

# O USO DO SIG NA ANÁLISE DA POPULAÇÃO COM ACESSO A COMPUTADOR, COM E SEM INTERNET NA REGIÃO METROPOLITANA DE BELÉM-PARÁ.

*PORFIRIO, B.K.A<sup>1</sup>; CUNHA, F.F<sup>1</sup>; SILVA, T.R.R<sup>1</sup>; SANTOS, T.C.M<sup>1</sup>;*

<sup>1</sup>Universidade Federal Rural da Amazônia, Belém-Pa, Brasil

## RESUMO

No que diz respeito a buscar formas de incluir a sociedade digitalmente, os dados estatísticos são indispensáveis, neste sentido, torna-se relevante a utilização das tecnologias de mapeamento digital e mineração de dados apoiadas em Sistemas de Informações Geográficas (SIG), a fim de auxiliar nas tomadas de decisões, na formulação e monitoramento de políticas governamentais na área da Região Metropolitana de Belém (RMB). Este estudo teve a finalidade de gerar dados quantitativos e qualitativos através de mapas temáticos utilizando o software ArcGIS 10.2, de acordo com a base de dados do IBGE, Censo 2010 e Atlas Brasil 2013, com o auxílio do software R para gerar inferências probabilísticas, demonstrando o índice de pessoas que usam computador com ou sem internet, na Região Metropolitana de Belém. Portanto, a visão estatística, geográfica e social dos dados, é de suma importância para esta análise, possibilitando a descoberta de conhecimentos ou a confirmação de hipóteses de cunho analítico sobre problemáticas referentes ao contexto escolhido.

**Palavras chave:** SIG, Arcgis, mapas

## ABSTRACT

With regard to the search for ways of depositing a society digitally, statistical data are indispensable, in this sense, the use of digital mapping and data mining technologies supported by Geographic Information Systems (GIS) is becoming relevant, in order to assist in decision-making, formulation and monitoring of government policies in the Belém Metropolitan Region (RMB). This study aimed to generate quantitative and qualitative data through thematic maps using ArcGIS 10.2 software, according to the IBGE database, Census 2010 and Atlas Brasil 2013, with the aid of software R to generate probability inferences, demonstrating the index of people using computers with or without internet in the Metropolitan Region of Belém. Therefore, the statistical, geographic and social view of the data is of paramount importance for this analysis, enabling a discovery of knowledge or a confirmation of hypotheses analytic about the problematic of the chosen context.

**Keywords:** GIS, Arcgis, maps

### 1- INTRODUÇÃO

Com o advento da internet, atualmente, o grande desafio é a ampliação do acesso à rede para aqueles que ainda estão excluídos digitalmente. No que diz respeito a buscar formas de incluir a sociedade digitalmente, os dados estatísticos são indispensáveis, uma vez que, os governantes podem ter um panorama da situação atual do país/estado/município em diferentes áreas.

Neste sentido, torna-se relevante, a utilização das tecnologias de mapeamento digital e mineração de dados apoiadas em Sistemas de Informações Geográficas (SIG), que conforme Fitz (2008), “é um sistema constituído por um conjunto de programas computacionais, o qual integra dados, equipamentos e

pessoas com objetivo de coletar, armazenar, recuperar, manipular, visualizar e analisar dados espacialmente referenciados a um sistema de coordenadas conhecido”. Com isso, servirá para fazer a aquisição de dados estatísticos e especializados, seleção, processamento, transformação, mineração, avaliação e geração de mapas temáticos, a fim de auxiliar nas tomadas de decisão, na formulação e monitoramento de políticas governamentais dentro da área da Região Metropolitana de Belém (RMB).

No contexto deste trabalho, o interesse pela RMB pode ser destacado em razão das enormes diferenças entre a posse de computador com Internet nas metrópoles e a posse em outros municípios, principalmente aqueles localizados nos municípios de interior das regiões Norte e Nordeste, reconhecidamente

as regiões com os piores percentuais de posse de computador e acesso à Internet no país.

Segundo o IPEA (2017), a Região Metropolitana de Belém foi criada em 1973 pela Lei Complementar Federal Nº. 14/73, que é composta por sete municípios: Ananindeua, Belém, Castanhal, Marituba, Santa Bárbara do Pará e Santa Izabel do Pará e possui área total de 3.566 Km<sup>2</sup>. Apesar da região metropolitana de Belém possuir só 2% do espaço territorial do estado do Pará, o mesmo concentra aproximadamente 40% da riqueza e 35% da população do estado em 2011. Sabendo-se que mesmo a RMB tendo aumento no número de pessoas com acesso à internet, o estado do Pará ainda possui uma das menores margens do país.

Dessa forma, este estudo teve a finalidade de gerar dados quantitativos e qualitativos através de mapas utilizando o software ArcGIS 10.2, de acordo com a base de dados do IBGE, Censo 2010 e Atlas Brasil 2013, com o auxílio do software R para gerar inferências probabilísticas, demonstrando o índice de pessoas que usam computador com ou sem internet, na Região Metropolitana de Belém. Destacando a importância da ferramenta do SIG, a mineração de dados e seus produtos no auxílio de decisões, no fornecimento de informações mais rápidas e precisas.

## 2- METODOLOGIA

Para estudo da população com acesso a computador, com e sem internet, foi adotada uma abordagem que associa a exclusão digital, mineração de dados e sistemas de informações geográficas. Para isso, a pesquisa iniciou com uma investigação sobre exclusão digital e as técnicas de mineração de dados, buscando entender seus conceitos, aplicações, suas etapas e modo de utilização, principalmente sua relação com a análise de dados estatísticos e geográficos.

Em seguida é realizada uma pesquisa sobre Sistemas de Informações Geográficas (SIG), que segundo Worboys (1995), “caracteriza os sistemas de informação que tornam possível a captura, modelagem, manipulação, recuperação, análise e apresentação de dados referenciados geograficamente (ou dados georreferenciados). De forma geral, um software de SIG é um sistema composto de quatro grandes componentes: componentes de captura de dados, armazenamento, análise e componentes de apresentação dos dados”. E linguagem de programação R, que conforme Loretto (2015), “R é uma linguagem e um ambiente de desenvolvimento voltado principalmente para computação estatística (inferência, simulações, data mining, etc.) e gráficos”. Buscando entender as técnicas estatísticas e espaciais, o raciocínio e as análises das incertezas, além da entrega de informações através de mapas temáticos.

Como fonte de dados bibliográficos foram utilizadas pesquisas de diversos autores, dentre eles,

Michael Worboys, Paulo Fitz, Silvana Rossy, Aleksandra Silva, Marcelo Neri e alguns tutoriais, todos sendo levados em consideração na análise e elaboração do mapa temático e das inferências estatísticas.

A base de dados do IBGE, Censo 2010, foi utilizada para alcançar informações a respeito da posse de computadores com ou sem acesso à Internet, nos domicílios. Assim, foi possível definir as estratégias para analisar os dados obtidos, tendo como foco a Região Metropolitana de Belém. Para alcançar o objetivo de incluir a relação entre a posse de computador com ou sem acesso à Internet com outras variáveis, foi utilizado o Atlas Brasil 2013 de onde foram coletados indicadores de IDHM, PIB e POPULACAO de cada município.

Logo após de definir as fontes bibliográficas e de dados que serão utilizados, são apontadas as ferramentas de software que a partir das quais os resultados são extraídos para análise. As características da posse de computador com ou sem acesso à internet são apresentadas e discutidas, baseando-se em técnicas de geoprocessamento para produção de mapas temáticos, através dos softwares ArcGIS versão 10.2 com sua ferramenta ArcMAP. Além de gerar inferências probabilísticas através do gráfico de correlação da linguagem R, através do software RStudio.

Dessa forma, através das variáveis: POPULACAO, PIB (Produto Interno Bruto), IDHM (Índice de Desenvolvimento Humano Municipal), pessoas com acesso a computador com e sem internet, utilizamos as seguintes etapas: Aquisição das bases de dados do IBGE, Censo 2010 e Atlas Brasil 2013, aplicados sobre o software ArcGIS 10.2 e sua ferramenta ArcMAP; a limpeza desses dados através do Excel, gerando assim um banco de dados; mineração de dados, que é o processo da análise e exploração das grandes quantidades de dados à procura de padrões consistentes, que foi realizado através do software RStudio; e a projeção de mapas temáticos, onde implica o reflexo do acesso de internet das pessoas, através dos resultados obtidos.

Portanto, este trabalho tem o intuito de analisar os dados do Censo 2010 (SIDRA, 2017), utilizando tecnologias de mapeamento da distribuição espacial e mineração de dados, adotando a metodologia proposta e utilizada em trabalhos prévios (BRITO et al. 2016 e SILVA et al., 2016), que combina a estratégia de busca de associação entre variáveis com o uso de Sistemas de Informação Geográfica (SIG), procurando identificar as diferenças entre os municípios da RMB.

## 3- RESULTADOS

A partir da análise dos dados, foram gerados dois mapas pelo software Arcgis e sua ferramenta Arcmaps, onde foi possível obter representações dinâmicas, na qual refere-se especificamente à manipulação interativa da informação espacial, no qual

certos fatores influenciam na demanda de acesso aos microcomputadores com e sem internet como a POPULACAO, IDHM e PIB.

(MC) e posse de Microcomputador com Internet (MC\_Internet), correlacionadas com as variáveis de PIB e IDHM.

### 3.1 – MAPAS TEMÁTICOS

A noção da desigualdade digital depende, portanto da compreensão dos conceitos de inclusão e de exclusão. Para quantificar essa desigualdade, autores usam métodos quantitativos, expressos através de indicadores e estatísticas. Essa tem sido a abordagem, por exemplo, utilizada no Mapa de Inclusão Digital (NERI, 2003).

Para que pudesse ter uma visão espacial dos dados, foram produzidos dois mapas temáticos considerando as variáveis posse de Microcomputador

O primeiro passo foi a criação de um mapa temático abordando a posse de microcomputador nos municípios da RMB. A Fig. 1 mostra uma clara tendência espacial nos municípios através de graduação de cores onde a cor vermelha representa os municípios de menor faixa de posse de microcomputador e a cor verde escuro a de maior faixa. Posteriormente, foi criado um gráfico da correlação com os fatores que influenciam diretamente a posse de microcomputador de um município.

Logo depois, foi criado um mapa temático abordando a posse de Microcomputador com Internet (Fig. 2) nos municípios da RMB, com a mesma distribuição espacial através da graduação de cores.

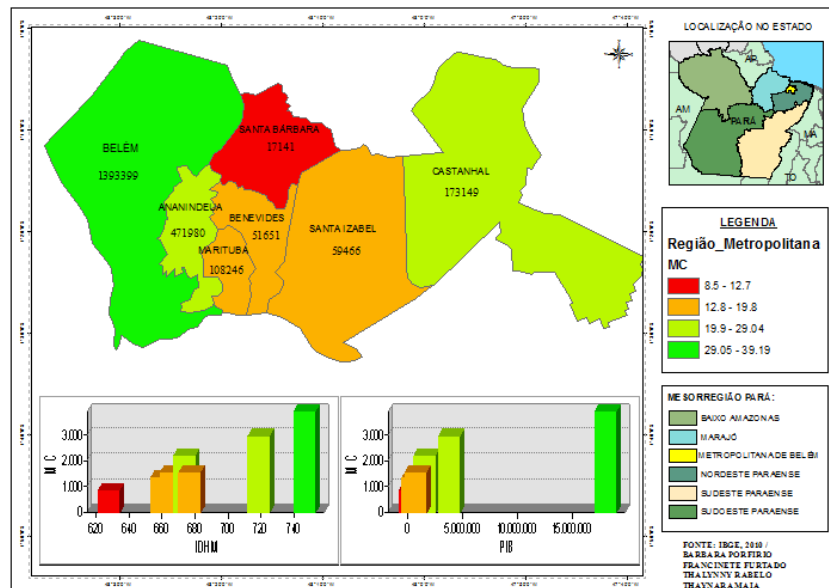


Fig. 1: Distribuição Espacial – municípios da RMB (MC e outros indicadores)

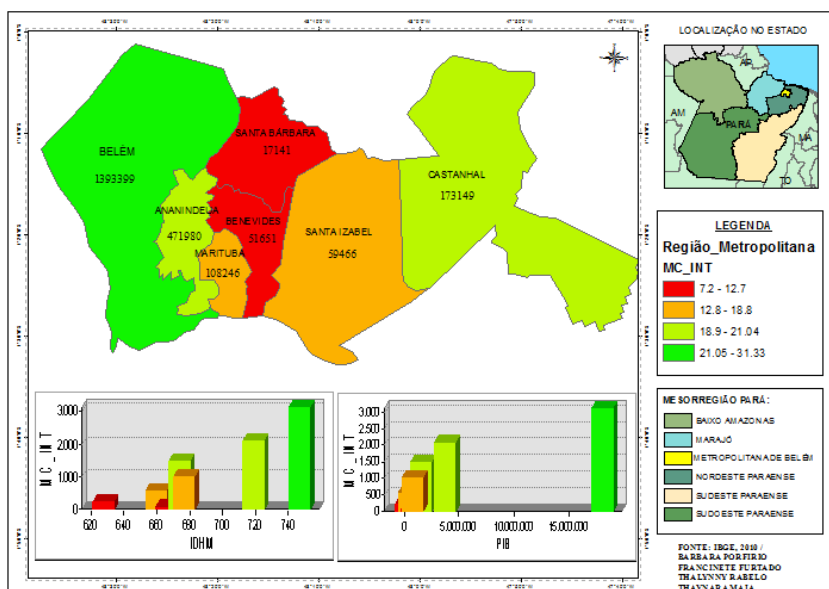


Fig. 2: Distribuição Espacial – municípios da RMB (MC\_Internet e outros indicadores)

De acordo com os mapas temáticos podemos perceber de modo geral que em Ananindeua 21,04% tem acesso à Internet e 29,67% não têm acesso à Internet. Em Belém 31,33% tem acesso e 39,19% não tem acesso. Em Benevides foi 7,2% e 15,57% com e sem acesso à Internet, respectivamente.

Na cidade de Castanhal foram 14,78% que tem acesso e 22,07% que não tem acesso à Internet. Em Marituba o resultado foi de 9,98% com acesso à Internet e 15,45% sem acesso da mesma. Em Santa Bárbara do Pará foi de 2,27% e 8,61% com e sem acesso à Internet, respectivamente. E por fim, Santa Isabel do Pará o resultado foi de 5,75% para os que têm acesso e 13,43% para os que não têm acesso à Internet.

### 3.2 - GRÁFICO DE CORRELAÇÃO NO R

Foi realizado no trabalho para obter coeficientes de correlação no software R entre duas variáveis, o método de Pearson. Que reflete a intensidade de uma relação linear entre duas variáveis, ou seja, é um índice que mede o quanto estão relacionadas as variáveis, com isso quanto mais próximo de 1 (um) mais forte é essa relação. Em princípio, como exemplo, para verificar a existência de alguma relação entre MC\_Internet e IDHM, construiu-se um Diagrama de Dispersão, com o acréscimo da linha de tendência para as duas variáveis, como mostrado na Fig. 3.

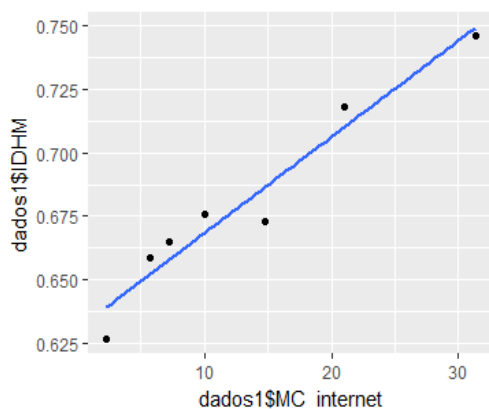


Fig. 3 - Diagrama de dispersão das variáveis MC\_Internet e IDHM

Por conseguinte, é possível obter o Coeficiente de Correlação Linear de Pearson entre as variáveis, esses coeficientes de correlação variam de -1 a 1. Portanto, foi obtido através do software R o coeficiente de correlação entre as variáveis MC\_Internet e IDHM de 0,9697869, um valor estatisticamente bem próximo ao 1(um), demonstrando assim, que houve uma correlação perfeita positiva entre as duas variáveis, ou seja, quanto maior o IDHM do município, maior será o valor da taxa de acesso à internet, nesse caso são variáveis diretamente proporcionais. É demonstrado na Fig. 4 todos os coeficientes de correlação de Pearson da base de dados obtida.

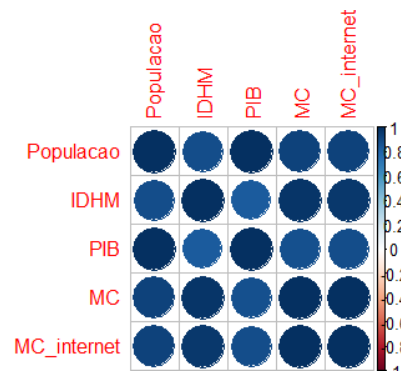


Fig. 4 – Correlações das variáveis

De acordo com a Fig. 4, quanto mais forte for o tom de azul das elipses, mais forte são as correlações, mais próximo do índice 1(um) estão. Dessa forma, podemos observar que as correlações entre a maioria das variáveis são consideradas forte, pois os coeficientes são bem próximos a 1 (um) e positivos, ou seja, são variáveis diretamente proporcionais.

### 4- DISCUSSÕES

De acordo com o IBGE (2010), a população da Região Metropolitana de Belém chega a 2.275.032 de habitantes, sendo considerada a segunda região metropolitana mais populosa da Região Norte do Brasil. Como observou Brito et al. (2016), o número de habitantes é um aspecto importante na influência da posse de computador com Internet. Para os autores, são os municípios com o maior número de habitantes os que apresentam os maiores percentuais de posse de computador com acesso à Internet — ou seja, são as metrópoles e os municípios de grande porte, considerados aqueles com mais de cem mil habitantes, que apresentam os maiores percentuais de computador com Internet.

Pode-se inferir então que Belém como município-sede possui 39% da sua população que tem acesso ao computador e 31% tem acesso ao computador e a Internet, sendo assim, representa aproximadamente em relação a população da RMB, mais de 887 mil de pessoas incluídas digitalmente.

Segundo Neri (2012), “As razões para os ‘sem rede’ diferem de lugar para lugar, revelando a importância de políticas públicas ajustadas à realidade local”.

Outras inferências que podem ser encontradas são em relação ao PIB e o IDHM, mesmo sendo considerados altos não significa totalmente que o município vai possuir maior acesso da população à internet.

Pois foram encontrados municípios com altos índices no PIB e no IDHM, e os mesmos não possuem índices com alto acesso à internet, por exemplo, o município de Castanhal, apesar de possuir o terceiro

maior PIB e o quarto maior IDHM da RMB, possui apenas 15% da sua população com acesso à internet, ou seja, 25.972 mil pessoas incluídas digitalmente.

As menores faixas de acesso à internet (faixa menor de 10%), concentram-se nos municípios Benevides, Marituba, Santa Izabel e Santa Bárbara. Somando esses menores índices de taxa de acesso à internet (25%) não chega ao índice do município-sede Belém (31%). Sabendo-se que a população de Belém chega a 1.393.399 de habitantes, esses 25% representam 348 mil de pessoas incluídas digitalmente. Em relação a RMB representa 553 mil pessoas incluídas digitalmente.

Os municípios de Benevides, Santa Izabel e Santa Bárbara também correspondem às menores faixas de desenvolvimento humano da RMB. Portanto, o desenvolvimento da renda, educação e longevidade tem uma grande interferência nessa pequena taxa de acesso à internet.

Sabendo-se que o município de Belém possui 62% da população da RMB, apenas 542.645 mil pessoas são inclusas digitalmente e 168.219 tem acesso à internet, ou seja, 850.753 mil são excluídos digitalmente. Portanto, o município-sede Belém é onde concentra-se o maior acesso de posse de computador com e sem internet comparando-se a população total da RMB.

Logo, podemos perceber que apesar de Belém ter a maior concentração da posse desses recursos, mesmo assim não chega atingir nem 50% da população no município de acesso a microcomputadores com internet.

Apesar dos outros seis municípios da RMB possuírem PIB e IDHM consideravelmente numa faixa de nível médio para alto, podemos perceber que o acesso a posse de computador com e sem internet são baixos.

## 5- CONSIDERAÇÕES FINAIS

A desigualdade que caracteriza o país também se traduz no acesso à internet, excluindo alguns grupos dependendo de suas características como o local, renda, educação, entre outras.

Dessa forma, os possíveis fatores que têm um alto grau de influência no baixo acesso à internet nos municípios da RMB são: localizado na zona rural, baixa renda, classes sociais D ou E, o baixo nível de escolaridade, fraco suporte a infraestruturas como escolas e centros públicos gratuitos e está localizado na região Norte.

As possíveis causas da ausência de microcomputadores com internet, pode ser pela expansão da área urbana que facilita a infraestrutura para a instalação de redes de internet, ou seja, as localidades com menor área urbana, tendem a possuir pouco acesso à internet.

Portanto, a visão geográfica e social dos dados, é de suma importância para esta análise, possibilitando a descoberta de conhecimentos ou a confirmação de hipóteses de cunho analítico sobre problemáticas referentes ao contexto escolhido, auxiliando nas soluções e políticas públicas que venham a ser implementadas.

## REFERÊNCIAS BIBLIOGRÁFICAS

AMARAL, G. D.; Silva, V. L.; REIS, E. A.; Análise de Regressão Linear no Pacote R. Disponível em: <http://www.est.ufmg.br/portal/arquivos/rts/RT-SE-2009.pdf>. Acesso em: 28/09/2017.

ATLAS. Índice de Desenvolvimento Humano do Brasil. Disponível em: [http://www.atlasbrasil.org.br/2013/pt/perfil\\_uf/para#caracterizacao](http://www.atlasbrasil.org.br/2013/pt/perfil_uf/para#caracterizacao)>. Acesso em: 15 de agosto de 2017.

BRITO, S. R. et al. Gravidez na adolescência e o acesso às Tecnologias de Informação e Comunicação na Amazônia. Mundo Amazônico, v. 6, n. 2, 2016.

FITZ, P. R. Cartografia básica. São Paulo: Oficina de textos, 2008. 143p.

IBGE. Relatório da Divisão de Contabilidade: Relatório de Gestão. Disponível em: <http://www.ibge.gov.br/home/presidencia/noticias/relatgestao.shtm>>. Acesso em: 25 de agosto de 2017.

IPEA. Região Metropolitana de Belém. Disponível em: [http://ipea.gov.br/agencia/images/stories/PDFs/livros/livros/141125\\_atlas\\_belem](http://ipea.gov.br/agencia/images/stories/PDFs/livros/livros/141125_atlas_belem)>. Acesso em: 01 de setembro de 2017.

LAURETTO, M. S. Introdução à Análise de Dados Utilizando o Ambiente R. Disponível em: <http://each.uspnet.usp.br/lauretto/cursoR2015/cursoR2015.pdf>. Acesso em: 28/09/2017.

NERI, M. C. Mapa da Inclusão Digital. Rio de Janeiro: FGV, CPS, 2012.

NERI, M. O Mapa da Exclusão Digital. Revista Conjuntura Econômica, v. 57, n. 5, p. 70-73, 2003.

SIDRA. Sistema IBGE de Recuperação Automática. Disponível em: <http://www2.sidra.ibge.gov.br/bda/tabela/listabl.asp?z=t&o=1&i=P&e=l&c=2249>>. Acesso em: 01 de setembro de 2017.

SILVA, A. S. et al. Gravidez na adolescência e associação com indicadores de renda, educação e acesso às Tecnologias de Informação e Comunicação no Brasil e na Amazônia Legal Brasileira. Saúde e Desenvolvimento Humano. ResearchGate, v. 4, n. 2, p. 21-33, nov. 2016.

WORBOYS, M.F. GIS: A Computing Perspective. 1. ed. London: CRC Press, 1995.