muitas maneiras. O Gráfico 8 e a Tabela 8 mostram o resultado da questão.

Tabela 8 – Resultado da Questão 8.

	Respostas	Quantidade	Porcentagem
Poucas maneiras	1	1	11%
	2	1	11%
	3	5	56%
	4	0	0%
Muitas maneiras	5	2	22%

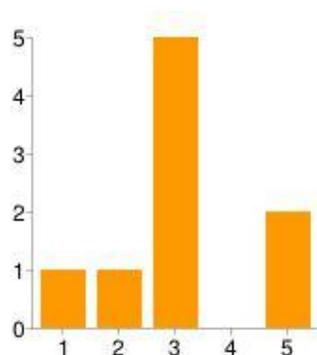


Gráfico 8 – Questão 8.

A Tabela 8 mostra que 56% (cinquenta e seis por cento) dos participantes deram a nota 3 (três), 22% (vinte e dois por cento) deram a nota 5 (cinco), 11% (onze por cento) deram a nota 1 e 2.

Segundo Mccall (1977) a operabilidade é facilidade de operação de um programa. Embora de acordo com as respostas obtidas por meio do questionário, tenha sido visualizado um nível médio de possibilidades para encontrar a informação que precisava, foi possível inferir que os participantes não tinham muitos meios de encontrar as informações que precisavam, mas o *Pentaho* atendeu a este conceito.

Questão 9 – Quão fácil foi se desorientar no sistema?

A nona pergunta tem por objetivo descobrir se o participante se perdia no sistema com facilidade. Na questão o participante escolhia uma nota de 1 (um) a 5 (cinco), a nota 1 (um) o mesmo achou fácil de perder no sistema, e a nota 5 (cinco) foi difícil se perder no sistema. O Gráfico 9 e a Tabela 9 mostram o resultado da questão.

Tabela 9 – Resultado da questão 9.

	Respostas	Quantidade	Porcentagem
Fácil	1	0	0%
	2	0	0%
	3	4	44%
	4	2	22%
Difícil	5	3	33%

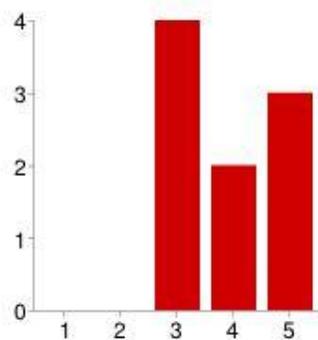


Gráfico 9 – Questão 9.

Como pode perceber na Tabela 9, 44% (quarenta e quatro por cento) dos participantes deram a nota 3 (três), 33% (trinta e três por cento) deram a nota 5 (cinco), e 22% (vinte e dois por cento) deram a nota 4 (quatro).

Assim sendo, é visível por meio das respostas dos participantes que a maioria deram a nota 3, sendo possível inferir que não é fácil se desorientar na interface do *Pentaho*. Segundo Jakob Nielsen (1994) a métrica de estética e *design* minimalista mostra ao usuário o que ele apenas precisa saber. Logo é possível concluir que a ferramenta atende ao presente critério heurístico.

3.1.4 Questões de avaliação dos relatórios do *Pentaho*

Neste grupo de questões todos os participantes puderam responder questões, com o intuito de avaliar a usabilidade dos relatórios fornecidos pelo *Pentaho*. Um participante deixou de responder a Questão 11 alegando não conhecer a funcionalidade abordada, um outro participante também não respondeu a Questão 13 alegando que não poderia comparar se a informação era exata pois não tinha onde comparar.

Questão 10 – Quão fácil foi usar os relatórios?

A questão 10 tem por objetivo descobrir o quanto foi fácil para o participante usar os relatórios, nela o participante tinha que dar uma nota entre 1 (um) e 5 (cinco), a nota 1 (um) refere-se que foi pouco fácil utilizar os relatórios, a nota 5 (cinco) foi muito fácil. Os resultados podem ser visualizados pela Tabela 10 e pelo Gráfico 10.

Tabela 10 – Resultado da questão 10.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	0	0%
	4	5	56%
Muito	5	4	44%

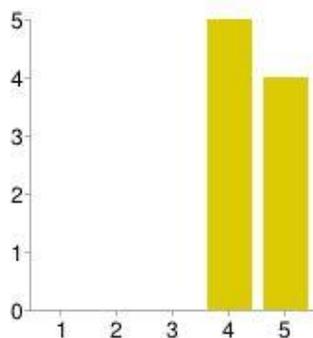


Gráfico 10 – Questão 10.

De acordo com a Tabela 10, 56% (cinquenta e seis por cento) deram a nota 4 (quatro), 44% (quarenta e quatro por cento) deram a nota 5 (cinco) para a facilidade de uso dos relatórios.

Para Sommerville (2010) para que haja usabilidade um *software* precisa ser usável. Por intermédio das repostas desta questão, é possível inferir que a maioria dos participantes acharam fácil utilizar os relatórios atendendo ao conceito pregado.

Questão 11 – Como você avaliaria o agendamento de relatórios?

A pergunta de número 11 (onze) calcula a qualidade do agendamento de relatórios. O participante dava uma nota entre 1 (um) e 5 (cinco), quanto mais próxima a nota fosse de 1 (um) refere-se que o participante achou ruim o agendamento dos relatórios, quanto mais de 5 (cinco) a nota o participante achou excelente. O Gráfico 11 junto a Tabela 11 mostram o resultado da questão.

Tabela 11– Resultado da Questão 11.

	Respostas	Quantidade	Porcentagem
Ruim	1	0	0%
	2	0	0%
	3	3	33%

Excelente	4	1	11%
	5	4	44%

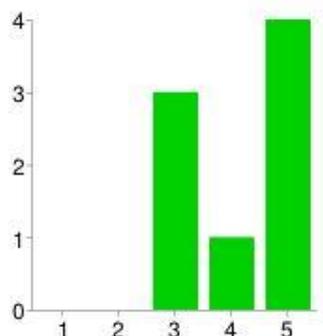


Gráfico 11 – Questão 11.

Como se pode verificar na Tabela 11 a maioria dos participantes deram a nota 5 ficando com 44% (quarenta e quatro por cento), 33% (trinta e três por cento) ficou a nota 3 (três) e 11% (onze por cento) deram a nota 4 (quatro).

Embora algumas as respostas dessa questão tenham sido acirradas, foi possível perceber que os entrevistados avaliaram bem o agendamento de relatórios. Um participante não respondeu esta questão argumentando não conhecer a funcionalidade. Pode-se inferir que a presente questão atendeu o conceito de operabilidade de *software* de Mccall (1977).

Questão 12 – Os formatos disponibilizados dos relatórios foram úteis?

A décima segunda questão tem por objetivo descobrir se os formatos disponibilizados dos relatórios foram úteis, nela o participante tinha que dar uma nota entre 1 (um) e 5 (cinco), a nota 1 (um) refere-se que participante achou os formatos dos relatórios pouco úteis, a nota 5 (cinco) refere-se que o participante achou muito útil os formatos dos relatórios. No Gráfico 12 e na Tabela 12 é possível visualizar os resultados da questão.

Tabela 12 – Respostas da Questão 12.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	1	11%
	4	0	0%
Muito	5	7	78%

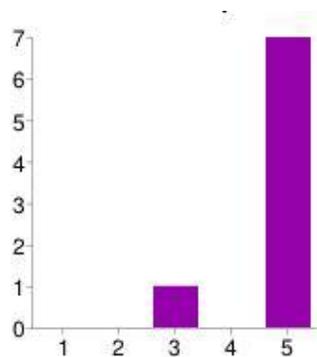


Gráfico 12 – Questão 12.

A maioria dos participantes que responderam o questionário deram a nota 5 (cinco) para os formatos disponibilizados dos relatórios que ficou com 78% (setenta e oito por cento) e 11% (onze por cento) dos participantes deram 3 (três).

Contudo apesar das respostas avaliarem bem a questão, não foi possível avaliar a usabilidade de um *software* através dessa questão.

Questão 13 – As informações exibidas pelos relatórios são exatas?

A questão de número 13 (treze) tem por finalidade avaliar se as informações disponibilizadas pelos relatórios são exatas, ou seja, se os valores batem com a realidade. Na questão o participante dava uma nota entre 1 (um) e 5 (cinco) para a questão, a nota 1 (um) refere-se que as informações exibidas foram pouco exatas, a nota 5 (cinco) refere-se que as informações são totalmente exatas.

Tabela 13 – Respostas da Questão 13.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	1	11%
	4	3	33%
Totalmente	5	4	44%

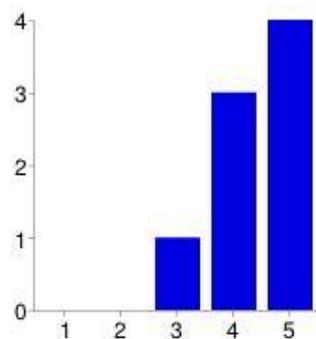


Gráfico 13 – Questão 13.

Como se pode verificar na Tabela 13 e no Gráfico 13, dos participantes que responderam o questionário, 4 (quatro) deram a nota 5 (cinco) que ficou com 44% (quarenta e quatro por cento), 3 (três) deram a 4 (quatro) que ficando com 33% (trinta e três por cento) e 11% (onze por cento) deram a nota 3 (três).

As respostas obtidas na questão comprovam um alto índice de aceitação na veracidade da informação. Um participante não respondeu a questão devido a ela não conseguir comparar a exatidão das informações por não ter onde comparar. Segundo Mccall (1977) a comunicatividade de um *software* deve prover saídas que podem ser interpretadas pelo usuário, é possível inferir que o *software* atendeu a este conceito.

Questão 14 – Os relatórios destacam as informações importantes?

A décima quarta questão tem por finalidade mensurar se as informações importantes são destacadas nos relatórios. O participante dava uma nota entre 1 (um) e 5 (cinco), a nota 1 (um) refere-se que as informações importantes são pouco destacadas, a nota 5 (cinco) refere-se que as informações importantes são muito destacadas.

Tabela 14 – Respostas da Questão 14.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	0	0%
	4	4	44%
Muito	5	5	56%

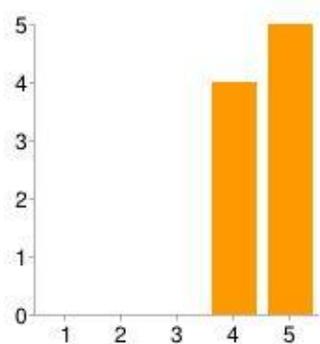


Gráfico 14 – Questão 14.

Analisando os dados da Tabela 14 e do Gráfico 14, percebe-se que a maioria dos participantes deram a nota 5 (cinco) que ficou com 56% (cinquenta e seis por cento), 44% (quarenta e quatro por cento) dos participantes deram 4 (quatro).

De acordo com a tabela acima, é possível inferir por meio das respostas que os relatórios destacam as informações importantes para o usuário. Porém não é possível avaliar a usabilidade de um *software* através dessa questão.

Questão 15 – Na sua opinião quão úteis foram os relatórios?

A pergunta de número 15 (quinze) tem por finalidade avaliar a satisfação do participante sobre aos relatórios. O participante tinha que escolher uma nota entre 1 (um) e 5 (cinco) para avaliar a questão, a nota 1 (um) refere-se que os relatórios foram pouco úteis, a nota 5 (cinco) refere-se eles foram muito úteis. O Gráfico 15 e a Tabela 15 mostra o resultado da questão.

Tabela 15 – Respostas da Questão 15.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	0	0%
	4	0	0%
Muito	5	9	100%

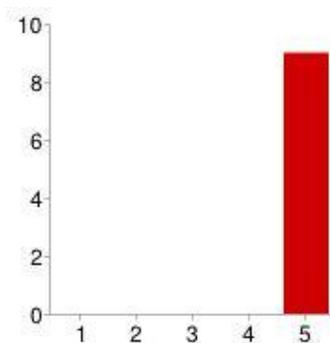


Gráfico 15 – Questão 15.

Como pode perceber no Gráfico 15, mostra que todos os participantes que responderam o questionário acharam os relatórios muito úteis.

Para McCall (1977) a comunicabilidade de um *software* deve prover saídas que podem ser interpretadas pelo usuário. Por meio das respostas é notório o alto índice de aceitabilidade dos participantes em relação à utilidade dos relatórios, logo é possível inferir que o *software* atendeu a este conceito.

3.1.5 Questões de avaliação dos *dashboards* do *Pentaho*

Neste grupo de questões apenas cinco participantes responderam, devido à disponibilidade de conteúdo mais específico para eles. Este grupo de questões tem por finalidade avaliar a usabilidade dos *dashboards*

Questão 16 – Quão fácil é usar os *Dashboards*?

A décima sexta pergunta tem por finalidade avaliar a facilidade de uso dos *Dashboards*. O participante tinha que escolher uma nota entre 1 (um) e 5 (cinco) para avaliar a questão, quanto mais próxima a nota for 1 (um) refere-se que os *Dashboards* foram pouco fáceis de usar, e quanto mais próximo da nota 5 (cinco) refere-se eles foram muito fáceis de usar. O Gráfico 16 junto com a Tabela 16 exibem o resultado da questão.

Tabela 16 – Respostas da Questão 16.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	0	0%

Muito	4	3	33%
	5	2	22%

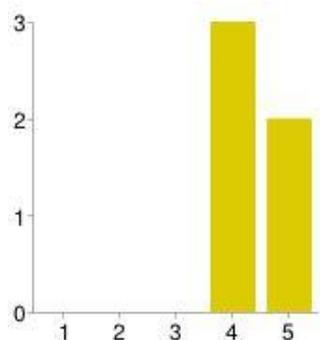


Gráfico 16 – Questão 16.

Como pode perceber no Gráfico 16 e na Tabela 16, 33% (trinta e três por cento) dos participantes deram a nota 4 (quatro) para a facilidade de uso dos *Dashboards*, 22% (vinte e dois por cento) deram a nota 5 (cinco).

Pode-se perceber de acordo com as respostas obtidas nesta questão, que há um índice alto de aceitação em relação à usabilidade dos *dashboards*. Segundo Mccall (1977) a operabilidade é uma medida da facilidade de operar um programa. É possível afirmar que segundo a opinião dos participantes que o conceito de operabilidade foi atendido.

Questão 17 – As informações exibidas nos painéis foram importantes e úteis?

A questão de número 17 (dezesete) tem por finalidade descobrir se as informações disponibilizadas pelos *Dashboards* foram importantes e úteis para o participante, ele tinha que dar uma nota entre 1 (um) e 5 (cinco) para avaliar a questão, a nota 1 (um) refere-se que as informações exibidas pelos *Dashboards* foram pouco úteis e importantes, a nota 5 (cinco) refere-se eles foram muito úteis e importantes. O Gráfico 17 e a Tabela 17 mostra os valores da questão.

Tabela 17 – Respostas da Questão 17.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	0	0%
	4	1	11%
Muito	5	4	44%

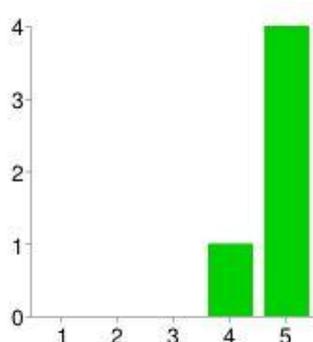


Gráfico 17 – Questão 17.

Pode-se perceber pela Tabela 17 que 4 (quatro) participantes deram a nota 5 (cinco) ficando com 44% (quarenta e quatro por cento), 11% (onze por cento) dos participantes que responderem o questionário deram a nota 4 (quatro). A maior parte dos participantes demonstrou que as informações disponibilizadas pelos *dashboards* foram úteis e importantes.

Segundo Mccall (1977) a comunicatividade de um *software* deve prover saídas que podem ser interpretadas pelo usuário. É possível inferir por meio das respostas que o *software* atendeu a este conceito. Assim sendo é possível inferir que a comunicação do sistema de BI com o usuário foi bem aceita atendendo ao conceito de comunicatividade.

3.1.6 Questões de avaliação da análise de informações do *Pentaho*

Neste grupo de questões apenas alguns participantes responderam, devido à disponibilidade de conteúdo para eles.

Questão 18 – Quão fácil foi achar as ferramentas de análise?

A questão de número 18 (dezoito) tem por finalidade avaliar a facilidade para acessar as ferramentas de análise. O participante tinha que dar uma nota entre 1 (um) e 5 (cinco) para avaliar a questão, a nota 1 (um) refere-se que foi pouco fácil achar as ferramentas, a nota 5 (cinco) refere-se que foi muito fácil achar. Os resultados podem ser visualizados pela Tabela 18 e pelo Gráfico 18.

Tabela 18 – Respostas da Questão 18.

Respostas	Quantidade	Porcentagem
-----------	------------	-------------

Pouco	1	0	0%
	2	0	0%
	3	0	0%
	4	3	33%
Muito	5	2	22%

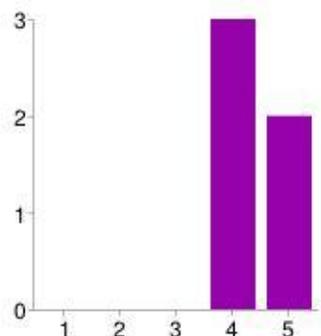


Gráfico 18 – Questão 18.

Como pode perceber pela Tabela 18 e pelo Gráfico 18, 33% (trinta e três por cento) dos participantes deram a nota 4 (quatro) e 22% (vinte e dois por cento) deram a nota 5 (cinco). Todos os participantes que responderam as questões acharam fáceis de encontrar as ferramentas de análise de informação disponibilizada pelo *Pentaho*.

Para Nielsen (1994) a métrica de reconhecimento em vez de recordação aborda que o sistema deve evitar que o usuário precise pensar em como executar uma tarefa. Por meio das respostas desta questão é possível inferir que o *software* proporciona facilidade em achar as ferramentas de análise atendendo assim a heurística mencionada.

Questão 19 – É fácil criar uma análise de informação?

A questão de número 19 (dezenove) tem por objetivo examinar a facilidade para criar análises de informações, na questão o participante tinha que dar uma nota de 1 (um) a 5 (cinco) para avaliá-la, a nota 1 (um) refere-se que foi pouco fácil criar as análises, a nota 5 (cinco) refere-se que foi muito fácil. Os resultados podem ser visualizados pelo Gráfico 19 e pela Tabela 19.

Tabela 19 – Respostas da Questão 19.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%

	3	0	0%
	4	3	33%
Muito	5	2	22%

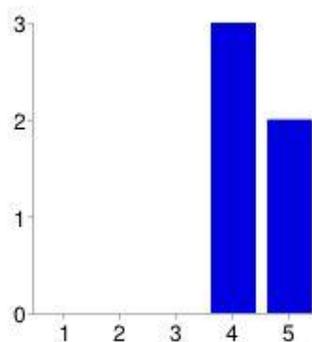


Gráfico 19 – Questão 19.

Com base nas respostas dos participantes, pode-se perceber que, 33% (trinta e três por cento) dos participantes deram a nota 4 (quatro) e 22% (vinte e dois por cento) deram a nota 5 (cinco). Evidentemente a construção de uma análise de informação não é fácil, mas por meio das respostas dos participantes foi possível perceber que a usabilidade desta ferramenta foi bem aceita.

Segundo Mccall (1977) a operabilidade é a facilidade de operar um *software*. Por meio das respostas desta questão é possível inferir que o conceito de operabilidade foi atendido.

Questão 20 – Os cubos possuem informações importantes e úteis?

A questão de número 20 (vinte) tem por objetivo avaliar as informações dos cubos, nela o participante tinha que dar uma nota entre 1 (um) e 5 (cinco) para avaliar a questão, a nota 1 (um) refere-se que as informações exibidas pelos cubos foram pouco úteis e importantes, a nota 5 (cinco) refere-se eles foram muito úteis e importantes. O Gráfico 20 e a Tabela 20 exibem o resultado da questão.

Tabela 20 – Respostas da Questão 20.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	0	0%
	4	2	22%
Muito	5	3	33%

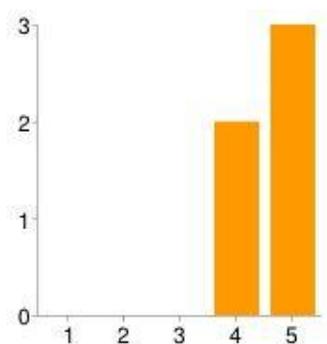


Gráfico 20 – Questão 20.

A maioria dos participantes acharam que as informações exibidas pelos cubos de informações foram boas, a Tabela 20 mostra os resultados, 33% (trinta e três por cento) dos participantes deram a nota 4 (quatro) e 22% deram a nota 5 (cinco). Por meio das respostas dos participantes os cubos de informações disponibilizadas foram úteis e importantes, logo é possível perceber que a usabilidade da ferramenta foi bem aceita.

Segundo Mccall (1977) a comunicatividade de um software deve prover saídas que podem ser interpretadas pelo usuário. Logo é possível inferir que o software atendeu a este conceito.

Questão 21 – Quão fácil foi criar relatórios personalizados?

A vigésima primeira questão teve por finalidade mensurar a facilidade de criar relatórios personalizados. Na questão o participante tinha que dar uma nota de 1 (um) a 5 (cinco), a nota 1 (um) refere-se que foi pouco fácil criar os relatórios, a nota 5 (cinco) refere-se que foi muito fácil. O Resultado pode ser visualizado na Tabela 21 e no Gráfico 21.

Tabela 21 – Respostas da Questão 21.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	1	11%
	4	2	22%
Muito	5	2	22%

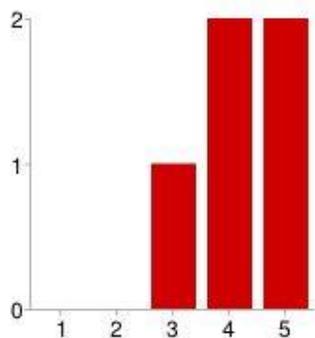


Gráfico 21 – Questão 21.

Como se pode verificar na Tabela 21 e no Gráfico 21, criar relatórios personalizados teve uma boa aceitação, 22% (vinte e dois por cento) dos participantes deram a nota 4 (quatro) e 5 (cinco) e 11% (onze por cento) deram a nota 3 (três). Fundamentando nas repostas da questão a facilidade de criar relatórios personalizados foi bem aceita pelos participantes.

Segundo Mccall (1977) a operabilidade é uma medida da facilidade de operação de um programa. A opinião dos participantes convergiu com o seu conceito, reafirmando a usabilidade do *software*.

3.1.7 Questões de avaliação geral do *Pentaho*

Neste grupo de questões todos os participantes responderam as questões, com a finalidade de avaliar o *Pentaho* de forma geral.

Questão 22 – As informações disponibilizadas são confiáveis?

A questão de número 22 (vinte e dois) tem por finalidade avaliar se as informações disponibilizadas pelo *Pentaho* são confiáveis. O participante tinha que escolher uma nota de 1 (um) a 5 (cinco) para avaliar a questão, a nota 1 (um) refere-se que a informação nunca foi confiável, a nota 5 (cinco) refere-se que sempre foi confiável. O Resultado da questão pode ser visualizado na Tabela 22 e no Gráfico 22.

Tabela 22 – Respostas da Questão 22.

	Respostas	Quantidade	Porcentagem
Nunca	1	0	0%
	2	0	0%
	3	1	11%

	4	3	33%
Sempre	5	5	56%

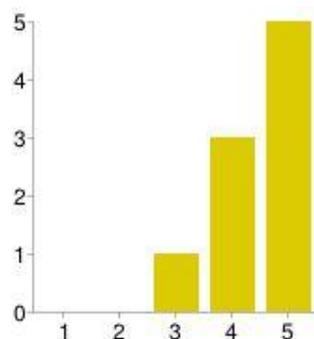


Gráfico 22 – Questão 22.

Pode-se verificar na Tabela 22, 56% (cinquenta e seis por cento) dos participantes deram a nota 5 (cinco), 33% (trinta e três por cento) deram a nota 4 (quatro) e 11% (onze por cento) deram a nota 3 (três). Com a análise das respostas desta questão através da Tabela 22, é seguro afirmar que as informações disponibilizadas pelo *Pentaho* são confiáveis.

Segundo Mccall (1977) a comunicatividade de um software deve prover saídas que podem ser interpretadas pelo usuário. De acordo com as respostas dos participantes é possível inferir que o software atendeu a este conceito.

Questão 23 – Na sua opinião o tempo de espera para visualizar as informações foi satisfatório?

A questão de número 23 (vinte e três) tem por objetivo avaliar se o tempo de espera para visualizar as informações. O participante tinha que dar uma nota de 1 (um) a 5 (cinco) para avaliar a questão, a nota 1 (um) refere-se que foi pouco satisfatório o tempo de espera, a nota 5 (cinco) refere-se que o tempo de espera foi muito satisfatório.

Tabela 23 – Respostas da Questão 23.

	Respostas	Quantidade	Porcentagem
Pouco	1	0	0%
	2	0	0%
	3	0	0%
	4	6	67%
Muito	5	3	33%

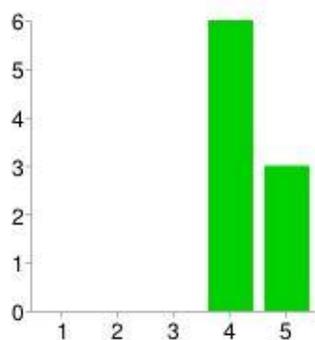


Gráfico 23 – Questão 23.

O Gráfico 23 e a Tabela 23 exibem os valores. Como pode-se verificar na Tabela 23, 67% (sessenta e sete por cento) dos participantes deram a nota 4 (quatro), 33% (trinta e três por cento) deram a nota 5 (cinco). De modo geral e de acordo com a Tabela 23, é possível observar que o tempo de espera para visualizar as informações foram satisfatória para os participantes.

Segundo Mccall (1977) a operabilidade é uma medida da facilidade de operação de um programa. Com base nas respostas dos entrevistados foi possível concluir que este conceito de operabilidade foi atingindo.

Questão 24 – Quão fácil foi achar os seus arquivos?

A vigésima quarta questão teve por objetivo avaliar se o participante encontrou facilmente os arquivos que lhe pertenciam. O participante tinha que escolher uma nota de 1 (um) a 5 (cinco) para avaliar a questão, quanto menor a nota fosse mais dificuldade o participante teve para encontrar os arquivos e quanto maior fosse mais fácil foi de encontrar. O Gráfico 24 e a Tabela 24 demonstram os resultados da questão.

Tabela 24 – Respostas da Questão 24.

	Respostas	Quantidade	Porcentagem
Péssimo	1	0	0%
	2	0	0%
	3	2	22%
	4	4	44%
Excelente	5	3	33%

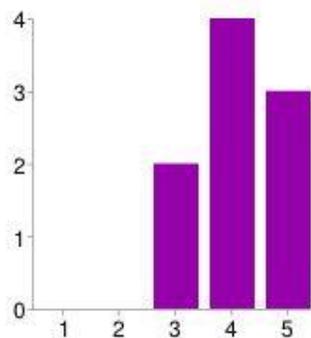


Gráfico 24 – Questão 24.

Os participantes acharam facilmente os seus arquivos como o Gráfico 24, 44% (quarenta e quatro por cento) deram a nota 4 (quatro) para a facilidade de achar os arquivos, 33% (trinta e três) deram a nota 5 (cinco) e 22% (vinte e dois por cento) deram a nota 3 (três). De acordo com o Gráfico 24, é visível que a ampla maioria dos entrevistados acharam facilmente os arquivos que procuravam.

Segundo Mccall (1977) a operabilidade é uma medida da facilidade de operação de um programa. Por meio dos resultados da questão foi possível concluir que este conceito de operabilidade foi atingindo.

Questão 25 – Na sua opinião qual fácil foi navegar entre os dados?

A questão de número 25 (vinte e cinco) tem por finalidade mensurar a facilidade de o participante navegar entre os dados que estão abertos. O participante avaliava a questão dando uma nota entre 1 (um) e 5 (cinco), a nota 1 (um) refere-se que foi péssimo para navegar nos dados abertos, a nota 5 (cinco) refere-se que foi excelente. O Gráfico 25 e a Tabela 25 exibem os valores.

Tabela 25 – Respostas da Questão 25.

	Respostas	Quantidade	Porcentagem
Péssimo	1	0	0%
	2	0	0%
	3	0	0%
	4	6	67%
Excelente	5	3	33%

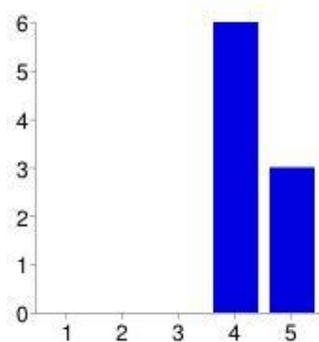


Gráfico 25 – Questão 25.

Como pode perceber na Tabela 25, 67% (sessenta e sete por cento) dos participantes deram a nota 4 (quatro), 33% (trinta e três por cento) deram a nota 5 (cinco). Assim sendo, é visível que os usuários navegaram com facilidade nos dados abertos pelo *Pentaho*.

Segundo Mccall (1977) a operabilidade é uma medida da facilidade de operação de um programa. Logo é possível inferir com base nas respostas dos entrevistados que este conceito de operabilidade foi atingindo.

Questão 26 – As informações disponibilizadas são exatas?

A questão de número 26 tem por finalidade avaliar se as informações exibidas são exatas, para serem exatas as informações devem bater com os valores reais. Nesta questão o participante escolhia uma nota de 1 (um) a 5 (cinco), quanto mais próximo de 1 (um) a nota for, as informações não eram exatas, e quanto mais próximo de 5 (cinco) as informações eram exatas. Os resultados podem ser verificados na Tabela 26 e no Gráfico 26.

Tabela 26 – Respostas da Questão 26.

	Respostas	Quantidade	Porcentagem
Nunca	1	0	0%
	2	0	0%
	3	0	0%
	4	4	44%
Sempre	5	5	56%

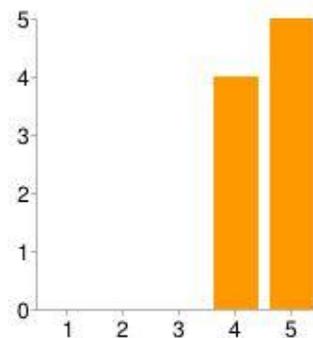


Gráfico 26 – Questão 26.

Analisando a Tabela 26, pode-se perceber que todos os participantes acharam que as informações disponibilizadas foram muito próximas de exata ou foram exata, 56% (cinquenta e seis por cento) deram a nota 5 (cinco), 44% (quarenta e quatro por cento) deram a nota 4 (quatro). Por intermédio das respostas desta questão, é possível perceber que os entrevistados acharam que as informações batem com os valores reais.

Segundo Mccall (1977) a comunicatividade de um software deve prover saídas que podem ser interpretadas pelo usuário. Com base nos resultados é possível inferir que o software atendeu a este conceito de comunicatividade.

Questão 27 – Como você avaliaria a mobilidade do *Pentaho*?

A vigésima sétima questão tinha por objetivo analisar a mobilidade do *Pentaho*. Para avaliar questão o participante dava uma nota de 1 (um) a 5 (cinco), a nota 1 (um) refere-se que a mobilidade do *Pentaho* foi péssima, a nota 5 (cinco) refere-se que foi excelente. O Gráfico 27 e a Tabela 27 exibem os valores.

Tabela 27 – Respostas da Questão 27.

	Respostas	Quantidade	Porcentagem
Péssimo	1	0	0%
	2	0	0%
	3	1	11%
	4	3	33%
Excelente	5	5	56%

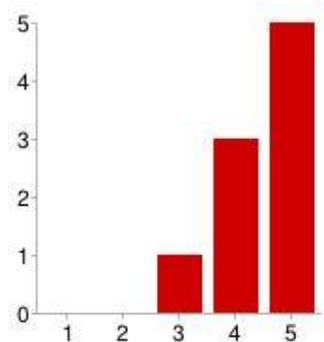


Gráfico 27 – Questão 27.

Como pode perceber na Tabela 27, 56% (cinquenta e seis por cento) dos participantes deram a nota 5 (cinco), 33% (trinta e três por cento) deram a nota 4 (quatro) e 11% (onze por cento) deram a nota 3 (três). Com base nas respostas desta questão, percebe-se que a mobilidade do *Pentaho* foi bem aceita, no entanto não é possível verificar a usabilidade de um *software* através desta questão.

Questão 28 – De forma geral qual a sua opinião sobre o *Pentaho*?

A vigésima oitava questão tem como objetivo avaliar de forma geral qual a satisfação do participante com o *Pentaho*. Nesta questão o participante teve que escolher uma nota entre 1 (um) e 5 (cinco), quanto mais próximo de 1 (um) a nota for, o participante não ficou satisfeito com o *Pentaho*, e quanto mais próximo de 5 (cinco) ele ficou satisfeito. A Tabela 28 mostra os resultados da questão junto com o Gráfico 28.

Tabela 28 – Respostas da Questão 28.

	Respostas	Quantidade	Porcentagem
Péssimo	1	0	0%
	2	0	0%
	3	0	0%
	4	3	33%
Excelente	5	6	67%

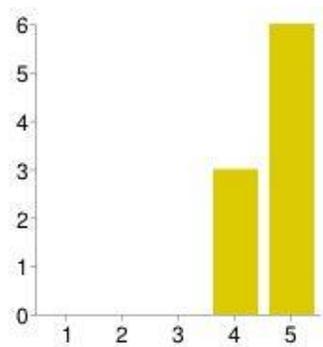


Gráfico 28 – Questão 28.

Todos os participantes ficaram satisfeito com o *Pentaho*, 67% (sessenta e sete por cento) deram a nota 5 (cinco), 33% (trinta e três por cento) deram a nota 4 (quatro). Com base nas respostas desta questão, percebe-se que ficaram satisfeitos com o *Pentaho*. Contudo não é possível verificar a usabilidade de um *software* através desta questão.

4 CONCLUSÃO

A proposta de analisar a usabilidade do Pentaho, que feita pelo presente estudo, mostrou-se bem eficiente, pois com base na presente pesquisa, que teve por objetivo avaliar a usabilidade da suíte *Pentaho*, e descobrir assim, o quanto é fácil utilizá-la, pôde-se perceber que o *Pentaho* foi bem aceito.

De acordo com grande parte dos entrevistados, foi possível observar que devido a sua interface intuitiva e fácil de ser manuseada juntamente com as facilidades para encontrar as informações que precisavam, os participantes obtiveram uma excelente experiência de usuário, ficando assim satisfeitos com a facilidade de usar o sistema e se familiarizarem com ele (tudo isto somado a resultados sólidos e precisos).

Por intermédio das repostas dos entrevistados, conclui-se que o *software* de *Business Intelligence* atende os conceitos de usabilidade apresentados neste trabalho e foi bem aceito na empresa, pois a ferramenta disponibiliza recursos de interface de fácil utilização e entendimento, tendo em vista que a interface é agradável e intuitiva.

Com a facilidade de uso do *Pentaho* comprovada neste trabalho, os processos de tomadas de decisões da área comercial da empresa DPC, ficaram mais fáceis e ágeis, por causa desses resultados positivos, alcançados por meio deste trabalho, o *Pentaho* hoje é utilizado pela maioria dos funcionários do setor comercial da empresa em que foi aplicado.

5 TRABALHOS FUTUROS

Uma possível continuação desse estudo poderia ser, levar informações para a tomada de decisão em outros setores da empresa DPC Distribuidor Atacadista S/A, com a finalidade de que no final toda a empresa tenha acesso interativo ao *Pentaho* e aos dados, para tornar os processos mais produtivos.

Também pode realizar um estudo na empresa aplicando outras ferramentas de *Business Intelligence* visando os benefícios que a nova ferramenta trará, comparando as diferenças e os benefícios de cada ferramenta.

Outra possível continuação desse estudo é utilizar a ferramenta *Pentaho Data Integration* para integrar os dados, de outros bancos de dados que a empresa utiliza, disponibilizando-os para eventuais processos de tomada de decisão.

6 REFERÊNCIAS

ANZANELLO, Cynthia. A. **OLAP Conceitos e Utilização**. Universidade Federal do Rio Grande do Sul, Porto Alegre - RS, 2008.

BALLARD, Chuck; FARRELL, Daniel; *et al.* **Dimensional Modeling**: In a Business Intelligence Environment, IBM, 2006.

BARBIERI, Carlos. **BI2 – Business Intelligence: modelagem e tecnologia**. 2nd ed. Rio de Janeiro: Axcel Books do Brasil Editora, 2001.

BARROS, Renata Pitta. **Evolução, Avaliação e Validação do Software RoboEduc**. Natal, 2011.

BOBSIN, Juliana. **Uma Solução de BI Utilizando Ferramentas Open Source**. Gravataí - SP, 2010.

CRAMER, R. **Estudo analítico de ferramentas open source para Ambientes OLAP**. Criciúma: Universidade do Extremo Sul Catarinense, 2006.

DPC. DPC Distribuidor Atacadista S/A. **Empresa**. Caratinga – MG, 2013. Disponível em: <<http://www.dpcnet.com.br/pagina/6-empresa>> Acessado em: 26 outubro. 2014.

ECKERSON, Wayne. **Smart Companies in the 21st Century: The Secrets of Creating Successful Business Intelligence Solutions**. Seattle, WA: The Data Warehousing Institute.

FERREIRA, Kátia Gomes. **Teste de Usabilidade**. Belo Horizonte, 2002.

FORTULAN, Marcos. **UMA PROPOSTA DE APLICAÇÃO DE BUSINESS INTELLIGENCE NO CHÃO-DE-FÁBRICA. GESTÃO & PRODUÇÃO**, São Carlos - SP v. 12, p. 12, 2005.

FOUCHÉ Guy; LANGIT Lynn. **Foundations of SQL Server 2008 R2: Business Intelligence**. Apress, 2 ed, 2011.

FUJIWARA, Daniel Kenji. **Data Warehouse como Instrumento de Suporte à Avaliação Acadêmica**. Brasília - DF: Universidade de Brasília, 2006.

HAMMERGREN, Thomas C.; SIMON, Alan R. **Data Warehouse for Dummies**. 2 ed. Indianapolis: Wiley Publishing, 2009.

HAN, Jiawei; KAMBER, Micheline. **Data Mining: Concepts and Techniques**. 2 ed. São Francisco, USA: Morgan Kaufmann Publishers, 2006.

INMON, William. H. **Building the Data Warehouse**. New York: John Wiley, 2005.

ITALIANO, Isabel. C.; ESTEVES, Luiz. A. **Modelagem de Data Warehouses e Data Marts** – Parte 1, SQL Magazine, Rio de Janeiro - RJ, 2004. Disponível em: <<http://www.devmedia.com.br/artigo-sql-magazine-13-modelagem-de-data-warehouses-e-data-marts-parte-1/5656>> Acessado em: 14 de dezembro 2014.

KIMBALL, Ralph; ROSS Margy. **Data Warehouse Toolkit: The Definite Guide to Dimensional Modeling**. 3 ed Canada: John Wiley, 2013.

MACHADO, Felipe Nery Rodrigues **Tecnologia e Projeto de Data Warehouse: Uma visão Multidimensional**. São Paulo: Érica, 2010.

MCCALL, A Jim. RICHARD, Paul K. WALTERS, F. Gene. **Factors in Software Quality: Concepts and Definitions of Software Quality**. 1977.

MOREIRA, E. **Modelo Dimensional para Data Warehouse**, 2006. Disponível em: <<http://imasters.com.br/artigo/3836/gerencia-de-ti/modelo-dimencional-para-data-warehouse/>>. Acessado em: 14 dezembro. 2014.

NIELSEN, Jakob. **10 Usability Heuristics for User Interface Design**. Usability Inspection Methods, John Wiley & Sons, New York, NY. 1994.

OLSZAK, Celina M.; ZIEMBA, Ewa. **Business Intelligence as a key to management of an enterprise**. Pori, Finland. Proceedings. Pori: Informing Science Institute, 2003. p. 855-863. Disponível em: <<http://proceedings.informingscience.org/IS2003Proceedings/docs/109Olsza.pdf>>

PENTAHO CORPORATION. **Pentaho Data Integration**. Disponível em: <<http://www.pentaho.com/product/data-integration>>. Acessado em: 20 agosto. 2014.

PRESSMAN, Roger S. **Engenharia de Software: Uma abordagem profissional**. 7ª ed. AMGH Editora Ltda, 2011.

PRIMAK, Fábio Vinicius. **Decisões com BI: (Business Intelligence)**. Rio de Janeiro: Ciência Moderna, 2008.

REIS, Eduardo. **Implementando uma solução de Business Intelligence com o Microsoft SQL Server 2005 – Parte 1**, SQL Magazine, Rio de Janeiro - RJ, 2009. Disponível em <<http://www.devmedia.com.br/artigo-sql-magazine-59-implementando-uma-solucao-business-intelligence-com-o-microsoft-sql-server-2005-parte-1/11077>> Acessado em: 14 de dezembro de 2014.

ROSINI, Alessandro M.; PALMISANO, Ângelo. **Administração de sistemas de informação e a gestão do conhecimento**. São Paulo – SP: Thomson, 2003.

SALLES, Fábio de. *et al.* **Pentaho na Prática: Um guia para construir soluções de inteligência de negócios com Pentaho**. São Paulo: Amazon, 2013.

SCHEPS, Swain. **Business Intelligence for Dummies**. EUA: Wiley Publishing, 2008.

SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. **SISTEMA DE BANCO DE DADOS**. 6 ed. São Paulo - SP: Elsevier Editora Ltda, 2011.

SOMMERVILLE, Ian. **Software Engineering**. 9th ed. Pearson, 2010.

SOUZA, Hudson. **Aplicação de Ferramentas de Business Intelligence para Consolidação de Dados Financeiros de uma Instituição de Ensino: Estudo de Caso com a Suíte Pentaho**. Caratinga, 2012.

TURBAN, Efraim. *et al.* **Business Intelligence: um enfoque gerencial para a inteligência do negócio**. São Paulo: Bookman, 2009.

7 ANEXOS

7.1 ANEXO I: QUESTIONÁRIO

Questionário de satisfação do usuário.

Este formulário tem por finalidade analisar o grau de satisfação do usuário com o sistema Pentaho.

1. Qual a sua idade?

- 18 a 25
- 26 a 35
- 36 a 45
- Acima 45

2. Qual cargo você ocupa na empresa?

- Diretor
- Gerente
- Coordenador
- Colaborador
- Outro

3. Com que frequência você usa o Pentaho?

- Diariamente
- Semanalmente
- Mensalmente
- Outro

Acesso ao Pentaho

4. Quão fácil foi acessar o Pentaho?

Muito difícil ()1 ()2 ()3 ()4 ()5 Muito fácil

5. O Pentaho ao seu ver o é seguro?

Pouco seguro ()1 ()2 ()3 ()4 ()5 Muito seguro

Interface do Pentaho

6. Como você avaliaria a interface do Pentaho?

Muito confusa ()1 ()2 ()3 ()4 ()5 Intuitiva e agradável

7. Quão fácil foi encontrar a informação que você desejava?

Pouco fácil ()1 ()2 ()3 ()4 ()5 Muito fácil

8. Havia muitas maneiras de encontrar a informação que você desejava?

Poucas maneiras ()1 ()2 ()3 ()4 ()5 Muitas maneiras

9. Quão fácil foi se desorientar no sistema?

Fácil ()1 ()2 ()3 ()4 ()5 Difícil

Relatórios do Pentaho

10. Quão fácil foi usar os relatórios?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

11. Como você avaliaria o agendamento de relatórios?

Ruim ()1 ()2 ()3 ()4 ()5 Excelente

12. Os formatos disponibilizados dos relatórios são úteis?

Os formatos PDF, XLS, HTML.

Pouco ()1 ()2 ()3 ()4 ()5 Muito

13. As informações exibidas pelos relatórios são exatas?

Os valores batem com a realidade

Pouco ()1 ()2 ()3 ()4 ()5 Totalmente

14. Os relatórios destacam as informações importantes?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

15. Na sua opinião quão úteis foram os relatórios?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

Dashboards

16. Quão fácil foi usar os *dashboards*?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

17. As informações exibidas nos painéis são importantes e úteis?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

Análise de informação

18. Quão fácil foi achar as ferramentas de análises?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

19. Foi fácil criar uma análise de informação?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

20. Os cubos possuem informações importantes e úteis?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

21. Quão fácil foi criar relatórios personalizados?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

Geral sobre o Pentaho

22. As informações disponibilizadas são confiáveis?

Nunca ()1 ()2 ()3 ()4 ()5 Sempre

23. Na sua opinião o tempo de espera para visualizar a informação é satisfatório?

Pouco ()1 ()2 ()3 ()4 ()5 Muito

24. Quão fácil foi achar os seus arquivos?

Péssimo ()1 ()2 ()3 ()4 ()5 Excelente

25. Na sua opinião quão fácil foi navegar entre os dados?

Péssimo ()1 ()2 ()3 ()4 ()5 Excelente

26. As informações disponibilizadas são exatas?

Nunca ()1 ()2 ()3 ()4 ()5 Sempre

27. Como você avaliaria a mobilidade do *Pentaho*.

Péssimo ()1 ()2 ()3 ()4 ()5 Excelente

28. De forma geral qual a sua opinião sobre o *Pentaho*?

Péssimo ()1 ()2 ()3 ()4 ()5 Excelente

7.2 ANEXO II: TERMO DE AUTORIZAÇÃO

**TERMO DE AUTORIZAÇÃO PARA DIVULGAÇÃO DE INFORMAÇÕES DA
EMPRESA****Empresa:** DPC Distribuidor Atacadista S/A**CNPJ:** 66471517000177**Inscrição Estadual:** 1348384310043**Endereço completo:** AV AUREA CARLOS LEITE DE MATOS**Representante da empresa:** Edson Nascimento Ferreira**Telefone:** (33) 8424-0000 **e-mail:** edsonferreira@dpcnet.com.br**Tipo de produção intelectual:** TCC - monografia de Curso de Graduação**Título/subtítulo:** USO DE FERRAMENTAS OPEN SOURCE PARA SOLUÇÕES
DE BUSINESS INTELLIGENCE: ESTUDO DE CASO COM A SUÍTE PENTAHO**Autor:** Hugo Morais Mendes**Código de matrícula:** 110300009**Orientador:** Msc. Fabrícia Pires Souza Tiola**Graduação:** Ciência da Computação

Como representante da empresa acima nominada, declaro que as informações e/ou documentos disponibilizados pela empresa para o trabalho citado podem ser publicados sem restrição.

Representante da empresa

Local e Data

7.3 ANEXO III: *VIEWS*

View `dpcv_com_metas_valores`

```
select mp.id_meta,  
       mp.cod_pessoa,  
       mp.tipo,  
       me.valor_meta,  
       me.codigo cod_tipoApurado,  
       me.evento,  
       mp.responsabilidade  
  
from poseidon.dpc_com_meta_pessoa mp inner join poseidon.dpc_com_meta_extra me  
on mp.id_meta_pessoa = me.id_meta_pessoa;
```

View `dpcv_com_metas`

```
select M.id_meta,  
       M.data_inicio,  
       M.data_fim,  
       case  
         when M.tipometa = 'FAT' then  
           'FATURAMENTO'  
         when m.tipometa = 'POS' then  
           'POSITIVACAO DE CLIENTES'  
         when m.tipometa = 'TIC' then  
           'TICKET MEDIO'  
         else  
           m.tipometa  
         end tipometa,  
       m.status,  
       m.nome  
  
from poseidon.dpc_com_meta m  
where m.status = 'A'
```

```
order by m.data_inicio desc,m.nome;
```

View dpcv_dw_meta

```
select m.id_meta,
       m.nome as descricao,
       m.tipometa as tipo_meta,
       m.data_inicio,
       m.data_fim
from poseidon.dpc_com_meta m
where trunc(sysdate - 1) between m.data_inicio and m.data_fim
and m.status = 'A';
```

View dpcv_dw_meta_positivacao_cli

```
select t.id_meta,
       t.cod_pessoa as codigo_funcionario,
       t.tipo as tipo_funcionario,
       t.valor_meta as meta
from poseidon.dpcv_com_metasp_valores t
inner join poseidon.dpcv_com_metasp m
on t.id_meta = m.id_meta
and trunc(sysdate - 1) between m.data_inicio and m.data_fim
and m.status = 'A'
and m.tipometa = 'POSITIVACAO DE CLIENTES'
and t.tipo = 'GER';
```

View dpcv_dw_meta_ticket_medio

```
select t.id_meta,
       t.cod_pessoa as codigo_funcionario,
       t.tipo as tipo_funcionario,
       t.valor_meta as meta
```

```
from poseidon.dpcv_com_metras_valores t
inner join poseidon.dpcv_com_metras m
  on t.id_meta = m.id_meta
  and trunc(sysdate - 1) between m.data_inicio and m.data_fim
  and m.status = 'A'
  and m.tipometa = 'TICKET MEDIO'
  and t.tipo = 'GER' ;
```

View dpcv_dw_meta_mix_producto

```
select t.id_meta,
       t.cod_pessoa   as codigo_funcionario,
       t.tipo         as tipo_funcionario,
       t.valor_meta   as meta
from poseidon.dpcv_com_metras_valores t
inner join poseidon.dpcv_com_metras m
  on t.id_meta = m.id_meta
  and trunc(sysdate - 1) between m.data_inicio and m.data_fim
  and m.status = 'A'
  and m.tipometa = 'MIX'
  and t.tipo = 'GER';
```

View dpcv_dw_meta_empresa

```
select e.id_meta_empresa,
       e.id_meta,
       e.cod_empresa,
       e.valor_meta
from poseidon.dpc_com_meta_empresa e;
```

View dpcv_dw_meta_geral

```
select t.id_meta,
       t.cod_pessoa    as codigo_funcionario,
       t.tipo          as tipo_funcionario,
       t.cod_tipoApurado as codigo_tipoApuracao,
       t.evento        as tipo_apuracao,
       t.responsabilidade,
       t.valor_meta    as meta
from poseidon.dpcv_com_metras_valores t
inner join poseidon.dpcv_com_metras m
  on t.id_meta = m.id_meta
 and trunc(sysdate - 1) between trunc(m.data_inicio) and
    trunc(m.data_fim)
 and m.status = 'A'
 and m.tipometa = 'FATURAMENTO';
```