



Pós-Graduação em Ciência da Computação

“UMA SOLUÇÃO DE BUSINESS INTELLIGENCE
BASEADA EM DATA WAREHOUSE PARA A
SECRETARIA DE SAÚDE DO ESTADO DE
PERNAMBUCO”

Por

ROBERTSON NOVELINO FERRAZ

Dissertação de Mestrado



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

RECIFE, FEVEREIRO/2009



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

ROBERTSON NOVELINO FERRAZ

"UMA SOLUÇÃO DE BUSINESS INTELLIGENCE BASEADA EM DATA
WAREHOUSE PARA A SECRETARIA DE SAÚDE DO ESTADO DE
PERNAMBUCO"

*ESTE TRABALHO FOI APRESENTADO À PÓS-GRADUAÇÃO EM
CIÊNCIA DA COMPUTAÇÃO DO CENTRO DE INFORMÁTICA DA
UNIVERSIDADE FEDERAL DE PERNAMBUCO COMO REQUISITO
PARCIAL PARA OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIA DA
COMPUTAÇÃO.*

ORIENTADOR(A): PROF. FREDERICO LUIZ GONÇALVES DE FREITAS

RECIFE, FEVEREIRO/2009

Ferraz, Robertson Novelino

**Uma solução de business intelligence baseada
em data warehouse para a secretaria de saúde do
estado de Pernambuco / Robertson Novelino
Ferraz. - Recife: O Autor, 2009.**

xxvii, 183 p. : il., fig., tab.

**Dissertação (mestrado) – Universidade Federal
de Pernambuco. CIn. Ciência da Computação,
2009.**

Inclui bibliografia e apêndices.

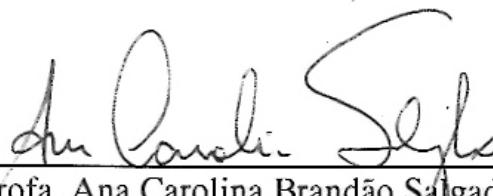
1. Engenharia de software. I. Título.

005.1

CDD (22. ed.)

MEI2009 - 160

Dissertação de Mestrado apresentada por **Robertson Novelino Ferraz** à Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, sob o título “**Uma Solução Business Intelligence Baseada em Data Warehouse para a Secretaria de Saúde do Estado de Pernambuco**”, orientada pelo **Prof. Frederico Luiz Gonçalves de Freitas** e aprovada pela Banca Examinadora formada pelos professores:



Profa. Ana Carolina Brandão Salgado
Centro de Informática / UFPE


Prof. Celso Roberto Perez
Associação de Empresas de Bases Tecnológica de Pernambuco



Prof. Frederico Luiz Gonçalves de Freitas
Centro de Informática / UFPE

Visto e permitida a impressão.
Recife, 5 de março de 2009.



Prof. FRANCISCO DE ASSIS TENÓRIO DE CARVALHO
Coordenador da Pós-Graduação em Ciência da Computação do
Centro de Informática da Universidade Federal de Pernambuco.

*Aos meus filhos e esposa, fontes de motivação e inspiração
para a vida.*

Agradecimentos

Agradeço a todos que contribuíram direta e indiretamente para o sucesso deste projeto. De forma especial gostaria de agradecer:

Aos meus filhos e esposa pelo apoio e horas roubadas de nossa prazerosa convivência.

Ao Professor e amigo Edson Costa de Barros Carvalho Filho pelas idéias e incentivos, inclusive, e principalmente, por me fazer acreditar na possibilidade da realização deste trabalho.

Ao amigo e orientador Frederico Luiz Gonçalves de Freitas pelas constantes presteza, atenção, dedicação, disponibilidade, incentivo, orientação e paciência sem o qual, certamente, não seria possível a concretização deste projeto.

Aos amigos da Secretaria de Saúde de Pernambuco, em especial a Ângelo Lins, Luiz Carlos e Ivan Garcia pelo atendimento permanente das diversas solicitações.

Ao amigo e consultor Manoel Jocelino de Oliveira, Supervisor de Informática do Hospital Getúlio Vargas, pela incansável presteza nos esclarecimentos das intermináveis dúvidas sobre os atendimentos nas emergências.

Ao Dr. João Veiga, médico do Estado e ex-Secretário Municipal de Saúde de Olinda, referência como médico e gestor público de saúde, por sua constante disponibilidade, refletindo em contribuição e direcionamento da pesquisa, principalmente viabilizando as análises críticas dos resultados, fundamental para o projeto.

A Dra. Darcy Freitas pela presteza no atendimento e incentivo na avaliação da viabilidade da ferramenta, permitindo a contextualização da gestão estratégica entre os vários níveis de gestão pública da saúde.

A Marta Agra pela constante presteza, sempre buscando atender qualquer solicitação, seja pessoal ou profissional.

Aos velhos amigos, onde citar nomes serve mais para os esquecidos que os lembrados, que compreenderam e respeitaram as ausências, o que fortalece cada vez mais nossa amizade fraterna.

Ao novos amigos das diversas disciplinas na UFPE, Adriana Simões, Iandé Coutinho, Rodrigo Freitas, Erick Gomes, Ricardo Scholtz e Sandra Lubambo pelas importantes contribuições.

AGRADECIMENTOS

Ao SERPRO, na pessoa da Gerente Sênior Helena Cristina e dos Chefes de Projetos Ézio Oliveira e Orlando Catarino, pelo incentivo e compreensão das necessárias ausências durante esta caminhada.

Ao meu pai, eterno incentivador dos estudos, embora perdêssemos o contato físico durante mais esta jornada...

*Impossível compor um poema a essa altura da evolução da humanidade.
Impossível escrever um poema - uma linha que seja - de verdadeira poesia.*

...

*Há máquinas terrivelmente complicadas para as necessidades mais
simples.*

Se quer fumar um charuto aperte um botão.

Paletós abotoam-se por eletricidade.

Amor se faz pelo sem-fio.

Não precisa estômago para digestão.

...

Os homens não melhoram e matam-se como percevejos.

Os percevejos heróicos renascem.

Inabitável, o mundo é cada vez mais habitado.

E se os olhos reprendessem a chorar seria um segundo dilúvio.

—CARLOS DRUMMOND DE ANDRADE (1930)

Resumo

A gestão dos recursos públicos enfrenta problemas associados à ineficiência na utilização de informações armazenadas em grandes volumes de dados que poderiam servir de base para soluções de suporte à tomada de decisão. No caso específico de saúde pública do Estado de Pernambuco, apesar do crescente volume de investimentos, o Estado não possui ferramentas integradas e de apoio à gestão estratégica que forneçam uma visão estrutural do seu funcionamento.

Por outro lado, soluções de *Business Intelligence* (BI) propõem-se a prover informações estratégicas, nas mais diversas formas de apresentação, a partir das fontes de dados de uma organização. Seguindo uma orientação do Governo Federal, que vem utilizando uma política de incentivo à utilização de software livre, nada mais oportuno técnica e financeiramente que a proposta de uma solução que atenda este direcionamento.

Neste trabalho utilizando-se como base de dados as informações disponibilizadas nas emergências dos quatro maiores hospitais públicos do Estado de Pernambuco, é proposto um sistema de suporte à decisão que passa pela elaboração de um *Data Warehouse*. Nesta fase experimental já estão sendo utilizados dados reais, finalizando com a disponibilização de uma solução de *Business Intelligence* para a Secretaria de Saúde do Estado de Pernambuco.

Os resultados obtidos em diversas análises realizadas e validadas pelos gestores públicos envolvidos demonstram a potencialidade, flexibilidade e facilidade de implementação desta solução de apoio à gestão estratégica na saúde pública, como uma parte de uma gestão integrada envolvendo todo o Estado e seus diversos processos de decisão visando garantir a universalização do atendimento como direito do cidadão.

Palavras-chave: Sistemas de Apoio à Decisão, *Business Intelligence*, *Data Warehouse*, Gestão Pública da Saúde.

Abstract

The administration of public resources faces problems associated to inefficiency in the use of information stored in large volumes of data that serve as the basis for the support to decision-making processes. In the specific case of public healthcare in the state of Pernambuco (Brazil), despite the growing volume of investments, the state does not yet have integrated tools to support administration strategies that provide a structural view of its functioning.

Business Intelligence (BI) solutions propose the promotion of strategic information in the most diverse forms of presentation, using data sources from an organization. Based on the guidelines of the federal government, which has been employing an incentive policy for the use of free software, there is nothing more opportune from a technical and financial standpoint than the proposal of a solution that addresses this issue.

Using information available at emergency rooms in four major public hospitals in the state of Pernambuco as the database, a system of support to decision-making is proposed, which involves the design of a Data Warehouse. In this experimental phase, real data are currently being used, with the availability of a BI solution for the Secretary of Health of the state of Pernambuco.

The results obtained from different analyses and validated by the public administrators involved demonstrate the potentiality, flexibility and ease of implementation of this support solution to strategic management in public healthcares as part of an integrated administration involving the entire state and its diverse decision-marking processes, with the aim of ensuring universal healthcare as the right of every citizen.

Keywords: Decision Support Systems, Business Intelligence, Data Warehouse, Public Healthcare Management.

Sumário

1	Introdução	1
1.1	Motivação	1
1.2	Objetivos	2
1.3	Contribuições	2
1.4	Estrutura da Dissertação	3
2	Business Intelligence	5
2.1	Introdução	5
2.2	Sistemas de Informação	5
2.2.1	Tipos de Sistemas de Informação	6
2.3	Soluções de Business Intelligence	8
2.3.1	Definição	9
2.3.2	Breve Histórico	10
2.3.3	Componentes de uma arquitetura de BI	10
2.3.4	Fatores-Chave de Sucesso para a implantação do BI	11
2.3.5	ETL - Extração, Transformação e Carga	13
2.3.6	<i>Data Warehouse</i>	14
2.3.6.1	Definição	14
2.3.6.2	<i>Data Mart</i>	15
2.3.6.3	A Análise dos Dados em uma Organização	15
2.3.6.4	Modelagem de dados em um Data Warehouse	17
2.3.6.5	Cubo de dados	20
2.3.7	Área de Apresentação	21
2.3.8	OLAP	22
2.3.8.1	Arquitetura OLAP	24
2.3.8.2	Operações de OLAP	27
2.3.9	Fases de Desenvolvimento do Data Warehouse	29
2.3.10	Periodicidade de Carga do <i>Data Warehouse</i>	30

2.3.10.1	Slowly Changing Dimension Tipo 1 - atualizar	31
2.3.10.2	Slowly Changing Dimension Tipo 2 - Mantendo o Histórico	31
2.3.10.3	Slowly Changing Dimension Tipo 3 - Realidades Alternativas	31
2.3.10.4	Slowly Changing Dimension Híbrida	32
2.4	Considerações finais	32
3	Trabalhos Relacionados	33
3.1	A Informática na Saúde Pública	33
3.2	Projeto de um <i>Data Warehouse</i> para a Saúde Pública	35
3.2.1	Análise Crítica	37
3.3	A Aplicação do BI no Segmento de Saúde Pública Ambulatorial	38
3.3.1	Análise Crítica	39
3.4	Sistema de Informação Gerencial Utilizando Solução de DW em Software Livre	40
3.4.1	Análise Crítica	41
3.5	Análise Comparativa e Considerações Finais	41
4	A Proposta de um BI para a Saúde Pública	43
4.1	O Custeio da Assistência à Saúde	43
4.2	A Secretaria Estadual de Saúde: Gestão, Custeio e Informatização	45
4.2.1	A Gestão Hospitalar	48
4.2.1.1	Gerenciamento da Atenção à Saúde	49
4.2.2	Conjuntura Atual da Informática na SES-PE	50
4.3	O Sistema de Atendimento aos Pacientes da Emergência (SAPE)	51
4.3.1	Uma proposta de <i>Data Warehouse</i> para a SES-PE	54
4.3.2	Análise dos Requisitos	55
4.3.3	O Modelo de Dados Dimensional	57
4.3.3.1	Definição das Dimensões	57
4.3.3.2	Definição dos Fatos	61
4.4	Considerações Finais	62
5	Implementação da Solução de BI	63
5.1	Introdução	63
5.2	Problemas das Ferramentas de BI	64
5.3	Análise das Ferramentas	64
5.3.1	Critérios de Seleção	65
5.3.2	As Ferramentas Analisadas	66

5.3.2.1	Mondrian	66
5.3.2.2	JPivot	67
5.3.2.3	OpenI	67
5.3.2.4	Palo	68
5.3.2.5	SpagoBI	68
5.3.2.6	Pentaho	68
5.3.3	Comparativo entre as Ferramentas	69
5.3.4	Porque o Pentaho?	70
5.4	A Plataforma Pentaho de BI	71
5.4.1	A Arquitetura	73
5.4.1.1	A Plataforma	73
5.4.1.2	Ferramentas	75
5.5	Extração, Transformação e Carga	77
5.5.1	A Integração dos Dados	77
5.5.2	A Extração	79
5.5.3	Transformação	82
5.5.4	Carga	86
5.6	Cubos OLAP	87
5.7	Considerações Finais	90
6	Resultados Obtidos	93
6.1	Metodologia	93
6.2	Análises e Resultados de Apoio à Decisão	94
6.2.1	Atendimentos por Hospital	95
6.2.2	Tipos de Ocorrências	96
6.2.3	Tipos de Ocorrências por Hospital	103
6.2.4	Domicílio dos Pacientes	106
6.2.5	Internações Hospitalares	112
6.3	Considerações Finais	113
7	Conclusão	115
7.1	Contribuições	115
7.2	Considerações Finais	116
7.3	Trabalhos Futuros	117

A Apêndice A	119
A.1 Modelo Entidade Relacionamento do Sistema SAPE-Emergência	119
A.2 Dicionário de Dados das Tabelas de Dimensão	121
A.2.1 Dimensão Paciente	121
A.2.2 Dimensão Tipo_Ocorrencia	122
A.2.3 Dimensão Procedencia	122
A.2.4 Dimensão Destino	122
A.2.5 Dimensão Medico_Liberador	123
A.2.6 Dimensão Hospital	123
A.2.7 Dimensão Data_Registro	124
A.2.8 Dimensão Emergencia_Origem	124
A.2.9 Dimensão Ind_Acid_Trabalho	125
A.2.10 Dimensão Ind_Pac_Nao_Identificado	125
A.2.11 Dimensão Forma_Transporte	125
A.2.12 Fato Atendimento	126
A.3 Modelo Relacional de Correspondência das dimensões	127
A.3.1 Dimensão Paciente	127
A.3.2 Dimensão Tipo_Ocorrencia	128
A.3.3 Dimensão Procedencia	128
A.3.4 Dimensão Destino	129
A.3.5 Dimensão Medico_Liberador	129
A.3.6 Dimensão Hospital	130
A.3.7 Dimensão Emergencia_Origem	130
A.3.8 Dimensão Data_Registro	131
A.3.9 Dimensão Ind_Acid_Trabalho	131
A.3.10 Dimensão Ind_Pac_Nao_Identificado	131
A.3.11 Dimensão Faixa_Etaria	131
A.3.12 Dimensão Faixa_Horario	131
A.3.13 Dimensão Forma_Transporte	131
A.3.14 Fato Atendimento	132
B Apêndice B	133
B.1 Descrição das Transformações e <i>Jobs</i> Gerados	133
B.1.1 Dimensão Emergência	133
B.1.1.1 Extração Emergência	133

B.1.1.2	Geração das <i>SurrogateKeys</i> das Emergências	136
B.1.1.3	Carga das Emergências no DW	138
B.1.2	Dimensão Hospital	140
B.1.2.1	Extração dos Hospitais	140
B.1.2.2	Geração das <i>SurrogateKeys</i> dos Hospitais	142
B.1.2.3	Carga das Hospitais no DW	144
B.1.3	Dimensão Forma Transporte	145
B.1.3.1	Extração dos Hospitais	145
B.1.3.2	Geração das <i>SurrogateKeys</i> das Formas de Transportes	146
B.1.3.3	Carga das Formas de Transporte	148
B.1.4	Dimensão Destino	149
B.1.4.1	Extração dos Destinos	149
B.1.4.2	Geração das <i>SurrogateKeys</i> dos Destinos	151
B.1.4.3	Carga das Destinos no DW	153
B.1.5	Tabela de Fatos Atendimentos	154
B.1.5.1	Extração dos Atendimentos	154
B.1.5.2	Transformação dos Fatos para o Esquema Estrela	156
B.1.5.3	Carga dos Fatos no DW	160
C	Apêndice C	163
C.1	Atendimentos e Óbitos por Dia da Semana	163
C.2	Faixas Etárias	164
C.3	Acidente de Trabalho	166
C.4	Paciente Não Identificado por Hospital	167
C.5	Procedência dos Pacientes	169
C.6	Transferências e Destinos dos Pacientes	170
C.7	Internações Hospitalares	173

Lista de Figuras

2.1	Tipos de sistemas de informação. Retirada de [Lau01].	6
2.2	Arquitetura típica de um solução de Business Intelligence. Retirada de [Sys07].	11
2.3	Modelo de Dados Multidimensional.	18
2.4	Cubo de Dados.	20
2.5	Arquitetura MOLAP. Retirada de [Cun04].	24
2.6	Arquitetura ROLAP. Retirada de [Cun04].	25
2.7	Arquitetura HOLAP. Retirada de [Cun04].	25
2.8	Arquitetura DOLAP. Retirada de [Cun04].	26
2.9	Arquitetura WOLAP. Retirada de [Cun04].	26
2.10	Análise <i>drill-down</i> e <i>roll-up</i> .	27
2.11	Diagrama do ciclo de vida dimensional do negócio. Retirado de [KR02].	29
3.1	Escopo Global do Projeto. Retirado de [SGFT08].	35
4.1	Distribuição geográfica das GERES.	45
4.2	Estrutura da Diretoria de Planejamento da SES-PE.	47
4.3	Fluxo do Processo de Gestão Hospitalar. Retirada de [ses06].	48
4.4	Tela Inicial do Sistema de Atendimento aos Pacientes da Emergência (SAPE).	52
4.5	Modelo Multidimensional do SAPE Emergência.	58
5.1	Diagrama de uma Solução de BI com a Pentaho Open BI Suite. Retirada de [pen08].	71
5.2	Estrutura de Componentes da Pentaho Open BI Suite. Retirada de [pen08].	74
5.3	Job de Extração DW SAPE.	79
5.4	Extração da Tabela de Hospital do DW SAPE.	80
5.5	Execução do <i>Job</i> da Extração dos Fatos.	82
5.6	Transformação para Gerar Versão da Carga.	82
5.7	Transformação da Dimensão Medico_Liberador.	83
5.8	Transformação da Dimensão Paciente.	84
5.9	Algoritmo de Eliminação de Redundância de Pacientes.	84

5.10 Job de Carga da Dimensão Procedência.	86
5.11 Job de Carga do DW SAPE Emergência.	88
5.12 Elaboração do Cubo OLAP.	89
6.1 Atendimentos por Hospital.	95
6.2 Atendimentos e Óbitos Anual por Hospital.	96
6.3 Atendimentos por Tipos de Ocorrências.	97
6.4 Natureza de Ocorrência Sintoma.	98
6.5 Atendimentos e Óbitos por Agressão.	100
6.6 Agressão Animal por Faixa Etária - Anual.	101
6.7 Agressão Física por Sexo - Anual.	101
6.8 Atendimentos e Óbitos por Agressão - Anual.	102
6.9 Agressão e Acidentes por Dia da Semana.	103
6.10 Tipo de Ocorrência por Hospital.	104
6.11 Ocorrências Otorrinolaringológicas por Hospital.	106
6.12 Tipos de Atendimentos por Cidades de Pernambuco.	107
6.13 Atendimentos por Cidades e Hospitais de Pernambuco.	109
6.14 Atendimentos por Cidades e Hospitais de Pernambuco.	110
6.15 Atendimentos Anuais por Cidade e Hospital.	111
6.16 Atendimentos e Internações por Hospital.	112
6.17 Internações por Hospital- Anual.	113
A.1 Modelo Entidade Relacionamento do SAPE Emergência.	120
A.2 Tabelas de origem no SAPE Emergência - Paciente.	127
A.3 Tabelas de origem no Sistema SAPE Emergência - Tipo Ocorrência.	128
A.4 Tabelas de origem no Sistema SAPE Emergência - Procedência.	128
A.5 Tabelas de origem no Sistema SAPE Emergência - Procedência.	129
A.6 Tabelas de origem no Sistema SAPE Emergência - Médico Liberados.	129
A.7 Tabelas de origem no Sistema SAPE Emergência - Hospitais.	130
A.8 Tabelas de origem no Sistema SAPE Emergência - Emergência Origem.	130
A.9 Tabelas de origem no Sistema SAPE Emergência - Forma de Transporte.	131
A.10 Tabelas de origem no Sistema SAPE Emergência - Atendimentos.	132
B.1 Componentes das Emergências.	134
B.2 Extração das Emergências.	134
B.3 Componentes das <i>SurrogateKeys</i> das Emergências.	136

B.4	Transformação <i>Surrogate</i> das Emergências.	137
B.5	Componentes das Carga das Emergências.	138
B.6	Transformação Carga das Emergências.	138
B.7	Componentes dos Hospitais.	140
B.8	Extração dos Hospitais.	140
B.9	Componentes das <i>SurrogateKeys</i> dos Hospitais.	142
B.10	Transformação <i>Surrogate</i> dos Hospitais.	143
B.11	Componentes das Carga dos Hospitais.	144
B.12	Transformação Carga dos Hospitais.	144
B.13	Componentes das Formas de Transporte.	145
B.14	Extração das Formas de Transporte.	145
B.15	Componentes das <i>SurrogateKeys</i> das Formas de Transportes.	146
B.16	Transformação <i>Surrogate</i> das Formas de Transportes.	147
B.17	Componentes das Carga das Formas de Transporte.	148
B.18	Transformação Carga das Formas de Transporte.	148
B.19	Componentes dos Destinos.	149
B.20	Extração dos Destinos.	150
B.21	Componentes das <i>SurrogateKeys</i> dos Destinos.	151
B.22	Transformação <i>Surrogate</i> dos Destinos.	152
B.23	Componentes das Carga dos Destinos.	153
B.24	Transformação Carga dos Destinos.	153
B.25	Componentes dos Atendimentos.	154
B.26	Extração dos Atendimentos.	154
B.27	Componentes das Transformações dos Atendimentos.	156
B.28	Transformação Transformações dos Atendimentos.	157
B.29	Execução da Transformação dos Fatos.	159
B.30	Componentes das Carga dos Atendimentos.	160
B.31	Transformação Carga dos Atendimentos.	161
B.32	Execução da Carga dos Atendimentos.	161
C.1	Atendimentos e Óbitos por Dia da Semana.	163
C.2	Faixas Etárias dos Pacientes.	165
C.3	Faixas Etárias dos Pacientes por Hospital.	165
C.4	Acidentes de Trabalho por Hospital.	166
C.5	Acidentes de Trabalho por Ocorrências.	167

C.6 Pacientes Não Identificados por Hospital.	168
C.7 Pacientes Não Identificados por Ocorrências.	168
C.8 Procedência dos Pacientes.	169
C.9 Transferências Entre as Emergências dos Hospitais.	171
C.10 Transferências das Emergências para Hospitais Conveniados.	171
C.11 Transferências por Tipo de Ocorrência por Hospital.	172
C.12 Evasão e Destino Ignorado por Hospital.	173
C.13 Destino Pacientes para Hospitais Conveniados.	174
C.14 Atendimentos e Internações por Ocorrências.	175

Lista de Tabelas

2.1	Categorias dos Sistemas de Processamento de Informação	8
2.2	Comparação entre fatos, dimensões e medidas. Adaptado de [Mac00]	19
4.1	Requisitos necessários para a implementação do DW.	55
4.2	Faixas Etárias	60
4.3	Faixas de Horários	61
5.1	Critérios Utilizados para Seleção da Ferramenta.	66
5.2	Comparação entre as Ferramentas Analisadas e Critérios Utilizados.	69
5.3	Quantidade de Pacientes Atendidos de 2005 a Junho de 2008.	80
5.4	Quantidade de Atendimentos Anual por Hospital	81
A.1	Dimensão Paciente	121
A.2	Dimensão Tipo_Ocorrencia	122
A.3	Dimensão Procedencia	122
A.4	Dimensão Destino	122
A.5	Dimensão Medico_Liberador	123
A.6	Dimensão Hospital	123
A.7	Dimensão Data_Registro	124
A.8	Dimensão Emergencia_Origem	124
A.9	Dimensão Forma_Transporte	125
A.10	Fato Atendimento	126

CAPÍTULO 1

Introdução

1.1 Motivação

O ramo de informática em saúde pública envolve muitos desafios, tais como o agrupamento de um conjunto muito heterogêneo de dados, distribuídos espacial e temporalmente, e cuja análise usualmente norteia as políticas de saúde pública nos governos que os empregam.

No Brasil, apesar de haver grandes bancos de dados disponíveis sobre o assunto, dos quais o SINTERSE (Sistema Integrado de Tratamento de Informações Estratégicas) constitui o maior exemplo, por vezes estes recursos carecem tanto de boas ferramentas de apoio à decisão, conhecidas como ferramentas de *Business Intelligence*, como também de boas escolhas de coleta e modelagem dos dados que possam realmente proporcionar subsídios para a redução de custos na saúde.

Em particular, no estado de Pernambuco, os atendimentos nas emergências dos principais hospitais públicos são atualmente registrados no módulo do Sistema de Atendimento de Pacientes (SAPE), desenvolvido e de propriedade da Secretaria de Saúde do Estado de Pernambuco (SES-PE), com utilização de um Sistema Gerenciador de Banco de Dados. O SAPE tem como principal finalidade o registro de entrada de pacientes nas emergências dos hospitais públicos, cadastrando os dados do paciente, a ocorrência do fato gerador da emergência e a posterior baixa neste atendimento associada a um motivo informado pelos médicos.

Implantado nos maiores hospitais públicos de Pernambuco, o SAPE vem servindo ao seu propósito de ser uma ferramenta basicamente operacional, não havendo ainda uma solução de apoio aos gestores da organização que propicie respaldo para as decisões em saúde pública baseadas em uma análise extensiva dos dados consolidados, permitindo uma visão administrativa de um conjunto de hospitais públicos.

Neste cenário, a implantação de um *Data Warehouse* (DW) surge como uma solução à necessidade crescente de informações integradas pela gerência estratégica, como parte do processo decisório do mais alto nível. A partir de uma análise histórica de dados, sistemas de *Data Warehouse* possibilitam também a descoberta de tendências que se traduzem em melhoria da aplicabilidade dos recursos com a definição de processos operacionais mais precisos.

Em complemento à implementação de um *Data Warehouse*, no desenvolvimento da tecnologia da informação, soluções de *Business Intelligence* (BI) visam oferecer os meios necessários para a transformação de dados em informação e para sua divulgação através de ferramentas analíticas, a fim de suportar o processo decisório [Bro00]. As ferramentas de processamento analítico para apoio à decisão permitem a criação de análises, ou seja, relatórios e cubos de dados, obedecendo diferentes perspectivas sobre os dados disponíveis [KR02].

Importante salientar que as ferramentas disponibilizadas na solução de BI não fornecem a "inteligência" que compõe seu nome (*Intelligence*) pois somente apresentam suas funcionalidades para a exploração dos dados armazenados no *Data Warehouse*. Cabe à organização, personalizada pelos seus gestores, utilizar seu conhecimento do negócio aplicado à tecnologia disponibilizada apresentando resultados, das mais diversas formas, que auxiliem a tomada de decisão.

1.2 Objetivos

Este trabalho insere-se nas áreas de informática pública e de saúde pública, utilizando-se de tecnologias de *Business Intelligence* e *Data Warehouse* para otimizar seus objetivos. Neste contexto, surge a oportunidade de apresentar técnicas inovadoras no tratamento dos dados, fornecer critérios para a identificação de uma melhor solução, elaborar um modelo de dados associado à situação atual e disponibilizar um ambiente para auxiliar a tomada de decisão gerencial dos gestores envolvidos com os atendimentos médicos aos pacientes nas emergências dos principais hospitais públicos do Estado de Pernambuco.

Para a viabilização do objetivo de suporte à tomada de decisão optou-se pela elaboração de um *Data Warehouse* bem como uma solução de *Business Intelligence* como ferramenta de apoio, baseada na análise de cruzamento das informações fornecidas no atendimento do paciente.

Considerando-se uma forte tendência utilizada pelo Governo Federal, a ferramenta a ser disponibilizada é gratuita e de código aberto, facilitando a sua aceitação como solução técnica.

1.3 Contribuições

Apresentando-se como uma solução inovadora no trabalho em hospitais públicos no Brasil, como contribuição inicial temos a própria identificação da riqueza de informações das fontes de dados disponíveis em locais geograficamente distribuídos e não integrados. Neste contexto,

surge também a oportunidade de contribuição na elaboração e automação das transformações e limpeza dos dados, bem como a execução de tarefa de conformidade dos domínios nos diversos hospitais públicos envolvidos, permitindo pela primeira vez à Secretaria de Saúde do Estado de Pernambuco (SES-PE) uma visão unificada e global das informações.

Vale ressaltar que um diferencial fundamental desta solução reside na escolha do tratamento de dados das emergências dos maiores hospitais públicos do Estado. Estes dados podem ser considerados primordiais na tomada de decisão gerencial e estratégica em termos de políticas de saúde pública, visto que podem tornar a gestão mais eficiente, impactando diretamente na redução dos diversos custos associados ao atendimento.

A partir de alguns critérios será selecionada, utilizada, demonstrada a eficiência e disponibilizada a solução de *Business Intelligence* mais viável dentro da realidade técnica e financeira da SES-PE.

Em seguida, será definido um novo modelo de dados baseados na unificação das fontes de dados que servirá de base para futuras evoluções de outras soluções complementares, apresentadas com viabilidade bem factíveis de aplicabilidade.

A realização de análises das informações, apresentadas detalhadamente e com avaliações críticas, com o auxílio de gestores de saúde pública, permitirá simulações onde a situação atual e histórica dos atendimentos nos hospitais públicos serão identificadas, e se for o caso, corroboradas as políticas de saúde do Estado. Por outro lado, estas informações poderão também servir como sugestões para o redirecionamento de investimentos e decisões estratégicas a partir de uma visão unificada dos hospitais públicos.

O repasse de toda a solução, com o comprometimento da Gerência de Informática da SES-PE, parte diretamente envolvida, contribuirão, de forma clara e objetiva, para embasar, demonstrar e reforçar as constantes justificativas da necessidade de investimento em implementação de soluções de vanguarda em informática pública e na gestão pública na área de saúde.

1.4 Estrutura da Dissertação

Esta dissertação está estruturada em capítulos da seguinte forma:

- Capítulo 1 - Introdução: Neste capítulo apresenta-se a motivação que gerou este trabalho, seus objetivos e as contribuições desejadas;
- Capítulo 2 - *Business Intelligence*: Este capítulo expõe os principais conceitos relacionados a *Business Intelligence* (BI), iniciando com o desenvolvimento dos sistemas de informação nas organizações. Seguem-se os fatores críticos de sucesso na elaboração de

uma solução e os componentes de BI, destacando-se as ferramentas analíticas de apoio à decisão. Destacam-se também as Fases de Desenvolvimento de um *Data Warehouse* que servem de referência durante os capítulos seguintes;

- Capítulo 3 - Trabalhos Relacionados: Os trabalhos relacionados ao tema da informática na saúde pública são detalhados neste capítulo. Apresenta-se as metodologias e ferramentas utilizadas finalizando com uma análise comparativa entre os trabalhos;
- Capítulo 4 - A Proposta de um BI para a Saúde Pública: No início do capítulo mostra-se a forma de custeio da saúde envolvendo os vários níveis governamentais, a Secretaria de Saúde do Estado de Pernambuco (SES-PE), conceitos sobre a gestão hospitalar e a estruturação da área de Tecnologia da Informação. Na sequência apresentam-se o sistema transacional que fundamenta o problema identificado, os requisitos necessários da solução que servem para a elaboração do modelo dimensional;
- Capítulo 5 - Implementação da Solução de BI: Apresenta-se uma análise comparativa entre as soluções em software livre e a solução escolhida, envolvendo as várias fases do processo e analisando o desempenho de cada uma das ferramentas da solução. Com base no modelo dimensional elaborado no Capítulo 4, é desenvolvido e avaliado o modelo em uma situação real. Há um detalhamento dos processos de elaboração e implementação da solução nas ferramentas selecionadas;
- Capítulo 6 - Resultados Obtidos. Os resultados das análises visam justificar o projeto, demonstrando como realmente auxilia no apoio à decisão estratégica de gestão pública da saúde;
- Capítulo 7 - Contribuições, Considerações Finais e Trabalhos Futuros: São apresentadas as conclusões sobre a proposta inicial, as viabilidades e propostas para fomentar trabalhos futuros;
- Apêndice A: Neste apêndice há um detalhamento do Modelo Entidade e Relacionamento do sistema transacional e um dicionário de dados da solução proposta;
- Apêndice B: Apresentam-se detalhadamente as outras transformações elaboradas no Capítulo 5;
- Apêndice C: Relacionam-se outras análises realizadas e disponibilizadas para os gestores que auxiliam à tomada de decisão estratégica na gestão da saúde pública.

CAPÍTULO 2

Business Intelligence

2.1 Introdução

Para analisarmos a solução proposta de desenvolvimento de uma solução de *Business Intelligence* (BI), neste capítulo é apresentada uma introdução dos principais conceitos relacionados a BI, iniciando com o desenvolvimento dos sistemas de informação nas organizações. Em seguida os principais componentes de soluções de BI são apresentados, destacando-se as ferramentas analíticas de apoio à decisão.

2.2 Sistemas de Informação

Segundo Laudon e Laudon [Lau01], um sistema de informação pode ser definido tecnicamente como um conjunto de componentes inter-relacionados que coleta (ou recupera), processa, armazena e distribui informação para dar suporte à tomada de decisão e ao controle da organização. Além de apoiar, coordenar e controlar a tomada de decisão, os sistemas de informação também podem ajudar os gerentes e trabalhadores a analisar problemas, visualizar assuntos complexos e criar novos produtos.

Toda empresa possui um Sistema de Informação (SI) que pode possuir níveis diferentes de informatização. As atividades da organização são acompanhadas por estes sistemas, sendo possível controlar os vários processos internos de custo, manufatura, contabilidade ou logística assim como interagir com os clientes, fornecedores e sociedade em geral. O fornecimento de informações corretas, precisas e no tempo correto, impulsionado pelo seu SI, determina fortemente o sucesso de uma organização.

Segundo Davenport [DP99] o processamento básico realizado por um sistema de informação é baseado em três etapas: coleta de dados, processamento e disponibilização da informação. A informação resultado do processamento é utilizada para a geração do conhecimento.

O dado coletado servirá como matéria-prima para o sistema de informação e o produto final a informação. A produção do conhecimento surge de um processo de transmutação da

informação, o conhecimento por si próprio é explicitado, transformado em informação, para que possa ser processado e produzir novo conhecimento. Esta inovação propicia a formação de um ciclo virtuoso de formação de conhecimento [Une05].

2.2.1 Tipos de Sistemas de Informação

Quatro principais tipos de sistemas de informação servem em níveis da organização diferentes: sistemas de nível operacional, sistemas de nível de conhecimento, sistemas de nível gerencial e sistemas de nível estratégico [Lau01].

Na Figura 2.1 são mostrados os tipos de Sistemas de Informação.

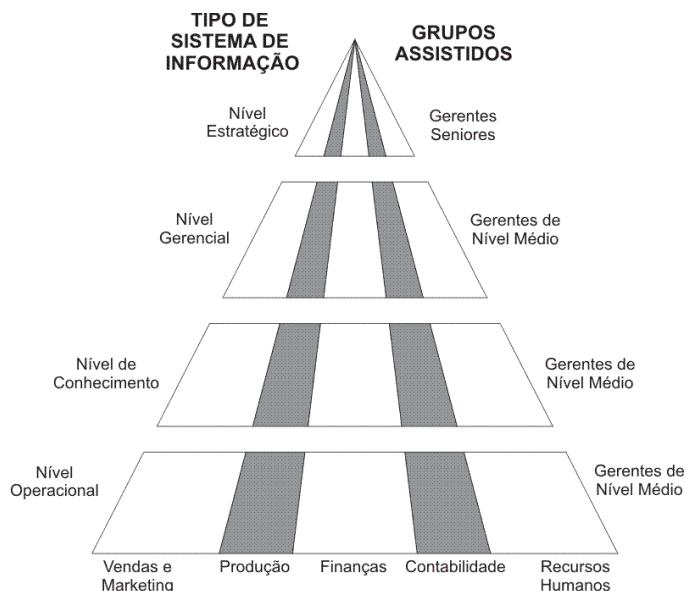


Figura 2.1 Tipos de sistemas de informação. Retirada de [Lau01].

Sistemas de nível operacional dão suporte aos gerentes operacionais no acompanhamento das atividades e transações elementares da organização, tais como vendas, receitas, folha de pagamento e logística. O principal propósito de sistemas desse nível é responder perguntas rotineiras e localiza o fluxo das transações através da organização. Também são chamados de OLTP (*On-Line Transactions Processing*) que caracterizam-se por serem sistemas operacionais de registro que capturam as transações diárias da empresa sendo responsáveis pela persistência dos dados que os aplicativos da empresa manipulam.

Sistemas de nível de conhecimento dão suporte aos trabalhadores do conhecimento e de dados em uma organização. O seu propósito é ajudar a empresa a integrar novos conhecimento no negócio e controlar o processo de encaminhamento de documentos. São as aplicações que

mais crescem nas empresas atualmente.

Sistemas de nível gerencial são projetados para servir de monitoramento, ao controle, à tomada de decisão e às atividades administrativas dos gerentes médios. Fornecem relatórios periódicos em lugar de informações instantâneas sobre as operações. Alguns sistemas de nível gerencial dão suporte a tomada de decisão não-rotineiras [Kee00]. Tendem a concentrar-se em decisões menos estruturadas para as quais as exigências de informação nem sempre são claras.

Sistemas de nível estratégico ajudam a administração sênior a atacar e enfocar assuntos estratégicos e tendências de longo prazo, tanto na empresa como no ambiente externo. Sua principal preocupação é adequar as mudanças no ambiente externo com a capacidade organizacional existente.

Sistemas de informação estratégicos mudam as metas, as operações, os produtos, os serviços ou os relacionamentos ambientais das organizações para ajudá-las a obter um diferencial estratégico que garanta melhoria tanto na prestação dos serviços quanto na garantia da subsistência da organização. Estes sistemas alteram profundamente o modo de condução do negócio da empresa, podendo ser utilizado como arma competitiva desde que entenda-se onde as oportunidades estratégicas para o negócio têm maiores chances de serem encontradas.

Segundo Laudon e Laudon [Lau01] as categorias específicas dos sistemas servem a cada nível organizacional e contribuem com o seu valor para a empresa.

A Tabela 2.1 resume as características dos seis tipos de sistemas de informação. Deve-se observar que cada um dos diferentes tipos de sistemas pode ter componentes que são utilizados por níveis organizacionais e outros grupos além daqueles que são seus principais clientes.

Cada um dos principais tipos de sistemas de informação é valioso para ajudar as empresas a resolverem um importante problema. Nas últimas décadas, alguns desses sistemas se tornaram especialmente vitais para o sucesso e a sobrevivência a longo prazo das organizações.

Com a evolução da tecnologia utilizada pelos sistemas de informação, tornou-se possível oferecer mais velocidade para o processamento das informações e apoio ao processo decisório [Inm02]. O surgimento da técnica de *data warehousing* e das soluções de BI permitiu a criação de sistemas de informação que possibilitam à organização um maior controle dos processos internos e a definição e o acompanhamento de estratégias de gestão dos seus negócios.

Sistema	Entradas	Processamento	Saída	Usuários
Suporte Executivo	Dados agregados; externos, internos	Gráficos, simulações; interações	Projeções; tabelas	Executivos
Supporte e Apoio à Decisão	Baixos volumes dados ou grande volume para análise	Interações, simulações; análise	Relatórios especiais, análises da decisão, tabelas	Gerentes, equipe da gerência
Informação Gerencial	Sumários de dados, grande volume de dados	Relatórios de rotina, análises de baixo nível	Sumário e relatórios de exceção	Gerentes médios
Trabalho do Conhecimento	Base de conhecimento	Modelagens e simulações	Modelos, gráficos	Equipe administração de conhecimento
Automação Escritório	Documentos, cronogramas	Gerência documentos e cronogramas	Documentos, cronogramas e cartas	Equipe administração de conhecimento
Transacional	Transações, eventos	Ordenação, agregação e atualização	Relatórios detalhados, listagens e sumários	Pessoal operacional e supervisores

Tabela 2.1 Categorias dos Sistemas de Processamento de Informação

2.3 Soluções de Business Intelligence

A transformação da sociedade industrial no século XX para a sociedade do conhecimento neste novo século, ressalta o valor da informação, na sua difusão e interpretação. A sociedade do conhecimento preconiza o desenvolvimento de meios para identificação, produção, processamento, transformação, disseminação e uso de informação para criar e aplicar conhecimento no desenvolvimento da sociedade [Une05].

Nesta nova sociedade as organizações vencedoras serão aquelas que conseguirem acúmulo de saber, contando com o empenho coletivo, a capacidade das pessoas envolvidas de se relacionarem umas com as outras, dentro de uma linguagem comum, de esforço conjunto [How00]. Se antes o conhecimento mostrava-se armazenado em benefício do seu proprietário, atualmente torna-se compartilhado, pois desta forma cresce, enquanto o conhecimento não utilizado torna-se obsoleto e perde o seu valor.

No contexto do desenvolvimento da tecnologia de informação, soluções de *Business Intelligence* (BI) visam oferecer os meios necessários para a transformação de dados em informação e para a sua divulgação através de ferramentas analíticas, a fim de suportar o processo decisório [Dev01]. Tais soluções são compostas tipicamente de ferramentas ETL (*Extraction, Transformation, Loading*).

mation and Loading ou Extração, Transformação e Carga), de um ou vários banco de dados e de ferramentas de processamento analítico e de mineração de dados [KR02]. As ferramentas de processamento analítico para apoio à decisão permitem a criação de análises, através de relatórios e cubos de dados, obedecendo a diferentes perspectivas sobre os dados disponíveis [KR02]. As funcionalidades exploratórias oferecidas por essas aplicações analíticas limitam-se normalmente a um conjunto de funcionalidades genéricas de navegação e filtragem de dados sobre o banco de dados.

As soluções de BI de fato não fornecem a "inteligência" referida no termo que cunha essas aplicações aos tomadores de decisão, pois se limitam a oferecer funcionalidades genéricas para exploração das montanhas de dados reunidas nos repositórios das organizações [Lie07].

2.3.1 Definição

O termo *Business Intelligence* (BI), pode ser traduzido como Inteligência de negócios, refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte a gestão de negócio [Wik08]. Para [Cor02], *Business Intelligence* é um conjunto de conceitos e metodologias que visa o apoio à tomada de decisões nos negócios a partir da transformação do dado em informação e da informação em conhecimento.

Segundo Barbieri [Bar01] *Business Intelligence* é um conjunto de ferramentas de processo contínuo e sistemático que produz conhecimento e perspicácia para uma companhia. Efetivamente, o BI aumenta a qualidade do planejamento estratégico e operacional e reduz o tempo para a tomada de decisões.

Segundo Cerqueira [Cer02] *Business Intelligence* é um conjunto de conceitos e metodologias que, fazendo uso de acontecimentos (fatos) e sistemas baseados nos mesmos, apóia a tomada de decisões em negócios. É o termo utilizado para denominar um conjunto de ferramentas de Tecnologia de Informações que proporciona suporte a decisão estratégica, tática e operacional. Este suporte se dá através da criação de uma série de indicadores de performance que podem estar relacionados diretamente à base de dados da corporação. O conceito se presta justamente ao objetivo de auxiliar as pessoas por meio do tratamento das bases de dados existentes, de maneira a aprimorar o processo de tomada de decisão.

Soluções de BI implantadas com sucesso provêem uma visão integrada do negócio, estendem as capacidades analíticas dos usuários e impulsionam a formação de *expertise* nas organizações. O foco de soluções de BI é facilitar o entendimento do negócio das organizações, fornecendo a todos os níveis das organizações informações relevantes sobre suas operações internas e o ambiente externo, incluindo clientes e competidores, parceiros e fornecedores

[Mos03]. O ambiente externo inclui ainda variáveis independentes que possam impactar no negócio, como tecnologia, leis e economia mundial, entre outros [Inm02].

2.3.2 Breve Histórico

Considerando-se a tecnologia em constante processo de crescimento, o BI está relacionado com a evolução do papel dos sistemas de informação nas organizações. Nos anos 70 e até meados dos anos 80, soluções de processamento e impressão de relatórios em lote, chamados de *batch*, surgiram como soluções de apoio à decisão. Trabalhava-se com extensos relatórios para extrair elementos básicos de informação. Com a proliferação dos terminais de acesso aos mainframes, o acesso aos relatórios digitais foi disseminado, mas o acesso à informação era dificultado devido à complexidade dos sistemas da época.

Uma segunda fase dos sistemas de apoio à decisão é marcada pelo surgimento do *Data warehouse* (DW), repositórios de dados integrados e preparados para o apoio à decisão, que em conjunto com a evolução das ferramentas analíticas ofereceu performance e poder analítico para o nível tático e executivo nas organizações [Inm02] e [KR02].

A terceira fase corresponde ao surgimento do BI. Segundo Almeida [Alm06], o foco do DW estava muito orientado à tecnologia de consolidação dos dados. As vantagens de projetos de BI em relação aos de DW são:

- soluções de BI não são orientadas unicamente à aplicação de tecnologia de informação de última geração, mas também ao fornecimento de soluções que integram pacotes verticais de aplicativos e metodologias para diversos segmentos de negócio;
- o foco das soluções de BI está no acesso e na distribuição de informação para o apoio à decisão; e
- soluções de BI suportam o acesso a todos os dados da organização, estruturados e não estruturados, e não somente aos armazenados no DW.

Sendo considerada uma metodologia, estruturalmente a solução de BI encontra-se dividido em componentes, que são introduzidos a seguir.

2.3.3 Componentes de uma arquitetura de BI

Segundo Cortez [Cor02], Kimball [KR02] e WGSystem [Sys07], a arquitetura típica de soluções de BI possui três componentes, apresentado na Figura 2.2, a saber:

- ETL (Extração, Transformação e Carga): conjunto de aplicativos e ferramentas que fazem a coleta de dados nos repositórios da organização, procedem com a limpeza e transformação para enfim carregar o DW [KR02] e [Inm02];
- Data Warehouse: repositório de dados integrado e não-volátil onde são armazenados os dados transformados pelo módulo ETL. Esse repositório deverá suportar as demandas analíticas das ferramentas de apoio à decisão e os aplicativos de extração de conhecimento [KR02] e [Inm02];
- Área de apresentação: diz respeito ao conjunto de instrumentos que serão utilizados pelos usuários na organização para navegar no DW. Esses instrumentos correspondem a relatórios previamente configurados, aplicativos para confecção de relatórios, ferramentas OLAP (*On-line Analytical Processing*), ferramentas de Data Mining (mineração de dados), entre outras [Tho02] e [Gon05].

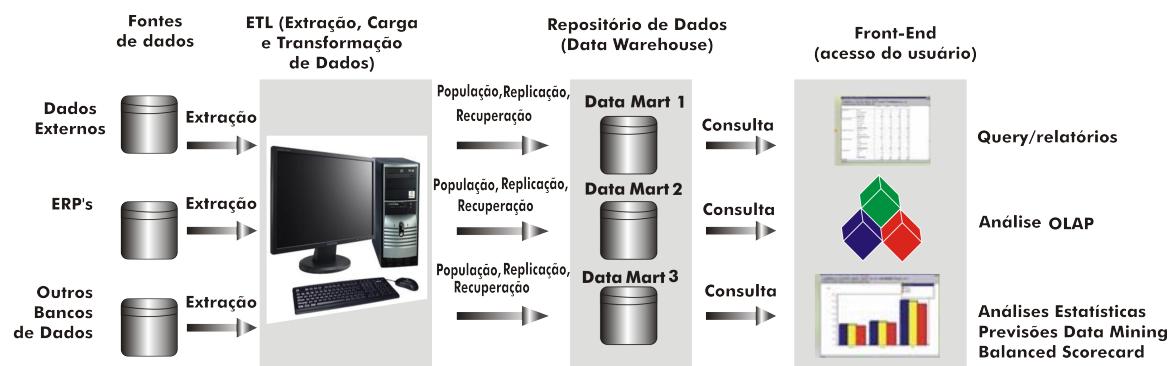


Figura 2.2 Arquitetura típica de um solução de Business Intelligence. Retirada de [Sys07].

2.3.4 Fatores-Chave de Sucesso para a implantação do BI

Devido ao recente desenvolvimento como solução tecnológica as ferramentas de BI possuem riscos de implantação associados que levam a uma apropriada discussão dos Fatores-Chave de Sucesso (FCS).

Segundo Furlan [Fur97] os FCS são aquelas poucas coisas que devem ocorrer de modo correto, ainda que em detrimento de outras que não estejam indo a contento, para alcançar os objetivos e garantir o sucesso. A lista dos FCS pode sofrer variações dependendo da ferramenta a ser abordada, mas, de maneira geral, é possível citar que é necessário que:

- se conheça a fundo o novo assunto sobre o qual se pretende trabalhar;

- possua planejamento e coordenação em todas as fases de duração do projeto;
- possua um canal de comunicação aberto entre todos os envolvidos no projeto;
- possua investimento em treinamento. Não se pode assumir que a equipe de desenvolvimento já saiba tudo sobre o novo projeto;
- haja o conhecimento da viabilidade técnica, de cronograma e financeira do projeto.

Um outro FCS fundamental é estipular um conjunto de medidas de desempenho que permitem avaliar os progressos obtidos na implantação, identificar os problemas e executar as ações corretivas necessárias. Por exemplo, quando o assunto é medida, um problema é definir o que medir, como utilizar os resultados e quais níveis de desempenho estipular, lembrando sempre que cada organização deve direcionar esta análise às suas necessidades e objetivos.

Segundo Poe, Klauer, Brobst [PKB98] consideram válidos outros FCS na construção de um *Data Warehouse* que aplicam-se quando tratamos de BI, a saber:

- projetar o BI com o foco no negócio, não na tecnologia. Não construir um BI apenas como uma tarefa técnica ou, simplesmente, para coletar todos os dados da empresa para dentro de um grande banco de dados. Deve haver um objetivo claro e definido que justifique sua construção, prevendo inclusive retorno do investimento, ROI (*Return On Investment*). Nos BI de sucesso, não existem decisões técnicas, somente decisões de negócio;
- executar a implantação em ciclos rápidos. Nos métodos tradicionais, o projeto é dividido em fases, em que a fase seguinte só pode ser iniciada após a conclusão da anterior. Já no BI, nem sempre é possível concluir totalmente uma fase antes de se partir para a próxima, por exemplo, a definição das telas, relatórios e consultas, pois tais especificações são interativas e não são conhecidas previamente. A solução é viabilizar um rápido retorno às fases usando uma metodologia de desenvolvimento iterativa, em que cada uma será repassada várias vezes numa série de refinamento;
- incluir a participação dos usuários finais na equipe de implantação. A participação dos usuários é necessária para projetar decisões que permitirão ao BI atingir os objetivos do negócio. O sucesso do BI está intimamente ligado aos usuários. Se estes não aceitarem o sistema ou não o utilizarem, o BI certamente fracassará.

Completando a lista, Meyer [MC98] expõem outros FCS garantindo que, se obedecidas as regras a seguir, a maioria das armadilhas que impossibilita um sistema de ser um sucesso estará sendo evitada:

1. Definir um responsável pela implantação do sistema. Este será o coordenador e promotor do projeto;
2. Desenvolver um protótipo rapidamente para que os usuários sintam o valor da ferramenta;
3. Começar a implantação passo a passo, permitindo mudanças no projeto inicial, protegendo assim seu investimento;
4. Identificar a estrutura e o escopo do seu projeto, bem como defina as expectativas dos usuários finais;
5. Criar uma estrutura flexível, incluindo software, hardware, sistemas operacionais;
6. Avaliar criticamente cada parte do hardware, software e todas as ferramentas envolvidas de forma a garantir o retorno sobre o investimento (ROI);
7. Acompanhar o projeto, envolver os usuários, medir e realizar as mudanças necessárias para o sucesso do sistema.

A seguir são detalhados cada um dos componentes da arquitetura de BI.

2.3.5 ETL - Extração, Transformação e Carga

Extração, transformação e carga (ETL) são os primeiros passos da criação e manutenção de um *data warehouse*. Ferramentas de ELT são oferecidas para o gerenciamento das três fases: extração seletiva de dados de um banco de dados, transformação em formatos para permitir as consultas e carga nas tabelas do *data warehouse* ou outro banco de dados. O alto volume de dados processados via ETL exige performance, horários críticos e desafios para os administradores de bancos de dados.

A extração é a primeira etapa do processo de obtenção de dados no ambiente de *data warehouse*. O processo de extração envolve a leitura e compreensão de dados de origem e cópia dos dados necessários ao *data warehouse* na *staging area* para que sejam manipulados posteriormente [KR02]. A *staging area* abrange tudo entre os sistemas operacionais de origem e a área de apresentação de dados.

Depois que os dados são extraídos para a *staging area* ocorrem transformações em potencial, como filtragem de dados (correções de erro de digitação, solução de conflitos de domínio, tratamento de elementos ausentes ou a divisão em formatos padrão), combinação de dados das várias origens, cancelamento de dados duplicados e atribuição de chaves de *data warehouse*.

Essas transformações são todas precursoras para carregar os dados na área de apresentação do *data warehouse* [KR02].

A *data staging area* é dominada pelas atividades simples de classificação e processamento sequencial. Em muitos casos esta área, não se baseia na tecnologia relacional, mas, em vez disso, pode ser formada por um sistema de arquivos simples. Em organizações com fonte de dados baseada em dados previamente normalizados em um banco de dados relacional, os dados chegam nesta etapa em um formato mais facilmente padronizado e habilitado para o passo seguinte, a carga.

O processo de carga no ambiente do *data warehouse* normalmente consiste de um processamento de apresentação de tabelas dimensionais garantidas por controle de qualidade para os vários recursos de carga. Após o banco de dados multidimensional carregado e validado por um controle de qualidade, os usuários passam a dispor do ambiente de consulta.

Uma vez realizada este último passo do ETL, a carga, os dados ficam organizados, armazenados e torna-se disponível para serem consultados diretamente pelos usuários, criadores de relatórios e outras aplicações de análise. O *data warehouse* está disponível porém a carga é efetivada em um novo modelo de dados, na maioria das vezes diferente do modelo transacional. Os conceitos de *data warehouse* ajudam a definir melhor esta nova estrutura.

2.3.6 *Data Warehouse*

Após o processo de extração dos dados no sistema de origem, realizada a transformação, inclusive com a possibilidade de realização de uma limpeza dos dados e a carga, tem-se um novo repositório de dados: o *Data Warehouse*.

2.3.6.1 Definição

Segundo Inmon [Inm02], *Data Warehouse* (DW) é uma coleção de dados, organizados por assunto, integrados, não voláteis e variável em relação ao tempo, cujo propósito é fornecer suporte à tomada de decisão. Trata-se de um banco de dados especialmente modelado para a produção de informações através de cruzamentos ilimitados de variáveis e valores.

Organizado por assunto pois são organizados em temas ou áreas específicas, envolvendo diversos sistemas de suporte à operação, provendo informações para os tomadores de decisões. Integrado pois dados residentes em diferentes ambientes, podem resultar em inconsistências na forma, conteúdo e conceito. Não volátil porque no repositório os dados são carregados em volume e não são alterados até a carga seguinte, mantendo assim a integridade das informações. Variável no tempo devido ao repositório criar registros históricos, conservando a informação,

para que esta possa ser analisada por períodos e só se adicionam informações na fase de carga do DW.

Para fornecer resposta as consultas realizadas, o acesso aos dados deve ser efetuado sobre uma base de dados consolidada, assim chamada por ser possível possuir informações de várias origens, ou fontes de dados, unificadas em um único local. Os dados são consolidados pelas ferramentas de ETL. Um vez consolidado, o repositório deve ser modelado para permitir que consultas, a análise dos dados, das mais diversas sejam processadas com alto desempenho pelas ferramentas analíticas.

2.3.6.2 *Data Mart*

Data Marts, muito utilizados em projetos pilotos, segundo Singh [Sin01], são subconjuntos do Data Warehouse da empresa inteira. Tipicamente desempenham o papel de um Data Warehouse departamental, regional ou funcional.

Uma empresa pode construir uma série de Data Marts ao longo do tempo e, eventualmente, vinculá-los por meio de um Data Warehouse lógico da empresa inteira. Os Data Marts são menores e de implantação mais rápida que um Data Warehouse.

Deve-se, no entanto, tomar cuidado com a proliferação dos Data Marts sem que se tenha por base um sistema que os consolide, entenda-se aqui um *Data Warehouse* corporativo. Do contrário, esses sistemas rapidamente estarão sem integração, gerando problemas de consolidação e inconsistência das informações. Não há problemas em se começar pequeno, mas não se pode perder a visão do sistema maior, o *Data Warehouse* corporativo [CAAT01].

Fazendo uma analogia com a realidade da Secretaria de Saúde do Estado de Pernambuco (SES-PE), o Sistema Gerenciador de Banco de Dados Relacional (SGBDR) disponibilizado nos hospitais não serve somente de repositório de dados para o Sistema de Atendimento aos Pacientes da Emergência (SAPE). Outros sistemas estão integrados a mesma base de dados, como o Ambulatório, Internação, Faturamento e Farmácia, para citar apenas alguns.

Os dados armazenados em um repositório passam a ser a fonte para diversos formas de solução de apoio à decisão. A maneira como analisar estes dados tem relação direta com os objetivos desejados de uma organização.

2.3.6.3 A Análise dos Dados em uma Organização

Um *data warehouse* é desenvolvido para prover um fácil acesso a origem dos dados de uma organização e também prover dados de qualidade para que possa melhorar a tomada de decisões gerenciais. O *data warehouse* é desenvolvido para um determinado fim, não devendo ser sim-

plesmente considerada a sua elaboração como o objetivo final. O objetivo do desenvolvimento é fundamentalmente a realização de análise e/ou produzir decisões através de uso dos dados.

Várias técnicas de análise de dados são utilizadas sobre a base de dados, atuando em níveis diferentes da organização, disponibilizando diferentes perspectivas dentro desta organização. Desta forma o conteúdo da base de dados deve ser analisado sob determinada perspectiva para proporcionar o conteúdo certo para cada nível da organização.

As técnicas de análise de dados podem influenciar o tipo de modelo de dados a ser selecionado e no conteúdo da base de dados. Estas técnicas seguem uma certa evolução no desenvolvimento de soluções corporativas e o entendimento dos dados e sua utilização facilitam a escolha do modelo a ser aplicado.

Se a necessidade da organização é uma necessidade mais operacional, ou seja, uma necessidade de consultar os dados e relatórios do resultado desta consulta com rapidez e com facilidade, necessita de um modelo que organize os dados de tal forma que possam ser agrupados, indexados e classificados. Para tal necessidade, utilizamos o modelo de entidade e relacionamento que tem demonstrado, na prática, grande eficiência e bons resultados.

Sendo a necessidade da empresa mais gerencial, ou seja, analisar as tendências e o desempenho da organização, o modelo mais recomendado passa a ser a análise multidimensional de dados, a qual requer que o modelo de dados suporte uma estrutura que permita mais rapidez em algumas combinações de dimensões e de fácil acesso à base de dados para produzir resultados que tenham significados para a tomada de decisões.

A análise multidimensional adequa-se a este tipo de modelo, pois para a tomada de decisões, os analistas das organizações devem ter várias visões e perspectivas dos dados.

Se a necessidade empresarial é a descobertas de novas oportunidades de negócio, o modelo mais recomendado é o *data mining*. Um modelo de *data mining* requer uma base de dados atemporal, pois o mesmo fará uma pesquisa pelo histórico dos dados da organização e tentará descobrir padrões de comportamento, com base no quais, proporá uma nova oportunidade de negócio.

No caso específico da Secretaria de Saúde, uma vez identificada a necessidade de apoio a gestão na tomada de decisão, partindo-se de informações existentes nas emergências dos hospitais públicos, a alternativa passa a ser a elaboração de um *data warehouse*. Para este fim, um novo modelo de dados necessita ser apresentado para dar suporte as ferramentas de BI: o modelo multidimensional.

2.3.6.4 Modelagem de dados em um Data Warehouse

A primeira fase no projeto de um repositório é a definição do modelo de dados. Nesta fase, são analisados os dados necessários e que podem ser obtidos junto aos sistemas transacionais. É então criado um modelo de dados corporativo, que difere do modelo corporativo operacional à medida que são retirados os dados que são utilizados apenas no ambiente operacional e incorporados elementos de tempo nas chaves, além de haver o agrupamento dos dados conforme sua utilização [KR02], [Sel01].

Quando tratamos de sistemas de informação, o dado é atualizado e normalmente não se mantém um registro histórico das mudanças. Diferentemente, em um DW o dado normalmente não é atualizado. Para uma atualização das informações uma nova carga necessita ser realizada.

Deve-se considerar também a legibilidade do modelo de dados, ou seja, a facilidade com que os usuários conseguem compreender como os dados estão sendo persistidos. Ferramentas gráficas que expõem um modelo entidade-relacionamento ao usuário para possibilitar pesquisas customizadas muitas vezes promovem uma enorme confusão com um emaranhado de tabelas e relacionamentos.

Com o objetivo de simplificar a visualização, os desenvolvedores de bancos de dados adotaram alguns conceitos para simplificar a apresentação de informações analíticas, apresentando um modelo mais adequado à realidade de um DW, o modelo dimensional.

Segundo Kimball [KR02], modelagem dimensional é um novo nome para uma técnica antiga que permitia tornar os bancos de dados fáceis e comprehensíveis. Caso após caso, a partir da década de 1970, as empresas de TI, as consultorias, os usuários finais e os fornecedores migraram para uma estrutura dimensional mais simples para atender a necessidade humana fundamental de simplicidade. A capacidade de visualizar algo abstrato como um conjunto de dados de forma concreta e tangível é o segredo da capacidade de compreensão.

Esse modelo dimensional lembra a idéia do cubo contendo três ou mais dimensões, cada uma representando um atributo diferente como apresentado na Figura 2.3.

Um modelo dimensional é constituído de uma ou várias tabelas de fato, que conterão as medidas do negócio a serem analisadas, e, em torno, várias tabelas descritivas, chamadas tabelas de dimensão.

Um pouco mais detalhadamente, fatos são tabelas que armazenam medidas numéricas totalizáveis relacionadas a um determinado assunto ou processo de negócio. Cada registro em uma tabela de fato está relacionado a um conjunto de dimensões que determinam a granularidade dos fatos armazenados e definem qual o escopo dessas medidas. Quanto menor a granularidade de um fato, maior será o nível de detalhe armazenado.

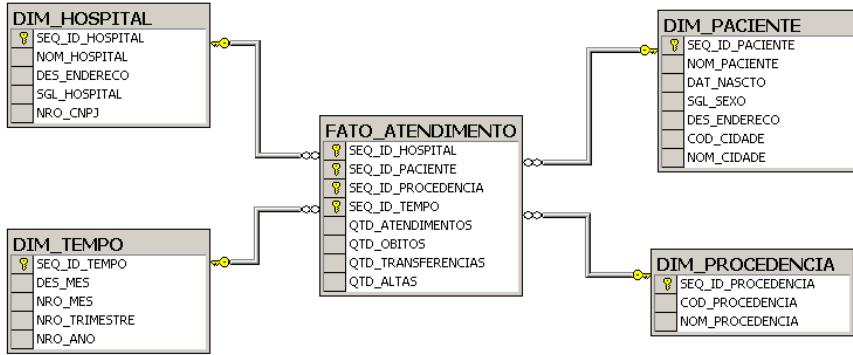


Figura 2.3 Modelo de Dados Multidimensional.

As tabelas de dimensões armazenam descrições textuais de um processo de negócio, podendo cada uma ser formada por um conjunto de atributos, denominados hierarquias, que servem como base para determinar regras de agrupamento, quebras e filtros em consultas a uma tabela de fato.

No exemplo ilustrado na Figura 2.3, dimensões descritivas de pacientes, procedência do paciente, hospital de atendimento e a data de registro. Essas dimensões estão associadas a uma tabela de fatos. A tabela de fato, fato-atendimento, representa quantidade de atendimentos, óbitos, transferências e altas médicas atendidos em determinado hospital em determinado período. O usuário pode "girar" o cubo para obter diferentes perspectivas sobre os fatos a partir das quatro dimensões ligadas à tabela de fatos. Cada tabela de dimensão possui uma única ligação com a tabela de fatos, a qual é feita através de chaves externas. Na Tabela 2.2 apresenta-se uma comparação entre fatos, dimensões e medidas.

O modelo dimensional, além de agilizar o processamento das consultas, permite uma melhor visualização dos dados devido à forma simples de organizá-los e propicia mais flexibilidade para eventuais ajustes no modelo [KR02]. Na segunda fase da modelagem é construído o modelo físico, onde são incluídas características físicas e chaves. Tendo-se definido os atributos que figurarão nas tabelas de fatos e nas dimensões, resta ainda decidir sobre importantes aspectos de organização dos dados e performance, representados inicialmente pela definição da granularidade, do particionamento e da definição de agregados.

O modelo dimensional apresentado na Tabela 2.2 utiliza o chamado Esquema Estrela, utilizando-se a analogia da tabela de fatos como um centro e as dimensões suas "estrelas".

	Fatos	Dimensões	Medidas
Escopo	Representam um item, uma transação ou um evento de negócio.	Determinam um contexto de um assunto de negócios, como, por exemplo, uma análise de produtividade dos grupos de pesquisa.	São os atributos numéricos que representam um fato e são determinados pela combinação das dimensões que participam deste fato.
Objetivo	Refletem a evolução do negócio.	São os balizadores de análise de dados.	Representam o desempenho de um indicador de negócios relativos às dimensões que participam de um fato.
Tipo de Dado	São representados por conjuntos de valores numéricos (medidas) que variam ao longo do tempo.	Normalmente não possuem atributos numéricos, pois são somente descritivas e classificatórias dos elementos que participaram de um fato.	Podem possuir uma hierarquia de composição de seu valor.

Tabela 2.2 Comparação entre fatos, dimensões e medidas. Adaptado de [Mac00]

As vantagens deste esquema, de acordo com Poe [PKB98] podem ser descritas como:

1. permitir a criação de um projeto de banco de dados que fornecerá respostas rápidas, com menos tabelas e índices;
2. permitir ao administrador do banco de dados trabalhar com projetos mais simples e assim produzir melhores planos de execução;
3. possuem uma estrutura mais intuitiva, assemelhando o projeto do banco de dados com a forma como o usuário final pensa e utiliza os dados;
4. simplificar o entendimento e a navegação dos metadados para os usuários e desenvolvedores;
5. ampliar as opções de escolha da ferramenta front-end de acesso aos dados devido a majoritária adoção deste esquema nas soluções.

Somente como motivo de comparação, uma alternativa ao Esquema Estrela é o chamado Foco de Neve, ou *Snowflake* que consiste na realidade em uma normalização do Esquema Estrela. No esquema *Snowflake*, as tabelas de Dimensão são estruturadas de modo que atendam a

terceira forma normal, mantendo as tabelas de Fato em sua estrutura inicial. Para atender a terceira forma normal, na tabela de Dimensão não pode haver campos que sejam dependentes de outros campos que não sejam a própria chave primária. Tabelas chamadas de Sub-Dimensões são criadas para atender à normalização.

No capítulo referente ao Estudo de Caso o modelo dimensional aplicado à Secretaria de Saúde, utilizando Esquema Estrela, será apresentado detalhadamente no intuito de servir de suporte a utilização de ferramentas no auxílio à gestão estratégica.

2.3.6.5 Cubo de dados

Segundo Machado [Mac00], Cubo de dados é uma estrutura multidimensional que expressa a forma na qual os tipos de informações se relacionam. É formado pela tabela de fatos e pelas tabelas de dimensão que a circundam e representam possíveis formas de visualizar e consultar os dados. O cubo armazena todas as informações relacionadas a um determinado assunto, de maneira a permitir que sejam montadas várias combinações entre elas, resultando na extração de várias visões sobre o mesmo tema.

A idéia fundamental é que praticamente todo o tipo de negócio pode ser representado por uma espécie de cubo de dados, onde as células do cubo contêm valores mensuráveis e as bordas do cubo definem as dimensões naturais dos dados.

Pode-se necessitar utilizar mais de três dimensões nos modelos, passando o cubo a ser chamado de hiper-cubo. De forma resumida, o cubo é uma estrutura multidimensional que armazena os dados tornando-os mais fáceis de se analisar. Cada cubo contém uma tabela de fatos e várias dimensões. Por exemplo, um cubo contendo informações de vendas poderá ser composto pelas dimensões tempo, região, produto, cliente, cenário (orçado ou real) e medidas como valor de vendas e custos, conforme a Figura 2.4

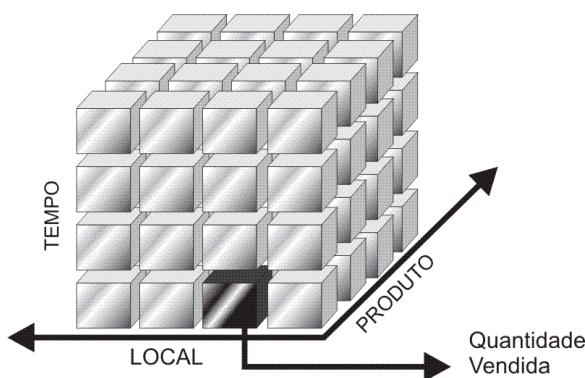


Figura 2.4 Cubo de Dados.

Algumas ferramentas de BI são utilizadas para a criação de cubos gerando os esquema dos dados de um modelo multidimensional a ser disponibilizado para uma ferramenta OLAP. No capítulo referente ao Estudo de Caso, uma ferramenta de software livre será apresentada para a realização desta funcionalidade.

Os cubos são construídos para permitir rápido acesso aos dados, sendo normalmente construídos a partir de subconjuntos do modelo de dados do DW, organizados dentro de estruturas multidimensionais definidas pelas dimensões e medidas.

2.3.7 Área de Apresentação

A área de apresentação de dados, também chamada de *Front-end*, é o local em que os dados ficam organizados, armazenados e tornam-se disponíveis para serem consultados diretamente pelos usuários, criadores de relatórios e por outras aplicações de análise [KR02].

A interface que apresenta os dados disponíveis ao usuário é normalmente composta, no ambiente de *Business Intelligence*, por ferramentas de criação de relatórios padronizados, de consultas *ad-hoc*, que são as Ferramentas OLAP (*On-Line Analytical Processing*) e de *Data Mining*.

O *Data Mining* é adequado para analisar grupos de dados que são difíceis de analisar usando apenas a função OLAP, visto que esses grupos são grandes demais para serem "navegados", ou explorados manualmente, ou ainda porque contêm dados muito densos ou não-intuitivos para serem compreendidos.

Cada uma destas tecnologias desempenha um papel fundamental nesta camada. As ferramentas de criação de relatórios padronizados possibilitam a criação de relatórios pré-definidos e que geralmente, são os mais comumente utilizados pelos usuários. As ferramentas de consulta OLAP possibilitam a criação de relatórios personalizados. Já as ferramentas de *Data Mining*, que são mais complexas, são utilizadas para a descoberta de padrões e relações entre os dados, informações estas, que geralmente não se consegue obter com as outras ferramentas.

Assim, a diferença básica entre OLAP e *Data Mining* está na maneira como a exploração dos dados é realizada. Na análise OLAP, a exploração é feita através da verificação, isto é, o analista conhece a questão, elabora uma hipótese e utiliza a ferramenta para refutá-la ou confirmá-la. Com o *Data Mining*, a questão é total ou parcialmente desconhecida e a ferramenta é utilizada para a busca de conhecimento.

Ambos os tipos de ferramentas normalmente contam com o apoio de algum tipo de metadado. Esse metadado, entretanto, normalmente restringe-se à descrição sintática do modelo de dados do DW e a configurações específicas dessas ferramentas, como a definição de hierarquias e

dimensões dos cubos em ferramentas OLAP e a definição dos parâmetros de entrada para os algoritmos de *Data Mining* [Ale02].

Motivado pela importância da ferramenta OLAP na solução de BI, seguem-se os conceitos de aplicações OLAP.

2.3.8 OLAP

O *Data Warehouse*, conforme citado, é basicamente uma ferramenta que visa o armazenamento de dados, porém que não dispõe de recursos para realizar consultas e análises muito sofisticadas ou cálculos mais complexos.

Para esses casos, existe uma outra ferramenta de BI, o OLAP (*On-Line Analytical Processing*), que possui a capacidade de processamento dos dados. O grande diferencial do OLAP está principalmente na sua capacidade de permitir consultas e análise dos dados de forma consistente, interativa, rápida e com uma grande variedade de possíveis visões dos dados. No OLAP, as respostas não são automáticas. Trata-se de um processo interativo em que o usuário formula hipóteses, faz consultas, recebe informações, verifica um dado específico em profundidade e faz comparações.

O OLAP ajuda os usuários a sintetizarem as informações sobre a empresa por meio de comparações, visões personalizadas, análises estatísticas, previsões e simulações. Assim como o *Data Warehouse*, o OLAP também pode ser usado em diversos departamentos organizacionais, como Finanças, Vendas, Marketing e Manufatura [Bis04].

Segundo Pendse [Pen00] o termo OLAP foi definido por E. F. Codd, que criou em 1993 um conjunto de doze regras que serviriam de referência para a avaliação e desenvolvimento da ferramenta. Em seu trabalho, [Pen00] apresenta e questiona essas regras e mais seis que foram incluídas em 1995, devido a limitações das doze iniciais, agora chamadas de características. Bispo [Bis04] também compartilha dessas doze regras ou características, e lista as capacidades que uma ferramenta OLAP deve possuir, a saber:

- permitir a visão multidimensional dos dados;
- realizar cálculos e análises estatísticas complexas, fazer comparações e analisar tendências;
- criar agregações e consolidações;
- construir cenários a partir de suposições e fórmulas aplicadas aos dados históricos;
- navegar com facilidade entre os diversos níveis de agregação dos dados;

- analisar os dados, por meio de qualquer combinação possível entre eles;
- analisar os dados de acordo com as regras do negócio, como crescimento, mudanças, alterações e análises de mercado;
- permitir aos usuários uma grande variedade de opções de relatórios;
- trabalhar com múltiplas fontes de dados, como *Data Warehouse*, bancos de dados operacionais e outras fontes externas de dados;
- dar suporte a ambientes de *workgroup* e ambientes distribuídos;
- possuir interface com diversas outras ferramentas de uso comum, como planilhas eletrônicas, geradores de relatórios, geradores de gráficos, *Web Browsers*, dentre outros.

Segundo Thomsen [Tho02] o objetivo das ferramentas OLAP é permitir análises para analistas de negócios, tendo requisitos bem particulares comparados a outros sistemas de informação:

- flexibilidade: Os analistas de negócio devem ter liberdade e facilidade para escolher os dados a serem analisados bem como o formato no qual eles devem ser visualizados;
- simplicidade: Ferramentas OLAP devem prover maneiras simples e intuitivas de confecção de análises para os analistas de negócio;
- expressividade: A linguagem de consulta utilizada pela ferramenta OLAP deve ser poderosa o suficiente para que o analista consiga extrair informações realmente úteis para o negócio;
- poder de análise: Ferramentas OLAP devem prever um conjunto de operações de agregação e exploração de dados que permitam análises de tendências e comparações complexas utilizando os dados que descrevem o negócio;
- velocidade: O processamento das consultas sobre o DW deve ser realizado em um tempo relativamente baixo. O tempo de processamento, dependendo do volume de dados e da complexidade do código da consulta pode em casos específicos chegar a horas.

Um dos desafios no projeto de aplicações OLAP refere-se a como aliar a simplicidade para o usuário à expressividade nas consultas e análises. Parte da resposta é obtida através da maneira como os dados são organizados no modelo dimensional [KR02]. Como descrito

anteriormente, uma das características desse modelo em relação aos modelos normalizados é a legibilidade. Porém, nas soluções atuais, usuários ainda assim precisam lidar com descrições técnicas do modelo de dados. Para definir suas análises, os usuários deverão entender como os dados estão organizados no modelo dimensional. Através do modelo dimensional, o usuário tem liberdade para visualizar diversos aspectos do negócio sob a perspectiva que entenda ser a mais interessante para sua análise [Mac00]. A amplitude de análise desse modelo reside no fato de que um grande número de aspectos do negócio pode ser expresso em termos desses dois conceitos [Ber07] e [Kim98].

2.3.8.1 Arquitetura OLAP

De acordo com Cunha [Cun04] e Cielio [Cie05], a arquitetura OLAP diz respeito ao método de armazenamento de dados utilizado para uma aplicação OLAP. Os métodos de armazenamento de dados são multidimensional, relacional, híbrido, *desktop* e *web*. Cada um deles tem uma função específica e deve ser utilizada quando melhor atender às necessidades de análise pela ferramenta de OLAP. As características destas arquiteturas são:

MOLAP (*Multidimensional On Line Analytical Processing*): Com um servidor multidimensional o acesso aos dados ocorre diretamente no banco de dados, ou seja, o usuário trabalha, monta e manipula os dados do cubo diretamente no servidor. Isso traz grandes benefícios aos usuários no que diz respeito à performance, mas tem problemas com escalabilidade além de ter um custo alto para aquisição. A Figura 2.5 demonstra esta arquitetura.

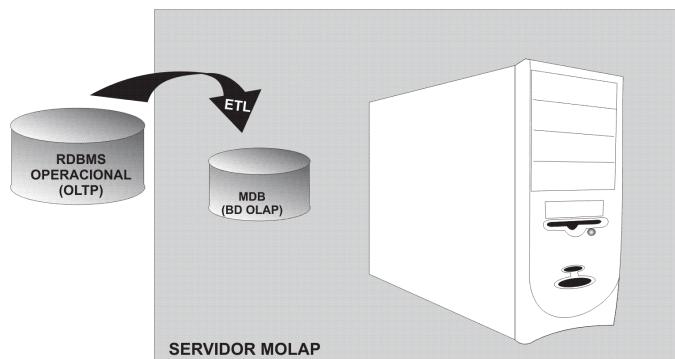


Figura 2.5 Arquitetura MOLAP. Retirada de [Cun04].

ROLAP (*Relational On Line Analytical Processing*): Possuem uma engenharia de acesso aos dados e análise OLAP com uma arquitetura um pouco diferente. Nesse caso a consulta é enviada ao servidor de banco de dados relacional e processada no mesmo, mantendo o cubo no servidor. O que se pode notar nesse caso é que o processamento OLAP se dará somente no

servidor. A principal vantagem dessa arquitetura é que ela permite analisar enormes volumes de dados, em contra partida uma grande quantidade de usuários acessando simultaneamente poderá causar sérios problemas de performance no servidor. A Figura 2.6 demonstra esta arquitetura.

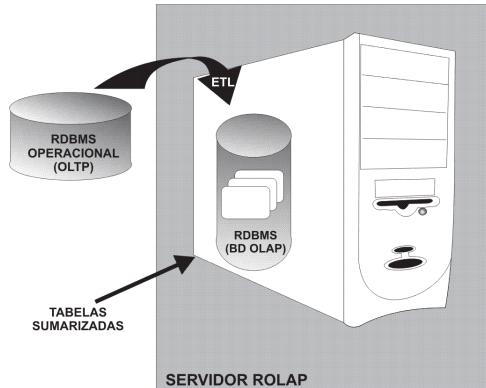


Figura 2.6 Arquitetura ROLAP. Retirada de [Cun04].

HOLAP (Hybrid On Line Analytical Processing): Essa forma de acessar os dados é uma mistura de tecnologias onde há uma combinação entre ROLAP e MOLAP. A vantagem é que com esta mistura de tecnologias pode-se extrair o que há de melhor de cada uma, ou seja, a alta performance do MOLAP com a melhor escalabilidade do ROLAP. A Figura 2.7 demonstra esta arquitetura.

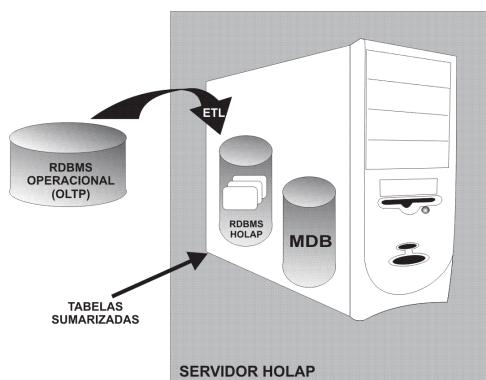


Figura 2.7 Arquitetura HOLAP. Retirada de [Cun04].

DOLAP (Desktop On Line Analytical Processing): São as ferramentas que, a partir de um cliente qualquer, emitem uma consulta para o servidor e recebem o cubo de informações de volta para ser analisado na estação cliente. O ganho com essa arquitetura é o pouco tráfego que se dá na rede, visto que todo o processamento OLAP acontece na máquina cliente, e a maior agilidade de análise, além do servidor de banco de dados não ficar sobrecarregado, sem incorrer

em problemas de escalabilidade. A desvantagem é que o tamanho do cubo de dados não pode ser muito grande, caso contrário, a análise passa a ser demorada e/ou a máquina do cliente pode não suportar em função de sua configuração. Esta arquitetura é exemplificada na Figura 2.8.

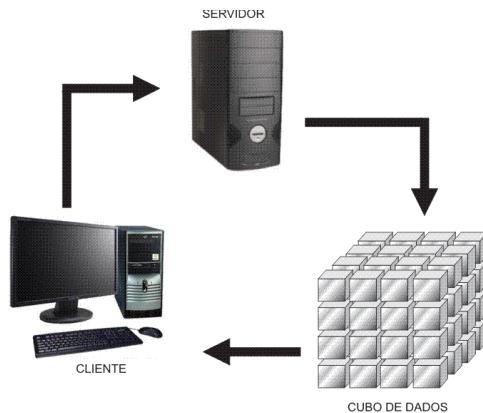


Figura 2.8 Arquitetura DOLAP. Retirada de [Cun04].

WOLAP (Web On Line Analytical Processing): É a utilização de uma ferramenta OLAP a partir de um navegador Web, *browser*. A arquitetura das ferramentas WOLAP é uma variação da cliente/servidor. A diferença está na utilização de um middleware do lado servidor que será o responsável pela comunicação entre o cliente e uma aplicação servidora, neste caso, o banco de dados. A Figura 2.9 exemplifica esta arquitetura.

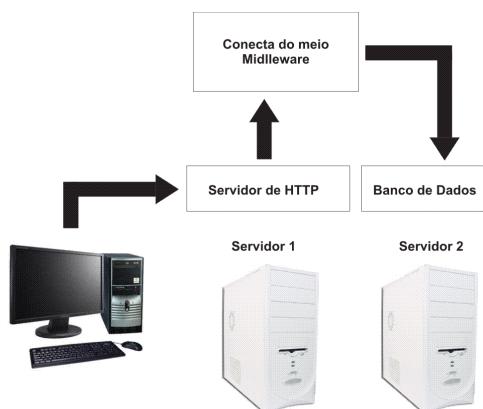


Figura 2.9 Arquitetura WOLAP. Retirada de [Cun04].

A ferramenta OLAP utilizada neste trabalho, que será apresentada no capítulo 4, utiliza-se dos métodos multidimensional, relacional e híbrido.

2.3.8.2 Operações de OLAP

A análise multidimensional fornece uma maneira de estender a capacidade de consultas e relatórios. Em uma pergunta do tipo: "Quanto de cada um dos produtos foram vendidos em determinado dia, por um determinado representante e em determinado local?", cada parte da pergunta corresponde a uma consulta em determinada dimensão do modelo de dados. A análise multidimensional permite ao usuário avaliar um grande número de fatores interdependentes envolvendo o problema associado ao negócio e visualizar dados em relacionamentos complexos.

Os gestores envolvidos com a solução estão interessados em explorar dados de diferentes níveis de detalhes. Bem melhor se disponibilizados dinamicamente e com tempo de resposta compatível com a evolução da análise. Estes relacionamento podem ser analisados através de um processo iterativo que incluem aprofundar-se (*drilling down*) nos níveis de detalhe ou subir ao nível mais alto de sumarização e agregação (*rolling up*), de acordo com a granularidade implementada

De acordo com Barbieri [Bar01] e Singh [Sin01], as ferramentas OLAP permitem ao usuário, navegar entre diferentes granularidades de um cubo de dados. O termo granularidade refere-se ao nível de detalhe do processo de negócio. Especificar a granularidade significa especificar exatamente o que uma linha da tabela de fatos representa. O grão exprime o nível de detalhe associado às medidas da tabela de fatos [KR02].

Antes de detalharmos cada operação possível, a Figura 2.10 demonstra que o usuário pode iniciar sua análise pelo total de vendas da empresa e executar um "*drill-down*" para visualizar as vendas por continente, região, país e finalmente cliente. Se desejar, é possível iniciar a análise pelo cliente e executar um "*drill-up*" através dos diferentes níveis para finalmente alcançar o total de vendas.

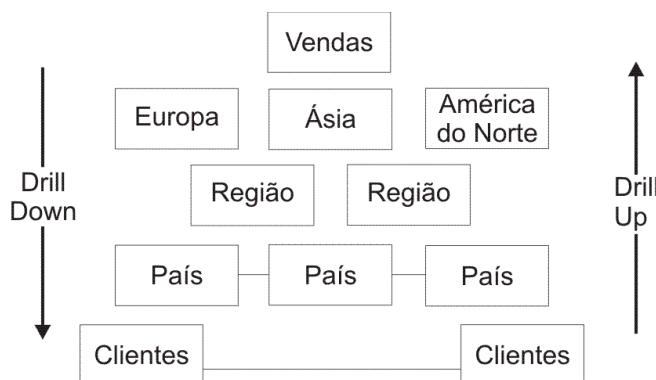


Figura 2.10 Análise *drill-down* e *roll-up*.

As mais freqüentes funcionalidades oferecidas pelas ferramentas OLAP são:

Drill-down: aumento do nível de detalhe da informação, por parte do usuário, para solicitar uma visão mais detalhada em um conjunto de dados. Por exemplo, se os dados de vendas estiverem sumarizados e o usuário desejar informações mais detalhadas, por vendedor que realizou a venda, por exemplo, realiza-se então um *drill-down* para que os dados sejam específicos de cada vendedor.

Drill-up: é o contrário do *drill-down*, ocorre quando o usuário diminui o nível de detalhe da informação, solicitando uma visão menos detalhada de um conjunto de dados. Por exemplo, se o usuário está visualizando a venda de um determinado vendedor e passa a analisar as vendas sumarizadas de todos os vendedores.

Drill-across: é a requisição de dados das tabelas de dimensão com o valor das consultas modificada, ocorre quando o usuário "pula"um nível intermediário dentro de uma mesma dimensão. Por exemplo, em uma dimensão tempo composta por ano, semestre, trimestre, mês e dia. O usuário estará executando um *drill-across* quando ele passar de ano direto para trimestre ou mês.

Drill-through: é quando se deseja uma informação em um nível de detalhe menor do que aquele armazenado na tabela de fato e permitido pela sua granularidade. Por exemplo, em um nível de granularidade de produto por dia e por loja deseja-se uma informação que está presente na nota fiscal correspondente. O *drill-through* é a operação que busca esta informação além do nível de granularidade existente na estrutura dimensional. Esta busca poderia ser realizada no próprio sistema transacional de origem, no caso de compatibilidade entre ambos.

Slice-dice: é a descrição padrão para a habilidade de acessar os dados do *data warehouse* através de qualquer uma das dimensões de forma igual. Serve para modificar a posição de uma informação, alterar linhas por colunas de maneira a facilitar a compreensão dos usuários e mudar de dimensão sempre que necessário. Esta operação equivale a "fatiar"o cubo de dados, ou seja, um determinado valor de uma dimensão é fixado. Por exemplo, podemos fixar o valor da região Sul para uma dimensão Localidade, e analisar as outras dimensões. Assim, pode ser conhecida a eficiência de entrega desta região em função dos produtos (explorando os valores da dimensão Produtos) e da forma de distribuição utilizada (dimensão Distribuição). Esta operação permite conhecer em detalhes o relacionamento de um valor de uma dimensão com as outras dimensões.

Pivoting (pivotamento): é a mudança do arranjo das linhas e colunas em um relatório tabular, onde freqüentemente as linhas ou as colunas são derivadas de dimensões diferentes. É a inversão dos eixos das dimensões, para obter-se novas visões de consultas. Em algumas ferramentas, para que seja realizada esta operação, podem-se simplesmente utilizar as funções de "arrastar e soltar"com o mouse.

Consultas *ad-hoc*: de acordo com [Inm02], são consultas com acesso casual único e tratamento dos dados segundo parâmetros nunca antes utilizados, geralmente executadas de forma iterativa e heurística. Isso tudo nada mais é do que o próprio usuário gerar consultas de acordo com suas necessidades de cruzar as informações de uma forma não vista e com métodos que o levem a descoberta daquilo que procura.

A elaboração de um *Data Warehouse*, como a maioria dos projetos nas mais diversas áreas, requer a execução de atividades ou tarefas, algumas concorrentes, outras seqüenciais. A seguir apresentam-se as fases deste projeto.

2.3.9 Fases de Desenvolvimento do Data Warehouse

A implementação de um *Data Warehouse* requer atividades intimamente integradas. Todos que façam parte da equipe, incluindo o analista de negócio, o arquiteto, o projetista de banco de dados, o testador de dados e o desenvolvedor de aplicação analítica, necessitam de um entendimento do ciclo de vida completo de um *Data Warehouse* [KR02].

O diagrama da Figura 2.11 serve para encapsular as atividades principais do ciclo de vida dimensional do negócio. O diagrama ilustra a seqüência de tarefas, a dependência e a concorrência (simultaneidade). Porém não reflete uma cronologia absoluta bem como o tamanho das caixas não refletem o esforço despendido.

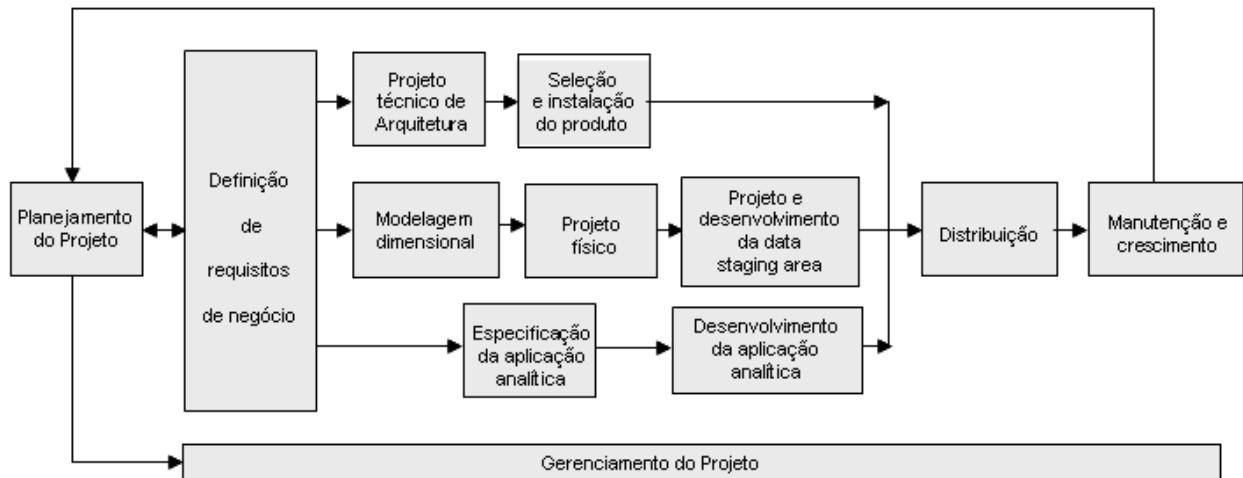


Figura 2.11 Diagrama do ciclo de vida dimensional do negócio. Retirado de [KR02].

O ciclo de vida começa com o planejamento do projeto, servindo como preparação da iniciativa de um *data warehouse* na organização. Esta fase serve para estabelecer o escopo e a justificativa preliminar, a obtenção dos recursos e o lançamento do projeto. O gerenciamento contínuo serve como base para manter o restante do ciclo de vida no caminho correto.

A segunda principal tarefa se concentra na definição dos requisitos de negócio. A seta bidirecional entre o planejamento do projeto e a definição de requisitos de negócio, pois há muita interação entre as duas atividades. O alinhamento do *data warehouse* com os requisitos de negócio é crucial. Por melhor que seja a tecnologia não se viabiliza um *data warehouse* que não esteja alinhado com o negócio da organização. Os projetistas devem entender as necessidades do negócio e traduzi-las em considerações de projeto. O usuário de negócio e seus requisitos têm um impacto sobre quase toda decisão de projeto e implementação tomada no curso de um projeto de *data warehouse*. Na Figura 2.11 as três linhas paralelas refletem esta importância.

A linha superior da Figura 2.11 trata da tecnologia. O projeto técnico da arquitetura estabelece a estrutura geral para suportar a integração de múltiplas tecnologias. Importante observar que a seleção do produto não é a primeira etapa do diagrama. Desta forma o produto impacta o menos possível no entendimento da solução desejada.

A linha central originada na definição dos requisitos de negócio se concentra nos dados. A fase inicial é a tradução dos requisitos em um modelo dimensional e a transformação deste modelo dimensional transformado em uma estrutura física. Nesta fase há uma concentração nas estratégias de ajuste como agregação, indexação e particionamento, durante as atividades físicas do projeto. Por último, mas não menos importante, o processo de Extração, Transformação e Carga (ETL) de dados são projetados e desenvolvidos.

O conjunto final de tarefas gerado pela definição dos requisitos de negócio é o projeto e o desenvolvimento de aplicações analíticas. O projeto de *data warehouse* não é realizado quando do processo de ETL. As aplicações analíticas, na forma de modelos e análises orientadas ao parâmetro, visam atender um grande percentual das necessidades analíticas dos usuários de negócio.

A tecnologia, os dados e a aplicação analítica somam-se a um amplo treinamento e suporte para a fase final de distribuição. A manutenção contínua faz-se necessária para garantir que o *data warehouse* e sua comunidade de usuários permaneçam parceiros e conquistem o retorno do investimento.

2.3.10 Periodicidade de Carga do *Data Warehouse*

Na manutenção de um DW a freqüência com que uma nova carga é executada deve seguir uma periodicidade que pode variar de acordo com a necessidade de atualização das informações para análise do usuário e a disponibilidade dos dados a serem carregados.

Uma vez definida a necessidade de uma nova carga nas dimensões e fatos uma estratégia

de carga precisa ser definida. Segundo Kimball e Caserta [KC04] a maneira como a carga é realizada depende das características das tabelas de dimensão. As dimensões que evoluem no tempo e que necessitam de pequenas adições para capturar a mudança ao longo do tempo são denominadas *Slowly Changing Dimensions* (SCD) definindo-se técnicas de três tipos, a saber:

2.3.10.1 Slowly Changing Dimension Tipo 1 - atualizar

O tipo 1 de SCD é simplesmente sobrescrever um ou mais atributos existentes em um registro de uma dimensão. O processo de ETL deve optar pelo tipo 1 se as fontes de dados estiverem sendo corrigidas ou se não existe interesse em manter um histórico de valores anteriores com emissão de consultas a informações de cargas anteriores [KC04].

Algumas ferramentas de ETL oferecem otimização que podem detectar quanto um registro necessita ser inserido ou alterado. Contudo este processo deve pesquisar na tabela a existência da chave primária do registro candidato, exigindo muitos recursos do banco de dados, inclusive aumentando sobremaneira os arquivos de log do SGBD.

Uma opção para o tipo 1 de uma carga com muitas alterações é preparar a nova dimensão em um arquivo em separado, deletar os registros da dimensão e carregando novamente.

2.3.10.2 Slowly Changing Dimension Tipo 2 - Mantendo o Histórico

O SCD tipo 2 é a técnica padrão para as buscas apuradas de mudanças nas tabelas de dimensão associando-as corretamente com a tabela de fatos [KC04]. A idéia básica é simples. Quando é identificada a necessidade de nova carga de uma dimensão, diferentemente de sobrescrever, um novo registro é inserido. A este novo registro é atribuída uma nova chave artificial. Esta nova chave será associada a partir deste momento a todos os fatos que apontam para o registro da dimensão. Como a nova chave é criada nenhum atributo na tabela de fatos precisam ser alterado.

Há a garantia do histórico porque cada versão de registro na dimensão está associada aos fatos anteriormente armazenados. Em algumas situações recomenda-se a inclusão de alguns atributos às dimensões como indicadores da situação do registro e datas de validade.

2.3.10.3 Slowly Changing Dimension Tipo 3 - Realidades Alternativas

O tipo 3 de SCD é utilizado quando as mudanças acontecem aos registros da dimensão porém os antigos valores das colunas permanecem válidos como uma segunda alternativa. Nesta técnica uma nova coluna é criada com o novo valor do registro e o antigo valor é colocado neste campo.

Novas chaves não são necessárias nas dimensões e na tabela de fatos [KC04].

Quando um novo registro é adicionado a dimensão a regra de negócio deve ser requisitada para decidir onde colocar o antigo valor. Valores nulos podem ser inseridos, se assim determinado. Com a inclusão de uma nova coluna o usuário pode escolher entre as duas versões dos dados. Esta realidade alternativa podem ser incrementadas a cada inserção de colunas. O controle de qual coluna representa as realidades requer mais processamento nas consultas.

2.3.10.4 Slowly Changing Dimension Híbrida

Uma outra alternativa de SCD corresponde a mudanças nos atributos das dimensões analisando campo a campo as opções. Torna a solução mais flexíveis e completas, porém aumentam consideravelmente a complexidade no processo de ETL de um DW.

2.4 Considerações finais

Por tratar-se de uma pesquisa com aplicação de um modelo de *Business Intelligence* (BI), coube a este capítulo os conceitos e métodos de trabalho deste tipo de solução.

O detalhamento sobre os tipos de sistema transacionais torna-se importante na busca de destacar suas particularidades, funcionalidades e evolução, até chegar aos atuais sistemas BI e a descrição de suas novas ferramentas: ETL, *Data Warehouse* e OLAP.

As Fases de Desenvolvimento do Data Warehouse permeiam todo o projeto de solução sendo constantemente referenciado ao longo dos capítulos. Um esclarecimento sobre a maneira de realizar o processo de ETL, o SCD, serve de base para a escolha do procedimento a ser adotado pela SES-PE.

No capítulo seguinte apresenta-se um modelo utilizando-se a modelagem multidimensional, iniciando com o desenvolvimento do *data warehouse* concluindo com a implantação de uma aplicação de software livre de BI.

A apresentação destes conceitos e da metodologia de uso das ferramentas servem de base desta aplicação e, fundamentalmente, a discussão sobre os Fatores-Chaves de Sucesso tem sua importância ressaltado pelo fato do compromisso assumido do produto final estar disponibilizado para a Secretaria Estadual de Saúde.

CAPÍTULO 3

Trabalhos Relacionados

O desenvolvimento de uma solução de *Business Intelligence* encontra investimentos nas diversas áreas, com maior ênfase nas empresas privadas. Com relação às instituições públicas percebe-se uma maior dificuldade de identificação de soluções.

Neste capítulo apresentam-se alguns trabalhos relacionados e suas características particulares identificando-se os diferenciais entre as soluções propostas.

Ao final, as considerações finais sugerem uma solução inovadora e diferenciada, única no contexto da informática na saúde pública nacional, a ser apresentada no restante do trabalho.

3.1 A Informática na Saúde Pública

As soluções de *Business Intelligence* envolvendo a área de saúde geralmente envolvem hospitais particulares e empresas envolvidas com planos e seguros de saúde de instituições privadas. Em ambos os casos apresentam-se preocupações semelhantes em relação à identificação e controle de gastos visando uma melhor gestão dos recursos financeiros.

Em relação à saúde pública, a Organização Mundial de Saúde (OMS) [oms09], publica anualmente um relatório sobre temas diversos. Na primeira e, até o momento, única comparação realizada entre os atendimentos públicos dos países, no ano de 2000, o relatório técnico disponibilizado examina e compara aspectos do sistema de saúde em todo o mundo, visando identificar situações e disponibilizar melhorias considerando-se os recursos disponibilizados [whr00].

No relatório, a OMS utilizou-se de cinco critérios para compor o índice de desempenho dos sistemas de saúde:

- Nível global de saúde da população;
- Desigualdades da saúde da população;
- Eficiência do sistema de saúde;
- Diferenças dessa eficiência de acordo com as classes sociais;

- Diferenças de gasto com saúde entre a população.

Na publicação, utilizando-se dos critérios citados, o Brasil encontrava-se na posição de número 125, dentre os 191 avaliados, quando o tema é desempenho do sistema de saúde. A França aparece em primeiro lugar, seguido da Itália e outros pequenos países como San Marino, Andorra e Malta.

Possivelmente esta avaliação encontra-se relacionada, dentre outros fatores, a uma gestão deficiente do Sistema Único de Saúde. Neste aspecto, uma melhoria na gestão em saúde pública passa obrigatoriamente por uma melhor análise e diagnóstico da efetividade dos recursos aplicados nos atendimentos das emergências dos hospitais públicos.

Devido às características peculiares do SUS, principalmente os dados relacionados aos custos envolvidos, a comparação com soluções de *Business Intelligence* de outros países encontra-se restrita a relatórios elaborados por fontes governamentais. Dentre os países considerados com as maiores taxas de desenvolvimento humano, e colocado na décima posição no relatório da Organização Mundial de Saúde com relação ao atendimento público e gratuito, destacam-se os trabalhos realizados no Japão.

Nas páginas da *Japan Public Health Association* [jpa09a], apresentam-se variados dados estatísticos correspondentes aos anos de 2007 e 2008. No relatório referente ao ano de 2008 [jpa09b] ressalta-se a nova política de privatização e desregulamentação dos serviços públicos.

Contudo, apesar da riqueza de informações e gráficos detalhando-se inclusive políticas públicas, ambientais, farmacêutica, faixas etárias de atendimento, epidemiologia, dentre outro, analisando-se os gráficos e planilhas apresentados no documento, não foram identificados projetos com o enfoque nos atendimentos das emergências dos hospitais públicos.

As peculiaridades do Sistema Único de Saúde do Brasil dificultam sobremaneira a tarefa de identificação de trabalhos relacionados no exterior, o que pode ser interpretado como ainda mais necessária uma solução de *Business Intelligence* completamente direcionada a nossa realidade política, social e econômica.

Nos trabalhos identificados no Brasil, a principal fonte de informação utilizada pelas soluções referem-se ao dados fornecidos pelo Ministério da Saúde através do Departamento de Informática do Sistema Único de Saúde (SUS), o DATASUS [dat08]. Estes dados são coletados nas Secretarias Estaduais e Municipais através da disponibilização gratuita do software, com a obrigatoriedade de entrega periódica de informações epidemiológicas, de nascimentos, atendimentos básicos de saúde, ambulatoriais e internações, dentre outros.

Por outro lado, este trabalho propõe uma solução de *Business Intelligence* baseado em *Data Warehouse* que forneça uma melhor gestão dos atendimentos das emergências, identificado como um fator primordial para a melhoria da saúde como um todo.

A seguir apresenta-se uma análise de alguns projetos voltados à saúde pública desenvolvidos no Brasil onde realiza-se, ao final de cada um dos projetos, uma análise crítica de sua aplicação e tecnologia envolvidas.

3.2 Projeto de um *Data Warehouse* para a Saúde Pública

No trabalho de Ricardo S. Santos e outros [SGFT08] é apresentado um projeto de *Data Warehouse* para a gestão da saúde pública atendendo à Secretaria de Estado de Saúde de São Paulo (SES-SP), o maior estado do país, com disponibilização de informações gerenciais utilizando fontes de dados de diversos sistemas isolados. Uma das metas da SES-SP é a construção de uma rede estadual de informações visando garantir acesso à informação a todos os gestores do SUS e aos cidadãos.

Como fonte de dados, o projeto utiliza as bases de dados do Departamento de Informática do SUS [dat08], DATASUS, centralizando, integrando e disponibilizando as informações provenientes de diversos sistemas do DATASUS, além de algumas planilhas e documentos da SES-SP, conforme mostra a Figura 3.1. Os dados contemplados pelo projeto incluem unicamente os dados do Sistema de Informações Ambulatoriais (SIA) [sia09].

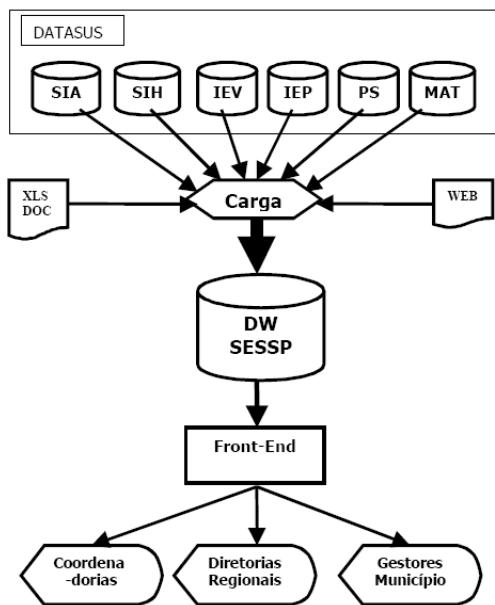


Figura 3.1 Escopo Global do Projeto. Retirado de [SGFT08].

Como metodologia de desenvolvimento estabelecida utilizou-se como base o estudo de Inmon [Inm02] que sugere definir inicialmente o modelo do DW e depois estudar os sistemas

fontes. Como as fontes pertencem a outra instituição, os modelo passa a ser projetado em função dos dados disponibilizados pelo DATASUS.

O desenvolvimento da solução sugere a criação de um banco de dados relacional além do dimensional, elaborado no esquema estrela, com o objetivo de manter os dados e fontes em um ambiente controlado, evitando-se dependência de outros sistemas quanto a disponibilidade das fontes de dados.

Os critérios na escolha das ferramentas utilizadas no desenvolvimento do projeto envolveram: robustez para suportar o volume de dados, existência de casos de sucesso implementados e custo acessível. Com base nestes critérios, adotou-se um conjunto de ferramentas da Oracle [ora09], a saber:

- Armazenamento do DW: Oracle 9i;
- Modelagem de Dados: *Oracle Designer*;
- Extração, Transformação e Carga (ETL): *Oracle Warehouse Builder*;
- Metadados e Ferramenta de OLAP: *Oracle Discoverer*.

No processo de carga, devido à grande quantidade e variedade de formato de arquivos, desenvolveu-se uma ferramenta para automação, em parceria com uma empresa privada, para realizar a busca na página da internet do DATASUS, *download*, descompactação, transformação e carga na base de dados relacional. A ferramenta também permite auditoria da tarefa realizada. As cargas no *Data Warehouse* são executadas por mapeamentos implementados na ferramenta *Oracle Warehouse Builder*.

Com referência à exibição dos dados, há uma mapeamento das descrições dos dados e das regras de negócio no metadados e desenvolvidos relatórios e consultas pré-definidos para atender os principais requisitos dos usuários, utilizando-se a ferramenta *Oracle Discoverer*.

O volume de dados considerado, nesta primeira fase relatada, é de cerca de 1.800.000 registros, correspondendo a 211 Mb, com volume anual correspondendo a 2,9 Gb com testes realizados em um servidor x-series 200 da IBM [ibm09] e duração de todo o processo de ETL de aproximadamente duas horas.

Quando da publicação do trabalho relacionado referente ao projeto no XI Congresso Brasileiro de Informática em Saúde promovido pela Sociedade Brasileira de Informática em Saúde [sbi08], em dezembro de 2008, não havia uma mensuração dos benefícios obtidos e satisfação dos usuários.

3.2.1 Análise Crítica

Como avaliação positiva, o projeto caracteriza-se por envolver uma grande integração de dados em formatos distintos, o que acarreta uma grande tarefa de integração e representa uma unificação de vários ambientes possibilitando uma visão simplificada para os usuários. O envolvimento da SES-SP certamente valida a necessidade e importância da solução.

A elaboração de uma solução específica de extração, mesmo que realizada por outra empresa envolvida no projeto, permite uma maior automação do processo de ETL, fator diferencial a ser almejado pelas soluções de BI.

A utilização de metadados disponibilizado para os usuários facilita a independência na elaboração de relatórios e minimiza a dificuldade de identificação da representação de cada uma das colunas das tabelas do modelo dimensional, uma vez que utiliza-se a linguagem do negócio, bastante familiar aos usuários envolvidos na solução.

A solução também apresenta uma performance satisfatória com relação ao tempo do processo de ETL e o volume de dados envolvidos.

Por outro lado, o projeto não analisa os dados oriundos especificamente das emergências dos hospitais públicos que representam um dos maiores custos associados à saúde pública.

Outro fator a ser analisado refere-se a pouca amplitude no critério de seleção das ferramentas, que não são citados no trabalho, e resulta em um custo elevado de aquisição e manutenção em uma solução proprietária de *Business Intelligence*.

A elaboração de uma solução envolvendo o processo automatizado de ETL indica que, a partir da necessidade de qualquer manutenção devido a alteração no formato e/ou disponibilização dos dados, que já encontra-se sob controle do DATASUS, ou seja, fora do controle da organização, acarretaria em necessidade de um novo envolvimento da empresa terceirizada, elevando-se, assim, o custo da solução.

A características de elaboração de análises pré-definidas limita a alternativa de uma livre simulação na busca de correlação entre as informações, criando uma dependência maior dos usuários com os desenvolvedores da solução, uma vez que qualquer nova análise representa uma nova manutenção com novos prazos para implementação.

Outra restrição identificada é a ausência de disponibilização da solução em um navegador na internet, limitando a portabilidade da aplicação tanto com relação à instalação quanto aos trabalhos associados à distribuição de novas versões.

Finalmente, a ausência de mensuração dos benefícios avaliados pelos usuários, apesar de uma forte indicação de aceitação apresentado pelos desenvolvedores da solução, torna, ainda assim, incerta a satisfação das expectativas, a utilização efetiva da solução e o retorno dos

investimentos realizados.

3.3 A Aplicação do BI no Segmento de Saúde Pública Ambulatorial

Valente [VA08] descreve o processo de desenvolvimento, a arquitetura e tecnologias utilizadas no projeto de *Business Intelligence* para apoio a gestão ambulatorial, tendo como fonte de dados o SIGA Saúde [ate09], um sistema de informação integrado de Gestão e Assistência em Saúde Pública implantado e em funcionamento nas Secretarias Municipais de Saúde de São Paulo e Diadema.

A arquitetura adotada no projeto precede a elaboração de um *Data Warehouse* composta pela fonte de dados dos sistemas transacionais, uma área de transformação e carga de dados, a *Data Staging Area*, a base de dados do *Data Warehouse* e a camada de apresentação.

À época, o processo de ETL ainda não automatizado e utilizando rotinas na linguagem PL/SQL da Oracle [ora09], a carga de um mês do sistema transacional demanda de cerca de dois dias para a execução de todo o processo.

Neste trabalho utilizou-se um processo de desenvolvimento próprio, seguindo as seguintes fases: Levantamento Preliminar, Planejamento, Análise do Negócio, Projeto, Construção, Testes, Homologação e Implantação.

No Levantamento Preliminar, realizado em reuniões com os usuários, identificaram-se os relatórios gerenciais e as áreas de interesse. A partir do protótipo de cada relatório gerencial especificado e aprovado, identificaram-se as áreas de interesse do projeto de BI. Um primeiro grupo, considerado básico envolve: Estabelecimentos, Profissionais, Usuários SUS e Procedimentos. O segundo grupo, considerado complexo envolve: Equipes de Saúde, Atividades Coletivas, Vacinação, Vagas, Agendamentos, Fila de Espera, Atendimentos e Regulação.

Na fase de Análise de Negócio elaborou-se um protótipo na ferramenta OLAP, no Projeto, o modelo de dados dimensional.

A ferramenta OLAP utilizada no trabalho apresentado, denominada Infomax, é de propriedade da Atech, customizada para atender a solução em saúde pública. Oferece recursos de navegação e publicação de relatórios, onde encontram-se registrados os metadados.

Um Sistema de Informação Geográfica (SIG) vem sendo implementado com os mapas geográficos do município integrado à ferramenta, permitindo a visualização colorida da distribuição de quantidades de agendamentos por divisões operacionais e/ou administrativas do município.

3.3.1 Análise Crítica

Semelhante ao trabalho anterior, a proposta envolve a integração de grande volume de dados, possuindo a característica de disponibilizar aos envolvidos a facilidade de um único ambiente para visualização da solução. Também como fator positivo a ser considerado é a utilização de um forte direcionamento à aplicação de uma metodologia, tendo como resultado final a elaboração de um modelo de dados dimensional.

A utilização de metadados facilita para os usuários a elaboração de relatórios, minimizando a dificuldade de identificação da representação de cada uma das colunas das tabelas do modelo dimensional, utilizando-se a linguagem do negócio, familiar aos usuários envolvidos na solução.

A proposta de elaboração de um Sistema de Informações Geográficas demonstra uma tendência pioneira na utilização desta nova tecnologia a ser implementada, que estão disponibilizadas nas melhores ferramentas de BI, e geralmente associadas à utilização de *dashboards*. A possibilidade de visualização em um navegador na internet representa o importante diferencial em portabilidade da solução.

No entanto, observa-se que o processo de desenvolvimento da solução não envolve uma metodologia com aplicação de melhores práticas, estipulados pela literatura acadêmica de referência na elaboração de um *Data Warehouse*, como Inmon [Inm02] ou Kimball [KC04].

Percebe-se também uma ausência de critérios na seleção das ferramentas associadas a solução de BI, visto que a proposta apresenta-se mais direcionada a uma solução de um modelo de dados dimensional em um ambiente de *Data Warehouse* sem um direcionamento a uma solução completa envolvendo as ferramentas para a implementação deste modelo sugerido.

Também não há sugestão de automação do processo de ETL, meta a ser necessariamente atingida, bem como a performance de dois dias de carga precisa ser melhor avaliada. A ferramenta OLAP apresentada, uma solução proprietária, envolve uma customização de outra ferramenta, gerando dependências de manutenção e evolução tecnológica.

Novamente a ausência de mensuração e avaliação dos resultados pelos usuários torna imprevisível a aceitação da solução pelos usuários envolvidos.

Finalmente, a não utilização de dados das emergências poderia melhorar a visão gerencial, possibilitando uma melhor gestão dos recursos públicos através de um ambiente de BI.

3.4 Sistema de Informação Gerencial Utilizando Solução de DW em Software Livre

Em seu trabalho, Grein [GC08] apresenta um sistema de informação gerencial desenvolvido pela Companhia de Informática do Paraná [cel09], utilizando dados fornecidos pela Secretaria Estadual de Saúde do Paraná (SESA-PR) visando estabelecer os controles operacionais e estratégicos necessários para a tomada de decisão pelos gestores.

Os principais requisitos apresentados para o sistema foram os seguintes:

- Documentar as fontes de informação;
- Adquirir ou implementar uma ferramenta de ETL;
- Automatizar as atividades de ETL;
- Processar cálculos dos Indicadores;
- Representar as informações de forma tabular e em gráficos;
- Representar as informações através de dados geoprocessados;
- Projetar um ambiente de Data Warehouse em Software Livre;
- Disponibilizar um ambiente para elaboração relatórios gerenciais.

O trabalho utiliza como base de informações as características referentes aos óbitos ocorridos no Estado do Paraná, com fontes de dados oriundos do Instituto de Medicina Legal do Paraná (IML), relacionando-se os serviços de corpo-delito, mortes violentas que necessitam de investigação, mortes naturais encaminhadas ao IML e informações laboratoriais, como toxicologia, patologia e sexologia.

Os critérios na escolha das ferramentas utilizadas no desenvolvimento do projeto envolveram software livre e adotou-se o seguinte conjunto de ferramentas:

- Armazenamento do DW: Postgres [pos08];
- Modelagem de Dados: DBDesigner [dbd09];
- Extração, Transformação e Carga (ETL), Metadados e Ferramenta OLAP: Pentaho [pen08];

A solução proposta envolve a solução Pentaho, que será detalhada posteriormente quando da comparação entre as ferramentas disponíveis no Capítulo 5.

Encontra-se disponibilizado no trabalho o modelo dimensional e são apresentadas algumas análises, inclusive em forma gráfica. Há uma integração com os dados fornecidos pelo DATASUS com relação a mortalidade do estado.

3.4.1 Análise Crítica

A proposta também envolve a integração de dados, possuindo a característica de disponibilizar aos envolvidos a facilidade de um único ambiente para visualização da solução. A utilização de ferramentas de software livre sinaliza uma diminuição de custos de aquisição e manutenção.

A automação das atividade de ETL, a elaboração de indicadores, utilização de gráficos e o geoprocessamento possibilitam uma maior aceitação por parte dos usuários, integrando os vários ambientes sugeridos por soluções voltadas às mais recentes propostas tecnológicas.

A disponibilização em um ambiente na internet, semelhante à proposta anterior, minimiza as tarefas de instalação e configuração de um ambiente, tornando a solução mais portável.

Em uma análise dos pontos negativos da proposta, encontram-se a ausência de critérios na seleção de ferramentas bem como de uma metodologia na elaboração da solução.

A restrição associada ao sistema tratar unicamente de dados relacionados à mortalidade restringe as decisões gerenciais da saúde pública em outros aspectos relevantes associados aos custos.

A falta de avaliação dos resultados obtidos pelos usuários torna questionável sua aplicabilidade à realidade apresentada.

3.5 Análise Comparativa e Considerações Finais

Na busca de pesquisas relacionadas a *Business Intelligence* identificam-se uma larga gama de aplicabilidade teórica e prática de soluções comerciais e acadêmicas.

Com relação à informática na saúde pública, na maioria dos casos no Brasil os trabalhos referem-se a análises realizadas nas áreas sanitárias, epidemiológicas e ambulatoriais, voltadas para a atenção básica à saúde. As fontes de dados largamente majoritária nestes trabalhos encontram-se disponibilizadas pelo Ministério da Saúde, no portal da empresa de informática vinculada ao Ministério, o DATASUS.

Outros trabalhos utilizando-se das mesmas fontes de dados do DATASUS, encontram-se

direcionados às internações hospitalares, produção ambulatorial, imunizações, Programa de Saúde da Família (PSF), acompanhamentos nutricionais, mortalidade e transferências financeiras realizadas entre os Estados e municípios.

Nos trabalhos relacionados detalhados anteriormente, duas das soluções de BI apresentadas utilizam-se de ferramentas proprietárias. No caso da Secretaria Estadual do Estado do Paraná, quando da utilização de ferramentas gratuitas, o objetivo torna-se direcionado ao problema de taxas de mortalidade e suas possíveis causas.

No caso do exterior as especificidades do sistema de saúde pública do Brasil refletem a necessidade de uma análise e tratamento diferenciados das informações adequados à nossa realidade.

O presente trabalho, segue-se à aplicação de uma metodologia indicada na literatura acadêmica abrangendo todas as suas atividades e fases. No início de cada um dos Capítulos há um indicativo de posicionamento das atividades no contexto acadêmico. Como resultado deste trabalho apresenta-se e justifica-se o modelo de dados dimensional, com propostas de trabalhos futuros indicando o direcionamento da evolução deste modelo proposto.

Apresenta-se também uma seleção das ferramentas utilizando-se critérios bem definidos em tabela comparativa de cada uma das soluções em software livre, visando a viabilidade econômica na sua aplicabilidade. Em todo o desenvolvimento das atividades, apresentam-se indicadores de performance das ferramentas.

Assim sendo, a proposta deste trabalho classifica-se como inovadora no âmbito do Sistema Único de Saúde (SUS), ao propor automação do processo de transformação, limpeza e conformidade dos dados, a elaboração de um modelo de dados evolutivo, a seleção da solução adequada e a identificação de alternativas à gestão hospitalar.

Um outro diferencial fundamental desta solução apresentada reside na escolha do tratamento de dados das emergências dos maiores hospitais públicos do Estado, negligenciados nas soluções relacionadas apresentadas.

Como complemento, diferentemente dos outros trabalhos relacionados, há também a sugestão de várias estratégias para a melhoria da gestão da saúde pública, com atenção especial às emergências dos hospitais públicos, inclusive com o diferencial da avaliação realizada pelos gestores envolvidos diretamente neste projeto.

CAPÍTULO 4

A Proposta de um BI para a Saúde Pública

Neste capítulo, justificando algumas análises a serem realizadas no modelo dimensional, apresenta-se a forma de custeio da assistência à saúde realizado pelo SUS para um entendimento sobre os repasses financeiros e suas implicações.

Na seqüência apresenta-se um detalhamento da Secretaria de Saúde do Estado de Pernambuco (SES-PE) mostrando sua estrutura organizacional e um exemplo de Gestão Hospitalar para visualização do processo de atendimento de emergência em uma unidade de saúde com um levantamento da situação da infra-estrutura tecnológica.

Ao final uma breve descrição sobre o Sistema de Atendimento aos Pacientes da Emergência (SAPE) e o seu modelo de dados que servirá para fundamentar o modelo de dados dimensional.

4.1 O Custeio da Assistência à Saúde

Segundo [msa08], a Constituição Federal de 1988 deu um importante passo na garantia do direito à saúde com a criação do Sistema Único de Saúde (SUS). Seus princípios apontam para a democratização nas ações e nos serviços de saúde que deixam de ser restritos e passam a ser universais, da mesma forma, deixam de ser centralizados e passam a nortearem-se pela descentralização.

Em 1990, o Congresso Nacional aprovou a Lei Orgânica da Saúde, que detalha o funcionamento do SUS. Foram mudanças profundas na Saúde Pública brasileira que exigiram, para sua implantação e funcionamento, o aprimoramento do sistema de informação em saúde.

Após esta lei, a missão do Ministério da Saúde passou a ser: Promover a saúde da população mediante a integração e a construção de parcerias com os órgãos federais, as unidades da Federação, os municípios, a iniciativa privada e a sociedade contribuindo para a melhoria da qualidade de vida e para o exercício da cidadania [msa08].

Nesta proposta de universalização do atendimento utilizando-se os vários tipos de financiamento, nem todos os municípios possuem recursos financeiros e materiais para manter um hospital ou serviços especializados. Por este motivo há um repasse financeiro dos recursos de custeio da esfera federal destinados à assistência hospitalar e ambulatorial que configuram

o Teto Financeiro de Assistência (TFA). Seus valores, considerados a fonte de receitas dos estados e municípios, ou seja, seu faturamento, tem sua principal maneira de execução pela Transferência Regular e Automática (Fundo a Fundo).

A Transferência Fundo a Fundo consiste na transferência de valores diretamente do Fundo Nacional de Saúde (FNS) aos fundos estaduais e municipais, independente de convênios ou instrumento congêneres. Cerca de 70% dos recursos do FNS são repassados fundo a fundo para estados e municípios [Lam07].

Nas diretrizes do Pacto pela Saúde de 2006 [msa08], publicado em Portaria do Ministério da Saúde e que contempla o Pacto firmado entre os gestores do SUS, estipularam-se cinco blocos de financiamento, visando proporcionar maior autonomia aos gestores para a alocação de recursos, a saber:

- Atenção Básica;
- Atenção de Média e Alta Complexidade Ambulatorial e Hospitalar;
- Vigilância em Saúde;
- Assistência Farmacêutica;
- Gestão do SUS.

Os atendimentos na emergências tem sua base de financiamento nos recursos repassados para a atenção básica e atenção de média e alta complexidade ambulatorial e hospitalar.

O montante de recursos financeiros destinado exclusivamente ao custeio de procedimentos e ações de atenção básica à saúde compõe o Piso de Atenção Básica Ampliado (PAB) que é composto de uma parte fixa destinada à assistência básica e uma parte variável, relativa a incentivos para o desenvolvimento de ações no campo específico da atenção básica.

Os valores referentes ao PAB serão transferidos aos municípios de forma regular e automática, do Fundo Nacional de Saúde ao Fundo Municipal de Saúde. Aos municípios não habilitados, a grande minoria, cerca de 1% no Brasil, os valores são transferidos ao Fundo Estadual de Saúde que devem definir critérios para o posterior repasse dos valores.

O PAB-Variável, por sua vez, remunera direta e automaticamente, estados e municípios habilitados à execução de ações previstas em programas e incentivos do Ministério da Saúde, de acordo com regulamentação específica. Contemplam as ações básicas de vigilância sanitária, assistência farmacêutica básica, programa de agentes comunitários de saúde e programa de saúde da família (PSF).

Dentre estes recursos há uma compensação devido as especificidades regionais que corresponde a 5% do valor mínimo do PAB fixo, multiplicado pela população do Estado, segundo alguns critérios pré-estabelecidos. Esta compensação caracteriza um repasse de parte dos recursos dos municípios ao Estado para atendimento de clínicas especializadas que, por algum motivo avaliado pela gestão como por exemplo custo do equipamento ou pessoal capacitado, não foram implementadas no município.

Importante ressaltar que cabe aos municípios a realização dos atendimentos relacionados ao PAB, onde devem ser realizadas estas ações de saúde. Os atendimentos de assistência básica realizados em unidades de saúde fora do domicílio do paciente caracteriza-se como gastos de obrigação de um município realizados em outro município, que necessitam de gestão do Estado sobre estas ocorrências.

4.2 A Secretaria Estadual de Saúde: Gestão, Custeio e Informatização

A rede de saúde do Estado de Pernambuco é composta pela Secretaria Estadual de Saúde (SES-PE) que possui diferentes unidades de saúde subordinadas distribuídas em todo o Estado. As unidades, de uma maneira geral, são geridas pelas onze Gerências Regionais de Saúde (GERES) que têm como função fornecer um apoio às unidades de saúde e à população local, estando geograficamente distribuídas nas várias regiões político-administrativas do Estado, conforme apresentado na Figura 4.1.

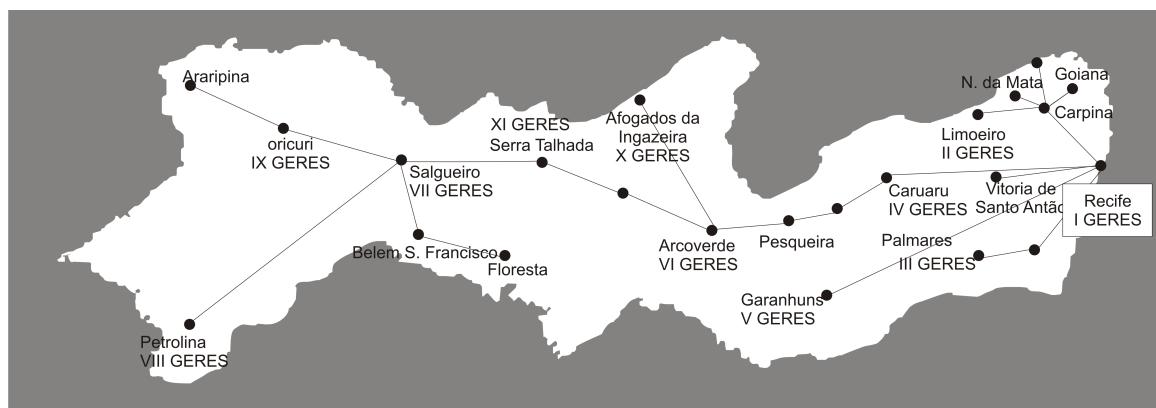


Figura 4.1 Distribuição geográfica das GERES.

A SES-PE é um órgão da administração direta, integrante do Sistema de Execução do Poder Executivo, possuindo entre seus objetivos institucionais:

- executar a política sanitária do Estado;

- promover e superintender as ações que vise o atendimento integral e equânime das necessidades de saúde de toda a população;
- orientar e controlar as ações que visem ao atendimento integral e equânime de saúde de toda a população;
- acompanhar e fortalecer o Sistema Estadual de Saúde.

Apesar da proposta, no final de 2008, de uma nova estrutura organizacional, a ser aprovado e implantado em 2009, durante todo este projeto de pesquisa considera-se a SES-PE organizada conforme uma estrutura organizacional onde a Diretoria de Planejamento em Saúde encontra-se diretamente vinculada às Unidades Hospitalares visando o controle, regulação e acompanhamento do funcionamento de cada unidade.

Subordinada à Diretoria de Planejamento, localizam-se os Hospitais Públicos do Estado, em número de vinte e dois, com onze localizados na região metropolitana e os outros onze no interior.

A rede estadual de saúde consta de hospitais de grande, médio e pequeno porte com o objetivo de atendimento à população. Na região metropolitana seis realizam atendimento de emergência e, entre eles, os quatro maiores possuem implantado e em funcionamento um sistema informatizado que realiza o acompanhamento dos atendimentos, a saber:

1. Hospital da Restauração (HR), o maior hospital público do Nordeste, localizado próximo ao centro da cidade do Recife no corredor rodoviário que interliga o Recife a Olinda e Jaboatão dos Guararapes, duas das maiores cidades da Região Metropolitana. Oferece os serviços especializados na emergência em Neurologia, Neurocirurgia, Clínica Médica (hemorragias digestivas, intoxicações e envenenamentos), Clínica Pediátrica, Traumatologia, Cirurgia Geral, Cirurgia Vascular, Buco-Maxilo-Facial, além de ser referência no atendimento ao grande queimado;
2. Hospital Agamenon Magalhães (HAM) que atende aos bairros da região norte do Recife com serviços oferecidos na emergência de Clínica Obstétrica, Clínica Médica, Odontológica, Urgência Cardiológica Pequenas Cirurgias e Otorrino;
3. Hospital Otávio de Freitas (HOF) localizado na região oeste da Região Metropolitana com serviços oferecidos na emergência de Clínica Médica, Clínica Pediátrica, Cirurgia Geral, UTI, Psiquiatria e Traumatologia;

4. Hospital Getúlio Vargas (HGV) localizado na região leste da Região Metropolitana com serviços oferecidos na Emergência de Clínica Médica, Cirurgia Geral, Neurologia, Buco-Maxilo-Facial, UTI, Traumato-Ortopedia, Politraumatismo e Cirurgia Vascular.

Os outros hospitais de grande porte, mas que não possuem o sistema informatizado de atendimento da emergência ou que encontram-se em fase de projeto de implantação, são o Hospital Barão de Lucena e o Hospital Regional do Agreste em Caruaru.

Na Figura 4.2 apresenta-se a estrutura da Diretoria de Planejamento onde encontram-se subordinados os Hospitais Públicos.

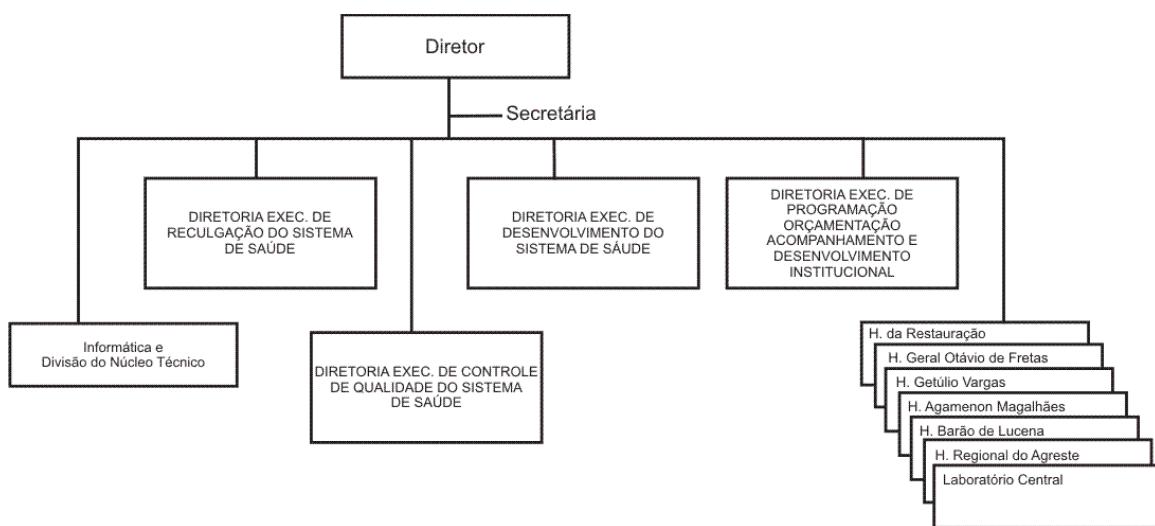


Figura 4.2 Estrutura da Diretoria de Planejamento da SES-PE.

Dentro da estruturação hierárquica, as unidades de saúde precisam de autorização e dados da SES-PE para tomar decisões relativas aos seus servidores, acompanhar o seu faturamento e recebimento e, principalmente, acompanhar e melhorar o sistema de saúde.

Por sua vez, a SES-PE baseia-se nas ocorrências realizadas nas unidades de saúde, como notificação de doenças, necessidade de leitos, óbitos, altas, transferências, internamentos, entre outras, para apoiar as decisões sobre investimentos e gestão no sentido de melhoria no atendimento à população e políticas públicas. Cabe também a SES-PE a responsabilidade pela prestação de informações ao Ministério da Saúde e intermediação de captação de recursos financeiros.

Cabe a cada unidade descrever como serão utilizados os recursos financeiros, produtividade dos servidores e produtividade das clínicas prestadoras de serviço, assim como licitações e controle do patrimônio, para apreciação da SES-PE.

Desta forma, cada uma das unidades de saúde possuem uma certa autonomia na sua gestão. Segue-se uma visão geral sobre o gerenciamento de uma unidade de saúde para um melhor entendimento da necessidade de informatização dos processos, permitindo uma contextualização do papel da emergência em uma gestão pública de saúde.

4.2.1 A Gestão Hospitalar

O funcionamento de um hospital público possui uma série de processos que envolvem a implantação de diversos sistemas que necessitam de integração para a geração de informações consolidadas de interesse da SES-PE. Torna-se necessário um entendimento dos processos na para detalhamento posterior da localização da emergência neste amplo contexto.

Em linhas gerais, o fluxo genérico do processo de gestão de uma unidade hospitalar de uma unidade de saúde envolve diversos setores como recepção, emergência e internamento conforme apresentado na Figura 4.3.

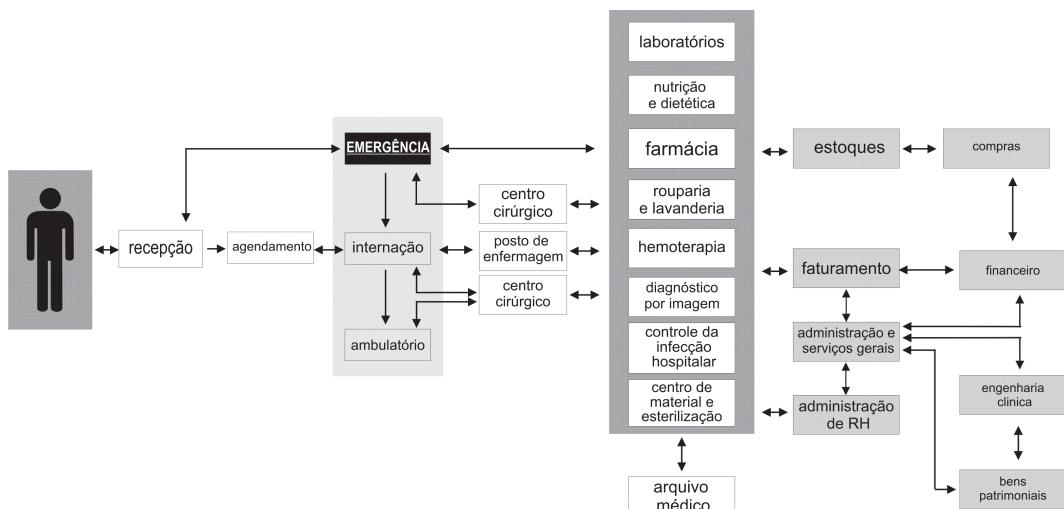


Figura 4.3 Fluxo do Processo de Gestão Hospitalar. Retirada de [ses06].

Contudo, devido à grande quantidade e diversidade de características das diferentes unidades de saúde da rede SUS no Estado de Pernambuco, deve-se considerar a possibilidade de adaptação do fluxo de processo apresentado.

Segundo [ses06] podemos considerar um sistema de gestão hospitalar baseado nos seguintes processos:

- gerenciamento do apoio técnico assistencial: relacionado aos serviços hospitalares prestados aos pacientes através de procedimentos médicos e de enfermagem. Inclui funcional-

lidades que estarão presentes desde o controle da admissão do paciente, registro clínicos até as solicitações realizadas para o paciente e anotações médicas e/ou de enfermagem;

- gerenciamento administrativo e financeiro e recursos humanos: relacionado com as atividades operacionais e administrativas de uma unidade de saúde. Envolve os setores que interagem entre si neste processo como compras, contabilidade, financeiro, patrimônio, custos e faturamento.
- gerenciamento da atenção à saúde: por envolver o atendimento na emergência, segue um detalhamento a seguir.

4.2.1.1 Gerenciamento da Atenção à Saúde

O gerenciamento da atenção à saúde está relacionado com as atividades e procedimentos aplicados ao paciente, estando em atendimento na emergência, internado com tratamento nas diversas clínicas médicas da unidade de saúde ou no ambulatório.

Pode-se agrupar as atividades e procedimentos nas seguintes gestões:

- gerenciamento de pacientes externo na emergência: está relacionado as ações empregadas na emergência, para recuperação de pacientes cujos agravos de saúde necessitam de assistência imediata por apresentarem risco de morte;
- gerenciamento dos pacientes no internados: está relacionado ao conjunto de ações empregadas na recuperação do paciente envolvendo diferentes clínicas disponibilizadas pela unidade hospitalar;
- gerenciamento de pacientes externos no ambulatório: quando do atendimento após 24 horas da ocorrência médica.

Este tipo de gerenciamento está diretamente relacionado com a solução de BI proposta neste trabalho, visto que servirá de auxílio ao gerenciamento de atenção à saúde.

Cada processo envolvido na Gestão Hospitalar pode representar um sistema a ser desenvolvido pela área responsável pela Tecnologia da Informação com o objetivo de melhoria na gestão da unidade de saúde. A emergência e o sistema que realiza o acompanhamento do atendimento do paciente, o SAPE, que encontra-se descrito neste capítulo, representa somente a informatização de um dos processos, o que demonstra-se fortemente a necessidade de integração com as partes restantes da grande estrutura organizacional de uma gestão hospitalar.

4.2.2 Conjuntura Atual da Informática na SES-PE

A SES-PE conta com um Comitê de Informática que busca estabelecer metas e objetivos para o processo de automação da rede hospitalar e da própria Secretaria, que possui as seguintes metas:

- democratizar as informações;
- descentralizar os serviços da SES-PE pelos hospitais com o objetivo de reduzir o processo operacional, tornando os hospitais mais autônomos, com potencial para gestão própria, porém sob a subordinação à GERES e SES-PE.
- melhorar a comunicação entre as diferentes Unidades de Saúde pública;
- avaliar e propor informatização da rede hospitalar do Estado.

Muitas das informações hoje centralizadas na SES-PE e nas unidades de saúde são de interesse para a população e diferentes órgãos da administração direta e indireta. Informações sobre epidemias, unidades de referência em determinado atendimento, campanhas de saúde, ações em situação de emergência de saúde pública e fornecimento de medicamentos envolvem diretamente o processo de gestão pública da saúde.

Existem diversos sistemas desenvolvidos e distribuídos por diversos fornecedores, inclusive o DATASUS, Departamento de Informática do SUS, órgão da Secretaria Executiva do Ministério da Saúde, que visam atender à integração entre as unidades hospitalares e a Secretaria de Saúde.

O acesso aos sistemas disponibilizados nos hospitais permitem interligações permitindo a troca de informações sobre:

- controle de recursos humanos;
- controle de epidemiologias;
- vigilância sanitária;
- controle de medicamentos;
- controle de autorizações de internamento e leitos;
- orientação, supervisão e fiscalização;
- regulação de leitos de gestantes;

- orçamentos e licitações.

Com o objetivo de exercer um controle da rede hospitalar, a SES-PE mantém um cadastro das unidades hospitalares com suas características e capacidades. Através de fiscalização sistemática e contato com as unidades mantém atualizadas as informações sobre tetos financeiros possibilitando relatos de pareceres sobre cada unidade de saúde.

Para a comunicação entre as unidades de saúde e destas com a SES-PE, existe um contrato de prestação de serviços técnicos especializados com a PE Multidigital para a implantação, operacionalização e manutenção de uma solução integrada de telecomunicações e informática, que permite a constituição de uma rede abrangendo todo o estado de Pernambuco. Esta rede permite o acesso a sistemas de informação da gestão pública, acesso à internet e conexão da Secretaria aos sistemas e bases de dados dos hospitais públicos interligados [ped08].

Na seqüência apresenta-se o sistema que realiza o atendimento aos pacientes nas emergências das unidades de saúde.

4.3 O Sistema de Atendimento aos Pacientes da Emergência (SAPE)

O SAPE, desenvolvido pela Empresa de Fomento de Informática de Pernambuco (FISEPE), empresa pública do Estado de Pernambuco, atual Agência de Tecnologia de Pernambuco (ATI), teve sua implantação inicial realizada no Hospital da Restauração no ano de 2001. Desenvolvido em linguagem Delphi, numa arquitetura tipicamente Cliente/Servidor utiliza o Sistema Gerenciador de Banco de Dados SqlServer 2000, da Microsoft. A Figura 4.4 apresenta a tela inicial do sistema implantado nas emergências dos quatro maiores hospitais do Estado de Pernambuco.

O sistema possui a finalidade básica realizar o registro de entrada dos pacientes nas emergências dos hospitais públicos. Ao solicitar atendimento em uma emergência, são cadastrados os dados do paciente, a ocorrência do fator gerador da emergência e a posterior baixa do atendimento, associada a um motivo informado pelos médicos.

Na recepção dos hospitais existem microcomputadores e impressoras onde cadastrase o paciente quando da primeira vez no hospital, ou, se necessário, atualizadas as informações no caso de atendimentos anteriores. Posteriormente é emitida uma ficha de atendimento que segue com o paciente, ou acompanhante, sendo preenchida pelo(s) médico(s) durante o atendimento.

Quando da liberação do paciente ou qualquer movimentação dentro do hospital, ou entre hospitais, novos dados da ficha de atendimento são digitados no sistema, informando a situação do atendimento do paciente.

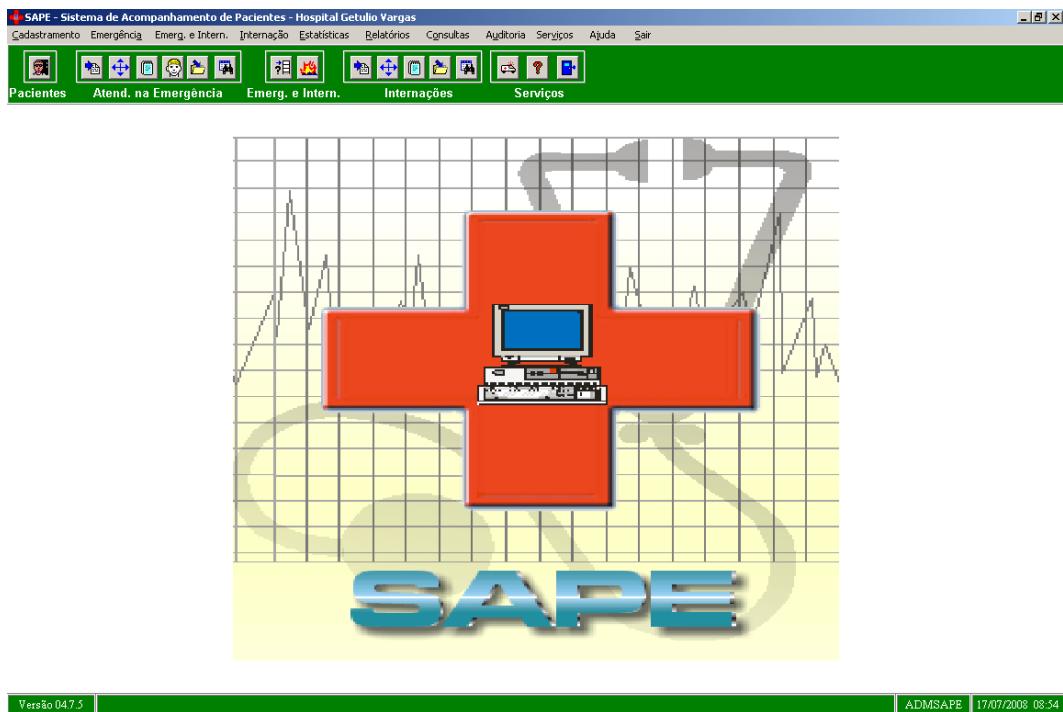


Figura 4.4 Tela Inicial do Sistema de Atendimento aos Pacientes da Emergência (SAPE).

As fichas de atendimento, apesar de não serem consideradas prontuários, servem de base para uma análise das ocorrências de crimes contra menores, mulheres e cidadãos em geral bem como acidentes de trânsito e outros motivos classificados em natureza da ocorrência. Cabe a Secretaria de Saúde a unificação compartilhada dos domínios de nomes das tabelas comuns a todos os hospitais da rede pública estadual onde o sistema encontra-se implantado.

O SAPE, inicialmente desenvolvido para atendimento às emergência, teve acrescido à sua funcionalidade inicial, durante o seu desenvolvimento, os seguintes módulos integrados à mesma base de dados:

- SAPE Ambulatório: responsável por gerenciar o agendamento das consultas realizadas no ambulatório das unidades de saúde;
- SAPE Internamento: realiza o gerenciamento de internamento dos pacientes oriundo das emergências, transferidos de outros hospitais ou de atendimentos previamente agendados (eletivas);
- SAPE Farmácia: módulo responsável pelo controle da prestação da assistência farmacêutica da unidade de saúde;

- SAPE Faturamento: controla a conta hospitalar através da integração de laudos de autorização de realização de procedimentos, internações, procedimentos de alto custo, controle das cotas da unidade de saúde.

Nem sempre todos os módulos são implementados em um hospital. Em face a realidade administrativa e financeira de cada unidade de saúde, a SES-PE determina, seguindo recomendações da equipe de Tecnologia da Informação, quais dos módulos serão implantados, negociando os prazos com os envolvidos. Por exemplo, no Hospital Geral de Areias (HGA) localizado no bairro da Estância e no Hospital Barão de Lucena (HBL) no bairro da Iputinga, ambos na região metropolitana somente encontram-se implantados os módulos de Ambulatório e Faturamento.

Por determinação do Ministério da Saúde o atendimento a ser considerado emergencial não deve exceder 24 horas. Ultrapassando este prazo, passa-se a ser considerada uma internação, sendo considerado encerrado o respectivo atendimento na emergência com o gerenciamento passando para o módulo de internamento. Uma rotina específica do sistema realiza a baixa dos atendimentos que ultrapassam as 24 horas.

O modelo de entidade e relacionamentos referente ao SAPE, módulo emergência, encontra-se no Apêndice A.

Dentre as funcionalidades do SAPE Emergência as principais são:

- realizar o cadastramento dos dados do paciente, como identificação, endereço, filiação, sexo e idade;
- registrar o tipo de ocorrência, ou seja, o fator gerador do atendimento;
- registrar a procedência do paciente;
- gerenciar a movimentação e localização do paciente na unidade de saúde;
- registrar o destino do paciente após o atendimento;
- identificar acidentes de trabalho;
- registrar atendimentos que gerem ocorrências policiais;
- registrar os motivos das baixas dos atendimentos, inclusive óbitos;
- identificar as transferências realizadas entre as unidades de saúde;
- permitir acesso à emergência de pacientes não identificados, ou seja, pacientes que devendo às condições clínicas encontrava-se impossibilitado de identificar-se;

- disponibilizar o acompanhamento do atendimento ao Serviço Social da unidade de saúde.

Apesar de atender ao que se propõe enquanto instrumento de operacionalização da emergência de uma unidade de saúde, o SAPE apresenta algumas limitações que identificam-se a seguir e servem de base para a proposta de uma outra solução de tecnologia da informação.

O SAPE Emergência mesmo atendendo satisfatoriamente as ocorrências nas unidade de saúde carece de ferramentas para o suporte à tomada de decisão necessárias para os gestores da área de saúde, inclusive de uma visão mais orientada aos recursos aplicados.

Com a transferência de recursos financeiro em sua parte fixa, os municípios recebem o PAB para a manutenção de vários serviços, entre eles os atendimentos das emergências. Pacientes enviados de um município para outro, o que ocorre principalmente na região metropolitana, geram gastos no atendimento do município destino, nem sempre havendo o repasse dos valores recebidos pelo Fundo Nacional de Saúde aos município de origem. Esta identificação é considerada de grande importância para o Estado.

Devido aos dados do SAPE não estarem integrados entre os hospitais, e também por não tratar-se de uma ferramenta de gestão estadual, surgem problemas como um melhor gerenciamento das transferências de pacientes entre as unidades de saúde.

O cadastramento de um mesmo paciente que tenha atendimento em várias das unidades de saúde tendem a causar uma distorção em estatísticas consolidadas.

A base de dados com informações diversas sobre o atendimento não apresenta-se de uma maneira que auxilie uma visão estratégica no direcionamento de ações de gestão pública, o que leva a proposta de um complemento ao sistema transacional.

4.3.1 Uma proposta de *Data Warehouse* para a SES-PE

Face aos problemas identificados e como oportunidade de melhoria, surge a proposta de uma solução utilizando-se de dados consolidados em um ambiente em que se permita simulações e análises com cruzamento das informações geradas nos atendimentos, caracterizando-se como uma alternativa de auxílio à gestão do processo de atendimento aos pacientes nas emergências garantindo, principalmente, o direito de atendimento dos pacientes.

Uma vez apresentado o modelo de dados, as funcionalidades existentes e com o problema identificado, novos requisitos são avaliados como um desafio à elaboração de uma solução de gestão, descritos a seguir.

4.3.2 Análise dos Requisitos

Seguindo as Fases de Desenvolvimento de um *Data Warehouse* explicitadas no Capítulo 2, após a fase de viabilização do projeto segue-se a Fase de Definição dos Requisitos de Negócio.

Nesta etapa, segundo Kimball [KR02], há a necessidade de decidir sobre quais processos de negócios serão modelados pela combinação de um entendimento dos requisitos de negócio com um entendimento dos dados disponíveis.

Segundo Sommerville [Som03], as descrições das funções e restrições são os requisitos para o sistema e o processo de descobrir, analisar, documentar e verificar estas restrições é chamado de engenharia de requisitos. Os requisitos funcionais para um sistema descrevem a funcionalidade ou os serviços que se espera que o sistema forneça. Os requisitos funcionais podem ser expressos de diferentes maneiras. Uma tabela de requisitos vem sendo utilizada pela Secretaria da Fazenda do Estado de Pernambuco (SEFAZ-PE) como forma de apresentação, que será utilizada também neste trabalho.

Baseando-se no modelo de entidade e relacionamento do sistema transacional e com o conhecimento das funcionalidades disponibilizadas pelo SAPE, após reuniões realizadas com gestores da SES-PE, surgiram um conjunto de informações gerenciais que caracterizam os novos requisitos funcionais de uma solução de apoio à decisão.

A Tabela 4.1 apresenta a lista de requisitos avaliados pelos gestores.

Requisito	Descrição do Requisito
01	Permitir análise dos Atendimentos por série de tempo: data do registro do paciente.
02	Permitir a análise dos Atendimentos por uma visão de Cidade e Estado de domicílio do Paciente.
03	Permitir a análise dos Atendimentos por Tipo de Ocorrência.
04	Permitir a análise dos Atendimentos por Procedência do Paciente.
05	Permitir a análise dos Atendimentos por Destino do Paciente.
06	Permitir a análise dos Atendimentos por Emergência de Origem.
07	Permitir a análise dos Atendimentos por Acidentes de Trabalho.
08	Permitir a análise dos Atendimentos por Forma de Transporte do Paciente até o Hospital.
09	Permitir a análise dos Atendimentos de Pacientes Não Identificados nas emergências.
10	Permitir a análise dos Atendimentos Baixados associados ao Médico.
11	Permitir a análise dos Atendimentos por Hospitais.

Tabela 4.1 Requisitos necessários para a implementação do DW.

O requisito de análise dos atendimentos por série de tempo especifica em que data, podendo

ser especificadas outras granularidades como ano, mês, dia da semana, semestre, o paciente realizou o atendimento servindo de análises relacionadas com outras informações como data de maior ocorrência de alguma agressão, transferência, internamento, entre outras.

Com relação ao requisito de localização do paciente, baseado no endereço informado consegue-se identificar em qual Estado e município reside o paciente e qual a ocorrência relativa ao seu atendimento.

As ocorrências são consideradas os fatos geradores do atendimento tornando-se importante seu relacionamento com as datas e localização dos domicílios dos pacientes, bem como seu destino após o atendimento.

A identificação e acompanhamento da procedência dos pacientes podem gerar informações sobre as ações a serem realizadas na tentativa de minimizar o tempo de transporte, que identifica-se utilizando o requisito forma de transporte.

O destino do paciente após o atendimento combinado com outros requisitos podem indicar que ocorrências, em que datas e localizações estão sendo direcionados os pacientes visando ações que minimizem o sobrecarregamento de outras unidades de saúde.

Os grandes hospitais públicos possuem várias emergências para a entrada dos pacientes, como emergências traumatólogica, pediátrica e cardiológica, entre outras. Uma associação com os outros requisitos podem indicar a necessidade de alguma especialidade médica associada a demanda das ocorrências ou que tipo de atendimento vem sendo realizado com maior quantidade.

Os acidentes de trabalho também merecem uma atenção especial, visto que há implicações legais para pacientes nesta condição. É de responsabilidade da unidade de saúde a emissão de declarações sobre o atendimento de pacientes vítimas de acidente de trabalho. Visando o atendimento deste direito do paciente, análises consolidadas com datas e tipos de ocorrência podem gerar soluções de gestão com tratamento específico para minimizar estas ocorrências.

O profissional médico que realizou algum procedimento na emergência também necessita ser identificado, permitindo-se auferir os tipos de ocorrências e destino do paciente por profissional, entre outras análises.

Todos os requisitos necessitam da identificação por unidade de saúde. Somente com esta informação torna-se possível identificar como cada ocorrência vem sendo correlacionada com os outros requisitos para possibilitar gestão por hospital através de uma visão geral integrada das emergências.

Após a análise de requisitos segue-se a elaboração do Modelo De Dados Dimensional da solução de BI.

4.3.3 O Modelo de Dados Dimensional

Seguindo a linha central das Fases Desenvolvimento explicitada no Capítulo 2, a fase seguinte consiste da tradução dos requisitos em um modelo dimensional e a transformação deste modelo dimensional em uma estrutura física.

Algumas etapas são realizadas nesta fase após a seleção do processo de negócio [KR02], a saber:

- declarar o grão: que nível de detalhe de dados deve ficar disponível no modelo dimensional;
- escolher as dimensões: uma determinação de granularidade cuidadosa determina a dimensionalidade principal da tabela de fatos;
- identificar os fatos: a interseção entre as dimensões relacionadas.

O modelo dimensional referente ao SAPE apresenta-se na Figura 4.5.

Todas as colunas das tabelas e suas características estão detalhadas no Apêndice A.

Para a elaboração das tabelas de dimensão e fatos os critérios mais relevantes encontram-se detalhados a seguir.

4.3.3.1 Definição das Dimensões

Na elaboração do modelo dimensional a definição das dimensões baseiam-se nos requisitos definidos pelos gestores envolvidos na proposta de solução.

Na maioria das dimensões, algumas colunas da tabela encontram-se desnormalizadas. No ambiente do DW esta situação não é considerada um grande problema pois as tabelas de dimensão não sofrem atualizações freqüentes e são geralmente pequenas em relação às tabelas de fatos. O considerável ganho de performance nas consultas justificam esta alternativa.

Dentre as diversas dimensões de um DW, a dimensão tempo, associada a tabela do modelo denominada Data_Registro, possui um alto nível de granularidade. Esta tabela possui os atributos detalhados da data de registro do atendimentos, permitindo a consulta por dia da semana, mês, ano, dia do ano, semana do ano, semestre, quadrimestre, trimestre e bimestre. Desta forma viabiliza-se uma análise em vários níveis de detalhe da data do atendimento atendendo ao requisito identificado.

Na tentativa de unificação dos valores de tabelas para a consolidação dos bancos de cada hospital na SES-PE, um dos problemas relatados anteriormente, houve, como parte da tarefa

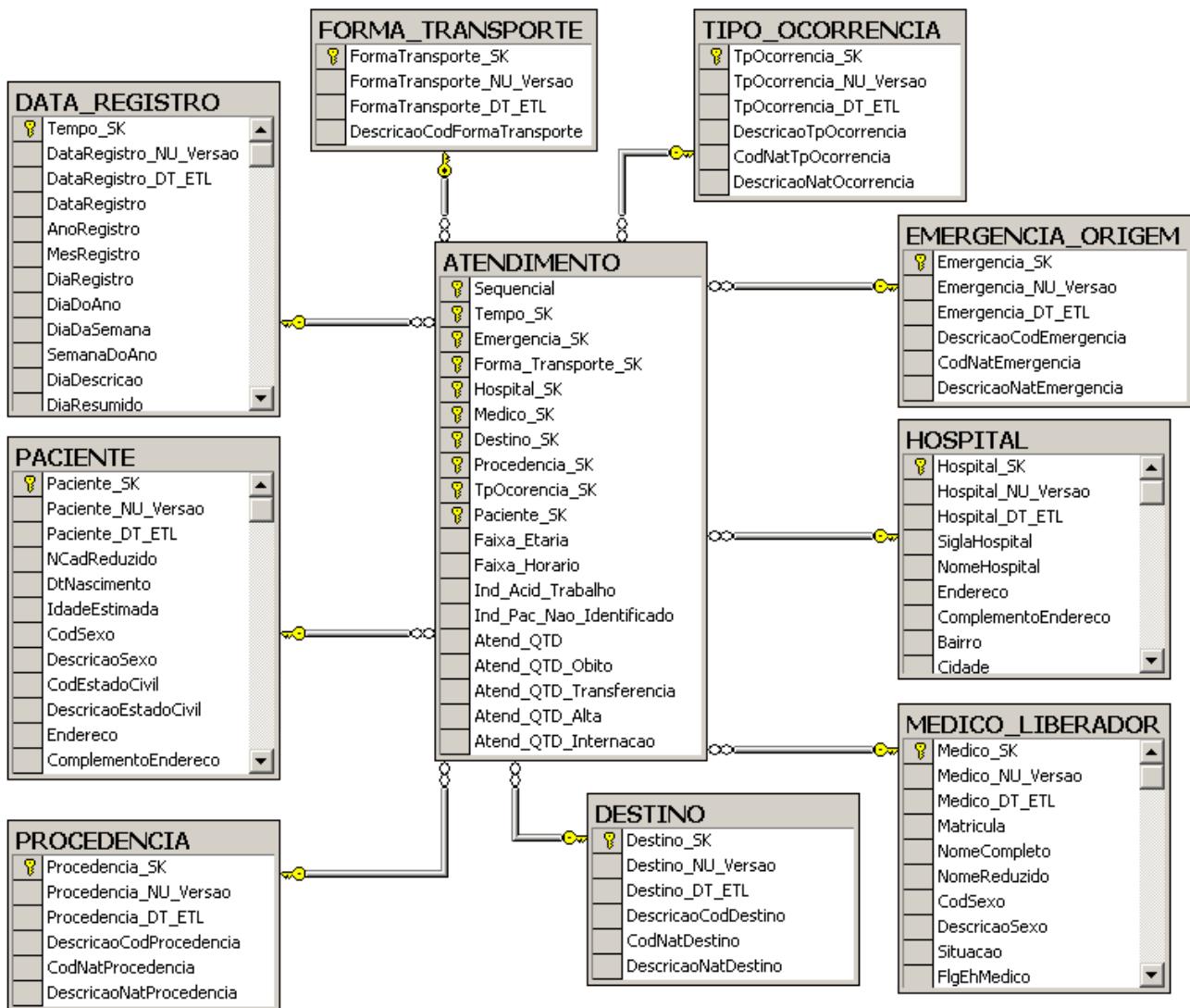


Figura 4.5 Modelo Multidimensional do SAPE Emergência.

deste trabalho, uma unificação em uma única tabela com a proposta que seja liberada pela Secretaria de Saúde, permitindo uma conformidade de domínio.

Desta forma, as tabelas Destino, Procedencia, Emergencia_Origem, Tipo_Ocorrencia e Forma_Transporte possuem domínios unificados, o que permite que a carga seja realizada sem a necessidade de uma tarefa de comparações. Cada uma destas tabelas tiveram seus domínios analisados servindo como referência para a criação de uma nova tabela.

Para citar um exemplo, o código de uma natureza de ocorrência "Agressão"encontra-se com o mesmo código em todos os hospitais. Na coluna Tipo_Ocorrencia "Arma de Fogo", associada a natureza "Agressão"estão cadastradas com um mesmo código em todos os hospitais. Cabe a SES-PE, se houver necessidade e através de acesso remoto implementado da PE Muldigital, inserir novos tipos e natureza de ocorrência nos bancos dos hospitais, visto que, propositalmente, não se permitirão alterações pelos usuários locais.

Na tabela Pacientes por devido ao sigilo das informações em relação ao tipo de ocorrência do atendimento, o nome do paciente não encontra-se disponibilizado, assim como carece de importância gerencial este nível de detalhamento.

Como a base de dados é disponibilizada por hospital, pode ocorrer a situação de um mesmo paciente estar cadastrado em mais de um hospital. No processo de ETL, na fase de limpeza dos dados, realizou-se um tratamento específico na transformação dos dados na tentativa de identificar estas repetições que podem gerar desvios no resultado de uma análise. No Capítulo 4 o processo de eliminação das duplicações será mais detalhado.

Os endereços dos pacientes nem sempre são fornecidos com o Código de Endereçamento Postal (CEP). O SAPE permite uma busca do CEP a partir da informação da Unidade da Federação e Cidade, obrigatórios, e uma descrição, mesmo que parcial, do logradouro. Uma vez informado o logradouro e solicitada uma pesquisa, alguns resultados são sugeridos com a apresentação do logradouro, bairro e, em casos de logradouros com mais de um CEP, o intervalo dos números das residências que compõem o CEP. Cabe ao atendente na recepção identificar o endereço do paciente a partir dos dados fornecidos. Há uma política de conscientização da importância da localização do endereço correto do paciente.

No processo de ETL realiza-se uma busca em uma tabela de CEP para identificar os dados do domicílio do paciente. Cerca de 70% dos paciente nas bases de dados encontram-se com o CEP cadastrado.

Em alguns atendimentos o paciente não encontra-se em condições de fornecer os dados pessoais para o cadastro. Nestas situações a ficha de atendimento é emitida com a geração de um código do paciente, porém como um paciente não identificado. Cabe ao Serviço Social da unidade de saúde, durante, ou após, o atendimento, a identificação do paciente e o seu registro

no SAPE.

Para as dimensões Acidente de Trabalho e Paciente Não Identificado, utilizou-se o conceito de dimensão degenerada. Neste conceito, não há a criação de uma dimensão no modelo dimensional. A indicação de acidente de trabalho é representado por uma coluna na tabela de Fatos com um indicador "S" para o caso de identificação positiva e "N" para a identificação negativa. Isto representa uma alternativa à criação de uma dimensão com unicamente dois registros onde serão necessárias junções de tabelas nas análises. Desta forma, a coluna que seria criada na tabela de fatos e seria a chave estrangeira para uma dimensão torna-se uma coluna específica, identificando a ocorrência de um acidente de trabalho no atendimento. A mesma solução aplica-se ao caso de Pacientes Não Identificados.

No cadastramento de pacientes na emergência, a data de nascimento nem sempre encontra-se digitada pois em algumas situações o próprio paciente e/ou o acompanhante desconhece a informação. O sistema transacional, neste caso, exige a informação de uma idade estimada a ser fornecida pelo atendente.

Nos requisitos elicitados, devido a grande variação de idade dos pacientes e visando facilitar a análise das informações, optou-se pelo enquadramento das faixas etárias apresentadas na Tabela 4.2. Devido a quantidade pequena de faixas e buscando melhoria na performance diminuindo-se as junções, utilizou-se novamente o conceito de Dimensão Degenerada, evitando-se a criação de uma dimensão. As faixas serão disponibilizadas como uma coluna na tabela de fatos sem a correlação em uma dimensão.

Faixa	Descrição da Faixa
01	0 a 5 anos
02	6 a 9 anos
03	10 a 19 anos
04	20 a 29 anos
05	30 a 39 anos
06	40 a 49 anos
07	50 a 60 anos
08	Mais de 60 anos

Tabela 4.2 Faixas Etárias

Utilizando-se novamente dos mesmos critérios e do conceito de Dimensão Degenerada, os horários dos atendimentos, que são os mais variados possíveis na entrada das emergências, foram enquadrados nas faixas apresentadas na Tabela 4.3

As tabelas de dimensão e as correspondências com o modelo relacional encontram-se detalhadas no Apêndice A. Um melhor entendimento dos conteúdos de cada uma das dimensões

Faixa	Descrição da Faixa
01	00:00 às 02:00
02	02:01 às 04:00
03	04:01 às 06:00
04	06:01 às 08:00
05	08:01 às 10:00
06	10:01 às 12:00
07	12:01 às 14:00
08	14:01 às 16:00
09	16:01 às 18:00
10	18:01 às 20:00
11	20:01 às 22:00
12	22:01 às 24:00

Tabela 4.3 Faixas de Horários

será possível quando da apresentação das análises, no Capítulo 4.

4.3.3.2 Definição dos Fatos

A tabela de fatos, Atendimento, utilizando os conceitos apresentados no Capítulo 2, é elaborada com as chaves primárias de todas as dimensões. Uma nova coluna, seqüencial torna-se necessária por haver a possibilidade, que tornou-se materializada na implementação, de um mesmo paciente realizar um atendimento com as mesmas características em uma mesma data.

As medidas representam as seguintes situações do atendimentos:

- Atend_QTD: quantidade de atendimentos realizados;
- Atend_QTD_Obito: quantidade de atendimentos cujos pacientes evoluíram ao óbito;
- Atend_QTD_Transferencia: quantidade de atendimentos com pacientes transferidos para outra unidade de saúde;
- Atend_QTD_Alta: quantidade de atendimentos com pacientes que receberam alta.

Nas análises a serem realizadas cabe observar que cada registro na tabela de fatos corresponde a um atendimento. Neste atendimento podem ter ocorrido o óbito, a transferência ou a alta. A tabela de Fato também encontra-se detalhada no Apêndice A.

4.4 Considerações Finais

Neste capítulo justifica-se a elaboração do Modelo Dimensional, resultado de um processo iterativo onde, mesmo quando considerado finalizado, sempre necessita de refinamentos que geram novas correções. Em algumas situações, somente na realização das análises, fase final do projeto, identificaram-se novas necessidades de ajustes, o que reforça que, semelhante a outras atividades, a fase de manutenção deve ter sua importância ressaltada no processo.

Importante salientar também que toda a construção do *Data Warehouse* do SAPE Emergência vem respeitando cada uma das Fases de Desenvolvimento descritas no Capítulo 2.

No próximo capítulo apresentam-se as Fases de Seleção do Produto, Projeto Físico, Projeto e Desenvolvimento da *data staging area*, Especificações e Desenvolvimento da Aplicação Analítica, citando-se a ferramenta utilizada, sua arquitetura e principais funcionalidades.

No final, os resultados das análises obtidas com a implementação deste modelo dimensional e a utilização da ferramenta, justifica-se enquanto solução de BI em uma proposta para a Secretaria de Saúde do Estado de Pernambuco.

CAPÍTULO 5

Implementação da Solução de BI

Inicialmente são elencados os critérios de seleção e realizada uma comparação que justificam a implementação da solução de *Business Intelligence* (BI) utilizando-se de uma solução gratuita e em código aberto.

Apresentam-se, na seqüência, as principais atividades realizadas no processo de ETL, com uma avaliação desta tarefa. Segue-se a elaboração do cubo OLAP e, no final, realiza-se uma avaliação sobre esta fase.

Seguindo-se todo um processo de desenvolvimento proposto no Capítulo 2, essas fases correspondem às atividades Projeto Técnico de Arquitetura, Seleção e Instalação do Produto, a atividade Projeto Físico e Projeto e Desenvolvimento da *Data Staging Area* e a atividade Especificação da Aplicação Analítica.

5.1 Introdução

A iniciativa de instalar um BI corporativo frequentemente torna-se uma coleção de soluções disjuntas usando uma combinação de aplicações comerciais caras, fechadas e de difícil manutenção.

Tal fato deve-se ao fato de que cada ferramenta é projetada para operar em uma tarefa muito específica. Os processos de negócios devem ser quebrado em pedaços e em tarefas distintas como geração de relatórios, análise, mineração de dados, *workflow*, entre outras. A grande dificuldade é identificar uma plataforma responsável pela gestão, verificação e coordenação destas tarefas. Pessoas e procedimentos são chamados para resolver estas deficiências.

Neste trabalho, elaboramos uma proposta de BI utilizando software livre, orientada a soluções que integra componentes de código e padrões abertos com um mecanismo de análise direcionado a processos. Assim os usuários têm liberdade de executar, copiar, distribuir, estudar, modificar e aperfeiçoar o software.

5.2 Problemas das Ferramentas de BI

Dentre as ferramentas que propõem soluções de BI disponibilizadas atualmente no mercado, encontramos os seguintes problemas a serem considerados no momento da escolha, para citar apenas alguns principais:

- Alto custo: além do alto preço de aquisição do produto, devem ser considerados os custos de manutenção, suporte e serviços relacionados ao software;
- Baixa usabilidade: diretamente ligado ao diálogo na interface e a capacidade do software em permitir que o usuário alcance suas metas de interação com o menor gasto de tempo;
- Dificuldade de implementação: problemas na transferência da solução do fornecedor para o cliente;
- Baixa customização: problemas de integração da solução com as regras de negócio da organização;
- Orientação a ferramentas: as soluções tendem a ser um conjunto de ferramentas e não um ambiente integrado;
- Baixa extensibilidade: as soluções são proprietárias e difíceis, algumas vezes impossível, para os clientes e consultores expandirem as funcionalidades da solução. Compra-se o direito de utilização e geralmente com pagamento no primeiro dia;
- Difícil auditoria: dificuldade na identificação de tarefas realizadas pelos usuários, ações realizadas, tempo gasto e desempenho no processo.
- Escolha sem prototipagem: contratos e valores financeiros geralmente são firmados antes de uma melhor avaliação através de projeto piloto com elaboração de um protótipo;
- Falta de flexibilidade da licença: Ferramentas para a apresentação do resultado de uma consulta no ambiente de BI normalmente requerem licenças de outros aplicativos para a apresentação do resultado de uma análise.

5.3 Análise das Ferramentas

Para identificação da melhor ferramenta de *Business Intelligence* é importante observar quais critérios devem ser levados em consideração no processo de escolha e realizar uma comparação

entre as ferramentas disponíveis, baseadas em software livre.

5.3.1 Critérios de Seleção

Alguns fatores servem de base na utilização de critérios de escolha da ferramenta. A necessidade específica do estudo de caso e os trabalhos relacionados serviram de fundamento para a seleção. Enquanto baseada em código aberto, a compatibilidade com diferentes sistemas operacionais torna-se fator determinante bem como a linguagem de desenvolvimento, a ser considerada em possíveis customizações solicitadas pela organização.

A realização de tarefas de forma prática, fácil e intuitiva, ou seja, a usabilidade, serve como mais um critério de comparação entre as ferramentas. Nesta avaliação não se utiliza um padrão especificamente definido, mas uma comparação quando da instalação das diversas ferramentas em um único ambiente com o objetivo de obtenção de resultados em análises semelhantes.

Em relação ao servidor OLAP deve ser considerada a compatibilidade com os bancos de dados, utilizando-se da mesma premissa de SGBD com código aberto. Considera-se também a linguagem de consulta OLAP e o protocolo padrão de acesso disponibilizado para os clientes da aplicação.

A ferramenta disponibilizada ao cliente que permite realizar as análises no modelo dimensional e a versatilidade na exportação dos resultados em vários formatos também são critérios a serem considerados, assim como a possibilidade de especificações de níveis de segurança na visualização das simulações realizadas.

A portabilidade de utilização em ambiente WEB, em uma rede interna ou em ambiente externo à organização, facilita sobremaneira a versatilidade na realização das análises pelos gestores envolvidos, dispensando a tarefa muitas vezes dispendiosa de instalação e configuração para um perfeito funcionamento.

Outro critério adotado refere-se as funcionalidades complementares como visualização em tabela, visualização em gráficos, visualização geográfica, geração de relatórios e *dashboard*, ou seja, uma possibilidade de disposição gráfica através de indicadores com intervalos pré-definidos, semelhante a um painel de instrumentos. São frequentemente um conjunto de gráficos que ilustram o comportamento ao longo do tempo de uma determinada característica, por exemplo, a evolução das vendas de um determinado produto em função de uma meta pré-estabelecida a ser alcançada.

O conceito de ferramenta integrada de um único fornecedor, embora seja desejável, não se considera um fator determinante na escolha, visto que a integração entre ferramentas e ambientes atualmente pode ser considerado um padrão na viabilidade de solução em tecnologia da

informação.

A periodicidade de atualização das ferramentas pela comunidade responsável, resultando em lançamentos de novas versões podem representar, em uma primeira análise, o grau de investimento intelectual e de aceitação no mercado. Mesmo não sendo um fator decisivo em uma escolha de solução, torna-se também um critério a ser considerado visto que garante uma certa tendência tanto de evolução da solução como de adaptação às novidades tecnológicas disponibilizadas.

Na Tabela 5.1, apresentam-se os critérios a serem analisados do cliente e do servidor OLAP.

Critério	Cliente OLAP	Servidor OLAP
Software Livre	X	X
Linguagem Desenvolvimento	X	X
Sistema Operacional	X	X
Usabilidade	X	X
SGBDs Disponibilizados		X
Linguagem de Consulta	X	X
Gerenciamento das Análises	X	
Ambiente WEB	X	
Visualização em Tabela	X	
Visualização de Gráficos	X	
Visualização Geográfica	X	
Geração de Relatórios	X	
<i>Dashboards</i>	X	
Periodicidade de Atualização	X	X

Tabela 5.1 Critérios Utilizados para Seleção da Ferramenta.

5.3.2 As Ferramentas Analisadas

Uma vez estabelecidos os critérios, faz-se necessário a instalação e configuração dos ambientes utilizando-se da ferramentas disponíveis analisadas, a saber:

5.3.2.1 Mondrian

O Mondrian [mon09] é um servidor OLAP desenvolvido em linguagem Java [jav09] que torna-se independente de sistema operacional por permitir a instalação em uma máquina virtualizada Java, permitindo consultas OLAP em banco de dados relacional.

Na comunidade que disponibiliza projetos em software livre [sou09] o Mondrian representa

a mais consolidada ferramenta em software livre como servidor OLAP [mon09].

O Mondrian executa consultas em uma linguagem padrão do modelo dimensional, o MDX (*Multidimensional Expressions*) [mdx09], obtendo os dados de um banco de dados relacional. Internamente, o Mondrian converte uma consulta MDX em consultas SQL submetidas ao SGBD relacional e compõe as respostas retornadas pelo SGBD em uma visão de cubo, para ser retornada a cliente OLAP.

5.3.2.2 JPivot

O Jpivot [jpi09] caracteriza-se como um cliente OLAP que renderiza tabelas OLAP utilizando navegações como *drill down*, *slice* e *roll-up* disponibilizadas pelo servidor OLAP Mondrian.

Desenvolvido na linguagem Java possibilita a instalação em todos os sistemas operacionais que possuam a máquina virtual Java (JVM).

Utiliza-se de componentes de um *framework* para a criação de gráficos, tabelas, opções de navegação e impressão dos relatórios em vários formatos. Com a utilização dos componentes, o usuário final pode navegar de uma forma fácil e intuitiva pelo modelo dimensional.

Cabe ao Jpivot enviar ao servidor de dados OLAP consultas em linguagem MDX resultantes das análises realizadas pelo usuário na interface OLAP no cliente WEB. Permite acesso diretamente à base de dados relacional utilizando-se a implementação da interface específica do Mondrian.

O Jpivot, enquanto cliente OLAP, compõe, de forma integrada, as várias ferramentas de código aberto para OLAP, como o próprio servidor OLAP Mondrian.

5.3.2.3 OpenI

O OpenI [ope09] caracteriza-se como uma aplicação WEB desenvolvida com tecnologia J2EE cumprindo o critério de possibilitar a instalação em uma máquina virtual Java (JVM) tornando-se, desta forma, multiplataforma.

Como um cliente OLAP, permite uma visualização simples e clara dos dados disponibilizados pelo servidor OLAP em consulta a banco de dados relacionais. Desta forma o usuário pode intuitivamente construir e armazenar relatórios, planilhas e gráficos de maneira bastante intuitiva.

Através da gestão dos projetos cadastrados em seu ambiente, gerencia também usuários, arquivos, conexão com servidores de dados OLAP, seleção de cubos, criação e publicação de relatórios.

O OpenI conta também com componentes dos projetos Jpivot para renderizar o retorno

obtido do servidor de dados OLAP utilizando recursos de segurança do servidor WEB, como por exemplo, o controle de acesso aos projetos usando definições de papéis. As consultas ao servidor OLAP são escritas na linguagem MDX.

5.3.2.4 Palo

O Projeto Palo [pal09] propõe uma solução de *Business Intelligence* que inclui ferramentas de análise, relatórios e ETL. Na solução encontra-se disponibilizado um servidor MOLAP, o PALO OLAP Server.

O Cliente WEB permite também a utilização em um servidor OLAP Mondrian. As análises também podem ser armazenadas com controle de acesso às informações. O cliente utiliza no desenvolvimento a tecnologia AJAX.

A Palo Web Client [pwc09] permite acesso ao servidor OLAP através de uma interface WEB. Encontra-se disponibilizada de forma gratuita, aberta ao público e de alteração livre. Utiliza uma navegação bastante intuitiva e de fácil utilização de forma a possibilitar o melhor das funcionalidades de *drill-down* e *roll-up* através das hierarquias dentro de uma dimensão.

5.3.2.5 SpagoBI

O SpagoBI [spa09] apresenta-se como uma plataforma livre de *Business Intelligence*, uma solução em código aberto para o desenvolvimento de projetos de BI. A solução propõe-se multiplataforma devido ao seu desenvolvimento em linguagem Java utilizando-se também da máquina virtual Java (JVM).

Utiliza-se do servidor OLAP Mondrian bem como do JPivot como cliente para a renderização e apresentação das informações. Outras ferramentas disponibilizadas pela solução incluem uma ferramenta para a geração de relatórios, repositório de conteúdo, ETL (Extração, Transformação e Carga), administração e mineração de dados.

5.3.2.6 Pentaho

O Pentaho [pen09] caracteriza-se como uma plataforma completa de *Business Intelligence* incluindo relatórios, análise OLAP, dashboards, *data mining* e ETL (Extração, Transformação e Carga). Coloca-se como uma solução completa permitindo a utilização de cada um de seus componentes individualmente todas acessíveis através de um ambiente WEB.

Desenvolvida em Java torna-se multiplataforma funcionando em uma máquina virtual Java (JVM). Utiliza-se também do servidor OLAP Mondrian e do cliente JPivot.

Atualmente na versão 3.0, vem apresentando uma atualização constante pela comunidade, refletindo-se em lançamento de versões das ferramentas quase que bimestralmente nos últimos dois anos.

5.3.3 Comparativo entre as Ferramentas

Utilizando-se dos critérios sugeridos em cada uma das ferramentas analisadas permite-se uma análise comparativa.

A Tabela 5.2 resume as ferramentas analisadas em relação aos critérios estabelecidos.

	Pentaho	SpagoBI	Mondrian	JPivot	OpenI	Palo
Open Source	Sim	Sim	Sim	Sim	Sim	Sim
Multi Plataforma	Sim	Sim	Sim	Sim	Sim	Sim
Linguagem	Java	Java	Java	Java	Java	Java
Usabilidade	Muito Alta	Alta	Baixa	Média	Alta	Muito Alta
Suporte SGBD	JDBC	JDBC	JDBC	-	JDBC	JDBC
Linguagem Consulta	MDX	MDX	MDX	MDX	MDX	MDX
Gestão Análises	Sim	Sim	-	Não	Sim	Sim
Ambiente WEB	Sim	Sim	-	Sim	Sim	Sim
Tabelas	Sim	Sim	-	Sim	Sim	Sim
Gráficos	Sim	Sim	-	Sim	Sim	Não
Geográficos	Sim	Sim	-	Sim	Não	Não
Relatórios	Sim	Sim	-	Sim	Sim	Não
Dashboards	Sim	Sim	-	Sim	Sim	Não
Servidor OLAP	Mondrian	Mondrian	-	-	-	Palo/Mondrian
Atualização	Muito Alta	Alta	Alta	Baixa	Média	Média

Tabela 5.2 Comparaçāo entre as Ferramentas Analisadas e Critérios Utilizados.

Ao analisar as diversas soluções e o modo de funcionamento observa-se um direcionamento comum na tentativa de possibilitar as análises dos dados bastante intuitiva e visualmente agradável. Percebe-se claramente uma preocupação crescente não somente no bom desempenho da solução mas também de uma boa apresentação.

A utilização de *dashboards* encontra-se cada vez mais em destaque pois caracteriza-se como uma forma muito simples de visualizar e monitorar os valores de algumas variáveis relacionadas com o desempenho da organização em um único ambiente, embora muitas vezes esta informação tenha origem em locais distintos.

Com referência à camada de análise de dados, tanto na suite SpagoBI como na Pentaho é utilizado o componente JPivot, biblioteca Java, o que mais uma vez revela a semelhança entre

as soluções. Devido ao ambiente de desenvolvimento comum entre as ferramentas analisadas, observa-se uma reutilização de bibliotecas e componentes, como o caso comum do Mondrian, servindo de servidor OLAP.

Com relação a geração de relatórios as opções variam quanto da solução proposta, cada uma utilizando uma solução diferenciada em software livre.

As ferramentas OpenI e Palo Web Client são apenas intermediários entre a camada de apresentação e a camada de dados, permitindo a conexão com um servidor de dados multidimensional através de um mesmo protocolo. Como clientes WEB diferenciam-se das suites completas SpagoBI e Pentaho.

Com relação a periodicidade de atualização, o Pentaho apresenta uma maior maturidade, com aparente maior capacidade de adaptação às novas tecnologias e correções de problemas identificados na comunidade.

5.3.4 Porque o Pentaho?

Após avaliação comparativa utilizando-se dos critérios estabelecidos, outros fatores servem de referendo a alternativa da solução Pentaho.

O Governo Federal vem utilizando uma política de incentivo a utilização de software livre, principalmente utilizando-se da maior empresa pública de informática da América Latina, o SERPRO, Serviço Federal de Processamento de Dados.

O SERPRO investe no desenvolvimento de soluções tecnológicas em Software Livre, como uma política estratégica que permite otimizar os recursos públicos, incentivar o compartilhamento de conhecimento e estimular a cooperação entre as esferas federal, estadual, municipal, iniciativas do segmento acadêmico e sociedade [ser08].

A Suite Pentaho permite que se crie soluções de BI perfeitamente alinhado com as metas de utilização de software livre do Governo Federal e do SERPRO, usando tecnologias atuais, completas e abertas [TdS08].

As soluções de BI utilizando o Pentaho vem sendo utilizada também pelo Ministério das Cidades, no GSAN, Sistema integrado de gestão de serviços de saneamento. O GSAN trata-se de um sistema, desenvolvido com ferramentas de software livre, de Gerência de Operações Comerciais e de Controle da execução de serviços internos, disponível gratuitamente para prestadores dos serviços de saneamento brasileiros e para atendimento de seus usuários [gsa08]. O GSAN foi criado com o objetivo de elevar o nível de desempenho e de eficiência das empresas de abastecimento de água e coleta de esgotos, e pode ser adaptado a empresas de pequeno, médio e grande portes. Atualmente encontra-se implementado em Pernambuco, Sergipe

e Rondônia.

Com casos de sucesso comercialmente implementados, o Pentaho coloca-se como uma real alternativa de solução e por ser *open source*, a suite recebe contribuições de todas as partes do mundo, sempre evoluindo e distribuindo esses avanços livremente. A adoção das mesmas ferramentas em várias instâncias e corporações do governo e da iniciativa privada, potencializa a redução de custos ao facilitar o desenvolvimento, aproveitando-se da facilidade de utilização de ferramentas integradas.

5.4 A Plataforma Pentaho de BI

A plataforma de BI do Pentaho é centrada em processos e orientada a soluções composta por componentes de BI que permitem as organizações desenvolverem uma solução completa e integrada [pen08].

A Pentaho Open BI Suite é um conjunto de softwares *open source* para criação de soluções de BI. Ela possui ferramentas para atender ao processo de criação de soluções de BI de ponta a ponta como mostrado na Figura 5.1. Somente o Sistema Gerenciador de Banco de Dados não é fornecido pela plataforma. Em todo o restante da solução apresentam-se ferramentas integradas.

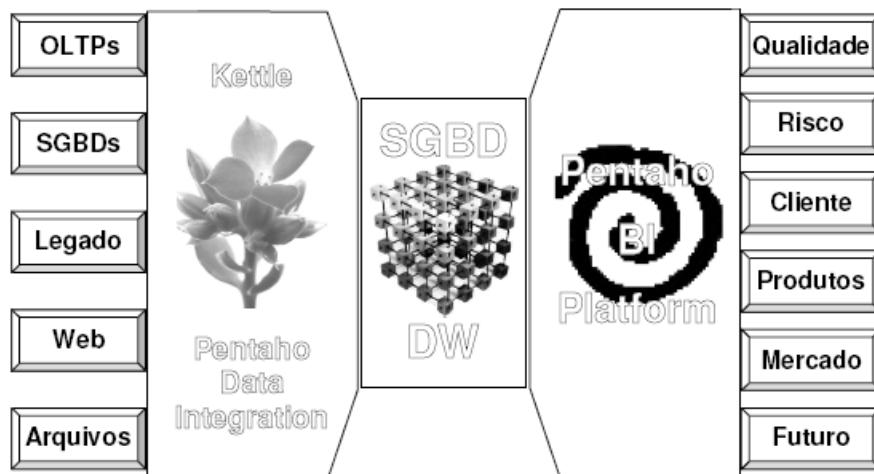


Figura 5.1 Diagrama de uma Solução de BI com a Pentaho Open BI Suite. Retirada de [pen08].

A plataforma conta com as seguintes facilidades para o usuário final:

- disponibiliza uma arquitetura orientada a serviços que inclui auditoria, segurança, escalonamento de tarefas, ETL, serviços de web, repositório de atributos e regras de negócios;

- capacidades do usuário final incluir relatórios, análise, *workflow*, *dashboards* e mineração de dados.
- O Pentaho Design Studio disponibiliza um conjunto de ferramentas para o projeto e gestão que são integradas ao ambiente Eclipse [ecl09]. Estas ferramentas permitem ao analista de negócios tanto desenvolver como criar relatórios, *dashboards*, modelos de análise, regras de negócios e processos de BI;
- As soluções de BI são projetadas usando o Pentaho Design Studio e desenvolvidas no Servidor Pentaho que é o mecanismo de execução em tempo real, com uma abordagem a *workflow*, o qual coordena a execução e a comunicação entre os componentes de BI.

A arquitetura é a combinação de código fonte original e componentes maduros, em código aberto, que integrados formam uma plataforma de BI completa e sofisticada. A plataforma Pentaho de BI é construída tendo como base, servidores, mecanismos e componentes. Eles disponibilizam servidores J2EE, segurança, portais, *workflow*, regras de negócios, colaboração, gestão de conteúdo, integração de dados, análise e ferramentas de modelagem do sistema em desenvolvimento. Muitos destes componentes são projetos consolidados em código aberto, porém podem ser facilmente substituídos por outros produtos [pen08].

Para criar uma solução totalmente integrada, o Pentaho adiciona os seguintes atributos:

- meta dados comuns para a definição de documentos orientados a soluções;
- interfaces comuns a todos os componentes;
- segurança;
- e-mail e notificações;
- instalação, integração e validação de componentes;
- exemplos de soluções;
- conectores de aplicações;
- ferramentas de utilização e diagnóstico;
- ferramentas de projeto;
- configuração e customização.

A Pentaho Open BI Suite utiliza componentes de código aberto desenvolvidos por parceiros. Os componentes de parceiros permitem a Pentaho focalizar nas funcionalidades que adicionam valor aos clientes, tais como: relatórios, análise, integração de *workflow*, regras de negócios, auditoria, interface e análise/modelagem de *workflow*.

Em contra-partida aos ganhos da Pentaho pela utilização destes componentes em sua plataforma de BI, a Pentaho contribui com recursos a estes projetos e lança componentes da plataforma de BI sob uma licença de código aberto [pen08].

Na Arquitetura que apresenta-se a seguir demonstram-se as relações entre os componentes de código aberto e os componentes da pentaho[pen08].

5.4.1 A Arquitetura

Para disponilizar a solução de BI, a plataforma Pentaho de *BI Platform* é composta do Servidor Pentaho, e um Design Studio baseado no Eclipse [pen08].

O Servidor Pentaho é constituído por uma plataforma de BI e bibliotecas que entregam funcionalidades de BI aos usuários e fornece os mecanismos e os componentes para gerar relatórios, análises, e-mail e fluxos de processos. Estes componentes são completamente integrados, podendo desta forma, serem utilizados para resolver um problema de *Business Intelligence*. A estrutura de componentes encontra-se representada na Figura 5.2.

A arquitetura da Suite divide-se em duas partes: a Pentaho BI Platform, a plataforma, e as ferramentas [TdS08]. A plataforma é responsável pela execução da Solução de BI, provendo controle de processos, visualização segurança e auditoria. As ferramentas fornecem produtividade na criação da solução e das estruturas usadas pela plataforma. Todos os softwares da Suite Pentaho são programas Java ou aplicações em Java Server Pages (JSP), e rodam em qualquer plataforma que tenha uma máquina virtual Java.

5.4.1.1 A Plataforma

A Plataforma é uma aplicação em JSP, executada sobre um servidor de aplicações Java, como o JBoss ou o Tomcat. Considerada o bloco central do conceito modular do Pentaho, conforme ilustrado na Figura 5.2, representa os três discos maiores. A Plataforma se divide em duas partes:

- A Solution Engine, ou mecanismo de soluções, que é o responsável pela execução e controle das soluções requisitadas. Funciona baseado em uma máquina de *workflow* interna;

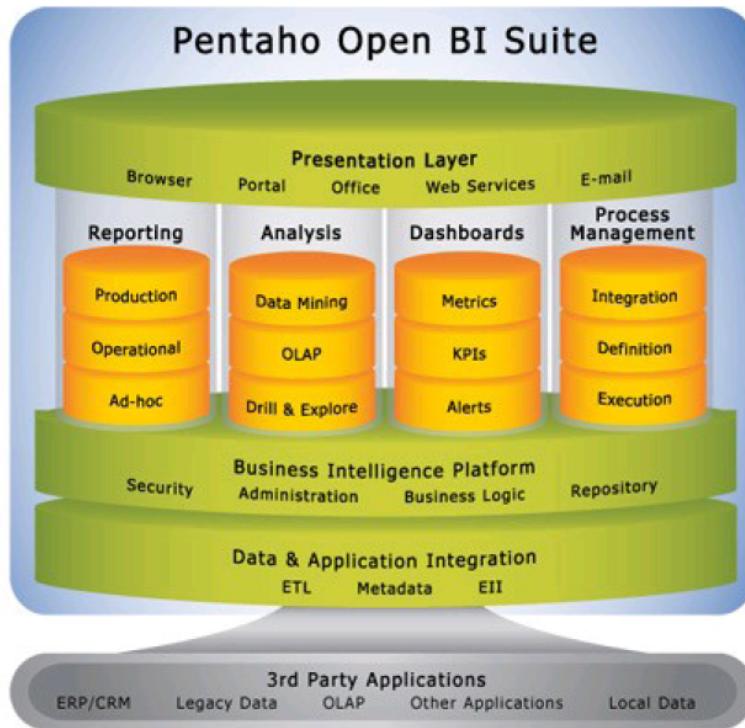


Figura 5.2 Estrutura de Componentes da Pentaho Open BI Suite. Retirada de [pen08].

- O Portal, a porção do Pentaho visível ao usuário final. Através dele o usuário acessa as soluções, sob controle de acesso e validação.

O portal oferece alguns serviços pré-configurados:

- Registro de soluções: cada nova solução é registrada dentro do portal, que se encarrega de exibi-la segundo o estilo visual definido pelo usuário;
- Controle de acesso: o nível de controle pode chegar a permitir ou negar acesso a colunas de tabelas usadas em consultas;
- Agendamento: ações podem ser agendadas através da plataforma, para execução sem supervisão. Os exemplos mais comuns são relatórios periódicos e alertas de eventos, tais como mudanças em indicadores de risco.

Outros serviços podem ser montados no portal, como *dashboards*, envio de emails programados, relatórios *ad-hoc*, entre outros. Em julho de 2008 o portal ganhou uma versão portátil, para o *iPhone* da Apple, que demonstra a versatilidade e agilidade das comunidades envolvidas no seu desenvolvimento e aperfeiçoamento.

5.4.1.2 Ferramentas

As ferramentas integradas disponibilizadas pela solução de BI Pentaho, divididas conforme suas funcionalidades, são as seguintes:

Extração, Transformação e Carga

O Kettle (*Kettle Extracction, Transference, Transformation and Loading Environment*) é a ferramenta da Suite Pentaho para realizar tanto integração de dados quanto criar processos de ETL que alimentam data warehouses. Possui capacidade de acessar, realizando leitura e escrita, mais de uma dezena de formatos de bancos de dados, como Oracle, PostgreSQL, MySQL, SQLServer, além de importar arquivos texto (csv, fixo ou estruturado), planilhas Excel e quaisquer bases que tenham interface ODBC, esse caso apenas em Windows.

O Kettle é um ambiente gráfico no qual conexões com fontes de dados são definidas e seqüências de passos executam a extração de dados e sua modificação até carga em um destino, sempre com nenhuma ou muito pouca programação.

O Kettle foi concebido para equiparar-se a ferramentas comerciais de alto valor agregado. Uma de suas utilizações previstas é a troca e integração de dados entre empresas e sistemas, substituindo o processo de criação de camadas de programas para integração por operações visuais, quase sem nenhuma programação.

Schema Workbench

O portal traz embutido o Mondrian, um projeto *open source* de gerador de cubos OLAP que apesar de criado independente do Pentaho, vem incorporado na Suite pela Pentaho. Com o Mondrian vem também outra ferramenta, o JPivot, uma interface gráfica para visualizar e explorar os cubos OLAP servidos pelo Mondrian.

O Schema Workbench cria os cubos que serão processados pelo Mondrian dentro do Portal e exibidos pelo JPivot. Possui uma interface visual para navegar entre as definições do cubo, permitindo criar, além das tradicionais métricas, dimensões e hierarquias, muitas das estruturas que agregam valor na exploração de um cubo OLAP, como métricas derivadas, cubos virtuais que são combinações de tabelas fato), permissões de acesso e atributos de dimensões. Alguns recursos importantes, porém, ainda não fazem parte do Schema Workbench, como edição de tabelas agregadas [TdS08].

Design Studio

O Design Studio é o ambiente de implementação de Soluções de BI. As Soluções de BI, são formadas por conjuntos de estruturas de execução chamadas *Action Sequences* (AS). Uma AS, é construída a partir de um conjunto de metacomandos, chamados Actions. Esses metacomandos ou Actions são executados pelo mecanismo interno de workflow.

Combinando AS em conjuntos podemos construir soluções de BI complexas. O Design Studio é o módulo onde o desenvolvedor cria as AS e as combina em Soluções, podendo também as testar e otimizar. Fornece ao desenvolvedor de soluções acesso de baixo nível aos recursos do Pentaho.

Aggregation Designer

Um dos maiores problemas a serem considerados na apresentação de uma solução utilizando-se a informatização é a performance e, neste caso específico, das análises realizadas no ambiente de BI. Como normalmente utilizam-se tabelas com conteúdo de milhões de linhas, há uma tendência a uma considerável demora no tempo de resposta. Uma alternativa para a melhoria da performance é a utilização de tabelas agregadas que podem reduzir drasticamente a performance das consultas utilizando-se de uma solução ROLAP.

O Pentaho disponibiliza o Aggregation Designer para simplificar a criação e publicação de tabelas agregadas que garantem melhoria na performance nos cubos OLAP.

Outras Ferramentas

Entre os outras ferramentas disponibilizadas estão:

- Report Design Wizard: permite a criação de relatórios com variados níveis de complexidade.
- Weka: utilizando-se de um ambiente gráfico permite ao usuários identificar comportamentos e realizar simulações em uma base de dados, ou seja, mineração de dados (*data mining*).
- GIS: elaborou-se uma integração da Suíte Pentaho com o Google Earth permitindo implementação de um Sistema de Informações Geográficas, *Geographic Information System* (GIS), obtendo-se o que se conhece como Inteligência em Negócios Georreferenciadas;
- Metadata Editor: permite a criação de mapeamentos entre tabelas que podem ser utilizados pela plataforma para gerar relatórios *ad-hoc* [TdS08]. Esta ferramenta também facilita o agrupamento de colunas de tabelas distintas, criando visões de negócio independentes de melhor visibilidade para os usuários da aplicação.
- Agendamento: ações podem ser agendadas através da plataforma, para execução sem supervisão. Os exemplo mais comuns são relatórios periódicos e alertas de eventos, tais como mudanças em indicadores de risco.

5.5 Extração, Transformação e Carga

Nesta seção apresenta-se os processos de ETL utilizando-se da ferramenta de BI selecionada, Pentaho, que encontra-se definida como parte integrante do Diagrama do Ciclo de Vida Dimensional do Negócio, citado no Capítulo 2, Fase Projeto de Desenvolvimento da *Data Staging Area*.

5.5.1 A Integração dos Dados

Um dos problemas identificados no levantamento atual da situação da informática na SES-PE é a estrutura atual dos bancos de dados não integrada das unidades de saúde. O SGBD onde será realizada a carga deve estar disponibilizado na SES-PE para acesso através da internet de usuários autorizados a realizar as análises no *BI Server*.

Nesta estrutura, a carga necessita ser realizada, recuperando-se as informações atualizadas de cada uma das fontes de dados geograficamente dispersas. O maior problema identificado não são necessariamente os registros inseridos, pois nesta situação bastaria um consulta partindo-se das últimas cargas realizadas. O problema maior é a identificação das alterações realizadas em registros pré-existentes.

Após discussão com os técnicos de informática da SES-PE, dentre as várias propostas, optou-se pela seguinte alternativa:

1. transferir, em um primeiro momento, através da conexão disponibilizada pela PE-Multi-digital ou até mesmo com o deslocamento de um técnico, o último *backup* do banco de dados de cada uma das fontes de dados das unidades de saúde envolvidas;
2. disponibilizar um servidor de dados com o SqlServer instalado na SES-PE e realizar a recuperação do backup em banco de dados distintos no mesmo SGBD;
3. elaborar uma rotina automática para a transferência através da internet somente dos *logs* dos bancos das unidades de saúde;
4. elaborar uma rotina automática para aplicar os *logs* no banco de dados centralizado na SES-PE;
5. realizar o processo de ETL com carga origem no SGBD do SqlServer e destino no banco de dados PostgreSQL com a periodicidade definida pelos gestores envolvidos nas análises.

Os *logs* dos bancos de dados, por tratarem-se de arquivos textos gerados periodicamente em cada um dos hospitais, possuem tamanho considerado pequeno sendo composto somente das inserções e atualizações realizadas no período, o que reduz o problema de gasto de tempo e recursos na transferência de dados entre a SES-PE e as unidades de saúde. A sugestão de utilização dos *logs* também elimina a necessidade de elaboração de sofisticadas rotinas para identificação das alterações realizadas no banco de dados.

Uma vez disponibilizado os SGBD's das unidades de saúde envolvidas na elaboração do *Data Warehouse* em um servidor na SES-PE o pode-se iniciar o processo de ETL que, de uma forma geral, seguem os seguintes passos:

1. Extração dos dados na fonte de origem;
2. Geração de um arquivo texto na *data staging area*;
3. Transformação e limpeza dos dados;
4. Criação da chave delegada ou artificial (*surrogate key*);
5. Carga na tabela do modelo dimensional.

Além disso, seguindo-se a teoria de modelagem dimensional de Kimbal [KR02] citada no Capítulo 2, precisa-se estabelecer a maneira como cada atributo varia com as mudanças da tabela origem ao longo do tempo, ou seja, se vamos adotar um SCD (*Slowly Changing Dimension*) tipo 1, 2 ou 3.

Neste trabalho, como o objetivo são as análises realizadas no DW com a solução de BI, a carga seguiu um critério simples de inserção de linhas nas tabelas de dimensão e fato. Os tratamentos das cargas diferenciais, SCD, estão sugeridas como trabalhos futuros.

Utilizando como base o modelo dimensional, criam-se primeiro as dimensões e em seguida os fatos. Deve-se seguir essa ordem pois na criação das dimensões é gerada a chave delegada ou artificial (*surrogate key*) que vai ser usada para representar o relacionamento do fato com as dimensões.

Importante ressaltar que mesmo tratando-se de dados estatísticos importantes, em uma análise dos gestores envolvidos, avaliando-se os custo e benefício como a importância da informação e a performance da ferramenta de consulta, considerou-se uma redução do volume de informação. Estipulou-se que seria suficiente a avaliação das informações entre os anos de 2005 e junho de 2008, podendo-se, assim, minimizar o tempo de carga e performance nas análise das informações não comprometendo, nesta avaliação inicial, os resultados analisados.

5.5.2 A Extração

Na ferramenta *Kettle* utilizam-se os conceitos Transformação e *Job*. Uma transformação caracteriza-se como uma tarefa a ser realizada, geralmente como acesso ao banco e passos lógicos a serem seguidos. O *Job* é utilizado para encadeamento das tarefas, no caso as transformações, ou até mesmo outros *jobs*.

Neste projeto utiliza-se um único job para a realização da extração de cada uma das tabelas do banco de dados Microsoft SqlServer 2000, como mostra a Figura 5.3, com a possibilidade de agendamento em horário pré-estabelecido e o resultado do processo sendo enviado por e-mail aos responsáveis pelo acompanhamento e execução.

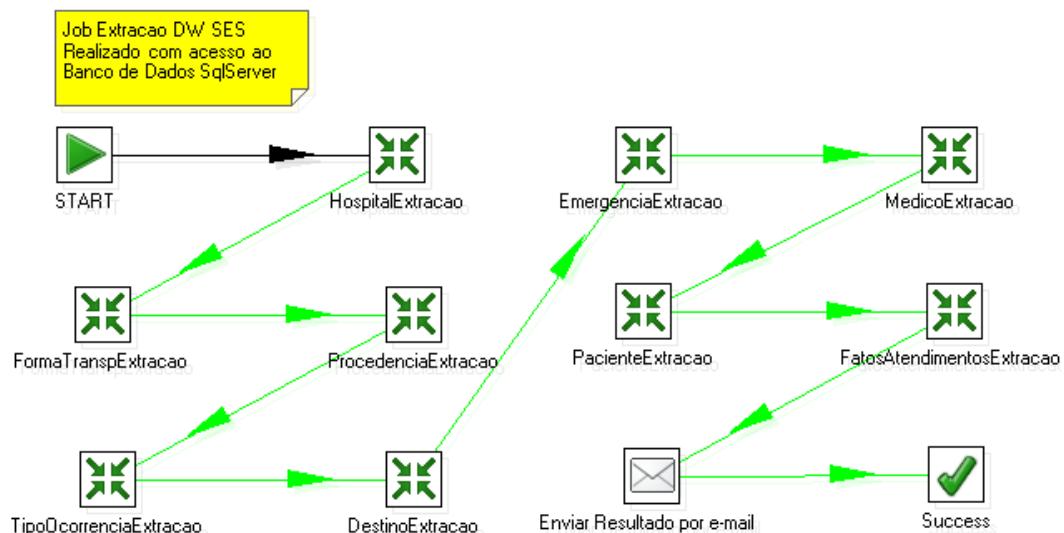


Figura 5.3 Job de Extração DW SAPE.

Cada uma das transformações referente as extrações utilizam o mesmo conceito apresentado para a dimensão Hospital apresentada na Figura 5.4. O detalhamento de todos os jobs e transformações encontram-se no Apêndice B.

Na extração de Pacientes, o problema de duplicação de pacientes que eventualmente estejam cadastrados em mais de um hospital no SAPE, necessita de uma solução que inicia-se na extração. No sistema transacional, para evitar a duplicação do cadastramento existe uma coluna na base de dados, NCADReducido, onde concatenam-se as três primeiras letras que iniciam os três primeiros nomes do paciente com a data de nascimento. Por exemplo, um paciente de nome Marcos Antonio Nascimento nascido em 12/12/1983, gera o NCAD "MAN19831212". Desta forma, se solicitado um novo cadastramento deste mesmo paciente, na confirmação da inserção, o sistema apresenta uma tela com todos os pacientes e o nome das respectivas mães,

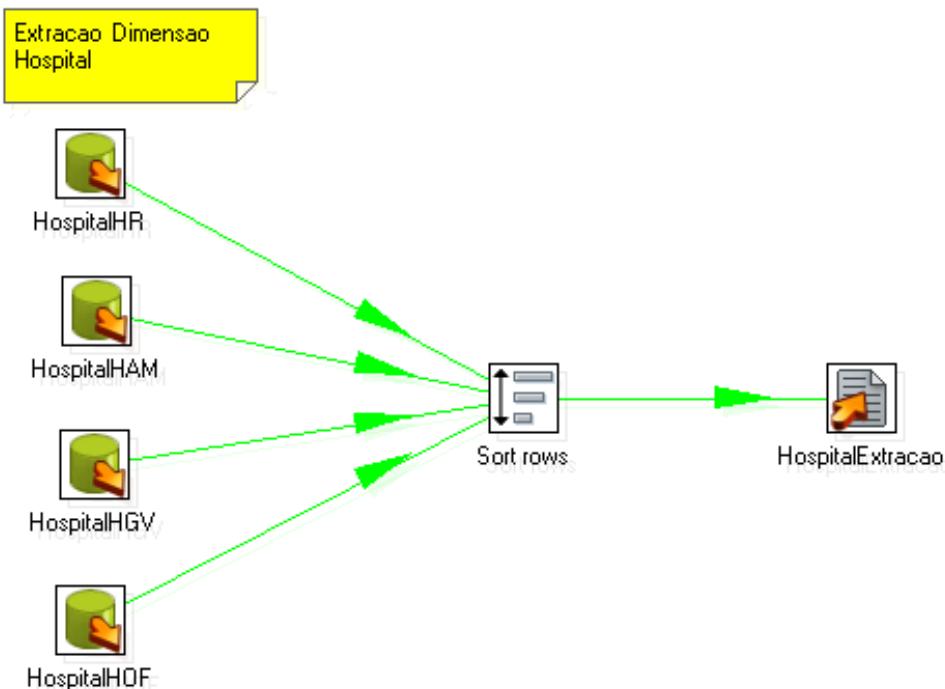


Figura 5.4 Extração da Tabela de Hospital do DW SAPE.

que possuem o mesmo NCADReducido, solicitando uma confirmação do atendente, evitando-se assim que um mesmo paciente seja inserido repetidamente no sistema.

Contudo, pela distribuição de bancos de dados distintos entre os hospitais esta rotina não se aplica. Na tentativa de minimizar o problema, na extração de pacientes, recupera-se o NCADReducido e o Nome da Mãe de cada paciente para que na fase de transformação estas duplicações sejam tratadas. Uma classificação por NCADReducido, Nome da Mãe e Data da Última Alteração torna-se necessária para facilitar o algoritmo que tenta realizar a repetição.

Com relação a quantidade de pacientes, sempre considerando-se do ano de 2005 a junho de 2008, ou seja, 47,97 % do total dos cadastrados, encontram-se distribuídos nos hospitais conforme apresentado na Tabela 5.3.

Hospital	Pacientes
HAM	260.006
HGV	294.524
HOF	268.947
HR	355.436
Total	1.178.913

Tabela 5.3 Quantidade de Pacientes Atendidos de 2005 a Junho de 2008.

Apesar de um considerável volume de dados da tabela de Pacientes, 1.178.913 linhas, e a necessidade de classificação por três colunas na tentativa de eliminação de repetições, o tempo de execução deste *job* de extração, nas diversas execuções, não ultrapassou cinco minutos.

Com relação à extração dos fatos relacionados aos atendimentos, a Tabela 5.4 apresenta a distribuição dos atendimentos cadastrados por Hospital e Ano do atendimento, quando da disponibilização das bases de dados. Estes valores corresponde a 42,99% dos atendimentos cadastrados.

Tabela 5.4 Quantidade de Atendimentos Anual por Hospital

Ano	Atendimentos				
	HR	HAM	HGV	HOF	TOTAL
2005	146.728	129.658	1190.691	157.723	553.800
2006	122.811	133.969	142.719	164.579	564.078
2007	108.269	129.345	135.944	107.015	480.573
Até jun/2008	55.398	56.399	62.388	52.291	226.476
Total	433.206	449.371	460.742	481.608	1.824.927

Nesta extração, visando a otimização na elaboração das faixas etárias, as colunas da tabela de pacientes referentes a data de nascimento e idade estimada foram incluídas. Desta forma, ao recuperar neste ponto da extração estes dados referentes aos pacientes, evita-se uma pesquisa a mais (*look-up*) na tabela de pacientes que possui 1.824.927 linhas, quando do processo de carga. Esta solução apresenta-se como fator determinante na redução do tempo de carga da tabela de fatos, visto que a junção é realizada pelo SGBD e não pela ferramenta de ETL, conforme será demonstrado.

O tempo de extração deste *job*, nas diversas execuções, não ultrapassa vinte e oito minutos, conforme apresentado na Figura 5.5.

Todo o processo de extração gerou arquivos na *data staging area* de tamanho de cerca de 2 *gigabytes* e o tempo total para a execução de todo o *job* de extração não ultrapassou quarenta e cinco minutos. Levando-se em consideração a quantidade de registros, a performance da ferramenta nesta tarefa, envolvendo expressivo volume de dados e algumas classificações, pode ser considerado mais do que perfeitamente aceitável quando da sua implantação na SES-PE.

Uma vez realizada a extração, o tratamento e limpeza dos dados gerados será apresentado a seguir.

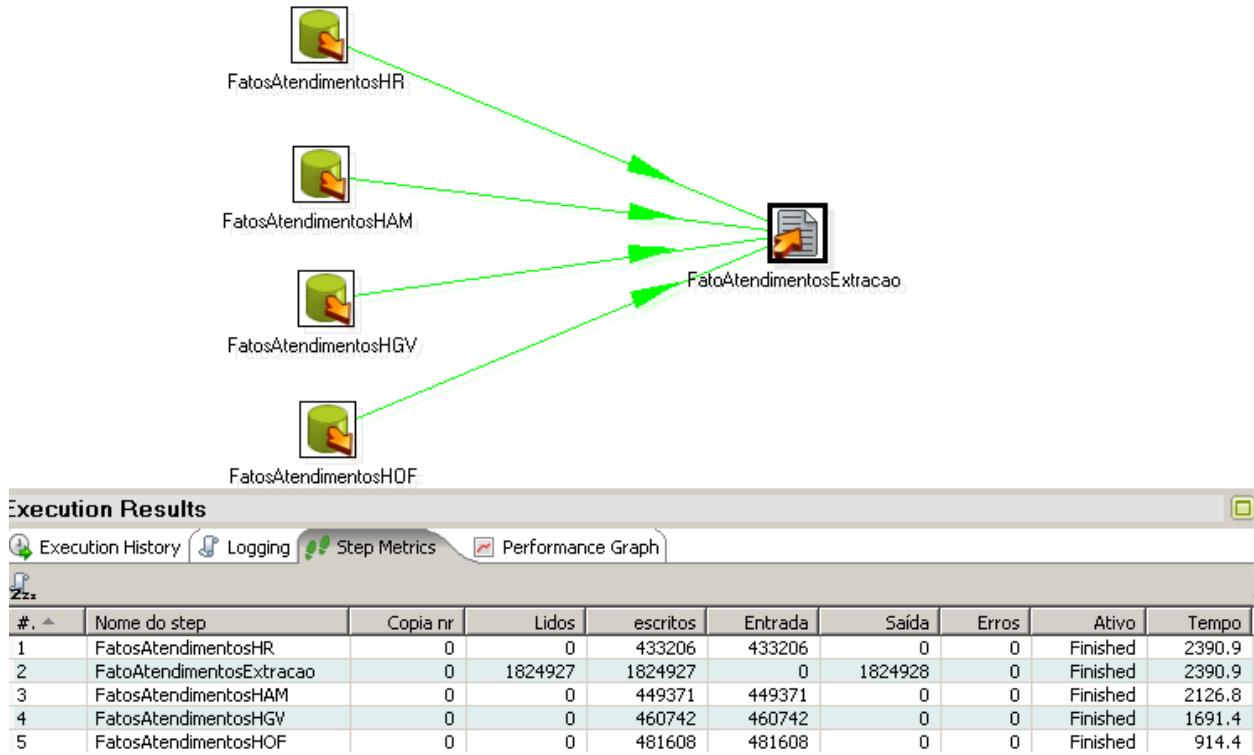


Figura 5.5 Execução do *Job* da Extração dos Fatos.

5.5.3 Transformação

Após a disponibilização dos arquivos na *data staging area* e seguindo-se o processo de ETL, antes da execução das transformações, objetivando um controle das versões carregadas no banco de dados, optou-se pela geração de um sequencial e a data da carga, conforme mostrado na Figura 5.6. Toda a carga é realizada no SGBD PostgreSQL.

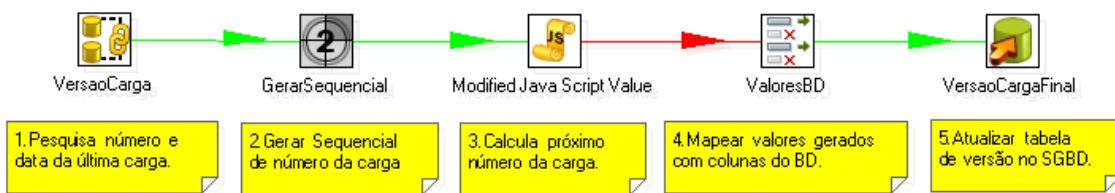


Figura 5.6 Transformação para Gerar Versão da Carga.

As transformações seguiram a mesma seqüência de passos:

1. leitura do arquivo extraído, disponibilizado na *data staging area*;
2. realização das transformações e limpeza, se necessário;

3. geração da chave delegada ou artificial (*surrogate key*) inserindo-se a correspondência entre a chave natural e a chave delegada em uma tabela do SGBD destino do DW;
4. geração do arquivo de carga na *data staging area*.

Na Figura 5.7 apresenta-se o padrão utilizado para as transformações. Utilizam-se, basicamente, os componentes da ferramenta *Kettle* e acrescenta-se um algoritmo simples, em linguagem *javascript* para que, após a geração do sequencial da chave artificial (*surrogate key*), seja realizada a atribuição dos valores na tabela de correspondência. Esta tabela é necessária para a recuperação da correspondência entre as chaves natural e artificial quando da carga da tabela de fatos.

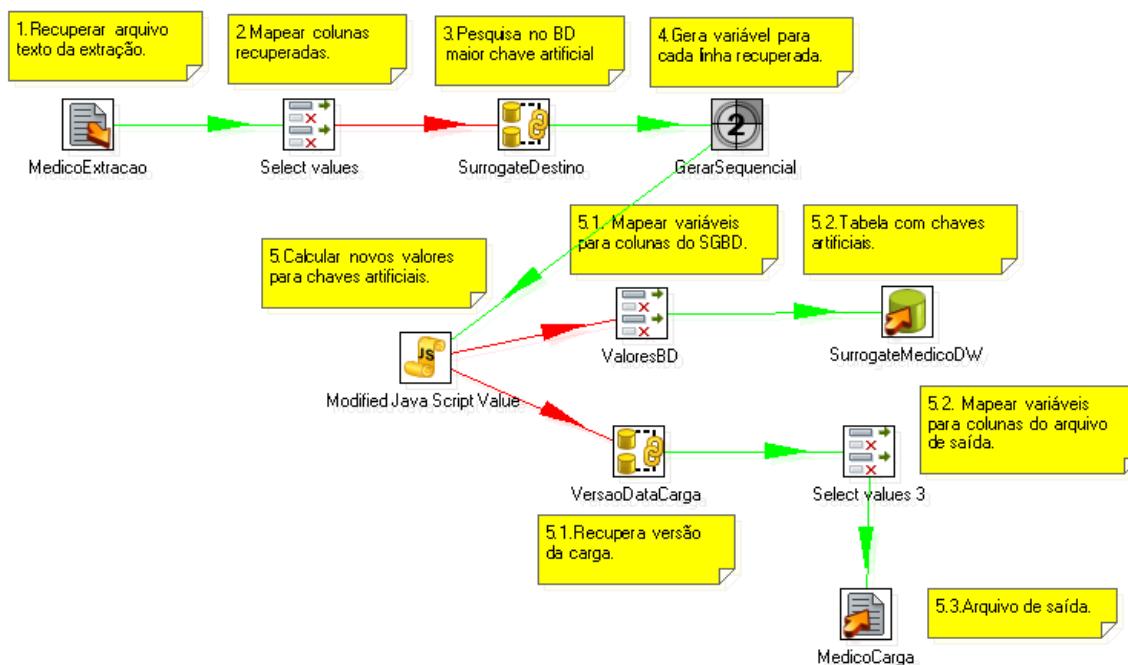


Figura 5.7 Transformação da Dimensão Medico_Liberador.

No caso específico de paciente, a geração da chave artificial segue um processo diferenciado. Na extração de pacientes gerou-se uma classificação crescente por NCADReducido e Nome da Mãe e decrescente pela Data de Atualização, levando em consideração que o registro a ser carregado de um paciente com atendimentos em vários hospitais, será o da última entrada na emergência, teoricamente o que constam as informações mais atualizadas. Identificada a ocorrência de repetição pelo mesmo NCADReducido e mesmo Nome da Mãe, tantos quantos forem os códigos registrados do paciente estarão associados a uma única chave artificial. Este

tratamento de limpeza é realizado no componente *Java Script* da transformação e apresentado na Figura 5.8.

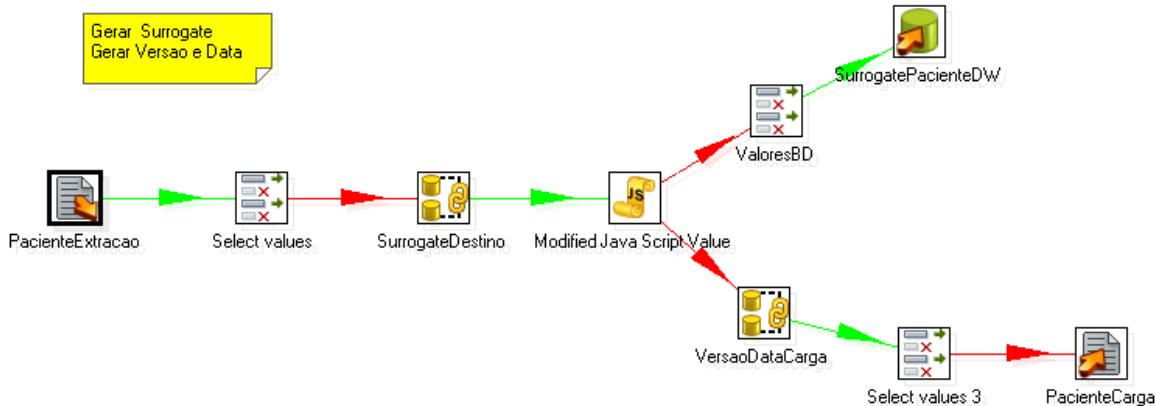


Figura 5.8 Transformação da Dimensão Paciente.

A Figura 5.9 apresenta o algoritmo para eliminação das repetições de paciente nas bases de dados das unidades de saúde.

```

Step name : Modified Java Script Value
Java script :
  Script 1
  //Script here
  var NCadAnterior;
  var NomeMaeAnterior;
  var newsk;
  if(NCadAnterior ==null){
    newsk=maxsk.getInteger()+1;
  } else {
    if (NomeDaMae.getString() != null && NomeMaeAnterior != null && NCadReduzido.getString().equals(NCadAnterior) ) {
      nomeAtual = trim(NomeDaMae.getString());
      nomeAntigo = trim(NomeMaeAnterior);
      k=nomeAtual.length<nomeAntigo.length ? nomeAtual.length-1:nomeAntigo.length-1;
      counter =0;
      for (i=0;i<k;i++){
        counter += nomeAtual.charAt(i)==nomeAntigo.charAt(i) ?0:1;
      }
      if (counter >3){
        newsk++;
      }
    } else {
      newsk++;
    }
  }
  NCadAnterior = NCadReduzido.getString();
  NomeMaeAnterior = NomeDaMae.getString();
  
```

Figura 5.9 Algoritmo de Eliminação de Redundância de Pacientes.

De forma simplificada, o algoritmo realiza os seguintes passos:

1. Identifica se o Nome Da Mãe atual está preenchido e o Nome da Mãe anterior, armazenado em variável, está preenchido e se o NCAD atual é igual ao NCAD anterior, armazenado em variável. Nesta situação tem-se o mesmo NCAD e nomes preenchidos a serem comparados.

2. Retira os caracteres "espaço" do NomeAtual e do NomeAntigo da Mãe, atribuindo à variáveis.
3. Obtem-se o tamanho da menor variável entre Nome Atual da Mãe e Nome Antigo da Mãe.
4. Pesquisando letra a letra até o tamanho do menor dos nomes, quando encontrar uma letra diferente, acrescenta 1 ao contador.
5. Após pesquisa, se quantidade de letras diferentes for maior que três, considera-se que mesmo possuindo o mesmo NCAD, o Nome da Mãe é diferente, e atribui-se uma nova chave artificial. Caso contrário, considera-se o mesmo paciente que receberá a mesma chave artificial do anterior.
6. Atribui às variáveis anteriores os nomes atuais para a repetição do algoritmo quando de novo registro de entrada para nova comparação.

Quando encontrado um mesmo NCADReducido, na comparação realizada no nome da mãe, por ser um valor digitado na recepção, estipulou-se um limite de até três letras diferentes para ser considerado outro nome, minimizando os erros de digitação do nome pelo atendente da emergência.

Devido ao grande volume de dados e ao tratamento das duplicidades, a transformação da dimensão Paciente resultou na maior demora de processamento, cerca de 48 minutos com eliminação de 107.000 pacientes repetidos, o equivalente a 9,08% dos cadastrados na emergência. Observando-se mais atentamente o resultado, percebe-se que houve eliminação não somente dos cadastros repetidos por atendimentos de um mesmo paciente em hospitais distintos mas também cadastramentos repetidos registrados no próprio hospital.

Estes valores representativos, demonstram que o trabalho de conscientização e o treinamento dos atendentes podem resultar em uma ação efetiva na redução na duplicidade do cadastramento no hospital. Entre os hospitais, devido aos bancos de dados distintos, somente nesta transformação as redundâncias serão eliminadas.

A fase de transformação encontra-se no mesmo *job* da fase de carga visto que utilizarão os arquivos disponibilizados pela extração em um SGBD e realizam as transformações para a carga em outro SGBD. As transformações utilizadas no processo de ETL encontram-se igualmente detalhadas no Apêndice B. A fase de carga apresenta-se a seguir.

5.5.4 Carga

Na seqüência do processo de ETL, realiza-se a carga dos dados transformados, disponibilizados na *data staging area* para a carga do DW no SGBD PostgreSQL.

O *job* de transformação e carga seguem o padrão apresentado na Figura 5.10.

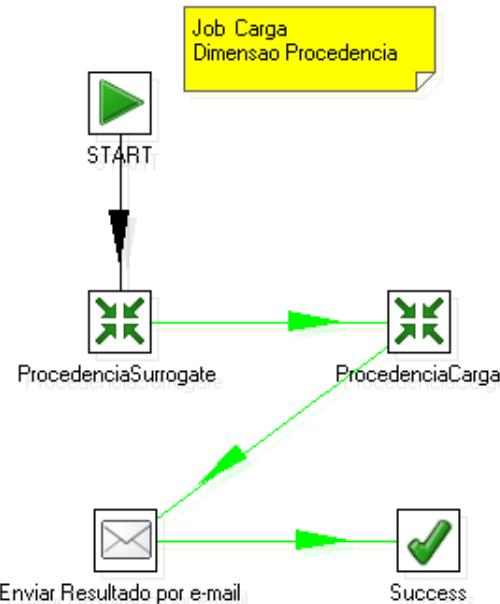


Figura 5.10 Job de Carga da Dimensão Procedência.

A carga da tabela de pacientes encontra-se otimizada com a eliminação dos pacientes que possuem a mesma chave artificial.

A carga da tabela de fatos, último passo da carga do DW, realiza uma busca nas tabelas onde constam as correlações entre a chave natural, do banco de dados transacional, e a chave delegada ou artificial, gerada no processo de transformação. Há também a identificação da Faixa Etária do Paciente, o mapeamento do Horário de Registro na Faixa de Horário e, ao final, encontra-se mapeados os valores referente a cada uma das medidas. A carga da tabela de fatos segue os seguintes passos:

1. Leitura do arquivo da extração de fatos;
2. Pesquisa na tabela de hospitais qual a chave artificial correspondente à chave natural, pesquisa de hospital;
3. Pesquisa de Emergência e Paciente;

4. Pesquisa de Data de Registro associando a tabela de Tempo;
5. Cálculo da Faixa Etária;
6. Mapeamento do Horário do Registro nas Faixas de Horários definidas;
7. Pesquisa de Procedência, Destino, Tipo de Ocorrência, Forma de Transporte e Médico;
8. Mapeamento de Médico para atribuição dos não identificados;
9. Mapeamento das medidas Óbito, Transferência, Internação e Atendimentos para valores numéricos;
10. Geração do arquivo de carga.

O Job de carga de todo DW, envolvendo as fases de geração das chaves artificiais e cargas, encontra-se apresentado na Figura 5.11. Um detalhamento de cada uma das transformações encontra-se disponibilizado no Apêndice B.

O processo de transformação e carga, nas diversas execuções realizadas, não ultrapassou 2 horas e 20 minutos.

5.6 Cubos OLAP

Os cubos OLAP são criados a partir da ferramenta integrada do Pentaho, *Schema Workbench*. Neste trabalho utiliza-se um único cubo e a partir deste cubo qualquer análise realizada pelos usuários podem ser armazenadas no ambiente do Pentaho. Em qualquer tempo, quando solicitada a apresentação da análise, seja em forma de gráfico e/ou tabelas com cruzamentos de informações, em qualquer das cargas realizadas, os valores serão atualizados, ou seja, uma vez realizada uma análise sobre o cubo, as visões dos cubos são armazenadas independentemente dos dados disponibilizados no DW. A Figura 5.12 apresenta a tela principal do ambiente da ferramenta de elaboração do cubo.

Interessante observar que as hierarquias que permitem as navegações (*drill*) entre os níveis e que servem de aprofundamento nas análises utilizam-se das desnormalizações realizadas no modelo dimensional. Por exemplo, na criação da dimensão do cubo Natureza da Ocorrência, o Tipo de Ocorrência encontra-se um nível abaixo, permitindo a navegabilidade da Natureza para o Tipo. Havendo a Natureza, por exemplo, Agressão, pode-se detalhar, realizando um

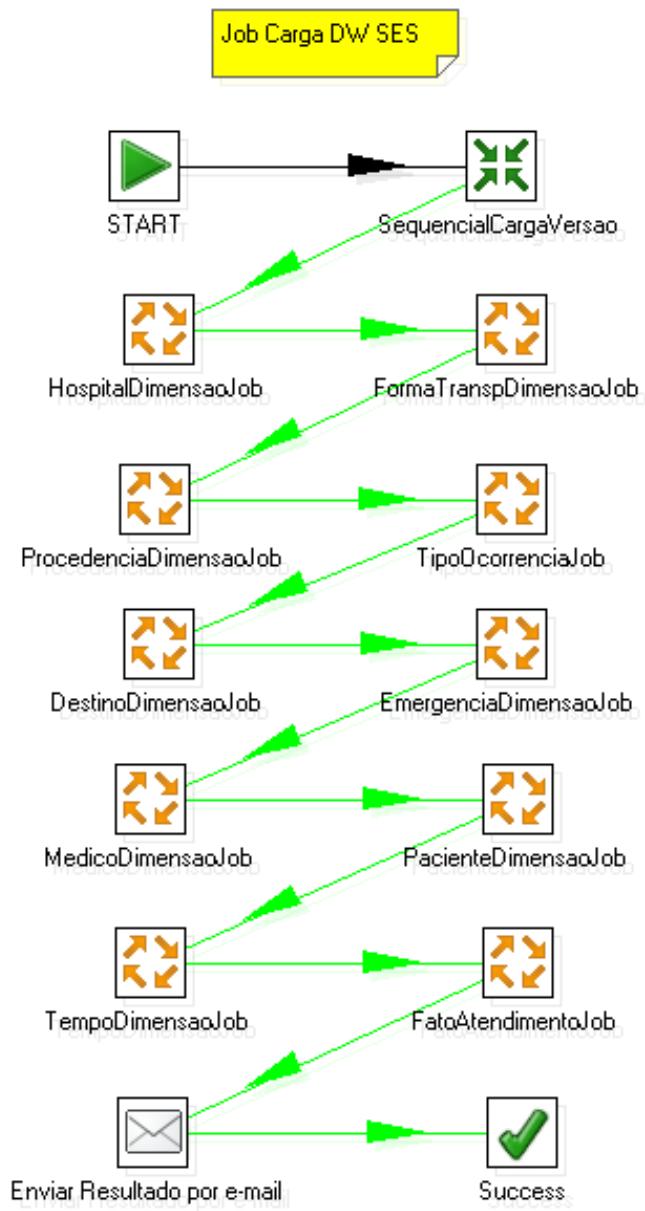


Figura 5.11 Job de Carga do DW SAPE Emergência.

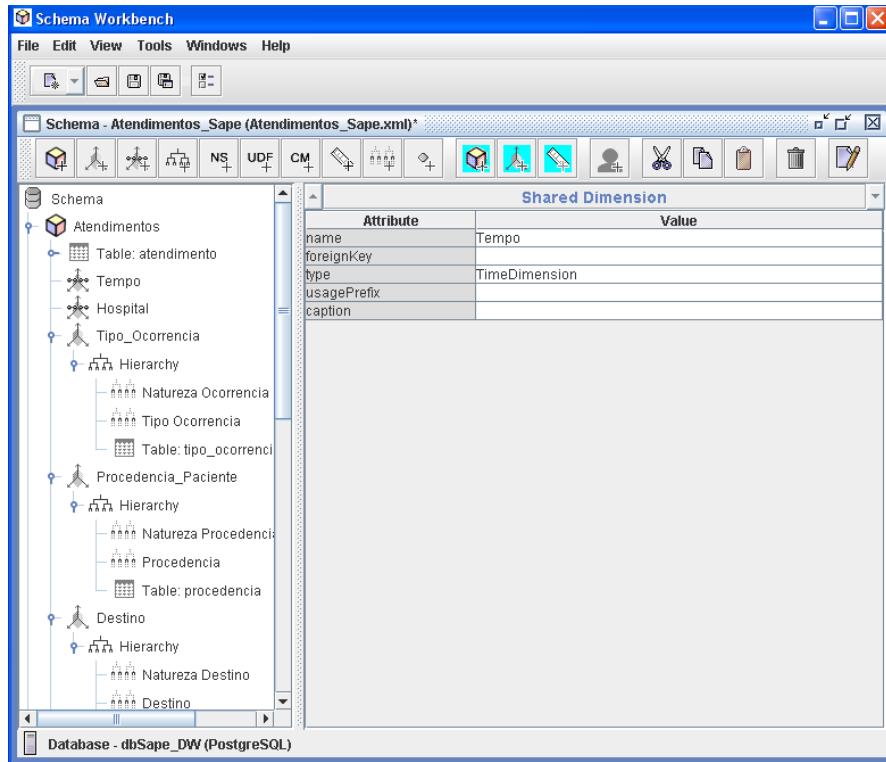


Figura 5.12 Elaboração do Cubo OLAP.

drill-down, as quantidades de Agressão Animal, Agressão Física e Arma Branca, por exemplo. Desta forma a construção do cubo está diretamente relacionada ao que será permitido na análise.

Para disponibilizar o cubo no *BI Server*, ambiente onde serão realizadas as análises para os gestores, devido as facilidades da integração das ferramentas, utilizando-se uma opção de publicação, *Publish*, e informando-se o usuário e a senha, o cubo encontra-se habilitado para as análises. Esta funcionalidade permite grande versatilidade nas inclusões ou alterações de novas dimensões, realizadas em ferramentas em ambientes independentes, porém com total integração.

Na elaboração do cubo OLAP há sempre a costumaz necessidade de associação entre as chaves primárias e chaves estrangeiras, neste caso no modelo dimensional. Alguns dos erros identificados nas análises referem-se a problemas relativos aos relacionamentos não informados no cubo OLAP. Desta forma torna-se necessário um cuidado especial na elaboração do cubo, assim como um bom conhecimento do modelo dimensional envolvido.

A implementação das dimensões degeneradas, uma vez que não existe a tabela física de dimensão no modelo, requer um pouco mais de cuidado, porém a versão mais recente do *Workbench* 3.0.4, torna a tarefa mais simplificada.

De modo semelhante às medidas, a necessidade de membros calculados refletindo, por exemplo, os percentuais de óbitos em relação aos atendimentos, solicitados pelos gestores, possuem implementações simples no cubo OLAP.

Uma outra alternativa é a implementação diretamente no arquivo fonte gerado pela ferramenta, em formato XML (*Extensible Markup Language*) [xml08]. Uma "engenharia reversa", ou seja, a implementação diretamente no código fonte e posterior acesso através do *Workbench* mostrou-se como uma alternativa a ser sempre considerada, sem apresentar nenhum problema quando da gravação no ambiente gráfico e posterior publicação no *BI Server*. Novamente a integração entre os ambientes e ferramentas mostrou-se ser um forte fator facilitador para a implementação das soluções.

5.7 Considerações Finais

Durante toda a tarefa de ETL utilizou-se do ambiente computacional onde será implementada a solução, na Secretaria Estadual de Saúde de Pernambuco, com um servidor System x3400 [ibm09], com processador Intel Xeon de núcleo duplo com 16 gb de memória RAM e sistema operacional Windows Server 2003 [mso09].

Baseando-se no tempo de extração em torno de 45 minutos e o no tempo da transformação e carga, em torno de 2 horas e 20 minutos, o tempo total do processo de ETL pode ser estimado em torno de 3 horas e 15 minutos. Este tempo pode ser considerado aceitável, visto que todo o processo encontra-se automatizado e a ferramenta permite o agendamento para execução em horário noturno.

Com o compromisso assumido de implantação de uma otimização das atualizações, utilizando-se das teorias de SCD, haverá certamente uma redução significativa em novas cargas.

Em relação à ferramenta Kettle, a facilidade de conexão com os diversos bancos sem a necessidade de realizar buscas dos chamados *drivers* facilitam sobremaneira a integração entre os bancos de dados envolvidos.

A quantidade de componentes disponibilizados demonstram grandes variedades de opções de recursos para a elaboração de tarefas que tornam-se simplificadas e com relativamente poucos passos, facilitando o entendimento e a manutenção.

O ambiente gráfico com a funcionalidade de "arrastar e soltar", *Drag-and-Drop*, associado a programação, mesmo com as dificuldades inerentes da linguagem *javascript*, aumentam a flexibilidade da ferramenta inclusive, facilitada com a implementação de novas funções da linguagem exclusivamente para o ambiente. Todos estes fatores agregados tornam a tarefa

simplificada e com alta produtividade.

Quanto à execução, apesar do *job* respeitar o sequenciamento das tarefas, uma das características da ferramenta que visivelmente implicam na redução do tempo de conclusão é o paralelismo de tarefas dentro de uma transformação. Observando-se a execução de cada tarefa, percebe-se que não há a serialização de leitura e escrita de cada base de dados. Os dados são lidos e gravados com independência, inclusive quando da necessidade de classificação.

As dificuldades encontradas em ETL referem-se mais a instabilidade de cada execução que, muitas vezes, apresentavam erros em uma primeira execução que desapareceram quando de novas execuções, sem haver sido realizado nenhuma alteração em seus passos. Este comportamento gera uma natural instabilidade na identificação da origem de um problema, visto que algumas vezes o problema nem mesmo existe.

A ausência de um *help on-line*, tanto na ferramenta de ETL quanto na de criação do cubo OLAP, é compensada pelos fóruns das diversas comunidades existentes em diversos países causando sempre a impressão de tratar-se de ferramentas de grande aceitação, alta produtividade e satisfatoriamente consolidada como solução tanto para o processo de ETL como para a elaboração do cubo OLAP.

CAPÍTULO 6

Resultados Obtidos

Este capítulo tem como objetivo ressaltar a utilidade da solução de *Business Intelligence* proposta como ferramenta de apoio à decisão gerencial na gestão da saúde, tendo como referências as emergências dos maiores hospitais públicos do Estado de Pernambuco.

Descreveremos a metodologia, alguns resultados e, principalmente, conclusões gerenciais e estratégicas da situação, atual e histórica, servindo como diagnósticos sobre a saúde e a gestão de saúde no Estado.

No processo de desenvolvimento do DW, detalhado no Capítulo 2, esta fase corresponde ao Desenvolvimento da Aplicação Analítica.

Outras análises implementadas encontram-se no Apêndice C.

6.1 Metodologia

Após a geração do DW com os dados das maiores emergências dos hospitais públicos do Estado referentes aos anos de 2005 a junho de 2008, ensejou-se validar a utilidade do sistema com seus reais usuários, ou seja, gestores da Secretaria de Saúde. Assim, conseguimos agendar entrevistas com dois gestores de saúde, participantes de políticas de governos um pouco distintas porém ambos com participação direta na gestão da saúde:

- Dr. João Veiga Filho, médico, ex-Secretário de Saúde de Olinda, na gestão da Prefeita Luciana Santos, do Partido Comunista do Brasil (PCdoB). Trabalhou como cirurgião, atendendo principalmente às emergências, durante 12 anos no Hospital da Restauração, onde publicou diversas pesquisas, participando, eventualmente, na gestão administrativa do hospital. Entrevistado em um tempo total de cerca de 12 horas.
- Dra. Darcy Freitas, médica, ex-secretária de Saúde de Olinda, na gestão da Prefeita Jacilda Urquiza, do Partido do Movimento Democrático Brasileiro (PMDB), e de Paulista, na gestão do Prefeito Antônio Speck, também do PMDB. Ex-diretora da I GERES, correspondente à região da cidade do Recife em duas gestões, Governador Miguel Arraes, do

Partido Socialista Brasileiro (PSB) e Governador Jarbas Vasconcelos do PMDB. Assessora da Secretaria de Saúde do Município de Ipojuca, Prefeito Pedro Serafim, do Partido Democrático Trabalhista (PDT). Entrevistada durante 3 horas.

Nestas entrevistas, realizaram-se apresentações da solução proposta de BI, com dados reais, efetuando-se consultas e simulações, algumas solicitadas pelos entrevistados. Após uma primeira versão da documentação dos resultados houve uma validação realizada com Dr. João Veiga, durante cerca de 3 horas.

A seguir apresentaremos alguns dos resultados destas consultas, bem como as respectivas análises, ou seja, a interpretação gerencial e estratégica destes dados, permitindo importantes diagnósticos e conclusões da saúde pública do Estado de Pernambuco.

6.2 Análises e Resultados de Apoio à Decisão

Uma vez elaborado o modelo dimensional e publicado o cubo OLAP, o Pentaho disponibiliza um ambiente para a realização das análises.

Nesta fase, apresentam-se os resultados que podem refletir em ações de melhoria nos dados, gestão hospitalar e gestão estratégica da saúde no Estado.

Todas as análises realizadas podem ser armazenadas em um diretório facilmente criado através da ferramenta. As opções envolvem as dimensões e os fatos disponibilizados na elaboração do cubo OLAP com filtros, seleção de linhas e colunas, ordenação, dezesseis tipos de gráficos e as mais variadas opções parametrizadas, adaptando-se facilmente à maneira de apresentação desejada pelo usuário.

Importante ressaltar que as análises extraídas no DW apresentam uma interpretação dos dados armazenados unicamente nas emergências no período de 2005 a junho de 2008. Diferentes interpretações podem ser realizadas por gestores utilizando-se desta ferramenta de análise.

Devido a baixas não realizadas, ou seja, atendimentos que ainda encontram-se em aberto, mesmo com o passar dos anos, ou que foram baixados automaticamente por rotina específica, em algumas situações o somatório de óbitos, internações, transferências e altas não correspondem ao total de atendimentos. Como correspondem a 7,24% do total de atendimentos e não representam informação relevante no contexto, foram desprezadas das análises.

6.2.1 Atendimentos por Hospital

Os atendimentos nas emergências dos hospitais públicos que utilizam o sistema SAPE Emergência são apresentados na Figura 6.1 com as respectivas quantidades total, óbitos, internação, transferência e alta e respectivos percentuais sobre os atendimentos.

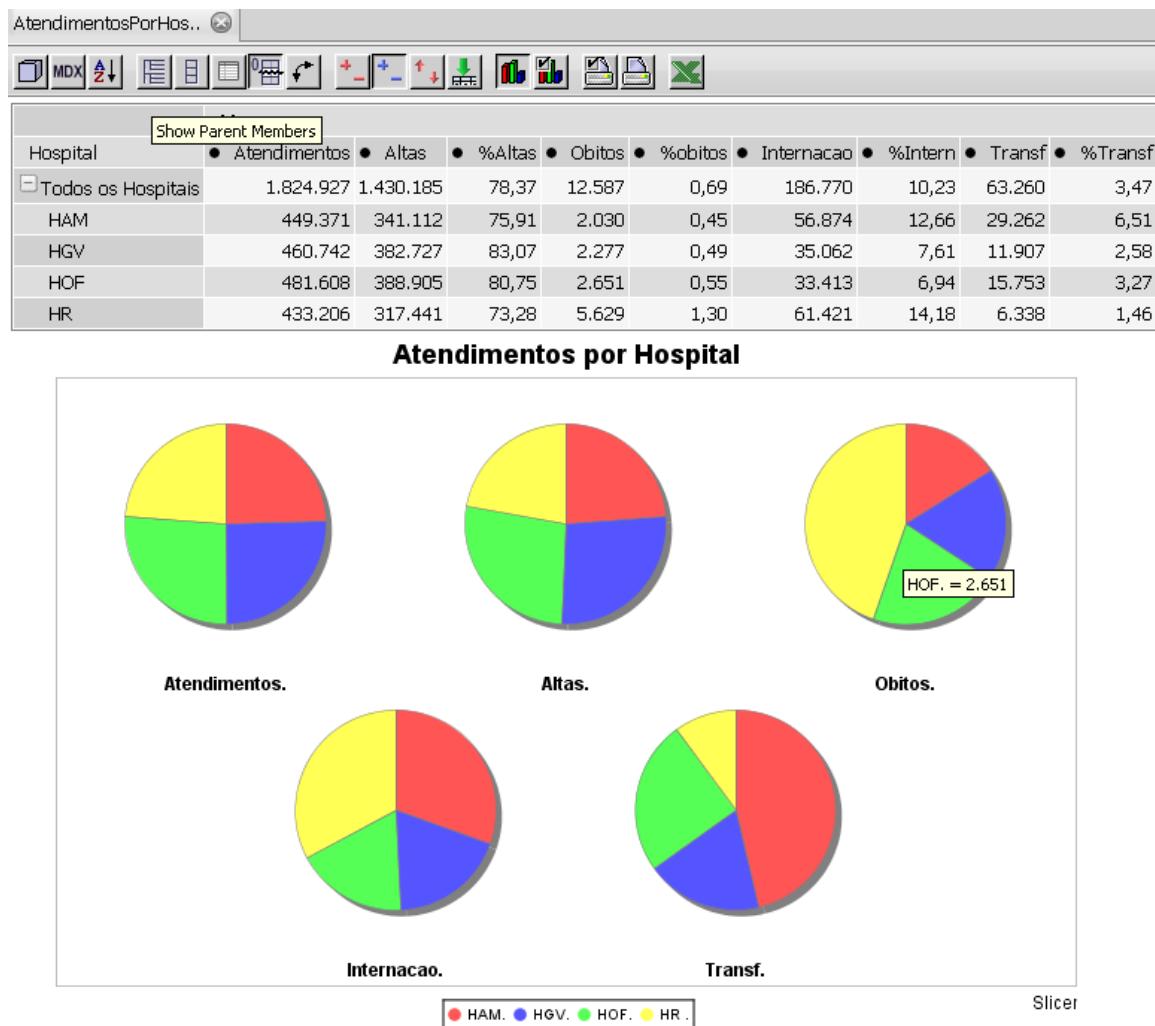


Figura 6.1 Atendimentos por Hospital.

Observa-se uma distribuição regular das quantidades de atendimentos por hospital. As maiores quantidades de óbitos e internamentos no Hospital da Restauração justificam-se pela referência nos atendimentos a politraumatizados e neurologia, assim como uma maior quantidade de leitos disponíveis, principalmente nas especialidades que historicamente geram mais óbitos. As transferências serão tratadas em uma seção específica com as análises e justificativas.

Em um acompanhamento anual, mostrado na Figura 6.2, percebe-se claramente uma dimi-

nuição nos atendimentos realizados no HR e no HOF. No caso do HR, mais especificamente no ano de 2007, houve um aumento considerável de óbitos mesmo com a diminuição dos atendimentos, tendência que vem se mantendo na mesma proporção em 2008. Em avaliação dos gestores, identificou-se que neste período houve um aumento das vagas nas Unidades de Terapia Intensiva (UTI), refletindo em um aumento dos óbitos devido a um aumento nas internações de casos de alta gravidade.

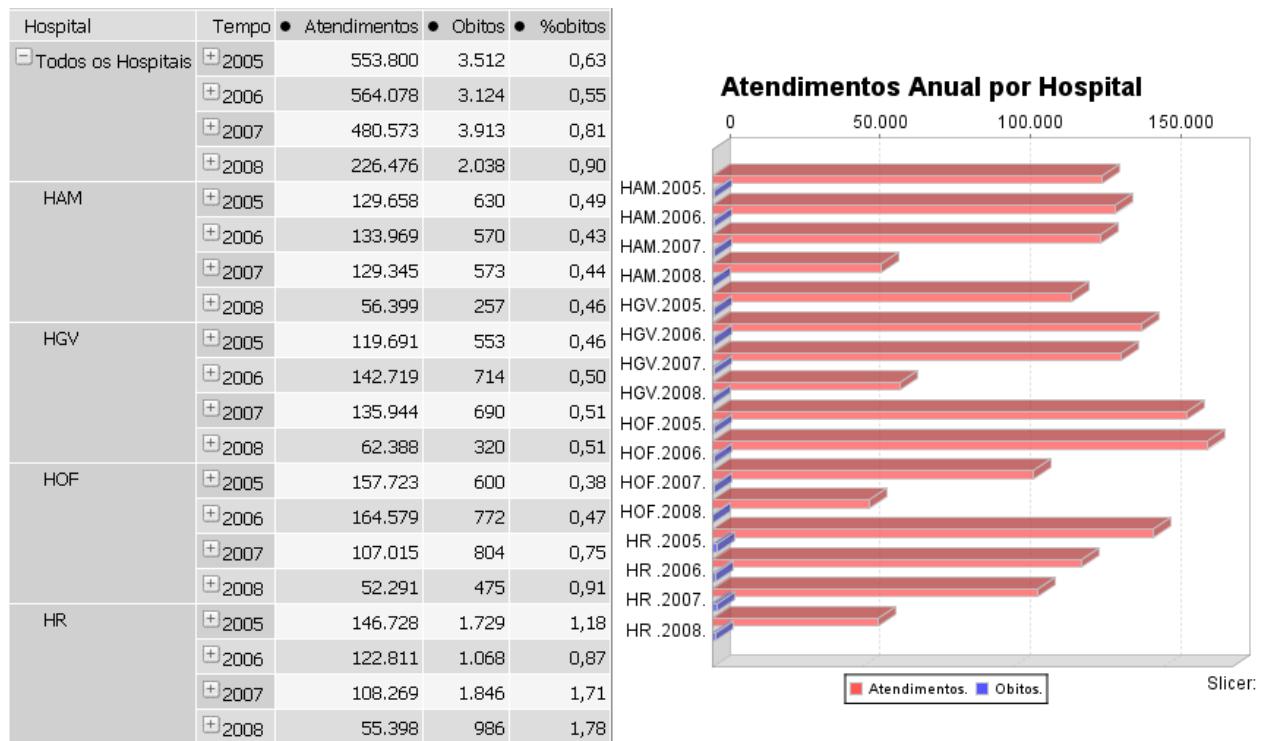
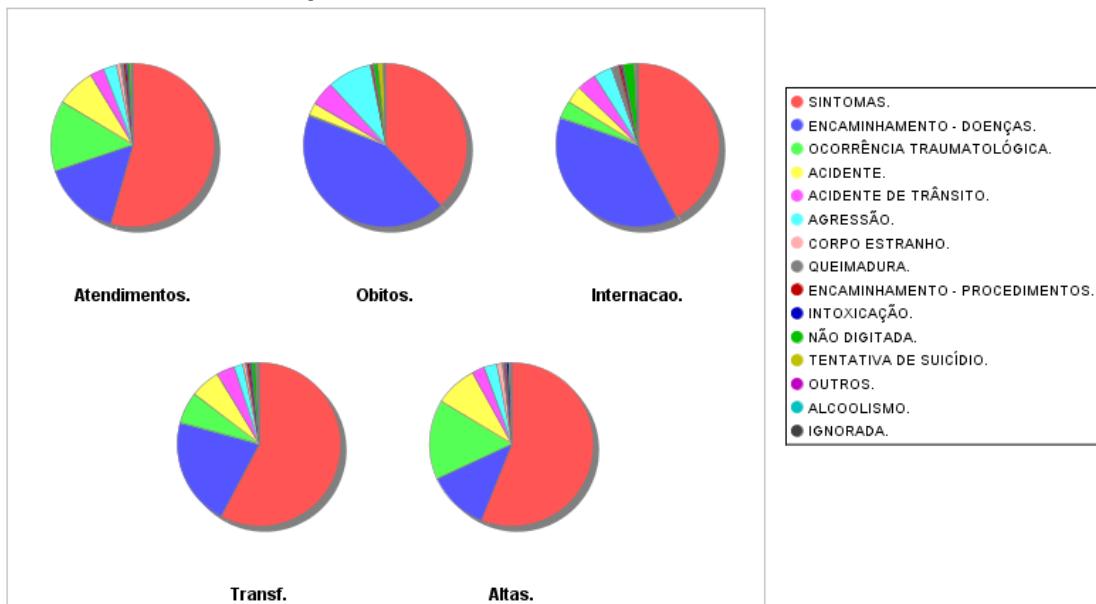


Figura 6.2 Atendimentos e Óbitos Anual por Hospital.

6.2.2 Tipos de Ocorrências

Nesta análise, identifica-se que a maior parte dos atendimentos, e o segundo em óbitos, referem-se à Natureza do Tipo de Ocorrência "Sintoma", conforme mostra a Figura 6.3. Como a identificação da ocorrência é realizada inicialmente pelos atendentes na recepção, apesar dos acertos decorrentes do convívio diário, o diagnóstico médico do atendimento não vem sendo informado no sistema quando da baixa do atendimento. Desta forma, ocorrências como falta de ar, cansaço, dor de vários tipos, desmaio, e outros, que podem ser identificados realizando-se um "drill-down", representam 54,42 % de todos os atendimentos.

Tipo_Ocorrencia	Atendimentos	Obitos	%obitos	Internacao	%Intern	Transf	%Transf	Altas	%Altas
SINTOMAS	993.165	4.805	0,48	78.852	7,94	36.521	3,68	800.997	80,65
ENCAMINHAMENTO - DOENÇAS	282.835	5.375	1,90	71.196	25,17	13.592	4,81	173.808	61,45
OCORRÊNCIA TRAUMATOLÓGICA	253.576	27	0,01	6.671	2,63	4.033	1,59	222.749	87,84
ACIDENTE	137.353	289	0,21	6.442	4,69	3.757	2,74	118.507	86,28
ACIDENTE DE TRÂNSITO	50.870	625	1,23	6.974	13,71	2.227	4,38	37.299	73,32
AGRESSÃO	49.027	1.111	2,27	6.913	14,10	1.188	2,42	36.134	73,70
CORPO ESTRANHO	15.919	1	0,01	255	1,60	360	2,26	14.566	91,50
QUEIMADURA	9.298	7	0,08	2.304	24,78	61	0,66	6.428	69,13
ENCAMINHAMENTO - PROCEDIMENTOS	8.532	50	0,59	1.167	13,68	331	3,88	5.752	67,42
INTOXICAÇÃO	7.932	16	0,20	655	8,26	286	3,61	6.570	82,83
NÃO DIGITADA	7.826	104	1,33	4.108	52,49	537	6,86	1.396	17,84
TENTATIVA DE SUICÍDIO	4.597	138	3,00	678	14,75	166	3,61	3.344	72,74
OUTROS	3.465	31	0,89	431	12,44	129	3,72	2.381	68,72
ALCOOLISMO	341	5	1,47	35	10,26	62	18,18	198	58,06
IGNORADA	191	3	1,57	89	46,60	10	5,24	56	29,32

Atendimentos por Natureza de Ocorrência**Figura 6.3** Atendimentos por Tipos de Ocorrências.

Ainda na Natureza "Sintoma", o Tipo de Ocorrência "Outros Sintomas"encontra-se com 241.088 casos, correspondendo a 24,27% do total da Natureza, registrando também a maior quantidade de óbitos, 1.335, conforme mostra a Figura 6.4. Identifica-se, desta maneira, uma necessidade de melhor classificação destes sintomas na recepção, ou posteriormente na informação do diagnóstico médico, para poder refletir em uma melhoria nas análises das informações.

Tipo_Ocorrencia	• Atendimentos	• Óbitos	• %Óbitos	• Internacao	• %Intern	• Transf	• %Transf	• Altas	• %Altas
SINTOMAS	993.165	4.805	0,48	78.852	7,94	36.521	3,68	800.997	80,65
Outros Sintomas	241.088	1.335	0,55	28.623	11,87	9.648	4,00	181.800	75,41
Mal Estar, Passando Mal	61.770	1.246	2,02	6.695	10,84	4.118	6,67	44.519	72,07
Desmaio	10.191	609	5,98	828	8,12	461	4,52	7.570	74,28
Falta de Ar, Cansaco, Dificuldade para Respirar	84.429	537	0,64	5.241	6,21	2.747	3,25	71.623	84,83
Vomito	40.298	162	0,40	2.520	6,25	879	2,18	34.616	85,90
Sangramento	15.230	105	0,69	1.838	12,07	1.047	6,87	11.321	74,33
Dor Abdominal	54.949	104	0,19	4.440	8,08	873	1,59	44.869	81,66
Frio, Tremor, Convulsao	6.261	92	1,47	1.409	22,50	151	2,41	4.158	66,41
Pressao Alta	26.903	69	0,26	1.661	6,17	1.028	3,82	21.280	79,10
Dor	70.919	58	0,08	7.776	10,96	3.831	5,40	56.171	79,20
Agonia, Nervosismo, Agitacao	35.145	56	0,16	1.385	3,94	4.931	14,03	24.487	69,67
Dor no Peito	16.124	56	0,35	1.074	6,66	1.320	8,19	12.395	76,87
Paralisia, Dormencia de Membros	1.935	46	2,38	437	22,58	70	3,62	1.258	65,01
Dores Generalizadas	12.292	39	0,32	679	5,52	142	1,16	9.994	81,30
Febre	65.689	34	0,05	1.973	3,00	683	1,04	59.896	91,18
Desorientacao, Confusao Mental, Alt. Comportamento	3.815	29	0,76	430	11,27	485	12,71	2.412	63,22
Inchaco, Edema	14.932	28	0,19	1.875	12,56	706	4,73	10.987	73,58
Fraqueza, Fadiga, Tontura	6.171	24	0,39	412	6,68	255	4,13	4.938	80,02
Dificuldade de Locomocao, Sem Coordenacao	734	23	3,13	264	35,97	34	4,63	371	50,54
Dor de Cabeça	51.213	22	0,04	1.911	3,73	1.018	1,99	42.954	83,87
Diarréia	13.868	20	0,14	607	4,38	178	1,28	12.249	88,33

Figura 6.4 Natureza de Ocorrência Sintoma.

O Tipo de Ocorrência "Desmaio", responsável pelo terceiro maior fator de óbito, refere-se, na maioria das vezes e através de identificação pelos gestores, a pacientes inconscientes (coma) com entrada na emergência, ou seja, com riscos de morte.

A Natureza "Encaminhamento-Doenças"refere-se ao atendimento nas emergências de pacientes procedentes de outras unidades de saúde, onde realizaram-se os primeiros atendimentos e são encaminhados aos hospitais. Os atendentes encontram dificuldades de identificação do Tipo de Ocorrência que originou o atendimento em outras unidades.

Visando a melhoria destas informações, uma ação de identificação necessita ser realizada visto que corresponde a maior quantidade de óbitos, 5.375 casos, e ao maior percentual, 1,90%. Em avaliação dos gestores, relatos sobre a forma de transporte entre as unidades pode ser considerado um dos fatores deste alto número de óbitos.

Outro indicador da necessidade de um tratamento melhor nas informações é a Natureza "Alcoolismo", que possui, estranhamente, cinco óbitos identificados.

Como ação importante de auxílio à gestão, o registro no sistema do diagnóstico médico

no final, ou durante, o atendimento, assim como o CID (Código Internacional de Doenças), permitirá uma avaliação mais determinante das causas dos atendimentos e dos óbitos.

Em uma outra análise, percebe-se que, quando realizado e informado um diagnóstico inicial no sistema, e retirando-se da quantidade total os Sintomas, tem-se as Naturezas Acidentes, Acidente de Trânsito e Agressão correspondendo a 28,52 % dos Tipos de Ocorrência.

A Ocorrência "Agressão", quando devidamente identificada, leva ao registro criminal, realizado pela unidade policial existente na recepção das emergências, cabendo-lhe o devido encaminhamento legal. Um estudo aprofundado sobre a relação entre as agressões identificadas no sistema e os registros policiais podem identificar se o Estado vem realizando o devido acompanhamento dos possíveis crimes, avaliando-se os índices de resolução dos casos.

A Natureza "Acidente de Trânsito" caracteriza-se como o quinto caso de atendimentos e quarto em óbitos, conforme apresentado na Figura 6.3. Nestes casos também estão associados aos registros implicações legais das empresas de seguros, envolvendo o seguro obrigatório dos veículos automotores. Por informação dos gestores, devido ao interesse financeiro envolvido nestas ocorrências, tratam-se dos registros entre os mais fidedignos nas emergências dos hospitais públicos. Isto somente comprova a possibilidade de gestão com o mesmo rigor e cuidado dos registros das outras ocorrências.

Por sua vez, as quantidades de óbitos da Natureza "Agressão" representam o maior percentual em relação aos atendimentos. Dos 49.027 atendimentos, 1.111 resultaram em óbito, correspondendo ao percentual de 2,27% sobre os atendimentos, tendo 6.913 casos, 14,10%, que levaram à internação, conforme mostra a Figura 6.5.

O Tipo de Natureza "Arma de Fogo" reflete a maior ocorrência desta Natureza, e o de maior letalidade. Dos 10.501 casos, 932, ou seja, 8,88%, levaram ao óbito e 3.983 casos, 37,93%, à internação.

Importante ressaltar que as agressões e os acidentes também refletem uma grande quantidade de internações, envolvendo um alto custo hospitalar, e com as restrições de leitos disponíveis na rede estadual. Mesmo em situações que não resultem em internação, os custos dos atendimentos das agressões podem ser considerados altos por geralmente envolverem um maior período de permanência na emergência, assim como procedimentos como suturas, imagens e imobilizações, entre outros.

Tipo_Ocorrencia	● Atendimentos	▼ Óbitos	● %Óbitos	● Internacao	● %Intern	● Transf	● %Transf	● Altas	● %Altas
ENCAMINHAMENTO - DOENÇAS	282.835	5.375	1,90	71.196	25,17	13.592	4,81	173.808	61,45
SINTOMAS	993.165	4.805	0,48	78.852	7,94	36.521	3,68	800.997	80,65
AGRESSÃO	49.027	1.111	2,27	6.913	14,10	1.188	2,42	36.134	73,70
Arma de Fogo	10.501	932	8,88	3.983	37,93	529	5,04	3.977	37,87
Arma Branca	5.610	92	1,64	1.636	29,16	179	3,19	3.298	58,79
Agressao Fisica	11.181	57	0,51	729	6,52	295	2,64	9.174	82,05
Agressao Objeto Contundente	3.993	20	0,50	282	7,06	93	2,33	3.338	83,60
Outras Agressoes	861	6	0,70	72	8,36	40	4,65	669	77,70
Agressao Animal - Mordida, Picada	16.777	4	0,02	202	1,20	43	0,26	15.609	93,04
Agressao Sexual - Estupro	95	0	0,00	9	9,47	8	8,42	61	64,21
Garrafada	3	0	0,00	0	0,00	1	33,33	2	66,67
Trauma na Face por Arma de Fogo	6	0	0,00	0	0,00	0	0,00	6	100,00
ACIDENTE DE TRÂNSITO	50.870	625	1,23	6.974	13,71	2.227	4,38	37.299	73,32
ACIDENTE	137.353	289	0,21	6.442	4,69	3.757	2,74	118.507	86,28
TENTATIVA DE SUICÍDIO	4.597	138	3,00	678	14,75	166	3,61	3.344	72,74
NÃO DIGITADA	7.826	104	1,33	4.108	52,49	537	6,86	1.396	17,84
ENCAMINHAMENTO - PROCEDIMENTOS	8.532	50	0,59	1.167	13,68	331	3,88	5.752	67,42
OUTROS	3.465	31	0,89	431	12,44	129	3,72	2.381	68,72
OCORRÊNCIA TRAUMATOLÓGICA	253.576	27	0,01	6.671	2,63	4.033	1,59	222.749	87,84
INTOXICAÇÃO	7.932	16	0,20	655	8,26	286	3,61	6.570	82,83
QUEIMADURA	9.298	7	0,08	2.304	24,78	61	0,66	6.428	69,13
ALCOOLISMO	341	5	1,47	35	10,26	62	18,18	198	58,06

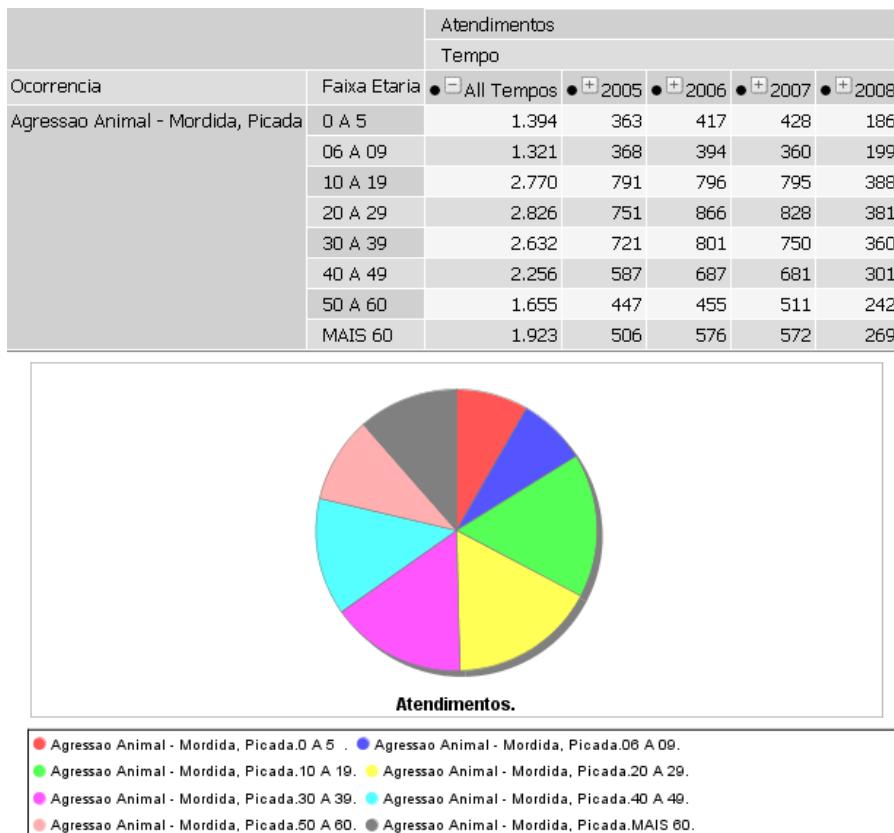
Figura 6.5 Atendimentos e Óbitos por Agressão.

As "Outras Agressões"correspondem a 1,76% dos casos, o que representa um baixo percentual em relação as agressões e uma melhor classificação realizada pelos atendentes, tornando as informações extraídas ainda mais representativas.

A "Agressão Animal-Mordida, Picada", respondem pela maior quantidade das agressões, 16.777, 34,22%, apresentando uma distribuição regular por faixas etárias, em um acompanhamento anual, conforme a Figura 6.6. Faz-se necessário uma sub-divisão nesta qualificação permitindo uma melhor identificação das ocorrências, viabilizando ações de prevenção dos órgãos públicos responsáveis.

A "Agressão Física", embora ambos com baixas taxas de óbitos, possuem grandes quantidades de atendimentos, 11.181, 22,81% das agressões, habilitando-se, geralmente, à ocorrências policiais. Em uma distribuição por sexo, o masculino corresponde a 70,64% e o feminino aos restantes 29,17%, conforme apresentado na Figura 6.7. A partir de 2005 até junho de 2008, a distribuição mostrou-se sem grandes alterações.

Podemos então concluir que, apesar das medidas recentemente adotadas nos diversos níveis governamentais visando tanto a inserção social quanto a diminuição da violência, houve pouca alteração no perfil dos pacientes referentes aos atendimentos refletidos nas emergências dos hospitais, que utilizam o sistema SAPE Emergência, em relação aos anos de 2005, 2006, 2007

**Figura 6.6** Agressão Animal por Faixa Etária - Anual.**Figura 6.7** Agressão Física por Sexo - Anual.

e primeiro semestre de 2008, conforme apresentado na Figura 6.8.

Um dos motivos indicados pelos gestores com relação aos índices mantidos pelo perfil destes pacientes, é ao chamado crime de proximidade, ou seja, cometido por pessoas conhecidas das vítimas e de difícil repressão policial.

Tipo_Ocorrencia	Tempo	● Atendimentos	● Óbitos	● %Óbitos	● Internacao	● %Intern	● Transf	● %Transf	● Altas	● %Altas
AGRESSÃO	+2005	13.413	298	2,22	1.979	14,75	324	2,42	9.898	73,79
	+2006	14.346	304	2,12	2.146	14,96	318	2,22	10.567	73,66
	+2007	14.130	320	2,26	1.855	13,13	358	2,53	10.501	74,32
	+2008	7.138	189	2,65	933	13,07	188	2,63	5.168	72,40
Agressao Fisica	+2005	3.077	16	0,52	195	6,34	79	2,57	2.558	83,13
	+2006	3.470	14	0,40	243	7,00	91	2,62	2.844	81,96
	+2007	3.096	16	0,52	205	6,62	82	2,65	2.526	81,59
	+2008	1.538	11	0,72	86	5,59	43	2,80	1.246	81,01
Agressao Objeto Contundente	+2005	1.067	9	0,84	67	6,28	17	1,59	902	84,54
	+2006	1.101	4	0,36	96	8,72	23	2,09	917	83,29
	+2007	1.218	2	0,16	79	6,49	43	3,53	1.016	83,42
	+2008	607	5	0,82	40	6,59	10	1,65	503	82,87
Arma Branca	+2005	1.482	19	1,28	459	30,97	49	3,31	873	58,91
	+2006	1.594	26	1,63	490	30,74	50	3,14	935	58,66
	+2007	1.645	24	1,46	454	27,60	51	3,10	988	60,06
	+2008	889	23	2,59	233	26,21	29	3,26	502	56,47
Arma de Fogo	+2005	2.980	250	8,39	1.185	39,77	156	5,23	1.141	38,29
	+2006	2.853	257	9,01	1.229	43,08	118	4,14	1.009	35,37
	+2007	3.031	275	9,07	1.044	34,44	159	5,25	1.197	39,49
	+2008	1.637	150	9,16	525	32,07	96	5,86	630	38,49

Atendimentos por Agressão - Anual

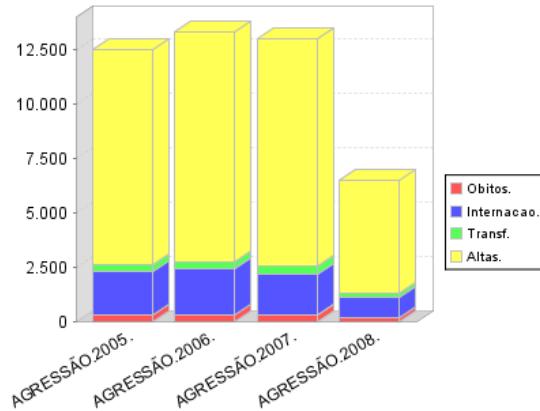


Figura 6.8 Atendimentos e Óbitos por Agressão - Anual.

Em uma outra análise nos Tipos de Ocorrência Agressão, Acidente de Trânsito e Acidente, em relação aos dias da semana, tanto os atendimentos quantos os óbitos e internações têm registros maiores no domingo e no sábado, conforme mostra a Figura 6.9.

Tipo_Ocorrencia	DiaDaSemana	Measures			
		● Atendimentos	▼ Óbitos	● Internacao	● Transf
+ AGRESSÃO	DOM	10.886	294	1.791	320
	SAB	8.671	208	1.355	268
	SEG	7.129	158	962	133
	SEX	6.088	143	823	134
	QUI	5.469	112	656	98
	TER	5.464	102	687	122
	QUA	5.320	94	639	113
+ ACIDENTE DE TRÂNSITO	DOM	11.445	149	1.721	518
	SAB	9.448	118	1.323	421
	SEX	6.442	83	867	294
	SEG	7.174	75	939	291
	TER	5.517	71	750	243
	QUA	5.394	71	672	230
	QUI	5.450	58	702	230
+ ACIDENTE	DOM	20.089	61	946	637
	SAB	20.573	47	960	646
	QUI	18.765	42	921	463
	TER	19.033	40	905	495
	SEG	21.306	37	861	519
	SEX	19.115	31	926	525
	QUA	18.472	31	923	472

Figura 6.9 Agressão e Acidentes por Dia da Semana.

Uma análise mais detalhada dos atendimentos e óbitos por dia da semana serão detalhados a seguir.

6.2.3 Tipos de Ocorrências por Hospital

Considerando-se os Tipos de Ocorrências associados aos hospitais, a Figura 6.10 demonstra que o perfil do paciente do Hospital da Restauração (HR) refere-se a Agressão, 66,71%, a Acidente de Trânsito, 52,26%, e a Acidente, 38,51%. Por tratar-se de um centro de referência em queimaduras, contando inclusive uma Unidade de Tratamento de Queimados (UTQ), realizaram-se atendimentos correspondentes a 96,27% do total de atendimentos nesta especialidade.

Em uma primeira análise, com esta quantidade de atendimentos dos Tipos de Ocorrência citados anteriormente, inclusive confirmando-se que a distribuição anual permanece proporcionalmente a mesma, o HR conta com cinco cirurgiões por plantão com 112.183 atendimentos enquanto no HAM possuem dois cirurgiões por plantão com 18.120 atendimentos, 13,91% do total do HR. Sugere-se um estudo sobre a distribuição de profissionais entre os hospitais, inclusive podendo ser considerado o histórico de atendimentos durante os dias da semana para a escolha do plantão historicamente com menos ocorrências. Torna-se interessante observar-se

Tipo_Ocorrencia	Hospital	● Atendimentos	● Altas	● %Altas	● Óbitos	● %obitos	● Internacao	● %Intern	● Transf	● %Transf
+ ACIDENTE	HAM	13.958	12.168	87,18	10	0,07	248	1,78	797	5,71
	HGV	39.477	34.404	87,15	55	0,14	1.946	4,93	1.345	3,41
	HOF	31.027	26.900	86,70	26	0,08	997	3,21	1.090	3,51
	HR	52.891	45.035	85,15	198	0,37	3.251	6,15	525	0,99
+ ACIDENTE DE TRÂNSITO	HAM	1.520	1.178	77,50	1	0,07	37	2,43	195	12,83
	HGV	16.540	12.789	77,32	73	0,44	1.863	11,26	963	5,82
	HOF	6.224	4.681	75,21	18	0,29	532	8,55	465	7,47
	HR	26.586	18.651	70,15	533	2,00	4.542	17,08	604	2,27
+ AGRESSÃO	HAM	2.635	1.903	72,22	77	2,92	58	2,20	424	16,09
	HGV	9.633	6.519	67,67	164	1,70	2.011	20,88	304	3,16
	HOF	4.053	2.355	58,11	141	3,48	905	22,33	315	7,77
	HR	32.706	25.357	77,53	729	2,23	3.939	12,04	145	0,44
+ ALCOOLISMO	HOF	341	198	58,06	5	1,47	35	10,26	62	18,18
+ CORPO ESTRANHO	HAM	8.223	7.670	93,27	0	0,00	86	1,05	276	3,36
	HGV	372	290	77,96	0	0,00	10	2,69	36	9,68
	HOF	450	389	86,44	0	0,00	8	1,78	27	6,00
	HR	6.874	6.217	90,44	1	0,01	151	2,20	21	0,31
+ ENCAMINHAMENTO - DOENÇAS	HAM	47.902	26.548	55,42	650	1,36	13.708	28,62	4.777	9,97
	HGV	109.367	78.056	71,37	930	0,85	17.741	16,22	4.667	4,27
	HOF	33.352	17.656	52,94	854	2,56	10.506	31,50	1.811	5,43
	HR	92.214	51.548	55,90	2.941	3,19	29.241	31,71	2.337	2,53
+ ENCAMINHAMENTO - PROCEDIMENTOS	HAM	2.522	1.595	63,24	15	0,59	571	22,64	172	6,82
	HGV	1.458	1.156	79,29	2	0,14	93	6,38	56	3,84
	HOF	1.261	882	69,94	7	0,56	166	13,16	19	1,51
	HR	3.291	2.119	64,39	26	0,79	337	10,24	84	2,55
+ IGNORADA	HAM	23	10	43,48	0	0,00	7	30,43	4	17,39
	HGV	13	4	30,77	0	0,00	6	46,15	0	0,00
	HOF	43	15	34,88	0	0,00	11	25,58	6	13,95
	HR	112	27	24,11	3	2,68	65	58,04	0	0,00
+ INTOXICAÇÃO	HAM	2.731	1.893	69,32	5	0,18	459	16,81	256	9,37
	HGV	1.601	1.524	95,19	0	0,00	5	0,31	11	0,69
	HOF	680	627	92,21	1	0,15	7	1,03	6	0,88
	HR	2.920	2.526	86,51	10	0,34	184	6,30	13	0,45
+ NÃO DIGITADA	HAM	424	137	32,31	3	0,71	121	28,54	33	7,78
	HGV	231	35	15,15	1	0,43	34	14,72	4	1,73
	HOF	6.549	1.166	17,80	100	1,53	3.532	53,93	498	7,60
	HR	622	58	9,32	0	0,00	421	67,68	2	0,32
+ OCORRÊNCIA TRAUMATOLÓGICA	HAM	7	5	71,43	0	0,00	1	14,29	1	14,29
	HGV	117.715	106.972	90,87	4	0,00	2.557	2,17	1.776	1,51
	HOF	82.122	71.524	87,09	20	0,02	2.661	3,24	1.618	1,97
	HR	53.732	44.248	82,35	3	0,01	1.452	2,70	638	1,19
+ OUTROS	HAM	1	1	100,00	0	0,00	0	0,00	0	0,00
	HGV	175	135	77,14	3	1,71	19	10,86	2	1,14
	HOF	2.292	1.579	68,89	20	0,87	271	11,82	117	5,10
	HR	997	666	66,80	8	0,80	141	14,14	10	1,00
+ QUEIMADURA	HAM	16	7	43,75	0	0,00	0	0,00	7	43,75
	HGV	41	28	68,29	0	0,00	0	0,00	11	26,83
	HOF	290	228	78,62	0	0,00	2	0,69	33	11,38
	HR	8.951	6.165	68,87	7	0,08	2.302	25,72	10	0,11
+ SINTOMAS	HAM	369.234	287.882	77,97	1.267	0,34	41.558	11,26	22.292	6,04
	HGV	163.592	140.441	85,85	1.032	0,63	8.711	5,32	2.685	1,64
	HOF	312.553	260.450	83,33	1.452	0,46	13.753	4,40	9.637	3,08
	HR	147.786	112.224	75,94	1.054	0,71	14.830	10,03	1.907	1,29
+ TENTATIVA DE SUICÍDIO	HAM	175	115	65,71	2	1,14	20	11,43	28	16,00
	HGV	527	374	70,97	13	2,47	66	12,52	47	8,92
	HOF	371	255	68,73	7	1,89	27	7,28	49	13,21
	HR	3.524	2.600	73,78	116	3,29	565	16,03	42	1,19

Figura 6.10 Tipo de Ocorrência por Hospital.

também a demanda de alocação de profissionais mais qualificados em suas especialidades nos plantões de maiores índices de atendimentos e, principalmente, óbitos.

Observa-se também que as classificações realizadas na Natureza "Sintoma" encontram-se em maior volume no Hospital Otávio de Freitas (HOF) e no Hospital Agamenon Magalhães (HAM), o que sugere a necessidade de um treinamento maior aos atendentes na classificação do Tipo de Ocorrência. Ação semelhante faz-se necessário na identificação da Natureza "Encaminhamento - Doenças" no Hospital Getúlio Vargas (HGV) e no Hospital da Restauração (HR).

A grande quantidade de atendimentos na Natureza "Ocorrência Traumatológica" no HGV e no HOF sugere a necessidade de um tratamento mais unificado nesta classificação entre os hospitais, apesar da diminuição dos casos em uma avaliação anual.

Identifica-se também no HOF uma grande quantidade de atendimentos "Outros", 66,15% em relação a mesma classificação nos outros hospitais, assim como "Não Digitada" com 83,68%, sugerindo-se, novamente, um treinamento específico. Os atendentes do HAM e o HGV mostram-se mais aplicados na identificação dos Tipos de Ocorrência.

Atendimentos da Natureza "Corpo Estranho" no HAM, conforme a Figura 6.11, refletem a especialização em Otorrinolaringologia contabilizando, respectivamente, Corpo Estranho no Aparelho Respiratório (Nariz e Garganta), 3.642 ocorrências e Corpo Estranho no Ouvido, com 4.017, totalizando 93,14% dos atendimentos desta natureza no hospital. Estes valores refletem uma resposta à política de especialização do HAM proposto pela SES-PE.

O baixo índice de transferências e alto número de altas neste Tipo de Ocorrência indicam alto grau de resolução associado aos profissionais envolvidos.

No total desta mesma Natureza, o atendimento do HR encontra-se uma quantidade de 1.682, 46,18% do HAM. Segundo os gestores envolvidos, estes atendimentos ocorrem devido ao HR possuir a única urgência em Endoscopia Digestiva dos hospitais públicos.

A incidência de "Corpo Estranho - Outros Locais", segundo tipo de ocorrências no HGV e HOF indicam a necessidade de uma melhor especificação desta ocorrência, facilitando a identificação e possibilitando uma melhor análise. No caso do HR, grande parte dos 1.259 casos referem-se, segundo os gestores, a Broncoscopia Rígida, retirada de corpos estranhos da traquéia e do pulmão, cujo aparelho de alto custo encontra-se disponível unicamente, em todo Estado, na emergência do HR.

Nestas situações, não se identificaram necessidade de realocação de especialidades e/ou especialistas entre os hospitais.

Tipo_Ocorrencia	HAM	HGV	HOF	HR
	Measures	Measures	Measures	Measures
● Atendimentos	● Atendimentos	● Atendimentos	● Atendimentos	
+ ACIDENTE	13.958	39.477	31.027	52.891
+ ACIDENTE DE TRÂNSITO	1.520	16.540	6.224	26.586
+ AGRESSÃO	2.635	9.633	4.053	32.706
+ ALCOOLISMO			341	
+ CORPO ESTRANHO	8.223	372	450	6.874
Corpo Est. Ap.Genital/Excretor (penis, vagina, anus)	8	9	53	64
Corpo Est. Aparelho Respiratorio (Nariz, Garganta)	3.642	89	99	1.682
Corpo Est. no Aparelho Digestivo (Ingestao)	291	44	97	2.872
Corpo Est. no Olho	12	1	1	943
Corpo Est. no Ouvido	4.017	6	52	54
Corpo Estranho - Outros Locais	253	222	148	1.259
Corpo Estranho - ver observacao		1		
+ ENCAMINHAMENTO - DOENÇAS	47.902	109.367	33.352	92.214
+ ENCAMINHAMENTO - PROCEDIMENTOS	2.522	1.458	1.261	3.291
+ IGNORADA	23	13	43	112
+ INTOXICAÇÃO	2.731	1.601	680	2.920
+ NÃO DIGITADA	424	231	6.549	622
+ OCORRÊNCIA TRAUMATOLÓGICA	7	117.715	82.122	53.732
+ OUTROS	1	175	2.292	997
+ QUEIMADURA	16	41	290	8.951

Figura 6.11 Ocorrências Otorrinolaringológicas por Hospital.

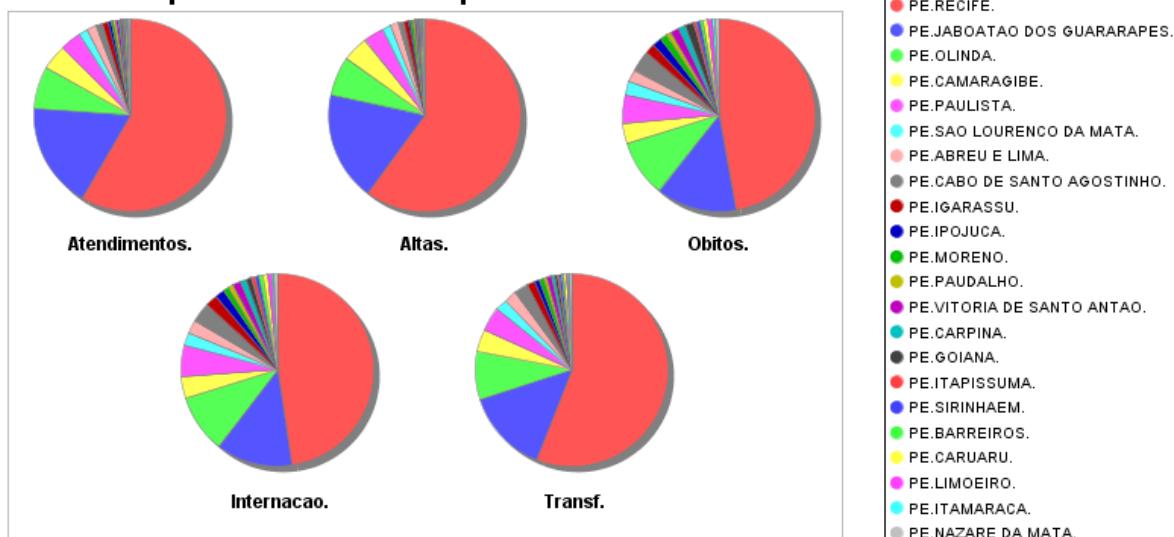
6.2.4 Domicílio dos Pacientes

Considerando-se uma análise por Estado, como pode-se esperar, Pernambuco corresponde a 99,44% dos atendimentos. Na seqüência, dos 10.204 atendimentos fora do estado, ou seja, 0,56%, aparecem os estados vizinhos Paraíba com 13,13% da quantidade de fora de Pernambuco, Alagoas com 6,62% e Bahia com 4,37%. Devido ao pequeno percentual de atendimentos com estas características entre 2005 e junho de 2008, não há grande representatividade para um aprofundamento da análise.

Entretanto, considerando-se as cidades de Pernambuco de domicílio dos pacientes e levando-se em conta a descentralização proposta pelo Ministério da Saúde, visando facilitar a análise, selecionamos os municípios que possuem mais de 3.000 atendimentos, no período envolvendo 2005 a junho de 2008, tendo-se um total de 22 municípios, correspondendo a 94,31% dos atendimentos, conforme apresentado na Figura 6.12. O Recife representa 54,90% dos atendimentos nos quatro hospitais, nenhum deles municipal, diferentemente de outras capitais do país.

Importante realizar também uma comparação em relação aos óbitos, visto que nestes casos tendem a ser transferências realmente necessárias, no sentido da gravidades da ocorrência.

Paciente	• Atendimentos	• Altas	• %Altas	• Obitos	▼ %obitos	• Internacao	• %Intern	• Transf	• %Transf
GOIANA	4.678	2.417	51,67	142	3,04	1.502	32,11	268	5,73
LIMOEIRO	3.106	1.546	49,77	90	2,90	1.027	33,07	180	5,80
CARPINA	5.302	2.931	55,28	123	2,32	1.560	29,42	282	5,32
CARUARU	3.125	1.257	40,22	72	2,30	1.272	40,70	232	7,42
VITORIA DE SANTO ANTÃO	6.981	3.652	52,31	144	2,06	2.006	28,74	482	6,90
NAZARE DA MATA	3.003	1.754	58,41	55	1,83	786	26,17	172	5,73
IPOJUCA	8.654	4.851	56,06	154	1,78	2.476	28,61	526	6,08
CABO DE SANTO AGOSTINHO	20.695	12.213	59,01	354	1,71	5.051	24,41	1.343	6,49
BARREIROS	3.390	1.794	52,92	55	1,62	1.020	30,09	216	6,37
SIRINHAEM	3.765	2.230	59,23	61	1,62	973	25,84	197	5,23
ITAMARACA	3.004	1.876	62,45	46	1,53	671	22,34	172	5,73
MORENO	8.035	5.364	66,76	115	1,43	1.440	17,92	415	5,16
PAUDALHO	7.210	5.015	69,56	88	1,22	1.299	18,02	306	4,24
ITAPISSUMA	4.282	2.825	65,97	51	1,19	895	20,90	197	4,60
IGARASSU	15.908	10.703	67,28	175	1,10	3.089	19,42	744	4,68
SAO LOURENCO DA MATA	26.818	20.218	75,39	233	0,87	3.203	11,94	1.119	4,17
OLINDA	121.237	91.615	75,57	973	0,80	14.904	12,29	4.563	3,76
PAULISTA	61.434	45.422	73,94	485	0,79	8.060	13,12	2.372	3,86
ABREU E LIMA	25.826	19.005	73,59	199	0,77	3.512	13,60	1.211	4,69
RECIFE	1.001.822	824.750	82,33	4.741	0,47	72.721	7,26	32.018	3,20
CAMARAGIBE	72.576	59.533	82,03	343	0,47	5.662	7,80	2.139	2,95
JABOATAO DOS GUARARAPES	304.245	252.321	82,93	1.346	0,44	19.605	6,44	8.106	2,66

Tipos de Atendimentos por Cidades de PE**Figura 6.12** Tipos de Atendimentos por Cidades de Pernambuco.

Nos casos de maiores índices, como Goiana com 3,04%, Limoeiro, 2,90%, Carpina, 2,32%, Caruaru, 2,30% e Vitória de Santo Antão, 2,06%, somente citando alguns, indicam que as transferências atendem a um critério de risco de morte do paciente. Estas informações possibilitem um estudo mais detalhado, por parte da SES-PE, na tentativa de viabilização de um pacto em relação as quantidade de pacientes a serem transferidos entre as cidades e os hospitais estaduais.

Em um detalhamento destas 22 cidades em quantidade de atendimentos por hospital, conforme Figuras 6.13 e 6.14, identificam-se que pacientes de Jaboatão dos Guararapes, segunda cidade em atendimentos, representam 70,01% das ocorrências do Hospital Otávio de Freitas.

Percebe-se também o HR como sendo o destino de grande parte dos pacientes da maioria das cidades relacionadas, exceto Jaboatão, citada anteriormente, Camaragibe, com 64,75% no HGV, São Lourenço da Mata, com 45,38% no HGV, Moreno, com 44,99% no HOF e Paudalho, com 44,01% no HGV.

O Governo Estadual vem realizando obras de construção de mais dois hospitais nas cidades periféricas da região metropolitana do Recife, o Hospital Miguel Arraes em Paulista, próximo a Abreu e Lima, e o Hospital Dom Hélder Câmara no Cabo de Santo Agostinho. Em atendendo média e alta complexidade, através da identificação das quantidades de pacientes das cidades circunvizinhas, podem vir a servir para direcionar pacientes das dez primeiras cidades em quantidade de atendimentos no Estado, minimizando os atendimentos nos hospitais localizados na cidade do Recife.

Outro fator a ser analisado é o decréscimo anual de atendimentos por parte do HR de pacientes de praticamente todas as cidades analisadas, conforme apresentado na Figura 6.15, apesar do aumento de profissionais nesta unidade de saúde. Mesmo com esta tendência não há reflexo nos meios de comunicação de melhorias nos atendimentos da emergências deste hospital.

A presença de pacientes domiciliados em Jaboatão dos Guararapes com atendimentos no HOF vem decrescendo, principalmente a partir de 2007, apesar de ainda corresponder a 63,06% em 2007. Pelas projeções do primeiro semestre de 2008 as quantidades no ano de 2008 tendem a manterem-se semelhantes ao mesmo percentual de 2007.

Em Olinda a distribuição dos atendimentos no HR vem decrescendo seguidamente de 2005 a 2007, e pelas projeções de 2008, mantendo-se com os mesmos valores de 2007.

	Atendimentos			
	Hospital			
Paciente	● HAM	● HGV	● HOF	● HR
ABREU E LIMA	9.206	5.695	1.922	9.003
BARREIROS	560	639	1.076	1.115
CABO DE SANTO AGOSTINHO	3.386	3.699	6.644	6.966
CAMARAGIBE	8.576	46.992	4.353	12.655
CARPINA	1.136	1.425	1.048	1.693
CARUARU	842	410	485	1.388
GOIANA	1.006	1.025	651	1.996
IGARASSU	3.742	3.464	2.154	6.548
IPOJUCA	1.373	1.572	2.984	2.725
ITAMARACA	673	614	589	1.128
ITAPISSUMA	1.022	945	812	1.503
JABOATAO DOS GUARARAPES	20.560	28.158	213.015	42.512
LIMOEIRO	613	833	451	1.209
MORENO	1.090	1.310	3.615	2.020
NAZARE DA MATA	571	831	734	867
OLINDA	40.025	16.941	6.008	58.263
PAUDALHO	1.072	3.173	1.075	1.890
PAULISTA	18.447	11.403	4.117	27.467
RECIFE	306.672	295.583	203.320	196.247
SAO LOURENCO DA MATA	4.064	12.170	4.205	6.379
SIRINHAEM	639	727	1.162	1.237
VITORIA DE SANTO ANTÃO	1.374	1.130	1.743	2.734

Figura 6.13 Atendimentos por Cidades e Hospitais de Pernambuco.

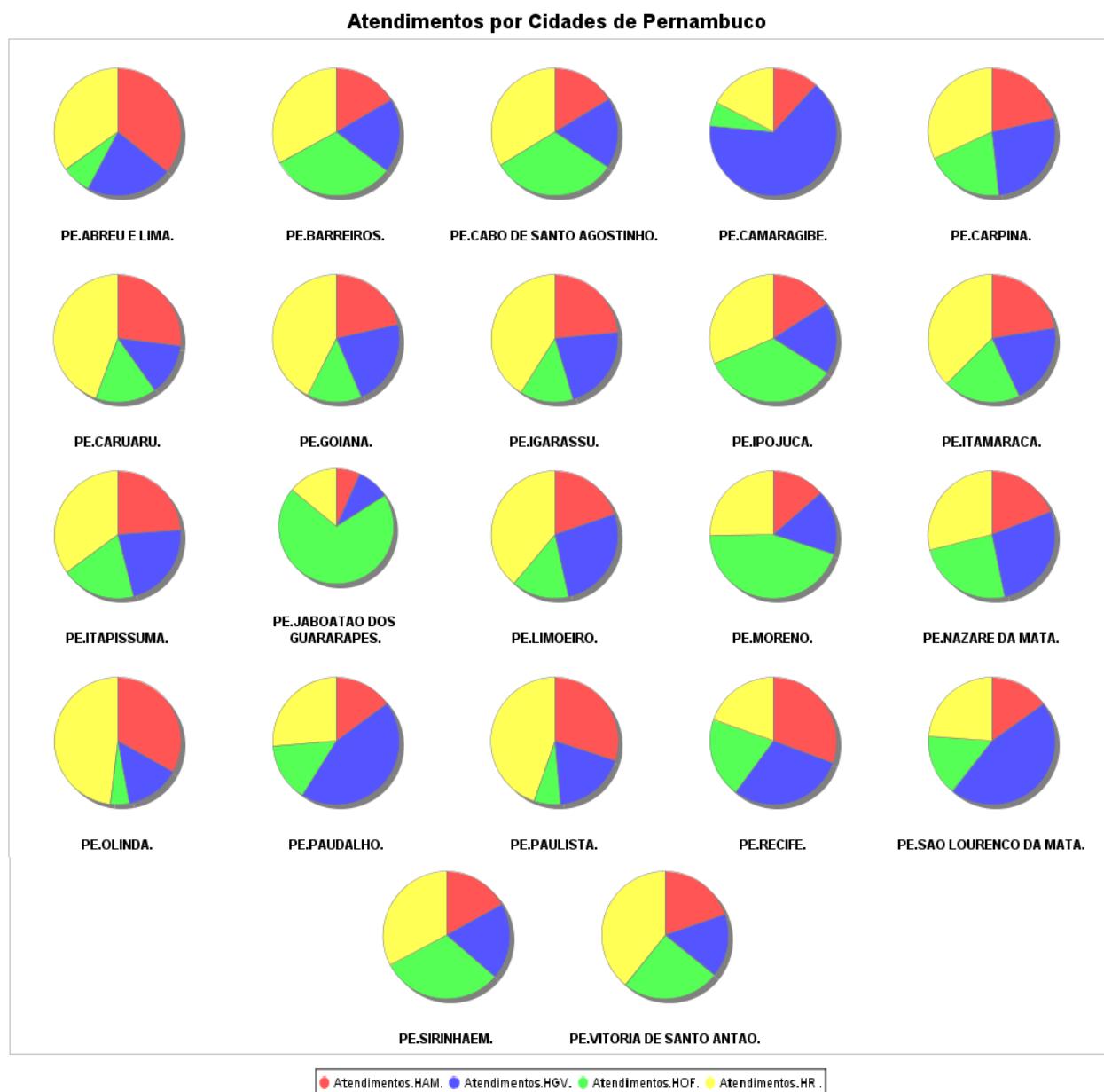


Figura 6.14 Atendimentos por Cidades e Hospitais de Pernambuco.

Paciente	Tempo	● HAM	● HGV	● HOF	● HR	Paciente	Tempo	● HAM	● HGV	● HOF	● HR
ABREU E LIMA	+2005	2.873	1.478	460	3.159	ITAPISSUMA	+2005	335	226	190	587
	+2006	2.820	1.723	578	2.629		+2006	316	312	288	416
	+2007	2.644	1.762	537	2.200		+2007	239	261	201	316
	+2008	869	732	347	1.015		+2008	132	146	133	184
BARREIROS	+2005	140	147	361	385	JABOATAO DOS GUARARAPES	+2005	5.894	6.539	71.366	14.003
	+2006	178	188	318	343		+2006	6.111	7.603	74.103	11.840
	+2007	191	216	271	249		+2007	5.965	9.524	45.339	11.073
	+2008	51	88	126	138		+2008	2.590	4.492	22.207	5.596
CABO DE SANTO AGOSTINHO	+2005	1.020	871	1.909	2.540	LIMOEIRO	+2005	185	188	126	389
	+2006	998	1.027	2.040	1.765		+2006	177	312	122	343
	+2007	955	1.235	1.822	1.733		+2007	163	234	121	300
	+2008	413	566	873	928		+2008	88	99	82	177
CAMARAGIBE	+2005	2.447	11.775	1.374	4.356	MORENO	+2005	317	314	1.072	702
	+2006	2.666	15.523	1.309	3.512		+2006	303	389	1.156	555
	+2007	2.397	13.646	1.090	3.127		+2007	331	403	905	478
	+2008	1.066	6.048	580	1.660		+2008	139	204	482	285
CARPINHA	+2005	291	342	258	479	NAZARE DA MATA	+2005	165	179	219	278
	+2006	313	447	325	502		+2006	142	248	201	240
	+2007	358	436	289	457		+2007	182	296	199	211
	+2008	174	200	176	255		+2008	82	108	115	138
CARUARU	+2005	293	123	133	425	OLINDA	+2005	11.192	4.165	1.714	19.273
	+2006	215	93	120	305		+2006	11.770	5.151	1.836	17.293
	+2007	221	131	139	427		+2007	11.727	5.274	1.620	14.462
	+2008	113	63	93	231		+2008	5.336	2.351	838	7.235
GOIANA	+2005	295	214	148	677	PAUDALHO	+2005	316	792	329	600
	+2006	299	291	192	544		+2006	326	1.041	308	484
	+2007	309	365	192	531		+2007	282	948	258	509
	+2008	103	155	119	244		+2008	148	392	180	297
IGARASSU	+2005	1.102	854	457	2.341	PAULISTA	+2005	5.353	2.913	1.095	9.076
	+2006	1.171	1.103	654	1.846		+2006	5.115	3.493	1.304	7.869
	+2007	1.036	1.044	725	1.568		+2007	5.529	3.422	1.077	7.054
	+2008	433	463	318	793		+2008	2.450	1.575	641	3.468
IPOJUCA	+2005	415	360	1.068	931	RECIFE	+2005	88.593	79.394	67.996	68.229
	+2006	404	478	904	706		+2006	92.001	91.706	70.273	56.331
	+2007	368	490	668	683		+2007	87.875	85.077	44.185	47.739
	+2008	186	244	344	405		+2008	38.203	39.406	20.866	23.948
ITAMARACA	+2005	196	156	167	410	SAO LOURENCO DA MATA	+2005	1.301	2.869	1.474	2.154
	+2006	190	172	178	318		+2006	1.093	4.113	1.267	1.763
	+2007	205	190	165	250		+2007	1.167	3.689	1.016	1.624
	+2008	82	96	79	150		+2008	503	1.499	448	838

Figura 6.15 Atendimentos Anuais por Cidade e Hospital.

6.2.5 Internações Hospitalares

Os procedimentos associados às internações são de grande custo financeiro para o Estado. Apesar do sistema tratar dos atendimentos nas emergências, no momento da internação é realizada a baixa no atendimento. Como mostra a Figura 6.16, o HR e o HAM, com 14,18% e 12,66% dos atendimentos, respectivamente, são os hospitais com mais internamentos, principalmente devido a quantidade de leitos e as especialidades envolvidas.

Hospital	Measures		
	● Atendimentos	▼ Internacao	● %Intern
Todos os Hospitais	1.824.927	186.770	10,23
HR	433.206	61.421	14,18
HAM	449.371	56.874	12,66
HGV	460.742	35.062	7,61
HOF	481.608	33.413	6,94

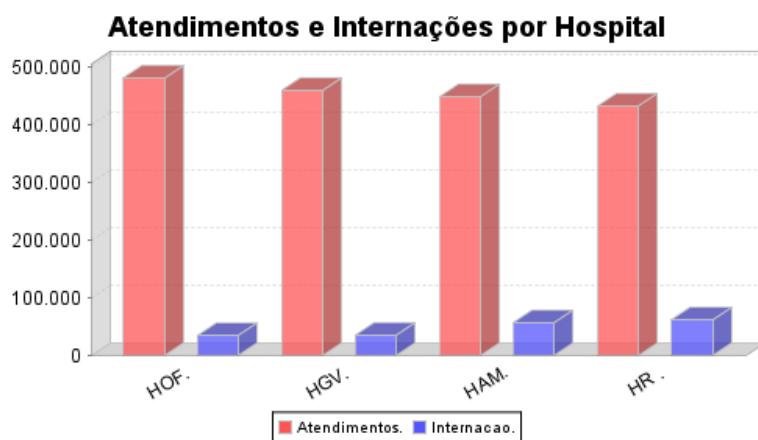


Figura 6.16 Atendimentos e Internações por Hospital.

Em uma análise de internações hospitalares por ano, Figura 6.17, observa-se uma regularidade entre os anos de 2005 as projeções para junho 2008, exceto no Hospital da Restauração onde do ano de 2006 para 2007 houve uma redução de 29,99%, mantendo-se no mesmo nível em projeções para 2008. Uma melhor análise pode ser realizada pela SES-PE para a identificação das causas desta redução.

No Apêndice C encontra-se um detalhamento das Ocorrências associadas aos internamentos com os percentuais referentes aos atendimentos destas mesmas Ocorrências.

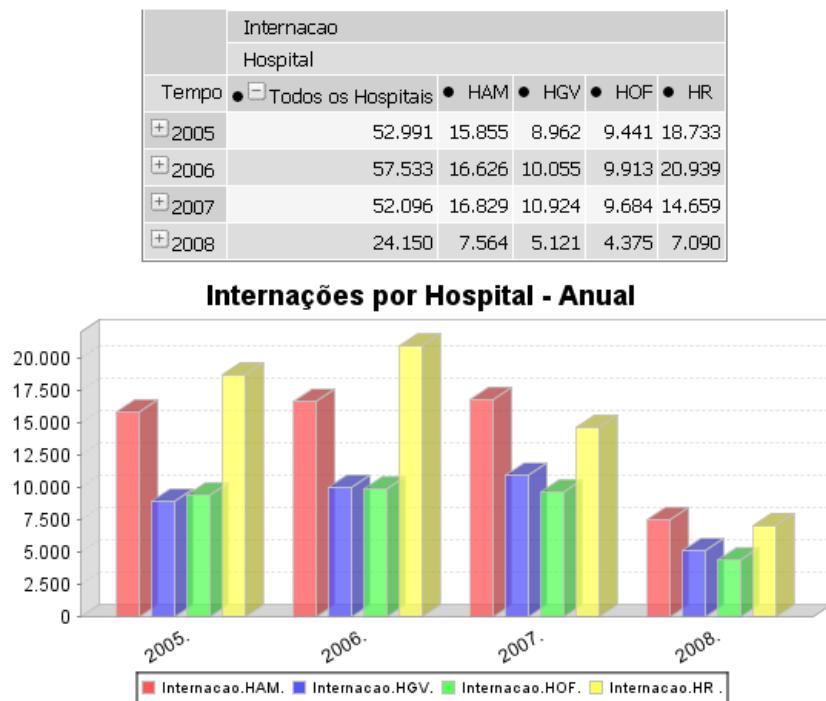


Figura 6.17 Internações por Hospital- Anual.

6.3 Considerações Finais

As possibilidade de análises claramente não se limitam às avaliações relatadas. Muitas outras análises podem decorrer quando da disponibilização da ferramenta para gestores de perfis diferenciados. Entretanto, a necessidade de documentar cada simulação neste trabalho limita bastante a riqueza e a versatilidade da ferramenta. Em algumas análises, resultados temporários ou pesquisas mais refinadas não puderam ser incorporadas à documentação, sendo somente relatadas suas conclusões.

Em outras análises, como a de Médico Liberador dos pacientes atendidos, 78,05% não encontram-se informados no sistema transacional, o que compromete o resultado da avaliação optando-se pela não apresentação.

De maneira semelhante, a Forma de Transporte do paciente quando da entrada nas emergências, "Outros"correspondem a 46,41%, "Não Digitado"23,54% e "Outros Veículos"13,40%, totalizando 83,35%, percentual considerado alto para ações conclusivas, sendo sugerida uma sub-divisão das classificações por parte da SES-PE.

Várias das análises evidenciam a necessidade de treinamento específico e uniforme entre os atendentes dos hospitais, sugerindo-se também um revezamento entre os atendentes mais qualificados, inclusive entre as unidades de saúde, na tentativa de disseminação e conformidade

do conhecimento.

Com relação à performance da ferramenta, de uma maneira geral, considerou-se o tempo de resposta satisfatório, porém, pelo ambiente disponibilizado, não foram realizados testes de concorrência de acesso. Nas consultas mais lentas, tabelas de agregação foram geradas resultando em ganhos consideráveis de performance, identificando-se, desta forma, como sendo um problema associado mais ao SGBD que propriamente à ferramenta de BI.

Em algumas consultas as operações de "*drill-through*" foram realizadas, permitindo acesso à menor granularidade sem maiores dificuldades. Esta característica agradou bastante aos gestores pois em alguns casos aparentemente incoerentes encontraram-se justificativas visualizando-se nível mais baixo da informação.

Outra característica importante refere-se ao arquivamento de cada uma das consultas o que facilita a recuperação de uma determinada avaliação, possibilitando aprofundamento nas análises, partindo-se de situações anteriores.

Na avaliação dos gestores a ferramenta apresenta-se perfeitamente viável do ponto de vista do negócio, sendo considerada estratégica sua implantação no ambiente da Secretaria de Saúde do Estado de Pernambuco.

CAPÍTULO 7

Conclusão

A motivação deste trabalho surgiu da percepção da disponibilidade de informações integradas servirem como insumo à decisão da gerência estratégica da Secretaria de Saúde do Estado de Pernambuco (SES-PE).

Identificando-se as emergências dos hospitais públicos e o sistema que realiza os registros dos pacientes, a implantação de um *Data Warehouse* (DW), utilizando-se dos dados reais, e a implementação de uma solução de *Business Intelligence* (BI), utilizando-se de software livre, o Pentaho, mostrou-se como um grande desafio, tornando sua implementação e implantação perfeitamente viável e de grande aceitação pelos gestores envolvidos.

7.1 Contribuições

Com a identificação da importância das fontes de dados, foram implementadas e executadas as tarefas de conformidade dos dados existentes, eliminando-se a duplicidade que não encontravam nenhum tipo de tratamento nas diversas emergências dos hospitais públicos. Em seguida, com a elaboração das técnicas de transformação com alguns algoritmos envolvidos, automatizam-se todo o processo de extração, transformação e carga necessário para a unificação da base de dados.

Após avaliadas algumas alternativas, houve a implementação da solução de *Business Intelligence*, demonstrando-se o funcionamento de suas diversas ferramentas, em software livre, comprovada a sua eficiência e eficácia em uma utilização voltada à realidade prática.

O modelo de dados dimensional disponibilizado apresentou-se viável e atendendo aos diversos requisitos elencados pelos gestores que especificaram a necessidade das informações.

Neste novo contexto, as análises apresentadas serviram de base para a comprovação de algumas percepções que permeavam o conhecimento de alguns gestores envolvidos nos atendimentos e que careciam de fatos para comprovação das afirmações, permitindo um maior respaldo nas tomada de decisões.

Por outro lado, algumas conclusões extraídas mostraram uma realidade diferente destas percepções, o que também fornece suporte a um redirecionamento da gestão.

Pelos sistemas pesquisados e apresentados, o pioneirismo identificado deste trabalho associado às emergências dos hospitais públicos torna evidente a importância do estudos dos dados das emergências que funcionam atualmente como uma importante, senão a principal, porta de entrada de resolução dos problemas dos pacientes do Sistema Único de Saúde no Brasil. Esta característica indica fortemente a necessidade de maior resolução nos atendimentos de atenção básica de saúde, de responsabilidade dos municípios, em uma grande ação compactuada das Secretarias Municipais de Saúde com a Secretaria Estadual, proposta pelo Ministério da Saúde.

Paralelamente às novas avaliações, torna-se claro que uma melhor qualidade das informações, com treinamento e padronização das ocorrências digitadas pelos atendentes, refletirão diretamente em uma melhor identificação do perfil dos principais hospitais públicos de Pernambuco.

Este ganho torna-se claro na avaliação dos gestores públicos entrevistados que confirmaram a utilidade da solução proposta para uma melhor tomada de decisão gerencial voltadas às políticas de saúde pública.

Com esta solução de BI, o gerenciamento da SES-PE pode alcançar outro patamar na utilização da informação. Informações que antes encontravam-se distribuídas, com formatações distintas e num nível de granularidade inadequado à boa decisão gerencial encontram-se agora consolidadas e unificadas sobre as dimensões de espaço e tempo, proporcionando a verificação de estatísticas que não estavam antes disponíveis e que podem ajudar a traçar um retrato mais acurado da saúde do Estado.

Ensejamos com este trabalho, portanto, incrementar agilidade, segurança, rapidez e eficiência a um só tempo, na decisão gerencial sobre a saúde do Estado. Cabe lembrar ainda que a solução proposta é bastante flexível: muitas estatísticas consolidadas diferentes podem ser conseguidas com sua utilização, e mesmo consultas particulares podem detectar um súbito aumento de determinadas ocorrências, correlacionando-o com determinadas variáveis, que podem representar a causa deste incremento, possibilitando sua rápida solução em benefício de toda a sociedade.

7.2 Considerações Finais

Após a universalização do direito público à saúde, garantido pela Constituição de 1988 através da criação do Sistema Único de Saúde (SUS), houve uma rápida democratização do atendimento que não recebeu o devido acompanhamento em sua infra-estrutura. É de conhecimento geral a precária situação da saúde pública no Brasil, principalmente nos atendimentos realiza-

dos nas emergências dos hospitais públicos.

A despeito das necessidades de alocação dos mais variados tipos de recursos públicos, a Tecnologia da Informação vem conquistando um papel de destaque tornando-se cada vez mais estratégica na medida em que possibilita a percepção da real situação do problema e viabiliza uma melhor aplicação dos recursos alocados.

Esta solução de BI, implementada e de fácil usabilidade, identifica e viabiliza o suporte à decisão nos níveis estratégicos da gestão pública de saúde. A inexistência do custo de aquisição, fator na maioria das vezes determinante nas avaliações de um produto, apresenta-se como um incentivo à implantação. Os investimentos restringem-se unicamente a treinamento e qualificação de pessoal na utilização das ferramentas.

De comprovada eficiência nas diversas avaliações realizadas pelos gestores envolvidos e fornecendo suporte à decisão em diversos níveis de gestão, a dúvida que surge, como sempre ocorre em situações semelhantes, é que mesmo após disponibilizada, o instrumento sofra problemas de descontinuidade. Resta-nos a certeza da tarefa cumprida em mais uma tentativa de melhoria dos serviços públicos na eterna busca de resgate da cidadania.

7.3 Trabalhos Futuros

Durante o desenvolvimento da solução, verifica-se a oportunidade de desenvolvimento de algumas propostas, a saber:

- Implementação de um Sistema de Informações Geográficas, ou *Geographic Information System* (GIS), integrado com a solução de BI melhorando visualmente a identificação das ocorrências;
- Área de apresentação com a utilização de *dashboard*, ou painel de controle, com a identificação e implementação de indicadores;
- Reformulação do processo de Extração, Transformação e Carga para a inclusão dos anos de 2000 a 2008;
- Melhoria no processo de Extração, Transformação e Carga possibilitando uma atualização com carga diferencial, identificando-se as alterações e inclusões (*Slowly Changing Dimension* tipo 2);
- Disponibilização para os usuários da ferramenta de relatórios *ad-hoc* do Pentaho, utilizando-se do Metadata Editor;

- Melhorias na performance com a implementação de tabelas agregadas utilizando-se do *Aggregation Designer*;
- Inclusão dos dados no *Data Warehouse* dos hospitais públicos dos municípios de Goiana, com o sistema transacional implantado em julho de 2008 e de Nazaré da Mata, implantado em outubro de 2008;
- Rastreamento entre as ocorrências criminais informados no sistema e os Boletins de Ocorrência registrados pelas unidades policiais nos hospitais, avaliando-se o índice de resolução do Estado;
- Integração das bases de dados do SAPE Internação e SAPE Faturamento ao DW, possibilitando uma maior riqueza nas informações financeiras envolvendo os hospitais públicos do Estado de Pernambuco.

APÊNDICE A

Apêndice A

Neste Apêndice são apresentados o Modelo Entidade Relacionamento do sistema transacional que atende às emergências dos quatro maiores hospitais públicos do Estado de Pernambuco, o Hospital da Restauração (HR), Hospital Agamenon Magalhães (HAM), Hospital Getúlio vargas (HGV) e Hospital Otávio de Freitas (HOF).

A.1 Modelo Entidade Relacionamento do Sistema SAPE-Emergência

Nesta seção apresenta-se o Modelo Entidade Relacionamento do Sistema SAPE-Emergência, Figura A.1.

APÊNDICE A APÊNDICE A

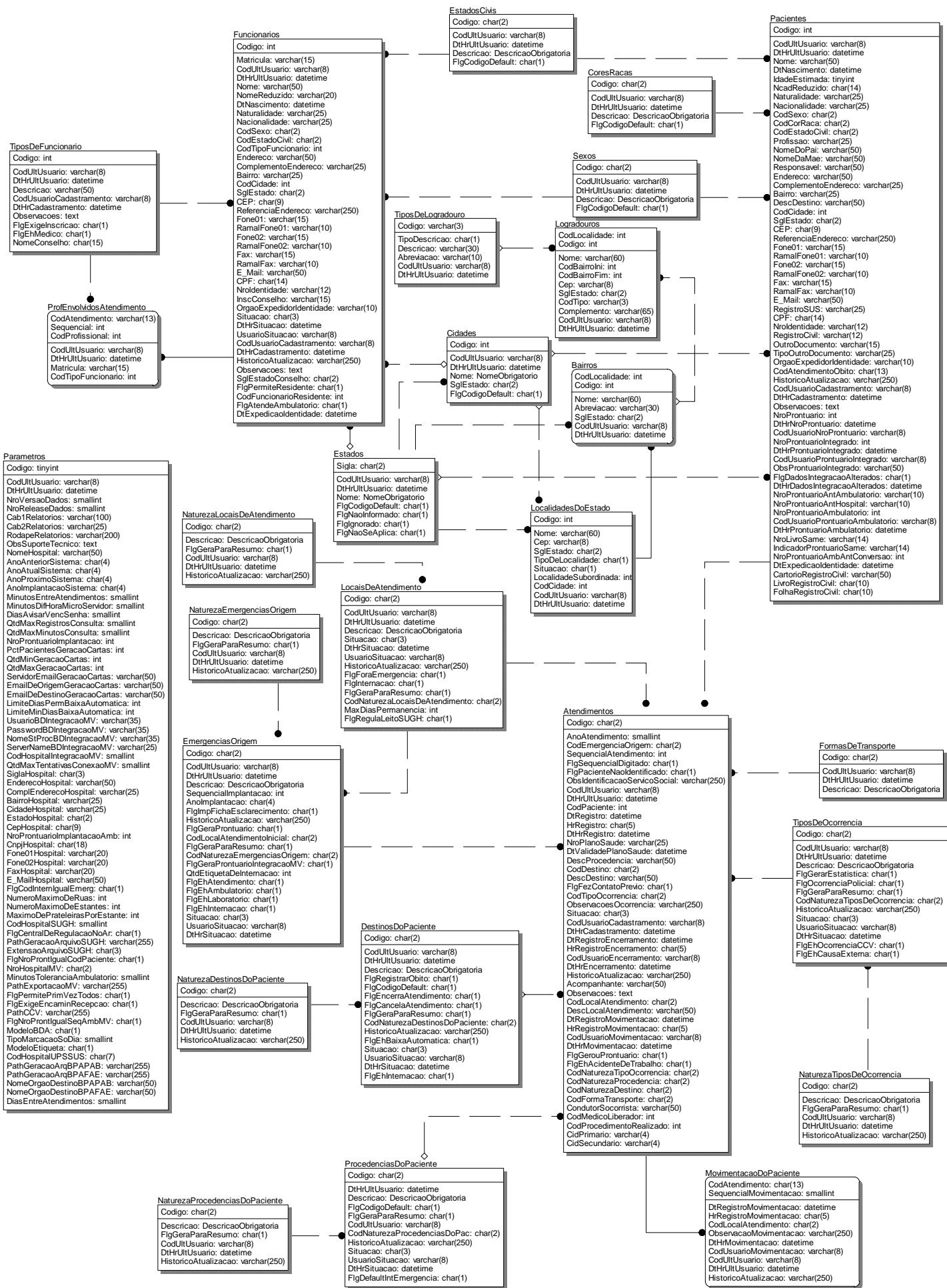


Figura A.1 Modelo Entidade Relacionamento do SAPE Emergência.

A.2 Dicionário de Dados das Tabelas de Dimensão

Nesta seção encontram-se detalhadas as colunas e os tipos de dados das tabelas de dimensão e de fato que compõem ao Modelo Multidimensional do *Data Warehouse*.

A.2.1 Dimensão Paciente

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	PACIENTE_SK	int		Não
	PACIENTE_NU_VERSAO	smallint		Não
	PACIENTE_DT_ETL	datetime		Não
	NCADREDUZIDO	char	13	Não
	DTNASCIMENTO	date		Sim
	IDADEESTIMADA	smallint		Sim
	CODSEXO	char	2	Não
	DESCRICAOSEXO	varchar	50	Não
	CODESTADOCIVIL	char	2	Não
	DESCRICAOSTADOCIVIL	varchar	50	Não
	ENDERECO	varchar	50	Sim
	COMPLEMENTOENDERECO	varchar	25	Sim
	BAIRRO	varchar	25	
	CODCIDADE	int		Não
	CIDADE	varchar	25	Não
	SGLESTADO	char	2	Não
	CEP	char	9	Sim
	SGLHOSPITAL	char	3	Não

Tabela A.1 Dimensão Paciente

A.2.2 Dimensão Tipo_Ocorrencia

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	TPOCORRENCIA_SK	int		Não
	TPOCORRENCIA_NU_VERSAO	smallint		Não
	TPOCORRENCIA_DT_ETL	date		Não
	DESCRICAOTPOCORRENCIA	varchar	50	Não
	CODNATTPOCORRENCIA	int		Não
	DESCRICAONATOCORRENCIA	varchar	50	Não

Tabela A.2 Dimensão Tipo_Ocorrencia

A.2.3 Dimensão Procedencia

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	PROCEDENCIA_SK	int		Não
	PROCEDENCIA_NU_VERSAO	smallint		Não
	PROCEDENCIA_DT_ETL	date		Não
	DESCRICAOCODPROCEDENCIA	varchar	50	Não
	CODNATPROCEDENCIA	char	2	Não
	DESCRICAONATPROCEDENCIA	varchar	50	Não

Tabela A.3 Dimensão Procedencia

A.2.4 Dimensão Destino

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	DESTINO_SK	int		Não
	DESTINO_NU_VERSAO	smallint		Não
	DESTINO_DT_ETL	date		Não
	DESCRICAOCODDESTINO	varchar	50	Não
	CODNATDESTINO	char	2	Não
	DESCRICAONATDESTINO	varchar	50	Não

Tabela A.4 Dimensão Destino

A.2.5 Dimensão Medico_Liberador

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	MEDICO_SK	int		Não
	MEDICO_NU_VERSAO	smallint		Não
	MEDICO_DT_ETL	date		Não
	MATRICULA	varchar	15	Não
	NOMECOMPLETO	varchar	50	Não
	NOMEREDUZIDO	varchar	20	Não
	CODSEXO	char	2	Não
	DESCRICAOSEXO	varchar	50	Não
	SITUACAO	char	3	Não

Tabela A.5 Dimensão Medico_Liberador

A.2.6 Dimensão Hospital

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	HOSPITAL_SK	int		Não
	HOSPITAL_NU_VERSAO	smallint		Não
	HOSPITAL_DT_ETL	date		Não
	SGLHOSPITAL	char	3	Não
	NOMEHOSPITAL	varchar	50	Não
	ENDERECO	varchar	50	Não
	COMPLEMENTOENDERECO	varchar	25	Não
	BAIRRO	varchar	25	Não
	CIDADE	varchar	25	Não
	SGLESTADO	char	2	Não
	CEP	char	9	Não
	CNPJ	char	18	Não
	ANOIMPLANTACAOSAPE	smallint		Não

Tabela A.6 Dimensão Hospital

A.2.7 Dimensão Data_Registro

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	TEMPO_SK	int		Não
	DATAREGISTRO_NU_VERSAO	smallint		Não
	DATAREGISTRO_DT_ETL	date		Não
	DATAREGISTRO	date		Não
	ANOREGISTRO	smallint		Não
	MESREGISTRO	smallint		Não
	DIAREGISTRO	smallint		Não
	DIADOANO	smallint		Não
	DIADASEMANA	smallint		Não
	SEMANADOANO	smallint		Não
	DIADESCRICAO	char	7	Não
	DIARESUMIDO	char	3	Não
	MESDESCRICAO	char	9	Não
	MESRESUMIDO	char	3	Não
	SEMESTRE	smallint		Não
	QUADRIMESTRE	smallint		Não
	TRIMESTRE	smallint		Não
	BIMESTRE	smallint		Não

Tabela A.7 Dimensão Data_Registro

A.2.8 Dimensão Emergencia_Origem

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	EMERGENCIA_SK	int		Não
	EMERGENCIA_NU_VERSAO	smallint		Não
	EMERGENCIA_DT_ETL	date		Não
	DESCRICAOCODEMERGENCIA	varchar	50	Não
	CODNATEMERGENCIA	char	2	Não
	DESCRICAONATEMERGENCIA	varchar	50	Não

Tabela A.8 Dimensão Emergencia_Origem

A.2.9 Dimensão Ind_Acid_Trabalho

Dimensão Degenerada utilizando-se de uma coluna na tabela de fatos com o indicador Sim ou Não (S/N).

A.2.10 Dimensão Ind_Pac_Nao_Identificado

Dimensão Degenerada utilizando-se de uma coluna na tabela de fatos com o indicador Sim ou Não (S/N).

A.2.11 Dimensão Forma_Transporte

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	FORMATTRANSPORTE_SK	int		Não
	FORMATTRANSPORTE_NU_VERSAO	smallint		Não
	FORMATTRANSPORTE_DT_ETL	date		Não
	DESCRICAOFORMATTRANSPORTE	varchar	50	Não

Tabela A.9 Dimensão Forma_Transporte

A.2.12 Fato Atendimento

Chave	Nome da Coluna	Tipo	Tamanho	Null
Sim	SEQUENCIAL	char	13	Não
Sim	TEMPO_SK	int		Não
Sim	EMERGENCIA_SK	int		Não
Sim	FORMA_TRANSPORTE_SK	int		Não
Sim	HOSPITAL_SK	int		Não
Sim	MEDICO_SK	int		Não
Sim	DESTINO_SK	int		Não
Sim	PROCEDENCIA_SK	int		Não
Sim	TPOCORRENCIA_SK	int		Não
Sim	PACIENTE_SK	int		Não
	FAIXA_ETARIA	char	7	Não
	FAIXA_HORARIO	char	14	Não
	IND_ACID_TRABALHO	char	1	Não
	IND_PAC_NAO_IDENTIFICADO	char	1	Não
	ATEND_QTD	smallint		Não
	ATEND_QTD_OBITO	smallint		Não
	ATEND_QTD_TRANSFERENCIA	smallint		Não
	ATEND_QTD_ALTA	smallint		Não
	ATEND_QTD_INTERNACAO	smallint		Não

Tabela A.10 Fato Atendimento

A.3 Modelo Relacional de Correspondência das dimensões

A seguir apresentam-se os modelos de entidades e relacionamentos das tabelas existentes no sistema transacional do SAPE que deram origem às dimensões e ao fato disponibilizados no *Data Warehouse*.

A.3.1 Dimensão Paciente

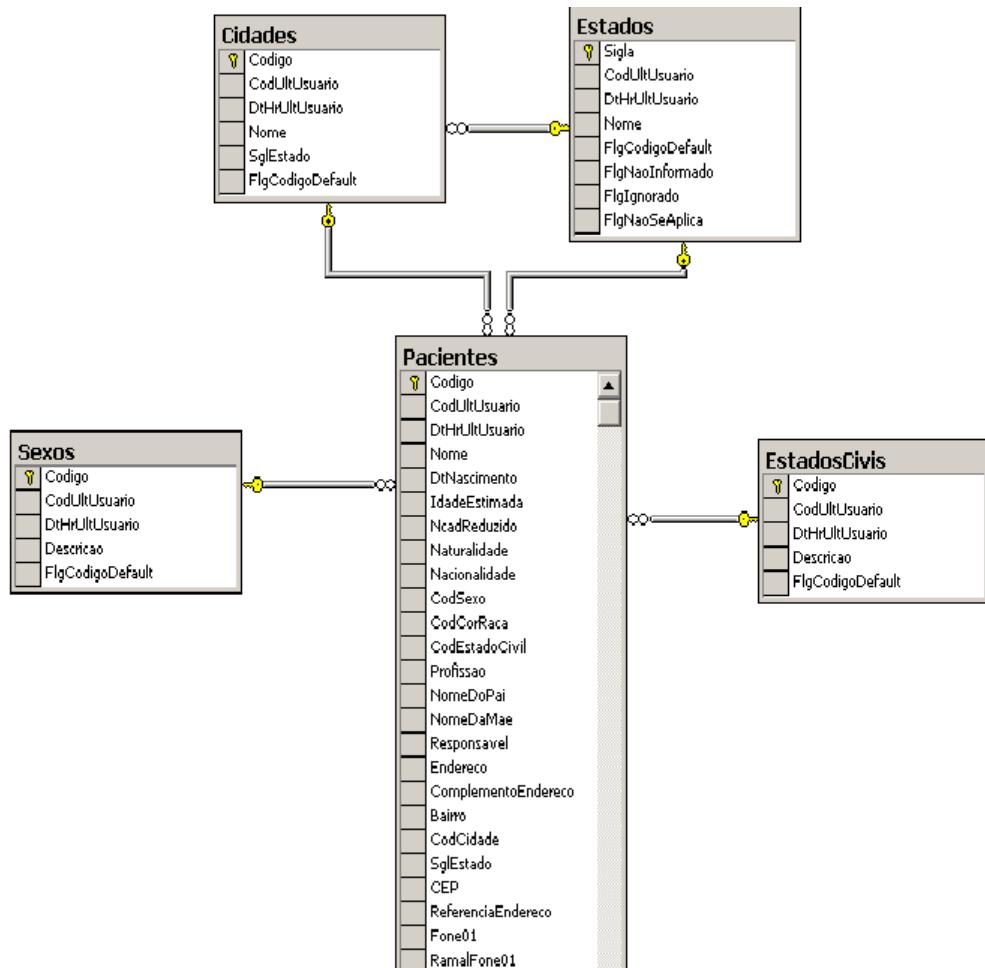


Figura A.2 Tabelas de origem no SAPE Emergência - Paciente.

A.3.2 Dimensão Tipo_Ocorrencia

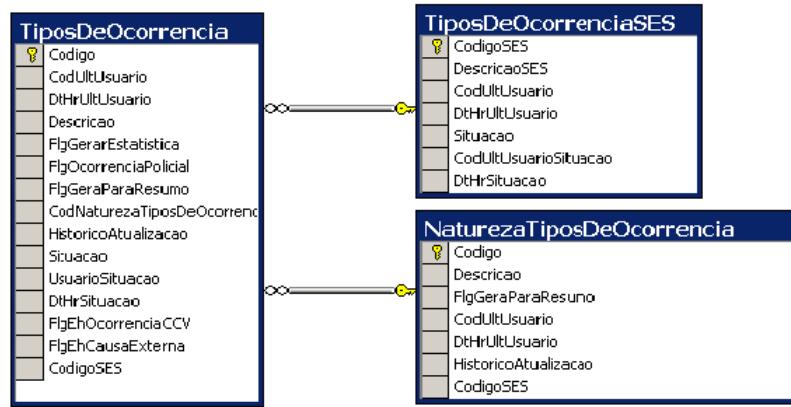


Figura A.3 Tabelas de origem no Sistema SAPE Emergência - Tipo Ocorrência.

A.3.3 Dimensão Procedencia

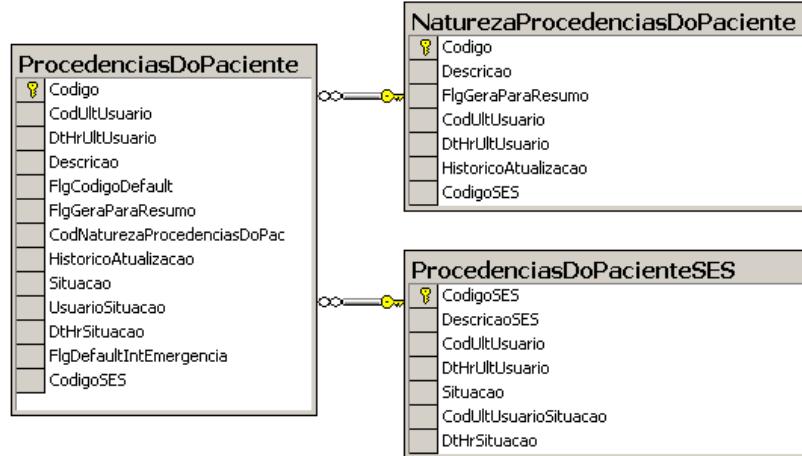


Figura A.4 Tabelas de origem no Sistema SAPE Emergência - Procedência.

A.3.4 Dimensão Destino

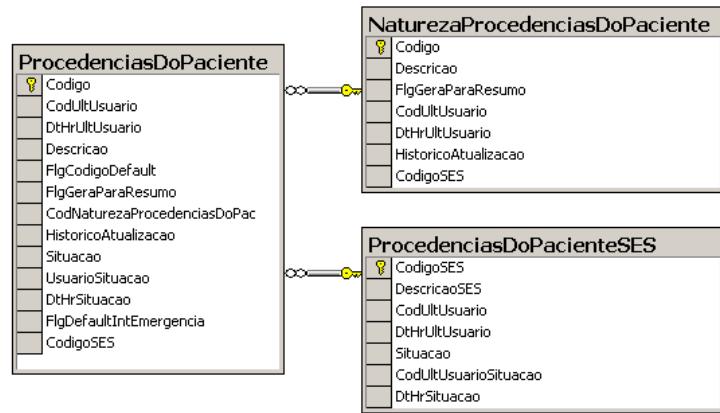


Figura A.5 Tabelas de origem no Sistema SAPE Emergência - Procedência.

A.3.5 Dimensão Medico_Liberador

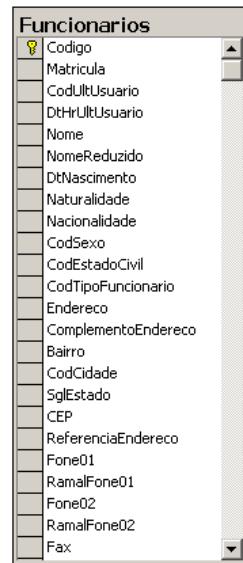


Figura A.6 Tabelas de origem no Sistema SAPE Emergência - Médico Liberados.

A.3.6 Dimensão Hospital

Parametros	
?	Codigo
	CodUltUsuario
	DtHrUltUsuario
	NroVersaoDados
	NroReleaseDados
	Cab1Relatorios
	Cab2Relatorios
	RodapeRelatorios
	ObsSuporteTecnico
	NomeHospital
	AnoAnteriorSistema
	AnoAtualSistema
	AnoProximoSistema
	AnoImplantacaoSistema
	MinutosEntreAtendimentos
	MinutosDifHoraMicroServidor
	DiasAvisarVencSenha
	QtdMaxRegistrosConsulta
	QtdMaxMinutosConsulta
	NroPronuarioImplantacao
	NroPronuarioImplantacaoAmb
	PctPacientesGeracaoCartas
	QtdMinGeracaoCartas
	QtdMaxGeracaoCartas
	ServidorEmailGeracaoCartas
	EmailDeOrigemGeracaoCartas
	EmailDeDestinoGeracaoCartas
	LimiteDiasPermBaixaAutomatica
	LimiteMinDiasBaixaAutomatica

Figura A.7 Tabelas de origem no Sistema SAPE Emergência - Hospitais.

A.3.7 Dimensão Emergencia_Origem

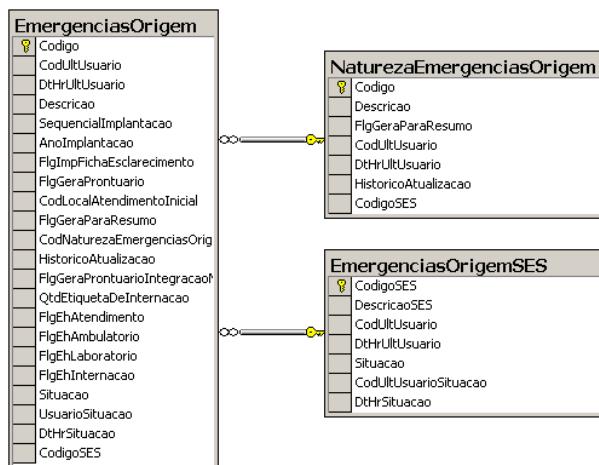


Figura A.8 Tabelas de origem no Sistema SAPE Emergência - Emergência Origem.

A.3.8 Dimensão Data_Registro

Tabela criada no Modelo Multidimensional.

A.3.9 Dimensão Ind_Acid_Trabalho

Dimensão Degenerada.

A.3.10 Dimensão Ind_Pac_Nao_Identificado

Dimensão Degenerada.

A.3.11 Dimensão Faixa_Etaria

Dimensão Degenerada.

A.3.12 Dimensão Faixa_Horario

Dimensão Degenerada.

A.3.13 Dimensão Forma_Transporte



Figura A.9 Tabelas de origem no Sistema SAPE Emergência - Forma de Transporte.

A.3.14 Fato Atendimento

Atendimentos	
Codigo	
AnoAtendimento	
CodEmergenciaOrigem	
SequencialAtendimento	
Atendimentos	Atendimentos
FlgPacienteNaoIdentificado	
CodUltUsuario	
DtHrUltUsuario	
CodPaciente	
DtRegistro	
HrRegistro	
DtHrRegistro	
CodPlanoSaude	
NroPlanoSaude	
DtValidadePlanoSaude	
CodProcedencia	
DescProcedencia	
CodDestino	
DescDestino	
FlgFezContatoPrevio	
CodTipoOcorrencia	
ObservacoesOcorrencia	
ObsIdentificacaoServicoSo	
Situacao	

Figura A.10 Tabelas de origem no Sistema SAPE Emergência - Atendimentos.

APÊNDICE B

Apêndice B

Neste Apêndice são apresentadas as Transformações e os *Jobs* utilizados no processo de Extração, Transformação e Carga (ETL) do *Data Warehouse* da Secretaria de Saúde do Estado de Pernambuco (SES-PE).

Este trabalho de documentação desenvolveu-se pelos alunos do mestrado, equipe de ETL, Adriana Carla Simões, David Carnaúba, Daniel Melo, Iandé Coutinho, Robertson Ferraz e Rodrigo Freitas, participantes da disciplina Tópicos Especiais em Banco de Dados 1 (Sistemas de Banco de Dados para Suporte à Decisão) ministrada pelo professor Dr. Robson Fidalgo.

B.1 Descrição das Transformações e *Jobs* Gerados

O detalhamento de cada transformação segue a sequência de Extração, Geração de chaves artificiais (*surrogate*) e carga.

B.1.1 Dimensão Emergência

A seguir são apresentadas as transformações necessárias para gerar a dimensão Emergência do esquema estrela do Data Warehouse da Secretaria da Saúde. Esta dimensão foi criada a partir da execução de três transformações comentadas nas próximas seções.

B.1.1.1 Extração Emergência

Requisitos: Nenhum.

Objetivo: Adicionar na *data staging area* os tipos de emergência.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Database	Este componente lê as tuplas de uma tabela do banco de dados e as passa para o próximo step.	Lê os tipos de emergências cadastrados na tabela EmergenciasOrigemSES de 4 bancos de dados distintos.
	Sort Rows	Este componente ordena as linhas de acordo com uma hierarquia de campos definidos pelo projetista.	Ordena as tuplas de forma ascendente pelos campos:CodigoSES, DescricaoSES, CodNaturezaEmergenciasOrigem, DescricaoNatureza.
	Unique Rows	Remove tuplas duplicadas.	Remove as tuplas que tenham o CodigoSES e DescricaoSES duplicados.
	Output Text	Grava dados em um arquivo.	Salva no arquivo (EmergenciaExtracao.txt) as transformações.

Figura B.1 Componentes das Emergências.

Descrição da Transformação:

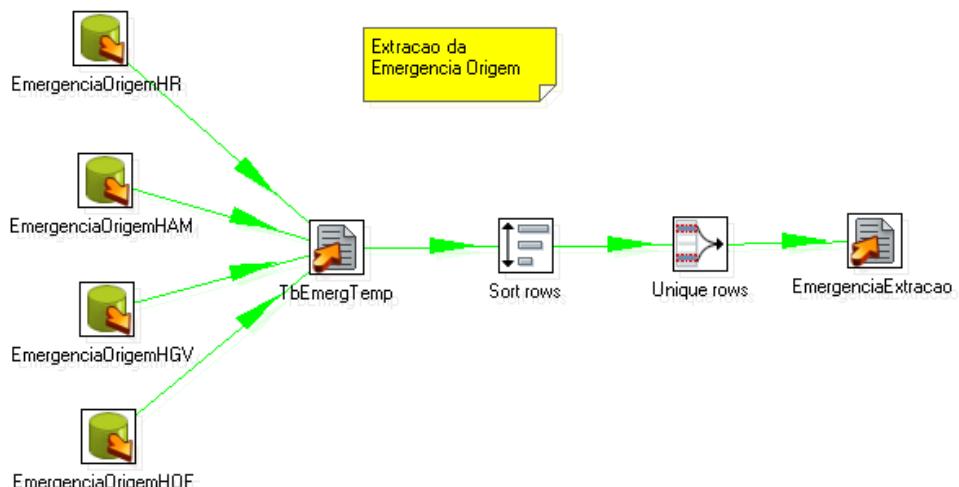


Figura B.2 Extração das Emergências.

As tabelas do banco de dados dos tipos de emergência são lidas salvas em um arquivo; Estes passos são realizados pelos passos: EmergenciaOrigemHR, EmergenciaOrigemHAM, EmergenciaOrigemHGV, EmergenciaOrigemHOF.

Este processo é indicado devido às operações de sort e distinct serem realizadas de forma mais eficientes quando retiradas do banco de dados e salvas em um arquivo.

O componente Sort Rows (referente ao order by de um SGBD), ordena os dados de forma ascendente de acordo com as colunas: CodigoSES, DescricaoSES, CodNaturezaEmergenciasOrigem, DescricaoNatureza. Esta operação é requisito fundamental para o próximo passo,

Unique rows, que elimina as tuplas duplicadas.

Este passo refere-se a um distinct realizado no banco de dados (distinct *CodigoSES*, *DescricaoSES*). E por último temos o Output de toda a transformação para um arquivo texto (*EmergenciaTransformacao.txt*).

B.1.1.2 Geração das *SurrogateKeys* das Emergências

Nome da Transformação: Geração das *SurrogateKeys* das Emergências

Requisitos: Extração das Emergências

Objetivo: Gerar as *SurrogateKeys* das Emergências no DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo EmergenciaExtracao.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: CodigoSES, DescricaoSES, CodNaturezaEmergencias, DescricaoNatureza.
	Database Joins	Faz um select no banco de dados utilizando o resultado de outras consultas como parâmetro.	Seleciona o maior código das SurrogateKey existentes para gerar o próximo valor.
	Add Sequence	Gera uma variável incrementada a cada registro passado pelo step.	Gera um número seqüencial.
	JavaScript	Executa uma transformação em JavaScript.	Evita que o SurrogateKey seja nulo.
	Text Output	Escreve dados em um arquivo.	Salva a transformação das emergências no arquivo EmergenciaCarga.txt.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva as relações das SurrogateKey e das NaturalKey na tabela SurrogateEmergencia do DW.

Figura B.3 Componentes das *SurrogateKeys* das Emergências.

Descrição da Transformação:

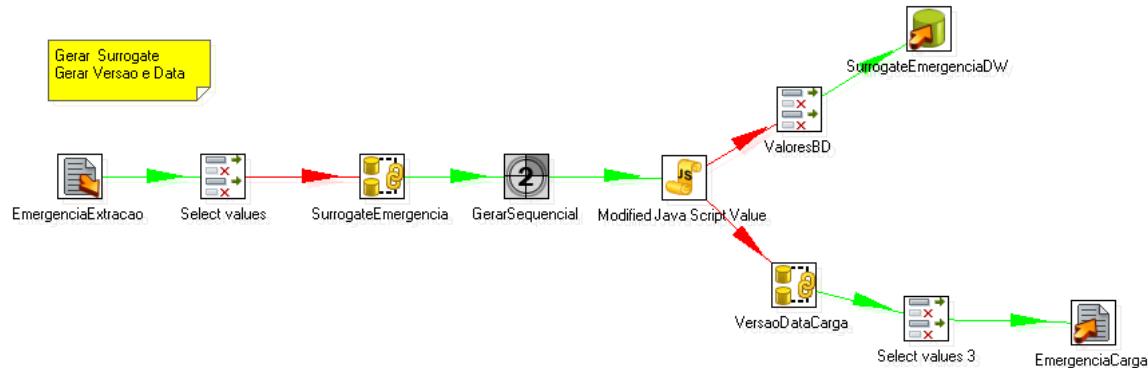


Figura B.4 Transformação *Surrogate* das Emergências.

O primeiro passo *EmergenciaExtracao* lê o arquivo que contém os dados necessários para gerar as *SurrogateKeys*, este arquivo foi previamente gerado pela transformação Extração de Emergência.

Os passos (*Select values*) removem e renomeiam campos desnecessários. O passo *SurrogateEmergencia* retira do banco o maior código existente para gerar uma nova *surrogate*, isto elimina a possibilidade de gerar *SurrogateKeys* duplicadas.

A seguir um seqüencial é gerado para adicionar ao maior *SurrogateKey* encontrado, criando assim, o novo código para o registro. Tem-se o problema de não existir nenhuma *SurrogateKey* no banco, sendo assim, é retornado um valor nulo. Para contornar este problema e evitar a soma do seqüencial com um valor nulo é realizada uma verificação em JavaScript.

Pode-se notar a cópia do fluxo dos dados que saem do passo *ModifiedJavaScript Value*, isto é para salvar as referências das *SurrogateKeys* geradas com as *NaturalKeys*. O outro fluxo que sai do passo vai para um arquivo que são os dados da dimensão *Emergencia*.

B.1.1.3 Carga das Emergências no DW

Requisitos: Extração das Emergências, Geração das *SurrogateKeys* das Emergências.

Objetivo: Realizar a carga das transformações realizadas para o DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo EmergenciaCarga.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: DescricaoSES, Emergencia_SK, NumeroVersao, DataETL, DescricaoNatureza, CodNaturezaEmergenciaOrigem.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva a transformação das emergências na tabela emergencia_origem do DW.

Figura B.5 Componentes das Carga das Emergências.

Descrição da Transformação:

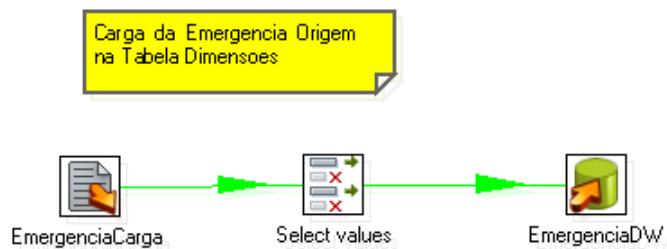


Figura B.6 Transformação Carga das Emergências.

O primeiro passo *EmergenciaCarga* lê o arquivo que contém os dados necessários para a carga, este arquivos foi previamente gerado pelas transformações: Extração de Emergência e Emergência Surrogate.

O próximo passo (*Select values*) foi necessário para o próximo passo que insere as transformações no Data Warehouse. *Select values* desempenha o papel de renomear e selecionar os

campos que existem na tabela de destino (*emergencia_origem*) apontado pelo passo *EmergênciaDW*.

B.1.2 Dimensão Hospital

A seguir são apresentadas as transformações necessárias para gerar a dimensão Hospital do esquema estrela do DW de saúde. Esta dimensão foi criada pela execução de três transformações comentadas nas próximas seções.

B.1.2.1 Extração dos Hospitais

Requisitos: Nenhum.

Objetivo: Adicionar na *data staging area* os tipos de hospitais.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Database	Este componente lê as tuplas de uma tabela do banco de dados e as passa para o próximo step.	Lê os hospitais cadastrados na tabela Parmetros de 4 bancos de dados distintos.
	Output Text	Escreve dados em um arquivo.	Salva no arquivo(HospitalExtracao.txt) a transformação.

Figura B.7 Componentes dos Hospitais.

Descrição da Transformação:

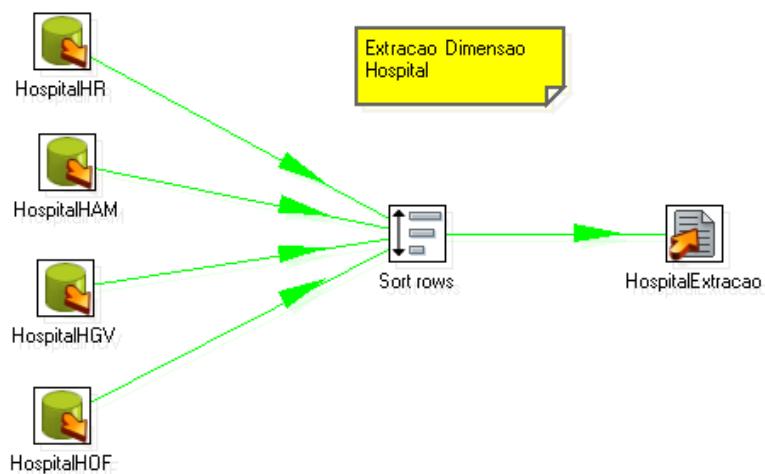


Figura B.8 Extração dos Hospitais.

Primeiro é lido das tabelas do banco de dados as informações os hospitais e salva em um arquivo, passos estes realizados pelos passos: HospitalHR, HospitalHAM, HospitalHGV, Hospi-

talHOF. Este processo é indicado devido às próximas serem realizadas de forma mais eficientes quando retiradas do banco de dados e salvas em um arquivo. E por último temos o Output de toda a transformação para um arquivo texto (HospitalExtracao.txt).

B.1.2.2 Geração das *SurrogateKeys* dos Hospitais

Nome da Transformação: Geração das *SurrogateKeys* dos Hospitais

Requisitos: Extração das Hospitais

Objetivo: Gerar as *SurrogateKeys* dos Hospitais no DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo HospitalExtracao.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: Codigo, NomeHospital, SiglaHospital, EnderecoHospital, ComplEnderecoHospital, BairroHospital, CidadeHospital, EstadoHospital, CepHospital, CnpjHospital, AnoImplantacaoSistema
	Database Joins	Faz um select no banco de dados utilizando o resultado de outras consultas como parâmetro.	Seleciona o maior código das SurrogateKey existentes para gerar o próximo valor.
	Add Sequence	Gera uma variável incrementada a cada registro passado pelo step.	Gera um número sequencial.
	JavaScript	Executa uma transformação em JavaScript.	Evita que o SurrogateKey seja nulo.
	Text Output	Escreve dados em um arquivo.	Salva a transformação das emergências no arquivo HospitalCarga.txt.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva a relação SurrogateKey e NaturalKey dos hospitais na tabela SurrogateHospital do DW.

Figura B.9 Componentes das *SurrogateKeys* dos Hospitais.

Descrição da Transformação:

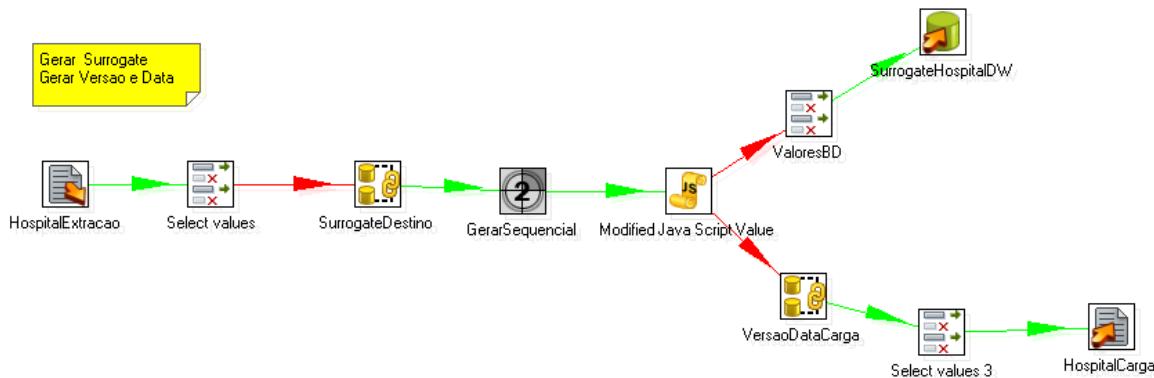


Figura B.10 Transformação *Surrogate* dos Hospitais.

O primeiro passo HospitalExtracao lê o arquivo que contém os dados necessários para gerar as *SurrogateKeys*, este arquivo foi previamente gerado pela transformação Extração dos Hospitais.

Os passos (*Select values*) removem e renomeiam campos desnecessários.

O passo SurrogateDestino retira do banco o maior código existente para gerar uma nova *surrogate*, isto elimina a possibilidade de gerar *SurrogateKeys* duplicadas.

A seguir um seqüencial é gerado para adicionar ao maior *SurrogateKey* encontrado, criando assim, o novo código para o registro. Tem-se o problema de não existir nenhuma *SurrogateKey* no banco, sendo assim, é retornado um valor nulo. Para contornar este problema e evitar a soma do seqüencial com um valor nulo é realizada uma verificação em JavaScript.

Pode-se notar a cópia do fluxo dos dados que saem do passo *ModifiedJavaScript Value*, isto é para salvar as referências das *SurrogateKeys* geradas com as *NaturalKeys*. O outro fluxo que sai do passo vai para um arquivo que são os dados da dimensão Hospital.

B.1.2.3 Carga das Hospitais no DW

Requisitos: Extração dos Hospitais, Geração das *SurrogateKeys* dos Hospitais.

Objetivo: Realizar a carga das transformações realizadas para o DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo HospitalExtracao.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: NomeHospital, SiglaHospital, Endereco, ComplementoEndereco, Bairro, Cidade, SglEstado, Cep, Cnpj, AnoImplantacaoSape, Hospital_SK, Hospital_NU_Versao, Hospital_DT_ETL
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva a transformação das emergências na tabela hospital do DW.

Figura B.11 Componentes das Carga dos Hospitais.

Descrição da Transformação:



Figura B.12 Transformação Carga dos Hospitais.

O primeiro passo HospitalCarga lê o arquivo que contém os dados necessários para a carga, este arquivos foi previamente gerado pelas transformações: Extração dos Hospitais e Geração das *SurrogateKey* dos Hospitais.

O próximo passo (*Select values*) foi necessário para o próximo passo que insere as transformações no DW. *Select values* desempenha o papel de renomear e selecionar os campos que existem na tabela de destino (Hospital) apontado pelo passo HospitalDW.

B.1.3 Dimensão Forma Transporte

A seguir são apresentadas as transformações necessárias para gerar a dimensão FormaDeTransporte do esquema estrela do DW de saúde. Esta dimensão foi criada pela execução de três transformações comentadas nas próximas seções.

B.1.3.1 Extração dos Hospitais

Requisitos: Nenhum.

Objetivo: Adicionar na *data staging area* os tipos de hospitais.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Database	Este componente lê as tuplas de uma tabela do banco de dados e as passa para o próximo step.	Lê as formas de transporte cadastradas da tabela FormaDeTransporte.
	Output Text	Escreve dados em um arquivo.	Salva no arquivo (FormasTransporteExtracao.txt) a transformação.

Figura B.13 Componentes das Formas de Transporte.

Descrição da Transformação:

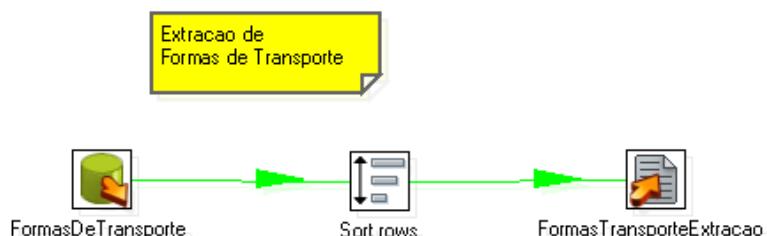


Figura B.14 Extração das Formas de Transporte.

Primeiro é lido da tabela do banco de dados as informações das Formas de Transporte e salva em um arquivo este passo é realizado pelo passo FormasDeTransporte. Este processo é indicado devido às próximas serem realizadas de forma mais eficientes quando retiradas do banco de dados e salvas em um arquivo. E por último temos o Output de toda a transformação para um arquivo texto (FormasTransporteExtracao.txt).

B.1.3.2 Geração das *SurrogateKeys* das Formas de Transportes

Nome da Transformação: Geração das *SurrogateKeys* das Formas de Transporte.

Requisitos: Extração das Formas de Transporte.

Objetivo: Gerar as *SurrogateKeys* das Formas de Transporte no DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo FormasTransporteExtracao.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: Código, Descrição.
	Database Joins	Faz um select no banco de dados utilizando o resultado de outras consultas como o parâmetro.	Seleciona o maior código das SurrogateKey existentes para gerar o próximo valor.
	Add Sequence	Gera uma variável incrementada a cada registro passado pelo step.	Gera um número seqüencial.
	JavaScript	Executa uma transformação em JavaScript.	Evita que o SurrogateKey seja nulo.
	Text Output	Escreve dados em um arquivo.	Salva a transformação das emergências no arquivo FormaTransporteCarga.txt.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva a relação SurrogateKey e NaturalKey das FormasDeTransporte na tabela SurrogateFormaTransporte do DW.

Figura B.15 Componentes das *SurrogateKeys* das Formas de Transportes.

Descrição da Transformação:

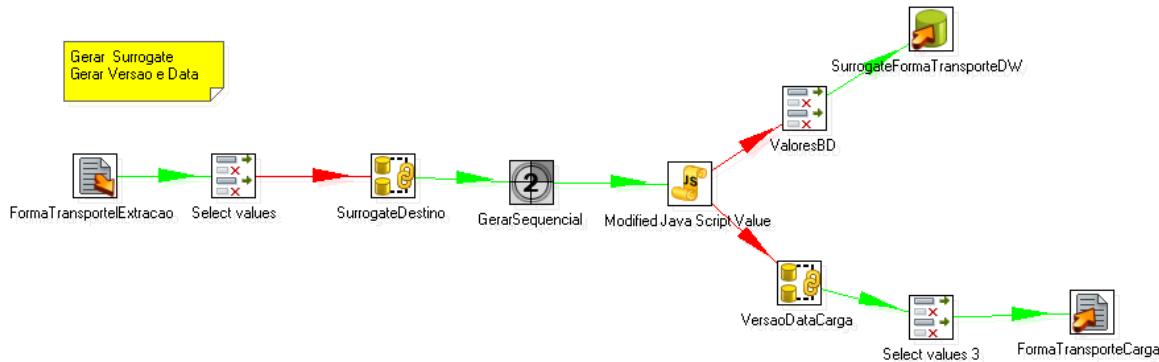


Figura B.16 Transformação *Surrogate* das Formas de Transportes.

O primeiro passo FormaTransporteExtracao lê o arquivo que contém os dados necessários para gerar as *SurrogateKeys*, este arquivo foi previamente gerado pela transformação Extração das Formas de Transporte.

Os passos (*Select values*) removem e renomeiam campos desnecessários.

O passo SurrogateDestino retira do banco o maior código existente para gerar uma nova *surrogate*, isto elimina a possibilidade de gerar *SurrogateKeys* duplicadas. A seguir um seqüencial é gerado para adicionar ao maior *SurrogateKey* encontrado, criando assim, o novo código para o registro.

Tem-se o problema de não existir nenhuma *SurrogateKey* no banco, sendo assim, é retornado um valor nulo. Para contornar este problema e evitar a soma do seqüencial com um valor nulo é realizada uma verificação em JavaScript.

Pode-se notar a cópia do fluxo dos dados que saem do passo *ModifiedJavaScript Value*, isto é para salvar as referências das *SurrogateKeys* geradas com as *NaturalKeys*. O outro fluxo que sai do passo vai para um arquivo que são os dados da dimensão FormasDeTransporte.

B.1.3.3 Carga das Formas de Transporte

Requisitos: Extração das Formas de Transporte, Geração das *SurrogateKeys* das Formas de Transporte.

Objetivo: Realizar a carga das transformações realizadas para o DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo FormaTransporteCarga.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: Descricao, DescricaoCodFormaTransporte, FormaTransporte_SK, NumeroVersao, FormaTransporte_NU_Versao, DataETL, FormaTransporte_DT_ETL.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva a transformação das emergências na tabela hospital do DW.

Figura B.17 Componentes das Carga das Formas de Transporte.

Descrição da Transformação:

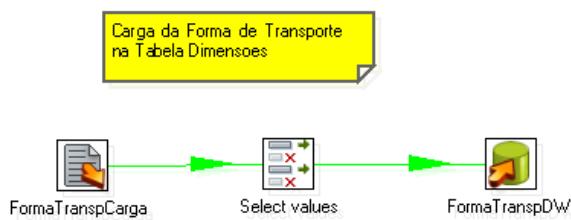


Figura B.18 Transformação Carga das Formas de Transporte.

O primeiro passo FormaTranspCarga lê o arquivo que contém os dados necessários para a carga, este arquivos foi previamente gerado pelas transformações: Extração das Formas de Transporte e Geração das *SurrogateKey* das Formas de Transporte.

O próximo passo (*Select values*) foi necessário para o próximo passo que insere as transformações no DW. *Select values* desempenha o papel de renomear e selecionar os campos que existem na tabela de destino (forma_transport) apontado pelo passo FormaTranspDW.

B.1.4 Dimensão Destino

A seguir são apresentadas as transformações necessárias para gerar a dimensão Destino do esquema estrela do DW de saúde. Esta dimensão foi criada pela execução de três transformações comentadas nas próximas seções.

B.1.4.1 Extração dos Destinos

Requisitos: Nenhum.

Objetivo: Adicionar na *data staging area* os tipos de hospitais.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Database	Este componente lê as tuplas de uma tabela do banco de dados e as passa para o próximo step.	Lê os destinos cadastrados na tabela DestinoDosPacientes de 4 bancos de dados distintos.
	Sort Rows	Este componente ordena as linhas de acordo com uma hierarquia de campos definidos pelo projetista.	Ordena as tuplas de forma ascendente pelos campos:CodigoSES, DescricaoSES, CodNaturezaDestinosDoPaciente, DescricaoNatureza.
	Unique Rows	Remove tuplas duplicadas.	Remover as tuplas que tenham o CodigoSES e DescricaoSES duplicados.
	Output Text	Escreve dados em um arquivo.	Salva no arquivo(DestinoExtracao.txt) as transformações.

Figura B.19 Componentes dos Destinos.

Descrição da Transformação:

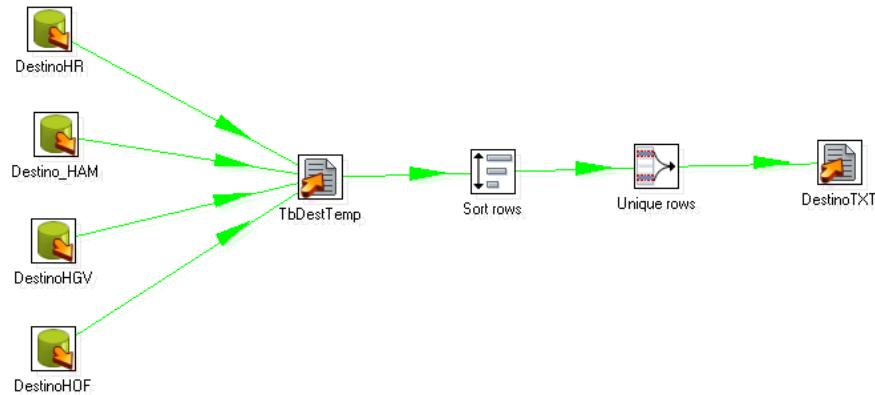


Figura B.20 Extração dos Destinos.

Primeiro é lido das tabelas do banco de dados as informações dos destinos do paciente e salva em um arquivo, passos estes realizados pelos passos: DestinoHR, DestinoHAM, DestinoHGV, DestinoHOF. Este processo é indicado devido às operações de sort e distinct serem realizadas de forma mais eficientes quando retiradas do banco de dados e salvas em um arquivo.

O componente Sort rows (referente ao order by de um SGBD), ordena os dados de forma ascendente de acordo com as colunas: CódigoSES, DescriçãoSES, CodNaturezaDestinoDoPaciente, DescriçãoNatureza. Esta operação é requisito fundamental para o próximo passo, Unique rows, que elimina as tuplas duplicadas. Este passo refere-se a um distinct realizado no banco de dados (distinct CódigoSES, DescriçãoSES). E por último temos o Output de toda a transformação para um arquivo texto (DestinoExtracão.txt).

B.1.4.2 Geração das *SurrogateKeys* dos Destinos

Nome da Transformação: Geração das *SurrogateKeys* dos Destinos

Requisitos: Extração das Destinos

Objetivo: Gerar as *SurrogateKeys* dos Destinos no DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo DestinoExtracao.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: CodigoSES, DescricaoSES, CodNaturezaDestinoDo Paciente, DescricaoNatureza.
	Database Joins	Faz um select no banco de dados utilizando o resultado de outras consultas como parâmetro.	Seleciona o maior código das SurrogateKey existentes para gerar o próximo valor.
	Add Sequence	Gera uma variável incrementada a cada registro passado pelo step.	Gera um número seqüencial.
	JavaScript	Executa uma transformação em JavaScript.	Evita que o SurrogateKey seja nulo.
	Text Output	Escreve dados em um arquivo.	Salva a transformação das emergências no arquivo DestinoCarga.txt.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva as relações das SurrogateKey e das NaturalKey na tabela SurrogateDestino do DW.

Figura B.21 Componentes das *SurrogateKeys* dos Destinos.

Descrição da Transformação:

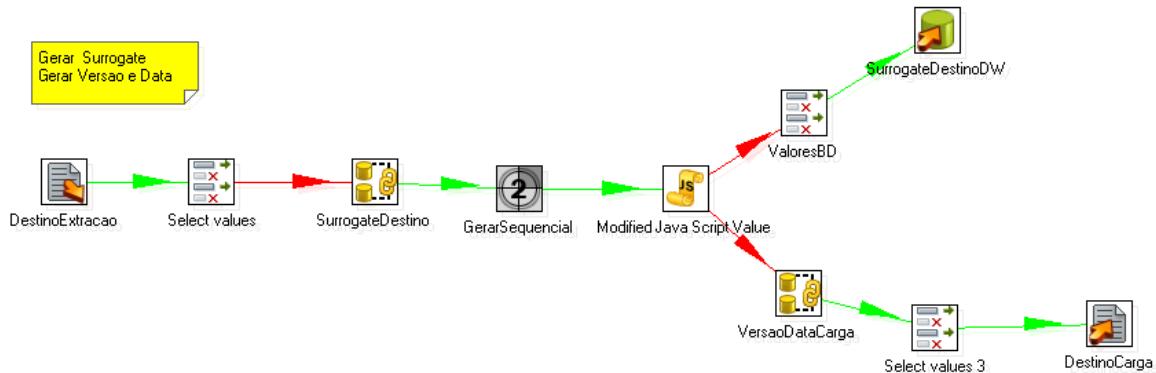


Figura B.22 Transformação *Surrogate* dos Destinos.

O primeiro passo DestinoExtracao lê o arquivo que contém os dados necessários para gerar as *SurrogateKeys*, este arquivos foi previamente gerado pela transformação Extração dos Destinos.

Os passos (*Select values*) removem e renomeiam campos desnecessários. O passo SurrogateDestino retira do banco o maior código existente para gerar uma nova *surrogate*, isto elimina a possibilidade de gerar *SurrogateKeys* duplicadas.

A seguir um seqüencial é gerado para adicionar ao maior *SurrogateKey* encontrado, criando assim, o novo código para o registro. Tem-se o problema de não existir nenhuma *SurrogateKey* no banco, sendo assim, é retornado um valor nulo. Para contornar este problema e evitar a soma do seqüencial com um valor nulo é realizada uma verificação em JavaScript.

Pode-se notar a cópia do fluxo dos dados que saem do passo *ModifiedJavaScript Value*, isto é para salvar as referências das *SurrogateKeys* geradas com as *NaturalKeys*. O outro fluxo que sai do passo vai para um arquivo que são os dados da dimensão Destino.

B.1.4.3 Carga das Destinos no DW

Requisitos: Extração dos Destinos, Geração das *SurrogateKeys* dos Destinos.

Objetivo: Realizar a carga das transformações realizadas para o DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo DestinoCarga.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: DescricaoSES, DestinoSK, NumeroVersao, DataETL, DescricaoNatureza, CodNaturezaDestinosDo Paciente.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva a transformação das emergências na tabela Destino do DW.

Figura B.23 Componentes das Carga dos Destinos.

Descrição da Transformação:



Figura B.24 Transformação Carga dos Destinos.

O primeiro passo DestinoCarga lê o arquivo que contém os dados necessários para a carga, este arquivos foi previamente gerado pelas transformações: Extração dos Destinos e Destino Surrogate.

O próximo passo (*Select values*) foi necessário para o próximo passo que insere as transformações no DW. *Select values* desempenha o papel de renomear e selecionar os campos que existem na tabela de destino (destino) apontado pelo passo DestinoDW.

B.1.5 Tabela de Fatos Atendimentos

A seguir são apresentadas as transformações necessárias para gerar a dimensão de Fato do esquema estrela do DW de saúde. Esta dimensão foi criada pela execução de três transformações comentadas nas próximas seções.

B.1.5.1 Extração dos Atendimentos

Requisitos: Nenhum.

Objetivo: Adicionar na *data staging area* os atendimentos.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Database	Este componente lê as tuplas de uma tabela do banco de dados e as passa para o próximo step.	Lê os Atendimentos cadastrados na tabela Atendimentos de 4 bancos de dados distintos.
	Output Text	Escreve dados em um arquivo.	Salva no arquivo(FatosAtendimentosExtracao.txt) as transformações.

Figura B.25 Componentes dos Atendimentos.

Descrição da Transformação:

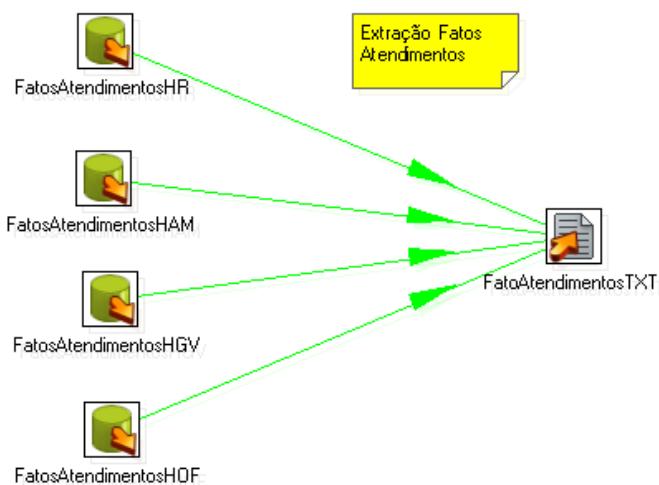


Figura B.26 Extração dos Atendimentos.

Primeiro é lido das tabelas do banco de dados as informações os hospitais e salva em um arquivo, passos estes realizados pelos passos: FatosAtendimentosHR, FatosAtendimentosHAM, FatosAtendimentosHGV, FatosAtendimentosHOF.

Este processo é indicado devido às próximas serem realizadas de forma mais eficientes quando retiradas do banco de dados e salvas em um arquivo. Nesta extração, o comando de extração realiza uma junção com a tabela de pacientes para retornar as colunas referentes a Data de Nascimento e Idade Estimada para melhoria da performance quando da transformação, no mapeamento das Faixas Etárias.

E por último temos o Output de toda a transformação para um arquivo texto (FatoAtendimentosExtracao.txt).

B.1.5.2 Transformação dos Fatos para o Esquema Estrela

Nome da Transformação: Extração dos Fatos para o Esquema Estrela

Requisitos: Carga dos Destinos, Carga dos Tipos de Emergência, Carga dos Hospitais, Carga das Formas de Transporte, Carga de Pacientes, Carga de Datas, Carga de Procedência, Carga de Ocorrência, Carga de Médico e Extração dos Atendimentos.

Objetivo: Adicionar na *data staging area* os Fatos transformados para o esquema estrela.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Database	Este componente lê as tuplas de uma tabela do banco de dados e as passa para o próximo step.	Lê os Fatos de Atendimento cadastrados na tabela Atendimentos de 4 bancos de dados distintos.
	Input Text	Este componente lê os registros armazenados como o CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo DestinoCarga.txt.
	Stream Lookup	Retorna o valor de uma outra stream se os campos chaves são iguais.	Extrai as SurrogateKey.
	Value Mapper	Mapeia determinados valores de um campo para outros valores.	Trata os valores nulos e inconsistências.
	Output Text	Escreve dados em um arquivo.	Salva no arquivo(FatosAtendimentosExtracao.txt) as transformações.
	JavaScript	Executa uma transformação em JavaScript.	Evita que o SurrogateKey seja nulo, realiza mapeamento de Faixa Etária e Horário de Atendimento.

Figura B.27 Componentes das Transformações dos Atendimentos.

Descrição da Transformação:

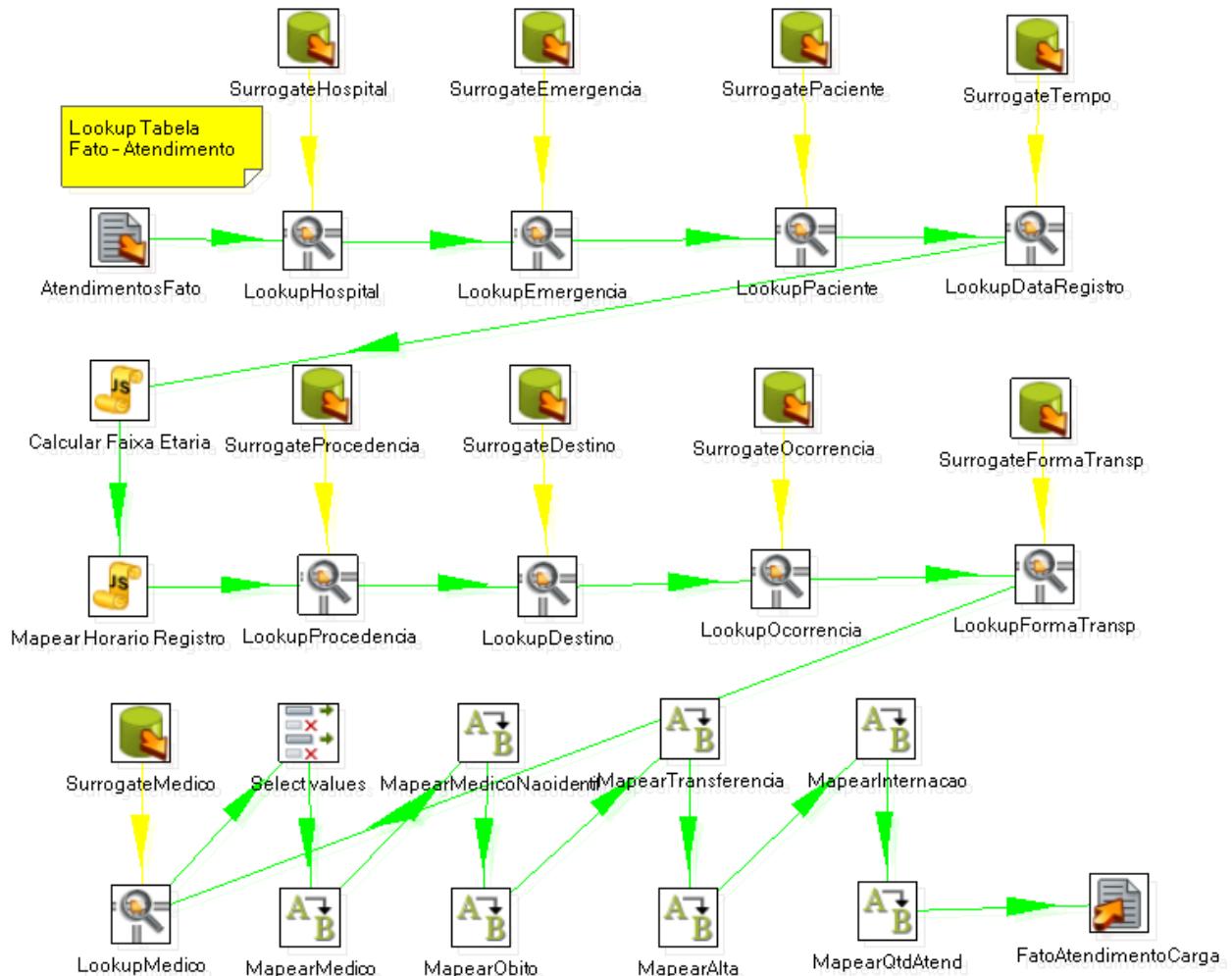


Figura B.28 Transformação Transformações dos Atendimentos.

O passo AtendimentoFato lê os registros de atendimento extraídos dos atendimentos.

Os passos do tipo lookup realizam uma procura nas *SurrogateKeys*. Pode-se notar nesta transformação, a importância do armazenamento da relação das *NaturalKeys* com as *SurrogateKeys*, processo realizado nas transformações citadas anteriormente, o que possibilitou recuperar as *SurrogateKeys* geradas nas tabelas de dimensão.

Um passo do tipo lookup deve ter dois passos como entrada, onde uma das entradas deve ser o stream principal, responsável por disponibilizar a chave para procurar no stream secundário onde é retornado o campo alvo. Na transformação exibida nesta seção o stream principal é indicado por uma flecha verde e o secundário onde é feita a procura, por uma flecha amarela.

O componente *JavaScript* Calcular Faixa Etária realiza o cálculo da faixa etárias do paciente

a partir da data de nascimento ou idade estimata, extraídas juntamente com os atendimentos. O outro componente Mapear Horário Registro identifica em que faixa de horário encontra-se o horário do atendimento.

O próximo passo do tipo Value Mapper é útil para trocar valores determinados por outros. Nesta transformação o Value Mapper foi utilizado para trocar valores nulos por outros valores padrões e realizar o mapeamento de todas as medidas, identificando os atendimentos, óbitos, internações e transferências.

E finalmente temos o passo FatoAtendimentoCarga que salva os fatos em um arquivo (FatoAtendimentoCarga) texto.

Execução da Transformação:

#	Nome do step	Copia nr	Lidos	escritos	Entrada	Saída	Atualizados	Rejected	Erros	Ativo	Tempo
1	AtendimentosFato	0	0	1824927	1824928	0	0	0	0	Finished	1676.7
2	LookupHospital	0	1824931	1824927	0	0	0	0	0	Finished	1676.7
3	SurrogateHospital	0	0	4	4	0	0	0	0	Finished	2.5
4	SurrogateEmergencia	0	0	66	66	0	0	0	0	Finished	2.6
5	LookupEmergencia	0	1824993	1824927	0	0	0	0	0	Finished	1677.0
6	SurrogatePaciente	0	0	1178914	1178914	0	0	0	0	Finished	1031.3
7	LookupPaciente	0	3003841	1824927	0	0	0	0	0	Finished	1677.2
8	LookupDataRegistro	0	1831927	1824927	0	0	0	0	0	Finished	1677.3
9	SurrogateTempo	0	0	7000	7000	0	0	0	0	Finished	3.1
10	SurrogateProcedencia	0	0	100	100	0	0	0	0	Finished	2.6
11	LookupProcedencia	0	1825027	1824927	0	0	0	0	0	Finished	1677.7
12	SurrogateDestino	0	0	134	134	0	0	0	0	Finished	2.5
13	LookupDestino	0	1825061	1824927	0	0	0	0	0	Finished	1677.7
14	SurrogateOcorrencia	0	0	184	184	0	0	0	0	Finished	2.5
15	LookupOcorrencia	0	1825111	1824927	0	0	0	0	0	Finished	1677.7
16	SurrogateFormaTransp	0	0	8	8	0	0	0	0	Finished	2.7
17	LookupFormaTransp	0	1824935	1824927	0	0	0	0	0	Finished	1677.7
18	SurrogateMedico	0	0	15635	15635	0	0	0	0	Finished	3.3
19	LookupMedico	0	1840562	1824927	0	0	0	0	0	Finished	1677.7
20	MapearMedico	0	1824927	1824927	0	0	0	0	0	Finished	1677.7
21	MapearMedicoNaoidentif	0	1824927	1824927	0	0	0	0	0	Finished	1677.7
22	Select values	0	1824927	1824927	0	0	0	0	0	Finished	1677.7
23	MapearObito	0	1824927	1824927	0	0	0	0	0	Finished	1677.8
24	MapearTransferencia	0	1824927	1824927	0	0	0	0	0	Finished	1677.8
25	MapearAlta	0	1824927	1824927	0	0	0	0	0	Finished	1677.9
26	MapearInternacao	0	1824927	1824927	0	0	0	0	0	Finished	1678.0
27	MapearQtdAtend	0	1824927	1824927	0	0	0	0	0	Finished	1678.0
28	FatoAtendimentoCarga	0	1824927	1824927	0	1824928	0	0	0	Finished	1678.2
29	Calcular Faixa Etaria	0	1824927	1824927	0	0	0	0	0	Finished	1677.4
30	Mapear Horario Registro	0	1824927	1824927	0	0	0	0	0	Finished	1677.6

Figura B.29 Execução da Transformação dos Fatos.

B.1.5.3 Carga dos Fatos no DW

Requisitos: Extração dos Atendimentos, Extração dos Fatos para o Esquema Estrela.

Objetivo: Realizar a carga das transformações realizadas para o DW.

Componentes Utilizados:

Step	Tipo	Descrição	Objetivo
	Input Text	Este componente lê os registros armazenados como CSV de um arquivo e os passa para o próximo step.	Lê os registros do arquivo FatoAtendimentoCarga.txt.
	Select Values	Este componente seleciona os campos definidos pelo projetista e descarta os que não foram selecionados. Também é possível renomear os campos selecionados.	Seleciona os campos: Código, Sequencial, Tempo_SK, CódigoEmergência_SK, Emergência_SK, Ind_AcidTrabalho_SK, Forma_Transporte_SK, Forma_Transporte_SK, PacienteNaoIdentificado_SK, Ind_PacienteNaoIdentificado_SK, SigaHospital_SK, Hospital_SK, Médico_SK, Destino_SK, Procedência_SK, Ocorrência_SK, TpOcorrência_SK, CódigoPaciente_SK, Paciente_SK, HrRegistro, Hora_Minuto, Atend_QTD, FlgRegistrarObito, Atend_QTD_Obito, FlgEhTransferencia, Atend_QTD_Transferencia, FlgEhAlta, Atend_QTD_Alta, FlgEhInternacao, Atend_QTD_Internacao.
	Table Output	Salva os registros em uma tabela do banco de dados.	Salva a transformação das emergências na tabela hospital do DW.

Figura B.30 Componentes das Carga dos Atendimentos.

Descrição da Transformação:

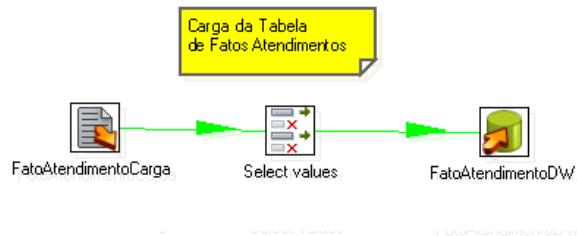


Figura B.31 Transformação Carga dos Atendimentos.

O primeiro passo FatoAtendimentoCarga lê o arquivo que contém os dados necessários para a carga, este arquivo foi previamente gerado pelas transformações: Extração dos Atendimentos e Extração dos Fatos para o Esquema Estrela.

O próximo passo (*Select values*) foi necessário para o próximo passo que insere as transformações no DW. *Select values* desempenha o papel de renomear e selecionar os campos que existem na tabela de destino (atendimento) apontado pelo passo FatoAtendimentoDW.

Execução da Transformação:

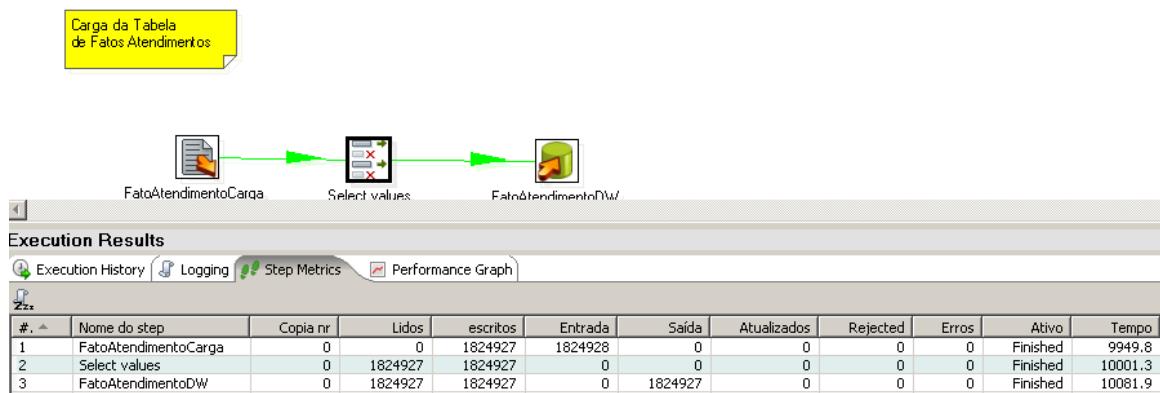


Figura B.32 Execução da Carga dos Atendimentos.

APÊNDICE C

Apêndice C

Neste Apêndice são apresentadas outros resultados obtidos e apresentados à Secretaria de Saúde do Estado de Pernambuco (SES-PE).

C.1 Atendimentos e Óbitos por Dia da Semana

A análise realizada anteriormente por Tipo de Ocorrência, reflete mais diretamente nos atendimentos e óbitos por dia da semana, apresentados na Figura C.1.

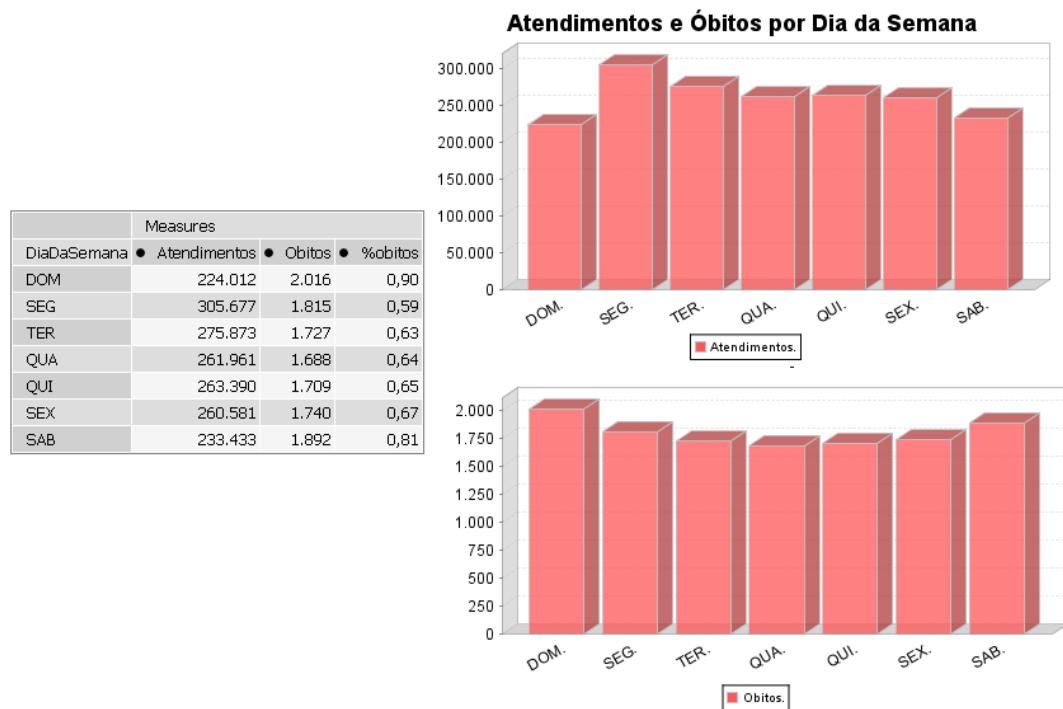


Figura C.1 Atendimentos e Óbitos por Dia da Semana.

Observa-se que os atendimentos seguem uma diminuição do atendimento durante a semana. O dia de maior atendimento é a segunda-feira.

Apesar de uma quantidade menor de atendimentos nos finais de semana, os óbitos vêm

ocorrendo em maior número, seguindo-se em ordem decrescente, nos domingos, sábado e segunda-feira. Associando-se aos Tipos de Ocorrências, apresentado na Figura 6.9, percebe-se uma relação direta entre os Tipos de Agressão e Acidentes com os óbitos. Com o aumento destas ocorrências, os médicos dos plantões nos finais de semana enfrentam casos mais graves, resultando em óbitos também na segunda-feira. Desta forma, profissionais mais qualificados em politraumatologias precisariam ser priorizados em plantões nestes dias da semana, onde a quantidade de profissionais pode não refletir em uma diminuição da mortalidade.

A análise realizada nas faixas dos horários, utilizando-se o mesmo critério, demonstram um aumento dos óbitos no período noturno, a partir das 22:00 horas, justificando-se pelo ciclo circadiano, ou seja, o período de 24 horas sobre qual se baseia o ciclo biológico do corpo humano, influenciado pela luz solar bem como as ocorrências de acidentes associados ao período noturno. Esta avaliação não apresenta grande importância neste contexto, visto que trata-se de uma informação bastante difundida na comunidade médica.

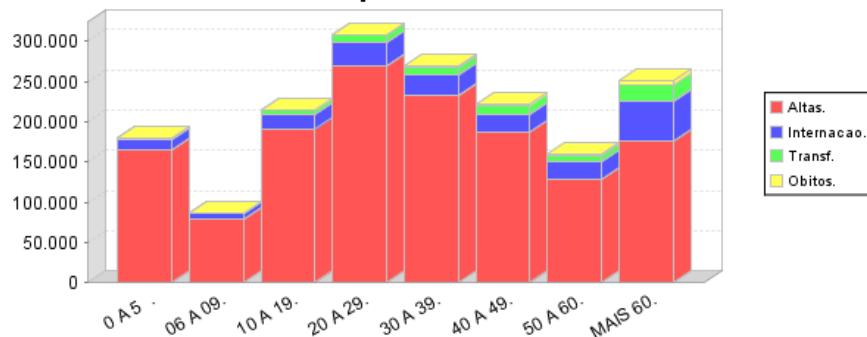
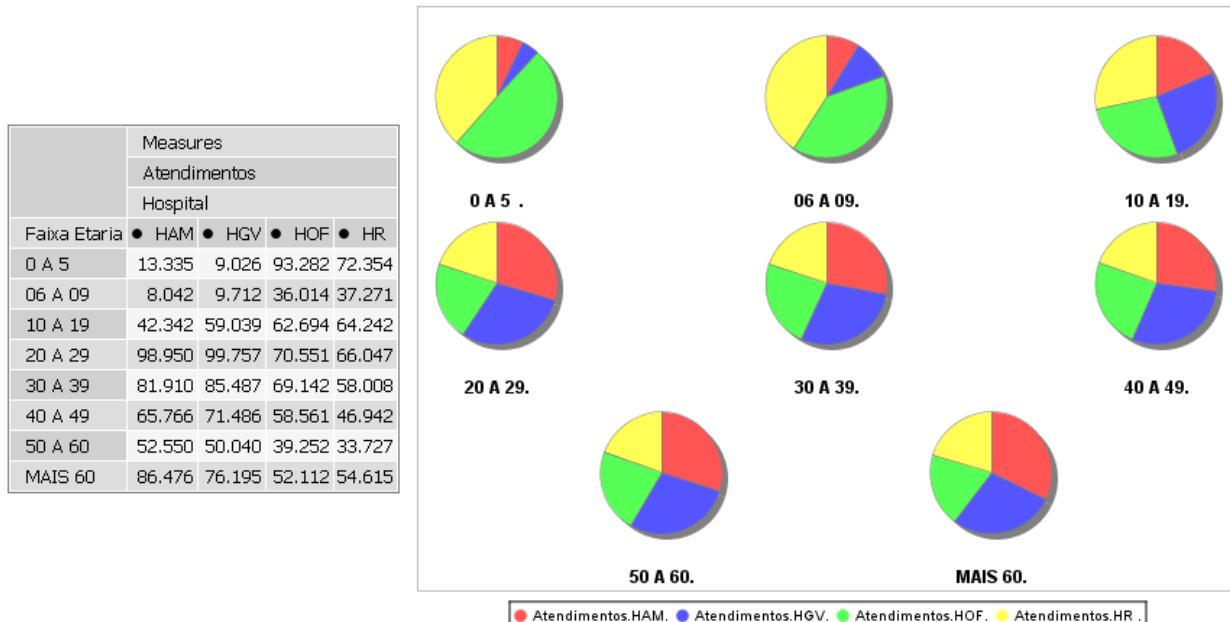
C.2 Faixas Etárias

A distribuição dos atendimentos por faixa etária são apresentados na Figura C.2, indicando que os atendimentos possuem maior quantidade na faixa de 20 a 29 anos, decaindo até a faixa de 50 a 60 anos, voltando a subir na faixa de mais de 60 anos. Os óbitos, com exceção da faixa de 0 a 5 anos, crescem proporcionalmente ao avanço das faixas.

Considerando-se as faixas etárias por hospital, apresentado na Figura C.3, nas faixas iniciais, 0 a 9 anos, o HR vem realizando um grande número de atendimentos, assim como o HOF. Nas demais faixas há uma distribuição mais regular. O aumento de ocorrências nas faixas etárias mais altas no HAM pode ser atribuído a especialização em doenças cardiológicas.

Uma sugestão para diminuição da quantidade de atendimentos no HR, inclusive viabilizando mais ocupação de leitos em outras faixas, é o direcionamento destas faixas etárias iniciais para outros hospitais, reforçando-se uma referência nesta especialidade, o Hospital Barão de Lucena (HBL) que já possui uma emergência pediátrica. A disponibilidade do espaço físico no HBL, viabilizando expansão de considerável magnitude contrastando com o espaço limitado do HR, possibilita um melhor direcionamento dos atendimentos nesta faixa etária.

Faixa Etaria	Measures									
	● Atendimentos	● Altas	● %Altas	● Internacao	● %Intern	● Transf	● %Transf	● Obitos	● %obitos	
0 A 5	187.997	165.814	88,20	12.007	6,39	2.119	1,13	121	0,06	
06 A 09	91.039	79.254	87,05	6.406	7,04	1.107	1,22	31	0,03	
10 A 19	228.317	190.013	83,22	19.396	8,50	4.499	1,97	453	0,20	
20 A 29	335.305	269.947	80,51	28.845	8,60	8.994	2,68	1.042	0,31	
30 A 39	294.547	233.658	79,33	25.167	8,54	9.534	3,24	1.129	0,38	
40 A 49	242.755	187.478	77,23	22.533	9,28	9.575	3,94	1.557	0,64	
50 A 60	175.569	128.626	73,26	21.355	12,16	8.745	4,98	1.926	1,10	
MAIS 60	269.398	175.395	65,11	51.061	18,95	18.687	6,94	6.328	2,35	

Atendimentos por Faixa Etária**Figura C.2** Faixas Etárias dos Pacientes.**Atendimentos por Faixa Etária e Hospital****Figura C.3** Faixas Etárias dos Pacientes por Hospital.

C.3 Acidente de Trabalho

Os acidentes de trabalho requerem do hospital um tratamento diferenciado pois estão associados tanto a causas trabalhistas como ao requerimento de licenças e aposentadorias à Previdência Social.

A distribuição dos atendimentos são apresentados na Figura C.4. Estes tipos de acidentes refletem em necessidades dos acidentados de comprovações de atendimentos e outros documentos legais solicitados tanto pelas empresas quanto pelos órgãos governamentais. Melhorias na prestação destes tipos de serviços nas unidades de saúde refletem diretamente no atendimento do direito de cidadania do paciente acidentado.

Acidente Trabalho	Hospital	Atendimentos	Obitos	%obitos	Internacao	%Intern	Transf	%Transf	Altas	%Altas
S	Todos os Hospitais	16.926	41	0,24	1.108	6,55	293	1,73	14.364	84,86
	HGV	8.708	13	0,15	545	6,26	163	1,87	7.571	86,94
	HR	6.067	26	0,43	455	7,50	44	0,73	4.973	81,97
	HOF	1.192	2	0,17	78	6,54	27	2,27	999	83,81
	HAM	959	0	0,00	30	3,13	59	6,15	821	85,61

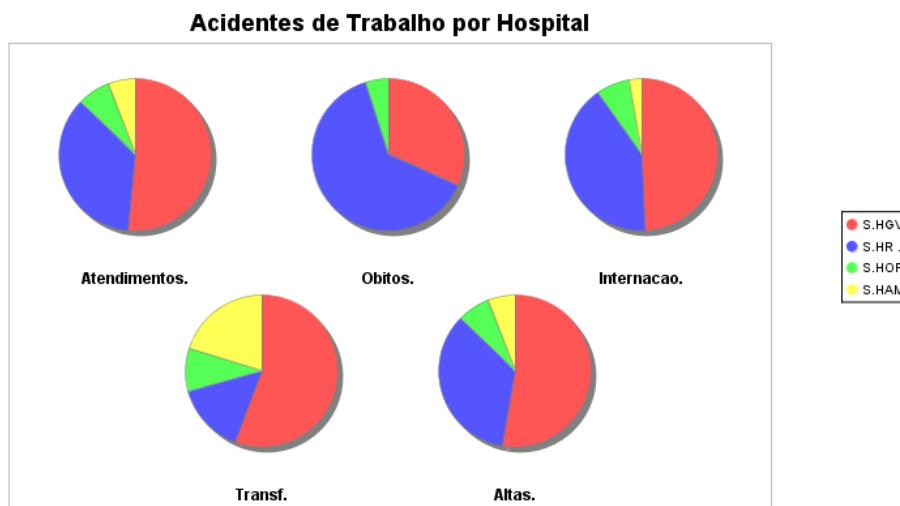


Figura C.4 Acidentes de Trabalho por Hospital.

Percebe-se que o HGV representa a maior quantidade de atendimentos, devido, principalmente, a maior proximidade das vias de acesso aos distritos industriais associado-se também à disponibilidade de tratamento especializado em traumatologia e ortopedia, conforme demonstrado na Figura C.5

Os acidentes associados a "Corte" na maioria das vezes refere-se a atividades relacionadas a abatedouros e frigoríficos, como resultado de acidentes com armas brancas. Há a necessidade de uma melhor classificação na tentativa de prevenção, visto que 2.353 casos, 13,90%, referem-se a "Outros Acidentes".

Acidente Trabalho	Ocorrencia	▼ Atendimentos	● Obitos	● Internacao	● Transf	● Altas
S	Corte	3.653	0	181	56	3.185
	Outros Acidentes	2.353	3	234	51	1.930
	Queda Fora de Casa	1.622	10	101	56	1.362
	Pancada	1.418	2	37	7	1.289
	Dor Membros Inferiores (Pés, Pernas, Joelho)	876	0	9	12	808
	Dor Membros Superiores (Mãos, Braços, Pulso)	845	0	20	9	770
	Outras Ocorrências Traumatológicas	659	0	19	12	549
	Acidente Moto - Queda, Colisão	540	3	32	12	455
	Esmagamento	401	1	85	6	284
	Corpo Est. no Olho	344	0	0	0	323
	Queimadura por Líquidos Quentes (Água, Café, Óleo)	342	0	20	0	300
	Encaminhamento Traumatologia	281	0	62	11	196
	Atropelamento	248	1	28	9	198
	Colisão com Outro Veículo	222	0	22	4	188
	Dor na Coluna	212	0	2	0	200
	Agressão Animal - Mordida, Picada	195	0	4	0	178
	Problema nos Olhos, Dificuldade/Alteração Visual	172	0	1	0	157
	Torção	147	0	1	0	124
	Choque Elétrico	144	7	20	3	107
	Outras Queimaduras	138	0	26	1	103
	Encaminhamento para Cirurgia	111	0	10	1	88
	Contusão	104	0	2	1	83
	Arma de Fogo	92	3	21	3	56
	Outros Acidentes de Trânsito	87	0	8	3	70

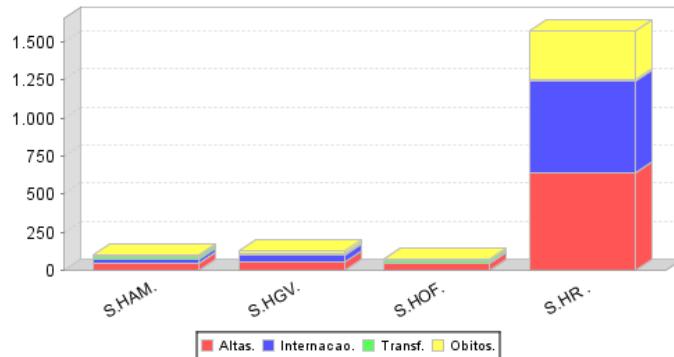
Figura C.5 Acidentes de Trabalho por Ocorrências.

C.4 Paciente Não Identificado por Hospital

Apesar de proporcionalmente representar somente 0,12% dos atendimentos, ou seja, 2.261 casos, constitui atribuição do Serviço Social de cada hospital a realização da identificação do paciente, seja durante ou após a realização do atendimento. A Figura C.6 apresenta a distribuição destas ocorrências e identifica-se o HR com as maiores incidências, 83,24%, inclusive com um alto índice de óbitos que permanecem sem a identificação.

Contudo, ao analisarem-se os Tipos de Ocorrências, a maioria relacionadas a acidentes e agressões, aparentemente explicam a grande incidência de não identificação na entrada, conforme apresentado na Figura C.7, porém não inviabilizando uma ação de gestão nestas situações.

Pac Nao Identificado	Hospital	Atendimentos	Altas	Internacao	Transf	Obitos
S	Todos os Hospitais	2.261	788	674	68	347
	HR	1.882	640	599	8	325
	HGV	145	53	46	16	12
	HAM	129	51	22	25	7
	HOF	105	44	7	19	3

Paciente Não Identificado por Hospital**Figura C.6** Pacientes Não Identificados por Hospital.

Pac Nao Identificado	Ocorrencia	Atendimentos	Obitos	Internacao	Transf	Altas
S	All Ocorrencia.Ocorrencias	2.261	347	674	68	788
	Atropelamento	378	68	153	6	106
	Arma de Fogo	327	134	98	15	31
	Agressao Fisica	156	12	42	3	76
	Acidente Moto - Queda, Colisao	121	33	42	2	32
	Frio, Tremor, Convulsao	113	4	26	0	68
	Outros Sintomas	96	9	24	4	24
	Desmaio	94	1	14	2	57
	Queda Fora de Casa	83	8	14	0	50
	Encaminhamento - Neuro Cirurgiao	73	8	27	0	28
	Outros Acidentes	64	7	19	6	26
	Mal Estar, Passando Mal	60	4	11	1	34
	Agonia, Nervosismo, Agitacao	59	0	5	12	28
	Arma Branca	58	4	21	2	19
	Colisao com Outro Veiculo	57	11	27	0	14
	Agressao Objeto Contundente	49	2	13	3	22
	Encaminhamento Neurologia	39	4	17	0	9
	Desorientacao, Confusao Mental, Alt. Comportamento	35	1	10	0	11
	Encaminhamento por TCE (Traumatismo Craneano Encef)	30	6	8	1	9
	Encaminhamento Clinica Medica	27	1	7	1	12
	Pancada	25	0	6	0	11
	Corte	24	0	3	4	13
	Capotamento	22	7	6	0	8
	Outros Acidentes de Transito	21	2	8	0	7

Figura C.7 Pacientes Não Identificados por Ocorrências.

C.5 Procedência dos Pacientes

Em uma análise das procedências dos pacientes com atendimentos realizados nas emergências detectou-se, como mostra a Figura C.8, que em sua grande maioria provem do domicílio, 75,99%.

Procedencia_Paciente	Hospital	Measures					
		• Atendimentos	• Obitos	• Internacao	• Transf	• Altas	
+ DOMICILIO	HAM	356.157	1.034	32.839	21.009	285.660	
	HGV	361.777	1.068	14.791	6.042	317.516	
	HOF	414.199	1.402	16.973	11.106	350.421	
	HR	254.470	1.057	17.102	2.757	207.334	
+ HOSPITAIS, MATERNIDADES	HAM	34.896	515	13.274	3.540	16.068	
	HGV	25.651	472	8.223	2.216	13.304	
	HOF	33.809	737	9.648	2.412	18.391	
	HR	92.253	3.010	30.482	2.313	50.089	
+ LOCAL DE TRABALHO	HAM	10.696	4	767	356	9.015	
	HGV	15.563	19	420	154	13.572	
	HOF	3.539	6	81	38	2.646	
	HR	10.377	34	372	51	8.065	
+ POLICLINICAS, S.P.A.s	HAM	14.666	185	3.876	1.605	8.456	
	HGV	11.737	165	2.930	745	7.389	
	HOF	7.136	111	1.879	342	4.330	
	HR	26.525	524	5.794	444	18.173	
+ VIA PUBLICA	HAM	4.387	61	210	592	3.253	
	HGV	19.673	204	2.217	1.061	15.079	
	HOF	8.338	106	713	868	5.833	
	HR	34.539	940	5.026	628	24.780	

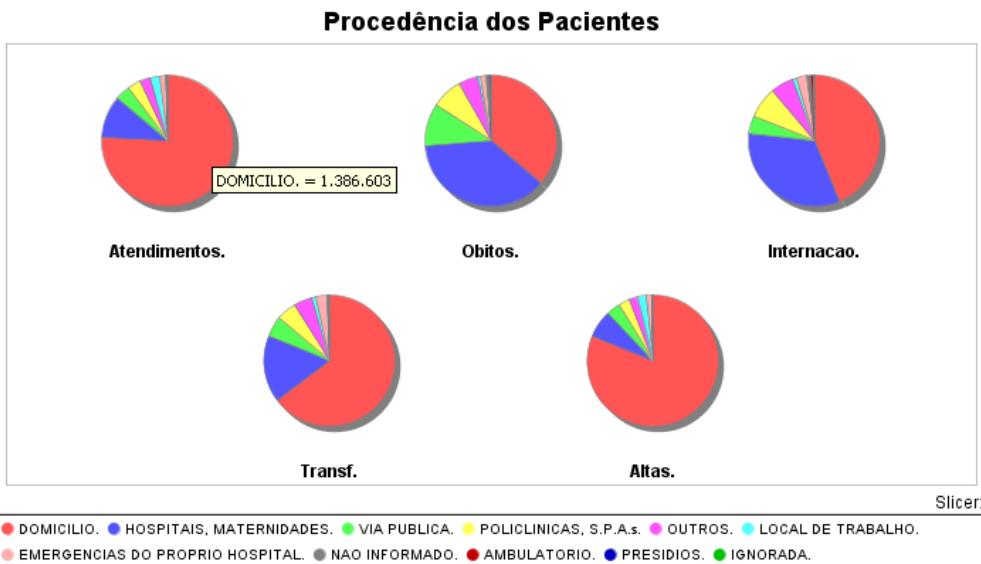


Figura C.8 Procedência dos Pacientes.

Esta quantidade desproporcional demonstra uma centralização dos atendimentos, com Pro-

cedimentos de Atenção Básica incluídos, realizados por hospitais estaduais.

Pela proposta de descentralização, os encaminhamentos, quando de ocorrências de casos de maior gravidade identificados nas unidades de saúde municipais, refletiriam em transferências para os hospitais de média e alta complexidade, reduzindo-se a procedência "Domicílio" dos pacientes e aumentando-se as procedências de Hospitais, Maternidades e Policlínicas.

A procedência Domicílio ocasiona a maior quantidade de óbitos dentre todas as procedências, enquanto Local de Trabalho e Via Pública representam os menores óbitos.

Com relação às Procedência das internações no Hospital da Restauração, os Hospitais e Maternidades corresponde a 50,54% e Policlínicas 59,98%. Estas informações comprovam o HR como o hospital de referência para as unidades de saúde.

C.6 Transferências e Destinos dos Pacientes

Parte dos pacientes das emergências, após um atendimento inicial, são transferidos para outros hospitais para a continuidade do tratamento em outra emergência, não se tratando necessariamente de internação.

A Figura C.9 demonstra que das 12.403 transferências realizadas, o HR corresponde a 3,32% do total. O maior índice refere-se ao HAM com 44,59%.

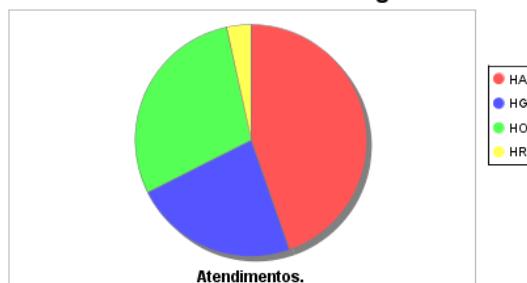
Com relação aos hospitais conveniados, o HOF realiza mais transferências, correspondendo a 43,33% do total, conforme apresentado na Figura C.10

Avaliando-se as Naturezas Acidente, Acidente de Trânsito e Agressão prevalece a transferência para o HR, conforme a Figura C.11, o que pode requerer uma atenção especial da SES-PE tanto na preparação para esta realidade quanto da opção por redimensionar as demandas de transferências entre as várias unidades de saúde.

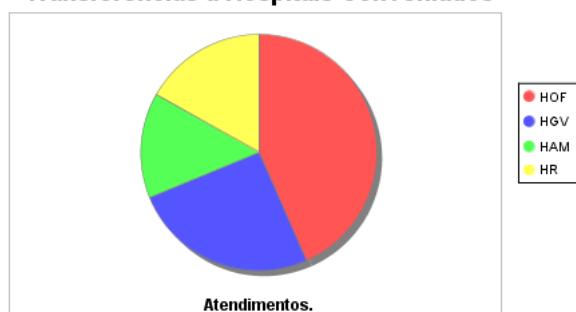
A Figura C.12 representa a quantidade de destinos ignorados e evasão de pacientes durante o atendimento nas emergências que necessitam de um tratamento mais rigoroso, pois há implicações legais nestes casos. O HAM, HOF e o HR representam o maior percentual, cabendo ao Serviço Social destes hospitais um melhor acompanhamento mesmo representando um percentual muito pequeno em relação aos atendimentos, 0,36% dos atendimentos totais.

A SES-PE também realiza convênios com outras unidades de saúde para a transferência de pacientes tanto para realização de tratamentos específicos como para liberação de leitos. A Figura C.13 apresenta os principais destinos conveniados, possibilitando uma visão geral das transferências. Para uma melhor análise, o gestor pode classificar por hospital, incluir a dimensão ano, tipo de ocorrência, faixa etária, sexo ou qualquer outra que desejar, facilitando

Hospital	Destino	Atendimentos
Todos os Hospitais	PACIENTES ENCAMINHADOS EMERG OUTROS HOSP	12.403
	TRANSFERENCIA HGA	284
	TRANSFERENCIA HGV	2.427
	TRANSFERENCIA HOF	1.832
	TRANSFERENCIA HR	7.860
HAM	PACIENTES ENCAMINHADOS EMERG OUTROS HOSP	5.530
	TRANSFERENCIA HGA	101
	TRANSFERENCIA HGV	1.543
	TRANSFERENCIA HOF	775
	TRANSFERENCIA HR	3.111
HGV	PACIENTES ENCAMINHADOS EMERG OUTROS HOSP	2.852
	TRANSFERENCIA HGA	69
	TRANSFERENCIA HOF	672
	TRANSFERENCIA HR	2.111
HOF	PACIENTES ENCAMINHADOS EMERG OUTROS HOSP	3.609
	TRANSFERENCIA HGA	87
	TRANSFERENCIA HGV	884
	TRANSFERENCIA HR	2.638
HR	PACIENTES ENCAMINHADOS EMERG OUTROS HOSP	412
	TRANSFERENCIA HGA	27
	TRANSFERENCIA HOF	385

Transferências entre as Emergências**Figura C.9** Transferências Entre as Emergências dos Hospitais.

Hospital	Destino	Atendimentos
Todos os Hospitais	PACIENTES ENCAMINHADOS HOSP CONVENIADOS	33.286
	HOF	14.422
	HGV	8.474
	HAM	4.807
	HR	5.583

Transferências à Hospitais Conveniados**Figura C.10** Transferências das Emergências para Hospitais Conveniados.

Tipo_Ocorrencia	Hospital	Destino	● Transf
+ ACIDENTE	HAM	TRANSFERENCIA HGA	8
		TRANSFERENCIA HGV	178
		TRANSFERENCIA HOF	74
		TRANSFERENCIA HR	319
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	16
	HGV	TRANSFERENCIA HGA	1
		TRANSFERENCIA HOF	46
		TRANSFERENCIA HR	363
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	921
	HOF	TRANSFERENCIA HGA	5
		TRANSFERENCIA HGV	127
		TRANSFERENCIA HR	489
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	450
	HR	TRANSFERENCIA HGA	4
		TRANSFERENCIA HOF	29
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	454
+ ACIDENTE DE TRÂNSITO	HAM	TRANSFERENCIA HGA	1
		TRANSFERENCIA HGV	50
		TRANSFERENCIA HOF	22
		TRANSFERENCIA HR	84
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	7
	HGV	TRANSFERENCIA HGA	1
		TRANSFERENCIA HOF	39
		TRANSFERENCIA HR	401
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	519
	HOF	TRANSFERENCIA HGV	56
		TRANSFERENCIA HR	188
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	220
	HR	TRANSFERENCIA HOF	27
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	564
+ AGRESSÃO	HAM	TRANSFERENCIA HGA	1
		TRANSFERENCIA HGV	121
		TRANSFERENCIA HOF	36
		TRANSFERENCIA HR	225
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	3
	HGV	TRANSFERENCIA HGA	4
		TRANSFERENCIA HOF	30
		TRANSFERENCIA HR	219
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	43
	HOF	TRANSFERENCIA HGA	4
		TRANSFERENCIA HGV	97
		TRANSFERENCIA HR	184
	HR	+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	26
		TRANSFERENCIA HGA	7
		TRANSFERENCIA HOF	24
		+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	93
+ OCORRÊNCIA TRAUMATOLÓGICA	HAM	TRANSFERENCIA HGV	1
		TRANSFERENCIA HGA	2
		TRANSFERENCIA HOF	59
		TRANSFERENCIA HR	71
	HGV	+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	1.639
		TRANSFERENCIA HGA	106
		TRANSFERENCIA HOF	149
	HOF	+ PACIENTES ENCAMINHADOS HOSP CONVENIADOS	1.360
		TRANSFERENCIA HGV	3
	HR	TRANSFERENCIA HOF	629

Figura C.11 Transferências por Tipo de Ocorrência por Hospital.

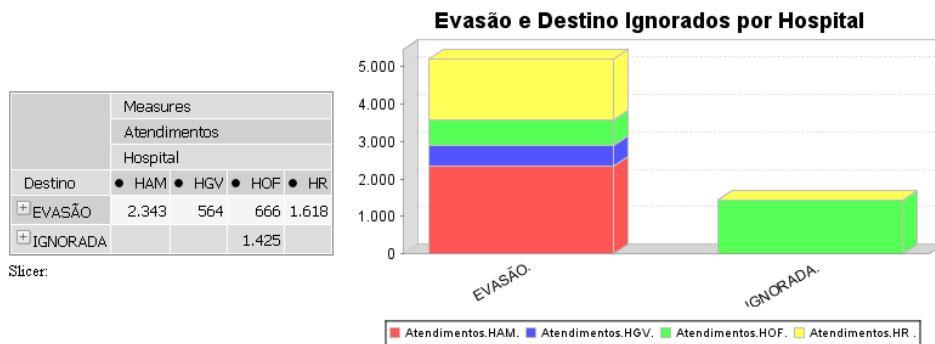


Figura C.12 Evasão e Destino Ignorado por Hospital.

as conclusões para ações estratégicas nos convênios compactuados pela Secretaria.

C.7 Internações Hospitalares

As Ocorrências associadas às internações estão apresentadas na Figura C.14. Apesar de "Outros Sintomas" comprometer uma melhor análise, com 15,33%, os encaminhamentos de outras unidades de saúde correspondem a 36,56% das internações. Análises sobre os casos de internamento, possibilitando prevenções, podem resultar em considerável economia para o Estado, visto que a identificação do problema pode minimizar sua ocorrência. Percebe-se também nestes casos a importância do registro do diagnóstico médico e do CID no sistema.

Destino	Atendimentos					
	Hospital					
	▼	Todos os Hospitais	● HAM	● HGV	● HOF	● HR
PACIENTES ENCAMINHADOS HOSP CONVENIADOS		33.286	4.807	8.474	14.422	5.583
TRANSFERENCIA HOSPITAL EVANGELICO		3.573	897	1.825	141	710
TRANSFERENCIA HOSPITAL SANATORIO RECIFE		3.435			3.435	
TRANSFERENCIA HOSPITAL GERAL CAMARAGIBE		3.173	1	1.759	854	559
CLINICA SANTO ANTONIO		2.655			2.655	
TRANSFERENCIA HOSPITAL MEMORIAL GUARARAPES		2.025	146	849	673	357
TRANSFERENCIA HOSPITAL N.S. DE LOURDES		1.986	333	1.069	369	215
TRANSFERENCIA HOSPITAL PADRE ANTONIO MANOEL		1.983	439	111	1.433	
TRANSFERENCIA H.P.P. DE BOA VIAGEM		1.944			1.944	
TRANSFERENCIA HOSPITAL SAO JOSE		1.913	26	677	772	438
TRANSFERENCIA HOSPITAL CENTRAL DE PAULISTA		1.305	274	503	260	268
TRANSFERENCIA HOSPITAL GERAL JABOATAO		1.020	108	307	369	236
TRANSFERENCIA HOSPITAL OSWALDO CRUZ		990	83	291	238	378
TRANSFERENCIA HBL		863	222	313	131	197
TRANSFERENCIA HOSPITAL SANTO AMARO		631	308	161	50	112
TRANSFERENCIA HOSPITAL CORREIA PICANCO		619	107	80	209	223
TRANSFERENCIA HOSPITAL DAS CLINICAS		546	96	96	78	276
TRANSFERENCIA HOSPITAL MATERNIDADE TRICENTENARIO		510	510			
TRANSFERENCIA HOSPITAL HELENA MOURA		495	18			477
TRANSFERENCIA IMIP		471	96	42	119	214
TRANSFERENCIA HOSPITAL MARIA LUCINDA		444	197	61	10	176
TRANSFERENCIA MATERNIDADE BARROS LIMA		396	396			
TRANSFERENCIA HOSPITAL DA MIRUEIRA		262		4	258	
TRANSFERENCIA APAMI		239	236	2	1	
TRANSFERENCIA HOSPITAL ULISSES PERNAMBUCANO		219	93	13	70	43
TRANSFERENCIA HOSPITAL IPSEP		218	49	76	33	60
TRANSFERENCIA HGV		201				201
TRANSFERENCIA CISAM		139	48	44	30	17
TRANSFERENCIA HOSPITAL MARTINIANO FERNANDES		118	4	88	10	16

Figura C.13 Destino Pacientes para Hospitais Conveniados.

Ocorrencia	● Atendimentos	▼ Internacao	● %Intern
Outros Sintomas	241.088	28.623	11,87
Encaminhamento Clinica Medica	66.885	13.242	19,80
Encaminhamento Neurologia	35.931	11.856	33,00
Dor	70.919	7.776	10,96
Outros Encaminhamentos por Doenca	21.459	7.296	34,00
Mal Estar, Passando Mal	61.770	6.695	10,84
Encaminhamento para Cirurgia	23.280	6.188	26,58
Encaminhamento para Parecer	26.447	6.070	22,95
Encaminhamento Emergencia Cardiologica	16.894	5.977	35,38
Falta de Ar, Cansaco, Dificuldade para Respirar	84.429	5.241	6,21
Dor Abdominal	54.949	4.440	8,08
Encaminhamento Traumatologia	27.467	4.329	15,76
Nao Informado	7.826	4.108	52,49
Arma de Fogo	10.501	3.983	37,93
Outras Ocorrencias Traumatologicas	97.917	3.063	3,13
Encaminhamento - Cirurgiao Geral	7.745	2.872	37,08
Encaminhamento Emergencia Geral	8.485	2.818	33,21
Acidente Moto - Queda, Colisao	20.843	2.704	12,97
Encaminhamento - Neuro Cirurgiao	9.995	2.692	26,93
Atropelamento	15.107	2.591	17,15
Vomito	40.298	2.520	6,25
Encaminhamento por HDA (Hemorragia Digestiva Alta)	5.215	2.440	46,79
Internacao	2.838	2.166	76,32
Febre	65.689	1.973	3,00
Dor de Cabeca	51.213	1.911	3,73
Inchaco, Edema	14.932	1.875	12,56
Sangramento	15.230	1.838	12,07
Pressao Alta	26.903	1.661	6,17
Arma Branca	5.610	1.636	29,16
Queda Fora de Casa	25.405	1.524	6,00
Encaminhamento - Cirurgiao Vascular	3.689	1.463	39,66
Frio, Tremor, Convulsao	6.261	1.409	22,50
Agonia, Nervosismo, Agitacao	35.145	1.385	3,94
Dor Membros Superiores (Maos, Bracos, Pulsos)	54.609	1.329	2,43
Queda no Domicilio	20.646	1.296	6,28
Outros Acidentes	17.584	1.260	7,17
Dor Membros Inferiores (Pes, Pernas, Joelho)	62.569	1.239	1,98
Corte	42.583	1.124	2,64
Dor no Peito	16.124	1.074	6,66
Queimadura por Liquidos Quentes (Agua,Cafe,Oleo)	4.710	1.060	22,51
Encaminhamento por AVC (Acidente Vascular Cerebral)	2.210	1.041	47,10
Desmaio	10.191	828	8,12
Encaminhamento por TCE (Traumatismo Craneano Encef	4.665	764	16,38
Outros Encaminhamentos para Procedimentos	3.194	758	23,73
Colisao com Outro Veiculo	5.635	745	13,22
Agressao Fisica	11.181	729	6,52

Figura C.14 Atendimentos e Internações por Ocorrências.

Referências Bibliográficas

- [Ale02] P. L. Alenquer. Regras de negócio para análise em ambientes olap. 2002. dissertação (mestrado em informática). Master's thesis, Instituto de Matemática - IM, Núcleo de Computação - NCE, Universidade Federal do Rio de Janeiro, 2002.
- [Alm06] S. et al Almeida. Getting started with datawarehouse and business intelligence. *IBM Corporation*, 2006.
- [ate09] Atech tecnologias críticas. disponível em <http://www2.atech.br/saude.atech>. acesso em 22 mar. 2009, 2009.
- [Bar01] Carlos Barbieri. *Business Intelligence: modelagem e tecnologia*. Rio de Janeiro: Axcel Books, 2001.
- [Ber07] A. Berson. *Data Warehousing, Data Mining & OLAP*. USA: McGraw-Hill, 2007.
- [Bis04] C. A. F. Bispo. Uma análise da nova geração de sistemas de apoio à decisão. 2004. dissertação (mestrado em engenharia de produção). Master's thesis, Escola de Engenharia de São Carlos, Universidade de São Paulo, 2004.
- [BPD08] Eder Luis Both, Edson Luis Padoin, and Sérgio Luis Dill. Um sistema de apoio à decisão para secretarias municipais de saúde. disponível em <http://www.unijui.edu.br/detec>. acesso em 02 mai. 2009, 2008.
- [Bro00] M. K. Brohman. The business intelligence value chain: Data-driven decision support in a data warehouse environment: an exploratory study. *The 37th Hawaii International Conference on System Sciences, IEEE, Hawaii*, 2000.
- [CAAT01] M Corey, M. Abbey, I. Abramson, and B. Taub. *Oracle 8i Data Warehouse*. Rio de Janeiro: Editora Campus, 2001.
- [cel09] Companhia de informática do paraná. disponível em <http://www.celepar.pr.gov.br/modules/conteudo/conteudo.php?conteudo=1>. acesso em 02 mai. 2009, 2009.

- [Cer02] J. N. C Cerqueira. Afinal de contas, o que é bi?, 2002. Acesso em 01/04/2007.
- [Cie05] Ivã Cielio. Arquiteturas olap. disponível em <http://www.datawarehouse.inf.br/artigos/olap.asp>. acesso em: 20 de set. 2005. 2005.
- [Cor02] S Cortes. Bi, data warehouse e data mining - como a tecnologia aumenta a inteligência do negócio., 2002.
- [Cun04] José Antonio da Cunha. Arquiteturas de ferramentas olap. disponível em: <http://www.engcomp.ufrn.br/~max/cefet/banco2/arquiteturaolap.ppt>. acesso em 20 set. 2007. 2004.
- [dat08] Datasus - ministério da saúde. disponível em <http://www.datasus.gov.br>. acesso em 05 mai. 2008., 2008.
- [dbd09] Dbdesigner: Fabulous force database tools. disponível em <http://fabforce.net/dbdesigner4/>. acesos em 06 mai. 2009, 2009.
- [Dev01] J. Devaney. Business intelligence strategies. *Microsoft Press International*, 2001.
- [DP99] T. Davenport and L. Prusak. *Conhecimento Empresarial*. Editora Campus, Rio de Janeiro, RJ, Brasil, 1999.
- [ecl09] About the eclipse foundation. disponível em <http://www.eclipse.org/home/newcomers.php>. acesso em 14 abr. 2009, 2009.
- [Fur97] J. D. Furlan. *Modelagem de negócio*. São Paulo: Makron Books, 1997.
- [GC08] Dirceu Grein and Edson Chitz. Sistema de informação gerencial utilizando solução de data warehouse em software livre. disponível em http://www.circuito.pr.gov.br/arquivos/file/anteriores2007/sig_usando_dw_em_sl.pdf. acesso em 02 mai. 2009. 2008.
- [Gio00] W. A. Giovinazzo. *Object-Oriented Data Warehouse Design - Building a Star Schema*. New Jersey: Prentice Hall, 2000.
- [Gon05] T. Gonzaga. Uma metodologia para o desenvolvimento de instrumentos de análise multidimensional da informação em projetos de governo eletrônico voltado ao cidadão. Dissertação (mestrado em engenharia de produção)., Universidade Federal de Santa Catarina, Florianópolis, 2005.

- [gsa08] Comunidade gsan. disponível em http://www.softwarepublico.gov.br/community_id=1593449. acesso em 28 set. 2008, 2008.
- [How00] Robert et al. Howard. *Aprendizado Organizacional: gestão de pessoas para inovação contínua*. In [How00], 2000.
- [ibm09] Servidores ibm em torre. disponível em <http://www.ibm.com/br/systems/x/tower/index.phtml>. acesso em 10 mar. 2009., 2009.
- [Inm02] W.H. Inmon. *Building the Data Warehouse*. John Wiley & Sons, New York, NY, USA, 2002.
- [jav09] Java language. disponível em <http://java.sun.com>. acesso em 02 mai. 2009., 2009.
- [jpa09a] Japan public health association. disponível em <http://www.jpha.or.jp/jpha/english/>. acesso em 04 mai. 2009, 2009.
- [jpa09b] Public health of japan 2008. disponível em <http://www.jpha.or.jp/jpha/english/apha/publiccpdf>. acesso em 10 mai. 2009, 2009.
- [jpi09] Jpivot - a jsp based olap client. disponível em <http://sourceforge.net/projects/jpivot/>. acesso em 25 abr. 2009, 2009.
- [KC04] Ralph Kimball and Joe Caserta. *The Data Warehouse ETL Toolkit*. New York: John Wiley & Sons, 2004.
- [Kee00] M.S. Keen, Peter G. W. e Morton. *Decision Support System: An Organization Perspective*. Addison-Wesley, San Antonio, CA, USA, 2000.
- [Kim98] R. et al. Kimball. *The Data Warehouse lifecycle toolkit: expert methods for designing, developing and deploying Data Warehouses*. New York: John Wiley & Sons, 1998.
- [KR02] R. Kimball and M. Ross. *The Data Warehouse Toolkit*. Editora Campus, Rio de Janeiro, RJ, Brasil, 2002.
- [Lam07] Lirce Lamounier. A organização do sus. *Faculdade de Farmácia da Universidade Federal de Goiás - NECAF*, 2007.
- [Lau01] J. Laudon, K. e Laudon. *Gerenciamento de Sistemas de Informação*. LTC-Livros Técnicos e Científicos, Rio de Janeiro, RJ, Brasil, 2001.

- [Lie07] J. Liebowitz. Where is the "intelligence"behind "business intelligence"? , 12 2007. Acesso em 20 dez. 2007.
- [Mac00] Fernando Numeriano Rebouças Machado. *Projeto de Data Warehouse: Uma Visão Multidimensional*. São Paulo: Érica, 2000.
- [MC98] D. Meyer and Cannon C. *Building a better data warehouse*. NJ: Prentice-Hall, 1998.
- [mdx09] Mdx overview. microsoft corporation. disponível em: <http://msdn.microsoft.com/library/default.asp?url=/library/enus/olapdmad/agmdxbasics90qg.asp>. acesso em 02 mai. 2009, 2009.
- [MGS05] Simone Marinai, Marco Gori, and Giovanni Soda. Artificial neural networks for document analysis and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(1):23–35, 2005.
- [mon09] Mondrian, olap engine. disponível em <http://sourceforge.net/projects/mondrian/>. acesso em 25 abr. 2009, 2009.
- [Mos03] L. T. Moss. *Business Intelligence roadmap: the complete project lifecycle for decision-support applications*. Addison-Wesley, San Antonio, CA, USA, 2003.
- [msa08] Atribuições do ministério da saúde. disponível em <http://portal.saude.gov.br/portal/saude>. acesso em 04 mai. 2008, 2008.
- [mso09] Windows server 2003. disponível em [http://technet.microsoft.com/pt-br/windowsserver/bb512919\(en-us\).aspx](http://technet.microsoft.com/pt-br/windowsserver/bb512919(en-us).aspx), acesso em 07 mai. 2009, 2009.
- [nob08] Norma operacional básica. disponível em: <http://portal.saude.gov.br/portal/arquivos/pdf/nob96.pdf>. acesso em 25 mai 2008, 2008.
- [oms09] World health organizacion. disponível em <http://www.who.int/en/>. acesso em 26 abr. 2009., 2009.
- [ope09] Openi: Bi web app for saas deployments. disponível em <http://sourceforge.net/projects/openi/>. acesso em 07 abr. 2009., 2009.
- [ora09] Oracle business intelligence discoverer. disponível em <http://www.oracle.com/technology/products/discoverer/index.html>. acesso em 20 mai. 2009, 2009.

- [pal09] Palo bi suite. disponível em <http://sourceforge.net/projects/palo/>. acesso em 10 mai. 2009, 2009.
- [ped08] Pe digital. disponível em <http://www.pemultidigital.com.br>. acesso em 20 out. 2008, 2008.
- [Pen00] N. Pendse. What is olap? disponível em <http://www.olapreport.com/fasmi.htm>. acesso em 04 set. 2007., 2000.
- [pen08] Pentaho artigo técnico. disponível em <http://www.pentaho.org>. acesso em 25 set. 2007., 2008.
- [pen09] Pentaho - business intelligence. disponível em <http://sourceforge.net/projects/pentaho>. acesso em 02 mai. 2009, 2009.
- [PKB98] V. Poe, P. Klauer, and S. Brobst. *Building a data warehouse for decision support*. NJ: Prentice-Hall PTR, 1998.
- [pos08] Postgresql 8.3.4 documentation. disponível em <http://www.postgresql.org/docs/>. acesso em 01 out. 2008. 2008.
- [pwc09] Jpalo - palo java world. disponível em http://www.jpalo.com/en/products/start_products.html. acesso em 28 jun. 2009, 2009.
- [sbi08] Sociedade brasileira de informática em saúde. disponível em <http://www.sbis.org.br.html>. acesso em 25 mar. 2009, 2008.
- [Sel01] D. Sell. Uma arquitetura para distribuição de componentes tecnológicos de sistemas de informação baseados em data warehouse. Dissertação (mestrado em engenharia de produção) programa de pós-graduação em engenharia de produção., Universidade Federal de Santa Catarina, Florianópolis, 2001.
- [ser08] O serpro. disponível em <http://www.serpro.gov.br/instituicao/quem>. acesso em 28 set. 2008, 2008.
- [ses06] Secretaria estadual de saúde de pernambuco. informações e processos identificados na rede de saúde do estado de pernambuco. 2006.
- [SGFT08] Ricardo S Santos, Marco Antonio Gutierrez, Sergio Furui, and Umberto Tachinardi. Projeto de um data warehouse para a saúde pública. departamento de informática em saúde. universidade federal de são paulo. serviço de informática,

- instituto do coração (incor). universidade de são paulo (usp). disponível em <http://www.sbis.org.br/cbis11/arquivos/846.pdf>. acesso em 20 mar. 2009. 2008.
- [sia09] Sia - sistema de informações ambulatoriais do sus. disponível em <http://w3.datasus.gov.br/datasus/index.php?area=040102>. acesso em 10/06/2009. 2009.
- [Sin01] H. S. Singh. *Data Warehouse*. São Paulo: Macron Books, 2001.
- [slb08] Projeto software livre brasil. disponível em <http://www.softwarelivre.org/the-project.php>. acesso em 28 set. 2008, 2008.
- [Som03] Ian Sommerville. *Engenharia de Software*. São Paulo: Pearson Education do Brasil, 2003.
- [sou09] Source forge, find and develop open-source software. disponível em <http://sourceforge.net>. acesso em 25 abr. 2009, 2009.
- [spa09] Spagobi. disponível em <http://forge.ow2.org/projects/spagobi/>. acesso em 07 abr. 2009, 2009.
- [Sys07] WG System. Arquitetura de um sistema de bi. 2007.
- [TdS08] Gerson Tessler and Fábio de Salles. Por dentro da pentaho open bi suite: Conceitos, arquitetura e componentes. disponível em <http://www.opentaho.org.br>. acesso em 30 set. 2008. 2008.
- [Tho02] E. Thomsen. *OLAP: Construindo Sistemas de Informação Multidimensionais*. Editora Campus, Rio de Janeiro, RJ, Brasil, 2002.
- [Une05] Unesco. Towards knowledge societies. disponível em: <<http://www.unesco.org/publications>>. acesso em: 20 nov. 2007. 2005.
- [VA08] Gilberto Capatina Valente and Newton Naoki Ahagon. A aplicação do business intelligence no segmento de saúde pública ambulatorial. atech tecnologias críticas, são paulo, sp. disponível em <http://www.sbis.org.br/cbis11/arquivos/846.pdf>. acesso em 22 mar. 2009. 2008.
- [whr00] The world health report 2000 - health systems: improving performance. disponível em <http://www.who.int/entity/whr/2000/en/whr00.pdf>. acesso em 26 abr. 2009, 2000.

- [Wik08] Wikipedia, 2008. Acesso em 02 mai. 2007.
- [xml08] Extensible markup language (xml). disponível em <http://www.w3.org/xml>. acesso em 28 nov. 2008, 2008.

Este volume foi tipografado em L^AT_EX na classe UFPEThesis (www.cin.ufpe.br/~paguso/ufpethesis).