

DECISÃO DE LOCALIZAÇÃO DE ESCOLAS COM USO DE SISTEMA DE INFORMAÇÃO GEOGRÁFICA E ANÁLISE DE AGRUPAMENTOS

F. C. C. MELO^{*1}, A. K. C. LIMA², J. S. C. FERREIRA³

¹Departamento de Ciências Econômicas - Universidade do Estado do Rio Grande do Norte - UERN,

²Departamento de Administração - UERN, ³Departamento de Ciências Contábeis – UERN

fcarlosdemelo@uol.com.br *

Submetido 14/08/2017 - Aceito 10/07/2018

DOI: 10.15628/holos.2018.6552

RESUMO

O conhecimento das características da população, aglutinadas em perfis homogêneos, pode indicar demandas e, por conseguinte, sugerir locais para instalação de empreendimentos educacionais. Este estudo desenvolve esforço metodológico ao propor a utilização conjunta da análise de agrupamentos e Sistemas de Informações geográficas como subsídio para a tomada de decisão de localização de escolas. Para análise da metodologia, foi realizada a segmentação dos setores censitários da área urbana da cidade de Mossoró/RN e definida a localização georreferenciada das escolas municipais do ensino fundamental. O agrupamento dos casos em estudo foi efetuado mediante análise hierárquica de *Clusters* com o método Ward, usando a distância euclidiana quadrada como medida de similaridade. O dendograma resultante sugere a formação de até 5 *Clusters*, quantidade confirmada pelo critério do R- Quadrado (R^2). Nas áreas representadas pelos *Clusters* 2 e 4, está instalado o

maior número de unidades educacionais. Nos *Clusters* 1 e 3, está instalado o menor número de escolas da rede pública, onde existe o melhor desempenho de renda e menor contingente populacional. O *Cluster* 5 detém um contingente populacional de 40.011 pessoas, sendo 5.346 delas na faixa de 6 a 14 anos, com percentual de alfabetização de 90%. Neste *Cluster*, a renda média é de R\$ 1514,25. Os resultados evidenciam que as escolas que compõem a Rede Municipal de Ensino apresentam boa distribuição no espaço geográfico da área urbana do município de Mossoró, concentrando-se em áreas da população com menor poder aquisitivo. O perfil do *Cluster* 5, contudo, sugere que há demanda por investimentos em infraestrutura educacional. A aplicação da metodologia proposta sugere que a realização de estudos com uso conjunto das técnicas SIG e análise de agrupamentos é promissora, sendo necessárias novas pesquisas para refinar a metodologia proposta.

PALAVRAS-CHAVE: Gestão da educação, localização de escolas, sistemas de informações geográficas, análise de agrupamentos (*Clusters*).

DECISION ON LOCALIZATION OF SCHOOLS USING GEOGRAPHIC INFORMATION SYSTEM AND CLUSTER ANALYSIS

ABSTRACT

Knowledge on characteristics of population, agglutinated in [homogeneous](#) profiles, may indicate needs and, consequently, suggest places to build schools. This research develops methodological effort to suggest to use combination of SIG and Cluster analysis for decision-making on the localization of schools. For the analysis of methodology, we accomplished a Cluster analysis from census sectors in the urban plan of Mossoró/RN and defined georeferenced localization from communal schools with primary education. The cases' Clusters was made through hierarchical analysis of Clusters using Ward method, using Euclidian distance as similarity measure. Resulting dendogram suggests forming at most five Clusters, number confirmed through R-square method (R^2). In the areas represented by Cluster 2 and 4, most schools are built. On

Clusters 1 and 3, the least number of public schools is located, where is registered the best profit performance and the lowest population contingent. Cluster 5 has a population contingent of 40.011 inhabitants, 5.346 of them in the age group from 6 to fourteen years and 90% of literacy. In this Cluster, rental average is R\$1514.25. Results show schools forming communal education network have a good allocation in geographical space of urban area in Mossoró, concentrated in sectors with lower buying power. However, the profile of Cluster five suggests a demand for public investments and possible private investments. The use of proposed methodology suggests the studies using techniques SIG and Cluster are promising, being necessary new investigations in order to refine proposed methodology.

KEYWORDS: Education management, Location of schools, geographic Information system, *Cluster* analysis.

1 INTRODUÇÃO

A informação para subsidiar a tomada de decisão sobre localização de empreendimentos é um recurso estratégico muito valioso para as organizações. Quando aliada aos recursos tecnológicos, a informação influencia estratégias e proporciona vantagens competitivas para organizações públicas e privadas, contribuindo para o alcance dos objetivos organizacionais.

Preocupações com localização de instalações despertam interesses investigativos nas áreas de engenharia, decisão, ciências da computação, matemática, ciências sociais e ciência ambiental (FARAHANI; STEADIESEIFI; ASGARI, 2010), nos setores privado e público (MAPA; LIMA, 2012). Para Chung; Tcha (1992), o problema de localização no setor privado se concentra em minimizar o custo total para o cumprimento dos requisitos das demandas dadas, ao passo que no setor público o problema visa ao maior grau de satisfação social pela utilização eficiente de recursos em um sistema de distribuição.

Farahani et al. (2010) identificaram cerca de 730 trabalhos envolvendo problemas de tomada de decisão ou decisão de localização. Esse campo de estudo compõe uma parte importante da pesquisa operacional e da ciência administrativa, podendo ser classificadas em: Tomada de Decisão Multiobjetivos, Tomada de Decisão de Múltiplos Atributos, cujas combinações possíveis possibilitam a Tomada de Decisão Multicritérios. A tomada de decisão envolve no mínimo uma função objetiva (como custo, lucro, receita, distância de viagem, serviço, tempo de espera, cobertura e partes de mercado).

As pesquisas nesse campo de estudo são amplas e diversificadas, abrangendo, por exemplo, transporte público (FERNANDES; MAURO; ALIXANDRINI JUNIOR; SANTOS, 2012), unidades básicas de saúde (GOUDARD; OLIVEIRA; GERENTE, 2015; AHMADI-JAVID; SEYEDI; SYAM, 2017) e resíduos sólidos (EISELT; MARIANOV, 2015). Estudos de localização de escolas também são relativamente comuns (PIZZOLATO; BARROS; BARCELOS; CANEN, 2004). Contudo, o interesse pelo tema continua (NEVES, 2015; MOURA; IGNACIO, 2016), provavelmente porque a demanda por educação não foi suprida, persistindo os desafios de acesso.

Estudos têm empreendido esforços para desenvolver ou aprimorar técnicas de localização de escolas utilizando Sistemas de Informações Geográficas - SIG (BATISTA; ORTHE BORTOLUZZI, 2011; VERRAM, 2014), Sistema Espacial de Apoio à Decisão com interface de SIG (CARVALHO; BARBIRATO; CIRILO; POLETO, 2016), localização de escolas públicas em áreas rurais e urbanas utilizando uma combinação de p-mediana, cobertura, p-centro e dos fatores de localização em sua construção (CARVALHO, 2011; SANTOS, 2012). Apesar desses esforços, a implantação de equipamentos urbanos ocorre de forma pouco sustentável, pouco acessível à população, mal locados e mal dimensionados (NEVES, 2015), o que indica a existência de lacunas a serem preenchidas.

Este estudo propõe uma metodologia, que pode ser classificada como recurso de Tomada de Decisão Multicritérios, para subsidiar a decisão de localização de escolas em municípios brasileiros, mediante uso conjunto de SIG e análise de agrupamentos. O esforço metodológico processa dados demográficos, socioeconômicos e geográficos disponíveis no último censo

demográfico realizado no Brasil, para estabelecer um perfil da população e obter indícios sobre suas demandas por investimentos em infraestrutura educacional.

Ao utilizar escolas públicas municipais do ensino fundamental como objeto de estudo, espera-se contribuir para o desenvolvimento de técnica de localização objetiva, confiável e acessível aos tomadores de decisão do setor de educação pública. Em especial, a metodologia pode ser particularmente útil para a maioria dos pequenos e médios municípios brasileiros, que dispõem de poucos recursos humanos e tecnológicos, o que dificulta a utilização das sofisticadas técnicas de localização disponíveis na literatura.

Este estudo segue o seguinte roteiro: revisa a literatura sobre SIG e análise de agrupamentos em estudos de localização; segmenta os setores censitários da área urbana do município de Mossoró/RN; elabora um modelo georreferenciado da segmentação dos setores censitários da área urbana do município de Mossoró/RN; elabora um mapa georreferenciado da localização das escolas do ensino fundamental da área urbana do município de Mossoró/RN, por setor censitário; apresenta indícios para orientar a tomada de decisão para localização de novas escolas públicas do ensino fundamental.

2 SIG E ANÁLISE DE AGRUPAMENTOS EM ESTUDOS DE LOCALIZAÇÃO

Importante instrumento no processo decisório, dados demográficos e socioeconômicos aliados a aspectos georreferenciados podem ser estudados conjuntamente com suporte dos SIGs, que são ferramentas de apoio ao geomarketing. Essas tecnologias da informação armazenam dados georreferenciados que também podem ser trabalhados com uso de recursos da matemática e da estatística, que servem de parâmetros para as análises espaciais e nos processos decisórios sobre a melhor localização para um empreendimento.

A população é a primeira força macroambiental. As mudanças demográficas (grandeza, composição, renda, posição social, nível educacional, crescimento ou declínios de etnias em determinadas áreas geográficas, ascensão e queda de classes sociais) se configuram como uma das principais e mais observadas mudanças externas a uma organização, pois não são ambíguas e são mais previsíveis (DRUCKER, 2008; KOTLER, 2013). Informações demográficas são, portanto, cruciais quando relacionadas a determinada localidade onde se pretende empreender e investir (SCHIFFMAN; KANUK, 2009).

O estudo do perfil e das projeções demográficas de uma população permite antecipar necessidades e desejos dos clientes, tendências como redistribuição geográfica, aumento da diversidade étnica e envelhecimento da população (SHETH; MITTAL; NEWMAN, 2001). O pleno conhecimento dessas tendências ajuda as organizações públicas e privadas a traçar suas estratégias mercadológicas, como o melhor *mix* de marketing (incluindo a localização), para atender aos anseios de determinado segmento da população.

A associação entre aspectos demográficos e sociodemográficos com características do espaço geográfico dá-se o nome de geomarketing. Essa disciplina é recente e se apóia também em outras disciplinas a relacionadas, como a ciência da computação, estatística e mapeamento (COSTA; NEVES, 2006; CHASCO, 2003), funcionando por meio de SIGs que facilitam o tratamento

de dados estatísticos e representações gráficas que dão suporte a tomada de decisões mercadológicas. Seus conceitos e recursos, frequentemente associados apenas ao setor privado, podem dar suporte também aos tomadores de decisões do setor público.

Os SIGs podem ser determinantes na consistência das decisões e escolhas de localização, tendo em vista que a amplitude dos métodos de coleta de informações proporciona convergência entre os dados armazenados, possibilitando a construção de parâmetros para a tomada de decisões e adoção de estratégias. A seguir, são apresentados estudos de localização de escolas, que utilizam SIGs aliados a outros recursos técnicos, como matemática e estatística.

Pizzolato et al. (2004) realizaram estudo com o objetivo de divulgar diversas experiências realizadas na área de localização de escolas. O estudo destaca o modelo da p-mediana como o mais popular em estudos de localização de escolas, com base no critério de menor distância, e simboliza o deslocamento do aluno até a escola mais próxima. Estudo seguindo a mesma metodologia foi realizado por Ndiaye; Ndiaye; Ly (2012); Maller; Gandolpho (2014).

Tagliatti (2010) verificou as relações espaciais entre distribuição de dados socioeconômicos e a localização de escolas públicas em Araraquara (SP) utilizando SIG e Mineração de Dados relacionados ao nível de renda, escolaridade e idade da população. Moura; Ignacio (2016), por sua vez, utilizaram um modelo clássico de programação linear/inteira, para localização de escolas, tendo como área de estudo a cidade de médio porte de Macaé (RJ). O estudo considerou os seguintes parâmetros: Demanda, dada ao nível de setor censitário; Distância, cuja matriz considera o percurso gerado do centro de massa à unidade escolar; Capacidade das escolas, dada em número de crianças atendidas em cada unidade escolar.

Moura; Ignacio (2016) apresentam diversos modelos consolidados de localização de escolas, inclusive citando que modelos matemáticos são importantes ferramentas da Pesquisa Operacional (PO), que, em geral, definem parâmetros como capacidade total da unidade escolar, custo, distância, etc. Consideram a situação real de infraestrutura da unidade de escolar, a partir de dados sobre demanda (população por faixa etária em cada setor censitário), distância e capacidade das escolas, utilizando programação linear/inteira.

A segmentação de mercado, por sua vez, é o processo de identificação de grupos de clientes existentes e potenciais que compartilham aspectos sociais, necessidades ou padrões de consumo, como idosos com alto poder aquisitivo, minorias jovens de baixo poder aquisitivo (LOVELOCK; WRIGHT, 2002; KOTLER, 2008). Pode ser chamada de “segmentação geodemográfica” quando identifica unidades domésticas específicas em determinado mercado por meio da geografia de uma área, bairro ou região para classificar bairros reais em que os consumidores vivem (PETER; OLSON, 2009), indicando que o mapeamento geográfico vinculado a dados socioeconômicos possibilita o conhecimento das necessidades de um determinado público.

Segmentação de mercado pode ser realizada com uso da análise de agrupamentos (MELO; ARAÚJO; CRISOSTOMO; ZARCAN, 2013), já utilizada em estudos na área de educação por Brandão; Ramos; Troccoli (2003); Gomes; Levy; Lachtermacher (2014); Mantovani; Gouvêa; Tamashiro (2015). Verica; Villwock; Johann (2015), por sua vez, utilizaram técnicas de mineração de dados e análise de agrupamentos para analisar o perfil de 399 municípios do Estado do Paraná, com base

de dados formada por 18 variáveis educacionais e o Produto Interno Bruto, que formaram 10 *Clusters* pelo método Ward, o qual utiliza a distância euclidiana como medida de dissimilaridade.

A decisão de localização para instalação de uma escola está relacionada à demanda da região, a uma série de informações que podem dar para este processo de forma, minimizando a possibilidade de erros. Informações relacionadas ao perfil da demanda estão disponíveis no censo demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE), fonte útil para prospectar e avaliar zonas de atuação e aproveitar oportunidades de mercado (FARIAS, 2012). Com os dados disponíveis no Censo demográfico, é possível construir indicadores sociais, importante instrumento operacional para monitoramento da realidade social para fins de formulação e reformulação de políticas públicas. Estes mesmos indicadores auxiliam no trabalho de planejamento, implementação, execução, avaliação dos programas, projetos e serviços sociais (JANNUZZI, 2009).

3 PROCEDIMENTOS METODOLÓGICOS

Este estudo, que tem caráter exploratório e quantitativo, elaborou a segmentação dos 242 setores censitários da área urbana do município de Mossoró/RN. Foram identificados cinco agrupamentos pela análise hierárquica de *Clusters* com o método Ward, usando a distância euclidiana quadrada com medida de similaridade, refinada pelo método K-means. Em seguida, os *Clusters* obtidos foram inseridos em um mapa georeferenciado da localização das escolas do ensino público fundamental da área urbana do município de Mossoró/RN.

Optou-se pela utilização do município como unidade de análise, na medida em que a gestão do ensino fundamental brasileiro acontece no nível municipal. A escolha da Rede Municipal de Ensino (RME) administrada pela Prefeitura Municipal de Mossoró/RN como objeto de análise, para aplicação da metodologia proposta, ocorre em virtude da disponibilidade da localização georreferenciada das escolas. Contudo, em princípio, qualquer município poderia ser objeto da metodologia proposta, executada com suporte em *softwares* apropriados para procedimentos de técnicas e análise estatísticas de dados e elaboração do SIG, conforme detalhado adiante.

Os dados demográficos e socioeconômicos utilizados são originários dos setores censitários da área urbana de Mossoró/RN, obtidos no censo realizado pelo IBGE no ano de 2010, último ano em que esses dados foram levantados. A RME composta por unidades escolares que desenvolvem atividades nas áreas de educação infantil e ensino fundamental. Na área urbana, a RME é composta por 67 Unidades Educacionais localizadas, sendo 35 na educação infantil e 32 no ensino fundamental. As escolas estão distribuídas nos diversos bairros da cidade de Mossoró-RN, conforme figura 1.

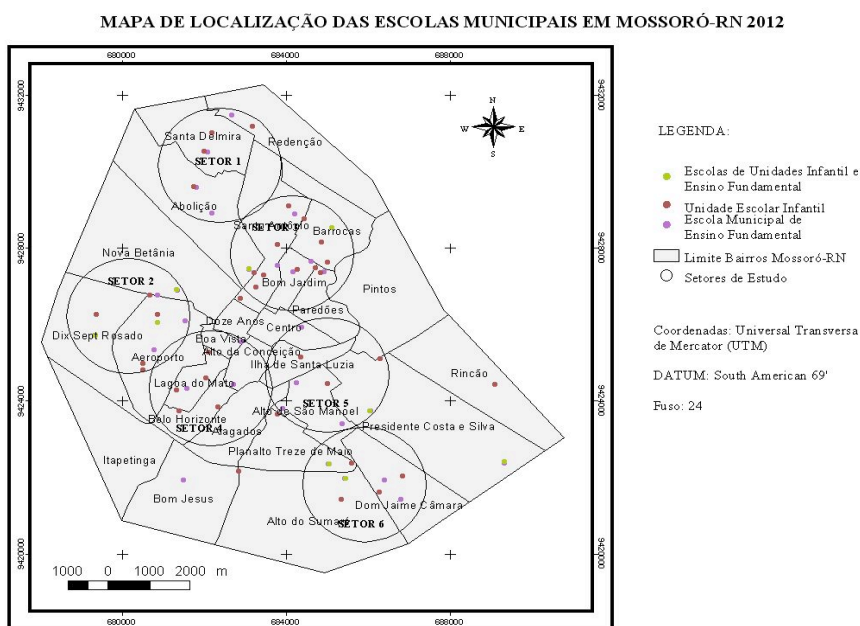


Figura 1: Mapa de Localização das Escolas Municipais Urbanas em Mossoró-RN. Fonte: Pesquisa dos autores (abril, 2013).

O desenho da amostra é estruturado de modo a analisar o perfil demográfico e socioeconômico por setores censitários da área urbana do município de Mossoró/RN, para subsidiar decisões de localização de escolas públicas do ensino fundamental. A amostra contém os 242 setores censitários da área urbana da cidade, conforme definido pelo IBGE para realização do censo brasileiro em 2010. Esses setores censitários formam os 28 bairros existentes na área urbana da cidade de Mossoró-RN, com 95% da população total do município. Após a realização da análise exploratória dos dados, optou-se pela retirada de dois *outliers*. A retirada desses dois casos, que representam setores censitários, gera espaços vazios no mapa.

O planejamento da análise de *Clusters* cumpriu os seguintes estágios (ALDENDERFER; BLASHFIELD, 1984 como citado em FIGUEIREDO FILHO; SILVA JÚNIOR; ROCHA, 2012): seleção da amostra; determinação das variáveis; definição da medida de similaridade; delimitação do número de *Clusters*, validação do resultado.

As variáveis foram selecionadas no Censo Demográfico de 2010 do IBGE, estando relacionadas à educação, renda e condições de vida em termos domiciliares e urbanísticos. Os valores das variáveis são disponibilizados pelo IBGE em termos absolutos, motivo pelo qual precisam ser discretizados, o que implicou utilização do conceito razão por domicílio (SILVA; GARCIA, 2012), à exceção da variável renda, já apresentado em termos de rendimento médio por pessoa. A Tabela 1 apresenta as variáveis selecionadas para pesquisa.

Tabela 1: Variáveis selecionadas para Pesquisa.

Variável	Descrição da Variável
V1	Código do Setor Censitário
V2	Nome do Setor Censitário (segmento do bairro)
V3	Total de pessoas de 0 a 5 anos
V4	Total de pessoas de 6 a 14 anos
V5	Total de pessoas de 5 anos ou mais
V6	Total de pessoas com 5 anos ou mais alfabetizadas
V7	Total de pessoas de 6 a 14 anos alfabetizadas
V8	Rendimento nominal médio mensal das pessoas
V9	Domicílios particulares e coletivos
V10	Domicílios particulares permanente com abastecimento da rede geral
V11	Dom. part. permanentes c/ banheiro de uso exclusivo dos moradores ou sanitário
V12	Dom. part. permanentes com banheiro de uso exclusivo ou sanitário e esgotamento sanitário via rede geral de esgoto ou pluvial
V13	Domicílios particulares permanentes com lixo coletado
V14	Dom. part. permanentes com energia elétrica de companhia Distribuidora
V15	Dom. part. Perm. com moradia adequada – Existe identificação do logradouro
V16	Dom. part. Perm. com moradia adequada – Não existe identificação do logradouro

Fonte: Dados IBGE (2010).

Após seleção das variáveis, se estabeleceu as razões entre as variáveis, na forma apresentada na Tabela 2.

Tabela 2: Variáveis razão utilizadas na análise de *Cluster*

Variável	Descrição da Variável
V1	Pessoas de 6 a 14 anos alfabetizadas / Total de pessoas de 6 a 14 anos
V2	Pessoas com 5 anos ou mais alfabetizadas / Pessoas de 5 anos ou mais
V3	Rendimento nominal médio mensal das pessoas
V4	Água / Domicílio
V5	Banheiro / Domicílio
V6	Esgotamento / Domicílio
V7	Lixo / Domicílio
V8	Energia / Domicílio
V9	Moradia adequada / Domicílio

Fonte: Pesquisa dos autores (abril, 2013).

As variáveis da Tabela 2 foram utilizadas na análise de *Cluster* que fundamenta este trabalho. Este método é uma técnica estatística que tem como objetivo agrupar elementos com características semelhantes. Explorar grupos homogêneos visa a identificar relações significativas entre os dados observados.



O agrupamento dos casos em estudo foi efetuado pela análise hierárquica de *Clusters* com o método Ward, usando a distância euclidiana quadrada com medida de similaridade, depois de testadas outras alternativas. O dendograma resultante sugere a formação de até cinco *Clusters*. Para definição final do número de *Clusters* retidos, utilizou-se o critério do R- Quadrado (R^2), descrito por Marôco (2011) como uma medida da variabilidade retida em cada uma das soluções dos *Clusters*. Essa solução sugere o número de *Clusters* a serem retidos e indica o percentual de explicação da variância total. Por esse critério, foram retidos 5 *Clusters* com 95% ($R\text{-sq} = 0,947$) de explicação da variância total.

Após a análise hierárquica, a classificação dos casos foi refinada pelo método K-means, com $k=5$, cuja distribuição dos casos está expressa na Tabela 3.

Tabela 3: Cluster Não-hierárquico (K-means).

Nº de <i>Cluster</i>	Número de setores censitários
1	14
2	72
3	9
4	104
5	41
Total	240

Fonte: Pesquisa dos autores (abril, 2013).

Conforme Malhota (2012), dados os vários julgamentos que a análise de *Cluster* acarreta, nenhuma solução de aglomeração deve ser aceita sem uma avaliação da sua confiabilidade e validade. Desta maneira, o último estágio, dedicado à validação dos resultados, foi realizado por meio da divisão da amostra de duas partes iguais (particionamento), cujos dendogramas indicam o mesmo número de *Clusters* e cujas estatísticas descritivas apresentam resultados semelhantes.

4 ANÁLISE E INTERPRETAÇÃO DOS RESULTADOS

Marôco (2011) sugere que as variáveis que mais contribuem para o significado dos *Clusters* são aquelas com maior valor da estatística F, com dados de saída do SPSS constantes na tabela da ANOVA ($F=QMC/QME$), semelhante à análise discriminante. Dessa maneira, pela ordem, os *Clusters* são definidos pelo Rendimento Médio Mensal, Razão da Alfabetização para pessoas com mais de 5 anos e Razão da Alfabetização para pessoas entre 6 e 14 anos.

A Tabela 4 apresenta a média das variáveis para cada *Cluster*. Sua análise contribui para caracterizarmos cada um dos *Clusters*.

Tabela 4 - Média das variáveis para cada Cluster.

	Cluster				
	1	2	3	4	5
Razão Alfabetização 6_14anos	,86	,89	,94	,85	,90
Razão Alfabetização 5anos	,86	,88	,92	,83	,89
Rendimento Médio Mensal	2288,94	994,94	3271,39	603,74	1514,25
Razão Água Domicílio	,98	,98	1,00	,96	,98
Razão Banheiro Domicílio	1,00	,99	1,00	,99	1,00
Razão Esgotamento Domicílio	,43	,44	,55	,37	,59
Razão Lixo Domicílio	,97	,98	1,00	,97	,97
Razão Energia Domicílio	,99	,99	1,00	,99	,99
Razão Moradia Domicílio	,73	,69	,66	,57	,74

Fonte: Pesquisa dos autores (abril, 2013).

A análise da Tabela 4 permite inferir as seguintes características para cada Cluster:

O Cluster 1, formado por 14 setores censitários, que apresentam um contingente de 10.783 pessoas, sendo 1.416 delas na faixa etária de 6 a 14 anos. Esse Cluster apresenta a segunda maior renda média mensal (R\$ 2.228,94), aproximadamente o dobro da renda média mensal da área urbana do município de Mossoró, R\$ 1.121,43. Os scores para alfabetização entre pessoas com mais de 5 anos e pessoas entre 6 e 14 anos são iguais a 86%. É o único Cluster no qual essas duas variáveis apresentam o mesmo resultado. Apenas cinco escolas da rede pública municipal estão instaladas nesse Cluster, o menor número entre todos os Clusters.

O Cluster 2, formado por 72 setores censitários, que apresentam um contingente de 73.890 pessoas, sendo 10.165 delas na faixa etária de 6 a 14 anos. Esse Cluster apresenta a segunda menor renda média mensal (R\$ 994,94), com scores para alfabetização entre pessoas com mais de 5 anos e pessoas entre 6 e 14 anos iguais a 88% e 89%, respectivamente. A distribuição espacial dos setores censitários que compõem esse Cluster é dispersa, com ocupação mais ampla de duas áreas periféricas da cidade.

O Cluster 3, formado por 9 setores censitários, que apresentam um contingente de 9.175 pessoas, sendo 1.167 delas na faixa etária de 6 a 14 anos. Esse Cluster apresenta a maior renda média mensal por pessoa entre os setores observados (R\$3.271,39). Também é o setor com melhor desempenho em seis outras variáveis, com destaque para a variável alfabetização, que apresenta percentual de 94% de alfabetização entre pessoas de 6 a 14 anos, percentual superior ao da segunda faixa etária, representada pela variável alfabetização entre pessoas de 5 anos. Apenas seis escolas da rede pública municipal estão instaladas nesse Cluster, o segundo menor número de escolas instaladas entre todos os Clusters.

Registra-se, contudo, que esse Cluster apresenta o segundo menor percentual de moradias adequadas, apesar da renda mensal mais elevada. Essa aparente contradição pode ter a seguinte explicação: no bairro Nova Betânia, onde estão localizados os setores censitários que formam este

Cluster, há uma área com moradias ocupadas majoritariamente por famílias de baixa renda, a favela do Ouro Negro, que passa por processo de urbanização.

O **Cluster 4**, formado por 104 setores censitários, distribuídos em praticamente todas as áreas da cidade, mas com concentração nas zonas Noroeste e Sudoeste, que reúnem um total de 101.245 pessoas, com sendo 15.270 na faixa de 6 a 14 anos. Este *Cluster* apresenta a menor renda média por pessoas (R\$ 610,28) e os menores percentuais de alfabetização, acesso à água, saneamento e moradias adequadas.

O **Cluster 5**, formado por 41 setores censitários, que compreendem um contingente de 40.011 pessoas, sendo 5.346 delas na faixa de 6 a 14 anos. Esse *Cluster* apresenta uma renda média mensal de R\$ 1.514,25, a terceira melhor entres os cinco *Clusters*. Apesar disso, seus domicílios apresentam o melhor acesso a esgotamento sanitário (59%), o melhor percentual de moradias consideradas adequadas pelo IBGE (74%) e *scores* para alfabetização entre pessoas com mais de 5 anos e pessoas entre 6 e 14 anos são iguais a 89% e 90%, respectivamente. A distribuição geográfica dos setores censitários que compõem esse *Cluster* é bastante dispersa.

Os *Clusters* dos setores censitários, com localização das unidades escolares na planta urbana da cidade de Mossoró/RN, estão representados na Figura 2. Esse mapa foi elaborado a partir do módulo *geoprocessing Wizard*, do programa Arcview® GIS 3.2, que possibilita o cruzamento entre dois mapas, como também *Arcview Projection Utility* para modificação dos dados de projeção dos mapas obtido do IBGE. Os dados foram transferidos e o formato de coordenadas foi alterado para *Universal Transverse de Mercator-UTM* no DATUM *South American 69'*, que é o Sistema Geodésico Brasileiro – SGB.

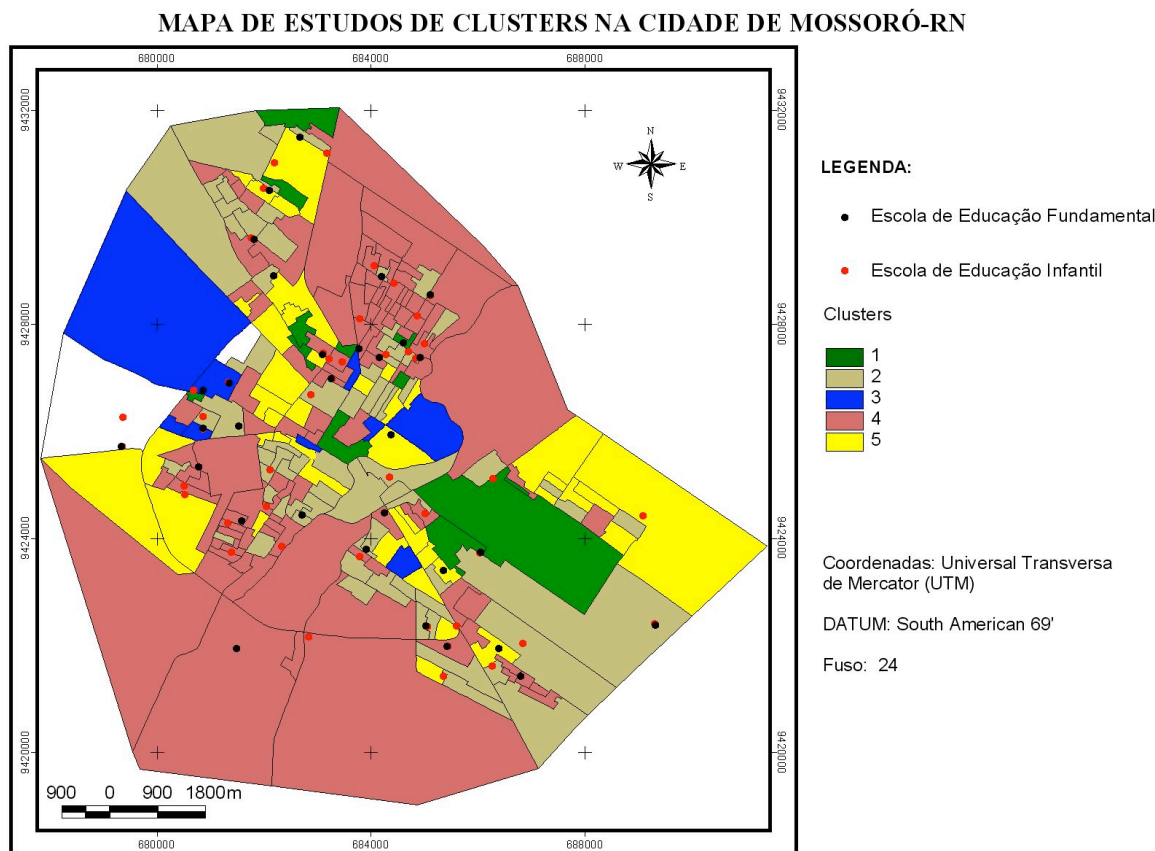


Figura 2 – Clusters dos setores censitários e localização de escolas na cidade de Mossoró/RN.

Fonte: Pesquisa dos autores (abril, 2013).

Observação: Os setores indicados com a cor branca se referem aos dois casos retirados da amostra.

A observação da distribuição geográfica dos Clusters, das unidades escolares e do contingente populacional das pessoas entre 6 e 14 anos está sintetizada na Tabela 5.

Tabela 5 - Localização das unidades escolares do ensino fundamental, por Cluster.

Clusters	Unidades de educação infantil	Escolas do ensino fundamental	Total de unidades escolares	Pessoas entre 6 e 14 anos	Razão Pessoas/unidades escolares
1	2	3	5	1.416	283,20
2	11	13	24	10.165	423,54
3	3	3	6	1.167	194,50
4	11	17	28	15.270	545,36
5	5	5	10	5.346	534,60

Fonte: Pesquisa dos autores (abril, 2013).

Nas áreas representadas pelos **Clusters 2 e 4**, está instalado o maior número de unidades educacionais, sendo também a área que detém o maior contingente populacional e as menores rendas médias mensais por pessoa. Percebe-se, então, que o investimento público foi dirigido

para o espaço demográfico que apresenta maior vulnerabilidade social, conforme pode ser observado, em comparação à tabela 4. Contudo, uma análise mais detalhada indica que 76,23% da população entre 6 e 14 anos residem nos setores censitários que formam os **Clusters 2 e 4**, nos quais estão instalados apenas 41% dos estabelecimentos educacionais.

Nas áreas representadas pelos **Clusters 1 e 3**, está instalado o menor número de escolas da rede pública, o que parece justificável por se tratar das áreas com melhor desempenho de renda e menor contingente populacional. A escolarização no *Cluster 3* alcança o percentual de 94%.

A área representada pelo **Cluster 5** detém um contingente populacional de 40.011 pessoas, sendo 5.346 delas na faixa de 6 a 14 anos e um percentual de alfabetização de 90%. A renda média de R\$ 1514,25, aliada a uma razão Pessoas entre 6 e 14 anos/Unidades escolares alta, sugere que este *Cluster* detém bom potencial para instalação de empreendimentos escolares pela iniciativa privada.

5 CONSIDERAÇÕES FINAIS

O conhecimento das características da população, aglutinadas em perfis homogêneos, indica demandas e orienta a identificação de locais para instalação de empreendimentos educacionais. Este estudo utilizou dados demográficos e socioeconômicos disponíveis no último censo demográfico realizado no Brasil, cujos indicadores foram trabalhados com uso de SIG e análise de agrupamentos, para identificar áreas com mais desenvolvimento e áreas com maiores necessidades de investimentos em equipamentos educacionais.

A aplicação da metodologia implicou segmentação dos setores censitários da área urbana do município de Mossoró/RN, obtendo-se 5 agrupamentos pela análise de *Cluster*. Esses *Clusters* georreferenciados foram utilizados para elaboração de mapa com a localização das escolas do ensino público fundamental da área urbana do município. Constatou-se que as escolas que compõem a Rede Municipal de Ensino estão razoavelmente bem distribuídas no espaço geográfico da área urbana do município de Mossoró, com localizações concentradas nas áreas de menor poder aquisitivo. No entanto, com uso apenas de dados da oferta do setor público, há espaço para ampliação de investimentos pelo poder público, pois a razão entre pessoas entre 6 a 14 anos/número de escolas ainda é bastante elevada.

Os setores censitários que formam os *Clusters 2 e 4*, por exemplo, apesar de abrigarem 76,23% das pessoas entre 6 e 14 anos, possuem apenas 41% dos estabelecimentos educacionais. Os dados sugerem que há demanda para instalação de unidades escolares da rede pública, uma vez que a razão Pessoas entre 6 e 14 anos/ Unidades escolares é maior no *Cluster 4* e muito acentuada no *Cluster 2*. Além disso, observa-se que a menor taxa de alfabetização entre pessoas de 6 a 14 anos (85%) se encontra no *Cluster 4*. Identificou-se que também há demanda para investidores do setor privado, dispostos a explorar o mercado formado pela classe C, relacionado no estudo ao *Cluster 5*, que detém renda média por pessoa igual a R\$ 1.514,25.

Constata-se que o uso conjunto da segmentação de mercado e SIG serve como fonte de informações contextualizadas, acessíveis e confiáveis sobre o perfil da demanda por unidades

educacionais. Essas técnicas podem orientar a decisão de localização de escolas públicas, além de oferecer subsídios para o investidor privado.

Este estudo apresenta a seguinte limitação: um único setor censitário pode abrigar pessoas com condições de moradia, faixa de renda e escolaridade bastante distintas. Essas diferenças não podem ser identificadas por meio dos dados dos setores censitários, cujo limite de desagregação não alcança a unidade domiciliar, mas o conjunto delas dentro de cada setor. Embora a metodologia proposta permitisse, este estudo não incorporou a oferta de ensino pelas escolas do setor privado, o que alteraria os resultados obtidos, a exemplo da razão pessoas/unidades escolares. Contudo, mesmo que parciais, os resultados demonstram a viabilidade da metodologia proposta.

A metodologia proposta permite a utilização de dados mais abrangente, com incorporação de dados de oferta e cobertura do setor privado e inclusão de outras variáveis, tais como a oferta do setor privado, a quantidade de alunos matriculados por unidade escolar e informações georreferenciadas dos endereços dos alunos, de maneira que as distâncias percorridas pelos alunos até as escolas também sejam consideradas. Também podem ser utilizadas informações com maior nível de desagregação, disponíveis, por exemplo, na Pesquisa por Orçamento Familiar (POF). Portanto, estudos posteriores devem ser desenvolvidos para refinar a metodologia proposta.

6 REFERÊNCIAS

- AHMADI-JAVID, A.; SEYEDI, P.; SYAM, S. S. A survey of healthcare facility location. **Computers & Operations Research**, London, v. 1, n. 79, p. 223-263, 2017. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0305054816301253>>. Acesso em: 02 mai. 2017.
- BATISTA, G. V.; ORTH, D. M.; BORTOLUZZI, S. D. Geoprocessamento para determinação de acessibilidade aos equipamentos educacionais como ferramenta de apoio aos Estudos de Impacto de Vizinhança: estudo de caso na Planície do Campeche - Florianópolis/SC - Brasil. *Anais XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, Curitiba, PR, Brasil. 2011. Disponível em: <<http://www.dsr.inpe.br/sbsr2011/files/p1149.pdf>>. Acesso em: 30 abr. 2011.
- BRANDÃO, M.F. R.; RAMOS, C. R. S., & TRÓCCOLI, B. T. Análise de agrupamento de escolas e Núcleos de Tecnologia Educacional: mineração na base de dados de avaliação do Programa Nacional de Informática na Educação. Rio de Janeiro/RJ, *Simpósio Brasileiro de Informática na Educação (SBIE)*, 12 a 14 de novembro de 2003. Disponível em: <<http://www.nce.ufrj.br/sbie2003/publicacoes/paper39.pdf>>. Acesso em: 16 set. 2017.
- CARVALHO, V. D. H.; BARBIRATO, J. C. C., & CIRILO, J. V. A., POLETO, T. Uma metodologia para sistemas espaciais de apoio à decisão aplicados à gestão da educação pública. *7º Congresso Luso Brasileiro para o Planejamento Urbano, Integrado e Sustentável. Contrastes, Contradições e Complexidades*. Maceió, Brasil, 05 a 7 de outubro de 2016.

- CARVALHO, W. L. **Metodologia de Análise para a Localização de Escolas em Áreas Rurais**. 2011. 215f. (Tese de Doutorado). Tese de Doutorado – Departamento de Engenharia Civil e Ambiental – Faculdade de Tecnologia – Universidade de Brasília, 2011.
- CHASCO, C. El geomarketing y la distribución comercial. **Investigación y marketing**, Madrid, n. 79, p. 6-13, 2003.
- CHUNG, K.; TCHA, D. A Fuzzy set-theoretic method for public facility location. **European Journal of Operational Research**, n. 58, p. 90-99, 1992. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0377221792902385>>. Acesso em: 02 mai. 2017.
- COSTA, A. M.; NEVES, J. A. D. 2006. Geomarketing e pequenas empresas: análise espacial dos postos de combustível de Fortaleza. In: II EMA – Encontro de Marketing da ANPAD. Anais... Rio de Janeiro-RJ.
- EISELT, H. A.; MARIANOV, V. Location modeling for municipal solid waste facilities. *Comput. Oper. Res.*, London, v. 62, p. 305-315, 2015. Disponível em: <<http://dl.acm.org/citation.cfm?id=2793621>>. Acesso em: 29 ago. 2017.
- FARAHANI, R. Z.; STEADIESEIFI, M.; ASGARI, N. Multiple criteria facility location problems: A survey. **Applied Mathematical Modelling**, London, v. 34, Issue 7, p. 1689-1709, 2010. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0307904X09003242>>. Acesso em: 02 mai. 2017.
- FARIAS, L. P. Boas perspectivas para o geomarketing em 2012 – Parte II. Disponível em: <<http://www.implantandomarketing.com/boas-perspectivas-para-o-geomarketing-em-2012-parte-ii/>>. Acesso em: 24 abr. 2017.
- FERNANDES, V.O.; MAURO, C. S. P.; ALIXANDRINI JUNIOR, J.; SANTOS, D. V. C. Aplicação conjunta de técnicas de análise multivariada de dados e análise espacial exploratória para avaliação do desempenho de transporte público. IV Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação, Recife - PE, 06 a 09 de maio de 2012, p. 01–09. Disponível em: <https://www.ufpe.br/cgtg/SIMGEOIV/CD/artigos/SIG/023_5.pdf>. Acesso em: 02 mai. 2017.
- FIGUEIREDO FILHO, D. B.; SILVA JUNIOR, J. A.; ROCHA, E. C. Classificando regimes políticos utilizando análise de conglomerados. **Opinião Pública**, Campinas, v. 18, p. 109-128, jun. 2012.
- GOMES, J. C.; LEVY, A.; LACHTERMACHER, G. (2014). Segmentação do censo educacional 2000 utilizando técnicas de mineração de dados. *SÃO João Del Rei/MG*. XXXVI SBPO, 23 a 26 de novembro de 2014. Disponível em: <<http://din.uem.br/sbpo/sbpo2004/pdf/arg0260.pdf>>. Acesso em: 16 set. 2017
- GOUDARD, B.; OLIVEIRA, F. H.; GERENTE, J. Avaliação de modelos de localização para análise da distribuição espacial de unidades básicas de saúde. Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto. **Revista Brasileira de Cartografia**, Rio de Janeiro, nº 67/1, p. 15-34, 2015.

- JANNUZZI, P. M. **Indicadores Sociais no Brasil**: conceitos, fontes de dados e aplicações. 4ª ed. Campinas: Alínea, 2009.
- KOTLER, P. **Marketing para o século XXI**: como criar, conquistar e dominar mercados. São Paulo: Futura, 1999.
- KOTLER, P. **Administração de Marketing**: Análise, Planejamento, Implementação e Controle. São Paulo: Atlas, 1999.
- LOVELOCK, C., WRIGHT, L. **Serviços**: marketing e gestão. São Paulo: Saraiva, 2002.
- MALHOTRA, N. K. **Pesquisa de Marketing**: uma orientação aplicada. Porto Alegre: Bookman, 2012.
- MALLER, R.; GANDOLPHO, A. A. Localização de escolas do ensino fundamental: Caso de Itaipava/RJ. **Revista da Universidade Católica de Petrópolis**, Petrópolis, v. 8, n. 2, p. 108-123, 2014.
- MANTOVANI, D. M. N.; GOUVÊA, M. A.; TAMASHIRO, H. R. S. Segmentação e qualidade em serviços educacionais: o caso de um curso de graduação à distância. **Revista de Administração da UNIMEP**, v. 13, n. 1, p. 91-116, jan./abr., 2015.
- MAPA, S. M. S.; LIMA, R. S. Uso Combinado de SIG para transportes e programação linear inteira mista em problemas de localização de instalações. *Revista Gestão e Produção*, São Carlos, v. 19, n. 1, 119-136, 2012.
- MARÔCO, J. **Análise estatística com o SPSS Statistics**. (5ª ed.) Pero Pinheiro: ReportNumber, 2011.
- MELO, F. C. C.; ARAUJO, J. A. R.; CRISOSTOMO, V. L.; ZARCAN, R. M. Identificação de nichos de mercado a partir da análise do orçamento familiar da população de baixa renda. **Revista ambiente contábil**, Natal, v. 9, p. 142-157, 2017. Disponível em: <<http://www.aten.org.br/revista/ojs2.2.308/index.php/Ambiente/article/viewFile/3086/2440>>. Acesso em: 12 set. 2017.
- MOURA, L. M.; IGNACIO, A. A. V. Cálculo de parâmetros para o problema de localização de escolas no município de Macaé, XVIII Simpósio de Pesquisa Operacional & Logística da Marinha, Blucher Marine Engineering Proceedings, v. 2, p. 739-750, 2016.
- NDIAYE, F.; NDIAYE, B. M.; LY, I. Application of the p-Median Problem in School Allocation. **American Journal of Operations Research**, Washington, p. 253-259, 2012. Disponível em: <http://file.scirp.org/pdf/AJOR20120200013_67444730.pdf>. Acesso em: 29 ago. 2017.
- NEVES, F. H. Planejamento de equipamentos urbanos comunitários de educação: algumas reflexões. **Cad. Metrop.**, São Paulo, v. 17, n. 34, p. 503-516, nov. 2015. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2236-99962015000200503>. Acesso em: 05 jun. 2017.
- PETER, J. P.; OLSON, J. C. **Comportamento do consumidor e estratégia de marketing**. Tradução de Beth Honorato. (8ª ed.) São Paulo: McGraw-Hill, 2009.

- PIZZOLATO, N. D.; BARROS, A. G.; BARCELOS, F. B.; CANEN, A. G. C. Localização de escolas públicas: síntese de algumas linhas de experiências no Brasil. **Pesquisa Operacional**, Rio de Janeiro, v. 24, n. 1, jan./abr. 2004.
- SANTOS, A. C. A. O. **Estudo de Localização de Escolas Públicas em Áreas Urbanas**. 2012. 103f. Dissertação de Mestrado (Universidade de Brasília), Brasília, Brasil.
- SANTOS, K. C. **Proposta conceitual de um modelo de localização de estabelecimentos de ensino**. 2005. 210f. Dissertação de mestrado (Universidade Federal de Santa Catarina Programa de Pós-Graduação em Engenharia Civil) Florianópolis - SC, Brasil.
- SCHIFFMAN, G. L.; KANUK, L. L. **Comportamento do Consumidor**. Rio de Janeiro: LTC, 2009.
- SHETH, J. N.; MITTAL, B.; NEWMAN, B. I. **Comportamento do Cliente: indo além do comportamento do consumidor**. São Paulo: Atlas, 2001.
- SILVA, P. E. A. B.; GARCIA, R. A. Integração entre dados sociodemográficos e ambientais: o Índice da Qualidade do Saneamento Ambiental obtido a partir de geotecnologias e análise multivariada. *XVIII Encontro Nacional de Estudos Populacionais, ABEP, Anais...* Águas de Lindóia-SP. 2012.
- TAGLIATTI, C. **Verificação das relações espaciais entre a destruição de dados sócio-econômicos e localização de escolas**. 2010. Dissertação (Mestrado em Engenharia Urbana), São Carlos, UFSCar, Brasil.
- VERICA, W. R.; VILLWOCK, R.; JOHANN, J. A. Uso de técnicas de mineração de dados para agrupamento e espacialização de dados educacionais no Estado do Paraná. *Anais XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, João Pessoa-PB, Brasil, 25 a 29 de abril de 2015, INPE. Disponível em: <<http://www.dsr.inpe.br/sbsr2015/files/p0337.pdf>>. Acesso em: 16 set. 2017.
- VERRAN, P. G. **Análise da acessibilidade a equipamentos públicos de educação em áreas de expansão urbana utilizando técnicas de geoprocessamento: bacia hidrográfica do Arroio do Salso em Porto Alegre/RS**. 2014. 55f. Monografia (Bacharelado em Geografia). Universidade Federal do Rio Grande do Sul. Porto Alegre, RS, Brasil. Disponível em: <<http://www.lume.ufrgs.br/handle/10183/115406>>. Acesso em: 06 jun. 2017.