



**UNIVERSIDADE DO SUL DE SANTA CATARINA**

**ANTÔNIO L P FILHO**

**FELLIPE L GUIMARÃES**

**APLICAÇÃO DO BI NA ÁREA DE RECURSOS HUMANOS  
OTIMIZANDO O RH EM PROL DA ORGANIZAÇÃO**

Palhoça

2015

**ANTÔNIO L P FILHO**  
**FELLIPE L GUIMARÃES**

**APLICAÇÃO DO BI NA ÁREA DE RECURSOS HUMANOS**  
**OTIMIZANDO O RH EM PROL DA ORGANIZAÇÃO**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Sistemas de Informação da Universidade do Sul de Santa Catarina, como requisito parcial à obtenção do título de Bacharel em Sistemas de Informação.

Orientador: Aran Bey Tcholakian Morales, Dr. Eng.

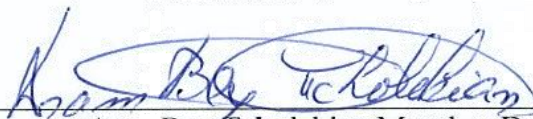
Palhoça  
2015

**ANTÔNIO L P FILHO**  
**FELLIPE L GUIMARÃES**

**APLICAÇÃO DO BI NA ÁREA DE RECURSOS HUMANOS**  
**OTIMIZANDO O RH EM PROL DA ORGANIZAÇÃO**

Este Trabalho de Conclusão de Curso foi julgado adequado à obtenção do título de Bacharel em Ciência da Computação e aprovado em sua forma final pelo Curso de Graduação em Sistemas de Informação da Universidade do Sul de Santa Catarina.

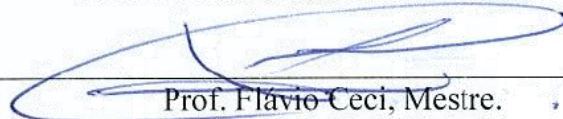
Palhoça, 08 de Junho de 2015.



Aran Bey Tcholakian Morales, Dr. Eng.  
Universidade do Sul de Santa Catarina



Prof. Daniella Pinto Viera, Mestre.  
Universidade do Sul de Santa Catarina



Prof. Flávio Ceci, Mestre.  
Universidade do Sul de Santa Catarina

Este trabalho de conclusão de curso será dedicado a nossos pais e amigos por toda a ajuda moral e o suporte dado durante nosso caminho pelo curso de Sistemas de Informação na UNISUL.

## **AGRADECIMENTOS**

Em primeiro lugar, gostaríamos de agradecer nossos pais, familiares e amigos por todo o suporte dado durante toda a trajetória da graduação.

Em seguida gostaríamos e agradecer o professor Aran, cujo é o nosso orientador, por toda a ajuda com o conteúdo didático e com todo o suporte nas análises e validações dos dados, bem como nossa professora de TCC, Maria Inês, por todo o apoio com a estrutura do trabalho e suas melhores práticas para construí-lo.

“Os problemas significativos que enfrentamos não podem ser resolvidos no mesmo nível de pensamento em que estávamos quando os criámos.” (Albert Einstein).

## RESUMO

Nas pequenas, médias e grandes empresas, é cada vez mais comum o setor de Recursos Humanos sendo responsável por uma grande parcela do gerenciamento dos funcionários e aquilo que englobam os seus perfis e seus indicadores para a organização. Esse setor possui uma representação muito, e talvez, a mais significativa para o sucesso e o diferencial competitivo entre as empresas no mercado de trabalho, tendo a tangibilidade de tomar decisões que impactam diretamente em ações e resultados que podem mudar o comportamento e a forma como são vistas e tratadas às pessoas e seus perfis dentro da organização, o que pode proporcionar uma estratégia focada ou, por exemplo, o aumento de desempenho em determinado setor.

Com todo esse embasamento, a proposta de encorpar ainda mais a responsabilidade do setor de RH e torna-lo capaz de tomar decisões que afetam toda a corporação, foi identificada a necessidade de pôr nas mãos desta área a capacidade de analisar dados de funcionários e melhorar a sua Gestão de Pessoas com o intuito de retratar ações com bases em informações construídas a partir de um *Data Warehouse* e uma solução de *Business Intelligence*.

Essas soluções possibilitam formas mais assertivas de se tomar de decisões que possam mudar o rumo de toda a visão das pessoas perante a organização, assim como promover a área e sua gestão perante toda a empresa, resultando em um poder analítico e diretivo de toda a gestão e a parte operacional que a mesma compõe.

Palavras-chave: Inteligência de Negócios, *Business Intelligence*, *BI*, *Data Warehouse*, Recursos Humanos, RH.

## LISTA DE ILUSTRAÇÕES

|  |    |
|--|----|
| Figura 1 Processo de um <i>Business Intelligence</i> .....                             | 19 |
| Figura 2 Tabela Fato .....   | 22 |
| Figura 3 Modelo Estrela .....  | 23 |
| Figura 4 Modelo <i>Snowflake</i> .....   | 24 |
| Figura 5 Segmentações de um <i>DW</i> em <i>Data Mart's</i> específicos .....          | 26 |
| Figura 6 Processo de uma consolidação de uma <i>ETL</i> .....                          | 28 |
| Figura 7 Transformação do <i>DW/Data Mart</i> para o cubo em <i>OLAP</i> .....         | 29 |
| Figura 8 Arquitetura da solução .....  | 33 |
| Figura 9 Requisitos Funcionais .....   | 38 |
| Figura 10 Requisitos Não Funcionais .....  | 39 |
| Figura 11 Arquitetura do Modelo Relacional .....                                       | 42 |
| Figura 12 Arquitetura do Modelo Dimensional .....                                      | 44 |
| Figura 13 Transformação completa do modelo relacional .....                            | 52 |
| Figura 14 Transformação dos dados dos funcionários .....                               | 53 |
| Figura 15 Transformação dos dados hierárquico dos funcionários .....                   | 54 |
| Figura 16 Transformação dos dados de registros de ponto .....                          | 55 |
| Figura 17 Transformação completa do modelo dimensional .....                           | 56 |
| Figura 18 Transformação dos dados dos funcionários .....                               | 57 |
| Figura 19 Transformação das agregações dos dados dos registros de pontos diários ..... | 58 |
| Figura 20 Transformação das agregações dos dados dos registros de pontos mensais ..... | 59 |
| Figura 21 Cubo de dados .....  | 60 |
| Figura 22 Análises dos dados de absenteísmo por faixa de idade .....                   | 63 |
| Figura 23 Análise do total de atrasos por município .....                              | 64 |
| Figura 24 Análise do total de demissões e admissões por departamento .....             | 65 |
| Figura 25 Análise da distribuição de cargos e salários da organização .....            | 66 |
| Figura 26 Análise do <i>turnover</i> de cada unidade .....                             | 67 |
| Figura 27 Estrutura de pastas do servidor .....  | 68 |
| Figura 28 Tela de <i>Login</i> .....   | 69 |
| Figura 29 Tela inicial do <i>Pentaho</i> .....   | 69 |
| Figura 30 Tela de Login do <i>Saiku Analytics</i> .....                                | 70 |



## **LISTA DE TABELAS**

|  |    |
|--|----|
| Tabela 1 - Relação de registros da guia de Funcionário.....  | 40 |
| Tabela 2 - Relação de registros da guia de Status Ponto..... | 41 |
| Tabela 3 - Relação de registros da guia de Ponto.....        | 41 |

## **LISTA DE ABREVIATURAS E SIGLAS**

BI – *Business Intelligence.*

CSV – *Comma-Separated Values*

ETL – *Extract Transform Load*

DW – *Data Warehouse*

OLAP – *Online Analytical Processing*

RH – Recursos Humanos

SGBD – Sistema de Gerenciamento do Banco de Dados

TI – Tecnologia da Informação

## SUMÁRIO

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>INTRODUÇÃO .....</b>                  | <b>12</b> |
| 1.1      | PROBLEMÁTICA .....                       | 13        |
| 1.2      | OBJETIVO .....                           | 13        |
| 1.3      | JUSTIFICATIVA .....                      | 14        |
| 1.4      | ESTRUTURA DO TRABALHO .....              | 15        |
| <b>2</b> | <b>REVISÃO BIBLIOGRÁFICA.....</b>        | <b>16</b> |
| 2.1      | A ÁREA DE RECURSOS HUMANOS.....          | 16        |
| 2.2      | OS SISTEMAS DE INFORMAÇÃO .....          | 17        |
| 2.3      | BUSINESS INTELLIGENCE.....               | 17        |
| 2.4      | DATA WAREHOUSE .....                     | 19        |
| 2.5      | DATA MINING .....                        | 26        |
| 2.6      | ETL .....                                | 27        |
| 2.7      | FERRAMENTAS DE FRONT-END .....           | 28        |
| <b>3</b> | <b>METODOLOGIA .....</b>                 | <b>31</b> |
| 3.1      | CARACTERIZAÇÃO DO TIPO DE PESQUISA ..... | 31        |
| 3.2      | ETAPAS METODOLÓGICAS .....               | 32        |
| 3.3      | ARQUITETURA DA SOLUÇÃO .....             | 32        |
| 3.4      | DELIMITAÇÕES.....                        | 33        |
| <b>4</b> | <b>MODELAGEM.....</b>                    | <b>35</b> |
| 4.1      | OBJETIVO GERAL DE NEGÓCIO.....           | 35        |
| 4.2      | CONTEXTO ATUAL DE NEGÓCIO .....          | 36        |
| 4.3      | CONTEXTO DO SISTEMA .....                | 37        |
| <b>5</b> | <b>DESENVOLVIMENTO .....</b>             | <b>48</b> |
| 5.1      | FERRAMENTAS E TECNOLOGIA .....           | 48        |
| 5.2      | APRESENTAÇÃO DA APLICAÇÃO .....          | 50        |
| 5.3      | UTILIZAÇÃO DA APLICAÇÃO.....             | 67        |
| 5.4      | AVALIAÇÃO DA APLICAÇÃO .....             | 70        |
| <b>6</b> | <b>CONCLUSÃO .....</b>                   | <b>72</b> |
| 6.1      | CONCLUSÕES E CONTRIBUIÇÕES.....          | 72        |
| 6.2      | TRABALHOS FUTUROS .....                  | 73        |
|          | <b>REFERÊNCIAS .....</b>                 | <b>74</b> |
|          | <b>APÊNDICES.....</b>                    | <b>76</b> |
|          | <b>APÊNDICE A – CRONOGRAMA.....</b>      | <b>77</b> |

## 1 INTRODUÇÃO

Cada vez mais os avanços tecnológicos beneficiam as áreas e departamentos das organizações atuais, e consequentemente, a área de Recursos Humanos (RH) está incluída nela. Entretanto, não é o que se tem visto em todas as pequenas, médias e grandes empresas do mercado, onde a área de RH, por muitas das vezes, possui processos e controles ainda indefinidos, manuais ou dependentes de tarefas que necessitam de um maior tempo para sua execução, o que para esta área, o tempo torna-se algo cada vez mais raro de se encontrar.

Outro fator, talvez o mais importante deles, é a cultura organizacional resistente do setor de RH. Esta resistência dos colaboradores da área se dá devido ao fato dos mesmos ainda acreditarem que o processo atual, o não informatizado, é de fato, o mais simples de se executar, mesmo sendo, muito das vezes, o mais trabalhoso e demorado de se realizar.

O excesso de demandas com os níveis de urgência altíssimos é outro ponto que impacta diretamente na informatização da área de Recursos Humanos, pois como cada operação da área envolve tarefas críticas e de impactos direcionados a organização como um todo, o nível de priorização das atividades é sempre levado de maneira absoluta, do qual contribui com este atrito, entre RH e a tecnologia na automatização de processos.

A Gestão de Recursos Humanos está associada diretamente a parte estratégica de uma empresa. Ter a governança sobre seu quadro de funcionários, por exemplo, tornou-se cada vez mais essencial e diferencial entre as organizações, visto que uma gestão de pessoas bem estruturada nesta área difere o sucesso e o fracasso de uma empresa.

Este potencial encontrado vem de muitas responsabilidades existentes, pois o volume de dados que circulam e englobam a área, é, na maioria das vezes, um volume altíssimo e nem sempre vem em uma organização, ou uma estrutura, que facilitam a integração entre si, como por exemplo, o recebimento de relatórios de outros setores, como o financeiro, o tecnológico, o administrativo, entre outros.

## 1.1 PROBLEMÁTICA

Segundo Ulrich (2003) o setor de recursos humanos tem um papel fundamental e estratégico para uma organização, contudo, o mesmo ainda encontra-se estagnado e pouco explorado, encontrando barreiras e mitos que o impedem os colaboradores de serem mais profissionais, tornando-se apenas contribuintes do lado operacional. Essas barreiras servem como travas para que os funcionários da área se desenvolvam e se tornem grandes planejadores estratégicos da gestão de pessoas.

Ainda de acordo com as conclusões de Ulrich (2003, p. 43) “como parceiros estratégicos, os profissionais de RH devem ser capazes de identificar as práticas que fazem com que a estratégia aconteça”. Esse perfil estratégico por parte do RH é escasso devido à falta de insumo para a análise de indicadores da organização, uma vez que esses indicadores estejam disponíveis para os profissionais da área, o lado estratégico pode fazer parte desta gestão, construindo para a tomada de decisões e mudando a forma que este setor é visto e incorporado na empresa.

Chiavenato (2009) afirma que as organizações estão passando por mudanças e transformações a cada dia, seja introduzindo novas e diferentes tecnologias, seja modificando seus produtos e serviços, seja alterando o comportamento das pessoas, seja mudando processos internos, entre outros. Por isso, essa constante mudança, faz com que a área de RH perca um pouco de sua identidade original, a de enxergar as pessoas como parte de um grande processo em potencial, viabilizando estratégias que possam incorporar ainda mais a organização como um todo.

## 1.2 OBJETIVO

O objetivo geral é implantar uma arquitetura de *BI (Business Intelligence)* para o auxílio na gestão dos principais processos da área de recursos humanos de uma empresa.

### 1.2.1 Objetivos Específicos

Para alcançar este objetivo geral devemos alcançar os seguintes objetivos específicos:

- Levantar os requisitos para analisar dos principais processos da gestão de RH;
- Identificar os principais indicadores que fazem parte da Gestão de Recursos Humanos;
- Construir o modelo dimensional que atenda os requisitos levantados;
- Manipular uma base de dados para a importação;
- Realizar a carga do modelo dimensional;

### 1.3 JUSTIFICATIVA

O conceito de *Data Warehouse (DW)*, segundo Inmon (1997, p. 33), “é um conjunto de dados baseados em assuntos, integrado, não volátil, e variável em relação ao tempo, de apoio às decisões gerenciais”, vem como forma de uma estrutura centralizada para o armazenamento de dados, o que facilita consultas e cruzamentos de diversas informações, gerando uma maior visibilidade de informações processuais e corporais a área.

A partir da utilização de um *DW*, as tomadas de decisões, por parte da gestão, passam a ser mais precisas e eficazes, pois todas as variações e tendências das informações passam a ser identificadas de maneiras mais ágeis, com o armazenamento e direcionamento de registros históricos das tarefas e demandas do RH.

A visão agregada dos dados também facilita em uma análise mais técnica e precisa, permitindo a construção de gráficos que agem de maneira mais visual e veloz na compreensão e na dedução de ofensores, por exemplo.

Essa arquitetura permite a construção de um sistema de análise com a metodologia de *BI (Business Intelligence)*, que visa tornar as decisões e as tarefas executadas pelos profissionais de RH, o mais simples possível, com um olhar ainda mais analítico e assertivo sobre os problemas recorrentes do dia-a-dia, a fim de resultar em um impacto direto na organização como um todo. Além disso, as oportunidades de melhorias tornam-se mais evidentes com base nas análises geradas pelo *BI*, servindo de espelho para o crescimento de atividades semelhantes nas demais áreas de toda a organização.

A automatização dos processos também é incorporada com a implantação destes conceitos, já que a análise das informações e tarefas que compõem os procedimentos serão mais visíveis e acessíveis. Controles como o de presenças e faltas são gerados de maneira mais simples e objetiva.

As análises, os estudos e os cruzamentos de dados dos colaboradores absenteístas são levantados de acordo com os perfis dos funcionários, o que acaba aprofundando no entendimento dos índices de absenteísmo da corporação. Segundo Chiavenato (2004), um funcionário dito problemático, é aquele responsável pelo aumento do índice de absenteísmo da organização, que acaba servindo de reflexo da gestão de pessoas que compõem a área do mesmo, refletindo em um desencadeamento de outros absenteístas, sejam através de atrasos ou faltas, demissões e desmotivações dos demais.

Além disso, estas análises permitem o controle de *turnover* em benefícios voltados ao controle de gestão de pessoas, controle de campanhas relacionadas à gestão de pessoas, a identificação e estruturação de novos e atuais perfis de funcionários, o que garante uma maior aderência aos setores, assim como a candidatos que buscam empregos na organização.

#### 1.4 ESTRUTURA DO TRABALHO

No primeiro capítulo é feita uma introdução ao projeto de conclusão de curso, abordando todas as práticas e métodos que são efetuados no trabalho, diagnosticando todos os problemas, objetivos e perguntas que encontramos no trabalho.

Já no segundo capítulo, há uma descrição de todos os conceitos já existentes referentes as problemáticas já levantadas no primeiro capítulo, abordando todas as melhores e até mesmo as piores práticas já encontradas.

O terceiro capítulo mostra quais são os métodos abordados no Sistema de *BI*, tal como sua arquitetura.

O quarto capítulo mostra as atividades desenvolvidas para a realização do projeto.

No quinto capítulo evidenciamos os resultados obtidos com a implementação do sistema de *BI*, tal como seus benefícios, problemas e oportunidades.

No sexto capítulo é feita a conclusão do projeto e possíveis próximos passos.

## 2 REVISÃO BIBLIOGRÁFICA

Este capítulo mostra os principais conceitos e as principais ferramentas que estão atreladas a inteligência de negócios.

Este capítulo está subdividido em sete subtópicos, retratando os seguintes assuntos: A área de Recursos Humanos, os Sistemas de Informação, *Business Intelligence*, *Data Warehouse*, *Data Mining*, *ETL* e *Front-End OLAP*.

### 2.1 A ÁREA DE RECURSOS HUMANOS

A área de Recursos Humanos surgiu a partir do crescimento e do nível de complexidade das tarefas organizações, tendo suas origens ao início do século XX, após a Revolução Industrial e toda a necessidade gerada pela mesma (CHIAVENATO, 2009).

Por volta da década de 50, a área que até então representava o RH, deixou de ser apenas responsável apenas por resolver conflitos e desavenças, passando também a administrar os funcionários de acordo com as legislações trabalhistas. Já nos anos 60, esse conceito passou novamente por transformações, no qual as pessoas passaram a ser vistas como recursos fundamentais para a organização. Foi então a partir da década de 1970 que surgiu de fato o conceito concreto de Recursos Humanos, embora ainda visasse os trabalhadores como recursos produtivos ou agentes controlados de acordo com as necessidades da organização. Desde a década de 70 até hoje, as grandes organizações deixaram de tratar as pessoas como recursos operacionais controlados de acordo a própria organização, e passaram a trata-los como parte fundamental de um todo, administrando com pessoas (CHIAVENATO, 2009).

Atualmente é praticamente indispensável não incluir esta área em uma organização, vista que seu grau de responsabilidade com as pessoas faz parte de sua gestão.

Segundo Chiavenato (2009, p. 3) “as pessoas constituem um poderoso ativo que impulsiona a competitividade organizacional, da mesma forma que o mercado e a tecnologia”.



## 2.2 OS SISTEMAS DE INFORMAÇÃO

Os sistemas de informação estão presentes em tudo aquilo que engloba o meio corporativo empresarial, assim como a tudo que cerca a vida urbana das pequenas e grandes cidades, associando máquinas com as pessoas em virtude de seus processos.

De acordo com Laudon e Laudon (1999, p. 4), “um sistema de informação pode ser definido como um conjunto de componentes interrelacionados trabalhando juntos para coletar, recuperar, processar, armazenar e distribuir informações com a finalidade de facilitar o planejamento, o controle, a coordenação, a análise e o processo decisório em organizações”.

Ao se utilizar sistemas de informações na Área de Recursos Humanos, por exemplo, é inevitável termos as entradas de dados e os surgimentos das informações. “O dado é a matéria-prima com que o sistema de informações vai trabalhar, e a informação é o produto final do sistema de informações e deve ser apresentada em forma, prazo e conteúdo adequado aos usuários” (GIL, 1995, p. 13).

A partir dos dados é que são iniciados os processamentos dos mesmos, a fim de resultar em uma visão ampliada do fator a ser analisado, gerando assim informações capazes de serem utilizadas em tomadas de decisões ou até relatórios personalizados (MELO, 1999).

## 2.3 BUSINESS INTELLIGENCE

A dinâmica dos negócios no mercado empresarial, cada vez mais competitivo e em uma linha crescente de evolução, fez com que as empresas busquem estar sempre bem informadas e atualizadas a respeito de informações que envolvem diretamente ou indiretamente o seu portfólio de negócios. Com isso, a empresa que conseguir entender a demanda pela qual o mercado vertigina ou o que se passa com o seu negócio e com o de seus concorrentes, será uma empresa com uma consistência suficiente capaz de sobrepor os seus concorrentes e impor-se no mercado, de forma técnica, objetiva e na grande maioria das vezes, eficaz.

O termo *BI* foi criado pelo *Gartner Group* em meados dos anos 90, porém seu conceito foi visto muito antes na geração de relatórios SIG dos anos 70, cujos relatórios eram estáticos, bidimensionais e não tinham nenhum recurso voltado para a análise em si (TURBAN, 2009).

No início dos anos 80, surgiu um novo conceito denominado ‘sistemas de informações executivas’ (EIS), do qual introduziu o recurso de sistemas de geração de relatórios dinâmicos multidimensionais, prognósticos e previsões, análise de tendências, fatores críticos de sucesso, entre outros. Esses mesmos recursos se estabeleceram até meados dos anos 90, onde os mesmos, e alguns outros novos, passaram a receber a nomenclatura de *BI* (TURBAN, 2009).

O *Business Intelligence* tem por objetivo o tratamento das informações que apoiam a gestão de um negócio, sendo indispensável à utilização de uma estrutura de armazenagem de dados em um modelo que irá otimizar e simplificar as buscas e cruzamentos nele encontrados, embora nem todos os sistemas de *Business Intelligence* necessitam obrigatoriamente de um *DW*.

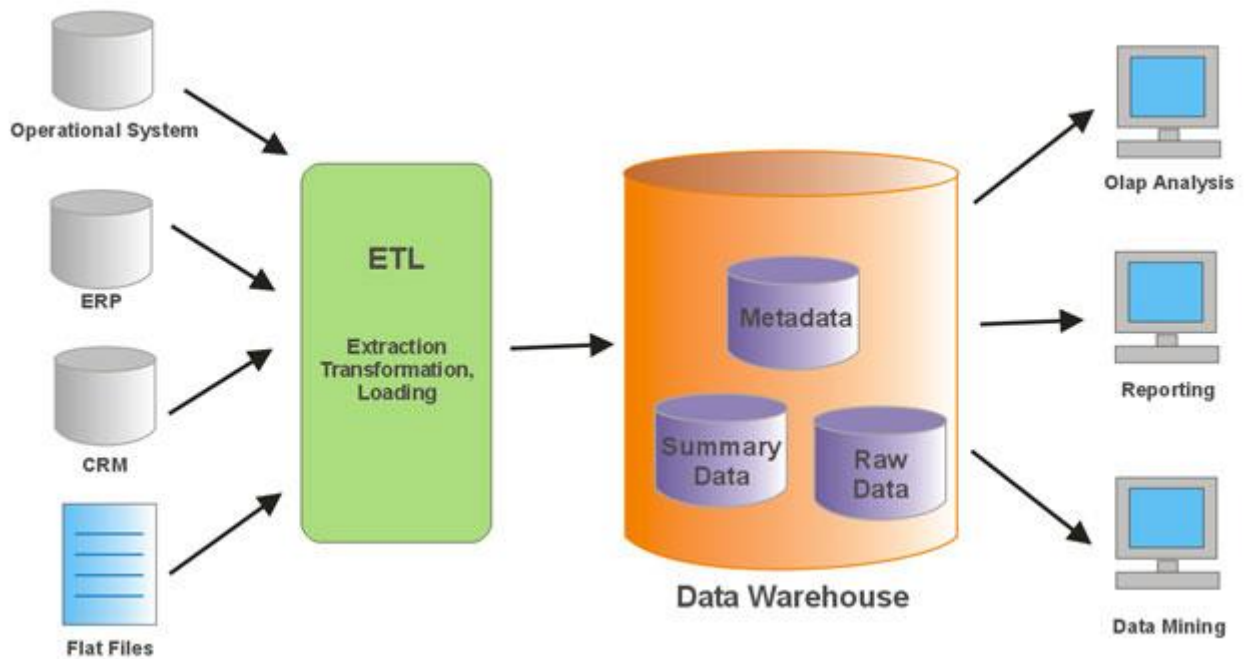
A implementação de uma ferramenta de *Business Intelligence* requer uma estruturação dos objetivos e melhorias propostas, tais como a relação entre os clientes e os riscos potenciais, alinhamento de metas, mensuração de objetivos propostos, e etc. Para Turban (2009) “hoje se reconhece que todas as informações de que os executivos necessitam podem estar em um bom sistema de informações empresariais baseado em *BI*”.

Segundo Barbieri (2011), pode-se considerar o *BI* como um mecanismo de suporte a tomada de decisão que contém três importantes passos: coleta de dados, geração de informação e tomada de decisão.

Um ponto da implementação de um sistema de *Business Intelligence* é a forma com que as organizações conseguem obter resultados através dela. Um sistema de *BI* pode ser entendido como a utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios de uma empresa (BARBIERI, 2011).

Na figura que se segue, é apresentado o processo de transformação de dados em uma visão analítica do *Business Intelligence*.

Figura 1 Processo de um *Business Intelligence*



Fonte: [www.dtech-education.com](http://www.dtech-education.com)

A figura 1 ilustra o processo de recolhimento das informações, sendo sequencialmente transformado e armazenado em um *Data Warehouse* capaz de gerir os dados em informações analisáveis.

Devido a necessidade de transformar dados em inteligência de negócio, o *Business Intelligence* surge como a maior ferramenta para apoio as decisões gerenciais através de técnicas e processos que auxiliam a isso.

## 2.4 DATA WAREHOUSE

Para o americano William H. “Bill” Inmon (1997), um cientista computacional, o *Data Warehouse* é uma coleção de dados relacionados a alguma área da empresa, organizados para dar suporte à decisão e baseados nas necessidades de um determinado departamento, ou seja, sem essa estrutura seria praticamente inviável a aplicação do conceito de *BI* em uma área de uma empresa. Um dos primeiros conceitos apresentados por Inmon (1997) foi que uma

estrutura de armazenamento central de dados, deve ser estruturada segundo um modelo Entidade-Associação (EA) e não questionável.

Segundo Turban (2009, p. 57) o *Data Warehouse* é “um conjunto de dados produzido para oferecer suporte à tomada de decisões”.

Na maioria dos bancos de dados atuais, a necessidade dos mesmos dá-se através do controle de dados, seja na entrada deles ou na armazenagem, até porque estes bancos estão sujeitos a uma constante modificação, inserção e remoção, todos efetuados por transações de aplicações ou softwares. No modelo *Data Warehouse* o foco é exclusivo para a saída dos dados, pois é através delas que se obtêm os resultados a serem analisados a fim de novas estratégias ou conclusões.

Um ponto característico vital do *Data Warehouse* é a utilização de uma tabela com dados relacionados a dimensão ‘tempo’, tais como hora, dia, mês, ano, bimestre, semestre, e etc. Geralmente os dados encontrados em um *DW* variam de um período de 5 a 10 anos, sendo determinados de acordo com os objetivos pretendidos pelas empresas ou clientes contratantes da ferramenta (INMON, 1997).

Para garantir o sucesso da aplicação do *DW*, deve se levar em conta algumas características:

- a) Orientado por assunto: Os dados são organizados de acordo com o assunto ou foco de uma empresa, por exemplo, o setor de marketing que irá constituir um modelo baseado em análise de vendas, análise de perfis de clientes, análise de vendas em potencial, e etc. Essa orientação por assunto permite uma visão abrangente sobre a organização (TURBAN, 2009);
- b) Não volátil: Assim que os dados forem inseridos no *DW* os mesmos devem estar ‘congelados’, sendo impossível de alterá-los ou atualizá-los, pois desta maneira todo o histórico sofreria com o impacto dessas modificações, prendendo consequentemente o histórico. (TURBAN, 2009);
- c) Em Tempo Real: “Os *Data Warehouses* mais novos oferecem acesso a dados e recursos de análise em tempo real, ou ativos” (TURBAN, 2009, pg. 58);
- d) Integrado: A padronização dos dados e suas respectivas unidades de medidas devem ser construídas a fim de um padrão em comum seja aplicado (INMON, 1997);

- e) Variável no tempo: A análise de um *DW* é decorrente da função tempo, logo que seu objetivo é a visualização de mudanças e variações de um determinado período de tempo (INMON, 1997).

### **2.4.1 Modelo Multidimensional**

Para Machado (2011, p. 79), “a modelagem multidimensional é uma técnica de concepção e visualização de um modelo de dados de um conjunto de medidas que descrevem aspectos comuns de negócios”.

O modelo multidimensional é uma forma de análise de informações através da compactação dos dados em dimensões, tem como principal objetivo, criar uma visão intuitiva para os usuários onde os dados ficarão armazenados em um ambiente de fácil localização e visualização.

Ele é caracterizado por três elementos principais, a tabela fato, as dimensões e as medidas.

#### **2.4.1.1 Tabela Fato**

A tabela fato é a principal parte do modelo multidimensional e é formada por um conjunto de chaves compostas de chaves estrangeiras, onde estão presentes todos os atributos numéricos e quantitativos do negócio analisado.

As tabelas de fatos contêm dados numéricos e somatórios, de forma agregada. Como possui um grande volume de dados, as tabelas de fatos fazem um agrupamento de registros para realizar e ter aumento de performance (KIMBALL, 1997).

De acordo com Machado (2010) cada fato representa um item, uma transação ou um evento de negócio e é utilizado para analisar o processo de negócio de uma empresa e tudo aquilo que reflete a evolução dos negócios do o dia a dia de uma organização.

A figura 2 mostra a representação da estrutura de uma tabela fato.

Figura 2 Tabela Fato



Fonte: <http://www.devmedia.com.br/introducao-ao-analysis-services-2005-parte-2/5730>.

A figura 2 mostra a tabela fato com as chaves estrangeiras oriundas de suas respectivas dimensões, contendo também as agregações encontradas em comum pelas mesmas.

#### 2.4.1.2 Tabela de Dimensões

As tabelas de dimensão representam atributos que formam o fato em si e é composta por um grande contexto textual, geralmente não possuindo atributos numéricos (MACHADO, 2010).

Estas tabelas também determinam um contexto de negócio como uma análise das vendas de produtos. Geralmente as dimensões utilizadas para a criação de um fato para produtos são.

- Tempo;
- Localização;
- Clientes;
- Vendedores;
- Cenários (Realizados e Projetados);

#### 2.4.1.3 Modelo *Star Schema* (Estrela)

O Modelo Estrela (*Star Schema*) foi criado e idealizado por Ralph Kimball, tendo como conceito uma espécie de transformação de modelos relacionais para modelos

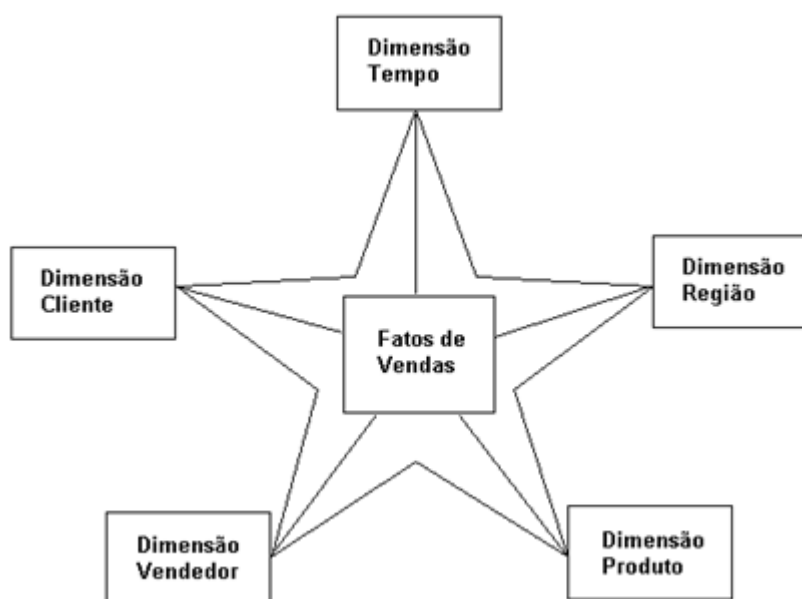
dimensionais, podendo ser implementado tanto em banco de dados relacionais quanto multidimensionais (KIMBALL e ROSS, 2002).

Machado afirma que (2010, p. 92):

Sua composição típica possui uma grande entidade central denominada fato (fact table) e um conjunto de entidades menores denominadas dimensões (dimension tables), arranjadas ao redor dessa entidade central, formando uma estrela.

Na ilustração abaixo é possível ver a simbologia de uma estrela e sua analogia ao modelo *Star Schema*, ou simplesmente, modelo estrela.

Figura 3 Modelo Estrela



Fonte: MACHADO 2010, p. 92.

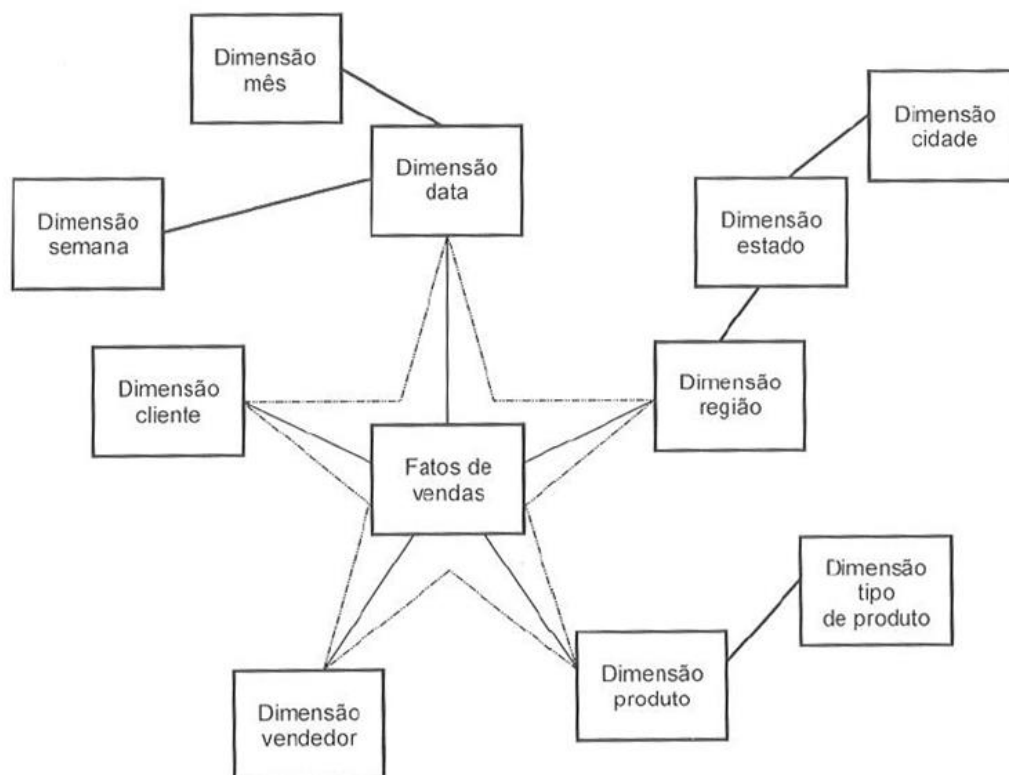
A figura 3 utiliza a simbologia de uma estrela para representar o meio como sendo a tabela fato, neste caso de vendas, com cada respectiva ponta para cada dimensão, nestes casos a dimensão de tempo, região, produto, vendedor e cliente.

#### 2.4.1.4 Modelo *SnowFlakes* (Flocos de Neve)

O modelo *SnowFlakes*, ou simplesmente 'Flocos de Neve', é também usualmente similar a uma estrela, porém com um complemento, ele é o resultado da decomposição de uma ou mais dimensões, formando uma hierarquia entre seus membros, ou seja, temos relacionamento também entre ambas (MACHADO, 2010).

Na figura 4 é possível ver as variações de uma dimensão no conceito “flocos de neve”.

Figura 4 Modelo *SnowFlake*



Fonte: MACHADO 2010, p. 94.

Assim como a figura 3, a figura 4 mostra a representatividade da estrela, porém, neste caso são exibidas as variadas dimensões oriundas de outras dimensões, por isso a analogia a um floco de neve.

#### 2.4.1.5 Granularidade

O mais importante aspecto do projeto de um *Data Warehouse* é a questão de sua granularidade. A granularidade diz respeito ao nível de detalhe ou de resumo contido nas unidades de dados existentes no *DW*. Quanto mais detalhe, mais baixo será o nível de granularidade e quanto menos detalhe, maior será o nível de granularidade (INMON, 1997, p. 45).



Inmon (1997, p. 143) afirma que “quando a granularidade de um *Data Warehouse* é apropriadamente estabelecida, os demais aspectos do projeto e implementação fluem tranquilamente, já quando ela não é estabelecida, todos os outros aspectos se complicam”.

### 2.4.2 Data Mart

Para Inmon (1997), um *Data Mart* é derivado de um *Data Warehouse*. Diferentemente do *DW*, o *Data Mart* é segmentado visando a organização e a especificação das áreas ou setores específicos da empresa, tais como a área de contabilidade, de finanças, de *marketing*, entre outros.

Diferentemente do *DW*, o *Data Mart* possui um alto índice de indexamento dos dados, pelo seu baixo volume específico para cada setor. Outro fator que o difere de um *Data Warehouse*, deve-se ao fato da pequena quantidade de dados e de sua estrutura ser modelada de acordo com o perfil do departamento do qual ela foi construída (INMON, 1997).

Outro grande precursor dos conceitos da armazenagem de dados e sistemas de apoio às decisões, o Professor PhD Ralph Kimball (1997), define que um *Data Warehouse* é o conglomerado de todos os *Data Mart's* dentro de uma empresa e que a informação deve ser armazenada sempre em um formato multidimensional.

Um *Data Mart* pode ser dividido em dois conceitos: *Data Mart Dependente* e *Data Mart Independente*.

O *Data Mart Dependente* é um subconjunto criado diretamente a partir da fonte de um *DW*, tendo a vantagem da utilização de um modelo já consistente, apresentando dados de qualidade (TURBAN, 2009).

Como forma de alternativa, muitas organizações utilizam uma versão de *Data Warehouse* reduzida em custo e escala, chamada *Data Mart Independente*. Esse conceito, segundo Turban (2009, pg. 58) é “um *Data Warehouse* pequeno, projetado para uma unidade estratégica de negócios (UEN) ou um departamento, mas cuja fonte não é um *EDW* (*Data Warehouse Empresarial*)”.

Na figura 5 se tem uma representação de alguns *Data Marts* específicos para cada setor de uma empresa.

Figura 5 Segmentações de um DW em *Data Mart's* específicos



Fonte: [www.chartio.com](http://www.chartio.com)

Na figura 5 é possível enxergar oito áreas, ou departamentos, de uma organização, sendo segmentados e, posteriormente, direcionados para um único DW centralizado.

## 2.5 DATA MINING

Devido ao crescente volume de dados que surgem diariamente nas empresas, o trabalho com certo nível de granularidade das informações torna-se um processo com um nível de complexidade enorme, sendo muita das vezes difícil de analisar ou manusear. Visando essa dificuldade, surge o conceito do *Data Mining*.

O *Data Mining* são um conjunto de técnicas que envolvem métodos matemáticos, algoritmos e heurísticos para o descobrimento de padrões e regularidades em grandes conjuntos de dados (WESTPHAL; BLAXTON, 1998).

Para realizar a interpretação desses dados é utilizado o conceito de *Data Mining* que segundo Carvalho (2001, p. 7) é “o uso de técnicas automáticas de exploração de grandes quantidades de dados de forma a descobrir novos padrões e relações que, devido ao volume de dados, não seriam facilmente descobertos a olho nu pelo ser humano”.

Em razão de termos, uma distinção de informações com um volume exorbitante de dados a serem armazenados, é necessário realizar análises precisas sobre como serão as principais fases do processo de *Data Mining*, por isso, podemos enumerar as fases do processo como (CARVALHO, 2001, p. 15):

1. Identificação de um problema ou definição de um objetivo a ser alcançado;
2. Descoberta de novas relações por técnicas de *Data Mining*;
3. Análise humana das novas relações descobertas;
4. Uso racional das novas relações descobertas;
5. Avaliação dos resultados;

## 2.6 ETL

Quando é implementado um *Data Warehouse* dentro de uma organização, são encontradas fontes de dados de diversos formatos, tipos e volumes. Em virtude disso, é necessário desempenhar um trabalho de unificação, transformação e padronização das informações já existentes dentro do modelo inicial da empresa.

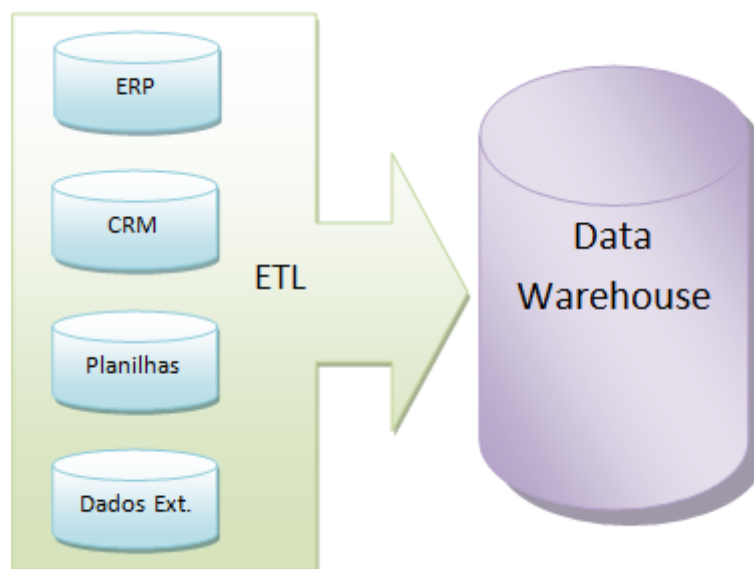
Para atuar sobre esse problema, utiliza-se uma ferramenta *ETL* (*Extract Transformation and Load*), que nada mais é que a transformação de bases iniciais em um modelo compacto de banco de dados que consiga reunir todas as informações necessárias de dentro de um *DW*.

Várias origens diferentes podem armazenar esses dados, e elas podem possuir formatos distintos, podendo ser um robusto SGBD ou até mesmo um simples arquivo texto ou arquivos DBF, entre outros possíveis (KIMBALL e ROSS, 2002).

Devido aos dados serem extraídos de fontes diferentes, a chance de termos uma incompatibilidade entre registros, ou dificuldade de padronização de informações é muito grande. Por isso, o esforço exigido na construção de um *DW* nesse momento é muito alto e oitenta por cento de todo esforço deve ser empregado neste processo (INMON, 1997).

Na figura 6 é possível ver a consolidação de variadas fontes pelo *ETL*.

Figura 6 Processo de uma consolidação de uma *ETL*



Fonte: [www.devmedia.com.br](http://www.devmedia.com.br)

O *ETL* utiliza diversas fontes diferentes como mostra na figura 6, originando uma centralização e agregação dos dados em um único *Data Warehouse*.

## 2.7 FERRAMENTAS DE FRONT-END

Ferramentas *Front-End* são ferramentas voltadas para o usuário final, a fim de gerar uma visão analítica e gerencial dos dados que compõem o *Data Warehouse* do qual a ferramenta estará integrada.

### 2.7.1 OLAP

O *OLAP* (*Online Analytical Processing*) é uma ferramenta que possibilita a análise de um grande volume de dados de diversas perspectivas diferentes. A utilização desta ferramenta pode ser feita por gestores de qualquer área ou setor de uma empresa, sempre com a finalidade da disponibilização de informações que irão auxiliar na tomada de decisão final. Um exemplo da aplicação do *OLAP* em uma empresa é na área financeira, desde análises dos fluxos de caixa, contas a pagar e a receber, orçamento, entre outros, assim como na área de

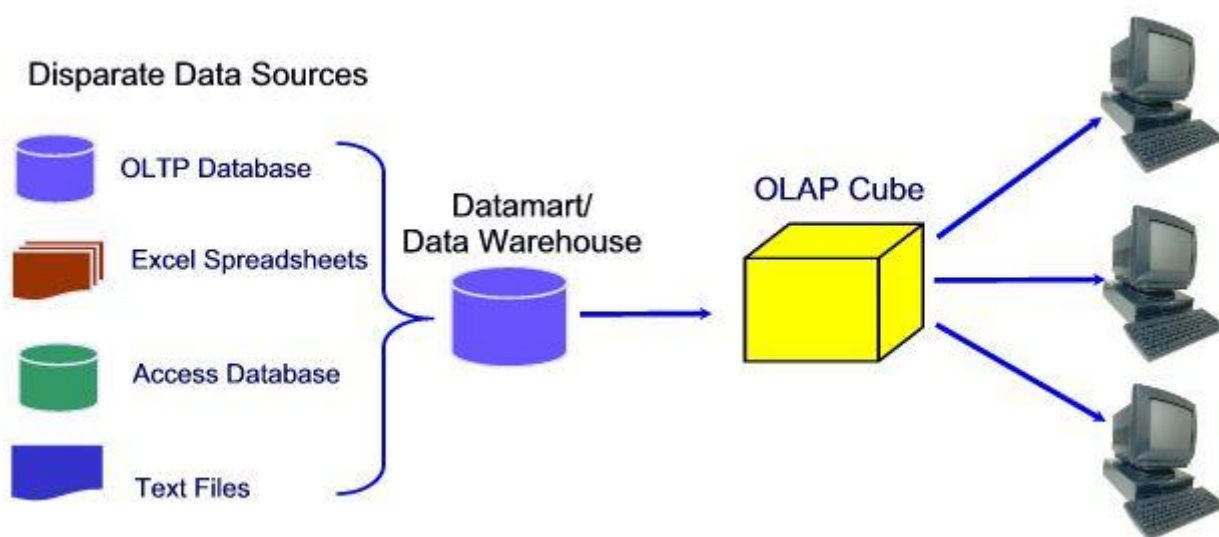
recursos humanos, como a análise de colaboradores, a aprovação de vagas oferecidas no mercado de trabalho de acordo com um perfil pré-determinado pela ferramenta de avaliação, entre outros (CRAMER, 2006).

Esta ferramenta foi criada com o principal objetivo de atender as variadas interações entre as várias dimensões, pois a mesma foi construída visando à análise de dados multidimensionais.

No *OLAP*, o usuário pode realizar a busca em diferentes níveis de profundidade ou granularidade, sendo que quanto menos específicos forem os filtros, menor será o tempo de resposta e no caso de um nível maior de granularidade, a um maior tempo de processamento da ferramenta.

A figura abaixo mostra os processos envolvidos para a consolidação de um cubo *OLAP*.

Figura 7 Transformação do *DW/Data Mart* para o cubo em *OLAP*



Fonte: [www.obs3.com](http://www.obs3.com)

Na figura 7, é possível enxergar diversas fontes como arquivos textos e planilhas eletrônicas, sendo consolidadas em um único centro de armazenagem, o *DW*, servindo de base e estrutura para a criação e alimentação dos dados do cubo *OLAP*, cujo é utilizado nas consultas analíticas pelos usuários finais.

Existem quatro arquiteturas distintas de *OLAP*, sendo em *ROLAP* (Relational Online Analytical Processing), *MOLAP* (Multidimensional Online Analytical Processing),

*HOLAP* (Hybrid Online Analytical Processing) e *DOLAP* (Desktop Online Analytical Processing):

*ROLAP*: Nesta arquitetura, os dados do *Data Warehouse* são armazenados nos próprios SGBD's relacionais, das quais as estruturas das tabelas são em modelos relacionais clássicas, funcionando em modelo semelhante a um software de uma aplicação simples (CARVALHO, 2004).

*MOLAP*: A requisição do usuário é feita diretamente ao servidor multidimensional com o seu respectivo tratamento específico (CARVALHO, 2004).

*HOLAP*: Como o próprio nome já diz, ele é híbrido, sendo a combinação entre o ROLAP e o MOLAP, permitindo a análise multidimensional dos dados armazenados em modelos multidimensionais, tanto quanto em modelos relacionais (THOMSEN, 2002).

*DOLAP*: A requisição do usuário é enviada para o servidor que o retorna em formato de um “cubo de dados”, para que o mesmo possa ser analisado na própria máquina do solicitante da requisição. Uma vantagem desse modelo é o processamento estar na mão do usuário solicitante e não do servidor que irá realizar os cruzamentos (CAVALCANTI; OLIVEIRA; MONTEIRO, 1998).

### 3 METODOLOGIA

Neste capítulo, são apresentadas todas as metodologias que foram baseadas para a construção deste projeto, visando referências e pesquisas que compõem uma consolidação de todo o material e ferramentas utilizadas para a construção do mesmo.

#### 3.1 CARACTERIZAÇÃO DO TIPO DE PESQUISA

Do ponto de vista da natureza, este projeto pode ser considerado uma pesquisa aplicada, pois tem como objetivo gerar uma aplicação prática, focando na solução de problemas específicos através da implementação de um sistema de *Business Intelligence*.

Com a relação da abordagem, de acordo com Silva e Menezes (2005), podemos classificá-la como uma pesquisa qualitativa, pois a interpretação dos dados por parte do pressuposto analisador é algo que não pode ser interpretada pelos números brutos, mas sim pelas conclusões, subjetivas ou não, do mesmo, embora este projeto esteja relacionado a números e análises de dados numéricos. Silva e Menezes (2005, p. 20) ainda destacam que “entre o mundo real e o sujeito, isto é um vínculo indissociável do mundo objetivo e a subjetividade do sujeito que não pode ser traduzida em números”.

Não podemos caracterizar essa pesquisa como quantitativa, pois não tem como objetivo, entender regras matemáticas e valores numéricos da aplicação proposta.

De acordo com os procedimentos técnicos, podemos ter as seguintes classificações:

- a) Pesquisa bibliográfica: pois é constituído de materiais já publicados como artigos ou livros;
- b) Estudo de caso: devido ao estudo exaustivo e profundo de um conjunto de dados, visando o máximo de detalhamento possível;

### 3.2 ETAPAS METODOLÓGICAS

As etapas metodológicas podem ser divididas das seguintes formas:

- a) Levantamento de informação para o processo decisório da área de RH;
- b) Construção do modelo dimensional que atenda os requisitos levantados;
- c) Mapeamento das fontes de dados e indicadores de origem que agregam valor a organização;
- d) Processo de extração, transformação e carga dos dados no *Data Warehouse*;
- e) Análise dos dados através de ferramentas *OLAP*, explanando os resultados obtidos;

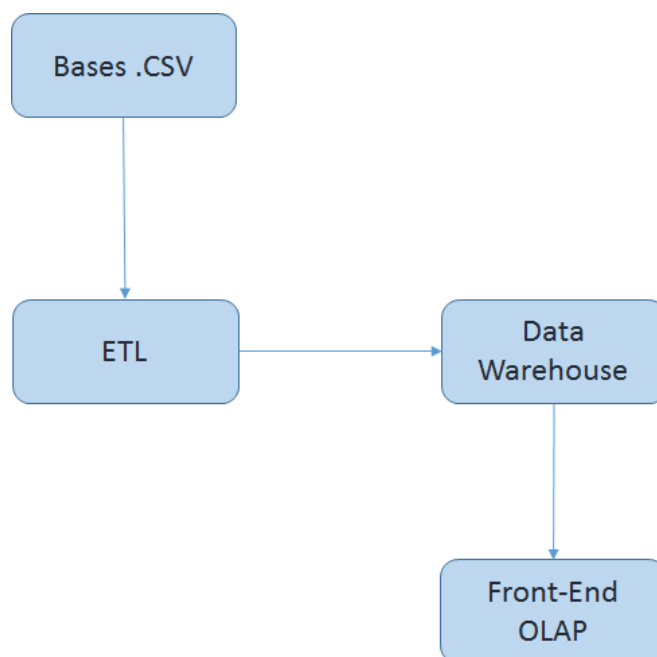
### 3.3 ARQUITETURA DA SOLUÇÃO

A arquitetura proposta visa o desenvolvimento da análise dos dados da gestão de recursos humanos. Essa análise consiste em quatro etapas: reunião das bases em *CSV*, construção de uma *ETL automatizada*, elaboração de um *Data Warehouse* e utilização de uma ferramenta *OLAP*.

Na figura 8 é possível ter uma visão sintética das principais etapas envolvidas na arquitetura da solução.



Figura 8 Arquitetura da solução



Fonte: Elaborada pelos autores.

**Bases CSV:** Dados compactados no formato *CSV* (separado por vírgulas), onde estes dados contêm as principais informações relacionadas aos colaboradores como seus respectivos dados pessoais, seu quadro na escala de ponto, rotatividade referente a demissões e admissões.

**ETL:** Extração e transformação da carga de dados *CSV*, cujos mesmos são tratados e preparados para a importação em um modelo de banco de dados.

**Data Warehouse:** Repositório de dados dos quais estão estruturados em um modelo multidimensional.

**OLAP:** Dados estruturados em uma visão voltada para a análise gerencial das informações extraídas do *Data Warehouse*.

### 3.4 DELIMITAÇÕES

Este projeto terá suas delimitações baseadas e especificadas apenas ao setor de RH de qualquer organização independente de sua área de atuação, tendo como objetivo o alcance de uma melhora na performance e no controle de gestão de pessoas, sem uma visão

voltada para as demais áreas da empresa, como o setor financeiro, administrativo ou até mesmo de TI.

Todas as ferramentas utilizadas para a construção da solução proposta são ferramentas *Open Source* (código aberto), ou seja, não há gastos com a parte de *softwares* para o desenvolvimento deste projeto.

## 4 MODELAGEM

O capítulo que segue, aborda toda a parte estrutural da solução proposta, tal como sua arquitetura e informações técnicas da aplicação e, além disso, são descritos todo o contexto e regras de negócios.

### 4.1 OBJETIVO GERAL DE NEGÓCIO

Visando a contemplação da área de RH, especificadamente a Gestão de Pessoas e seus indicadores operacionais, este projeto tem como princípio a coleta de dados relacionados aos atributos e indicadores dos funcionários, registros de ponto, aderência à jornada de trabalho, rotatividade dos colaboradores, entre outros, a fim de transformá-los em informações suscetíveis a análises, gerando valor as ideias corporativas e agregando uma análise diretiva a tomada de decisões e as segmentações de perfis dos funcionários.

Para isso, a sistematização das informações de Recursos Humanos dá-se de forma assertiva a adição de valores a área, sendo capaz de agregar mais responsabilidade e credibilidade aos profissionais do setor de recursos humanos, assim como aumentar o grau de importância do mesmo na organização como um todo.

A utilização desse modelo de controle das informações dá ao profissional da área a consciência de que além da simples visão de se ter um RH de forma mais sistêmica e informatizada, a área de recursos humanos poderá direcionar as prioridades da organização, assim como a do meio em si, em virtude de análises concluídas com base neste modelo automatizado, alavancando o potencial da área na contribuição de estratégias que visam atender a toda a empresa.

Com relação à importância de estratégias construídas no RH, o autor americano David Olson Ulrich (2003, p. 43) garante que:

Os profissionais da área se tornam parceiros estratégicos quando participam do processo de definição da estratégia empresarial, quando fazem perguntas que convertem a estratégia em ação e quando concebem práticas de RH que se ajustam à estratégia empresarial. Como parceiros estratégicos, os profissionais de RH devem ser capazes de identificar as práticas que fazem com que a estratégia aconteça.

## 4.2 CONTEXTO ATUAL DE NEGÓCIO

A área de Recursos Humanos tem um papel fundamental e estratégico para a empresa. Porém, muitas das mesmas ainda possuem a simples visão de que o RH tem um papel pouco decisivo na empresa, direcionando suas atividades apenas a contratação e recrutamento de profissionais. Dave Ulrich afirma que (2003, p. 35):

Tradicionalmente, a função recursos humanos passou mais tempo professando do que sendo profissional. Foi contaminada por mitos que a impedem de ser profissional. Não importa se esses mitos se originam do pessoal ou dos gerentes de RH. É hora de serem superados. De falar menos e fazer mais. Adicionar valor, não de redigir declarações de valor. Construir organizações competitivas, e não confortáveis. É hora de desempenho, e não pregação.

Com base nesta visão, pode-se concluir que o RH possui um papel que ainda é incapaz de tomar suas próprias decisões, sendo mediadores entre a ação e a estratégia inicial. Contudo, e baseado no grau de responsabilidade que a área tem, ou deveria ao menos possuir, os profissionais ainda são, na grande maioria, resistentes a mudanças, acreditando em velhos paradigmas culturais corporativos.

Ulrich (2003) afirma que as maiorias das atividades operacionais da área também são importantes para a corporação. A abordagem atual da área recursos de humanos visa atender o operacional de uma empresa, ignorando ou evitando o lado estratégico do qual deveria também, assumi-lo, se não incorporá-lo.

A estratégia, bem como o planejamento estratégico, atualmente é um ponto altamente relevante para as organizações, assim como para o mercado empresarial, tornando-se um diferencial competitivo entre as mesmas.

Com isso, o planejamento estratégico passou de centralizado na mão do setor administrativo e passou a ser incluído nas demais áreas que fazem parte de toda uma corporação, assumindo um papel contínuo e reciclável, perdendo a sua originalidade anual ou semestral que havia em tempos atrás.

### 4.3 CONTEXTO DO SISTEMA

A fim de auxiliar as decisões que envolvem a Gestão dos Recursos Humanos, o sistema possui módulos que tenham como foco principal, a análise do comportamento dos indicadores mapeados, visando estudar a influência deles, no resultado final do quadro de funcionários da organização.

Para realizar essa análise, será necessário realizar o processo de extração, através de bases compactadas que possuam todos os dados separados com informações variáveis a tempos de cada colaborador, ou seja, será necessário ter a informação diária de seus funcionários, o qual será o mais profundo grau de granularidade.

Essa separação de dados com a granularidade máxima por usuário e ‘tempo’, tem como finalidade a efetuação de cálculos ponderados, gerando valores quantitativos que consigam ser importados em um banco de dados em formato dimensional.

Após todos os cálculos efetuados, o sistema irá resgatar as informações do modelo dimensional através de uma ferramenta *OLAP*, da qual é responsável pela visualização em um formato analítico e detalhado de fatores que possam influenciar na análise de pontos específicos da organização com o objetivo de encontrar suas maiores oportunidades ou ameaças.

#### 4.3.1 Requisitos

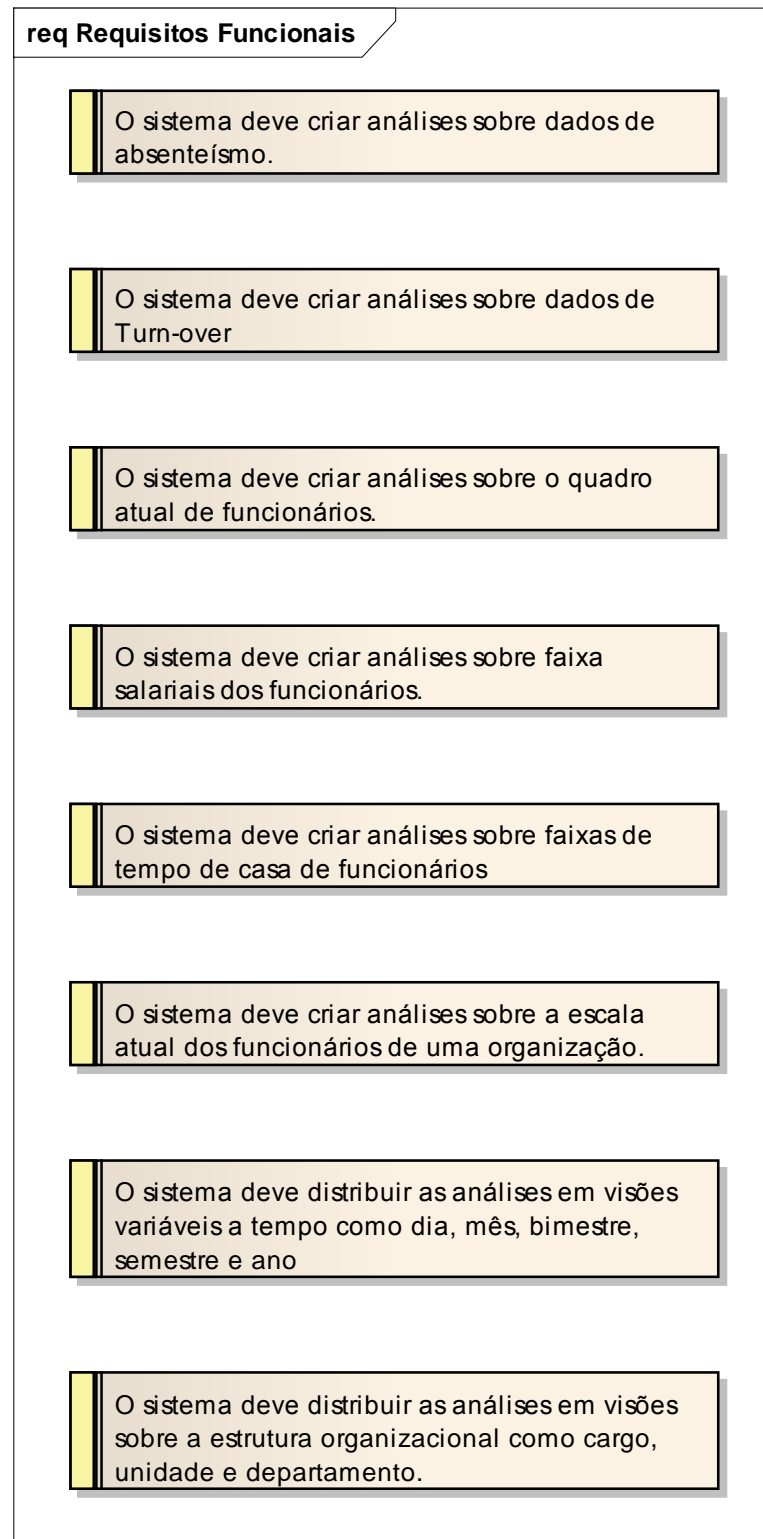
Nesta seção, é feita uma listagem de todos os requisitos que compõe a aplicação proposta. As descrições das funções e das restrições são os requisitos para o sistema; e o processo de descobrir, analisar, documentar e verificar essas funções é chamado de análise de requisitos (SOMMERVILLE, 2003, p.82).

##### 4.3.1.1 Requisitos Funcionais

Requisitos funcionais é a declaração de funções que o sistema deve fornecer como o sistema deve reagir a entradas específicas e como deve se comportar em determinadas situações. Em alguns casos, os requisitos funcionais podem também explicitamente declarar o que o sistema não deve fazer (SOMMERVILLE, 2003, p. 83).

Na figura 9 todos os requisitos funcionais da estrutura de *Business Intelligence* estão listados.

Figura 9 Requisitos Funcionais



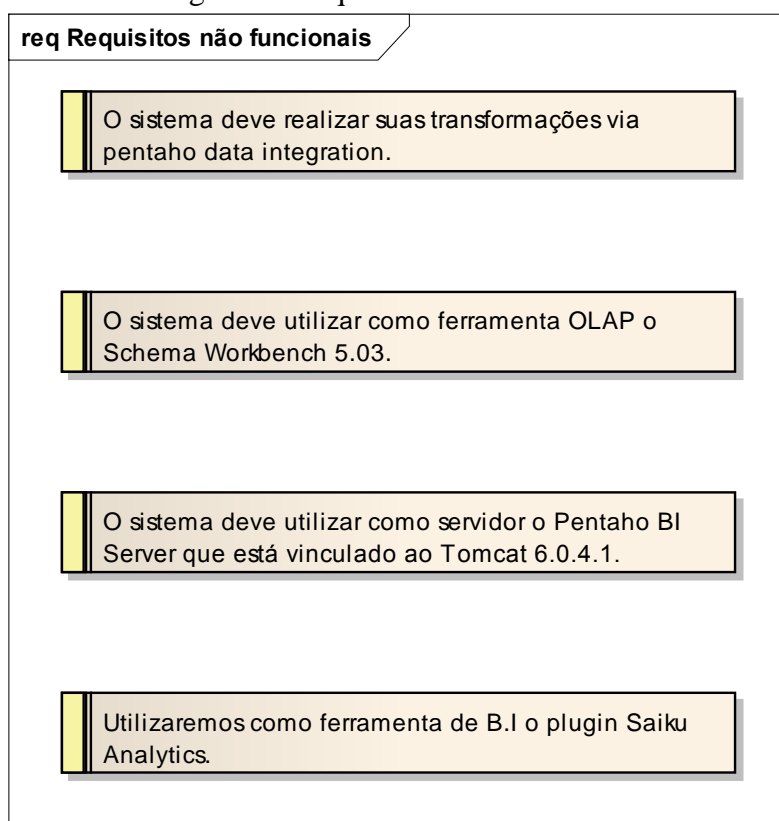
Fonte: Elaborada pelos autores.

#### 4.3.1.2 Requisitos Não Funcionais

São restrições sobre os serviços ou as funções oferecidas pelo sistema. Entre eles destacam-se restrições de tempo, restrições sobre o processo de desenvolvimento, padrões, entre outros (SOMMERVILLE, 2003, p.83)

Na imagem que segue abaixo os requisitos não funcionais, ou seja, os requisitos relacionados a aplicação, como por exemplo *softwares* de execução das tarefas de disponibilidade, manutenibilidade, e etc., estão representados da seguinte forma.

Figura 10 Requisitos Não Funcionais



Fonte: Elaborada pelos autores.

#### 4.3.2 Base de Dados

A base de dados utilizada está compactada no formato CSV, gerado por qualquer planilha eletrônica, neste caso, o *Calc* do pacote *LibreOffice*. Esse arquivo está dividido em sete guias onde cada uma representa um tipo de informação oriunda da área de recursos humanos em geral.

As informações das cargas utilizadas em nossa aplicação de *BI*, podem se encaixar em qualquer tipo de organização, pois possuem dados genéricos existentes em toda e qualquer área de qualquer empresa.

Abaixo segue uma especificação de cada guia presente nas cargas:

a) Guia de Funcionário: Relação com os principais dados de todos os funcionários. A tabela 1 mostra esses campos de forma mais detalhada.

Tabela 1 - Relação de registros da guia de Funcionário.

| Campo                        | Tipo              | Max. Caracteres | Aceita Nulos? |
|------------------------------|-------------------|-----------------|---------------|
| Matrícula                    | Texto             | 10              | Não           |
| Nome                         | Texto             | 50              | Não           |
| Salário                      | Numérico          | 20              | Não           |
| Matrícula do primeiro Gestor | Numérico          | 20              | Não           |
| Matrícula do segundo Gestor  | Numérico          | 20              | Não           |
| Matrícula do terceiro Gestor | Numérico          | 20              | Não           |
|                              | Texto             | 50              | Não           |
| Departamento                 | Texto             | 50              | Não           |
| Unidade                      | Texto             | 50              | Não           |
| Naturalidade                 | Texto             | 50              | Não           |
| Endereço                     | Texto             | 50              | Não           |
| Município                    | Texto             | 50              | Não           |
| Estado                       | Texto             | 50              | Não           |
| Data de Nascimento           | Data (dd/mm/aaaa) | 10              | Não           |
| Data de Admissão             | Data (dd/mm/aaaa) | 10              | Não           |
| Data de demissão             | Data (dd/mm/aaaa) | 10              | Sim           |
| Sexo                         | Texto             | 1               | Não           |

Fonte: Elaborada pelos autores.

b) Guia de Status Ponto: Listagem de todos os status possíveis de um colaborador em um determinado dia. Esse status determina se o colaborador veio trabalhar ou não, por isso devemos ter duas colunas indicando se ele será considerado como falta e como presença esperada de trabalho, esses campos devem estar na estrutura a da tabela 2.



Tabela 2 - Relação de registros da guia de Status Ponto.

| Campo    | Tipo  | Max. Caracteres | Aceita Nulos? |
|----------|-------|-----------------|---------------|
| Status   | Texto | 50              | Não           |
| Presença | Texto | 3               | Sim           |
| Falta    | Texto | 3               | Sim           |

Fonte: Elaborada pelos autores.

c) Guia de Ponto: Relação com as informações de ponto dos funcionários, a tabela 3 ilustra esses campos.

Tabela 3 - Relação de registros da guia de Ponto.

| Campo                        | Tipo              | Max. Caracteres | Aceita Nulos? |
|------------------------------|-------------------|-----------------|---------------|
| Matrícula                    | Texto             | 10              | Não           |
| Horário de Início da Entrada | Hora (HH:mm:ss)   | 8               | Não           |
| Horário de Fim da Escala     | Hora (HH:mm:ss)   | 8               | Não           |
| Data                         | Data (dd/mm/aaaa) | 10              | Não           |
| Horário de Entrada           | Hora (HH:mm:ss)   | 8               | Não           |
| Horário de Saída             | Hora (HH:mm:ss)   | 8               | Não           |
| Status do Ponto              | Texto             |                 | Não           |
| Início da Pausa              | Hora (HH:mm:ss)   | 8               | Não           |
| Início da Pausa              | Hora (HH:mm:ss)   | 8               | Não           |

Fonte: Elaborada pelos autores.

Os dados coletados nesse formato foram retirados de uma empresa de Telemarketing, sendo que todos eles foram manipulados com o intuito de preservar a integridade das informações originais, tanto da organização quanto dos colaboradores da mesma.

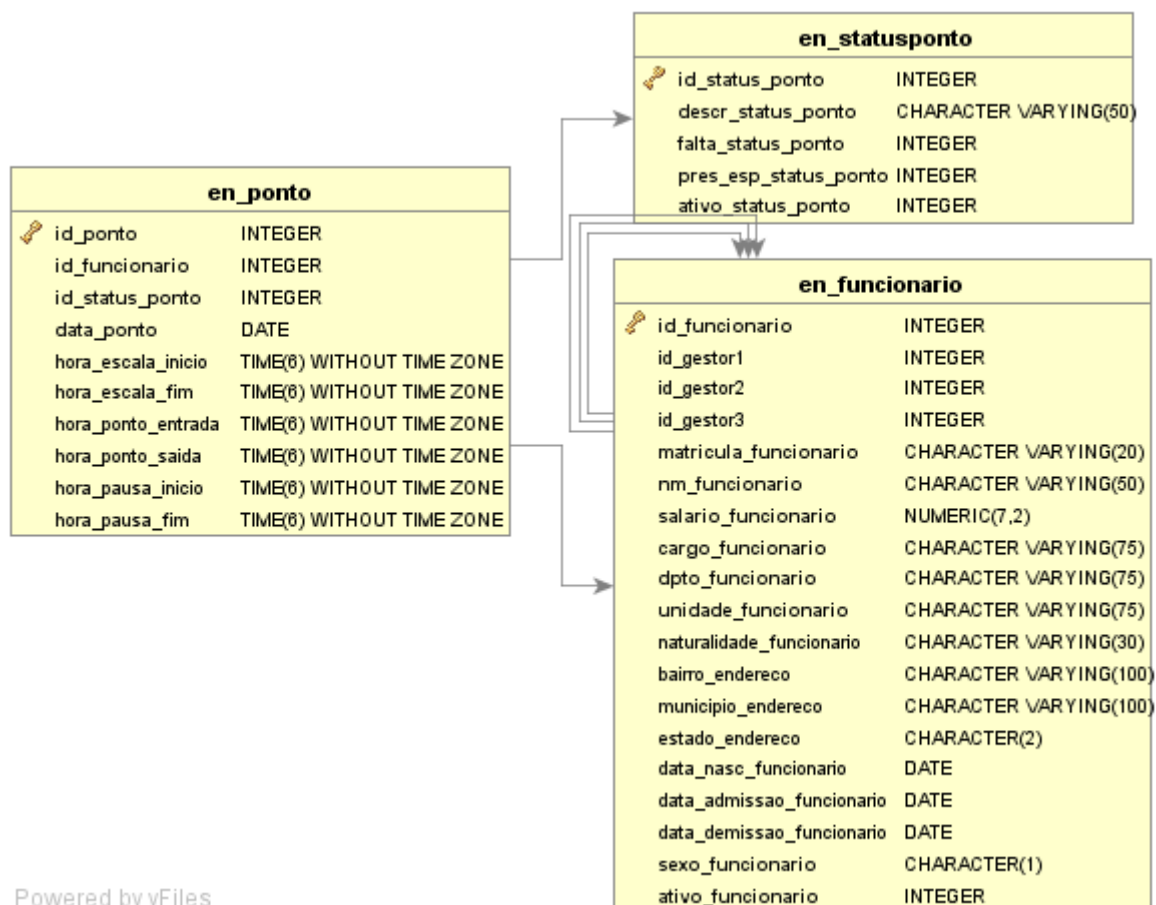
A base foi extraída de uma intranet corporativa, onde está disponível a todos os colaboradores essas informações e ela são constituídas com cerca de aproximadamente 3.000 funcionários.

### 4.3.3 Modelo Relacional

Para facilitar o processo de análise e construção de um modelo dimensional, iremos inicialmente manipular os dados em um formato de estrutura relacional, visto que este modelo mais simples possibilita a abertura a quaisquer levantamentos e agregações de valores com suas respectivas granularidades.

Na figura 11 são apresentadas as estruturas de cada tabela presente no modelo relacional.

Figura 11 Arquitetura do Modelo Relacional



Fonte: Gerada pelo aplicativo *DbVisualizer Free*.

Neste modelo relacional, há apenas uma tabela com uma estrutura relativamente simples, conceituando apenas a referência ao dado em si:

- en\_StatusPonto:** Responsável por armazenar todos os *status* de ponto para cada dia de trabalho de um funcionário, avaliando a sua presença, falta, folga, férias, atestado médico, ausência justificada e etc. Além disso, nessa tabela teremos a correspondência de cada status com o seu significado para uma presença esperada e falta;

Além desta, ainda há outras com um peso mais relevante para a consolidação das informações, das quais são extraídas as informações relacionadas ao ponto e aos usuários. Estas tabelas são:

- a) en\_Funcionario: Responsável por armazenar todos os dados de um usuário, possuindo atributos como as referências as outras tabelas de listagens, aos seus respectivos Gestor Imediato 1, Gestor Imediato 2, Gestor Imediato 3, a sua matrícula, nome, salário bruto, cargo, departamento, unidade, naturalidade, atributos relacionados ao endereço, data de nascimento, data de admissão na empresa, a data de demissão, sexo e o campo ativo. No caso da data de demissão, se o funcionário estiver ativo, este campo irá constar como vazio (*NULL*) e caso contrário, o mesmo irá armazenar a data de desligamento do mesmo;
- b) en\_Ponto: Responsável por armazenar os registros de ponto dos funcionários, fazendo uma relação ao funcionário, a escala atribuída, o status ponto registrado de acordo com o ponto, a data do registro, o horário de entrada e saída do ponto e o horário inicial e final das pausas registradas pelo funcionário;

Essas estruturas são praticamente equivalentes às apresentadas nas cargas, porém com uma tratativa já realizada e validada. Além disso, os dados em estruturas no banco de dados permitem um melhor gerenciamento e controle sobre os mesmos.

#### **4.3.4 Modelo Dimensional**

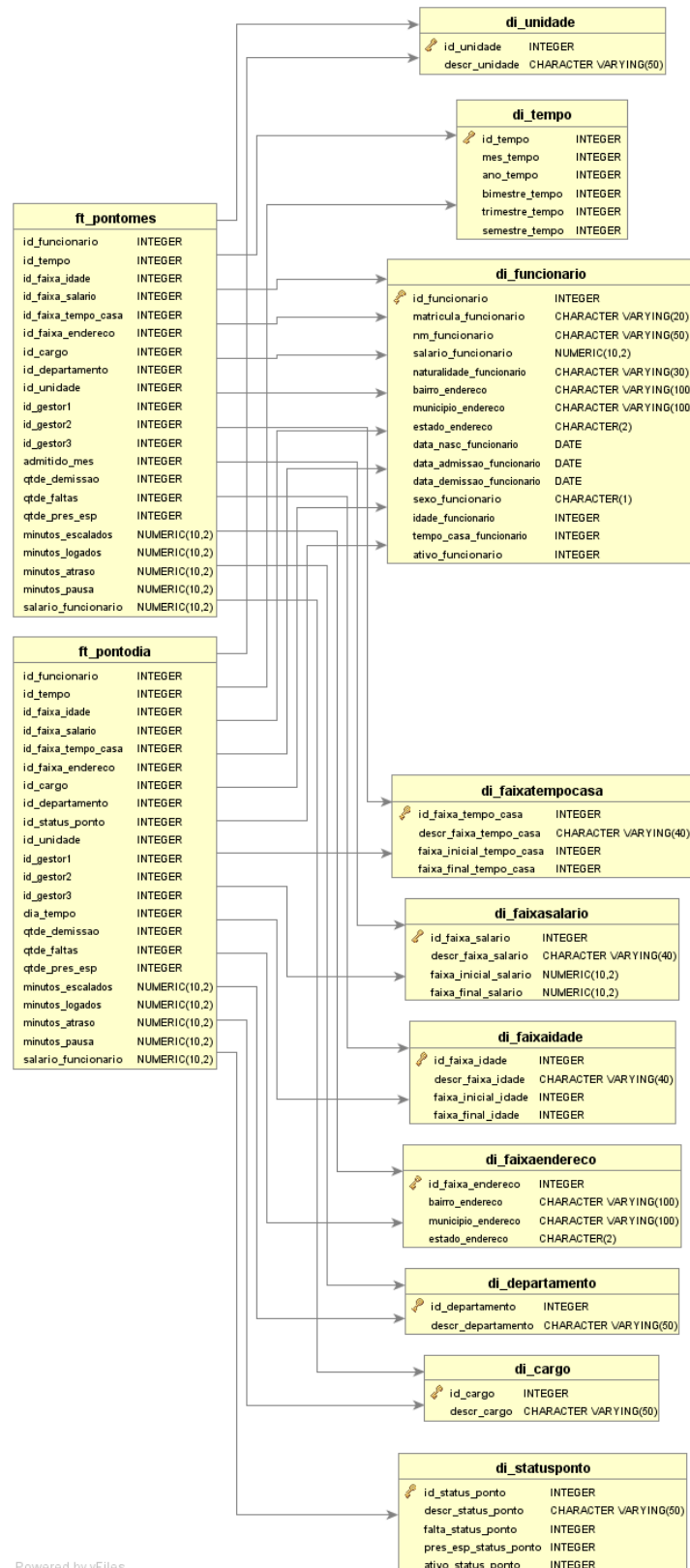
O modelo dimensional implantado será o modelo estrela (*Star Schema*) devido à adequação a diferentes tipos de negócios desde pequeno até grande porte com um ótimo desempenho de processamento e uma fácil rastreabilidade.

Segundo Turban “o esquema estrela oferece tempo extremamente rápido de resposta a consultas, simplicidade e facilidade de manutenção para estruturas de bancos de dados somente leitura” (TURBAN, 2009, p. 77).

Com base nesse conceito, a estrutura dimensional é construída em virtude de uma estrutura relacional já elaborada e alimentada com os dados diários.

Na figura 12 são apresentadas as estruturas de cada tabela presente no modelo dimensional.

Figura 12 Arquitetura do Modelo Dimensional



Powered by yFiles

Fonte: Gerada pelo aplicativo *DbVisualizer Free*.

No modelo dimensional, assim como no relacional, também há tabelas (dimensões) com suas respectivas estruturas simples, conceituando apenas a referência ao dado em si, como as tabelas:

- a) di\_Cargo: Dimensão responsável pela armazenagem de todos os cargos de todos os funcionários, permitindo ter a visão dos dados agregados por cargos da empresa;
- b) di\_Departamento: Dimensão responsável pela armazenagem de todos os departamentos que a empresa possui, como por exemplo, administrativo, financeiro, operacional e etc., permitindo ter a visão dos dados agregados por departamento;
- c) di\_FaixaEndereco: Dimensão responsável pela armazenagem de dados relacionados ao endereço, contemplando todas as possíveis associações de bairro, município e estado a todos os funcionários presentes no modelo relacional;
- d) di\_FaixaIdade: Dimensão responsável pela armazenagem das principais faixas etárias dos funcionários, permitindo a análise de informações relacionadas as faixas etárias dos funcionários, como por exemplo, qual a idade que possui um percentual de *turnover* maior;
- e) di\_FaixaSalário: Dimensão responsável pela armazenagem das principais faixas salariais dos funcionários, permitindo a análise de informações relacionadas as faixas salariais dos funcionários, como por exemplo, qual a faixa salarial de funcionários tende a possuir um número menor de faltas;
- f) di\_FaixaTempoCasa: Dimensão responsável pela armazenagem das principais faixas de tempo de casa em total de dias dos funcionários, permitindo a análise de informações relacionadas as faixas por tempo de casa dos funcionários, como por exemplo, qual a variação da quantidade de dias que um funcionário possui na empresa;
- g) di\_Funcionário: Dimensão responsável pela armazenagem de todos os funcionários da empresa, com seus principais campos como matrícula, nome, naturalidade, salário bruto, naturalidade, atributos relacionados ao endereço, data de nascimento, data de admissão na empresa, a data de demissão, sexo e o campo ativo. No caso da data de demissão, se o funcionário estiver ativo, este

campo irá constar como vazio (*NULL*) e caso contrário, o mesmo irá armazenar a data de desligamento do mesmo;

- h) di\_StatusPonto: Dimensão responsável pela armazenagem de todos os *status* de ponto para cada dia de trabalho de um funcionário, possuindo o mesmo conceito da tabela estrutural no modelo relacional;
- i) di\_Tempo: Dimensão responsável pela armazenagem de todos os fatores envolvidos a relação tempo, com os campos mês, ano, bimestre, trimestre e semestre, permitindo assim visões agregadas a mês e ano, a algum bimestre, entre outros.;
- j) di\_Unidade: Dimensão responsável pela armazenagem de todas as unidades da empresa, permitindo ter a visão dos dados agregados por unidades;

Para a agregação das informações, há duas tabelas ‘fatos’ que são responsáveis pelo armazenamento dos dados agregados e possuem relação com todas as dimensões, permitindo quaisquer cruzamentos de dados com suas dimensões e consolidações, gerando assim informações precisas em diversas visões. As tabelas ‘fatos’ são:

- a) ft\_PontoDia: Tabela responsável pelas consolidações dos dados de todas as dimensões envolvidas na estrutura dimensional. Esses dados estão agregados nas granularidades máximas, cujas são o funcionário e o dia de referência dos dados, permitindo uma visão diária de cada funcionário, assim como uma visão por períodos de intervalos inferiores a um mês, como semana e intervalos específicos;
- b) ft\_PontoMes: Essa tabela possui exatamente o mesmo conceito de agregação que a ‘ft\_PontoDia’. Contudo, o campo ‘dia’ não faz parte desta estrutura o que acaba tendo um intervalo de agregação maior, na visão mínima de mês. Essa visão com inferioridade menor que mês, permite cruzamentos e em intervalos anuais e semestrais, além de possuir um tempo de processamento e resposta menor do que se compararmos com intervalos idênticos a tabela ‘fato’ com a visão diária.

#### 4.3.5 Indicadores Mapeados

Para realizar a análise dos dados na ferramenta *OLAP* foram mapeados alguns indicadores que agregam valor para a área de Gestão dos Recursos Humanos:

- Tempo de casa;
- *Turnover*;
- Absenteísmo;
- Distribuição dos Cargos;
- Quadro de Funcionários;
- Administração das escalas de trabalho;
- Atrasos;
- Pausas;

Esses indicadores podem ser vistos com mais ênfase no capítulo de desenvolvimento, onde são cruzados e gerando informações para auxiliar a tomada de decisão da organização através de análises de Business Intelligence como já foi falado anteriormente.

## 5 DESENVOLVIMENTO

O capítulo abrange sobre todo o desenvolvimento da aplicação de *Business Intelligence*, assim como seu escopo e sua estrutura de abrangência.

### 5.1 FERRAMENTAS E TECNOLOGIA

Para o desenvolvimento desta aplicação de *Business Intelligence*, foram utilizadas ferramentas gratuitas, ou seja, *Open Source* (código aberto), a fim de não gerar nenhum custo durante este processo, possibilitando assim o alcance de um maior potencial de avaliações e melhorias das mesmas. As ferramentas utilizadas durante todo o processo de desenvolvimento e suas respectivas funções são exibidas a seguir.

#### 5.1.1 PostgreSQL

É um SGBD (sistema de gerenciamento de banco de dados) gratuito para desenvolvedores, porém não para empresas, capaz de armazenar dados e manter uma estrutura relacional e dimensional para a armazenagem dos mesmos e a extração de informações.

Um ponto forte do *PostgreSQL* é que o mesmo possui uma *interface* amigável, denominada *pgAdmin* da qual torna mais fácil e intuitiva sua apresentação dos dados.

Link de acesso: <http://www.postgresql.org/>.

#### 5.1.2 Pentaho Data Integration (PDI)

Também conhecido apenas como *kettle*, é uma ferramenta *ETL* da qual faz parte de um conjunto, ou pacote, de outras aplicações pertencentes ao grupo *Pentaho*. O *PDI* é gratuito para desenvolvedores de sua comunidade, porém para utilização corporativa o mesmo passa a ser pago. Essa poderosa ferramenta permite a conversão de dados de diversas fontes e origens, em dados trabalhados de acordo com a necessidade e o desenvolvimento das



mesmas, sendo capaz de possuir uma vasta forma de armazenagem, como por exemplo, um banco de dados.

Além disso, o *Pentaho Data Integration* possui uma vasta gama de relacionamentos com os dados, possuindo integrações com linguagens de programação, como é o caso do *Javascript*.

No website da empresa, é possível encontrar toda a documentação referente os componentes encontrados na ferramenta, exibindo o seu objetivo, seus parâmetros, suas premissas e funções.

Essa ferramenta pode ser instalada individualmente das demais ferramentas do pacote, porém possui um único requisito que é a obrigatoriedade de se ter instalado na máquina o *Java Runtime* (máquina virtual *Java*).

Link de acesso: <http://community.pentaho.com/projects/data-integration/>.

### **5.1.3 DBVisualizer Free**

Ferramenta de extração, geração e visualização dos modelos de estruturas do banco de dados em imagens.

Devido ao seu simples manuseio e seu *design* intuitivo, essa aplicação é eficaz em termos de fácil visualização de um banco de dados, bem como as relações entre tabelas.

Link de acesso: <https://www.dbvis.com/>.

### **5.1.4 LibreOffice Calc**

Um arquivo *CSV* nada mais é que um simples arquivo de texto com suas separações de colunas representadas por algum caractere, neste caso o ponto e vírgula.

Com isso, para uma melhor visualização e manutenibilidade destes dados, resumidos em cargas com formatos *CSV*, utilizamos a ferramenta *Calc* do pacote *LibreOffice* da qual é totalmente gratuita tanto para desenvolvedores e usuários comuns, quanto para organizações empresariais.

O *Calc* é uma planilha eletrônica com a finalidade de criação de tabelas e planilhas, permitindo ao seu utilizador a agregação de funções matemáticas a mesma, gerando assim análises e gráficos, relatórios e insumos para quem o utiliza.

Link de acesso: <https://pt-br.libreoffice.org/>.

### **5.1.5 Mondrian Schema Workbench**

O *Schema WorkBench* é uma ferramenta da plataforma *Pentaho* onde pode ser feita a preparação dos dados oriundos da arquitetura dimensional para o cubo de dados, nela mapeamos as medidas, níveis e hierarquia dos dados importados.

Link de acesso: <http://mondrian.pentaho.com/documentation/workbench.php>.

### **5.1.6 Pentaho BI Server**

O *Pentaho BI Server* é uma ferramenta onde é possível realizar as análises de negócios, além de que ela possibilita a criação de tabelas '*JPIVOT*' que é uma estrutura dinâmica de apresentação de dados nesse formato e também de gráfico, todos os dados são oriundos do que foi mapeado no *Schema WorkBench*. O sistema roda a partir do *Apache Tomcat 6.0* e serve também como servidor da aplicação.

Link de acesso: <http://community.pentaho.com/>.

### **5.1.7 Saiku Analytics**

O *Saiku Analytics* é um *plugin* da plataforma *Pentaho* integrado com o *BI Server*, ele se destaca na usabilidade para o usuário final, dos dados gerados no *Mondrian Schema Workbench*.

Link de acesso: <http://www.meteorite.bi/>.

## **5.2 APRESENTAÇÃO DA APLICAÇÃO**

Como esta aplicação envolve dois processos, um de construção dos dados e outro da apresentação dos mesmos, abaixo é apresentado dois tópicos com o detalhamento e as principais funções executadas em cada etapa.

### 5.2.1 Construção das Cargas

Através de um mapeamento das necessidades de informações requeridas pelo setor de Recursos Humanos, uma carga no formato *CSV*, separado por ponto e vírgula, é o primeiro passo que ocorre em todo o processo de transformação dos dados.

Estes arquivos, denominado cargas, precisam estar em um determinado formato e estrutura já considerada e desenvolvida. Entretanto, as fontes dos dados originais não fazem qualquer alteração de impacto nestes arquivos, visto que os mesmos serão tratados somente pela ferramenta de *ETL*, ou seja, se as bases de dados de um cliente de RH forem armazenadas em bancos, planilhas eletrônicas ou até mesmo em papéis, isso não será um problema para a aplicação, desde que haja algum responsável pela construção e disponibilização destas cargas, previamente formatadas, a aplicação de *ETL*. As cargas precisam possuir as formatações estruturais apresentadas no capítulo 4, no modelo relacional.

Além disso, a consolidação destas cargas, bem como seu envio e disponibilização das mesmas para a importação, fica de responsabilidade por parte da área de RH, sendo responsável pelos dados, porém não pelas informações geradas.

### 5.2.2 Construção e Desenvolvimento dos Dados Relacionais

O modelo relacional é o primeiro modelo do qual os dados passam a ser estruturados antes de serem transformados em dados com estruturas multidimensionais. Esta estrutura relacional viabiliza uma fácil compreensão dos mesmos, possuindo uma maior escalabilidade e facilidade na conversão dos dados.

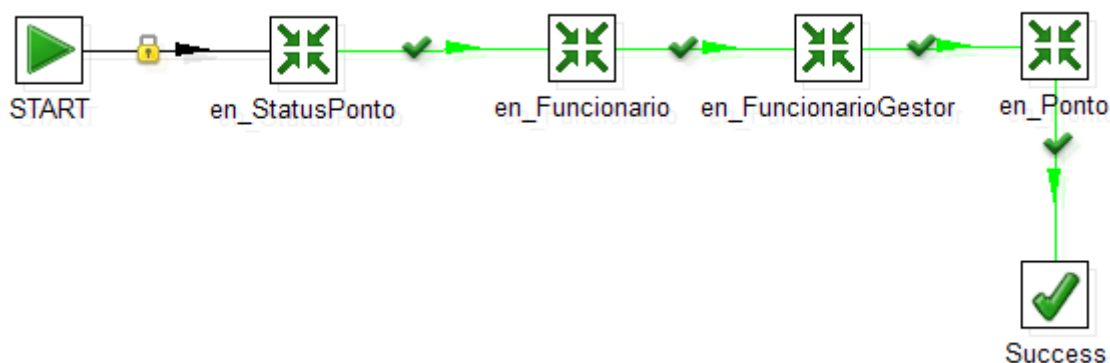
Além disso, as manutenções das transformações tornam-se mais simples e assertivas, uma vez identificado um problema relacionado na transformação, checa-se primeiramente se as tabelas relacionais foram previamente populadas em uma única *database*, sem a necessidade de olhar individualmente cada arquivo.

Entretanto, caso um problema esteja relacionado com a estrutura relacional, as cargas deverão ser checadas, analisadas e apresentadas ao usuário responsável pelas extrações e construções das mesmas.

O modelo relacional é alimentado pela própria *ETL*, da qual busca os dados nas cargas e os converte nesta estrutura.

Na figura 13 se tem uma imagem de todas as transformações envolvidas no processo da construção do modelo relacional e inclusas no *Job* da mesma.

Figura 13 Transformação completa do modelo relacional



Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

Dentre as transformações vistas acima que a ferramenta realiza, ficam em destaques as referentes aos funcionários e as referentes aos pontos registrados.

#### 5.2.2.1 Transformação dos Dados dos Funcionários

As transformações dos dados de funcionários são divididas em duas: uma que contempla os funcionários em si com seus respectivos atributos individuais e outra que contempla sua hierarquia, possuindo três níveis hierárquicos acima do seu.

O processo de funcionário ocorre inicialmente pela leitura e validação dos campos da carga pela *ETL*, posteriormente por uma seleção específica dos dados e uma ordenação e unificação dos funcionários pelo campo matrícula, da qual permite que não haja duplicidade de funcionários na carga gerada.

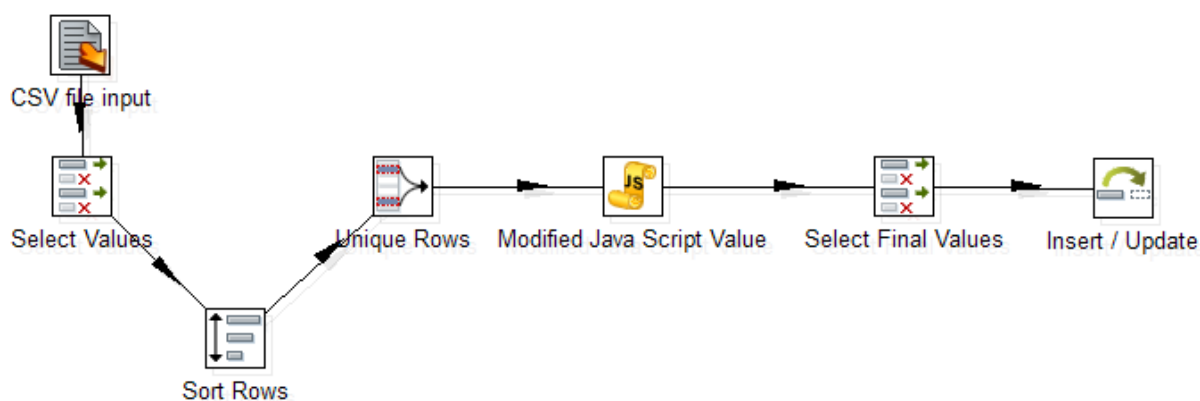
Feito isso, uma busca pelas chaves primárias correspondentes no banco de dados são realizadas, buscando a chave primária das tabelas de cargo, departamento e unidade. Caso uma chave primária não seja localizada pelo fato de não haver uma correspondência do campo no banco de dados, o registro será descartado da importação e só voltará a ser incluso após a correção do mesmo na carga e o processo de transformação executado novamente.

Após estas buscas, um código desenvolvido em *Javascript* checa se o usuário a ser importado está ativo ou não, com base em sua data de demissão, aonde um usuário que não possua nenhum registro neste campo será considerado ativo e caso tenha alguma data de demissão registrada, o mesmo passa a estar inativo.

Por último, outra seleção dos dados filtra todo o procedimento realizado e inicia o processo de inclusão dos registros no modelo relacional. Esta inclusão pode ser uma nova inclusão ou uma alteração, visto que usuários que já possuem sido cadastrados, serão atualizados e os demais inseridos como novos.

Na imagem abaixo todo o processo de transformação de funcionário é representada pelo procedimento desenvolvido no *Pentaho Data Integration*.

Figura 14 Transformação dos dados dos funcionários

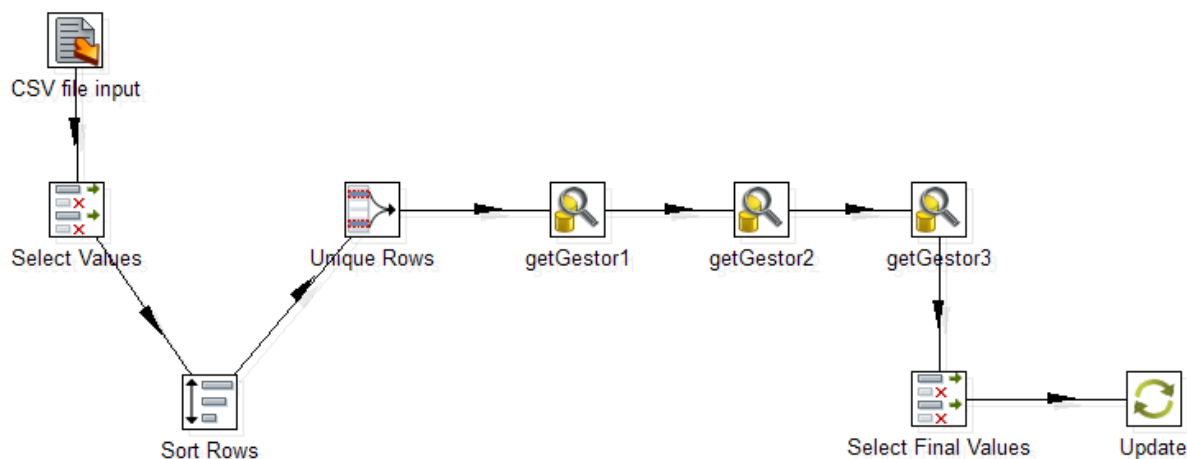


Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

Após esse procedimento, outro relacionado à hierarquia dos funcionários é iniciado, começando pela leitura e validação da carga, seleção dos campos, ordenação e unificação dos registros, retorno das chaves primárias dos funcionários, dos gestores dos funcionários inseridos na carga, seleção final dos valores retornados e a atualização hierárquica no banco de dados da tabela de funcionários.

Na figura 15, todo o processo de transformação da hierarquia dos funcionários é representada pelo procedimento desenvolvido no *Pentaho Data Integration*.

Figura 15 Transformação dos dados hierárquico dos funcionários



Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

#### 5.2.2.2 Transformação dos Dados de Registros de Ponto

Assim que os dados referentes aos funcionários forem transformados, inicia-se o procedimento de transformação dos respectivos registros de pontos, visto que cada registro estará vinculado diretamente a um funcionário e uma data específica.

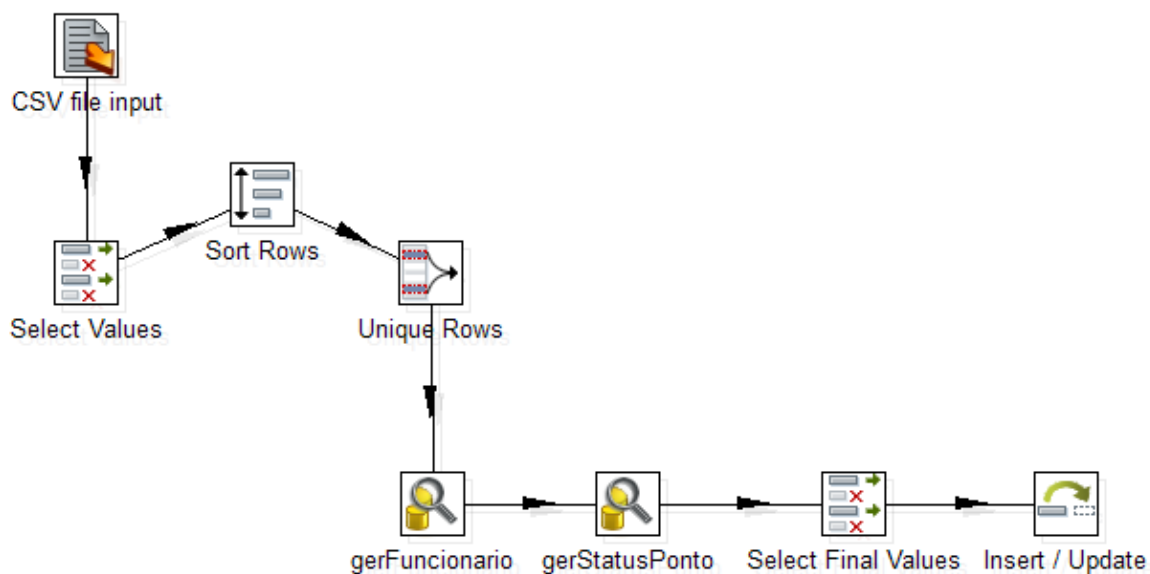
O processo dá-se inicialmente pela leitura e validação da carga CSV, posteriormente seguida pela seleção dos principais campos, assim como a ordenação e a unificação dos registros, garantindo assim a unicidade dos registros de ponto na respectiva tabela.

Em seguida, a ferramenta de transformação inicia suas buscas pelas chaves primárias dos funcionários, através do campo matrícula, a busca pelas chaves dos status de ponto, através dos status registrados na carga e a busca pelas chaves das escalas atribuídas no arquivo, localizando assim qual o horário inicial e final das escalas de trabalhos dos funcionários para as respectivas datas.

Por último, uma nova seleção ocorre seguida por uma inserção, ou atualização baseada nos campos de *ID* do funcionário, *ID* da escala, *ID* do status de ponto, o horário de entrada e saída registradas para o ponto e para as pausas pelo funcionário.

Na figura 16, todo o processo de transformação dos registros de pontos dos funcionários é representada pelo procedimento desenvolvido no *Pentaho Data Integration*.

Figura 16 Transformação dos dados de registros de ponto



Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

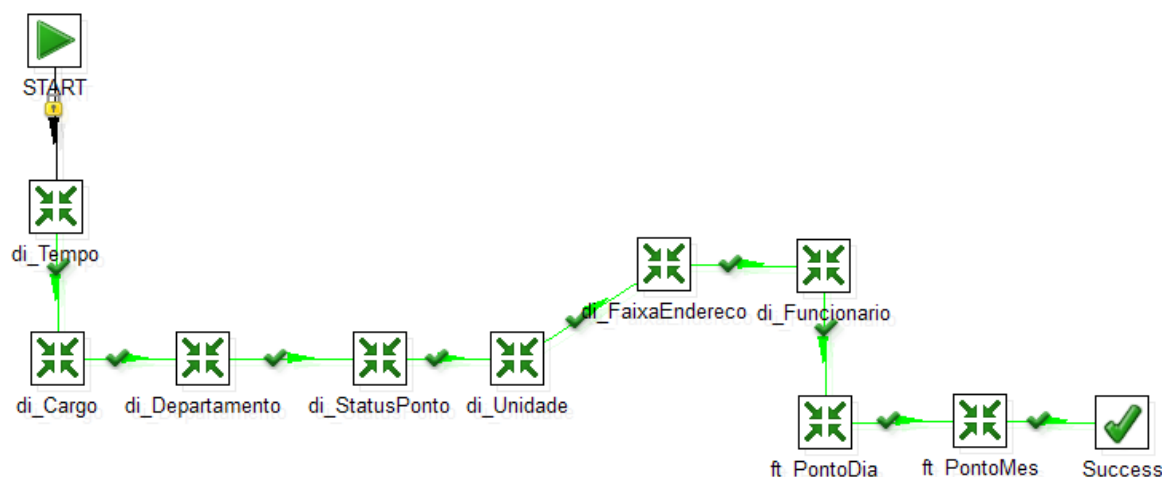
### 5.2.3 Construção e Desenvolvimento dos Dados Dimensionais

Após todos os procedimentos de transformação dos dados relacionais ocorrerem, é que os procedimentos dos dados dimensionais entram em vigor.

Desta forma o tempo de execução da ferramenta de *ETL* diminui, visto que os dados lidos em arquivos físicos requerem uma maior ocupação da memória com relação ao servidor de banco de dados, e a implementação torna-se mais simples e eficaz, pois os dados já estão estruturados em tabelas e estruturas semelhantes, com campos e formatos reconhecidos pelo próprio SGBD, posteriormente pelo *Pentaho Data Integration*.

Na figura 17, todas as transformações dos dados relacionais em dimensionais está representada pelo *Job* executado pela ferramenta *ETL*.

Figura 17 Transformação completa do modelo dimensional



Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

De todas as transformações englobadas neste processo, ficam em destaques as referentes aos funcionários e as agregações dos registros de pontos diários e mensais.

### 5.2.3.1 Transformação dos Dados dos Funcionários

Esta transformação é relativamente simples, vista que todo o procedimento de seus atributos foi realizado no modelo relacional.

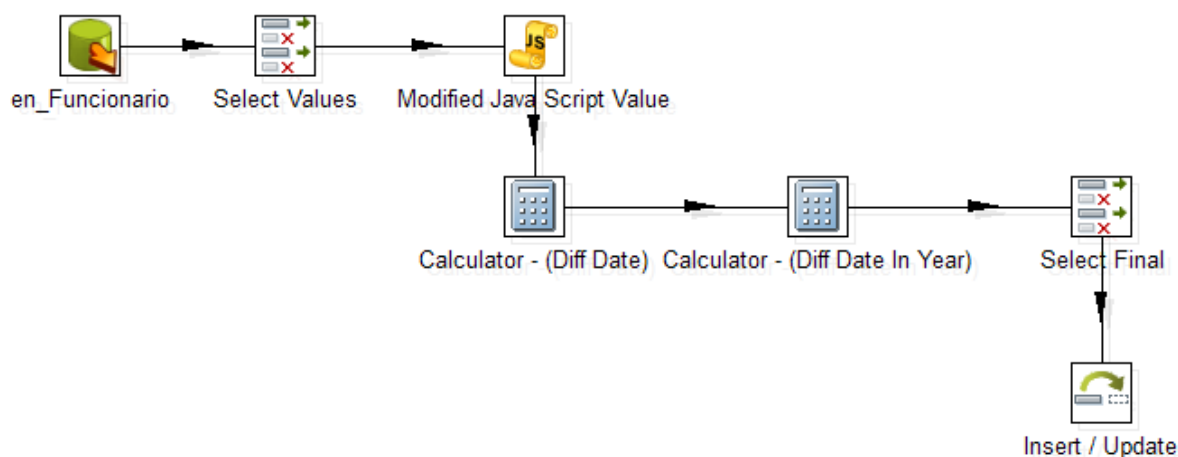
Inicialmente a ferramenta de *ETL* faz a leitura dos registros da tabela relacional de funcionários selecionando os seus principais campos. Em seguida é realizado o cálculo do tempo de admissão em dias e o cálculo da idade em anos de cada um dos colaboradores, porém serão apenas por via de armazenamentos dinâmicos, sem intuito de análises, pois neste caso serão utilizadas as tabelas fatos.

Posteriormente os dados serão inseridos ou atualizados, de acordo com as chaves primárias dos mesmos, na tabela dimensional de funcionário.

Na figura 18, todo o processo de transformação da ferramenta de *ETL* é exibida.



Figura 18 Transformação dos dados dos funcionários



Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

### 5.2.3.2 Transformação das Agregações dos Dados dos Registros de Pontos Diários

Esta transformação pode ser considerada a principal de todos os processos dimensionais, assumindo o ponto de maior importância para a leitura diária dos dados e, posteriormente, os mensais.

Inicialmente a ferramenta seleciona a tabela relacional de ponto com algumas junções (*joins*) de suas respectivas chaves estrangeiras.

Após essa seleção é realizado o cálculo dos minutos escalados, logados e em pausa que um funcionário teve.

Em seguida, o *ETL* faz as relações da data com a dimensão de tempo, cuja é responsável pela determinação do mês, ano, bimestre, trimestre e semestre da mesma.

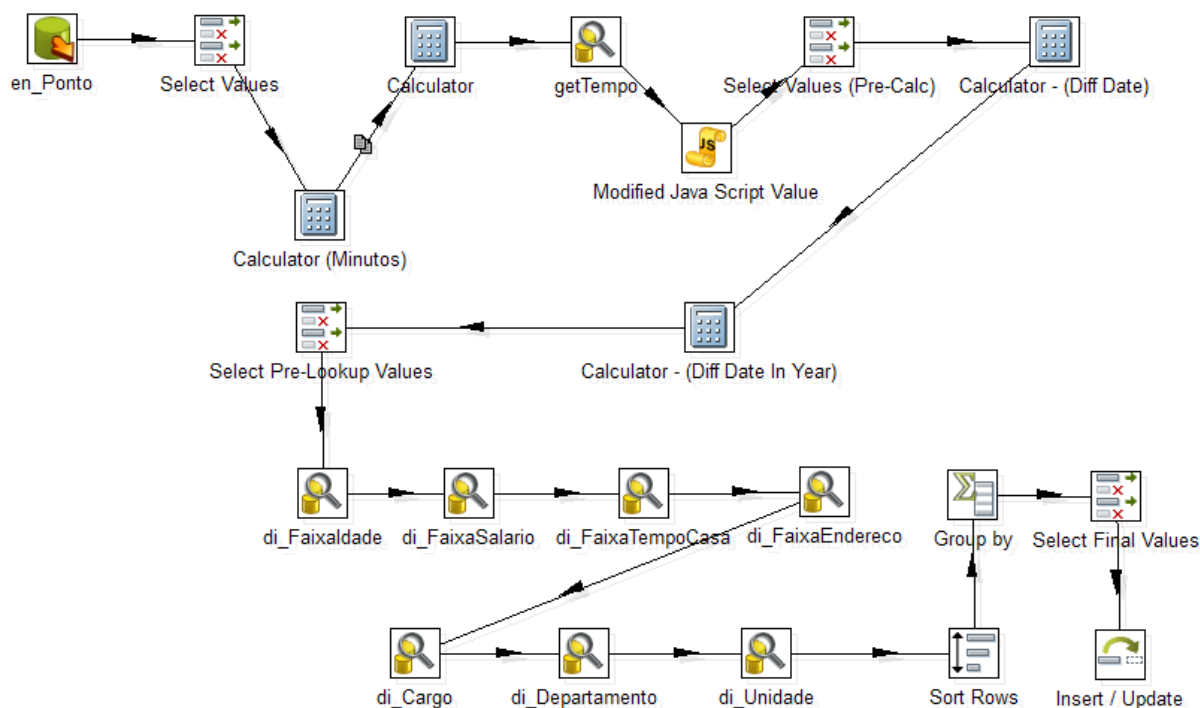
Novamente ocorrem cálculos, porém desta vez são referentes ao tempo de casa e a idade dos colaboradores. Estes dois campos, assim como o campo salário, serão utilizados como referência para buscar suas respectivas chaves primárias nas tabelas de faixa de idades, faixa de salários e faixas de tempo de casa.

Posteriormente ocorrem as ordenações e agrupações dos registros por todos os campos de chaves estrangeiras e pelo campo dia, do qual fica responsável apenas pelo armazenado do dia do campo de data de registro de ponto.

Assim que concluídas essas unificações, é realizado o procedimento de inserção ou atualização dos registros, considerando todas as chaves estrangeiras e o campo dia.

Na figura abaixo todas as transformações dos dados relacionais em dimensionais dos registros de pontos diários executadas pela ferramenta *ETL* é exibida.

Figura 19 Transformação das agregações dos dados dos registros de pontos diários



Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

### 5.2.3.3 Transformação das Agregações dos Dados dos Registros de Pontos Mensais

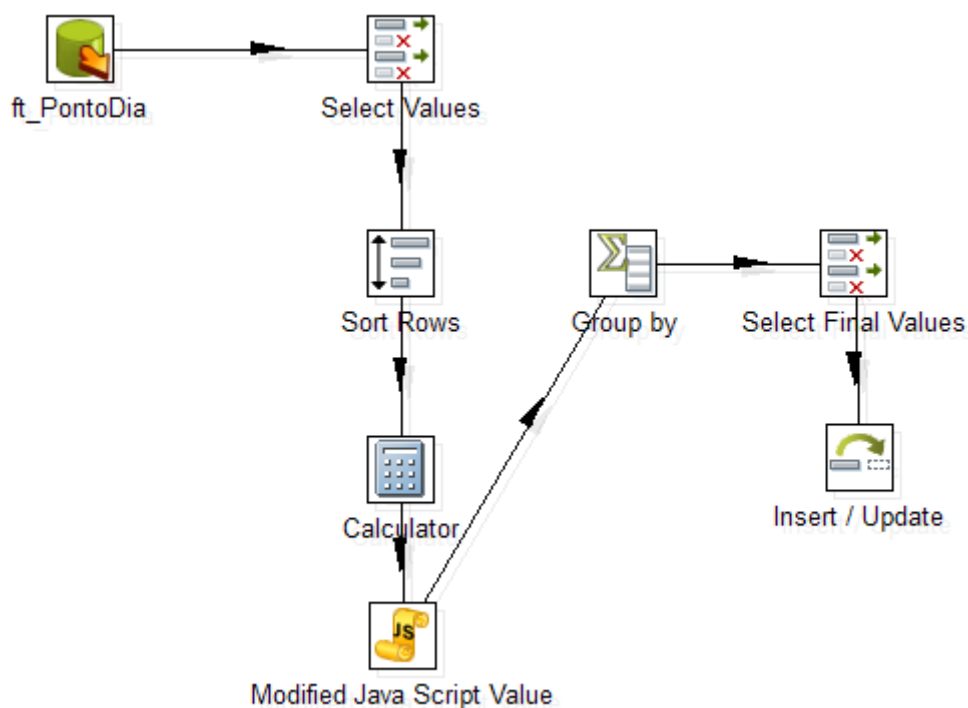
Esta transformação é realizada após a transformação diária, e com isso ela tem um tempo curto de processamento, já que sua visão principal no conceito de tempo é apenas mensal.

A leitura dos dados ocorre na tabela fato de registros de pontos diários, selecionando-os, ordenando-os e agrupando-os, para posteriormente inseri-los ou atualizá-los de acordo com os campos de chaves estrangeiras.

A estrutura desta tabela é idêntica a tabela fato diária, porém ela não engloba o campo dia, o que faz dela ser agregada com visões mensais, permitindo assim uma análise mensal mais eficaz, sem a necessidade de reprocessar os dados diários toda vez em que uma requisição de análise mensal for realizada.

Na figura abaixo todas as transformações dos dados da tabela fato dos registros de pontos diários em mensais executadas pela ferramenta *ETL* é exibida.

Figura 20 Transformação das agregações dos dados dos registros de pontos mensais



Fonte: Gerada pelo aplicativo *Pentaho Data Integration*.

#### 5.2.4 Criação do Cubo de Dados

Para realizar a criação do modelo de cubo de dados, foi utilizada a ferramenta *Schema Workbench*, com ela foi possível realizar todos os cálculos dos indicadores mapeados na modelagem.

O desenvolvimento dos cubos de dados é similar ao que já foi construído no *Data Warehouse*. O modelo é constituído por duas tabelas fatos, uma mensal e outra diária, seguidas por hierarquias e medidas onde são mapeados os indicadores e as dimensões no qual devem ser constituídas as análises na ferramenta *OLAP*.

Na figura 21 são apresentados todos os cubos de dados criados na ferramenta *Schema Workbench*.

Figura 21 Cubo de dados



Fonte: Gerada pelo aplicativo *Schema Workbench*

#### 5.2.4.1 Dimensões

As dimensões mapeadas no cubo são agrupadas por hierarquias, no qual trazem as informações que são utilizadas como base dados para as análises que são realizadas.

#### 5.2.4.2 Medidas

Nas medidas temos uma representação numérica dos dados da organização analisada, após os resultados serem gerados, eles agregam com as hierarquias formando um maior nível de detalhamento da informação.

Para realizar as análises, definimos algumas medidas:

- a) Minutos Logados: Soma total dos minutos logados;
- b) Minutos Escalados: Soma total dos minutos escalados;
- c) Atrasos: Para calcular os atrasos, é realizada uma subtração do total de minutos escalados com o total de minutos logados;
- d) Pausa: Soma total do tempo em pausa;
- e) Admissões: Soma total de admissões;
- f) Demissões: Soma total de demissões;
- g) Quantidade de Funcionários: Contagem distinta da quantidade de registros na tabela fato;

#### 5.2.4.3 Indicadores Calculados

Os indicadores calculados são gerados a partir dos resultados já existentes nas medidas, para a área de gestão dos recursos humanos, temos dois que são de extremo valor, o percentual de *absenteísmo* e o percentual de *turnover*.

##### 5.2.4.3.1 Absenteísmo

Para Chiavenato (2005, p. 86) “o absenteísmo constitui a soma dos períodos em que os funcionários se encontram ausentes no trabalho, seja por falta, atraso ou algum motivo interveniente”.

Chiavenato também afirma que o absenteísmo deve ser calculado com a seguinte fórmula:

$$\text{Índice de absenteísmo} = \frac{\text{Total de pessoas/horas perdidas}}{\text{Total de pessoas/horas de trabalho}}$$

Com base nesse conceito, o indicador é calculado através da divisão de duas medidas do cubo de dados, a soma de atrasos, ou seja, todas as horas perdidas na jornada de trabalho e a soma de minutos escalados que é o período total que o colaborador deve trabalhar.

#### 5.2.4.3.2 Turnover

Em relação ao cálculo de *turnover*, Chiavenato (2005, p. 89) afirma que “a fórmula mais utilizada é o número de pessoas que se desligaram durante um determinado período de tempo (um mês ou um ano) em relação número médio de funcionários existentes”.

Abaixo o desenho de como deve ser a fórmula de *turnover* segundo Chiavenato:

$$\text{Índice de Turnover} = \frac{\text{Nº de funcionários desligados}}{\text{Efetivo médio da organização}}$$

Para atender a fórmula do *turnover* no cubo de dados, é utilizada a divisão da medida quantidade de demissões pela quantidade de funcionários.

### 5.2.5 Análise e Interação das Informações

Na seção que segue, podem ser encontradas análises que tem como objetivo oferecer maneiras de apoio a Gestão dos Recursos Humanos de qualquer organização. Para ilustrar todas elas, é possível realizar a gestão através da ferramenta *OLAP*, utilizando uma base de dados fictícia manipulada especificamente para isso.

A apresentação é feita através de um *plugin* do *Pentaho BI Server* chamado *Saiku Analytics*, onde o utilizamos para cruzar todas as dimensões e medidas construídas no cubo de dados do *Schema Workbench* interativamente pela ferramenta. As análises são expostas por gráficos e também por tabelas ‘*JPIVOT*’.

Na arquitetura atual, é possível realizar cruzamentos pela ferramenta em todos os tipos de separação de dimensões que qualquer gestor achar necessário, a seguir, as principais formas de gestão através dela.

#### 5.2.5.1 Análises de Absenteísmo Por Faixa de Idade

A primeira análise mostra uma divisão por faixas de idade de cada unidade importada no *Data Warehouse*, fazendo uma interseção com o percentual de absenteísmo dos seus respectivos colaboradores em cada mês.

A Figura 22 ilustra uma simulação entre o cruzamento de dados através da uma tabela *JPIVOT*.

Figura 22 Análises dos dados de absenteísmo por faixa de idade

| NívelUnidade  | NívelIdade | %Absenteísmo |        |        | %Absenteísmo Grand Total |  |
|---------------|------------|--------------|--------|--------|--------------------------|--|
|               |            | 1            | 2      | 3      |                          |  |
| FLORIANÓPOLIS | 0-17 Anos  | 15,19%       | 12,21% | 9,87%  | 12,54%                   |  |
|               | 18-25 Anos | 10,16%       | 9,94%  | 7,35%  | 9,16%                    |  |
|               | 26-35 Anos | 10,23%       | 9,25%  | 6,98%  | 8,84%                    |  |
|               | 36-45 Anos | 10,84%       | 8,84%  | 6,33%  | 8,72%                    |  |
|               | 46-55 Anos | 5,60%        | 8,07%  | 8,87%  | 7,48%                    |  |
|               | 56-65 Anos | 8,82%        | 7,00%  | 4,75%  | 6,93%                    |  |
|               | 66-75 Anos | 0,16%        | 0,16%  | 0,03%  | 0,12%                    |  |
| SANTOS        | 18-25 Anos | 12,24%       | 10,23% | 7,03%  | 9,91%                    |  |
|               | 26-35 Anos | 15,47%       | 13,11% | 9,16%  | 12,66%                   |  |
|               | 36-45 Anos | 10,65%       | 8,20%  | 4,52%  | 7,87%                    |  |
|               | 46-55 Anos | 10,19%       | 7,97%  | 5,94%  | 8,09%                    |  |
|               | 56-65 Anos | 4,87%        | 9,11%  | 7,05%  | 7,06%                    |  |
| SÃO PAULO     | 18-25 Anos | 10,19%       | 10,57% | 7,47%  | 9,47%                    |  |
|               | 26-35 Anos | 14,56%       | 11,26% | 8,52%  | 11,50%                   |  |
|               | 36-45 Anos | 17,33%       | 14,79% | 8,16%  | 13,28%                   |  |
|               | 46-55 Anos | 28,90%       | 20,16% | 6,52%  | 18,53%                   |  |
|               | 56-65 Anos | 13,82%       | 15,23% | 8,91%  | 12,95%                   |  |
|               | 66-75 Anos |              | 12,79% | 17,16% | 14,97%                   |  |
|               |            | 13,64%       | 11,93% | 8,00%  |                          |  |
| Grand Total   |            | 11,56%       | 10,27% | 7,37%  |                          |  |

Fonte: Gerada pelo aplicativo *Saiku Analytics*.

O principal objetivo da análise é expor como cada faixa etária pode ter resultados diferentes de comprometimento com a carga horária de trabalho, ou seja, identificar se um público mais jovem ou mais velho está tendo resultados superiores ou inferiores em um determinado período.

Com base nesses dados, as organizações podem criar estratégias de contratação de perfis específicos de profissionais, descartando talvez um público-alvo que por experiências já vividas em outros meses, acarreta um grande histórico de fatores ruins.

### 5.2.5.2 Análises Sobre os Atrasos

A próxima análise exhibe todos os municípios dos funcionários de cada unidade, com uma soma de seus dados de atrasos nos últimos meses.

A figura 23 exibe uma simulação entre os resultados encontrados a partir de uma tabela ‘JPIVOT’.

Figura 23 Análise do total de atrasos por município

| NivelUnidade  | Município            | Atrasos |         |         | Atrasos Grand Total |  |
|---------------|----------------------|---------|---------|---------|---------------------|--|
|               |                      | 1       | 2       | 3       |                     |  |
| FLORIANÓPOLIS | AGUAS MORNAS         | 11.503  | 9.880   | 11.762  | 33.145              |  |
|               | ANTONIO CARLOS       | 195     | 805     | 310     | 1.110               |  |
|               | BIGUACU              | 71.057  | 80.080  | 43.188  | 174.285             |  |
|               | FLORIANOPOLIS        | 192.957 | 146.759 | 125.008 | 464.722             |  |
|               | FPOLIS               | 324     | 182     | 800     | 1.106               |  |
|               | GAROPABA             | 804     | 17      | 21      | 842                 |  |
|               | GOV CELSO RAMOS      | 307     | 1.508   | 1.790   | 3.603               |  |
|               | GOVERNADOR CELSO RAM | 7.513   | 994     | 2.389   | 10.896              |  |
|               | GOVERNADOS CELSO RAM | 3.100   | 3.908   | 1.140   | 8.146               |  |
|               | PALHOCA              | 311.158 | 294.253 | 217.040 | 822.449             |  |
|               | PAULO LOPES          | 1.280   | 1.954   | 5.939   | 9.173               |  |
|               | SANTO AMARO DA IMPER | 13.077  | 13.008  | 16.881  | 42.964              |  |
|               | SAO BONIFACIO        | 189     | 222     | 147     | 558                 |  |
|               | SAO JOAO BATISTA     | 1.545   | 951     | 41      | 2.537               |  |
|               | SAO JOSE             | 218.880 | 199.139 | 140.034 | 558.033             |  |
|               | SAO JOSE             | 2.045   | 1.220   | 1.370   | 4.635               |  |
|               | SAO PEDRO DE ALCANTA | 688     | 830     | 181     | 1.679               |  |
|               | SOAJO SE             | 541     | 370     | 109     | 1.020               |  |
| SANTOS        | CARUARA              | 287     | 19      | 135     | 421                 |  |
|               | CUBATAO              | 119.235 | 68.889  | 59.538  | 245.662             |  |
|               | GUARUJA              | 10.919  | 2.598   | 2.542   | 16.059              |  |
|               | JACAREI              | 330     | 432     | 889     | 1.651               |  |
|               | JUNDIAI              | 557     | 808     | 992     | 2.355               |  |

Fonte: Gerada pelo aplicativo *Saiku Analytics*.

As informações geradas têm como propósito relatar o foco de onde a empresa tem maiores dificuldades de deslocamento, apontando se algum determinado grupo de colaboradores que moram em lugares específicos estão sendo influenciados diretamente em seu resultado por não conseguir chegar a tempo na empresa.

Depois de identificar o foco, as organizações podem criar ações de auxílio ao grupo que residem em uma localidade de difícil acesso, como fretamento de ônibus, escalas alternativas ou qualquer forma que reduza o índice.

### 5.2.5.3 Cruzamento de Admissões versus Demissões



A análise que segue é formada pela quantidade de admissões e demissões de cada departamento de uma organização nos meses importados na carga separada por unidade.

Na figura 24, é exibida uma simulação com o cruzamento de admissões e demissão, geradas pelo *Saiku Analytics*.

Figura 24 Análise do total de demissões e admissões por departamento

| NivelUnidade  | NivelDepartamento           | 1         |           | 2         |           | 3         |           | Admissões Grand Total | Demissões Grand Total |  |
|---------------|-----------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------------------|-----------------------|--|
|               |                             | Admissões | Demissões | Admissões | Demissões | Admissões | Demissões |                       |                       |  |
| FLORIANÓPOLIS | DEPARTAMENTO ADMINISTRATIVO | 0         | 1         | 0         | 0         | 0         | 0         | 0                     | 1                     |  |
|               | DEPARTAMENTO FINANCEIRO     | 0         | 0         | 0         | 0         | 0         | 0         | 0                     | 0                     |  |
|               | DEPARTAMENTO OPERACIONAL    | 56        | 92        | 28        | 74        | 82        | 47        | 166                   | 213                   |  |
| SANTOS        | DEPARTAMENTO ADMINISTRATIVO | 0         | 0         | 0         | 0         | 0         | 0         | 0                     | 0                     |  |
|               | DEPARTAMENTO FINANCEIRO     | 0         | 0         | 0         | 1         | 0         | 0         | 0                     | 1                     |  |
|               | DEPARTAMENTO OPERACIONAL    | 44        | 48        | 0         | 43        | 18        | 31        | 62                    | 122                   |  |
| SÃO PAULO     | DEPARTAMENTO ADMINISTRATIVO | 0         | 0         | 0         | 0         | 0         | 1         | 0                     | 1                     |  |
|               | DEPARTAMENTO OPERACIONAL    | 13        | 17        | 34        | 34        | 18        | 9         | 65                    | 60                    |  |

Fonte: Gerada pelo aplicativo *Saiku Analytics*.

Na figura, pode ser identificado como foi o processo de contratação de cada departamento nos últimos meses, assim como o de demissão, analisando se algum departamento excedeu a quantidade devida desses indicadores.

Com essa análise, o gestor pode agir criando ações de contratação caso um determinado setor perca um grande número de funcionários ou parar de realizar qualquer contratação caso já possua um número demasiado de admissões.

#### 5.2.5.4 Análises de Cargos e Salários

Uma importante avaliação nas empresas é detectar como está a distribuição salarial e dos cargos de seu quadro de funcionários.

Na figura 25 é ilustrada a faixa salarial de todos os cargos das unidades, assim como a quantidade de cada colaborador.

Figura 25 Análise da distribuição de cargos e salários da organização

| NívelUnidade  | NívelCargo              | NívelSalário                | 1                          | 2                          | 3                          |     |
|---------------|-------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|-----|
|               |                         |                             | Quantidade de Funcionários | Quantidade de Funcionários | Quantidade de Funcionários |     |
| FLORIANÓPOLIS | ANALISTA ADMINISTRATIVO | R\$ 1000,00 - R\$ 1500,00   | 1                          | 1                          | 1                          | ■■■ |
|               |                         | R\$ 1500,00 - R\$ 2500,00   | 1                          | 1                          |                            | ■■  |
|               | ANALISTA FINANCEIRO     | R\$ 1500,00 - R\$ 2500,00   | 18                         | 18                         | 17                         | ■■■ |
|               |                         | R\$ 2500,00 - R\$ 3500,00   | 7                          | 7                          | 8                          | ■■■ |
|               | AUXILIAR ADMINISTRATIVO | R\$ 1000,00 - R\$ 1500,00   | 4                          | 3                          | 3                          | ■■■ |
|               |                         | R\$ 500,00 - R\$ 1000,00    | 1                          | 1                          | 1                          | ■■■ |
|               | COORDENADOR DE PRODUÇÃO | R\$ 3500,00 - R\$ 5000,00   | 4                          | 4                          | 5                          | ■■■ |
|               | GERENTE ADMINISTRATIVO  | R\$ 5000,00 - R\$ 8000,00   | 2                          | 2                          | 1                          | ■■■ |
|               | OPERADOR DE PRODUÇÃO    | R\$ 1000,00 - R\$ 1500,00   | 647                        | 653                        | 639                        | ■■■ |
|               |                         | R\$ 1500,00 - R\$ 2500,00   | 291                        | 281                        | 270                        | ■■■ |
|               |                         | R\$ 500,00 - R\$ 1000,00    | 53                         | 52                         | 49                         | ■■■ |
|               | SUPERVISOR DE PRODUÇÃO  | R\$ 1500,00 - R\$ 2500,00   | 16                         | 15                         | 13                         | ■■■ |
|               |                         | R\$ 2500,00 - R\$ 3500,00   | 26                         | 27                         | 29                         | ■■■ |
| SANTOS        | ANALISTA ADMINISTRATIVO | R\$ 1000,00 - R\$ 1500,00   | 4                          | 4                          | 4                          | ■■■ |
|               | ANALISTA FINANCEIRO     | R\$ 1500,00 - R\$ 2500,00   | 19                         | 20                         | 16                         | ■■■ |
|               |                         | R\$ 2500,00 - R\$ 3500,00   | 6                          | 5                          | 6                          | ■■■ |
|               | AUXILIAR ADMINISTRATIVO | R\$ 1000,00 - R\$ 1500,00   | 11                         | 12                         | 10                         | ■■■ |
|               |                         | R\$ 500,00 - R\$ 1000,00    | 2                          | 2                          | 2                          | ■■■ |
|               | COORDENADOR DE PRODUÇÃO | R\$ 3500,00 - R\$ 5000,00   | 4                          | 3                          | 3                          | ■■■ |
|               | GERENTE ADMINISTRATIVO  | R\$ 5000,00 - R\$ 8000,00   | 2                          | 2                          | 2                          | ■■■ |
|               |                         | R\$ 8000,00 - R\$ 999999,00 | 1                          | 1                          | 1                          | ■■■ |
|               | GERENTE GERAL           | R\$ 8000,00 - R\$ 999999,00 | 1                          | 1                          | 1                          | ■■■ |
|               | OPERADOR DE MÁQUINA     | R\$ 1500,00 - R\$ 2500,00   | 32                         | 32                         | 29                         | ■■■ |

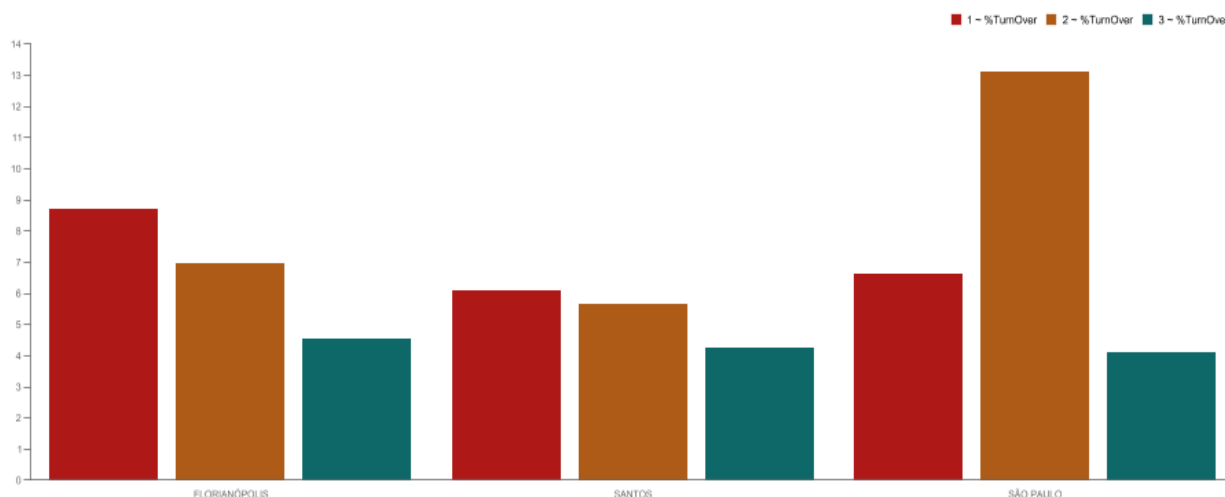
Fonte: Gerada pelo aplicativo *Saiku Analytics*.

Nessa avaliação, é possível verificar como está remuneração de cada perfil da organização, assim como a evolução salarial dos funcionários.

Com esse resultado, podem ser efetuadas análises sobre a necessidade de promoção, reformulação, demissão e diversos outros fatores que podem ser identificados a partir da distribuição de cargos e salários da empresa.

#### 5.2.5.5 Análises de *Turnover*

A última avaliação aborda o percentual de *turnover* nos últimos três meses da organização analisada. A figura 26 ilustra esse número.

Figura 26 Análise do *turnover* de cada unidade

Fonte: Gerada pelo aplicativo *Saiku Analytics*.

A partir dos resultados encontrados, é possível identificar como esse indicador está sendo distribuído em cada unidade, tal como se alguma delas está influenciando diretamente no resultado final de organização como um todo e a evolução dela em si.

O objetivo dessa análise é mostrar graficamente os números levantados de *turnover*, trazendo uma margem de análise com mais usabilidade em relação às tabelas, conseguindo facilmente enxergar a evolução ou queda do número.

### 5.3 UTILIZAÇÃO DA APLICAÇÃO

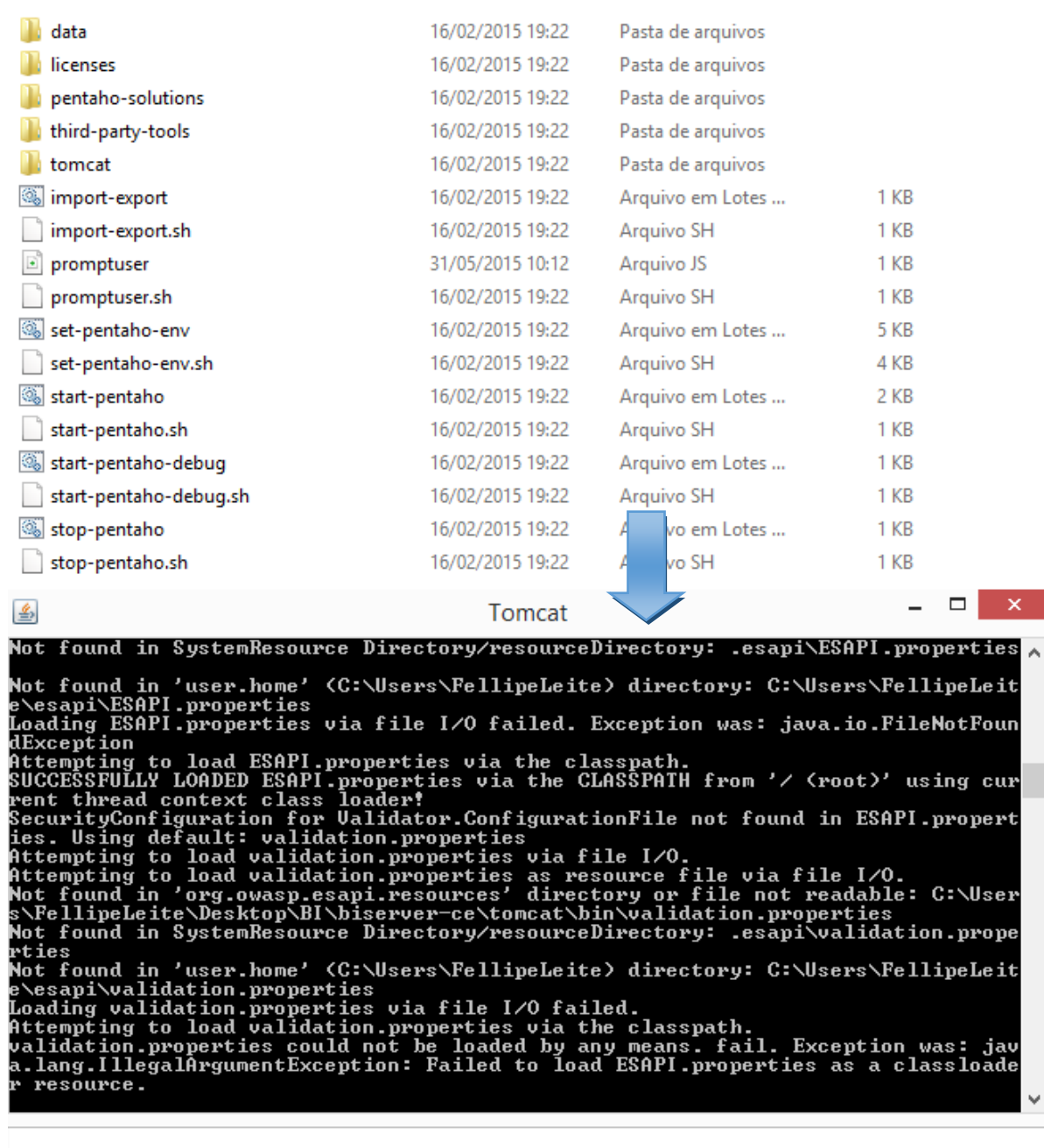
A seção que prossegue refere-se a um tutorial de como o usuário deve utilizar a aplicação, desde o acesso, escolha das variáveis e apresentação dos resultados a partir da ferramenta *Saiku Analytics*.

#### 5.3.1 Hospedagem e acesso

Como já foi abordada nas seções anteriores, a aplicação utiliza o *Pentaho BI Server* como ferramenta para hospedagem do servidor, que pode ser inicializado em uma máquina local ou hospedado em qualquer ambiente propício à produção.

A figura abaixo mostra a estrutura de pastas do servidor que deve ser iniciado a partir do arquivo *start-pentaho*.

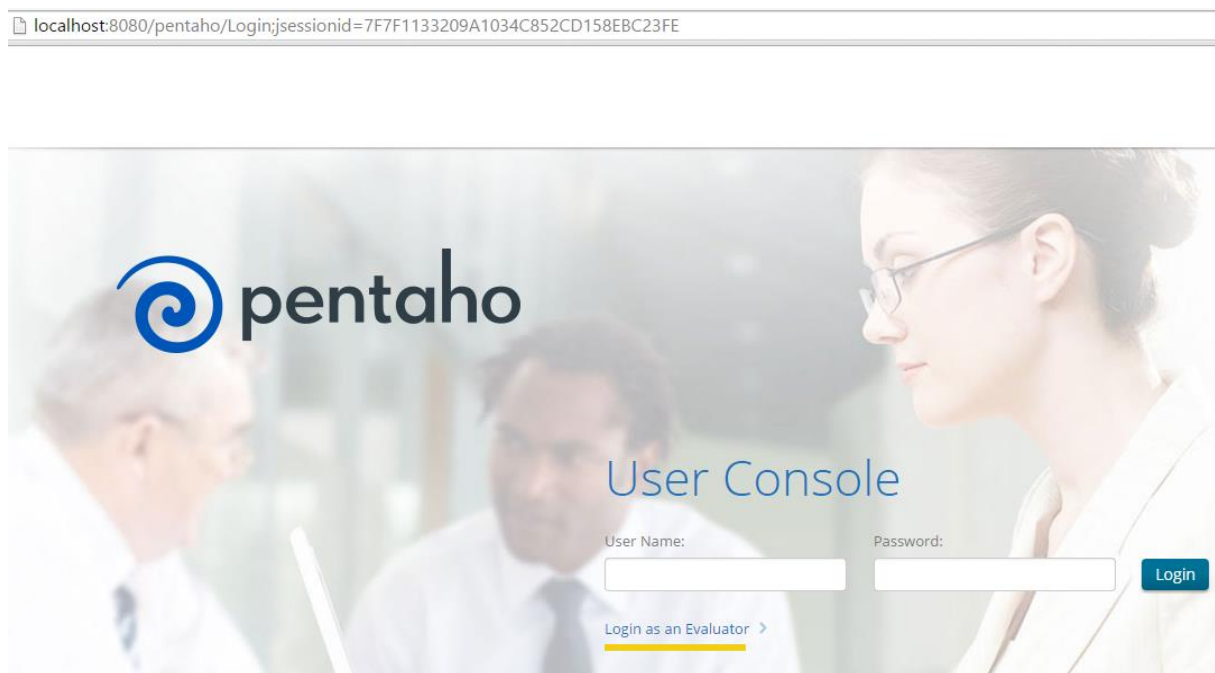
Figura 27 Estrutura de pastas do servidor



Fonte: Gerada pelos autores.

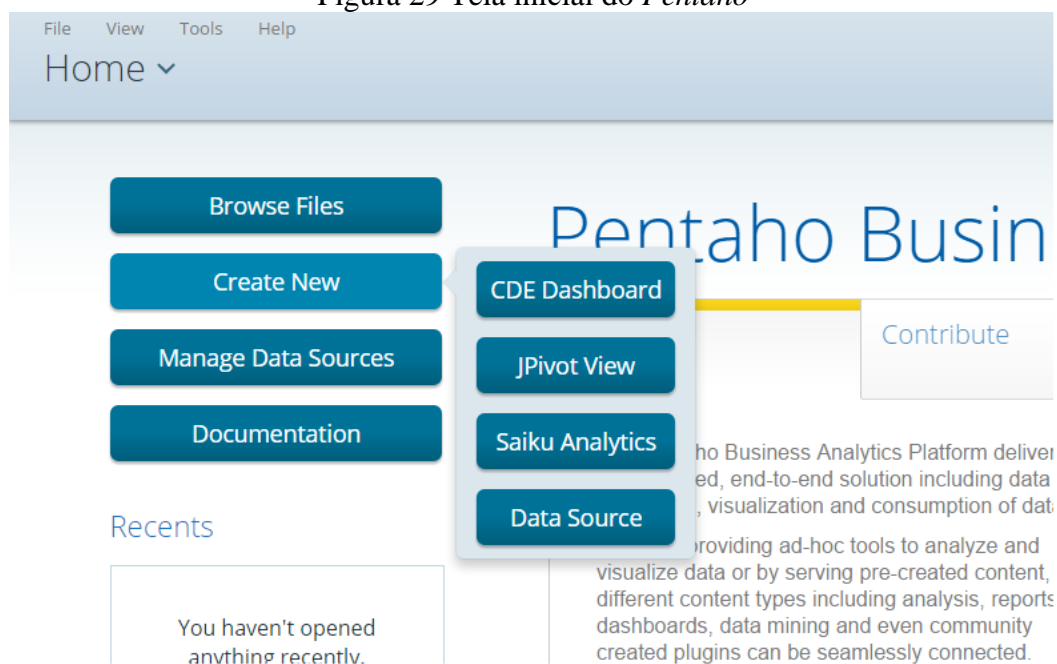
Após o servidor estar no ar, o usuário deve acessar a ferramenta através do seu navegador, seja na conexão local ou no ambiente de produção e realizar a autenticação na tela de *login*.

A figura 28 exibe a tela de *login* da aplicação.

Figura 28 Tela de *Login*

Fonte: Gerada pelos autores.

Com a autenticação realizada, será aberta a tela inicial, o usuário deve selecionar a opção 'Create New' e em seguida *Saiku Analytics*. A figura abaixo mostra a tela inicial da aplicação.

Figura 29 Tela inicial do *Pentaho*

Fonte: Gerada pelos autores.

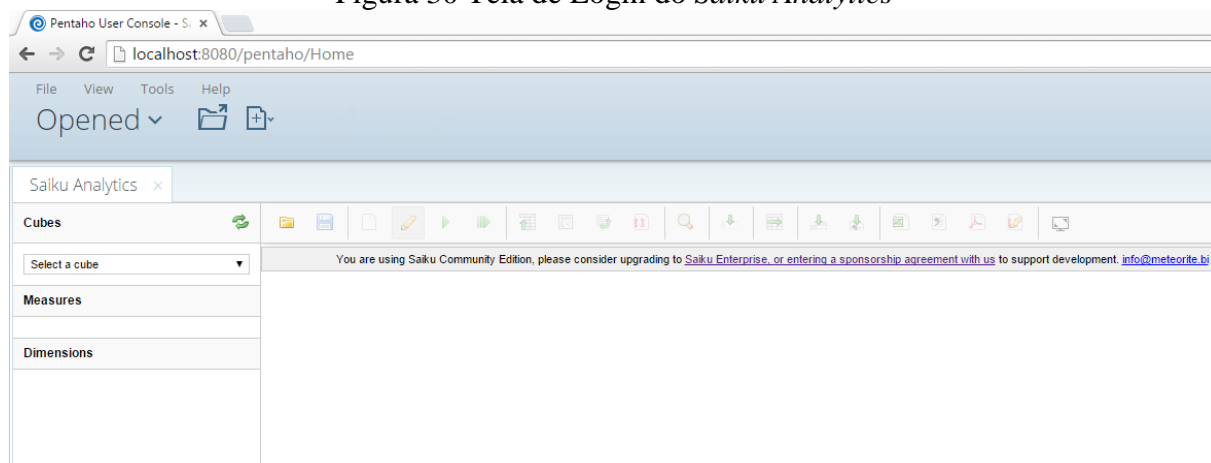
### 5.3.1 Escolha das variáveis e apresentação dos resultados

Depois do acesso ao *Saiku Analytics* feito, o usuário tem em mãos quais variáveis deseja escolher para realizar cruzamentos de dados transformando em informações relevantes que auxiliem na tomada de decisão.

Primeiramente o usuário deve selecionar qual cubo será utilizado, após isso, a ferramenta carrega na esquerda da tela, as dimensões e medidas no qual o usuário seleciona e automaticamente já são exibidas no formato de tabela.

A figura abaixo mostra a tela inicial do *Saiku Analytics*.

Figura 30 Tela de Login do *Saiku Analytics*



Fonte: Gerada pelos autores.

## 5.4 AVALIAÇÃO DA APLICAÇÃO

A ferramenta construída trouxe formas de gerir um grupo de funcionários de uma organização, a partir de informações genéricas que qualquer organização, desde pequeno até grande porte pode levantar sem grandes dificuldades.

Através do desenvolvimento da aplicação de *BI*, os objetivos específicos da monografia foram alcançados, desde o mapeamento das informações do setor de recursos humanos, passando pelo processo de *ETL* e carga no modelo dimensional, até as análises na ferramenta *OLAP*.

Para alcançar os objetivos específicos, foi essencial o desempenho de todas as ferramentas utilizadas da plataforma *Pentaho* que nos propôs desenvolver a aplicação sem restrição nenhuma ao que estava mapeado.

É de grande valia enfatizar que todas as análises geradas na monografia, são sugestões de uso da aplicação, cada empresa possui um cenário de negócio diferente e caso seja benéfico, vários outros cruzamentos de informações podem ser realizados a partir do sistema.

A ferramenta utilizada para a visualização dos dados é de fácil aprendida e bem didática, cores limpas e boas para longas horas de trabalho e análises, fácil compreensão das ações e cruzamentos, bem como toda a parte de filtragem é simples e bem diretiva. Apesar de ter exclusivamente no idioma inglês, o que requer um nível técnico de conhecimento nesta linguagem.

## 6 CONCLUSÃO

Este capítulo retrata as conclusões finais tiradas pelo nosso projeto, assim como próximos passos e estudos futuros, a fim de darmos continuidade em tudo àquilo que foi construído e embasado de boas práticas para a obtenção do resultado proposto.

### 6.1 CONCLUSÕES E CONTRIBUIÇÕES

Este projeto de aplicação de um *BI* dentro de um setor de Recursos Humanos possui uma grande relevância para as organizações que buscam mudar a forma como seus funcionários são vistos, tratados e procurados dentro da própria organização ou no mercado de trabalho em si. Embora esta implementação necessite de um investimento de tempo de adaptação e mudança de cultura para que os resultados possam aparecer e se tornarem eficazes a médio e longo prazo, pois para maiores e melhores análises, devem ser considerados períodos maiores.

A aplicação possui um poder de pôr nas mãos do RH à capacidade de torna-lo mais representativo a organização como um todo e tomar decisões que possam afetar diretamente o percurso e o diferencial competitivo no mercado de trabalho.

Visto que este setor já é responsável por todos os processos envolvidos na Gestão de Pessoas, nada seria mais eficaz que otimizar grande parte de suas ações através de um cubo de dados capaz de diagnosticar possíveis causas de problemas e remediar futuros *turnovers* ou absenteísmos, como por exemplo, o caso de um perfil de um funcionário recém contratado que não possui as mesmas características de funcionários que já fazem parte do quadro da empresa e representam uma gama de sucesso ou de um bom colaborador.

Todas as etapas do o processo de implementação do *Data Warehouse* no setor de Recursos Humanos é feito através de ferramentas *Open Sources*, livres e sem custos, o que não oneram em não-aplicações das mesmas para uma corporação e nem afetam o setor financeiro, por isso é excelente para uma empresa que queira testá-lo ou utilizá-lo como um precursor de boas práticas para outros setores ou outros funcionários.



Outra vantagem desta aplicação para a organização é a sua fácil utilização, sendo capaz de possuir um nível de aprendizado fácil, através de toda a forma didática que a mesma trabalha, neste caso tanto a parte visual do *Pentaho Data Integration*, responsável pela construção e a geração dos dados em informações, quanto o *Pentaho BI Server* e seus *plugins*, responsáveis por toda as visualizações das análises e os cruzamentos de informações requeridas pelo seu utilizador ou por suas necessidades.

## 6.2 TRABALHOS FUTUROS

Com o intuito de darmos continuidade em nossos estudos sobre a utilização de uma aplicação de *Business Intelligence* no setor de RH nas empresas, pretendemos inicialmente implementar toda a aplicação em um cenário real e uma empresa de pequeno porte, afim de coletarmos experiências e necessidades não mapeadas ou não levantadas, já que o *feedback* real pode abrir novos horizontes na forma como as informações são construídas e os dados gerados.

Além disso, ainda podemos aprimorar nossa aplicação com o armazenamento dos dados na nuvem, *Cloud Computing*, o que nos permite ter uma imagem e um controle maior sobre todo o processo envolvido na consolidação dos dados e na mineração dos mesmos, uma vez que possuindo este controle, será possível ter um *feedback* imediato sobre o processo em si. O *cloud computing* vem se tornando uma prática cada vez mais recorrente em organizações que pretendem não ser responsáveis pela parte de TI, preferindo a terceirização da mesma, o que neste caso seria todo o processo envolvido na aplicação do *BI* no setor de RH.

## REFERÊNCIAS

- BARBIERI, Carlos. **BI - Business Intelligence - Modelagem & Tecnologia**. Rio de Janeiro. Axcel Books do Brasil Editora, 2001.
- BARBIERI, Carlos. **BI2 - Business Intelligence - Modelagem & Qualidade**. Rio de Janeiro. Editora Elsevier, 2011.
- CARVALHO, B. F. **Arquiteturas de Ferramentas OLAP**. Rio de Janeiro. 2004.
- CARVALHO, Luis Alfredo Vidal de. **Data Mining – A mineração de dados no marketing, medicina, economia, engenharia e administração**. São Paulo. Editora Érica, 2001.
- CAVALCANTI, M. C.; OLIVEIRA, P. C. S.; MONTEIRO, R. R.; SOARES, V. J. **Análise Comparativa de Ferramentas OLAP**. Rio de Janeiro. UFRJ, 1998.
- CHIAVENATO, Idalberto. **Gestão de Pessoas**. Rio de Janeiro. Elsevier, 2004.
- CHIAVENATO, Idalberto. **Gestão de Pessoas**. Rio de Janeiro. Elsevier, 2005.
- CHIAVENATO, Idalberto. **Recursos Humanos**. Rio de Janeiro. Elsevier, 2009.
- CRAMER, Renato. **Estudo Analítico de ferramentas open source para ambientes OLAP**. Criciúma, 2006. Disponível em: <  
<http://www.datawarehouse.inf.br/Academicos/OpenSorce.pdf> > Acessado em: 19 out. 2014.
- GIL, Antônio L. **Sistemas de Informações Contábil/Financeiros**. São Paulo. Editora Atlas S.A., 1995.
- INMON, William H. **Como Construir o Data Warehouse**. Rio de Janeiro. Editora Campus, 1997.
- KIMBALL, Ralph. **Digging into Data Mining – Your Data Warehouse is your Data Mining Plataforma**. DBMS and Internet Systems. 1997.
- KIMBALL, Ralph; ROSS, Margy. **The Data Warehouse Toolkit: Guia Completo para Modelagem Dimensional**. Rio de Janeiro. Editora Campus. 2002.
- LAUDON, Kenneth C.; LAUDON, Jane Price. **Sistemas de Informação com Internet**. 1999.
- MACHADO, Felipe N. R. **Tecnologia e Projeto de Data Warehouse**. São Paulo. Editora Érica. 2010.
- MELO, Ivo S. **Administração de Sistemas de Informação**. São Paulo. Editora Guazzelli S.A.. 1999.

SILVA, E. Lucia; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração de Dissertação**. Florianópolis. UFSC. 2005.

SOMMERVILLE, D. **Engenharia de Software**. São Paulo. Pearson Education, 2003.

THOMSEN, E. **Construindo Um Sistema de Informação Multidimensional**. São Paulo. Editora Campus, 2002.

TURBAN, E. **Business Intelligence – Um Enfoque Gerencial**. Porto Alegre. Bookman. 2009.

ULRICH, D. **Os Campeões de Recursos Humanos: Inovando Para Obter Os Melhores Resultados**. São Paulo. Editora Futura, 2003.

WESTPHAL, Cristopher; BLAXTON, Teresa,. **Data Mining Solutions**. Canada. John Wiley & Sons Inc, 1998.

## **APÊNDICES**

## APÊNDICE A – Cronograma

**Tabela 3.1** Cronograma do Projeto.

| <b>Etapas</b>   | <b>Descrição</b>                           | <b>DATA DE INÍCIO<br/>(PREVISÃO)</b> | <b>DATA DE<br/>TÉRMINO<br/>(PREVISÃO)</b> |
|-----------------|--|--------------------------------------|---|
| Pesquisa        | Introdução ao Tema                         | 20/08/2014                           | 04/09/2014                                |
|                 | Revisão Bibliográfica                      | 10/09/2014                           | 23/10/2014                                |
|                 | Métodos de Pesquisa                        | 25/10/2014                           | 30/10/2014                                |
|                 | Modelagem - Introdução                     | 03/11/2014                           | 27/11/2014                                |
| Modelagem       | Levantamento dos requisitos                | 26/01/2015                           | 08/02/2015                                |
|                 | Criação das cargas                         | 08/02/2015                           | 20/02/2015                                |
|                 | Criação do Modelo Dimensional              | 20/02/2015                           | 27/02/2015                                |
|                 | Importação da carga.                       | 27/02/2015                           | 20/03/2015                                |
| Desenvolvimento | Criação da Arquitetura BI.                 | 21/03/2015                           | 31/03/2015                                |
|                 | Análise dos dados gerados.                 | 01/04/2015                           | 17/04/2015                                |
| Documentação    | Homologação da Estrutura Normativa da ABNT | 18/04/2015                           | 30/04/2015                                |

Fonte: Elaborada pelos autores.