

UNIVERSIDADE CANDIDO MENDES - UCAM
MESTRADO EM PESQUISA OPERACIONAL E INTELIGÊNCIA
COMPUTACIONAL

MARCELO TADEU MONTEIRO FREITAS

**DESENVOLVIMENTO DE UM DATA WAREHOUSE ESPACIAL PARA O
SETOR TRIBUTÁRIO MUNICIPAL COM USO DE TECNOLOGIA ABERTA**

CAMPOS DOS GOYTACAZES
2007

UNIVERSIDADE CANDIDO MENDES - UCAM
MESTRADO EM PESQUISA OPERACIONAL E INTELIGÊNCIA
COMPUTACIONAL

Marcelo Tadeu Monteiro Freitas

DESENVOLVIMENTO DE UM DATA WAREHOUSE ESPACIAL PARA O SETOR
TRIBUTÁRIO MUNICIPAL COM USO DE TECNOLOGIA ABERTA

Dissertação apresentada ao PROGRAMA de
Pós-Graduação em Pesquisa Operacional e
Inteligência Computacional, da Universidade
Candido Mendes – Campos/RJ, para a
obtenção do grau de MESTRE EM
PESQUISA OPERACIONAL E
INTELIGÊNCIA COMPUTACIONAL.

Orientadores: Prof. Astério Kiyoshi Tanaka, Ph.D.
Profa. Sahudy Montenegro González, D.Sc.

CAMPOS DOS GOYTACAZES, RJ
2007

MARCELO TADEU MONTEIRO FREITAS

DESENVOLVIMENTO DE UM DATA WAREHOUSE ESPACIAL PARA O SETOR
TRIBUTÁRIO MUNICIPAL COM USO DE TECNOLOGIA ABERTA

Dissertação apresentada ao PROGRAMA de
Pós-Graduação em Pesquisa Operacional e
Inteligência Computacional, da Universidade
Candido Mendes – Campos/RJ, para a
obtenção do grau de MESTRE EM
PESQUISA OPERACIONAL E
INTELIGÊNCIA COMPUTACIONAL.

Aprovada em _____

BANCA EXAMINADORA

Prof. Astério Kiyoshi Tanaka, Ph.D. – Orientador
Universidade Federal do Estado do Rio de Janeiro - UNIRIO

Profª. Sahudy Montenegro González, D.Sc. – Co-Orientadora
Universidade Candido Mendes

Profª. Elzira Lúcia de Oliveira, D.Sc.
Universidade Candido Mendes

Prof. Rogério Atem de Carvalho, D.Sc.
Centro Federal de Educação Tecnológica de Campos (CEFET-Campos)

Campos dos Goytacazes
2007

AGRADECIMENTOS

Agradeço a Deus por ter me concedido a oportunidade de participar de um curso desse nível, pleno de grandes mestres amigos, que não hesitaram em compartilhar seus conhecimentos e direcionar meus estudos.

A minha mãe, Maria da Penha, pelo amor e dedicação permanentes, e aos meus irmãos, Carlos Magno e Maria Inês, pelo apoio e incentivo que sempre me mantiveram motivado para essa jornada.

À minha namorada e companheira Carmélia, pelo carinho, compreensão e apoio na parte tributária sempre que precisei.

Aos meus orientadores, o Prof. Tanaka pela confiança, por acreditar em mim, pela paciência diante das minhas limitações e principalmente por ter compartilhado seus conhecimentos, despertando em mim, através de suas aulas contagiantes, a vontade de me aprofundar na área de banco de dados; e à Profa. Sahudy pela sua cumplicidade, companheirismo e apoio incondicional que foram vitais nos momentos difíceis, além de orientar meus passos.

Aos meus colegas de turma que, com o tempo, passaram a ser amigos inesquecíveis, guardados para sempre na memória e no meu coração. Em especial, aos amigos Celso e Rodrigo, pela amizade marcante.

A todos os profissionais da Prefeitura Municipal de Cachoeiro de Itapemirim, que colaboraram nas entrevistas realizadas durante a etapa de levantamento sempre que foram requisitados e à Empresa DATACI, na figura do Diretor Superintendente Alcione Dias da Silva, pela disponibilização da estrutura quando foi necessário.

“Não há como administrar aquilo que não se pode medir”

Peter Drucker

RESUMO

DESENVOLVIMENTO DE UM DATA WAREHOUSE ESPACIAL PARA O SETOR TRIBUTÁRIO MUNICIPAL COM USO DE TECNOLOGIA ABERTA

A importância dos Sistemas de Suporte à Decisão (SSD) nos tempos atuais é um fato incontestável, principalmente em face da crescente necessidade de ferramentas capazes de auxiliar na análise de grandes quantidades de dados e promover a Inteligência do Negócio. Nesta pesquisa, a base para um SSD foi o desenvolvimento de um Data Warehouse (DW), que é um conjunto de ferramentas capaz de prover consultas voltadas para os processos de negócio de uma forma mais objetiva do que um ambiente de produção, uma vez que é composto por um banco de dados estruturado por assunto, que armazena os dados históricos da organização.

Um ponto a destacar é a utilização de dados geográficos, o que promoveu mais riqueza nas informações armazenadas e processadas. Essa abordagem foi viável graças ao uso de uma ferramenta para On Line Analytical Processing (OLAP), PostGeoOlap, que possibilitou uma funcionalidade adicional: consultas analíticas envolvendo tipos de dados espaciais.

Tanto a ferramenta PostGeoOlap quanto o sistema de gerenciamento de banco de dados (SGBD) utilizado, PostgreSQL, são Softwares Livres (com código fonte aberto e licença livre), tornando a solução ainda mais atraente em virtude do seu menor custo de prototipação.

O objeto deste trabalho é o desenvolvimento de um protótipo de Sistema de Suporte à Decisão com integração de informações gerenciais e geográficas em um Data Warehouse espacial, para aplicação na área de gestão tributária municipal. O objetivo é disponibilizar para os gestores públicos, dos níveis tático e estratégico, um instrumento de trabalho bastante eficaz, capaz de permitir a eles uma visão temporal e espacial dos fatos armazenados. E, assim, contribuir para uma gestão pública pró-ativa.

PALAVRAS-CHAVE: Data Warehouse, PostGeoOlap, Tributação, Informações geográficas

ABSTRACT

DEVELOPMENT OF A SPATIAL DATA WAREHOUSE FOR THE MUNICIPAL TRIBUTARY SECTION WITH USE OF OPEN TECHNOLOGY

The importance of Decision Support Systems (DSS) nowadays is an uncontested fact, mainly in face of the increasing need of tools capable of aiding in analysis of great amount of data and of promoting the Business Intelligence. In this research, the basis for a DSS was the development of a Data Warehouse (DW), which it is a tool set capable of providing business rule oriented queries in a more objective way than a production environment, since it is composed by a subject structured database that stores the organization's historic data.

A point to highlight is the use of geographical data, which promoted more richness on information stored and processed. This approach became feasible thanks to the use of an On Line Analytical Processing (OLAP) tool, PostGeoOlap, which made it possible a new functionality: analytical queries involving spatial data types.

Both the PostGeoOlap tool and the database management system (DBMS) used, PostgreSQL, are Free Software (with open source code and free license), making the solution even more attractive by its virtue of smaller prototyping cost.

The object of this work is the development of a Decision Support System prototype with management and geographical information integration in a spatial Data Warehouse, for application on the area of municipal tributary administration. The goal is to make it available for public managers, at tactical and strategic levels, a quite effective instrument, capable of allowing then a temporal and spatial view of stored facts. And so contribute for a pro-active public administration.

KEY WORDS: Data Warehouse, PostgeoOlap, Tributary, Geographical information

LISTA DE FIGURAS

Figura 1:	Um ambiente de Data Warehouse.....	20
Figura 2:	Comparação com a imagem de um cubo.....	22
Figura 3:	Esquema Estrela ou Star Schema.....	23
Figura 4:	Conjunto de tabelas fato e dimensão.....	24
Figura 5:	Conjunto de tabelas fato e dimensão normalizadas (Snowflake).....	25
Figura 6:	Representação vetorial e matricial.....	35
Figura 7:	Hierarquia de Classes de Tipos de Dados Geométricos.....	37
Figura 8:	Metamodelo da ferramenta PostGeoOlap.....	41
Figura 9:	Arquitetura da Ferramenta PostGeoOlap.....	43
Figura 10:	Estrutura Organizacional da Secretaria da Fazenda.....	49
Figura 11:	Diagrama de caso de uso de negócios utilizando dependência entre pacotes	51
Figura 12:	Plataforma utilizada na PMCI.....	57
Figura 13:	Diagrama de Classes do DW da Prefeitura Municipal.....	58
Figura 14:	Dimensão Tempo.....	60
Figura 15:	Dimensão Contribuinte.....	61
Figura 16:	Dimensão Natureza.....	61
Figura 17:	Dimensão Débito.....	63
Figura 18:	Dimensão Origem do Débito.....	64
Figura 19:	Dimensão Local de Pagamento.....	65
Figura 20:	Dimensão Tipo de Pagamento.....	66
Figura 21:	Fato Item do Débito.....	67
Figura 22:	Fato Pagamento.....	67
Figura 23:	Seleção do Esquema para Conexão.....	71
Figura 24:	Criação do Cubo.....	71
Figura 25:	Seleção da Tabela Fato.....	72
Figura 26:	Atributos da Tabela Fato para operações.....	73
Figura 27:	Seleção da Tabela Dimensão Débito.....	74
Figura 28:	Tela Principal do PostGeoOlap.....	75
Figura 29:	Resultado dos Imóveis Quitados em 2004.....	76
Figura 30:	Inadimplência do IPTU com adoção de critérios.....	79

LISTA DE TABELAS

Tabela 1:	Abordagem relacional entre o Uso e a Área.....	21
Tabela 2:	Abordagem relacional entre o Padrão e a Área.....	21
Tabela 3:	Abordagem dimensional.....	21
Tabela 4:	Metadados de um atributo da Dimensão Contribuinte.....	29
Tabela 5:	Exemplo de uma Operação <i>Drill up</i>	30
Tabela 6:	Exemplo de uma Operação de <i>Drill down</i>	30
Tabela 7:	Tabela antes da Operação de <i>Drill across</i>	31
Tabela 8:	Tabela depois da Operação de <i>Drill across</i>	31
Tabela 9:	Tabela antes da Operação de <i>Slice and dice</i>	31
Tabela 10:	Tabela depois da Operação de <i>Slice and dice</i>	31
Tabela 11:	Tabela antes da Operação de Pivotamento.....	31
Tabela 12:	Tabela depois da Operação de Pivotamento.....	32
Tabela 13:	Quadro comparativo entre as representações matricial e vetorial.....	36
Tabela 14:	Funções para construção de valores geométricos estabelecidos pelo (WKT)..	38
Tabela 15:	Estereótipos usados para os tipos de dados do OpenGIS.....	38
Tabela 16:	Relação de Casos de Uso da Ferramenta PostGeoOlap.....	42
Tabela 17:	Quadro ilustrativo da formação do tributo.....	47
Tabela 18:	Total de registros inseridos.....	68
Tabela 19:	Comparativo de Lançamentos e Retorno do IPTU.....	77
Tabela 20:	Comparativo de Lançamentos e Retorno do ISS.....	77
Tabela 21:	Inadimplência no IPTU por Categorias.....	80
Tabela 22:	Inadimplência entre Terrenos e Edificações.....	82
Tabela 23:	Total Arrecadado por Faixa de Valor do Tributo.....	83
Tabela 24:	Quantidade de Pagamentos por Mês (2000 a 2004).....	84
Tabela 25:	Quantidade de Pagamentos em Cota Única.....	85
Tabela 26:	Créditos Realizados Após o Vencimento.....	85

LISTA DE SIGLAS

ACI – Associação Cartográfica Internacional
BSC – Balanced Scorecard
BSD – Berkeley Software Distribution
BI – Business Intelligence
CTN – Código Tributário Nacional
CRM – Customer Relationship Management
DATACI – Empresa Processamento de Dados do Município de Cachoeiro de Itapemirim
DDL – Data Definition Language
DML – Data Manipulation Language
DER – Diagrama de Entidade e Relacionamento
DOLAP – Desktop OLAP
DSS – Decision Support System
DW – Data Warehouse
EIS – Execute Information System
ETL – Extract, Transformation and Load
FEBRABAN – Federação Brasileira de Bancos
HD – Hard Disk
HOLAP – Hybrid OLAP
IBAM – Instituto Brasileiro de Administração Municipal
IBGE – Instituto Brasileiro de Geografia e Estatística
IPTU – Imposto Predial e Territorial Urbano
ISS – Imposto sobre Serviços de Qualquer Natureza
JDBC – Java Database Connectivity
MOLAP – Multi-dimensional OLAP
OGC – Open Geospatial Consortium
OLAP – On Line Analytical Processing
OMG – Object Management Group
PMCI – Prefeitura Municipal de Cachoeiro de Itapemirim
ROLAP – Relational OLAP
SGBD – Sistema Gerenciado de Banco de Dados
SGBDOR – Sistema Gerenciador de Banco De Dados Objeto-Relacional

SIG – Sistemas de Informações Geográficas

SQL – Structured Query Language

SSD – Sistema de Suporte à Decisão

UML – Unified Modeling Language

XML – Extensible Mark Language

SUMÁRIO

1 INTRODUÇÃO	13
1.1 MOTIVAÇÃO	14
1.2 OBJETIVOS DO TRABALHO	15
1.3 ORGANIZAÇÃO DO TRABALHO.....	15
2 DATA WAREHOUSING E OLAP	16
2.1 CONCEITOS E DEFINIÇÕES	16
2.1.1 Histórico.....	16
2.1.2 Data Warehouses	18
2.2 A MODELAGEM DIMENSIONAL.....	21
2.2.1 Esquema Estrela	22
2.2.2 Esquema Snowflake.....	25
2.2.3 Granularidade	25
2.2.4 Atualização das Dimensões	26
2.2.5 Escolha das Chaves de Identificação.....	27
2.2.6 Agregações.....	27
2.2.7 Hierarquias	27
2.3 PROCESSO DE EXTRAÇÃO, TRANSFORMAÇÃO E CARGA	28
2.4 METADADOS.....	29
2.5 OLAP.....	29
3 INTEGRAÇÃO DE DADOS GEOGRÁFICOS EM DATA WAREHOUSES	33
3.1 INTRODUÇÃO A SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS.....	33
3.2 PADRONIZAÇÃO DE INFORMAÇÕES GEOGRÁFICAS	36
3.2.1 Tipos de Dados Geométricos.....	37
3.2.2 Estereótipos	38
3.3 POSTGRESQL E POSTGIS	39
3.4 POSTGEOOLAP.....	40
4 DESCRIÇÃO DO PROBLEMA: GESTÃO DE TRIBUTOS	44
4.1 O TRIBUTO	45
4.2 CENÁRIO ATUAL: O SETOR TRIBUTÁRIO DA PREFEITURA DE CACHOEIRO DE ITAPEMIRIM.....	48
5 CONCEPÇÃO E ELABORAÇÃO DO DATA WAREHOUSE	52
5.1 DESCRIÇÃO DO ESTUDO DE CASO.....	52

5.2 LEVANTAMENTO DE REQUISITOS	53
5.2.1 Entrevistas e questionário com usuários.....	53
5.2.2 Engenharia reversa dos bancos de dados transacionais	56
5.3 ESQUEMA ESTRELA DO SISTEMA DE GESTÃO DE TRIBUTOS DA PMCI.....	57
5.3.1 Tabelas Dimensão.....	59
5.3.1.1 Dimensão Tempo.....	59
5.3.1.2 Dimensão Contribuinte.....	60
5.3.1.3 Dimensão Natureza.....	61
5.3.1.4 Dimensão Débito	62
5.3.1.5 Dimensão Origem do Débito	63
5.3.1.6 Dimensão Local de Pagamento.....	64
5.3.1.7 Dimensão Tipo de Pagamento	65
5.3.2 Tabelas Fato	66
5.3.2.1 Fato Item do Débito	66
5.3.2.2 Fato Pagamento	67
5.4 PROCESSO DE ETL	67
6 CONSTRUÇÃO DO DATAWAREHOUSE USANDO O POSTGEOOLAP	70
6.1 UTILIZAÇÃO DA FERRAMENTA POSTGEOOLAP	70
6.2 ANÁLISE DOS RESULTADOS	75
6.2.1 Análise dos lançamentos tributários.....	75
6.2.2 Análise dos créditos tributários.....	82
7 CONCLUSÃO	86
7.1 CONTRIBUIÇÕES DO PRESENTE TRABALHO	87
7.2 SUGESTÕES PARA TRABALHOS FUTUROS	87
8 REFERÊNCIAS	89
APÊNDICE A - OS 67 TRIBUTOS BRASILEIROS	93
APÊNDICE B - CÓDIGO FONTE UTILIZADO NO PROCESSO DE ETL.....	95
APÊNDICE C - QUESTIONÁRIO UTILIZADO NO LEVANTAMENTO	110

1 INTRODUÇÃO

Afirmar que a fronteira entre a história da humanidade e dos tributos sempre foi muito tênue não é um absurdo, visto que o passado atesta tal afirmação através de seus registros. A partir do momento em que o homem deixou sua vida itinerante e fixou sua residência num espaço de terra, ele começou a desenvolver valores materiais que, associados a uma vida em sociedade, acabaram por culminar numa necessidade de melhorias constantes da qualidade de vida e da manutenção do grupo, tornando necessária a divisão dos custos desse desenvolvimento entre os beneficiados. Desde então, o tributo sempre esteve presente, muitas vezes financiando guerras, em outras, sendo o pilar de revoluções que marcaram o passado e contribuíram para o presente. No decorrer desta mesma história e em tempos mais recentes, após a revolução industrial, as atividades profissionais passaram a ser o foco de constantes aprimoramentos visando ao aumento da produtividade.

Com o advento da informática e seus recursos, muitas dessas atividades passaram por uma verdadeira revolução: o papel deixa de ser o meio de controle, cedendo lugar para os registros armazenados em sistemas que possuíam uma capacidade de processamento, recuperação e armazenamento infinitamente superiores. Uma considerável massa de dados passa a ser gerada numa progressão de crescimento geométrica.

Em tempos mais recentes, esse montante de dados começou a ser utilizado para a extração de informações capazes de influenciar no processo decisório. Os Sistemas de Suporte à Decisão (SSD) começaram a ser projetados e novas técnicas foram sendo criadas na busca por informações como banco de dados e ferramentas de pesquisa especialmente desenvolvidos para esse fim. A tecnologia mais recente aponta não só para o tratamento dos dados mas também para o tratamento das regras de negócios como um importante objeto de formalização, estruturação e medida de performance (AMORIM; TANAKA, 1998, desta forma, reforçando a

idéia de agregar inteligência ao negócio, o que vai ao encontro do conceito de Business Intelligence (BI), que pode ser entendido como a utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios da empresa (BARBIERI, 2001). O emprego de dados espaciais foi outro fator que veio somar a essas tecnologias, pois permitiu uma visão sobre um novo prisma, o geográfico.

Na década de 90, idealizou-se o conceito de Data Warehouse (DW), que foi definido como uma projeção de um banco de dados organizado por assuntos, que contivesse todos os fatos ocorridos e disponíveis para consultas específicas.

Este trabalho visa à integração de algumas das principais tecnologias voltadas para área de sistemas de suporte à decisão como Data Warehouse, ferramentas OLAP (*On-Line Analytical Processing*) e Sistemas de Informações Geográficas (SIG), através da ferramenta PostGeoOlap (COLONESE,2004), de modo a utilizá-las no setor tributário municipal.

1.1 MOTIVAÇÃO

Dentre os motivos que influenciaram a escolha deste tema de dissertação está a participação constante dos tributos no cotidiano dos cidadãos, reforçando assim sua importância. A experiência profissional do autor, adquirida através da participação em projetos de desenvolvimento e manutenção de sistemas na área tributária, desde 1990, também foi um fator decisivo, visto que a base de dados atual proporciona mais de uma década de registros tributários municipais, criando, dessa forma, um ambiente favorável através da associação entre a pesquisa e as situações práticas vindas de necessidades bastante heterogêneas.

Outro fator de grande relevância foi a carência de ferramentas voltadas para o apoio ao setor tático e estratégico do setor público, bem como o patrocínio proporcionado pelos gestores que, além do apoio à iniciativa, colaboraram através da participação em entrevistas com sugestões e críticas que enriqueceram muito esta pesquisa.

Também foi causa determinante a utilização de um sistema de geoprocessamento que embora esteja paralisado na fase de implantação, serviu de fonte para o levantamento dos dados geográficos dos imóveis.

E por fim, a possibilidade de utilizar a ferramenta PostGeoOlap (COLONESE *et al.*,2005), de código aberto (*open source*), que proporciona a união de dados tabulares e geográficos dentro de um único contexto, tendo como diferencial a integração desses dois tipos desde o nível conceitual até o nível físico. Tal ferramenta foi tema de dissertação de mestrado e

encontra-se atualmente sendo migrada para uma plataforma universal através da linguagem Java (MANHÃES,2007). Portanto, uma solução de baixo custo para a municipalidade.

1.2 OBJETIVOS DO TRABALHO

Primeiramente, o foco está na modelagem e desenvolvimento de um Data Warehouse voltado unicamente para a gestão de tributos, ou seja, a criação de um ambiente, distinto do transacional, capaz de reunir todos os fatos tributários ocorridos no município e organizados de uma forma objetiva, de maneira que permita, por exemplo, a análise de séries históricas, a projeção de perfis dos contribuintes, baseados em seus comportamentos e a apuração de resultados de políticas tributárias, como o retorno de anistias concedidas no decorrer dos anos.

Outro objetivo é a inserção de dados geográficos vindos de um sistema de geoprocessamento, que proporcionará uma projeção espacial dos tributos municipais; o que certamente trará um recurso a mais para a análise. A união entre dados tabulares e espaciais facilita muito a observação de tendências e a criação de indicadores que conseqüentemente auxiliarão no direcionamento de políticas de desenvolvimento.

Vê-se, por isso, a intenção de disponibilizar para o gestor público uma ferramenta de apoio gerencial capaz de proporcionar liberdade na maneira de agrupar e analisar os dados de acordo com suas necessidades. Ademais, será o início de um projeto que futuramente servirá de base para expansão de novas ferramentas de suporte à decisão.

1.3 ORGANIZAÇÃO DO TRABALHO

Este trabalho encontra-se organizado na seguinte seqüência: o Capítulo 2 aborda os conceitos e características de Data Warehouses e da tecnologia OLAP (*On-Line Analytical Processing*). No Capítulo 3 são abordados os conceitos sobre Sistemas de Informações Geográficas, os padrões internacionais, sua integração no DW e a ferramenta PostGeoOlap. No Capítulo 4 são expostas as fundamentações sobre os tributos, tanto na forma jurídica como em sua aplicabilidade na esfera municipal. O Capítulo 5 descreve a concepção e elaboração do modelo de DW proposto. O capítulo 6 contempla a construção deste DW através da ferramenta PostGeoOlap. Por fim, o Capítulo 7 traz as conclusões finais sobre os resultados obtidos, as contribuições do presente trabalho e as sugestões para trabalhos futuros. A dissertação é complementada como os Apêndices A, B e C, que contêm material suplementar do trabalho realizado.

2 DATA WAREHOUSING E OLAP

As tecnologias de DW e OLAP, além de se complementarem, são relativamente novas, se comparadas com a história da informática; e sua aplicação tem produzido resultados nas mais diversas áreas, principalmente na área da Gestão da Informação, onde os executivos devem estar cientes de que as necessidades e a importância da informação podem crescer de maneira exponencial em relação ao crescimento das empresas (OLIVEIRA,2005). Neste capítulo, é realizada uma revisão dos conceitos sobre Data Warehouse e OLAP, visando a sua aplicação ao domínio de administração de tributos municipais.

2.1 CONCEITOS E DEFINIÇÕES

2.1.1 Histórico

Com o advento da computação e sua posterior acessibilidade, uma jornada foi iniciada rumo à informatização de diversas atividades humanas. Naquele momento, a utilização do computador se resumia em armazenar os registros resultantes de atividades que outrora eram manuais. Por exemplo, uma nota fiscal que antes era preenchida a mão tinha seus dados lançados num programa que a imprimia e automaticamente atualizava a quantidade do estoque. Este procedimento não estava errado e nem deixou de ser utilizado atualmente, mas o foco da época se limitava apenas em automatizar o serviço.

Contudo, não havia uma preocupação em analisar se um processo poderia ser ajustado, substituído ou otimizado antes de ser informatizado, ou seja, se um procedimento estivesse sendo feito de forma inadequada ou com alguma espécie de erro, ele teria sua propagação potencializada, uma vez que o usuário lançava os dados no sistema em grandes quantidades e num curto espaço de tempo.

Como não havia padrões, métricas ou formalismos, os sistemas eram concebidos de acordo com o conhecimento de cada profissional. Aliada a este fator estava a falta de integração, que proporcionava a redundância de dados e rotinas; por exemplo, uma empresa (comércio ou prestação de serviços) em situação regular tinha seus dados lançados no cadastro de ISS (Imposto Sobre Serviços) e as mesmas informações ou parte delas aparecia também no setor de cadastro imobiliário, caso fosse sediada num imóvel próprio.

As limitações dos recursos físicos (*hardware*) bem como seus custos elevados geravam uma forte dependência nos ambientes informatizados, propiciando o aparecimento de ilhas de informações, o que era um fato bastante comum na era dos computadores tipo “*mainframe*”. Antes de criar uma tabela era necessário calcular a quantidade de espaço que um registro ocuparia, uma vez que o armazenamento de dados era uma preocupação constante. Em consequência disso, campos eram gravados com tamanho menor que o necessário, como foi o caso da data que armazenava somente os dois últimos dígitos do ano, resultando o chamado Bug do Milênio; também vale ressaltar os campos com valores, que muitas vezes truncavam por terem sido concebidos com um tamanho de acordo com a necessidade inicial, e que não estavam preparados para uma expansão.

Logo após a criação do Modelo Relacional de banco de dados, em 1970, por Egdard Codd, um marco na história do processamento de dados, Peter Chen, na segunda metade da década de 70, criou o Modelo de Entidades e Relacionamentos, ou melhor, o Diagrama de Entidades e Relacionamentos (DER). Esta criação possibilitou uma compreensão mais efetiva dos bancos de dados relacionais, que foi percebido na década de 80, uma vez que em sua concepção fundamental estava a eliminação da redundância dos dados. As tabelas, chamadas de Entidades, eram representadas por retângulos e os relacionamentos, que representavam a associação entre as entidades, por losangos. Esse modelo mostrou-se apropriado para sistemas transacionais que necessitavam de estrutura normalizada e que lhes garantisse um bom desempenho. Em virtude deste alto grau de normalização, o diagrama de um sistema ficava com um grande número de tabelas, o que dificultava a compreensão e utilização como documentação do projeto junto ao usuário.

Aliada a este paradigma, a necessidade de se obter informações relevantes a partir dos dados armazenados começou a ganhar força, visto que o dado apresentava um comportamento meramente estático e de pouca importância quando visto isoladamente, enquanto a informação demonstrava seu dinamismo através de sua participação nos processos de tomada de decisão.

Primeiramente, surgiram os EIS (*Executive Information Systems*), que eram sistemas com uma interface amigável, utilizando-se de mapas e planilhas como indicadores para os gerentes.

Os setores tático e estratégico das corporações começaram a trabalhar a massa de dados gerada pelo setor operacional e a observar que comportamentos, padrões e tendências poderiam ser identificados e extraídos deste conjunto. Começaram a surgir os Management Information Systems ou Sistemas de Informações Gerenciais e os DSS (*Decision Support System*), ou SSD (Sistemas de Suporte à Decisão). O primeiro fornecia informações para serem analisadas e comparadas, enquanto o segundo gerava um resultado que poderia influenciar na decisão a ser tomada.

Na década de 90, com a redução dos custos e a evolução dos meios de comunicação, o esforço voltou-se para o sentido de estabelecer uma padronização, como por exemplo: a UML (*Unified Modeling Language*), que segundo a especificação da OMG (*Object Management Group*) é uma linguagem gráfica para visualização, especificação, construção e documentação de artefatos para distribuição de objetos num sistema, a XML (*Extensible Mark Language*) criado pela W3C para permitir a troca de informações através de uma forma semi-estruturada utilizando uma linguagem de padrão universal e o padrão CORBA, também especificado pela OMG como uma arquitetura que permite a comunicação entre objetos, independente da linguagem e sistema operacional adotados. Em resumo, todos com o objetivo de criar uma normalização que facilite a integração e a interoperabilidade de ambientes e plataformas.

2.1.2 Data Warehouses

Com a a necessidade de disponibilizar um ambiente em que os executivos pudessem ter acesso aos dados, porém dispostos de uma forma mais objetiva, ou seja, organizados de acordo com suas necessidades, nasce o conceito de Data Warehouse.

Um Data Warehouse é um banco de dados específico, concebido através de uma técnica distinta que visa à consolidação de dados, de forma estrutural, oriundos de diversas fontes como os bancos utilizados diretamente pelos sistemas transacionais e/ou sistemas legados. Entretanto, com um foco distinto: reunir e disponibilizar a informação de maneira fácil e clara, apresentando sempre uma consistência e mantendo a característica de flexibilidade com relação às alterações naturais dos fatos, servindo assim como base sólida para auxiliar nas tomadas de decisão, mas acima de tudo ser aceito por aqueles que lidam diretamente com as regras de negócios.

Segundo William H. Inmon (1997:33), “Um Data Warehouse é um conjunto de dados baseado em assuntos, integrado, não volátil, e variável em relação ao tempo, de apoio à tomada de decisão”.

O DW é orientado por assuntos porque um de seus objetivos está no foco das principais áreas de negócios da empresa, uma vez que nos sistemas transacionais, devido ao fato de estarem altamente normalizados, um determinado tema geralmente necessita de várias tabelas para sua manutenção. E quem vai utilizar um DW precisa visualizar os dados de forma global e organizada.

A integração se traduz no fato de que os dados que povoam um DW serão provenientes de diversas fontes (sistemas). E um mesmo tipo de dado, vindo de várias origens para ser armazenado, deverá estar integrado através de uma padronização. Exemplificando, num determinado sistema que abastece o DW o conteúdo do campo que identifica pessoa física e jurídica é “PF” e “PJ” respectivamente. Já numa outra fonte como um sistema legado, a pessoa física é representada pelo código 1 e a pessoa jurídica pelo código 2. Para que sejam carregados num DW todos estes conteúdos deverão ser convertidos para um conteúdo único.

A partir do momento em que um registro é inserido num DW, significa que o seu estado naquele momento foi registrado, pois, caso ele venha a sofrer alguma alteração no nível transacional, um novo ponto no tempo será considerado, permitindo assim, o acompanhamento de sua evolução, ou seja, como o dado foi inserido, que modificações ele sofreu e qual sua situação atual (variação no tempo).

Um DW é projetado para atender as necessidades gerenciais, ou seja, ser uma fonte confiável onde serão realizados todos os tipos de consultas, das elementares às mais complexas, envolvendo muitas variantes e condições. E esta confiabilidade só é possível através da preservação de todos os eventos ocorridos com os dados. A informação, uma vez inserida no DW, não sofrerá manutenção (alteração ou exclusão); caso isso aconteça, não será possível analisar a informação historicamente. Fazendo uma comparação: seria como fotografar um dado e mantê-lo registrado no tempo. Se fosse alterado ou excluído, sua referência original se perderia, por isso a característica de não volatilidade.

Embora algumas correntes recentemente venham questionando este aspecto e sugerindo mudanças, os projetos de DW continuam respeitando este procedimento.

Conforme ilustrado na Figura 1, um ambiente de DW é composto de:

- Fontes de dados, de onde são retirados, regularmente, os dados que serão armazenados no DW. Como por exemplo, os bancos de dados operacionais, os sistemas legados, as planilhas de dados e os arquivos do tipo texto;
- Software para a extração, transformação e carga dos dados, que desempenha o papel de retirar os dados necessários das fontes, padronizá-los e adicioná-los ao DW;
- Data Mart, que pode ser entendido como um DW departamental (de menor volume), em geral correspondente a um processo do negócio e que pode servir de fonte para um DW central;
- Um banco de dados central, que é o repositório principal, onde são armazenados todos os dados do DW;
- Metadados, que são os dados relacionados ao próprio conteúdo que é trafegado (armazenado), a fim de promover um suporte semântico para auxiliar na compreensão do ambiente do DW;
- Saídas, que podem ser geradas através de ferramentas de mineração de dados (*Data Mining*), consultas analíticas (*OLAP*) e relatórios.

Alguns dos componentes serão descritos sucintamente em seções seguintes.

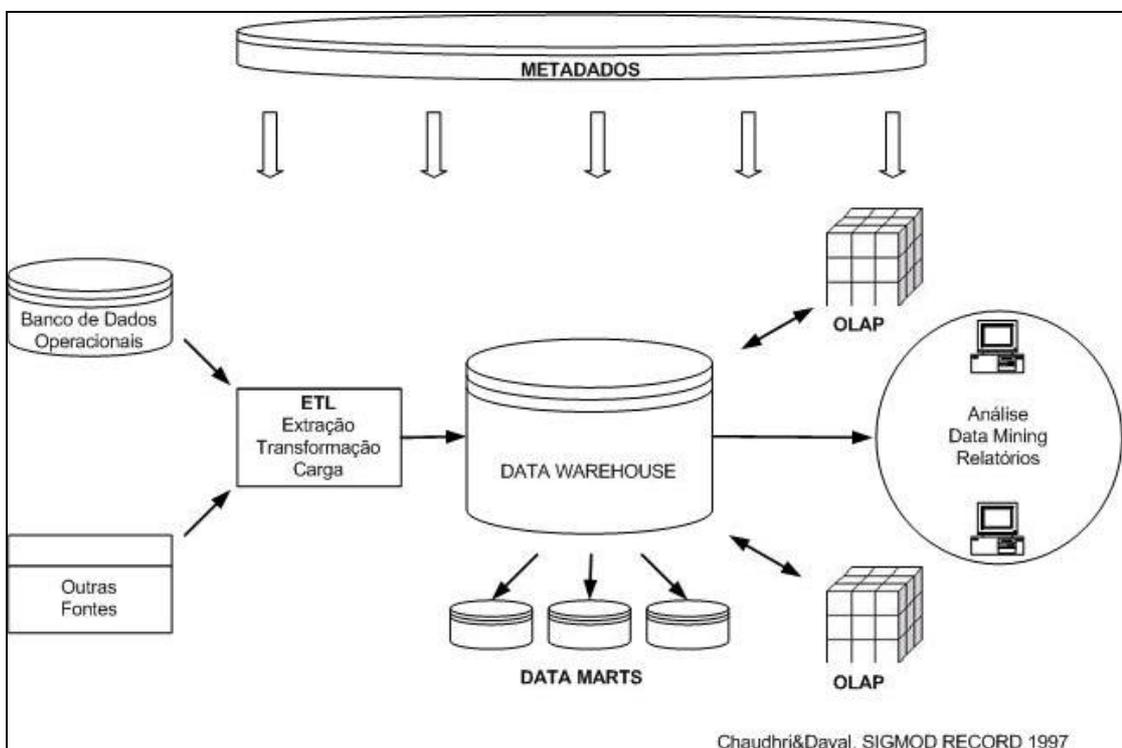


Figura 1 Um ambiente de Data Warehouse

2.2 A MODELAGEM DIMENSIONAL

A Modelagem Dimensional, introduzida por Ralph Kimball, é uma técnica que permite modelar os dados de forma unificada e organizada para uma melhor visualização e concepção de consultas. Comparando-se com o modelo relacional, a característica que mais se acentua é a forma como os dados são dispostos, agrupados por temas ligados aos processos do negócio e permitindo visões de vários “ângulos”. No sentido figurado, é como se no modelo relacional tivesse apenas os eixos “X” e “Y” (representando a largura e o comprimento) e na modelagem dimensional tivesse também o eixo “Z” (representando a profundidade) e outros eixos representando mais dimensões. A visão analítica se torna mais aprimorada e permite a descoberta de fatos que antes se encontravam dispersos. A seguir, uma comparação simples entre o modelo relacional (Tabelas 1 e 2) e o modelo dimensional (Tabela 3), que mostram as quantidades de imóveis por área, padrão e uso.

Uso do Imóvel / Área(m ²)	<= 300	> 300 e <= 1000	> 1000
Residencial	36847	327	6
Comercial	5480	309	60
Industrial	342	116	49

Tabela 1 – Abordagem relacional entre o Uso e a Área

Padrão do Imóvel / Área (m ²)	<= 300	> 300 e <= 1000	> 1000
A	286	42	0
B	3946	276	16
C	26692	220	34
D	10846	129	28
E	900	85	37

Tabela 2 - Abordagem relacional entre o Padrão e a Área

Uso / Padrão / Área (m ²)		<= 300	> 300 e <= 1000	> 1000
Residencial	A	93	31	0
	B	2870	220	2
	C	23137	67	2
	D	10010	8	0
	E	737	1	2
Comercial	A	193	10	0
	B	1054	52	14
	C	3394	134	23
	D	737	74	11
	E	102	39	12
Industrial	A	0	1	0
	B	22	4	0

Uso / Padrão / Área (m ²)		<= 300	> 300 e <= 1000	> 1000
	C	161	19	9
	D	99	47	17
	E	60	45	23

Tabela 3 - Abordagem dimensional

A Tabela 1 apresentou um simples relacionamento entre o tipo de uso dado aos imóveis (*residencial, comercial e industrial*) e três parâmetros de área em metros quadrados. Na Tabela 2 mantiveram-se os mesmos parâmetros de metragem, contudo, o relacionamento dessa vez foi feito com os padrões de construção (*A, B, C, D e E*). Na Tabela 3, ocorreu a união dos três vetores envolvidos: área, uso e padrão. Para se fazer uma analogia às dimensões contidas numa nessa consulta, foi utilizada a imagem de um cubo, conforme a Figura 2.

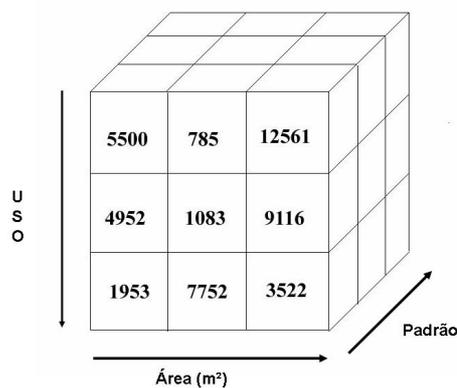


Figura 2 Comparação com a imagem de um cubo

Como as dimensões não se limitam a 3, foi criada a noção de hipercubo, isto é, um cubo com N dimensões.

2.2.1 Esquema Estrela

Um DW é composto por dois tipos de tabelas: fato e dimensão. Assim como todo sistema possui uma tabela principal, aquela que sofre uma manutenção mais freqüente e que é fundamental para as consultas e relatórios mais importantes, na modelagem dimensional é a tabela fato que ocupa a posição central do modelo, armazenando indicadores e medições numéricas de desempenho dos negócios. Vale ressaltar que as tabelas fato são altamente normalizadas, tanto que a denominação de “*esquema estrela*” ou “*star schema*” vem justamente da sua disposição central, recebendo todas as chaves das tabelas dimensão que com ela possuem

ligação, de modo que sua chave principal seja a composição de todas as chaves estrangeiras. Sua origem geralmente se encontra nas tabelas que, no modelo relacional, possuem relacionamentos de “muitos para muitos”. Nesta pesquisa, cujo foco principal é a área de tributos municipais, a tabela de tributos, denominada “Débito”, ocupa a posição central no modelo, conforme a Figura 3. Tal posicionamento se deve ao fato de ela receber todas as chaves principais das tabelas que se relacionam com ela. A tabela “Tempo” representa a questão de **quando** o fato ocorreu; a tabela “Contribuinte” identifica **quem** é o responsável pelo fato; a tabela “Origem do Débito” aponta de **onde** veio o fato, se é de origem imobiliária, mobiliária ou utilização de serviços e a tabela “Natureza” assume o papel de mostrar a composição do fato, ou seja, **o que** é o fato.

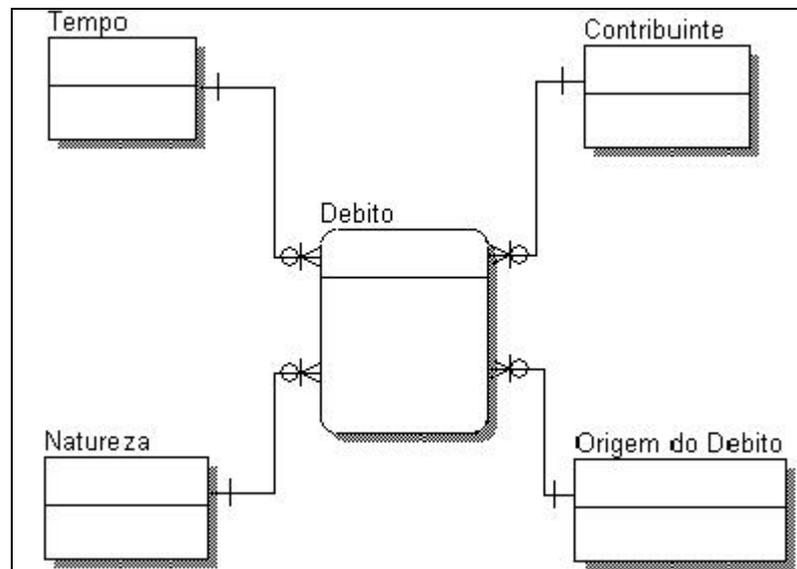


Figura 3 Esquema Estrela ou Star Schema

Um modelo dimensional pode conter mais de uma tabela fato, contudo, o excesso de emprego deste tipo de tabela pode ser um indicativo de falha no projeto, uma vez que um dos objetivos dessa técnica é a consolidação dos dados vindos de tabelas consideradas chaves no modelo transacional.

Numa tabela fato os atributos que podem ser utilizados como indicadores, sendo calculados e adicionados a qualquer tabela dimensão que com ela se relacione: são denominados *fatos aditivos*. Um exemplo deste tipo de fato são os valores dos impostos calculados, que podem ser utilizados diretamente em somas e levantamentos, pois possuem

relacionamento como o tempo, contribuinte, natureza e débito, conforme será visto no modelo disposto no capítulo 3.

Os fatos considerados *não-aditivos* são aqueles cuja utilização como adição não faz sentido algum, citando como exemplo, percentual de quitação de um parcelamento. Nesse caso só terá algum sentido se essa informação for utilizada como parte de uma média aritmética através do tempo.

Entretanto, outros atributos só podem ser adicionados em algumas dimensões; são os chamados *fatos semiaditivos*, como por exemplo, a inadimplência, que pode ser adicionada a determinadas tabelas, mas não em todas.

As tabelas dimensão são responsáveis pelo armazenamento da parte descritiva do modelo, ou seja, descrevem o significado dessas tabelas. Suas estruturas são altamente desnormalizadas uma vez que esse tipo de tabela carrega para dentro de si campos considerados descritivos, como por exemplo, uma tabela que possua em sua estrutura os campos código e descrição seria incorporada como atributo de uma dimensão fazendo com que a descrição apareça em todas as linhas que façam referência àquele código. Isto não ocorreria no modelo relacional e normalizado, pois apenas o código seria levado para tabela hierarquicamente superior, como chave estrangeira. Essa desnormalização gera estruturas com muitos atributos, visto que uma tabela dimensão resulta da junção de várias tabelas do modelo relacional, através de suas chaves estrangeiras.

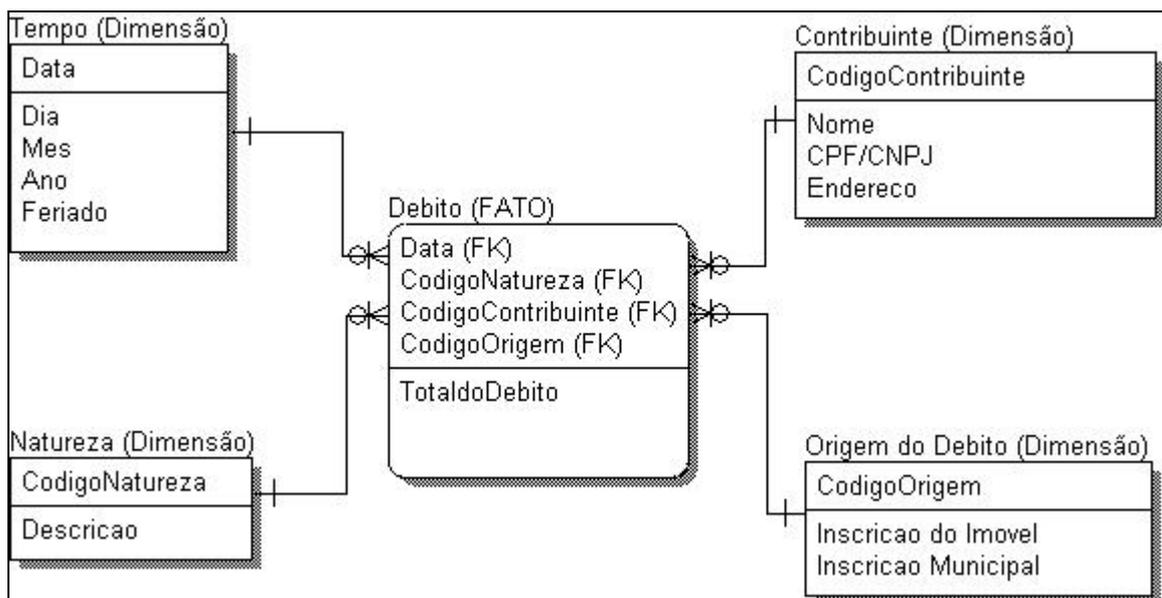


Figura 4 Conjunto de tabelas fato e dimensão

Nota-se que a composição da chave que identifica a tabela fato é formada pela união de todas as chaves principais das tabelas dimensão (Data, CodigoContribuinte, Natureza e Origem do Débito).

2.2.2 Esquema Snowflake

Snowflaking é o nome dado quando se aplica a técnica da normalização nas tabelas dimensão. Dessa forma, os atributos redundantes são removidos para tabelas secundárias, ou seja, aplica-se a terceira forma normal (3FN), conforme a Figura 5. Embora o objetivo seja economizar espaço, essa técnica traz mais desvantagens que benefícios, uma vez que dificulta a visualização por parte do usuário, tornando o modelo mais complexo e prejudicando o processo de pesquisa no banco, devido ao fato de exigir um maior número de junções para atender às necessidades de busca.

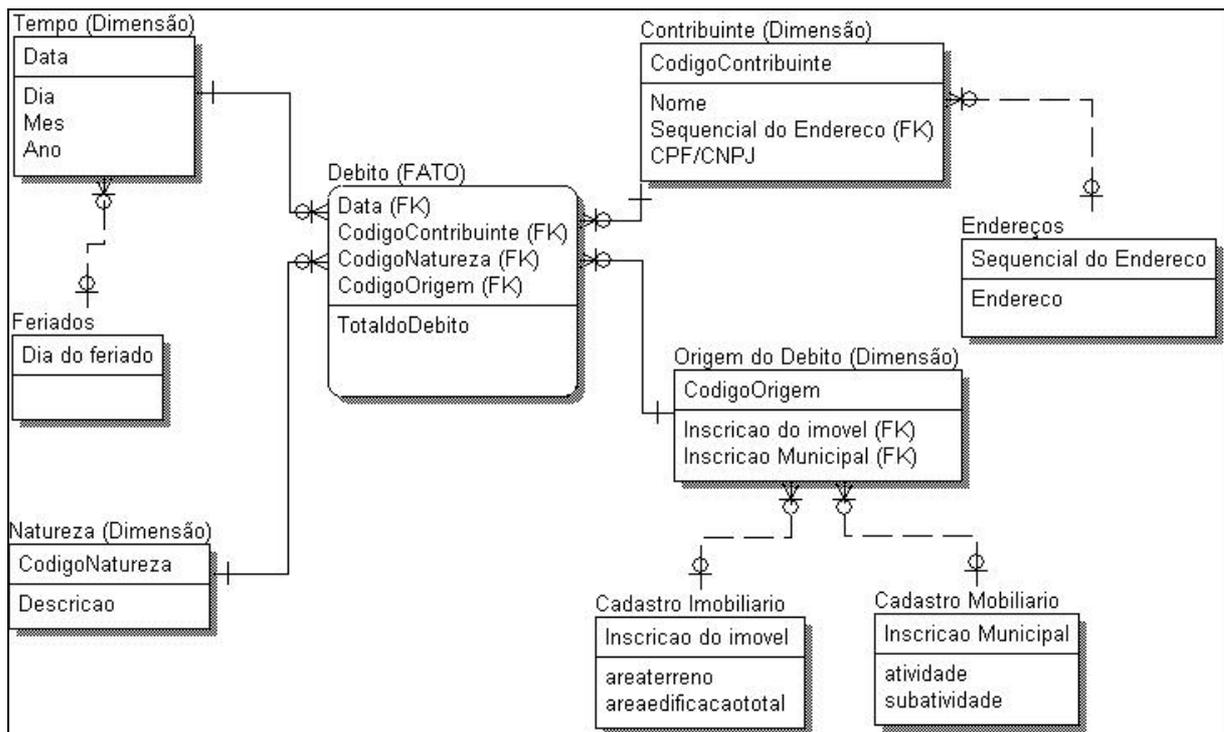


Figura 5 Conjunto de tabelas fato e dimensão normalizadas (Snowflake)

2.2.3 Granularidade

A granularidade se traduz no nível de detalhamento em que uma abordagem dimensional será empregada. Não há um nível granular padrão para se utilizar, pois cada

aplicação possui particularidades e será construída sobre os processos de negócio de cada cenário.

Quanto mais se respeitar a atomicidade dos dados, ou seja, menor granularidade (grão menor), melhor será a capacidade de análise dos dados e liberdade nas consultas. Quanto maior a granularidade (grão maior), a utilização de resumos e totalizadores será maior, criando uma limitação quando as consultas exigirem uma análise dos dados independente de sua totalização. Portanto, estabelecer o grão é um fator determinante para o sucesso de um projeto de DW.

2.2.4 Atualização das Dimensões

Assim como todo sistema, o DW, apesar de não sofrer exclusão devido a sua não volatilidade, também sofre manutenção. Mas como é realizada a carga de um determinado dado que já foi inserido num DW, ou seja, uma atualização de dados? Essa pergunta, segundo KIMBALL (1998), pode ser respondida através de três formas de atualização: a primeira, também chamada de *tipo 1* consiste no método da substituição dos valores, que funciona com a simples troca do valor antigo do atributo pelo novo valor. A vantagem desse método está na rapidez e facilidade de sua utilização associada ao seu valor que sempre estará atualizado, contudo, surge uma desvantagem bastante significativa, pois o histórico dos acontecimentos é perdido. A chave principal não é alterada e dessa forma não é possível analisar e acompanhar a evolução dos dados em virtude da mudança dos valores.

A segunda forma, conhecida como *tipo 2*, é a que melhor vai ao encontro das necessidades de preservação histórica dos dados dentro de um DW. Nesse tipo, toda vez que um registro sofre alguma alteração ao ser adicionado, uma nova linha será criada para receber o atributo alterado e repetir os que não foram modificados. Nesse caso tem-se um controle perfeito das alterações realizadas e sem limitação de ocorrências, uma vez que, a cada inserção, uma nova chave principal será utilizada. A desvantagem está no aumento do espaço, o que nos tempos atuais, não chega a ser um fator tão preocupante frente às novas tecnologias de armazenamento disponíveis. Mesmo com esse fator, essa modalidade foi a escolhida como forma de atualização para o modelo em questão.

E por último, a alteração do *tipo 3*, quando a alteração de um registro ocorrer pela segunda vez, é criada uma nova coluna para receber o novo valor e não uma linha. Ou seja, o valor atual do campo é transferido para a coluna duplicada e o novo valor é colocado na coluna

original. Desse modo, apenas os dois últimos estados do registro serão armazenados, o que é uma desvantagem semelhante à apresentada pela alteração do *tipo 1*.

2.2.5 Escolha das Chaves de Identificação

Como todas as alterações ocorridas com os dados são armazenadas nas tabelas dimensão, não se pode utilizar como chave única as mesmas chaves utilizadas no modelo transacional, pois nesse caso haveria um conflito no momento da inserção. Portanto, no modelo de DW as chaves consideradas únicas deverão ser chaves artificiais auto-incrementais (*surrogate keys*), permitindo assim a inclusão de um mesmo registro quantas vezes forem necessárias, como acontece nas alterações do *tipo 2*.

Com relação às tabelas fato, sua chave única será a composição de todas as chaves estrangeiras vindas das tabelas dimensão, mas podem ocorrer situações em que uma manutenção não resulte obrigatoriamente na vinda de novos valores por parte de todas as tabelas dimensão envolvidas. Para esta situação, uma chave única que represente o status “não aplicável” deverá ser empregada, pois uma tabela fato não pode conter chaves nulas.

2.2.6 Agregações

O termo agregação é designado para representar atributos que são resultantes de operações de outros atributos, portanto seu armazenamento na tabela fato é desnecessário. A sua transposição para uma tabela auxiliar faz com que as consultas ganhem performance, pois caso esse atributo seja necessário, ele já estará disponível na tabela fato agregada.

Exemplificando: no caso dos tributos municipais, que é o escopo desta pesquisa, um carnê de IPTU (Imposto Predial e Territorial Urbano), que normalmente é composto por até três itens, não necessitaria ter o atributo “total do débito” alocado na tabela fato uma vez que todos os itens estão armazenados.

A agregação é um dos recursos que se destaca como característica da ferramenta PostGeoOLAP devido seu emprego, conforme será visto no capítulo seguinte.

2.2.7 Hierarquias

Para serem analisados, os dados precisam estar organizados dentro de suas dimensões. Esta afirmação significa sua classificação ou sua posição dentro da hierarquia de uma dimensão.

A dimensão Tempo exemplifica mais claramente este conceito, uma vez que sua estrutura é composta por ano, semestre, quadrimestre, trimestre, bimestre, mês e dia.

Tal disposição faz com que se tenha a capacidade de referenciar fatos de acordo com sua posição ao longo da hierarquia, o que é muito útil para o controle de aplicações no mundo real (THOMSEN, 2002). Este recurso permite que a análise seja realizada de diversas formas e perspectivas, trazendo maior precisão na tomada de decisões.

2.3 PROCESSO DE EXTRAÇÃO, TRANSFORMAÇÃO E CARGA

Mesmo com o mercado atual oferecendo inúmeras plataformas, linguagens de programação e bancos de dados, ainda é muito comum a presença de empresas que utilizam sistemas baseados em linguagens de gerações passadas como o Cobol que, pelo fato de estarem atendendo às necessidades, continuarão em produção por um bom tempo; são os chamados *sistemas legados*.

Independentemente da ferramenta utilizada, todo sistema pode vir a ser uma fonte de dados para um DW. Periodicamente, dependendo do projeto, os sistemas transacionais podem fornecer dados e a via para esta exportação pode variar de diversas maneiras como a utilização de *scripts* executados a partir do “*prompt Shell*” do sistema operacional, execução de ferramentas de exportação e geração de arquivos do tipo Texto (TXT). Essa fase é chamada de *extração (extract)*, na qual os dados são copiados das fontes de origem para uma área onde serão analisados e tratados.

E por mais que contenha procedimentos de críticas e filtros, cada sistema pode ter sido desenvolvido em épocas diferentes, por diversos profissionais e seguindo metodologias distintas, ocasionando divergências na concepção de seus conteúdos. Como por exemplo, os sistemas que trabalham com valores indexados, como a UPF (Unidade Padrão Fiscal) e os sistemas que utilizam valores em moeda corrente (Real). Para que esses dados sofram operações de soma, por exemplo, eles deverão ser padronizados numa unidade de valor comum (UPF ou Real). Nesse momento os dados passam por uma fase de *transformação (transformation)*. Uma vez codificadas e retiradas todas as inconsistências, os dados estarão prontos para serem inseridos no DW, ou seja, a fase de *carga (load)*. Segundo KIMBALL (2004), a elaboração dos procedimentos de ETL (*Extract, Transform and Load*), freqüentemente, é responsável por 70% do tempo e esforço gastos na construção de um DW.

2.4 METADADOS

Como um DW é composto por dados vindos de diversas fontes, a única maneira de acompanhar todo esse trâmite é dispor de informações suficientes sobre esses dados durante todo o processo. Ou seja, um metadado é todo tipo de informação que possa ser armazenada sobre aquele determinado dado ou do banco de dados; enfim, tudo que possa caracterizá-lo ou identificá-lo, uma vez que, pelo fato de ser a junção de diversas tabelas, um catálogo com informações individualizadas é de fundamental importância para sua compreensão por parte de todos que estejam participando da gestão do modelo, do profissional que executa a extração ao usuário final. A participação dos metadados ocorre durante todo o processo.

Um exemplo de metadado é a origem do dado, isto é, o sistema de produção ou fonte de dados de onde foi transportado o dado para o DW. No caso em estudo, um exemplo está descrito na Tabela 4.

Tipo	Dimensão
Tabela	Contribuinte
Atributo	Nome
Tabela de Origem	UCCADA
Atributo de Origem	Cadanome
Sistema Transacional	Cadastro Único
Sigla do Sistema	UC

Tabela 4 – Metadados de um atributo da Dimensão Contribuinte

2.5 OLAP

Ter todos os dados relevantes, armazenados de forma temática numa base de dados não é o suficiente, sendo, portanto necessário acessá-los de forma inteligente para que sejam transformados em informação e conhecimento. As tecnologias OLAP vêm justamente com este propósito, permitir ao usuário explorar os dados (consultas, somas, comparações, etc.) dentro de um DW e dessa forma, proporcionar a descoberta de novas tendências, comportamentos e cenários. Essas ferramentas possuem derivações que atendem a diversos segmentos na área de banco de dados como:

- ROLAP (*Relational OLAP*) – neste caso, a ferramenta atua sobre uma base de dados relacional;
- MOLAP (*Multi-dimensional OLAP*) – empregada em bancos de dados multidimensionais;

- HOLAP (*Hybrid OLAP*) – de maneira híbrida permite trabalhar tanto com bancos de dados relacionais como dimensionais;
- DOLAP (*Desktop OLAP*) – sua utilização está mais voltada para computadores e banco de dados pessoais;

No seu funcionamento, as ferramentas OLAP possuem diversas operações básicas que permitem a mudança na granularidade das consultas bem como na forma de visualização. São elas:

Drill down: quando ocorre o aumento do nível de detalhe da informação e dessa forma, diminui a granularidade. A Tabela 5 apresenta uma consulta com o total de contribuintes inadimplentes, agrupados por ano. A Tabela 6 mostra uma consulta com o mesmo resultado, porém, separando os contribuintes em pessoas físicas e pessoas jurídicas. A operação de *Drill down* pode ser observada na mudança da Tabela 5 para Tabela 6.

Ano	Contribuintes Inadimplentes
2000	27542
2001	23455
2002	21258
2003	24547
2004	28501

Tabela 5 – Exemplo de uma Operação *Drill up*

Ano	Contribuintes Inadimplentes	
	Pessoa Física	Pessoa Jurídica
2000	20154	7388
2001	21925	1530
2002	19105	2153
2003	22154	2393
2004	24220	4281

Tabela 6 – Exemplo de uma Operação de *Drill down*

Drill up: ocorre o inverso do item anterior, exigindo um aumento da granularidade em virtude da redução do grau de detalhamento da informação; a sumarização é muito utilizada. Neste caso, este tipo de operação pode ser observado na mudança da Tabela 6 para a Tabela 5.

Drill across: nesse caso, dentro de uma dimensão o nível de solicitação de uma consulta é alterado sem respeitar a ordem de um determinado atributo. Por exemplo, na dimensão tempo, uma consulta realizada por ano é alterada diretamente para um resultado mensal sem passar pelos semestres, trimestres ou bimestres, conforme as Tabelas 7 e 8.

Valor Arrecadado (R\$)	Imposto Predial		Imposto sobre Serviços	
	2000	2001	2000	2001
Banco do Brasil	50.254,78	45.254,89	61.587,54	57.688,91
Caixa Econômica Federal	85.548,36	82.547,74	90.986,53	87.699,11
Banestes	90.155,87	89.458,14	98.744,96	95.201,66

Tabela 7 – Tabela antes da Operação de *Drill across*

Valor Arrecadado (R\$)	Imposto Predial		Imposto sobre Serviços	
	Jan/2000	Jan/2001	Jan/2000	Jan/2001
Banco do Brasil	2.258,36	2.101,68	3.755,26	4.101,56
Caixa Econômica Federal	3.522,94	3.225,96	3.158,77	3.900,25
Banestes	3.201,88	3.198,55	5.525,52	6.224,62

Tabela 8 – Tabela depois da Operação de *Drill across*

Slice and dice: traduzindo para uma linguagem simples significa filtrar o resultado de uma consulta (cubo) e mudar a perspectiva da visão, com possibilidade do surgimento de sub-cubos. Na transposição da Tabela 9 para Tabela 10 pode ser observado uma filtragem nos resultados.

Quantidade de Pagamentos	Cota Única e Parcelamentos			
	Janeiro	Fevereiro	Março	Abril
Imposto Predial	7201	8658	8901	6202
Imposto sobre Serviços	5204	6204	5888	7687

Tabela 9 – Tabela antes da Operação de *Slice and dice*

Quantidade de Pagamentos	Cota Única			
	Janeiro	Fevereiro	Março	Abril
Imposto Predial	3652	5986	6601	3157
Imposto sobre Serviços	2785	1954	2269	3101

Tabela 10 – Tabela depois da Operação de *Slice and dice*

Rotation ou Pivotamento: nesse aspecto o que muda é a forma de visualização dos dados, pois nenhuma posição é retirada ou adicionada, isto é, as linhas podem se inverter com as colunas, mas o conjunto permanece o mesmo. Conforme por ser visto comparando-se as Tabelas 11 e 12.

Quant. de Imóveis	Padrão A		Padrão B		Padrão C	
	2000	2001	2000	2001	2000	2001
Residencial	1500	1550	3785	3882	8221	8955
Comercial	575	601	942	956	1051	1157
Industrial	356	422	487	503	225	297

Tabela 11 – Tabela antes da Operação de Pivotamento

Quant. de Imóveis		Residencial	Comercial	Industrial
Padrão A	2000	1500	575	356
	2001	1550	601	422
Padrão B	2000	3785	942	487
	2001	3882	956	503
Padrão C	2000	8221	1051	225
	2001	8955	1157	297

Tabela 12 – Tabela depois da Operação de Pivotamento

3 INTEGRAÇÃO DE DADOS GEOGRÁFICOS EM DATA WAREHOUSES

Neste capítulo, serão descritas as tecnologias e as ferramentas utilizadas no desenvolvimento do presente estudo, a saber: Sistemas de Informação Geográficas, padrões OpenGIS para dados geográficos, SGBD PostgreSQL e sua extensão PostGIS, e a ferramenta PostGeoOlap para processamento OLAP espacial.

3.1 INTRODUÇÃO A SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS

Os Sistemas de Informações Geográficas permitem a manipulação de bancos de dados variáveis espacialmente e que utilizam dados geograficamente referenciados (georeferenciados) e dados não espaciais, incluindo operações que dão suporte às análises espaciais, ou seja, são sistemas computacionais utilizados para armazenar e manipular informações geográficas. Em sua concepção é um conjunto composto por *software*, *hardware*, dados, pessoas e métodos. Segundo BURROUGH (1998), “...um poderoso conjunto conjunto de ferramentas para coleta, armazenamento, recuperação, transformação e apresentação de dados espaciais do mundo real”.

Um SIG é projetado para a coleta, armazenamento e análise de objetos e fenômenos em que a localização geográfica é uma característica importante ou fundamental para as análises.

Os tipos de dados podem ser:

- *Dados temáticos* - admitem tanto representação matricial quanto vetorial;
- *Dados cadastrais e Redes* – sua parte gráfica é armazenada em forma de coordenadas vetoriais e seus atributos não gráficos são guardados em um banco de dados;
- *Imagens de sensoriamento remoto*: armazenadas em representação matricial;

Em um SIG, o elemento de dado geográfico, também denominado localizacional, é usado para fornecer uma referência para o elemento de dado atributo, também chamado descritivo ou não localizacional.

GOODCHILD (1988) apud MAGUIRE (1991) define a análise espacial como um conjunto de métodos analíticos que requerem acesso aos atributos dos objetos em estudo e sua informação localizacional.

Segundo CALIJURI (1999), os principais componentes de um SIG são:

- Banco de Dados Espaciais e de Atributos – é o núcleo do sistema e pode ser entendido como uma coleção de mapas e informações associadas na forma digital. É composto por um banco de dados espaciais e um banco de dados de atributos;
- Sistema de Apresentação Cartográfica – permite a extração de elementos selecionados do banco de dados;
- Sistema de Digitalização de Mapas – permite que mapas existentes em papel, sejam convertidos para a forma digital;
- Sistema de Gerenciamento de Banco de Dados – gerencia os componentes espaciais e de atributo;
- Sistema de Análise Geográfica – amplia a capacidade de consultas tradicionais ao banco de dados;
- Sistema de Processamento de Imagens – possibilita a análise de imagens de sensoriamento remoto;
- Sistema de Análise Estatística – oferece procedimentos estatísticos tradicionais e rotinas especializadas para a análise estatística do dado espacial;

Um SIG deve possuir um perfil que abranja interatividade, multiusuário, gráfico, volume e velocidade, memória virtual, sistemas de gerenciamento de banco de dados e custo. Segundo DANGERMOND (1990), as informações geográficas devem refletir três características:

- O fenômeno, ou característica propriamente dita, como uma variável, sua classificação, seu valor, seu nome;
- Sua localização espacial, ou seja, a localização do fenômeno no espaço geográfico;
- O tempo.

O modo espacial trata da variação geográfica; o modo temporal trata da variação em intervalos de tempo e o modo temático trata da variação de características.

Os pontos, linhas e polígonos são comumente definidos, nos mapas, usando-se um sistema de coordenadas cartesianas X, Y, como latitude e longitude, baseado nos princípios de geometria euclidiana.

Dentro de um SIG existem duas técnicas de representação (Figura 6): *vector* (vetorial) e *raster* (matricial). Na representação *vector*, os limites das características são definidos por uma série de pontos, que, quando interligados com retas, formam a representação gráfica daquela característica. Os pontos são codificados com um par de números que dão as coordenadas X e Y no sistema (latitude/longitude). Na segunda, a representação gráfica das características e dos atributos que elas possuem é armazenada em arquivos de dados unificados.

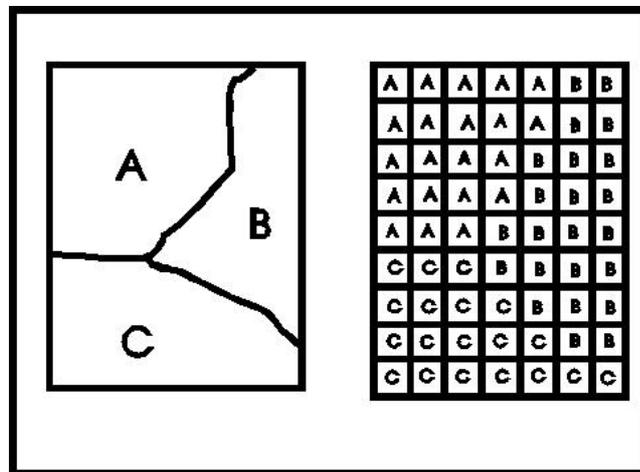


Figura 6 Representação vetorial e matricial
Fonte: (Câmara e Davis, 2000)

A vantagem do sistema raster é que o espaço geográfico é uniformemente definido em um simples e previsível uso. Além disso, sua estrutura está mais próxima da arquitetura dos computadores digitais. Outras características entre os dois sistemas encontra-se na Tabela 13.

A ACI (Associação Cartográfica Internacional), em 1964, estabeleu a cartografia como:

Conjunto de estudos e operações científicas, artísticas e técnicas, baseado nos resultados de observações diretas ou de análise de documentação, com vistas à elaboração e preparação de cartas, projetos e outras formas de expressão, assim como a sua utilização.

Aspecto	Representação Vetorial	Representação Matricial
Relações espaciais entre objetos	Relacionamento topológico entre objetos disponíveis	Relacionamentos espaciais devem ser inferidos
Ligação com banco de dados	Facilita associar atributos a elementos gráficos	Associa atributos apenas a classes do mapa
Análise, Simulação e Modelagem	Representação indireta de fenômenos contínuos Álgebra de mapas é limitada	Representa melhor fenômenos com variações contínuas no espaço. Simulação e modelagem mais fáceis
Escalas de trabalho	Adequado tanto a grandes quanto a pequenas escalas	Mais adequado para pequenas escalas (1:25.000 e menores)
Algoritmos	Problemas com erros geométricos	Processamento mais rápido e eficiente
Armazenamento	Por coordenada (mais eficiente)	Por matrizes

Tabela 13 – Quadro comparativo entre as representações matricial e vetorial
Fonte: (Câmara e Davis, 2000).

Escalas – são as relações entre a distância de dois pontos quaisquer do mapa com a correspondente distância na superfície da Terra.

Projeção de mapas – formulam matematicamente uma superfície terrestre representando-a, mas distorcendo algumas de suas características em função da projeção.

O SIG utiliza a cartografia digital como formadora de sua base cartográfica.

A digitalização é o processo de conversão de feições gráficas de um mapa convencional (pontos, linha e polígonos) para um formato compatível para uso em computador.

3.2 PADRONIZAÇÃO DE INFORMAÇÕES GEOGRÁFICAS

O OGC (*Open Geospatial Consortium*) é uma aliança internacional, envolvendo instituições governamentais e de pesquisa para o estabelecimento de conceitos, padrões, requisitos e especificações relacionadas ao tratamento de dados geográficos. Todo esse consenso resultou numa especificação denominada *OpenGIS*, criada para normatizar publicamente e operacionalizar este tipo de tecnologia. No documento *OpenGIS Simple Features Specification for SQL* estão as especificações e padrões para o tratamento de objetos geográficos, que serão abordados a seguir.

3.2.1 Tipos de Dados Geométricos

De acordo com o OpenGIS os elementos geográficos foram organizados dentro de uma hierarquia conforme a Figura 7.

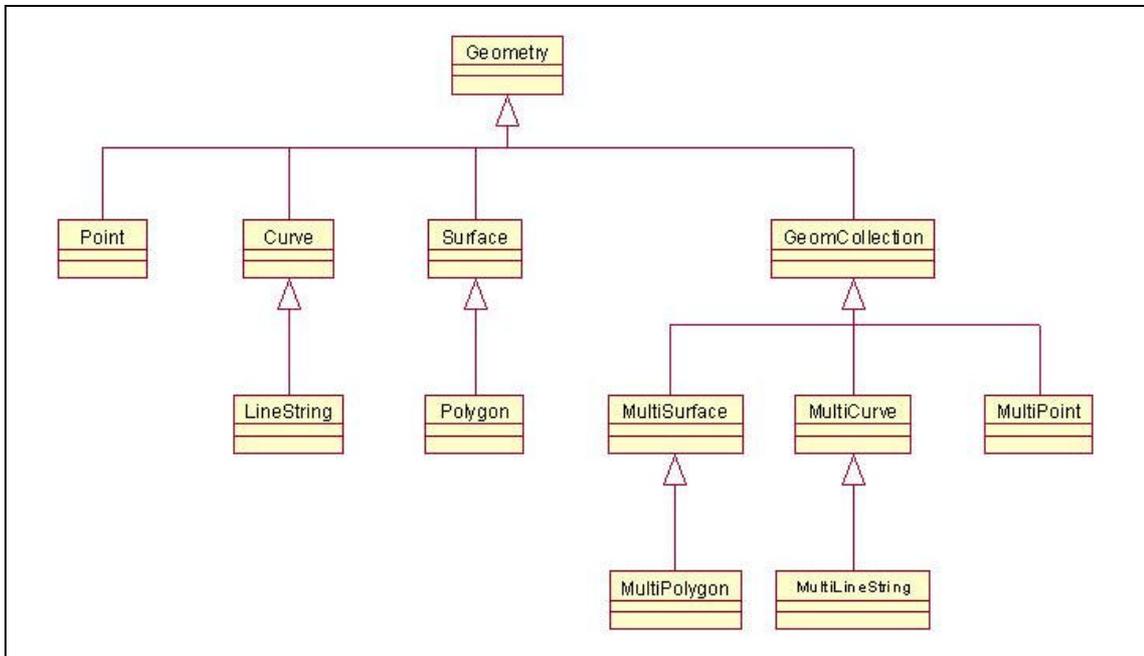


Figura 7 Hierarquia de Classes de Tipos de Dados Geométricos (OpenGIS)

Na raiz da árvore está a classe *Geometry*, que tem subtipos para *Point*, *Curve*, *Surface* e *Geometry Collection*. Cada objeto é associado a um sistema de referência espacial, o qual descreve uma coordenada espacial em cada objeto geométrico definido.

Um Ponto (*Point*) é um objeto geométrico de dimensão zero e representa uma simples localização espacial tendo um valor na coordenada X e outro na coordenada Y. *Multipoint* também possui dimensão zero e estão restritos aos pontos, que não estão conectados ou ordenados. A Curva (*Curve*) é um objeto de dimensão um e normalmente é armazenado com uma seqüência de pontos. *Linestring* é uma curva (*Curve*) com interpolação linear entre os pontos e cada par de pontos define uma *Linestring*. *Multicurve* é um objeto da *Geometrycollection* cujos elementos são *Curves* (Curvas) e não são instanciáveis; *Multilinestring* são *Multicurve* cujos elementos são *LineString*; *Surface* (superfície) é um objeto de dimensão dois; *Polygon* (polígono) é uma superfície plana; *MultiSurface* é um objeto de dimensão dois cujos elementos são as superfícies e *MultiPolygon* são *MultiSurface* nos quais os elementos são polígonos. Além dos tipos de dados, o OPENGIS define a representação dos dados de acordo com o formato KWT (*Well-Known Text*) mostrado na Tabela 14.

Tipo Geométrico	Representação Literal SQL Text	Comentário
Point	'POINT(10 10)'	Um ponto
LineString	'LINESTRING(10 10,20 20,30 40)'	Um linestring com três pontos
Polygon	'POLYGON(((10 10,10 20, 20 20, 20 15, 10 10)))'	Um Polygon com 1 anel exterior e nenhum interior
Multipoint	'MULTPOIN(10 10, 20 20)'	Um Multipoint com 2 pontos
MultiLineString	'MULTILINESTRING((10 10, 20 20),(15 15, 30 15))'	Um multilinestring com 2 linestring
MultiPolygon	'MULTIPOLYGON(((10 10, 10 20, 20 20, 20 15, 10 10,)), ((60 60, 70 70, 80 60, 60 60)))'	Um multipolygon com 2 polygons
GeomCollection	'GEOMETRYCOLLECTION(POINT (10 10), POINT(30 30),LINESTRING(15 15, 20 20))'	Uma coleção de geométrica composta de 2 pontos e 1 linestring

Tabela 14 – Funções para construção de valores geométricos estabelecidos pelo (WKT)

3.2.2 Estereótipos

Uma das facilidades proporcionadas pela notação UML é a utilização de estereótipos, que substituem a representação de associações com extensas hierarquias de objetos, sem perdas semânticas, permitindo a criação de diagramas ricos em conceitos distintos (COLONESE,2004), conforme a representação exposta na Tabela 15 para representar os tipos de dados do OpenGIS. Neste trabalho, apenas o estereótipo de “ponto” foi empregado, mas a ferramenta PostGeoOlap suporta todos os tipos.

Tipo de Dado	Estereótipo
Point (Ponto)	
LineString (Linha)	
Polygon (Polígono)	

Tipo de Dado	Estereótipo
MultiPoint (Multiponto)	
MultiLineString (Multilinha)	
MultiPolygon (MultiPolígono)	
GeometryCollection (Coleção Geométrica)	

Tabela 15 – Estereótipos usados para os tipos de dados do OpenGIS

3.3 POSTGRESQL E POSTGIS

O PostgreSQL é um SGBDOR (Sistema Gerenciador de Banco De Dados Objeto-Relacional) que está fundamentado nos padrões SQL ANSI-92 e 99; possui o código fonte aberto e licença livre BSD (Berkeley Software Distribution), o que permite sua implementação de acordo com novas necessidades. Sua concepção original foi para ambiente UNIX e sua instalação tanto neste ambiente quanto no LINUX é simples. Também pode ser executado utilizando-se o Windows, contudo, é necessária a utilização de dois softwares, o cygwin (emulador para Unix) e o cygpic (efetua a troca de informação entre processos). Entretanto, com a versão 8, lançada em janeiro de 2005, foi desenvolvido um pacote para instalação no windows que dispensa a instalação desses emuladores. Tal pacote, além de conter o banco de dados propriamente dito, contempla também a instalação do JDBC (*Java Database Connectivity*), o PGAdminIII (ferramenta gráfica para gerenciamento do banco) e como opcional um *template* para utilização de dados espaciais via PostGIS.

O SGBD PostgreSQL foi escolhido para este projeto devido ao fato de ser um banco de dados robusto, de administração simples, que suporta grande quantidade de transações concorrentes; possui arquitetura multi-processos, o que é uma característica fundamental para o mercado de banco de dados; suporta os mais variados tipos de dados existentes e tem sua distribuição gratuita, o que o torna mais acessível. Outro fator de vital importância é a comunidade (com integrantes em diversos países) que vem trabalhando constantemente na evolução do SGBD PostgreSQL, fazendo com que freqüentemente seja possível baixar novas versões, com correções, novas rotinas e extensões. Dentre outras características marcantes, foi o

primeiro banco com tecnologia aberta a utilizar uma extensão que permitisse a utilização de dados espaciais, o PostGIS.

O PostGIS é uma extensão criada para SGBD PostgreSQL a fim de suportar objetos geográficos; permitindo a armazenagem de dados vetoriais e de objetos de um SIG no banco de dados, além de suportar atributos básicos definidos pelo OGC (*OpenGIS Simple Features Specification for SQL*): *Point*, *LineString*, *Polygon*, *MultiPoint*, *MultiLineString*, *MultiPolygon* e *GeomCollection* (Kolodziej, 2003).

De acordo com a especificação, a integração entre o PostGIS e o banco de dados pode ser percebida através de duas tabelas: “SPATIAL_REF_SYS” e “GEOMETRY_COLUMNS”. A primeira armazena os indicadores numéricos e descrições textuais dos sistemas de coordenadas usados na base de dados. A segunda tabela armazena um metadado contendo a definição de cada coluna de dados do tipo geométrico.

3.4 POSTGEOOLAP

A grande dificuldade de projetos que envolviam a integração de ferramentas OLAP e SIG estava justamente na utilização de duas aplicações para sua realização. A ferramenta PostGeoOLAP vem justamente ao encontro dessa necessidade, pois reúne essas duas aplicações simplificando a modelagem. É uma ferramenta do tipo ROLAP, que segundo COLONESE (2004) “possibilita tornar mais simples a modelagem e implementação de um sistema de suporte à decisão que integre características analíticas de um DW e geográfica de um SIG, fazendo com que tais conceitos possam coexistir no mesmo modelo, tanto em nível conceitual quanto em nível de implementação e, mais ainda, que possa ser o mais direto possível tal mapeamento entre os citados níveis de abstração”. O metamodelo da ferramenta encontra-se na Figura 8.

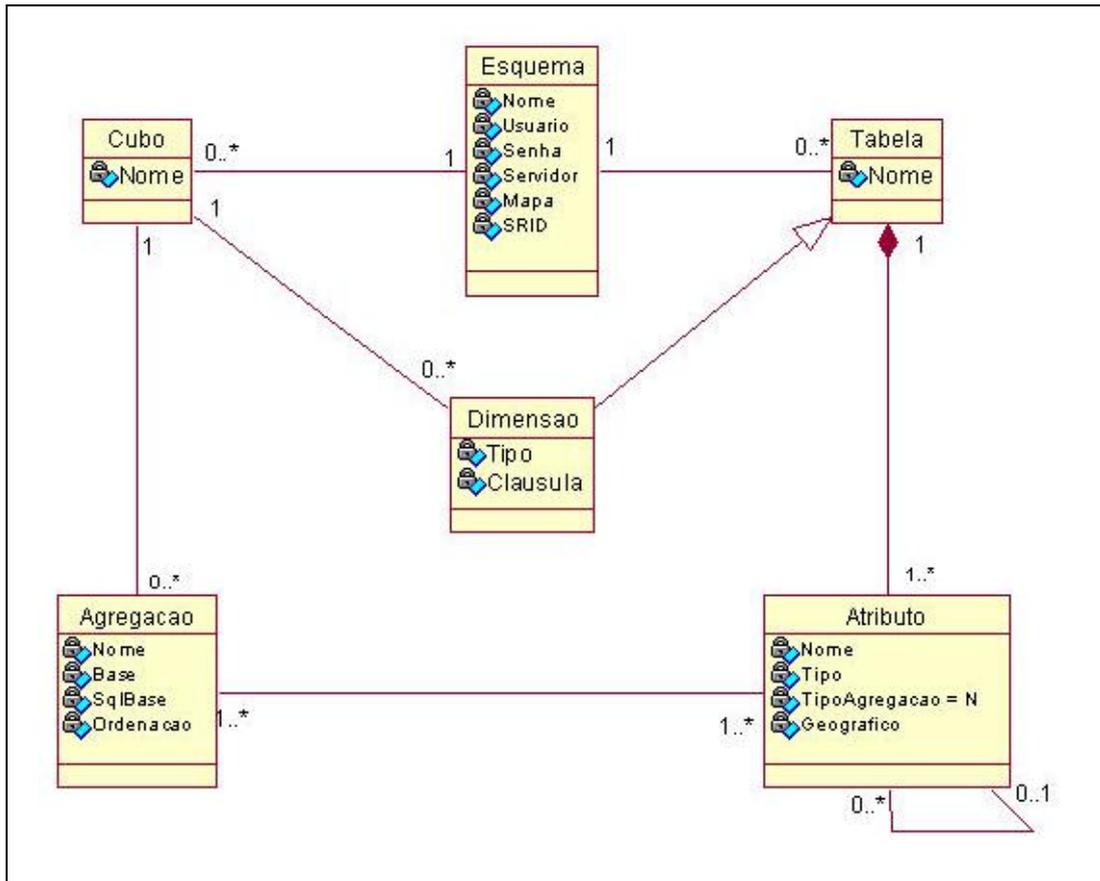


Figura 8 Metamodelo da ferramenta PostGeoOlap
 Fonte: (Colonese, 2004)

A classe *Esquema* contempla o banco de dados, neste caso, o SGBD PostgreSQL, contendo atributos (parâmetros) necessários para conexão com o banco onde está localizado o Data Warehouse; a classe *Cubo* armazena as informações sobre a consulta solicitada; a classe *Tabela* contém todas as tabelas envolvidas na consulta do cubo; já a *Dimensão*, sub-classe de *Tabela*, representa todos os componentes de um Cubo (incluindo tabelas fato e dimensão); a classe *Atributo* é responsável por representar todos os dados das tabelas (dimensões), bem como todas as hierarquias existentes entre os atributos; o auto-relacionamento descreve uma possibilidade de um atributo servir de rótulo para mais de um tipo geográfico e finalmente a classe *Agregação*, que será responsável pelo ganho de performance do sistema através do armazenamento das agregações geradas. O mecanismo de funcionamento da ferramenta PostGeoOlap pode ser melhor representado através dos casos de uso expressos na Tabela 16.

Caso de Uso	Finalidade
Criar Esquema	Cria uma conexão com um certo Banco de dados do PostgreSQL onde serão posteriormente definidos os Cubos
Criar Cubo	Cria um Cubo dentro de determinado Esquema, selecionando também uma tabela como Fato e determinando seus itens numéricos e as operações desejadas sobre tais itens
Adicionar Dimensão	Cria uma perspectiva para análise dos dados contidos na tabela Fato, por meio da seleção a partir de uma das tabelas do Banco de Dados. Determina também a hierarquia dentro da dimensão, atribuindo um nível para cada um de seus atributos
Criar dimensão não-agregável	Cria uma dimensão que, embora não se constitua em perspectiva para os dados da tabela Fato (e, por tal motivo, não participe da geração das agregações), serve como referência para comparativos geográficos com as demais dimensões que possuam atributos geográficos.
Processar Cubo	Verifica a massa de dados armazenados, por meio de seus metadados, buscando inferir sobre a velocidade de execução de consultas. As consultas avaliadas como de baixa performance (em termos de execução), são otimizadas por meio da criação de agregações, o que deixa o Cubo em condições de ser analisado sob qualquer perspectiva, dentro de tempo aceitável.
Analisar Dados	Provê uma interface capaz de permitir a seleção de atributos de uma consulta, sob restrições convencionais ou geográficas e de exibir o resultado de tal consulta sob duas visões: uma tabular para análise dos dados não-espaciais e outra em um mapa para a visualização dos dados espaciais.

Tabela 16 – Relação de Casos de Uso da Ferramenta PostGeoOlap
Fonte: (Colonese, 2004)

Ainda com relação ao funcionamento da ferramenta PostGeoOlap, presente no trabalho de Colonese (2004), dois fatores não poderiam passar sem um comentário, mesmo que de forma sucinta. Primeiramente a adoção de hierarquias (item 2.2.7) para cada atributo da dimensão que são atribuídas de acordo com a incidência granular do mesmo dentro da tabela.

Por último, a forte utilização de agregações (item 2.2.6) que são criadas com base nas hierarquias e de acordo com as necessidades das consultas.

Na Figura 9 pode-se observar a arquitetura da ferramenta PostGeoOlap, através de um *deployment diagram* (notação UML) proposto por MANHÃES (2007), que reescreveu o código

da ferramenta uma plataforma mais padronizada através da utilização da linguagem Java e de um framework JUMP, para visualização dos dados.

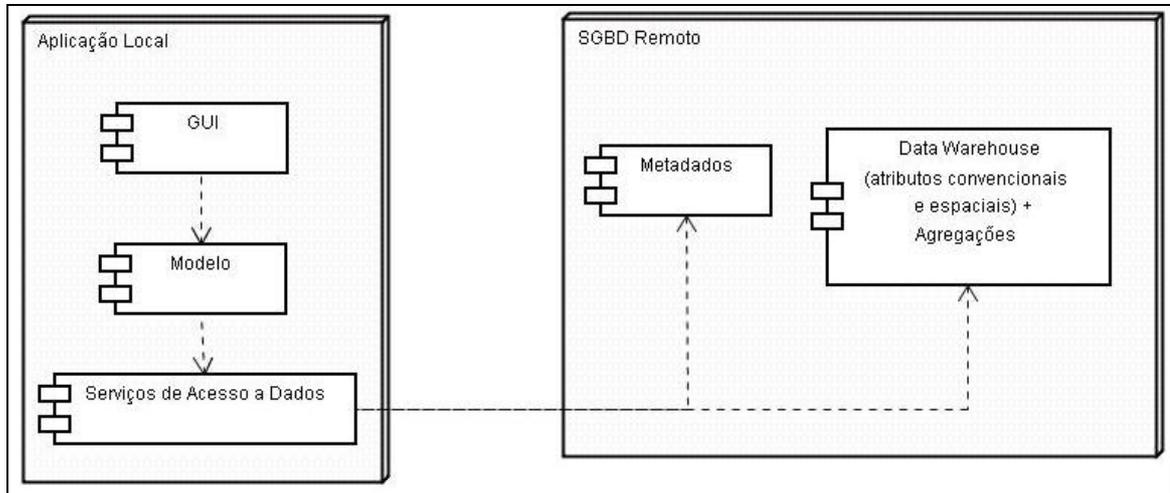


Figura 9 Arquitetura da Ferramenta PostGeoOlap
Fonte: (Manhães, 2007)

4 DESCRIÇÃO DO PROBLEMA: GESTÃO DE TRIBUTOS

Desde que o homem deixou sua vida de nômade nos tempos remotos e fixou residência, os valores materiais começaram a surgir. No início, sua função era apenas manter os líderes do grupo, que tinham suas necessidades cada vez mais aumentadas, e para atendê-las, os subordinados partiam para dominar as aldeias vizinhas e aplicar tributos sobre as mesmas. Os primeiros manuscritos que fazem referência à forma como os impostos eram cobrados na antiguidade, foram encontrados no ano 2350 a.C. Com a evolução das sociedades, o tributo passou a ter um papel social e não mais individual. A primeira forma de tributo, semelhante à que conhecemos atualmente, foi instituída na Grécia no ano 477 a.C.

No decorrer da história, muitos acontecimentos marcantes tiveram a participação dos tributos, como por exemplo, o financiamento das cruzadas e a “Revolução Francesa”, em 1789, que teve como um de seus pilares a revolta quanto à cobrança abusiva dos impostos.

No Brasil, a “Inconfidência Mineira”, em 1789, foi um manifesto contra a cobrança do “quinto” sobre as pedrarias. Em 1834, foram traçados limites e fundamentos de um Direito Tributário Nacional, do qual se originou o atual Código Tributário Nacional, em 1966.

Com o passar do tempo, o tributo foi sofrendo modificações até culminar num sistema composto por impostos, taxas e contribuições cobrados pelo Estado para desempenhar suas funções e com uma finalidade social nos governos democráticos. Aliás, a promoção de benfeitorias e prestação dos serviços coletivos seria impossível sem essa cobrança.

Este capítulo aborda a concepção do Tributo baseada no sistema tributário brasileiro, bem como suas fundamentações e características, ambos com o foco na esfera dos tributos municipais, que é o escopo deste trabalho.

4.1 O TRIBUTO

“No Brasil vigora a regra da liberdade de iniciativa na ordem econômica. A atividade econômica é entregue à iniciativa privada. A não ser nos casos especialmente previstos na Constituição, o exercício direto da atividade econômica só é permitido ao Estado quando necessário aos imperativos da segurança nacional, ou em face de relevante interesse coletivo, conforme definidos em lei (Constituição Federal, art. 173). Não é próprio do Estado, portanto, o exercício da atividade econômica, que é reservada ao setor privado, de onde o Estado obtém os recursos financeiros de que necessita. Diz-se que o Estado exercita apenas *atividade financeira*, como tal entendido o *conjunto de atos que o Estado pratica na obtenção, na gestão e na aplicação dos recursos financeiros de que necessita para atingir os seus fins.*” (MACHADO,1992).

Através deste pensamento chega-se à conclusão de que o tributo é a principal fonte de recursos (receitas) da qual o Estado se nutre para o cumprimento de suas atividades primordiais.

Apesar de não estar explícito, o tributo encontra-se presente no cotidiano. Seja numa simples compra, numa prestação de serviço ou numa movimentação financeira, sua presença sempre é caracterizada através de um percentual ou valor.

No Brasil, o conceito de tributo encontra-se definido no art. 3º do CTN (Código Tributário Nacional), “Art. 3º. Tributo é toda prestação pecuniária compulsória, em moeda ou cujo valor nela se possa exprimir, que não constitua sanção de ato ilícito, instituída em lei e cobrada mediante atividade administrativa plenamente vinculada”.

Ou seja, o tributo é um encargo financeiro, no qual o contribuinte entrega o dinheiro ao Estado através de um pagamento que deve ser em valor monetário direto e que não pode ser confundido com punição. Sua criação e modificação só poderão ocorrer mediante lei e sua cobrança só poderá ser realizada por uma autoridade administrativa competente que seguirá normas legais específicas.

Todo tributo, segundo a legislação tributária, deve possuir os seguintes elementos:

- **Lei:** é o ato que institui o tributo, que define o fato gerador, fixa a alíquota, determina as penalidades do não cumprimento e declara as hipóteses de exclusão (isenção).
- **Fato Gerador:** é o elemento chave da obrigação tributária, ou seja, é a situação definida em lei como necessária e suficiente para que sua cobrança ocorra.
- **Sujeito ativo:** é a quem deve ser pago.

- **Sujeito passivo:** é quem deve pagar.
- **Base de Cálculo:** é o montante sobre o qual será aplicada a alíquota para determinar o valor a ser cobrado.

Um bom exemplo é o IPTU (Imposto Predial e Territorial Urbano):

Fato gerador: a posse do imóvel.

Sujeito ativo: o município.

Sujeito passivo: o proprietário do imóvel.

Base de cálculo: o valor venal do imóvel (alíquotas previstas na legislação local).

Segundo o Art. 5. do CTN os tributos são os impostos, taxas e contribuição de melhorias. Nos impostos não há uma justificativa para sua cobrança, ou seja, seu único objetivo é a captação de recursos para o Estado. Sua principal característica é simplesmente a exigência do seu recolhimento por parte da comunidade aos cofres públicos. A definição de imposto que consta no CTN estabelece: “Art. 16. Imposto é o tributo cuja obrigação tem por fato gerador uma situação independente de qualquer atividade estatal específica, relativa ao contribuinte”.

Os impostos, quanto à sua base econômica, podem ser aplicados sobre o comércio exterior, sobre o patrimônio e renda, ou ainda, sobre a produção e circulação. Quanto à alíquota, podem ser fixas ou proporcionais quando variam de acordo com a base de cálculo. Quanto à forma de percepção, podem ser diretos quando incidem diretamente sobre o contribuinte ou indiretos quando são repassados. E por último, quanto ao objeto de incidência, podem ser reais, quando relacionados diretamente aos bens palpáveis; ou pessoais, quando incidem sobre a pessoa do contribuinte.

Com relação às taxas, sua concepção difere dos impostos no sentido de o contribuinte ter o conhecimento sobre o que, ou para que está pagando (serviço público). As taxas são caracterizadas pela sua utilização que pode ser *efetiva* quando o contribuinte usa de fato o serviço ou *potencial* quando o serviço fica a sua disposição; o **serviço específico** é outro aspecto, pois o contribuinte sabe exatamente pelo que está pagando, e por fim, o fato de ser **divisível**, ou seja, quando pode ser dividido entre os usuários. O CTN giza:

Art. 77. As taxas cobradas pela União, pelos Estados, pelo Distrito Federal ou pelos Municípios, no âmbito de suas respectivas atribuições, têm como fato gerador o exercício regular do poder de polícia, ou a utilização, efetiva ou potencial, de serviço público específico e divisível, prestado ao contribuinte ou posto à sua disposição.

E, finalmente, sobre a contribuição de melhoria que, como as taxas, também têm o motivo de sua cobrança claro e específico. O principal gerador deste tipo de tributo são as obras públicas, que conseqüentemente trazem benfeitorias (benefícios ou valorização). E nada mais justo que dividir a conta entre os beneficiados. O seu objetivo é cobrir os custos da obra entre os beneficiados e não captar mais recursos. E esse custo, por lei, não poderá ser superior à soma das valorizações adquiridas. O CTN dita:

Art. 81. A contribuição de melhoria cobrada pela União, pelos Estados, pelo Distrito Federal ou pelos Municípios, no âmbito de suas respectivas atribuições, é instituída para fazer face ao custo de obras públicas de que decorra valorização imobiliária, tendo como limite total a despesa realizada e como limite individual o acréscimo de valor que da obra resultar para cada imóvel beneficiado.

Os impostos pertencem à categoria dos tributos não-vinculados, uma vez que não há uma contrapartida, ou seja, o contribuinte simplesmente é obrigado a recolher. As taxas e contribuições de melhoria pertencem ao grupo dos tributos vinculados, pois neste caso, a contrapartida aparece, visto que o contribuinte pagou por um serviço solicitado ou foi beneficiado através de alguma obra pública e pagará proporcionalmente pelo benefício. A Tabela 17 expõe mais claramente essa relação de composição.

TRIBUTOS		
Não-vinculados: impostos diretos sobre renda, ganhos e patrimônio; e indiretos sobre o consumo de bens e serviços	Vinculados	
<ul style="list-style-type: none"> • Gerais • Restituíveis (empréstimos compulsórios) • Especiais ou finalísticos 	<ul style="list-style-type: none"> • Taxas <ul style="list-style-type: none"> ➤ de serviços ➤ de polícia 	<ul style="list-style-type: none"> • Contribuições <ul style="list-style-type: none"> ➤ de melhoria ➤ previdenciária

Tabela 17 – Quadro ilustrativo da formação do tributo (COELHO,2002)

No que concerne a sua competência, os tributos podem ser:

- **Federais:** imposto sobre importação (II), imposto sobre exportação (IE), imposto sobre a renda de proventos de qualquer natureza (IR), imposto sobre produtos industrializados (IPI), imposto sobre operações de crédito, câmbio, seguro ou relativas a títulos ou valores mobiliários (IOF), imposto sobre a propriedade territorial rural (ITR) e imposto sobre grandes fortunas (IGF)..

- **Estaduais:** imposto sobre a propriedade de veículos automotores (IPVA), imposto sobre operações relativas à circulação de mercadoria e sobre prestação de serviços de transporte interestadual e intermunicipal e de comunicações (ICMS) e imposto sobre transmissão “causa mortis” e doação de quaisquer bens ou direitos a eles relativos (ITD).
- **Municipais:** imposto sobre a propriedade predial e territorial urbana (IPTU), imposto sobre serviços de qualquer natureza (ISS) e imposto sobre a transmissão inter vivos de bens imóveis a título oneroso (ITBI).

Foram citados apenas os principais impostos, uma vez que o sistema tributário brasileiro é composto de 67 (sessenta e sete) tributos que estão discriminados no apêndice A desta pesquisa. Esta pesquisa aborda apenas os tributos da esfera municipal e com ênfase no IPTU e no ISS. No ano de 2003, de acordo com informações do Sistema de Contas Nacionais, do IBGE (IBGE,2005), os dois impostos foram responsáveis por 70% da arrecadação própria dos municípios brasileiros, sendo com 38%, para o ISS e 32%, para o IPTU.

As variação nas receitas tributárias fornecem indicadores valiosos que podem auxiliar na tomada de decisão do gestor público. Como exemplo, a variação de impostos prediais por faixa de valores pode indicar um crescimento de uma determinada região do município. Em outro caso, o índice de inadimplência num determinado bairro pode revelar a necessidade de mudanças na política de incentivos naquela região. Ou ainda, verificar se um projeto de lei, criado para conceder um desconto sobre determinados tributos de contribuintes de uma determinada faixa de renda, teve seu objetivo (retorno) alcançado através dos pagamentos efetuados num determinado período de tempo após a sua data da publicação. E esses tributos associados a dados geográficos podem prover mais uma forma de visualização, análise e interpretação, promovendo assim, mais um insumo para o gestor público. Outro aspecto positivo é possibilidade de criação de mapas sociais, uma vez que é possível a identificação de áreas onde residem contribuintes mais carentes.

4.2 CENÁRIO ATUAL: O SETOR TRIBUTÁRIO DA PREFEITURA DE CACHOEIRO DE ITAPEMIRIM

Este trabalho utiliza o setor tributário da Prefeitura Municipal de Cachoeiro de Itapemirim para desenvolver um SSD focado em tributos. A estrutura organizacional da PMCI encontra-se disposta conforme a Figura 10.

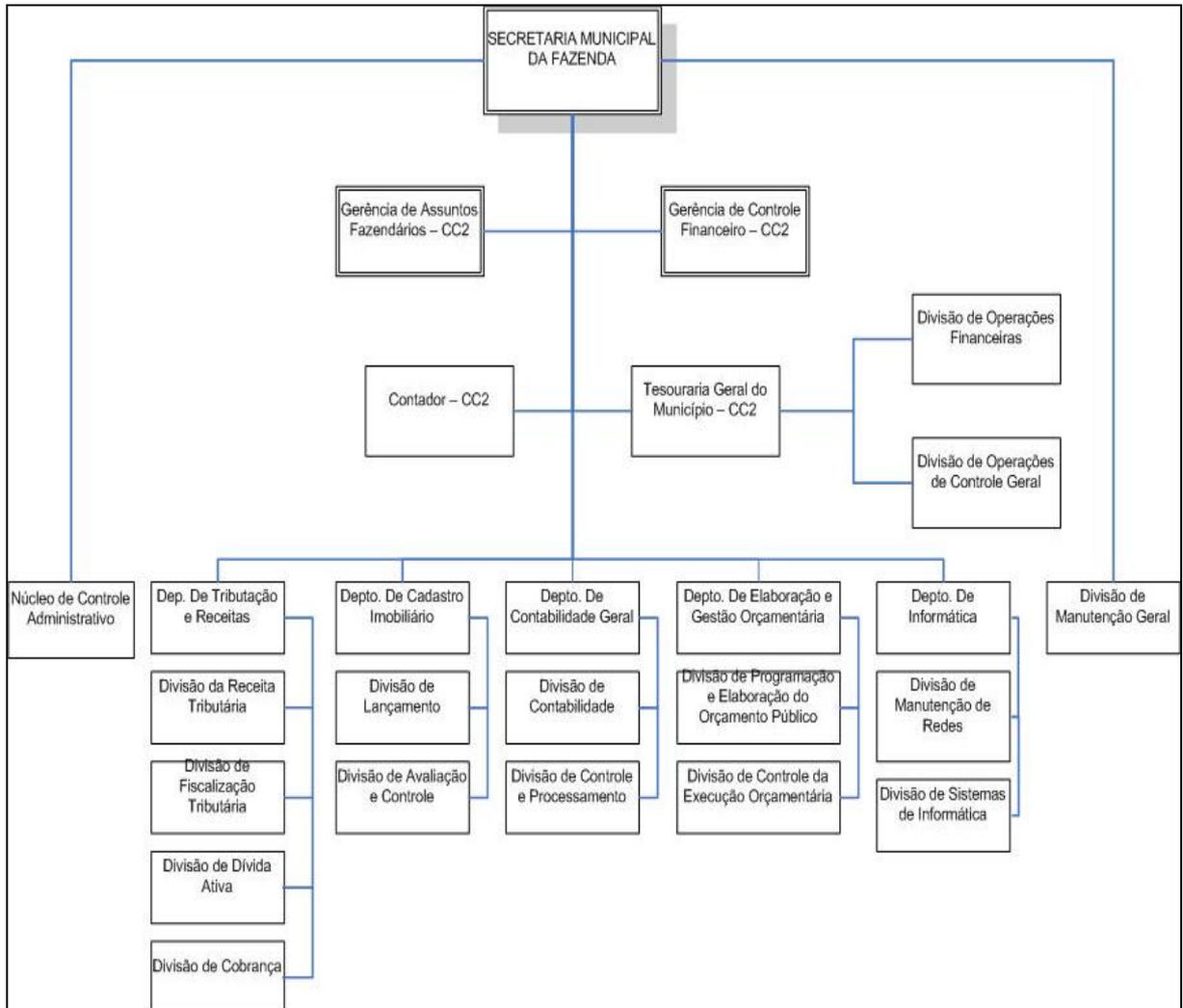


Figura 10 Estrutura Organizacional da Secretaria da Fazenda

A Secretaria da Fazenda é atendida por vários sistemas (transacionais) e que estão integrados; todos foram desenvolvidos pela DATACI (Empresa Processamento de Dados do Município de Cachoeiro de Itapemirim), uma empresa criada pela prefeitura em 1988 com o objetivo de fornecer este apoio tecnológico como:

- **Cadastro Único** – como o próprio nome já sugere, neste sistema encontram-se cadastrados todos os contribuintes inscritos na Prefeitura, tanto pessoa física quanto jurídica, bem como todos os dados a eles inerentes. Todos os sistemas acessam este cadastro único, evitando, assim, a existência de dados redundantes.
- **Cadastro Imobiliário** – responsável pelo armazenamento e manutenção de todos os imóveis do município, bem como todos os insumos necessários para a estipulação do valor venal que será utilizado como base no cálculo do tributo correspondente, neste caso, o IPTU (imposto predial e territorial urbano). Um outro sistema, denominado

SIGMUN (sistema de informações geográficas do município), foi adquirido pela Prefeitura. Este recebeu uma carga de dados oriunda do sistema transacional atual. Em seguida, um serviço de recadastramento de todo o município foi iniciado e as informações geográficas foram acrescidas.

- **Cadastro Mobiliário** – neste módulo ficam concentradas informações sobre os prestadores de serviços e comércio do município, ou seja, todas as empresas e profissionais autônomos. Todas as atividades comerciais exercidas no município estão no Código Tributário Municipal, cada qual com uma alíquota que será utilizada no cálculo do tributo. O imposto sobre a prestação de serviço pode ser cobrado de duas formas: a primeira, através de um valor fixo anual; a segunda, através de uma alíquota sobre seu faturamento mensal, que deve ser pago no mês posterior.
- **Fiscalização Tributária** – é utilizado pelo setor de fiscalização de rendas do município. Seu objetivo é fiscalizar e autuar, quando necessário, as empresas ou prestadores de serviços que deixaram de recolher seus tributos ou recolheram um valor menor que o devido.
- **Taxa Avulsa** – é utilizado para emitir taxas diversas que são cobradas pela prefeitura, como por exemplo, a taxa de expediente para algum serviço solicitado por parte do contribuinte.
- **Dívida Ativa** – quando todos os tributos gerados pelos demais sistemas deixam de ser pagos, e ultrapassam a data limite para sua quitação, estes são transferidos para o sistema de dívida ativa. De fato, não há um deslocamento do registro, e sim uma mudança de titularidade e de autoridade. Uma vez transferidos os tributos para dívida ativa, o carnê contendo o tributo que o contribuinte recebeu não será mais aceito pelos bancos para pagamento e o contribuinte deverá comparecer ao setor de dívida ativa para negociar a quitação dos tributos.
- **Débito Único** – todos os tributos, gerados pelos sistemas acima referenciados, quando calculados, são gerados neste sistema, bem como todos os pagamentos que forem sendo efetuados. Diariamente, os bancos enviam arquivos no formato Texto (TXT) contendo todos os pagamentos efetuados em suas agências no dia anterior. Estes arquivos são processados e abastecem o banco de dados com os créditos.

A Figura 11 apresenta uma visão conceitual de todos os sistemas de gestão tributária existentes na prefeitura através de um diagrama de caso de uso de negócios. Cada um dos sistemas aparece na forma de um pacote (*package*).

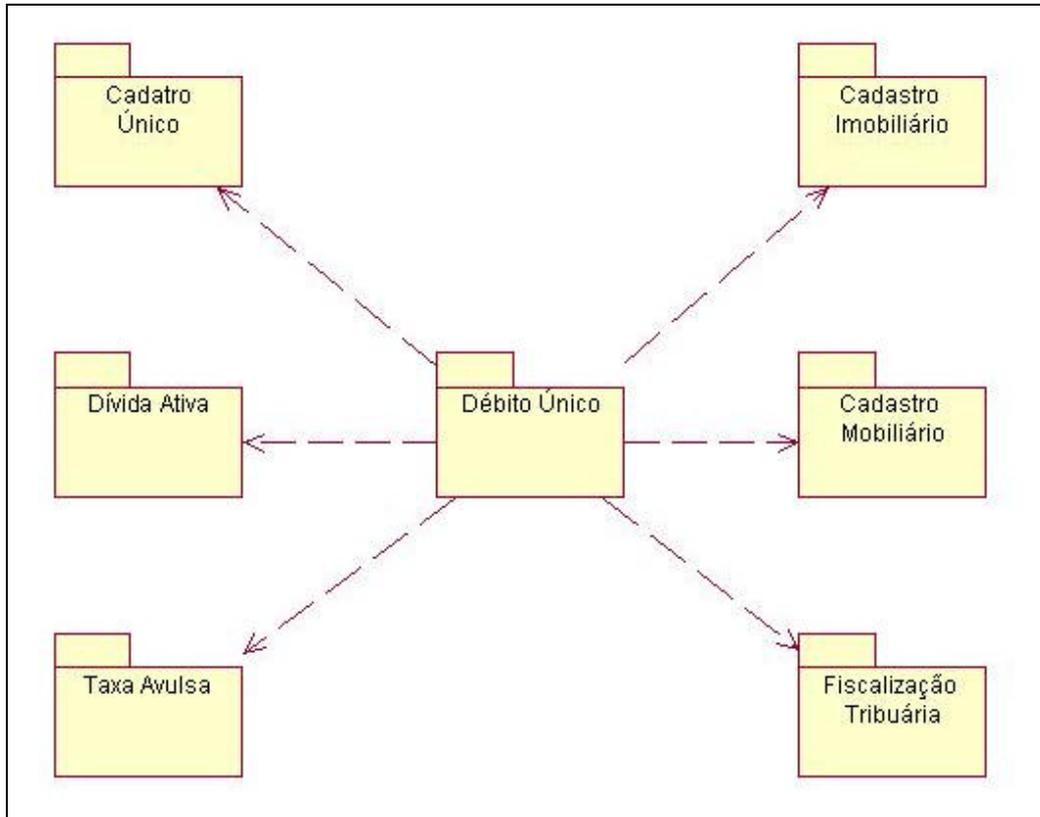


Figura 11 Diagrama de caso de uso de negócios utilizando dependência entre pacotes.

5 CONCEPÇÃO E ELABORAÇÃO DO DATA WAREHOUSE

O objetivo deste capítulo é expor e detalhar as fases de modelagem do SSD desde o levantamento de requisitos até a elaboração do DW para atender a gestão de tributos da Prefeitura Municipal de Cachoeiro de Itapemirim (PMCI).

5.1 DESCRIÇÃO DO ESTUDO DE CASO

A PMCI, como todas as prefeituras, é o órgão responsável pela manutenção da infraestrutura de uma cidade e necessita de recursos para cumprimento desse papel. Os recursos possuem duas origens principais: repasses vindos da União e Estado e os tributos arrecadados pelo município, conforme visto no capítulo 4. No âmbito municipal os recolhimentos podem ser provenientes da propriedade de bens imóveis (prediais e territoriais), da prestação de serviços (empresas e profissionais autônomos) e taxas recolhidas junto à Prefeitura para fins diversos como licença para demolição, inscrição para o funcionamento de empresas e outros.

No início do ano, são gerados aproximadamente 55.000 carnês de IPTU e aproximadamente 10.000 carnês de ISS. Diariamente, são gerados carnês relacionados a autos de infração, licenças e parcelamentos.

No decorrer do ano, pode ser aprovado um projeto de lei na câmara concedendo descontos para débitos de anos anteriores ou para uma determinada faixa de valores. Nesse caso, com a ferramenta PostGeoOlap, será possível, dentre outras vantagens, saber se essa lei teve o retorno esperado e em quais regiões ocorreram concentrações destes pagamentos.

Todos os documentos emitidos pela PMCI seguem o padrão FEBRABAN (Federação Brasileira de Bancos). Desta forma, 100% dos pagamentos são recebidos através dos bancos e transmitidos, posteriormente, via meio eletrônico, eliminando a intervenção humana dentro da

prefeitura para registrar os pagamentos e, com isso, minimizando a possibilidade de falhas. Os arquivos são recebidos diariamente via Internet no formato texto e processados.

5.2 LEVANTAMENTO DE REQUISITOS

O primeiro passo para a construção do modelo de DW no setor tributário foi a busca por patrocínio junto ao setor tático e estratégico da PMCI, pois é notório que a ausência do “*sponsor*”, ou seja, de um patrocinador para a idéia, constitui um dos fatores de maior risco para os projetos de DW. Inicialmente, foram identificados três setores responsáveis pela gerência dos tributos: a Direção do Departamento de Tributação e Receitas, a Direção do Departamento de Cadastro Imobiliário e a Chefia de Divisão da Fiscalização Tributária. Todos esses gestores possuem o conhecimento sobre os processos de negócio, exercendo essas funções há bastante tempo e o mais relevante, serão os futuros usuários da ferramenta.

Mesmo a prefeitura tendo muitas outras áreas de atuação (saúde, educação e segurança), o objeto deste trabalho está focado na administração dos tributos. Entretanto, esse ambiente (DW) será muito útil para todos os setores da PMCI, uma vez que a concentração dos dados tributários servirá como um repositório de informações e todas as demais áreas tendem a ser beneficiadas com uma gestão eficiente do tributo. Ademais, ações como a Lei de Responsabilidade Fiscal (Lei 101/00) norteia o setor público, cada vez mais, para a busca de ferramentas analíticas para auxiliar a gestão pública.

5.2.1 Entrevistas e questionário com usuários

Primeiramente, foram realizadas reuniões (não ultrapassando quarenta e cinco minutos) com cada usuário, a fim de esclarecer os objetivos, os benefícios do projeto de DW e o levantamento dos requisitos. O apoio foi incondicional, uma vez que o setor gerencial encontra-se carente de ferramentas para análise da informação e que auxiliem na tomada de decisão. A base de dados da PMCI possui registros a partir do ano de competência de 1989, o que indica a existência de um repositório considerável como fonte de abastecimento para o sistema.

Além das reuniões, foi aplicado um questionário com o intuito de coletar as informações mais discretas e documentar os anseios por parte dos entrevistados; para o preenchimento e devolução do questionário foi dado o prazo de uma semana. O questionário utilizado é composto de nove questões, encontra-se no Apêndice C e as respostas encontram-se abaixo na ordem dos usuários supra mencionados:

Pergunta 1 – Atualmente, de quais informações gerenciais, referentes a tributos, tem-se maior necessidade e conseqüentemente maior dificuldade de obtê-las?

Usuário 1 – “Relatórios gerenciais com informações de valores recolhidos, sendo que este valor tem de ser igual ao valor registrado na contabilidade (ex: ISS recolhido no exercício de 2004, por mês e atividade). Valores lançados e recolhidos de um tributo em determinado período (ex: qual o valor do ISS de autônomo lançado e recolhido em 2003). Quantidade de contribuintes cadastrados por atividade que recolheram imposto em 2003 (ex: contribuintes na atividade de manutenção de veículos e que recolheram ISS em 2003). Os relatórios acima, hoje, são feitos pela Dataci”.

Usuário 2 – “Atualmente, a maior necessidade do departamento é a espacialização do débito e pagamento do IPTU no município”.

Usuário 3 – “Firmas fiscalizadas por período levando em conta a última fiscalização na empresa. Relatório das empresas nunca fiscalizadas”.

Pergunta 2 – Que tipo de resumos, ou totais, seriam de maior utilidade? Qual a melhor forma de visualizá-los?

Usuário 1 – “Seria interessante que a ferramenta pudesse dar a opção de controlar todas as informações gravadas no sistema, de modo que o usuário informasse quais dados seriam incluídos. A forma de visualização poderia ser feita conforme o usuário definisse: Tabela, gráfico, texto, etc”.

Usuário 2 – “Totais por itens e valores do débito e pagamento por cores. Isenção de tributos por setores e/ou bairros. Imóveis sem georeferência”.

Usuário 3 – “Recolhimento mensal: por faixa de recolhimento, por ordem alfabética. Imóveis avaliados mensalmente. Ao consultar um imóvel para avaliação, faz-se necessária uma crítica se este já foi avaliado e em que período”.

Pergunta 3 – Qual o nível de detalhamento do tributo que você considera importante do ponto de vista gerencial? () por itens que compõem o débito () por débito

Usuário 1 – “Por itens que compõem o débito”.

Usuário 2 – “Por itens que compõem o débito”.

Usuário 3 – “Por itens que compõem o débito”.

Pergunta 4 – Com relação a pagamento de tributos, qual a forma de agrupamento que você considera mais relevante? () por débito () diário () semanal () mensal

Usuário 1 – “Por débito”

Usuário 2 – “Por débito”

Usuário 3 – “Mensal”

Pergunta 5 – **Como são tratadas (tributação) as filiais das empresas?**

Usuário 1 – “Cada empresa tem que ter a sua inscrição própria e pagar seus tributos individualmente”.

Usuário 2 – Em branco

Usuário 3 – “Como contribuinte único, com obrigações distintas para cada uma: matriz e filial”.

Pergunta 6 – **Com relação a anos anteriores, qual a melhor forma de visualização dos dados? Ou seja, que tipo de agrupamento atenderia às necessidades gerenciais?**

Usuário 1 – “Entendo que o tipo de pesquisa deve ser aberto a diversas formas, tais como: por contribuinte, por número do débito, por documento, por inscrição municipal e por tipo de tributo”.

Usuário 2 – “Comparativos de inadimplência e comparativos de acréscimo de valores”.

Usuário 3 – “Por atividade; por endereço; por faixa de recolhimento; por período e por fiscalização no período”.

Pergunta 7 – **Qual a importância de vincular informações tributárias com informações geográficas?**

Usuário 1 – “Todo administrador precisa de informações precisas para tomar decisões, por isso, tudo que se refere ao tributo tem de ter sua localização identificada. Por exemplo, recentemente, Vitória concedeu redução da alíquota do ISS para as empresas localizadas no centro da cidade”.

Usuário 2 – “A importância maior está na análise das informações de forma espacializada, facilitando a tomada de decisões quanto a valores de venda, isenções, legalidade de lançamentos, intervenção para aumentar a quantidade de pagamentos, etc.”.

Usuário 3 – “Verificar a atividade predominante na região. Direcionar a fiscalização. Atribuir valor venal de forma justa para fins de cálculo do ITBI”.

Pergunta 8 – **Nesse caso, objeto do tributo seria tratado como um “ponto” ou “polígono”?**

Usuário 1 – “Entendo que a informação tem que ser armazenada da forma mais precisa possível”.

Usuário 2 – “Ponto, porque polígonos são utilizados quando há necessidade de cálculo de áreas, o que não é o caso da ferramenta proposta”.

Usuário 3 – “Como um ponto”.

Pergunta 9 – **Há uma necessidade de análise do endereço do contribuinte do ponto de vista espacial?**

Usuário 1 – “Entendo que se possível, quanto melhor a informação a ser fornecida, melhor será a avaliação do trabalho prestado”.

Usuário 2 – “Sim. A PMCI sempre necessita localizar o contribuinte e poder espacializar esta informação acelera todo o processo de localização”.

Usuário 3 – “Sim. Quando envolve cálculos para o imóvel”.

Observações:

Usuário 1 – “A maior dificuldade do usuário é que na maioria das vezes ele sabe na prática o que quer, mas não tem o conhecimento necessário da parte de informática”.

Usuário 2 – “Em branco”.

Usuário 3 – “Considerando que com cinco anos ocorre a prescrição para cobrança do ISS, faz-se necessário criar relatórios que “avise” esta ocorrência. As sugestões do item I atendem a esta necessidade”.

5.2.2 Engenharia reversa dos bancos de dados transacionais

Em seguida, foi iniciada a etapa de levantamento dos modelos adotados nos sistemas transacionais a fim de compreender toda a estrutura empregada no setor tributário. Entretanto, alguns sistemas simplesmente não possuíam documentação nem um modelo de Entidade e Relacionamento (E-R), possuindo apenas o dicionário de dados.

Dessa forma, foi utilizado o recurso da engenharia reversa, ou seja, uma técnica que permite a construção de modelos lógico e físico a partir de uma base de dados já existente ou de arquivos no padrão SQL (*Structured Query Language*) com comandos de criação das tabelas do banco de dados, permitindo assim, a documentação de estruturas, chaves primárias (*primary key*) e chaves estrangeiras (*foreign key*). Foi utilizado o software ERwin devido a sua versatilidade entre os modelos lógico e físico, além da capacidade de geração automática dos comandos de definição de estrutura dos dados (*Data Definition Language - DDL*).

No início desta pesquisa, o ambiente (banco de dados e aplicações) existente na PMCI utilizava o *software ZIM (ZIM Corporation)*, um ambiente de desenvolvimento de sistemas composto por um banco de dados e uma linguagem de programação (compatível com a linguagem SQL), ambos baseados no modelo “Entidade-Relacionamento” e sendo executado no sistema operacional Unix HP-UX. Em outubro de 2004, foi realizada uma conversão dos dados para o banco de dados em Oracle, sobre o mesmo sistema operacional (UNIX) e as aplicações

foram mantidas no ambiente ZIM, porém convertidas para uma versão gráfica da ferramenta, conforme a Figura 12.

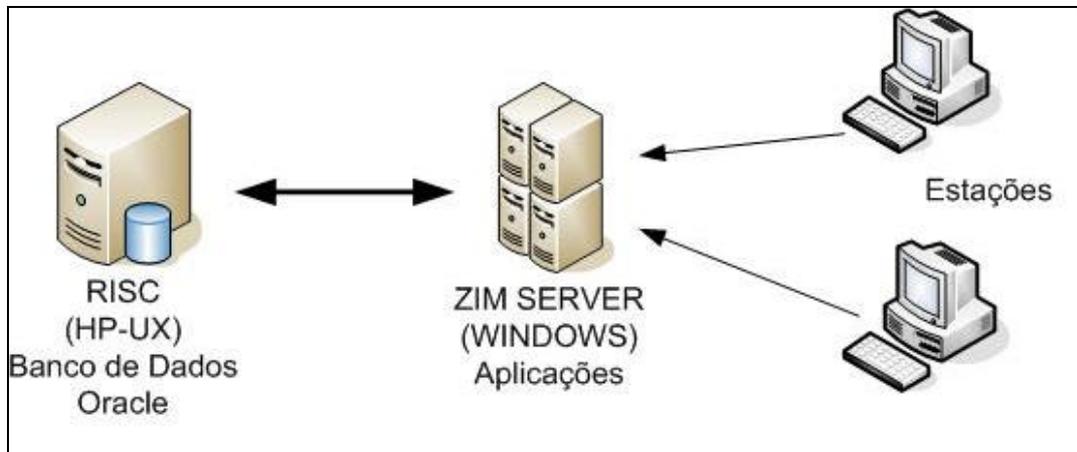


Figura 12 Plataforma utilizada na PMCI

5.3 ESQUEMA ESTRELA DO SISTEMA DE GESTÃO DE TRIBUTOS DA PMCI

A elaboração do DW seguiu os passos segundo KIMBALL (1998), no qual o primeiro deles é a definição do processo de negócio, que neste caso foi a gestão de tributos municipais. É importante ressaltar que o cumprimento dessa etapa é de vital importância para o sucesso do projeto, pois nela são definidos os objetivos e o escopo, ou seja, o que será tratado.

O segundo passo é o estabelecimento do nível de granularidade a ser utilizado, ou seja, o nível de detalhamento que as informações deverão atingir. No terceiro passo são identificadas as tabelas dimensão, contendo a parte descritiva do modelo e no quarto passo são especificadas as tabelas fato, responsáveis pelas características métricas do modelo.

Baseado nos levantamentos e nos modelos físicos dos sistemas transacionais vigentes foi proposto um esquema estrela conforme o diagrama de classes exibido na Figura 13. Por questões de espaço e para não prejudicar a legibilidade, foram citados apenas os principais atributos, ficando os demais para serem citados nos próximos tópicos, onde cada tabela será detalhada.

A generalização utilizada na classe *Origem do Debito* foi utilizada para destacar as origens dos tributos que pode variar entre um Imóvel, uma Empresa ou uma solicitação de algum serviço para a municipalidade. Também pelo fato de ser a única classe que contém atributos geográficos.

Outro ponto de destaque no esquema projetado é a existência de duas tabelas fato, mas com algumas tabelas dimensão em comum. Tal coexistência ocorreu devido à diferença no contexto, uma vez que os pagamentos de um tributo (débito) não são distribuídos entre os itens (naturezas) que compõem o débito, mas referenciados diretamente ao valor total.

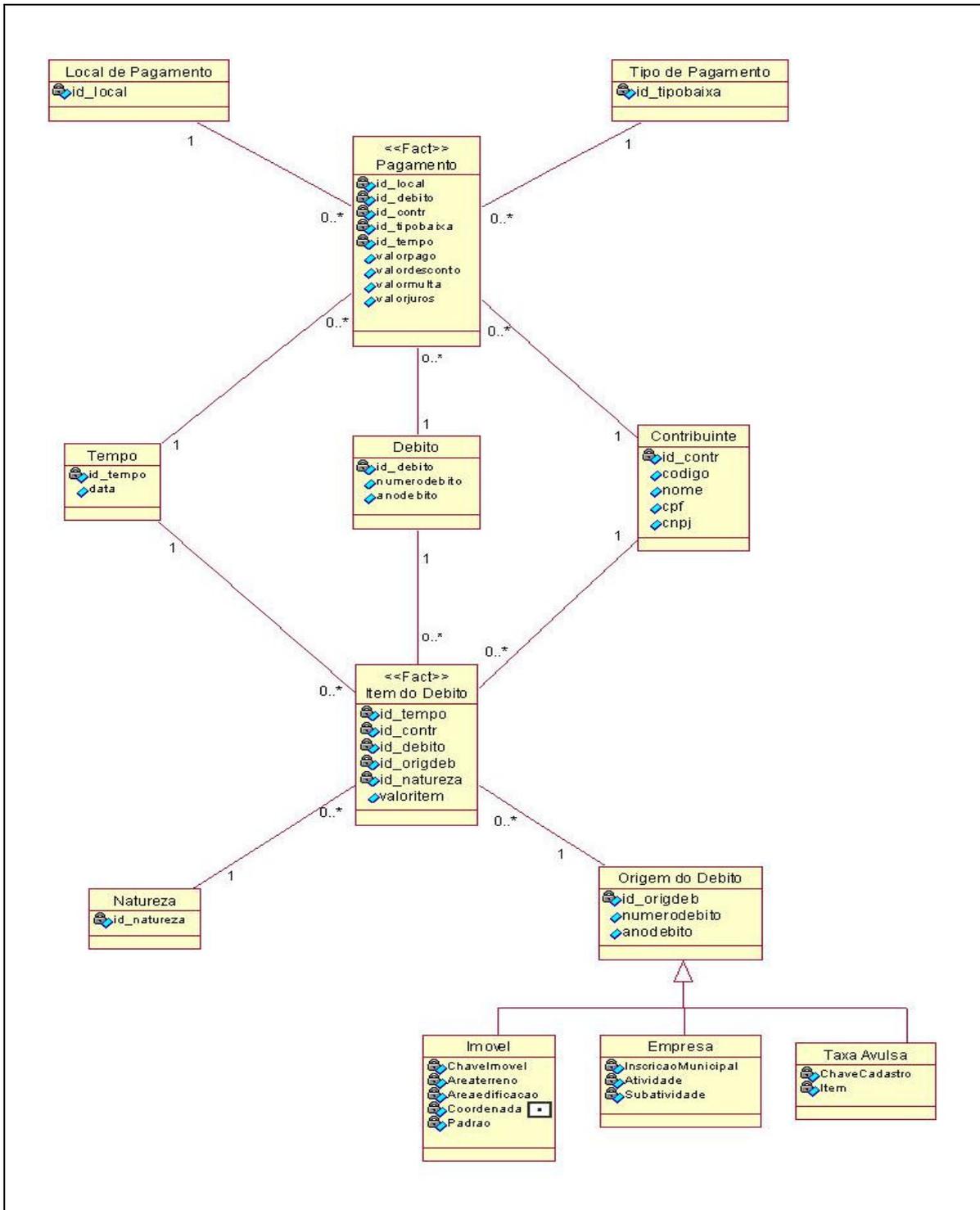


Figura 13 Diagrama de Classes do DW da Prefeitura Municipal

5.3.1 Tabelas Dimensão

As tabelas dimensão, dispostas através das classes *tempo*, *debito*, *contribuinte*, *natureza*, *origem do débito*, *local* e *tipo de pagamento*, reúnem a parte descritiva do esquema.

5.3.1.1 Dimensão Tempo

Como em todo DW, a existência da tabela dimensão Tempo é fundamental, pois a sua finalidade é a de registrar o momento em que os dados foram inseridos ou alterados (aspecto temporal). No esquema proposto acima, a tabela Tempo (Figura 14) mantém um relacionamento com as duas tabelas fato existentes no esquema: Item do debito e Pagamento. Com relação ao fato Item do Debito, seu objetivo é a marcação do tempo em que o débito foi concebido através da data do cálculo do tributo. Já a sua vinculação com o fato Pagamento registra o momento em que o contribuinte efetua o crédito da quantia referente ao valor calculado. A sua cardinalidade com ambas as tabelas é do tipo um para muitos, em virtude da ocorrência de vários cálculos e pagamentos numa determinada data, ao mesmo tempo em que ambos possuem uma única data. A massa de dados utilizada no sistema reuniu a movimentação tributária dos últimos quinze anos da prefeitura, ou seja, todas as datas de cálculos e pagamentos.

Além de utilizar uma chave artificial como única (*id_tempo*), a tabela não possui muitos atributos (*data*, *dia*, *mês*, *ano*, *bimestre*, *trimestre*, *quadrimestre* e *semestre*), mas estes encontram-se uma estrutura hierárquica (item 2.2.7) bem definida. A decomposição desta através dos demais atributos foi implementado pois facilitará a identificação, por exemplo, do aumento da arrecadação num determinado trimestre. Os atributos “*diasemana*” e “*mesano*” armazenarão a descrição do dia e mês respectivamente. O atributo denominado “*feriado*” será utilizado como um sinalizador para marcar as datas especiais, pois um contribuinte pode perfeitamente procurar um terminal de auto-atendimento bancário num feriado para efetuar seus pagamentos.



Figura 14 Dimensão Tempo

5.3.1.2 Dimensão Contribuinte

Esta tabela registra a quem pertence os tributos. Bem como seus pagamentos, todos os contribuintes da PMCI estão reunidos num único cadastro, que é acessado por todos os demais sistemas transacionais. Isto facilitou muito o processo de carga do DW e proporcionou um cadastro com o mínimo de redundância, uma vez que não há incidência de ilhas de dados. Esta tabela (Figura 15) é o resultado da união de muitas tabelas do sistema transacional específicas do cadastro único. O atributo “*tipo*”, com domínio de um a seis, revela uma classificação dos contribuintes descrita no atributo “*tipodesc*” em pessoa jurídica local (1), pessoa física local (2), pessoa jurídica externa (3), pessoa física externa (4), pessoa jurídica com localização externa (5) e pessoa física com localização externa (6). A diferença entre externa e com localização externa encontra-se no endereço, pois no primeiro o endereço veio a partir da digitação de um “*cep*”, enquanto no segundo tipo foi digitado livremente, sem verificações. Sua cardinalidade também é do tipo “1 para muitos” uma vez que um contribuinte pode ter vários débitos e pagamentos.



Figura 15 Dimensão Contribuinte

5.3.1.3 Dimensão Natureza

Todo imposto possui no mínimo uma natureza (Figura 16) ou várias que o identificam e que são utilizadas para o controle da origem de sua emissão e distribuição no plano de contas da contabilidade. Este fato ocorre, pois cada receita deve ter um destino específico, por exemplo, o imposto predial possui uma natureza distinta do imposto territorial. Ambos são considerados do grupo “IPTU”, mas são direcionados para contas separadas. Dessa forma, é possível executar consultas analíticas envolvendo apenas terrenos ou edificações separadamente. Sua cardinalidade também é do tipo “1 para muitos”, uma vez que vários débitos podem ter uma determinada natureza em sua composição. Apesar de poucos atributos sua importância é altamente relevante no contexto tributário.

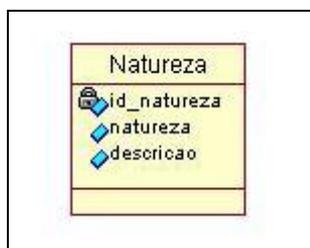


Figura 16 Dimensão Natureza

5.3.1.4 Dimensão Débito

Essa dimensão concentra todos os atributos descritivos com relação ao débito propriamente dito, conforme Figura 17. O campo “*numerodebito*” é um número único vindo dos sistemas transacionais. O “*anodebito*” indica o exercício em que o mesmo foi gerado. O “*valortotal*” representa o montante do débito. O atributo “*parcelas*” representa a quantidade de vezes em que o débito poderá ser pago. Um débito pode possuir o valor da primeira parcela maior que as demais; devido a este fato, foram transportados os atributos “*valor1parcela*” e “*valorparcela*”. O atributo “*data1venc*”, contendo a data do primeiro vencimento, servirá para o cálculo dos vencimentos das outras parcelas. O atributo “*tipovencimento*” serve como referência para não deixar que uma data de vencimento coincida como um fim de semana (1 – antes / 2 – depois). A “*situação*” registra o estado atual do débito (0 – não pago / 1 – parcelado no setor de origem / 2 – quitado / 3 – carnê de dívida ativa / 4 – parcelado em dívida ativa / 5 – anistiado / 6 – cálculo diferenciado / 7 – carnê de tributação / 8 – parcelado em tributação de receitas e 9 – cancelado). O “*setor*” aponta a localização atual do débito (DA – dívida ativa / CI – cadastro imobiliário / TB – tributação e receitas / TA – taxas avulsas e FT – fiscalização tributária). O “*tipo*” representa o tipo de chave que estará presente (0 – vazio pois é da dívida ativa, 1 – inscrição municipal, 2/3/4 – inscrição imobiliária, 5 – código único do contribuinte e 6 – autos de infração). O campo “*siglaorgao*” indicará o órgão no qual o débito foi originado; assim como a coluna “*usuário*” armazena quem gravou o débito. O atributo “*data lanc divat*” estabelece o dia em que o débito será transferido para dívida ativa (caso não sofra quitação), por exemplo, um débito de IPTU localizado no setor “CI” e que não tenha sido pago até a data estabelecida terá seu setor alterado de “CI” para “DA”. O atributo “*numeroautoinfr*” será utilizado de forma restrita pela fiscalização de rendas, pois este armazenará o número do auto de infração. Os atributos “*numerocarne*” e “*anocarne*” são preenchidos quando ocorre um parcelamento em dívida ativa. Neste caso, o contribuinte escolhe os débitos que deseja parcelar. Um novo débito será criado com o somatório dos débitos escolhidos, e este novo número será gravado no campo (*numerocarne*) dos débitos selecionados.



Figura 17 Dimensão Débito

5.3.1.5 Dimensão Origem do Débito

Assim como os contribuintes estão numa única base, os débitos também estão centralizados numa base única e no seu processo de carga para o DW todas as características foram unificadas, de forma que os atributos exclusivos de imóveis sejam preenchidos quando o débito tiver sua origem no IPTU. Os atributos referentes a empresas e profissionais autônomos serão preenchidos quando sua origem estiver no ISS. Esta dimensão funciona como uma complementação da dimensão “Débito”, vista anteriormente.

Os atributos “*numerodebito*”, “*anodebito*” e “*tipo*” possuem a mesma função que foi mencionada na dimensão supra citada. O campo “*chavedebito*” conterà a inscrição imobiliária ou a inscrição municipal, dependendo da procedência do débito. As colunas “*atividade*”, “*descatividade*”, “*subatividade*”, “*descsubatividade*” e “*aliquota*” somente estarão preenchidas quando o tributo tiver sua procedência do ISS. Paralelamente, o atributo “*coordenada*” armazenará uma coordenada espacial, referente a um ponto do imóvel, para que seja possível a sua localização geográfica. De acordo com a Figura 18, foi utilizada uma generalização justamente para caracterizar as particularidades de cada tipo de tributo. Sua cardinalidade é do

tipo “1 para muitos”, pois um imóvel possui débitos em vários exercícios, assim como as empresas.

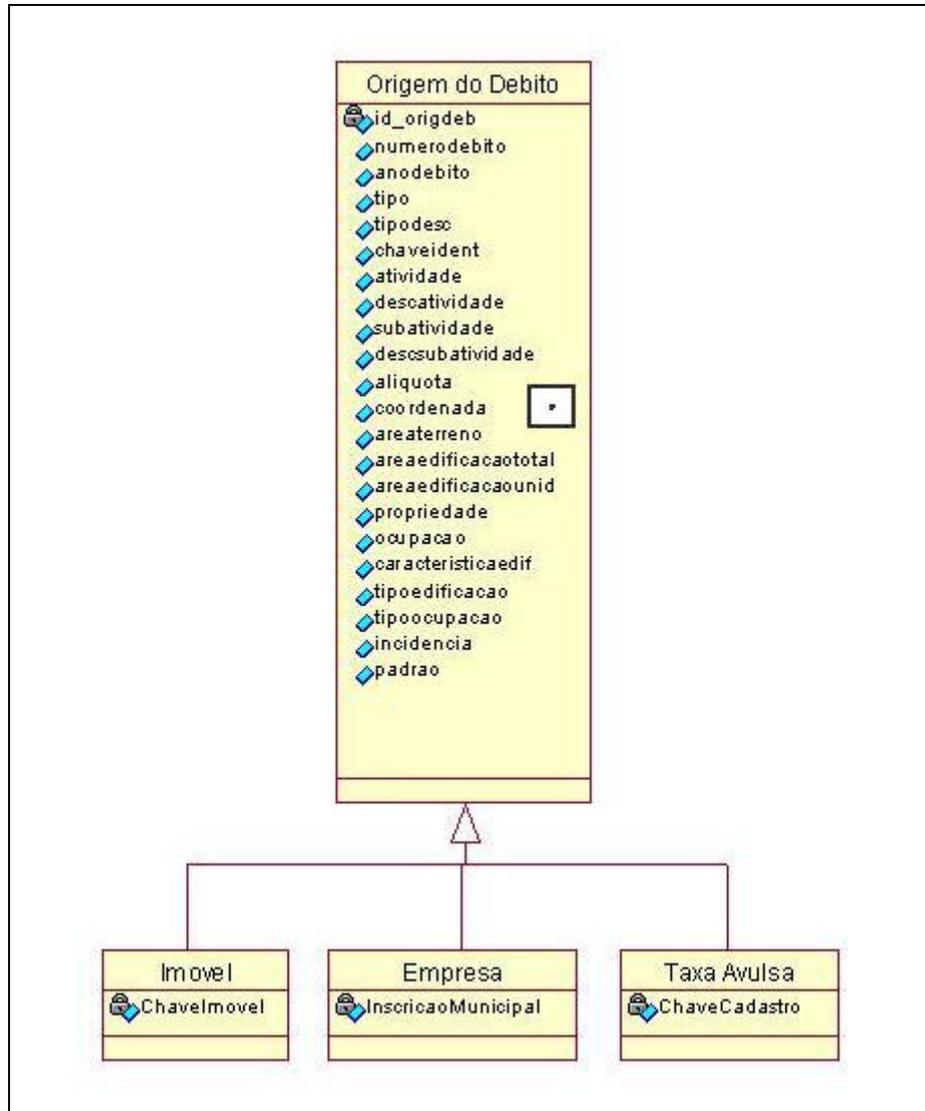


Figura 18 Dimensão Origem do Débito

5.3.1.6 Dimensão Local de Pagamento

Esta mantém informações sobre os bancos (Figura 19) que recebem os pagamentos dos tributos. Conforme já foi mencionado, eles são responsáveis pelo envio diário de arquivos contendo a arrecadação bancária que será processada e armazenada no banco de dados da PMCI. Esta dimensão permite a análise da preferência dos contribuintes, com relação ao local

de pagamento, no momento do recolhimento do tributo. Dessa forma, auxiliando numa melhor negociação no momento da criação ou renovação de convênios bancários.

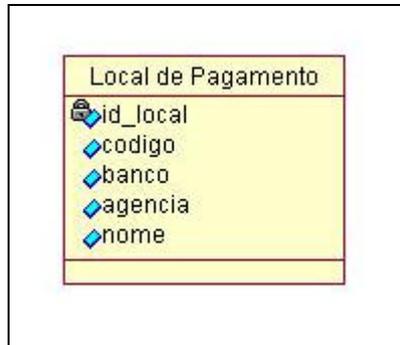


Figura 19 Dimensão Local de Pagamento

5.3.1.7 Dimensão Tipo de Pagamento

Esta dimensão possui informações macros a respeito dos pagamentos (Figura 20), ou seja, informações que caracterizam a transação bancária. Novamente, os atributos “*numerodebito*” e “*anodebito*” aparecem como identificadores do tributo. O campo “*parcelaseq*” complementa a coluna “*parcela*”, pois um contribuinte pode pagar uma parcela mais de uma vez por engano, gerando um segundo seqüencial. A “*datacredito*” geralmente é dois dias úteis após a data de pagamento. O “*valorcalculado*” constitui o valor fixado até o vencimento (campo “*datavencimento*”). Todo pagamento ou abatimento efetuado tem um código que esclarece se o valor foi recebido ou foi um acerto realizado no débito; por exemplo, um contribuinte que comprove o erro na constituição do seu débito pode entrar com um processo e, uma vez deferido, sofrerá uma redução no valor. Esta redução tem que ficar documentada e para este fim será utilizado um código de baixa específico, que será diferente do código utilizado para registrar o pagamento em dinheiro realizado nos bancos. Outro exemplo a ressaltar é o caso de anistias; se for votado um decreto que quite débitos abaixo de R\$ 250,00, por exemplo, uma baixa deverá ser criada para justificar a quitação, utilizando um código específico para cancelamentos. Sua cardinalidade é de “1 para muitos”, uma vez que vários pagamentos são comuns para um débito.



Figura 20 Dimensão Tipo de Pagamento

5.3.2 Tabelas Fato

O modelo proposto possui duas tabelas fato. Uma representando a composição do tributo (Item do Débito) e a outra representando os lançamentos dos créditos realizados (Pagamentos). A diferença de granularidade entre essas duas tabelas reside no fato de que a primeira informa a composição do tributo, enquanto a segunda registra os valores pagos com relação ao total do tributo, independente de sua composição.

5.3.2.1 Fato Item do Débito

Um débito pode ser composto por mais de um item. Por exemplo, um débito proveniente do IPTU pode ter em sua composição um valor referente ao “imposto predial” ou “territorial” (dependendo se for casa ou terreno, respectivamente), um valor relativo à “taxa de conservação de vias e logradouros” e um valor pertencente à “taxa de expediente”. Cada um proveniente de uma natureza distinta, mas todos formando um único débito. E de acordo com o levantamento realizado junto aos usuários, através de entrevistas e questionários, a necessidade gerencial seria no âmbito dos itens que compõem o débito.

A composição da tabela, além das chaves artificiais oriundas das dimensões como “*id_tempo*”, “*id_contribuinte*”, “*id_debito*”, “*id_origemdebito*” e “*id_natureza*”, possui o atributo “*valoritem*” que tem o valor de cada item que compõe o débito, conforme Figura 21.



Figura 21 Fato Item do Débito

5.3.2.2 Fato Pagamento

Esta tabela (Figura 22) fato reúne todo tipo de pagamento ou acerto ocorrido com um débito. Pode ser considerada uma complementação da dimensão “*Tipo de Pagamento*”, pois contém os campos que podem ser sumarizados. Além das chaves artificiais que a compõem, possui o valor efetivamente pago (*valorpago*), descontos e juros serão armazenados separadamente.

No que tange ao pagamento dos tributos, sua quitação, independentemente da quantidade de parcelas, ocorre com relação ao total do débito, e não aos itens que o formaram, o que resultou na existência de duas tabelas fato no modelo.



Figura 22 Fato Pagamento

5.4 PROCESSO DE ETL

A extração dos dados para carga no DW foi realizada da seguinte forma: primeiramente, foi elaborado um programa utilizando a linguagem oferecida pelo SGBDR ZIM (disposto no

Apêndice C deste trabalho) para acessar o banco de dados em Oracle e gerar os arquivos do tipo texto com as colunas necessárias para serem adicionadas nas tabelas fato e dimensão.

Em seguida, foi utilizada a ferramenta *Dbtools*, versão gratuita, para adicionar os dados no SGBD PostgreSQL. As coordenadas geográficas, que se encontravam num banco de dados separado da base da PMCI, tiveram de ser inseridas de uma maneira distinta, pois como o tipo de dado criado do lado do DW era do tipo geométrico e a ferramenta de conversão não dispunha de tal tipo, foi necessário transformar o arquivo do tipo texto em um arquivo contendo o comando DML:

```
(insert into d_coordenada (codigogeo, coordenada) values
('0100050094000500010101', GeomFromText('POINT(272330.7517670008
8 7694070.4773363546)', -1));)
```

Sendo que “*d_coordenada*” é uma tabela temporária criada para armazenar os dados do tipo geométrico e a inscrição do imóvel, que em seguida foram inseridos na tabela definitiva (Dimensão Origem do débito).

Após a retirada de todas as inconsistências e ajustes, feitos através da verificação de campos com conteúdo nulo ou inválido, foi possível quantificar o número de registros válidos inseridos conforme a Tabela 18.

Tipo da Tabela	Nome da Tabela	Quantidade de Registros (Total)	Quantidade de Registros (Lançados)
Dimensão	Tempo	2.664	2.664
Dimensão	Contribuinte	53.296	53.296
Dimensão	Débito	678.588	390.559
Dimensão	Natureza	95	83
Dimensão	Origem do Débito	678.588	390.559
Dimensão	Local	18	6
Dimensão	Tipo de Baixa	677.385	403.276
Fato	Item do Débito	1.645.063	925.521
Fato	Pagamento	677.385	403.276

Tabela 18 - Total de registros inseridos

O processo de carga no banco de dados foi realizado uma única vez, visto que o escopo desta dissertação está no emprego da visão analítica dos dados e não no seu processo de extração e carga.

Apesar do apoio, o ambiente foi montado fora das instalações da PMCI porque não havia equipamento disponível. Além disso, pelo fato de ser um protótipo e fazer parte de um

trabalho acadêmico, ou seja, não poderia ser implementado em horário de expediente, optou-se pelo ambiente externo.

6 CONSTRUÇÃO DO DATAWAREHOUSE USANDO O POSTGEOOLAP

Neste capítulo é apresentada a aplicação da ferramenta PostGeoOlap, integrada a um DW, abordando desde a preparação dos dados até a execução de consultas analíticas. Dessa forma, promover um sistema de suporte à decisão voltado para a área de tributos municipais.

Para a implementação do ambiente foi utilizado um computador com o processador Intel® Pentium® IV 3.00 GHz, com 1 GByte de memória RAM, um HD (*Hard Disk*) de 120 GBytes e um HD de 200 GBytes de armazenamento, utilizando o sistema operacional Windows XP Professional. Com relação aos softwares adotados, foi utilizado o pacote PostgreSQL 8.0.3 (composto pelo SGBD + pgAdminIII 1.2.1 + PostGIS 1.0) e o software PostGeoOlap versão 1.0. Um arquivo do tipo Arc View Shape File, com a imagem georeferenciada, foi usado como o mapa da cidade.

6.1 UTILIZAÇÃO DA FERRAMENTA POSTGEOOLAP

O primeiro passo para a utilização da ferramenta é conectar-se ao banco de dados, o que é realizado através da seleção do “Esquema” (Figura 23). Nele estão contidos todos os dados (*nome do banco, usuário, senha, porta para acesso*) necessários para realizar a conexão junto ao banco.

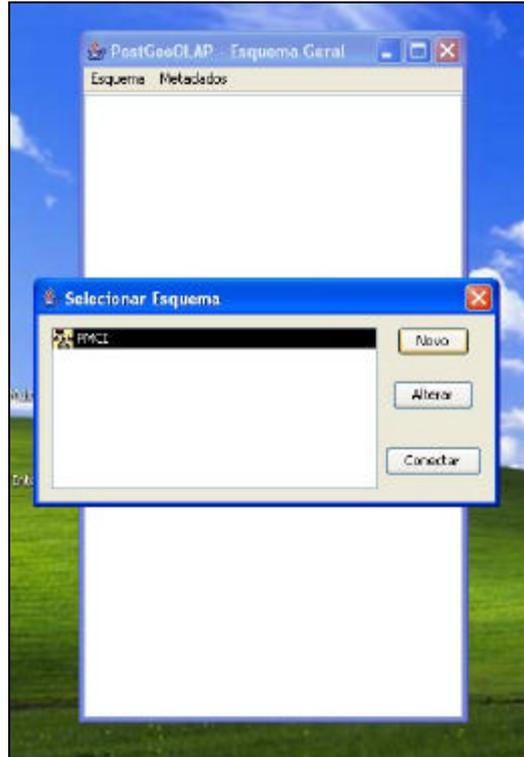


Figura 23 Seleção do Esquema para Conexão

A etapa seguinte é a criação do cubo onde serão realizadas as consultas analíticas, conforme a Figura 24. A tela, além de exibir o nome do esquema escolhido, possui um campo para receber o nome para o cubo e um parâmetro com o número mínimo de linhas para a geração de agregações (parâmetro para verificar o custo da operação).

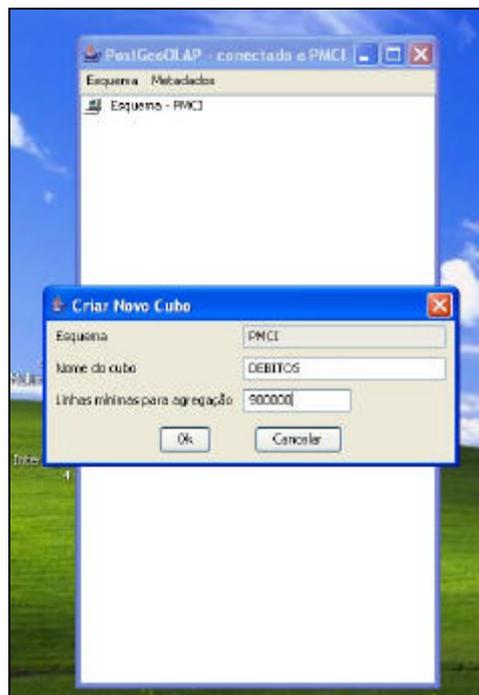


Figura 24 Criação do Cubo

Após a criação do cubo, o programa exibe uma janela (Figura 25) com todas as tabelas do banco, na qual irá selecionou-se a tabela fato e, por conseguinte, a seleção dos atributos dessa tabela que serão utilizados em operações como contagem, soma, média, máximo e mínimo (Figura 26).



Figura 25 Seleção da Tabela Fato

Uma vez selecionada a tabela fato, o programa exibe novamente todas as tabelas contidas no banco, porém, dessa vez para a identificação das tabelas dimensão. Para cada uma das tabelas, serão exibidos todos os seus atributos, que por sua vez, receberão um número (entre 1 e 9) que determinará seu posicionamento hierárquico dentro da dimensão.

Segundo o princípio de funcionamento da ferramenta, os atributos de menor granularidade devem receber o número 9 e os de maior granularidade devem receber o números inferiores, até o limite de 1. Como mais de um atributo pode receber o mesmo nível hierárquico, cada agrupamento deve ter apenas um atributo predefinido, através da identificação na coluna “*padrão*”. A Figura 27 apresenta, como exemplo, a tabela dimensão Debito. Uma vez que as tabelas fato e dimensão foram selecionadas e preparadas, o próximo passo é o processamento do cubo.

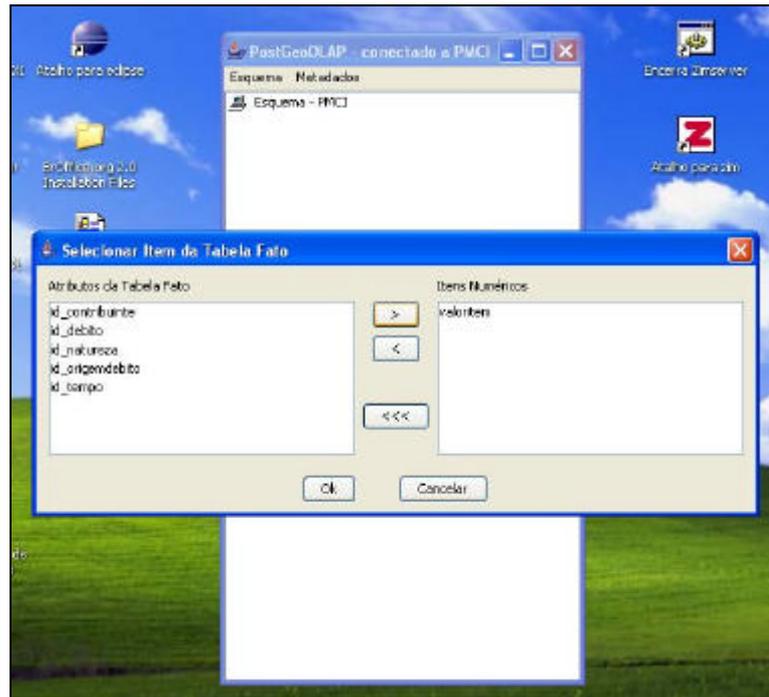


Figura 26 Atributos da Tabela Fato para operações

Segundo (COLONESE,2004), o processamento do cubo precisa analisar o custo-benefício da criação de novas estruturas de agregação para cada combinação dos níveis presentes em cada uma das hierarquias, ou seja, significa fazer o processamento de “k” hierarquias , cada uma delas com “n” níveis.

Durante essa etapa, foram encontrados alguns problemas na utilização da ferramenta, sendo o principal deles a ocupação de todo o espaço físico do HD, interrompendo o processamento. Como a máquina dispunha de apenas um HD de 120 Gbytes, foi adquirido um segundo disco rígido, de 200 Gbytes para o experimento. A questão de ocupação do espaço total, inclusive no HD adquirido, persistiu. Dessa forma, a única alternativa foi a mudança de estratégia, seguida de novos testes. Dentre as modificações realizadas está a redução da base de dados para aos exercícios de 2000 até 2005; a utilização de 3 níveis hierárquicos (0,9,8), ao invés de 4 níveis (0,9,8,7) e o aumento do número mínimo de linhas para agregação para o patamar de 900.000. Após 25 horas e 22 minutos, o processamento terminou com sucesso, gerando 137 agregações no banco de dados e ocupando 181 Gbytes no HD de tamanho maior.

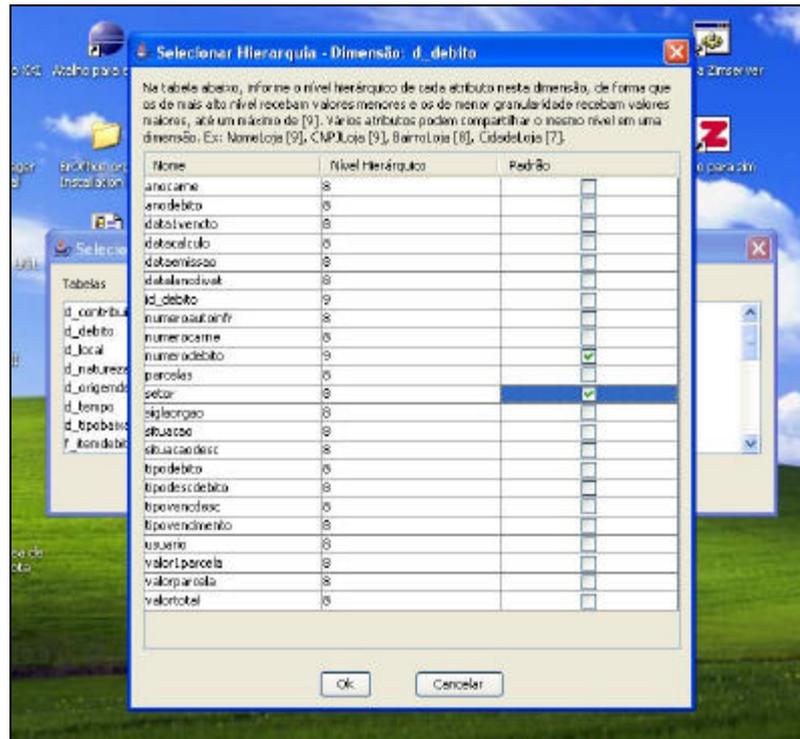


Figura 27 Seleção da Tabela Dimensão Débito

Depois de processado, o cubo está pronto para receber as consultas analíticas. A tela principal (Figura 28) da ferramenta PostGeoOlap possui quatro áreas distintas. Na parte superior, à esquerda, encontra-se uma janela onde são exibidas todas as tabelas dimensão e fato que foram selecionadas no processo de criação e montagem do cubo. Na mesma parte, porém à direita, o mapa georeferenciado da cidade, onde serão plotados os resultados das consultas que contiverem dados vetoriais. A barra à esquerda possui dois recursos disponíveis (ícones). O primeiro (desenho de uma lupa) para ampliação do mapa (*zoom*) e o segundo (desenho de uma mão) para movimentar o mapa em qualquer direção. Na parte inferior, à esquerda, outra janela onde serão exibidos os critérios selecionados para a montagem da instrução SQL. E por fim, ao lado desta, a tela onde serão exibidos os dados (tabulares) resultantes da execução das consultas.

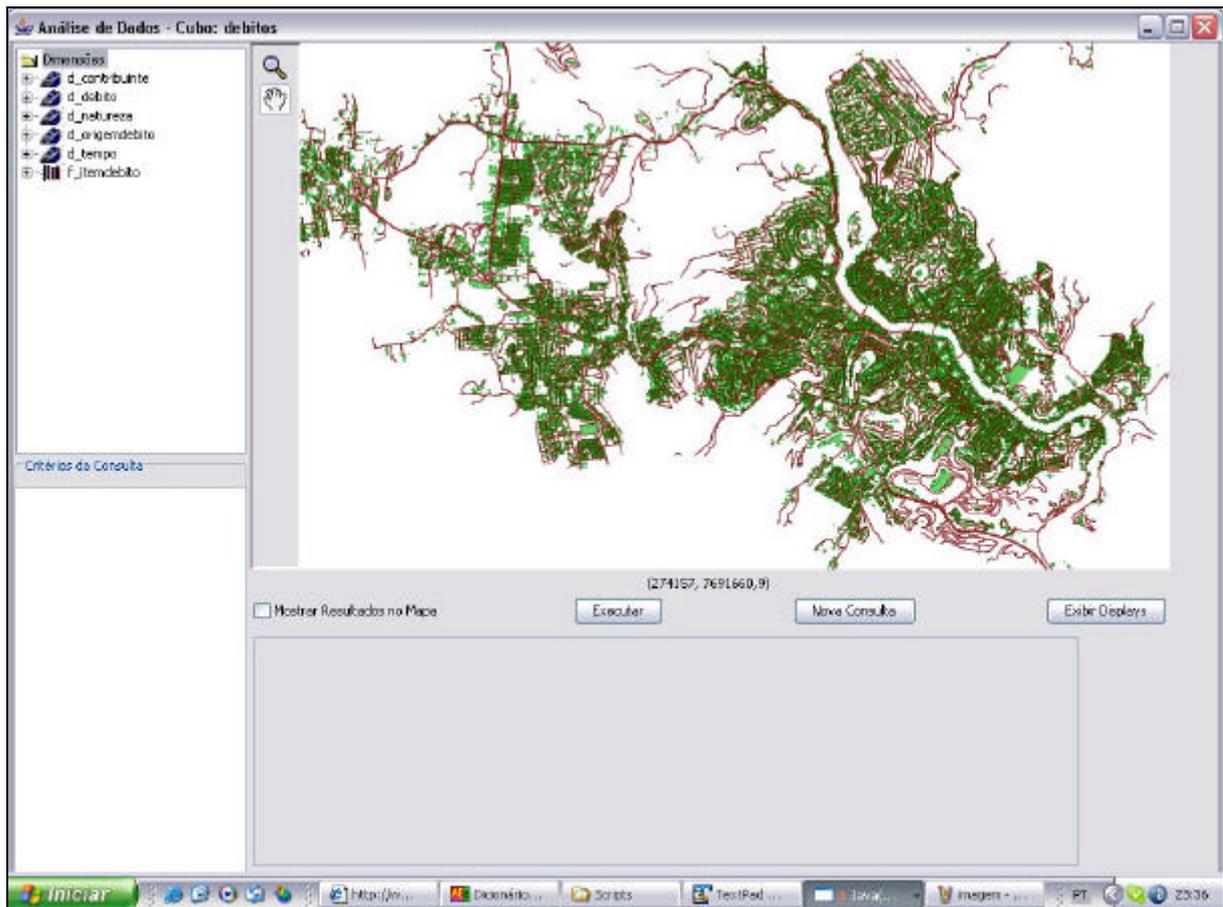


Figura 28 Tela Principal do PostGeoOlap

6.2 ANÁLISE DOS RESULTADOS

O cubo, depois de processado, está pronto para receber as consultas analíticas, bem como para visualizar resultados no mapa (geográficos) e na tabela (convencionais). A atual versão da ferramenta possui algumas limitações, o que demandará a necessidade de complementar alguns resultados com consultas realizadas no próprio SGBD, quando necessário, para auxiliar a interpretação dos resultados.

6.2.1 Análise dos lançamentos tributários

A primeira consulta (Figura 29), de cunho relativamente simples, busca mostrar no mapa os imóveis que tiverem seus tributos quitados dentro do próprio exercício (fato gerador), que é o ano em que o tributo foi calculado. A utilização de apenas um exercício (2004), dentro

dos critérios dessa consulta, ocorreu devido ao fato de a ferramenta ainda não oferecer recursos como pontilhar cada exercício com identificadores ou cores diferentes.

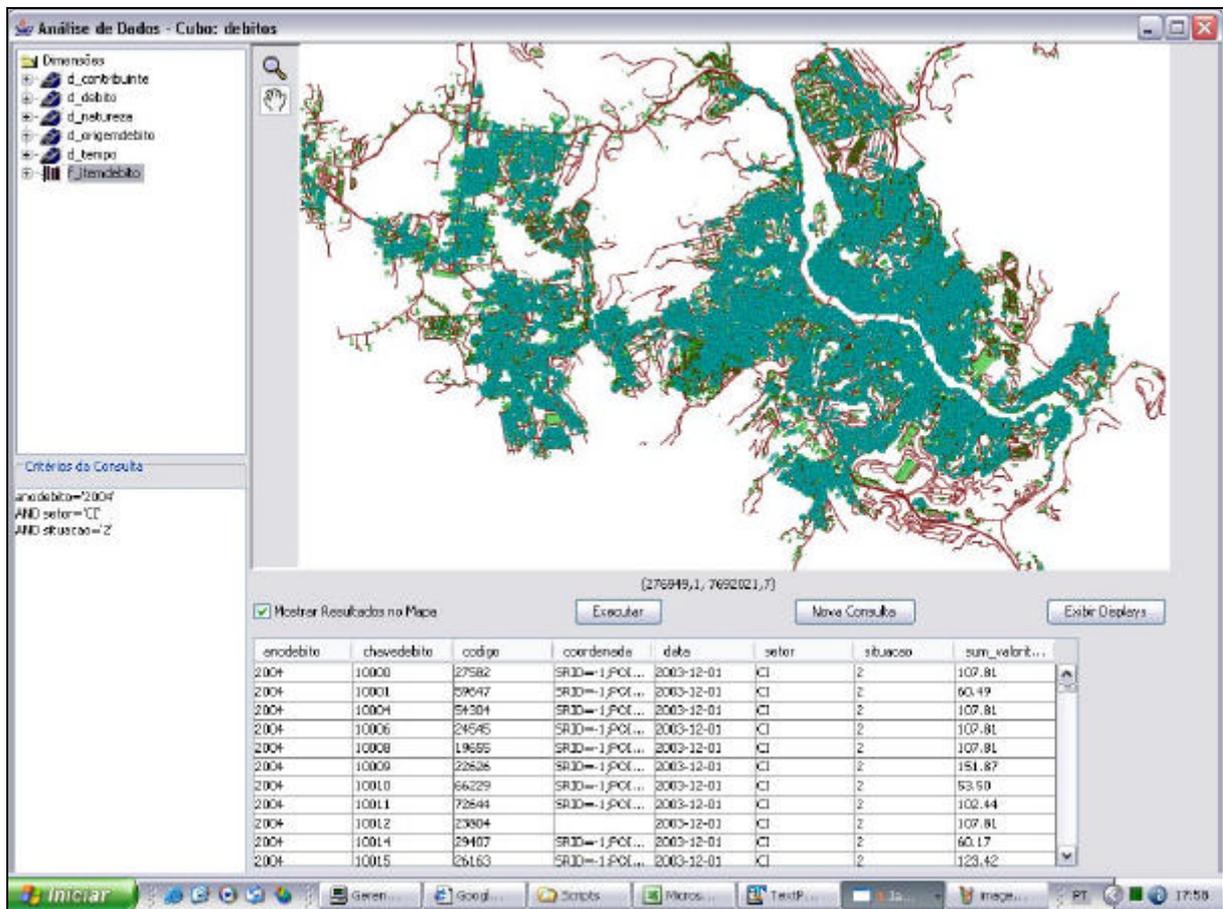


Figura 29 Resultado dos Imóveis Quitados em 2004

Os critérios utilizados nessa consulta foram os seguintes: o exercício do tributo, o setor (IPTU) e a situação (2-quitado). Baseado nesses itens e uma vez confirmada a execução, a ferramenta montou uma a seguinte instrução SQL para a busca de dados no cubo:

```
(SELECT debitos98899.anodebito, debitos98899.chavedebito,
debitos98899.codigo, debitos98899.coordendada,
debitos98899.data, debitos98899.setor, debitos98899.situacao,
SUM(debitos98899.valoritem) AS SUM_valoritem FROM debitos98899
WHERE (anodebito='2004' AND setor='CI' AND situacao='2') GROUP
BY debitos98899.anodebito, debitos98899.chavedebito,
debitos98899.codigo, debitos98899.coordendada,
debitos98899.data, debitos98899.setor, debitos98899.situacao;).
```

As colunas exibidas como saída para os dados convencionais foram o ano do tributo, a inscrição do imóvel, o código do proprietário, a coordenada geográfica (necessária para a identificação), a data que o tributo foi calculado, o seu setor de origem, a situação de quitado e o valor do tributo.

No entanto, além do resultado visual, que abordou apenas um exercício, uma consulta utilizando os mesmos critérios foi executada diretamente no banco de dados, porém abrangendo os exercícios de 2000 a 2004, trazendo o resultado expresso na Tabela 19.

CI	Lançados		Quitados			Inadimplentes		
	Ano	Qtd.	Valor	Qtd.	Valor	%	Qtd.	Valor
2000	39929	6.919.352,60	21968	4.087.301,00	55,02	17952	2.829.860,89	44,96
2001	40175	6.918.187,23	22151	3.984.447,85	55,14	18021	2.930.022,18	44,86
2002	40190	7.210.055,45	22061	4.063.321,95	54,89	18124	3.140.102,19	45,10
2003	49038	8.690.319,28	26006	4.480.806,67	53,03	23032	4.209.512,61	46,97
2004	52904	9.112.768,61	27610	4.613.374,30	52,19	25291	4.498.642,60	47,81

Tabela 19 – Comparativo de Lançamentos e Retorno do IPTU

O resultado apresentou uma queda no percentual de pagamentos com relação aos lançamentos, fato que revelou o aumento da inadimplência.

Segundo François E. J. de Bremaeker (BREMAEKER,2001), em relatório emitido pelo Instituto Brasileiro de Administração Municipal (IBAM), nos municípios entre 100 mil e 500 mil habitantes, nada menos de 37,1% arrecadou menos de 40% do valor do IPTU e para aqueles com população acima de 500 mil habitantes, o percentual chega a 26,9%. Levando-se os índices supracitados em consideração, a situação tributária do município de Cachoeiro de Itapemirim, mesmo apresentando o aumento na inadimplência ao longo dos anos, ainda se encontra acima da média apresentada pelos municípios com o mesmo perfil demográfico.

A mesma consulta, porém voltada para os tributos de ISS, revelou um quadro (Tabela 20) de estabilidade com uma pequena variação positiva na quantidade dos créditos tributários.

TB	Lançados		Quitados			Inadimplentes		
	Ano	Qtd.	Valor	Qtd.	Valor	%	Qtd.	Valor
2000	8959	2.842.301,06	3193	957.782,11	35,64	5571	1.822.781,41	62,18
2001	9290	4.915.464,13	3056	931.283,30	32,90	5980	3.790.765,18	64,37
2002	9194	4.495.821,51	3342	902.901,19	36,35	5850	3.591.671,06	63,63
2003	9086	2.968.157,02	3289	869.370,00	36,20	5786	2.095.057,49	63,68
2004	9385	2.285.612,20	3434	925.067,37	36,59	5925	1.353.155,63	63,13

Tabela 20 – Comparativo de Lançamentos e Retorno do ISS

Na próxima consulta, ainda com relação à cobrança do IPTU, outras variáveis foram acrescentadas, como a ocupação do tipo “*edificada*”, os terrenos vagos (sem construção) foram descartados; o tipo de ocupação igual a “*residencial*”; o padrão do imóvel do tipo “*B*” e com uma área edificada menor ou igual a 300m². Nesta consulta, a instrução SQL foi construída da seguinte forma:

```
(SELECT
                                debitos99099.areaedificacaounid,
debitos99099.chavedebito,                                debitos99099.codigo,
debitos99099.coordenada,                                debitos99099.data,
debitos99099.numerodebito,                                debitos99099.ocupacao,
debitos99099.padrao,                                    debitos99099.setor,
debitos99099.tipocupacao,    SUM(debitos99099.valoritem)    AS
SUM_valoritem FROM debitos99099 WHERE (anodebito_o = '2004' AND
setor = 'DA' AND tipodebito_o = '2' AND tipocupacao =
'RESIDENCIAL' AND padrao = 'B' AND ocupacao = 'EDIFICADO' AND
areaedificacaounid    <=    300)    GROUP    BY
debitos99099.areaedificacaounid,    debitos99099.chavedebito,
debitos99099.codigo,                                debitos99099.coordenada,
debitos99099.data,                                    debitos99099.numerodebito,
debitos99099.ocupacao, debitos99099.padrao, debitos99099.setor,
debitos99099.tipocupacao;).
```

E o programa apresentou o resultado conforme a Figura 30.

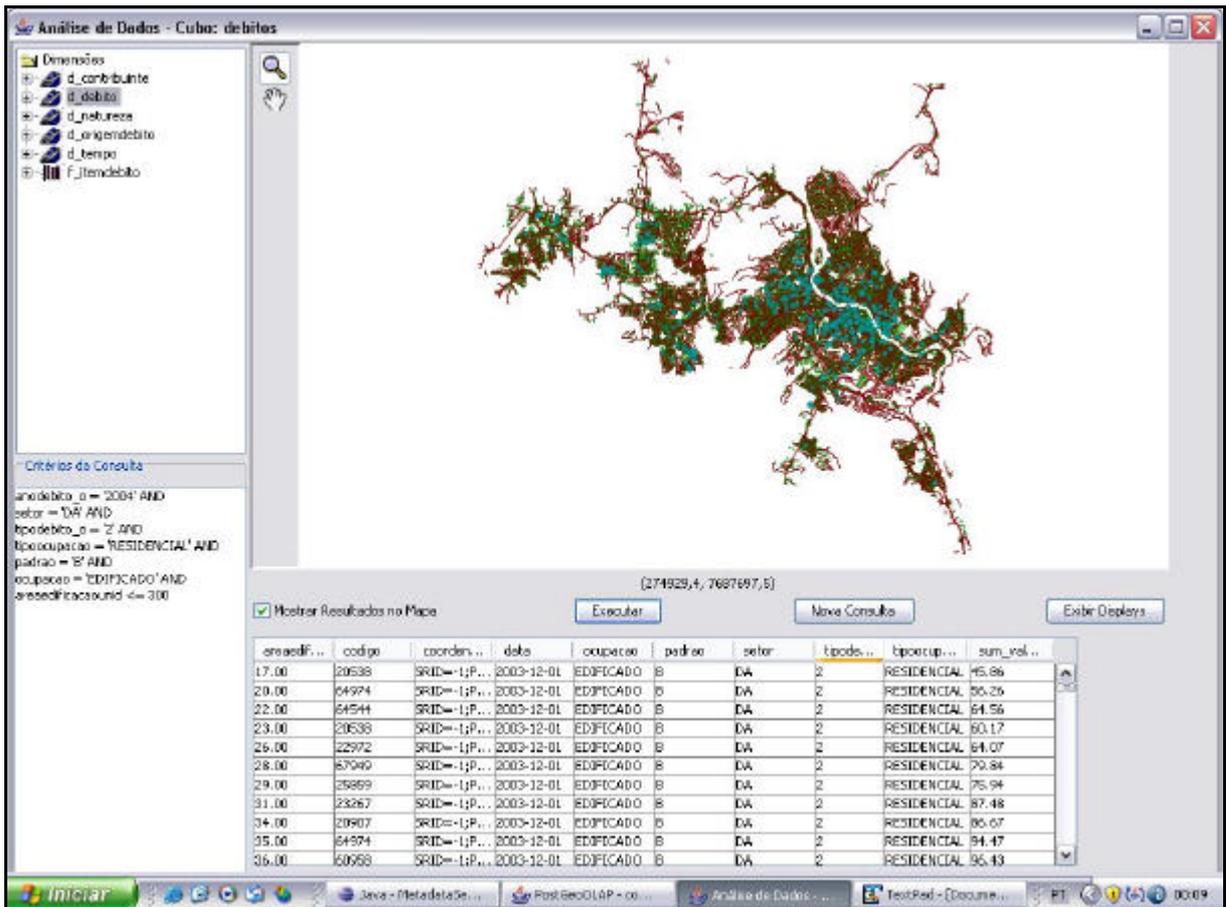


Figura 30 Inadimplência do IPTU com adoção de critérios

Como mostra o mapa, os imóveis que atenderam à condição estão situados, em sua maior parte, no centro da cidade, próximos do rio. Baseado nesta observação, talvez seja necessário o desenvolvimento de um trabalho direcionado, visando a reduzir a inadimplência, uma vez que essa região possui toda infra-estrutura.

Devido à limitação da ferramenta e para uma análise temporal dos dados, a mesma consulta foi realizada no SGBD, porém agrupando-se os resultados por exercício. O resultado dos dados tabulares encontra-se distribuído na Tabela 21. Nessa mesma tabela, por questões de disposição física, a coluna “tipo” teve seu conteúdo abreviado (*R* – residencial / *C* – comercial / *I* – Industrial) e a coluna “valor(*R*\$)” permaneceu apenas com a vírgula da casa decimal.

Ano	Tipo	Padrão	Área (m ²)							
			Até 300,00		300,01 a 500,00		500,01 a 1000,00		Acima de 1000,00	
			Quantidade / Valor(R\$)		Quantidade / Valor(R\$)		Quantidade / Valor(R\$)		Quantidade / Valor(R\$)	
2000	R	A	18	7907,90	3	1450,56	0	0,00	0	0,00
		B	547	140310,75	36	20840,94	9	6277,99	1	1964,37
		C	9501	1023971,52	43	14377,59	2	1776,08	4	16029,27
		D	287	27330,16	0	0,00	0	0,00	0	0,00
		E	16	2211,21	0	0,00	0	0,00	0	0,00
	C	A	36	16891,52	2	414,63	0	0,00	0	0,00
		B	161	43443,76	5	6297,98	0	0,00	2	3324,03
		C	1022	225184,64	21	15101,16	16	25711,63	7	16654,64
		D	294	50715,42	12	7269,35	6	5096,52	2	3645,01
		E	5	2845,17	0	0,00	1	2541,60	1	529,26
	I	A	0	0,00	0	0,00	0	0,00	0	0,00
		B	4	4498,15	0	0,00	1	5305,48	0	0,00
		C	52	15828,15	1	464,06	0	0,00	0	0,00
		D	41	11526,37	9	5067,27	5	12066,55	2	5261,25
		E	3	664,87	0	0,00	0	0,00	0	0,00
2001	R	A	26	13502,34	4	3893,78	1	1642,59	0	0,00
		B	644	165401,75	44	26599,25	7	5845,96	4	2674,77
		C	9444	1012074,20	36	14267,83	2	537,07	1	15443,91
		D	173	18180,80	2	542,11	0	0,00	0	0,00
		E	14	1449,42	0	0,00	0	0,00	0	0,00
	C	A	46	26946,24	2	471,54	0	0,00	0	0,00
		B	201	52853,07	4	16681,30	2	4248,26	1	2963,55
		C	1144	250812,27	24	20909,41	18	33444,55	7	19899,32
		D	335	56248,75	11	6713,42	8	8389,67	2	4097,36
		E	1	82,76	0	0,00	1	2541,60	2	5033,28
	I	A	0	0,00	0	0,00	0	0,00	0	0,00
		B	4	3371,50	0	0,00	1	5305,48	0	0,00
		C	62	18686,15	1	561,95	1	2126,16	0	0,00
		D	45	12890,58	9	5870,84	9	16395,78	4	9351,61
		E	2	288,29	0	0,00	0	0,00	0	0,00
2002	R	A	24	12659,23	3	2878,02	0	0,00	0	0,00
		B	688	183783,86	39	23586,63	1	1419,17	1	2674,77
		C	9675	1040619,41	34	14486,03	5	7372,29	3	16021,14
		D	87	10571,11	1	232,84	1	187,48	0	0,00
		E	7	1368,12	0	0,00	0	0,00	0	0,00
	C	A	47	19495,72	2	471,54	0	0,00	0	0,00
		B	261	64808,19	4	7226,43	2	5885,64	4	16363,25
		C	1177	256897,45	20	18834,46	18	36264,19	10	17095,61
		D	341	57838,53	14	9061,19	9	11638,42	2	4097,36
		E	1	82,76	0	0,00	0	0,00	1	6186,93
	I	A	0	0,00	1	2754,28	0	0,00	0	0,00
		B	7	5970,02	1	1309,26	1	5305,48	0	0,00

Ano	Tipo	Padrão	Área (m ²)							
			Até 300,00		300,01 a 500,00		500,01 a 1000,00		Acima de 1000,00	
			Quantidade / Valor(R\$)		Quantidade / Valor(R\$)		Quantidade / Valor(R\$)		Quantidade / Valor(R\$)	
		C	61	19592,14	1	561,95	1	2126,16	0	0,00
		D	47	13563,75	8	5327,28	7	15615,95	5	12806,21
		E	2	475,28	0	0,00	0	0,00	0	0,00
2003	R	A	22	12422,83	1	1720,47	0	0,00	0	0,00
		B	717	261354,01	44	46132,85	6	12734,51	1	5034,10
		C	10417	1301923,09	27	20981,26	0	0,00	2	23473,26
		D	4279	213369,70	5	1888,44	0	0,00	0	0,00
		E	249	8454,93	0	0,00	0	0,00	1	861,29
	C	A	40	10333,09	0	0,00	0	0,00	0	0,00
		B	230	57266,96	5	7968,70	1	3497,36	3	19817,20
		C	1060	208037,61	21	25194,55	19	42498,78	1	7608,38
		D	324	47988,59	10	7042,86	8	9382,03	3	5491,48
		E	34	5616,88	6	3159,15	7	4604,17	5	24098,79
	I	A	0	0,00	1	1796,57	0	0,00	0	0,00
		B	8	7724,16	1	1693,48	1	3813,13	0	0,00
		C	70	21377,31	2	1443,89	1	1906,81	1	8075,53
		D	47	11930,45	11	7192,94	7	11740,20	6	13261,82
		E	33	8551,47	8	3657,69	10	7658,46	12	34991,18
2004	R	A	22	12093,54	1	1697,06	0	0,00	0	0,00
		B	712	257223,12	44	48205,05	5	10905,42	1	4984,02
		C	10190	1272929,70	21	14167,17	0	0,00	2	23277,00
		D	6471	344954,33	4	2735,42	0	0,00	0	0,00
		E	563	17587,50	0	0,00	0	0,00	1	1218,69
	C	A	45	11591,16	0	0,00	0	0,00	0	0,00
		B	192	46771,17	5	8275,69	2	6679,29	3	19580,95
		C	1034	193331,06	19	23193,74	15	30214,50	2	10393,07
		D	329	44714,57	10	7062,38	8	10583,63	2	3512,16
		E	37	9704,48	3	1564,71	6	5236,70	7	39358,62
	I	A	0	0,00	1	1773,32	0	0,00	0	0,00
		B	9	6244,98	1	1674,29	1	3770,37	0	0,00
		C	56	18159,65	1	842,76	1	1240,64	1	7975,37
		D	43	12451,58	12	7334,07	9	14763,10	7	15295,47
		E	25	9418,95	5	2761,44	7	9169,34	13	45982,63

Tabela 21 – Inadimplência no IPTU por Categorias

Com relação à Tabela 21, alguns comportamentos foram descobertos, a saber: no âmbito residencial, a inadimplência foi observada nos imóveis de padrão “B” e “C” com área construída de até 300 m²; na esfera comercial, a falta do recolhimento foi percebida, através de um desvio significativo, apenas nos imóveis do padrão “B” e com até 300 m²; entretanto, no

aspecto industrial, uma leve variação foi observada nos padrões “B” e “C”, contudo uma forte elevação no padrão “E”, mas todos dentro da faixa de até 300 m².

Outro fator relevante encontrado foi quando da comparação entre a inadimplência ocorrida em terrenos vagos e edificações (Tabela 22), demonstraram comportamentos divergentes. Com relação ao primeiro tipo, o levantamento apontou para uma queda bastante discreta na quantidade, diferentemente do segundo, que apresentou um aumento desproporcional entre os anos de 2002 e 2003.

IPTU	Terrenos		Edificações	
	Ano	Quantidade	Valor	Quantidade
2000	5450	954.133,63	12178	1.750.796,81
2001	5500	1.005.368,38	12347	1.869.240,47
2002	5380	1.025.932,93	12624	1.935.515,53
2003	5196	1.653.383,51	17767	2.546.772,41
2004	5343	1.856.038,74	19948	2.642.603,86

Tabela 22 – Inadimplência entre Terrenos e Edificações

6.2.2 Análise dos créditos tributários

A consulta a seguir (resultado na Tabela 23) teve como objetivo o levantamento do montante arrecadado dividido por faixa do valor total calculado do tributo e agrupados por setor e exercício, conforme a instrução SQL:

```
(SELECT 'ANO' || ' ' || anodebito || ' - ' || setor || ' - ' || sum(valorpago) from f_pagamento, d_debito where f_pagamento.id_debito = d_debito.id_debito AND d_debito.situacao = 2 AND d_debito.valortotal BETWEEN 200.01 AND 200.00 GROUP BY anodebito, setor ORDER BY anodebito, setor asc);
```

A coluna “setor” foi utilizada para identificar a origem do tributo, a saber: “CF” para identificar os impostos do IPTU, “DA” para os débitos que estão em Dívida Ativa, “TA” para as Taxas Avulsas e “TB” para tributos do tipo ISS.

No total dos créditos tributários por exercício, o comportamento do IPTU, mesmo com uma inadimplência em ascensão, se manteve como a maior fonte de receitas durante todos os exercícios pesquisados, chegando a ultrapassar os 50% (cinquenta por cento) em 2004. Ainda

com relação a este tributo, o maior montante arrecadado ficou situado na faixa dos tributos com valor total de até R\$ 200,00 (duzentos reais).

ANO	SETOR	FAIXA (R\$) - Valor Total do Tributo Calculado				TOTAL
		<=200,00	200,01 e 500,00	500,01 e 1000,00	> 1000,00	
2000	CI	1.387.305,94	911.132,97	345.161,20	647.471,08	3.291.071,19
	DA	534.682,52	385.887,78	396.646,99	2.514.289,57	3.831.506,86
	TA	275.769,17	188.266,89	201.489,79	714.393,56	1.379.919,41
	TB	150.372,79	327.459,14	316.907,92	271.575,78	1.066.315,63
	Sub-total	2.348.130,42	1.812.746,78	1.260.205,90	4.147.729,99	9.568.813,09
2001	CI	1.510.164,97	975.242,05	371.995,76	599.630,31	3.457.033,09
	DA	567.256,62	300.265,64	250.889,14	1.925.997,60	3.044.409,00
	TA	245.469,00	168.579,31	147.254,75	724.314,70	1.285.617,76
	TB	141.451,68	296.961,68	322.268,95	354.298,81	1.114.981,12
	Sub-total	2.464.342,27	1.741.048,68	1.092.408,60	3.604.241,42	8.902.040,97
2002	CI	1.457.706,04	977.194,57	368.962,37	650.569,72	3.454.432,70
	DA	142.306,13	289.461,54	336.362,31	1.065.105,69	1.833.235,67
	TA	241.063,60	181.614,61	173.698,26	609.831,85	1.206.208,32
	TB	170.407,95	380.134,44	232.478,60	157.989,13	941.010,12
	Sub-total	2.011.483,72	1.828.405,16	1.111.501,54	2.483.496,39	7.434.886,81
2003	CI	1.670.160,72	928.716,32	497.170,47	697.914,26	3.793.961,77
	DA	141.029,48	321.609,29	375.220,19	1.381.133,77	2.218.992,73
	TA	199.315,60	167.106,36	184.473,65	719.134,00	1.270.029,61
	TB	191.220,30	484.060,72	110.745,78	165.448,66	951.475,46
	Sub-total	2.201.726,10	1.901.492,69	1.167.610,09	2.963.630,69	8.234.459,57
2004	CI	1.708.679,51	951.497,03	523.683,30	708.195,48	3.892.055,32
	DA	142.092,26	294.745,26	268.060,03	597.565,80	1.302.463,35
	TA	218.648,01	225.610,91	231.230,95	957.965,40	1.633.455,27
	TB	194.702,87	528.744,52	122.389,96	70.064,31	915.901,66
	Sub-total	2.264.122,65	2.000.597,72	1.145.364,24	2.333.790,99	7.743.875,60

Tabela 23 – Total Arrecadado por Faixa de Valor do Tributo

A Dívida Ativa despontou em segundo lugar, apresentando o maior volume de recolhimentos na faixa dos tributos acima de R\$ 1.000,00 (um mil reais). Este fato se deve, em grande parte, aos parcelamentos de dívidas que não foram recolhidas dentro do exercício do fato gerador. Dessa forma, quando o contribuinte comparece à Prefeitura para negociar a dívida, vários débitos são reunidos para o parcelamento.

As Taxas Avulsas, que são solicitações específicas de serviços feitos à municipalidade (alvará de construção, habite-se, licenças ambientais e outros) apresentaram um aumento substancial no contexto, chegando até a ultrapassar a Dívida Ativa no exercício de 2004.

Com relação ao ISS, sua participação foi uniforme durante o período analisado, sem oscilações significativas, o que do ponto de vista administrativo, poderia ser um indício de uma necessidade de mudança ou ajuste da política tributária nesse setor, uma vez que a atividade econômica base do município encontra-se voltada para a extração e beneficiamento de rochas (mármore e granitos), comércio e serviços.

Um outro levantamento foi realizado sobre os meses de maior arrecadação, conforme Tabela 24, no qual pode ser observado um pico de arrecadação nos meses de Fevereiro a Maio, quando geralmente são fixados os primeiros vencimentos do IPTU e do ISS Fixo.

Após este período, uma queda nos recolhimentos seguido de pequenas variações descendentes, que volta a ter uma pequena ascensão em dezembro, quando finda o prazo para o recolhimento dentro do exercício, evitando sua passagem para Dívida Ativa, onde está sujeito à execução judicial.

Esse comportamento não se aplica às Taxas Avulsas, uma vez que as mesmas não estão sujeitas à transferência para a Dívida Ativa, tanto que o seu comportamento durante os meses permanece uniforme, sem grandes variações. Ressalte-se que os números contidos na tabela representam o somatório de todos os exercícios analisados, conforme instrução SQL a seguir:

```
(SELECT COUNT(f_pagamento.id_debito) FROM f_pagamento, d_tempo,
d_debito WHERE f_pagamento.id_tempo = d_tempo.id_tempo AND
f_pagamento.id_debito = d_debito.id_debito GROUP BY setor,mes
ORDER BY setor,mes asc);
```

Mês	IPTU	Dívida Ativa	Taxas Avulsas	ISS Fixo
Janeiro	5.123	6.854	2.298	356
Fevereiro	40.613	7.081	2.340	1.475
Março	51.647	9.217	2.840	11.491
Abril	52.096	8.931	2.784	6.381
Maio	23.063	9.128	3.140	4.746
Junho	17.753	8.219	3.332	3.855
Julho	9.672	7.063	3.225	2.732
Agosto	5.507	6.554	2.939	1.379
Setembro	9.358	6.143	2.563	1.080
Outubro	3.199	5.929	2.484	888
Novembro	4.293	1.581	2.266	1.239
Dezembro	11.547	8.881	2.621	1.095

Tabela 24 – Quantidade de Pagamentos por Mês (2000 a 2004)

Ainda com relação aos créditos tributários, uma outra informação a ser gerada foi a quantidade de pagamentos realizados em cota única ao longo dos anos (Tabela 25). De acordo com o resultado, o número de contribuintes que tem quitado seus débitos de IPTU num único parcelamento vem aumentando consideravelmente. O mesmo vem acontecendo, de forma discreta com os débitos de ISS e as Taxas Avulsas. Contudo, os pagamentos à vista referentes à Dívida Ativa vêm caindo bruscamente, atingindo percentuais próximos de 50% (cinquenta por cento).

Ano	IPTU	ISS	Dívida Ativa	Taxas Avulsas
2000	17930	2066	7468	5837
2001	17026	2175	8033	5180
2002	17099	2213	3060	5112
2003	19933	2089	3927	5235
2004	21868	2391	3782	5750

Tabela 25 – Quantidade de Pagamentos em Cota Única

Um último levantamento efetuado foi o de pagamentos efetuados fora do vencimento, como disposto na Tabela 26. O IPTU, assim como na consulta de créditos realizados em cota única, também vem apresentando uma elevação nos pagamentos com atraso. Para os demais tributos (ISS e Taxas Avulsas) foram documentadas pequenas variações e ficando a Dívida Ativa com oscilações maiores, mas mantendo-se na média.

ANO	IPTU	ISS	Dívida Ativa	Taxas Avulsas
2000	9871	1893	4683	312
2001	11952	1513	3905	237
2002	11229	2078	6047	348
2003	15004	2042	5934	338
2004	18724	1971	4690	373

Tabela 26 – Créditos Realizados Após o Vencimento

Todos os levantamentos realizados até o momento tiveram como objetivo principal a demonstração de alguns comportamentos da base tributária e do município, servindo de ferramenta para o usuário analisar sobre a necessidade de ajustes nos processos de negócio do município.

7 CONCLUSÃO

Assim como todo processo tecnológico, as necessidades das instituições públicas e privadas também vêm sofrendo modificações e cada vez mais buscando ferramentas capazes de auxiliar as gestões e aprimorar os processos de tomada de decisão. No caso do setor público, abordado nesse contexto, o acesso a esses recursos ainda se encontra num estágio muito restrito, visto que dentro de um universo de aproximadamente 5.500 prefeituras, menos de 10 % utilizam algum recurso tecnológico. É um setor muito específico, que além de não visar ao lucro tem a função de promover o bem estar social através do recolhimento de tributos, que foi o objeto de estudo desta pesquisa, juntamente com o emprego de tecnologia aberta (software livre) em sua totalidade.

Inicialmente, foi construído um Data Warehouse com dados tributários captados da base de dados transacional. O SGBD utilizado foi o PostgreSQL que, além de ser uma ferramenta livre, suporta dados do tipo geográfico através do PostGIS. Os dados geográficos foram levantados com base num sistema de geoprocessamento que atualmente se encontra desativado. Como o sistema não possuía uma documentação, alguns dados não foram localizados, como as coordenadas do rio que corta a cidade.

Em seguida, foi utilizada a ferramenta analítica sobre esses dados, PostGeoOlap, um software de código aberto, desenvolvido na Universidade Candido Mendes, por alunos do curso de Mestrado em Pesquisa Operacional e Inteligência Computacional, com o objetivo de desenvolver um Sistema de Suporte à Decisão. A penúltima versão foi a utilizada nesse trabalho.

Na etapa de criação do cubo foram encontrados alguns problemas durante os processamentos como a geração de agregações inválidas e utilização de todo o espaço físico nos discos de armazenamento, levando a uma redução no tamanho da base de dados analisada.

Uma vez criado o cubo, várias consultas foram realizadas inicialmente como, por exemplo, a visão geral dos tributos lançados e a inadimplência, ambas com resultados plotados no mapa e dados tabulares.

Algumas funções ainda não foram implementadas na versão utilizada da ferramenta PostGeoOlap. Por exemplo, uma consulta que envolva vários exercícios (anos) deveria diferenciar cada um deles com uma cor ou símbolo diferentes no mapa, facilitando a visualização e compreensão. A ferramenta também não apresenta totalizadores nas consultas. Em virtude deste fato, as consultas (com os mesmos critérios) também tiveram de ser executadas diretamente no SGBD para a totalização dos dados e sua disposição em tabelas para complementar os levantamentos executados através da ferramenta analítica. Apesar dessas limitações, a ferramenta permitiu o alcance do objetivo, no presente trabalho, de criar visões multidimensionais e com uma abordagem espacial. Mesmo ainda não tendo sido implantada na PMCI, ficou comprovada a sua aplicabilidade.

7.1 CONTRIBUIÇÕES DO PRESENTE TRABALHO

Esta dissertação beneficiou o ramo de pesquisa da tecnologia da informação e a municipalidade de Cachoeiro de Itapemirim através dos seguintes recursos:

- A utilização de um ambiente capaz de suportar as consultas analíticas e gerar resultados que contribuirão para um ajuste na política tributária municipal.
- Permitiu o uso da visão espacial, como um recurso adicional, que facilitou a identificação de possíveis distorções que certamente não seriam percebidas em relatórios tabulares convencionais.
- A ferramenta PostGeoOlap foi submetida a mais um cenário (aplicação) que comprovou sua eficiência e atestou seus princípios de funcionalidade, evidenciando algumas limitações e oportunidades de melhoria de funcionalidade e desempenho.
- Ratificou o uso de software livre como uma alternativa viável para o setor público, uma vez que não acarretou custos para a municipalidade que certamente usufruirá dessa ferramenta como objeto de suporte à tomada de decisão.

7.2 SUGESTÕES PARA TRABALHOS FUTUROS

Este trabalho, que atingiu o seu propósito, poderá ser utilizado em possíveis implementações de trabalhos futuros, a saber:

Criação de um recurso dentro da ferramenta PostGeoOlap que implemente o conceito de “business-view”, ou seja, permitir que os usuários criem sinônimos para atributos e tabelas, criando uma nomenclatura própria dentro do modelo de negócios.

Desenvolvimento de um processo de Datamining, como preconizado na arquitetura de um ambiente de DW, para pesquisar comportamentos e desvios, localizados na base de dados e que dificilmente seriam observados através de relatório convencionais.

Utilização do DW criado para o desenvolvimento de um programa de CRM (*Customer Relationship Management*), para melhorar o gerenciamento do relacionamento com o cliente. Contudo, adaptado para a gestão do relacionamento com o contribuinte, sem deixar de atender às quatro premissas do CRM (cliente certo, oferta certa, canal certo e hora certa) .

Essa base de dados também poderá ser útil na criação de indicadores de performance da gestão pública através do desenvolvimento de um BSC (*Balanced Scorecard*), programa que cria pontos para mensuração do desempenho institucional seguindo quatro perspectivas (financeiro, clientes, processos, aprendizado e crescimento) que visam auxiliar no planejamento estratégico.

8 REFERÊNCIAS

ABRANTES, Graça. *Sistemas de Informações Geográficas – Conceitos*. Disponível em: <<http://www.isa.utl.pt/dm/sig/sig/SIGconceitos.html>>. Acesso em: 01/03/2005.

AMORIM, M. Viana; TANAKA, Astério. *Ambiente de Desenvolvimento de Aplicações de Banco de Dados Ativos voltadas para Regras de Negócio*. Relatório Técnico no. 038/DE9/99. Instituto Militar de Engenharia. Rio de Janeiro, 1998.

BALEEIRO, Aliomar. *Direito Tributário Brasileiro*, 5ª ed., Rio de Janeiro: Forense, 1977.

BARBIERI, Carlos. *BI- Business Intelligence – Modelagem & Tecnologia*. Rio de Janeiro: Axcel Books, 2001.

BEUREN, Ilse Maria. *Gerenciamento da Informação: um recurso estratégico no processo de gestão empresarial*. São Paulo: Atlas, 2000.

BORBA, Cláudio. *Direito tributário: teoria e 600 questões*. 7.ed. Rio de Janeiro: Impetus, 2001.

BREMAEKER, François E. J. de. *A Iluminação Pública no Brasil: situação e custeio*. IBAM / APMC / NAPI / IBAMCO, 2001. 2.ed. Disponível em: <<http://www.ibam.org.br/publique/media/ESP034P2.pdf>>. Acesso em: 22/07/2006.

BURROUGH, P.A. *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford: Oxford University Press, 1998.

CALIJURI, Maria Lúcia. *A Tecnologia da Informação a Serviço do Desenvolvimento Municipal: programa de capacitação técnica. Núcleo Sigeo*. Viçosa, 1999. (Material do treinamento realizado na Prefeitura Municipal de Cachoeiro de Itapemirim. Apostila)

CÂMARA, Gilberto; DAVIS, Clodoveu; CASANOVA, Marco. *Geoprocessamento: Teoria e Aplicações*. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/>>. Acesso em: 04/04/2004.

CINTRA, Marcos. *Os Tributos Brasileiros*. Disponível em: <<http://www.marcoscindra.org/politica/padrazo.asp?id=83>>. Acesso em: 18/08/2004.

COELHO, Sacha Calmon Navarro, *1940 – Manual do direito tributário*. Rio de Janeiro: Forense, 2002.

COLONESE, Giovani. *Uma ferramenta aberta para desenvolvimento integrado de sistemas de informação para processamento analítica e geográfico*. Dissertação (Mestrado em Informática) - Universidade Candido Mendes, Campos dos Goytacazes, 2004.

COLONESE, Giovani; MANHÃES, Rodrigo; GONZÁLEZ, Sahudy; CARVALHO, Rogério; TANAKA, Astério: *PostGeoOlap: An Open Source Tool for Decision Support*. In: Anais do II Simpósio Brasileiro de Sistemas de Informação (SBSI 2005). Florianópolis, Brazil, 2005.

COLONESE, Giovanni et al. "*Uma Ferramenta de Baixo Custo para o Desenvolvimento de Sistemas de Suporte à Decisão*". In: Anais do XXXVIII Simpósio Brasileiro de Pesquisa Operacional (SOBRAPO). Goiânia, 2006.

COME, Gilberto de. *Contribuição ao estudo da implementação de Data Warehousing: um caso no setor de telecomunicações*. Dissertação (Mestrado) - FEA/USP, São Paulo, 2001.

CTN: Código Tributário Nacional. Disponível em:
<http://www.presidencia.gov.br/ccivil_03/Leis/L5172.htm>. Acesso em 18/08/2004.

DANGERMOND, J. *A Classification of Software Components Commonly Used in Geographic Information Systems*, In: Introductory readings in Geographic Information Systems, Donna Peuquet and Duane F. Marble (eds.), Taylor & Francis, 1990.

FOWLER, Martin. *UML essencial: um breve guia para a linguagem-padrão de modelagem de objetos*/Martin Fowler e Kendal Scott; trad. Vera Pezerico e Christian Thomas Price. 2.ed. Porto Alegre: Bookman, 2000.

GALANTE, Alan Carvalho. *Sistema de suporte à decisão no planejamento urbano municipal: um estudo de caso do município de Macaé*. Dissertação (Mestrado em Informática) - Universidade Candido Mendes, Campos dos Goytacazes, 2004.

GOODCHILD, M. F. *Towards an Enumeration and Classification of GIS Functions*. In *International Geographic Information Systems (IGIS) Symposium: The research agenda* (Aageenbrug R. T. e Schiffman Y. M. eds.), pp. 67-77. AAG, Falls Church Virginia, E.U.A, 1988.

GORNIK, Davor. *Entity Relationship Modeling with UML*. Disponível em:
<www.uml.org.cn/UMLTools/pdf/ermodeling.pdf> . Acesso em: 17/07/2004.

GRIMBERG, Carl. *História Universal*. Ed. Azul, vols. 1, 2 e 3.

IBGE: Instituto Brasileiro de Geografia e Estatística. *Finanças Públicas no Brasil 2002 – 2003*. IBGE. Rio de Janeiro, 2006. Disponível em:

<http://www.ibge.gov.br/home/estatistica/economia/despesaspublicas/financaspublicas_2003/financaspublicas_2003.pdf>. Acesso em: 22/07/2006.

IBGE: Instituto Brasileiro de Geografia e Estatística. *Perfil dos Municípios Brasileiros: gestão pública 2004/IBGE*, Coordenação de População e Indicadores Sociais. Rio de Janeiro: IBGE, 2005. 133p. Disponível em:
<<http://www.ibge.gov.br/home/estatistica/economia/perfilmunic/2004/munic2004.pdf>>. Acesso em: 15/08/2006.

INMON, William H. *Como construir o Data Warehouse*. Rio de Janeiro: Campus, 1997.

JÚNIOR, Methanias Colaço. *Projetando Sistemas de Apoio à Decisão Baseados em Data Warehouse*. Rio de Janeiro: Axcel Books, 2004.

KIMBALL, Ralph. *The Data Warehouse Toolkit*. São Paulo: Makron Books, 1998.

KIMBALL, Ralph. *The Data Warehouse Toolkit: o guia para modelagem dimensional*. Rio de Janeiro: Campus, 2002.

KIMBALL, Ralph. *Three little letters-E, T, and L- obscure the reality of 38 subsystems vital to successful data warehousing*. Dezembro, 2004. Disponível em:
<<http://www.intelligententerprise.com/showArticle.jhtml?articleID=54200319>>. Acesso em: 15/02/2005.

KOŁODZIEJ, Kris. *OpenGIS Web Map Server Cookbook*. Version 1.0. OpenGis Project Document 03-050. Publicado em 28/04/2003. Disponível em:
< <http://xml.coverpages.org/OGC-WMS-Cookbook200305.pdf> > . Acesso em: 04/04/2005.

LIMA, Adilson da Silva. *Erwin 4.0: Modelagem de dados*. São Paulo: Érica, 2002.

MACHADO, Felipe Nery Rodrigues. *Tecnologia e projeto de Data Warehouse: uma visão multidimensional*. 1. ed. São Paulo: Érica, 2004.

MACHADO, Hugo de Brito. *Curso de Direito Tributário*, 5ª ed. Rio de Janeiro: Forense, 1992.

MAGUIRE, D. J. *An Overview and Definition of GIS*. In *Geographical Information Systems: Principal and Applications* (D. J. Maguire, M. F. Goodchild e D. W. Rhind, eds.), Vol. 1, 1, pp. 9-20. Longman Scientific & Technical, Londres, Grã-Bretanha, 1991.

MANHÃES, Rodrigo Soares. *Uma ferramenta em software livre para processamento analítico em sistemas de bancos de dados convencionais e geográficos*. Dissertação (Mestrado em Informática) - Universidade Candido Mendes, Campos dos Goytacazes, 2007.

NAVATHE, Shamkant B.; Elmasri, Ramez. *Sistemas de Banco de Dados: Fundamentos e aplicações*. 3. ed. Rio de Janeiro: LTC, 2000.

NETO, Álvaro Pereira. *PostgreSQL: técnicas avançadas: versões open source: soluções para desenvolvedores e administradores de bancos de dados*. São Paulo: Erica, 2003.

OGC: *Open Geospatial Consortium*. Disponível em:
<<http://www.opengeospatial.org/resources/?page=cookbooks>>. Acesso em: 02/04/2005.

OGC (OPEN GEOSPATIAL CONSORTIUM). OpenGIS Simple Especifications for SQL Revision 1.1. OpenGis Project Document 99-049. Publicado em 05/05/1999.

OLIVEIRA, Djalma de Pinho Rebouças de. *Sistemas de informações gerenciais: estratégias, táticas, operacionais*. 10 ed. São Paulo: Atlas, 2005.

OMG(OBJECT MANAGEMENT GROUP). *Unified Modeling Language*. Disponível em: <<http://www.http://www.uml.org/>>. Acesso em: 04/04/2005.

POSTGIS. *Documentação da extensão espacial para o SGBD PostgreSQL*. Disponível em: <<http://postgis.refractor.net/>>. Acesso em: 15/07/2004.

POSTGRESQL.ORG. *Documentação do SGBD PostgreSQL 8.0*. Disponível em: <<http://www.postgresql.org>>. Acesso em: 03/05/2004.

Surajit Chaudhuri and Umesh Dayal, *An overview of data warehousing and OLAP technology*, ACM SIGMOD Record, 26:1, 1997. Disponível em: <<http://www.ondelette.com/OLAP/dwbib.html>>. Acesso em: 15/07/2004

THOMSEN, Erik. *OLAP: construindo sistemas de informações multidimensionais*. Tradução [da 2.ed. original] de Daniel Vieira. Rio de Janeiro: Campus, 2002.

UCHOA, E. N.; FERREIRA, P. R. *Geoprocessamento com Software Livre. Versão 1.0*. Rio de Janeiro. Publicado em 26/10/2004. Disponível em <<http://www.geolivres.org.br/downloads/>>. Acesso em 19/08/2005.

APÊNDICE A - OS 67 TRIBUTOS BRASILEIROS

- 1-Adicional de Frete para Renovação da Marinha Mercante - AFRMM (Lei 10206/2001)
- 2-Contribuição à Direção de Portos e Costas (DPC)
- 3-Contribuição ao Fundo Nacional de Desenvolvimento Científico e Tecnológico - FNDCT (Lei 10168/2000)
- 4-Contribuição ao Fundo Nacional de Desenvolvimento da Educação (FNDE)
- 5-Contribuição ao Funrural
- 6-Contribuição ao Instituto Nacional de Colonização e Reforma Agrária (INCRA)
- 7-Contribuição ao Seguro Acidente de Trabalho (SAT)
- 8-Contribuição ao Serviço Brasileiro de Apoio a Pequena Empresa (Sebrae)
- 9-Contribuição ao Serviço Nacional de Aprendizado Comercial (SENAC)
- 10-Contribuição ao Serviço Nacional de Aprendizado dos Transportes (SENAT)
- 11-Contribuição ao Serviço Nacional de Aprendizado Industrial (SENAI)
- 12-Contribuição ao Serviço Nacional de Aprendizado Rural (SENAR)
- 13-Contribuição ao Serviço Social da Indústria (SESI)
- 14-Contribuição ao Serviço Social do Comércio (SESC)
- 15-Contribuição ao Serviço Social do Cooperativismo (SESCOOP)
- 16-Contribuição ao Serviço Social dos Transportes (SEST)
- 17-Contribuição Confederativa Laboral (dos empregados)
- 18-Contribuição Confederativa Patronal (das empresas)
- 19-Contribuição de Intervenção do Domínio Econômico - CIDE Combustíveis (Lei 10336/2001)
- 20-Contribuição para Custeio do Serviço de Iluminação Pública (Emenda Constitucional 39/2002)
- 21-Contribuição para o Desenvolvimento da Indústria Cinematográfica Nacional - CONDECINE (art. 32 da Medida Provisória 2228-1/2001 e Lei 10.454/2002)
- 22-Contribuição Provisória sobre Movimentação Financeira (CPMF)
- 23-Contribuição Sindical Laboral
- 24-Contribuição Sindical Patronal
- 25-Contribuição Social Adicional para Reposição das Perdas Inflacionárias do FGTS (Lei Complementar 110/2001)
- 26-Contribuição Social para o Financiamento da Seguridade Social (COFINS)
- 27-Contribuição Social sobre o Lucro Líquido (CSLL)
- 28-Contribuições aos Órgãos de Fiscalização profissional (OAB, CRC, CREA, CRECI etc.)
- 29-Contribuições de Melhoria: asfalto, calçamento, esgoto, rede de água, rede de esgoto, etc.
- 30-Fundo Aeronáutico (FAER)

- 31-Fundo de Universalização dos Serviços de Telecomunicações - FUST (art. 6 da Lei http://www.marcoscindra.org/editorial/intagna/vejatambem_muda.asp?ID=829998/2000)
- 32-Fundo de Fiscalização das Telecomunicações - FISTEL (Lei 5070/1966 com novas disposições da Lei 9472/1997)
- 33-Fundo de Garantia por Tempo de Serviço (FGTS)
- 34-Fundo Especial de Desenvolvimento e Aperfeiçoamento das Atividades de Fiscalização (Fundaf) - (art.6 do Decreto-lei 1.437/1975 e art. 10 da IN SRF 180/2002)
- 35-Imposto sobre Circulação de Mercadorias e Serviços (ICMS)
- 36-Imposto sobre a Exportação (IE)
- 37-Imposto sobre a Importação (II)
- 38-Imposto sobre a Propriedade de Veículos Automotores (IPVA)
- 39-Imposto sobre a Propriedade Predial e Territorial Urbana (IPTU)
- 40-Imposto sobre a Propriedade Territorial Rural (ITR)
- 41-Imposto sobre a Renda e Proventos de Qualquer Natureza (IR - pessoa física e jurídica)
- 42-Imposto sobre Operações Financeiras (IOF)
- 43-Imposto sobre Serviços de Qualquer Natureza (ISS)
- 44-Imposto sobre Transmissão Bens Intervivos (ITBI)
- 45-Imposto sobre Transmissão Causa Mortis e Doação (ITCMD)
- 46-INSS Autônomos e Empresários
- 47-INSS Empregados
- 48-INSS Patronal
- 49-IPi (Imposto sobre Produtos Industrializados)
- 50-Programa de Integração Social (PIS) e Programa de Formação do Patrimônio do Servidor Público (PASEP)
- 51-Taxa de Autorização do Trabalho Estrangeiro
- 52-Taxa de Coleta de Lixo
- 53-Taxa de Combate a Incêndios
- 54-Taxa de Conservação e Limpeza Pública
- 55-Taxa de Controle e Fiscalização Ambiental - TCFA (Lei 10.165/2000)
- 56-Taxa de Controle e Fiscalização de Produtos Químicos (Lei 10357/2001, art. 16)
- 57-Taxa de Emissão de Documentos (níveis municipais, estaduais e federais)
- 58-Taxa de Fiscalização de Vigilância Sanitária (Lei 9782/1999, art. 23)
- 59-Taxa de Licenciamento Anual de Veículo
- 60-Taxa de Licenciamento para Funcionamento e Alvará Municipal
- 61-Taxa de Pesquisa Mineral DNPM (Portaria Ministerial 503/99)
- 62-Taxa de Serviços Administrativos - TSA - Zona Franca de Manaus (Lei 9960/2000)
- 63-Taxas ao Conselho Nacional de Petróleo (CNP)
- 64-Taxas CVM (Comissão de Valores Mobiliários)
- 65-Taxas de Outorgas (Radiodifusão, Telecomunicações, Transporte Rodoviário e Ferroviário, etc.)
- 66-Taxas de Saúde Suplementar - ANS (Lei 9.961/2000, art. 18)
- 67-Taxas do Registro do Comércio (Juntas Comerciais)

APÊNDICE B - CÓDIGO FONTE UTILIZADO NO PROCESSO DE ETL

```
%Programa...: Script Para Carga De Dados Para Prototipo Do Data
Warehouse De Tributos
%Autor.....: Marcelo Tadeu
%Data.....: 10/01/2005
%Observacoes:
```

```
localprocedure gera_arquivos()
  find all dwdimtempo
  out $messagebox($concat('DIMENSAO TEMPO: ', $setcount, \
    ' REGISTROS'), 3, 1, 1)
  pgtb913("AGUARDE, GERANDO ARQUIVO TXT (TEMPO)...", 1)
  find all dwdimtempo
  set out tmp5
  li all format $concat($mask(id_tempo, 'zzzzzzzz'), ', ', \
    $mask(data, 'YYYY-MM-DD'), ', ', \
    $mask(dia, '99'), ', ', $mask(mes, '99'), ', ', \
    $mask(ano, '9999'), ', ', \
    diasemana, ', ', mesano, ', ', $mask(bimestre, '9'), ', ', \
    $mask(trimestre, '9'), ', ', \
    $mask(quadrimestre, '9'), ', ', $mask(semester, '9'), ', ', \
    $mask(feriado, '9'))
  set out terminal
  pDPSDocCopy('', \
    $concat($trim($dbpath), '\\..\\RELAT\\tmp5'), \
    $concat('\\\\Dataaci_mtadeu\\PROJETO_DW\\dim_tempo', '.txt'))
  pgtb913($null, 3)
  find all dwdimcontr
  out $messagebox($concat('DIMENSAO CONTRIBUINTE: ',
    $setcount, ' REGISTROS'), 3, 1, 1)
  pgtb913("AGUARDE, GERANDO ARQUIVO TXT (CONTRIBUINTE)...", 1)
  find all dwdimcontr
  set out tmp5
  li all format $concat($mask(id_contr, 'zzzzzzzz9'), ', ', \
    $mask(codigo, 'zzzzzzzz9'), ', ', \
    nome, ', ', $mask(tipo, '9'), ', ', tipodesc, ', ', cpf, \
    cnpj, ', ', ende, ', ', \
```

```

    $mask(numero,'zzz9'),'',' bairro','','complemento','',' \
    cidade','',' \
    estado','',' telefone','',' cep','','email)
set out terminal
pDPSDocCopy('',' \
    $concat($trim($dbpath),'\\..\RELAT\$tmp5'),'',' \
    $concat('\\\\Dataaci_mtadeu\\PROJETO_DW\\dim_contribuinte',\
    '.txt'))
pgtb913($null,3)
find all dwdimdebito
out $messagebox($concat('DIMENSAO DEBITO: ','$setcount,\
    ' REGISTROS'),3,1,1)
pgtb913("AGUARDE, GERANDO ARQUIVO TXT (DEBITO)...",1)
find all dwdimdebito
set out tmp5
li all format $concat($mask(id_debito,'zzzzzzz9'),'','\
    $mask(numerodebito,'zzzzzzz9'),'','\
    $mask(anodebito,'zzz9'),'','\
    $mask(valortotal,'zzzzzzzzzz9.99'),'',' \
    $mask(parcelas,'z9'),'',' \
    $mask(valor1parcela,'zzzzzzzzzz9.99'),'',' \
    $mask(valorparcela,'zzzzzzzzzz9.99'),'',' \
    $mask(datacalculo,'YYYY-MM-DD'),'',' \
    $mask(dataemissao,'YYYY-MM-DD'),'',' \
    $mask(data1vencto,'YYYY-MM-DD'),'',' \
    $mask(tipovencimento,'9'),'',' tipovencdesc','','\
    $mask(situacao,'z9'),'',' situacaodesc','',' setor','','\
    $mask(tipo,'9'),'',' \
    tipodesc','',' siglaorgao','',' usuario','',' \
    {$mask(data1ancldivat,'YYYY-MM-DD') wh \
    data1ancldivat is not $null,0}','',' \
    $mask(numeroautoinfr,'zzzzzzzzz9'),'',' \
    $mask(numerocarne,'zzzzzzz9'),'',' \
    $mask(anocarne,'zzz9'))
set out terminal
pDPSDocCopy('',' \
    $concat($trim($dbpath),'\\..\RELAT\$tmp5'),'',' \
    $concat('\\\\Dataaci_mtadeu\\PROJETO_DW\\dim_debito',\
    '.txt'))
pgtb913($null,3)
find all dwdimorigdeb
out $messagebox($concat('DIMENSAO ORIGEM DO DEBITO: ','\
    $setcount,' REGISTROS'),3,1,1)
pgtb913("AGUARDE, GERANDO ARQUIVO TXT (ORIGEM DO DEBITO).."\
    ,1)
find all dwdimorigdeb
set out tmp5
li all format $concat($mask(id_origdeb,'zzzzzzz9'),'','\
    $mask(numerodebito,'zzzzzzz9'),'',' \

```

```

    $mask(anodebito,'zzz9'),'',' $mask(tipo,'z9'),'',' \
    tipodesc,'',' \
    $mask(chaveident,'zzzzzzzzzzzzzzzz9'),'',' \
    $mask(atividade,'zzzz9'),'',' \
    descatividade,'',' $mask(subatividade,'zzzz9'),'',' \
    descsubatividade,'',' $mask(aliquota,'zz9.99'))
set out terminal
pDPSDocCopy('',' \
    $concat($trim($dbpath),'\\..\RELAT\$tmp5'),'',' \
    $concat('\\\\Dataaci_mtadeu\PROJETO_DW\dim_origemdebito',\
    '.txt'))
pgtb913($null,3)
find all dwdimnatureza
out $messagebox($concat('DIMENSAO NATUREZA: ', $setcount, '
REGISTROS'),3,1,1)
pgtb913("AGUARDE, GERANDO ARQUIVO TXT (NATUREZA)...",1)
find all dwdimnatureza
set out tmp5
li all format $concat($mask(id_natureza,'zzzzzzz9'),'',' \
    $mask(natureza,'zz9'),'','descricao)
set out terminal
pDPSDocCopy('',' \
    $concat($trim($dbpath),'\\..\RELAT\tmp5'),'',' \
    $concat('\\\\Dataaci_mtadeu\PROJETO_DW\dim_natureza',\
    '.txt'))
pgtb913($null,3)
find all dwfatitemdeb
out $messagebox($concat('FATO ITEM DEBITO: ', $setcount, '
REGISTROS'),3,1,1)
pgtb913("AGUARDE, GERANDO ARQUIVO TXT (ITEM DEBITO)...",1)
find all dwfatitemdeb
set out tmp5
li all format $concat($mask(id_tempo,'zzzzzzz9'),'',' \
    $mask(id_contr,'zzzzzzz9'),'',' \
    $mask(id_debito,'zzzzzzz9'),'',' \
    $mask(id_origdeb,'zzzzzzz9'),'',' \
    $mask(id_natureza,'zzzzzzz9'),'','valoritem)
set out terminal
pDPSDocCopy('',' \
    $concat($trim($dbpath),'\\..\RELAT\tmp5'),'',' \
    $concat('\\\\Dataaci_mtadeu\PROJETO_DW\fat_itemdeb',\
    '.txt'))
pgtb913($null,3)
find all dwdimlocal
out $messagebox($concat('DIMENSAO LOCAL: ', $setcount, '
REGISTROS'),3,1,1)
pgtb913("AGUARDE, GERANDO ARQUIVO TXT (LOCAL)...",1)
find all dwdimlocal
set out tmp5

```

```

li all format $concat($mask(id_local,'zzzzzzz9'),'','\
    $mask(codigo,'zz9'),'','banco','','agencia','',' nome)
set out terminal
pDPSDocCopy('', \
    $concat($trim($dbpath),'\\..\RELAT\$tmp5'),'',' \
    $concat('\\\\Dataci_mtadeu\PROJETO_DW\dim_local',\
        '.txt'))
pgtb913($null,3)
find all dwdimtipobaixa
out $messagebox($concat('DIMENSAO TIPO DE BAIXA: ',\
    $setcount,' REGISTROS'),3,1,1)
pgtb913("AGUARDE, GERANDO ARQUIVO TXT (TIPO DE BAIXA)...",1)
find all dwdimtipobaixa
set out tmp5
li all format $concat($mask(id_tipobaixa,'zzzzzzz9'),'','\
    $mask(numerodebito,'zzzzzzz9'),'',' \
    $mask(anodebito,'zzz9'),'','\
    $mask(parcela,'z9'),'',',$mask(parcelaseq,'z9'),'',' \
    $mask(datacredito,'YYYY-MM-DD'),'',' \
    $mask(datavencimento,'YYYY-MM-DD'),'','\
    valorcalculado,'',',$mask(codigobaixa,'zz9'),'','descricao)
set out terminal
pDPSDocCopy('', \
    $concat($trim($dbpath),'\\..\RELAT\$tmp5'),'',' \
    $concat('\\\\Dataci_mtadeu\PROJETO_DW\dim_tipobaixa',\
        '.txt'))
pgtb913($null,3)
find all dwfatpagto
out $messagebox($concat('FATO PAGAMENTO: ', $setcount,'
REGISTROS'),3,1,1)
pgtb913("AGUARDE, GERANDO ARQUIVO TXT (PAGAMENTOS)...",1)
find all dwfatpagto
set out tmp5
li all format $concat($mask(id_tempo,'zzzzzzz9'),'','\
    $mask(id_local,'zzzzzzz9'),'',' \
    $mask(id_debito,'zzzzzzz9'),'',' \
    $mask(id_contr,'zzzzzzz9'),'',' \
    $mask(id_tipobaixa,'zzzzzzz9'),'','valorpago,'','\
    valordesconto,'','valormulta,'','valorjuros)
set out terminal
pDPSDocCopy('', \
    $concat($trim($dbpath),'\\..\RELAT\$tmp5'),'',' \
    $concat('\\\\Dataci_mtadeu\PROJETO_DW\fat_pagto',\
        '.txt'))
pgtb913($null,3)
endprocedure

localprocedure grava_tempo()
    let vudcontrolador[1] = vudcontrolador[1] + 1

```

```

add 1 dwdimtempo let id_tempo = vudcontrolador[1] \
                    data      = vdadebidtcalc \
                    dia       = $substring(vdadebidtcalc,7,2)\
                    mes       = $substring(vdadebidtcalc,5,2)\
                    ano       = $substring(vdadebidtcalc,1,4)\
                    diasemana = $dayname(vdadebidtcalc) \
                    mesano    = $monthname(vdadebidtcalc) \
bimestre = {1 wh $substring(vdadebidtcalc,5,2) in(01,02),\
            2 wh $substring(vdadebidtcalc,5,2) in(03,04),\
            3 wh $substring(vdadebidtcalc,5,2) in(05,06),\
            4 wh $substring(vdadebidtcalc,5,2) in(07,08),\
            5 wh $substring(vdadebidtcalc,5,2) in(09,10),\
            6 wh $substring(vdadebidtcalc,5,2) in(11,12),0}\
trimestre = {1 wh $substring(vdadebidtcalc,5,2) in(01,02,03),\
             2 wh $substring(vdadebidtcalc,5,2) in(04,05,06),\
             3 wh $substring(vdadebidtcalc,5,2) in(07,08,09),\
             4 wh $substring(vdadebidtcalc,5,2) in(10,11,12),0}\
quadrimestre = {1 wh $substring(vdadebidtcalc,5,2) \
                in(01,02,03,04),\
                2 wh $substring(vdadebidtcalc,5,2) \
                in(05,06,07,08),\
                3 wh $substring(vdadebidtcalc,5,2) \
                in(09,10,11,12),0}\
semester = {1 wh $substring(vdadebidtcalc,5,2) \
            in(01,02,03,04,05,06),\
            2 wh $substring(vdadebidtcalc,5,2) \
            in(07,08,09,10,11,12),0}\
feriado = {1 wh $substring(vdadebidtcalc,5,4) \
           in('0101','0421','0501','0526','0907',\
             '1012','1102','1115','1225','1231'),0}\
let vudidtempo = vudcontrolador[1]
endprocedure

localprocedure grava_contribuinte()
let vudcontrolador[2] = vudcontrolador[2] + 1
find 1 uccada uccadatpca uctpca wh cadacodigo =
vdacadacodigo -> scada
let vdatpcatcontr = cadatipo
busca_end()
set currentset scada
add 1 dwdimcontr from scada\
    let id_contr      = vudcontrolador[2] \
    codigo           = vdacadacodigo \
    nome             = $translate(cadanome,',','.') \
    tipo             = $tonumber(uccada.cadatipo,0) \
    tipodesc        = tpcadescr \
    cpf = {uccada.cadacpfnum wh uccada.cadacpfnum <>
           '',''} \
    cnpj = {uccada.cadacgc wh uccada.cadacgc <> ''}

```

```

        '           ' } \
ende = $concat({vuetplgcod wh vuetplgcod is not
               $null,''},' ', \
{vuetituabr wh vuetituabr is not $null,''},' ', \
{vuedescr  wh vuedescr  is not $null,''})\
numero = {vgmnumrua wh vgmnumrua is not $null,0}\
bairro = {vgmbair wh vgmbair <> ' ',' '}\
complemento = {$translate(vgmcomple,' ',' ') wh
               vgmcomple <> ' ',' ' } \
cidade      = {vgmcidade wh vgmcidade <> ' ',' '}\
cep         = {vdacepr wh vdacepr <> ' ',' '}\
estado      = {vgmcauf wh vgmcauf <> ' ',' '}\
telefone = $concat({$translate(cadatelr,' ',' '.'')
                   wh cadatelr <> ' ',' '},\
                   {$translate(cadatelc,' ',' ')
                   wh cadatelc <> ' ',' '})\
email       = {$translate(cadaemail,' ',' ') wh
               cadaemail <> ' ',' ' }

let vudidcontr = vudcontrolador[2]
endprocedure

localprocedure busca_end()
let vucnumrua = $null vuccomple = $null vuezonacod = $null\
vuelogrcod = $null vuetplgcod = $null vuetituabr = $null\
vuedescr  = $null vucbair = $null vdacepr      = $null \
vuclogrcep = $null vuctpcatcontr = $null vdaflag2 = 0
let vgmprioridade[1] = "3" vgmprioridade[2] = "2" \
vgmprioridade[3] = "1" vgmprioridade[4] = "9" \
vuccadacodigo = vdacadacodigo \
vuctpcatcontr = vdatpcatcontr \
vgmtipoend = "L" vgmteletipo[1]= "2" vgmteletipo[2]= "9"
pggm903
let vucnumrua=vgmnumrua vuccomple=vgmcomple vucbair=vgmbair\
vdacepr = vgmlogrcep vuclogrcep = vgmlogrcep
let vdacepr = $tocharacter($substring(vuclogrcep,5,8),8)
if vdacepr is $null or vdacepr = "" or vdacepr = 0
let vdacepr = "29300000"
endif
if vucnumrua is $null
let vucnumrua = 0
endif
if vuetplgcod is $null
let vuetplgcod = " "
endif
if vgmcidade is $null or vgmcidade = ""
let vgmcidade = "CACH. DE ITAPEMIRIM"
endif
if vgmcauf is $null or vgmcauf = ""
let vgmcauf = "ES"

```

```

endif
endprocedure

```

```

localprocedure grava_debito()
  find 1 uddebi wh debicarne = vudcarne -> setdebi2
  let vudcontrolador[3] = vudcontrolador[3] + 1
  add 1 dwdimdebito from setdebi2\
    let id_debito      = vudcontrolador[3] \
       numerodebito   = debicarne\
       anodebito      = $tonumber(uddebi.debiexerc,0) \
       valortotal     = {(debitotdeb * 16.26) wh
       debitotdeb is not $null,0} \
       parcelas =
         {$tonumber((debipardeb+debiparpag),0) wh \
         debipardeb is not $null and debiparpag is\
         not $null,0} \
         valorparcela = {(debivallpar * 16.26) wh\
         debivallpar is not $null,0} \
         valorparcela = {(debivalpar * 16.26) wh\
         debivalpar is not $null,0} \
       datacalculo = {debitdcalc wh debitdcalc is not \
         $null,$null} \
       dataemissao = {debittemis wh debittemis is not \
         $null,debitdcalc} \
       dataalvencito = {debitdvenc wh debitdvenc is not\
         $null,debitdcalc} \
       tipovencimento = {$tonumber(debitpvenc,0) wh \
         debitpvenc is not $null and \
         debitpvenc <> ' ',0}\
       tipovencdesc = {'ANTES DO FIM DE SEMANA/FERIADO'\
         wh debitpvenc = '1', \
         'APOS O FIM DE SEMANA/FERIADO ' \
         wh debitpvenc = '2',' '}\
       situacao = {$mask($tonumber(debistatus,0),'z9')\
         wh debistatus is not $null and
         debistatus <> ' ',0} \
       situacaodesc = {'NAO PAGO' wh debistatus = '0', \
         'PARCELADO' wh debistatus = '1',\
         'QUITADO' wh debistatus = '2', \
         'CARNE DIVAT' wh debistatus='3',\
         'PARCELADO DIVAT' wh debistatus = '4',\
         'ESPECIAL' wh debistatus = '5', \
         'ANISTIA' wh debistatus = '6', \
         'CARNE TRIBUTARIO' wh debistatus = '7',\
         'PARCELADO TRIBUTARIO' wh debistatus = '8',\
         'CANCELAMENTO TOTAL' wh debistatus = '9',' '}\
       Setor = {aplicodigo wh aplicodigo is not $null\
         and aplicodigo <> ' ',' ' } \
       tipo = {$tonumber(debitipo,0) wh \

```



```

        uddebi.debichave <> ' ',0}\
atividade = {vftativcodigo wh vftativcodigo <> 0 and\
            vftativcodigo is not $null,0}\
descatividade = {$translate(vftativdescri,',','.')}\
            wh vftativdescri <> ' ' and \
            vftativdescri is not $null,' '}\
subatividade = {vftsaticodigo wh vftsaticodigo <> 0\
            and vftsaticodigo is not $null,0}\
descsubatividade = {$translate(vftsatidescri,',','.')}\
            wh vftsatidescri <> ' ' and \
            vftsatidescri is not $null,' '}\
aliquota = {vtbaliquota wh vtbaliquota <> 0 \
            and vtbaliquota is not $null,0}

    let vudidorigdeb = vudcontrolador[4]
endprocedure

localprocedure busca_atividade()
    find all tbtrib wh tribinscri = vtbinscricao1 -> nstb072
    find all nstb072 tbtribatsa tbativ tbativsati \
        tbsati -> nstb115
    compute 1 nstb115 eval(let vftativcodigo=ativ.ativcodigo) \
        (let vftativdescri=tativ.ativdescr) \
        (let vftsaticodigo=bsati.saticodigo) \
        (let vftsatidescri = tbsati.satidescr)

    if $membercount<> 0
        find all tbexer wh ativcodigo = vftativcodigo and \
            exermes is not $null and exermes <> "00" \
            sorted by exercicio descending \
            exermes descending -> nstb002
        find 1 nstb002 -> nstb002
        if $setcount <> 0
            let vtbaliquota = tbexer.exeraliquo
        else
            let vtbaliquota = 5.00
        endif
    endif
endprocedure

localprocedure grava_itemdebi()
    let vipcontrole = 0
    find all uditem wh debicarne = vudcarne -> setitem
    let vipcontrole = $setcount
    while (let vipcontrole = vipcontrole - 1) >= 0
        set currentset setitem
        let vudnatucodigo = uditem.natucodigo
        find 1 dwdimnatureza wh natureza = vudnatucodigo
        if $setcount = 0
            grava_natureza()
        else

```

```

        let vudidnatureza = id_natureza
    endif
    set currentset setitem
    let vudcontrolador[6] = vudcontrolador[6] + 1
    add 1 dwfatitemdeb from setitem let id_tempo=vudidtempo \
        id_contr = vudidcontr \
        id_debito = vudiddebito\
        id_origdeb = vudidorigdeb \
        id_natureza = vudidnatureza \
        valoritem = {(uditem.itemvalor * 16.26) wh
                    uditem.itemvalor is not $null,0}
    set currentset setitem
    next 1 setitem
endwhile
endprocedure

```

```

localprocedure grava_natureza()
    let vudcontrolador[5] = vudcontrolador[5] + 1
    find 1 udnatu wh natucodigo = vudnatucodigo -> setnatu
    add 1 dwdimnatureza from setnatu let \
        id_natureza = vudcontrolador[5] \
        natureza = udnatu.natucodigo \
        descricao = $translate(udnatu.natudescr,',','.')
    let vudidnatureza = vudcontrolador[5]
endprocedure

```

```

localprocedure grava_pagto()
    let vipcontrole = 0 \
        vudlocacodigo = ''
    find all udpaga wh debicarne = vudcarne and \
        pagadtpgto is not $null -> setpaga
    let vipcontrole = $setcount
    while (let vipcontrole = vipcontrole - 1) >= 0
        set currentset setpaga
        let vudlocacodigo = udpaga.locacodigo \
            vdadebidtcalc = udpaga.pagadtpgto \
            vudbaixcodigo = udpaga.baixcodigo
        find 1 dwdimlocal wh codigo = vudlocacodigo
        if $setcount = 0
            grava_local()
        else
            let vudidlocal = dwdimlocal.id_local
        endif
        set currentset setpaga
        compute 1 udbaix wh baixcodigo = vudbaixcodigo \
            eval(let vudbaixdescr = $translate(baixdescr,',','.'))
        if $membercount = 0
            let vudbaixcodigo = '00' \
                vudbaixdescr = 'NAO INFORMADO'

```

```

endif
set currentset setpaga
find 1 dwdimtempo wh data = vdadebidtcalc
if $setcount = 0
    grava_tempo()
else
    let vudidtempo = dwdimtempo.id_tempo
endif
set currentset setpaga
verifica_eventos()
set currentset setpaga
let vudcontrolador[8] = vudcontrolador[8] + 1
add 1 dwdimtipobaixa from setpaga
    let id_tipobaixa = vudcontrolador[8] \
        numerodebito = debicarne \
        anodebito = $tonumber(udpaga.debiexerc,0)\
        parcela = $tonumber(udpaga.pagaparcels,0)\
        parcelaseq = udpaga.pagasequen \
        datacredito = {udpaga.pagadtcred wh \
            pagadtcred is not $null, \
            $adddays(pagadtpgto,2)} \
        datavencimento = {udpaga.pagadtvenc wh \
            pagadtvenc is not $null, \
            udpaga.pagadtpgto} \
        valorcalculado = {(udpaga.pagaupfcal * 16.26)\
            wh pagaupfcal is not $null,0}\
        codigobaixa = $tonumber(vudbaixcodigo,0) \
        descricao = vudbaixdescr
let vudidtipobaixa = vudcontrolador[8]
set currentset setpaga
let vudcontrolador[9] = vudcontrolador[9] + 1
add 1 dwfatpagto from setpaga let id_tempo= vudidtempo \
    id_debito = vudiddebito \
    id_contr = vudidcontr \
    id_tipobaixa = vudidtipobaixa
    id_local = vudidlocal \
    valorpago = {(udpaga.pagaupfpag * 16.26) wh \
        pagaupfpag is not $null,0} \
    valordesconto = vudvalorevento[1]\
    valormulta = vudvalorevento[2] \
    valorjuros = vudvalorevento[3]
set currentset setpaga
next 1 setpaga
endwhile
endprocedure

```

```

localprocedure grava_local()
    let vudcontrolador[7] = vudcontrolador[7] + 1
    find 1 udloca wh locacodigo = vudlocacodigo -> setloca

```

```

add 1 dwdimlocal from setloca let \
    id_local = vudcontrolador[7] \
    codigo = $tonumber(vudlocacodigo,0) \
    banco = locabanco \
    agencia = locaagenc \
    nome = $translate(locanome,',','.')
let vudidlocal = vudcontrolador[7]
endprocedure

localprocedure verifica_eventos()
let vudvalorevento[1] = 0 vudvalorevento[2] = 0 \
vudvalorevento[3] = 0
compute all udevenpaga wh debicarne = vudcarne eval \
    (let vudvalorevento[1]= \
    {($total(udevenpaga.evpavalor)*16.26) wh tpevcodigo\
in('01','02','07'),0}) \
    (let vudvalorevento[2]= \
    {($total(udevenpaga.evpavalor)* 16.26) wh tpevcodigo\
in('03','08','11','13','15'),0}) \
    (let vudvalorevento[3]= \
    {($total(udevenpaga.evpavalor) * 16.26) wh tpevcodigo\
in('04','09','12','14','16'),0})
endprocedure

%-----INICIO-----%
clear
set info off
let vudcontrolador[1]=0 vudcontrolador[2]=0 vudcontrolador[3]=0
vudcontrolador[4]=0 vudcontrolador[5]=0 vudcontrolador[6]=0 \
vudcontrolador[7]=0 vudcontrolador[8]=0 vudcontrolador[9]=0 \
vudcontrolador[10]=0 vudcontador=0 vdadebidtcalc= $null \
vdacadacodigo=0 vipcontrole=0 vudcarne=0 vudnatucodigo=$null \
vudidtempo=0 vuddtpgto=$null vftativcodigo=0 vftsaticodigo = 0\
vftativdescri = '' vftsaticodescri = '' vtbaliquota = 0 \
vtbinscricao1 = 0 vuddebiexerc = ''
dc_dialog("Atencao!","PREPARAR TABELAS", \
2,"SIM","NAO",2,1,14,17,vtbfunc2)
if vtbfunc2 = 1
pgtb913("AGUARDE, LIMPANDO DIMENSOES...",1)
del all dwcontador
del all dwdimtempo
del all dwdimlocal
del all dwdimnatureza
del all dwdimcontr
del all dwdimdebito
del all dwdimorigdeb
del all dwdimtipobaixa
del all dwfatitemdeb
del all dwfatpagto

```

```

        out $messagebox('TABELAS PRONTAS PARA PROCESSAMENTO',3,1,1)
    endif
    let vuddebiexerc = '1982'
    find 1 dwdimdebito
    if $setcount <> 0
        while
            find 1 dwdimdebito wh anodebito = vuddebiexerc
            if $setcount <> 0
                let vuddebiexerc = vuddebiexerc + 1
                continue
            else
                compute 1 dwcontador eval \
                    (let vudcontrolador[1] = id_tempo      ) \
                    (let vudcontrolador[3] = id_debito    ) \
                    (let vudcontrolador[2] = id_contr     ) \
                    (let vudcontrolador[5] = id_natureza  ) \
                    (let vudcontrolador[4] = id_origdeb   ) \
                    (let vudcontrolador[7] = id_local     ) \
                    (let vudcontrolador[8] = id_tipobaixa)

                break
            endif
        endwhile
    endif
    while
        dc_dialog("Atencao!", $concat("Processar Exercicio de: ", \
            vuddebiexerc, " ?"), 2, "SIM", "NAO", 1, 1, 14, 17, vtbfunc2)
        if vtbfunc2 = 1
            pgtb913($null, 3)
            pgtb913("AGUARDE, PROCESSANDO...", 1)
            find all uddebi wh debiexerc = vuddebiexerc -> setdebi
            let vudcontador = $setcount
            while (let vudcontador = vudcontador - 1) >= 0
                let frum007.msg3 = $concat("CARREGANDO DIMENSOES E
FATOS ", $mask(vudcontador, "9999999"))
                form display frum007.msg3
                set currentset setdebi
                let vdacadacodigo = uddebi.cadacodigo \
                    vudcarne      = uddebi.debicarne
                if uddebi.debidtcalc is $null or \
                    uddebi.debitotdeb is $null or \
                    uddebi.debitotdeb = 0 or \
                    uddebi.debivalpar is $null or uddebi.debivalpar = 0
                    set currentset setdebi
                next 1 setdebi
                continue
            endif
            let vdadebidtcalc = uddebi.debidtcalc
            find 1 dwdimtempo wh data = vdadebidtcalc
            if $setcount = 0

```

```

        grava_tempo()
    else
        let vudidtempo = dwdimtempo.id_tempo
    endif
    find 1 dwdimcontr wh codigo = vdacadacodigo -> scadadw
    if $setcount = 0
        grava_contribuinte()
    else
        let vudidcontr = dwdimcontr.id_contr
    endif
    grava_debito()
    grava_origdeb()
    grava_itemdebi()
    find all udpaga wh debicarne = vudcarne -> setpaga
    if $setcount <> 0
        grava_pagto()
    endif
    set currentset setdebi
    next 1 setdebi
endwhile
else
    find dwcontador -> setconta
    if $setcount = 0
        add 1 dwcontador let \
            id_contador = 1 \
            id_tempo = vudcontrolador[1] \
            id_debito = vudcontrolador[3] \
            id_contr = vudcontrolador[2] \
            id_natureza = vudcontrolador[5] \
            id_origdeb = vudcontrolador[4] \
            id_local = vudcontrolador[7] \
            id_tipobaixa = vudcontrolador[8]
    else
        ch 1 setconta let \
            id_tempo = vudcontrolador[1] \
            id_debito = vudcontrolador[3] \
            id_contr = vudcontrolador[2] \
            id_natureza = vudcontrolador[5] \
            id_origdeb = vudcontrolador[4] \
            id_local = vudcontrolador[7] \
            id_tipobaixa = vudcontrolador[8]
    endif
    break
endif
let vuddebiexerc = vuddebiexerc + 1
endwhile
pgtb913($null,3)
dc_dialog("Atencao",$concat("Gerar Arquivos de Exportacao?"), \
    2,"SIM","NAO",1,1,14,17,vtbfunc2)

```

```
if vtbfunc2 = 1
    gera_arquivos()
endif
pgtb913($null,3)
set info on
return
```

APÊNDICE C - QUESTIONÁRIO UTILIZADO NO LEVANTAMENTO

Levantamento de Requisitos para uma Ferramenta Gerencial

Nome _____ Setor _____

O objetivo deste formulário é de auxiliar no levantamento das necessidades gerenciais do setor de tributos municipais. A solução que será proposta é construção de um Data Warehouse com dados geográficos que auxiliará nas tomadas de decisão. Para este fim, estas perguntas deverão ser respondidas com a maior proximidade possível da realidade diária.

Atualmente, quais informações gerenciais, referentes a tributos, têm-se maior necessidade e conseqüentemente maior dificuldade de obtê-las?

Que tipo de resumos, ou totais, seria de maior utilidade? Qual a melhor forma de visualizá-los?

Qual o nível de detalhamento do tributo que você considera importante do ponto de vista gerencial?

() por itens que compõem o débito

() por débito

Com relação a pagamento de tributos, qual a forma de agrupamento que você considera mais relevante?

- por débito
- diário
- semanal
- mensal

Como são tratadas (tributação) as filiais das empresas?

Com relação há anos anteriores, qual a melhor forma de visualização dos dados? Ou seja, que tipo de agrupamento atenderia as necessidades gerenciais?

Qual a importância de vincular informações tributárias com informações geográficas?

Neste caso, objeto do tributo seria tratado como um “ponto” ou “polígono”?

Há uma necessidade de análise do endereço do contribuinte do ponto de vista espacial?

Observações
