

# APLICAÇÃO DA TECNOLOGIA DE SISTEMAS DE SUPORTE À DECISÃO NO PLANEJAMENTO URBANO MUNICIPAL: UMA PROPOSTA PARA O MUNICÍPIO DE MACAÉ

**Alan Carvalho Galante (UCAM) (1)**  
**Jorge Luís Nunes e Silva Brito (UERJ) (2)**

**(1) Universidade Cândido Mendes - Campos (UCAM)**  
Rua Anita Peçanha, nº 100 - Parque São Caetano - Campos dos Goytacazes/RJ  
alancarvalho@gmail.com

**(2) Universidade do Estado do Rio de Janeiro**  
**Programa de Pós-Graduação em Engenharia da Computação – Geomática**  
Rua São Francisco Xavier, 524 – Maracanã  
Rio de Janeiro – RJ – CEP 20559-900  
jnunes@uerj.br

## RESUMO

O município de Macaé representa uma das maiores economias do Estado do Rio de Janeiro. Com o uso da tecnologia da informação é possível criar uma ferramenta poderosa que dê suporte no processo de tomada de decisão do município, ajudando na melhoria da qualidade de vida de seus habitantes. Baseado nesta idéia, o uso de Sistema de Suporte à Decisão permite fornecer diferentes tipos de informação sobre áreas do município, como saúde e educação. Para a união de todas as informações é usada a tecnologia de *datawarehouse* e para as consultas são usadas em conjunto as tecnologias de OLAP e SIG. Com essas tecnologias unidas foi possível construir uma ferramenta automatizada de auxílio ao processo decisório do município de Macaé.

Palavras chaves: Suporte a Decisão, Sistemas de Gestão do Conhecimento; Datawarehouse; On-line Analytical Processing (OLAP); Sistema de Informação Geográfica (SIG)

## ABSTRACT

The Macaé County is one of the greatest economy of the state of Rio de Janeiro. With the use of the information technology is possible to create a powerful tool for supporting the decision making processing for this County, aiding the process of improvement of life quality. For that one, intends to use a Decision Support System able to give different kind of information of County areas, like health and education. For the union of all information the datawarehouse technology will be used. For query implementation the technologies of OLAP and GIS are used together. Therefore, those technologies together make a powerful tool for aiding the decision making process of the Macaé County.

Keywords: Decision Support System, Knowledge-Management System; DataWarehouse, On-line Analytical Processing(OLAP); GIS

## 1 INTRODUÇÃO

O objetivo deste artigo é projetar um Sistema de Suporte à Decisão (SSD) que integre as tecnologias de *Datawarehouse* (DW), *On-line Analytical Processing* (OLAP) e SIG (Sistema de Informação Geográfica) para o auxílio nas tomadas de decisão no âmbito do município de Macaé.

Segundo Elmasri e Navathe (2001), um Sistema de Suporte à Decisão é um sistema que leva dados aos decisores de alto nível de uma organização para apoiar decisões complexas e importantes. O SSD possui conhecimentos específicos sobre o assunto e, mediante as normas de avaliação introduzidas pelo usuário, apresenta solução otimizada para o problema em questão, auxiliando o processo decisório. Pode-se destacar entre os SSD: o *Datawarehouse*, o *On-line*

*Analytical Processing* (OLAP) e os Sistemas de Informação Geográfica.

Segundo Inmon (1992) *Datawarehouse* corresponde a “uma coleção de dados orientada por assuntos, integrada, não volátil, variante no tempo, e que tem por objetivo dar suporte aos processos de tomada de decisão”. O *Datawarehouse* é orientado por assunto porque contém informações sobre temas específicos importantes para a empresa, tais como: produtos, vendas e clientes. A integração possibilita a obtenção de bases de dados diferentes, tanto tecnologicamente quanto geograficamente, num mesmo repositório de dados, de forma que a área gerencial da empresa tenha uma visão não somente deste ou daquele setor, mas da organização como um todo. A não volatilidade se refere ao fato de que o *Datawarehouse* recebe apenas uma carga de dados inicial e então fica pronto somente para consultas. A atualização é periódica e incremental, de acordo com uma política pré-definida em função de seu propósito. A variabilidade no tempo se refere ao fato de que o propósito do *Datawarehouse* é mostrar de que forma as informações contidas nele se comportam em função do tempo.

OLAP se refere a um conjunto de ferramentas que, de forma amigável e flexível, fornece ao usuário acesso e análise de dados *ad-hoc*, permitindo que *datawarehouses* sejam usados efetivamente para análise *on-line*, provendo respostas rápidas e interativas em consultas analíticas complexas.

Com o DW, o OLAP é capaz de aplicar suas técnicas de agregação, organização e sumarização em grandes volumes de dados, provendo ainda uma visualização *on-line* com o uso de ferramentas gráficas.

No contexto deste artigo, o SIG será entendido como um Sistema de Suporte à Decisão que tem como principal característica de análise o dado geográfico. Para isso, a tecnologia SIG utiliza-se de um conjunto de disciplinas e técnicas como geografia, cartografia, sensoriamento remoto, fotogrametria, modelagem de dados, etc. As técnicas utilizadas em SIG são denominadas de geoprocessamento.

Num SIG, os dados são organizados em camadas ou planos de informação. Aronoff (1995) conceituou camada de dados como “um conjunto de características e atributos geográficos relacionadas logicamente”. Desta forma, pode-se encontrar camadas do tipo rodoviária, ferroviária, hidrográfica, distritos, bairros, ruas, etc.

Para Navathe e Elmasri (2001) um SIG deve ser capaz de realizar as seguintes operações espaciais: interpolação, interpretação, análise de proximidade, processamento de imagens e análises de rede.

Um dos fatores com que se deve ter muita cautela no tratamento das informações, segundo Dias et. al. (2005) é a preservação dos dados individuais.

Para resolver esse problema usa-se a agregação dos dados, neste caso, dados geográficos.

## 2 O MUNICÍPIO DE MACAÉ

O município de Macaé está localizado na região Norte Fluminense, interior do estado do Rio de Janeiro, distando 182 quilômetros da capital. Este município teve grande projeção nacional após a instalação da sede da Petrobrás, que trouxe para a região um grande número de empresas prestadoras de serviços na área petrolífera. Graças aos recursos provenientes principalmente dos “*royalties*” e da presença da Petrobrás, Macaé hoje é considerado o quarto município em qualidade de vida no Estado, tendo, portanto, grande potencial para investimentos.

Localizado entre duas importantes capitais de Estado, Rio de Janeiro e Vitória, Macaé conta com boas malhas rodoviária e ferroviária, um aeroporto e um porto - hoje operado pela Petrobrás. Macaé tem como acessos principais a BR-101 e a Rodovia Amaral Peixoto (RJ 106). Estima-se sua população em cerca de 132.461 habitantes, de acordo com o Censo Demográfico realizado em 2000 pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

A Prefeitura de Macaé atualmente encontra-se organizada administrativamente em Órgãos e Unidades, dentre elas: Gabinete do Prefeito, Gabinete do Vice Prefeito, Gerência Municipal, Procuradoria Geral, Secretarias Municipais, Guarda Municipal de Macaé, etc. Cada Órgão ou Unidade Administrativa possui a liberdade de desenvolver ou contratar empresas para o desenvolvimento de sistemas para o controle e gerenciamento de suas atividades cotidianas.

Além das informações já existentes em seus órgãos ou Unidades Administrativas, a Prefeitura Municipal, com o propósito de levantar dados específicos, avaliar os serviços prestados pela Administração Municipal e direcionar os futuros investimentos do governo, desenvolveu um programa denominado *Macaé Cidadão*. Além desse programa, a Prefeitura Municipal está desenvolvendo o Projeto Geo-Macaé, que tem como propósito adequar o Município de Macaé às novas tecnologias no trato das informações e georreferenciar as informações municipais.

O Programa *Macaé Cidadão* consiste em uma pesquisa de campo realizada pela própria prefeitura, cujo propósito é o levantamento de informações sobre a população, com o objetivo de enriquecer as informações vindas do IBGE. Este programa foi motivado pela necessidade da Prefeitura de Macaé obter de sua população informações sobre diversos temas, tais como: educação, saúde, infraestrutura, trabalho e renda, esporte e lazer, indicadores sociais, indicadores urbanos, território, divisão administrativa, dados domiciliares e dados demográficos. A consolidação dos dados desses

diversos temas serve como base para elaboração de estrutura de dados (tabelas) capazes de identificar as carências e necessidades da população.

A proposta do projeto *Geo-Macaé* é a de adequar o município de Macaé às novas tecnologias no trato da informação. Trata-se de um Sistema de Informações Cartográficas (SIC), que tem como proposta georreferenciar as informações do município, de modo que se obtenha uma base cartográfica digital com a estrutura territorial da região, a partir da base de setores censitários do IBGE e do programa *Macaé-Cidadão*. Com o projeto *Geo-Macaé* ficam disponibilizados um conjunto de Mapas, com a Estrutura Territorial e Administrativa do Município, além de Mapas Temáticos.

### 3 O SSD DE MACAÉ

Para o desenvolvimento do SSD, foram usados os dados das Unidades Administrativas do município, do *Projeto Macaé Cidadão* e do *Projeto Geo-Macaé*.

Quanto às áreas de atuação ou de interesse imediato foram selecionadas a Saúde e a Educação, por terem sido julgadas como as áreas de maior interesse pela Prefeitura Municipal. As informações sobre saúde provêm do *Projeto Macaé Cidadão* (dados da população) e da Secretaria de Saúde (Postos de Saúde). As informações sobre a Educação são oriundas do *Projeto Macaé Cidadão* (dados da população) e da própria Secretaria Municipal de Educação (Escolas). Além dessas informações, também foram usados os cadastros mobiliário e imobiliário que também foram retirados do *Projeto Macaé Cidadão*. Todas essas informações foram unidas através do *Projeto Geo-Macaé*, que tem várias dessas informações georreferenciadas. Desta forma, o SSD estará atendendo ao propósito de trazer à tona as necessidades da população e os resultados já obtidos pelos programas implantados.

### 4 A DIVISÃO GEOGRÁFICA DO MUNICÍPIO

A menor unidade de análise a ser usada neste trabalho é o setor censitário que, para o Instituto Brasileiro de Geografia e Estatística (IBGE), corresponde à capacidade de levantamento do recenseador, variando por região em torno de 200 a 400 domicílios. Uma visão da estratégia para a modelagem do espaço geográfico macaense é apresentada na figura 1.

Pode-se considerar o Bairro como um conjunto de Setores Censitários, que também possui informações sobre tendências tão importantes quanto o Setor Censitário. Com esta camada geográfica é possível obter informações importantes acerca da população macaense. Pelos dados informados pela prefeitura, o município de Macaé possui hoje 27 bairros.

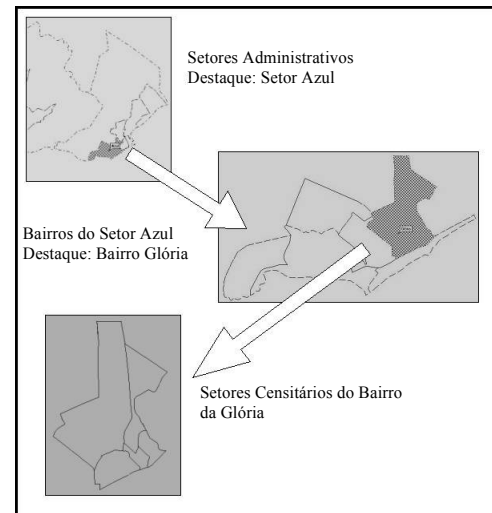


Fig. 1 - Representação do espaço geográfico do município de Macaé

Uma outra camada geográfica importante administrativamente para o município de Macaé denomina-se “Setor Administrativo”. Esta foi criada para facilitar a administração do município. Trata-se dos setores gerenciais dos quais existem 10 setores nomeados por cores.

### 5 Modelagem do SSD

#### 5.1 Modelando as dimensões geográficas agregáveis

Entende-se por dimensão geográfica agregável a toda aquela dimensão do modelo multidimensional do *Datawarehouse* que possui ligação direta com a tabela fato e, além disso, possui algum atributo geográfico.

Usando a UML, as três camadas geográficas são representadas conforme a figura 2, onde um Bairro é composto de vários setores censitários, e um setor administrativo composto de vários bairros. Conforme o tratamento usado no Projeto Geo-Macaé, considera-se todas as três camadas como figuras geométricas do tipo polígono. Além disso, usa-se o conceito de estereótipos para representar o tipo de figura geométrica que cada classe representa.

Para se evitar complexidades no modelo do *Datawarehouse* e evitar degenerações nas dimensões, que poderiam tornar o modelo pouco eficiente em termos de consulta, foi adotada a união de todas as classes em uma única classe como recurso para a construção do modelo físico da dimensão geográfica, conforme pode ser visto na figura 3. Apesar desse recurso gerar repetição de uma mesma informação várias vezes nesta classe, deve-se destacar que o propósito de um DW não é a facilidade de inclusão e exclusão como em um modelo para sistemas transacionais, e sim a eficiência das consultas.

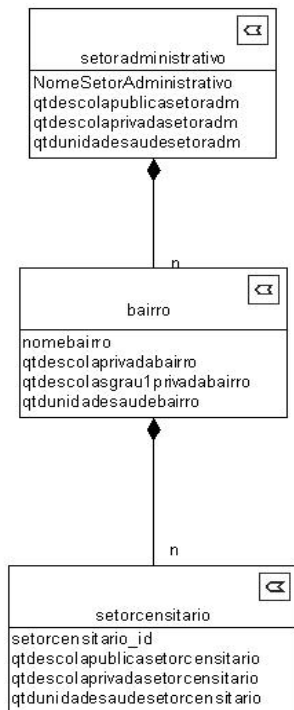


Fig. 2 - Representação com UML do espaço geográfico do município de Macaé

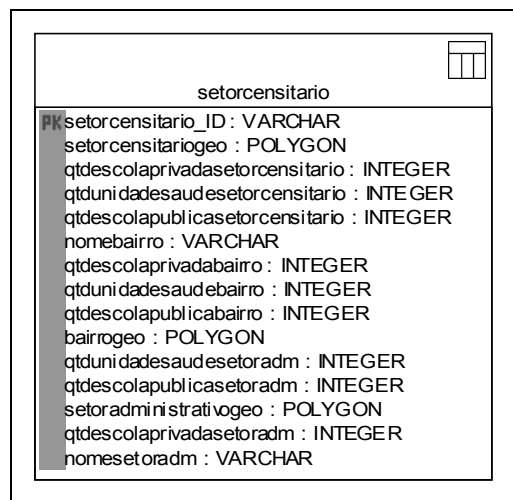


Fig. 3 - Modelo Físico das Classes Geográficas

Outro ponto que merece destaque é o surgimento de três atributos: setorcensitariogeo, bairrogeo e setoradministrativogeo. Estes três atributos correspondem ao armazenamento em banco de dados dos polígonos referentes a cada um setor censitário.

### 5.2 Modelando as dimensões geográficas não agregáveis

Consideram-se como classes geográficas não agregáveis as classes que não têm ligação direta com a tabela fato do modelo multidimensional. Essas classes, na verdade, estarão associadas a uma dimensão que seja geográfica, uma vez que a “chave” de ligação

entre a dimensão geográfica e a classe geográfica não agregável é o posicionamento geográfico.

No modelo usado para o DW implementado para a Prefeitura de Macaé, foram usadas as seguintes classes geográficas não agregáveis: escola e UnidadeSaúde. As Escolas e as Unidades de Saúde estão sendo consideradas do tipo ponto. Essas duas classes estão associadas ao modelo através da dimensão geográfica denominada setorcensitario.

### 5.3 Modelando as dimensões não geográficas agregáveis

Além das dimensões geográficas agregáveis, das dimensões geográficas não agregáveis e da tabela fato, existem as dimensões que estão ligadas à tabela fato e que não possuem nenhuma relação direta com o espaço geográfico. Essas dimensões serão explicadas a seguir.

A dimensão “condicaomoradia” serve para exibir as condições nas quais o cidadão está morando no município, tais como: a moradia possui filtro de água, saneamento básico, etc? A dimensão “tempo”, pertinente a praticamente todos os DW, existe nesse modelo com o propósito de armazenar o tempo em que as informações foram carregadas na base de dados.

Com essa dimensão será possível avaliar as mudanças no município ao longo do tempo. A dimensão “trabalho” exibe as informações do cidadão no que se refere a sua característica trabalhista, tais como: se trabalha, se é funcionário público, etc. A dimensão educacao abrange características do ponto de vista dos estudos do cidadão, tais como se ele estuda, qual o grau de escolaridade, etc. A dimensão queixasistemaude, exibe quais as principais queixas que um cidadão possui em relação à área de saúde do município. Na dimensão problemasaude, estão representadas as informações referentes aos problemas de saúde dos municípios. Na dimensão deficienciafisica, estão representados os dados referentes as deficiências físicas mais comuns entre os municípios.

### 5.4 O Modelo Multidimensional

A figura 4 exibe o modelo multidimensional físico, onde se deve dar destaque à tabela fato (indivíduo) que tem como sua chave primária as chaves de todas as dimensões agregáveis do modelo e a quantidade de pessoas, que corresponde a um contador, quando da ocorrência de indivíduos com as mesmas características.

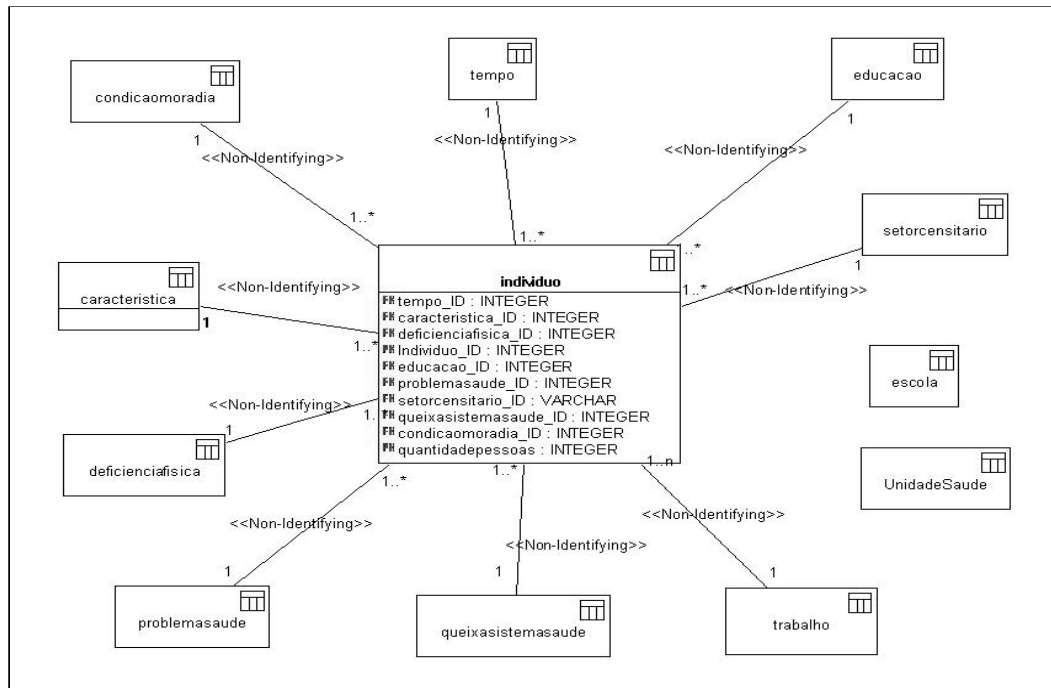


Fig. 4 - Modelo Multidimensional Físico

5.5 Tecnologia utilizada na implementação do SSD

Para o emprego do SSD há a necessidade do uso de tecnologias que implementam *Datawarehouse*, OLAP e SIG de maneira integrada. Uma das preocupações para a construção dessa ferramenta foi no uso de *softwares* livre, uma vez que poderia ser provado que qualquer município ou empresa é capaz de fazer um trabalho de tamanha importância a baixo custo.

O *PostgreSQL* corresponde a um banco de dados objeto-relacional Espacial, ou seja, é um banco de dados capaz de armazenar dados convencionais e dados geográficos. Quanto aos dados geográficos, este Sistema de Gerência de Banco de Dados (SGBD) é capaz não só de armazenar estes tipos de dados, mas também de executar funções geográficas sobre esses dados. O *PostgreSQL* se tornou um banco de dados muito interessante para ser utilizado neste trabalho uma vez que, além de todas essas capacidades ditas anteriormente, é uma ferramenta de distribuição gratuita o que viabiliza o seu uso não só para a aplicação corrente mas também para muitas aplicações do mesmo tipo. O *PostgreSQL* é um SGBD desenvolvido para funcionar também em ambientes *Linux*. Isto proporciona mais uma grande economia na implantação de um sistema com um banco de dados como este, uma vez que o *Linux* é um sistema operacional gratuito e livre com grande aceitação no mercado nos dias de hoje. Este reconhecimento se dá

em termos de funcionalidades, gerenciamento de redes de computadores e facilidade na interface com o usuário no que se refere ao uso nas estações de trabalho.

Na verdade, o *PostgreSQL* não possui nativamente todos os tipos e operações geográficas contempladas pelo Open GIS, que segundo Buhler e Mckee (2003) é uma organização internacional responsável pela criação de novas padronizações técnicas e comerciais para garantir interoperabilidade em SIG, buscando criar uma especificação de software e novas estratégias empresariais a fim de tornar os sistemas de geoprocessamento abertos e integrar completamente os dados geográficos e as operações necessárias para manipulá-los.

Esses tipos e operações são implementados através de uma extensão do *PostgreSQL* conhecida por *PostGIS*. Logo, este implementa as definições do *OpenGIS* para a SQL, permitindo desta forma que a SQL tenha funções sobre tipos geográficos.

Uma deficiência do *PostgreSQL* é a ausência de uma ferramenta OLAP integrada ao mesmo. Para isso, foi utilizado neste projeto o *PostGeoOLAP* que é uma ferramenta desenvolvida numa dissertação de mestrado da Universidade Cândido Mendes, cujo propósito, segundo Colonese (2004) é “a possibilidade de tornar mais simples a modelagem e implementação de um sistema de suporte à decisão que integre características analíticas de um *Datawarehouse* e geográficas de um SIG, fazendo com que tais conceitos possam coexistir no mesmo modelo, tanto em nível conceitual quanto em nível da implementação e, mais

ainda, que possa ser o mais direto possível tal mapeamento entre os citados níveis de abstração<sup>7</sup>. Na verdade, o *PostGeoOLAP* é uma ferramenta OLAP que permite consultas espaciais sobre dimensões geográficas. Não existe sumarização sobre os dados geográficos, ou seja, não é contemplada a possibilidade de se ter uma tabela fato com dados geográficos. Tem-se apenas dimensões (sejam agregáveis ou não agregáveis) com este tipo de característica.

Uma dificuldade dessa ferramenta é o tratamento dos dados. Deve haver um primeiro momento onde os dados armazenados no SGBD devem ser também armazenados no banco de dados multidimensional (*PostgreSQL*). Existe também um segundo momento, em que os mapas devem estar preparados, dentro de uma ferramenta denominada *Planet GIS*. Esta ferramenta é gratuita, com um potencial geográfico, mas com pouca ajuda disponível sobre como utilizá-la. As principais características do *software Planet GIS* são: interpretação dos principais formatos de troca (SHP, MIF, DXF, DGN, FEA, SDL, E00), aquisição e edição de todos os tipos de entidades geográficas, inclusive no formato *raster*; produção de mapas com grande potencial estético, criação de mapas temáticos de acordo com os atributos desejados, edição e visualização de atributos vinculados com um banco de dados, vinculação das entidades geográficas a uma tabela em diversos tipos de banco de dados, possibilitando exibir características completas de cada entidade.

De posse de toda a tecnologia necessária para a criação do SSD, o próximo passo é a migração para o *Datawarehouse* dos dados vindos da Prefeitura Municipal de Macaé.

### 5.6 A geração do Banco de Dados Multidimensional

A base de dados relacional foi gerada normalmente, através de um *script* com comandos SQL (DDL), de forma a gerar as tabelas e índices num banco de dados denominado *dwprefeitura*. É importante lembrar que, para agilizar o processo de consultas, foram criados índices para cada atributo de cada tabela. Cada tabela possui um número de índices sempre igual ao número de atributos mais um, pois além de existir um para cada atributo, também existe a chave primária.

Os dados convencionais das dimensões agregáveis foram retirados do Projeto Macaé Cidadão. Estes dados foram disponibilizados pela prefeitura, em *Microsoft SQL Server*. Para alimentar o modelo multidimensional, fez parte deste trabalho o desenvolvimento de um *software* que realizou todo o procedimento de limpeza, formatação dos dados e armazenamento. Este *software* realiza uma interação com pessoas identificadas no *Projeto Macaé Cidadão*, oriundos da base de dados da Prefeitura de Macaé, observando suas características em todas as dimensões do modelo do *Datawarehouse* (Condição de Moradia, Característica, Trabalho, Educação, Saúde, Queixa do

Sistema de Saúde, Deficiência Física). Caso essa pessoa possua alguma característica que ainda não tenha sido armazenada em alguma dimensão, ele cria aquela característica no modelo multidimensional para só então armazenar na tabela fato todas as chaves estrangeiras dessas dimensões e atribuir à quantidade de pessoas com aquelas características o valor 1.

Quanto à migração das dimensões geográficas não agregáveis, houve duas preocupações: o tratamento das informações geográficas e das informações não-geográficas ou transacionais. As informações não-geográficas foram provenientes dos órgãos responsáveis, que disponibilizaram os dados, no formato *Microsoft Excel*, sem nenhum referenciamento geográfico, tendo este sido realizado neste trabalho.

As informações sobre as escolas macaenses foram disponibilizadas pela prefeitura no formato *Microsoft Excel*, contendo as seguintes informações: Nome da escola, Endereço e foco de ensino da escola (educação infantil, pré-escolar, ensino fundamental, etc.). Essas informações estavam presentes em três planilhas distintas, uma das quais apresentava dados de escolas particulares, outra de escolas públicas municipais e outra de escolas públicas estaduais. Manualmente, foi gerada uma planilha única com todas as escolas e criado um atributo que identificava qual o tipo de escola (público municipal, público estadual ou particular). Nesta mesma planilha, foi feita uma análise sobre os focos de ensino das escolas e identificados os seguintes atributos: adultos, creche, maternal, pré-escolar, fundamental 1a. Série, Fundamental 1a. a 4a. Série, Fundamental 5a. a 8a. série, Ensino Médio, Técnico e Educação Especial. Desta forma, foi feita uma leitura linha a linha de cada escola, atribuindo-se sim ou não para cada uma dessas características. Uma vez formada essa planilha contendo 138 registros, bastou usar a própria aplicação do *Microsoft Excel* para exportar os dados para o *PostgreSQL* usando o ODBC para acessar os dados no servidor de banco de dados.

O procedimento usado no tratamento dos dados das Unidades de Saúde não foi diferente do adotado em relação às Escolas. As informações foram disponibilizadas numa planilha do *Microsoft Excel* constando o nome da Unidade de Saúde, o tipo (Ambulatórios Especializados, Núcleos de Atenção à Saúde, Unidades de Saúde da Família e Unidades Básicas de Saúde, Mista, Coordenadorias, Laboratórios, Ambulatórios, Hospitais), se é ou não associada ao SUS e o endereço. Para a exportação da planilha disponibilizada pela prefeitura foi usada a ferramenta de exportação de dados do *Microsoft Excel* que, através de um driver ODBC, levou os 62 registros para o modelo multidimensional.

O próximo passo correspondeu à transformação dos "endereços" que estavam registrados nas planilhas em pontos do espaço geográfico para que esses dados estivessem georreferenciados. Como a menor granularidade geográfica usada no modelo deste trabalho é o setor

cenitário, a preocupação era que as escolas e unidades de saúde estivessem posicionadas corretamente dentro do setor censitário respectivo. Usando os mapas disponibilizados pela Prefeitura Municipal de Macaé, o *software Planet GIS* e as localizações das dimensões não-agregáveis dadas no formato convencional de endereço, cada escola e unidade de saúde foram locadas manualmente uma a uma no mapa, gerando, mais duas camadas geográficas: *UnidadesSaude* e *Escolas*. O modelo do mapa usado no *software Planet GIS* passou a possuir cinco camadas geográficas distintas: Setor Administrativo, Bairro, Setor Censitário, Escolas e Unidades de Saúde. Os três primeiros foram disponibilizados pela Prefeitura Municipal de Macaé no formato *ArcView Shape File*, enquanto que os dois subseqüentes foram locados e inseridos manualmente integrando um único arquivo no formato do *software Planet GIS*.

O próximo passo foi levar as informações geográficas geradas no *Planet GIS* para o *Datawarehouse*. Para isso, foi usada uma ferramenta do *Planet GIS* que exporta uma camada geográfica para o formato de banco de dados *Dbase IV*. Dois arquivos *Dbase* foram gerados, um com as Unidades de Saúde e outro com as Escolas. Nestes arquivos, consta o nome de cada ponto e os valores de X e Y, correspondentes ao posicionamento geográfico planimétrico em coordenadas cartesianas. Usando o *Visual Basic 6.0*, foi criado um programa que le esses arquivos e gera um arquivo texto com comandos SQL capaz de criar uma tabela e preencher essa tabela com o nome e o ponto geográfico usando os X e Y's existentes no arquivo no formato *Dbase IV*. Uma vez executados esses comandos SQL, duas tabelas temporárias foram geradas no *PostgreSQL* com as informações necessárias. Com uma operação SQL simples foi possível a atualização das dimensões não-agregáveis escola e unidadesaude usando os nomes como chave de comparação para ambas.

Faltou ainda migrar a dimensão geográfica agregável: a dimensão *setorcensitario*. Esta dimensão migrou em três etapas: os dados convencionais genéricos, os dados geográficos dos polígonos de cada setor e os totalizadores das dimensões geográficas não-agregáveis.

Para a migração dos dados convencionais foi usada a base de dados disponibilizada pela prefeitura em *Microsoft SQL Server*, em três tabelas diferentes. Para a gravação dos dados no modelo multidimensional foi utilizada a ferramenta de exportação do próprio *SQL Server*, via *ODBC*.

Setor Administrativo, Setor Censitário e Bairro são dimensões geográficas do tipo polígono. Os mapas geográficos de cada um foram disponibilizados no formato "ArcView Shape File" pela Prefeitura Municipal de Macaé, oriundos do *Projeto Geo-Macaé*.

Para incorporar as informações geográficas contidas nesses arquivos ao Banco de Dados

multidimensional, foi usado o aplicativo "shp2pgsql" do próprio *PostGeo*, capaz de, a partir de um arquivo *shape*, criar um arquivo com comandos SQL para construir e povoar uma tabela no *PostgreSQL* e com os dados contidos no formato *ArcView Shape*. Três arquivos com comandos SQL foram criados: um para Bairro, outro para Setor Administrativo e mais um para Setor Censitário. Esses comandos SQL, executados no *PostgreSQL*, geraram tabelas temporárias. Deste ponto, apenas foram executados comandos SQL de atualização que comparavam a tabela temporária com a tabela *setorcensitario*.

Primeiro, foram migrados os Setores Censitários, comparando a tabela temporária com a tabela *setorcensitario* através da chave *setorcensitario\_id*. Depois, foram migrados os Bairros, comparando a tabela temporária com a tabela *setorcensitario*, através da chave *nomebairro*. Por fim, foram migrados os Setores Administrativos, comparando-se a tabela temporária com a tabela *setorcensitario* através da chave *nomesetoradministrativo*. Os atributos *setorcensitariogeo*, *bairrogeo* e *setoradmgeo* foram instanciados com o polígono que materializa a sua imagem geométrica.

A última etapa da atualização dos Setores Censitários é a atualização dos valores referentes às dimensões geográficas não agregáveis. Na verdade, esses totalizadores existem somente para agilizar o processo de consulta. Para isso, foram utilizadas funções geográficas presentes no *PostgreSQL* através da extensão *PostGeo*, que associam os pontos das Escolas e Unidade de Saúde com os Setores Censitários, Bairros e Setores Administrativos.

## 5.7 A preparação do cubo OLAP

O próximo passo é a configuração do cubo OLAP na ferramenta *PostGeoOLAP*, que foi denominado "pmm". Neste cubo foram adicionadas as dimensões com a preocupação na definição dos níveis de consolidação que se pode ter ao longo de cada dimensão. Esta operação é denominada "hierarquização".

O software define como Item Numérico todo atributo da tabela fato sobre o qual serão realizadas operações, tais como: soma, média, etc. No caso do cubo pmm, a tabela fato possui um único item numérico que é denominado "quantidadepessoas". Este atributo representa a quantidade de pessoas que possuem uma determinada característica. Sobre esse atributo o cubo deve ser capaz de realizar operações de soma.

Depois de definida a tabela fato, passou-se a definir as dimensões agregáveis, juntamente com as hierarquias internas de cada dimensão. Por exemplo, na dimensão *setorcensitario* foram definidos os atributos relativos a setor administrativo como de nível hierárquico mais alto, os de bairro como de nível

hierárquico intermediário e o nível mais baixo com os atributos de setor censitário.

Após a definição da tabela fato e das dimensões, o próximo passo é a definição das dimensões geográficas não agregáveis (escola e *unidadesaude*). Em seguida, realizam-se as pré-agregações, baseado na operação definida (soma sobre o atributo *quantidadepessoas*), em todas as dimensões com suas devidas hierarquias. Isto permite que o *software* desenvolvido realize a geração de tabelas com as previsões de consultas sobre os dados para agilizar o processo de busca de informações do usuário. De posse dessas pré-agregações realizadas, a ferramenta OLAP espacial está pronta para a realização de consultas *on-line* para a validação e busca de informação para o gestor municipal.

6 TESTES E ANÁLISES DE RESULTADOS

Serão abordadas neste item algumas consultas, com o propósito de validar o SSD proposto. Para isso foram levantadas algumas questões junto à Prefeitura de Macaé para poder dar mais autenticidade e valor aos testes aqui realizados. Para responder a cada questão, primeiramente serão explicados quais atributos de cada dimensão estão envolvidos e, em seguida, apresentada a resposta dada pelo sistema.

6.1 Quantidade de pessoas que possuem mais de 45 anos e estão desempregadas, por Setor Administrativo.

Para a realização desta consulta, estão envolvidas as dimensões característica, trabalho, *setorcensitario*, além da tabela fato.

	nomesetoradm	soma_quantidadepessoas
▶	Amarelo	2893
	Azul	977
	Bege	193
	Cinza	513
	Laranja	658
	Marrom	2251
	Rosa	1790
	Verde	2162
	Vermelho	3192
*		

Fig. 5 - Pessoas com mais de 45 anos, desempregadas por Setor Administrativo.

Sobre a dimensão *caracteristica* deve-se informar que o atributo idade deve ser maior que 45 anos. Sobre a dimensão trabalho, deve-se informar que trabalha possui o valor lógico falso. Na dimensão *setorcensitario* não será realizada qualquer seleção, exceto a escolha do atributo *nomesetoradm* para a exibição, já que assim será realizado o agrupamento dos dados. Finalmente, na tabela fato, é selecionado o atributo *quantidadepessoas* para a realização da

operação de totalização. O resultado pode ser visto na figura 5.

6.2 Quantidade de brancos e outras raças, que estejam cursando o 3o grau, estejam ou não trabalhando.

Esta é uma consulta que envolve a tabela fato e as dimensões *educacao*, *caracteristica* e *trabalho*.

A afirmativa “que estão trabalhando e os que não estão trabalhando” deve ser vista como uma forma de agrupamento do resultado pelo atributo *trabalho* da dimensão *trabalho*. Além desse agrupamento, entende-se que as raças também são formas de agrupamento.

Desta forma, devem ser exibidos no resultado os atributos *quantidadepessoa* da tabela fato, o atributo trabalha da dimensão trabalho e o atributo *cor\_raca* da dimensão característica, conforme a figura a seguir.

	cor_raca	trabalha	soma_quantidadepessoas
▶	Amarela	1	2
	Amarela	0	1
	Branca	0	939
	Branca	1	1052
	Parda	0	191
	Parda	1	377
	Preta	1	76
	Preta	0	19

Fig. 6 - Pessoas cursando o 3o grau por raça e por estar ou não trabalhando

6.3 Quantidade de pessoas que recebem educação especial e moram a mais de 1000 metros de alguma Unidade de Ensino Público.

Esta consulta exige alguns cuidados envolvendo a tabela fato e as dimensões escola, *setorcensitario* e *educacao*, uma vez que somente o seu enunciado não traz todos os filtros realmente existentes. Usa-se o atributo *cursoquefrequenta* para informar que devem ser envolvidas nesta consulta somente as pessoas que recebem educação especial. Na busca das pessoas que moram a 1000 metros de alguma unidade de ensino público há duas dimensões envolvidas: a dimensão agregável *setorcensitario* e a dimensão não-agregável *escola*. Sobre o atributo geográfico *setorcensitariogeo* realiza-se a operação de comparação de distância sobre o atributo *escolageo* da dimensão escola, desde que as escolas possuam o atributo *educacaoespecial* verdadeiro e o atributo *tipoadministracao* seja diferente de particular. O resultado desta consulta pode ser visto na figura 7.



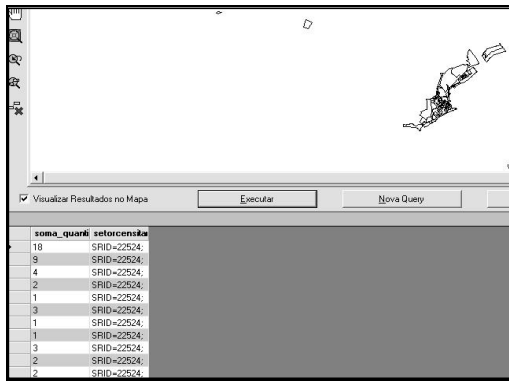


Fig. 7 - Setores Censitários que possuem pessoas com educação especial morando a mais de 1Km de alguma Unidade de Ensino Público

## 7 CONCLUSÃO

É importante salientar que o sistema é de elevado potencial de aplicações públicas no que se refere ao uso pelo administrador municipal. Para colocá-lo em prática, basta dar aos administradores municipais um treinamento que o deixe envolvido com as nomenclaturas usadas nos atributos e também para o aprendizado no uso do *PostGeoOLAP*.

Uma dificuldade no emprego desta ferramenta está na migração dos dados, uma vez que qualquer alteração realizada nos dados do *Projeto Macaé Cidadão* repercutirá em mudanças na ferramenta de migração dos dados. Mas esse é um trabalho transparente ao usuário final. O importante, neste tipo de ferramenta, é a facilidade para este usuário do sistema que não necessita ter conhecimentos profundos de informática para ter as suas informações em mãos.

As consultas realizadas destinam-se somente à validação do sistema desenvolvido. Espera-se que, com o uso diário, os usuários apresentem mais necessidades que um sistema com tamanha flexibilidade é capaz de lhes oferecer.

Espera-se que a pesquisa possa apresentar diversas contribuições, dentre as quais pode-se destacar: os benefícios para o município de Macaé, em particular, que terá nas mãos de seus administradores uma ferramenta capaz de responder muitas das perguntas das quais eles necessitavam de respostas de forma rápida, prática e eficiente; mostrar que um Sistema de Suporte a Decisão tem grande importância no âmbito municipal, principalmente na organização de informações “espalhadas”, para que decisões a favor do município e de seus habitantes sejam tomadas de forma mais inteligente, precisa e consciente; a utilização de ferramentas de distribuição gratuita no desenvolvimento de um Sistema de Suporte à Decisão, tentando colocar essa tecnologia cada vez mais disponível para as mais diversas empresas, sejam de pequeno ou de grande porte, pública ou privada; a integração de ferramentas de suporte à decisão, dando

um valor muito maior à informação. SIG, OLAP e *Datawarehouse* juntos podem trazer informações mais precisas e completas ao tomador de decisão; o georreferenciamento das escolas e unidades de saúde do município que até o momento da realização deste trabalho ainda não estava disponível pelo *Projeto Geo-Macaé*; o uso de uma ferramenta de interface amigável para o usuário que não tem necessidade de conhecimento tecnológico para a busca de informação que até então seria de difícil acesso e baseada em conhecimento profundo de informática.

## 8 REFERÊNCIA BIBLIOGRÁFICA

- Aronoff, S. Geographic Information Systems: A Management Perspective. WDL Publications, Ottawa, Canadá, 1995.
- Buehler, K.; Mckee, L. The OpenGIS Reference. Disponível em: <http://www.opengis.org/info/orm/03-040.doc>. Último acesso: Setembro de 2003.
- Colonese, G, 2004. Uma Ferramenta Aberta de Desenvolvimento Integrado de Sistemas de Informação para Processamento Analítico e Geográfico. Tese de Mestrado, Universidade Cândido Mendes, Campos dos Goytacazes, Brasil, 2004.
- Dias, T.L.; Oliveira, M.P.G., Câmara, G., Sá, M. de. Problemas de Escala e a Relação Área-Indivíduo em Análise Espacial de Dados Censitários. Disponível em: <http://www.ip.pbh.gov.br/sum0401.html>. Último acesso: Maio, 2005.
- Inmon, W.H. Building the Data Warehouse, Wiley, 1992.
- Navathe, S.B. Elmasri, R. Fundamental of Database Systems, Addison-Wesley, 2001.
- R. Kimball. The Datawarehousing Toolkit. John Wiley & Sons, Nova York, Estados Unidos, 1996.
- OpenGIS Simple Specifications for SQL Revision 1.1. OpenGis Project Document 99-049. Publicado em 05/05/1999.
- PlanetGIS guide, versão 2.29. Disponível em <http://www.planetgis.co.za/>. Último acesso em: 20/10/2003.
- PostGIS online documentation, 0.8 version. Disponível em: <http://postgis.refrations.net/>. Último acesso: 05/02/2004.
- PostgreSQL 7.4 documentation. Disponível em <http://www.postgresql.org>. Último acesso: 05/02/2004.