

Análise de Agrupamentos para a Taxa de Resíduos Sólidos de Fortaleza via Sistema Visual TAX

RESUMO

Este trabalho mostra as fases de desenvolvimento de um projeto para cobrança da taxa de resíduos sólidos de Fortaleza (TRS99), que foi desenvolvido e aplicado a uma população de mais de 2 milhões de habitantes ou 500 mil imóveis, sobre um ambiente de alta hostilidade e forte questionamento da mídia a qualquer tipo de cobrança tributária pela prefeitura de Fortaleza aos municípios, após um aumento do Imposto Predial e Territorial Urbano (IPTU). Mostramos aqui a técnica de agrupamento utilizada para caracterizar imóveis-municípios, unidade adotada para taxação e a forma como abordamos o problema, os resultados de classificação, alguns aspectos do sistema Visual TAX, e o resultado da experiência obtida com este projeto. Ênfase seja dada ao alto envolvimento da Prefeitura de Fortaleza e suas principais secretarias, pelo montante pretendido de R\$ 39 milhões a serem arrecadados.

Por:

Marcos José Negreiros Gomes
UNIVERSIDADE ESTADUAL DO CEARÁ
Departamento de Estatística e Computação
e-mail: negreiro@uece.br

Paul Gerhard de Almeida
EMPRESA MUNICIPAL DE LIMPEZA URBANA - EMLURB
e-mail: paul@truenet_ce.com.br

Anchieta Guarany Fernandes Bezerra
SECRETARIA DE FINANÇAS (SEFIN)
e-mail: abezerra@secrel.com.br

Adilson Elias Xavier
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
COPPE/ENG. DE SISTEMAS E COMPUTAÇÃO
Centro de Tecnologia
e-mail: adilson@cos.ufrj.br

PALAVRAS CHAVES:

Taxação/Tarifação Pública, Coleta de Lixo, Análise Multivariada.

1. Introdução

A taxa de resíduos sólidos é um ônus aplicado ao município, que visa cobrir o custeio com o sistema de coleta, destino final e administração do lixo domiciliar. É um recurso necessário e importante ao município, uma vez que divide a responsabilidade deste serviço entre as prefeituras e o cidadão. Apesar de sua importância e necessidade, é uma contribuição bastante questionada pela população de algumas cidades brasileiras nos últimos anos. Mesmo com o questionamento algumas das principais cidades brasileiras, como Recife (Taxa de Limpeza Pública) e Curitiba, já adotam formas de se taxar a limpeza pública.

Em maio de 1998, após a suspensão da Taxa de Coleta de Lixo (TCL) provocada pela OAB-Ceará, devido à inconstitucionalidade da fórmula de cobrança, fomos colocado um desafio para torná-la legítima e factível de ser cobrada ao município de Fortaleza. Uma equipe composta de consultores em sistemas, consultores de tributação (SEFIN), advogados, juristas e o pessoal da administração do lixo da cidade (EMLURB) desenvolveram uma nova e audaciosa lei que utilizou uma metodologia científica para atingir os objetivos preliminares de legalidade e constitucionalidade, num ambiente totalmente hostil à sua aprovação.

Na "seção 2" apresentamos os critérios de modelagem adotados para o problema, onde colocamos brevemente os aspectos técnicos da lei que ajudamos a elaborar, na "seção 3" mostramos um exemplo de geração da taxa, por meio da utilização da metodologia proposta, na "seção 4" apresentamos as experiências obtidas com a formulação e aplicação da lei.

2. Modelagem para Categorização dos Grupos Taxáveis

Apesar da taxa ter que ser cobrada a todo município, pelo princípio da divisibilidade (taxa igual para todos), legalmente a cobrança somente pode ser feita no imóvel e a seu proprietário. Alguns pré-requisitos devem ser analisados.

2.1. Pré-requisitos

- **Divisibilidade da Taxa** - Taxa única para todo município de Fortaleza, porém municípios podem ser subsidiados pela prefeitura;
- **Local de cobrança o imóvel**, sendo o proprietário responsável pelo pagamento da taxa referente ao número aproximado de pessoas que o ocupam em média durante o ano, de acordo com os indicadores censitários.

• **A quantidade de imóveis pertencentes aos grupos categorizados** compõe o denominador da taxa assim como a média de habitantes por imóvel de cada grupo. A soma do produto destes valores representa uma estimativa factível da população de Fortaleza;

• **Os custos da coleta do lixo estimados para o ano** deverão ser completamente satisfeitos e justificados por método de previsão adequado, utilizando os números da **EMLURB**. A estimativa poderá ser otimista, desde que preserve o município;

• **O valor da taxa** deve ser factível para cada grupo, de modo a não comprometer a arrecadação total.

2.2. A fórmula na lei da TRS

§1. A taxa do lixo será cobrada de grupos distintos de imóveis-municípios, os quais serão categorizados de acordo com as seguintes unidades de informação: consumo censitário, benefícios da prefeitura, produção do lixo local, categoria do imóvel e dados censitários de ocupação regional.

O valor da taxa será cobrado conforme a seguinte fórmula:

$$\text{Taxa} = \frac{EC(\text{ano})}{\sum_{i=1}^n G_i}$$

Onde:

$EC(\text{Ano})$ é a estimativa oficial do Custo Total da Coleta, Destino Final e Administração de Resíduos Sólidos no ano vigente.

n é o número de grupos a serem categorizados (valor definido como sendo igual a 48).

$$\sum_{i=1}^n G_i = \text{Estimativa oficial da população de Fortaleza}$$

e/ou a soma de municípios por grupo categorizado G_i ;

Sendo:

$$G_i = \bar{p}_i * q_i$$

\bar{p}_i é o número médio de pessoas produtoras de lixo do grupo G_i ;

q_i é o número de imóveis categorizados no grupo G_i

2.3. Metodologia empregada

No que compete a este trabalho, desenvolveu-se uma metodologia que analisa as diversidades de características existentes entre imóveis e calcula automaticamente as ta-

xas que se ajustam aos diversos perfis existentes. Ou seja, o objetivo da Análise de Grupamentos é reunir elementos distintos em grupos semelhantes. O conceito de semelhança deve, portanto, receber um tratamento matemático.

Uma forma de medir a semelhança entre indivíduos, é fazer uma medida da dispersão interna dos elementos dentro de um grupo. As principais medidas de dispersão utilizadas são:

- soma dos quadrados dentro do grupo;
- variância interna de grupo;
- diâmetro de grupo e
- dispersão via mediana de grupo.

Trabalhamos com a primeira forma, ou seja, a soma dos quadrados dentro do grupo a qual passamos a descrever, [Hartigan, 1971].

Considerando um conjunto E de indivíduos que se deseja particionar em m grupos, $\{g_1, g_2, \dots, g_m\}$, de forma a reunir os indivíduos mais semelhantes dentro de cada grupo. Definimos, portanto, uma função objetivo que traduz matematicamente tal semelhança e, para isso, utilizamos o conceito de 'Soma dos quadrados dentro do grupo', como segue. Seja:

$$\bar{X} = \frac{1}{n_i} \sum_{i=1}^{n_i} x_i \quad [2.1]$$

$$W_i = \sum_{i=1}^{n_i} \left(X_i - \bar{X}_i \right)^T \left(X_i - \bar{X}_i \right) \quad [2.2]$$

Onde, nas equações [2.1] e [2.2], X é dito ser a média ou centróide de grupo e n_i é o número de elementos do grupo g_i .

Assim, W_i representa a soma dos quadrados das distâncias euclidianas entre cada elemento do grupo e o centróide (ponto médio) do mesmo. Poderíamos, portanto, tomar para função objetivo $f(P_m) = (W_i)$, resolvendo um problema de minimização. Teríamos, por conseguinte, a situação:

Minimizar $f(P_m)$

s.a:

P_m viável

Por trás desta formulação genérica, vários modos de formulações matemáticas existem. Os métodos exatos do tipo "branch-and-bound", que teoricamente resolvem o problema, somente permitem soluções para instâncias de pequeno porte, a menos que os grupos estejam definidos quase que diretamente. Métodos exatos

resolvem hoje problemas clássicos com até 150 objetos. Apesar disto, muitos casos reais necessitam de resolvidores para tratar instâncias de grande porte, o que justifica o uso de heurísticas para o problema, [Hansen & Mladenovic, 1999].

O problema de "categorização" ou de agrupamento pode ser visto na construção de um dado número de grupos que reúnem o maior número de características em comum, como já dissemos. Isso é muito importante, pois a lei de formação de grupos, a partir de determinados critérios, evita discrepâncias na formação de grupos com indivíduos com relativo grau de semelhança. Dentre os métodos explorados existentes, avaliamos para este projeto dois dos mais conhecidos da literatura, *Forgy* e *Kmeans*, [Forgy, 1965] e [Mac Queen, 1967].

2.3.1. O método de Forgy

Trata-se de uma heurística interativa para obtenção de um ótimo local, que busca identificar rapidamente aqueles k -grupos (onde k é inicialmente informado) que melhor se ajustam aos dados pela proximidade aos centros previamente encontrados, [Forgy, 1954].

Passo 0: Faça $k=0$

Passo 1: Se $k=0$, tome uma partição P_m^0 de E ;

Senão, forme uma nova partição P_m^k de E , alocando cada elemento ao grupo cujo

Ponto semente esteja mais próximo e faça $k=k+1$.

Passo 2: Calcule as centróides dos grupos formados, que serão os pontos sementes para os novos grupos.

Passo 3: Se $f(P_m^k) = f(P_m^{k-1})$, pare;

Senão, vá ao passo 1.

A solução inicial é fornecida para o algoritmo juntamente com os centróides dos m grupos, sendo estes, os pontos sementes para a geração de novos grupos. A função objetivo utilizada, $f(P_m^k)$, é a soma dos quadrados dentro dos grupos, definida anteriormente, objetivando minimizar as distâncias dos elementos aos centróides de seus grupos. As seqüências de passos 1 e 2 geram novas partições tais que: $f(P_m^k) \leq f(P_m^{k-1})$

A convergência do método de Forgy é garantida. Uma particularidade interessante deste método é a rapidez com que o algoritmo converge para a solução ótima local.

2.3.2. O método de KMeans

Esse algoritmo, bem mais elaborado e custoso, do ponto de vista computacional, busca resolver o problema das realocações entre "clusters" (grupos), de uma maneira gulosa, porém mais global do que a estratégia proposta por Forgy, [Jansey, 1966] e [Mac Queen, 1967].

Passo 1: Assuma uma solução inicial de alocação dos objetos a seus respectivos "k-clusters".

Passo 2: Assuma que um objeto do "cluster" i , é reatribuído a um outro grupo (j), onde os centróides destes novos grupos podem ser facilmente obtidos pela seguinte fórmula de atualização, [Späth, 1980]:

$$\bar{x}_i \leftarrow \frac{n_i x_i - x_j}{n_i - 1}; \quad \bar{x}_j \leftarrow \frac{n_i \bar{x}_i - x_j}{n_i + 1};$$

Onde, $n_i = |C_i|$ e $n_j = |C_j|$

Passo 3: A mudança na função objetivo causada devido a esse movimento pode ser recalculada pela seguinte fórmula, para todas as reatribuições possíveis:

$$v_{ji} \leftarrow \frac{n_i}{n_i + 1} \|\bar{x}_i - x_j\|^2 - \frac{n_i}{n_i - 1} \|\bar{x}_i - x_j\|^2, x_j \in C_i;$$

Passo 4: Se não ocorrer nenhum movimento de um objeto de um "cluster" para outro, então termine, (um ótimo local foi atingido). Caso contrário a reatribuição que provocou a maior redução é feita, retorne ao passo 2.

2.5. O método escolhido

Considerando o tipo de instância que trataríamos em nosso sistema, com ordem superior a 500mil imóveis e objetos com até 12 variáveis, preferimos utilizar o método de Forgy. É importante ressaltar que a escolha do método resultou da necessidade de obtenção de respostas em tempo hábil para posterior modificação da solução obtida, via facilidades do sistema **Visual TAX**. É claro que refinamentos interativos poderiam ser feitos no método usado, porém não utilizamos esta técnica para isto, pela brevidade do tempo que dispunhamos (três meses, entre elaboração, testes e impressão da taxa). Avaliações de outros métodos interessantes podem ser encontradas em [Hansen & Mladenovic, 1999], [Hartigan, 1971], [Mac Queen, 1967].

3. Exemplo de formação da taxa

Para compreender e visualizar melhor a categorização dos grupos adotada, a seguir exemplificamos uma situação de categorização utilizada para o cálculo da TRS de Fortaleza.

Supondo que existam as seguintes informações de proprietários de imóveis, disponíveis na secretaria de finanças (SEFIN), e a ligamos a informações de produção do lixo da prefeitura (EMLURB), dados censitários do IBGE, consumo de água e luz. Além disso, considerando que um munícipe pagaria R\$ 22,00 por ano (= Despesa Total/População). Seja:

