



FUNDAÇÃO  
GETULIO VARGAS



Escola de Administração  
de Empresas de São Paulo

# Inteligência Geográfica e Indicadores Econômicos baseados em Consumo de Energia Elétrica

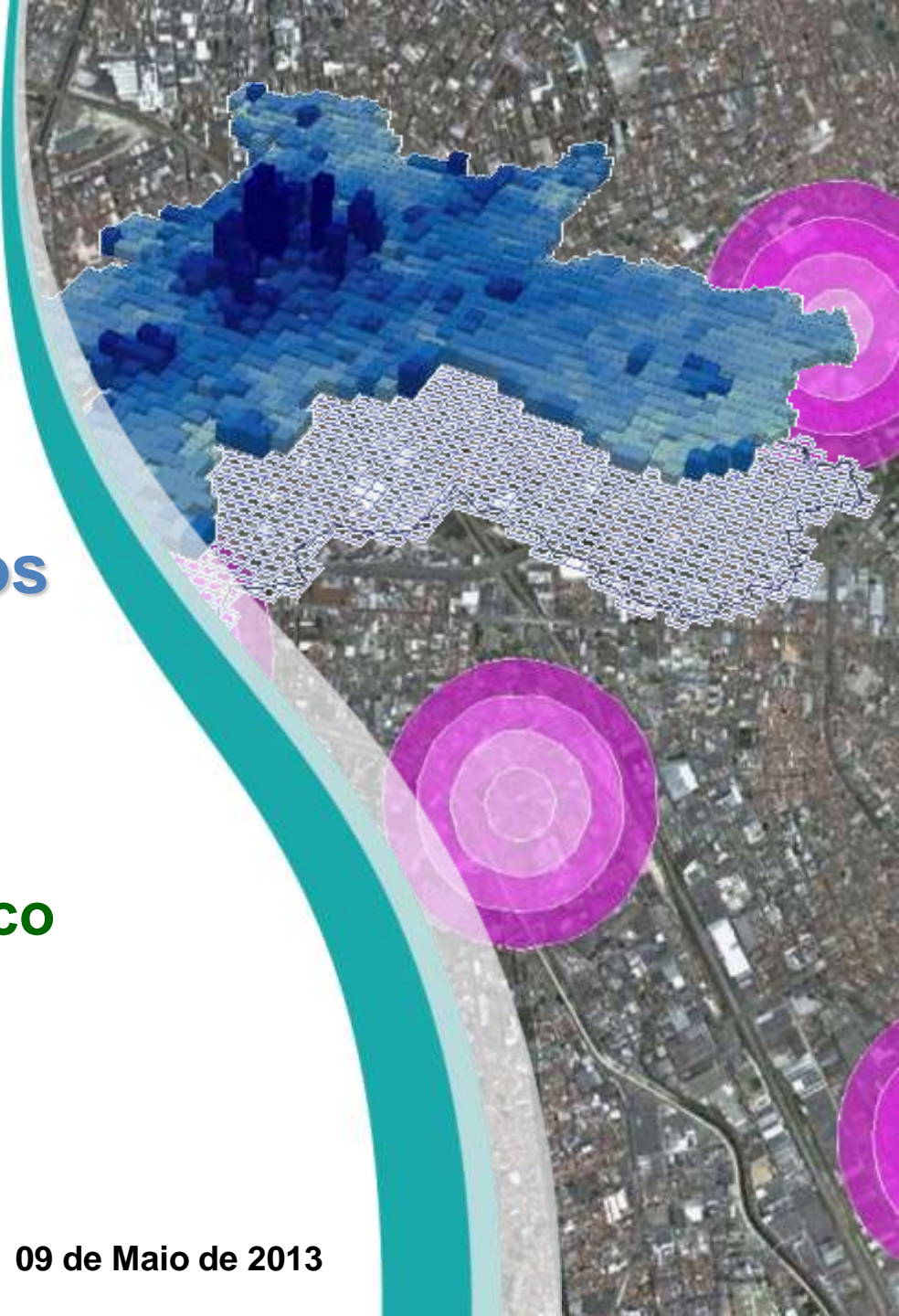
**Eduardo de Rezende Francisco**

[eduardo.francisco@fgv.br](mailto:eduardo.francisco@fgv.br)

IMQ - FGV-EAESP



09 de Maio de 2013





FUNDAÇÃO  
GETULIO VARGAS



Escola de Administração  
de Empresas de São Paulo



**SIRC** | Spatial Information  
Research Centre

Dunedin, Nova Zelândia



- Renda
  - Indicador tradicional adotado em estudos sobre Condições de Vida, Pobreza e Mercado (poder de compra)
- Dificuldade em conseguir informações precisas
  - 1 Altos Custos para obter cobertura e detalhamento
    - Depende do Censo Demográfico (IBGE, a cada 10 anos) para que se obtenha informação por setor censitário (conjunto de 300 domicílios)

- Renda

Indicador tradicional adotado em estudos sobre Condições de Vida, Pobreza e Mercado (poder de compra)

- Dificuldade em conseguir informações precisas

- Altos Custos para obter cobertura e detalhamento
- Sub ou superdeclaração, recusa (principalmente com Baixa Renda)
  - Critérios alternativos, como o Critério Brasil não são adequados para extremos da distribuição e necessitam constante atualização

Itens	Quantidade de itens				
	0	1	2	3	4 ou +
Televisão em cores	0	1	2	3	4
Rádio	0	1	2	3	4
Banheiro	0	4	5	6	7
Automóvel	0	4	7	9	9
Empregada mensalista	0	2	4	4	4
Aspirador de pó	0	0	0	0	0
Máquina de lavar	0	2	2	2	2
Videocassete e/ou DVD	0	2	2	2	2
Geladeira	0	4	4	4	4
Freezer (aparelho independente ou parte da geladeira duplex)	0	2	2	2	2

Analfabeto / Primário incompleto	Analfabeto / Até 3ª. Série Fundamental	0
Primário completo / Ginásial incompleto	Até 4ª. Série Fundamental	1
Ginásial completo / Colegial incompleto	Fundamental completo	2
Colegial completo / Superior incompleto	Médio completo	4
Superior completo	Superior completo	8

### Critério de Classificação Econômica Brasil

Classe	Pontos
A1	42-46
A2	35-41
B1	29-34
B2	23-28
C1	18-22
C2	14-17
D	8-13
E	0-7

- Renda

Indicador tradicional adotado em estudos sobre Condições de Vida, Pobreza e Mercado (poder de compra)

- Dificuldade em conseguir informações precisas

- 1 Altos Custos para obter cobertura e detalhamento
- 2 Sub ou superdeclaração, recusa (principalmente com Baixa Renda)
  - Critérios alternativos, como o Critério Brasil não são adequados para extremos da distribuição e necessitam constante atualização

Utilização de variáveis e indicadores que não têm estabilidade ao longo do tempo e que são pouco discriminadores dos estratos da população (PEREIRA, 2004)

Não adequado para a caracterização de famílias posicionadas nos extremos da distribuição de renda – em particular, para o consumidor de Baixa Renda (MATTAR, 1996; SILVA, 2004)

Estudos aprofundados necessitam de especializações do Critério Brasil: Indicadores de consumo de grande abrangência, capilaridade e que não necessitem de constante atualização podem ser úteis

Viés: Famílias mais Velhas têm mais condições de ascender economicamente do que as Famílias recém-configuradas

## Classe Social

- Estilo de Vida
- Comportamento



## Classe Econômica

- Posse de Bens duráveis
- Itens de conforto

- **Diferentes Propósitos**

- Padrões de Compra e Segmentação de Mercado
- Condições de Vida e Pobreza & Políticas Sociais

CHAPIN (1933), GUTTMAN (1942), MARTINEAU (1958), COLEMAN (1960), LEVY (1966), MYERS & GUTTMAN (1974), SCHANINGER (1981), MATTAR (1994, 1996)

- **Diferentes “Percepções” pela opinião pública**

- Classe Média ≈ “Classe C”
- Base da Pirâmide × “Classes D e E”
  - Distribuição por Faixas de Salários Mínimos
  - Segmentos Econômicos

## Classe Social

- Estilo de Vida
- Comportamento



## Classe Econômica

- Posse de Bens duráveis
- Itens de conforto

- “Classes D e E” versus “Classes D e E”

Critério Brasil (2000)	
Classe Econômica Critério Brasil	Distribuição da População (%)
A (A1 + A2)	6
B (B1 + B2)	23
C	36
D	31
E	4

Números de Prahalad (2002)		
Renda Familiar (em SMs)	“Segmento Econômico”	Distribuição da População (%)
Acima de 25	A	4,1
De 10 a 25	B	12,3
De 4 a 10	C	27,7
De 2 a 4	D	25,1
Até 2	E	30,8

Fonte: adaptado de ABEP, 2004; PRAHALAD, 2005 p. 119

➤ **Quem é afinal a Baixa Renda ?**



Revista Veja - 02/Abril/2008

Especial

## Ela empurra o crescimento

Em dois anos, 20 milhões de brasileiros saíram da pobreza e emergiram para a classe C. Esse fenômeno catapultou o consumo e expandiu a classe média, deixando o país a um salto do desenvolvimento

Julia Duailibi e Cíntia Borsato

A história de sucesso das nações hoje desenvolvidas possui um traço comum e inequívoco: a ascensão política e econômica da classe média. Essa evolução foi primeiro observada na Europa, onde sociedades estratificadas e sem mobilidade deram lugar a países dinâmicos após a irrupção do capitalismo. Esse fenômeno chegou tardiamente ao Brasil e nunca fincou raízes sólidas e duradouras. Houve fases de rápido crescimento, como na industrialização do início do século passado ou no milagre econômico dos anos 70. Mais recentemente, o país chegou a vislumbrar um salto rumo ao desenvolvimento com o Plano Cruzado. Mas, em todos esses casos, o aumento no padrão de vida dos mais pobres foi transitório e abortado pouco depois por crises econômicas. Agora, com a retomada do crescimento econômico, o país volta a se ver diante da oportunidade de romper, de maneira definitiva, com o subdesenvolvimento. Nos dois últimos anos, mais de 20 milhões de brasileiros saíram das camadas sociais mais baixas – as chamadas classes D e E – e alcançaram a classe C, a porta de entrada para a sociedade de consumo.

### VEJA TAMBÉM

#### Nesta reportagem

- Quadro: Ladeira acima
- Quadro: o recheio da pirâmide

#### Nesta edição

- Artigo: A era do crédito



Revista Istoé - 25/Agosto/2010

## A classe C vai ao paraíso

Eles já são 90 milhões de pessoas, respondem por quase 50% da renda nacional e não querem parar de subir na escada social brasileira

Adriana Nicacio



### RAIMUNDA DA SILVA GONÇALVES

Faxineira há mais de duas décadas, ela finalmente conseguiu construir sua casa própria, colocou um dos filhos na universidade, matriculou o caçula no curso de inglês e, aos 46 anos de idade, conquistou seu primeiro diploma. No ano passado, concluiu o primeiro grau do ensino fundamental. Agora que lê, escreve e está craque nas operações básicas de matemática, seu sonho é ir para a faculdade

Raimunda conseguiu mudar seu padrão de vida não só pelos seus espetaculares esforços pessoais. Sem o alicerce econômico alcançado pelo País nos últimos 15 anos, talvez pouco adiantaria seu empenho. Como muita gente, ela pôde ganhar mais dinheiro porque o Brasil vive um momento excepcional. Em cinco anos, 32 milhões de pessoas, o equivalente à metade da França, ascenderam socialmente. O fenômeno mais impressionante ocorreu com a **antiga classe média baixa, hoje chamada de classe C**, que se multiplicou e passou a representar a metade da população do País. Cerca de 90 milhões de brasileiros agora possuem renda familiar mensal entre R\$ 1.115 e R\$ 4.807 e se tornaram uma força tão poderosa que já é apontada por alguns especialistas como a classe dominante, no sentido econômico. Com o R\$ 1,3 mil que ganha fazendo faxinas nas casas de classe média alta de São Paulo, Raimunda passou a fazer parte dessa nova força econômica.

- Renda

Indicador tradicional adotado em estudos sobre Condições de Vida, Pobreza e Mercado (poder de compra)

- Dificuldade em conseguir informações precisas

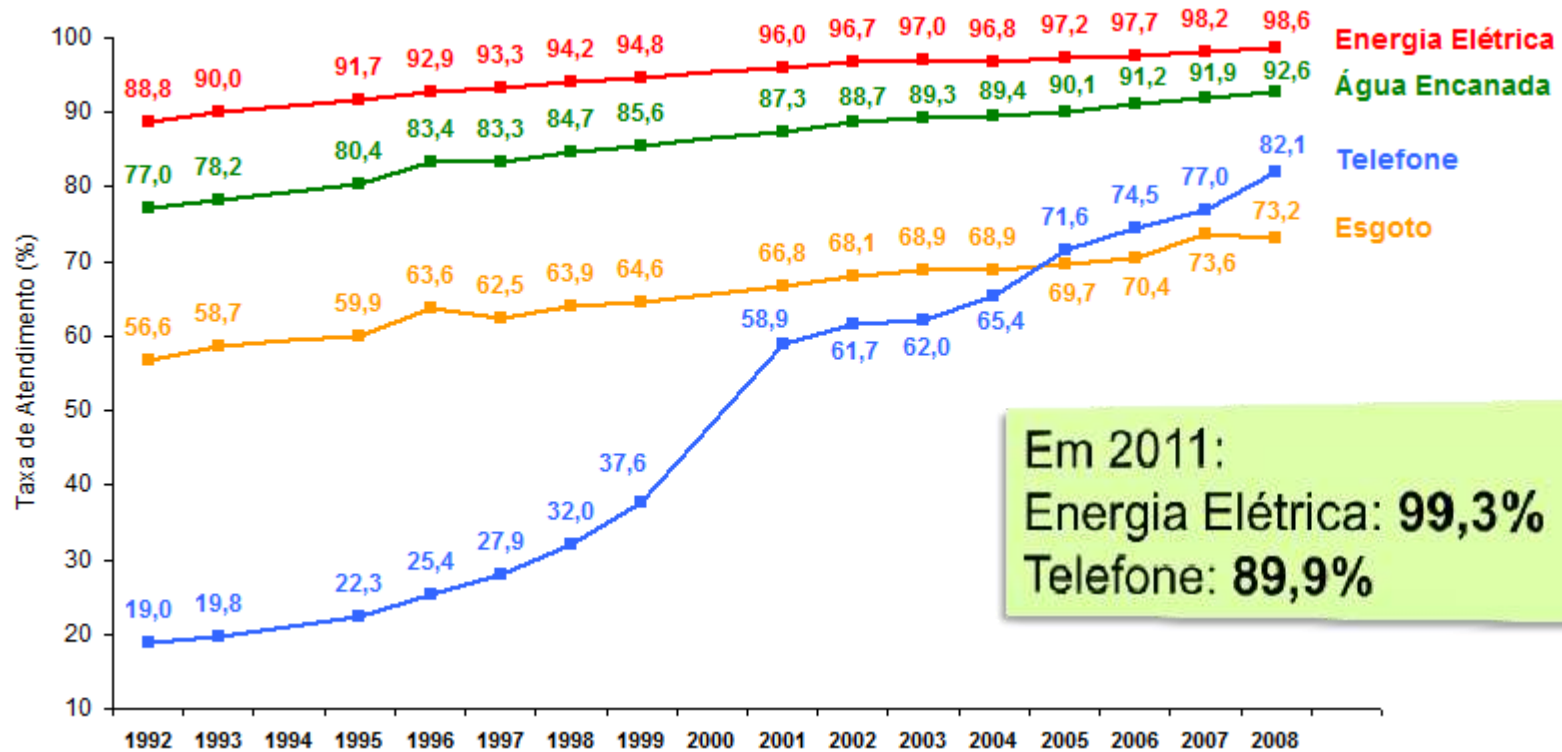
- 1 Altos Custos para obter cobertura e detalhamento

- 2 Sub ou superdeclaração, recusa (principalmente com Baixa Renda)

- Critérios alternativos, como o Critério Brasil não são adequados para extremos da distribuição e necessitam constante atualização

O Consumo de Energia Elétrica pode ser um bom Indicador Econômico para auxiliar processo de caracterização do consumidor

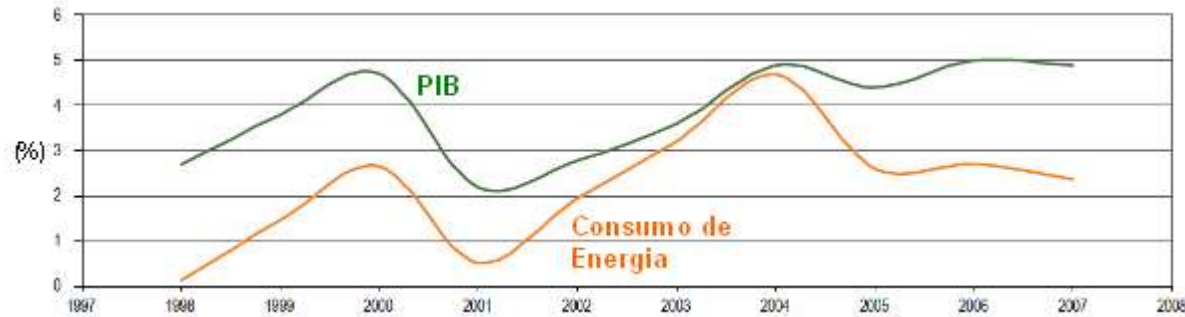
- **Serviço Essencial, de maior cobertura no país**
  - **99,3%** dos domicílios brasileiros ( **99,82%** na área urbana )
- **Coletado mensalmente, com precisão geográfica**
- **Histórico de faturamento e arrecadação**
- **Pode ser publicado de forma agregada**
  - **sem ferir sigilo ou confidencialidade**



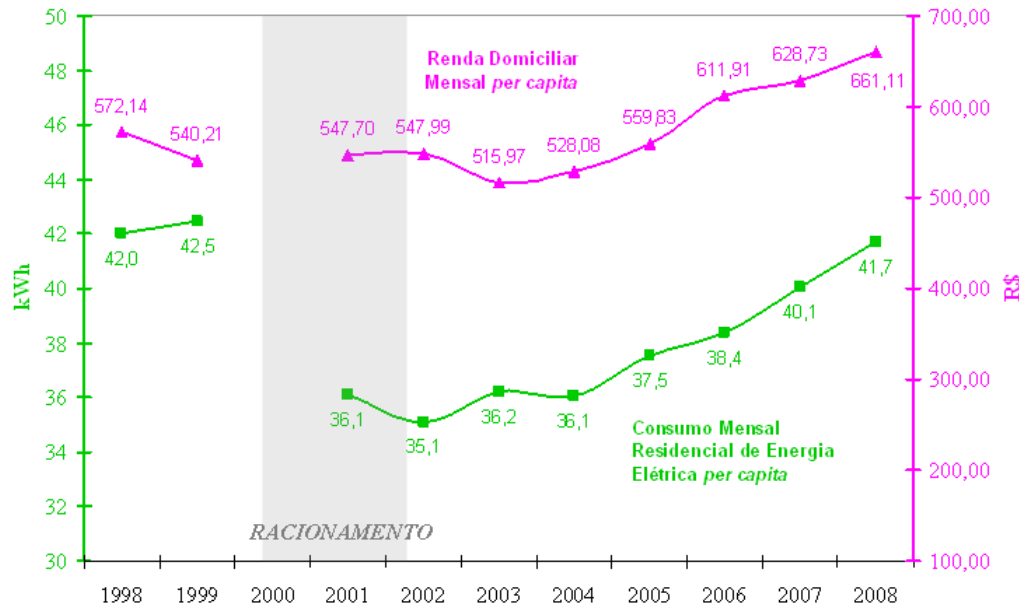
Em 2011:  
Energia Elétrica: **99,3%**  
Telefone: **89,9%**

- Serviço Essencial, de maior cobertura no país
  - **99,3%** dos domicílios brasileiros ( **99,82%** na área urbana )
- Coletado mensalmente, com precisão geográfica
- Histórico de faturamento e arrecadação
- Pode ser publicado de forma agregada
  - **sem ferir sigilo ou confidencialidade**

### Mundo: Variação do PIB × Variação Consumo de Energia

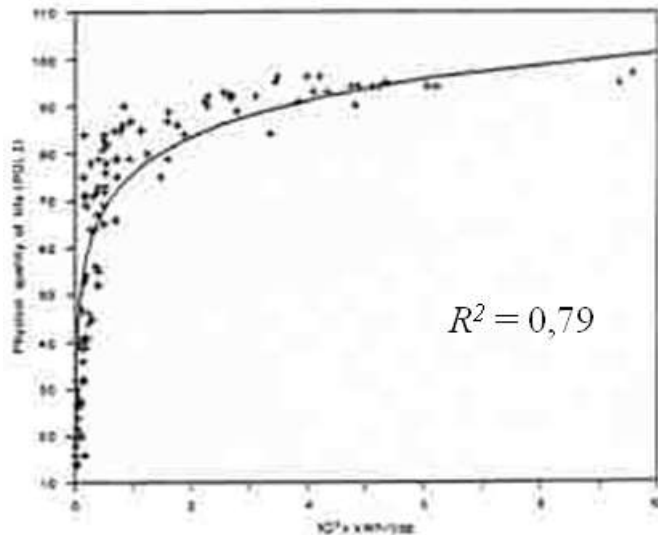


### Brasil: Renda Domiciliar *per capita* × Consumo Residencial de Energia Elétrica *per capita*



**Muitos estudos mostram alta aderência do Consumo Residencial ou do Consumo Industrial com o PIB, com Renda e com Qualidade de Vida**

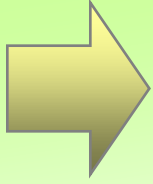
- **SPI – Sinalizador de Produção Industrial**
  - Indicador econômico mensal (AES Eletropaulo/FGV) que projeta a evolução da produção física mensal da Indústria Paulista com um mês de antecedência
- ***Physical Quality of Life Indicator (PQLI)*** (MARTIN, 1980)
  - 3 Dimensões: Expectativa média de Vida, Mortalidade Infantil e Taxa de Alfabetização (precursor do IDH)
  - Associação de 79% ( $R^2$ ) com log do Consumo de Energia *per capita* análise de dados de 112 países



## Índice Paulista de Responsabilidade Social (IPRS)

- Indicador municipal desenvolvido pela Fundação SEADE para o Estado de SP
- 3 Dimensões: Riqueza, Educação e Longevidade
- Tipologia (segmentação) de 5 grupos

Dimensão	Componentes (Variáveis)		Contribuição para o Indicador Sintético
Riqueza	Riqueza do Município	Consumo de energia elétrica na agricultura, no comércio e nos serviços	23,71%
		Valor adicionado fiscal <i>per capita</i>	13,90%
	Renda Familiar	<b>Consumo residencial de energia elétrica</b>	<b>43,51%</b>
		Remuneração média dos empregados com carteira assinada e do setor público	19,42%
Longevidade	Mortalidade perinatal		30,00%
	Mortalidade infantil		30,00%
	Mortalidade de adultos de 15 a 39 anos		20,00%
	Mortalidade de adultos de 60 anos ou mais		20,00%
Escolaridade	% de jovens de 15 a 17 anos que concluíram o ensino fundamental		36,08%
	% de jovens de 15 a 17 anos com pelo menos quatro anos de escolaridade		8,38%
	% de jovens de 18 a 19 anos que concluíram o ensino médio		35,37%
	% de crianças de 5 e 6 anos que freqüentam a pré-escola		20,17%



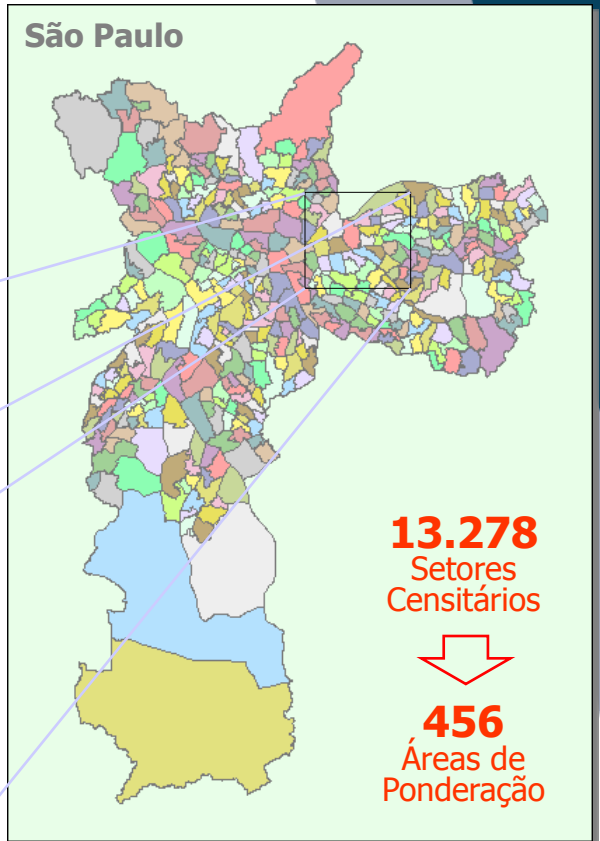
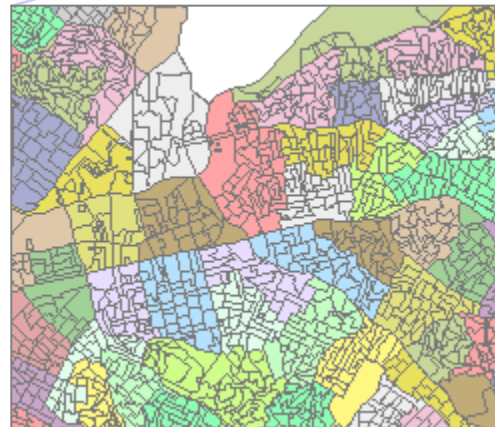
- Examinar a Relação entre Consumo de Energia Elétrica e Renda Domiciliar
- Propor a construção de um **indicador de Classificação Econômica** dos domicílios **baseado em Consumo de Energia Elétrica**, de forma a estender ou substituir o Critério Brasil



Objeto de Estudo:  
**Município de São Paulo**

Censo Demográfico de 2000  
(IBGE) + Dados AES  
Eletropaulo

Indicadores Agregados por  
conjunto de Setores  
Censitários



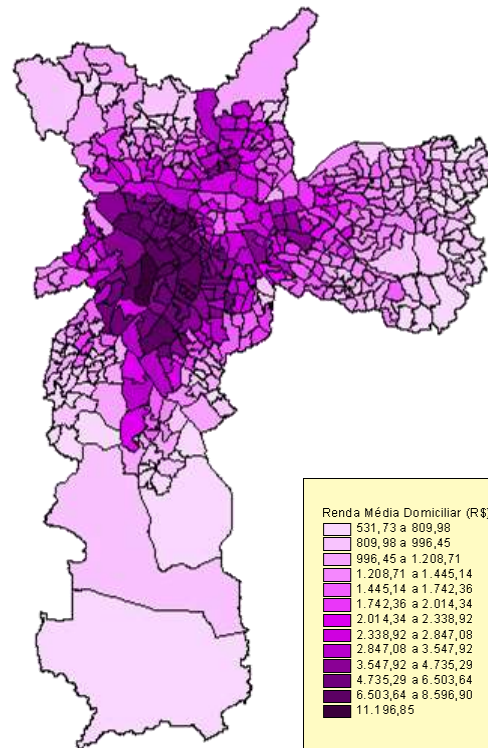
- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR



Objeto de Estudo:  
**Município de São Paulo**

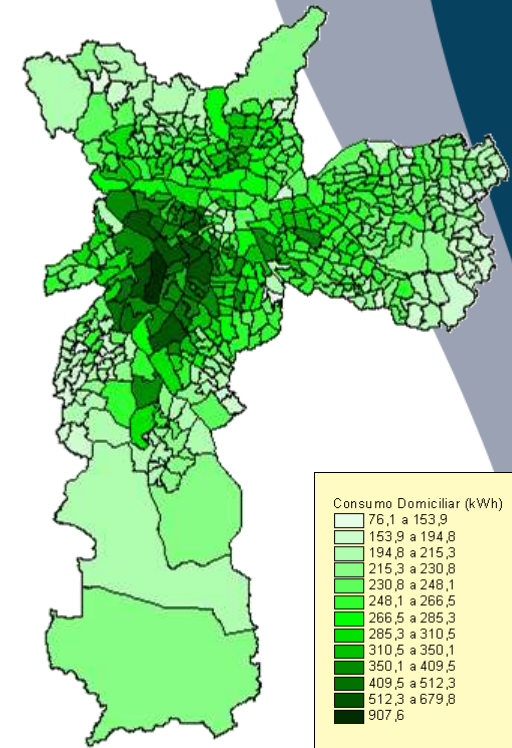
Censo Demográfico de 2000  
(IBGE) + Dados AES  
Eletropaulo

Indicadores Agregados por  
conjunto de Setores  
Censitários



**Renda Familiar**

(IBGE)



**Consumo Eletricidade**

(AES Eletropaulo)

Descreve e modela distribuições espaciais, padrões espaciais,

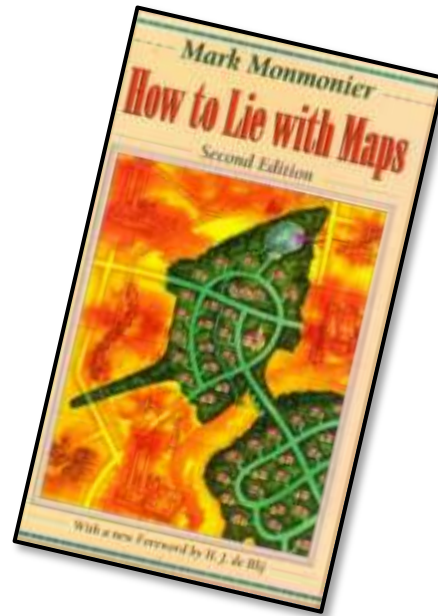
Incorpora o espaço (área, comprimento, proximidade, orientação e/ou relacionamentos espaciais) diretamente em seus modelos estatísticos

De muitas formas, a estatística espacial estende o que nossos olhos e mentes fazem intuitivamente para avaliar padrões, tendências e relações espaciais.

## Por que utilizar ?

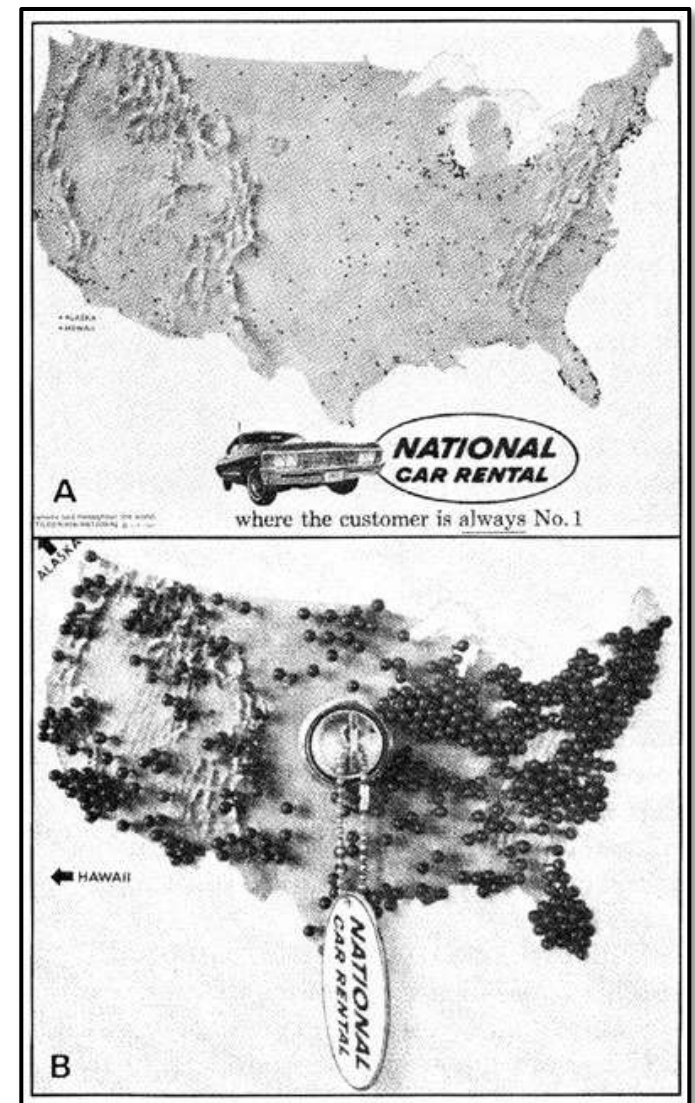
A Estatística Espacial ajuda-nos a avaliar:

- Padrões
- Relações
- Tendências

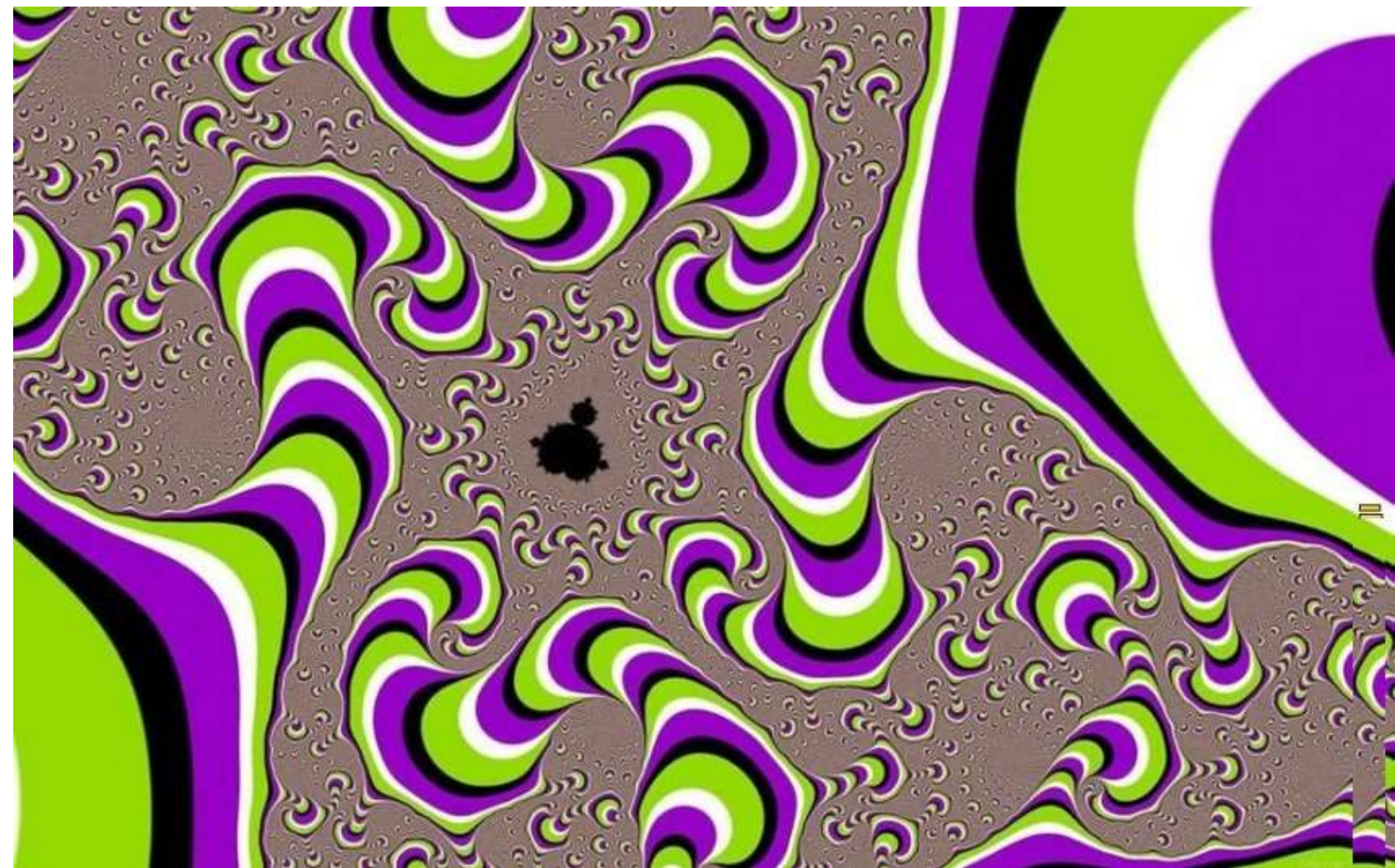


Como nós apresentamos e mapeamos nossos resultados (classes de legendas, símbolos, cores) pode destacar ou obscurecer a informação que queremos comunicar.

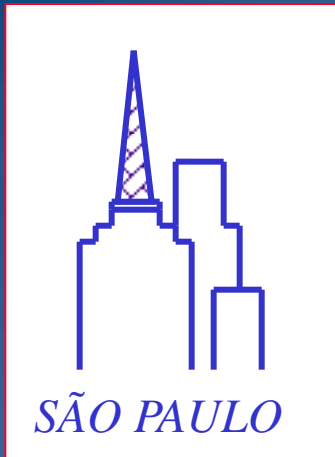
**Nossa INTUIÇÃO ESPACIAL  
pode nos pregar peças**



**Dois anúncios do National Car Rental. O mapa de baixo substituiu o de cima um ano depois do início da campanha.**



Nenhum movimento. Apenas uma ilusão de ótica...  
A imagem está estática, mas seus olhos e seu cérebro dimensionam uma ilusão de movimento...

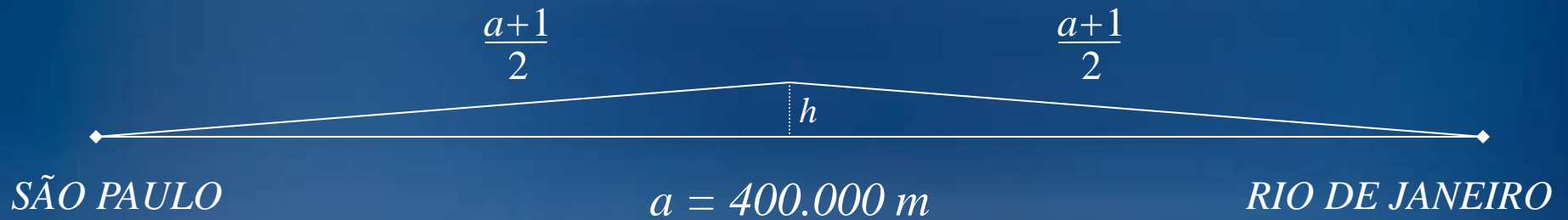


**A altura do triângulo formado é  
MAIOR, IGUAL ou MENOR  
que 1 metro ?**



**50 Quarteirões**  
**126 Andares**

# Intuição Espacial



$$h^2 = \left[ \frac{a+1}{2} \right]^2 - \left[ \frac{a}{2} \right]^2 = \frac{1}{4}(2a+1)$$

$$h = \frac{1}{2} \sqrt{2a+1}$$

Sendo  $a = 400.000 \text{ m}$ , temos

$$h = \frac{1}{2} \sqrt{800.001} \approx 447 \text{ m}$$

- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR
  - Utilizando Estatística Espacial em Modelos de Regressão



Objeto de Estudo:  
**Município de São Paulo**

Censo Demográfico de 2000  
(IBGE) + Dados AES  
Eletropaulo

Indicadores Agregados por  
conjunto de Setores  
Censitários

❖ **Incorpora o Espaço e as Relações Espaciais nos Modelos Estatísticos**

- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR
  - Utilizando Estatística Espacial em Modelos de Regressão



Objeto de Estudo:  
**Município de São Paulo**

Censo Demográfico de 2000  
(IBGE) + Dados AES  
Eletropaulo

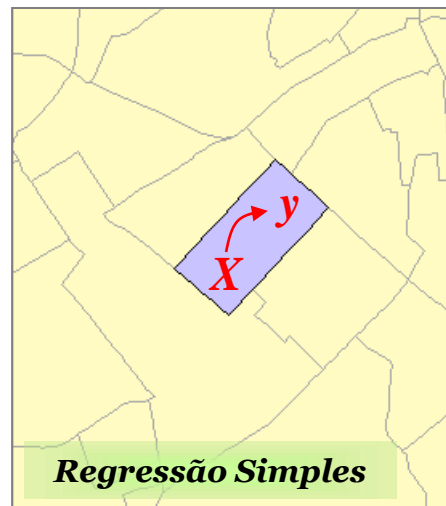
Indicadores Agregados por  
conjunto de Setores  
Censitários

❖ **Incorpora o Espaço e as Relações Espaciais nos Modelos Estatísticos**

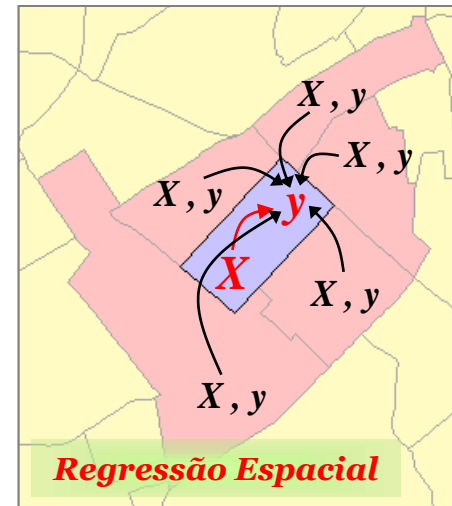
$y$  : Renda (R\$)  
 $X$  : Consumo de Energia (kWh)

- *Spatial Auto-Regression ( SAR )*
- *Geographically Weighted Regression ( GWR )*
- **SAR+GWR**

$$y = X\beta + \varepsilon$$



$$y = \rho W y + X\beta + \varepsilon$$

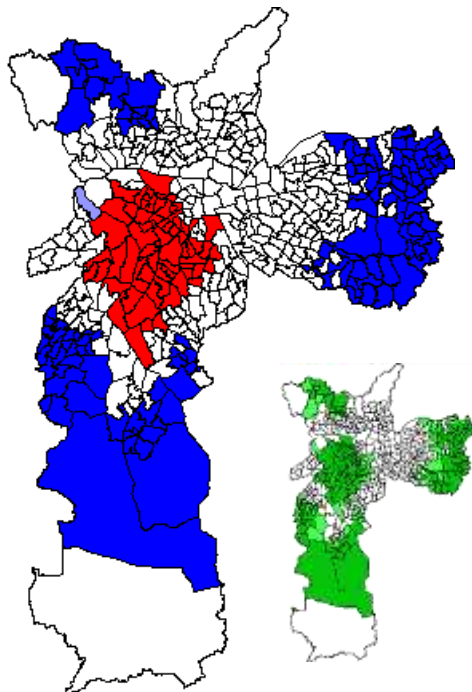


- Auto-Correlação Espacial dos Construtos

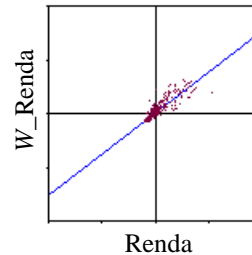
- $I$  de Moran da Renda  $> 0,77$  e  $I$  de Moran do Consumo de Energia  $> 0,58$   
na maioria das matrizes de vizinhança adotadas
- LISA maps : alta renda no centro, áreas de transição no entorno e heterogeneidade das baixas rendas nas regiões periféricas
- Quanto maior o número médio de vizinhos por área, menor a magnitude do  $I$  de Moran, maior a significância das áreas Baixo-Baixo (principalmente) e Alto-Alto e menores as áreas de transição de regimes

$W$ : Contiguidade de 1ª ordem do tipo *Queen*

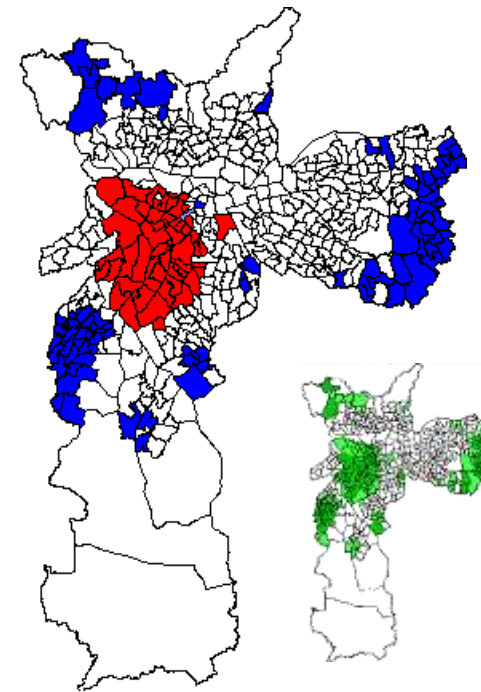
Renda Domiciliar



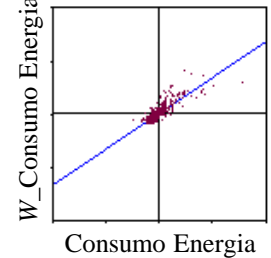
$I = 0,7722$



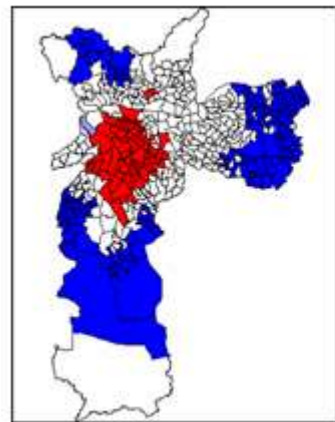
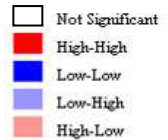
Consumo de Energia Elétrica



$I = 0,6661$

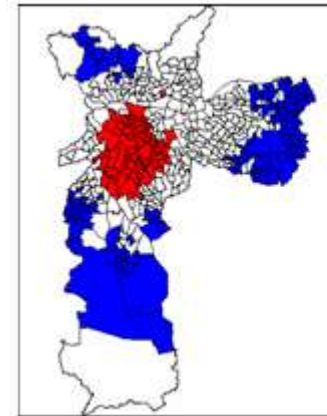
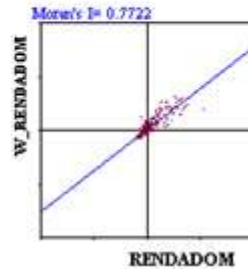


- *Moran's I* da Energia apresentou altos valores (0,77+) para diferentes “tipos” de vizinhança
- Sugere alta influência da vizinhança no comportamento da Renda Domiciliar
- *LISA maps*: Aumento da concentração de renda na direção Periferia - Centro (o mesmo fenômeno para o Consumo de Energia Elétrica)



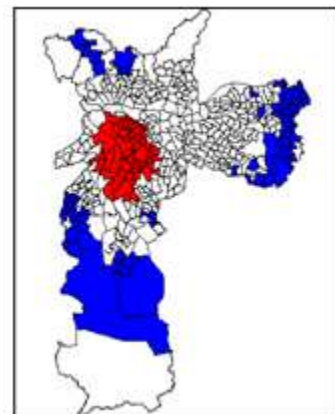
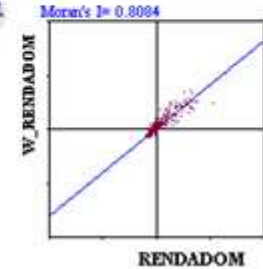
Neighborhood Graph:  
*Queen Graph*

Moran's I = **0,7722**



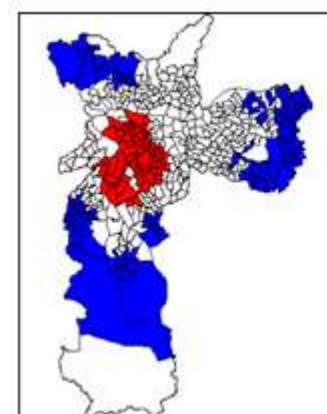
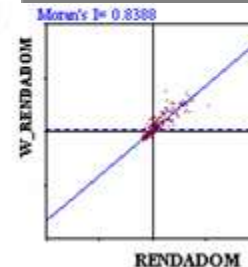
Neighborhood Graph:  
*Gabriel's Graph*

Moran's I = **0,8084**



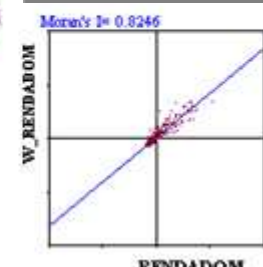
Neighborhood Graph:  
*Relative Neighbor Graph*

Moran's I = **0,8388**



Neighborhood Graph:  
*Sphere of Influence Graph*

Moran's I = **0,8246**



# Indicadores de Renda baseados em Consumo de Energia Elétrica

## RESULTADOS – INVESTIGAÇÃO TERRITORIAL



OLS

SAR  
Adjacência de 1ª ordem

GWR  
Kernel adaptável gaussiano –  $k=9$

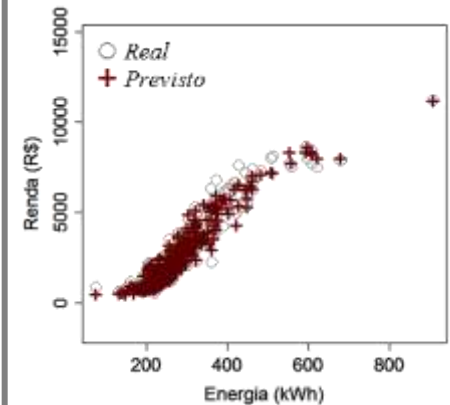
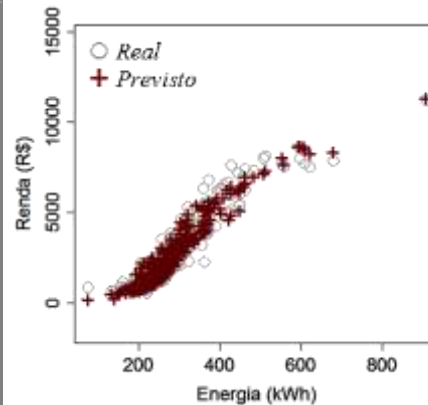
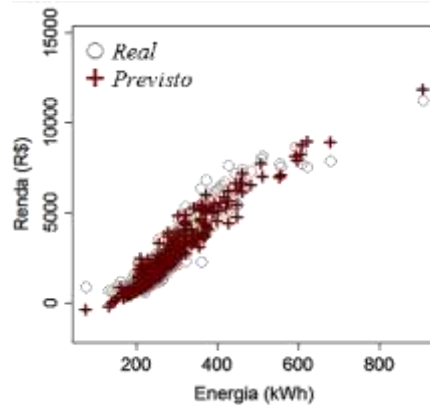
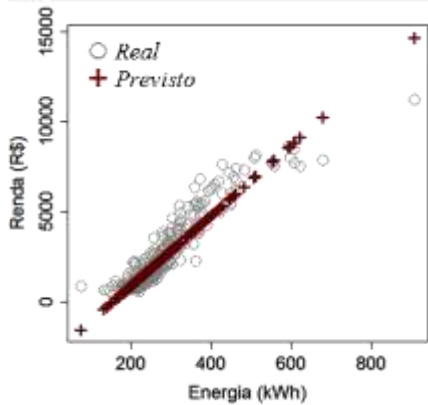
GWR+SAR  
Kernel adaptável gaussiano –  $k=9$

$R^2 = 86,80\%$

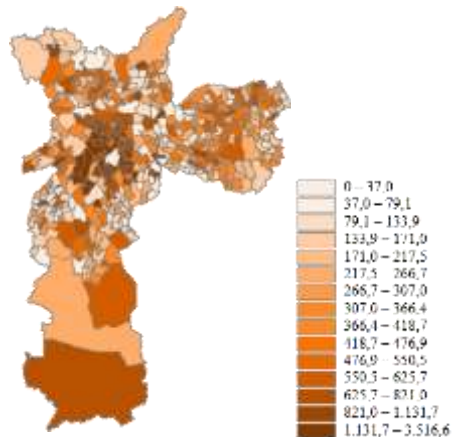
$R^2 = 94,44\%$

$R^2 = 96,80\%$

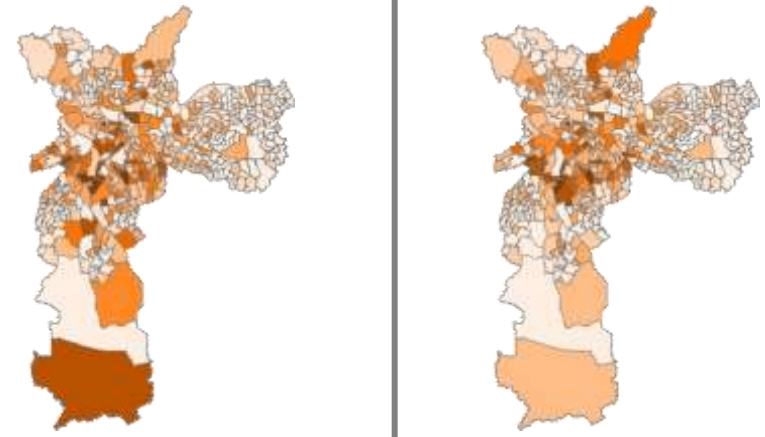
$R^2 = 98,01\%$



Mapa Resíduos Absolutos



Mapa  $R^2$  Local (3D)



- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR
  - Utilizando Estatística Espacial em Modelos de Regressão

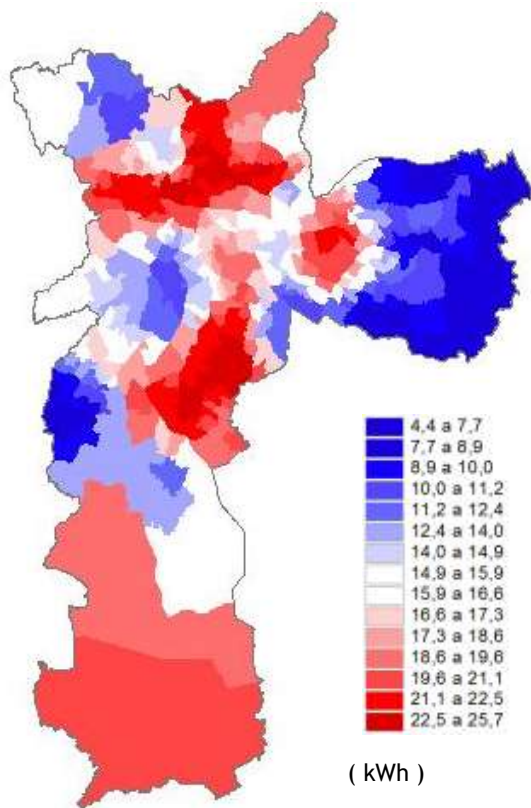


**RESULTADO TERRITORIAL:**  
O Consumo de Energia Elétrica explica  
**93 a 98%** da Renda Familiar

Objeto de Estudo:  
**Município de São Paulo**

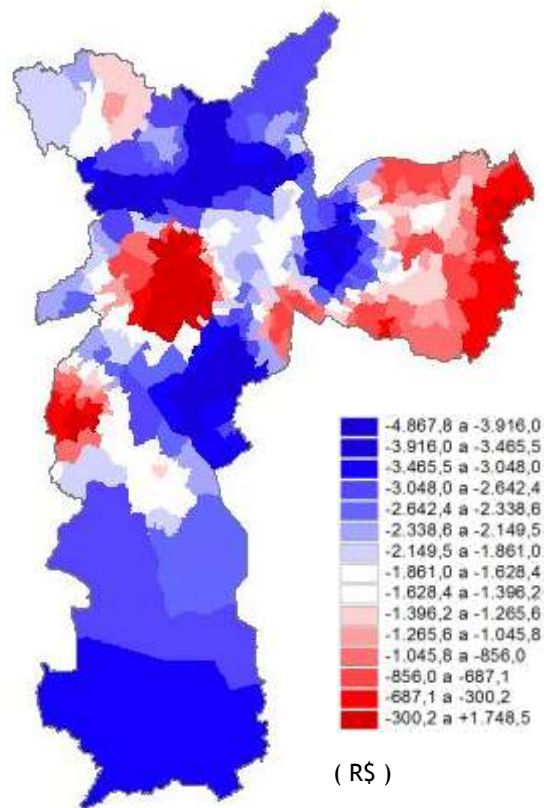
Censo Demográfico de 2000  
(IBGE) + Dados AES  
Eletropaulo

Indicadores Agregados por  
conjunto de Setores  
Censitários



### Energia ( $\beta_1$ )

Nas áreas centrais, o **parâmetro Energia** tem uma contribuição mais baixa que nas periferias.



### Intercepto ( $\beta_0$ )

Ocorre o oposto com o **Intercepto** (valor inicial da **Renda** considerada na predição).

**k = 3**

**k = 9**

**k = 25**

**k = 50**

**k = 100**

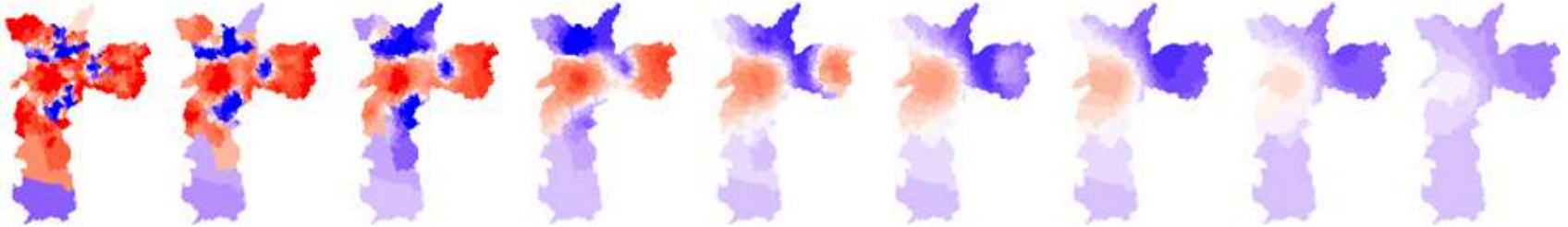
**k = 150**

**k = 200**

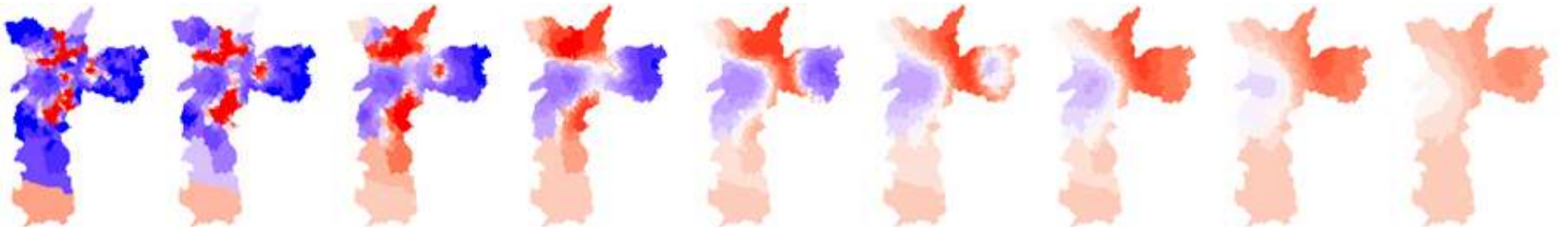
**k = 300**

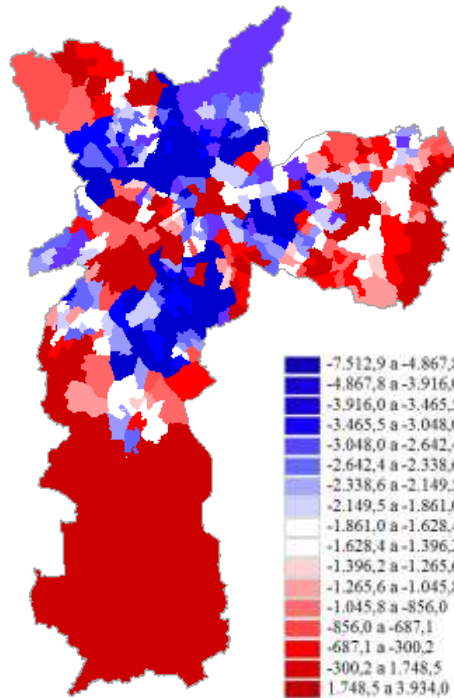
**k = 400**

Intercepto ( $\beta_0$ )

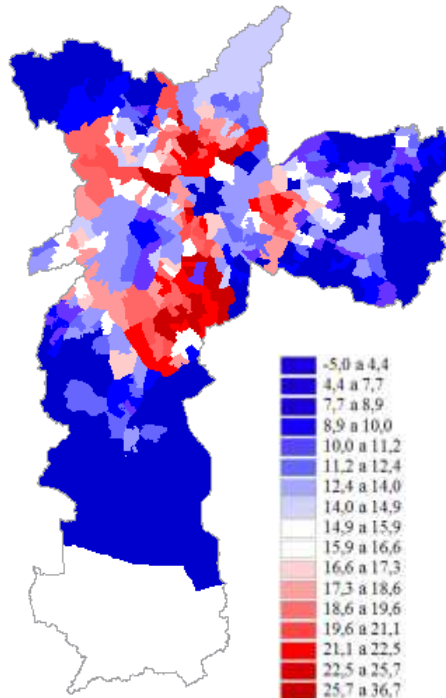


Consumo de Energia Elétrica ( $\beta_1$ )

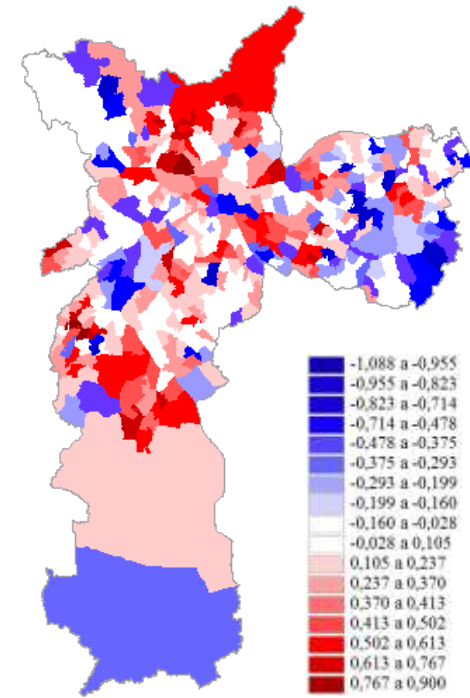




Intercepto ( $\beta_0$ )



Consumo de  
Energia Elétrica ( $\beta_1$ )

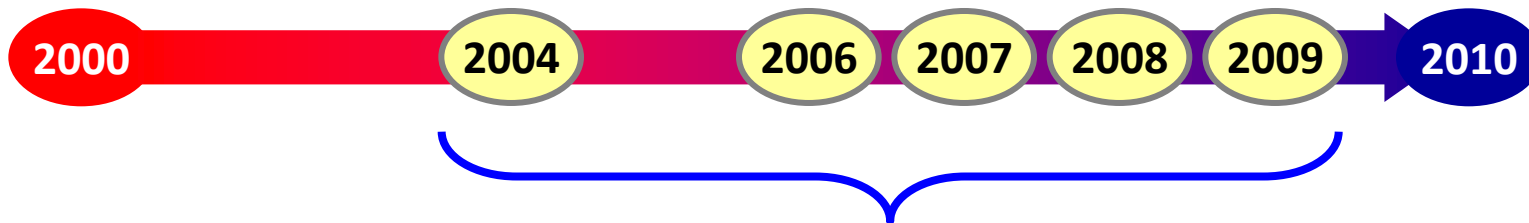


W\_Renda ( $\rho$ )

- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR
  - Utilizando Estatística Espacial em Modelos de Regressão



**RESULTADO TERRITORIAL:**  
O Consumo de Energia Elétrica explica  
**93 a 98%** da Renda Familiar



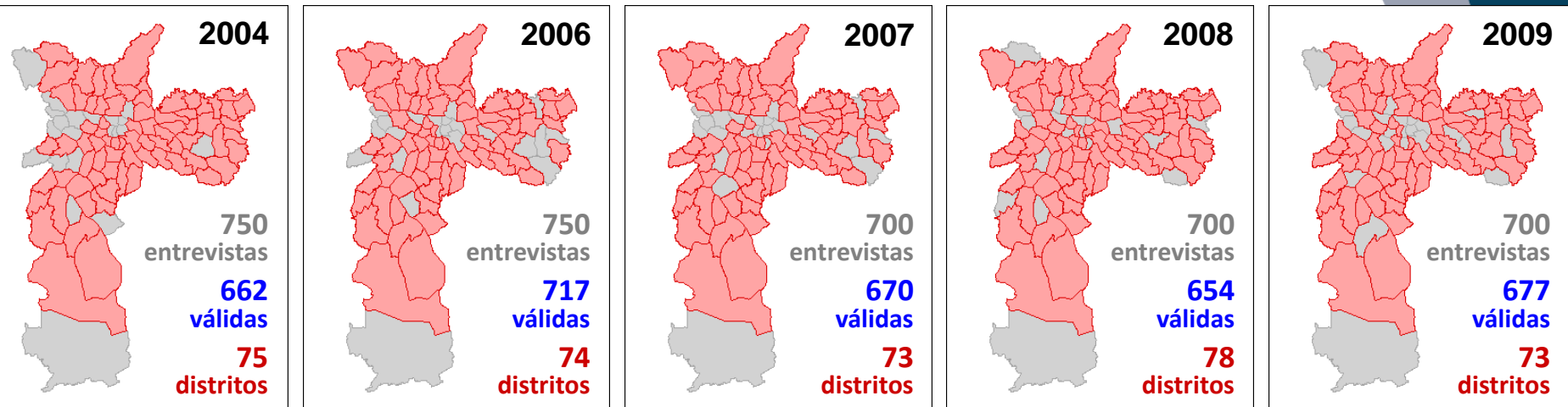
Avaliação a partir de dados domiciliares  
coletados através da Pesquisa ABRADÉE

- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR
  - Utilizando Estatística Espacial em Modelos de Regressão



Avaliação Histórica de 2004 a 2009, utilizando Informações da Pesquisa ABRADÉE :

- Renda Familiar (R\$)
- Valor Médio da Conta de Luz (R\$)



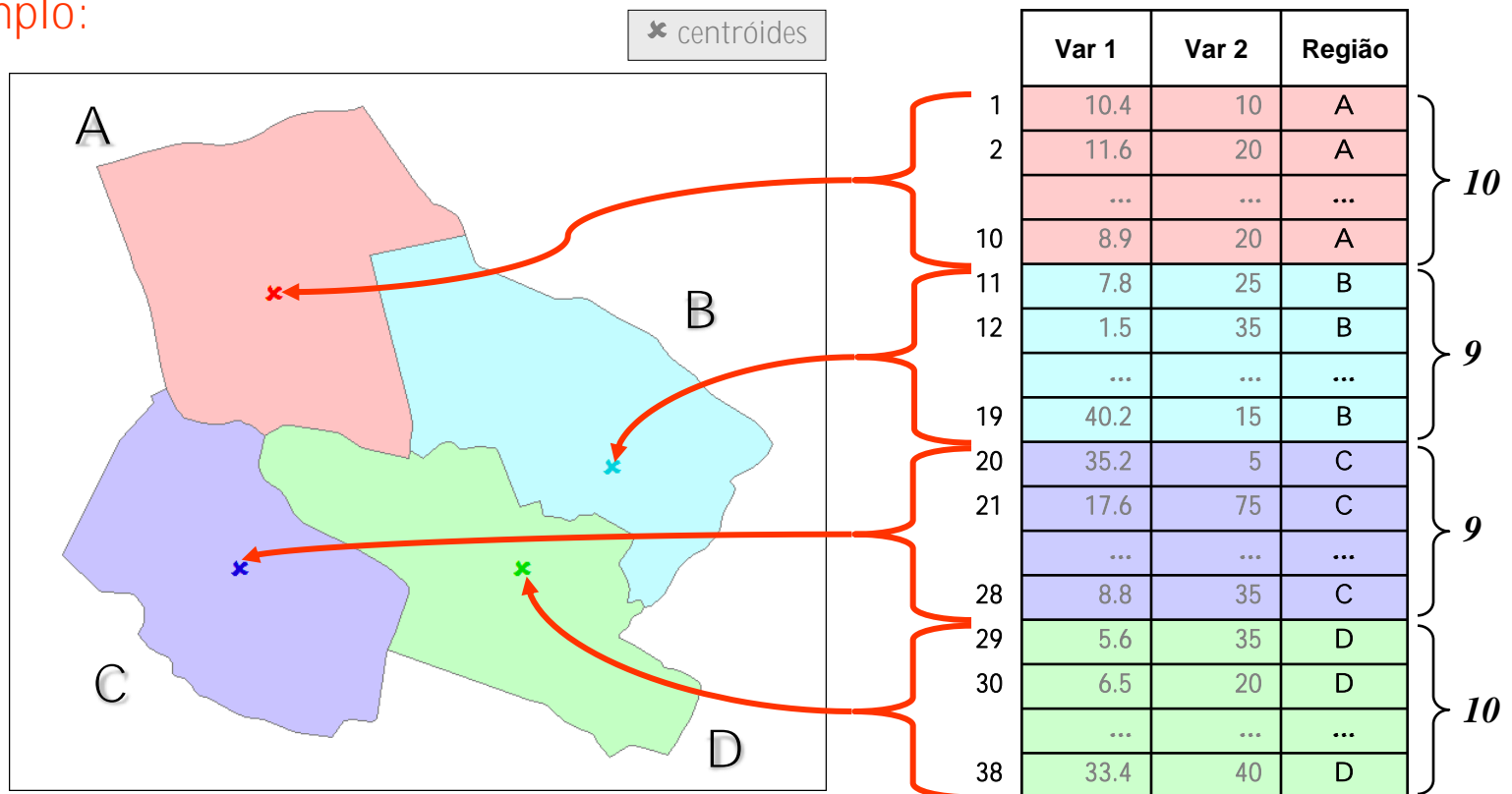
Dentro de cada distrito de São Paulo,  
ONDE foram realizadas as entrevistas ?

# Indicadores de Renda baseados em Consumo de Energia Elétrica

## MOTIVAÇÃO METODOLÓGICA

- Pesquisas de Campo (*Surveys*)
  - Em muitos casos, Dados “Pontuais” não estão disponíveis
    - Apenas informações sobre a região onde a observação está localizada (município, distrito, código postal etc)
  - Não apropriado para algumas análises espaciais:  
especialmente considerando ponderação baseada em distâncias (GWR, SAR)

Exemplo:



# Indicadores de Renda baseados em Consumo de Energia Elétrica

## MOTIVAÇÃO METODOLÓGICA

GWR:  $\hat{y}_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i) x_i$

$$\beta(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) Y$$

$$w_{ij} = \left(1 - \left(d_{ij}^2 / h^2\right)\right)^2 \quad \text{Bisquare}$$

$$w_{ij} = e^{-(d_{ij}/h)^p} \quad \text{Gaussiano}$$

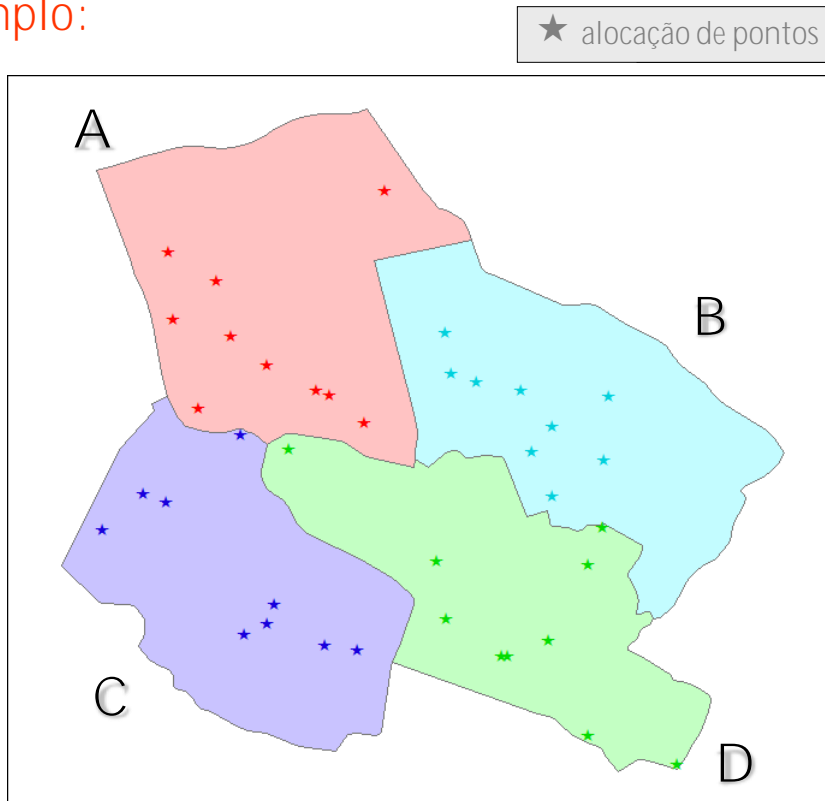
se  $d_{ij} < h$ , caso contrário  $w_{ij} = 0$

$d_{ij}$ : distância entre observações  $i$  e  $j$

$h$ : largura de banda

Essas podem ser distribuições mais realísticas  
(ou talvez estejamos adicionando alguma “perturbação” aos dados...)

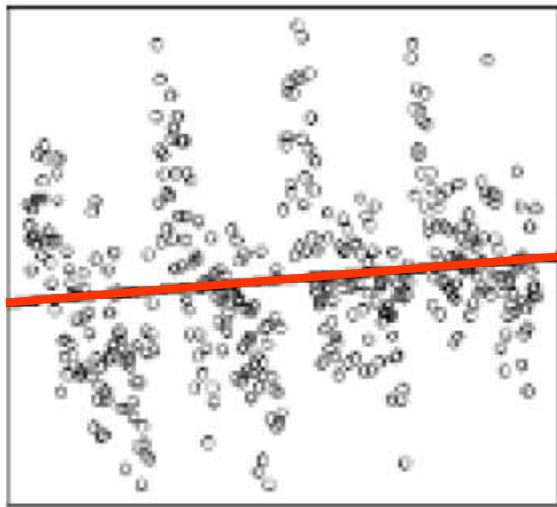
Exemplo:



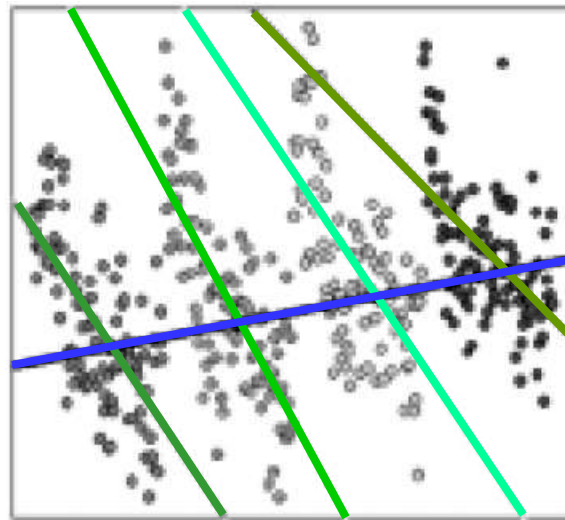
Dados Coletados (domicílios)

	Var 1	Var 2	Região	
1	10.4	10	A	10
2	11.6	20	A	
...	...	...	...	
10	8.9	20	A	9
11	7.8	25	B	
12	1.5	35	B	
...	...	...	...	9
19	40.2	15	B	
20	35.2	5	C	
21	17.6	75	C	10
...	...	...	...	
28	8.8	35	C	
29	5.6	35	D	10
30	6.5	20	D	
...	...	...	...	
38	33.4	40	D	

- Oposto do “Modifiable Areal Unit Problem” (MAUP)
  - **Descreve o efeito das relações espaciais observadas nos dados devido a escala e zoneamento** (OPENSHAW & TAYLOR 1979; JELINSKI & JIANGUO 1996)
    - **Diferentes resultados podem ser obtidos para a mesma análise estatística para diferentes níveis de resolução espacial ou a partir de reagrupamento de regiões**
  - **Exemplo (LEMOS DIAS *et al.* 2001) :**



LM para todos os indivíduos:  
 $R^2 = +0,15$



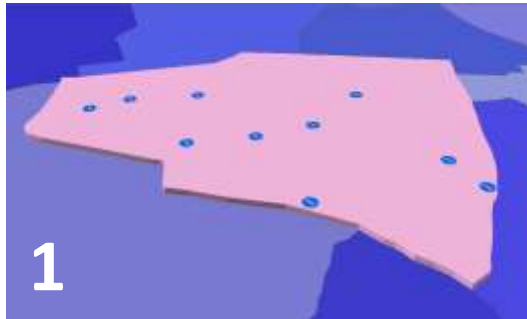
LM (por grupo):  $R^2$  varia de -0,5 a -0,8

LM (apenas considerando a média de cada grupo):  $R^2 = +0,49$

- **Agregação em regiões reduz a variabilidade e tende a aumentar a correlação entre as variáveis**

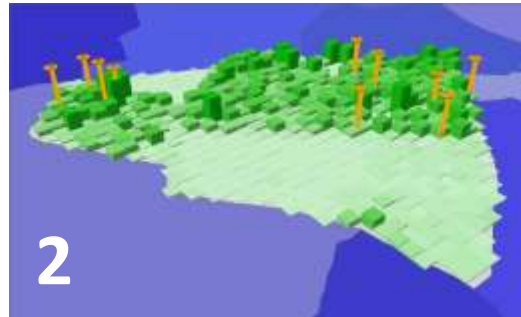
Dentro de cada distrito de São Paulo,  
ONDE foram realizadas as entrevistas ?

6 Algoritmos de Espalhamento de  
Pontos nos Polígonos dos Distritos



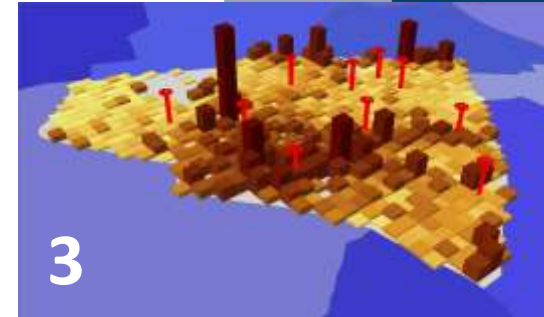
1

Entrevistas espalhadas aleatoriamente



2

Baseadas na densidade de domicílios



3

Baseadas na densidade de consumo de energia



4

Sorteio de Domicílios baseado no o de domicílios de consumo de energia



5

Entrevistas espalhadas aleatoriamente em setores censitários urbanos



6

Baseada no Plano Amostral da Pesquisa ABRADEE

➤ 1000 Aplicações Iteradas dos Algoritmos

# Investigação Domiciliar: Estrutura Metodológica

**Amostra Pesquisa ABRADDEE**  
2004, 2006, 2007, 2008, 2009



**Pontos =**

- Centróides
- Sedes dos Distritos

**1**  
Aplicação de WLM, SAR e GWR

1 x 5 amostras

**2** Aplicação Iterativa das Alternativas de Alocação Espacial das Entrevistas nos Distritos, para cada Amostra da Pesquisa ABRADDEE

- Alternativa 1
- Alternativa 2
- Alternativas 3A e 3B
- Alternativas 4A e 4B
- Alternativa 5
- Alternativa 6

1.000 Iterações por Alternativa por Pesquisa



**Pontos alocados pelas Alternativas**

1.000 x 8 x 5 amostras



**3**

Aplicação de GWR

Para cada Amostra  $i$ , aplicação com *kernel* adaptável (fixo em número de vizinhos) considerando:

$k_i$  = sugerido pela minimização do AIC da Amostra  $i$

**4**

Comparação com regressões com a Localização Real das Entrevistas

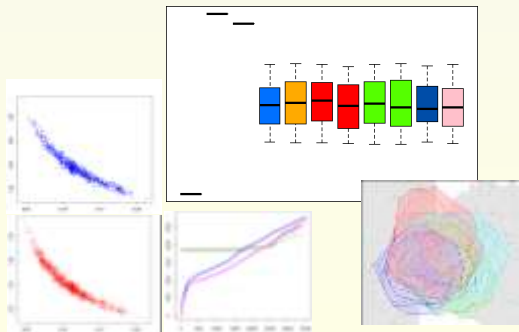
para Amostra Pesquisa ABRADDEE 2009



Localização Real das Entrevistas

**5**

**Análises dos Resultados** (*Box plots*  $R^2$ , Associação entre  $k$  e  $R^2$ , homogeneidade e clusterização espacial das amostras locais GWR, entre outras)

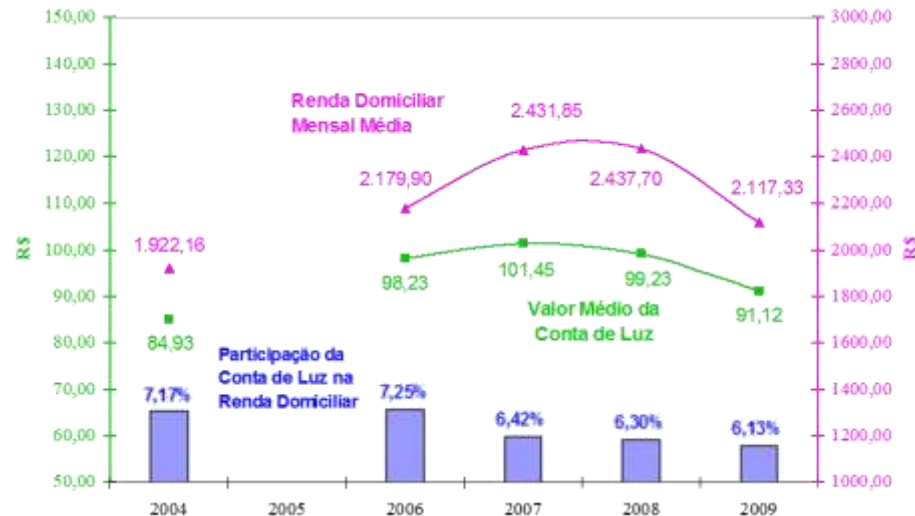


- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR
  - Utilizando Estatística Espacial em Modelos de Regressão



Avaliação Histórica  
de 2004 a 2009, utilizando  
Pesquisa ABRADEE

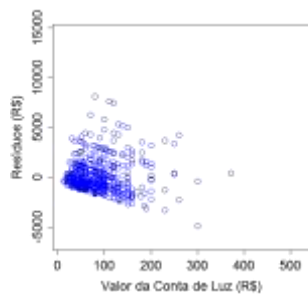
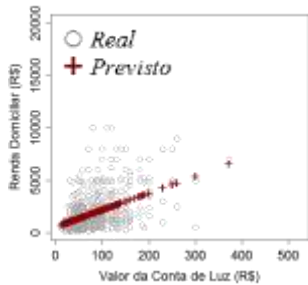
**Relação Estável** entre  
Renda e Consumo de  
Energia nos últimos anos,  
condizente com a  
tendência de evolução  
dessas variáveis (IBGE,  
IPEA)



### WLM

$$\hat{y} = 618,19 + 16,05x$$

$R^2 = 19,35\%$



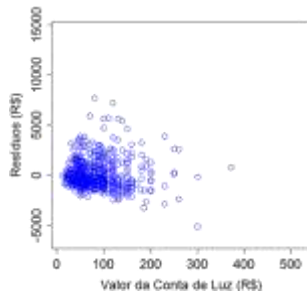
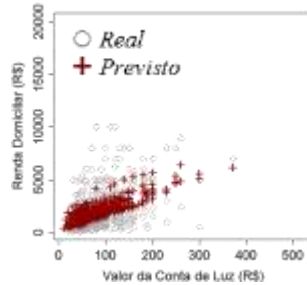
### SAR

Vizinhança mínima de 4 domicílios

#### Centróides

$$\hat{y} = -135,06 + 13,63x + 0,49W_y$$

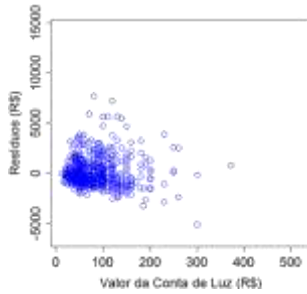
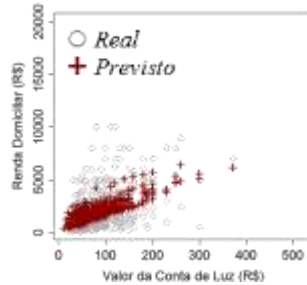
$R^2 = 33,06\%$



#### Sedes

$$\hat{y} = -145,74 + 13,58x + 0,49W_y$$

$R^2 = 33,39\%$



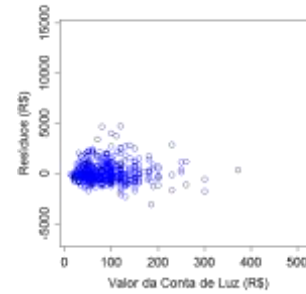
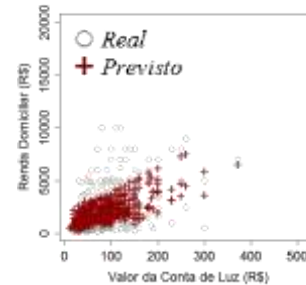
### GWR

Kernel adaptável gaussiano

#### Centróides (k=57)

	Mín	1Q	Med	3Q	Máx
$\beta_0$ :	-781	486	876	1334	3150
$\beta_1$ :	-9,18	6,55	10,8	17,7	32,6

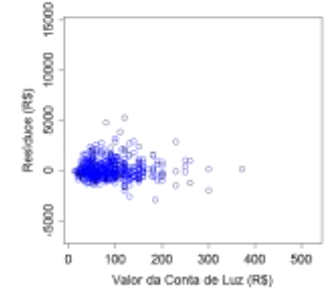
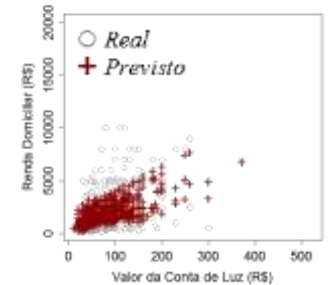
$R^2 = 43,59\%$



#### Sedes (k=47)

	Mín	1Q	Med	3Q	Máx
$\beta_0$ :	-631	573	868	1391	2943
$\beta_1$ :	-9,18	6,55	10,8	17,4	31,1

$R^2 = 45,63\%$



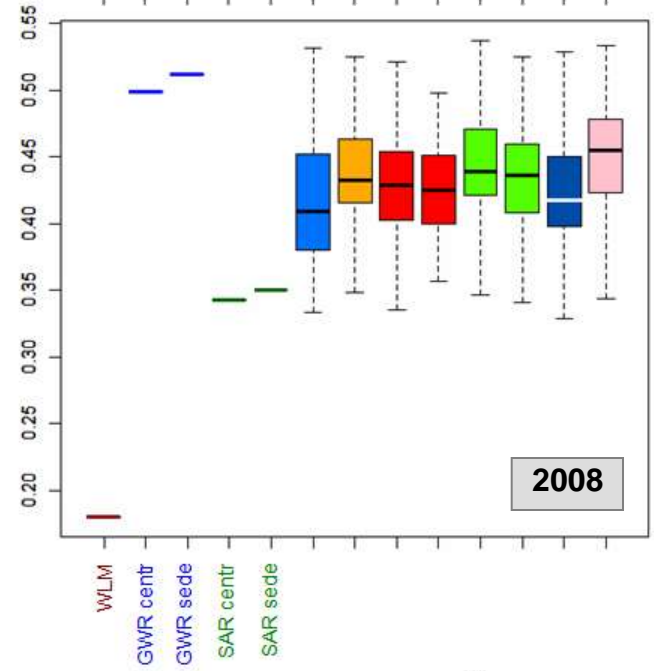
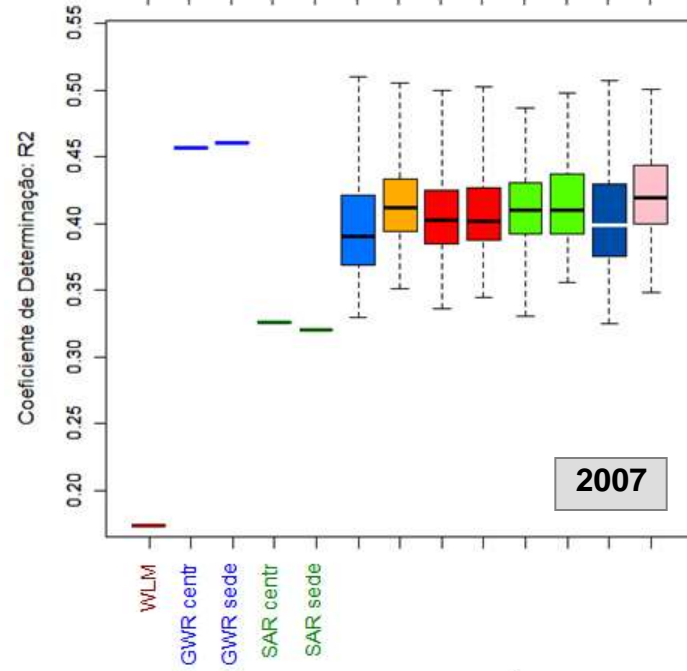
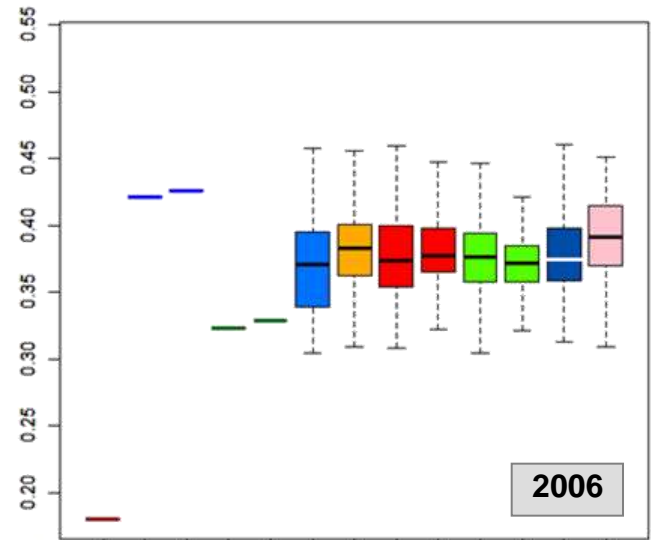
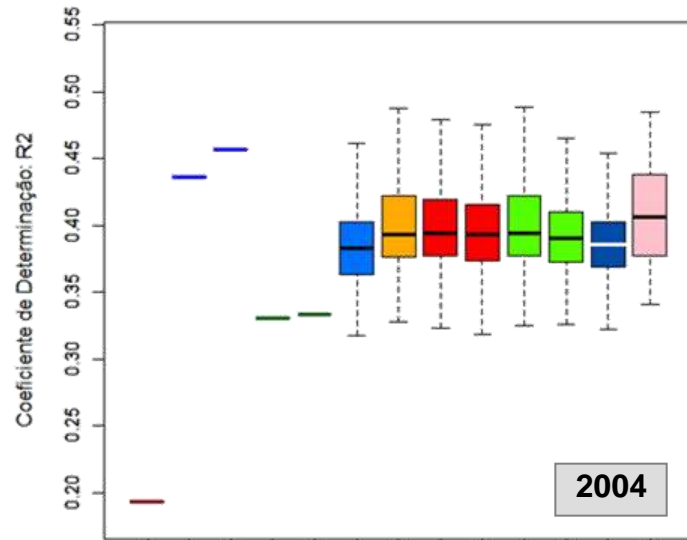
### Pesquisa ABRADÉE 2004 (n = 662)

- Resultados dos Modelos WLM, SAR e GWR muito similares entre as rodadas



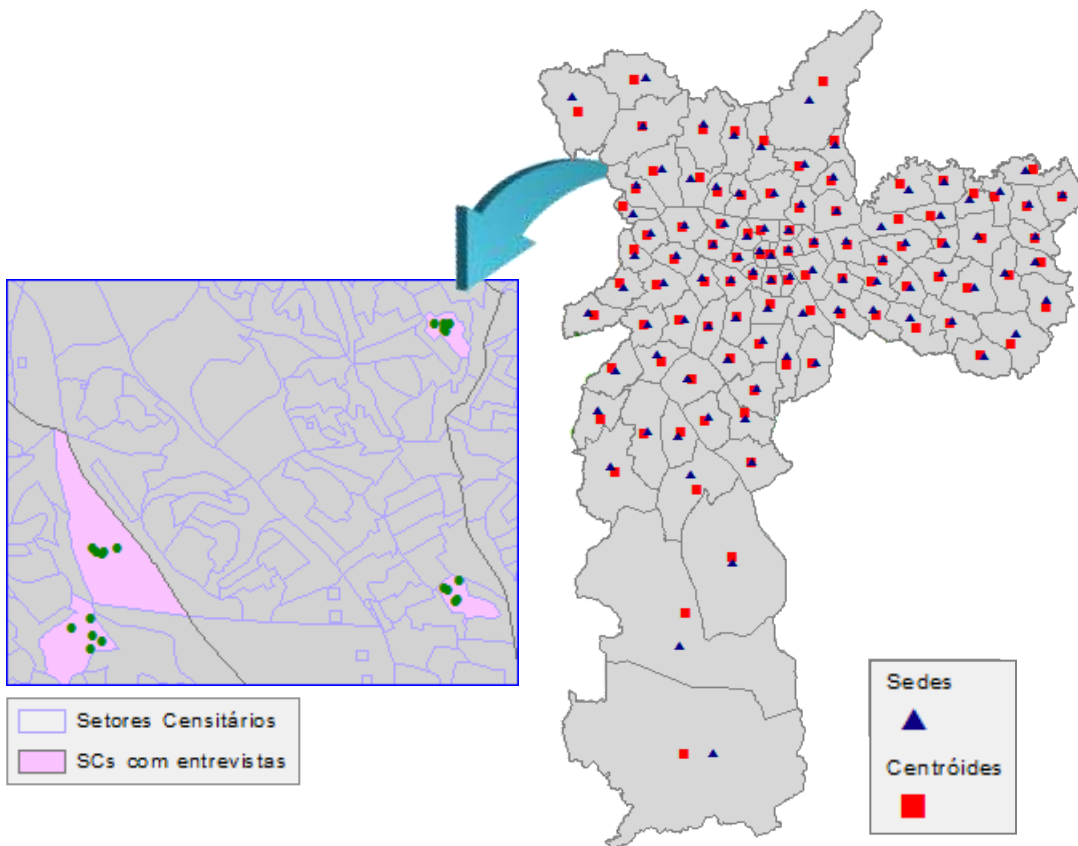
$R^2$  dos Modelos Preditivos da Renda a partir do Valor da Conta de Luz (WLM, GWR, SAR)

Dispersão do  $R^2$  da GWR das 1.000 iterações das Alternativas de Alocação de Pontos nos Polígonos





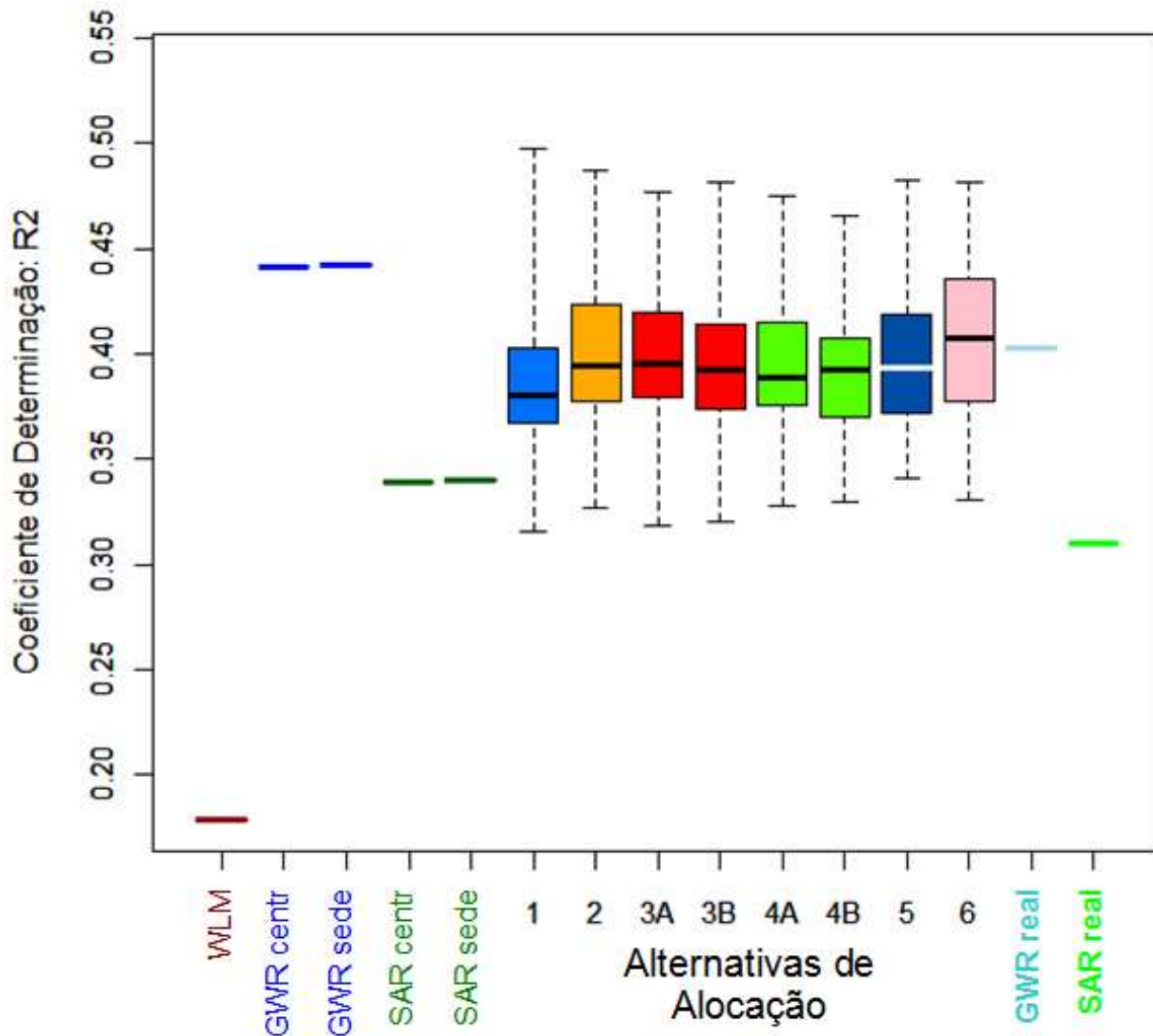
Pesq. ABRADÉE 2009  
Localização Real dos  
Domicílios Entrevistados





Pesq. ABRADÉE 2009  
**Localização Real dos  
Domicílios Entrevistados**

**Comparação dos  
Resultados**

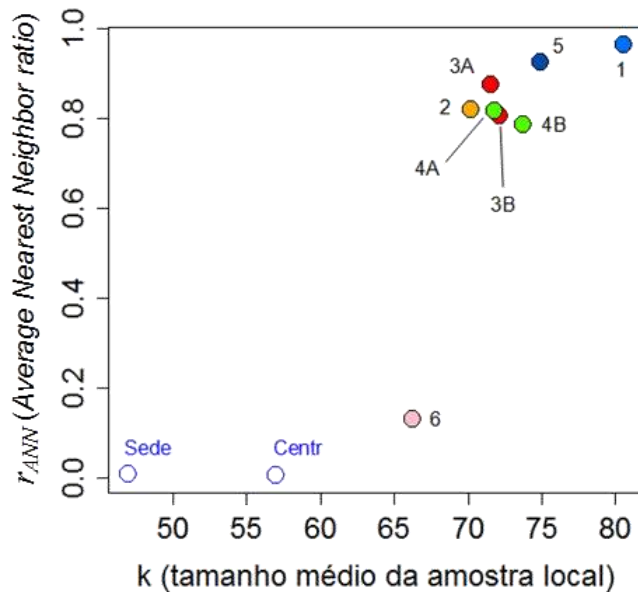




### Algoritmos de Alocação de Pontos em Polígonos

- Altíssima Associação do  $R^2$  com o tamanho  $k+1$  das amostras locais
- $r_{ANN}$ : Algoritmos 1 a 5 com Padrão quase completamente Aleatório e Algoritmos 6 com Padrão bastante agrupado
- Avaliação da Homogeneidade Espacial ( $K$  de Ripley e  $L$  transformada): Algoritmo 6 é o que mais se aproxima da Localização Real das entrevistas para a Pesquisa ABRADDEE 2009

Pesq. ABRADDEE 2009  
**Análises Espaciais Exploratórias**



	$\bar{d}_E$ (m)	$\bar{d}_O$ (m)	$r_{ANN}$	$k$
Centróide		3,8887	0,0059	57
Sede		4,6154	0,0070	47
Alt. 1		638,40	0,9643	80,565
Alt. 2		542,14	0,8189	70,172
Alt. 3A	649,90	579,54	0,8754	71,562
Alt. 3B		534,04	0,8067	72,158
Alt. 4A		540,58	0,8166	71,827
Alt. 4B		520,69	0,7865	73,474
Alt. 5		612,29	0,9249	74,938
Alt. 6		86,41	0,1305	66,266

- $n = 662$ ; 1.000 iterações para cada alternativa
- $\bar{d}_O$ ,  $r_{ANN}$  e  $k$  são valores médios para as iterações das alternativas

$$\bar{d}_O = \frac{\sum d_i^1}{n}$$

$$\bar{d}_E = \frac{0,5}{\sqrt{n/A}}$$

$$r_{ANN} = \frac{\bar{d}_O}{\bar{d}_E}$$

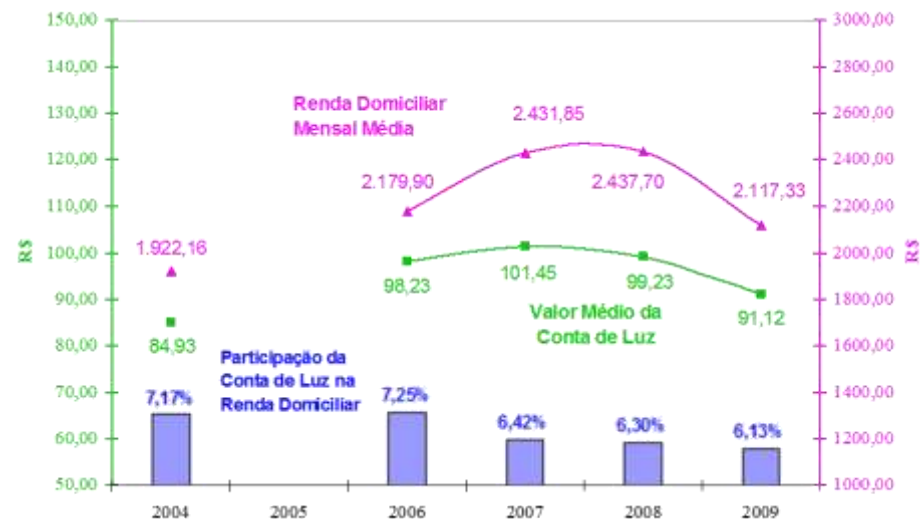
- Modelos Preditivos de Renda a partir do Consumo de Energia Elétrica ou do Valor da Conta de Luz
  - 2 Níveis: TERRITORIAL e DOMICILIAR
  - Utilizando Estatística Espacial em Modelos de Regressão



**RESULTADO DOMICILIAR:**  
O Consumo de Energia Elétrica explica  
**40 a 45%** da Renda Familiar

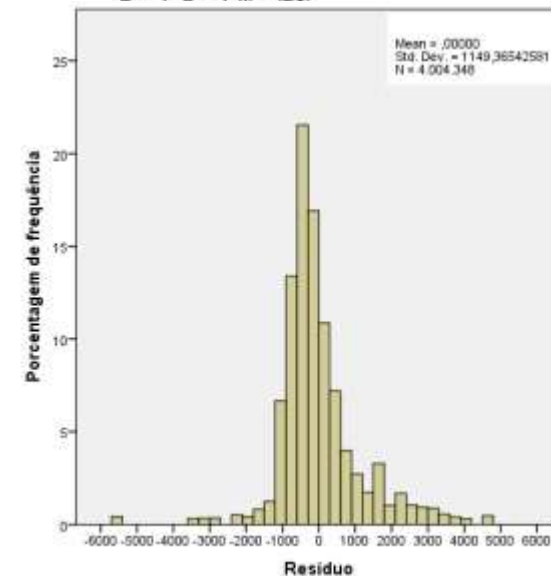
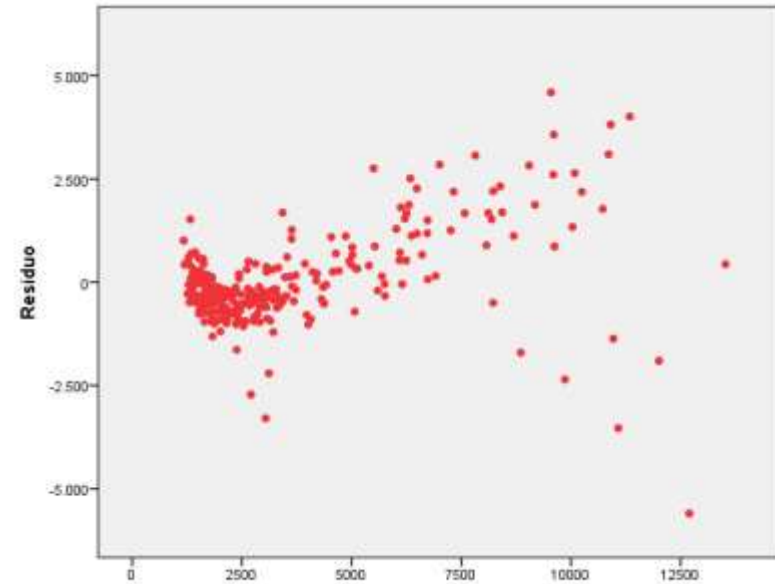
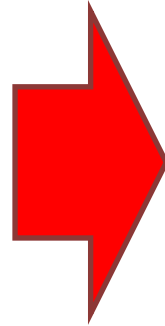
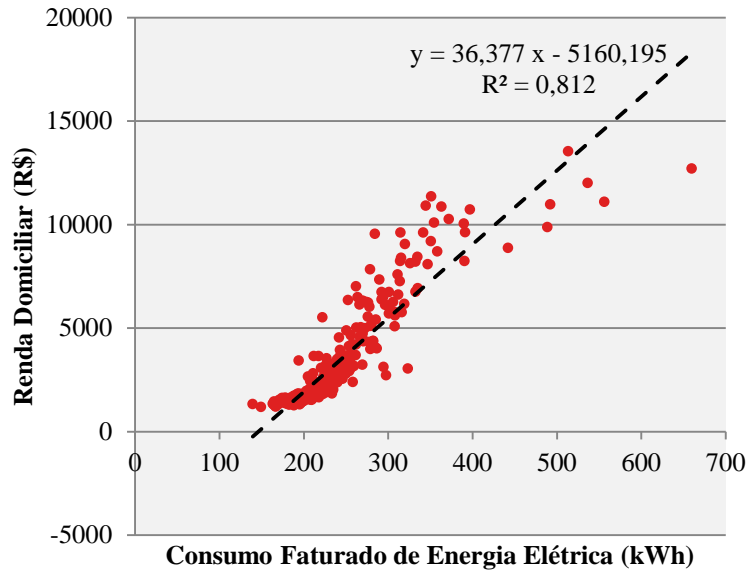
Avaliação Histórica de 2004 a 2009, utilizando Pesquisa ABRADEE

**Relação Estável** entre Renda e Consumo de Energia nos últimos anos, condizente com a tendência de evolução dessas variáveis (IBGE, IPEA)



# Relação entre Consumo de Energia Elétrica e Renda Domiciliar: Um Estudo Censitário no Município de São Paulo com Dados de 2010

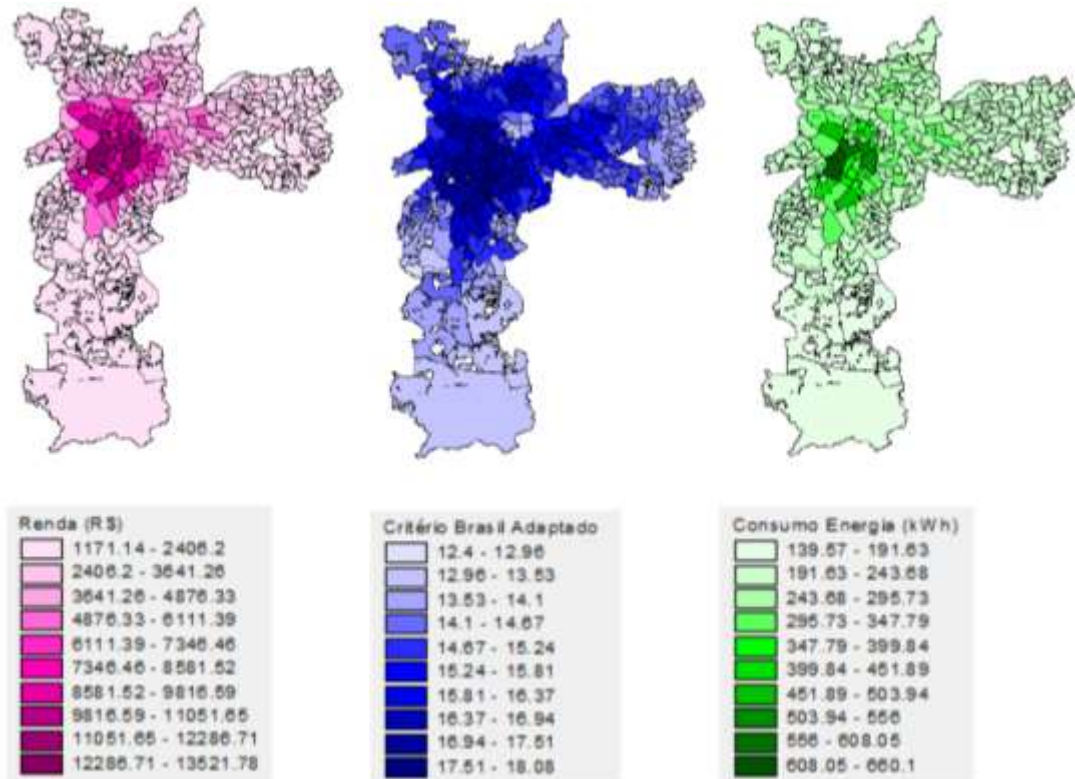
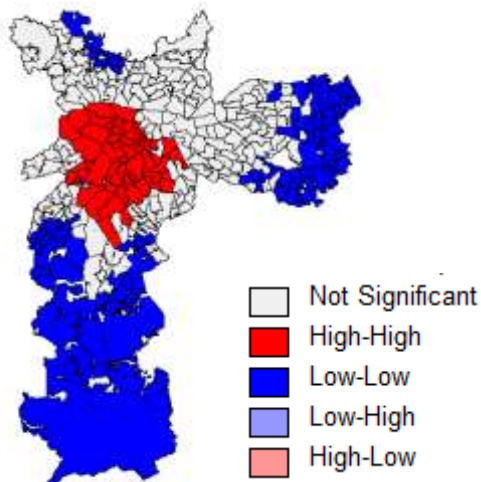
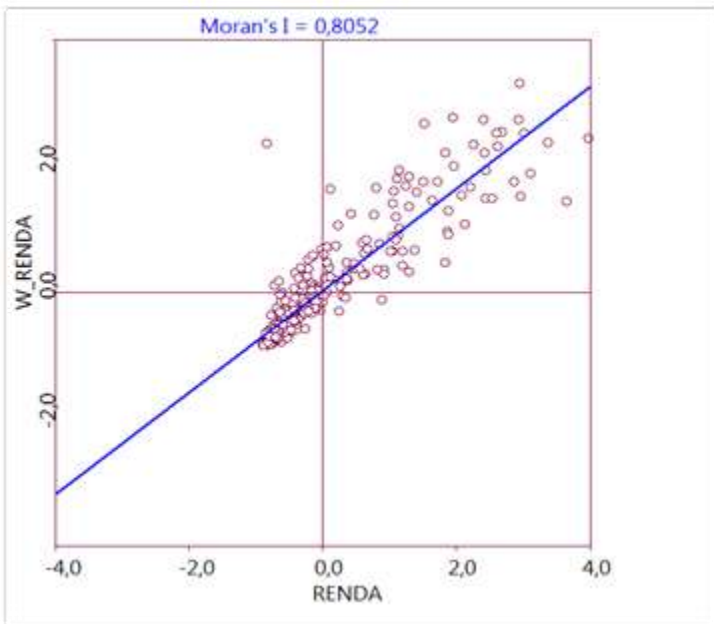
TCC – Aluno Rafaell Marcello Jabur Casella (Graduação AE, FGV-EAESP), 2012



- $R^2 = 0,812$
- Teste de significância  $F \rightarrow 0,000$
- Distribuição residual normal e homoscedástica

# Relação entre Consumo de Energia Elétrica e Renda Domiciliar: Um Estudo Censitário no Município de São Paulo com Dados de 2010

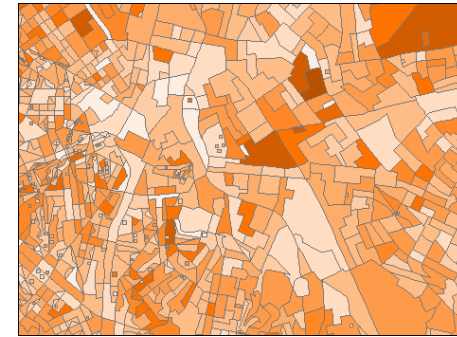
TCC – Aluno Rafaell Marcello Jabur Casella (Graduação AE, FGV-EAESP), 2012



- S.A.R.  $\rightarrow R^2 = 0,929$
- R.L.S.  $\rightarrow R^2 = 0,812$
- I de Moran = 0,8052

Indicadores de Renda  
baseados em  
Energia Elétrica

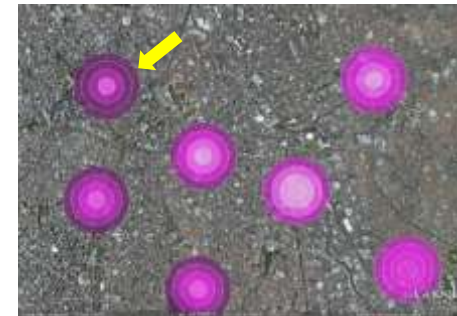
### Setores Censitários



Indicadores sistemáticos de Renda por Setor Censitário, granularidade convencional do Mercado

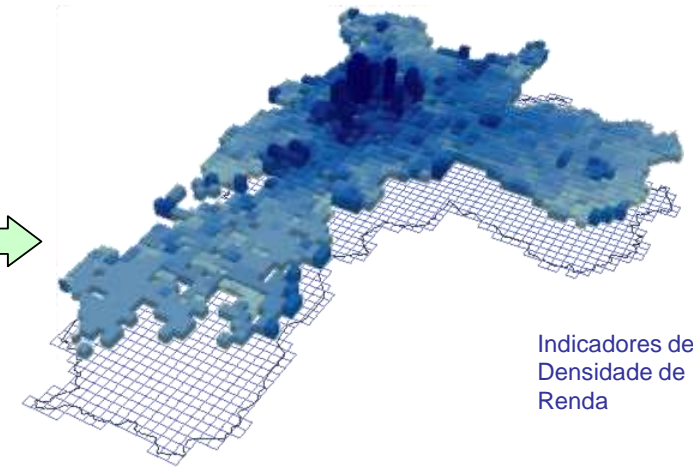
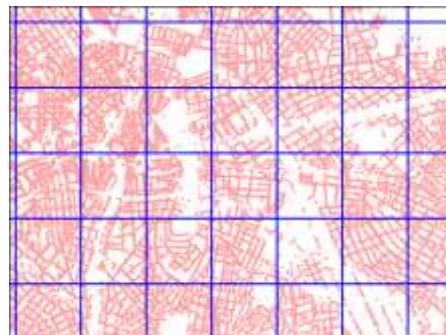
### Círculos Concêntricos

(raios progressivos de 125 m)



Estudos *Ad-hoc* (p.e., escolha do melhor ponto comercial para expansão ou reposicionamento no varejo)

### Quadrículas de 1 km<sup>2</sup>



Indicadores de Densidade de Renda

### Domicílios



Atual, abrangente, flexível,  
e atualizável mensalmente

**Um Potencial Novo Negócio  
para as Distribuidoras  
de Energia Elétrica**

- **Ferramenta Estatística R 2.11.0**

- Gratuito – download em [www.r-project.org](http://www.r-project.org)
- Completa, com extensões para Geoestatística, Estatística Espacial, Modelos Multi-nível
- Recomendação: Trabalhar junto com o editor Tinn-R
- Extensões: MAPTOOLS, SPDEP (Modelos SAR), SPGWR (Modelos GWR), SPATSTAT (Análises Exploratórias com Estatística Espacial), GeoR (Geoestatística), SPLANCS (Spatial and Space-Time Point Pattern Analysis)

- **GeoDA – Geo Data Analysis 0.9.5 e OpenGeoDa 0.9.8**

- Gratuito – download em [geodacenter.asu.edu](http://geodacenter.asu.edu)
- Auto-correlação Espacial e Modelos SAR

- **GWR3X**

- Modelos GWR
- Maiores informações em [ncg.nuim.ie/ncg/GWR/software.htm](http://ncg.nuim.ie/ncg/GWR/software.htm)

- **ArcView GIS 3.2**

- **ArcGIS 9.3**

- Spatial Analyst
- Geostatistical Analyst
- Spatial Statistical Toolbox – Estatística Espacial Exploratória e Modelo GWR
  - Incorporação da Topologia da Malha Viária nos Modelos de Vizinhança

- **Validação da Investigação Territorial com base em dados mais atuais**
  - Censo Demográfico 2010 – Renda Familiar por Setor Censitário (avaliação da influência do MAUP) e itens relacionados à Energia Elétrica
  - Avaliação do efeito Racionamento (ACHÃO, 2003) e avanço do consumo residencial de Gás Natural e GLP em São Paulo (SÃO PAULO, 2005)
    - Estabilidade da Relação Consumo de Energia e Renda, conforme rodadas da Pesquisa ABRADEE, sugerem manutenção da relação
- **Replicação da Investigação Territorial em outras regiões**
  - Extensão para outros municípios da AES Eletropaulo e outras áreas urbanas brasileiras
  - Avaliação da influência do perfil urbano regional
  - Comparações Internacionais, outras unidades de análises
- **Replicação da Investigação Domiciliar e dos Algoritmos de Espalhamento de Pontos em Polígonos**
  - Utilização dos dados nacionais da Pesquisa ABRADEE e dados da América Latina (Pesquisa CIER) para Avaliação Comparativa
  - Pesquisa ABRADEE 2010: publicação do setor censitário do entrevistado
  - Outras pesquisas de interesse acadêmico aplicado

- **Investigação da Estrutura Familiar e outros construtos na Relação entre Renda e Consumo de Energia Elétrica**
- **Investigação com outros pressupostos e modelos estatísticos**
  - Integração das Investigações Territorial e Domiciliar: Modelos multi-nível
    - Promoção a Modelos multi-nível geográficos
  - Estatística Bayesiana no contexto espacial: estimação em pequenas áreas
- **Potencial de Aprofundamento em Projetos de P&D**
  - Estudos Comparativos para Distribuidoras de Energia
  - Aprofundamento dos Modelos de Cálculo para efeito de aprimoramento das modalidades tarifárias de gestão da distribuição de energia
  - Subsídio para Argumentação a respeito de Tarifa Social perante ANEEL
- **Prova de Conceito para Programas de Microcrédito**
  - Diminuição da Assimetria de Informação
  - Substituição da Renda Domiciliar declarada pelo Consumo Médio de Energia Elétrica individual combinado com o da região de moradia do tomador de crédito

## MICROCRÉDITO PRODUTIVO ORIENTADO – PÚBLICO ALVO

### DEFINIÇÃO

*Pessoas físicas ou jurídicas, formais ou informais, excluídos do acesso ao crédito pelo sistema financeiro tradicional, ou que por razões de sua fragilidade socioeconômica têm dificuldades de acessar as linhas de financiamentos geralmente existentes*

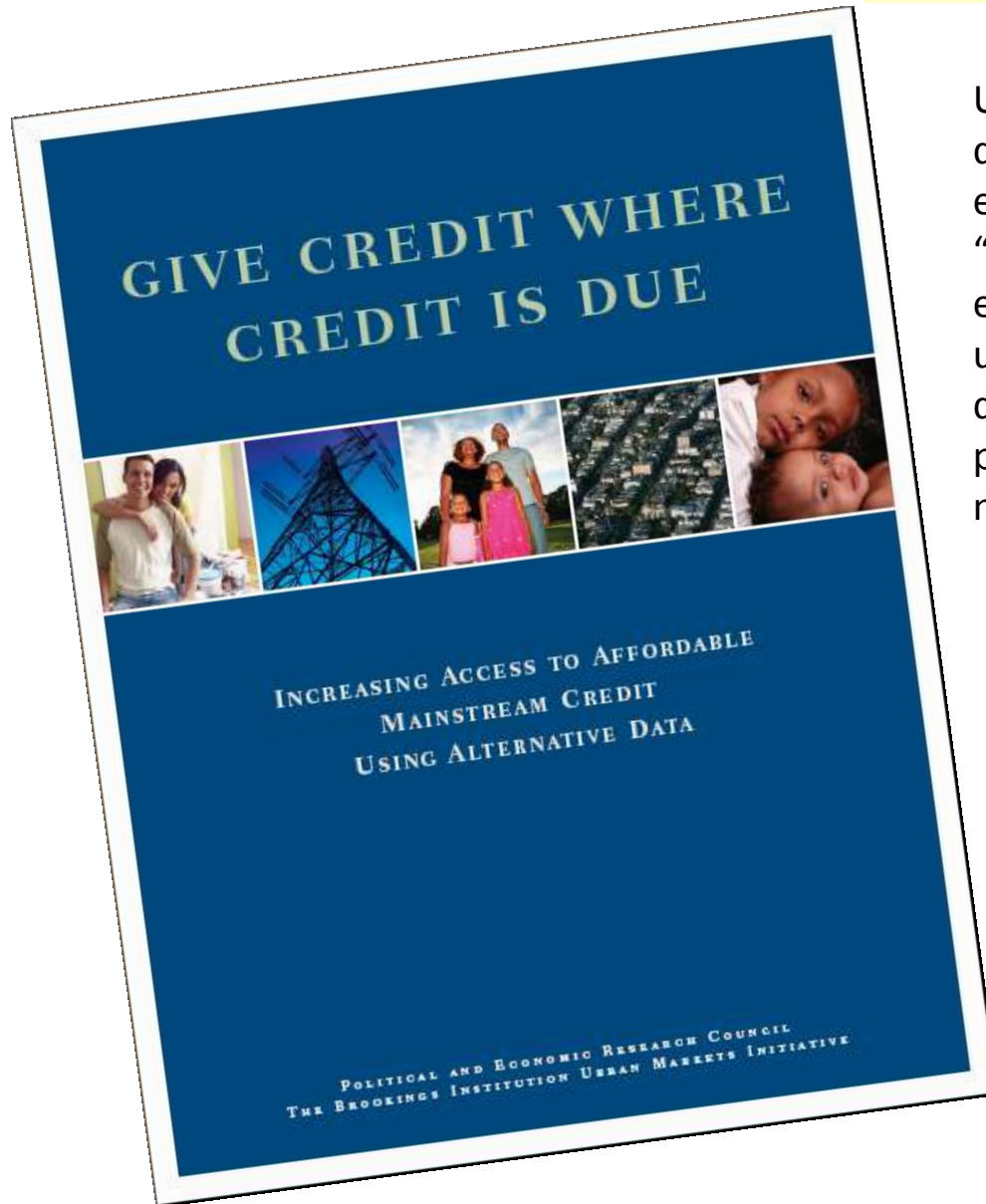
Na grande maioria,  
**INFORMAIS**

- *Costureiras, Artesões, Comércio e produção de docinhos e salgadinhos, Marcenarias, Sapatarias, Carpintarias, Alfaiatarias, Pequenas Gráficas, Padarias, Ambulantes, Vendedores em geral, Mercadinhos, Papelarias, Armarinhos, Bazares, Farmácias, Armazéns, Pequenos Restaurantes, Lanchonetes, Feirantes, Açougues, Pequenos Lojistas, Vendedores de cosméticos, Salões de beleza, Oficinas mecânicas, Borracharias, etc*

## MOTIVAÇÃO

“O fornecimento de crédito, de uma forma geral, é uma atividade sujeita a dois problemas de difícil solução: a **assimetria de informação** e os **custos de transação**. Esses problemas são ainda maiores no caso do microcrédito”

“As microfinanças se dirigem para esses microempreendedores com poucos recursos próprios, onde a operação de crédito é baseada em **mecanismos não convencionais** que buscam contornar os problemas de **assimetria de informação e custos de transação**”



Utilizando uma base de dados de 8 milhões de potenciais solicitantes de crédito em diversos estados norte-americanos, mostra que dados “alternativos” ou “não tradicionais” (energia elétrica, aluguel, gás, seguros) podem ser utilizados no processo de concessão de crédito, diminuindo a assimetria de informação e potencializando maior assertividade nos modelos de *credit scoring*.

This study offers a feasible market solution to bring those outside the mainstream credit fold within it. Mainstream lenders can use “alternative” or “nontraditional” data, including payment obligations such as rent, gas, electric, insurance, and other recurring obligations, to evaluate the risk profile of a potential borrower.<sup>1</sup> Our findings indicate that alternative data, if widely incorporated into credit reporting, can bridge the information gap on financial risk for millions of Americans. More concretely, considering that many of these millions outside the credit mainstream are poorer, less advantaged Americans, the information can direct markets toward a faster alleviation of poverty in this country.

## PROJETO: *MICROCREDIT SCORE* - Proposta:

- Construir e manter um *score* de Microcrédito regional (setores censitários, quadrículas, unidades *ad hoc*)
- Utilizar o *score* de Microcrédito de forma individualizada, caso o solicitante de crédito autorize, a partir do acesso à conta de luz
- Incorporar Informações da Conta de Luz nos Modelos de Crédito



## Etapas:

*Análise conjunta das informações de energia elétrica (consumo e perfil de pagamento da conta de luz) e informações externas (birôs de informação)*

*Proposição e Aprimoramento de Modelos de Crédito híbridos (dados convencionais + dados alternativos + dados coletados em campo)*

*Futuro: Prova de Conceito (testar o score em campo)*


- **Orientação: diminuir assimetria de informação e custo de transação; buscar ganhos de escala nas operações de microcrédito**

## PROJETO: *MICROCREDIT SCORE* - Hipóteses:

- O acesso ao dado de consumo de energia elétrica, histórico de pagamentos e comportamento fraudulento no uso de energia permitem diminuir a assimetria de informações para a concessão do crédito.
- Este fato beneficiaria principalmente :
  - Clientes com pouco histórico de crédito em *bureaux*, não bancarizados e alguns segmentos demográficos específicos (jovens, informais, baixa renda, mulheres)
  - IMFs (Instituições de Micro Finança) – ganho de escala e diminuição de custo em todo o processo do ciclo de crédito
- Modelos alternativos, considerando o Consumo de Energia e a Estatística Espacial, podem trazer muitos benefícios ao negócio através do aprimoramento do conhecimento do cliente/solicitante de crédito
- Outro Uso Potencial: Avaliação do EFEITO dos Programas de Microcrédito na Renda dos solicitantes e/ou na Renda agregada por áreas censitárias

PROJETO: *MICROCREDIT SCORE* - Modelos Criados:

- Criação de **3 Tipos de Modelos**:

- Titulares da conta de energia elétrica, com históricos de crédito no mercado;
- Não-titulares da conta de energia elétrica, com históricos de crédito no mercado; e
-  Pessoas sem histórico de crédito no mercado (tipicamente, donas-de-casa cujas operações bancárias eram feitas através do marido, e jovens recém chegados ao mercado de trabalho).

- Os modelos (i) e (ii) foram criados utilizando dados combinados da AES Eletropaulo e da Boa Vista Serviços (BVS), grande birô de informação do grupo Equifax.
- O modelo (iii) foi feito com apenas os dados da AES Eletropaulo

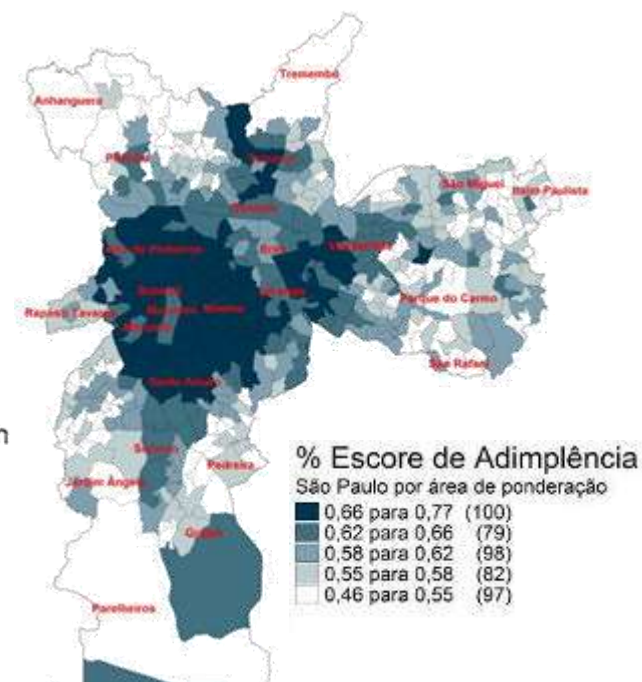
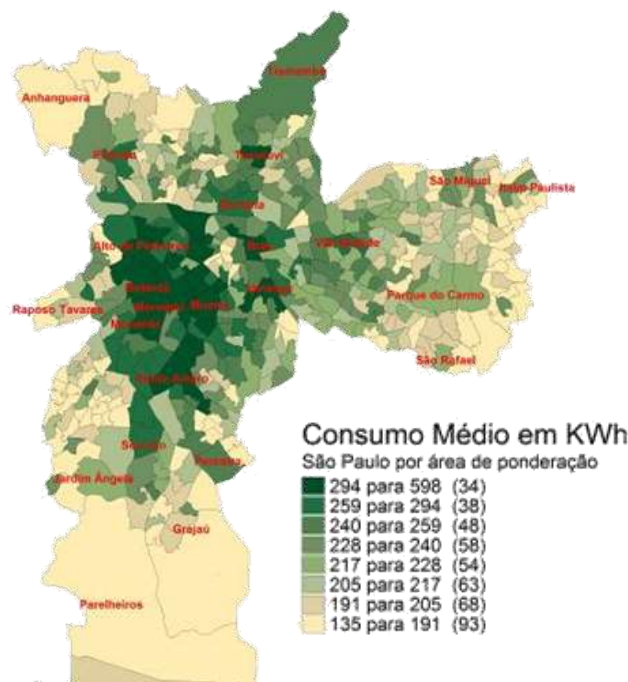
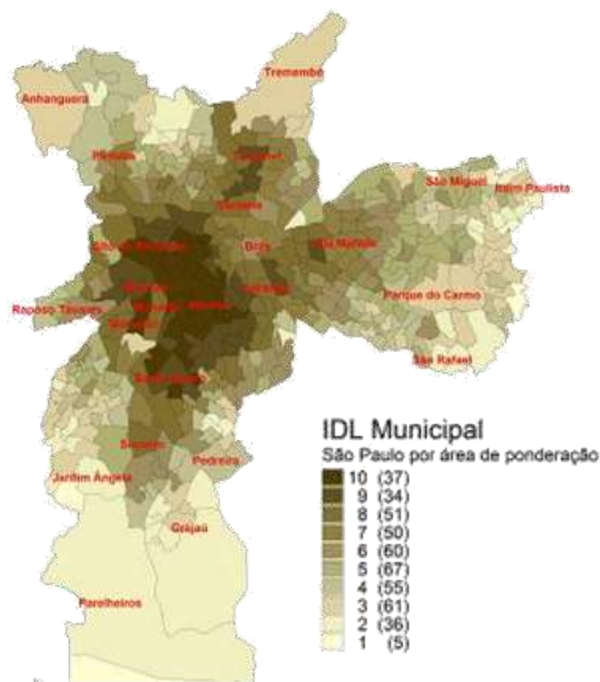
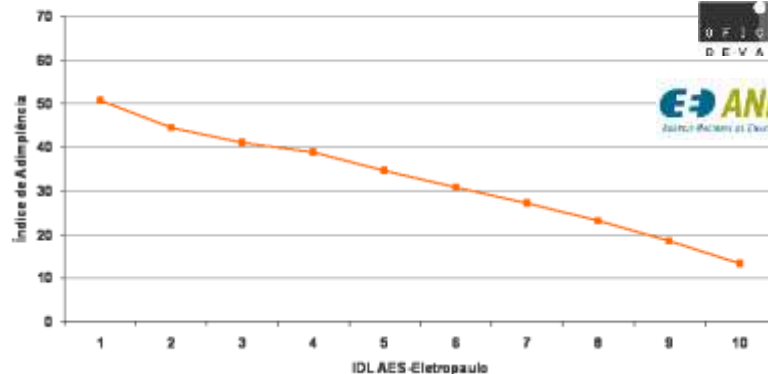


**PROJETO: MICROCREDIT SCORE - Highlights**

- Comportamento do Pagamento da Conta de Luz é aparentemente compatível com a capacidade de pagamento típica do mercado
- Mapas muito similares entre IDL, Consumo de Energia e Escore de Adimplência



Índice de Adimplência X IDL AES-Eletropaulo  
(Perfil de adimplência com relação a conta de luz)



## PROJETO: *MICROCREDIT SCORE* - Resultados:

- **Modelos (i) e (ii):**
  - Melhoria significativa dos modelos tradicionais de mercado da BVS (altos índices de desempenho [KS] e melhorias de 4 a 9 pontos percentuais)
- **Modelo (iii):**
  - Altamente eficaz na inovação – endereça donas-de-casa e jovens, tipicamente sem histórico de crédito.
- Modelos testados com dados reais em uma operação de microcrédito (Santander Microcrédito) confirmaram as expectativas.
- Modelos regionais (indicadores para pequenas regiões, setores censitários, distritos, etc) também foram criados.

Indicadores de Renda e Crédito baseados em Energia Elétrica

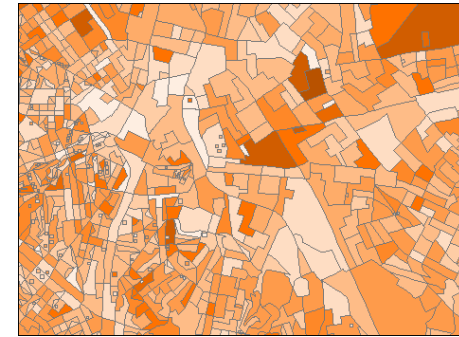
### Conjunto de Indicadores

- Renda Presumida
- Classe Econômica
- Taxa de crescimento da renda
- Total de Domicílios
- Taxa de crescimento domiciliar
- Total de Empreendimentos Comerciais
- Taxa de crescimento comercial
- Total de Indústrias
- Taxa de Crescimento Industrial
- *Total de inadimplência*
- *Taxa de crescimento da inadimplência*
- ...

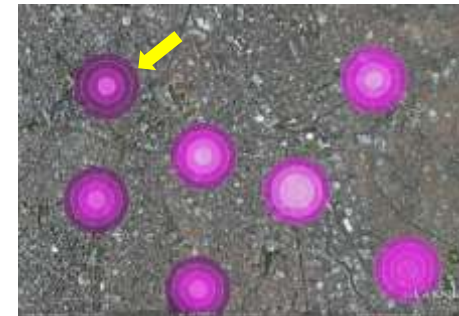
Fonte de Receita para as Distribuidoras de Energia Elétrica



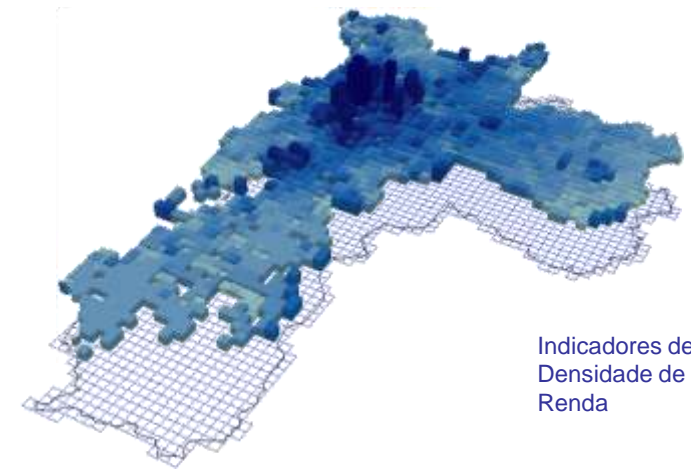
Ferramenta para Relacionamento Estratégico com os Clientes (Cli Corp, Poder Público, etc)



Indicadores sistemáticos de Renda por Setor Censitário, granularidade convencional do Mercado



Estudos *Ad-hoc* (p.e., escolha do melhor ponto comercial para expansão ou reposicionamento no varejo)



Indicadores de Densidade de Renda



FUNDAÇÃO  
GETULIO VARGAS



Escola de Administração  
de Empresas de São Paulo

# Obrigado !!!!

**Eduardo de Rezende Francisco**  
[eduardo.francisco@fgv.br](mailto:eduardo.francisco@fgv.br)

**FGV-EAESP**



09 de Maio de 2013

