

SEGREGAÇÃO SOCIAL EM BELO HORIZONTE: UMA COMPARAÇÃO ENTRE DOIS MÉTODOS DE *FUZZY CLUSTERING*

Adriana de Miranda-Ribeiro ♥

Ricardo Alexandrino Garcia *

Resumo

Este trabalho tem como objetivo comparar resultados obtidos pela aplicação do método FANNY (*Fuzzy Analysis*) com resultados obtidos pela aplicação do método GoM (*Grade of Membership*) na construção de um indicador de segregação social para Belo Horizonte. Para tanto, serão utilizados microdados do universo do Censo Demográfico 2000, desagregados por setores censitários. O método FANNY presta-se à modelagem multidimensional de dados contínuos e foi desenvolvido a partir da teoria dos conjuntos nebulosos de Zadeh (1965). O método de GoM, desenvolvido a partir da mesma teoria, presta-se à modelagem de dados discretos. A grande vantagem na utilização das metodologias consiste na capacidade de análise de dados heterogêneos e de representação matemática de conceitos vagos e imprecisos. Os dois métodos permitem que se estime o grau de associação de cada elemento a diversos *clusters*. Para a construção do indicador de segregação social, serão determinados dois *clusters*, representando os dois extremos da condição social. O ‘indicador de segregação social’ será o grau de associação do setor ao *cluster* de baixa condição social. Resultados preliminares indicam que os métodos caminham na mesma direção e que são ferramentas poderosas na composição de perfis, a partir de um número grande de variáveis e para objetos heterogêneos. No entanto, neste caso, o método FANNY apresentou maior diferenciação dos graus de pertencimento dos setores aos dois perfis criados, mostrando-se mais eficiente.

Palavras-chave: Segregação social; Análise de *cluster*; *Fuzzy Analysis*; *Grade of Membership*.

Sessão Temática: Relações raciais e in(ex)clusão social em Minas Gerais

♥ Secretaria de Estado de Educação de Minas Gerais.

* Instituto de Geociências / UFMG.

Segregação social em Belo Horizonte: uma comparação entre dois métodos de *fuzzy clustering*

1) Introdução e objetivos

O processo de urbanização no Brasil teve início na década de 1930, com a quebra da hegemonia agrícola, cresceu nas décadas seguintes, mas foi durante a década de 1980 que ocorreu a verdadeira inversão quanto ao local de residência da população brasileira. Ao final dos anos 1970, encerrou-se a trajetória de crescimento e diversificação das estruturas produtivas, com um certo nível de modernização da agropecuária no país e com a consolidação da inserção do país no contexto industrial. A partir daquele momento, a população urbana superara a população rural em tamanho (Santos, 1996). Aliado ao processo de urbanização, o processo de industrialização gerou a expansão do mercado de trabalho, criando um enorme número de empregos urbanos, o que permitiu a integração de grande parte da população à sociedade urbano-industrial. Esse dinamismo atraiu enormes parcelas da população, que passaram a se concentrar nas grandes cidades, principalmente nas metrópoles.

O modelo econômico adotado no país – concentrador e excludente – promoveu a expansão urbana apoiada em uma sociedade com distribuição de renda bastante desigual. Isso aconteceu porque os maiores ganhos, que são incorporados ao preço da terra urbana, ao capital e ao salário, beneficiam predominantemente os proprietários de terras, os donos do capital e os trabalhadores com qualificações valorizadas, enfim, devido a todos aqueles fatores que não contemplam os segmentos inferiores da distribuição de renda ou de riqueza (Fava, 1984). A pobreza surgiu, então, como um fenômeno generalizado, principalmente nas regiões metropolitanas – ainda que de forma diferenciada entre o nordeste e o sul do país –, revelando de maneira indiscutível as desigualdades sociais (Santos, 1996).

Nas grandes cidades latino-americanas, de um modo geral, e especialmente no Brasil, o espaço urbano vem sendo configurado a partir das relações mercantis, com nítido favorecimento das classes de renda mais alta. Ao mesmo tempo, as relações não-mercantis, baseadas na reciprocidade e/ou na lógica redistributiva da ação estatal têm um papel importante na produção e apropriação da moradia, por grande parte da população de baixa renda. Em grandes metrópoles como Rio de Janeiro, São Paulo, Belo Horizonte, Salvador e Recife, as ocupações ilegais em áreas centrais definiram uma configuração socioespacial particular, caracterizada pela proximidade física entre ricos e pobres. Além da ocupação de áreas centrais, a expansão urbana dessas cidades foi marcada pela expansão das periferias como o *locus* privilegiado da carência e da ilegalidade (Lago, 2003).

De acordo com Vainer (1998), a pobreza e as desigualdades acarretaram uma forma de ocupação em que os habitantes se fecham em determinadas áreas segundo o poder, o *status* ou a riqueza que detêm. Condomínios fechados de luxo, conjuntos habitacionais populares, favelas e a bairros de classe alta convivendo em espaços próximos são exemplos claros e frequentes nas grandes metrópoles. Visíveis ou não, as barreiras que separam os habitantes de uma mesma cidade fragmentam o espaço construído e definem o que se pode chamar de segregação espacial.

Caracterizada ou não pela distância geográfica, esta forma de segregação expressa a distribuição espacial das classes sociais determinada pelo mercado imobiliário.

A distribuição espacial dos grupos sociais se dá de forma a garantir que os grupos mais fortes se apropriem dos recursos materiais da cidade com maior facilidade. Os recursos necessários à reprodução dos grupos sociais, desigualmente distribuídos no espaço urbano, refletem chances desiguais de acesso. Desta forma, a construção do espaço urbano é hierarquizada, segregada. A segregação espacial, aliada ao contexto de pobreza e desigualdades faz surgir segmentos socialmente segregados, compostos por indivíduos sem acesso aos serviços básicos de infra-estrutura urbana e com acesso limitado aos serviços sociais – como saúde e educação – e acesso marginal ao mercado de trabalho (Caiado, 1998).

A sociologia brasileira tem analisado com bastante interesse a questão da segregação urbana desde a década de 1970. Os espaços urbanos ocupados pelos grupos sociais menos favorecidos foram caracterizados como periferias e descritos como espaços socialmente homogêneos, esquecidos pelas políticas públicas, e localizados em geral nas extremidades da área metropolitana. Nestes locais, grande parte das casas é construída pelos próprios moradores, em loteamentos irregulares ou ilegais, sem o cumprimento das normas e leis para a aprovação do assentamento no município (Torres et al, 2003).

Em períodos mais recentes, em que o desenvolvimento tecnológico permitiu o acesso a grandes bases de dados, mapas digitalizados etc., uma série de estudos têm mostrado, de maneira mais detalhada e localizada, a questão da pobreza urbana. O trabalho aqui proposto se insere nesta perspectiva. Pretende-se analisar a desigualdade no município de Belo Horizonte, a partir da aplicação de métodos multivariados a variáveis demográficas, socioeconômicas e de infra-estrutura urbana, em níveis bastante desagregados, buscando captar a heterogeneidade da ocupação urbana, mesmo em áreas muito próximas.

O município de Belo Horizonte nasceu planejado, ao final do século XIX, durante a recém proclamada República para ser a nova capital de Minas Gerais. A cidade representava a necessidade de se ter uma capital moderna, e mostrar a emergência de novas idéias e influências, o que se refletiu no traçado regular, nas grandes avenidas e nas construções imponentes.

A ocupação da cidade deveria ocorrer na área interna ao perímetro da atual Avenida do Contorno e seria ocupada pela população economicamente mais favorecida (os funcionários públicos, comerciantes etc); às classes menos favorecidas (os operários, empregados etc), destinar-se-ia a área suburbana, ou seja, fora do perímetro central. A ocupação de áreas distantes da área central tinha como principal objetivo dotar de infra-estrutura as áreas externas que ficavam mais próximas da área central, tornando-as mais valorizadas e visando futuras ocupações.

No entanto, o crescimento da cidade e o desenvolvimento das atividades econômicas – o que certamente atraiu trabalhadores de outras localidades – fizeram com que populações de classe mais baixa fossem ocupando os vazios urbanos de forma irregular e desordenada, na busca de um lugar para morar e da proximidade do local de trabalho. Com o passar do tempo, o crescimento e empobrecimento da população consolidaram este tipo de ocupação. Belo Horizonte teve e tem sua expansão urbana apoiada numa sociedade com distribuição de renda desigual, seu desenvolvimento marcado pelo crescimento da pobreza e da violência urbana, além de uma queda na qualidade de vida e degradação do espaço construído. Assim como nas demais metrópoles brasileiras, revela de maneira indiscutível as desigualdades sociais. Esse contexto de

pobreza e desigualdades faz surgir segmentos excluídos da ordem social, sem acesso aos serviços básicos de infra-estrutura urbana, acesso limitado aos serviços sociais, como saúde e educação, e acesso marginal ao mercado de trabalho (Caiado, 1988).

Miranda-Ribeiro e Garcia (2001) construíram um indicador de segregação social para Belo Horizonte, utilizando o método *Fuzzy Analysis – FANNY*. Para tanto, foram utilizados os microdados da Contagem Populacional de 1996,¹ do IBGE, desagregados por setores censitários. A opção de se utilizar dados desagregados por setores censitários deveu-se ao fato de esta ser a menor unidade espacial para a qual estão disponíveis os dados e baseou-se no fato de que Belo Horizonte apresenta, mesmo em áreas próximas, diferenças significativas.

O objetivo principal deste trabalho é elaborar um indicador de segregação social para Belo Horizonte a partir da utilização do método *GoM* e comparar os resultados com os obtidos por Miranda-Ribeiro e Garcia (2001), da aplicação do método *FANNY*. A principal diferença entre as duas metodologias é que, enquanto o *FANNY* aplica-se a dados contínuos, o *GoM* trabalha com dados discretos.

2) O Indicador de Segregação Social

Um indicador pode ser entendido como uma medida sintética, uma forma de se reduzirem informações, preservando-se, no entanto, a variabilidade das informações originais. A maior vantagem na utilização de um indicador reside na sua capacidade de representar uma realidade multidimensional e muitas vezes complexa em uma medida única. A maior limitação na utilização de um indicador está relacionada à sua interpretação, que requer cuidados, além de uma boa fundamentação teórica. A construção dos indicadores de segregação social e espacial proposta neste trabalho buscou, a partir das variáveis disponíveis, a melhor reprodução da sua definição. Para tanto, foram utilizados dois métodos de análise de *clusters*, desenvolvidos a partir da teoria dos conjuntos nebulosos.

De acordo com a literatura, pode-se dizer que a *segregação social* está ancorada em dois fatores: “desigualdade” e “pobreza”. Diante disso e das variáveis disponíveis, foi possível trabalhar alguns aspectos que, direta ou indiretamente, traduzem a base da definição. Além disso, outras variáveis também foram utilizadas, para efeito de melhor caracterização e diferenciação dos setores por parte do método empregado.

As variáveis foram divididas em três grupos, como mostra o Diagrama 1.²

¹ Em função do caráter da contagem populacional de 1996, foram levantados apenas dados referentes a sexo, idade, escolaridade, posição na família, migração interestadual, município de residência e situação de domicílio. Deste modo, a análise das características dos setores censitários esteve restrita a essas variáveis.

² A divisão não altera os resultados, ela serve apenas como um instrumento de exposição.

Diagrama 1: Variáveis que compõem o Indicador de Segregação Social

SEGREGAÇÃO SOCIAL

ASPECTOS SOCIAIS

percentual de empregados domésticos por domicílio
razão de sexo de chefia de domicílio

ASPECTOS DEMOGRÁFICOS

percentual da população residente do setor
razão de sexo
percentual da população masculina migrante
percentual da população feminina migrante

ASPECTOS DE ESCOLARIDADE

percentual da população que frequenta escola, por grupo etário
média de anos de estudo da pop. masculina, por grupo etário e total
média de anos de estudo da pop. feminina, por grupo etário e total

O primeiro grupo, ‘aspectos sociais’, é composto por duas variáveis. A primeira, ‘percentual de empregados domésticos por domicílio’ é a razão entre o número total de empregados domésticos³ e o número de domicílios do setor. A variável ‘razão de sexo de chefia de domicílio’ é a razão entre homens e mulheres chefes de domicílio. Sem entrar em discussões teóricas, normalmente, a denominação chefe de domicílio está relacionada à renda e à capacidade de prover os habitantes do domicílio, ou seja, geralmente é dito ‘chefe do domicílio’ aquele ou aquela que contribui com maior parcela da renda familiar.

O segundo grupo de variáveis, ‘aspectos demográficos’ é composto por variáveis que dizem respeito a fenômenos demográficos, que podem ter comportamentos bastante específicos, de acordo com o nível sócio-econômico da população. A variável ‘população total do setor’, é a porcentagem da população do setor em relação à população total do município. Dado que o setor é definido com base no número de domicílios, em torno de 300, uma grande diferença nesta variável pode representar diferenciais significativos de habitantes por domicílio. As variáveis ‘percentual de migrantes masculinos’ e ‘percentual de migrantes femininos’ dizem respeito apenas às migrações interestaduais ocorridas no período 91-96. São aqueles que não moravam em Belo Horizonte e que moravam fora de Minas Gerais em 01/09/1991. Isto significa, na maioria das vezes, que o migrante percorreu uma distância considerável, o que representa custos mais elevados e maiores riscos.

Os ‘aspectos educacionais’ estão representados no terceiro grupo. A frequência à escola a partir de determinada idade e, principalmente, os anos médios de estudo, são características fundamentais na determinação do nível de qualificação, da ocupação e da renda dos indivíduos. Em outras palavras, nível educacional diferenciado provoca níveis sócio-econômicos diferenciados.

³ Empregados domésticos que residem no domicílio.

3) Metodologia

O método FANNY – *Fuzzy Analysis* – presta-se à modelagem multidimensional de dados contínuos, tendo sido desenvolvido a partir da teoria dos conjuntos nebulosos, de Zadeh (1965). O método de GoM, desenvolvido a partir da mesma teoria de conjuntos, presta-se à modelagem de dados discretos. A grande vantagem na utilização de tais metodologias consiste na capacidade de análise de dados heterogêneos.

De um modo geral, a teoria dos conjuntos nebulosos permite que elementos distintos possuam graus de pertinência a vários conjuntos, de modo a tornar possível a representação matemática de conceitos vagos e imprecisos. O grau de pertinência de um elemento de um universo a um determinado conjunto ou partição nebulosa, é representado por um número real no intervalo $[0,1]$, que representa o quão verdadeira é a afirmação de que esse elemento pertence àquela partição (GARCIA, 2000). Enquanto que em determinados métodos de *análise de cluster* (*crisp sets*), cada elemento analisado pertence a um único *cluster*, gerando uma divisão clara dos elementos, os métodos FANNY e GoM permitem que se estime o grau de associação entre cada elemento e os diversos clusters. Em outras palavras, os métodos FANNY e GoM associam um objeto a diversos *clusters*.

Nos conjuntos discretos, em que se baseia a maioria dos métodos estatísticos, um elemento ou objeto pode pertencer ou não a um determinado conjunto com características específicas. Nos conjuntos nebulosos um elemento ou objeto pode pertencer parcialmente - ou seja, apresentar “graus” de pertinência ou pertencimento – a múltiplos conjuntos. Tais conjuntos generalizam os modelos estatísticos baseados na lógica discreta, visando aperfeiçoar a abordagem da heterogeneidade inerente aos elementos ou objetos a serem classificados (Cerqueira, 2004).

No método FANNY, para cada objeto i e cada cluster v , existe uma associação, $u_{i,v}$, a qual indica o grau de pertencimento do objeto ao cluster. As associações são definidas através de processos iterativos, buscando a minimização da função f :

$$f = \sum_{v=1}^k \frac{\sum_{i,j=1}^n u_{iv}^2 u_{jv}^2 d(i,j)}{2 \sum_{j=1}^n u_{jv}^2} \quad (1),$$

onde $d(i, j)$ representa as distâncias (ou dissimilaridades) conhecidas entre os objetos i e j ; e $u_{i,v}$ é a associação desconhecida dos objetos i ao cluster v . As funções de associação estão sujeitas às condições:

$$u_{i,v} \geq 0, \text{ para } i= 1, \dots, n; v= 1, \dots, k$$

$$\sum_v u_{i,v} = 1, \text{ para } i=1, \dots, n,$$

as quais mostram que as associações entre cada objeto e os diversos *clusters* tem necessariamente que ser nulas ou positivas, e que a soma das associações entre cada objeto e os diversos clusters é constante e igual a 1.

O método GoM estima, com base em um modelo de probabilidade multinomial, dois tipos de parâmetros: um de associação, g_{ik} , e outro de estrutura, λ_{kji} ; ou seja, os graus de pertencimento (g) de cada elemento (i) à cada subconjunto, ou perfil, ou tipo (k); e as probabilidades de cada

categoria (l) de cada variável (j) em cada diga-se, perfil (k) que, por sua vez, definem esses perfis. Segundo Manton (1994), esse o modelo de probabilidade multinomial é dado por:

$$L_{(y)} = \prod_{i=1}^I \prod_{j=1}^J \prod_{i=1}^{L_j} (\sum g_{ik} \lambda_{kjl})^{Y_{ijl}},$$

onde: Y_{ijl} representa o valor (zero ou um) que cada elemento (i) assumiu na categoria (l) da variável (j), sendo que, para todas as combinações de i e j, Y_{ijl} será sempre igual a zero, exceto para l, que será igual a unidade.

A partir da determinação do número de *clusters*, são criados ‘perfis ideais’. Assim, para cada *cluster*, cada variável passa a ter um ‘valor ideal’. Um objeto terá grau de associação a um *cluster* tanto mais próximo de 1, quanto mais próximos estiverem os valores de suas variáveis dos valores estabelecidos para o “perfil ideal”. Pode ocorrer, mas não necessariamente, que, dentre os objetos, haja tipos puros, ou seja, objetos que pertençam totalmente a um único cluster.

Miranda-Ribeiro e Garcia (2001) determinaram a criação de dois perfis, representando os dois extremos da condição social. Seguindo os objetivos deste trabalho, determinou-se o mesmo número de perfis; a partir de agora, os dois perfis passam a ser denominados *perfil de nível alto* e *perfil de nível baixo*.⁴ Para cada setor censitário, haverá um grau de associação a cada um dos dois clusters, cuja soma é igual à unidade. Uma maior associação ao ‘nível alto’ significa uma menor associação ao ‘nível baixo’ e se traduz em melhores condições sociais. Uma menor associação ao ‘nível alto’ se traduz em condições sociais menos favoráveis. O ‘indicador de segregação social’ (*ISS*) será o grau de associação do setor ao *perfil de nível baixo*. Deste modo, quanto mais próximo da unidade estiver o *ISS* de um setor, piores as condições sócio-econômicas do mesmo, e vice-versa.

4) Resultados

Para se chegar aos resultados, as 58 variáveis utilizadas por Miranda-Ribeiro e Garcia (2001) foram categorizadas, e aplicou-se o método *GoM* diversas vezes, com o intuito de aprimorar os resultados. Na primeira tentativa, foram utilizadas as variáveis categorizadas segundo os ‘quintis’ de suas distribuições, com ‘inlambda=random’ e ‘ingamma=default’. Na segunda tentativa, as frequências marginais de cada categoria das variáveis foram utilizadas como valores iniciais de ‘lambda’, e manteve-se ‘ingamma=default’. Na terceira tentativa, repetiu-se a primeira tentativa, mas com as variáveis categorizadas segundo os ‘decis’ de suas distribuições. Na quarta e quinta tentativas, ainda com as variáveis categorizadas segundo os ‘decis’ de suas distribuições, foram utilizadas as frequências marginais de cada categoria das variáveis para os ‘lambdas’ iniciais, mantendo os ‘gamas’ como *default*. Nas cinco tentativas, os resultados finais foram quase idênticos, com os graus de pertencimento aos dois perfis variando a partir da segunda casa decimal. Uma sexta tentativa foi feita a partir de uma categorização diferente das variáveis: ao invés de utilizar ‘decis’, cada variável foi dividida em 10 categorias de igual tamanho, de acordo com a sua amplitude. Para efeito de comparação entre as duas metodologias, a sexta tentativa não trouxe contribuição significativa, motivo pelo qual seus resultados não serão comentados.

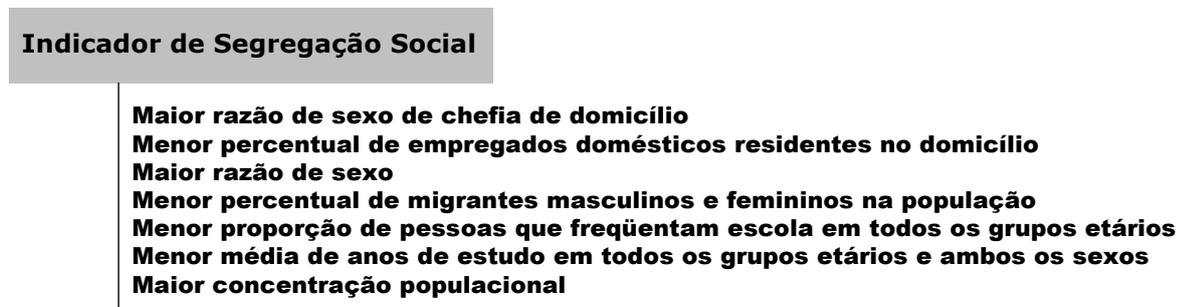
⁴ A denominação é apenas por questões operacionais, sem a pretensão de entrar discussões teóricas sobre o tema.

4.1) Caracterização dos perfis

Segundo os resultados encontrados e apresentados no Quadro 1, os dois perfis apresentaram, de fato, características opostas em todas as variáveis, o que os confere a possibilidade de representar os dois extremos da condição social. O perfil de “nível baixo” é caracterizado, em relação ao perfil de “nível alto” por: (i) maior razão de sexo de chefia de domicílio, o que indica que há mais mulheres chefiando domicílios nos setores de ‘nível baixo’; (ii) menor percentual de empregados domésticos residentes no domicílio, que indica menor poder aquisitivo dos setores de ‘nível baixo’; (iii) maior razão de sexo, o que reforça a questão das mulheres chefiando domicílios nos setores de ‘nível baixo’; (iv) menor percentual de migrantes (de data fixa) masculinos e femininos na população, mostrando que, proporcionalmente, mais residentes dos setores de ‘nível baixo’ residiam em Belo Horizonte em 1991; isto pode estar indicando uma menor mobilidade espacial para os residentes nos setores de ‘nível baixo’; (v) menor proporção de pessoas que freqüentam escola, em todos os grupos etários; (vi) menor média de anos de estudo em todos os grupos etários, para homens e mulheres; (vii) maior concentração populacional, indicada por uma média maior de pessoas residentes no setor (os setores censitários são definidos de acordo com o número de domicílios, em média, o mesmo para todos os setores; um número maior de pessoas residentes no setor indica que há mais pessoas por domicílio).

Miranda-Ribeiro e Garcia (2001) encontraram resultados semelhantes no que diz respeito à caracterização dos perfis de ‘nível baixo’ e ‘nível alto’. Neste aspecto, pode-se dizer que as duas metodologias em questão apresentaram resultados satisfatórios, dentro daquilo que se esperava.

Diagrama 2: Caracterização do Indicador de Segregação Social.



Quadro 1: Caracterização dos perfis de ‘nível alto’ e ‘nível baixo’, segundo o método GoM.

Variáveis	Decis	lambda / frequência marginal		Variáveis	Decis	lambda / frequência marginal	
		perfil nível alto	perfil nível baixo			perfil nível alto	perfil nível baixo
População Total	1	1,0297	0,9693	Razão de Chefia de Domicílio	1	1,6728	0,5010
	2	1,5000	0,6222		2	1,5750	0,5820
	3	1,7119	0,4475		3	1,3910	0,7080
	4	1,5465	0,5828		4	1,1470	0,8890
	5	1,2683	0,7950		5	0,7510	1,1760
	6	1,0869	0,9364		6	0,7250	1,1950
	7	0,6460	1,2750		7	0,8840	1,0870
	8	0,6446	1,2673		8	0,6160	1,2720
	9	0,3010	1,5277		9	0,4770	1,3720
	10	0,2596	1,5636		10	0,7120	1,2020
Porcentagem de Empregados Domésticos	1	0,0000	0,0000	Razão de Sexo	1	2,3440	0,0000
	2	0,0000	1,7353		2	2,2458	0,0979
	3	0,0000	1,7466		3	1,8404	0,3952
	4	0,0000	1,7450		4	1,4798	0,6511
	5	0,3300	1,4910		5	0,9510	1,0481
	6	0,8100	1,1370		6	0,4030	1,4404
	7	1,7510	0,4235		7	0,3471	1,4814
	8	2,3727	0,0000		8	0,2282	1,5495
	9	2,3850	0,0000		9	0,0464	1,7010
	10	2,3400	0,0000		10	0,1720	1,6100
Porcentagem de Homens Migrantes	1	0,3248	1,4762	Porcentagem de Mulheres Migrantes	1	0,2052	1,5536
	2	0,4104	1,4188		2	0,3796	1,4224
	3	0,7050	1,2000		3	0,5537	1,3083
	4	0,7844	1,1490		4	0,7191	1,1936
	5	0,8918	1,0735		5	0,6859	1,2182
	6	0,9869	1,0131		6	0,8037	1,1315
	7	1,0484	0,9763		7	1,0742	0,9581
	8	1,3667	0,7552		8	1,5269	0,6548
	9	1,6535	0,5455		9	2,0245	0,2888
	10	1,8424	0,4000		10	2,0168	0,2792
Porcentagem da população de 5 a 9 anos de idade que freqüenta escola	1	0,0000	1,7670	Porcentagem da população de 10 a 14 anos de idade que freqüenta escola	1	0,0000	1,7576
	2	0,0000	1,7510		2	0,2194	1,5765
	3	0,0000	1,7431		3	0,2875	1,5317
	4	0,1950	1,6110		4	0,4414	1,4222
	5	0,4426	1,4228		5	0,7440	1,2000
	6	1,1800	0,8710		6	1,1960	0,8653
	7	1,6378	0,5255		7	1,5198	0,6099
	8	2,0392	0,2294		8	1,5653	0,5853
	9	1,9630	0,2667		9	1,9975	0,2299
	10	2,2908	0,0000		10	0,0000	0,0000
Porcentagem da população de 15 a 19 anos de idade que freqüenta escola	1	0,0000	1,7380	Porcentagem da população de 20 a 24 anos de idade que freqüenta escola	1	0,0000	1,7040
	2	0,0000	1,7400		2	0,0000	1,7010
	3	0,0000	1,7450		3	0,0000	1,7040
	4	0,0000	1,7280		4	0,0000	1,7109
	5	0,5100	1,3650		5	0,0000	1,7131
	6	0,9250	1,0590		6	0,6600	1,2430
	7	1,4970	0,6250		7	2,1040	0,2060
	8	2,3424	0,0000		8	2,4390	0,0000
	9	2,3792	0,0000		9	2,4390	0,0000
	10	2,3460	0,0000		10	2,3580	0,0000
Média de anos de estudo da população masculina	1	0	1,673	Média de anos de estudo da população feminina	1	0,0000	1,6750
	2	0	1,650		2	0,0000	1,6550
	3	0	1,661		3	0,0000	1,6723
	4	0	1,664		4	0,0000	1,6776
	5	0	1,673		5	0,0000	1,6733
	6	0	1,662		6	0,0242	1,6646
	7	2,511	0		7	2,5030	0,0000
	8	2,511	0		8	2,5030	0,0000
	9	2,511	0		9	2,4901	0,0000

Fonte: Estimativas próprias

4.2) A segregação social

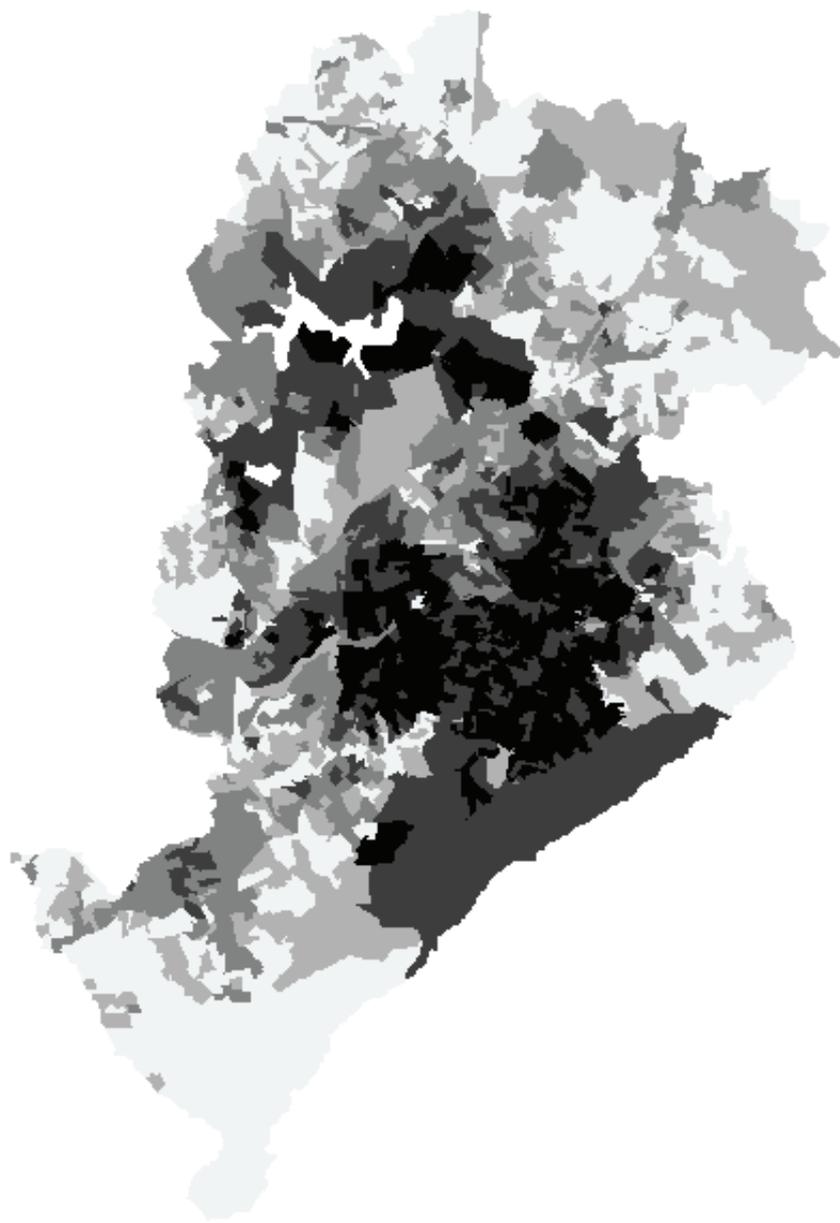
Para a análise dos resultados obtidos para os $ISS - ISS_{FANNY}$ e ISS_{GoM} – foram adotadas duas estratégias: na primeira fez-se uma comparação dos valores absolutos dos ISS_{FANNY} e ISS_{GoM} – embora saiba-se que não se deve considerar os valores absolutos sem que se tenha um cuidado especial; na segunda, analisou-se a distribuição espacial dos valores encontrados. Mesmo caminhando na mesma direção, os resultados apresentam-se bastante diferenciados.

Em termos absolutos, o que chama mais a atenção é o fato de que o ISS_{GoM} apresenta muitos setores com valores extremos, ou seja, com graus de pertencimento iguais a zero ou um. A rigor, isso indica que o setor pertence totalmente a um ou outro perfil, ou em outras palavras, ele pode ser considerado como um “tipo puro”. Ao todo, 809 setores foram considerados como totalmente de ‘nível alto’, enquanto que 580 foram considerados totalmente de ‘nível baixo’.

Uma rápida análise do ISS_{FANNY} indica que nenhum setor foi considerado um tipo puro. Os valores de ISS_{FANNY} variaram entre 0,1774 e 0,7823, indicando que o setor mais segregado socialmente pertence aproximadamente 78% ao perfil de ‘nível baixo’, enquanto que o setor menos segregado socialmente pertence cerca de 82% ao perfil de ‘nível alto’. Nestes dois setores, os valores de ISS_{GoM} foram iguais a 0 (zero) e 1 (um), respectivamente, o que não os diferencia dos outros setores que também foram considerados totalmente de nível baixo e alto.

Os Mapas 1 e 2 apresentam a distribuição espacial dos ISS , segundo os quintis de suas distribuições. Em ambos os mapas, as cores mais claras indicam maior *segregação social*. É possível observar que, embora haja uma menor concentração de setores segregados na região centro-sul e na Pampulha e uma maior concentração de setores segregados nas regiões Norte, Leste e próximo ao Barreiro, o Mapa 1, que mostra os resultados do método *FANNY* apresenta uma maior diferenciação dos valores do ISS do que o Mapa 2, que mostra os resultados do método *GoM*. Visualmente, isso se reflete pela maior proporção de tons intermediários no Mapa 1. Por exemplo, o Mapa 2 apresenta, nas regiões centro-sul e Pampulha, uma forte concentração da cor preta, o Mapa 1 mostra a mesma região representada por tons de cinza.

Mapa 1: Indicador de Segregação Espacial, segundo o método *FANNY*



Indicador de Segregação Social
FANNY

■	0,705 to 0,783	(415)
■	0,525 to 0,705	(427)
■	0,312 to 0,525	(423)
■	0,238 to 0,312	(415)
■	0,177 to 0,238	(428)

Fontes: IBGE: Malha Municipal Digital do Brasil
Contagem Populacional, 1996

Mapa 2: Indicador de Segregação Social, segundo o método *GoM*



Indicador de Segregação Social
GoM

■	0,61 to 1	(837)
■	0,04 to 0,61	(373)
■	0,02 to 0,04	(68)
■	0,01 to 0,02	(13)
■	0 to 0,01	(817)

Fontes: IBGE: Malha Municipal Digital do Brasil
Contagem Populacional, 1996

5) Considerações finais

Os resultados apresentados neste trabalho indicam que ambas as metodologias utilizadas – métodos *FANNY* e *GoM* – são ferramentas poderosas na composição de perfis a partir de um número grande de variáveis e para objetos heterogêneos. Apesar das diferenças entre os métodos, foi possível observar que os resultados caminham na mesma direção, o que, de certo modo, valida as duas metodologias.

No caso específico dos dados utilizados neste trabalho, o método *FANNY* mostrou-se mais adequado à construção do indicador de segregação social, por ter apresentado maior diferenciação dos ‘graus de pertencimento’ dos 2108 setores censitários de Belo Horizonte aos dois perfis criados. Essa constatação não necessariamente implica que o método *FANNY* seja melhor que o método *GoM*, mas pode ser um indício de que o método *FANNY* se aplica com mais eficiência a dados contínuos. Por outro lado, os resultados podem estar relacionados exatamente ao fato de os dados originais serem de natureza contínua, fazendo com que a categorização das variáveis não tenha gerado a heterogeneidade necessária à perfeita aplicação do método *GoM*.

Contudo, quaisquer que sejam as explicações, a aplicação preliminar em um conjunto de dados não pode ser conclusiva. Seria necessário, em primeiro lugar, trabalhar um pouco mais os dados, no sentido de tentar melhorar os resultados; em segundo lugar, seria fundamental repetir a aplicação dos dois métodos para outro conjunto de dados; por fim, poder-se-ia tentar a aplicação de outros métodos de *cluster* na comparação de resultados.

6) Referências Bibliográficas

BRITO, F. Mobilidade espacial e expansão urbana: o caso da Região Metropolitana de Belo Horizonte. In: X Encontro Nacional de Estudos Populacionais, 1996, Caxambu. **Anais...** Belo Horizonte, ABEP, 1996.

BRITO, F., SOUZA, J. A metropolização da pobreza. In: XI Encontro Nacional de Estudos Populacionais, 1998, Caxambu. **Anais...** Belo Horizonte, ABEP, 1998.

CAIADO, M. C. S. O Padrão de urbanização brasileiro e a segregação espacial da população na região de Campinas: o papel dos instrumentos de gestão urbana. In: XI Encontro Nacional de Estudos Populacionais, 1998, Caxambu. **Anais...** Belo Horizonte, ABEP, 1998.

CD-ROM “BH100 - Nossa História”. Estado de Minas, Belo Horizonte, Novembro de 1996.

CERQUEIRA, C.A. *Tipologias e características dos estabelecimentos escolares brasileiros*. 2004. 316 p. Tese (Doutorado em Demografia) – CEDEPLAR, Universidade Federal de Minas Gerais, Belo Horizonte, 2004. (disponível em <http://www.cedeplar.ufmg.br>)

FAVA, V. L. **Urbanização, custo de vida e pobreza no Brasil**. São Paulo: IPE/USP, 1984.

FERREIRA, F.P.M. Evolução urbana e demográfica do envelhecimento em Belo Horizonte. In: Seminário Mineiro de Economia, 9, 2000, Diamantina. **Anais...** Belo Horizonte, CEDEPLAR/UFMG, 2000.

GARCIA, R.A. **Modernização e Declínio da Fecundidade no Nordeste Brasileiro: um estudo micro-regional**. Belo Horizonte, CEDEPLAR/UFMG, 2000 (dissertação de mestrado).

KAUFMAN, L. ROUSSEEUW, P.J. **Finding groups in data: an introduction to cluster analysis**. New York, John Wiley & Sons Inc., 1990.

LAGO, L.C. (1998) Estrutura urbana e mobilidade espacial: alterações no quadro de desigualdades socioespaciais na metrópole do Rio de Janeiro. Encontro Nacional de Estudos Populacionais, XI., 1998, Caxambu. *Anais...* Belo Horizonte: ABEP, 1998. p.383-419.

MIRANDA-RIBEIRO, Adriana, GARCIA, Ricardo Alexandrino. “Segregação Sócio-Espacial em Belo Horizonte: uma aplicação de modelos difusos”. In: XXIV IUSSP General Conference, Salvador, agosto de 2001. *Anais ...* Campinas, ABEP, agosto de 2001 (disponível em meio magnético).

PASTORE, J. ZYLBERSTAIN, H., PAGOTTO, C. S. **Mudança social e pobreza no Brasil: 1970-1980 (o que ocorreu com a família brasileira?)**. São Paulo, Pioneira/FIPE, 1983.

SABÓIA, J. **Distribuição de renda e pobreza metropolitana no Brasil**. Rio de Janeiro: UFRJ/IEI, 1991.

SANTOS, M. **A urbanização desigual**. Petrópolis, Editora Vozes, 1980.

_____. **Urbanização Brasileira**. São Paulo, Editora Hucitec, 1996.

SECRETARIA MUNICIPAL DE PLANEJAMENTO DE BELO HORIZONTE. Revista Planejar BH, agosto de 2000.

SINGER, Paul. **Desenvolvimento Econômico e Evolução Urbana**. São Paulo, Cia Editora Nacional, 1977.

SOUZA, J.M., TEIXEIRA, J.G. Desigualdade sócio-espacial e migração intra-urbana na Região Metropolitana de Belo Horizonte, 80-91. (Trabalho apresentado no XXI Encontro Anual da Associação Nacional de Pós-Graduação e Pesquisa em Ciências Sociais, 1997, Caxambu, mimeo) .

TORRES, H.G. et al. Pobreza e espaço: padrões de segregação em São Paulo. *Estudos Avançados*, São Paulo: USP, v.17, n. 47, p.97-128.

VAINER, C. B. Cidades, cidadelas e a utopia do reencontro – uma reflexão sobre tolerância e urbanismo. In: **Cadernos IPPUR**, ano XII, n. 1, 1998, p. 33-46.

ZADEH, L. A. Fuzzy Sets, **Inf. Control** **8**, 338-353, 1965.