



UNIVERSIDADE DO SUL DE SANTA CATARINA
LEONARDO BRAND BIANCO PEREIRA

BUSINESS INTELLIGENCE APLICADO AO TRANSPORTE COLETIVO

Palhoça
2008

LEONARDO BRAND BIANCO PEREIRA

BUSINESS INTELLIGENCE APLICADO AO TRANSPORTE COLETIVO

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

Orientador: Prof. Aran Morales, Dr.

Palhoça

2008

LEONARDO BRAND BIANCO PEREIRA

BUSINESS INTELLIGENCE APLICADO AO TRANSPORTE COLETIVO

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

Palhoça, 21 de novembro de 2008.

Orientador Prof. Aran Morales, Dr.
Universidade do Sul de Santa Catarina

Prof. Mauro Madeira, Dr.
Universidade do Sul de Santa Catarina

José Aloísio Cavallieri
Solução Tecnologia em Movimento

AGRADECIMENTOS

Ao Prof. Aran Morales, Dr. que me conduziu e orientou com excelência durante todas as etapas do desenvolvimento deste trabalho.

Ao Sr. Humberto Arantes diretor da empresa de transporte coletivo Transul, por ter viabilizado a realização do estudo caso.

À minha família e noiva pelo grande apoio e carinho, fundamentais para conclusão deste trabalho.

A todos que direta ou indiretamente contribuíram para a realização deste trabalho.

“Muitos dos fracassos desta vida estão concentrados nas pessoas que desistiram por não saberem que estavam muito perto da linha de chegada.” (Thomas Alva Edison).

RESUMO

O objetivo deste trabalho é apresentar e implementar uma arquitetura de *Business Intelligence* (BI) em uma empresa de transporte coletivo. O foco nesse segmento empresarial se deve à sua grande importância social e econômica, pois está relacionado ao principal meio de transporte da população dos médios e grandes centros urbanos. A referência bibliográfica concentra-se nos conceitos e nos tipos de sistemas de informação, seguindo pela evolução dos sistemas de apoio gerencial até o conceito de BI e seus principais elementos. Por fim, são apresentados os conceitos de *Data Warehouse* (DW), sua arquitetura e a ferramenta de *On-Line Analytical Processing* (OLAP) utilizada. A modelagem do sistema desenvolvido partiu do levantamento de requisitos e da modelagem dimensional do DW. O sistema-protótipo desenvolvido consiste em um DW e na visualização do seus dados através de uma ferramenta OLAP. Utilizou-se na implementação, o banco de dados Microsoft SQL Server 2005 Standard para armazenamento, o Microsoft Analysis Services como servidor OLAP, o Microsoft Integration Services para realizar a extração, transformação e carga do DW e, por fim, o Microsoft Excel 2007 para visualização e análise dos dados. Para validar o sistema-protótipo, comparou-se a quantidade de passageiros transportados por mês obtida através da ferramenta OLAP, com o mesmo indicador obtido através de uma consulta SQL (*Structured Query Language*) no banco de dados operacional. O estudo de caso realizado comprovou a aplicabilidade de uma arquitetura de BI em uma empresa de transporte coletivo.

Palavras-chave: *Business Intelligence*. *Data Warehouse*. Transporte Coletivo.

ABSTRACT

The purpose of this work is to present and implement an architecture of Business Intelligence in a public transport company. The focus on this enterprise segment it's related to its great social and economic importance, therefore it is related to main kind of transport of the population in the medium and the great urban centers. The bibliographical reference is concentrated in the concepts and the types of the information systems, followed by the evolution of support systems management until the concept of Business Intelligence and its main elements. Finally, are presented the data warehouse concepts, architecture and the used OLAP (On-Line Analytical Processing) tool. The developed system modeling begun from the requirements survey, which contemplates the functional and not functional rules, business requirements, use cases and the dimensional model. The prototype system developed consists of a Data Warehouse and his data visualization through an OLAP tool. To accomplish this objective it was used the Microsoft SQL database Server 2005 Standard for storage, the Microsoft Analysis Services as an OLAP server, Microsoft Integration Services to carry through the operational systems ETL and, finally, the Microsoft Excel 2007 for visualization and analysis of the data available in the OLAP server. To validate the developed prototype system, an amount of month carried passengers obtained through the OLAP tool, was compared with the same indicator obtained through a SQL query in the operational database. The case study proved the applicability of an architecture of Business Intelligence in a public transport company

Keywords: Business Intelligence. Data Warehouse. Public Transport.

LISTA DE FIGURAS

Figura 1 - Proposta de solução.	22
Figura 2 - Componentes de um sistema de informação.....	26
Figura 3 - Tipos de sistemas de informação.	27
Figura 4 - Arquitetura <i>Data Warehouse</i> global.	36
Figura 5 - Arquitetura de <i>Data Marts</i> independentes.....	37
Figura 6 - Arquitetura de <i>Data Marts</i> integrados.....	38
Figura 7 - Fluxo de dados em um ambiente de DW com ODS.....	43
Figura 8 - Os patamares de granularidade.....	44
Figura 9 - Composição básica de uma tabela fato.....	45
Figura 10 - Modelo <i>Star Schema</i> aplicado a tabela fato vendas.....	47
Figura 11 - Modelo <i>Snowflake</i>	48
Figura 12 - Operação <i>Drill Down</i> na dimensão geográfica.....	51
Figura 13 - Operação <i>Roll Up</i> na dimensão tempo.....	51
Figura 14 - Fatias (<i>slices</i>) de um cubo de dados.....	52
Figura 15 - Representação dos principais comandos dimensionais.....	53
Figura 16 - Regras de negócio.....	55
Figura 17 - Requisitos funcionais.....	59
Figura 18 - Requisitos não funcionais.....	60
Figura 19 - Diagrama de casos de uso.....	62
Figura 20 - Modelo lógico do banco de dados.....	64
Figura 21 - Modelo físico do banco de dados.....	65
Figura 22 - Solução proposta.....	67
Figura 23 – Interface de controle de fluxo do <i>Integration Services</i>	70
Figura 24 – Interface do <i>data flow</i> para carga da dimensão linhas no <i>Integration Services</i>	71
Figura 25 - Interface do <i>data flow</i> para carga da tabela fato no <i>Integration Services</i>	71
Figura 26 - Interface principal do <i>Analysis Services</i>	72
Figura 27 - Interface da ferramenta OLAP Excel 2007.....	72
Figura 28 - Quantidade passageiros por mês do DW, via OLAP.....	73
Figura 29 - Quantidade de passageiros por mês do banco de dados operacional (OLTP).....	74
Figura 30 - Índice de passageiros transportados por tipo de pagamento.....	75
Figura 31 - Índice de passageiros transportados por faixa de horários.....	76

Figura 32 - Índice de passageiros por linha.	76
Figura 33 - Índice de passageiros transportados por mês.	77
Figura 34 - Índice de estudantes transportados por faixa etária.	78
Figura 35 - Índice de passageiros estudantes transportados por sexo.	78
Figura 36 - Índice de passageiros que utilizam o vale-transporte por sexo.	79
Figura 37 - Índice de passageiros que utilizam o vale-transporte por linha.	80

LISTA DE GRÁFICOS

Gráfico 1 - Motivos da substituição do ônibus/microônibus por classe socioeconômica.....	18
Gráfico 2 - Principais modos utilizados por classe socioeconômica.....	19

LISTA DE QUADROS

Quadro 1 - Características dos SI por nível.....	27
Quadro 2 - Características do <i>Data Warehouse</i>	34
Quadro 3 - Tipos de implementação.....	39
Quadro 4 - Vantagens e desvantagens das abordagens de implementação.	39
Quadro 5 - Fatores a serem considerados na definição do tipo de DW.....	41
Quadro 6 - Regras de negócio do DW.....	58
Quadro 7 - Documentação requisitos funcionais.....	60
Quadro 8 - Requisitos não funcionais.....	61
Quadro 9 - Documentação dos casos de uso.	63
Quadro 10 - Descrição da ilustração da solução proposta.....	68
Quadro 11 - Passos da implementação do DW.	69

LISTA DE SIGLAS

BI - *Business Intelligence*
CASE - *Computer-Aided Systems Engineering*
CI - *Competitive Intelligence*
DASD - *Direct Access Storage Device*
DDS.- *Dynamic Data Storage*
DM - *Data Marts*
DM - *Data Warehouses*
DSS - *Decision Support Systems*
EIS - *Executive Information System*
ER – *Entidade Relacionamento*
ERP - *Enterprise Resource Planning*
ESS - *Executive Support System*
ETL - *Extract Transform Load*
KMS - *Knowledge Work Systems*
MIS - *Management Information System*
OAS - *Office Automation Systems*
ODS - *Operational Data Storage*
OLAP - *On-Line Analytical Processing*
OLTP - *On-line Transaction Processing*
SA- *Staging Area*
SAD - *Sistema de Apoio à Decisão*
SAE - *Sistemas de Automação de Escritório*
SI – *Sistemas de Informação*
SGBD – *Sistema Gerenciador de Banco de Dados*
SIE - *Sistema de Informação Executiva*
SIG - *Sistemas de Informação Gerenciais*
SPT - *Sistemas de Processamento de Transações*
SQL - *Structured Query Language*
SSD - *Sistemas de Suporte de Decisões*
SSE - *Sistemas de Suporte Executivo*

STC - Sistemas de Trabalho do Conhecimento

TPS - *Transaction Processing Systems*

UML - *Unified Modeling Language*

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	APRESENTAÇÃO DO PROBLEMA	17
1.2	JUSTIFICATIVA	19
1.3	OBJETIVOS	20
1.3.1	Objetivo geral	20
1.3.2	Objetivos específicos	20
1.4	METODOLOGIA	20
1.5	PROPOSTA DE SOLUÇÃO	21
1.6	DELIMITAÇÕES	22
1.7	ESTRUTURA DO TRABALHO	22
2	REVISÃO BIBLIOGRÁFICA	24
2.1	SISTEMAS DE INFORMAÇÃO	24
2.1.1	Recursos dos sistemas de informação.....	25
2.1.2	Tipos de sistemas de informação.....	27
2.1.3	A evolução dos sistemas de apoio à decisão	29
2.2	BUSINESS INTELLIGENCE	31
2.3	EXTRACT, TRANSFORM AND LOAD - ETL.....	32
2.4	DATA WAREHOUSE.....	33
2.4.1	Definição de um Data Warehouse.....	33
2.4.2	Características do Data Warehouse	33
2.4.3	Justificativa de um Data Warehouse.....	34
2.4.4	Data Mart.....	35
2.4.5	Arquitetura de um Data Warehouse	35
2.4.6	Metadados.....	40
2.4.7	Tipos de Data Warehouse.....	40
2.4.8	Operational Data Storage.....	42
2.4.9	Granularidade	43
2.4.10	Modelagem multidimensional	45
2.5	ON-LINE ANALYTICAL PROCESSING - OLAP	49
2.5.1	Conceito.....	49

2.5.2	Tipos de OLAP	49
2.5.3	Operadores OLAP.....	50
2.6	CONSIDERAÇÕES FINAIS.....	53
3	MODELAGEM.....	54
3.1	INTRODUÇÃO	54
3.2	LEVANTAMENTO DE REQUISITOS	54
3.2.1	Regras de negócio	54
3.2.2	Requisitos funcionais.....	58
3.2.3	Requisitos não funcionais	60
3.2.4	Casos de uso	61
3.3	MODELAGEM DIMENSIONAL	63
3.3.1	Modelo Lógico	64
3.3.2	Modelo Físico	65
3.4	CONSIDERAÇÕES FINAIS.....	66
4	SISTEMA IMPLEMENTADO	67
4.1	INTRODUÇÃO	67
4.2	SOLUÇÃO PROPOSTA	67
4.3	FERRAMENTAS UTILIZADAS.....	68
4.4	PASSOS DA IMPLEMENTAÇÃO	69
4.5	PROTÓTIPO DESENVOLVIDO.....	70
4.6	VALIDAÇÃO.....	73
4.7	RESULTADOS E CONCLUSÕES	74
4.7.1	Índice de passageiros transportados por tipo de pagamento	75
4.7.2	Índice de passageiros transportados por faixa de horário	75
4.7.3	Índice de passageiros transportados por linha	76
4.7.4	Índice de Passageiros transportados por mês.....	77
4.7.5	Índice de estudantes transportados por faixa etária.....	77
4.7.6	Índice de passageiros estudantes transportados por sexo	78
4.7.7	Índice de passageiros que utilizam o vale-transporte por sexo.....	79
4.7.8	Índice de passageiros usuários do vale-transporte por linha	79
4.8	CONSIDERAÇÕES FINAIS.....	80
5	CONCLUSÕES E TRABALHOS FUTUROS.....	82
5.1	CONCLUSÕES GERAIS	82

5.2 TRABALHOS FUTUROS.....	83
REFERÊNCIAS.....	84
APÊNDICES	86
APÊNDICE A - CÓDIGO DDL PARA CRIAÇÃO FÍSICA DO BANCO DE DADOS GERADO COM FERRAMENTA SQL MANAGER	87
APÊNDICE B - CÓDIGO SQL DA VIEWQTDPA XPAGANTES CARTÃO DO SISTEMA OPERACIONAL	95
APÊNDICE C - CÓDIGO SQL DA VIEWQTDPA XPAGANTES DINHEIRO DO SISTEMA OPERACIONAL	97

1 INTRODUÇÃO

1.1 APRESENTAÇÃO DO PROBLEMA

O transporte coletivo de passageiros é um dos principais meios de locomoção da população nos médios e grandes centros urbanos. Para Bicalho (1998, apud AZAMBUJA, 2002), o transporte coletivo tem grande importância social, pois permite que as pessoas acessem locais de trabalho, equipamentos sociais e de lazer, e oportunidades de consumo, principalmente nas cidades de médio e grande porte, onde é maior a dependência da população dos meios de transporte coletivo. Segundo o mesmo autor, o transporte coletivo possui grande importância do ponto de vista econômico, pois viabiliza o crescimento das cidades, organizando e condicionando a ocupação dos espaços urbanos e do uso do solo através da articulação entre as atividades desenvolvidas em diferentes locais. Trata-se de um serviço público essencial, que deve atender as necessidades sociais e dar suporte a atividades econômicas.

Mesmo sendo essencial nos médios e grandes centros urbanos, o transporte coletivo tem sofrido com a evasão de passageiros. Segundo a Associação Nacional das Empresas de Transportes Urbanos (2006), em diversas cidades brasileiras, o sistema de ônibus perdeu quantidades significativas de passageiros desde meados dos anos 90. Os motivos para a substituição do transporte coletivo por outros meios variam segundo o poder aquisitivo. Para as pessoas das classes A e B a falta de conforto e privacidade no transporte coletivo são os principais motivos. Em seguida, a falta de segurança, a demora na viagem, a demora na espera e a tarifa cara. As pessoas de classes D e E indicam como principal motivo a tarifa cara e, em seguida, demora na espera; sendo que a falta de conforto, privacidade e demora na viagem estão num nível mais abaixo. Os motivos da classe C encontram-se aproximadamente entre os da classe B e da classe D.

Motivos da substituição do uso de ônibus/microônibus por classe socioeconômica

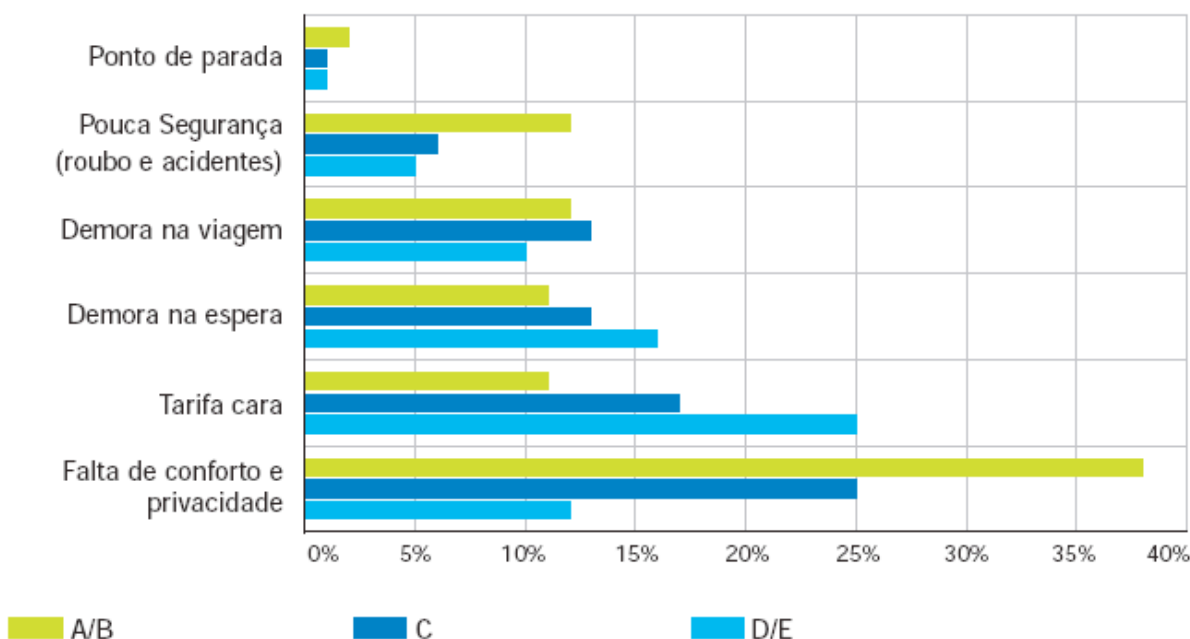


Gráfico 1 - Motivos da substituição do ônibus/microônibus por classe socioeconômica.

Fonte: Associação Nacional das Empresas de Transportes Urbanos. (2006, p. 33).

Juntamente com a evasão de passageiros, as empresas operadoras do transporte coletivo tiveram um aumento significativo nos custos operacionais com as altas constantes do preço do petróleo e seus derivados.

Na tentativa de reduzir seus custos, as empresas de transporte coletivo passaram a utilizar sistemas de informações para auxiliar suas operações, dando mais agilidade e controle aos serviços prestados.

Conseqüentemente, uma grande quantidade de informações passou a ser gerada diariamente pelos sistemas de informações e é fundamental a definição de critérios para selecionar e organizar os dados que são importantes para o apoio à tomada de decisão. (SOUZA, 2007).

Para dar suporte à tomada de decisão nas empresas de transporte coletivo, será feita a abordagem do conceito de *Business Intelligence* (BI) e sua importância para o apoio à tomada de decisões nas organizações, assim como o detalhamento dos principais itens de sua arquitetura.

O problema de pesquisa apresentado é representado pela seguinte questão: Como disponibilizar informações estratégicas e relevantes que auxiliem a gerência no processo de tomada de decisão em ambientes dinâmicos e complexos como o do transporte coletivo?

1.2 JUSTIFICATIVA

O tema utilizado para a elaboração do estudo foi escolhido por ser um assunto de grande importância social e econômica, pois está relacionado ao principal meio de transporte da população dos médios e grandes centros urbanos. (BICALHO, 1998 apud AZAMBUJA, 2002).

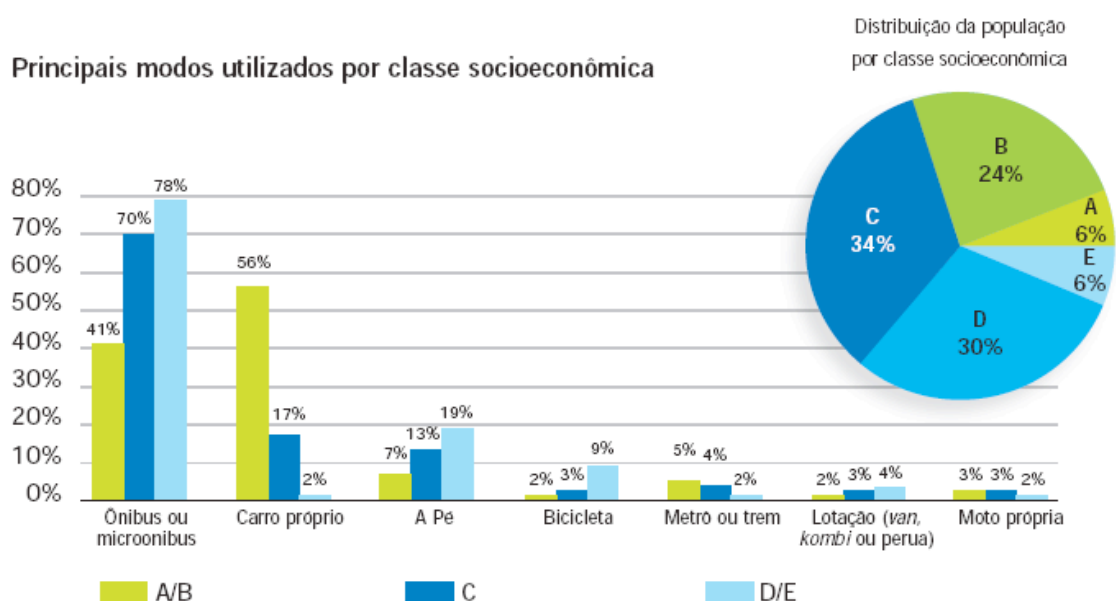


Gráfico 2 - Principais modos utilizados por classe socioeconômica.
 Fonte: Associação Nacional das Empresas de Transportes Urbanos. (2006, p. 13).

Para o sucesso das organizações em um segmento dinâmico e complexo como o do transporte, as informações possuem grande importância, pois é a partir delas que são tomadas as decisões e definidas as estratégias de negócios. No entanto, é necessário ter em mãos informações corretas, confiáveis e que retratem exatamente a situação real da empresa.

O presente trabalho tem o intuito de auxiliar as empresas de transporte coletivo na extração de informações importantes ao processo decisório, com a aplicação dos conceitos de BI.

1.3 OBJETIVOS

Com base na apresentação do problema, apresentam-se neste capítulo os objetivos gerais e específicos do presente trabalho.

1.3.1 Objetivo geral

Definir e implementar uma arquitetura de BI para auxiliar as empresas de transporte coletivo na extração de informações importantes ao processo decisório.

1.3.2 Objetivos específicos

- a) Apresentar os conceitos da arquitetura de BI;
- b) Construir um modelo dimensional para uma organização;
- c) Construir o sistema de extração, transformação e carga dos dados para realizar periodicamente a carga do *Data Warehouse*;
- d) Exibir as informações através de uma ferramenta OLAP (*On-Line Analytical Processing*);
- e) Realizar um estudo de caso para uma empresa de transporte coletivo.

1.4 METODOLOGIA

Para Gil e outros (1999 apud SILVA; MENEZES, 2005), método científico é um conjunto de processos ou operações mentais que devem ser utilizados durante uma investigação. É a linha de raciocínio adotada no processo de pesquisa.

Pesquisa é um conjunto de ações e propostas para encontrar a solução para um determinado problema, que têm por base procedimentos racionais e sistemáticos. É utilizada

quando se tem um problema e não se têm informações para solucioná-lo. (SILVA; MENEZES, 2005).

O tipo de pesquisa escolhido para o desenvolvimento deste projeto pode ser classificado como aplicado e de abordagem exploratória. Uma pesquisa aplicada de abordagem exploratória tem como objetivos gerar conhecimentos para aplicação prática e dirigidos à solução de problemas específicos, e visa proporcionar maior familiaridade com o problema com a intuito de torná-lo explícito ou a construir hipóteses. (SILVA; MENEZES, 2005).

A metodologia para elaboração do trabalho consiste das seguintes etapas:

- a) Levantamento teórico;
- b) Definição das necessidades de informação de uma empresa de transporte coletivo;
- c) Desenvolvimento da modelagem dimensional do sistema de DW;
- d) Desenvolvimento de um sistema-protótipo.

1.5 PROPOSTA DE SOLUÇÃO

A proposta de solução é desenvolver um DW que auxilie as empresas operadoras do transporte coletivo no processo de tomada de decisão, fornecendo informações úteis aos seus gestores. Para isso, são extraídos dados operacionais dos sistemas da empresa, inclusive de planilhas eletrônicas. Na sequência, esses dados são transformados e agrupados de forma que somente as informações relevantes sejam inseridas no DW para, por fim, permitir aos usuários a visualização das informações através de uma ferramenta OLAP. A figura 1 ilustra de forma gráfica a proposta de solução supracitada.

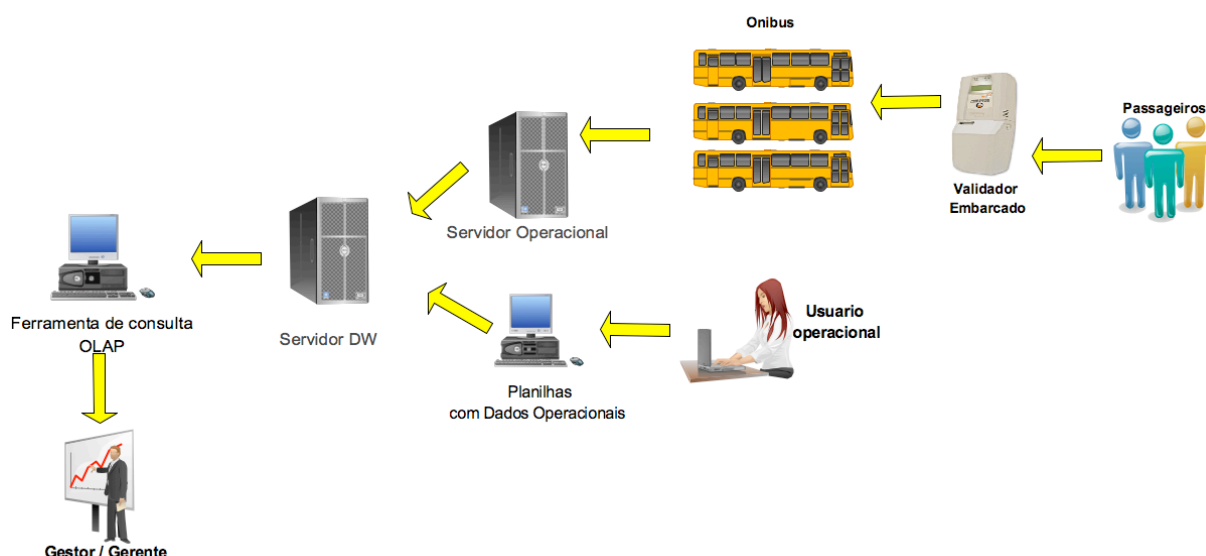


Figura 1 - Proposta de solução.
Fonte: Elaboração do autor, 2008.

1.6 DELIMITAÇÕES

O projeto aborda a definição e a construção do DW com as ferramentas disponíveis no mercado para a realização da extração, transformação e carga dos dados e a apresentação dos dados do DW através de uma ferramenta OLAP.

Serão realizadas análises para que sejam selecionados somente os dados importantes para o auxílio à tomada de decisão de um determinado assunto de negócios. Portanto, nem todos os dados existentes na base operacional serão extraídos para o DW.

A ferramenta OLAP apenas fornecerá informações aos tomadores de decisão, não tomando decisões pela gerência.

1.7 ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em 5 capítulos. O primeiro, denominado introdução, dispõe das subseções: apresentação do problema, justificativa, objetivos, metodologia da pesquisa, solução proposta e delimitações.

O segundo capítulo aborda as tecnologias de informação no processo decisório, abrangendo os tipos de sistemas de informação, a evolução dos sistemas de apoio gerencial para o conceito de BI e seus principais elementos. Também enfatiza os conceitos de DW sua arquitetura e a ferramenta OLAP utilizada.

O terceiro capítulo apresenta a modelagem utilizada para o desenvolvimento do DW, a definição dos requisitos de negócios e a modelagem do DW utilizando a linguagem de modelagem unificada (UML).

O quarto capítulo segue com a proposta de solução, seu desenvolvimento, sua validação e os resultados obtidos.

Por fim, no quinto capítulo, são apresentadas as conclusões obtidas com a realização deste projeto e sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo apresentam-se os fundamentos teóricos que irão sustentar o presente trabalho, iniciando pelo conceitos e os tipos de sistemas de informação, a evolução dos sistemas de apoio gerencial para o conceito de *Business Intelligence* e seus principais elementos. Também são enfatizados os conceitos de *Data Warehouse*, sua arquitetura e a ferramenta OLAP utilizada.

2.1 SISTEMAS DE INFORMAÇÃO

Para Mallack (2000, apud CLEMES, 2001), um sistema de informação é um sistema que tem como propósito armazenar, processar e compartilhar informações.

Um sistema de informação (SI) pode ser definido como um conjunto ou grupo de componentes inter-relacionados que captura, processa, armazena e distribui informações para dar apoio à tomada de decisão e ao controle de uma organização. Além de fornecer suporte, coordenar e controlar a tomada de decisões, os sistemas de informações freqüentemente ajudam gestores e trabalhadores a analisar problemas, visualizar assuntos complexos e na criação de novos produtos. (LAUDON; LAUDON, 2001).

Os sistemas de informações armazenam informações sobre pessoas, lugares e coisas úteis para a organização ou do ambiente que a cerca. Informação significa que os dados foram organizados ou ordenados de uma forma que é útil para as pessoas. Dados, em contrapartida, são uma seqüência de fatos brutos que representam eventos que acontecem dentro de organizações ou no seu ambiente físico antes de serem moldados de uma forma compreensível as pessoas. (LAUDON; LAUDON, 2001).

Os componentes ou funções básicas de um sistema de informação são: entrada, processamento e saída. A entrada envolve a reunião e coleta de dados brutos que entram no sistema da organização para serem processados. O processamento converte e transforma os dados brutos em informações úteis. Durante o processamento podem ser realizados cálculos, comparações, escolhas de ações alternativas e o armazenamento dos dados para uso posterior. Em sistemas de informação, a saída transfere a informação processada às pessoas em forma

de informação útil, geralmente no padrão de relatórios e/ou documentos. As saídas podem ser utilizadas em alguns casos para entrada de outros sistemas. (STAIR; REYNOLDS, 2002).

O conceito de sistema se torna mais útil com a inclusão de dois componentes adicionais: *feedback* e controle. *Feedback* são dados referentes ao desempenho de um sistema, como, por exemplo, os dados sobre desempenho de vendas, que são um *feedback* para um gerente de vendas. O controle envolve acompanhamento e avaliação do *feedback* para determinar se um sistema está se dirigindo rumo a sua meta. Na sequência, o controle tem a função de realizar os ajustes necessários nos componentes de entrada e processamento de um sistema, a fim de assegurar que produção adequada seja alcançada. (O'BRIEN, 2003).

2.1.1 Recursos dos sistemas de informação

Para O'Brien (2003), um sistema de informação é composto de cinco recursos principais:

- a) recursos humanos – são as pessoas necessárias para a operação dos sistemas de informações. Estes recursos humanos incluem os usuários finais e os especialistas em SI. Usuários finais ou clientes são as pessoas que utilizam um sistema de informação ou a informação que ele gera. Os especialistas em SI são pessoas que desenvolvem e operam sistemas de informação. Incluem os analistas de sistemas, programadores e operadores de computador entre outros;
- b) recursos de *hardware* – os dispositivos físicos e equipamentos utilizados no processamento de informações definem o conceito de recursos de *hardware*. Especificamente, o conceito não inclui apenas máquinas, como computadores e outros equipamentos, mas também todas as mídias de dados, ou seja, objetos tangíveis nos quais são armazenados dados. Exemplos de *hardware* em sistemas de informações computadorizados são sistemas de computadores e periféricos de computador. Sistemas de computadores consistem em unidades de processamento central contendo microprocessadores e uma grande variedade de dispositivos periféricos interconectados. Já os periféricos de computador, são dispositivos de entrada e saída como teclados, mouses, monitores de vídeo, impressoras e discos ópticos para armazenamento de dados;
- c) recursos de *software* - a definição de recursos de *software* envolve todos os conjuntos de instruções de processamento da informação. Essa definição genérica de *software* não

inclui só os conjuntos de instruções operacionais chamados programas, que gerenciam e controlam o *hardware*, mas também os conjuntos de instruções de processamento das informações requisitadas por pessoas, chamados procedimentos.

- d) recursos de dados - os dados são a matéria-prima dos sistemas de informação e constituem um valioso recurso organizacional. Podem-se ter diversas formas de dados, entre as quais, a de dados alfanuméricos tradicionais, descrições de transações de negócios e outros eventos e atividades. Os recursos de dados dos sistemas de informação são geralmente armazenados processados e organizados em banco de dados, e em bases de conhecimento que guardam o conhecimento em uma grande variedade de formas.
- e) recursos de rede - as redes de telecomunicações como a Internet, intranets e extranets são essenciais para o sucesso das operações de todos os tipos de organizações e seus sistemas de informações computadorizados. As redes de telecomunicações consistem em computadores, processadores de comunicações e outros dispositivos interconectados por mídia de comunicação e controlados por *software*.

A figura 2 apresenta em forma esquemática os principais componentes e recursos dos sistemas de informação.

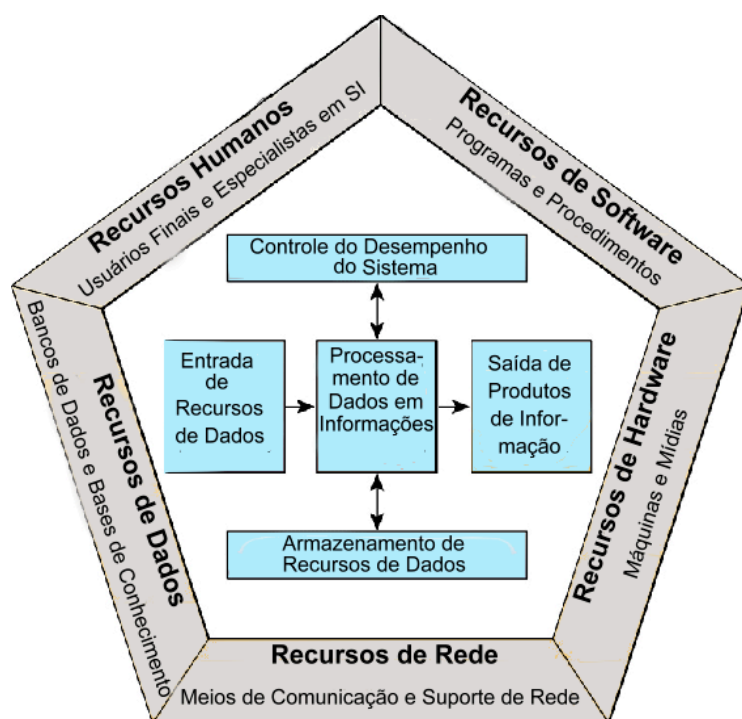


Figura 2 - Componentes de um sistema de informação.

Fonte: O'Brien (2003).

2.1.2 Tipos de sistemas de informação

Laudon e Laudon (2001) classificam os sistemas de informações em quatro tipos, que servem a diferentes níveis na organização, conforme a figura 3.

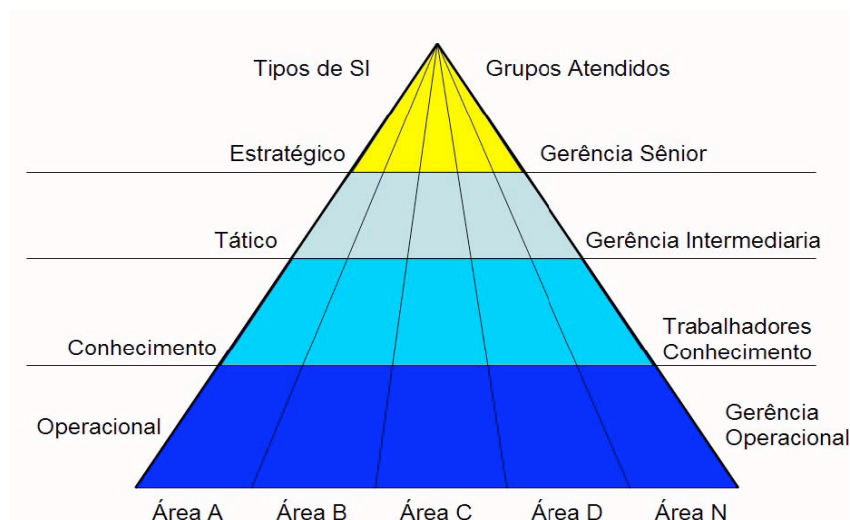


Figura 3 - Tipos de sistemas de informação.

Fonte: Laudon e Laudon (1998, apud CLEMES, 2001).

O quadro 1 apresenta um resumo das características dos quatro níveis de um sistemas de informações.

Nível	Características	Tipo de SI
Operacional	Responder questões rotineiras e acompanhar o fluxo das transações através da organização	<i>Transaction Processing Systems</i> (TPS) ou sistemas de processamentos transacionais
Conhecimento	Permitir aos funcionários de escritório a criação de produtos, racionalização de serviços e acompanhamento do fluxo de documentos na organização	<i>Knowledge Work Systems</i> (KWS) ou sistema de controle de conhecimento <i>Office Automation Systems</i> (OAS) ou sistemas de automação de escritório
Tático	Suportar o monitoramento, controle, tomadas de decisão e avaliação das atividades da gerência intermediária	<i>Management Information System</i> (MIS) ou sistemas de administração das informações <i>Decision Support Systems</i> (DSS) ou sistemas de suporte à decisão
Estratégico	Ajudar a gerência sênior no planejamento das atividades de longo prazo	<i>Executive Support System</i> (ESS) ou sistema de suporte aos executivos

Quadro 1 - Características dos SI por nível.

Fonte: Adaptado de Laudon e Laudon (1998, apud CLEMES, 2001).

Na seqüência, descrevem-se os seis principais tipos de sistemas de informações:

- a) sistemas de processamento de transações (SPT) - são os sistemas empresariais básicos que servem o nível operacional da organização. (LAUDON; LAUDON, 2001). Stair e Reynolds (2002) definem como um conjunto organizado de pessoas, procedimentos, *softwares*, banco de dados e dispositivos com o propósito de registrar as transações empresarias realizadas;
- b) sistemas de trabalho do conhecimento (STC) e sistemas de automação de escritório (SAE) – provêem as necessidades de informação no nível de conhecimento da empresa. Os sistemas de trabalho do conhecimento auxiliam os trabalhadores do conhecimento, enquanto os sistemas de automação de escritório ajudam principalmente os trabalhadores de dados. Trabalhadores do conhecimento são pessoas que possuem níveis universitários e, na maioria das vezes, possuem uma profissão reconhecida como engenheiros, médicos e advogados. Os projetos de trabalho do conhecimento, tais como o projeto científico ou de engenharia de estações de trabalho, promovem a criação de conhecimento novo e asseguram que esse novo conhecimento e especialidades técnicas sejam corretamente agregados ao negócio. Os sistemas de automação de escritório são aplicações de informática projetadas para aumentar a produtividade dos trabalhadores de dados, dando suporte à coordenação e às atividades de um escritório. Os trabalhadores de dados possuem, normalmente, níveis educacionais menos formais e tendem a processar em lugar de criar informação (LAUDON; LAUDON, 2001) ;
- c) sistemas de informação gerenciais (SIG) – envolve um conjunto organizado de pessoas, procedimentos, *softwares*, banco de dados e equipamentos que fornecem informação do dia-a-dia aos gerentes e aos tomadores de decisão. (STAIR; REYNOLDS, 2002). Para Laudon e Laudon (2001), os SIG dão suporte ao nível gerencial da empresa, fornecendo aos gerentes relatórios e, em alguns casos, acesso *On-Line* para o desempenho atual e registros históricos da empresa. Na maioria dos sistemas de informações gerenciais seu foco é praticamente exclusivo para dentro e não para eventos externos ou ambientais. Esses sistemas dão suporte às funções de planejamento, controle e tomada de decisão no nível gerencial da organização através de relatórios. O'Brien (2003), define os SIG como uma concentração de sistemas de informação com a função de fornecer informações e apoio à tomada de decisão pelos gerentes através de relatórios;
- d) sistemas de suporte de decisões (SSD) – para Stair e Reynolds (2002, p. 19), “[...] é uma coleção organizada de pessoas, procedimentos, *software*, bancos de dados e dispositivos

usados para dar suporte a um problema específico na tomada de decisão.”. Os sistemas de suporte à decisões auxiliam os gerentes a tomarem decisões que são semi-estruturadas, únicas, ou velozmente modificadas e dificilmente especificadas com antecipação (LAUDON; LAUDON, 2001) ;

- e) sistemas de suporte executivo (SSE) – fornecem suporte ao nível estratégico da organização. Eles priorizam decisões não estruturadas e geram um ambiente computadorizado e de comunicações generalizado, em vez de fornecerem qualquer aplicação fixa ou capacidade específica. (LAUDON; LAUDON, 2001).

2.1.3 A evolução dos sistemas de apoio à decisão

Para Sprague e Watson (1991, apud BISPO, 1998), a evolução dos sistemas de apoio à decisão teve início na década de 70. Diversas empresas e vários grupos de pesquisas começaram a analisar e desenvolver sistemas de apoio à decisão, caracterizados como sistemas computacionais interativos que auxiliavam no processo decisório de problemas considerados não estruturados. Contudo, na década seguinte, diversos pesquisadores e desenvolvedores de sistemas ampliaram essa definição de modo a incluir qualquer sistema capaz de disponibilizar alguma contribuição ao processo decisório, desde que possuam as seguintes características:

- a) desenvolvidos para problemas menos estruturados e menos específicos que os gerentes vivenciam;
- b) mesclam a utilização de modelos ou técnicas analíticas e funções tradicionais de acesso e recuperação de informações;
- c) com foco especificamente em recursos que facilitem seu uso para pessoas não especializadas em computação;
- d) destaquem a flexibilidade e a adaptabilidade de acomodar mudanças no ambiente e na abordagem ao processo decisório.

Pearson e outros (1995, apud BISPO, 1998) afirmam que os primeiros sistemas de apoio à decisão (SAD) apareceram nos anos 60 e 70 para dar suporte aos gerentes na solução de problemas gerenciais não estruturados. Estes SADs tinham custo elevado, eram de uso específico e difíceis de se operar. Naquela época, o mercado era estável para as empresas, a concorrência era menor e existia a fidelidade dos clientes. A principal dificuldade, nesse

período, era que as bases de dados dos sistemas computacionais não possuíam a arquitetura necessária para realização de pesquisas típicas de sistemas de apoio à decisão e, devido à falta de dados históricos, havia muita dificuldade na criação de relatórios e na realização de análises necessárias ao gerenciamento dos negócios. A ênfase dos SADs não estava no processo decisório, mas no apoio computacional para o desenvolvimento rápido das aplicações.

Nessas décadas, os sistemas computacionais para o suporte à decisão eram desenvolvidos na própria empresa ou por terceiros. Esses sistemas tinham primeiramente o intuito de auxiliar na solução de problemas gerenciais específicos na empresa e, depois, eram aperfeiçoados para envolver outros problemas gerenciais. Contudo, nenhum destes sistemas conseguiu fornecer um bom suporte ao dinâmico processo de tomada de decisão, disponibilizando as informações necessárias no momento certo. Quando os gerentes ou os analistas precisavam realizar consultas ou confeccionar relatórios que envolviam muitos dados, era necessário desenvolvê-los através de alguma linguagem de programação, o que acarretava um grande tempo para a sua realização. (BISPO, 1998).

Ainda segundo Bispo (1998), surgiram nos anos 80 os primeiros sistemas de gerenciamento de banco de dados (SGBD), o que possibilitou um melhor acesso aos dados disponíveis, à sua formatação e à construção de consultas e relatórios de uma forma mais prática, rápida e barata. Entretanto, toda a análise mais detalhada dos dados era realizada de forma manual, ou seja, os dados eram colecionados, formatados e só então realizadas as análises necessárias e, muitas vezes, sem o auxílio de qualquer sistema computacional. Mesmo com o advento de novos *softwares*, como planilhas eletrônicas e os programas de visualização gráfica dos dados, era difícil elaborar um relatório que contivesse uma análise cruzando informações de sistemas diferentes, devido ao grande problema de que a modelagem de dados era baseada na estrutura dos processos ao invés de na estrutura dos negócios. Ainda nos anos 80, começaram a surgir os primeiros sistemas especialmente desenvolvidos para os gerentes, denominados sistemas de informação para executivos (*Executive Information System* - EIS). Porém, o gerenciamento das empresas e dos negócios estava crescendo mais velozmente que esses sistemas.

Para The e Weldon (1998, apud BISPO, 1998), no início da década de 90 surgiram as ferramentas CASE (*Computer-Aided Systems Engineering*) e as linguagens de quarta geração. Estas ferramentas se propunham a resolver os problemas dos usuários finais que necessitavam de informações rápidas e não tinham tempo a perder com o desenvolvimento de sistemas específicos. Com a utilização das linguagens de quarta geração

foi possível confeccionar relatórios e realizar consultas de uma maneira bem mais rápida e prática. E, com o uso das ferramentas CASE, foi possível desenvolver sistemas de uma maneira mais rápida e simples. Contudo, ambas as ferramentas não eram flexíveis o suficiente para atender a todas as necessidades gerenciais. Mesmo com a utilização de planilhas eletrônicas para análise dos dados, os gerentes continuavam gastando mais tempo manipulando dados e construindo modelos do que realizando as análises necessárias para execução de suas atividades. Ainda na década de 90, foram criadas diversas ferramentas para dar suporte à solução dos problemas das empresas. Entre as novas ferramentas de tecnologia da informação, pode-se destacar o ERP (*Enterprise Resource Planning*) e “[...] a nova geração de sistemas de apoio à decisão (SAD): o data warehouse, o OLAP e o data mining [...]”. (BISPO; CAZARINI, 1998 apud BISPO, 1998, p.31).

Com a utilização das novas ferramentas SADs chegamos a era do BI, onde as informações são extraídas a partir dos dados e aplicadas diretamente aos negócios das organizações. (BARBIERI, 2001).

2.2 BUSINESS INTELLIGENCE

O conceito de BI pode ser compreendido como a utilização de variadas fontes de informações para se definir estratégias de competitividade nos negócios da empresa. Atualmente as organizações geram grandes quantidades de dados, mas enfrentam dificuldades na extração de informações. Esse crescimento das informações dificulta o processo de tomada de decisão, na medida em que a alta e a média gerência sentem-se inúteis no processo de sua procura e recuperação. Os sistemas legados e os ERP não disponibilizam as informações gerenciais na sua forma mais agradável. Ao contrário, as informações cruciais para a tomada de decisões estratégicas estão ocultas em milhares de tabelas e arquivos inacessíveis aos leigos na área de computação, conectadas por relacionamento e correlações transacionais em uma forma inadequada para os tomadores de decisão. Desta forma, o conhecimento corporativo e as informações externas não estão prontamente disponíveis para sua utilização. Neste contexto, o BI tem com seu principal objetivo a definição de técnicas para a formatação adequada destes volumes de dados, visando transformá-los em depósitos estruturados de informações, independente de sua fonte. Os dados poderão ser extraídos usando técnicas emergentes de garimpo de informações via CI (*Competitive Intelligence*), ou de fontes

conceituais como KMS (*Knowledge Work Systems*). Ambas as formas, a definição de estruturas modeladas dimensionalmente armazenados em *Data Warehouses* ou DM (*Data Marts*) e interpretadas pela visão analítica das ferramentas OLAP (*On-Line Analytical Processing*) ou pelas ferramentas de *Data Mining*, atingem o objetivo proposto pelas premissas de BI. (BARBIERI, 2001).

2.3 EXTRACT, TRANFORM AND LOAD - ETL

Um dos processos de um DW consiste na extração dos dados dos sistemas operacionais, na organização e integração desses dados de forma consistente. A extração, organização e integração dos dados devem ser realizadas com a intenção de garantir a consistência e integridade das informações, construindo desta forma uma base de dados de alta qualidade e confiabilidade, que retrate a realidade do negócio da empresa. Esses sistemas ou aplicações são responsáveis pela filtragem, limpeza, sumarização e concentração dos dados espalhados pelas fontes externas e nos sistemas operacionais. A maior parte dessas ferramentas possuem interfaces gráficas e interativas que auxiliam na realização do mapeamento de dados e a automação do processo de extração, limpeza e carga, sendo fortemente baseadas na linguagem SQL (*Structured Query Language*). (MACHADO, 2000).

Para Barbieri (2001), os conceitos de extração dos dados e de seu tratamento podem ser divididos em:

- a) filtro de dados - relaciona os procedimentos e condições para se eliminar os elementos indesejáveis no modelo dimensional;
- b) integração de dados - caracteriza a forma de se correlacionar informações existentes em fontes distintas, e que deverão ser integrados ao sistema gerencial;
- c) condensação de dados - define a maneira de reduzir o volume dos dados com o objetivo de obter informações resumidas e sumarizadas. Geralmente essas sumarizações acontecem nas dimensões dos dados, como tempo e localização geográfica;
- d) conversão de dados - especifica os procedimentos para se transformar dados em unidades, formatos e dimensões diferentes;
- e) derivação de dados - são os meios e fórmulas para gerar dados virtuais, a partir de dados existentes.

Segundo Machado (2000), o desenvolvimento de programas de extração requer, dos analistas envolvidos, um bom conhecimento das bases de dados das quais as informações serão extraídas e da base em que serão armazenadas. O mesmo autor afirma que o conhecimento necessário consiste no acesso aos catálogos dos bancos de dados mapeados como fontes de informação, assim como o pleno domínio dos metadados do modelo multidimensional do DW ou do DM.

2.4 DATA WAREHOUSE

2.4.1 Definição de um Data Warehouse

O conceito de *Data Warehousing* não é recente. Foi inicialmente apresentado pela IBM como uma infra-estrutura. Porém, a definição mais utilizada foi criada por Bill Inmon, que apresenta o DW como orientado ao assunto, integrado, variável com o tempo e não volátil, e que fornece suporte ao processo de tomada de decisão. (SINGH, 2001 apud CLEMES, 2001).

Um DW pode ser definido como um banco de dados voltado a sistemas de apoio à decisão e cujos dados foram armazenados em estruturas lógicas dimensionais, permitindo o seu processamento analítico por ferramentas OLAP e *Data Mining*. (BARBIERI, 2001).

Segundo Serra (2002), um DW é um conjunto de diversas tecnologias, como ferramentas de extração e conversão, banco de dados voltados para consultas complexas, ferramentas inteligentes de prospecção e análise de dados, ferramentas de administração e gerenciamento.

2.4.2 Características do Data Warehouse

O quadro 2 descreve, segundo Inmon (2000), as principais características de um DW.

Características	Descrição
Orientação por assunto	Significa que um DW armazena as informações agrupadas por assuntos de maior interesse da empresa, em contraste com os sistemas operacionais que são orientados a processos desenvolvidos para manter as transações realizadas diariamente.
Integração	Essa é uma das mais importantes características do DW. Através dela é possível padronizar uma representação única para os dados de todos os sistemas que formarão a base de dados do DW.
Variação de tempo	Os dados de um DW são valiosos em relação ao tempo e significam resultados operacionais em determinado momento de tempo. Os dados de um DW são uma imagem, um conjunto estático de registros de uma ou mais tabelas, capturados em um determinado momento de tempo predeterminado. Logo, os dados do DW não podem ser atualizados. Por exemplo, os dados sobre as vendas de um determinado mês jamais terá seus valores modificados.
Não volátil	No ambiente operacional, os dados são atualizados freqüentemente. Todavia, os dados existentes no DW apresentam características diferentes. Os dados do DW são somente carregados (inseridos) e acessados (visualizados). Raramente ocorre uma atualização de dados em um ambiente de DW.

Quadro 2 - Características do *Data Warehouse*.

Fonte: Adaptado de Machado (2000); Serra (2002); Inmon (2000).

2.4.3 Justificativa de um Data Warehouse

As organizações geram e captam dados sobre todos os aspectos dos seus negócios. Em grande parte estes dados estão sendo armazenados e manipulados para suportar suas atividades críticas com o uso dos sistemas de processamento de transações *On-Line* (OLTP), que têm como característica principal o baixo tempo de resposta para processar as suas requisições. (CLEMES, 2001).

Segundo Machado (2000), a principal justificativa para utilização da tecnologia de DW em uma organização é a existência na empresa de:

- diversas plataformas de *hardware* e *software*;
- modificações constantes nos sistemas transacionais corporativos;
- existência de diversos sistemas fechados de fornecedores diferentes;
- inexistência de padronização e integração dos dados nos diversos sistemas;

- e) deficiência de documentação e segurança no armazenamento dos dados;
- f) dificuldade de implantação de sistemas de informação executiva (EIS) ou de apoio a decisão (DDS) devido a dependências múltiplas de sistemas corporativos.

2.4.4 Data Mart

Um *Data Mart* (DM) pode ser representado como um subconjunto de dados do DW, possibilitando um acesso descentralizado que serve de fonte para os dados que integrarão os bancos de dados individuais. Sua utilização tem como principal vantagem o retorno rápido, assegurando um maior envolvimento com o usuário final, o que possibilita uma melhor avaliação dos benefícios extraídos de seu investimento. (MACHADO, 2000).

“Os marts diferem dos warehouse no sentido de que, enquanto o warehouse se concentra nas necessidades da empresa inteira, os marts são dedicados a áreas de assuntos específicas ou às necessidades de um departamento ou função empresarial em particular. [...]” (COREY; ABBEY; TAUB, 2001, p. 172).

2.4.5 Arquitetura de um Data Warehouse

Segundo Machado (2000), a arquitetura deve ser determinada de acordo com o local em que o DW ou DM estarão localizados. As três arquiteturas básicas são a global, independente e integrada.

2.4.5.1 Arquitetura global

Nesta arquitetura o DW é projetado de acordo com as necessidades da empresa como um todo. Pode ser definido como um repositório comum de dados de suporte à decisão, disponível para toda a empresa. O termo global reflete o escopo de acesso e utilização das informações da empresa, não sendo necessariamente centralizado. Esta arquitetura pode ser

fisicamente centralizada ou distribuída na estrutura de uma empresa. Geralmente a centralização física é escolhida quando a empresa existe em um único local e o DW é administrado por um setor de tecnologia da informação. A distribuição física de um *Data Warehouse* é empregada quando a empresa possui sua estrutura física em diversos locais e os dados em múltiplas instalações, possuindo também um setor de administração de tecnologia da informação. A arquitetura global possibilita aos usuários finais utilizar visões corporativas de dados. Porém, esse tipo de ambiente possui um custo de implementação muito alto devido ao grande tempo necessário para desenvolvimento e administração. (MACHADO, 2000).

Na seqüência, a figura 4 ilustra a arquitetura de um *Data Warehouse* global.

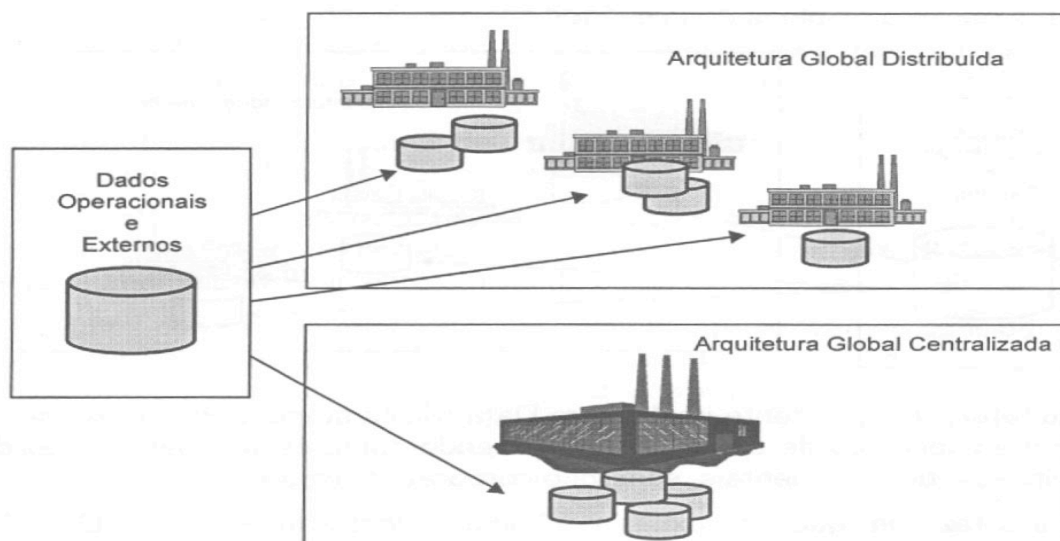


Figura 4 - Arquitetura *Data Warehouse* global.
Fonte: Machado (2000).

2.4.5.2 Arquitetura de Data Marts independentes

Para Machado (2000), essa é a arquitetura favorita dos fornecedores de *software* para consulta de informações de DW. Esse tipo de arquitetura requer DM separados controlados por grupos específicos de usuários, e que atendem somente às suas necessidades específicas e departamentais, sem qualquer foco corporativo. Conseqüentemente, não existe nenhuma conectividade desses *Data Marts* com os de outros departamentos ou áreas de negócio. A arquitetura independente demanda os mesmo perfis técnicos para a implantação,

mas os recursos e pessoal operacional podem ser administrados pelo grupo de trabalho ou departamento. Sua restrição é que possui um mínimo de integração corporativa e não permite nenhuma visão global da empresa

A arquitetura de *Data Marts* independentes é representada na figura 5.

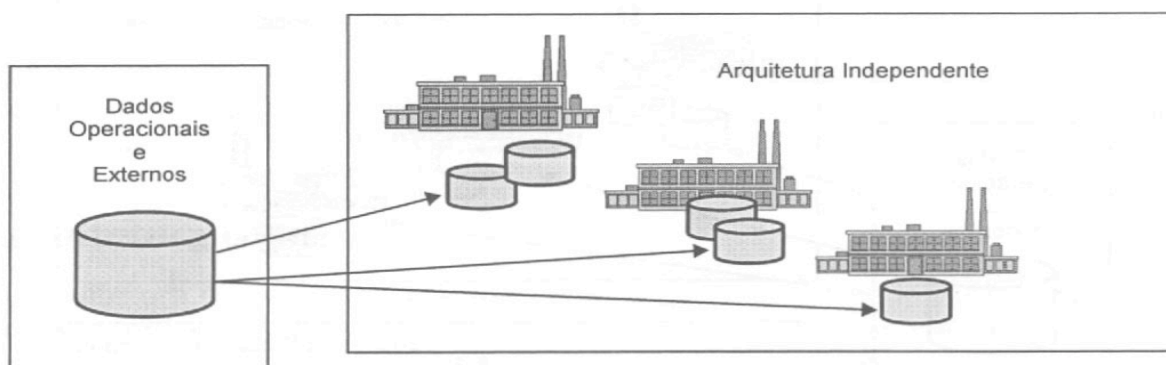


Figura 5 - Arquitetura de *Data Marts* independentes.
Fonte: Machado (2000).

2.4.5.3 Arquitetura de Data Marts integrados

A arquitetura de *Data Marts* integrados é basicamente uma distribuição de implementação. Embora os DM sejam implementados separadamente por grupos de trabalho ou departamentos, eles são integrados, disponibilizando maior visão corporativa dos dados e informações. Pode-se comparar com a arquitetura global devido ao alto nível de integração. Esse tipo de arquitetura possibilita muitas outras funções e capacidades de informação em relação à arquitetura independente. Conseqüentemente, a integração aumenta sensivelmente o nível de complexidade de requisitos. Em compensação ocorre um aumento da capacidade e de qualidade de visão corporativa de informações quando optamos por compartilhar dados entre os *Data Marts*. Nesta situação a atuação da área de tecnologia da informação deve ser superior a da arquitetura independente. (MACHADO, 2000).

A seguir, a figura 6 ilustra a arquitetura de *Data Marts* integrados.

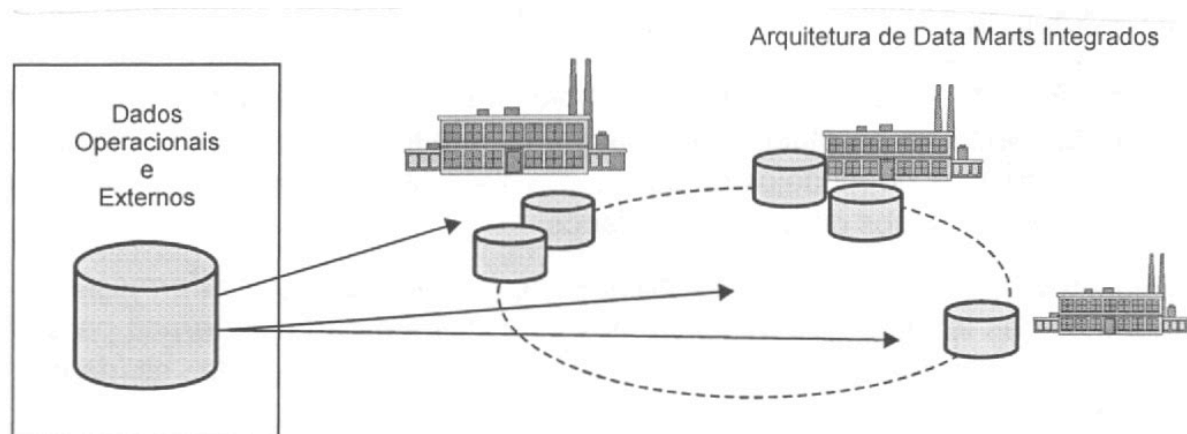


Figura 6 - Arquitetura de *Data Marts* integrados.
Fonte: Machado (2000).

2.4.5.4 Implementação

A escolha por um tipo de abordagem é influenciada por diversos fatores entre eles: recursos de infra-estrutura de tecnologia, a arquitetura escolhida, o escopo da implementação, os recursos disponíveis, velocidade de implementação e, principalmente, para uma possível necessidade de acessos aos dados, bem como o retorno do investimento pretendido. Basicamente, existem três maneiras de implementação: *Top down*, *Bottom Up* e combinada. (MACHADO, 2000).

O quadro 3 apresenta a descrição dos três tipos de abordagens.

Implementação	Descrição
<i>Top Down</i>	Essa abordagem foi sugerida por Inmon e baseia-se em um DW corporativo central, fundamentado no modelo relacional e totalmente normalizado. O processo de extração, transformação, carga e, conseqüentemente, a área de estágio de dados, são implementados de forma única e integrada. Este DW corporativo contempla todos os dados organizacionais e tem o objetivo servir de base de dados para os diversos <i>Data Marts</i> departamentais implementados de acordo com o modelo dimensional.
<i>Bottom Up</i>	A implementação <i>Bottom Up</i> se baseia na construção de DM dimensionais independentes, cada um com sua respectiva área de estágio e processos de extração, transformação e limpeza e metadados. Neste tipo de implementação o DW é composto pelo conjunto de todos os <i>data marts</i> construídos, criando um DW incremental lógico.
Combinada	O objetivo das técnicas <i>Top Down</i> e <i>Bottom Up</i> combinadas é aproveitar ao máximo as características positivas de cada uma das implementações. Embora seja difícil de implementar, alguns autores afirmam que um bom

	gerente de projeto pode executar esta tarefa. A monitoração cuidadosa do processo de implementação, o acompanhamento e gerenciamento dos novos requerimentos auxiliam a alcançar melhores benefícios de ambas as técnicas
--	---

Quadro 3 - Tipos de implementação.

Fonte: Adaptado de Morales (2008); Machado (2000); IBM (1998, apud CLEMES, 2001).

Na seqüência, o quadro 4 mostra as vantagens e desvantagens para cada tipo de implementação.

Implementação	Vantagens	Desvantagens
<i>Top Down</i>	Regras de negócio e definição dos dados de forma mais consistente	Precisa da participação de toda organização
	Permite uma fácil manutenção por utilizar a arquitetura de DW para originar outros DM	Tempo de implementação muito longo. Em média, são necessários mais de 15 meses para o desenvolvimento e produção da primeira área de assunto
	Devido ao desenvolvimento da extração de dados ser feito uma única vez, evita-se o retrabalho	Os gastos iniciais com aquisição do <i>hardware</i> e infra-estrutura e formação da numerosa equipe técnica são excessivamente altos
		Elevada taxa de risco por não existirem garantias neste tipo de investimento
<i>Botton Up</i>	Rápida implementação com a construção direcionada dos DM	Falta de planejamento de integração entre os DM
	A apresentação dos primeiros resultados é feita de modo rápido e barato	Maior chance de incoerência nos dados, devido a diferentes representações e fontes de dados
	Menor taxa de risco possibilitando investimentos adicionais	A integração dos DM em um DW global pode ser difícil até uma determinada fase do projeto
Combinada	Garantia da consistência dos dados com a utilização da modelagem de dados única para os DM	Complicações políticas em decorrência da determinação da seqüência de implementação dos DM e das prioridades de manutenção
	As ferramentas de extração são desenvolvidas uma única vez	Metadados complexos para gerenciar

Quadro 4 - Vantagens e desvantagens das abordagens de implementação.

Fonte: Adaptado de Morales (2008); Machado (2000); IBM (1998, apud CLEMES, 2001).

2.4.6 Metadados

Para Serra (2002), os metadados podem ser definidos como dados sobre dados. Mais exatamente como dados de mais alto nível que descrevem dados de um nível inferior.

Os metadados podem ser comparados com uma enciclopédia para o DW que referencia todas as informações no seu ambiente, inclusive os que não são propriamente dados. Metadados que envolvem os sistemas de gerenciamento de banco de dados (SGBD) do DW são responsáveis por itens como tabelas de sistemas, configurações de partições, índices, definições de exibição e concessões de privilégios de segurança do SGBD. (KIMBALL, 2002).

Serra (2002) afirma que, de uma maneira geral, existem três camadas de metadados em um DW:

- a) metadados operacionais – definem a estrutura dos dados mantidos pelo bancos de dados operacional, usado pelas aplicações de produção da organização;
- b) metadados centrais do DW – São armazenados no catálogo do DW. Caracterizam-se por serem orientados por assunto, definindo como os dados transformados devem ser interpretados. Contêm definições de agregados e campos calculados, assim como visões sobre cruzamento de assuntos;
- c) metadados do nível do usuário – mapeiam os metadados do DW para características que sejam familiares e apropriados aos usuários finais.

Segundo Cledes (2001), as fontes de metadados definem de onde os dados serão extraídos para compor os metadados. O mesmo autor classifica como de fontes formais, os que documentam a origem dos dados e fazem parte da documentação da empresa, e os de fontes informais que existem dentro da empresa mas não fazem parte da documentação formal.

2.4.7 Tipos de Data Warehouse

Para Singh (2001, apud CLEMES, 2001), a construção e implementação de um ambiente de DW deve ser única e centralizada pelos seguintes motivos:

- a) o volume de dados pode ser facilmente gerenciado;

- b) os dados de toda a empresa são integrados e apenas essa visualização é usada;
- c) existe dificuldade em integrar e acessar os dados em um único local, quando eles estão separados em vários locais na organização.

De acordo com Cledes (2001), outros motivos devem ser levados em consideração antes da definição do tipo de DW a ser implementado. O quadro 5 apresenta um resumo destes fatores.

Fatores	Descrição
Objetivos do negócio	Imprecisão na definição das prioridades ou opções que influenciam o modelo de DW, como tamanho, localização, frequência de uso e manutenção
Localização dos dados atuais	Compreender onde os dados estão, quais são as suas características e atributos, e a forma para acessá-los, são desafios para o fornecimento desses dados aos usuários finais
Necessidade de transferência de dados	No início de um projeto de DW, normalmente supõe-se que haverá transferência de dados. No primeiro momento será mais fácil deixar os dados em sua localização inicial, a menos que seja realmente necessário realizar sua transferência
Transferência de dados	Mesmo com a disponibilidade de ferramentas para transferir dados para qualquer lugar o desconhecimento dos atributos dos dados dificulta esta tarefa
Local para onde transferir dados	Deverá ser realizada uma análise do ambiente operacional (servidor, redes locais, etc.) onde os dados serão armazenados
Requisitos de suporte à decisão	Deverá ser capaz de atender às atividades de suporte à decisão e sistemas de informações gerenciais
Preparação dos dados	Somente substituir os dados existentes em um campo por novas informações não refletirá o histórico da mudança de dados ao longo do tempo. Deve-se optar pela substituição dos dados ou procurar formas de atualizá-los, com base em alterações incrementais
Requisitos de consulta e relatório	É uma das etapas mais importantes na implementação do DW. Deve-se analisar quais ferramentas serão necessárias e quais já estão disponíveis no mercado
Integração do modelo de dados	Utilizando os modelos de dados existentes na organização, deve-se integrar os resultados ao desenvolvimento do DW e as ferramentas utilizadas pelos usuários finais
Gerenciamento e administração	São as atividades que possibilitam a continuidade do DW após a sua construção

Quadro 5 - Fatores a serem considerados na definição do tipo de DW.

Fonte: Adaptado de Singh (2001, apud CLEMES, 2001).

2.4.8 Operational Data Storage

Um *Operational Data Storage* (ODS) ou *Staging Area* (SA) representa um armazenamento intermediário dos dados, promovendo a integração dos dados do ambiente operativo antes de sua atualização no DW. Inicialmente, um ODS era considerado um repositório temporário que armazenava apenas informações correntes antes de serem carregadas para o DW, similar a uma cópia dos ambientes de sistemas transacionais em uma empresa. Atualmente, alguns autores passaram a denominá-lo *Dynamic Data Storage* (DDS). Essa concepção se diferencia do conceito original pela sua periodicidade de armazenamento e pelo fato de não somente armazenar dados temporários para a carga do DW. Por não ser volátil, seus dados são armazenados ao longo do tempo e passam por alterações incrementais que ao longo do tempo, podendo se tornar um DW. (MACHADO, 2000).

Para Kimball (2002), um ODS é frequentemente implementado para produzir relatórios operacionais, principalmente quando nem os sistemas legados nem os sistemas de processamento de transações *On-Line* mais recentes fornecem relatórios operacionais adequados. Em outras situações, os ODSs são criados para suportar integrações em tempo real, principalmente aplicações de gerenciamento de relacionamentos com clientes.

Um ODS é o único lugar para se definir quais valores devem ser efetivamente importados dos sistemas legados. O ODS deve ser sempre utilizado para limpar dados desnecessários que entram no processo de extração e transformação, pois o quanto antes forem removidos, menor será possibilidade de erros. Como não é um componente indispensável em um ambiente de DW, sua criação é uma decisão de projeto. (MACHADO, 2000).

A figura 7, a seguir, apresenta o fluxo de dados em um ambiente de DW com a utilização de ODS.

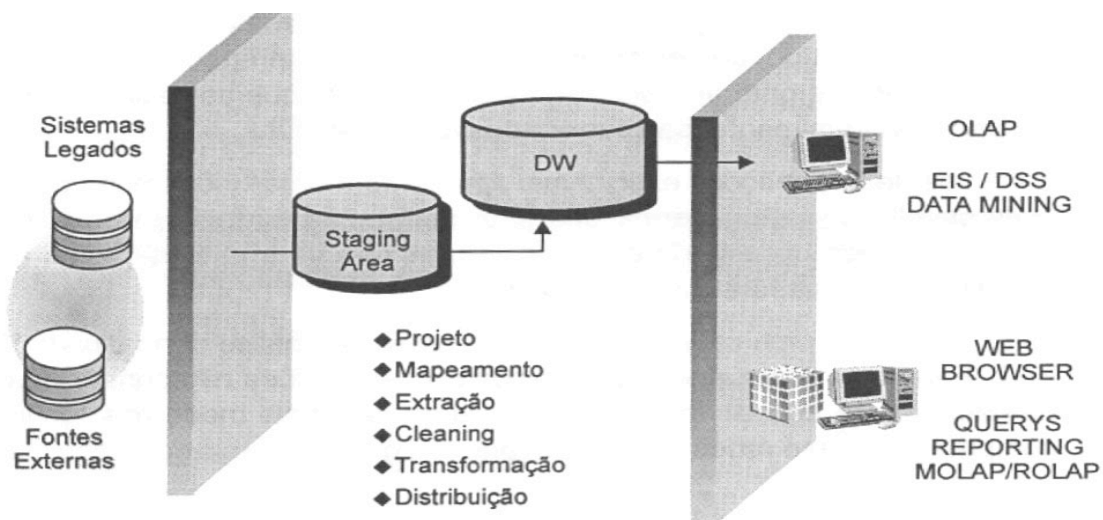


Figura 7 - Fluxo de dados em um ambiente de DW com ODS.
Fonte: Machado (2000).

2.4.9 Granularidade

Para Inmon (2000), a granularidade é a questão mais importante que um desenvolvedor precisa enfrentar em um projeto de DW. Quando a granularidade de um DW é definida corretamente, os demais aspectos de projeto e implementação seguem tranquilamente.

Granularidade de dados de um DW refere-se ao nível de sumarização dos elementos de dados e ao nível de detalhe disponível nos dados. Quanto mais detalhes disponíveis nos dados, menor é a granularidade e, conseqüentemente, quanto menor o nível de detalhes dos dados, maior é a granularidade. (MACHADO, 2000).

A primeira etapa para a definição do nível apropriado de granularidade consiste em fazer uma estimativa bruta do número de linhas de dados e do *Direct Access Storage Device* (DASD) que o DW conterà. Na seqüência, deve-se identificar as tabelas que serão criadas e estimar o tamanho da linha de cada tabela. Caso não seja possível identificar o tamanho exato, estimativas do limite inferior e superior são suficientes. Depois de realizadas as projeções de linhas de dados, as projeções do espaço de índices de dados são calculadas. Neste ponto, os números máximos e mínimos de ocorrências de linhas das tabelas são multiplicados respectivamente pelos comprimentos máximos e mínimos dos dados. Por fim, multiplicam-se os números mínimos e máximos de ocorrências de linhas nas tabelas pelos comprimentos máximos e mínimos dos dados, respectivamente. (INMON, 2000).

De acordo com Inmon (2000), o próximo passo consiste em comparar o número total de linhas do ambiente do DW aos mapas apresentados na figura 8.

horizonte de 1 ano		horizonte de 5 anos	
10.000.000	níveis duais de granularidade e projeto cauteloso	20.000.000	níveis duais de granularidade e projeto cauteloso
1.000.000	níveis duais de granularidade	10.000.000	níveis duais de granularidade
100.000	praticamente qualquer técnica funciona	1.000.000	projeto cauteloso
10.000	praticamente qualquer técnica funciona	100.000	praticamente qualquer técnica funciona

Figura 8 - Os patamares de granularidade.
Fonte: Inmon (2000).

Conforme o número de linhas que existirão em um ambiente de DW, diferentes enfoques de projeto e desenvolvimento deverão ser utilizados. Em um horizonte de um ano, se a perspectiva for de um total de dez mil linhas ou menos, então qualquer técnica de projeto e implementação poderá ser utilizada. Ainda para o horizonte de um ano, se o total for de cem mil linhas ou menos, deve-se conduzir o projeto com cautela. Caso o esperado para um ano ultrapasse o total de um milhão de linhas, níveis duais de granularidade deverão ser utilizados. E, caso o número total de linhas do ambiente DW ultrapasse dez milhões de linhas, será necessário conduzir de forma extremamente cautelosa o projeto e sua implementação, e níveis duais de granularidade devem ser adotados. (INMON, 2000).

Machado (2000) afirma que, para o sucesso de um projeto de DW, é vital a escolha correta do nível ou níveis de granularidades. Para o mesmo autor, o método mais indicado para esse fim é o uso do bom senso e da análise meticulosa das necessidades de informação levantadas para o projeto. Deve-se ouvir o usuário, discutir e sugerir alternativas afim de atingir uma correta granularidade de projeto.

2.4.10 Modelagem multidimensional

A modelagem multidimensional é uma técnica de concepção e visualização de um modelo de dados de um grupo de medidas que apresentam aspectos semelhantes de negócios. É usada especialmente para sumarizar e reestruturar dados e apresentá-los em visões que suportem a análise dos valores desses dados. O modelo multidimensional possui como conceitos gerais três elementos básicos: fatos, dimensões e medidas. (MACHADO, 2000).

2.4.10.1 Fatos

Machado (2000), define as tabelas fato como uma coleção de itens de dados, composta de dados de medidas e contexto. Cada fato representa um item de negócio, uma transação ou um evento de negócio, e é utilizado para analisar o processo de negócio de uma empresa. De acordo com o mesmo autor, a característica básica de uma tabela fato é que ela é representada por valores numéricos.

Para Barbieri (2001), as tabelas fatos são utilizadas para armazenar medidas numéricas associadas a eventos de negócio. Possuem como chave primária um campo de chave múltipla, formado pelas chaves primárias das dimensões que com ela se relacionam. De acordo com o mesmo autor, deve-se ter cuidado com o crescimento de dados das tabelas fatos, pois normalmente armazenam muito mais dados que as tabelas dimensão.

A figura 9 mostra a composição básica de uma tabela fato.

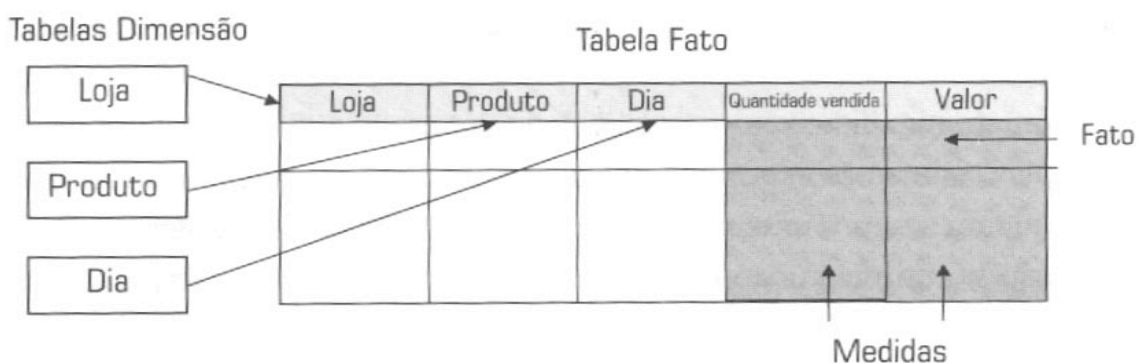


Figura 9 - Composição básica de uma tabela fato.
Fonte: Barbieri (2001).

2.4.10.2 Dimensões

As tabelas dimensão representam entidades de negócios e constituem as estruturas de entrada que servem para armazenar informações como tempo, geografia, produto, cliente, etc. Essas tabelas possuem uma relação 1:N (um para muitos) com a tabela fato, e a quantidade de dados é bem menor que nas fatos. Apresentam sempre uma chave primária para garantir a unicidade dos dados e participa da tabela fato como parte de sua chave múltipla. (BARBIERI, 2001).

Segundo Machado (2000), dimensões são conceitualmente os elementos que participam de um fato. Determinam o contexto de um assunto de negócios, por exemplo, um banco de dados que analisa as vendas de produtos possuem geralmente as dimensões tempo, localização, clientes, vendedores e cenários. O mesmo autor afirma que as dimensões normalmente não possuem atributos numéricos por serem somente descritivas e classificatórias dos elementos que participam de uma tabela fato.

2.4.10.3 Medidas

As medidas são os atributos numéricos que representam o desempenho de um indicador de negócios relativo às dimensões que participam desse fato. Por exemplo, medidas são valor em reais de vendas, o número de unidades de produtos vendidos, a quantidade em estoque, o custo de venda, etc. Uma medida é determinada pela combinação das dimensões que participam de um fato e que estão localizados como atributos. (MACHADO, 2000).

2.4.10.4 Modelo *Star Schema*

O termo *Star Schema* (modelo estrela) é geralmente utilizado para a designação de modelos de dados multidimensionais. Esse modelo é composto por uma grande entidade central denominada tabela fato e um conjunto de entidades menores denominadas dimensões,

posicionadas ao redor desta entidade central, formando uma estrela, como mostra a figura 10. (MACHADO, 2000).

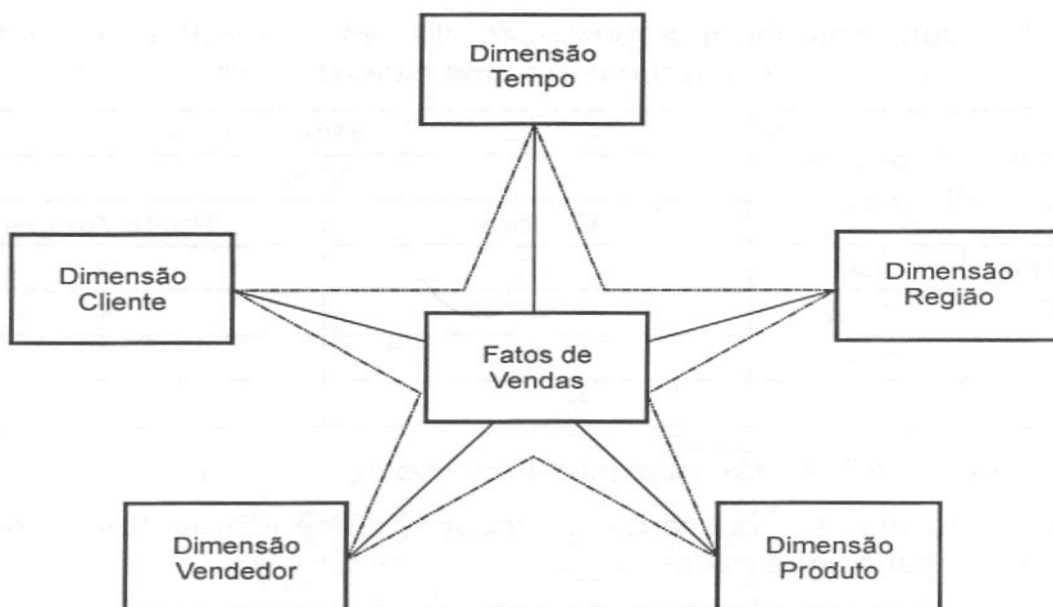


Figura 10 - Modelo *Star Schema* aplicado a tabela fato vendas.
Fonte: Machado (2000).

Para Corey e outros (2001), o modelo *Star* é a forma mais popular de construir estruturas de dados de DM de alto desempenho em um ambiente relacional. Para conseguir esses ganhos deve-se limitar o número de uniões que terão de ser realizadas e a complexidade de cada união.

Devido à maneira simples de organizar os dados, o modelo *Star* além de facilitar o processamento das consultas, permite uma melhor visualização dos dados. (MORALES, 2008).

2.4.10.5 Modelo *Snowflake*

O modelo de dados multidimensional geralmente tem seu formato semelhante a uma estrela, onde a tabela fato está no centro e as tabelas dimensões ao seu redor formando as pontas. O modelo *Snowflake* (flocos de neve) é o resultado da decomposição de uma ou mais

dimensões que possuem hierarquias entre seus membros. Esse modelo é de fácil compreensão pela equipe de desenvolvedores de sistemas OLTP, devido a aplicação das formas normais como em projeto relacional. Por ser um modelo normalizado, ele impede a redundância de valores textuais em uma tabela. (MACHADO, 2000).

Na sequência a ilustração apresenta o modelo *Snowflake*.

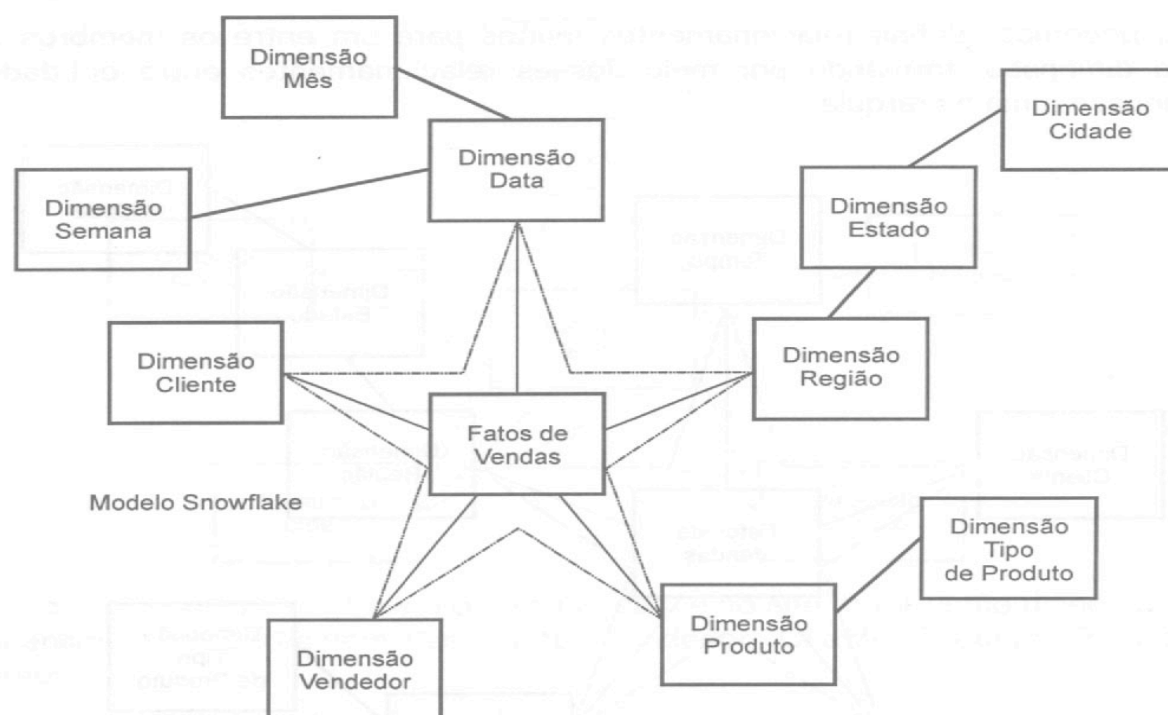


Figura 11 - Modelo *Snowflake*.
Fonte: Machado (2000).

2.4.10.6 Agregados

Segundo Barbieri (2001), um passo importante no projeto lógico de um DW ou DM é a etapa que envolve a definição de agregados. Os valores agregados representam, paradoxalmente, uma solução e alguns problemas. Solução, pois estabelecem a definição e a criação de tabelas prontas, trabalhadas e sumarizadas em várias dimensões, o que facilita os acessos aos dados e agilizam os processos decisórios. Problemas, porque de certa maneira ferem as regras de não redundância, estabelecidas nos princípios de banco de dados desde sua criação. Além disso, os valores agregados utilizam mais espaço por possuir uma coleção de

tabelas fato ou dimensão que agora estão dedicadas ao armazenamento de dados num estado já pré-processado.

2.5 ON-LINE ANALYTICAL PROCESSING - OLAP

2.5.1 Conceito

Para Thomsen (2002), o termo OLAP possui diversos significados. Isso ocorre porque os elementos essenciais do OLAP são manifestáveis em diversas camadas de tecnologia, desde armazenamento e acesso até as camadas de linguagem. De acordo com o mesmo autor, os conceitos de OLAP incluem a noção ou idéia de múltiplas dimensões hierárquicas e que podem ser utilizados por qualquer pessoa como auxílio a compreensão de um mundo multidimensional e com múltiplos níveis.

Um OLAP se refere ao tipo de processamento e ferramentas voltadas para análise de dados típica do suporte à decisão, onde os dados são exibidos através de uma visão multidimensional independente de como os dados estão armazenados. (SERRA, 2002).

A tecnologia OLAP é normalmente implementada em ambiente multiusuário e cliente/servidor, proporcionando respostas rápidas às consultas, não importando o tamanho do banco de dados nem sua complexidade. (SOUZA, 2003).

2.5.2 Tipos de OLAP

Para SOUZA (2003), a classificação das ferramentas OLAP é uma tarefa imprecisa, uma vez que não existe nenhuma característica peculiar que especifique como a ferramenta deva ser construída, qual tecnologia deva ser usada em sua construção, nem mesmo que funcionalidades devem ser implementadas.

De acordo com Barbieri (2001), a estratégia de armazenamento do DW/DM, ou de seus cubos¹ extraídos, possibilitam as seguintes opções:

- a) ROLAP – técnica onde são usados os próprios sistemas de gerenciamento de banco de dados, com as tabelas sendo implementadas como estruturas relacionais clássicas;
- b) MOLAP - estratégia onde são usados os gerenciadores de banco de dados proprietários com características de armazenamento especiais, e ferramentas para tratamento dimensional de dados. Podem ser entendidos como uma planilha multidimensional.
- c) HOLAP - representa uma abordagem de uso misto das duas estratégias anteriores, onde as estruturas relacionais são normalmente utilizadas para os dados de maior granularidade e as estruturas dimensionais nativas são dedicadas ao armazenamento de agregados;
- d) DOLAP - é uma abordagem onde as estruturas dimensionais ou relacionais são transferidas do DW/DM, e armazenadas nas estações dos clientes com o objetivo de melhorar a performance de certas análises, reduzindo o tráfego de informações entre os ambientes cliente e servidor.

2.5.3 Operadores OLAP

No modelo entidade relacionamento (ER) tradicional o conhecimento das operações básicas de álgebra relacional é obrigatório. No modelo de dados multidimensional existem as operações OLAP que suportam tais operações. A seguir descrevem-se as operações em um OLAP. (MACHADO, 2000).

2.5.3.1 *Drill Down e Roll Up*

Segundo Barbieri (2001), o conceito de *Drill Down* está diretamente relacionado com o fato de sairmos de um nível mais alto de hierarquia para buscar informações mais detalhadas, ou seja, com uma granularidade menor. O inverso é o conceito de *Roll Up* ou *Drill*

¹ A metáfora CUBO é apenas uma forma de visualização de como os dados estão organizados. (MACHADO 2000).

Up, que possibilita sair do nível mais baixo de hierarquia até o mais alto nível de sumarização.

Na figura 12 pode-se visualizar uma operação *Drill Down* na dimensão localização geográfica.

Volume de Produção (em milhares)		Telefone Celular		Pagers	
		1001	1002	2001	2002
Região Sul	RGS	33	12	8	12
	SC	45	34	20	23

Drill Down

Dimensão Localização Geográfica
Membro RGS

Volume de Produção (em milhares)		Telefone Celular		Pagers	
		1001	1002	2001	2002
RGS	Canoas	13	4	2	5
	Porto Alegre	20	8	6	7

Figura 12 - Operação *Drill Down* na dimensão geográfica.
Fonte: Machado (2000).

A operação *Roll Up* na tabela dimensão tempo é apresentada na figura 13.

Volume de Produção (em milhares)		1999			
		Trim. 1	Trim. 2	Trim. 3	Trim. 4
Região Sul	RGS	78	67	22	56
	SC	90	67	88	99

Roll Up

Dimensão Tempo

Volume de Produção (em milhares)		Trimestre 1		
		Janeiro	Fevereiro	Março
Região Sul	RGS	30	26	22
	SC	28	30	32

Figura 13 - Operação *Roll Up* na dimensão tempo.
Fonte: Machado (2000).

2.5.3.2 Slices

São operações para realizar navegação através dos dados na visualização de um cubo. *Slice* é a operação que corta o cubo mantendo a mesma perspectiva de visualização de dados, já a troca de perspectiva de visão é denominada *dice* (figura 14). (MACHADO, 2000).

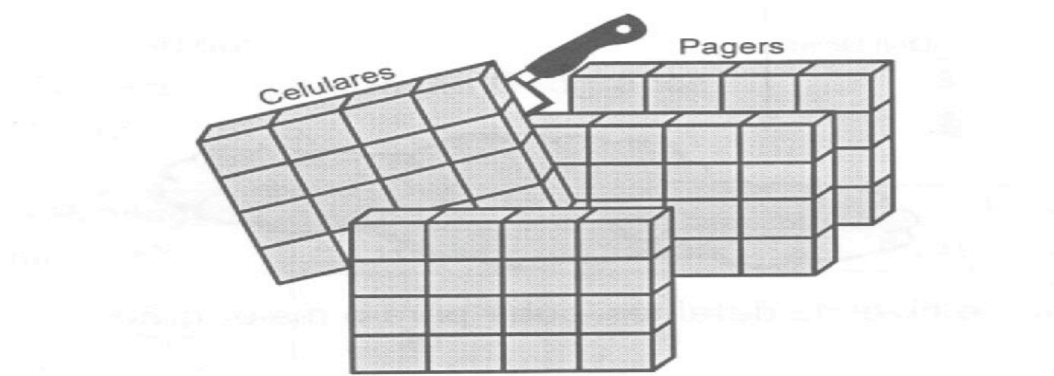


Figura 14 - Fatias (*slices*) de um cubo de dados.
Fonte: Machado (2000).

2.5.3.3 Drill Across

O conceito de *Drill Across* está relacionado com a possibilidade da troca de um esquema para outro, desde que ambos tenham dimensões em conformidade. Pode ser considerado como uma espécie de *join* (junção) dimensional entre estruturas relacionadas. (BARBIERI, 2001).

Os principais comandos dimensionais são ilustrados pela figura 15, na seqüência.

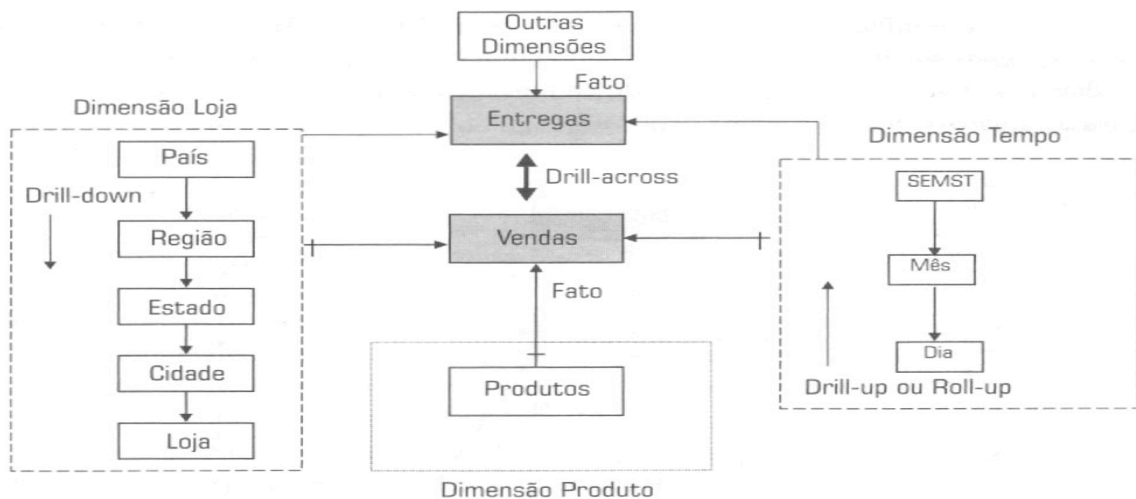


Figura 15 - Representação dos principais comandos dimensionais.
Fonte: Barbieri (2001).

2.6 CONSIDERAÇÕES FINAIS

Neste capítulo, apresentou-se uma revisão teórica dos sistemas de informações, de seus conceitos e sua evolução até chegar ao conceito de *Business Intelligence*, que é a evolução do *Executive Information System* (EIS) ou Sistema de Informação Executiva (SIE) e do *Decision Support System* (DSS) ou Sistema de Apoio à Decisão (SAD).

Assim pode-se conhecer toda a arquitetura de um sistema de BI e as tecnologias de informação a ele relacionadas, com o propósito de fundamentar o presente trabalho.

Na sequência, o capítulo 3 apresenta a modelagem do protótipo do DW desenvolvido.

3 MODELAGEM

3.1 INTRODUÇÃO

Neste capítulo apresenta-se a modelagem do DW. Iniciando-se com o levantamento dos requisitos que contemplam as regras de negócio, os requisitos funcionais e não funcionais e casos de uso. Por fim, é apresentada a modelagem dimensional do DW.

3.2 LEVANTAMENTO DE REQUISITOS

Segundo Silva (2008), levantamento de requisitos é toda a atividade realizada para identificar, analisar, especificar e definir as necessidades de negócio que um sistema deve fornecer para a solução de um problema.

3.2.1 Regras de negócio

Para Booch e outros (1997 apud NETO, 2008), uma regra de negócio pode ser definida segundo duas perspectivas: a visão do negócio e a visão dos sistemas de informação. Para a perspectiva do negócio, uma regra de negócio é uma diretiva destinada a nortear o comportamento do negócio, como suporte à política de negócio que é formulada em resposta a uma oportunidade. Já na perspectiva dos sistemas de informação, uma regra de negócio é uma sentença que define ou restringe algum aspecto do negócio.

As regras de negócios para este projeto de DW foram elaboradas em conjunto com a empresa de transporte coletivo na qual foi realizado o estudo caso e são apresentadas na sequência, figura 16.

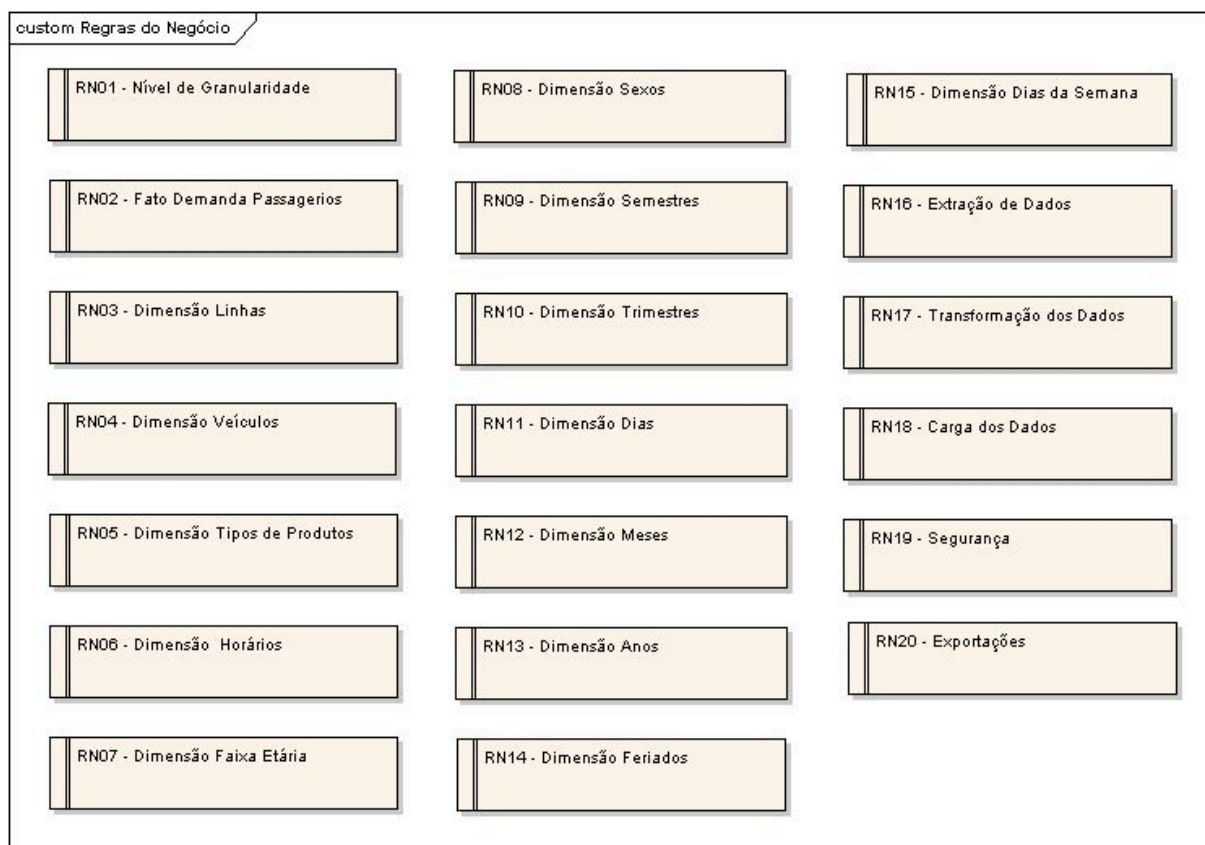


Figura 16 - Regras de negócio.
Fonte: Elaboração do autor, 2008.

O quadro 6 apresenta a documentação dos requisitos de negócio do DW ilustrados pela figura 16.

Regra de negócio	Descrição
RN01 – Nível de granularidade	O DW deve ter níveis duais de granularidade. <ul style="list-style-type: none"> • Mensal • Diário • Hora
RN02 – Fato demanda passageiros	O sistema deverá conter a seguinte medida na tabela fato: <ul style="list-style-type: none"> • Quantidade de Passageiros Entidades de negócio relacionadas (Dimensões): <ul style="list-style-type: none"> • Linhas, veículos, produtos, horários, faixa etária, sexos, semestres, trimestres, dias, meses, anos, dia da semana e feriados.

RN03 – Dimensão linhas	<p>O sistema deve conter uma tabela de linhas sincronizada com os sistemas operacionais (inclusive planilhas eletrônicas).</p> <p>A tabela de linhas deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Código; • Descrição • kmIda; • kmVolta • TempoIda • TempoVolta • Tipo (Circular, Radial, Alimentadora e Troncal) • Vigente (Sim e Não)
RN04 – Dimensão veículos	<p>O sistema deve conter uma tabela de veículos sincronizada com os sistemas operacionais (inclusive planilhas eletrônicas).</p> <p>A tabela de veículos deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Prefixo • Fabricante • Potência • Consumo Médio • Qtd. Passageiros em Pé • Qtd. Passageiros Sentados • Tipo (Convencional, Microônibus e Articulado) • Idade (Anos)
RN05 – Dimensão tipos de produtos	<p>O sistema deve conter uma tabela de Tipos de Produtos sincronizada com os sistemas operacionais.</p> <p>A tabela de Tipos de Produtos deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Descrição • Valor • Código Produto • Vigente (Sim/Não)
RN06 – Dimensão horários	<p>O sistema deve conter uma tabela com Horários.</p> <p>A tabela de Horários deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Descrição • Hora (00:00, 01:00..23:59) • Período (Matutino, Vespertino e Noturno)

RN07 – Dimensão faixa etária	<p>O sistema deve conter uma tabela com Faixas Etárias.</p> <p>A tabela de Faixa Etária deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Descrição • Idade • Categoria (Criança, Adolescente, Adulto e Idoso) • <p>Faixa Etárias (Categorias):</p> <ul style="list-style-type: none"> • Crianças: de 0 a 14 anos • Adolescente: de 15 a 20 anos • Adulto: de 21 a 59 anos • Idosos: Mais de 60 anos
RN08 - Dimensão sexos	<p>O sistema deve conter uma tabela com Sexos.</p> <p>A tabela de Sexos deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Descrição (FEMININO, MASCULINO) • Sexo (F, M)
RN09 - Dimensão semestres	<p>O sistema deve conter uma tabela com os dois semestres do ano.</p> <p>A tabela de Semestres deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Descrição (Primeiro Semestre, Segundo Semestre) • Semestre (1,2)
RN10 – Dimensão trimestres	<p>O sistema deve conter uma tabela com os quatro semestres do ano.</p> <p>A tabela de Semestres deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Descrição (Primeiro Trimestre, Segundo Trimestre, Terceiro Semestre e Quarto Semestre) • Trimestre (1,2,3,4)
RN11 – Dimensão dias	<p>O sistema deve conter uma tabela com os dias do mês.</p> <p>A tabela de dias deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Dia (1,2,3,4,5..30,31)
RN12 – Dimensão meses	<p>O sistema deve conter uma tabela com os doze meses do ano.</p> <p>A tabela de Meses deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Descrição (Janeiro, Fevereiro, Março..Dezembro) • Mês (1,2,3,4,5,6,7,8,9,10,11,12)
RN13 – Dimensão anos	<p>O sistema deve conter uma tabela com anos.</p> <p>A tabela de Anos deverá contemplar os seguintes dados:</p> <ul style="list-style-type: none"> • Ano (De 2007 até 2020)

RN14 – Dimensão feriados	O sistema deve conter uma tabela de feriados sincronizada com os sistemas operacionais. A tabela de feriados deverá contemplar os seguintes dados: <ul style="list-style-type: none"> • Descrição • Data
RN15 – Dias da semana	Descrição: O sistema deve conter uma tabela com os dias da semana. A tabela de dias da semana deverá contemplar os seguintes dados: <ul style="list-style-type: none"> • Descrição (Domingo, Segunda, Terça, Quarta, Quinta, Sexta e Sábado) • Dia da Semana (1,2,3,4,5,6,7)
RN16 – Extração de dados	O sistema deverá extrair apenas dados operacionais referente a demanda da passageiros.
RN17 – Transformação dos dados	O sistema deverá transformar os dados operacionais em dados que permitam a carga das tabelas de dimensões e da tabela fato do Data Warehouse. Os dados operacionais devem ser agrupados por: dia, mês, ano, feriado, dia semana, semestre, trimestre, faixa de horário, faixa etária, sexo, linha, prefixo.
RN18 – Carga dos dados	O sistema deverá realizar a inserção dos dados extraídos dos sistemas operacionais nas tabelas de dimensões e na tabela fato do DW. Os dados das tabelas de dimensões devem ser atualizadas quando novos dados são inseridos ou alterados nos sistemas operacionais.
RN19 – Segurança	O sistema deverá prover ferramentas de controle de acesso as informações do Data Warehouse.
RN20 – Exportações	O sistema deverá possibilitar a exportação das informações das consultas realizadas. Formatos: XLS, PDF e XML

Quadro 6 - Regras de negócio do DW.
Fonte: Elaboração do autor, 2008.

3.2.2 Requisitos funcionais

Os requisitos funcionais são a descrição das diversas funções que clientes e usuários desejam ou precisam que o software execute. O termo função é usado no sentido genérico da operação que pode ser realizada pelo sistema, seja através comandos dos usuários ou pela ocorrência de eventos internos ou externos ao sistema. (MACORATTI, 2008a).

A figura 17, a seguir, apresenta o diagrama de requisitos funcionais levantados para este projeto de DW.



Figura 17 - Requisitos funcionais.
Fonte: Elaboração do autor, 2008.

Na sequência, o quadro 7 documenta os requisitos funcionais apresentados pela figura 17.

Requisitos Funcionais	Descrição
RF01 – Consulta OLAP	O sistema deverá prover funcionalidades que permitam ao usuários acessar, visualizar, e analisar uma grande quantidade de dados corporativos com alta flexibilidade e performance. Regras de Negócio Relacionadas: <ul style="list-style-type: none"> • RN02 - Fato Demanda Passageiros • RN18 - Segurança
RF02 – Extração	O sistema deverá prover funcionalidades que permitam ao usuário realizar a extração de dados de diversos sistemas. Regras de Negócio Relacionadas: <ul style="list-style-type: none"> • RN15 - Extração de Dados
RF03 – Transformação	O sistema deve realizar a transformação dos dados extraídos conforme regras de negócios. Regras de Negócio Relacionadas: <ul style="list-style-type: none"> • RN16 - Transformação dos Dados

RF04 – Carga de Dados	O sistema deverá realizar após a extração e a transformação, a carga dos dados em um <i>data warehouse</i> . Regras de Negócio Relacionadas: • RN17 - Carga dos Dados
RF05 – Exportação de Dados	O sistema deverá prover funcionalidades que permitam ao usuário fazer exportação das informações para a plataforma Office e em PDF. Regras de Negócio Relacionadas: RN18 - Segurança
RF06 – Impressão	O sistema deverá prover ferramentas de impressão das informações obtidas nas consultas OLAP e dos relatórios. Regras de Negócio Relacionadas: RN18 - Segurança

Quadro 7 - Documentação requisitos funcionais.
Fonte: Elaboração do autor, 2008.

3.2.3 Requisitos não funcionais

Requisitos não funcionais são as qualidades globais de um software, como usabilidade, desempenho, custos e várias outras. (MACORATTI, 2008a).

Os requisitos não funcionais definidos para o DW são ilustrados pelo diagrama de requisitos, figura 18.

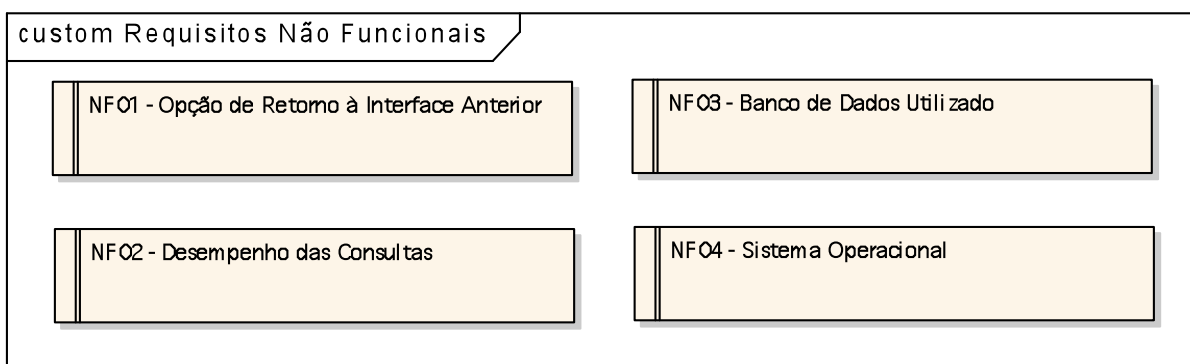


Figura 18 - Requisitos não funcionais.
Fonte: Elaboração do autor, 2008.

O quadro 8, descreve os requisitos não funcionais ilustrados pela figura 18.

Requisito não funcional	Descrição
NF01 – Opção de retorno à interface anterior	Toda interface deverá conter uma opção (um botão ou um link) que possibilite o retorno para a interface anterior. Critério: Usabilidade
NF02 – Desempenho das consultas	Nenhuma consulta realizada pelo sistema deve demorar mais que 15 segundos para que se obtenha uma resposta. Critério: Desempenho
NF03 – Banco de dados utilizado	O banco de dados deverá utilizar o SQL Server 2005 Std. por ser atualmente utilizado nos sistemas operacionais da empresa. Critério: Requisito do sistema
NF04 – Sistema operacional	O Sistema deverá ser compatível com os sistemas operacionais Windows XP e Windows 2003 Server Std. Critério: Requisito do sistema.

Quadro 8 - Requisitos não funcionais.
Fonte: Elaboração do autor, 2008.

3.2.4 Casos de uso

Segundo Macoratti (2008b), um caso de uso é uma técnica de modelagem utilizada para descrever o que um sistema deve realizar. Ele é construído através de um processo iterativo no qual as reuniões entre o cliente e os desenvolvedores do sistema conduzem a uma especificação do sistema.

Um caso de uso tem por objetivo apresentar uma visão geral das funções e serviços que o sistema deverá oferecer aos usuários, sem focar como essas funções serão implementadas. (GUEDES, 2004).

O diagrama de casos de uso para o projeto do DW é apresentado na sequência, figura 19.

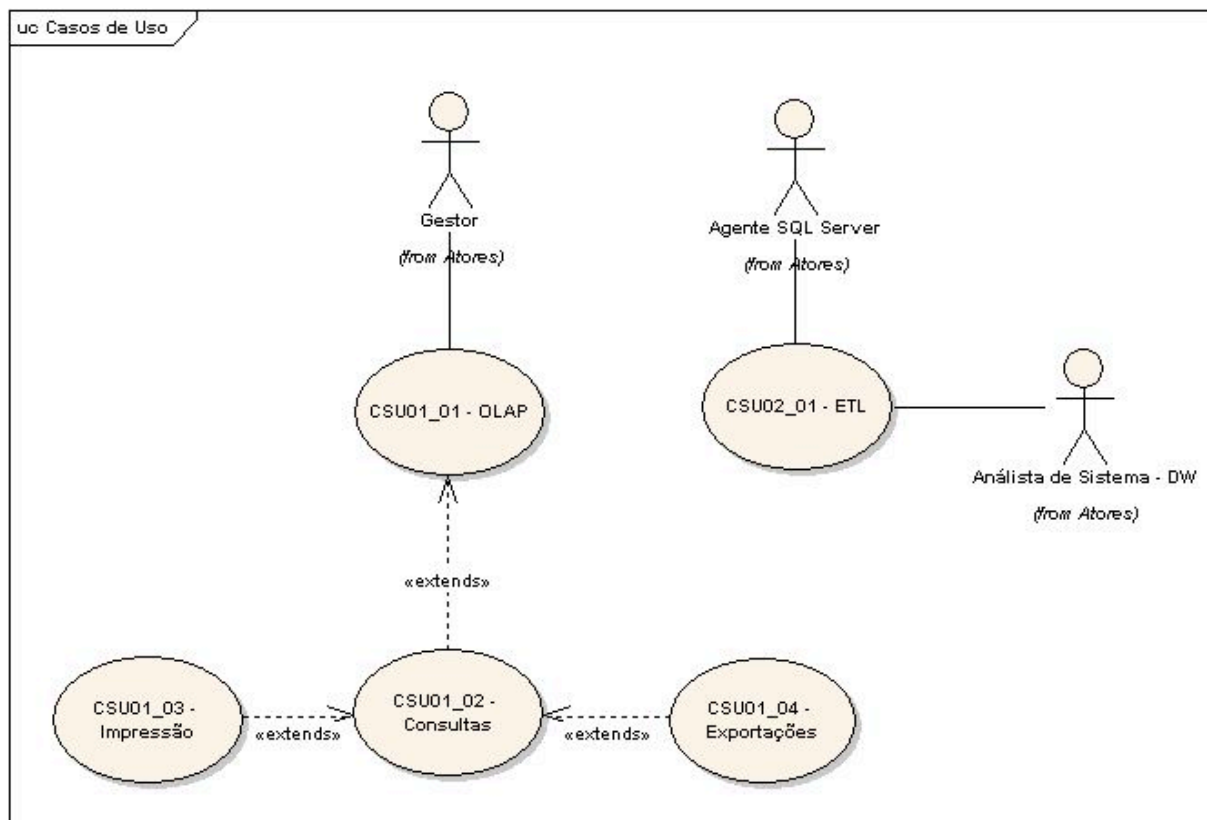


Figura 19 - Diagrama de casos de uso.

Fonte: Elaboração do autor, 2008.

A documentação dos casos de uso apresentado pela figura 19 é detalhado na sequência pelo quadro 9.

Caso de uso	Descrição
USU01_01 OLAP	<p>Este caso de uso tem por objetivo permitir ao usuário consultar, imprimir, exportar e criar relatórios com as informações dos cubos de dados.</p> <p>ATORES: Gestor</p> <p>PRÉ-CONDIÇÕES: não tem</p> <p>PÓS-CONDIÇÕES: Relatórios, Consultas, impressão e exportação</p> <p>REGRAS DE NEGÓCIO:</p> <ul style="list-style-type: none"> • RN18, RN02 e RN19 <p>REQUISITOS FUNCIONAIS:</p> <ul style="list-style-type: none"> • RF01, RF05, RF06 e RF07
USU01_02 Consultas	<p>Este caso de uso tem por objetivo permitir ao usuário navegar e visualizar as informações dos cubos de dados.</p> <p>ATORES: Gestor</p> <p>PRÉ-CONDIÇÕES: USU01_01 OLAP</p> <p>PÓS-CONDIÇÕES: USU01_03 Impressão e USU01_04 Exportações</p>

	<p>REGRAS DE NEGÓCIO:</p> <ul style="list-style-type: none"> • RN02, RN03, RN04, RN05, RN06, RN07, RN08, RN09, RN10, RN11, RN12, RN13, RN14 e RN18 <p>REQUISITOS FUNCIONAIS:</p> <ul style="list-style-type: none"> • RF01 e RF06
USU01_03 Impressão	<p>Este caso de uso tem por objetivo permitir ao usuário imprimir as informações obtidas nas consultas aos cubos de dados.</p> <p>ATOES: Gestor</p> <p>PRÉ-CONDIÇÕES: USU01_02 Consultas</p> <p>PÓS-CONDIÇÕES: Não tem.</p> <p>REGRAS DE NEGÓCIO:</p> <ul style="list-style-type: none"> • RN18 e RN19 <p>REQUISITOS FUNCIONAIS:</p> <ul style="list-style-type: none"> • RF01, RF05 e RF7
USU01_04 Exportações	<p>Este caso de uso tem por objetivo permitir ao usuário exportar as informações obtidas nas consultas aos cubos de dados.</p> <p>ATOES: Gestor</p> <p>PRÉ-CONDIÇÕES: USU01_02 Consultas</p> <p>PÓS-CONDIÇÕES: Não tem.</p> <p>REGRAS DE NEGÓCIO:</p> <ul style="list-style-type: none"> • RN18 e RN19 <p>REQUISITOS FUNCIONAIS:</p> <ul style="list-style-type: none"> • RF06
USU02_01 ETL	<p>Este caso de uso tem por objetivo realizar a extração, transformação e carga de dados nas dimensões e na tabela fato do DW.</p> <p>ATOES: Analista DW e Agente SQL Server</p> <p>PRÉ-CONDIÇÕES: Não tem.</p> <p>PÓS-CONDIÇÕES: Não tem.</p> <p>REGRAS DE NEGÓCIO:</p> <ul style="list-style-type: none"> • RN15, RN16 e RN17 <p>REQUISITOS FUNCIONAIS:</p> <ul style="list-style-type: none"> • RF02, RF03 e RF04

Quadro 9 - Documentação dos casos de uso.

Fonte: Elaboração do autor, 2008.

3.3 MODELAGEM DIMENSIONAL

A modelagem dimensional para esse projeto de DW foi baseada no modelo estrela ou *Star Schema*, afim de proporcionar um melhor desempenho no processamento dos dados.

3.3.1 Modelo Lógico

Na sequência, a figura 20 apresenta o modelo lógico do banco de dados utilizado no DW.

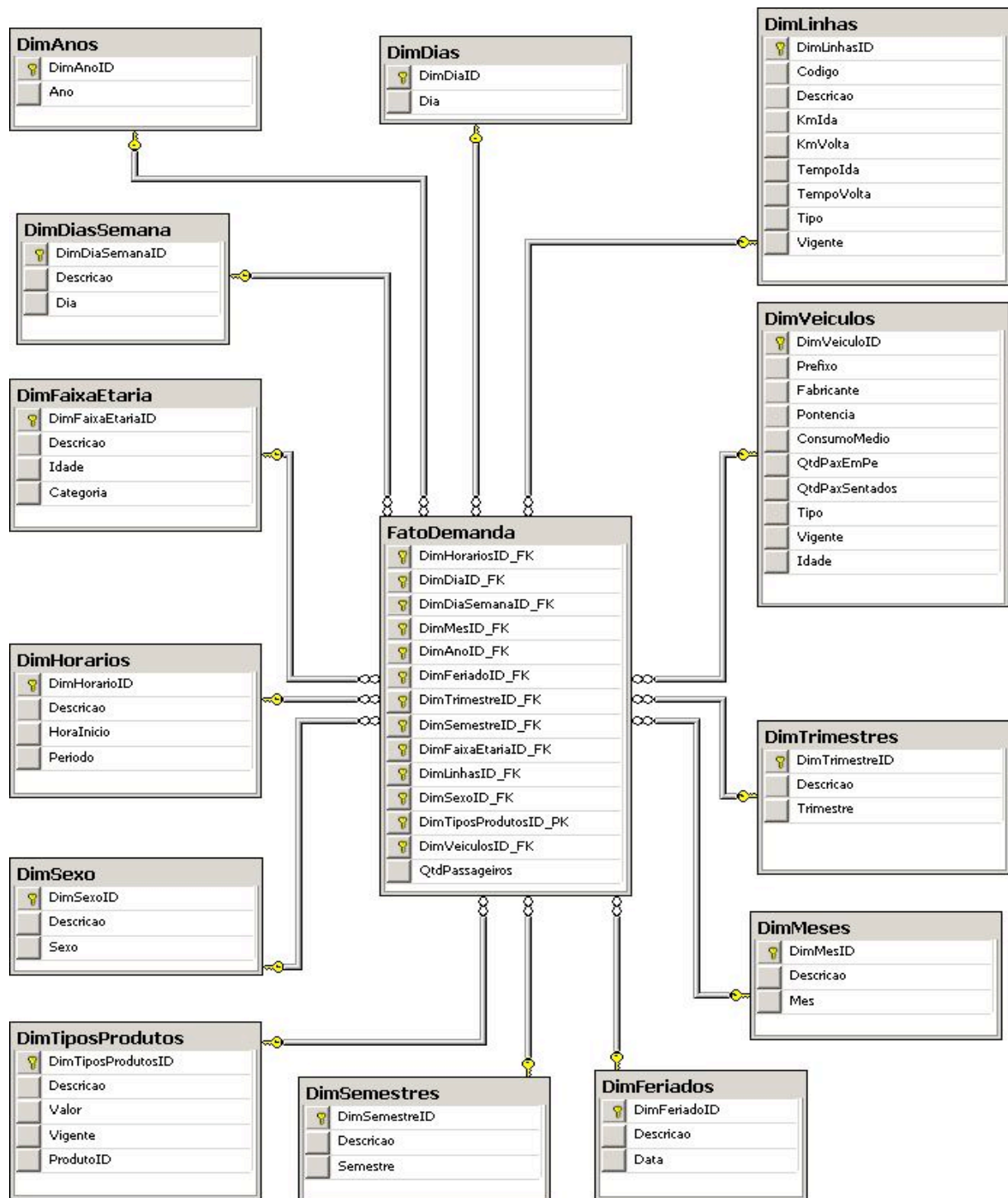


Figura 20 - Modelo lógico do banco de dados.

Fonte: Elaboração do autor, 2008.

3.3.2 Modelo Físico

O modelo físico do banco de dados é ilustrado na seqüência, figura 21.

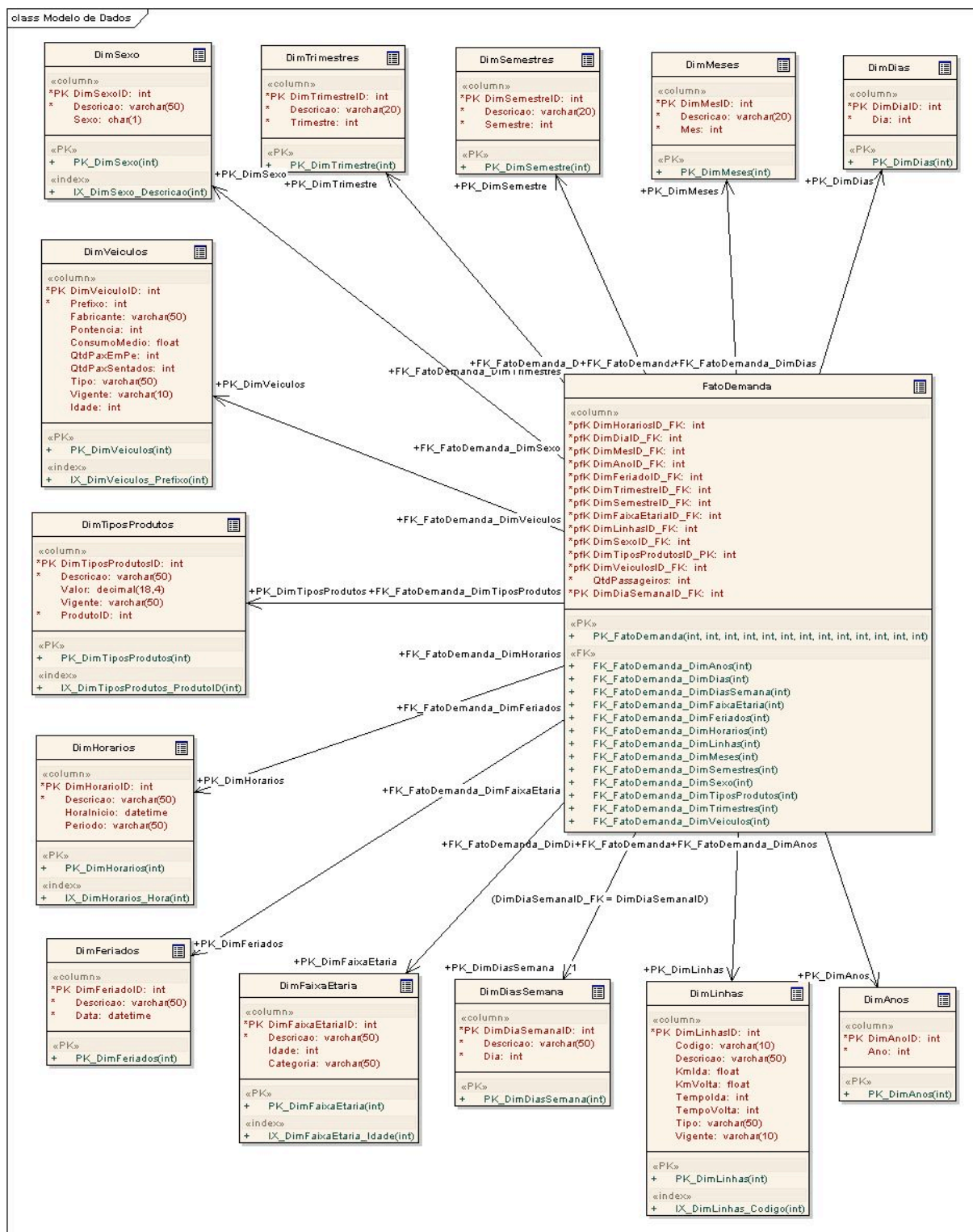


Figura 21 - Modelo físico do banco de dados.

Fonte: Elaboração do Autor, 2008.

3.4 CONSIDERAÇÕES FINAIS

Neste capítulo, apresentou-se o levantamento de requisitos e a modelagem dimensional do DW. Para isso, foi necessário conhecer pessoalmente o *modus operandi* da empresa de transporte coletivo acima citada afim de definir suas reais necessidades de informações. A ferramenta *Enterprise Architect*, desenvolvida pela Sparx Systems, possibilitou a criação de forma visual dos diagramas UML e da documentação acima apresentados.

Na sequência, o capítulo 4 apresenta o sistema implementado em conformidade com as informações deste capítulo.

4 SISTEMA IMPLEMENTADO

4.1 INTRODUÇÃO

Neste capítulo apresenta-se a solução proposta, as ferramentas utilizadas para implementar o DW, o processo de implementação, o protótipo desenvolvido, a validação dos dados e os resultados e conclusões obtidos.

4.2 SOLUÇÃO PROPOSTA

A solução para o problema descrito anteriormente no capítulo 1, consiste no desenvolvimento de um DW e na visualização de suas informações através da ferramenta OLAP. Para isso, utiliza-se o banco de dados Microsoft *SQL Server 2005 Standard* para armazenamento dos dados, o Microsoft *Analysis Services* como servidor OLAP, o Microsoft *Integration Services* para realizar o ETL dos sistemas operacionais e, por fim, o Microsoft Excel 2007 para visualização e análise dos dados disponibilizados pelo servidor OLAP.

A figura 22 ilustra a arquitetura da solução proposta.

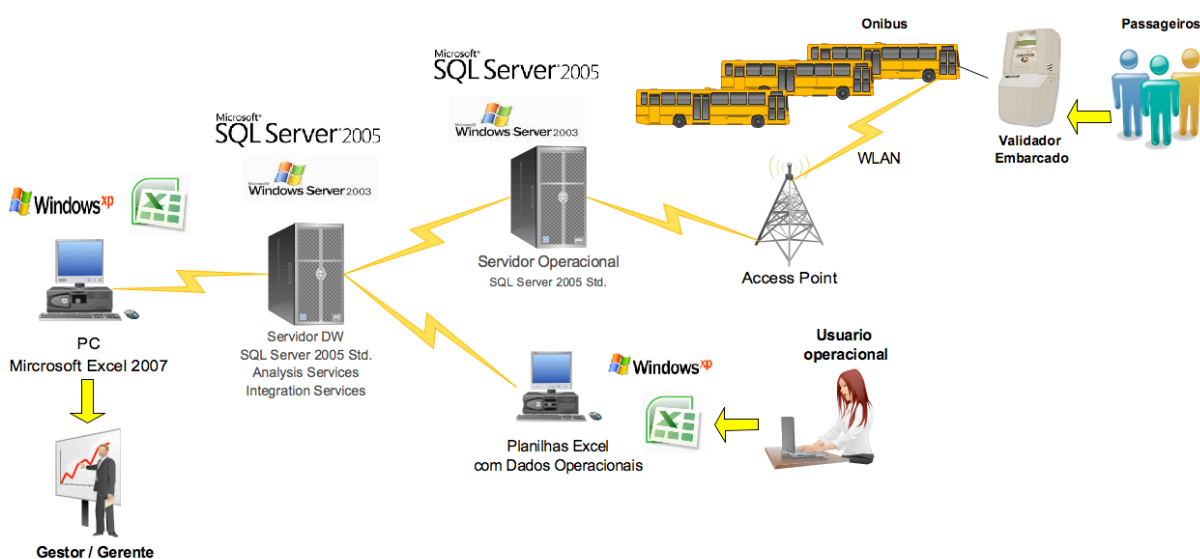


Figura 22 - Solução proposta.
Fonte: Elaboração do autor, 2008.

O quadro 11 descreve, em detalhes, a solução proposta apresentada na figura 22.

Item	Descrição
Passageiros	Os passageiros embarcam nos ônibus e transpõem a catraca apresentando seu cartão ou pagando a passagem em dinheiro ao cobrador.
Validador Embarcado	É o equipamento que efetua créditos ou débitos nos cartões dos passageiros e controla os passageiros pagantes em dinheiro liberados pelo funcionário cobrador dentro do veículo. Todos os eventos são armazenados na memória do equipamento.
Ônibus	Veículo de transporte coletivo de passageiros dotado de uma catraca eletromecânica e um equipamento validador.
<i>Access Point (AP)</i>	Antena WI-FI que realiza a comunicação sem fio com os validadores embarcados nos ônibus para coletar os eventos registrados.
Servidor Operacional	Os dados recebidos dos validadores embarcados são inseridos no banco de dados do sistema operacional da empresa.
Planilhas Eletrônicas Operacionais	São utilizadas como ferramentas de apoio pelos funcionários operacionais.
Servidor DW	Servidor utilizado para executar o Microsoft SQL <i>Server</i> e suas ferramentas de BI. Armazena os dados extraídos pelo <i>Integration Services</i> e fornece um servidor OLAP através do <i>Analysis Services</i> .
PC – Estação de Trabalho	Computador com o sistema operacional Windows e Excel 2007 instalados. Permite, ao usuário responsável pela tomada de decisões, realizar consultas dimensionais com as informações disponibilizadas pelo servidor OLAP (<i>Analysis Services</i>).

Quadro 10 - Descrição da ilustração da solução proposta.

Fonte: Elaboração do autor, 2008.

4.3 FERRAMENTAS UTILIZADAS

Foram utilizadas, na implementação deste protótipo, as ferramentas de BI da Microsoft integradas ao seu banco de dados *SQL Server 2005 Standard*. Dentre as razões para escolha da solução de BI destacam-se:

- a) Redução de custos com aquisição de ferramentas adicionais de BI;
- b) Maior facilidade de integração com os sistemas operacionais utilizados pela empresa fonte dos dados com os quais foi realizado o estudo de caso;

- c) Ferramenta de fácil aprendizado (no caso específico deste protótipo, dedicaram-se menos de 3 meses para aprendizado da ferramenta e sua implementação);
- d) Integração com o Excel 2007 e sua ferramenta de consulta OLAP.

4.4 PASSOS DA IMPLEMENTAÇÃO

Nesta seção descrevem-se os passos que foram necessários para implementação do DW com o uso do *SQL Server 2005 Standard*.

O quadro 11 descreve, detalhadamente, todos os passos realizados para implantação do DW.

Passo	Descrição
1	Utilizou-se, para a criação ² do esquema físico do DW, o <i>SQL Manager</i> . Esta ferramenta permitiu criar visualmente o esquema físico do modelo dimensional no servidor de banco de dados, bem como criar a integridade referencial, os <i>constraints</i> e as chaves estrangeiras.
2	Para popular as tabelas do DW com dados, utilizou-se o <i>Integration Services</i> . Esta ferramenta permitiu a extração de dados de diversas fontes externas como de planilhas eletrônicas e do banco de dados operacional ³ . Após a extração, realizaram-se transformações e agrupamento dos dados, finalizando com a sua inserção nas tabelas do DW.
3	Após a carga de dados para o DW utilizou-se a ferramenta <i>Analysis Services</i> para gerenciar os dados do modelo multidimensional. Essa ferramenta possibilitou organizar os dados do DW em cubos dimensionais permitindo consultas complexas e o processamento de uma grande quantidade de informações.
4	Para realizar consultas e análises dos dados disponibilizados pelo <i>Analysis Services</i> foi utilizado o Microsoft Excel 2007. Esta ferramenta permitiu a criação de consultas de uma forma visual pelo usuário, facilitando seu aprendizado.

Quadro 11 - Passos da implementação do DW.

Fonte: Elaboração do autor, 2008.

² O código DDL (*Data Definition Language*) gerado pela ferramenta *SQL Manager* é apresentado no apêndice A.

³ Os apêndices B e C apresentam o comando SQL utilizado para criação das *Views* na base operacional.

4.5 PROTÓTIPO DESENVOLVIDO

O protótipo foi desenvolvido com a utilização conjunta das ferramentas *Integration Services*, *Analysis Service* e Excel. Os dados operacionais utilizados foram disponibilizados pela empresa de transporte coletivo Transul.

A figura 23 apresenta a interface principal da ferramenta *Integration Services* utilizada para a extração, transformação e carga (ETL) dos dados no DW.

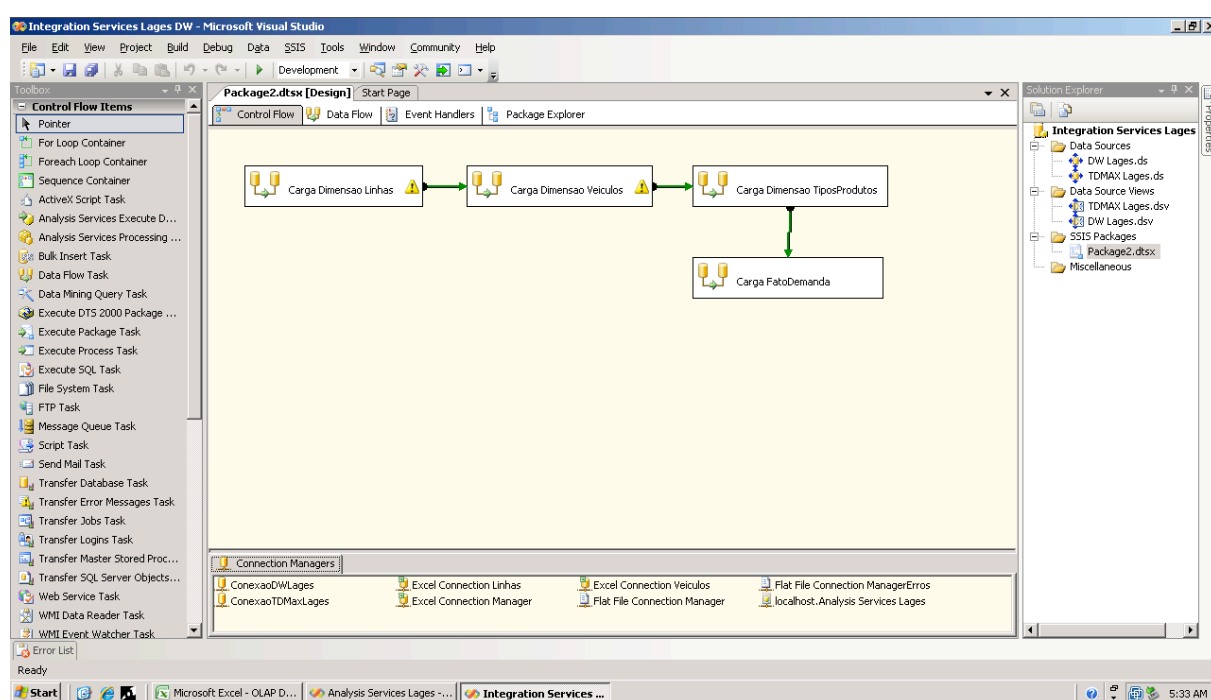


Figura 23 – Interface de controle de fluxo do *Integration Services*.
Fonte: Elaboração do autor, 2008.

Os componentes do *Integration Services*, utilizados para realizar a carga da dimensão com os dados das linhas operadas pela empresa acima citada, são apresentados na figura 24.

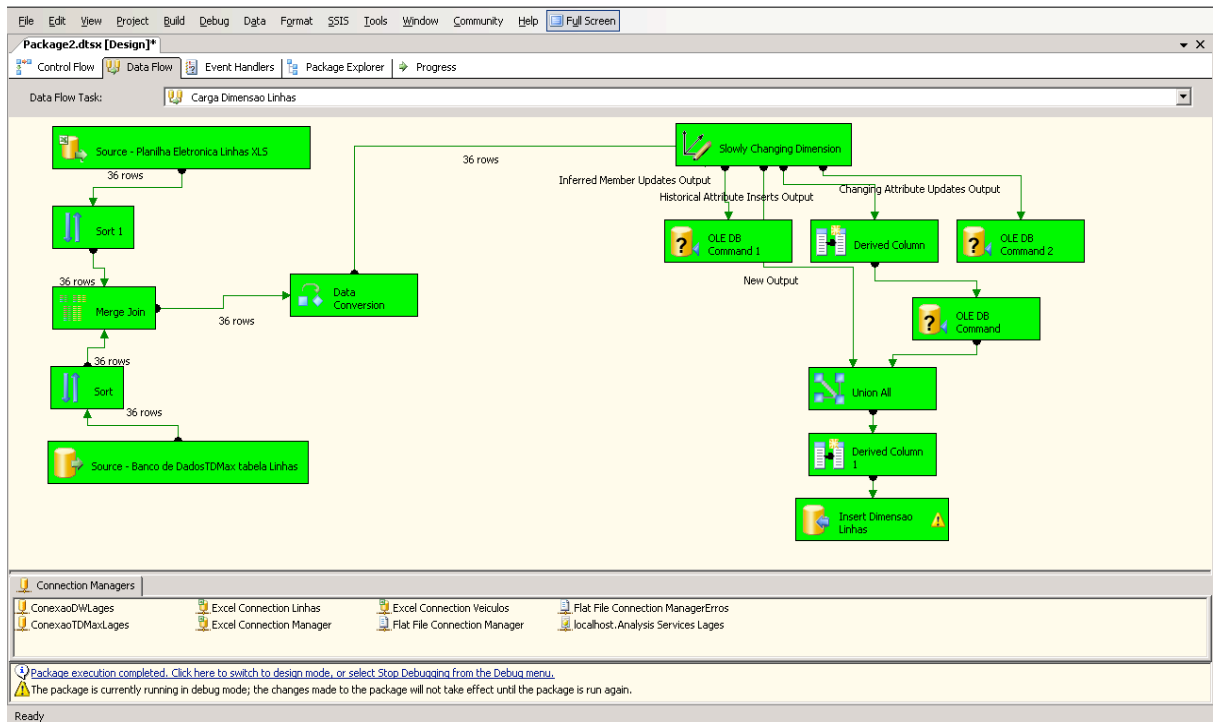


Figura 24 – Interface do *data flow* para carga da dimensão linhas no *Integration Services*.
Fonte: Elaboração do autor, 2008.

Na figura 25 são apresentados os componentes utilizados no *Integration Services* para realizar a carga da tabela fato.

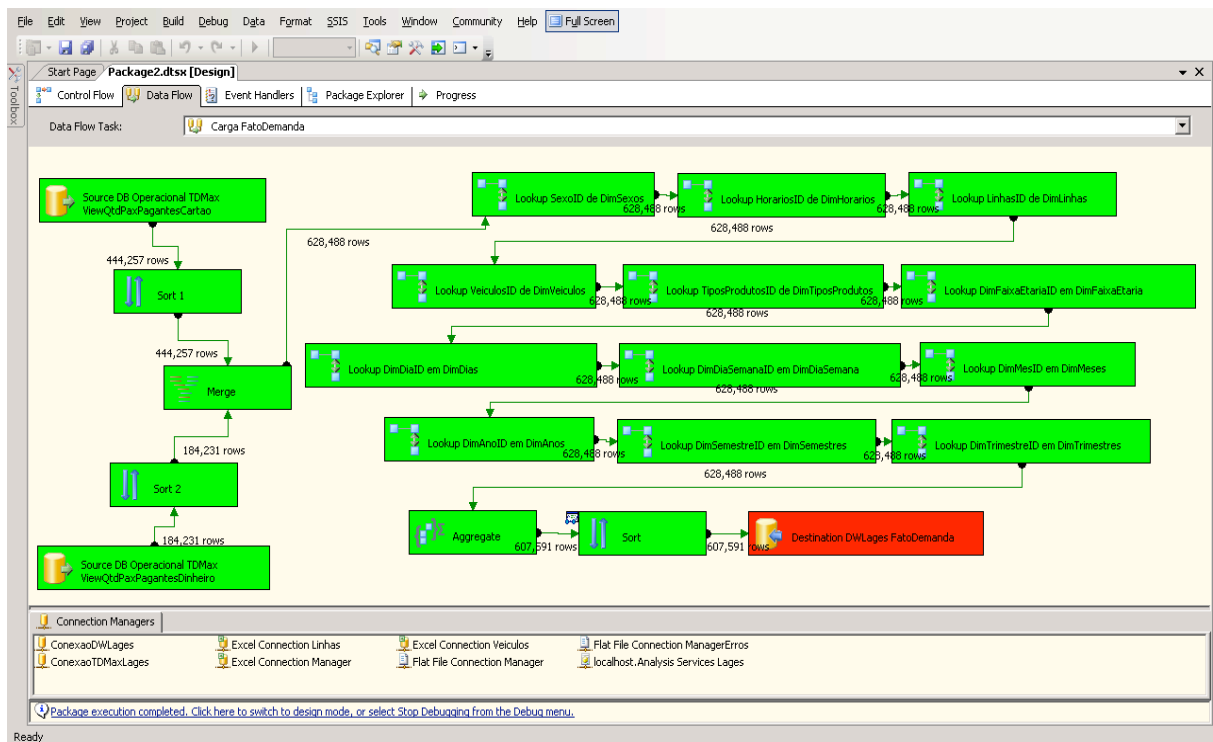


Figura 25 - Interface do *data flow* para carga da tabela fato no *Integration Services*.
Fonte: Elaboração do autor, 2008.

A figura 26 apresenta a tela da estrutura multidimensional obtida com a ferramenta *Analysis Services*.

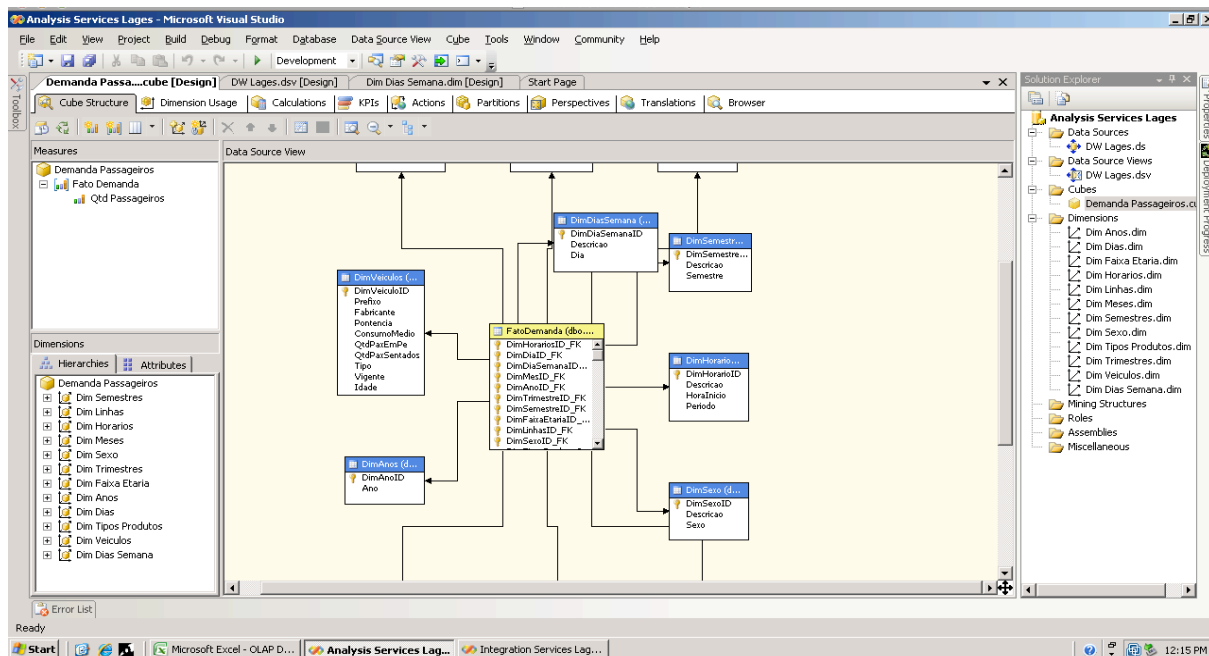


Figura 26 - Interface principal do *Analysis Services*.

Fonte: Elaboração do autor, 2008.

A interface da ferramenta OLAP, do Excel, utilizada para visualização dos dados do DW é apresentada na sequência (figura 27).

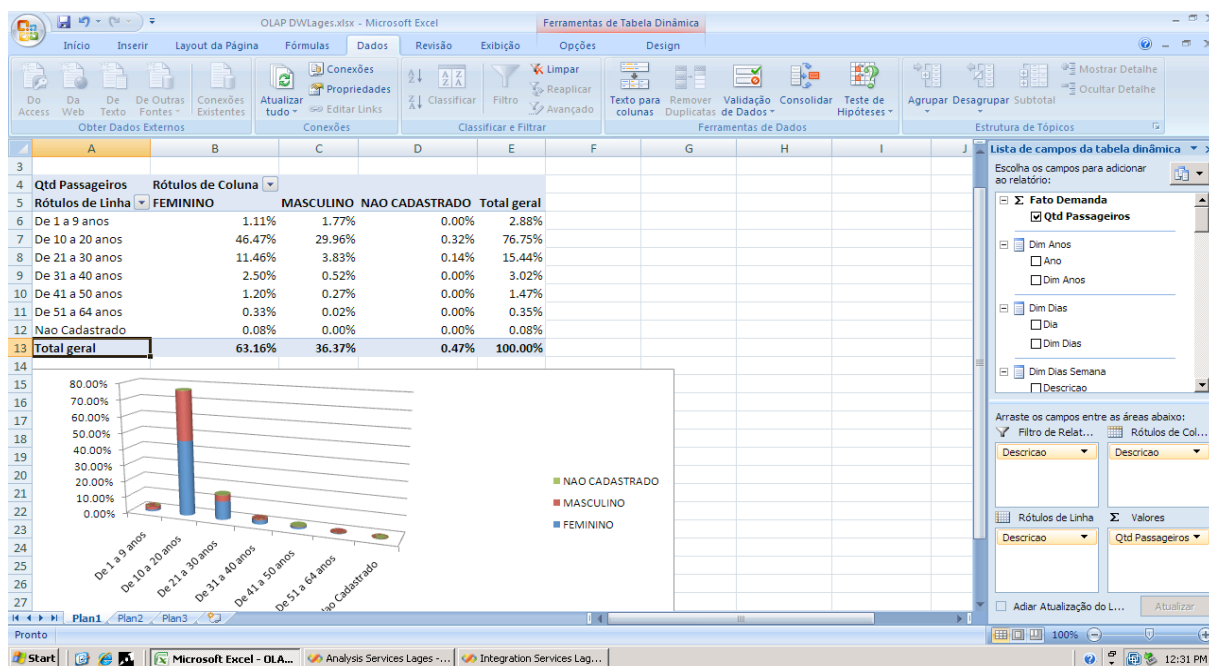


Figura 27 - Interface da ferramenta OLAP Excel 2007.

Fonte: Elaboração do autor, 2008.

4.6 VALIDAÇÃO

Para validar a solução proposta optou-se por realizar um comparativo entre a quantidade de passageiros transportados por mês exibida através da ferramenta OLAP (EXCEL), com os dados obtidos através de uma consulta SQL (*Structured Query Language*) feita diretamente no banco de dados operacional.

A figura 28 apresenta os resultados obtidos com a consulta da quantidade de passageiros transportados por mês, realizada através da ferramenta OLAP do Excel aos dados do DW.

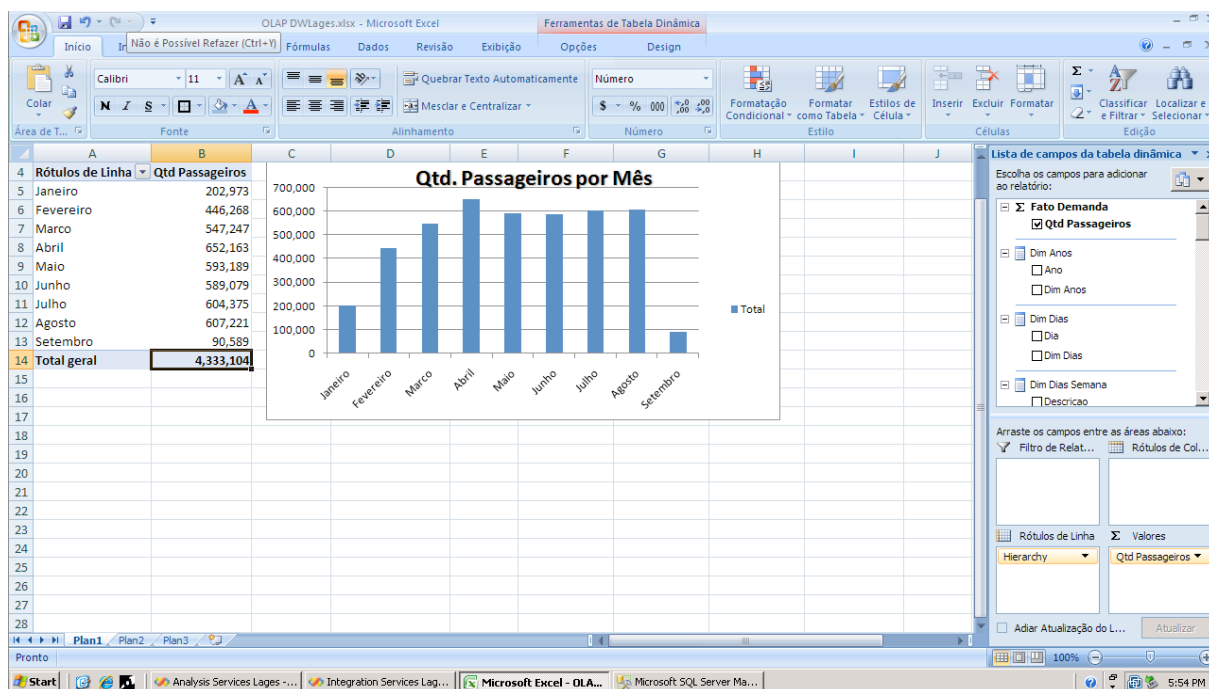


Figura 28 - Quantidade passageiros por mês do DW, via OLAP.

Fonte: Elaboração do autor, 2008.

A quantidade de passageiros por mês, obtida com a consulta SQL ao banco de dados operacional da empresa, é apresentada na figura 29.

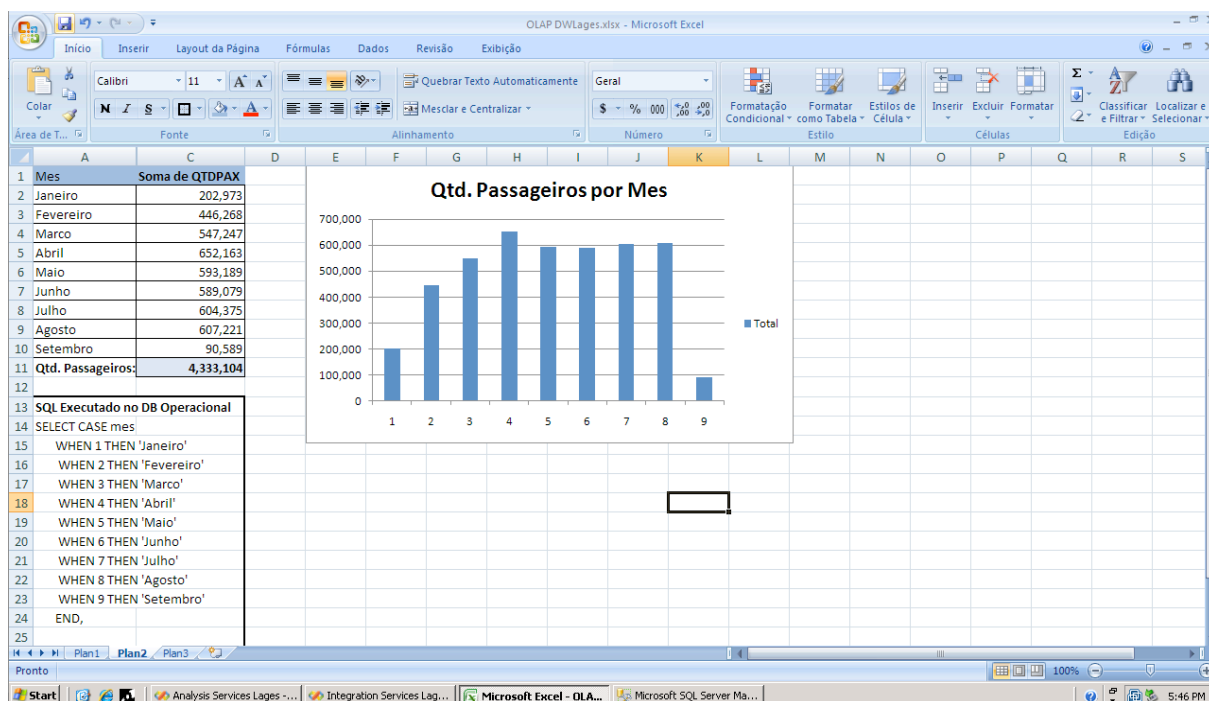


Figura 29 - Quantidade de passageiros por mês do banco de dados operacional (OLTP).
Fonte: Elaboração do autor, 2008.

Como ambas as consultas apresentaram o mesmo número de passageiros por mês, comprovou-se a integridade dos dados do protótipo desenvolvido.

4.7 RESULTADOS E CONCLUSÕES

Com base no levantamento de requisitos descrito anteriormente no capítulo 3, são apresentados, a seguir, alguns exemplos de como o protótipo poderá servir de apoio ao processo decisório dos gestores de empresas de transporte coletivo.

Afim de preservar os dados operacionais da empresa de transporte coletivo, fonte dos dados utilizados no protótipo desenvolvido, são apresentados aqui somente índices referentes ao total de passageiros transportados no período de 1 de Janeiro a 4 de Setembro de 2008.

4.7.1 Índice de passageiros transportados por tipo de pagamento

Este primeiro exemplo apresenta o percentual de passageiros transportados por tipo de forma de pagamento. Observa-se que apenas 36% do usuários do transporte coletivo utilizam o cartão eletrônico como forma de pagamento.

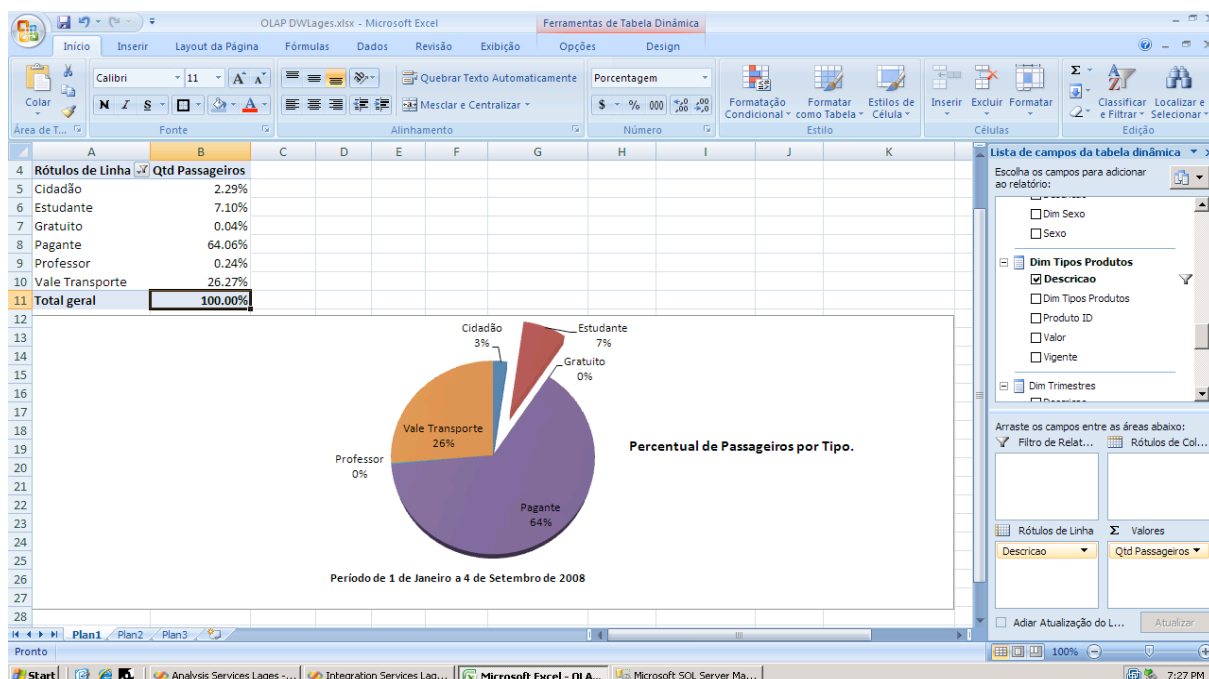


Figura 30 - Índice de passageiros transportados por tipo de pagamento.

Fonte: Elaboração do autor, 2008.

4.7.2 Índice de passageiros transportados por faixa de horário

Neste exemplo são exibidos os percentuais de passageiros transportados por faixa de horário. Observa-se que os maiores volumes de passageiros encontra-se nos horários entre sete e oito horas da manhã, das onze às doze horas da manhã e das seis às sete horas da tarde. Na figura 31, abaixo, são apresentados estes dados e o gráfico obtido.

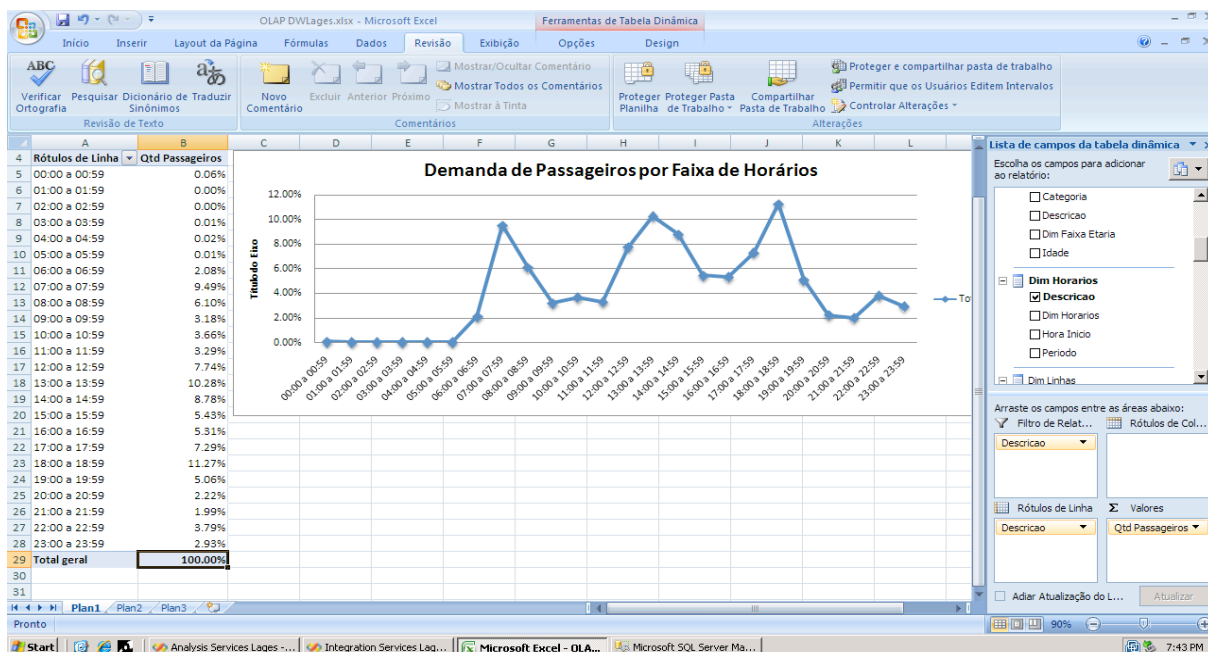


Figura 31 - Índice de passageiros transportados por faixa de horários.
Fonte: Elaboração do autor, 2008.

4.7.3 Índice de passageiros transportados por linha

Na consulta ao número de passageiros transportados por linha observa-se que as linhas “PARQUE - CONTE”, “PARQUE – JUDAS”, “VILA – MARIZA” e “JARDIM - CEPAR” possuem um maior percentual de passageiros transportados (figura 32).

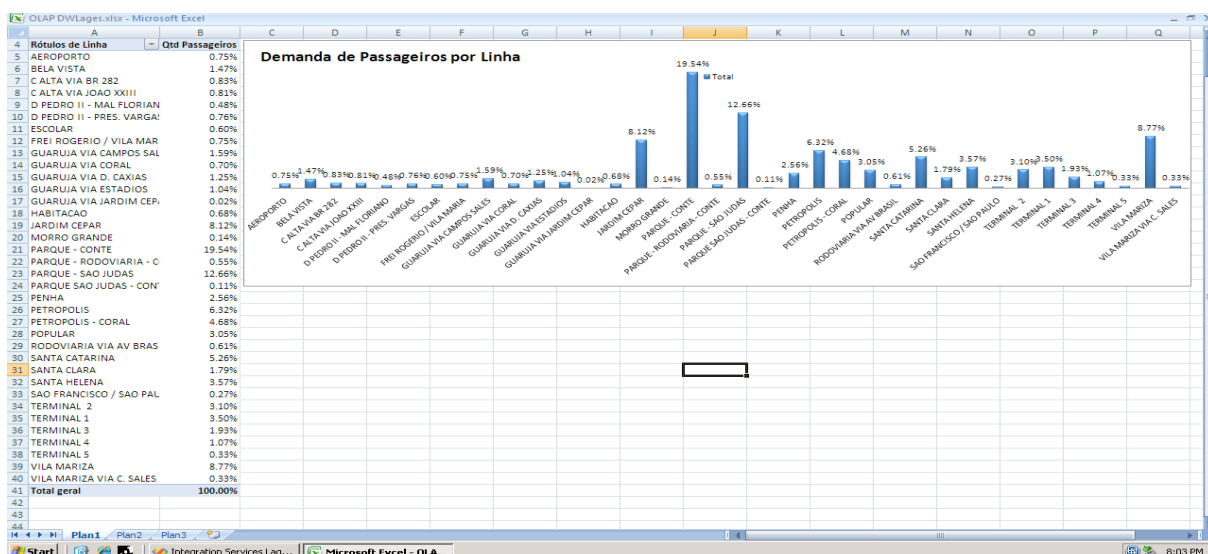


Figura 32 - Índice de passageiros por linha.
Fonte: Elaboração do autor, 2008.

4.7.4 Índice de Passageiros transportados por mês

Neste exemplo é possível observar uma queda no índice de passageiros transportados nos meses de janeiro, fevereiro e março se comparados com os demais meses. A figura 33 ilustra, de forma gráfica, estes índices.

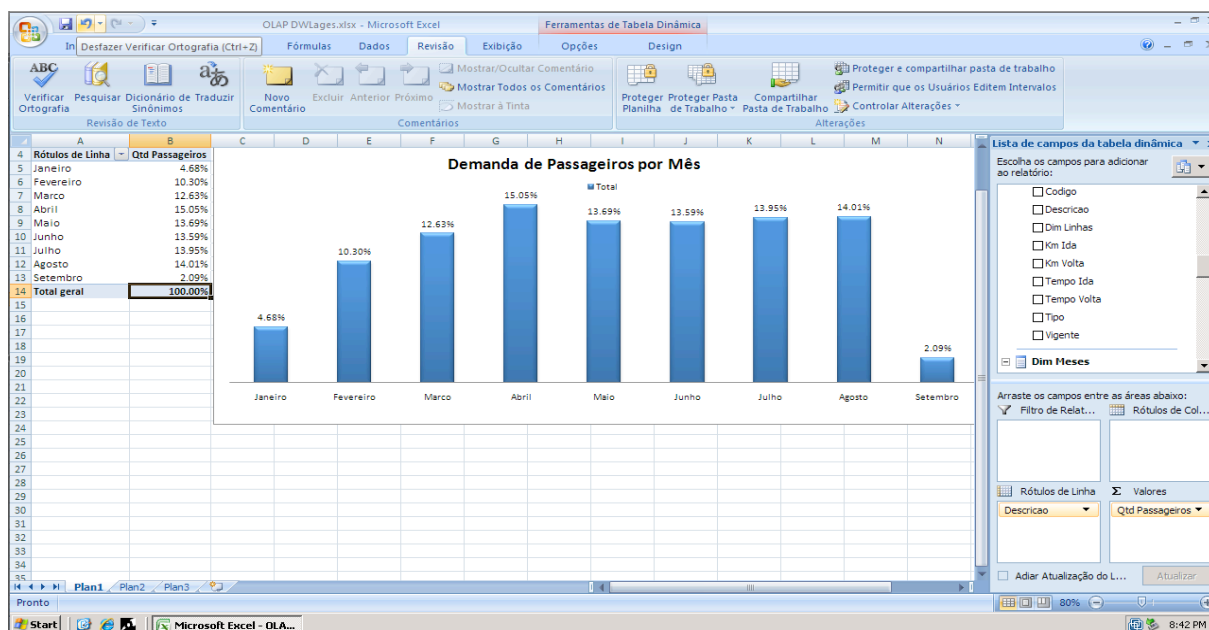


Figura 33 - Índice de passageiros transportados por mês.

Fonte: Elaboração do autor, 2008.

4.7.5 Índice de estudantes transportados por faixa etária

Observa-se, na ilustração 34 a seguir, que 76% dos estudantes transportados tem entre dez e vinte anos de idade.

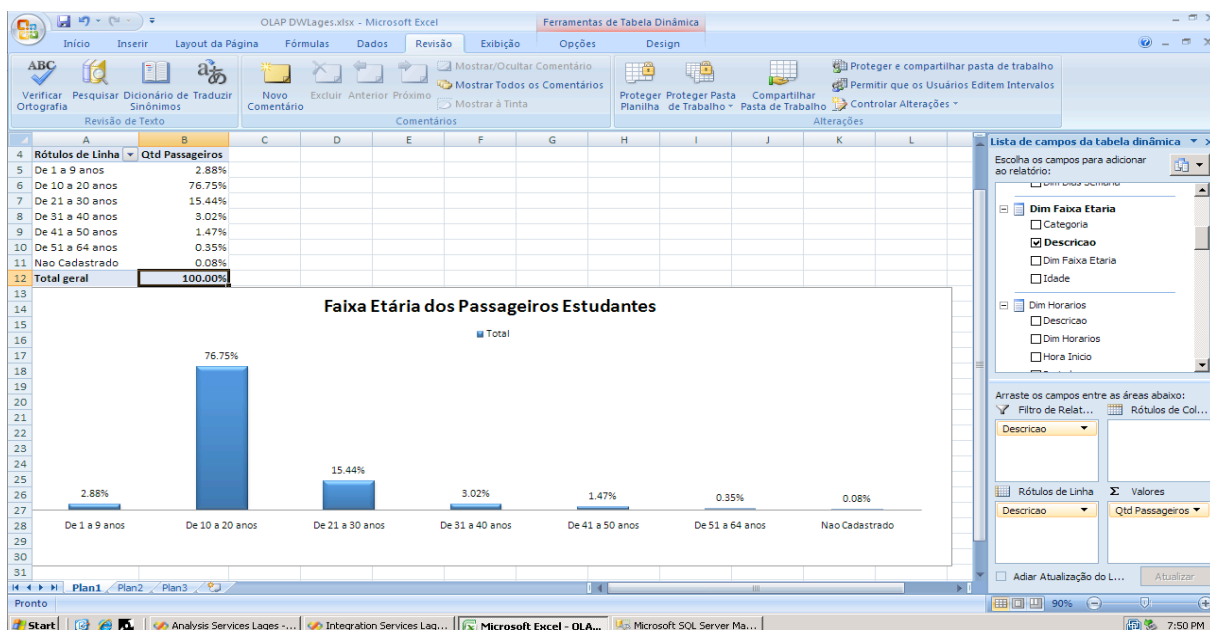


Figura 34 - Índice de estudantes transportados por faixa etária.
Fonte: Elaboração do autor, 2008.

4.7.6 Índice de passageiros estudantes transportados por sexo

Neste exemplo observa-se um maior percentual de mulheres entre os passageiros estudantes transportados. A figura 35 ilustra estes resultados de forma gráfica.

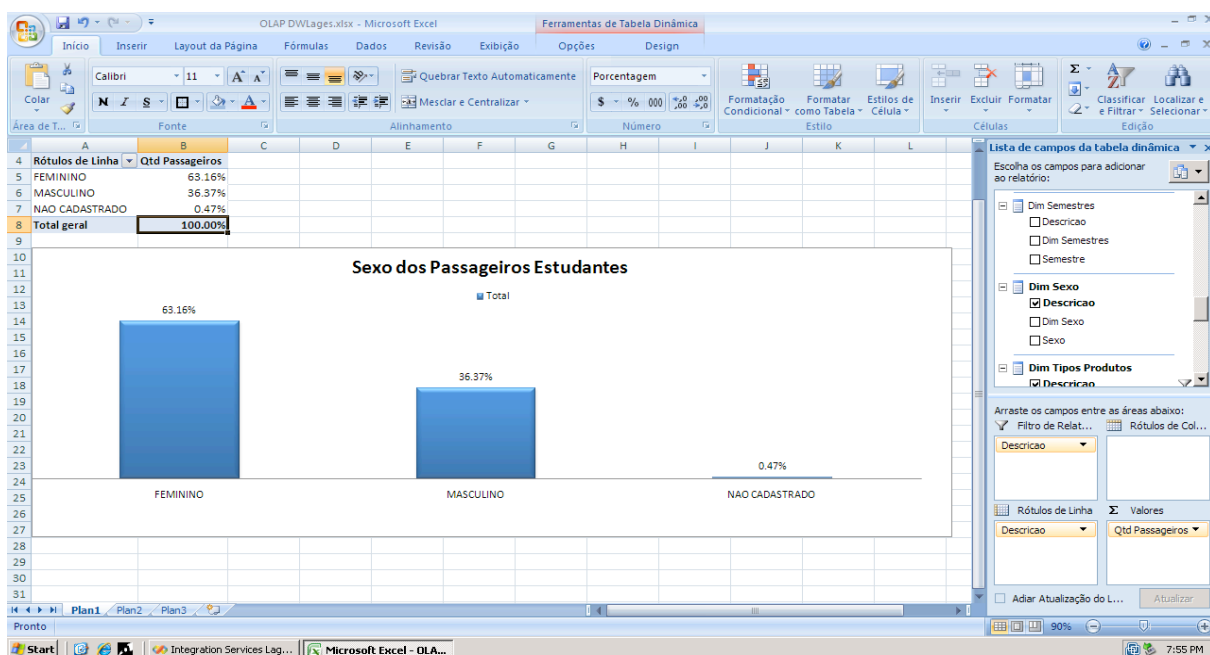


Figura 35 - Índice de passageiros estudantes transportados por sexo.
Fonte: Elaboração do autor, 2008.

4.7.7 Índice de passageiros que utilizam o vale-transporte por sexo

Aqui pode-se observar que é maior o percentual de mulheres do que homens utilizando o vale-transporte para se deslocar até seu local de trabalho. A seguir a figura 36 apresenta esses resultados.

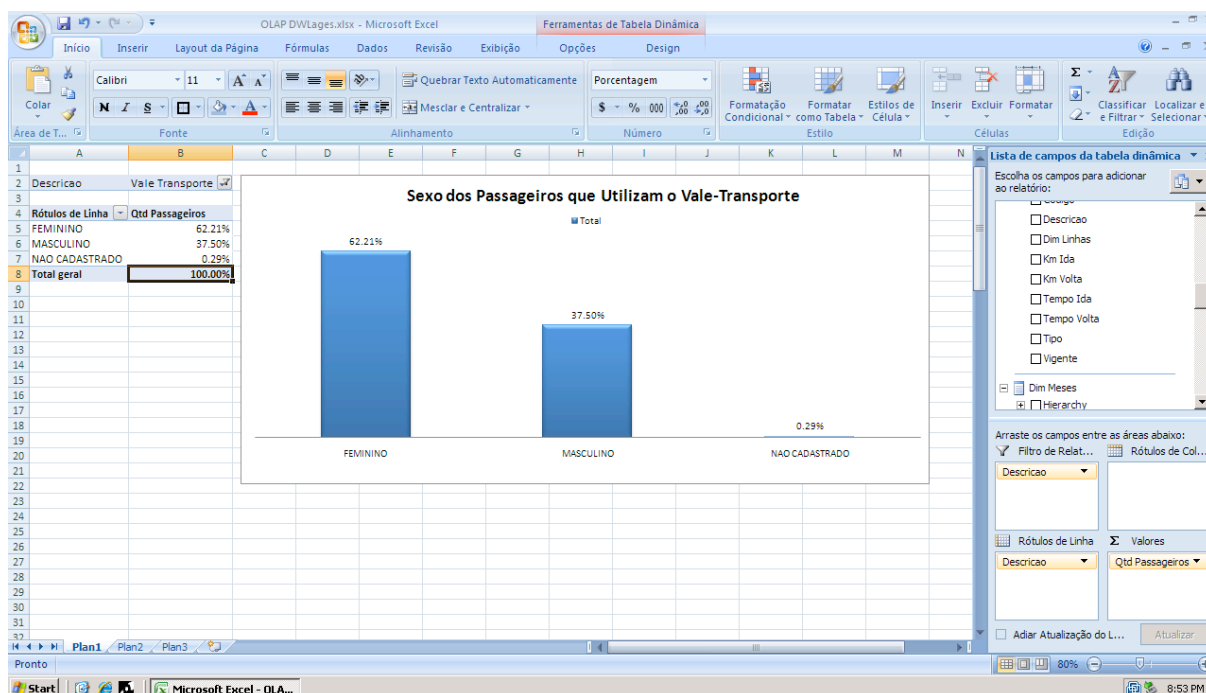


Figura 36 - Índice de passageiros que utilizam o vale-transporte por sexo.
Fonte: Elaboração do autor, 2008.

4.7.8 Índice de passageiros usuários do vale-transporte por linha

Observa-se no gráfico a seguir (figura 37), que o maior volume de passageiros usuários do vale-transporte ocorre na linha “TERMINAL1”, ou seja, no terminal localizado no centro da cidade de Lages/SC.

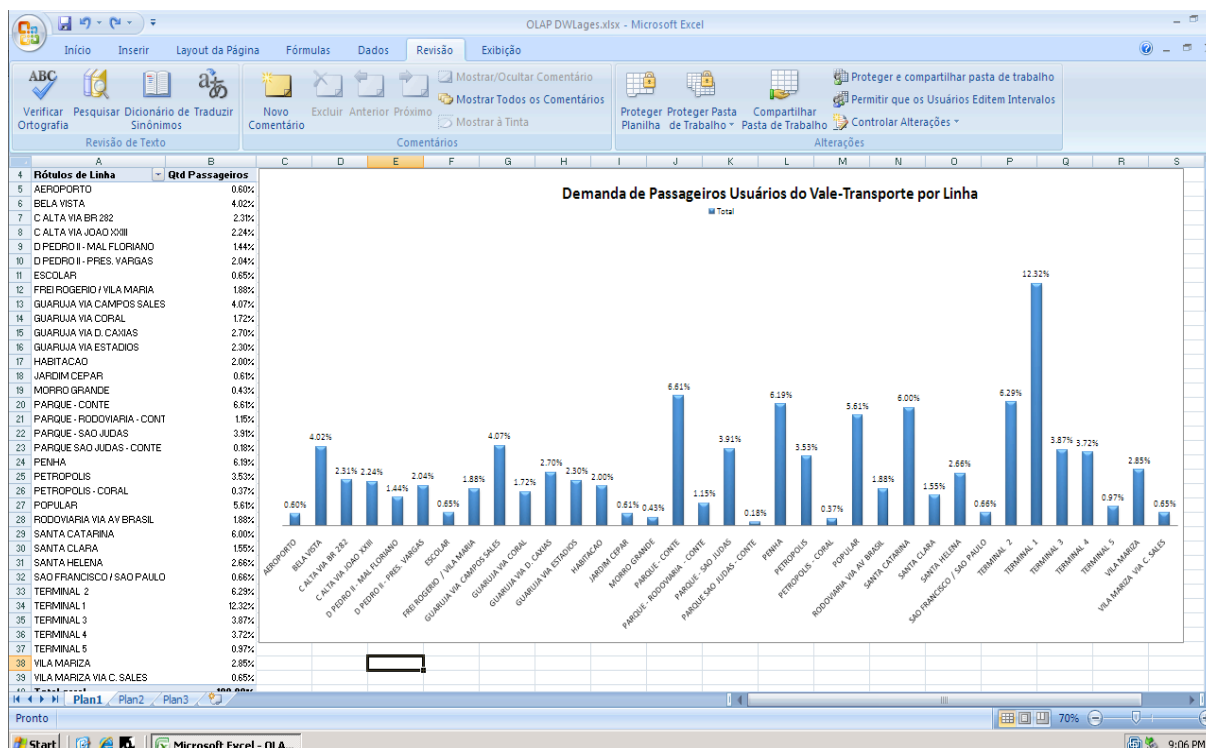


Figura 37 - Índice de passageiros que utilizam o vale-transporte por linha.
Fonte: Elaboração do autor, 2008.

4.8 CONSIDERAÇÕES FINAIS

Para demonstrar a aplicabilidade prática do modelo proposto foram selecionados alguns exemplos que comprovam de forma inequívoca, como uma ferramenta específica para *Data Warehouse* pode auxiliar os usuários no processo de tomada de decisão. Deve-se destacar que a utilização da ferramenta de consulta OLAP do Excel 2007 foi de fácil assimilação pelo usuário final, permitindo-o extrair informações para o processo de tomada de decisão apenas com um rápido treinamento.

As ferramentas *SQL Server Manager*, *Analysis Services*, *Integration Service* e Excel 2007, aplicadas na fase de implementação, são consagradas no mercado de BI e têm grande aceitação em nível mundial. Sua utilização facilitou de sobremaneira o desenvolvimento deste protótipo.

Muitos outros exemplos poderiam ter sido desenvolvidos, mas os escolhidos demonstram, por si só, e de forma fiel os fatos ocorridos na empresa de transporte coletivo Transul localizada na cidade de Lages/SC.

Os dados utilizados para validação do protótipo foram extraídos dos sistemas operacionais da empresa, principalmente do sistema de bilhetagem eletrônica, popularmente conhecido como “catraca eletrônica”.

5 CONCLUSÕES E TRABALHOS FUTUROS

Apresentam-se neste capítulo as conclusões obtidas no decorrer da implementação deste projeto, seguidas de algumas sugestões de trabalhos futuros que possam complementar este trabalho.

5.1 CONCLUSÕES GERAIS

Embora os exemplos elencados neste projeto sejam uma pequena amostra das possibilidades, eles demonstram de forma inequívoca o enorme potencial que as ferramentas de BI têm para auxiliar os gestores das empresas de transporte coletivo. Com o uso deste instrumento ficou evidente, por exemplo, a maior presença feminina entre os passageiros dos ônibus. Isto poderia suscitar aos empresários deste setor mudanças visando melhor atender a este público específico. Foi possível, também, observar que determinadas linhas têm maior demanda de passageiros se filtrados por tipo de pagamento. Os gerentes de empresas de ônibus poderiam, por exemplo, alocar veículos mais adequados a cada perfil de usuário. Também se observou um baixo índice de passageiros utilizando o cartão eletrônico como forma de pagamento. Considerando a grande vantagem da utilização do cartão, este indicador poderia, por exemplo, advertir à empresa da necessidade de maior investimento em publicidade.

O estudo de caso realizado comprovou a aplicabilidade de uma arquitetura de BI em uma empresa de transporte coletivo, fornecendo um completo ambiente de apoio à tomada de decisão aos gestores.

5.2 TRABALHOS FUTUROS

Devido ao limitado tempo para o desenvolvimento deste trabalho, não foi possível explorar todas as ferramentas que compõem uma arquitetura de BI, assim como a inclusão de outras medidas e tabelas fato. Por isso, sugere-se que sejam aplicadas técnicas e ferramentas de *Data Mining*, e a criação de outras medidas e tabelas fato em projetos futuros afim de complementar este trabalho.

REFERÊNCIAS

ASSOCIAÇÃO NACIONAL DAS EMPRESAS DE TRANSPORTE URBANO. **Pesquisa: Mobilidade da População Urbana**. Disponível em: <<http://www.ntu.org.br/novosite/arquivos/RelatorioMobilidade2006.pdf>> Acesso em: 20 jun. 2008.

AZAMBUJA, Ana Maria Volkmer. **Análise de eficiência na gestão do transporte urbano por ônibus em municípios brasileiros**. 2002. 385f. Tese (Doutorado em Engenharia de Produção) – Curso de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

BARBIERI, Carlos. **BI – Bussiness intelligence**. Rio de Janeiro: Axcel, 2001.

BISPO, Carlos. **Uma análise da nova geração de sistemas de apoio à decisão**. 1998. 174f. Dissertação (Mestrado em Engenharia de Produção) – Curso de Pós-graduação em Engenharia de Produção, Universidade Federal de São Paulo, São Carlos, 1998.

CLEMES, Márcio. **Data warehouse como suporte ao sistema de informações gerenciais em uma instituição de ensino superior: estudo de caso na UFSC**. 2001. 118. Dissertação (Mestrado em Engenharia de Produção) – Curso de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2001.

COREY, M.; ABBEY, M.; TAUB, B. **Oracle 8i data warehouse**. Rio de Janeiro: Campus, 2001.

DWBRASIL. **Business intelligence**. Disponível em: < <http://www.dwbrasil.com.br>>. Acesso em: 2 out. 2007.

GUEDES, Gilleanes. **UML Uma abordagem prática**. São Paulo: Novatec, 2004.

INMON, W.H., **Como construir o data warehouse**. 4. ed. Rio de Janeiro: Campus, 1997

LAGO, Sandra Mara Stocker. **Uma metodologia para avaliação da necessidade e da viabilidade de implantação de ferramentas de business intelligence**. 2002. 92f. Dissertação (Mestrado em Engenharia de Produção) – Curso de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

LAUDON, K. C.; LAUDON, J. P. **Gerenciamento de sistemas de informação**. Rio de Janeiro: LTC, 2001.

MACHADO, Felipe. **Projeto de data warehouse**. São Paulo: Érica, 2000.

MARCORATTI, José. **A Gestão de Requisitos**. Disponível em <http://imasters.uol.com.br/artigo/3860/des_de_software/a_gestao_de_requisitos/>. Acesso em: 08 out. 2008a.

_____. **Modelando sistemas em UML - Casos de uso**. Disponível em <http://imasters.uol.com.br/artigo/2753/uml/modelando_sistemas_em_uml_-_casos_de_uso/>. Acesso em: 13 out. 2008b.

MORALES, Aran. **Data warehouse: a memória da organização**. Disponível em <http://www2.stela.ufsc.br/aran/sad/sad_aula2.htm>. Acesso em: 20 jun. 2008.

SERRA, Laércio. **A essência do business intelligence**. 1. ed. São Paulo: Berkeley, 2002

SILVA, Edna; MENEZES, Estera. **Metodologia da pesquisa e elaboração de dissertação**. 4 ed. Disponível em <<http://inf.unisul.br/%7Eines/pccsi/metPesq27jun05.zip>>. Acesso em: 20 jun. 2008.

SILVA, Márcio. **A importância do levantamento de requisitos no sucesso dos projetos de software**. Disponível em <<http://www.linhadecodigo.com.br/ArtigoImpressao.aspx?id=1685>>. Acesso em: 08 out. 2008.

STAIR, Ralph; REYNOLDS, George. **Princípios de sistemas de informação**. Rio de Janeiro: LTC, 2002.

SOUZA, Michel. **BI - O hoje e o amanhã da inteligência organizacional**. São Paulo, nov. 2003. Disponível em <<http://www.imasters.com.br>> Acesso em: 02 de out. 2007.

_____. **Ferramentas olap**. Disponível em <http://imasters.uol.com.br/artigo/1498/gerencia/ferramentas_olap> Acesso em: 25 de jun. 2008.

O'BRIEN, James. **Sistemas de informações e as decisões gerenciais na era da Internet**. São Paulo: Saraiva, 2003.

APÊNDICES

APÊNDICE A - Código DDL para criação física do banco de dados gerado com ferramenta SQL Manager

```

USE [DWLages]
GO
/***** Object:  Schema [DW]      Script Date: 11/07/2008 00:06:33 *****/
CREATE SCHEMA [DW] AUTHORIZATION [dbo]
GO
/***** Object:  Table [dbo].[DimSemestres]      Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimSemestres] (
    [DimSemestreID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](20) NOT NULL,
    [Semestre] [int] NOT NULL,
    CONSTRAINT [PK_DimSemestre] PRIMARY KEY CLUSTERED
(
    [DimSemestreID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object:  Table [dbo].[DimTrimestres]      Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimTrimestres] (
    [DimTrimestreID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](20) NOT NULL,
    [Trimestre] [int] NOT NULL,
    CONSTRAINT [PK_DimTrimestre] PRIMARY KEY CLUSTERED
(
    [DimTrimestreID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object:  Table [dbo].[DimVeiculos]      Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON

```

```

GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimVeiculos] (
    [DimVeiculoID] [int] IDENTITY(1,1) NOT NULL,
    [Prefixo] [int] NOT NULL,
    [Fabricante] [varchar](50) NULL,
    [Pontencia] [int] NULL,
    [ConsumoMedio] [float] NULL,
    [QtdPaxEmPe] [int] NULL,
    [QtdPaxSentados] [int] NULL,
    [Tipo] [varchar](50) NULL,
    [Vigente] [varchar](10) NULL,
    [Idade] [int] NULL,
    CONSTRAINT [PK_DimVeiculos] PRIMARY KEY CLUSTERED
(
    [DimVeiculoID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimFeriados]      Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimFeriados] (
    [DimFeriadoID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](50) NOT NULL,
    [Data] [datetime] NOT NULL,
    CONSTRAINT [PK_DimFeriados] PRIMARY KEY CLUSTERED
(
    [DimFeriadoID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimHorarios]      Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimHorarios] (
    [DimHorarioID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](50) NOT NULL,
    [HoraInicio] [datetime] NULL,
    [Periodo] [varchar](50) NULL,
    CONSTRAINT [PK_DimHorarios] PRIMARY KEY CLUSTERED
(
    [DimHorarioID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =

```



```

OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimDiasSemana]      Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimDiasSemana] (
    [DimDiasSemanaID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](50) NOT NULL,
    [Dia] [int] NOT NULL,
    CONSTRAINT [PK_DimDiasSemana] PRIMARY KEY CLUSTERED
(
    [DimDiasSemanaID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimTiposProdutos]    Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimTiposProdutos] (
    [DimTiposProdutosID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](50) NOT NULL,
    [Valor] [decimal](18, 4) NULL,
    [Vigente] [varchar](50) NULL,
    [ProdutoID] [int] NOT NULL,
    CONSTRAINT [PK_DimTiposProdutos] PRIMARY KEY CLUSTERED
(
    [DimTiposProdutosID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimLinhas]          Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimLinhas] (
    [DimLinhasID] [int] IDENTITY(1,1) NOT NULL,
    [Codigo] [varchar](10) NULL,
    [Descricao] [varchar](50) NULL,

```

```

        [KmIda] [float] NULL,
        [KmVolta] [float] NULL,
        [TempoIda] [int] NULL,
        [TempoVolta] [int] NULL,
        [Tipo] [varchar](50) NULL,
        [Vigente] [varchar](10) NULL,
    CONSTRAINT [PK_DimLinhas] PRIMARY KEY CLUSTERED
(
    [DimLinhasID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimFaixaEtaria]      Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimFaixaEtaria] (
    [DimFaixaEtariaID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](50) NOT NULL,
    [Idade] [int] NULL,
    [Categoria] [varchar](50) NULL,
    CONSTRAINT [PK_DimFaixaEtaria] PRIMARY KEY CLUSTERED
(
    [DimFaixaEtariaID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimDias]      Script Date: 11/07/2008 00:06:33
*****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[DimDias] (
    [DimDiaID] [int] IDENTITY(1,1) NOT NULL,
    [Dia] [int] NOT NULL,
    CONSTRAINT [PK_DimDias] PRIMARY KEY CLUSTERED
(
    [DimDiaID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table [dbo].[DimSexo]      Script Date: 11/07/2008 00:06:33
*****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO

```

```

CREATE TABLE [dbo].[DimSexo] (
    [DimSexoID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](50) NOT NULL,
    [Sexo] [char](1) NULL,
    CONSTRAINT [PK_DimSexo] PRIMARY KEY CLUSTERED
(
    [DimSexoID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimMeses]      Script Date: 11/07/2008 00:06:33
*****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[DimMeses] (
    [DimMesID] [int] IDENTITY(1,1) NOT NULL,
    [Descricao] [varchar](20) NOT NULL,
    [Mes] [int] NOT NULL,
    CONSTRAINT [PK_DimMeses] PRIMARY KEY CLUSTERED
(
    [DimMesID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
SET ANSI_PADDING OFF
GO
/***** Object: Table [dbo].[DimAnos]      Script Date: 11/07/2008 00:06:33
*****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[DimAnos] (
    [DimAnoID] [int] IDENTITY(1,1) NOT NULL,
    [Ano] [int] NOT NULL,
    CONSTRAINT [PK_DimAnos] PRIMARY KEY CLUSTERED
(
    [DimAnoID] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table [dbo].[FatoDemanda]    Script Date: 11/07/2008
00:06:33 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[FatoDemanda] (
    [DimHorariosID_FK] [int] NOT NULL,
    [DimDiaID_FK] [int] NOT NULL,
    [DimDiaSemanaID_FK] [int] NOT NULL,
    [DimMesID_FK] [int] NOT NULL,

```

```

[DimAnoID_FK] [int] NOT NULL,
[DimTrimestreID_FK] [int] NOT NULL,
[DimSemestreID_FK] [int] NOT NULL,
[DimFaixaEtariaID_FK] [int] NOT NULL,
[DimLinhasID_FK] [int] NOT NULL,
[DimSexoID_FK] [int] NOT NULL,
[DimTiposProdutosID_PK] [int] NOT NULL,
[DimVeiculosID_FK] [int] NOT NULL,
[QtdPassageiros] [int] NOT NULL,
CONSTRAINT [PK_FatoDemanda_1] PRIMARY KEY CLUSTERED
(
    [DimHorariosID_FK] ASC,
    [DimDiaID_FK] ASC,
    [DimDiaSemanaID_FK] ASC,
    [DimMesID_FK] ASC,
    [DimAnoID_FK] ASC,
    [DimTrimestreID_FK] ASC,
    [DimSemestreID_FK] ASC,
    [DimFaixaEtariaID_FK] ASC,
    [DimLinhasID_FK] ASC,
    [DimSexoID_FK] ASC,
    [DimTiposProdutosID_PK] ASC,
    [DimVeiculosID_FK] ASC
) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY =
OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimAnos]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimAnos] FOREIGN KEY([DimAnoID_FK])
REFERENCES [dbo].[DimAnos] ([DimAnoID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT [FK_FatoDemanda_DimAnos]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimDias]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimDias] FOREIGN KEY([DimDiaID_FK])
REFERENCES [dbo].[DimDias] ([DimDiaID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT [FK_FatoDemanda_DimDias]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimDiasSemana]  Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimDiasSemana] FOREIGN KEY([DimDiaSemanaID_FK])
REFERENCES [dbo].[DimDiasSemana] ([DimDiaSemanaID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT
[FK_FatoDemanda_DimDiasSemana]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimFaixaEtaria]  Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimFaixaEtaria] FOREIGN KEY([DimFaixaEtariaID_FK])
REFERENCES [dbo].[DimFaixaEtaria] ([DimFaixaEtariaID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT
[FK_FatoDemanda_DimFaixaEtaria]
GO

```

```

/***** Object: ForeignKey [FK_FatoDemanda_DimHorarios]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimHorarios] FOREIGN KEY([DimHorariosID_FK])
REFERENCES [dbo].[DimHorarios] ([DimHorarioID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT
[FK_FatoDemanda_DimHorarios]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimLinhas]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimLinhas] FOREIGN KEY([DimLinhasID_FK])
REFERENCES [dbo].[DimLinhas] ([DimLinhasID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT [FK_FatoDemanda_DimLinhas]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimMeses]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimMeses] FOREIGN KEY([DimMesID_FK])
REFERENCES [dbo].[DimMeses] ([DimMesID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT [FK_FatoDemanda_DimMeses]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimSemestres]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimSemestres] FOREIGN KEY([DimSemestreID_FK])
REFERENCES [dbo].[DimSemestres] ([DimSemestreID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT
[FK_FatoDemanda_DimSemestres]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimSexo]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimSexo] FOREIGN KEY([DimSexoID_FK])
REFERENCES [dbo].[DimSexo] ([DimSexoID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT [FK_FatoDemanda_DimSexo]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimTiposProdutos]      Script
Date: 11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimTiposProdutos] FOREIGN KEY([DimTiposProdutosID_PK])
REFERENCES [dbo].[DimTiposProdutos] ([DimTiposProdutosID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT
[FK_FatoDemanda_DimTiposProdutos]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimTrimestres]      Script Date:
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimTrimestres] FOREIGN KEY([DimTrimestreID_FK])
REFERENCES [dbo].[DimTrimestres] ([DimTrimestreID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT
[FK_FatoDemanda_DimTrimestres]
GO
/***** Object: ForeignKey [FK_FatoDemanda_DimVeiculos]      Script Date:

```

```
11/07/2008 00:06:33 *****/
ALTER TABLE [dbo].[FatoDemanda] WITH CHECK ADD CONSTRAINT
[FK_FatoDemanda_DimVeiculos] FOREIGN KEY([DimVeiculosID_FK])
REFERENCES [dbo].[DimVeiculos] ([DimVeiculoID])
GO
ALTER TABLE [dbo].[FatoDemanda] CHECK CONSTRAINT
[FK_FatoDemanda_DimVeiculos]
GO
```

APÊNDICE B - Código SQL da ViewQtdPaxPagantesCartao do sistema operacional

```

USE [TDMAXLages]
GO
/***** Object:  View [dbo].[ViewQtdPaxPagantesCartao]      Script Date:
11/07/2008 00:15:22 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER VIEW [dbo].[ViewQtdPaxPagantesCartao]
AS
SELECT      TOP (100) PERCENT CAST(CAST(DATEPART(HOUR, dbo.Viagens.DataFim)
AS VARCHAR) + ':' + CAST(00 AS VARCHAR) AS DATETIME) AS Hora,
            DATEPART(DAY, dbo.Viagens.DataFim) AS Dia,
            DATEPART(WeekDay, dbo.Viagens.DataFim) AS DiaSemana, DATEPART(MONTH,
dbo.Viagens.DataFim)
            AS Mes, dbo.Feriados.Descricao AS Feriado, CASE WHEN
DATEPART(Month, VIAGENS.DATAFIM) <= 6 THEN 1 ELSE 2 END AS Semestre,
            CASE WHEN DATEPART(Month, VIAGENS.DATAFIM) <= 3 THEN
1 WHEN DATEPART(Month, VIAGENS.DATAFIM) >= 4 AND DATEPART(Month,
VIAGENS.DATAFIM) <= 6 THEN 2 WHEN DATEPART(Month,
VIAGENS.DATAFIM) >= 7 AND DATEPART(Month, VIAGENS.DATAFIM)
<= 9 THEN 3 WHEN DATEPART(Month, VIAGENS.DATAFIM) >=
10 AND DATEPART(Month, VIAGENS.DATAFIM) <= 12 THEN 4 END AS Trimestre,
            DATEPART(YEAR, dbo.Viagens.DataFim) AS Ano,
            dbo.CartoesTransacoes.ProdutoID, dbo.Linhas.Codigo AS CodigoLinha,
            dbo.Turnos.Prefixo AS PrefixoVeiculo, CASE WHEN
ViewComplementoCadastro.Sexo = 'F' OR
            ViewComplementoCadastro.Sexo = 'M' THEN
ViewComplementoCadastro.Sexo ELSE NULL END AS Sexo,
            dbo.ViewComplementoCadastro.Idade,
            dbo.ViewComplementoCadastro.Naturalidade, COUNT(*) AS
QTDPAx
FROM          dbo.CartoesTransacoes INNER JOIN
            dbo.Cadastrados ON dbo.Cadastrados.CadastroID =
            dbo.CartoesTransacoes.CadastroID INNER JOIN
            dbo.CartoesAcessos ON dbo.CartoesAcessos.NumSerie =
            dbo.CartoesTransacoes.NumSerie AND
            dbo.CartoesAcessos.Transacao =
            dbo.CartoesTransacoes.Transacao INNER JOIN
            dbo.Turnos ON dbo.Turnos.TurnoID =
            dbo.CartoesAcessos.TurnoID INNER JOIN
            dbo.Viagens ON dbo.Viagens.ViagemID =
            dbo.CartoesAcessos.ViagemID AND dbo.Viagens.TurnoID =
            dbo.CartoesAcessos.TurnoID INNER JOIN
            dbo.Linhas ON dbo.Linhas.LinhaID = dbo.Turnos.LinhaID
LEFT OUTER JOIN
            dbo.ViewComplementoCadastro ON
            dbo.ViewComplementoCadastro.CadastroID = dbo.Cadastrados.CadastroID LEFT
OUTER JOIN
            dbo.Feriados ON DATEPART(DAY, dbo.Viagens.DataFim) =
            dbo.Feriados.Dia AND DATEPART(MONTH, dbo.Viagens.DataFim) =
            dbo.Feriados.Mes AND
            DATEPART(YEAR, dbo.Viagens.DataFim) =
            dbo.Feriados.Ano
WHERE          (dbo.CartoesTransacoes.EventoID = 2) OR
            (dbo.CartoesTransacoes.EventoID = 27) OR

```

```

(dbo.CartoesTransacoes.EventoID = 20)
GROUP BY CAST(CAST(DATEPART(HOUR, dbo.Viagens.DataFim) AS VARCHAR) + ':' +
CAST(00 AS VARCHAR) AS DATETIME), DATEPART(DAY,
dbo.Viagens.DataFim), DATEPART(WeekDay,
dbo.Viagens.DataFim), DATEPART(MONTH, dbo.Viagens.DataFim),
dbo.Ferriados.Descricao,
CASE WHEN DATEPART(Month, VIAGENS.DATAFIM) <= 6 THEN
1 ELSE 2 END, CASE WHEN DATEPART(Month, VIAGENS.DATAFIM)
<= 3 THEN 1 WHEN DATEPART(Month, VIAGENS.DATAFIM) >=
4 AND DATEPART(Month, VIAGENS.DATAFIM) <= 6 THEN 2 WHEN DATEPART(Month,
VIAGENS.DATAFIM) >= 7 AND DATEPART(Month,
VIAGENS.DATAFIM) <= 9 THEN 3 WHEN DATEPART(Month, VIAGENS.DATAFIM) >= 10
AND
DATEPART(Month, VIAGENS.DATAFIM) <= 12 THEN 4 END,
DATEPART(YEAR, dbo.Viagens.DataFim), dbo.CartoesTransacoes.ProdutoID,
dbo.Linhas.Codigo, dbo.Turnos.Prefixo, CASE WHEN
ViewComplementoCadastro.Sexo = 'F' OR
ViewComplementoCadastro.Sexo = 'M' THEN
ViewComplementoCadastro.Sexo ELSE NULL END,
dbo.ViewComplementoCadastro.Idade,
dbo.ViewComplementoCadastro.Naturalidade
ORDER BY Hora, Dia, DiaSemana

```


APÊNDICE C - Código SQL da ViewQtdPaxPagantesdinheiro do sistema operacional

```

USE [TDMAXLages]
GO
/***** Object: View [dbo].[ViewQtdPaxPagantesDinheiro] Script Date:
11/07/2008 00:19:14 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER VIEW [dbo].[ViewQtdPaxPagantesDinheiro]
AS
SELECT TOP (100) PERCENT CAST(CAST(DATEPART(HOUR, dbo.Viagens.DataFim)
AS VARCHAR) + ':' + CAST(00 AS VARCHAR) AS DATETIME) AS Hora,
DATEPART(DAY, dbo.Viagens.DataFim) AS Dia,
DATEPART(WeekDay, dbo.Viagens.DataFim) AS DiaSemana, DATEPART(MONTH,
dbo.Viagens.DataFim)
AS Mes, dbo.Feriados.Descricao AS Feriado, CASE WHEN
DATEPART(Month, VIAGENS.DATAFIM) <= 6 THEN 1 ELSE 2 END AS Semestre,
CASE WHEN DATEPART(Month, VIAGENS.DATAFIM) <= 3 THEN
1 WHEN DATEPART(Month, VIAGENS.DATAFIM) >= 4 AND DATEPART(Month,
VIAGENS.DATAFIM) <= 6 THEN 2 WHEN DATEPART(Month,
VIAGENS.DATAFIM) >= 7 AND DATEPART(Month, VIAGENS.DATAFIM)
<= 9 THEN 3 WHEN DATEPART(Month, VIAGENS.DATAFIM) >=
10 AND DATEPART(Month, VIAGENS.DATAFIM) <= 12 THEN 4 END AS Trimestre,
DATEPART(YEAR, dbo.Viagens.DataFim) AS Ano,
dbo.Viagens_qtd.ProdutoID, dbo.Linhas.Codigo AS CodigoLinha,
dbo.Turnos.Prefixo AS PrefixoVeiculo,
CAST(NULL AS CHAR(1)) AS Sexo, CAST(NULL AS INT) AS
Idade, CAST(NULL AS VARCHAR(50)) AS Naturalidade,
SUM(dbo.Viagens_qtd.Quantidade)
AS QTDPA
FROM dbo.Viagens_qtd INNER JOIN
dbo.Turnos ON dbo.Turnos.TurnoID =
dbo.Viagens_qtd.TurnoID INNER JOIN
dbo.Viagens ON dbo.Viagens.ViagemID =
dbo.Viagens_qtd.ViagemID AND dbo.Viagens.TurnoID = dbo.Viagens_qtd.TurnoID
INNER JOIN
dbo.Linhas ON dbo.Linhas.LinhaID =
dbo.Viagens_qtd.ViagemID LEFT OUTER JOIN
dbo.Feriados ON DATEPART(DAY, dbo.Viagens.DataFim) =
dbo.Feriados.Dia AND DATEPART(MONTH, dbo.Viagens.DataFim) =
dbo.Feriados.Mes AND
DATEPART(YEAR, dbo.Viagens.DataFim) =
dbo.Feriados.Ano
WHERE (dbo.Viagens_qtd.ProdutoID = 3)
GROUP BY CAST(CAST(DATEPART(HOUR, dbo.Viagens.DataFim) AS VARCHAR) + ':' +
CAST(00 AS VARCHAR) AS DATETIME), DATEPART(DAY,
dbo.Viagens.DataFim), DATEPART(WeekDay,
dbo.Viagens.DataFim), DATEPART(MONTH, dbo.Viagens.DataFim),
dbo.Feriados.Descricao,
CASE WHEN DATEPART(Month, VIAGENS.DATAFIM) <= 6 THEN
1 ELSE 2 END, CASE WHEN DATEPART(Month, VIAGENS.DATAFIM)
<= 3 THEN 1 WHEN DATEPART(Month, VIAGENS.DATAFIM) >=
4 AND DATEPART(Month, VIAGENS.DATAFIM) <= 6 THEN 2 WHEN DATEPART(Month,
VIAGENS.DATAFIM) >= 7 AND DATEPART(Month,
VIAGENS.DATAFIM) <= 9 THEN 3 WHEN DATEPART(Month, VIAGENS.DATAFIM) >= 10
AND

```

```
DATEPART(Month, VIAGENS.DATAFIM) <= 12 THEN 4 END,  
DATEPART(YEAR, dbo.Viagens.DataFim), dbo.Viagens_qtd.ProdutoID,  
dbo.Linhas.Codigo,  
dbo.Turnos.Prefixo  
ORDER BY Hora, Dia, DiaSemana
```