

# Definição de Metodologia para Segmentação de Clientes, Identificação de Grupos e Conhecimento do Mercado “Clusters BT”

Eduardo de Rezende Francisco, *AES Eletropaulo*

Claudia Sciortino de Reina, *Fundação CPqD*

**Resumo** - Este trabalho tem por objetivo apresentar os resultados do desenvolvimento de uma metodologia específica de segmentação, que visa formar e caracterizar segmentos de clientes de baixa tensão da AES Eletropaulo, maximizando o aproveitamento de dados técnicos e comerciais advindos dos diversos processos de gestão da empresa, de forma integrada às informações externas (Censo Demográfico 2000, Pesquisa ABRADEE 2005 do Cliente Residencial Urbano, e localização geográfica dos clientes da AES Eletropaulo) para geração de novos conhecimentos e estimativas do Mercado, definição de soluções, produtos e serviços adequados a cada segmento e auxílio em diversos programas de diminuição de perdas financeiras e melhoria de receita daquela concessionária. De forma complementar, a identificação dos hábitos e padrões de consumo de cada segmento de clientes permite um contínuo aperfeiçoamento do conhecimento do perfil de consumo de clientes-chave, possibilitando a adoção de ações comerciais específicas e/ou oferta de serviços diferenciados.

**Palavras-chave:** *Data Mining, Estatística, Segmentação de Clientes, Pesquisa de Satisfação, Sistema de Informação Geográfica e Suporte à tomada de decisão.*

## I. INTRODUÇÃO

As grandes concessionárias de distribuição de energia elétrica, em especial a AES Eletropaulo, possuem um cadastro técnico e comercial de seus clientes totalmente georreferenciado, de acordo a [1]. A localização do medidor de energia elétrica que alimenta esses clientes é armazenada nos sistemas técnicos das distribuidoras, suportados, em geral, por tecnologias de Sistemas de Informação Geográfica. Todos os clientes da concessionária estão, pois, localizados no espaço geográfico, associados a um dispositivo ou ativo elétri-

co (poste, transformador) e a um circuito alimentador (ver [19]).

Nacionalmente, o serviço de fornecimento de energia elétrica abrange 97,0% dos domicílios brasileiros, índice que aumenta para 99,6% na área urbana, para 99,4% na região Sudeste, ver [13], e para 99,9% no município de São Paulo, conforme [13] e [19]. Adicionalmente, as bases de dados das distribuidoras de energia elétrica contêm informações de consumo de cada um de seus clientes, conforme [1].

Dessa forma, explorar o conhecimento sobre os perfis comportamentais dos clientes da concessionária garante a cobertura de praticamente todos os domicílios residenciais da área de concessão e implica em melhorar a política de relacionamento com o cliente, criando canais personalizados de atendimento, reduzir custos operacionais e diminuir perdas financeiras.

A busca por padrões e tendências (*Data Mining*), ver [2],[5] e [6], em conjunto com informações externas permitem uma melhor caracterização sócio-econômico-demográfica dos clientes de baixa tensão da AES Eletropaulo, constituindo um repositório rico para a criação de modelos preditivos aplicáveis às diversas áreas de negócio, em particular, para Fraude, Inadimplência, Serviços Não Regulados, *Call Center* e Ouvidoria. Adicionalmente, auxilia na contínua correção de inconsistências na base cadastral da empresa.

Os direcionadores de marketing de uma organização devem perseguir uma estratégia realista e devem estar, primordialmente, alinhados com os objetivos financeiros e de construção e valorização da imagem e marca da organização. A busca desse alinhamento deve ser suportada por um conjunto de objetivos de natureza quantitativa ao longo do tempo, a saber: *redução das perdas comerciais, identificação e administração da inadimplência, aumento da rentabilidade oriunda da base atual de clientes e redução de custos*. Assim, a AES Eletropaulo poderá tomar decisões baseada na integração das diversas fontes internas e externas de informações, possibilitando a geração de novos e eficientes elementos para o suporte à gestão de marketing, incrementando, otimizando e conduzindo ao reposicionamento estratégico dos processos de negócios junto ao mercado nos diversos cenários que se apresentem.

Os elementos aqui explicitados formam um cenário que se apresenta como expectativas que precisam se concretizar para atender aos crescentes desafios, sejam eles originados

---

Este trabalho foi desenvolvido em um projeto de Pesquisa e Desenvolvimento, suportado pela ANEEL, a Agência Nacional de Energia Elétrica, e conduzido em parceria com a Fundação CPqD. Seu ciclo foi de 12 meses, finalizado em Março de 2006.

Este projeto recebeu o prêmio de “Melhor Trabalho de Negócios” do CATI 2006 – 3º. Congresso Anual de Tecnologia da Informação da Fundação Getúlio Vargas em 28 de Julho de 2006, em São Paulo.

Eduardo de Rezende Francisco é Analista de Negócios Sênior da AES Eletropaulo, especialista em Estratégias de Marketing, Pesquisas de Satisfação de Clientes e Geomarketing (e-mail: [eduardo.francisco@aes.com](mailto:eduardo.francisco@aes.com)).

Claudia Sciortino de Reina é Estatística e Analista de Data Mining da Gerência de Business Intelligence da Fundação CPqD (e-mail: [csreina@cpqd.com.br](mailto:csreina@cpqd.com.br)).

do arcabouço regulatório, ou das exigências de gerenciamento, dos negócios de distribuição e comercialização de energia, o que outras indústrias eventualmente já experimentaram mais cedo.

Dado o contexto acima, a Figura 1 apresenta a visão geral do projeto mostrando a importância de dispor de uma base de clientes atualizada e bem estruturada para subsidiar as futuras tomadas de decisões estratégicas a partir da adequada análise de dados relevantes conforme os propósitos do projeto definidos anteriormente.

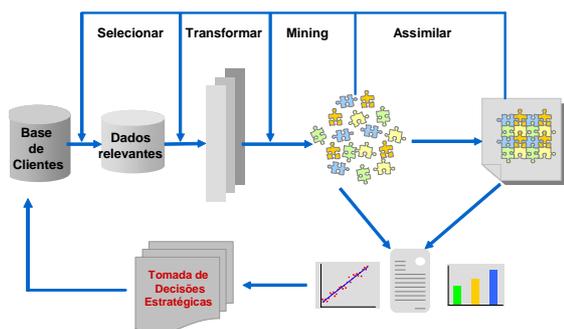


Figura 1. Visão Geral do Projeto  
Fonte: Reina et al. (2005)

O presente projeto visa atingir aos seguintes objetivos:

a) Apoio em Projetos de Melhoria da Imagem e Valorização da Marca da AES Eletropaulo, tais como:

- Melhor direcionamento das estratégias de Marketing;
- Orientar a formação de novas políticas de relacionamento e fidelização de clientes;
- Reduzir os índices de reclamações, de ações de indenização e de penalizações impostas segundo critérios da ANEEL;
- Administrar efetivamente as perdas comerciais e inadimplências.

b) Apoiar os diversos processos de gestão da AES Eletropaulo, através da formação de segmentos/clusters de clientes de baixa tensão.

## II. DESENVOLVIMENTO

Esta seção descreve a metodologia utilizada para o desenvolvimento do trabalho, a construção do modelo de segmentação e a utilização de fontes externas para ampliação das interpretações dos resultados da modelagem de segmentação de clientes de baixa tensão.

### A. Metodologia

A metodologia descrita em [5], sobre o processo de descoberta do conhecimento (do inglês *KDD* – *Knowledge Discovery in Database*), apresentada na Figura 2, visa alcançar conhecimento, a partir de grandes bases de dados focando os objetivos de negócio pré-definidos. Assim, esta metodo-

logia pode ser adequadamente adaptada a qualquer concessionária do setor elétrico pois inova desde a forma de identificação dos sistemas de dados (base de dados) até as aplicações analíticas, em particular as técnicas de *Data Mining*, aplicadas para descoberta de padrões, associações e perfis dos clientes.

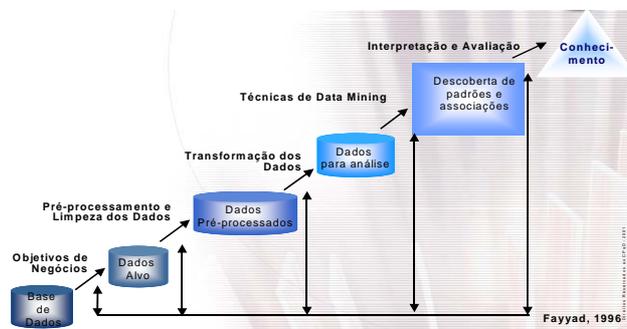


Figura 2. Adaptação da Metodologia de KDD  
Fonte: Adaptado da Metodologia de KDD, Fayyad (1996)

Analisando a Figura 3, pode-se notar que o modelo de dados que subsidiou o desenvolvimento do presente trabalho foi formado por várias fontes da AES Eletropaulo que armazenam informações sobre os clientes de baixa tensão, tais como o Protocolo, Primeiro Contato, Cobrança, entre outras. Considerando [14], os objetivos de negócio visam, em linhas gerais, entender, conhecer e caracterizar melhor a unidade de referência (cliente) de baixa tensão da AES Eletropaulo para diminuir a distância hoje existente entre a concessionária e seus clientes, bem como identificar as necessidades e os problemas que estes últimos têm perante o consumo e importância da energia elétrica.

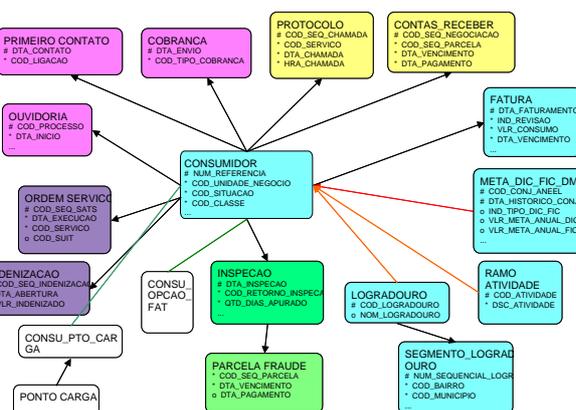


Figura 3. Desenvolvimento do Modelo de Dados  
Fonte: Neto (2005)

Diante de tais objetivos e dos dados existentes nos sistemas da AES Eletropaulo, iniciou-se o processo de Mapeamento Analítico, que consiste em identificar as informações (variáveis) necessárias para alcançar, através de técnicas de *Data Mining*, os objetivos anteriormente definidos. Ao final deste processo, obtém-se uma base de dados, normalmente, chamada de *dados alvo*. Esta base de dados passou a ser uma referência para alcançar os objetivos de negócio, onde os dados receberam tratamentos estatísticos e transforma-

ções matemáticas, sempre que necessários, conforme descrito em [18]. Como resultado final desta etapa, obteve-se uma nova base, contendo os dados pré-processados.

Até este momento, nenhuma das etapas descritas acima pode ser esquecida pois elas são indispensáveis para garantir o êxito das próximas etapas do processo de *KDD*, que visa descobrir os perfis comportamentais dos clientes de baixa tensão da AES Eletropaulo.

Segundo [2], outro fato a destacar, é que os dados a serem modelados também devem estar sumarizados no mesmo nível de granularidade. Feito isto, obtivemos os dados finais para serem analisados (dados para análise).

Para iniciar as etapas posteriores, foi utilizada a técnica estatística de amostragem, descrita em [15], com o intuito de obter uma amostra representativa da população de 4.659.441 unidades de referência (clientes) de baixa tensão, considerando o período de observação histórica de Abril a Setembro de 2005. A partir da amostra aleatória estratificada, obtida em 2 etapas e formada por 512.527 unidades de referência, foi realizada uma análise exploratória dos dados, conforme descrição detalhada em [16], obtendo uma base de dados consistente para início da construção do modelo de segmentação dessas unidades de referência (clientes).

Diante dos objetivos, desenvolveu-se uma modelagem de segmentação que visa descrever os perfis comportamentais dos clientes de baixa tensão da AES Eletropaulo, a partir de informações de consumo e de relacionamento, conforme [16].

O desenvolvimento da modelagem de segmentação se deu através da técnica *Two Step Cluster*, ver [10]. Esta técnica apresenta um algoritmo inovador quanto ao fato de formar *clusters* em grandes bases de dados compostas por variáveis contínuas e categóricas. Isto dá um forte caráter inovador ao presente trabalho por ser uma técnica de *clusters* que considera simultaneamente dados categóricos e contínuos na construção e definição do modelo. As demais técnicas, como *K-means*, não permitem analisar em conjunto variáveis categóricas e contínuas. E, considerando que para este trabalho o volume destas variáveis era grande, a técnica *Two Step Cluster* inova quanto à forma estatística de analisar os dados.

A análise *Two Step Cluster* também oferece a flexibilidade de especificar o número de agrupamentos ou permitir que o algoritmo descubra automaticamente o número mais adequado de *clusters* para a base de clientes em questão. As similaridades entre os registros com relação a cada uma das variáveis definem os *clusters*, onde estes serão caracterizados por aquelas variáveis mais significantes, conforme [10].

Por fim, os *clusters* criados geraram conhecimento que permitiram ser interpretados e avaliados de forma inovadora haja visto que, de uma só vez, informações contínuas (tais como: consumo nos últimos 6 meses, consumo a recuperar por fraude, etc.) e informações categóricas (tais como: ter ou não débito automático, ter ou não serviço não regulado, etc.) permitiram identificar com clareza os segmentos com que a concessionária deveria passar a interagir e direcionar a evolução de seus serviços, produtos e canais de comunicação com o cliente, criando uma nova cultura dentro da AES Ele-

tropaulo quanto à forma de entender e caracterizar, bem como de se relacionar com o cliente. Uma abordagem de segmentação coerente e adaptada pode servir de base para as demais ações mercadológicas, bem como direcionar o atendimento aos diversos segmentos/*clusters* das unidades de referência (clientes) que compartilham necessidades e expectativas semelhantes.

Em [2] se descreve a importância de segmentar estatisticamente a base de clientes, considerando o histórico do relacionamento existente entre empresa-cliente, bem como todas as informações relevantes desse relacionamento. Portanto, o presente trabalho também inova quanto à forma de entender o comportamento frente ao consumo e a importância da energia elétrica para as unidades de referência (clientes) de baixa tensão da AES Eletropaulo. Conseqüentemente, tem-se um alicerce sólido para novas investigações sobre estas unidades de referência (clientes) cada vez mais preocupadas em serem mais bem entendidas e atendidas por esta concessionária.

## B. Desenvolvimento do Modelo de Segmentação

Para a construção do modelo foi considerada a seguinte etapa:

- **Determinação do Número de *Clusters*:** Esta etapa é responsável pela determinação do “melhor” número de *clusters* a aplicar aos dados internos dos clientes de baixa tensão da AES Eletropaulo.

Para os dados utilizados (ver descrição no parágrafo 6 do item anterior (*Metodologia*)), a técnica calcula a formação de um modelo de segmentação com até 15 *clusters* a partir do Critério de Informação de Bayes (BIC). Este valor foi escolhido por se entender que mais de 15 *clusters* gera uma pulverização muito grande da base de clientes, dificultando a identificação, caracterização e diferenciação dos mesmos. Em [2] se sugere usar de bom senso na escolha do número máximo de *clusters*, e considerando também os resultados das medidas apresentadas na Figura 4, onde não se vê variações significativas entre as quantidades 9 e 15 de *clusters*. Sabe-se que uma quantidade enorme de *clusters* dificulta suas caracterizações e um número muito pequeno também não permite uma interpretação adequada dos mesmos.

A Figura 4 mostra graficamente os valores da Medida "Razão entre Distâncias" (ou em inglês, *Ratio of Distance Measures*) obtida pela técnica *Two Step Cluster* que possibilita comparar as maiores medidas, identificando os seus respectivos *clusters*. Em [10] sugere usar as duas maiores medidas e escolher a maior, mas aqui optou-se por analisar as 5 maiores medidas, fixando assim 5 números de *clusters* a comparar (2, 4, 6 e 14 *Clusters*) pois vemos que perante uma amostra (512.527 unidades de referência) e população (4.659.441 unidades de referência) tão grandes, criar apenas 2 *Clusters* gerariamos grandes distorções nas interpretações e caracterizações destes *Clusters*. Assim, optou-se analisar as 5 maiores medidas já que valores maiores a este também levariam a interpretações pouco enriquecedoras.

Diante destas medidas e considerando [10], construiu-se a Tabela 1, a fim de determinar o número de *clusters*.

Auto-Clustering				
Number of Clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change <sup>a</sup>	Ratio of BIC Changes <sup>b</sup>	Ratio of Distance Measures <sup>c</sup>
1	23699574,140			
2	20691840,344	-3007733,8	1,000	1,784
3	19007264,398	-1684575,9	,560	1,300
4	17712383,531	-1294880,9	,431	1,553
5	16880019,824	-832363,706	,277	1,078
6	16108066,256	-771953,568	,257	1,308
7	15518827,647	-589238,609	,196	1,202
8	15029385,138	-489442,509	,163	1,345
9	14666345,970	-363039,168	,121	1,135
10	14346940,759	-319405,212	,106	1,303
11	14102714,209	-244226,549	,081	1,038
12	13867663,444	-235050,766	,078	1,030
13	13639473,186	-228190,258	,076	1,100
14	13432315,264	-207157,921	,069	1,327
15	13277131,218	-155184,046	,052	1,050

a. The changes are from the previous number of clusters in the table.  
b. The ratios of changes are relative to the change for the two cluster solution.  
c. The ratios of distance measures are based on the current number of clusters against the previous number of clusters.

Figura 4. Bayesian Information Criterion para escolha do Número de Clusters

Fonte: Ferramenta SPSS versão 13.0

TABELA 1

COMPARAÇÃO ENTRE AS 5 (CINCO) MAIORES RAZÕES ENTRE DISTÂNCIAS (2, 4, 6, 8 E 14)

Incremento de Clusters	Razão entre Ration of Distance Measure	Seleção do Número de Clusters
2 → 4	1,1487	4
4 → 6	1,1873	4
4 → 8	1,1546	4
4 → 14	1,17	4
6 → 8	0,972	8
6 → 14	0,9857	14
8 → 14	1,013	14

Fonte: Os 5 Clusters com as maiores Ration entre distâncias obtidos pela técnica Two Step Clusters, via a Ferramenta SPSS versão 13.0

Considerando o descrito acima e a Tabela 1, foram selecionados os seguintes três valores: 4, 8 e 14, para construção do modelo de segmentação. Perante os três modelos (4, 8 e 14) foram analisadas e comparadas cada uma das variáveis, observando suas médias e desvios padrões quando contínuas, e valores reais quando discretas, a fim de perceber que o modelo de 14 clusters proporcionava a AES Eletropaulo um maior número de informações detalhadas sobre as unidades de referência (clientes) de baixa tensão quando comparado aos modelos com 4 e 8 clusters. Desta forma, este é o modelo recomendado a AES Eletropaulo aplicar sobre sua população de baixa tensão a fim de segmentá-la segundo suas informações cadastrais, canais de relacionamentos, recebíveis e perdas, serviços não regulados e metas ANEEL. Todas estas informações (variáveis) analisadas simultaneamente permitiram a formação da Figura 5, de onde se pode visualizar a formação nominal de cada um dos 14 clusters conforme os graus de importância dessas variáveis.

Diante do modelo recomendado, a próxima seção amplia as interpretações sobre cada um dos 14 clusters do modelo de segmentação indicado.

### C. Análise de Sobreposição

Visando ampliar as interpretações sobre cada um dos 14

clusters do modelo de segmentação recomendado, optou-se por sobrepor aos resultados dessa modelagem algumas análises oriundas de informações externas à base de dados construída na metodologia descrita anteriormente (e apresentada na Fig. 5).

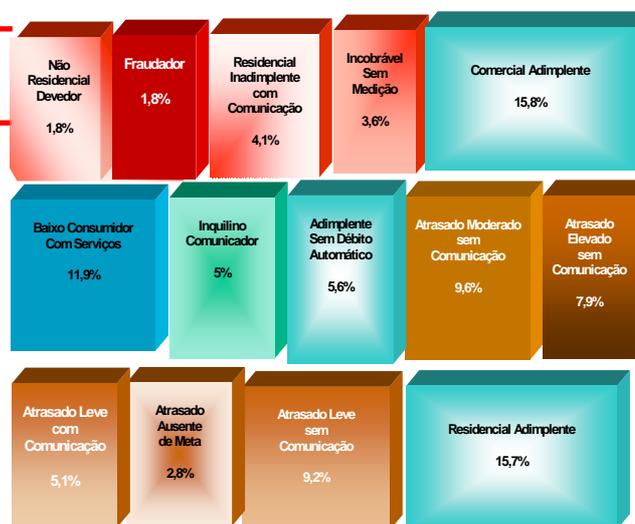


Figura 5. Modelo de 14 Clusters

Fonte: Reina & Petrielli (2006)

Basicamente, as informações utilizadas para esse enriquecimento foram: (i) informações do Score de Risco calculado pela AES Eletropaulo em Dezembro/2005, (ii) variáveis do Censo Demográfico 2000, realizado pelo IBGE, e (iii) Indicadores da Pesquisa ABRADÉE 2005 de Satisfação do Cliente Residencial Urbano da AES Eletropaulo, aplicada em Março de 2005. Estas informações foram obtidas em instantes de tempo diferentes ao observado na construção do citado modelo de segmentação, conforme [16], e, por esta razão, foram utilizadas para o enriquecimento posterior dos segmentos e não como origem de informação para a construção do modelo.

### (i) Score de Risco

O Score de Risco foi o resultado de um projeto desenvolvido pela AES Eletropaulo, que visou criar um indicador (ordinal) de risco de recebimento na Gestão de Arrecadação dos clientes da concessionária. O indicador dá origem à discriminação em cinco classes do risco de recebimento (A, B, C, D e E) com risco crescente. Tal indicador já está incorporado ao sistema de faturamento dos clientes de baixa tensão, sob gestão da Gerência de Recebíveis, e sua utilização, a partir de uma visão histórica do perfil de arrecadação e faturamento, permite que estratégias diferenciadas de cobrança e relacionamento sejam estabelecidas, conforme a classe ou pontuação do score, para cada unidade de referência (cliente).

Nesse contexto, cada cliente de baixa tensão da AES Eletropaulo já possui um score de risco de recebimento associado, o que permitiu sua incorporação na análise de sobreposição das unidades de referência da amostra que deram origem ao modelo.

## (ii) Sobreposição Geográfica e Junção Espacial

A Figura 6 a seguir apresenta a estratégia de sobreposição geográfica e junção espacial utilizada para a integração das informações de fontes externas (em particular, dados do Censo Demográfico 2000 e dados da Pesquisa ABRADDEE 2005) às unidades de referência (clientes) da amostra que originou os 14 *clusters*.

A sobreposição é possibilitada pela existência de uma localização geográfica de cada unidade de referência (cliente) da AES Eletropaulo na base de dados técnicos da empresa. A coordenada espacial que efetivamente está armazenada nos Sistemas de Informação Geográfica da AES Eletropaulo é a coordenada do ativo elétrico mais próximo do medidor de energia. Comumente, é o poste ou o transformador. Este fica tipicamente situado do lado de fora da unidade de referência, geralmente próximo ao logradouro. Este pode situar-se, em alguns casos, do outro lado da rua de endereço da unidade de referência (cliente). Essas considerações serão levadas em conta durante a análise e discussão dos resultados.

Dessa forma, pudemos associar as coordenadas geográficas do ponto de entrega das unidades de referência (ativo elétrico que alimenta a unidade consumidora) com os polígonos dos setores censitários. Já a sobreposição dos Indicadores ABRADDEE 2005 foi realizada associando as informações das unidades de referência às dez regiões pesquisadas.

Cientes AES Eletropaulo (Unidades Consumidoras)

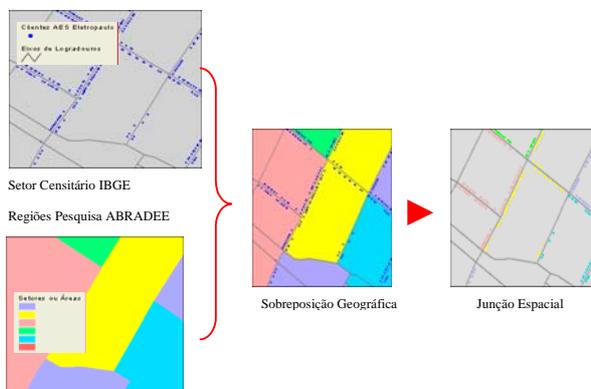


Figura 6. Estratégia de Associação das Informações do Censo Demográfico 2000 e da Pesquisa ABRADDEE 2005 às Unidades de Referência  
Fonte: Francisco (2006)

## (iii) Censo Demográfico 2000

O IBGE (Instituto Brasileiro de Geografia e Estatística) é o órgão oficial responsável pelas pesquisas censitárias, estatísticas sociodemográficas, econômicas, agropecuárias, de preços e por algumas pesquisas especiais no Brasil. Sua missão é “retratar o Brasil com informações necessárias ao conhecimento de sua realidade e ao exercício da cidadania”, conforme descrito em [11].

A coleta do Censo Demográfico 2000 foi realizada no período de 1º de Agosto a 30 de Novembro de 2000. Foram pesquisados 54.265.618 domicílios, nos 5.507 municípios existentes em 2000, de todo o território nacional, tendo mobilizado mais de 200 mil pessoas, segundo [12]. Destina-se

a dois propósitos, que se traduzem operacionalmente em duas unidades de investigação: (i) domiciliar – quantifica e levanta informações sobre todos os domicílios do país, e (ii) populacional – quantifica e levanta informações de toda a população que vive em domicílios.

Para o planejamento do censo, o território nacional foi dividido em 215.811 áreas contíguas, respeitando-se os limites da divisão político-administrativa, do quadro urbano e rural legal, de outras estruturas territoriais de interesse e de dimensão adequada para sua operação. Essa unidade territorial é denominada setor censitário, de acordo com [12], e engloba de 200 a 300 domicílios. Setor censitário é a unidade territorial de coleta e de controle cadastral percorrida por um único recenseador, ver [19]. Para os 24 municípios da área de concessão da AES Eletropaulo, 21 deles estão divididos operacionalmente em 18.547 setores censitários.

A Figura 7 ilustra a divisão da área de concessão da AES Eletropaulo em setores censitários.

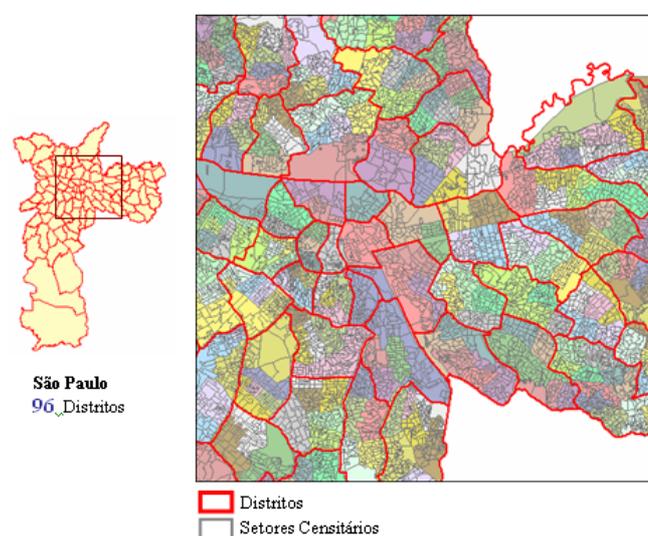


Figura 7. Área de concessão da AES Eletropaulo dividida em distritos e setores censitários  
Fonte: Francisco (2006)

## (iv) Pesquisa ABRADDEE 2005

A Pesquisa ABRADDEE é realizada no setor elétrico brasileiro desde 1999, aplicando-se simultaneamente para muitas das concessionárias brasileiras de distribuição de energia elétrica que são associadas da ABRADDEE (40 em 2004, 43 em 2005), conforme [8] e [9]. Sua realização conjunta, além de garantir a comparabilidade dos resultados devido à uniformidade metodológica e à simultaneidade de aplicação, possibilita redução de custos devido à economia de escala, ver [7]. Essa pesquisa visa avaliar a satisfação geral dos clientes, itens de qualidade do serviço fornecido pelas concessionárias e imagem da mesma. Além de permitir a classificação das empresas e a premiação das melhores colocadas nas diversas categorias.

A pesquisa pode ser dividida em duas partes: (i) uma pesquisa de importância, na qual o respondente deve ordenar e quantificar a importância de cada um dos atributos dentro de suas respectivas áreas e de cada uma das áreas, e (ii) a pes-

quisa de satisfação e fidelidade, na qual são obtidas as avaliações referentes à satisfação com os atributos de Qualidade e Preço e à fidelidade do cliente em relação à distribuidora, ver [19].

São cinco áreas de Qualidade avaliadas: *Fornecimento de Energia, Informação e Comunicação com o Cliente, Conta de Luz, Atendimento ao Cliente e Imagem*. Essas áreas totalizam um conjunto de 26 atributos. A ponderação de importância atribuída pelos respondentes e a satisfação declarada dos atributos e das áreas originam o *Índice de Satisfação com a Qualidade Percebida (ISQP)*, e é utilizado para classificação e premiação das empresas do setor, segundo categorias por região e porte. A área de Valor Percebido compreende a avaliação do Preço, conforme [7], contempla 4 atributos e origina o *Índice de Satisfação do Cliente com o Preço Percebido (ISCP)*.

A pesquisa é domiciliar, com planejamento amostral probabilístico estratificado, garantindo a representatividade da distribuidora de energia elétrica pesquisada. A população alvo é composta de clientes residenciais da zona urbana da área de concessão. O respondente é o chefe de família, ou cônjuge, de um domicílio de residência permanente do entrevistado e de uso exclusivamente residencial, cujo medidor de energia seja próprio, com fornecimento regularizado, com o valor da conta de energia diferente de zero, alfabetizado, e que saiba informar a renda mensal total da família e do chefe da família. Além disso, o respondente ou nenhum membro de sua família pode trabalhar na concessionária que está sendo pesquisada.

A AES Eletropaulo contratou a Pesquisa ABRADDEE 2005 em expansão para 10 regiões, que subdividem suas atuais 6 Diretorias Regionais. A Diretoria Regional Leste foi dividida em duas áreas (Leste–Portal e Leste–Trabalhadores) e a Regional Oeste foi dividida em quatro áreas (Oeste 1, 2, 3 e 4), conforme apresentado na Figura 8. Foram realizadas 1.500 entrevistas.

### Informações Sobrepostas

De maneira sintética, as informações utilizadas para essa análise foram, ver [16]:

- *Score de Risco*
- *Censo 2000*: Número médio de pessoas por domicílios particulares permanentes, Renda média das pessoas responsáveis pelos domicílios particulares permanentes, Grau de escolaridade das pessoas por domicílios particulares permanentes e Rendimento Nominal Mensal das pessoas responsáveis pelos domicílios particulares permanentes.
- *ABRADDEE 2005*: Índice de Aprovação do Cliente (IAC), Índice Intermediário de Satisfação (IIS), Índice de Satisfação Geral (ISG), Grau de Fidelidade (GF), Iluminação Pública (IP), Índice de Satisfação da Qualidade Percebida (ISQP), Área da Qualidade Percebida (Fornecimento de Energia, Informação e Comunicação com o Cliente, Conta de Luz, Atendimento ao Cliente, Imagem), Área de Valor Percebido (Preço) e Relacionamento.

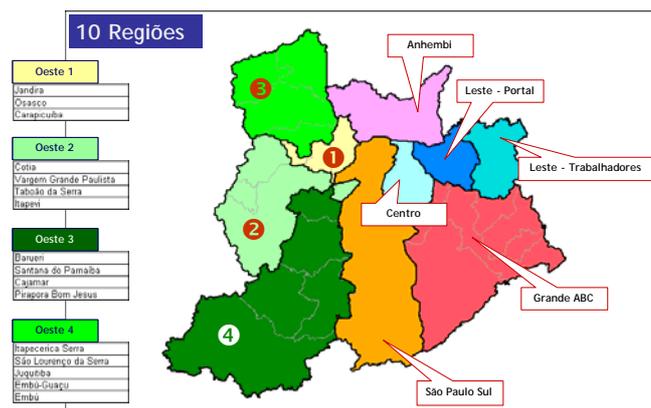


Figura 8. Área de concessão da AES Eletropaulo dividida em 10 regiões na Pesquisa ABRADDEE 2005

### D. Expansão dos Dados

Após todo o desenvolvimento do modelo de segmentação, tem-se um arquivo em *pmml* (*predictive model markup language*) que permite realizar a expansão (ou “escoragem”) para o resto da população de baixa tensão da AES Eletropaulo que não fez parte da amostra. As consequências imediatas desta expansão são:

- Realizar novas “escoragens” sobre a população analisada após certo período de tempo, permitindo avaliar o volume de unidades de referência (clientes) que migram de segmento.
- Avaliar a aderência deste modelo durante certo período de tempo, detectando o melhor instante de reajustar os parâmetros do modelo ou de rever a construção do próprio modelo de segmentação.

### III. CONCLUSÕES E RECOMENDAÇÕES SOBRE OS 14 CLUSTERS

Diante da Figura 5 e analisando todas as variáveis presentes em cada um dos Clusters, esta seção apresenta algumas das principais características que contribuíram na formação de cada um dos 14 clusters. Vale destacar que as recomendações foram estabelecidas a partir do estudo detalhado do conjunto de variáveis (contínuas e discretas) de todos os *Clusters*, que de forma única permite uma visão ampla e estruturada das características e necessidades dos clientes (unidades de referência).

#### a) CLUSTER 1 – NÃO RESIDENCIAL DEVEDOR

**Conclusões:** Elevados valores médio de Consumo e Fatura, e de unidades de referência com mais de R\$2.100,00 de dívida.

**Recomendações:** Manutenção da política de gestão e tratamento diferenciado das unidades segundo as faixas da dívida; Melhoramento das ações de pagamento parcelado da dívida e Contínuo monitoramento dos pagamentos.

#### b) CLUSTER 2 – FRAUDADOR

**Conclusões:** Realiza chamadas ao *Call Center* para Serviços e Informações; Apresenta Serviço Não Regulado e Débito Automático.

**Recomendações:** Tratamento Diferenciado nos Canais de Relacionamentos; Antecipar ações fraudulentas; Contínuo monitoramento dos pagamentos das unidades com Marcação de Fraude; Restrição de concessão aos Serviços Não Regulados e Débito Automático.

**c) CLUSTER 3 – RESIDENCIAL INADIMPLENTE COM COMUNICAÇÃO**  
**Conclusões:** Realiza chamadas ao *Call Center* para Serviços e Informações; Apresenta Serviço Não Regulado e Débito Automático.

**Recomendações:** Manutenção e aprimoramento de tratamento diferenciado nos Canais de Relacionamentos; Monitoramento das Ações de Cobrança, visando diminuição da inadimplência; Avaliação mais rigorosa para concessão aos Serviços Não Regulados e Débito Automático.

**d) CLUSTER 4 – INCOBRÁVEL SEM MEDIÇÃO**

**Conclusões:** Baixo percentual de Débito Automático, Serviços Não Regulados e chamadas ao *Call Center*.

**Recomendações:** Atualizar as Ações de Cobrança, visando recuperar aquelas unidades de referência com *status* Incobrável; Ampliar os Canais de Relacionamento; Antecipar a Inadimplência; Avaliação mais profunda quanto à concessão aos Serviços Não Regulados e Débito Automático.

**e) CLUSTER 5 – COMERCIAL ADIMPLENTE**

**Conclusões:** Alto percentual de Débito Automático; Não apresenta Serviços Não Regulados e chamadas ao *Call Center*.

**Recomendações:** Melhorar os Canais de Relacionamentos; Campanhas para oferecimento dos Serviços não Regulados e outros benefícios; Investigar necessidades a curto e médio prazo e Antecipar Atrasos, evitando futuras Inadimplências.

**f) CLUSTER 6 – BAIXO CONSUMIDOR COM SERVIÇOS**

**Conclusões:** Alto percentual de Débito Automático e de Serviços Não Regulados; Não realiza chamadas ao *Call Center*.

**Recomendações:** Realizar campanhas para melhorar os Canais de Relacionamentos; Fidelizar os Serviços não Regulados e Débito Automático; Antecipar Atrasos e futuras Inadimplências.

**g) CLUSTER 7 – INQUILINO COMUNICADOR**

**Conclusões:** Alto percentual de Reavisos e chamadas ao *Call Center*; Apresenta Serviços Não Regulados.

**Recomendações:** Tratamento Diferenciado nos Canais de Relacionamentos e Monitorar as atuais formas de pagamento, estimulando aos bons pagadores o Débito Automático e realizando ações que minimizem os atrasos e futuras inadimplências.

**h) CLUSTER 8 – ADIMPLENTE SEM DÉBITO AUTOMÁTICO**

**Conclusões:** Baixo volume de chamadas ao *Call Center* e de Serviço Não Regulado.

**Recomendações:** Realizar campanhas de Comunicações e de Relacionamentos; Monitorar as formas de pagamento, estimulando o pagamento via Débito Automático; Gerar campanhas para aquisição dos Serviços não Regulados; Antecipar Atrasos e futuras Inadimplências.

**i) CLUSTER 9 – ATRASADO MODERADO SEM COMUNICAÇÃO**

**Conclusões:** Alto percentual de reavisos; Baixo volume de Débito Automático e Serviço Não Regulado.

**Recomendações:** Antecipar novos atrasos de pagamento, evitando futuras Inadimplências; Personalizar as formas de negociação da dívida, monitorando-as constantemente; Novas campanhas de Relacionamento com foco na garantia do

pagamento da fatura, ressaltando as facilidades do pagamento via Débito Automático e Ações de Marketing para aquisição dos Serviços Não Regulados.

**j) CLUSTER 10 – ATRASADO ELEVADO SEM COMUNICAÇÃO**

**Conclusões:** Baixíssimo volume de Débito Automático.

**Recomendações:** Antecipar os atrasos de pagamento, monitorando-os continuamente e realizando comunicações sempre que necessário; Novas campanhas de Comunicações e Relacionamentos, e de estímulo ao pagamento via Débito Automático.

**k) CLUSTER 11 – ATRASADO LEVE COM COMUNICAÇÃO**

**Conclusões:** Discreto volume de Débito Automático e Serviços Não Regulados; Baixíssimos percentuais em Saldo Incobrável.

**Recomendações:** Fidelizar os meios de Comunicações, antecipando e evitando atrasos nos pagamentos das faturas; Gerar ações de Marketing para oferecer o pagamento via Débito Automático e Serviços Não Regulados.

**l) CLUSTER 12 – ATRASADO AUSENTE DE META**

**Conclusões:** Baixo volume de comunicação com o *Call Center*; Baixíssimo volume de Saldo Incobrável.

**Recomendações:** Gerar novas campanhas de Comunicações; Contínuo monitorando dos atrasados, evitando futuras inadimplências e melhorando o Relacionamento; Ações de Marketing para estimular o pagamento via Débito Automático e Serviços Não Regulados.

**m) CLUSTER 13 – ATRASADO LEVE SEM COMUNICAÇÃO**

**Conclusões:** Baixo volume de Débito Automático; Baixo volume em Saldo Incobrável.

**Recomendações:** Investigar as razões dos atrasos no pagamento da fatura; Estreitar os Canais de Relacionamentos permitindo gerar ações que evitem futuros atrasos e inadimplências; Ações de Marketing orientando o pagamento via Débito Automático.

**n) CLUSTER 14 – RESIDENCIAL ADIMPLENTE**

**Conclusões:** Não se comunica com o *Call Center*.

**Recomendações:** Gerar novas campanhas de Comunicações e Relacionamentos; Ações de Marketing para estimular o pagamento via Débito Automático e os Serviços não Regulados; Monitorar as formas de pagamento, sinalizando Atrasos e futuras Inadimplências.

#### IV. CONCLUSÕES SOBRE AS FONTES DE INFORMAÇÕES USADAS NA ANÁLISE DE SOBREPOSIÇÃO

Ampliando as conclusões apresentadas na seção anterior, nesta seção são apresentados os resultados dos 14 segmentos/*clusters* perante as fontes de informações usadas na análise de sobreposição (*Score* de Risco, Censo 2000 e Pesquisa ABRADÉE).

##### **Score de Risco:**

- Os *clusters* com *Score* de Risco A são: Cluster 1 – Não Residencial Devedor (76% das unidades); Cluster 2 – Fraudador (50%); Cluster 4 – Incobrável sem Medição (60%); Cluster 5 – Comercial Adimplente (93%); Cluster 6 – Baixo Consumidor com Serviços (81%); Cluster 7 – Inquilino Consumidor (53%); Cluster 8 – Adimplente sem Débito Automático (98%); Cluster 11 – Atrasado Leve com Comunicação (43%); Cluster 12 – Atrasado Ausente de Meta (60%); Cluster 13 – Atrasado Leve

sem Comunicação (57%) e Cluster 14 – Residencial Adimplente (85%).

- Os *clusters* com *Score* de Risco E são: Cluster 3 – Residencial Inadimplente com Comunicação (69% das unidades); Cluster 9 – Atrasado Moderado sem Comunicação (54%); Cluster 10 – Atrasado Elevado sem Comunicação (59%) e Cluster 11 – Atrasado Leve com Comunicação (33%). Pelo descrito acima, pode-se notar que o Cluster 11 (Atrasado Leve com Comunicação) é o único que apresenta consideráveis percentuais tanto na faixa A como na faixa E do *Score* de Risco desenvolvido pela AES Eletropaulo.

#### Censo 2000:

- Número médio de pessoas por domicílios particulares permanentes

A Tabela 2 mostra algumas medidas estatísticas do número médio de pessoas por domicílios particulares permanentes registrados no Censo 2000, onde se pode notar que os *clusters* 4 (Incoabrável sem Medição) e 13 (Atrasado Leve sem Comunicação) apresentam as maiores quantidades (em média) de pessoas por domicílios.

TABELA 2

ESTATÍSTICAS DO NÚMERO MÉDIO DE PESSOAS POR DOMICÍLIOS PARTICULARES PERMANENTES PARA CADA UM DOS 14 *CLUSTERS*

Cluster	Média	Desvio Padrão	Mediana	Percentil 95
1	3,02	0,75	3,17	3,84
2	3,40	0,70	3,50	4,10
3	3,34	0,71	3,49	4,00
4	3,69	0,64	3,80	4,28
5	3,05	0,73	3,22	3,77
6	3,07	0,74	3,25	3,82
7	3,14	0,72	3,30	3,84
8	3,12	0,76	3,31	3,82
9	3,22	0,73	3,38	3,89
10	3,40	0,98	3,64	4,13
11	3,51	0,82	3,68	4,30
12	3,22	1,07	3,55	4,11
13	3,69	0,59	3,57	4,17
14	3,38	0,90	3,59	4,09

- Renda média das pessoas responsáveis pelos domicílios particulares permanentes

A Tabela 3 mostra algumas medidas estatísticas da renda média das pessoas responsáveis pelos domicílios particulares permanentes observadas pelo Censo 2000, e a Figura 9 traz o histograma da distribuição dessa renda (até 20 salários mínimos) em cada um dos 14 *clusters*. Os *clusters* 1 (Não Residencial Devedor) e 6 (Baixo Consumidor com Serviços) apresentam as maiores rendas médias. Vale ressaltar que em todos os *clusters* o desvio padrão não é tão pequeno com relação à média.

- Rendimento Nominal Mensal das Pessoas Responsáveis pelos domicílios particulares permanentes.

As Tabela 4 e 5 apresentam o número médio de pessoas para cada uma das faixas de renda dentro de cada um dos 14 *clusters*, em números absoluto e relativo ao total de pessoas por *cluster*. Notadamente, há um maior número de pessoas nos *clusters* 10, 11, 12 e 14, e uma concentração maior de pessoas nas faixas de 3 a 10 salários mínimos (média de 47,5%) em todos os *clusters*.

TABELA 3

ESTATÍSTICAS DA RENDA MÉDIA DAS PESSOAS RESPONSÁVEIS PELOS DOMICÍLIOS PARTICULARES PERMANENTES PARA CADA UM DOS 14 *CLUSTERS*

Cluster	Média	Desvio Padrão	Mediana	Percentil 95
1	7,5	2,4	7,3	11,4
2	5,7	2,1	5,2	10,1
3	5,9	2,1	5,4	10,2
4	4,1	1,7	3,6	7,9
5	7,6	2,4	7,4	11,4
6	7,1	2,3	6,8	11,2
7	6,8	2,2	6,3	11,0
8	6,7	2,1	6,3	10,8
9	6,3	2,0	5,8	10,3
10	5,0	1,8	4,6	8,7
11	5,1	1,9	4,5	9,3
12	5,8	2,5	5,1	10,9
13	4,6	1,8	4,1	7,9
14	5,8	2,2	5,1	10,5

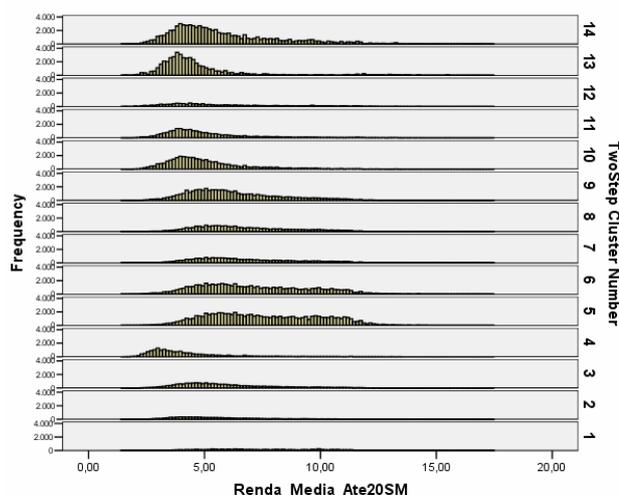


Figura 9. Histograma da Renda Média das Pessoas Responsáveis pelos Domicílios Particulares Permanentes (até 20 SMs) por *Cluster*  
Fonte: Ferramenta SPSS versão 13.0

#### Indicador da ABRADÉE – ISQP:

A Figura 10 mostra o Índice de Satisfação da Qualidade Percebida (ISQP) para cada um dos 14 *clusters*, onde se sabe que para o ano de 2005 este índice para a AES Eletropaulo foi de 77,9%. Esse indicador representa o percentual de clientes satisfeitos.

TABELA 4

NÚMERO MÉDIO DE PESSOAS PARA CADA UMA DAS FAIXAS DE RENDA DENTRO DE CADA UM DOS 14 *CLUSTERS*

Cluster	Até ½	½ a 1	1 a 2	2 a 3	3 a 5	5 a 10	10 a 15	15 a 20	Mais de 20	Sem SC
1	2,79	112,95	192,22	201,56	327,65	365,59	19,32	17,83	41,85	179,68
2	3,53	125,58	227,49	241,96	377,09	372,98	15,01	11,16	17,78	220,82
3	3,41	134,07	240,40	259,54	412,10	433,03	16,47	12,10	18,29	228,26
4	3,82	130,72	252,41	264,51	385,91	336,34	8,27	5,09	6,77	249,66
5	1,66	88,36	136,80	144,70	255,84	347,73	22,03	20,33	44,51	114,26
6	1,65	86,34	134,25	142,40	250,91	339,55	21,16	18,25	34,95	111,71
7	1,59	82,44	128,57	136,58	239,84	322,64	20,48	16,77	29,59	106,46
8	1,86	97,76	151,67	160,97	283,57	382,19	20,26	16,51	28,08	124,69
9	1,78	90,85	142,25	152,49	268,02	362,16	18,92	14,31	21,32	115,76
10	13,36	493,34	931,35	996,40	1511,04	1332,14	12,47	7,96	9,53	944,33
11	8,69	273,83	514,96	545,25	821,70	725,45	11,95	7,93	12,22	517,04
12	13,50	492,73	918,71	979,65	1495,26	1356,57	15,13	12,05	23,85	922,92
13	4,72	42,20	76,51	69,50	90,24	72,59	8,70	5,31	10,63	67,10
14	11,23	416,55	785,50	840,53	1276,82	1131,87	15,20	11,47	20,47	797,29

TABELA 5

PERCENTUAL DE PESSOAS PARA CADA UMA DAS FAIXAS DE RENDA DENTRO DE CADA UM DOS 14 CLUSTERS

Cluster	Até ½	½ a 1	1 a 2	2 a 3	3 a 5	5 a 10	10 a 15	15 a 20	Mais de 20	Sem SC
1	0,2	7,7	13,2	13,8	22,4	25,0	1,3	1,2	2,9	12,3
2	0,2	7,8	14,1	15,0	23,4	23,1	0,9	0,7	1,1	13,7
3	0,2	7,6	13,7	14,8	23,4	24,6	0,9	0,7	1,0	13,0
4	0,2	8,0	15,4	16,1	23,5	20,5	0,5	0,3	0,4	15,2
5	0,1	7,5	11,6	12,3	21,8	29,6	1,9	1,7	3,8	9,7
6	0,1	7,6	11,8	12,5	22,0	29,8	1,9	1,6	3,1	9,8
7	0,1	7,6	11,9	12,6	22,1	29,7	1,9	1,5	2,7	9,8
8	0,1	7,7	12,0	12,7	22,4	30,2	1,6	1,3	2,2	9,8
9	0,1	7,6	12,0	12,8	22,6	30,5	1,6	1,2	1,8	9,7
10	0,2	7,9	14,9	15,9	24,2	21,3	0,2	0,1	0,2	15,1
11	0,3	8,0	15,0	15,9	23,9	21,1	0,3	0,2	0,4	15,0
12	0,2	7,9	14,7	15,7	24,0	21,8	0,2	0,2	0,4	14,8
13	1,1	9,4	17,1	15,5	20,2	16,2	1,9	1,2	2,4	15,0
14	0,2	7,8	14,8	15,8	24,1	21,3	0,3	0,2	0,4	15,0

A Figura 11 mostra a média de satisfação dos seguintes índices: Índice de Aprovação do Cliente (IAC), Índice Intermediário de Satisfação (IIS), Índice de Satisfação Geral (ISG) e Iluminação Pública (IP), para cada um dos 14 clusters. Média AES Eletropaulo: IAC = 7,44 ; ISG = 7,34 e IP = 6,85

Os clusters 5, 6, 7, 8 e 9 apresentam os melhores ISQPs médios, enquanto que para os clusters 4, 10, 11, 12, 13 e 14 os valores de satisfação média do ISQP são relativamente baixos. Tal fenômeno se mantém, em menor grau, para os demais indicadores de satisfação. À exceção do Cluster 14 (Residencial Adimplente), a satisfação aparenta estar associada a experiências negativas de pagamento com a distribuidora.

## V. CONCLUSÕES GERAIS

Diante do modelo de segmentação de 14 clusters, caracterizado pela Figura 5, foi possível concluir, de forma sucinta, que as seguintes estratégias são potencialmente válidas:

- Possibilidades de diferenciar os atrasados e seus respectivos níveis de atrasos.
- Possibilidade de diferenciar os devedores (não residencial devedor, fraudador, residencial inadimplente, inco-brável sem medição).

- Percepção mais detalhada dos perfis dos Adimplentes, possibilitando direcionar campanhas para ampliação de mercado e de satisfação.
- Percepção mais detalhada do perfil do Baixo Consumidor e dos acessos aos Canais de Relacionamento com a AES Eletropaulo.

A partir da população de baixa tensão segmentada, as áreas de negócio: Gestão de Recebíveis, Gestão de Perdas Comerciais, e Relacionamento com o Cliente, poderão utilizar estes resultados como insumos para ações estratégicas diferenciadas por clusters, a fim de melhorar, sedimentar e ampliar o relacionamento empresa-cliente. Adicionalmente, a metodologia aqui descrita permite a AES Eletropaulo melhor entender seus consumidores, conseqüentemente, gerar ações pró-ativas capazes de reduzir os índices de reclamações, ações de indenizações e penalizações impostas pela ANEEL, pois tais ações deverão antecipar insatisfações e reclamações dos consumidores.

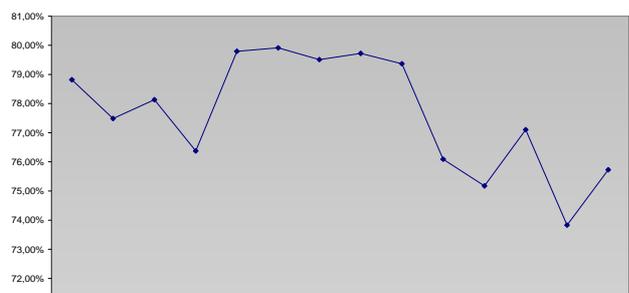


Figura 10. Índice de Satisfação da Qualidade Percebida (ISQP) da Pesquisa ABRADDEE 2005 para cada um dos 14 clusters

## Indicadores da ABRADDEE – IAC, IIS, ISG e IP:

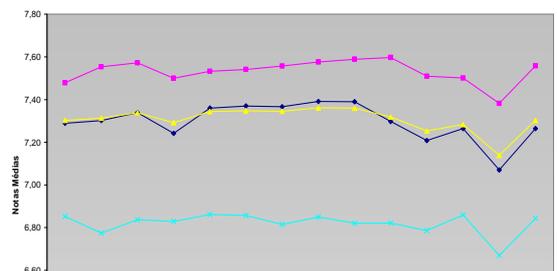


Figura 11. IAC, IIS, ISG e IP da Pesquisa ABRADDEE 2005 para cada um dos 14 clusters

A incorporação de informações de Fontes Secundárias, a criação de modelos de influência espacial (GIS) e a busca de padrões e tendências (*Data Mining*) permitem uma melhor caracterização sócio-econômico-demográfica dos clientes e do Mercado da Concessionária, o que propicia melhor eficácia em estudos de Fraude, Análise de Inadimplência, Contratação de Demanda por região e Identificação de área de incidência de medição estimada, Identificação de vocações econômicas das regiões (caracterizando setores comerciais e industriais), melhor ação em Planejamento de Rede, entre

outros, e, inclusive, auxílio na correção de inconsistências nas bases cadastrais da empresa.

Devido à segmentação da população de baixa tensão, áreas de negócios como Gestão de Recebíveis, Gerenciamento de Perdas Comerciais e Relacionamento com o Cliente estão habilitadas a utilizar esses resultados como *inputs* para ações estratégicas diferenciadas e focadas por segmentos, para que a relação empresa-cliente seja aprimorada.

Os direcionadores principais de marketing de uma organização devem ser perseguidos por uma estratégia realista e devem estar, primordialmente, alinhados com os objetivos financeiros e de construção de imagem da organização. A busca desse alinhamento deve ser suportada por um conjunto fundamental de objetivos de marketing de natureza quantitativa no curto, médio e longo prazos, a saber: redução das perdas comerciais, identificação e administração da inadimplência, aumento da rentabilidade oriunda da base atual de clientes, ampliação da base segmentada de clientes, e redução de custos.

Assim, a instrumentalização da AES Eletropaulo com uma sofisticada ferramenta de suporte à tomada de decisão baseada na integração das suas diferentes fontes internas e externas de informações, possibilita a geração de novos e eficientes elementos para o suporte à gestão de marketing, incrementando, otimizando e conduzindo ao reposicionamento estratégico dos processos de negócios junto ao mercado nos diversos cenários que se apresentem. Adicionalmente, a criação de uma inteligência diferenciada através da geração de banco de dados estruturado sobre perfis de consumo energético/comportamento também é possibilitada. Este trabalho inovou desde a forma de identificação dos sistemas de dados até a técnica de *Data Mining* aplicada para a descoberta dos perfis das unidades de referência de baixa tensão, e pode ser adaptado adequadamente a qualquer concessionária do setor elétrico.

O desenvolvimento deste projeto proporcionou o direcionamento do novo sistema de gestão de consumidores da empresa e as gerências de gestão de recebíveis, identificação de perdas comerciais e regularização de clandestinos estão utilizando os resultados em cada um dos seus processos para o desenvolvimento de estratégias diferenciadas a partir do modelo de segmentação comportamental.

## VI. AGRADECIMENTOS

Os autores agradecem a toda a equipe da AES Eletropaulo e da Fundação CPqD que trabalharam arduamente neste projeto para atingir os resultados almejados.

## VII. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] FRANCISCO, E.R., Customer Franchise: A Mina de Ouro do Geomarketing. *InfoGEO*, 56-57, 25/08/2002.
- [2] BERRY, M. J. A. & LINOFF, G. *Mastering Data Mining: the art and science of customer relationship management*, John Wiley & Sons, 2000.
- [3] BUSSAB, W. O & BOLFARINE, H., *Elementos de Amostragem*, 11o SINAPE/ABE, 1994.
- [4] DINIZ, C.A.R. & NETO, F.L., *Data Mining: Uma Introdução*, 14o. SINAPE/ABE, 2000.
- [5] FAYYAD, U. M., *Advances in Knowledge Discovery & Data Mining*, AAAI Press, 1996.
- [6] WEISS, S. M. & INDURKYA, N., *Predictive Data Mining: A Practical Guide*, Morgan K.Publishers, 1998.
- [7] ABRADDEE – Associação Brasileira de Distribuidores de Energia Elétrica. *Pesquisa ABRADDEE de Satisfação do Cliente Residencial 2003. Relatório Comparativo Eletropaulo*. Brasília, 2003.
- [8] ABRADDEE – Associação Brasileira de Distribuidores de Energia Elétrica. *Pesquisa ABRADDEE de Satisfação do Cliente Residencial 2004. Relatório Comparativo Eletropaulo*. Brasília, 2004.
- [9] ABRADDEE – Associação Brasileira de Distribuidores de Energia Elétrica. *Pesquisa ABRADDEE de Satisfação do Cliente Residencial 2005. Relatório Comparativo Eletropaulo*. Brasília, 2005.
- [10] BACHER, J. WENZIG, K. & VOGLER, M., *SPSS Two Step Cluster – A First Evaluation*, Universität Erlangen –Nürnberg, 2004.
- [11] BITTENCOURT, N. *Uso de Bancos de Dados Oficiais. Trabalho apresentado na Primeira Escola de Inverno da FGV-EAESP*. São Paulo, 2005.
- [12] IBGE – Instituto Brasileiro de Geografia e Estatística. *Censo Demográfico 2000. Documentação dos Microdados da Amostra*. Rio de Janeiro, 2002.
- [13] IBGE – Instituto Brasileiro de Geografia e Estatística. *Pesquisa Nacional por Amostra de Domicílios: Síntese de Indicadores 2004*. Rio de Janeiro, 2005.
- [14] REINA, C.S., NETO, J.O.R. & ARDENGHE, A.L.M., *Relatório de Visão Geral do Projeto de P&D ANEEL – Definição de Metodologia para Segmentação de Clientes, Identificação de Grupos e Conhecimento do Mercado, AES Eletropaulo e Fundação CPqD, Campinas,SP, 02/05/2005*.
- [15] REINA,C.S.R., PETRIELLI, A. & NETO, J.O.R., *Relatório de Planejamento Amostral do Projeto P&D ANEEL – AES Eletropaulo e Fundação CPqD, Campinas, SP, 15/07/2005*.
- [16] REINA,C.S.R. & PETRIELLI, A, *Relatório das Análises Exploratórias de Dados das Fontes Internas da Eletropaulo, Projeto P&D ANEEL – AES Eletropaulo e Fundação CPqD, Campinas, SP, 09/01/2006*.
- [17] REINA,C.S.R. & PETRIELLI, A, *Relatório Modelagem de Segmentação, Projeto P&D ANEEL – AES Eletropaulo e Fundação CPqD, Campinas, SP, 8/03/2006*.
- [18] NETO, J.O.R., *Relatório de Implantação, Projeto P&D ANEEL – Eletropaulo e Fundação CPqD, Campinas, SP, 06/01/2006*.
- [19] FRANCISCO, E.R., *Relação entre o Consumo de Energia Elétrica, a Renda e a Caracterização Econômica de Famílias de Baixa Renda do Município de São Paulo. Tese de Mestrado, Escola de Administração de Empresas de São Paulo, Fundação Getulio Vargas, São Paulo, Brasil, 2006*.