

# Os dados dos censos demográficos como base para análise de problemas complexos: uma avaliação dos países Latino-Americanos

José Diego Gobbo Alves  
jdgobboalves@gmail.com  
Faculdade de Ciências Aplicadas (FCA/UNICAMP)

## Introdução

Em um contexto de Big Microdata (RUGGLES, 2014) com dados cada vez mais detalhados e sistematizados, pensar em como estes estão sendo coletados e divulgados pelos países torna-se fundamental. As pequenas amostras sobre população não são mais suficientes para uma leitura sobre o comportamento da dinâmica da população ao longo dos anos e, muito menos, para apreender como a população está distribuída no espaço.

O objetivo desse texto é destacar a importância dos dados dos Censos Demográficos para a composição de análises de problemas da realidade cada vez mais complexos e multiescalares. Para isso, questiona as limitações encontradas a partir das diferentes formas de agregação desses dados nos países latino-americanos e como elas se expressam quando aplicado técnica de estatística espacial para análise dos padrões de distribuição da população. Nesse sentido, a hipótese trabalhada nesse texto é que as diferentes unidades geográficas afetam os resultados quanto ao padrão de distribuição da população, bem como outras análises que utilizam os dados censitários.

Ao trabalhar com um grande volume de dados, surgem diversos problemas relacionados com as unidades geográficas de agregação. O conhecido Problema de Unidade de Área Modificável (Modifiable Areal Unit Problem) ou MAUP, foi amplamente discutido por Openshaw (1984) em que “[..] regarded the MAUP as a fundamental geographical problem inherent in all studies of spatially aggregated data because the results of such studies are Always affected by the areal units used.” (JELINKS e WU, 1996, p.138). Isto é, acaba por alterar todo o fenômeno estudado através de uma suavização deste pela unidade geográfica utilizada, podendo gerar distorção nas análises e uma homogeneização na distribuição dos dados, além de interpretações equivocadas dos fenômenos.

O MAUP gera dois principais efeitos: 1) o efeito escala ou de agregação, em que se tem uma perda na resolução espacial dos dados na medida em que estes passam a ser agrupados em unidades maiores (BUENO e D’ANTONA, 2014) e 2) a falácia ecológica, que parte do pressuposto de que os resultados dispostos em unidades geográficas correspondem aos resultados obtidos em nível individual (DIAS et al, 2002).

## Metodologia

A metodologia desse trabalho consistiu no levantamento bibliográfico, documental e cartográfico sobre os Censos Demográficos dos países latino-americanos, especificamente, Argentina, Brasil, México e Uruguai. A escolha desses países se deu pelos fatores: disponibilidade dos dados do Censo em arquivos digitais e suas diferenças quanto à agregação dos dados. Buscou-se o menor nível no qual esses países agregavam

seus dados, ou seja, a maior resolução espacial de cada país. As cidades utilizadas para a análise foram: Aguascalientes (México), Buenos Aires (Argentina), Durazno (Uruguai) e Piracicaba (Brasil).

Nesses países foram aplicadas duas técnicas de estatística espacial presentes no *software* ArcGis, versão 10.5, sendo elas: *Standard Distance* e *Directional Distribution (Standard Deviational Ellipse)*. Elas oferecem, respectivamente, medidas da concentração e desconcentração das feições em torno de um centro médio e da orientação da distribuição das feições utilizadas. Nesse texto, as feições utilizadas foram os setores censitários, as quadras (quarteirões) e as células da grade estatística.

A Grade Estatística, lançada em 2016 pelo Instituto Brasileiro de Geografia e Estatística (IBGE), consiste em um conjunto de células regulares dispostas em formas de grade que agregam dados sobre residentes, domicílios e sexo da população do Censo Demográfico de 2010 (BUENO,2016; IBGE, 2016).

A resolução espacial consiste no maior ou menor detalhamento da representação espacial, isto é, da realidade com base no tamanho das unidades geográficas. É semelhante com o que ocorre com a escala cartográfica, ou seja, é inversamente proporcional: quanto menor a unidades geográfica, maior será o nível de detalhamento. Para calcular a resolução espacial média, usa-se a extração da raiz quadrada do quociente entre a área de uma região e o número de unidades que compõem esta região (TOBLER et al, 1997).

Nesse sentido, comparando duas unidades geográficas de um mesmo município como os setores censitários e a grade estatística por exemplo, a grade oferece uma maior resolução espacial do que os setores censitários, haja vista que as unidades geográficas nas quais os dados estão agregados são menores, salvo exceção nas áreas centrais muito adensadas, onde os limites dos setores e da grade são semelhantes.

### Resultados preliminares

Os países pesquisados utilizam diferentes metodologias e bases operacionais para a agregação e disseminação dos dados dos Censos Demográficos. A tabela 1 apresenta a menor unidade geográfica de agregação de dados encontrada para cada país, o ano do último censo disponível e o órgão aplicador.

Tabela 1: Nível de agregação dos dados dos países latino-americanos.

<b>País</b>	<b>Unidade geográfica</b>	<b>Último Censo</b>	<b>Aplicador</b>
Argentina	Radio Censal	2010	Instituto Nacional de Estadística y Censos
Brasil	Grade Estatística	2010	Instituto Nacional de Geografia e Estatística
México	Manzana Urbana	2010	Instituto Nacional de Estadística y Geografía
Uruguai	Zona Censal (manzana urbana)	2011	Instituto Nacional de Estadística

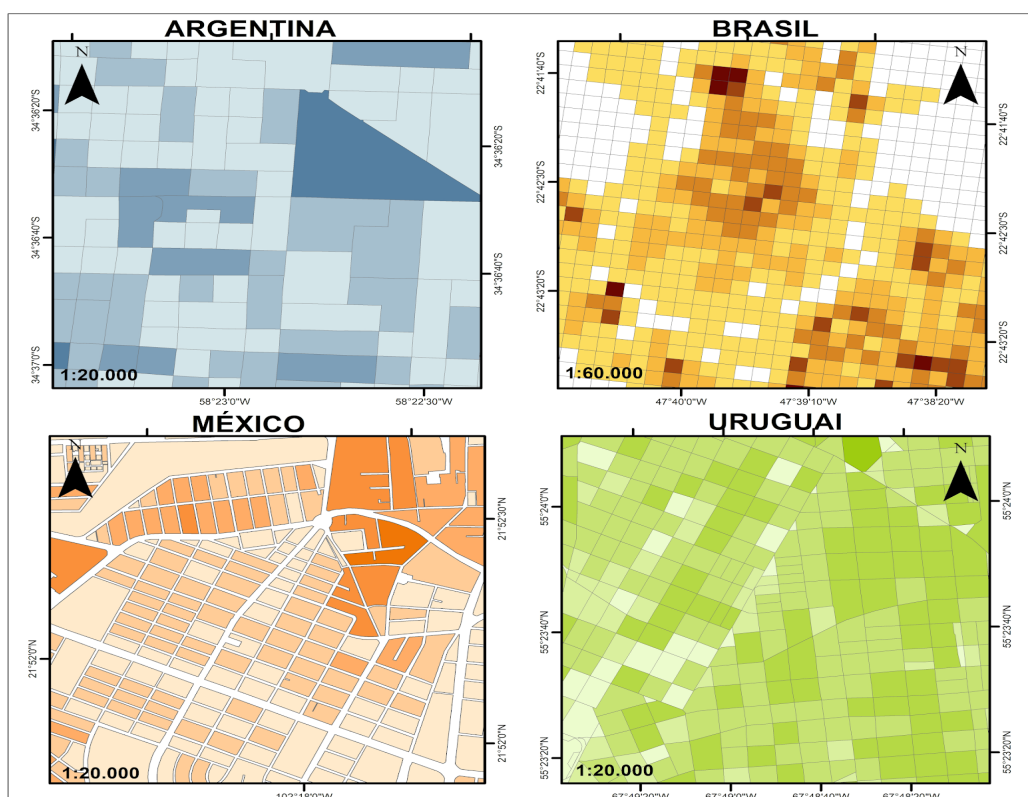
**Fonte:** IBGE, 2016. INDEC, 2015; INEGI, 2015; INE, 2015; INE, 2011.

Pensando em como essas bases de dados podem auxiliar o planejamento urbano para cidades mais sustentáveis, observa-se que a agregação dos dados da Argentina, Chile e da Bolívia são as piores para avaliar a forma urbana e a distribuição da população, já que é impossível identificar uma localização mais precisa da população, tendo os efeitos de MAUP mais presentes nessas unidades geográficas.

As Manzanas Urbanas do México e as Zonas Censais do Uruguai, por mais que estejam mais desagregadas em unidades geográficas menores, comparado com a Argentina, não podem oferecer uma garantia espaço-temporal da espacialização dos dados, já que seus limites podem ser alterados nas constantes transformações espaciais, impossibilitando a comparação a longo prazo. Além disso, não há uma padronização no tamanho das quadras.

Em todos os países analisados, com exceção da Grade Estatística brasileira, as áreas rurais são as que mais sofrem com a falta de dados precisos geograficamente. Por terem uma área de agregação de dados bem maior do que na área urbana, há uma significativa distorção da distribuição da população nessas áreas. Do modo como está posto nos censos, apresenta uma homogeneização da população em toda a área de agregação, o que de fato não representa a realidade. Mesmo que a população não esteja concentrada em uma área do setor, ela não está distribuída de forma homogênea por todo ele como os dados apontam.

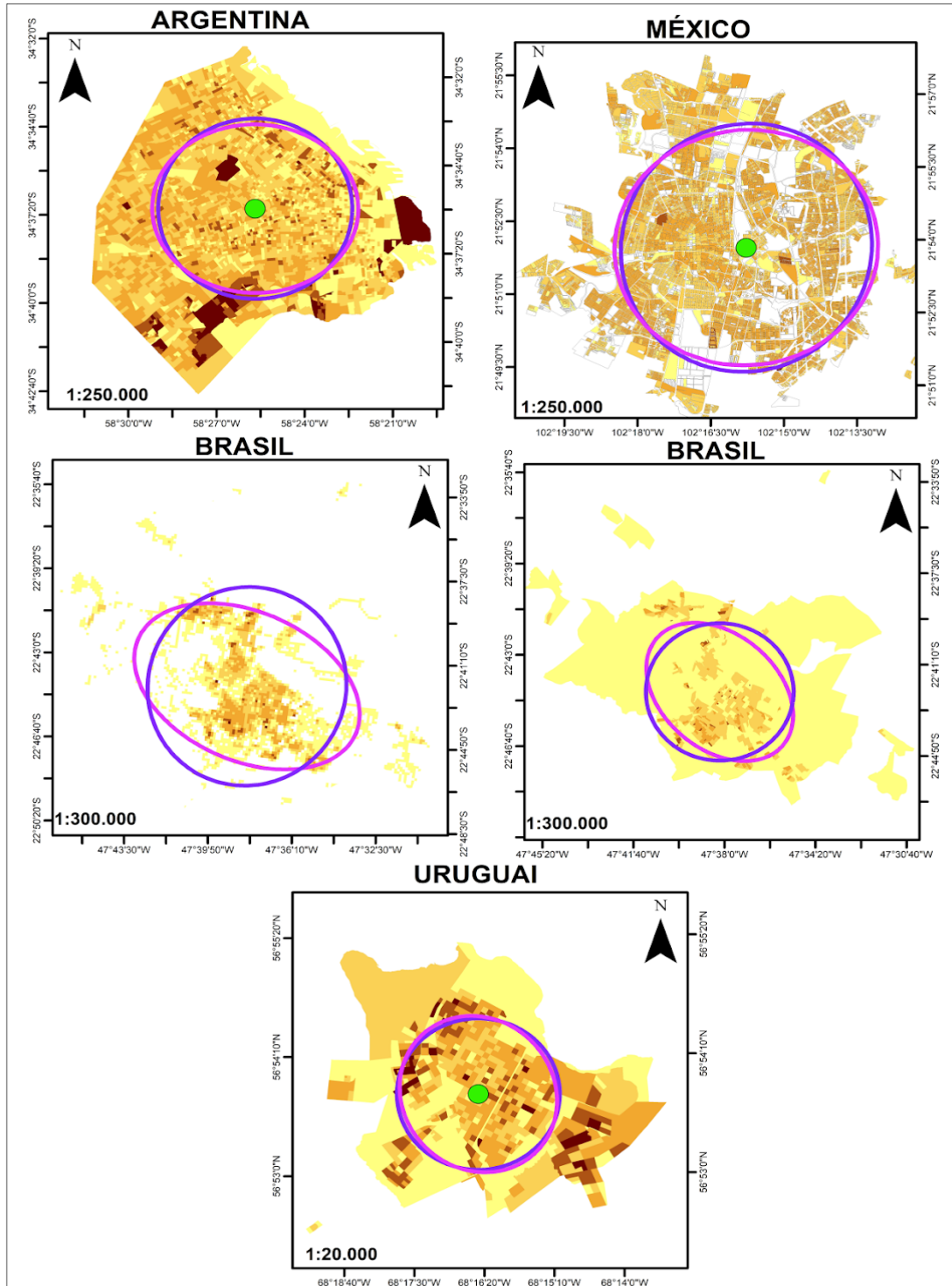
Figura 1. Apresentação a menor unidade geográfica utilizada por cada país para a agregação dos dados do Censo.



**Figura 1:** Menor unidade geográfica encontrada para o Censo de cada país. **Fonte:** IBGE, 2016. INDEC, 2015; INEGI, 2015; INE, 2015; INE, 2011;

A aplicação das técnicas de estatística espacial apontou para reafirmar a importância da resolução espacial nas análises da distribuição espacial da população.

Figura 2 apresenta a aplicação das mesmas nos quatro países selecionados.

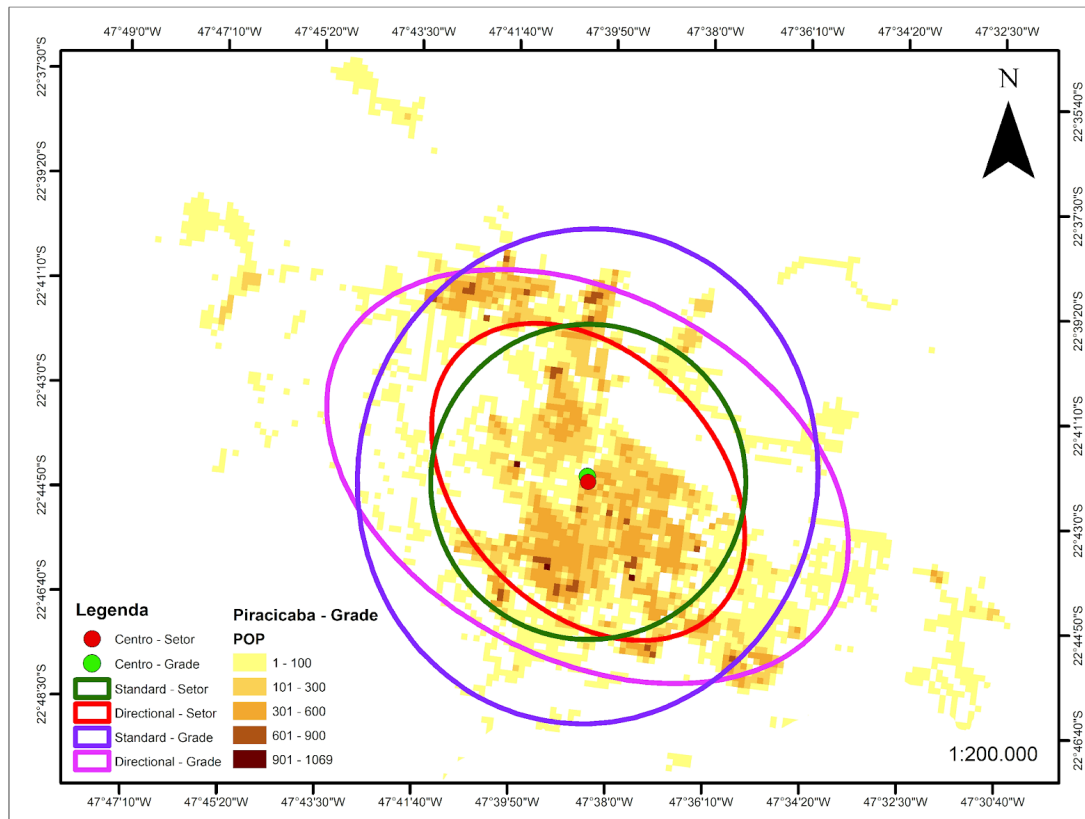


**Figura 2:** Resultado da aplicação das técnicas de estatística espacial *Standard Distance* e *Directional Distribution (Standard Deviation Ellipse)*. **Fonte:** IBGE, 2016. INDEC, 2015; INEGI, 2015; INE, 2015; INE, 2011; ERSI, s.d.

Partindo da hipótese desse texto de que as unidades geográficas de agregação dos dados interferem nos resultados acerca da distribuição espacial, a Figura 02, apresenta os resultados das técnicas de estatística espacial (*Standard Distance* e *Directional Distribution*). A distribuição da população em setores censitários, *radio* e *zona censal* e *manzana urbana* tarem problemas para analisar a distribuição da população. Dois principais problemas encontrados são: 1) a não regularidade na dimensão das unidades geográficas de análise, com a variação devido a critérios de quantidade de população e os contornos das pistas de rolamento (ruas, avenidas, estradas, etc.) e 2) a suposta continuidade e homogeneização da população.

Esses dois fatores influenciaram nos resultados, indicando que a distribuição espacial da população nas cidades analisadas (com exceção do uso da grade estatística) é mais compacta e contínua do que realmente é. Para justificar tal narrativa, foi feita uma comparação da distribuição da população em grade estatística e em setores censitários, utilizando o Brasil como exemplo. A escolha se deu pela disponibilidade dos dados agregados em dois níveis: os setores censitários e na grade estatística, o que não era possível de ser feito para os outros países.

A figura a seguir apresenta tal comparação.



**Figura 3:** Comparação do resultado da aplicação de técnicas de estatística espacial na grade estatística e em setores censitários em Piracicaba, cidade brasileira. Fonte: IBGE, 2012, 2016; ERSI, s.d.

Nota-se na figura que a aplicação das técnicas de concentração e direção da população diferenciam-se de acordo com a unidade geográfica utilizada. Isso ocorre devido

a dois principais fatores: a regularidade da dimensão das células em detrimento à diversidade de tamanhos do setor censitário e, conseqüentemente, maior precisão da ocupação urbana nas áreas mais distantes do centro da cidade onde a área dos setores censitários são maiores.

Como bem sabido, quanto mais distante da área central e quanto menor a quantidade de população, há uma tendência de que os setores censitários possuam uma dimensão maior, vide a diferença entre os setores rurais e os setores urbanos, por exemplo. A agregação e disseminação dos dados nesses setores sofrem muito com o MAUP e a falácia ecológica, principalmente com o primeiro. Devido sua extensão, a distribuição espacial da população é apresentada de forma homogênea ao longo de todo setor, perdendo as nuances e as descontinuidades internas nessas áreas.

Comparando a distribuição em setores censitários e na grade estatística e, conseqüentemente, a aplicação das técnicas espaciais (Figura 3), observa-se que os setores apresentam resultados que indicam que a população é mais concentrada do que observado na realidade. Para análises populacionais e espaciais sobre a forma urbana ou a distribuição espacial da população, por exemplo, a grade oferece um melhor resultado, se comparada com os setores.

### **Considerações finais**

A análise desse trabalho que está restrita a quatro países pois alguns deles apresentam características semelhantes quanto a forma de agregação dos dados, caso da Bolívia que usa áreas operacionais semelhantes aos setores censitários brasileiros ou argentinos. Há países que não possuem os dados divulgados em bases cartográficas digitais, sequer agregados em unidades geográficas. Além disso, o levantamento desses dados nas plataformas dos institutos nacionais é dificultoso, ponto problemático para pesquisadores de outros países que não estão habitados com sítio.

A dificuldade ao acesso dos dados dos Censos de outros países na plataforma é um indicador relevante como barreira para estudos comparativos entre países, isso porquê, as plataformas de acesso aos dados são confusas para pesquisadores que não possuem contato frequente com as mesmas. Em um período que estimula estudos comparativos entre os países, esse é um entrave significativo para a comparação dos fenômenos.

A complexidade dos problemas contemporâneos, bem como sua simultaneidade em diversas escalas geográficas e cartográficas, exige uma quantidade de dados maiores e agregações de dados com maior resolução espacial. Como as regras de sigilo estatístico impedem que os dados sejam desagregados a nível individual, de todos os países analisados nesse trabalho, a grade estatística brasileira se mostrou como a que melhor possui uma resolução espacial dos dados e respondendo melhor à aplicação de técnicas de estatística espacial.

### Refrências bibliográficas

BUENO, M. C. D. **Grade estatística: uma abordagem para ampliar o potencial analítico de dados censitários**. Tese (Doutorado em Demografia) - Universidade Estadual de Campinas, 2014.

BUENO, M. C. D.; D'ANTONA, A. O. Avaliação de métodos de desagregação para geração de grades de população. **Revista Espinhaço**, v. 3, p. 127-137, 2014.

DIAS, T. L.; et al. PROBLEMAS DE ESCALA E A RELAÇÃO ÁREA-INDIVÍDUO EM ANÁLISE ESPACIAL DE DADOS CENSITÁRIOS. IP (Belo Horizonte), **Belo Horizonte**, v. 01, n.04, p. 89-104, 2002.

ESRI. ARCMAP. **Directional Distribution (Standard Deviation Ellipse)**. Disponível em <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/directional-distribution.htm> S.d (b)

ESRI. ARCMAP. **Standard Distance**. Disponível em <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/standard-distance.htm> S.d

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE (BRASIL). **Grade Estatística**. Rio de Janeiro, 2016.

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS – INDEC (ARGENTINA). **CENSO 2010**. Argentina, 2015. Disponível em:

[https://www.indec.gov.ar/nivel4\\_default.asp?id\\_tema1=2&id\\_tema\\_2=41&id\\_tema\\_3=135](https://www.indec.gov.ar/nivel4_default.asp?id_tema1=2&id_tema_2=41&id_tema_3=135)

INSTITUTO NACIONAL DE ESTADÍSTICA INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA – INEGI (MÉXICO). **Censo 2010**. Disponível em: <http://www.inegi.org.mx>

INSTITUTO NACIONAL DE ESTADÍSTICA - INE (URUGUAI). **Censos 2011**. Disponível em: <http://www.ine.gub.uy/censos-2011>

JELINKS, D. E.; WU, J. The modifiable areal unit problem and implications for landscape ecology. **Landscape Ecology** vol. 11no. 3 pp 129-140. 1996.

OJIMA, R. **Análise comparativa da dispersão urbana nas aglomerações urbanas brasileiras: elementos teóricos e metodológicos para o planejamento urbano e ambiental**. Tese. 2007. UNICAMP –SP. 166p

RUGGLES, S. Big microdata for population research. **Demography**. 2014 Feb;51(1):287-97. doi: 10.1007/s13524-013-0240-2.

TOBLER, W.; DEICHMANN, U.; GOTTSEGEN, J.; MALOY, K. World population in a grid of spherical quadrilaterals. **International Journal of Population Geography**, v. 3, p. 203-225, 1997.