



ALTERNATIVAS NA RENDERIZAÇÃO AUTOMÁTICA DE CARTOGRAMAS TEMÁTICOS

Luiz Paulo do Nascimento

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientador: Ricardo Cordeiro de Farias

Rio de Janeiro
Junho de 2025

ALTERNATIVAS NA RENDERIZAÇÃO AUTOMÁTICA DE CARTOGRAMAS
TEMÁTICOS

Luiz Paulo do Nascimento

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Orientador: Ricardo Cordeiro de Farias

Aprovada por: Prof. Claudio Esperança

Prof. Luiz Arthur Silva de Faria

Prof. Luiz Carlos Agner Caldas

Prof. Helga Szpiz

Prof. Esteban Walter Gonzalez Clua

RIO DE JANEIRO, RJ – BRASIL

JUNHO DE 2025

do Nascimento, Luiz Paulo

Alternativas na renderização automática de cartogramas temáticos/Luiz Paulo do Nascimento. – Rio de Janeiro: UFRJ/COPPE, 2025.

XIII, 85 p.: il.; 29, 7cm.

Orientador: Ricardo Cordeiro de Farias

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2025.

Referências Bibliográficas: p. 79 – 85.

1. cartogramas. 2. mapas temáticos. 3. visualização de dados. I. Cordeiro de Farias, Ricardo. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

*A minha família e a Ana Paula,
que suportaram minhas
ausências e inquietações
durante a elaboração deste
trabalho.*

Agradecimentos

A meu orientador Ricardo Cordeiro de Farias, à equipe do PESC-COPPE, e aos colegas do IBGE que me incentivaram e deram suporte nesta trajetória acadêmica.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

ALTERNATIVAS NA RENDERIZAÇÃO AUTOMÁTICA DE CARTOGRAMAS TEMÁTICOS

Luiz Paulo do Nascimento

Junho/2025

Orientador: Ricardo Cordeiro de Farias

Programa: Engenharia de Sistemas e Computação

Apresentam-se, nesta tese, conceitos alternativos para a renderização de cartogramas temáticos, utilizando formas primitivas dispostas sob diferentes conceitos de distribuição espacial, visando atender de forma mais intuitiva aos critérios explicitados na proposta inicialmente defendida por BARRETO *et al.* (2018). São também apresentados algoritmos de renderização gráfica, desenvolvidos para a geração automática de cartogramas temáticos conforme os conceitos aqui apresentados, e seguindo a metodologia do trabalho de SPECKMANN e VERBEEK (2010).

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

ALTERNATIVES FOR THEMATIC CARTOGRAMS'S AUTOMATIC RENDERING

Luiz Paulo do Nascimento

June/2025

Advisor: Ricardo Cordeiro de Farias

Department: Systems Engineering and Computer Science

This thesis presents alternative concepts for rendering thematic cartograms, using primitive forms arranged under different concepts of spatial distribution, with the aim of meeting, in a more intuitive way, the criteria explained at the initially defended proposal by BARRETO *et al.* (2018). Graphic rendering algorithms, developed for the automatic generation of thematic cartograms according to the concepts presented here, and following the methodology of SPECKMANN e VERBEEK (2010)'s work, are also presented.

Sumário

Lista de Figuras	ix
Lista de Abreviaturas	xiii
1 Introdução e Contexto	1
2 Motivação para um mapa temático do Brasil	14
3 Conceitos e algoritmos rumo a um mapa temático do Brasil	27
3.1 Mapas de Colar	27
3.2 Simplificação Automatizada de Formas Geográficas	38
3.3 Grade Estatística	42
4 Novas propostas para um mapa temático do Brasil	51
4.1 Carto-centric	51
4.2 Carto-step	58
4.3 Step-centric	63
5 Implementações iniciais	66
6 Implementações para renderização automatizada de cartogramas	71
7 Trabalhos Futuros e Considerações Finais	77
Referências Bibliográficas	79

Listas de Figuras

1.1	Exemplo de mapa de pontos. Densidade populacional de Illinois, 2010. Extraído de https://commons.wikimedia.org/wiki/File:Population_Density_of_Illinois_2010_Wikipediamap.svg	7
1.2	Exemplo de mapa de símbolos proporcionais. Consumo de Petróleo nos EUA. Extraído de Leistiar / Ty. https://www.flickr.com/photos/89769525@N08/8242929652/	8
1.3	Mapa coroplético. População dos EUA por milha quadrada por estado (números do censo de 2000). Extraído de https://commons.wikimedia.org/wiki/File:US_2000_census_population_density_map_by_state.svg	9
1.4	Cartogramas de áreas contíguas. Imagem de um mapa mundial da agricultura orgânica. (PAULL, J. e HENNIG, B. 2011). Extraído de https://commons.wikimedia.org/wiki/File:Paul1Hennig2016WorldMap.OAha.CC-BY-4.0.jpg	10
1.5	Exemplo de cartograma de Dorling. Protovis. Extraído de http://mbostock.github.io/protovis/ex/cartogram.html	10
2.1	Mapa do Brasil, com destaque para os maiores e menores estados. Extraído de BARRETO <i>et al.</i> (2018)	15
2.2	Primeiro layout de mapas temáticos usando hexágonos. Extraído de BARRETO <i>et al.</i> (2018)	17
2.3	Mapa geográfico do Brasil. Extraído de BARRETO <i>et al.</i> (2018)	17
2.4	Segundo layout de mapas temáticos usando hexágonos. Extraído de BARRETO <i>et al.</i> (2018)	18
2.5	Relações de vizinhança de arestas e vértices em ladrilhamentos hexagonais do plano. Extraído de BARRETO <i>et al.</i> (2018)	19
2.6	Relações de vizinhança de arestas e vértices em ladrilhamentos triangulares do plano. Extraído de BARRETO <i>et al.</i> (2018)	19
2.7	Layout do mapa temático utilizando triângulos com unidades distribuídas por Grandes Regiões. Extraído de BARRETO <i>et al.</i> (2018)	20
2.8	Mapa geográfico do Brasil. Extraído de BARRETO <i>et al.</i> (2018)	20

2.9	Layout do mapa temático usando triângulos. Extraído de BARRETO <i>et al.</i> (2018)	21
2.10	Mapa temático do Brasil com hexágonos contíguos. Extraído de BARRETO <i>et al.</i> (2018)	22
2.11	Mapa temático do Brasil com hexágonos não-contíguos organizados por Grandes Regiões. Extraído de BARRETO <i>et al.</i> (2018)	22
2.12	Mapa temático do Brasil com hexágonos não-contíguos organizados por Grandes Regiões, com maior espaçamento. Extraído de BARRETO <i>et al.</i> (2018)	23
2.13	Grafo dos estados brasileiros posicionados de acordo com a distância entre as capitais dos estados. Extraído de BARRETO <i>et al.</i> (2018)	24
2.14	Grafo dos estados brasileiros posicionados em uma interpolação de 50 por cento entre a distância real e um valor fixo das capitais dos estados. Extraído de BARRETO <i>et al.</i> (2018)	24
2.15	Grafo dos estados brasileiros posicionados segundo um valor fixo de distância entre as capitais dos estados. Extraído de BARRETO <i>et al.</i> (2018)	25
2.16	Mapas temáticos do Brasil resultantes, com escalas monocromáticas representando notas médias do ENEM por Unidade da Federação. Extraído de BARRETO <i>et al.</i> (2018)	26
3.1	Exemplos de "necklace maps". Extraído de SPECKMANN e VERBEEK (2010)	28
3.2	À esquerda, intervalos de região sobre um colar; à direita, consumo de energia proveniente de recursos renováveis em 2007. Extraído de SPECKMANN e VERBEEK (2010)	29
3.3	À esquerda, intervalos por centroides; à direita, intervalos por margens. Extraído de SPECKMANN e VERBEEK (2010)	32
3.4	Intervalos dependentes da densidade. Extraído de SPECKMANN e VERBEEK (2010)	33
3.5	Raio de cobertura. À esquerda, sobre o círculo; à direita, sobre a curva spline. Extraído de SPECKMANN e VERBEEK (2010)	34
3.6	À esquerda, tamanhos de símbolos otimizados com intervalos por centroides; à direita, tamanhos de símbolo otimizados com intervalos por margem. Extraído de SPECKMANN e VERBEEK (2010)	36
3.7	À esquerda, símbolos com buffers empurrados para o meio; à direita, símbolos com buffers afastados uns dos outros. Extraído de SPECKMANN e VERBEEK (2010)	37

3.8	Exemplos de evolução discreta de curva; cada linha mostra apenas alguns estágios. Extraído de BARKOWSKY <i>et al.</i> (2000).	40
3.9	Exemplos de casos em dados cartográficos que requerem a extensão do método de evolução discreta de curva. Extraído de BARKOWSKY <i>et al.</i> (2000).	40
3.10	Captura de tela da aplicação Grade Estatística 2022, disponível em https://mapasinterativos.ibge.gov.br/grade2022/default.html	43
3.11	Grade Estatística oficial do Brasil no nível L0 (composta por células que apresentam 500 km de lado), disponível em https://dadosabertos.social/t/que-tal-o-uso-mais-amplo-da-nossa-grade-estatistica-oficial/317	46
3.12	Estrutura hierárquica de células. Extraído de BUENO (2014).	48
4.1	Níveis de concentricidade com suas respectivas Unidades da Federação.	52
4.2	Conjunto de tuplas para cartograma de concentricidade relacionado às Unidades da Federação.	54
4.3	Cartograma de concentricidade (carto-centric) representando o Brasil e as Unidades da Federação.	55
4.4	Cartograma de concentricidade (carto-centric) representando os municípios do estado do Rio de Janeiro, com as regiões de governo identificadas por cores distintas e o centro do cartograma sobre a capital.	56
4.5	Cartograma de concentricidade (carto-centric) representando os municípios do estado do Rio de Janeiro, com as regiões de governo identificadas por cores distintas e o centro do cartograma sobre o município de Cachoeiras de Macacu.	57
4.6	Colunas de líquido iluminadas por LED. Extraído de http://www.fsbaojiali.cn/	58
4.7	Sugestões de escalas para formas concéntricas básicas.	59
4.8	Conjunto de tuplas para o cartograma de degraus referente às Unidades da Federação.	60
4.9	Cartograma de degrau (carto-step) representando o Brasil e as Unidades da Federação.	61
4.10	Cartograma de degrau (carto-step) aplicado separadamente às Grandes Regiões brasileiras.	62
4.11	Cartograma de degrau (carto-step) representando os municípios do estado do Rio de Janeiro, com as regiões de governo identificadas por cores distintas.	63

4.12 Conjunto de tuplas para o cartograma 'step-centric' referente às Unidades da Federação.	64
4.13 Cartograma 'step-centric' representando o Brasil e as Unidades da Federação, utilizando os valores de posicionamento produzidos para a tupla correspondente ao cartograma de concentricidade (carto-centric) e os valores de dimensionamento produzidos para a tupla correspondente ao cartograma de degrau (carto-step).	65
5.1 Carto-centric apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos às Unidades da Federação.	67
5.2 Carto-centric apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos aos municípios do Estado do Rio de Janeiro, com centro do cartograma sobre o município de Cachoeiras de Macacu.	67
5.3 Carto-centric apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos aos municípios do Estado do Rio de Janeiro, com centro do cartograma sobre a capital do Estado.	67
5.4 Carto-step apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos às Unidades da Federação.	68
5.5 Carto-step apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos aos municípios do Estado do Rio de Janeiro. . . .	68
5.6 Carto-step tridimensional apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos às Unidades da Federação. . . .	69
6.1 Formato de tabela bidimensional utilizada pelo KNIME para intercâmbio de dados entre os módulos do fluxograma. Elaborado pelo autor.	72
6.2 Fluxograma do KNIME para a proposta carto-centric. Elaborado pelo autor.	73
6.3 Fluxograma do KNIME para a proposta carto-step. Elaborado pelo autor.	74
6.4 Tela capturada do KNIME para a proposta carto-centric. Elaborado pelo autor.	75
6.5 Tela capturada do KNIME para a proposta carto-step. Elaborado pelo autor.	75

Lista de Abreviaturas

API	Application Programming Interface.....	69
CSS	Cascading Style Sheets.....	66
ENEM	Exame Nacional do Ensino Médio	15
GPGPU	Unidade de Processamento Gráfico de Propósito Geral.....	77
GPU	Unidade de Processamento Gráfico	77
HTML	HyperText Markup Language	66
IBGE	Instituto Brasileiro de Geografia e Estatística.....	4
MAUP	Problema da Unidade de Área Modificável.....	44
PNG	Portable Network Graphics.....	75
SVG	Scalable Vector Graphics.....	75
URL	Universal Resource Locator	66
WKT	Well-Known Text Format	72

Capítulo 1

Introdução e Contexto

A visualização de dados é um problema de pesquisa relevante; quando os dados visualizados são georreferenciados, ou seja, estão associados a uma localização conhecida no espaço geográfico, o resultado é um mapa.

Segundo PINHEIRO e ESPERANÇA (2018), mapas são abstrações de um mundo tridimensional representado na superfície bidimensional de uma folha de papel ou de uma tela de vídeo. Como tal, não pretendem ser vistos como realidade, mas como representações ou projeções de pontos ou regiões do globo sobre um plano. Embora a maioria dos mapas afirme oferecer uma imagem precisa e objetiva do mundo, cada um é construído com base em convenções arbitrárias que atendem às necessidades e desejos de seus criadores.

Existem muitos tipos de mapas temáticos, incluindo "mapas coropléticos", "cartogramas", "mapas de símbolos proporcionais" e "mapas de pontos", entre os mais utilizados. Todos os tipos apresentam vantagens e desvantagens dependendo da região a ser coberta. A maioria deles, quando aplicados ao território brasileiro, apresentará limitações e distorções. Um mapa temático para os estados brasileiros, em particular, apresenta desafios interessantes, uma vez que os estados são fortemente diversificados em área, desenvolvimento econômico e densidade populacional. Por exemplo, um "mapa coroplético" onde as áreas geográficas são retratadas de forma realista pode erroneamente equiparar a importância visual de um estado com o tamanho de sua população ou significado econômico. Da mesma forma, "cartogramas" usando deformação podem levar ao desaparecimento visual de pequenas áreas, "mapas de símbolos proporcionais" podem gerar símbolos sobrepostos em pequenas regiões, e "mapas de pontos" podem impossibilitar o cálculo de números reais, ou ainda fornecer uma falsa impressão de vazio. Além disso, a deformação do mapa pode prejudicar a comparação entre mapas que mostram diferentes temas para a mesma região.

Considerando que a maioria dos mapas temáticos foi criada e pesquisada com mais frequência em países europeus e norte-americanos, é razoável supor que esses

tipos de mapas foram projetados de acordo com os requisitos e necessidades desses países. Embora essa hipótese possa ser investigada, digamos, conduzindo um grande levantamento de exemplos produzidos durante muitos anos, uma abordagem mais prática foi escolhida por BARRETO *et al.* (2018), qual seja: desenhar um mapa temático do Brasil adequado para apresentar dados referentes às unidades federativas brasileiras. Assim, a presente tese seguirá tal abordagem, na tentativa de analisar as necessidades especiais de um mapa temático do Brasil por estados, em face de suas peculiaridades geográficas.

Conforme visto em TVERSKY (2000), mapas podem simplificar e regularizar a informação removendo informações sem importância para o público-alvo e que poderiam prejudicar a carga cognitiva: desde a antiguidade, o homem tem desenvolvido artefatos cognitivos para ampliar a memória e facilitar o processamento de informações. Destacam-se entre tais artefatos os gráficos, que utilizam elementos e as relações visuais entre eles para representar mundos, real ou metaforicamente, espaciais. Crianças e adultos utilizam, com pouca dificuldade, o espaço e as relações espaciais, sugerindo que esses mapeamentos são cognitivamente naturais.

Mapas parecem compartilhar duas características, embora não de forma estrita; mapas geralmente retratam uma "visão geral", e reduzem e aproximadamente preservam a "escala". A seguir, são listadas as maneiras pelas quais os diagramas, mapas e gráficos visualizam o não-visual.

- Mapas são bidimensionais: além das restrições do meio, existem razões cognitivas para retratar mapas como visões gerais em duas dimensões. Primeiro, parece que as pessoas facilmente concebem ambientes tridimensionais como visões gerais em duas dimensões. Além disso, diagramas tridimensionais são difíceis de construir e compreender. Além dessas razões, para muitos propósitos, informações tridimensionais sobre ambientes simplesmente não são necessárias, e podem até mesmo causar interferências na compreensão.
- Mapas omitem informações: a própria utilidade de um mapa advém da redução de espaço; reduções de tamanho exigem reduções de informação para serem úteis. Mapas são tipicamente projetados com um propósito comunicativo; esse propósito determina quais informações devem ser mantidas e quais informações podem ser eliminadas.
- Mapas regularizam: outra característica prática dos mapas é que eles simplificam e regularizam as informações. Essa simplificação, no entanto, facilita a compreensão das informações desejadas, das orientações gerais e das conexões; e as distorções produzidas pela regularização não causam erros suficientes para criar problemas.

- Mapas usam escala e perspectiva inconsistentes: estradas, rios, ferrovias e outras informações ambientais importantes retratadas em mapas simplesmente não seriam visíveis se a escala fosse adotada de forma consistente. Além disso, muitos mapas violam também a consistência da perspectiva.

Mapas exageram, fantasiam, e transmitem mensagens estéticas, políticas, espirituais e humorísticas: mapas influenciados por elementos políticos, míticos, históricos ou fantásticos são comuns em todo o mundo; mapas que misturavam geografia, crenças e história não são exclusivos dos europeus. Mapas humorísticos avivam jornais, livros e periódicos; há muitos exemplos de mapas projetados para transmitir muito mais do que geografia, e que sacrificam a precisão geográfica em prol de outras mensagens. Mapas esquemáticos são criados para um objetivo ou objetivos específicos, geralmente comunicativos; eles destilam e destacam as informações relevantes para esses propósitos; simplificam e até mesmo exageram essas informações. As mentes dos indivíduos também esquematizam informações espaciais, bem como outros tipos de informação; de fato, muitas das maneiras pelas quais as mentes esquematizam correspondem à maneira como os mapas esquematizam. A correspondência entre modelos de esquematizações externas e internas também pode facilitar o processamento das informações presentes nos mapas.

Assim como os mapas, visualizações gráficas de conceitos e processos abstratos, como os de economia, biologia, química, física, matemática e outras áreas, devem trazer benefícios semelhantes. Segundo outro trabalho de TVERSKY (1995), uma análise de gráficos produzidos por crianças e adultos ao longo da história e em diferentes culturas lista algumas características gerais da maneira como o espaço foi e é utilizado por eles, e os elementos nele contidos para transmitir significado.

- Relações espaciais: enquanto mapas usam espaço para representar espaço, gráficos podem usar espaço para representar outros conceitos, como tempo, quantidade e preferência.
- Direção dos mapeamentos espaciais: crianças de todas as culturas linguísticas mapearam aumentos em quantidade ou preferência da esquerda para a direita, da direita para a esquerda e de baixo para cima. O viés de mapear para cima para mais, melhor ou mais forte não se restringe a crianças. Dizemos que alguém está na "cabeça" ou "topo" da classe, ou "abaixo" da média. Também para gestos, "para cima" é geralmente bom ou forte ou poderoso ou bem-sucedido, e "para baixo" o oposto.
- Elementos: gráficos que retratam informações que não são espaciais, mas sim abstratas, usam elementos posicionados no espaço, assim como o próprio espaço, para transmitir essas informações.

- Expressando significados usando o espaço: o espaço, portanto, assim como os recursos visuais e espaciais, pode ser usado para transmitir conceitos abstratos de maneiras cognitivamente naturais. A similaridade pode ser representada por aparências semelhantes; agrupamentos, baseados, é claro, na similaridade, podem ser representados por proximidade no espaço; conexões entre grupos são facilmente transmitidas por dispositivos que indicam caminhos; ordenações podem ser transmitidas de diversas maneiras; para indicar direção, um recurso popular é a seta; o conceito de extensão pode ser facilmente representado pela extensão espacial, comprimento ou área, e o conceito de proporção pode ser facilmente representado pela proporção espacial.

Representações refletem concepções da realidade, não a realidade; isso vale tanto para representações de elementos no mundo, quanto para elementos dentro da mente. Mapas, desenhos, gráficos e diagramas são formas de comunicação; como na maioria das interações sociais, a fidelidade à verdade pode não ser a principal preocupação.

Um mapa esquemático, conforme o sugerido nos parágrafos anteriores, pode por exemplo ajudar a visualizar, por unidade federativa, os dados do Censo Demográfico realizado a cada década pelo Instituto Brasileiro de Geografia e Estatística (IBGE), evitando as enormes diferenças de área ou população.

Em relação à qualidade das representações cartográficas, podemos citar o conceito de excelência desenvolvido por TUFTE (2001). A excelência em gráficos estatísticos consiste em ideias complexas comunicadas com clareza, precisão e eficiência. Apresentações gráficas devem:

- mostrar os dados
- induzir o espectador a pensar sobre a substância em vez de sobre a metodologia, o design gráfico, a tecnologia de produção gráfica ou algo mais
- evitar distorcer o que os dados têm a dizer
- apresentar muitos números em um espaço pequeno
- tornar grandes conjuntos de dados coerentes
- encorajar o olho a comparar diferentes pedaços de dados
- revelar os dados em vários níveis de detalhe, desde uma visão geral ampla até a estrutura fina
- servir a um propósito razoavelmente claro: descrição, exploração, tabulação ou decoração

- estar intimamente integrado com as descrições estatísticas e verbais de um conjunto de dados.

Assim, ao enxergarmos os mapas de dados, também chamados de "mapas temáticos", como um dos designs gráficos fundamentais segundo TUFTE (2001), um mapa temático contendo dados estatísticos deve, entre outros aspectos, permitir muitos níveis diferentes de análise, variando desde uma visão geral, passando pela contemplação de padrões globais gerais, até os pequenos detalhes.

O mapeamento temático, ou mapas que indicam a variação de um ou vários atributos estatísticos no espaço, é um tópico de interesse primário tanto para a ciência quanto para a prática da Cartografia. O tema tem sido o foco de grande parte da investigação científica do século XX no âmbito da Cartografia (por exemplo, MCMMASTER e MCMMASTER (2002); MONTELLO (2002)) e um tópico de discussão em inúmeros livros didáticos de cartografia (por exemplo, DENT (1999); FISHER (1982); KRYGIER e WOOD (2024); ROBINSON *et al.* (1969); SLOCUM *et al.* (2008); TYNER (2014)). Mapas temáticos são agora um meio familiar de visualização de informações, visíveis em jornais, revistas, artigos de periódicos e outros lugares (MONMONIER, 1989).

A literatura do mapa temático detalha as considerações e melhores práticas sobre o design do mapa temático; assuntos como a escolha adequada de esquemas de cor (BREWER (1989); BREWER *et al.* (1997); BREWER *et al.* (1997)), os meios de atribuição de dados em classes (OLSON e BREWER (1997); JENKS e CASPALL (1971)) e algoritmos para escala perceptual de símbolos proporcionais (BREWER e CAMPBELL (1998); FLANNERY (1971)). Cada um desses estudos trata de decisões de projeto específicas dentro um tipo particular de mapa selecionado para representar um tipo particular de fenômeno geográfico. Comparativamente menos se tem falado sobre uma das escolhas mais fundamentais que um designer de mapas deve fazer: qual estilo de simbolização selecionar para seu mapa. Por exemplo, mapa coroplético, mapa de símbolo graduado, mapa isoplético, e mapa de densidade de pontos são soluções comuns para mapas univariados, e todas podem ser usadas para representar a mesma informação numérica e enumerada (MACEACHREN e DIBIASE, 1991).

Segundo ELMER (2012), mapeamento temático refere-se a mapas que indicam a variação de um ou vários atributos estatísticos em todo o espaço, e hoje em dia é um meio familiar de visualizar informações com base na área ou localização. Mapas temáticos não servem apenas para mostrar locais, como em "mapas de propósito geral", mas sim para mostrar atributos ou estatísticas sobre lugares, padrões espaciais desses atributos, e relacionamentos entre lugares. Temas comuns usados em mapas temáticos são densidade populacional, renda, e votos por região ou país.

Existem muitos tipos de mapas temáticos que atendem a diferentes propósitos. "Mapas de pontos", "mapas de símbolos proporcionais", "mapas coropléticos" e "cartogramas" estão entre os mais usados. Estes são brevemente descritos abaixo, juntamente com um resumo de suas vantagens e desvantagens.

Mapas de pontos usam pontos do mesmo tamanho para representar uma unidade ou um valor fixo sobre um mapa (Figura 1.1). Por exemplo, em um mapa de densidade populacional, cada ponto pode representar 5000 habitantes, ou qualquer outra quantidade. Além do número de registros, os mapas de pontos apresentam a localização e distribuição espacial desses registros em uma área, preservando sua forma geográfica. Assim, podem apresentar densidades comparativas entre regiões, ao considerar a concentração de pontos. Eles são, no entanto, de elaboração demorada, especialmente quando construídos manualmente. Além disso, quantidades grandes de pontos causam excessiva concentração, e são difíceis de contar, impossibilitando o cálculo dos números reais. É difícil comparar valores em duas áreas que têm a mesma proporção de pontos. Quando a quantidade mínima é menor que o valor de um ponto, eles podem dar uma falsa impressão de vazio.

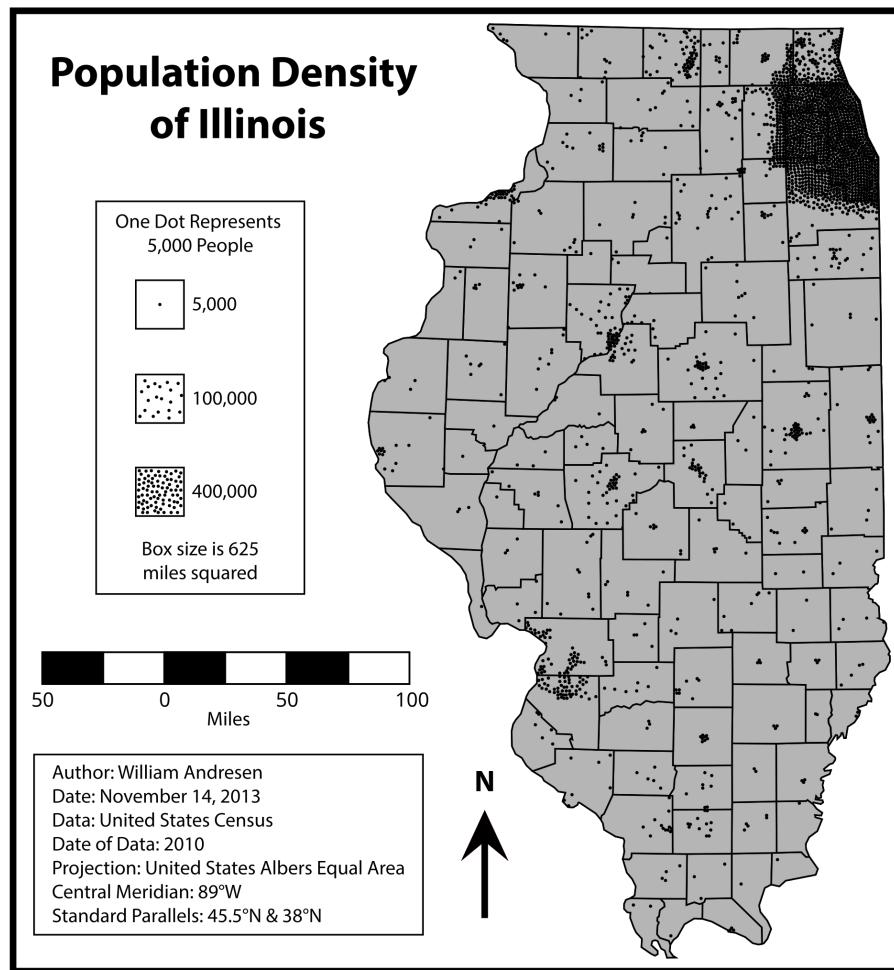


Figura 1.1: Exemplo de mapa de pontos. Densidade populacional de Illinois, 2010. Extraído de https://commons.wikimedia.org/wiki/File:Population_Density_of_Illinois_2010_Wikipediamap.svg.

Os "mapas de símbolos proporcionais", também conhecidos como "mapas de símbolos graduados", usam símbolos de tamanhos diferentes para representar os dados (Figura 1.2). O tamanho do símbolo reflete sua magnitude. Círculos ou quadrados, às vezes transparentes, são colocados diretamente no mapa, e permitem a comparação razoavelmente bem. Este tipo de mapa preserva a forma geográfica das áreas, mas pode levar a problemas em pequenas regiões do mapa. Nesse caso, símbolos podem se sobrepor, dificultando a relação com suas respectivas áreas e gerando confusão visual.

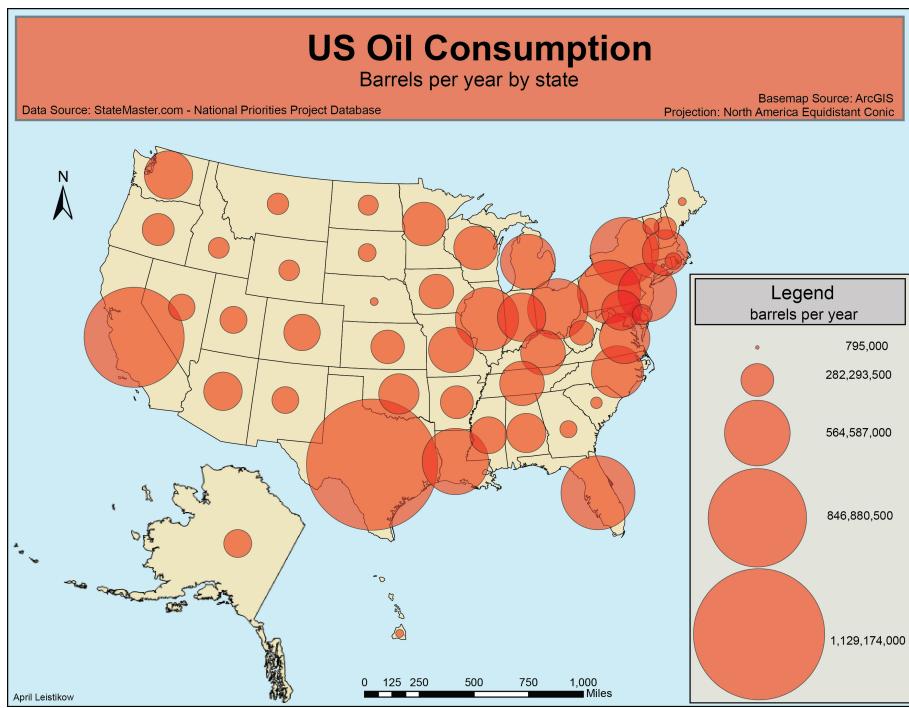


Figura 1.2: Exemplo de mapa de símbolos proporcionais. Consumo de Petróleo nos EUA. Extraído de Leistiar / Ty. <https://www.flickr.com/photos/89769525@N08/8242929652/>.

Um "mapa coroplético" apresenta áreas geográficas colorizadas ou padronizadas em proporção à medição da variável estatística exibida, como renda per capita ou distritos eleitorais. Frequentemente, um esquema de graduação de cores é usado, onde valores baixos aparecem levemente colorizados, e altos são indicados com tons mais escuros. "Mapas coropléticos" são indicados para regiões de tamanho e forma quase uniformes, e também para dados uniformemente distribuídos (SPECKMANN e VERBEEK, 2010). Embora os estados dos Estados Unidos da América apresentem alguma disparidade de áreas, particularidades podem justificar por quais razões esse tipo de mapa temático tem sido amplamente utilizado. Em primeiro lugar, as formas dos estados americanos são bastante regulares, facilitando a comparação visual de suas áreas (Figura 1.3). Embora existam claras diferenças visuais entre os maiores estados (Texas e Califórnia) e os cinco menores, o conjunto restante (cerca de 80 por cento dos estados) exibe áreas que parecem mais ou menos semelhantes. Por outro lado, quando este tipo de layout é aplicado a, digamos, alguma variável relacionada à população, então grandes áreas ainda serão visualmente mais importantes, mesmo que pouco povoadas.

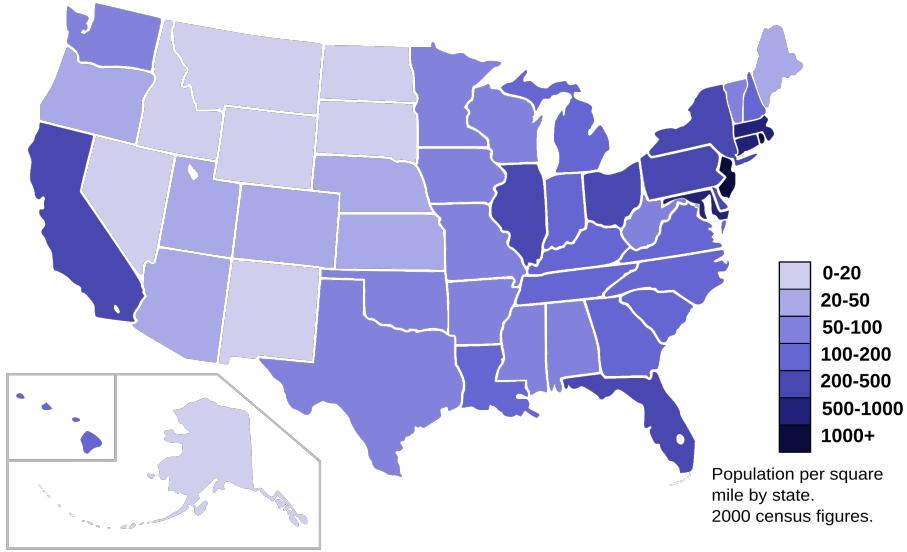


Figura 1.3: Mapa coroplético. População dos EUA por milha quadrada por estado (números do censo de 2000). Extraído de https://commons.wikimedia.org/wiki/File:US_2000_census_population_density_map_by_state.svg.

Cartogramas, também conhecidos como mapas de valor por área, são mapas nos quais uma variável mapeada, como população ou produto nacional bruto, substitui a área ou a extensão de terra. Existem muitos tipos diferentes de cartogramas, como o tipo padrão, os cartogramas não-contíguos, e os cartogramas circulares ou de Dorling, entre outros. No tipo padrão, também chamados de cartogramas de área contígua, a topologia é mantida, conforme os objetos (países ou estados) permanecem conectados uns com os outros. No entanto, como o valor do atributo estatístico é codificado como o tamanho da área, podem ocorrer deformações, ou seja, valores dos dados significativamente menores ou maiores podem afetar a forma original da região, renderizando o mapa de modo irreconhecível (SPECKMANN e VERBEEK, 2010). Como um exemplo, observe as formas dos países africanos no cartograma de produção de alimentos orgânicos (Figura 1.4). No cartograma não contíguo, a topologia (a conectividade entre objetos) é sacrificada para preservar sua forma. Livres para se afastarem de seus objetos adjacentes, as regiões representadas podem crescer ou diminuir de tamanho e ainda manter suas formas, mas a desvantagem é que as formas podem se sobrepor. Nos cartogramas de Dorling ou circulares (Figura 1.5), cada região é representada por um círculo não sobreposto, que é dimensionado de acordo com o valor dos dados. A distribuição dos círculos visa se assemelhar à topologia original - suas adjacências e posições relativas. Da mesma forma, cartogramas retangulares usam retângulos para representar cada região. Assim como cartogramas circulares, retângulos permitem comparações com baixo esforço cognitivo; no entanto, essas formas impõem limitações no layout, a fim de evitar sobreposições. Eles também tornam difíceis de respeitar as adjacências das regiões.

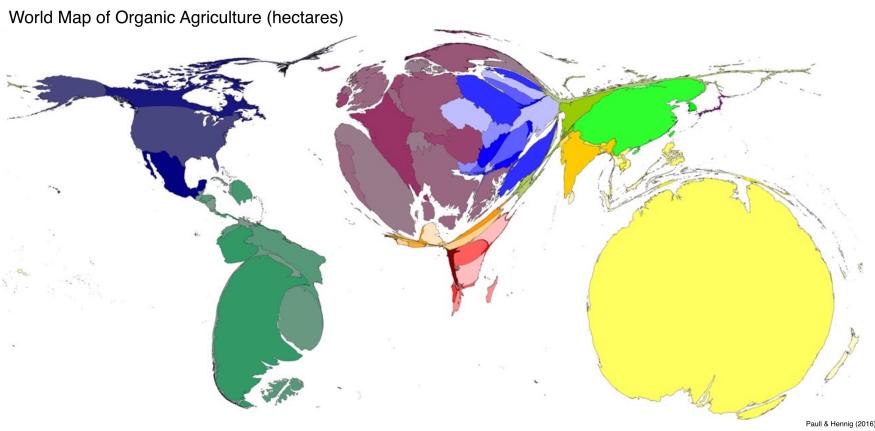


Figura 1.4: Cartogramas de áreas contíguas. Imagem de um mapa mundial da agricultura orgânica. (PAULL, J. e HENNIG, B. 2011). Extraído de <https://commons.wikimedia.org/wiki/File:PaullHennig2016WorldMap.OAha.CC-BY-4.0.jpg>.

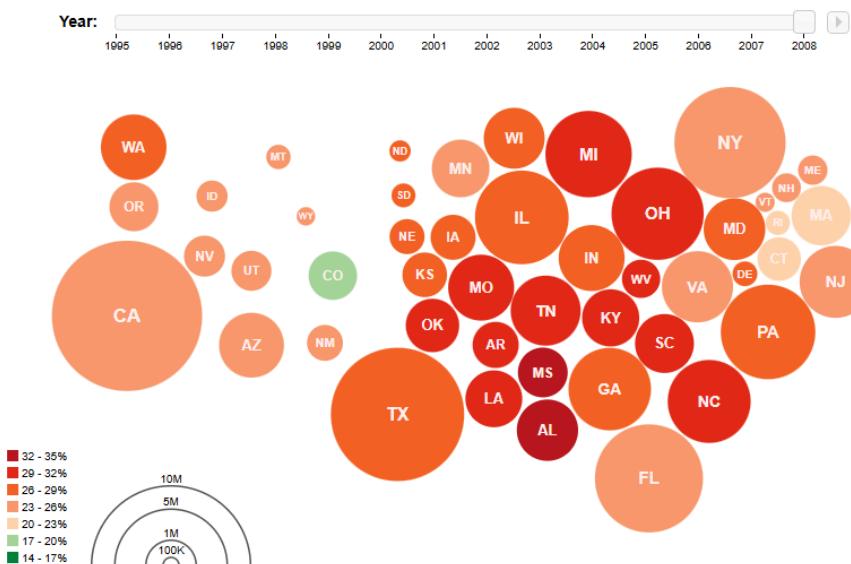


Figura 1.5: Exemplo de cartograma de Dorling. Protovis. Extraído de <http://mbostock.github.io/protovis/ex/cartogram.html>.

Quanto aos problemas de visualização em mapas temáticos, são muitas as maneiras já relacionadas de resolver esses problemas. Podemos citar o trabalho de BERTIN (1983), que estudou os problemas das representações gráficas e listou uma sequência histórica de mapas que permite compreender a evolução em paralelo com a tecnologia.

Segundo BARRETO *et al.* (2018), mapas são abstrações de um mundo tridimensional representado na superfície bidimensional de uma folha de papel ou de uma tela de vídeo. Como tal, não pretendem ser vistos como realidade, mas como representações ou projeções de pontos ou regiões do globo sobre um plano. Embora

a maioria dos mapas afirme oferecer uma imagem precisa e objetiva do mundo, cada um é construído com base em convenções arbitrárias que atendem às necessidades e desejos de seus criadores.

Segundo BERTIN (1983), a representação gráfica constitui um dos básicos sistemas de signos concebidos pela mente humana com o propósito de armazenar, compreender e comunicar informações essenciais. Como uma "linguagem" para o olho, a representação gráfica se beneficia das propriedades ubíquas da percepção visual. Como um sistema monossêmico, ela constitui a parte racional do mundo de imagens.

De quais variáveis o sistema gráfico de signos dispõe? BERTIN (1983) afirma que o olho é o intermediário em um grande número de percepções. Nem todas elas dizem respeito ao sistema que estamos estudando; dentro desses limites, consideraremos que o sistema gráfico dispõe de oito variáveis. Uma marca visível que expresse uma correspondência pertinente pode variar em relação às duas dimensões do plano; ela pode ainda variar em tamanho, valor, textura, cor, orientação e forma. A grande diversidade de construções gráficas, dentro de um grupo ou mesmo de um grupo para outro, se deve à aparente liberdade do designer em representar um determinado componente usando qualquer uma das oito variáveis visuais ou uma combinação de várias delas.

No entanto, certas escolhas se tornam convincentes devido à sua maior "eficiência": a eficiência está ligada ao grau de facilidade que caracteriza cada etapa da leitura de um gráfico. O conjunto de observações que conduzem às regras de construção constitui a teoria da imagem: as construções mais eficientes são aquelas em que qualquer pergunta, seja qual for o seu tipo ou nível, pode ser respondida em um único instante de percepção, em uma única imagem.

Por sua vez, a imagem, a forma visual significativa perceptível no instante mínimo de visão, é construída sobre três variáveis homogêneas e ordenadas: as duas dimensões do plano e uma variável retinal. Todas as informações com três componentes ou menos podem ser representadas como uma única imagem.

É necessário e suficiente que as regras de construção sejam respeitadas: neste caso, qualquer que seja o tipo ou nível da pergunta, a resposta será vista em um único instante de percepção; podemos dizer que a representação gráfica, e mapas e cartogramas aqui se incluem, é uma imagem.

Uma imagem não acomodará mais de três variáveis significativas: como resultado, todas as informações com mais de três componentes não podem ser construídas como uma única imagem; isso significa que, para certas perguntas, a compreensão da resposta necessitará de vários instantes de percepção e, consequentemente, várias imagens.

A eficiência visual é inversamente proporcional ao número de imagens necessárias

para a percepção dos dados: informações que envolvam mais de três componentes devem ser construídas de forma diferente, de acordo com a função pretendida, ou seja, de acordo com a natureza das perguntas preferenciais.

Existem tantos tipos de perguntas quantos componentes na informação, mas dentro de cada tipo há inúmeras perguntas possíveis, categorizadas conforme BERTIN (1983) em três níveis:

- Elementar: perguntas introduzidas por um único elemento de um componente, por exemplo: "Em uma determinada data..." e resultando em uma única correspondência.
- Intermediário: perguntas introduzidas por um grupo de elementos no componente, por exemplo: "Nos primeiros três dias, qual foi a tendência do preço?" Resposta: "O preço subiu". Essas perguntas são bastante numerosas, visto que podemos formar grupos bastante diversos.
- Geral ou Global: uma pergunta introduzida pelo componente inteiro: "Durante todo o período, qual foi a tendência do preço?" Resposta: "Movimento ascendente geral".

BERTIN (1983) sustenta que informações com três componentes ou menos, construídas como uma única imagem, podem cumprir todas as três funções de representação gráfica:

- Registro de informações: criar um mecanismo de armazenamento que evite o esforço de memorização. O gráfico utilizado para esse fim deve ser abrangente, e pode ser não memorizável em sua totalidade.
- Comunicação de informações: criar uma imagem memorizável que gravará a informação na mente do espectador. O gráfico usado aqui deve ser memorizável, e pode não ser abrangente; a imagem deve ser simples.
- Processamento de informações: fornecer uma figuração que permita simplificação e sua justificativa. O gráfico deve ser memorizável (para comparações) e abrangente (para escolhas).

Atualmente, serviços de mapas na internet para dispositivos móveis já são muito comuns, popularizando seu uso em nosso cotidiano; por esta razão, entre diversas outras, impõe-se como um tema de pesquisa relevante o desenvolvimento de soluções que cumpram de modo eficiente as funções da representação gráfica descritas acima e sejam adequadas às necessidades de visualização de dados georreferenciados em tempo real.

A presente tese parte, portanto, da necessidade de avanço nos estudos para solucionar e atender aos requisitos já expostos para representação gráfica, a fim de apresentar duas propostas de renderização automática de cartogramas que utilizam formas geométricas básicas, explicitando suas metodologias e oferecendo diversas sugestões de implementação, com aplicações práticas que utilizam dados coletados pelo IBGE no Censo Demográfico 2022 e outras fontes. A estrutura deste trabalho está organizada como se segue:

- o capítulo 2 apresenta as motivações para o desenvolvimento de um mapa temático adequado às especificidades do território brasileiro;
- o capítulo 3 apresenta os conceitos e algoritmos utilizadas como referência para o desenvolvimento das soluções;
- o capítulo 4 apresenta conceitualmente as propostas de cartogramas especiais;
- o capítulo 5 apresenta a implementação inicial das propostas sugeridas, utilizando dados estatísticos reais;
- o capítulo 6 apresenta uma nova implementação das propostas sugeridas visando a geração automatizada dos cartogramas;
- por fim, o capítulo 7 apresenta as considerações finais, bem como oportunidades de desenvolvimentos adicionais futuros para as soluções propostas no presente trabalho.

Capítulo 2

Motivação para um mapa temático do Brasil

Existem muitos tipos de mapas temáticos, incluindo "mapas coropléticos", "cartogramas", "mapas de símbolos proporcionais" e "mapas de pontos", entre os mais utilizados: todos os tipos apresentam vantagens e desvantagens dependendo da região a ser coberta; a maioria delas, quando aplicadas ao território brasileiro, apresentará limitações e distorções. Um mapa temático para os estados brasileiros, em particular, apresenta desafios interessantes, uma vez que os estados são fortemente diversificados em área, desenvolvimento econômico e densidade populacional. Por exemplo, um "mapa coroplético" onde as áreas geográficas são retratadas de forma realista pode erroneamente equiparar a importância visual de um estado com o tamanho de sua população ou significado econômico. Da mesma forma, "cartogramas" usando deformação podem levar ao desaparecimento visual de pequenas áreas, "mapas de símbolos proporcionais" podem gerar símbolos sobrepostos em pequenas regiões, e "mapas de pontos" podem impossibilitar o cálculo de números reais, ou ainda fornecer uma falsa impressão de vazio. Além disso, a deformação do mapa pode prejudicar a comparação entre mapas que mostram diferentes temas para a mesma região.

Considerando que a maioria dos mapas temáticos foi criada e pesquisada com maior frequência em países europeus e norte-americanos, é razoável supor que esses tipos de mapas foram projetados de acordo com os requisitos e necessidades desses países. Embora essa hipótese possa ser investigada, digamos, conduzindo um grande levantamento de exemplos de muitos anos, uma abordagem mais prática foi escolhida por BARRETO *et al.* (2018), qual seja: projetar um mapa temático do Brasil que fosse adequado para exibir dados referentes aos estados brasileiros. Assim, tal estudo apresentou uma tentativa de analisar as necessidades especiais de um mapa temático do Brasil por estados, em face de suas peculiaridades geográficas.

O Brasil é uma República Federativa que reúne 27 Unidades da Federação: 26

estados, e um Distrito Federal para a capital do país, Brasília. As unidades têm áreas extremamente díspares: Amazonas é o maior estado com 1.570.745,7 km², enquanto o Distrito Federal (DF) possui 5.822,1 km² e Sergipe, o menor dos estados, tem 21.910,3 km² (Figura 2.1). O Brasil é geopoliticamente dividido em cinco regiões; cada região é composta por três a nove estados. A população brasileira não está distribuída uniformemente entre as unidades, e o PIB varia muito de estado para estado (IBGE, 2025a).

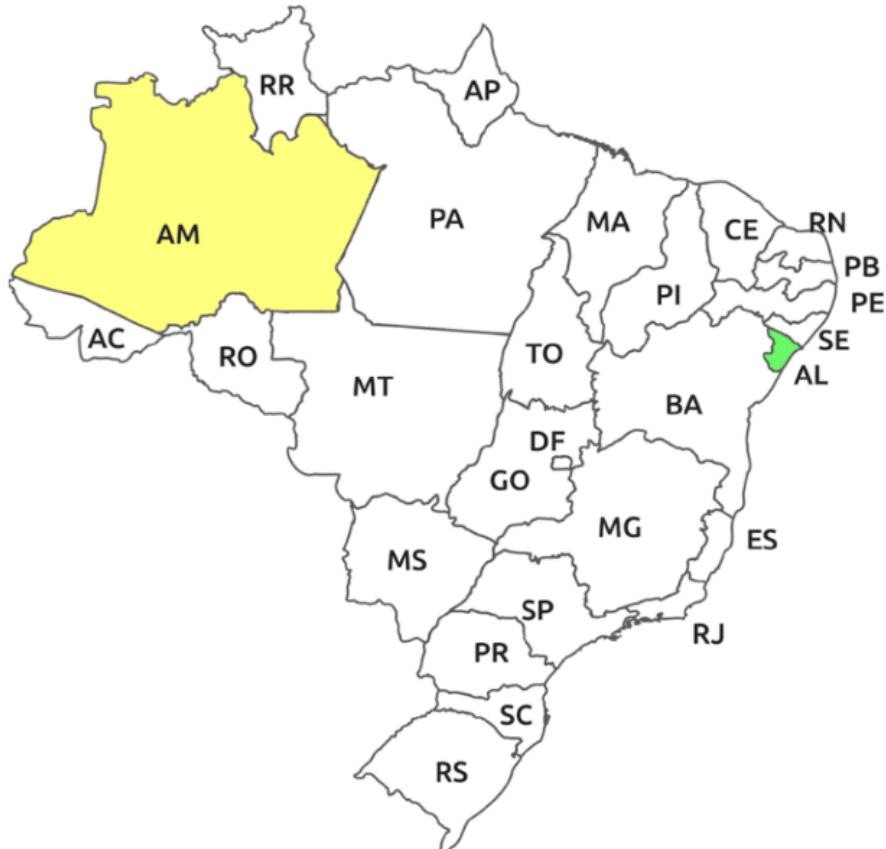


Figura 2.1: Mapa do Brasil, com destaque para os maiores e menores estados. Extraído de BARRETO *et al.* (2018).

Com o objetivo de avaliar as necessidades especiais de um "mapa temático" do Brasil por estado em face de suas peculiaridades geográficas, o trabalho de BARRETO *et al.* (2018) enfocou um estudo de caso que mapeia as notas médias do Exame Nacional do Ensino Médio (ENEM), organizado pelo Ministério da Educação do Brasil, por tipo de disciplina curricular, para cada estado brasileiro. Ao projetar um mapa temático para as notas médias do ENEM por disciplina e por estado, foram estabelecidas algumas exigências iniciais: a primeira dentre elas era permitir a comparação entre as notas médias de cada uma das cinco disciplinas curriculares, a fim de identificar padrões e correlações, sem perder a situação geral do país. Uma vez que tal comparação requer mapas que mantenham mais ou menos a

mesma forma, o uso de cartogramas pode ser descartado para esse fim.

Um segunda exigência era preservar o formato do país, que sugeria um mapa coroplético do Brasil como um objeto apropriado de estudo. No entanto, um mapa geograficamente preciso da divisão política do Brasil põe em evidência problemas bem conhecidos: áreas reduzidas, como Sergipe (SE) e o Distrito Federal (DF), são tão pequenas em relação ao tamanho total do mapa que, quando comparadas através das cores, não são claramente discerníveis; estados como Amazonas e Pará, apesar de sua baixa densidade populacional, seriam imediatamente superenfatizados, independentemente da variável sendo comparada. Esse problema pode ser agravado caso os valores sejam aproximadamente similares.

Para melhor comparar valores entre estados brasileiros que têm tamanhos muito diferentes, um terceiro requisito foi estabelecido: que os estados fossem renderizados em tamanho tão uniforme quanto possível. Assim, evitar o viés associado às diferenças de tamanho das áreas era uma questão central neste estudo, e seu terceiro requisito. Para esse fim, foi decidido que cada um dos 26 estados e o Distrito Federal seriam representados por formas de área constante.

O quarto requisito era usar paletas de cores monocromáticas para representar a magnitude da nota média de cada disciplina por estado. Embora muitos estudos (YAU, 2013) (MUNZNER, 2014) tenham enfatizado os limites da percepção ao comparar valores com cores, observou-se que usar uma paleta de cores monocromática separada para cada disciplina poderia personalizar os mapas de assuntos. Uma escala de cores não é apenas uma escolha estética, mas uma ferramenta útil e intuitiva para visualizar magnitude: além disso, as cores podem ser usadas para representar atributos ordenados; neste caso, as diferentes disciplinas escolares.

Finalmente, o quinto requisito era preservar as relações de vizinhança entre os estados, a fim de facilitar a busca visual por um estado específico para aqueles usuários familiarizados com a distribuição territorial das unidades da Federação em um mapa tradicional.

Tendo decidido usar elementos de tamanho fixo para representar os estados, preservar as relações de vizinhança entre eles e também reter o formato do país, o primeiro passo consistiu em escolher a melhor forma geométrica a ser empregada. Os hexágonos foram então considerados (Figura 2.2), uma vez que podem ser montados em um ladrilhamento do plano, de forma a se assemelhar ao formato do Brasil (apresentado para comparação na Figura 2.3). Além disso, através de suas seis arestas e vértices, as relações de vizinhança seriam preservadas.



Figura 2.2: Primeiro layout de mapas temáticos usando hexágonos. Extraído de BARRETO *et al.* (2018).



Figura 2.3: Mapa geográfico do Brasil. Extraído de BARRETO *et al.* (2018).

Na tentativa de manter o formato geral do Brasil e o relacionamento topológico entre as unidades, algumas modificações foram feitas nas posições dos hexágonos (Figura 2.4). No entanto, alguns relacionamentos de vizinhança foram rompidos, enquanto alguns estados costeiros foram transferidos para o interior. Os estados vizinhos do Piauí (PI) e Maranhão (MA) foram separados uns dos outros, enquanto os estados costeiros de Pernambuco (PE), Piauí (PI) e Bahia (BA) foram deslocados para o interior.

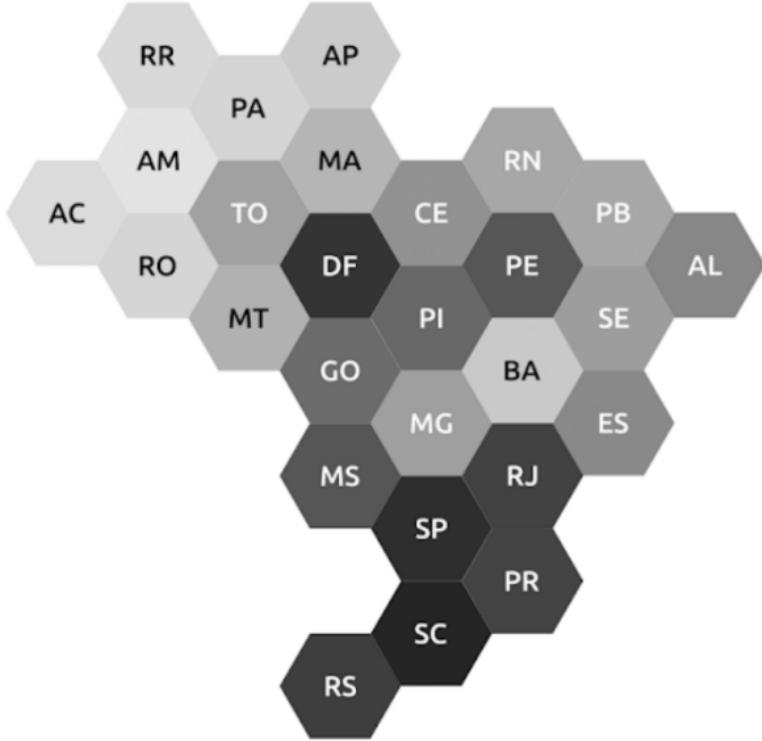


Figura 2.4: Segundo layout de mapas temáticos usando hexágonos. Extraído de BARRETO *et al.* (2018).

Considerando os problemas decorrentes da construção hexagonal do mapa, outras formas foram investigadas. Uma vez que outros polígonos regulares que ladrilham o plano são quadrados e triângulos, o uso de quadrados foi brevemente considerado, mas imediatamente descartado, principalmente porque os quadrados se distribuem automaticamente ao longo de linhas e colunas, o que inviabiliza a preservação de relações de vizinhança em um país de forma tão irregular como o Brasil.

A forma triangular foi considerada em seguida, uma vez que foi observado que, onde cada hexágono é adjacente a seis outros hexágonos por aresta em um ladrilhamento hexagonal do plano (Figura 2.5), cada triângulo é adjacente a três outros triângulos por aresta e nove outros triângulos por vértice em um ladrilhamento triangular do plano (Figura 2.6), o que ajudaria a resolver o problema do afastamento

indesejado de unidades vizinhas.

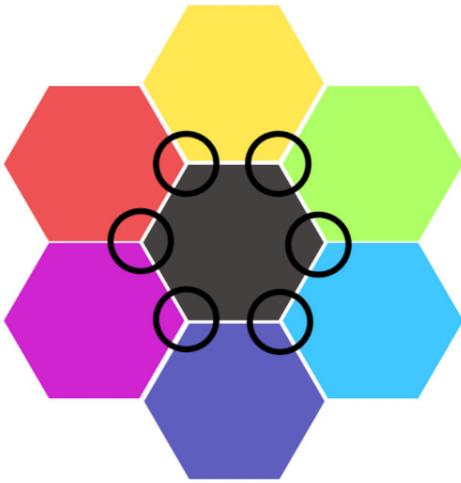


Figura 2.5: Relações de vizinhança de arestas e vértices em ladrilhamentos hexagonais do plano. Extraído de BARRETO *et al.* (2018).

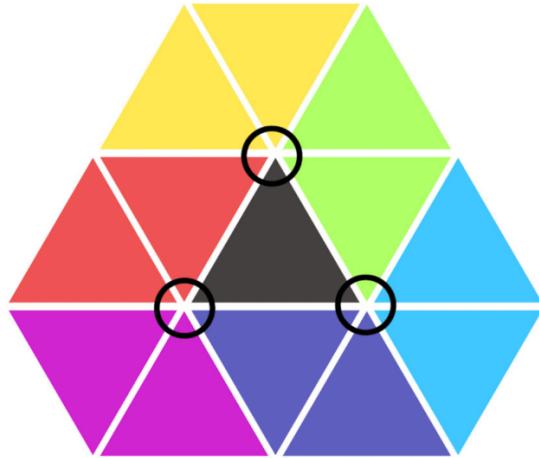


Figura 2.6: Relações de vizinhança de arestas e vértices em ladrilhamentos triangulares do plano. Extraído de BARRETO *et al.* (2018).

Muitas organizações diferentes foram tentadas, algumas distribuindo as unidades por Grandes Regiões (Figura 2.7), outras tentando focar no formato externo do país (Figura 2.9). Embora a maioria das adjacências de regiões fosse mantida, o formato final divergia muito do mapa geográfico (repetido para fins de comparação na Figura 2.8). Além disso, a variação entre as posições do triângulo, voltadas para cima ou para baixo, cria um problema adicional, pois o conjunto dos triângulos não apresenta a regularidade observada no desenho hexagonal.

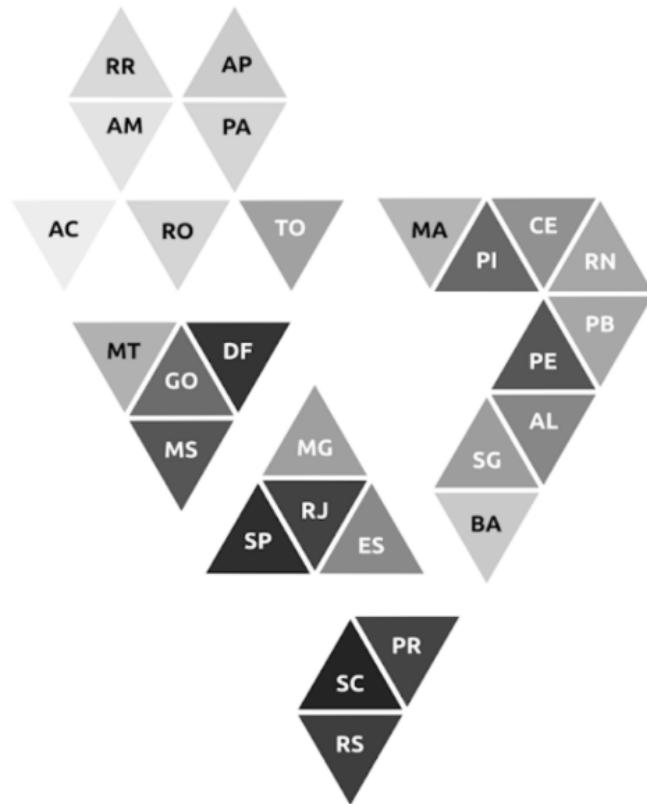


Figura 2.7: Layout do mapa temático utilizando triângulos com unidades distribuídas por Grandes Regiões. Extraído de BARRETO *et al.* (2018).



Figura 2.8: Mapa geográfico do Brasil. Extraído de BARRETO *et al.* (2018).

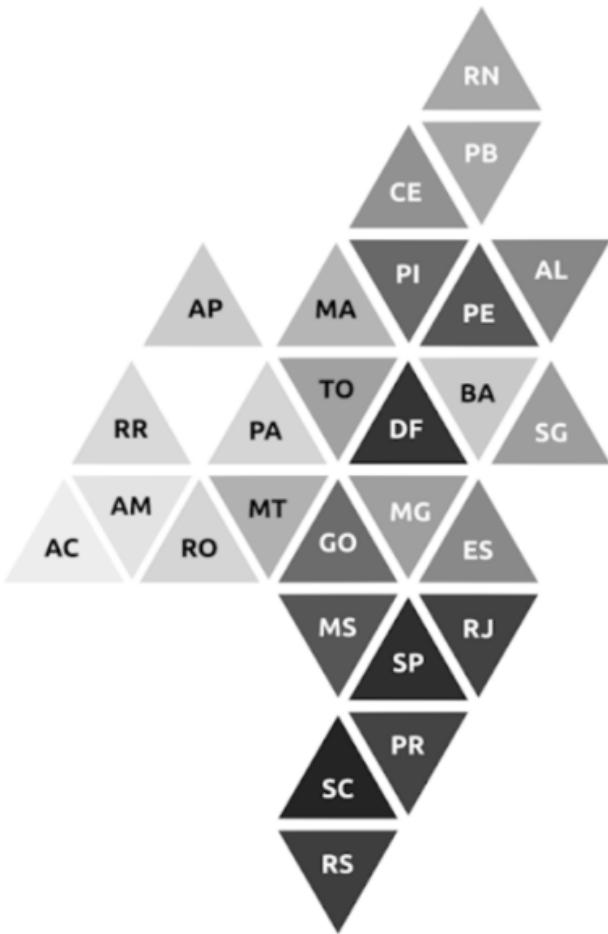


Figura 2.9: Layout do mapa temático usando triângulos. Extraído de BARRETO *et al.* (2018).

Assim, o uso de triângulos foi descartado para uso em um mapa temático do Brasil, restando os hexágonos como a escolha mais adequada. Outro layout foi tentado, onde as unidades foram redistribuídas de forma não-contígua, agrupando-as por Grandes Regiões e adicionando espaço entre as unidades, visando a minimização dos problemas topológicos observados (Figuras 2.11 e 2.12). A princípio, o posicionamento não-contíguo das unidades parecia visualmente melhorar o design do mapa, fazendo-o assemelhar-se ao formato final do país: o aproveitamento do espaço entre as unidades também minimiza a sensação desagradável de ter unidades litorâneas deslocadas para o interior do país; no entanto, conforme a distância entre os hexágonos aumentava, parecia diminuir a semelhança com o formato do país.

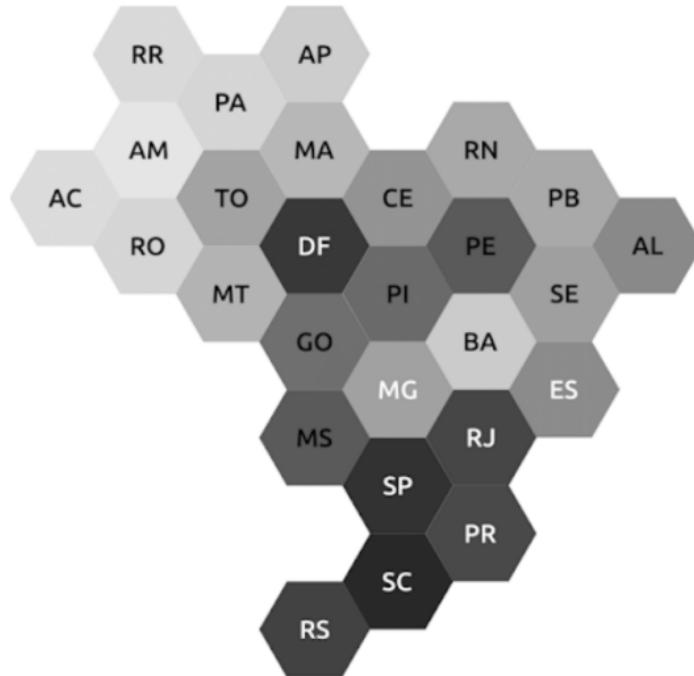


Figura 2.10: Mapa temático do Brasil com hexágonos contíguos. Extraído de BARRETO *et al.* (2018).

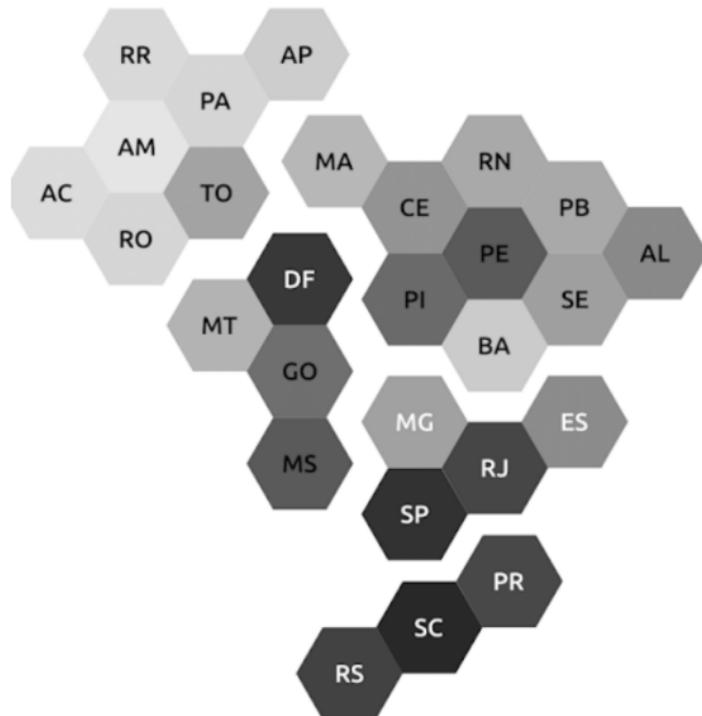


Figura 2.11: Mapa temático do Brasil com hexágonos não-contíguos organizados por Grandes Regiões. Extraído de BARRETO *et al.* (2018).

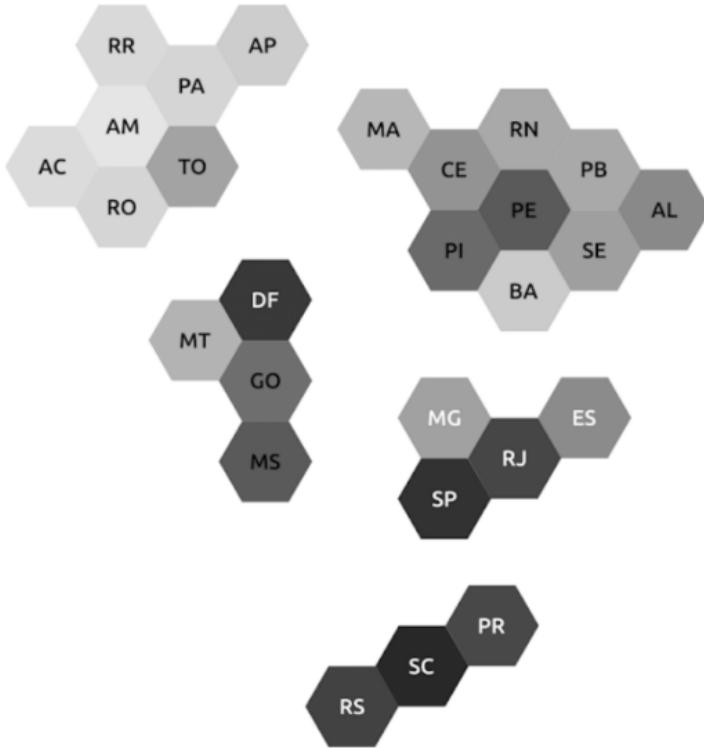


Figura 2.12: Mapa temático do Brasil com hexágonos não-contíguos organizados por Grandes Regiões, com maior espaçamento. Extraído de BARRETO *et al.* (2018).

Depois de muitas tentativas de tentar representar as unidades federativas do Brasil com diferentes polígonos regulares, observou-se que, independentemente da forma utilizada, o problema principal ainda residia na preservação da vizinhança entre os estados. Assim, foi criado um algoritmo para testar exaustivamente as possibilidades de mudar as posições da unidade, mantendo a vizinhança topológica. Para isso, foi elaborado um grafo, onde os vértices representam os estados e as arestas representam as relações de vizinhança, ou seja, se dois estados têm uma fronteira comum, então uma aresta do grafo é definida entre os dois vértices correspondentes. Um algoritmo de layout de grafos direcionado por força (EADES, 1984) foi então codificado, onde o comprimento de cada aresta podia variar continuamente entre a distância geográfica entre as capitais dos estados e um valor médio fixo. Nas Figuras 2.13 e 2.15, os grafos representam esses dois extremos, respectivamente, enquanto a Figura 2.14 corresponde a uma interpolação de 50 por cento. Cores foram atribuídas aos vértices, para representarem a alocação das unidades da Federação nas Grandes Regiões correspondentes.

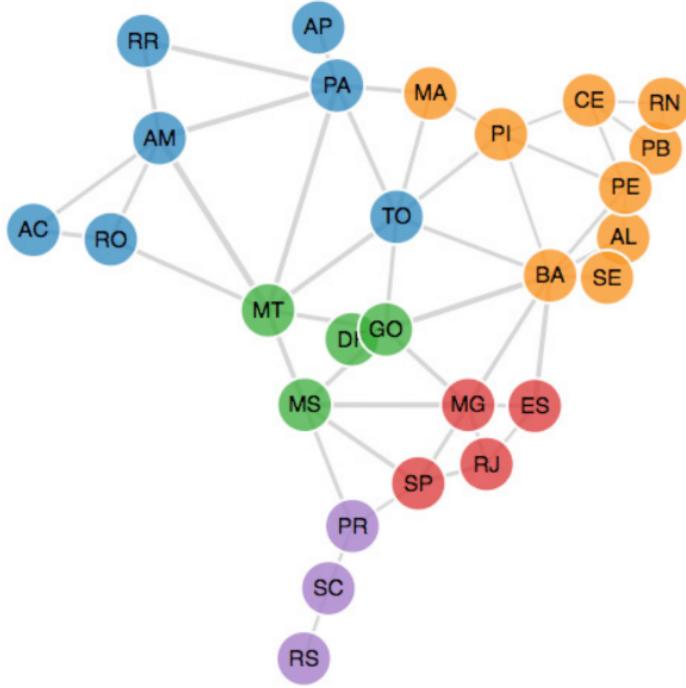


Figura 2.13: Grafo dos estados brasileiros posicionados de acordo com a distância entre as capitais dos estados. Extraído de BARRETO *et al.* (2018).

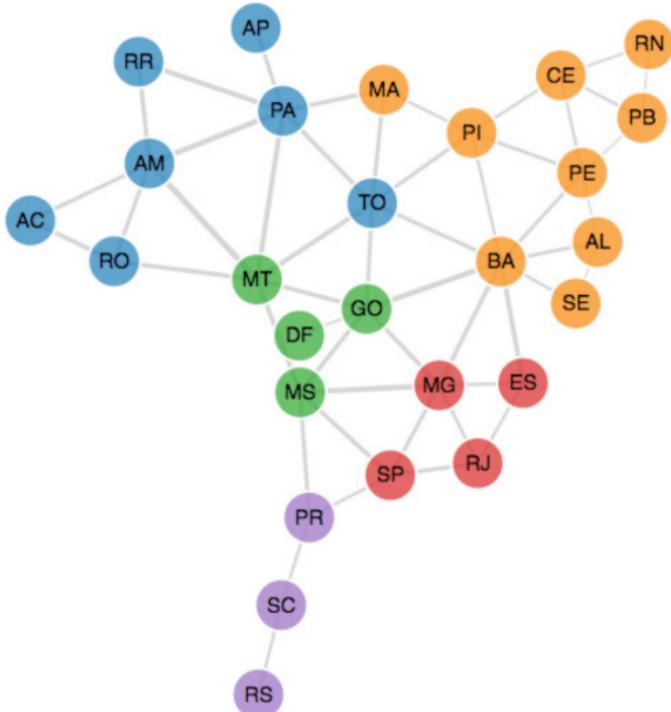


Figura 2.14: Grafo dos estados brasileiros posicionados em uma interpolação de 50 por cento entre a distância real e um valor fixo das capitais dos estados. Extraído de BARRETO et al. (2018).

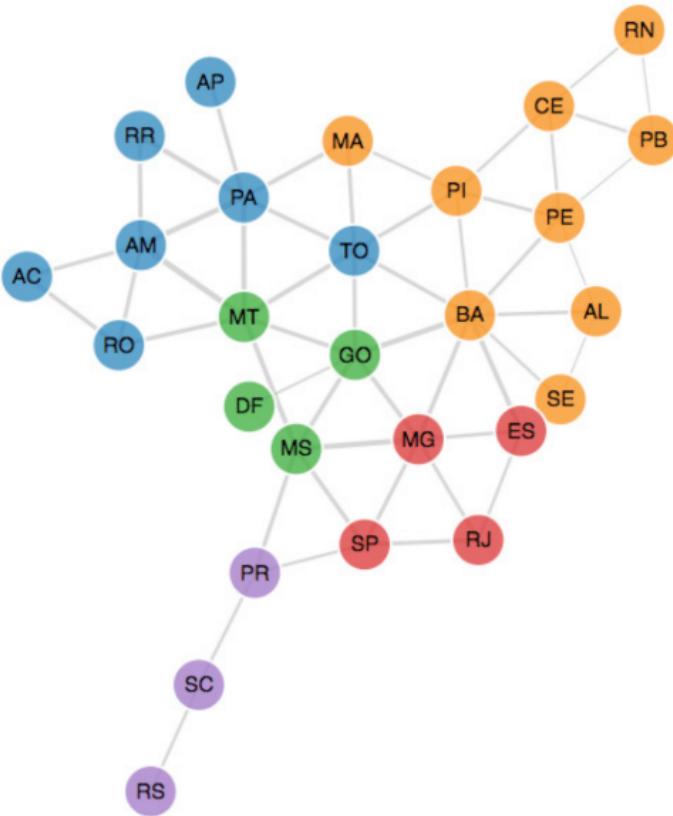


Figura 2.15: Grafo dos estados brasileiros posicionados segundo um valor fixo de distância entre as capitais dos estados. Extraído de BARRETO *et al.* (2018).

Pode-se notar que no desenho do gráfico da Figura 2.13, a forma geográfica do país é aproximadamente mantida. Por outro lado, o desenho na Figura 2.15 mostra cada estado em uma distância quase uniforme de seus vizinhos. Os estados que possuem áreas semelhantes e não muitos vizinhos fronteiriços, como os estados da região Sul (em roxo), sofrem pouco, preservando sua posição na forma do país. No entanto, muitos outros problemas tornaram-se visíveis: o primeiro deles é observado no que diz respeito ao Distrito Federal; sua área está totalmente localizada no interior do estado de Goiás, e possui apenas um reduzido trecho fronteiriço ao estado de Minas Gerais, resultando em um formato incomum. O segundo problema é com o estado da Bahia, que apresenta a situação oposta: possui muitos vizinhos, e por este motivo é empurrado para o centro do mapa. O último e pior problema observado é a deformação do formato do país: assim, pode-se concluir que qualquer organização das unidades da Federação que procure respeitar a vizinhança topológica e utilizar uma distância uniforme entre as unidades irá necessariamente corromper o formato geral do país.

Por este motivo, os autores do estudo optaram por preservar o formato do país, aceitando o risco de mover algumas unidades para uma posição não desejável no

mapa: observou-se também que o uso dos rótulos contendo as siglas das respectivas unidades da Federação minimizou problemas na localização visual de estados que se moveram ligeiramente em relação à sua posição geograficamente ótima; as diferentes paletas de cores monocromáticas usadas permitiram comparações entre as notas médias das disciplinas (Figura 2.16).

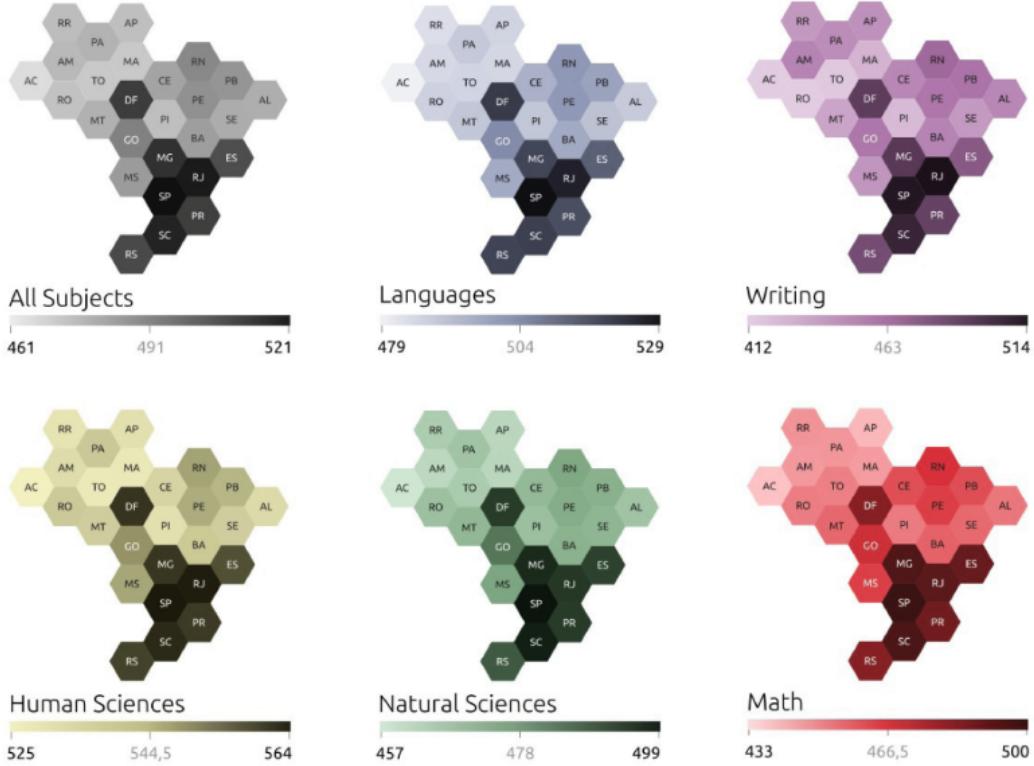


Figura 2.16: Mapas temáticos do Brasil resultantes, com escalas monocromáticas representando notas médias do ENEM por Unidade da Federação. Extraído de BARRETO *et al.* (2018).

O resultado final atingiu quatro dos cinco requisitos estabelecidos, a saber: ter a área do estado normalizada, permitir comparação entre as notas médias escolares por estado, preservar o formato do território do país, e utilizar paletas de cores monocromáticas para representar magnitude. Depois de todos os experimentos visuais realizados, percebeu-se que preservar as relações de vizinhança entre os estados (o quinto requisito) só poderia ser totalmente obtido sob o risco de perder o formato do país. O uso de rótulos minimizou a dificuldade de localizar visualmente estados que haviam sido deslocados de suas posições geograficamente ideais.

De modo geral, o trabalho de BARRETO *et al.* (2018) confirmou a importância do desenvolvimento de mais estudos em design da informação, focados na elaboração de mapas que considerem as necessidades regionais.

Capítulo 3

Conceitos e algoritmos rumo a um mapa temático do Brasil

3.1 Mapas de Colar

Na busca pelo desenvolvimento de alternativas para elaboração de mapas temáticos, chegamos ao trabalho de SPECKMANN e VERBEEK (2010): segundo os autores, dados estatísticos associados a regiões geográficas estão atualmente disponíveis em grandes quantidades e em alcance global, portanto, métodos automatizados para exibir visualmente esses dados estão em alta demanda. Existem vários tipos de mapas temáticos bem estabelecidos para dados quantitativos associados a regiões utilizando alguma escala proporcional: mapas coropléticos, cartogramas e mapas de símbolos proporcionais. No entanto, todos esses mapas sofrem com limitações, especialmente se valores grandes de dados estiverem associados a pequenas regiões. Para superar essas limitações, é proposto um novo tipo de mapa temático quantitativo, o "mapa de colar"(necklace map): em um mapa de colar, regiões do mapa bidimensional subjacente são projetadas em intervalos correspondentes sobre uma curva unidimensional (o colar), que circunda as regiões do mapa. Os símbolos são dimensionados de forma que suas áreas correspondam ao valor dos dados de sua região, e aplicados sem sobreposições entre eles dentro do intervalo correspondente no colar. Os mapas de colar oferecem visualizações claras e organizadas, e permitem tamanhos de símbolo comparativamente grandes. Oferecem boa visualização a conjuntos de dados que não são proporcionais aos tamanhos das regiões. A ordem linear dos símbolos ao longo do colar facilita uma comparação natural entre os tamanhos dos símbolos. Um mapa pode conter vários colares aninhados ou separados, a fim de visualizar dados agrupados. As vantagens dos mapas de colar têm um preço: a associação entre um símbolo e sua região é mais fraca do que em outros tipos de mapas; se necessário, a interatividade pode ajudar a fortalecer essa associação.

É apresentada uma abordagem automatizada para a geração de mapas de colar, a qual permite ao usuário controlar interativamente o posicionamento final dos símbolos. Tal abordagem é validada no estudo, com experimentos utilizando diversos conjuntos de dados e mapas (Figura 3.1).

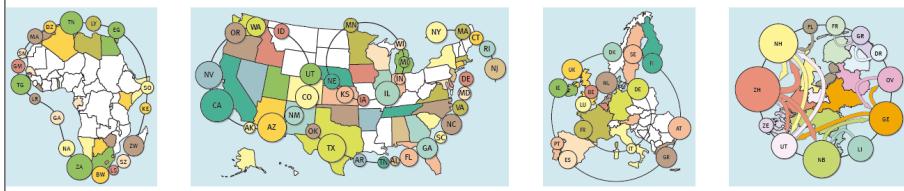


Figura 3.1: Exemplos de "necklace maps". Extraído de SPECKMANN e VERBEEK (2010).

Um mapa de colar é um mapa temático e, como tal, consiste em dois componentes: um mapa geográfico subjacente, e a sobreposição temática. Mapas de colar se destinam a exibir dados quantitativos na escala proporcional associada às regiões; dados na escala proporcional possuem um zero significativo (SLOCUM *et al.*, 2008). A sobreposição temática consiste em símbolos associados às regiões do mapa de entrada: assim como nos mapas de símbolos, cada símbolo – mais comumente um círculo ou um quadrado, embora outras formas sejam possíveis – é dimensionado, de forma que sua área corresponda ao valor dos dados associados à sua região; os símbolos são então colocados, sem sobreposição, em um ou mais "colares", as curvas que envolvem as regiões do mapa de entrada. Um bom mapa de colar preserva o máximo possível da relação espacial entre um símbolo e sua região: cada região (bidimensional) do mapa original é projetada sobre um intervalo contíguo (unidimensional) em um colar, de forma que a localização global da região em relação ao colar seja preservada (ver Figura 3.2, esquerda); os símbolos estarão restritos a posicionar seus centros dentro de seu intervalo correspondente no colar. Observando a Figura 3.2, direita, o centro do colar está sobre a França; os símbolos são encontrados aproximadamente na mesma ordem e direção que suas regiões correspondentes, em relação ao centro desse colar.

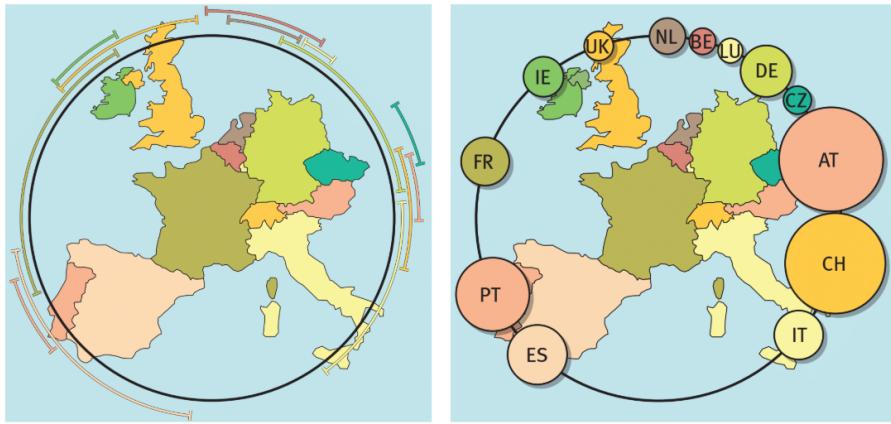


Figura 3.2: À esquerda, intervalos de região sobre um colar; à direita, consumo de energia proveniente de recursos renováveis em 2007. Extraído de SPECKMANN e VERBEEK (2010).

A forma dos colares é um fator importante na aparência visual de um mapa de colares: círculos e elipses são formas geométricas altamente simétricas, que desviam pouca atenção dos tamanhos dos símbolos; deixam o interior do mapa amplamente desocupado e, consequentemente, não interferem com a visibilidade da informação geográfica do mapa original. Além disso, tais formas de colar permitem uma fácil integração com mapas ou fluxos coropléticos; no entanto, também pode ser desejável usar formas de colar que sejam semelhantes à forma global das regiões de entrada, ou que não sejam fechadas: em princípio, qualquer curva aberta ou fechada em "forma-de-estrela" pode ser usada como um colar. Uma curva é em forma de estrela se existe pelo menos um ponto que pode "ver" a curva completa - este ponto, o "centro" do colar, que para o caso de uma curva fechada se situa em seu interior obrigatoriamente, é necessário para definir as projeções das regiões de entrada no colar.

Um mapa de colar comunica sua mensagem através dos tamanhos de seus símbolos - tanto seu tamanho real, quanto a proporção entre os tamanhos. Um corpo significativo de trabalho teórico e de estudos com usuários discute qual escala comunica a diferença entre as quantidades de forma mais eficaz; citamos aqui como exemplo os trabalhos de DENT (1999) e de SLOCUM *et al.* (2008). Os três tipos básicos de escala são os seguintes: a "escala matemática" dimensiona as áreas dos símbolos em relação direta aos dados; a "escala perceptiva" amplia símbolos maiores além de seu tamanho matematicamente correto para compensar a incapacidade humana de julgar os tamanhos relativos dos símbolos de área com precisão; e a "classificação de intervalo" subdivide os dados em classes. Todos os mapas de colar no estudo citado usam escala matemática.

A qualidade de um mapa de colar é determinada por vários fatores; cabe lem-

brar que cada símbolo está restrito a posicionar seu centro dentro de seu intervalo correspondente sobre o colar.

- Boas posições de símbolos - os intervalos no colar capturam bem a relação espacial das regiões e do colar.
- Tamanho máximo do símbolo.
- Símbolos disjuntos.
- Ordem adequada de símbolos ao longo do colar - símbolos vizinhos nos colares correspondem a regiões vizinhas.

Existe um claro balanceamento entre os dois primeiros critérios de qualidade: grandes intervalos permitem grandes tamanhos de símbolos, em troca de posições menos precisas de símbolos. Em uma condição extrema, cada intervalo cobre o colar inteiro, e assim a colocação do símbolo é arbitrária, e o tamanho do símbolo pode ser facilmente maximizado; no entanto, a maioria das associações espaciais entre símbolos e áreas será perdida. No outro extremo, o posicionamento de cada símbolo no colar é especificado o mais próximo possível de sua região; como resultado, os símbolos geralmente permanecerão bem pequenos.

Em seguida, é apresentado um algoritmo que computa mapas de colar de alta qualidade: várias abordagens são fornecidas para mapear as regiões para seus intervalos de colar, e é mostrado como maximizar os tamanhos dos símbolos, com o cuidado de manter todos os símbolos disjuntos.

Antes de descrever o algoritmo, primeiro é necessário introduzir algumas definições e notações. A entrada consiste em um conjunto de polígonos $\mathcal{P} = \{P_1, \dots, P_n\}$ representando as regiões geográficas, e um conjunto de dados $\mathcal{Z} = \{z_1, \dots, z_n\}$ onde cada valor z_i ($z_i > 0$) está associado ao polígono P_i . No caso de regiões com vários polígonos, como os EUA ou a Indonésia, escolhe-se um polígono representativo (na parte contígua dos EUA), ou usamos o envoltório convexo de todos os polígonos. Assume-se que os valores dos dados são normalizados, ou seja, $\sum_{i=1}^n z_i = 1$. Também temos o colar: uma curva \mathcal{C} em forma de estrela com centro $v = (x, y)$. Para facilitar a explicação, é assumido que os símbolos S_i são círculos; o algoritmo se estende a outras formas de símbolos de maneira direta. Representa-se a localização do centro de um círculo S_i por um ângulo α_i relativo ao centro v do colar \mathcal{C} . O fator de escala global para todos os símbolos é denotado por ρ . Como é usada a escala matemática, o raio de cada símbolo S_i no mapa final é $\rho\sqrt{z_i}$.

Para cada polígono P_i um único intervalo contíguo $I_i = [a_i, b_i]$ é computado no colar com um dos métodos descritos a seguir. O intervalo viável I_i representa todos os ângulos aceitáveis para posicionar S_i , no sentido anti-horário de a_i para b_i . Diz-se

agora que um posicionamento que consiste em α_i ($1 \leq i \leq n$) e ρ é "viável" se o seguinte for válido: (i) todos os S_i são disjuntos, e (ii) $\alpha_i \in I_i$ para todo $1 \leq i \leq n$. O objetivo é, portanto, encontrar um posicionamento viável que maximize ρ . Dado tal posicionamento viável, é utilizada uma variedade de abordagens para melhorar a localização de cada símbolo, mantendo ρ fixo. O algoritmo pode ser resumido da seguinte forma:

1. calcular intervalos viáveis
2. otimizar tamanhos de símbolos
3. otimizar posicionamentos de símbolos

Dependendo do tipo de dados de entrada e do mapa original, diferentes projeções de regiões em intervalos de colar podem ser adequadas. Abaixo estão descritas três abordagens: as duas primeiras estão implementadas no sistema apresentado e levam a resultados muito bons com vários conjuntos de dados e mapas; a terceira impõe uma maior complexidade ao algoritmo resultante.

Intervalos por Centroídes: Considere o raio do centro $v = (x, y)$ do colar \mathcal{C} através do baricentro (x_i, y_i) de um polígono P_i . A interseção desse raio com o colar \mathcal{C} no ângulo β_i pode ser considerada uma escolha lógica para posicionar S_i . O intervalo viável I_i é então definido como um intervalo de tamanho constante de tamanho c em torno de β_i (ver Figura 3.3, esquerda). O valor c pode ser uma constante, ou depender do número de regiões n ; portanto, usando esse método, os intervalos viáveis são calculados da seguinte forma:

$$I_i = \left(\beta_i - \frac{c}{2}, \beta_i + \frac{c}{2}\right), \text{ onde } \beta_i = \arctan 2(y_i - y, x_i - x)$$

Observe que, com esse método, o comprimento de todos os intervalos é o mesmo: no entanto, é razoável assumir que, para regiões maiores, um intervalo maior é aceitável; se for esse o caso, deve-se usar um dos outros dois tipos de intervalos.

Intervalos por Margem: Para fortalecer a relação entre os símbolos e as regiões, podemos definir os intervalos viáveis I_i para obter um significado geométrico claro. Dizemos que um ângulo α é aceitável para S_i se o raio do centro de \mathcal{C} no ângulo α passa por P_i : assim, I_i é o menor intervalo contendo todos os ângulos aceitáveis para S_i ; em outras palavras, I_i representa a menor fatia de \mathcal{C} contendo P_i (ver Figura 3.3, direita). Observe que com essa definição todos os ângulos são aceitáveis para o símbolo S_j pertencente ao polígono P_j que contém o centro de \mathcal{C} ; como isso geralmente é indesejável, usamos o intervalo do centroíde para esse símbolo específico S_j .

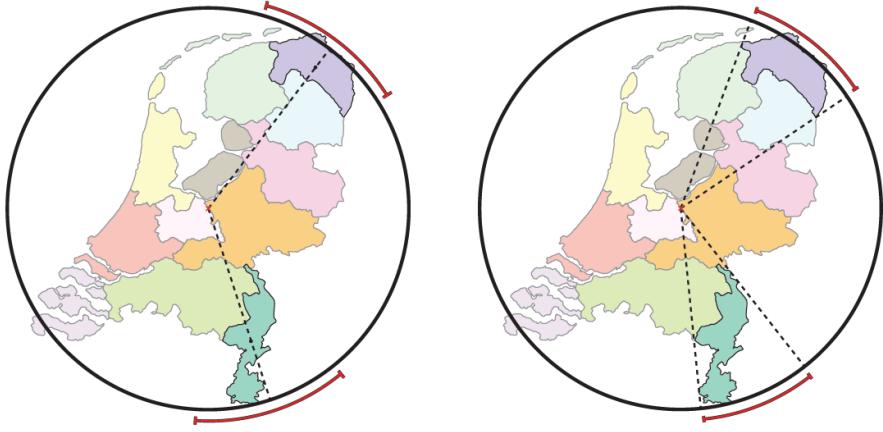


Figura 3.3: À esquerda, intervalos por centroides; à direita, intervalos por margens. Extraído de SPECKMANN e VERBEEK (2010).

Com esse método, a relação entre símbolos e regiões é relativamente clara. No entanto, intervalos por margem não levam em consideração o conjunto de dados \mathcal{Z} : intervalos por margem para regiões de pequenas dimensões são reduzidos, o que pode levar a mapas abaixo do ideal se grandes valores de dados forem atribuídos a pequenas regiões; por tal motivo, intervalos dependentes de densidade também devem ser considerados.

Intervalos Dependentes de Densidade: pode-se adaptar os intervalos dependendo do conjunto de dados \mathcal{Z} . Considere um intervalo de ângulos I : se um intervalo I_i está completamente contido em I ($I_i \subseteq I$), então S_i deve ser colocado sobre o intervalo I . Define-se a densidade de I da seguinte forma:

$$\delta(I) = \frac{1}{|I|} \sum_{i|I_i \subseteq I} \sqrt{z_i}$$

Aplica-se um limite superior na densidade máxima $\max_I \delta(I) \leq c$, onde c é um parâmetro que pode ser definido pelo usuário. Para isso, os tamanhos dos intervalos são alterados iterativamente: sempre que for encontrado um intervalo I para o qual $\delta(I) > c$, aumentam-se os intervalos contidos em I até que $\delta(I) \leq c$ (ver Figura 3.4); continua-se esse processo até que $\max_I \delta(I) \leq c$. Observe que o processo completo exige que verifiquemos apenas um número de intervalos da ordem $O(n^2)$. É claro, aumentar os intervalos enfraquece a relação espacial entre regiões e símbolos: o parâmetro c , de fato, controla o balanceamento entre os tamanhos dos símbolos e as posições dos símbolos. Pode-se melhorar ainda mais esse procedimento considerando o fato de que os dois símbolos extremos posicionados sobre I não precisam estar completamente contidos em I ; na verdade, metade desses dois símbolos pode estar fora de I , e isso deve ser considerado para exclusão ao computar $\delta(I)$.



Figura 3.4: Intervalos dependentes da densidade. Extraído de SPECKMANN e VERBEEK (2010).

Depois de calcular os intervalos, pode-se prosseguir para encontrar uma solução viável que maximize ρ ; para isso, usa-se uma busca binária sobre ρ . Desta forma, para um determinado ρ , é necessário verificar se há uma solução viável para aquele ρ : uma vez que todos os símbolos devem estar posicionados sobre \mathcal{C} , este é essencialmente um problema de dimensão 1; no entanto, é preciso saber qual parte de \mathcal{C} está coberta por um símbolo S_i com raio $\rho\sqrt{z_i}$. Caso \mathcal{C} seja um círculo com raio r , é computado o seguinte.

$$z'_i = \arcsin\left(\frac{\rho\sqrt{z_i}}{r}\right)$$

É fácil verificar que um símbolo com raio $\rho\sqrt{z_i}$ cobre exatamente uma cunha de ângulo $2z'_i$ de \mathcal{C} , se \mathcal{C} for um círculo (ver Figura 3.5, à esquerda); chamamos z'_i de "raio de cobertura" de S_i .

Se \mathcal{C} não for um círculo, o cálculo do raio de cobertura é um pouco mais complexo: se \mathcal{C} for um círculo, então a parte de \mathcal{C} coberta por S_i independe da posição de S_i ; não é esse o caso se \mathcal{C} for uma elipse ou uma curva B-spline cúbica.

Para uma dada posição e tamanho de um símbolo S_i , determina-se a parte do

colar coberto por S_i , ou o raio de cobertura z'_i , como se segue: esteja S_i em um ângulo γ do centro do colar; são traçadas linhas partindo do centro de S_i ao longo do colar em ambas as direções. Sejam γ_1 e γ_2 os ângulos onde o colar sai de S_i : assim, a parte do colar coberta por S_i se estende de γ_1 a γ_2 ; como não é possível computar γ_1 e γ_2 com exatidão, o colar é percorrido para aproximar os dois ângulos. Podemos então definir o raio de cobertura no ângulo γ como $z'_i(\gamma) = \max(\gamma_2 - \gamma, \gamma - \gamma_1)$. Observe que, usando esta definição, a cunha de γ_1 a γ_2 não contém necessariamente o símbolo inteiro (consulte a Figura 3.5, à direita): como S_i pode ser posicionado em qualquer lugar sobre seu intervalo correspondente I_i , usa-se o maior dos raios de cobertura para todas as possíveis colocações de S_i .

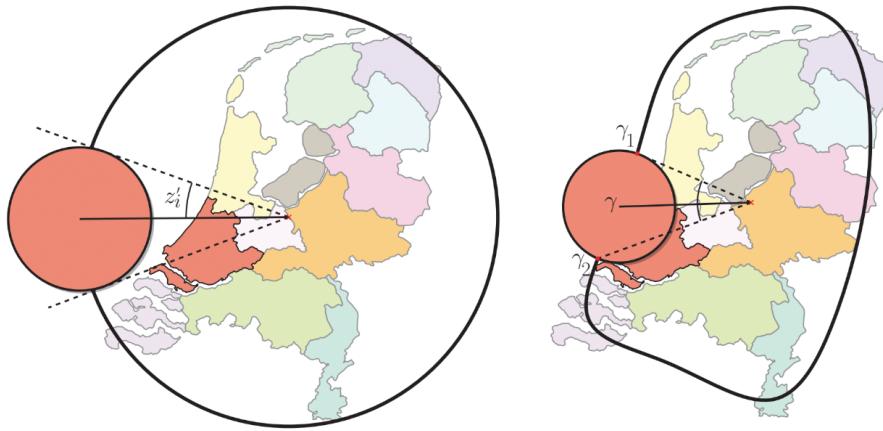


Figura 3.5: Raio de cobertura. À esquerda, sobre o círculo; à direita, sobre a curva spline. Extraído de SPECKMANN e VERBEEK (2010).

Assim, o raio de cobertura de um símbolo S_i é definido como se segue:

$$z'_i = \max_{\gamma \in I_i} z'_i(\gamma)$$

Também neste caso, passa-se por vários valores de γ para aproximar z'_i : embora isso sempre superestime a parte do colar coberta por um símbolo, este erro é limitado, desde que a forma do colar seja bem aproximada a de um círculo.

Alternativamente, em vez de usar o z'_i pré-computado, é possível calcular o raio de cobertura (exato) $z'_i(\gamma)$ quando necessário para um ângulo γ durante os estágios subsequentes do algoritmo: isso daria melhores resultados, mas tornaria o algoritmo muito mais complexo, e resultaria em uma ligeira melhoria apenas; portanto, tal aprimoramento não foi considerado no estudo citado.

Com os raios de cobertura, pode-se descrever o problema como o seguinte problema de dimensão 1: encontre ângulos α_i ($1 \leq i \leq n$) tais que:

- $\alpha_i \in I_i$ ou $a_i \leq \alpha_i \leq b_i$ para $1 \leq i \leq n$

- Os intervalos $[\alpha_i - z'_i, \alpha_i + z'_i]$ são disjuntos.

Pode-se agora considerar duas variantes do problema: deixar que os centros dos intervalos I_i determinem uma ordem, e exigir que os símbolos sigam essa ordem (isso faria sentido com os intervalos por centroides); ou permitir que os símbolos estejam em qualquer ordem. Fazemos essa distinção porque existe um algoritmo simples e rápido para a primeira variante, mas a segunda variante é NP-completa: no entanto, a segunda variante é menos restritiva, apresenta melhores resultados e, portanto, é mais interessante: o estudo citado oferece algoritmos para ambas as variantes.

Para a primeira variante, pode-se usar a simples abordagem a seguir: inicialmente, faz-se $\rho = 0$, e todos os símbolos são posicionados no início de seu intervalo; em seguida, procede-se aumentando ρ , fazendo os símbolos crescerem. Quando dois símbolos se chocarem, são afastados um do outro. Como todos os símbolos começam no início de seu intervalo, pode-se afastá-los em apenas uma direção (sentido anti-horário): esse processo continua até que um símbolo seja afastado para o final de seu intervalo; neste ponto, o ρ ideal é encontrado, e todos os símbolos são posicionados no colar sem sobreposição. Os autores detalham este procedimento em um artigo complementar (SPECKMANN e VERBEEK, 2015).

Infelizmente, a abordagem fácil para a primeira variante não funciona para a segunda variante: em vez disso, usa-se programação dinâmica para resolver essa variante, resultando em um algoritmo tratável de parâmetro fixo na "espessura" dos intervalos $I_1 \dots I_n$. Para um dado ângulo α , seja $k(\alpha)$ o número de intervalos I_i que contêm α . Define-se a "espessura" K de um conjunto de intervalos como sendo o máximo valor de $k(\alpha)$ para todos os α . Observe que, para intervalos de cunha, a espessura K tem um significado geométrico claro: K é o número máximo de regiões atravessadas por um mesmo raio único que se origina do centro do colar. O tempo de execução do algoritmo é $O(nK2^K)$. Até agora ignoramos o fato de que os símbolos devem ser posicionados sobre um círculo, não sobre uma linha: conforme SPECKMANN e VERBEEK (2015), estender a abordagem para um círculo impõe algum custo em relação ao tempo de execução. O algoritmo de programação dinâmica é bastante flexível, e permite impor uma ordem parcial aos símbolos; embora o tempo de execução ainda seja exponencial, K será bastante pequeno para a maioria das instâncias de entrada ($K \leq 10$). É claro que é possível construir facilmente uma instância de entrada que forneça um K muito maior, mas o mapa de colar resultante não seria muito claro; nesse caso, seria melhor mostrar menos símbolos, ou usar vários colares. A Figura 3.6 mostra os mapas de colar computados com a segunda variante, com intervalos por centroides (à esquerda) e intervalos por margem (à direita); na verdade, todos os mapas de colar no estudo citado foram computados com a segunda variante.

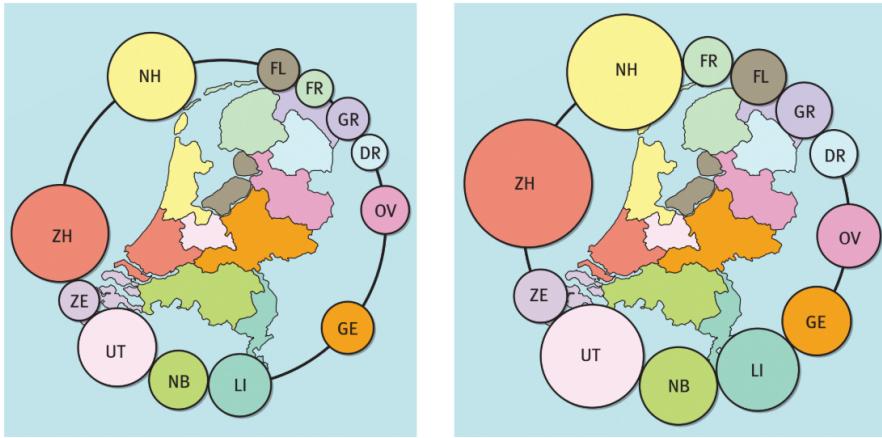


Figura 3.6: À esquerda, tamanhos de símbolos otimizados com intervalos por centroides; à direita, tamanhos de símbolo otimizados com intervalos por margem. Extraído de SPECKMANN e VERBEEK (2010).

Com o algoritmo descrito acima, é obtida uma solução viável onde os símbolos são os maiores possíveis: porém, devido à forma como o algoritmo funciona, as posições dos símbolos não são ótimas, mesmo se corrigirmos ρ . Isso ocorre porque o algoritmo tende a定位 os símbolos nas extremidades de seus intervalos, embora fosse melhor定位 os símbolos no meio de seus intervalos. Também seria visualmente mais atraente se os símbolos não estivessem aglomerados, mas ao invés disso tivessem algum espaço entre eles. Para remediar esse problema, usamos uma etapa de pós-otimização, para定位 os símbolos em uma melhor posição; a fim de não desfazer completamente o trabalho feito pelo algoritmo de posicionamento, não são alterados os tamanhos dos símbolos, sendo realizadas apenas pequenas alterações na ordem dos símbolos.

Para empurrar os símbolos para uma posição melhor, é utilizado um método baseado em força. Assumindo que a ordem dos símbolos é fixa, um símbolo S_i sobre o ângulo α_i tem dois símbolos vizinhos, S_{i-1} e S_{i+1} : seja d_i a distância entre S_{i-1} e S_i , e d_{i+1} a distância entre S_i e S_{i+1} ; aqui queremos significar a distância entre os símbolos no espaço de dimensão 1, então $d_i = \alpha_i - z'_i - (\alpha_{i-1} + z'_{i-1})$. Seja m_i o meio do intervalo I_i ; com isso, podemos definir as duas forças atuando sobre o símbolo S_i conforme se segue.

$$F_{rep}(i) = f_r \left(\frac{1}{d_i} - \frac{1}{d_{i+1}} \right)$$

$$F_{mid}(i) = f_m(m_i - \alpha_i)$$

As constantes f_r e f_m podem ser ajustadas para obter um balanceamento entre empurrar um símbolo para o meio de seu intervalo e afastar os símbolos uns dos outros: pode-se definir a força total como $F(i) = F_{rep}(i) + F_{mid}(i)$; o objetivo é

obter $F(i) = 0$ para todo i . É utilizada uma abordagem muito básica para conseguir isso: passa-se várias vezes por todos os símbolos e, para cada símbolo S_i , assume-se que S_{i-1} e S_{i+1} são fixos; então, resolve-se a equação $F(i) = 0$. Isso agora é um polinômio cúbico em α_i , que pode ser facilmente resolvido: tal abordagem pode ser relacionada ao método de Gauss-Seidel para resolver um sistema linear de equações; o algoritmo calcula rapidamente uma boa solução.

Conforme mencionado acima, são permitidas pequenas alterações na ordem dos símbolos: para cada dois símbolos vizinhos, verifica-se se os dois símbolos podem ser trocados de forma que o resultado ainda seja uma solução viável; troca-se dois símbolos de forma que $F_{rep}(i)$ permaneça o mesmo para todo i . Se a força total $F_{mid}(i - 1) + F_{mid}(i)$ for menor após a troca, são trocados os dois símbolos; caso contrário, essa troca não ocorre.

Usando a abordagem baseada em força, as posições dos símbolos podem ser otimizadas para um ρ fixo; no entanto, como ρ é maximizado, os símbolos podem ter pouco espaço para melhorar sua posição. Para permitir uma melhor compensação entre o tamanho do símbolo e a posição do símbolo, permite-se que o usuário defina um buffer para cada símbolo: esse buffer é incluído ao calcular o tamanho ideal do símbolo, de modo que, posteriormente, cada símbolo tenha algum espaço vazio ao seu redor, correspondendo ao tamanho do buffer; tal espaço pode ser utilizado para melhorar as posições dos símbolos durante esta etapa do algoritmo. A Figura 3.7 mostra os mapas de colar resultantes, depois de adicionar um buffer e empurrar os símbolos para o meio de seus intervalos (à esquerda), ou afastar os símbolos uns dos outros (à direita).

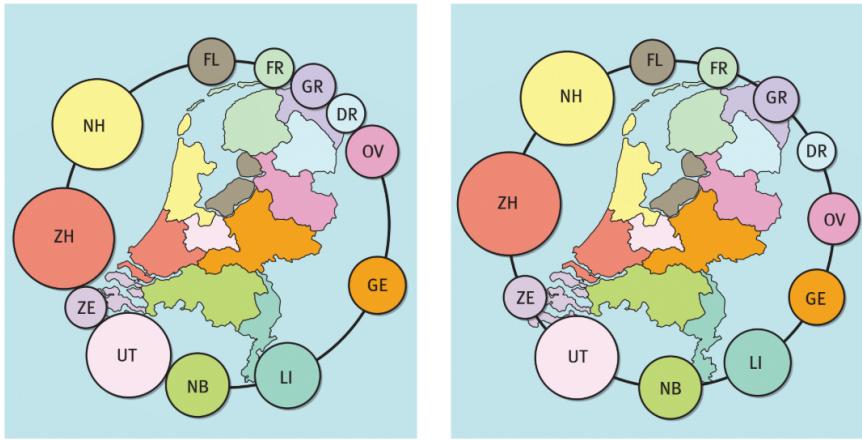


Figura 3.7: À esquerda, símbolos com buffers empurrados para o meio; à direita, símbolos com buffers afastados uns dos outros. Extraído de SPECKMANN e VERBEEK (2010).

A principal vantagem dos mapas de colar é sua aparência clara e organizada: a

ordem linear dos símbolos ao longo dos colares facilita a estimativa e a comparação correta dos tamanhos dos símbolos; mapas de colar visualizam bem conjuntos de dados que não são proporcionais aos tamanhos das regiões, e que não contêm dados para todas as regiões do mapa original.

O algoritmo proposto no estudo citado computa mapas de colares de alta qualidade: a posição espacial relativa de símbolos e regiões captura bem a relação espacial entre regiões e colares, e permite que os usuários associem rapidamente um símbolo à região correta; os símbolos são disjuntos, e aparecem ao longo dos colares em uma ordem que espelha as relações de vizinhança de suas regiões. Além disso, a abordagem de otimização produz mapas de colar cujos símbolos são tão grandes quanto possível, dada a entrada.

Existe uma clara compensação entre os tamanhos dos símbolos e a relação espacial dos símbolos e regiões, que também é influenciada pelo número e distribuição dos símbolos por colar; de um modo geral, é melhor usar vários colares em vez de apenas um colar. São mapeados dados bidimensionais em um domínio unidimensional: intuitivamente, vários colares aumentam as chances de um símbolo ser colocado próximo à sua região, e dimensionado para um tamanho apropriado. Em uma situação extrema, muitos colares levarão de volta aos mapas de símbolos; no entanto, mapas de colares com poucos colares ainda preservam sua aparência estruturada, permitindo um bom posicionamento e dimensionamento de símbolos.

Claramente, o formato e a localização exata dos colares são muito importantes para um bom mapa de colares. No momento do estudo citado, os colares eram criados manualmente, e adicionados ao mapa de entrada: os autores esperavam ser capazes de gerar automaticamente colares adequados para determinados mapas de origem e conjuntos de dados; no entanto, parece ser um problema algorítmico bastante desafiador por si só. Além disso, mapas de colar não possuíam no estudo uma legenda, como é comumente visto em mapas de símbolos; portanto, os usuários podiam apenas julgar tamanhos relativos de símbolos, mas não podiam ler valores absolutos nos mapas.

3.2 Simplificação Automatizada de Formas Geográficas

Um outro estudo avaliado como contribuição para a proposta desta tese foi produzido por BARKOWSKY *et al.* (2000), no qual um algoritmo pode auxiliar na simplificação de formas georreferenciadas, a fim de gerar um mapa esquemático em um grau de abstração adequado ao usuário, e que não viole a ordenação espacial dos elementos cartográficos.

A simplificação de formas em representações cartográficas é utilizada por dois motivos: para abstrair detalhes irrelevantes, a fim de reduzir a carga cognitiva do usuário do mapa, ou para simplificar informações quando um mapa de menor escala é derivado de um mapa detalhado de referência. O método de evolução discreta de curvas, desenvolvido para simplificar as características perceptuais de forma, é explicado; apresentamos um algoritmo que, por um lado, simplifica os dados espaciais, até o grau de abstração pretendido pelo usuário; e que, por outro lado, não viola a ordenação espacial local.

Mapas e representações cartográficas são meios comuns para transmitir conhecimento sobre ambientes espaciais que geralmente não podem ser mapeados como um todo; dependendo da sua escala, um mapa topográfico genérico pretende representar o máximo de informações espaciais possível, já que a finalidade específica para a qual o mapa será usado não é conhecida com antecedência. A execução de uma determinada tarefa, no entanto, geralmente requer apenas um subconjunto bastante pequeno de aspectos; portanto, mapas esquemáticos de uso específico são gerados, adequados apenas para fins restritos, mas que, por outro lado, facilitam sua interpretação, ao se concentrarem em aspectos relevantes da informação, abstraindo-os de outros.

Pretendendo derivar um mapa esquemático com informações reduzidas de formas a partir de dados espaciais detalhados, precisamos de técnicas para reduzir a precisão espacial dos aspectos de forma, a fim de abstrairmos do percurso exato das entidades lineares ou da forma detalhada de objetos geográficos de área. Um passo importante na generalização cartográfica é a simplificação de detalhes de características espaciais (HAKE e GRÜNREICH, 1994); a generalização cartográfica automatizada é uma questão importante de pesquisa na área de sistemas de informação geográfica (SIGs) (para uma visão geral, veja LAGRANGE *et al.* (1995); JONES (2014)).

O feito principal do processo de evolução discreta de curva descrito em BAR-KOWSKY *et al.* (2000) é a simplificação automática de curvas poligonais que permite negligenciar distorções menores, preservando a aparência perceptual (Figura 3.8). Uma definição algorítmica detalhada deste processo é dada em LATECKI e LAKÄMPER (1999a). Uma definição teórica de conjuntos recursivos pode ser encontrada em LATECKI e LAKÄMPER (1999b).

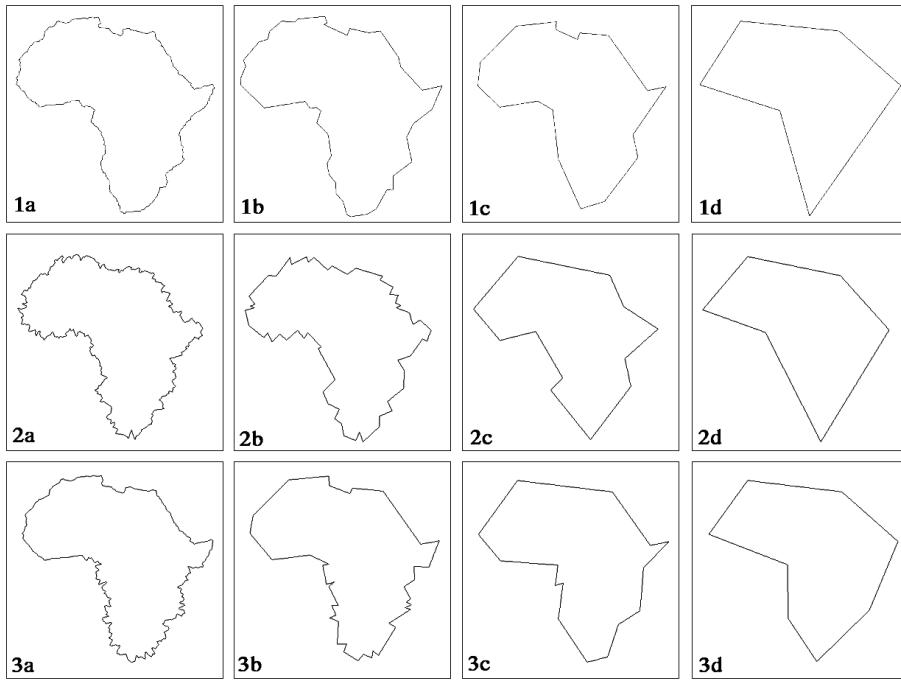


Figura 3.8: Exemplos de evolução discreta de curva; cada linha mostra apenas alguns estágios. Extraído de BARKOWSKY *et al.* (2000).

Entretanto, ao lidar com informações geográficas, o método de evolução da curva discreta precisa ser estendido em vários aspectos. Se os pontos pertencem a mais de um objeto, talvez não queiramos eliminá-los, pois isso pode alterar seriamente a informação espacial: portanto, são introduzidos "pontos fixos", como pontos que não devem ser considerados pelo procedimento de simplificação; seja porque não é possível atribuir-lhes uma medida de relevância, seja porque não devem ser eliminados. Em dados geográficos, objetos pontuais frequentemente estão localizados sobre objetos lineares: nesse caso, a posição dos objetos pontuais terá que ser movida quando o objeto linear ao qual pertencem for modificado (Figura 3.9).

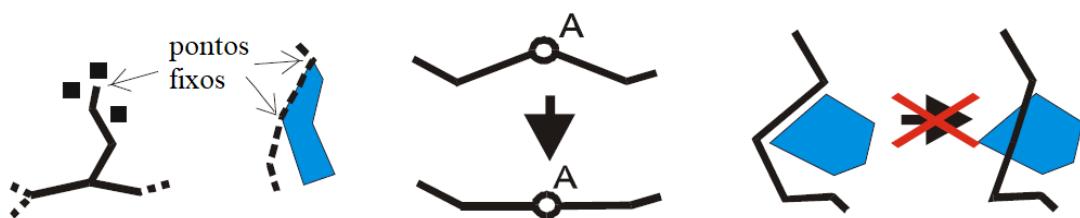


Figura 3.9: Exemplos de casos em dados cartográficos que requerem a extensão do método de evolução discreta de curva. Extraído de BARKOWSKY *et al.* (2000).

Assim, obtemos a seguinte classificação de pontos em um determinado mapa: pontos fixos (pontos que representam características geográficas que não podem ser

removidas, e cuja posição não pode ser alterada), pontos móveis (pontos que representam entidades geográficas que não podem ser removidas, mas cuja posição pode ser alterada) e pontos removíveis (pontos que podem ser excluídos; eles geralmente ocorrem como partes de curvas poligonais e não representam entidades geográficas individuais).

O algoritmo para executar a tarefa de simplificação com base na evolução discreta de curva requer algumas considerações adicionais; no entanto, ainda é bastante simples. Entidades pontuais são consideradas pontos fixos, conforme declarado na seção anterior; no que diz respeito a objetos lineares e de área, começamos calculando as medidas de relevância para todos os pontos (exceto aqueles pontos terminais, que representam a extremidade de alguma entidade linear), e os ordenamos. Em seguida, consideramos o primeiro ponto na lista ordenada (ou seja, o ponto que possui a menor medida de relevância):

Caso A: Coordenadas únicas. Os pontos terminais de objetos lineares são tratados como pontos fixos neste caso, pois nenhuma medida de relevância pode ser computada para eles. Para pontos não-terminais:

1. Se o ponto em questão for um ponto removível, podemos excluí-lo, a menos que isso não viole o arranjo local de entidades espaciais. Portanto, verificamos se há algum outro ponto dentro do triângulo formado pela linha original e pela nova linha retificada: se houver, o ponto em questão é tratado como ponto fixo neste caso; se não houver ponto no triângulo, a linha é retificada, removendo-se o ponto em questão.

2. Se o ponto em questão for um ponto móvel (ou seja, precisar ser preservado), ele será tratado como descrito anteriormente: em vez de removê-lo, no entanto, suas novas coordenadas na reta são calculadas, e o ponto é projetado para essa posição.

Caso B: Múltiplos objetos sobre a coordenada dada; se as coordenadas do ponto pertencem a mais de um objeto, há interações entre pelo menos dois objetos no ponto em questão. Três casos devem ser identificados:

1. Se as coordenadas dos pontos imediatamente vizinhos de todos os objetos envolvidos são idênticas entre si aos pares, então os objetos aos quais os pontos pertencem têm coordenadas comuns em pelo menos três vértices subsequentes; neste caso, o vértice do meio (a posição do ponto em questão) é um ponto removível, desde que a simplificação seja realizada simultaneamente para todos os objetos envolvidos. Portanto, podemos proceder em analogia aos pontos removíveis, conforme descrito para o caso A. Pontos móveis (que não devem ser eliminados de nenhum dos objetos) também podem ser tratados em analogia aos pontos móveis no caso A. Neste caso, todos os pontos nas coordenadas do vértice do meio são transferidos simultaneamente para sua nova posição comum na linha reta.

2. A posição do ponto atual coincide com o ponto terminal de um objeto e, para

os outros objetos, as coordenadas dos pontos imediatamente vizinhos são idênticas entre si aos pares. O respectivo ponto terminal é projetado na reta: assim, embora os pontos terminais ainda não possuam medida de relevância, eles são movidos de acordo com os pontos com os quais coincidem.

3. Em qualquer outro caso, o ponto em questão deve ser tratado como um ponto fixo.

A variante do algoritmo de evolução discreta de curva apresentada nesta contribuição também pode ser usada como um pré-filtro para medidas de similaridade de representações geográficas. O fato de a variante de evolução discreta de curva modificada preservar propriedades qualitativas relevantes é de fundamental importância para esta tarefa, uma vez que a semelhança das representações geográficas não depende apenas das características métricas, mas também de propriedades qualitativas. Medidas de similaridade de representações geográficas possuem muitas aplicações potenciais, por exemplo, em sistemas de informação geográfica.

3.3 Grade Estatística

Uma outra forma de lidar cartograficamente com os obstáculos trazidos pelas formas geográficas representativas dos territórios é apresentada pelas grades regulares. Segundo BUENO (2014), a proposta de criação de um sistema de grades regulares para disseminação de dados estatísticos surge a partir da necessidade de integração de dados de origens diversas e agregados em unidades geográficas incompatíveis, além da necessidade de se ter dados agregados em unidades menores e sem variação ao longo do tempo. Todas essas necessidades se tornam mais evidentes a partir da utilização de ferramentas de geoprocessamento nas análises socioeconômicas e demográficas, devido principalmente à facilidade de se criar unidades artificiais de análise nesses ambientes, como buffers baseados em pontos, linhas ou áreas (RUSANEN *et al.*, 2001). Este sistema não é recente, apesar de ainda ser relativamente pouco utilizado e explorado, mas a sua produção tem se tornado cada vez mais comum, tendo em vista as tecnologias espaciais atualmente disponíveis e empregadas na realização dos censos, como utilização de GPS, geoprocessamento e sensoriamento remoto (UNSD (2000); UNSD (2009)).



Figura 3.10: Captura de tela da aplicação Grade Estatística 2022, disponível em <https://mapasinterativos.ibge.gov.br/grade2022/default.html>.

Um sistema de células regulares dispostas em forma de grade pode ser utilizado como uma base para a geração de um suporte geográfico estável para a disseminação e visualização de dados. Geográfica e socialmente falando, essas unidades são totalmente arbitrárias, não tendo um significado que possa ser transportado para o mundo real, pois não consideram a distribuição de nenhum processo ou fenômeno subjacente (GRASLAND *et al.*, 2006). No entanto, as células servem perfeitamente ao propósito de receptáculo, permanecendo estáveis ao longo do tempo, apresentando uma forma regular e simples, com dimensões suficientemente pequenas para atuar como tijolos na construção de qualquer recorte geográfico desejado (NORMAN *et al.* (2003); GUZMÁN *et al.* (2013)), além de atender às demandas de disseminação de dados para pequenas áreas.

O sistema de grade apresenta uma série de vantagens quando comparado com o sistema tradicional de disseminação de dados censitários através de unidades geográficas que representam divisões político-administrativas ou operacionais, a saber:

- Estabilidade espaço-temporal;
- Adaptação a recortes espaciais;
- Hierarquia e flexibilidade;
- Versatilidade;
- Interpretação cartográfica;

- Identificação simples;
- Utilização em modelagens e
- Minimização dos efeitos do MAUP (Problema da Unidade de Área Modificável).

Uma característica do sistema de grades é a sua independência de recortes político-administrativos comumente utilizados para a agregação de dados censitários, o que traz vantagens relacionadas com a estabilidade temporal. Esta estabilidade dos limites das unidades ao longo do tempo é vista como uma das principais vantagens deste sistema (MARTIN (2000); TAMMILEHTO-LUODE *et al.* (2000); RUSANEN *et al.* (2001); TAMMILEHTO-LUODE (2011)), coisa que não acontece com unidades geográficas diretamente ligadas com a administração do território (municípios, distritos e outras), nem com unidades operacionais ou funcionais (setores censitários). No primeiro caso, isso acontece porque essas unidades estão sujeitas a alterações nos seus limites físicos, que pode ser devido à criação de novas unidades, que surgem à medida que as regiões vão se tornando mais desenvolvidas e, consequentemente, almejam uma maior autonomia no gerenciamento de seu território, ou mesmo pela redefinição dos limites em decorrência de questionamentos legais ou imprecisões conhecidas e que necessitam de correção. No segundo caso, a causa das alterações é inteiramente operacional, uma vez que os limites dos setores censitários é definido de acordo com o quantitativo de população existente em uma determinada área. Assim, à medida que essas áreas vão crescendo em número de moradores, é necessário que haja uma revisão nos limites para que os parâmetros que regulam a coleta de dados nessas unidades sejam mantidos. O caso mais comum é a divisão de um setor em dois ou mais setores, mas, ocasionalmente, pode acontecer a junção de setores.

Esta característica de estabilidade do sistema de grades pode ser utilizada de duas maneiras diferentes na realização de análises temporais. A primeira requer que a agência nacional de estatística adote este sistema e o mantenha nos próximos censos. Hoje, se quisermos realizar uma análise temporal utilizando a malha de setores censitários, primeiro teremos que ajustar esta malha utilizando os dados de comparabilidade divulgados juntamente com os resultados do censo. Esta tarefa não é simples nem rápida, necessitando do emprego de uma grande quantidade de horas de muitos técnicos habilitados. Com a utilização das grades associadas aos dados censitários, esta tarefa pode ser realizada pronta e rapidamente, sem a necessidade de um processamento de dados demorado ou pesado. A segunda maneira de explorar a estabilidade do sistema de grades é utilizá-lo para transformar os dados passados e criar uma base de dados histórica comparável em termos de unidade espacial. Em

ambos os casos, a consequência direta da estabilidade espaço-temporal das unidades espaciais é uma maior rapidez na identificação de mudanças ocorridas em qualquer variável analisada ao longo do tempo.

Outra característica das grades que oferece vantagem em relação às unidades geográficas tradicionais é a sua forma simples, regular e de pequenas dimensões, o que permite que cada célula funcione como um “ladrilho” que pode ser agrupado de diversas formas para se adaptar a diferentes recortes geográficos (TAMMILEHTO-LUODE, 2011) (Figura 3.11). A necessidade de ter dados socioeconômicos e demográficos em unidades espaciais que não sejam as tradicionais unidades político-administrativas é cada vez maior, tendo em vista a crescente necessidade de realização de avaliações e pesquisas envolvendo áreas vulneráveis com relação a algum aspecto físico, ambiental e/ou social. Essa demanda vem sendo feita desde o final dos anos 80 e começo dos anos 90, quando começa a surgir uma maior consciência do relacionamento entre população e meio ambiente, e continua até hoje, uma vez que a mesma não foi devidamente atendida em todos os níveis de precisão necessários. A área ambiental utiliza com frequência unidades espaciais derivadas de características físicas do meio ambiente, como bacias hidrográficas ou unidades de relevo. Isso causa alguns problemas quando se deseja integrar a dimensão humana aos estudos ambientais, pois geralmente os indicadores sociais são disponibilizados em unidades político-administrativas. Referências a essas questões podem ser vistas, por exemplo, em HOGAN (1992), DEICHMANN (1996), OJIMA e MARTINE (2012) e MARTINE e SCHENSUL (2013).



Figura 3.11: Grade Estatística oficial do Brasil no nível L0 (composta por células que apresentam 500 km de lado), disponível em <https://dadosabertos.social/t/que-tal-o-uso-mais-amplo-da-nossa-grade-estatistica-oficial/317>.

Independentemente das relações entre população e ambiente, existe a necessidade de integração de dados disponibilizados em unidades não coincidentes em quase todas as áreas do conhecimento que utilizam dados estatísticos, sejam eles derivados de pesquisas domiciliares ou de registros administrativos. A área da saúde sofre bastante com estas questões de integração de dados, uma vez que geralmente suas análises envolvem unidades espaciais que representam regiões de abrangência dos serviços de saúde ou de prevalência ou ocorrência de determinadas doenças e que não apresentam nenhuma relação com unidades político-administrativas. A área da saúde sempre investiu muito em modelos espaço-temporais, pois desde muito tempo se sabe da estreita relação entre a disseminação de doenças e a distribuição e o movimento das populações no tempo e no espaço (HAGERSTRAND, 1968). Talvez por isso, esta área do conhecimento esteja entre aquelas que buscam avançar tecnicamente na busca de soluções metodológicas para enfrentar os problemas aqui discutidos, como pode ser visto em LINARD *et al.* (2010) e em LINARD e TATEM (2012).

Os sistemas de grade com base em células regulares oferecem uma estrutura hierárquica, de forma que a união de um conjunto de células adjacentes leve à formação de uma célula maior, pertencente a um nível hierárquico superior. Esta estrutura

hierárquica dá flexibilidade ao sistema, permitindo que, caso se queira, diversos níveis hierárquicos possam conviver dentro do mesmo sistema, ou seja, células de dimensões diferentes possam fazer parte do mesmo sistema. No caso de uma grade estatística, é usual que se tenham células com dimensões diferentes cobrindo áreas urbanas e áreas rurais, com o objetivo de melhor representar a distribuição da população. Esta característica hierárquica do sistema de grade permite também que células menores das áreas urbanas possam ser agregadas até atingirem as dimensões das células maiores das áreas rurais e que, então, este novo arranjo celular seja utilizado nas análises; e mesmo as células maiores podem ser agregadas em outras ainda maiores, se assim for necessário. Isso demonstra o quanto o sistema pode ser flexível, com capacidade de se adaptar a diferentes escalas, além de todas as outras características vantajosas do sistema, como comparabilidade temporal e espacial.

A Figura 3.12 mostra uma estrutura hierárquica de células. Nesta figura, podemos observar que o agrupamento de células adjacentes de um mesmo nível forma uma célula de maior dimensão, pertencente ao nível imediatamente superior. Por exemplo, as células menores no canto superior direito 1, 2, 3 e 4 somadas formam uma célula maior, a célula D. Desta maneira, pode ser construído um sistema de grades que conjugue células de dimensões diferentes, de modo que as células menores se encaixem perfeitamente nas maiores. Grades com células de tamanhos diferentes podem ser utilizadas para diferenciar áreas rurais e urbanas, permitindo que regiões mais densas sejam mais detalhadas do que regiões menos povoadas. Isto também oferece ao usuário a opção de ter toda a região em estudo representada por células de mesmo tamanho, bastando que para isso, seja feita uma agregação das células menores, inclusive dos seus atributos, até que atinjam a dimensão das células maiores.

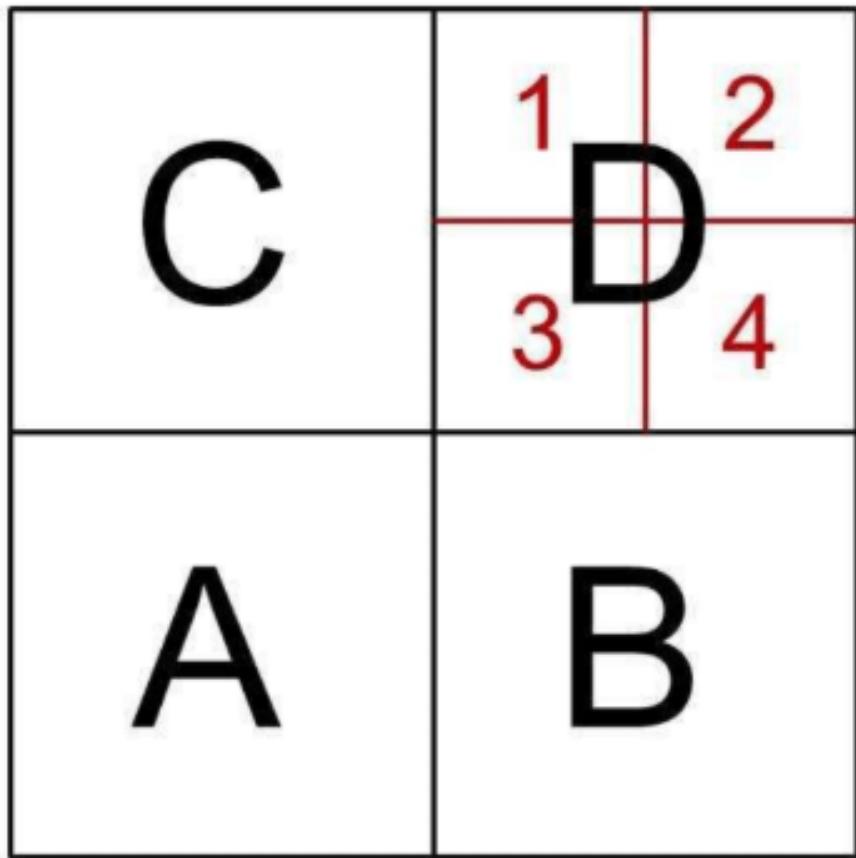


Figura 3.12: Estrutura hierárquica de células. Extraído de BUENO (2014).

Um sistema de grade pode ser construído utilizando tanto uma estrutura de dados vetorial quanto matricial, e também pode ser facilmente convertido de uma estrutura para outra. Quando a abordagem metodológica utilizada para a geração da grade é a agregação de dados pontuais, a estrutura usualmente empregada é a vetorial; quando a abordagem metodológica utilizada é a desagregação de dados disponibilizados em unidades areais, a estrutura comumente empregada no produto final é a matricial ou raster. A vantagem de se ter a grade em uma estrutura vetorial é que todas as variáveis existentes podem ser alocadas na tabela de atributos e relacionadas com a geometria, facilitando as análises conjuntas dessas variáveis (STRAND e BLOCH, 2009). No caso da estrutura matricial, é necessária a geração de um arquivo para cada variável ou da associação aos dados espaciais de apenas uma variável a cada vez.

Em termos cartográficos, mas sem esquecer que a escala é um fator importante, a utilização de células regulares e com dimensões iguais ou comparáveis em vez da utilização de unidades irregulares e com tamanhos diversos oferece ao leitor uma forma mais legível de interpretação do fenômeno que está sendo representado (GRASLAND et al., 2006). Isto é uma grande vantagem, principalmente quando o leitor não está

familiarizado com as nuances que a interpretação de um mapa pode oferecer.

A identificação única das células de uma grade pode ser feita de diferentes maneiras, tendo sempre como objetivo a simplicidade e, preferencialmente, permitindo que o posicionamento espacial esteja explícito na própria nomenclatura. No caso de o sistema ser baseado em Sistemas Cartográficos, existem regras e normas estabelecidas, e elas podem ser replicadas ou adaptadas à grade estatística. Outro modo de aplicar uma nomenclatura é a utilização direta das coordenadas de um dos cantos da célula, indicando o tamanho da mesma; esse método é semelhante ao sistema utilizado para definir o posicionamento espacial de uma imagem orbital em programas que trabalham com Sensoriamento Remoto (world file); ou de modo similar, porém utilizando as coordenadas do centro da célula ao invés de um dos cantos. O Infrastructure for Spatial Information in the European Community - INSPIRE sugere a seguinte composição para identificação da célula: a dimensão da célula, seguida pelas coordenadas vertical e horizontal do canto inferior esquerdo; por exemplo, 1kmN2599E4695, que identifica uma célula de 1 km, com coordenadas do canto inferior esquerdo $Y = 2.599.000\text{ m}$ e $X = 4.695.000\text{ m}$ (INSPIRE, 2023).

Muitos modelos de simulação que utilizam dados socioeconômicos ou demográficos no seu processamento, o fazem sob a forma de grade e, usualmente, no formato matricial. Isso se deve ao fato de que a estrutura de grades tem uma geometria simples e computacionalmente eficiente, e sob a forma de matriz, facilita os numerosos cálculos interativos que o modelo necessita executar, poupando tempo de máquina. Este fato já era visto como uma vantagem da utilização do sistema de grades no final dos anos 90 (TOBLER *et al.*, 1997), quando ainda não eram amplamente utilizadas as técnicas de simulação espacial hoje aplicadas. Temos, por exemplo, os modelos Cellular Automata que têm sido bastante utilizados na simulação de fenômenos urbanos, como crescimento regional, urban sprawl e evolução do uso das terras. Esses modelos, como o próprio nome indica, utilizam células distribuídas num espaço gradeado como um de seus componentes.

O problema da unidade de área modificável (MAUP) é uma fonte de viés estatístico que pode impactar significativamente os resultados de testes de hipóteses estatísticas. O MAUP afeta os resultados quando medidas pontuais de fenômenos espaciais são agregadas em partições espaciais ou unidades de área (como regiões ou distritos), como, por exemplo, densidade populacional ou taxas de doenças. Os valores resumidos resultantes (por exemplo, totais, taxas, proporções, densidades) são influenciados tanto pela forma quanto pela escala da unidade de agregação.

Por exemplo, dados censitários podem ser agregados em distritos de condados, setores censitários, áreas de código postal, delegacias de polícia ou qualquer outra partição espacial arbitrária. Assim, os resultados da agregação de dados dependem da escolha do cartógrafo sobre qual "unidade de área modificável" usar em sua aná-

lise. Um mapa coroplético censitário que calcula a densidade populacional usando limites estaduais produzirá resultados radicalmente diferentes de um mapa que calcula a densidade com base nos limites de condados. Além disso, os limites dos distritos censitários também estão sujeitos a alterações ao longo do tempo, o que significa que o MAUP deve ser considerado ao comparar dados anteriores com dados atuais.

OPENSHAW e RAO (1995) contestam que a agregação de dados no menor nível geográfico possível seja uma solução para minimizar os impactos causados pelo problema da área modificável. No entanto, esse procedimento é defendido por outros estudiosos, como GOODCHILD (1992), SOUZA e TORRES (2003), SCHUURMAN *et al.* (2006) e MONTEIRO *et al.* (2004), apenas para citar alguns. Por outro lado, uma das soluções propostas para o tratamento do MAUP pelo próprio OPENSHAW (1977) discorre sobre o desenvolvimento de uma metodologia que permita criar regiões adequadas para o estudo de determinado fenômeno, as unidades ideais ou ótimas. Entendendo o sistema de grade como um conjunto de unidades básicas de pequenas dimensões, ele pode ser utilizado para realizar a agregação que conduz à geração de uma unidade maior e mais adequada ao fenômeno estudado (GRASLAND *et al.*, 2006); assim, a utilização de grades pode ser considerada uma forma de minimizar os efeitos do MAUP.

De acordo com WRIGLEY (1995), os dois aspectos do MAUP ficam mais críticos quando se diminui o número de unidades, ou seja, à medida que a quantidade de unidades diminui e, consequentemente, a dimensão das mesmas aumenta, a severidade do MAUP tende a crescer. Isso significa que, ao compararmos um sistema de grade que tem uma grande quantidade de pequenas células com um sistema de polígonos administrativos com poucas unidades de grande extensão, podemos deduzir que o MAUP se apresenta menos crítico no primeiro caso do que no segundo.

Assim, os estudos citados neste capítulo apresentam conceitos fundamentais e implementações de algoritmos que fornecerão a base para o desenvolvimento de novas soluções de cartogramas temáticos, conforme as propostas a serem descritas no capítulo a seguir.

Capítulo 4

Novas propostas para um mapa temático do Brasil

Devido às características específicas de sua representação em termos de vizinhança e disparidade de tamanho das unidades da federação, o ponto de partida na busca por novas formas de representação foi o próprio mapa do território brasileiro, com o objetivo de formular soluções para tal problema representacional singular. Posteriormente, buscou-se aplicar as soluções propostas a outros territórios e suas respectivas divisões político-administrativas, a fim de validar sua aplicação em diferentes condições, considerando que, devido às limitações impostas por cada proposta, não há como apresentar uma solução única que ofereça simultaneamente a todos os territórios mapeáveis um cartograma de representação perfeita conforme os critérios estabelecidos por BARRETO *et al.* (2018).

4.1 Carto-centric

A primeira proposta se baseia no conceito de concentricidade geográfica: escolhe-se uma unidade geográfica adequada para ocupar a posição central do cartograma e, com base na vizinhança e na distância, as demais unidades são distribuídas ao longo do perímetro de um ou mais círculos, concêntricos a essa unidade central de referência.

Sob tal proposta, tendo em vista uma adequada adaptação ao território brasileiro e o conceito de centralidade geográfica expresso no projeto e na construção da nova capital, Brasília, fundada em 1960, o Distrito Federal foi identificado quase que de maneira natural como a Unidade da Federação mais adequada para ocupar a posição central em um cartograma que corresponda ao território do Brasil e suas respectivas unidades administrativas.

O passo seguinte é arbitrar as unidades que comporão as camadas ou níveis

concêntricos a partir do entorno da posição central, as quais neste exemplo foram determinadas conforme os critérios listados a seguir, e apresentadas visualmente na Figura 4.1:

- Unidades geográficas mais próximas da unidade central, mas sem litoral ou fronteiras com outros países. Esta camada é composta por três unidades: Goiás, Minas Gerais e Tocantins;
- Unidades geográficas que fazem divisa com as unidades anteriores. Esta segunda camada é composta por quinze unidades: Alagoas, Bahia, Ceará, Espírito Santo, Maranhão, Mato Grosso, Mato Grosso do Sul, Pará, Paraíba, Pernambuco, Piauí, Rio de Janeiro, Rio Grande do Norte, São Paulo, Sergipe;
- Camadas subsequentes, compostas por unidades geográficas que fazem fronteira com as unidades anteriores. Assim, a terceira camada será composta por Amapá, Amazonas, Paraná e Rondônia; a quarta camada será composta por Acre, Roraima e Santa Catarina; e a quinta camada será composta por uma única unidade federativa, o Rio Grande do Sul.

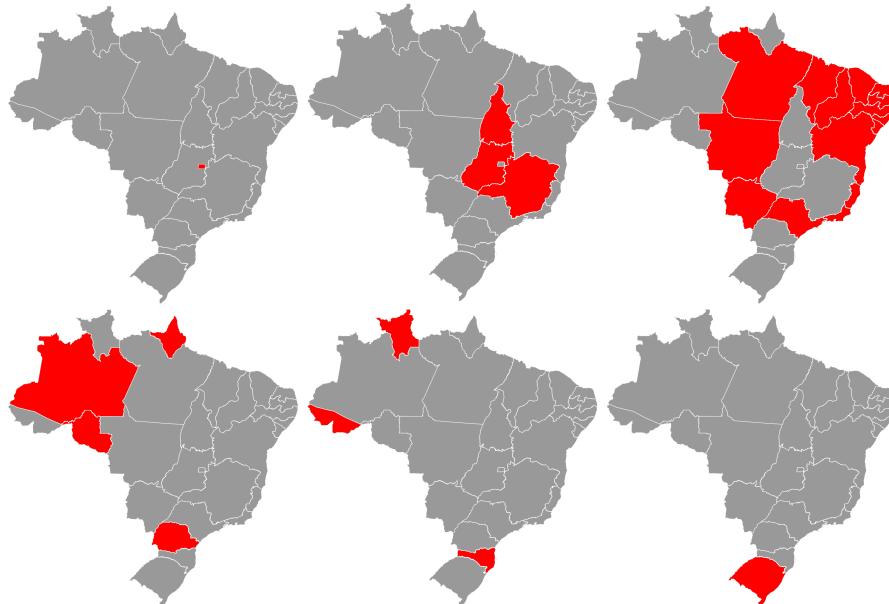


Figura 4.1: Níveis de concentricidade com suas respectivas Unidades da Federação.

Após a coleta e o estabelecimento desses aspectos, realiza-se a distribuição espacial dos elementos sobre o perímetro do círculo que define cada camada do cartograma, a fim de aproximá-los de suas posições relativas aos elementos vizinhos. Assim, para fins computacionais, algumas variáveis gerais precisam ser estabelecidas para a construção do cartograma:

- a forma e o tamanho básicos de cada elemento;
- a posição do elemento central do cartograma;
- o passo de distância para cada nível externo, em relação ao elemento central.

Para representar computacionalmente os aspectos relacionados a cada elemento, foi utilizado um conjunto de tuplas: uma tupla, entendida aqui em sua definição amplamente aceita na programação, é uma estrutura de dados leve que possui um número específico e finito, e uma sequência de valores ordenados por um índice numérico e inteiro; pode também ser entendida como uma lista ordenada, com um número finito de elementos. Além disso, os valores da tupla geralmente podem ser duplicados, e ser de qualquer tipo de dado. Os elementos de uma tupla também são tipicamente imutáveis — não podem ser alterados individualmente.

O objetivo do uso de um conjunto de tuplas é tornar a representação mais econômica em termos de espaço de memória, quando comparada aos formatos tradicionais de mapas vetoriais: essa característica tende a ser interessante quando pensamos em dispositivos de campo para tarefas de coleta de dados de pesquisa, que dispõem de capacidade limitada de armazenamento e processamento de dados, bem como telas de dimensões reduzidas e de baixa resolução.

O conjunto de tuplas gerado para o exemplo do território brasileiro seguindo o procedimento descrito acima é apresentado na Figura 4.2. Cada uma das tuplas correspondentes aos elementos do cartograma é composta pelos seguintes valores numéricos:

- um identificador único (o qual pode ser, em alternativa a um número inteiro, um caractere ou uma string de texto);
- o número correspondente ao nível de concentricidade sobre o qual o elemento será posicionado;
- a posição angular do elemento sobre aquele nível; e, opcionalmente,
- os dados a serem representados pelo cartograma.

```

const ufs =[[11,"RO",3,180],
[12,"AC",4,180],
[13,"AM",3,150],
[14,"RR",4,150],
[15,"PA",2,150],
[16,"AP",3,120],
[17,"TO",1,120],
[21,"MA",2,120],
[22,"PI",2,100],
[23,"CE",2,80],
[24,"RN",2,60],
[25,"PB",2,40],
[26,"PE",2,20],
[27,"AL",2,0],
[28,"SE",2,340],
[29,"BA",2,320],
[31,"MG",1,240],
[32,"ES",2,290],
[33,"RJ",2,265],
[34,"SP",2,240],
[41,"PR",3,240],
[42,"SC",4,240],
[43,"RS",5,240],
[50,"MS",2,210],
[51,"MT",2,180],
[52,"GO",1,180],
[53,"DF",0,0]
];

```

Figura 4.2: Conjunto de tuplas para cartograma de concentricidade relacionado às Unidades da Federação.

A partir deste ponto, denominaremos esta primeira proposta, cujo cartograma gerado é apresentado na Figura 4.3, como “carto-centric”, em vista de sua principal característica visual: a concentricidade.

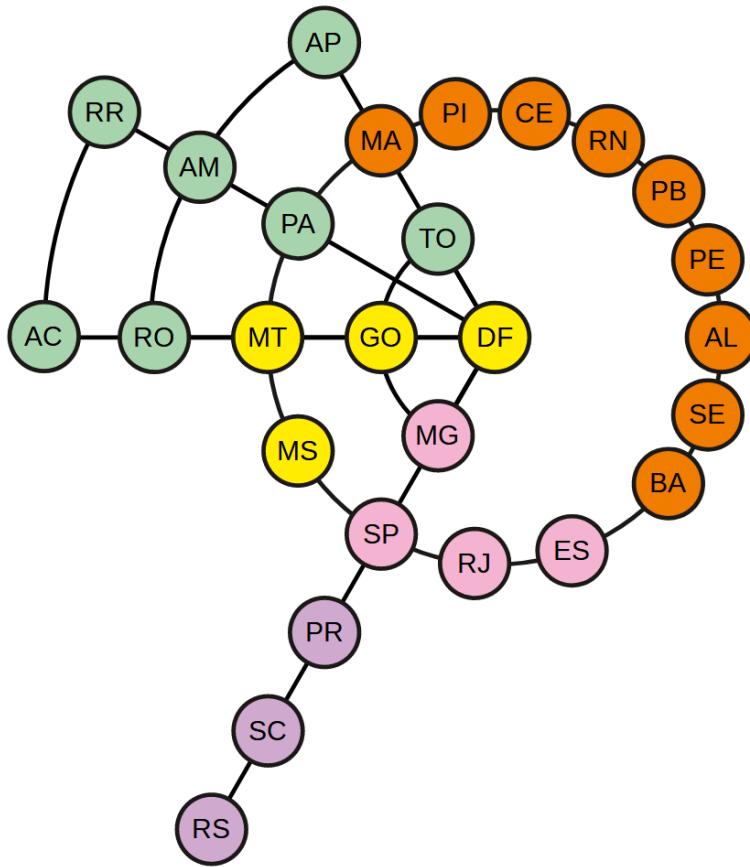


Figura 4.3: Cartograma de concentricidade (carto-centric) representando o Brasil e as Unidades da Federação.

Para verificar a validade desta proposta em cartogramas referentes a outras "áreas geográficas", o mesmo conceito de concentricidade foi experimentado com o mapa do estado do Rio de Janeiro: além do maior número de unidades geográficas (92 municípios) em comparação ao território brasileiro (27 Unidades da Federação), ocorre aqui a situação mais frequente em que, diferentemente do território brasileiro, a unidade administrativa central (a capital) não corresponde a uma posição territorial central. Nesse caso, dois cartogramas diferentes podem ser criados, a fim de avaliar qual será visualmente mais apropriado: um cartograma concêntrico

à capital (Figura 4.4), e outro cartograma concêntrico a alguma unidade arbitrária mais próxima do centroide do território (Figura 4.5), tornando-se necessário neste segundo caso utilizar algum recurso visual adicional para distinguir a capital em relação às outras unidades; nos exemplos apresentados, uma cor distinta de contorno foi aplicada à unidade específica da capital.

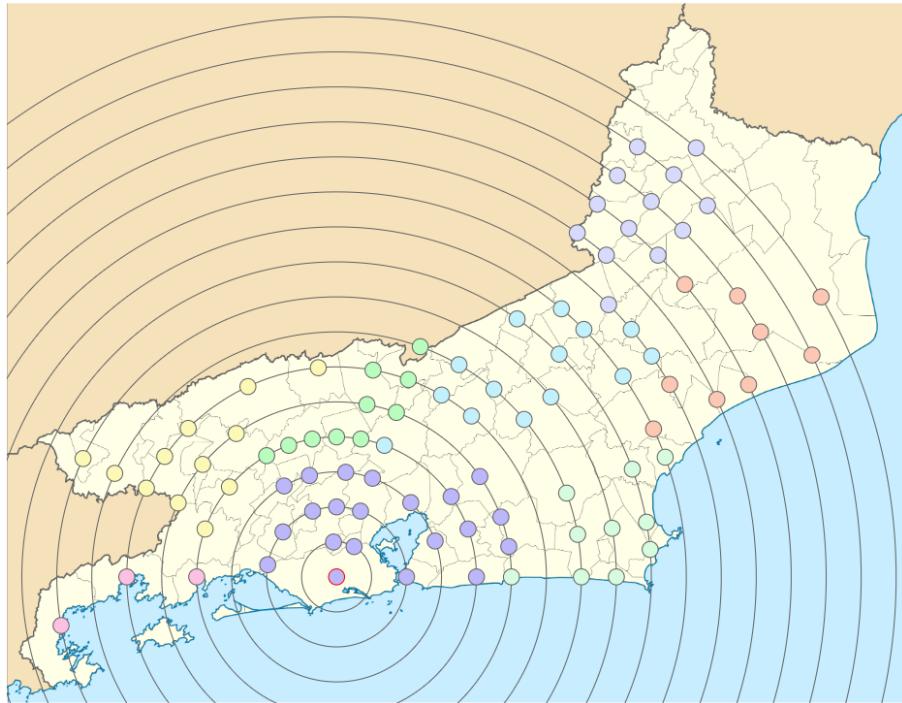


Figura 4.4: Cartograma de concentricidade (carto-centric) representando os municípios do estado do Rio de Janeiro, com as regiões de governo identificadas por cores distintas e o centro do cartograma sobre a capital.

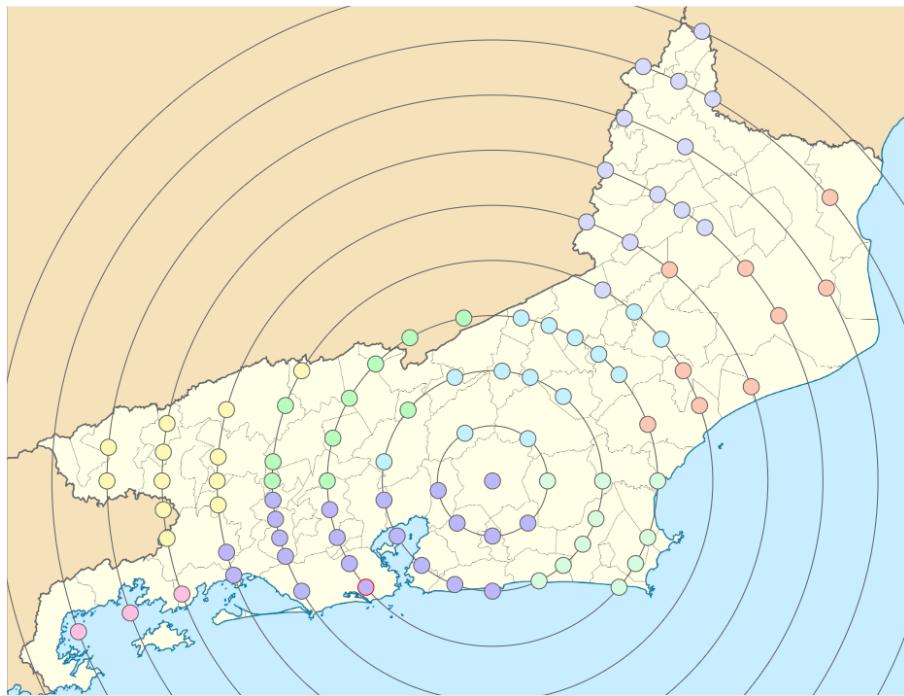


Figura 4.5: Cartograma de concentricidade (carto-centric) representando os municípios do estado do Rio de Janeiro, com as regiões de governo identificadas por cores distintas e o centro do cartograma sobre o município de Cachoeiras de Macacu.

Observa-se que a concentricidade centrada na capital produziu um cartograma representado por mais níveis, visto que em diversos territórios a capital tende a agregar, em seu entorno, unidades de menor tamanho e maior quantidade.

É possível também vislumbrar a aplicação do conceito de elementos concêntricos em termos de meios de visualização menos convencionais do que os suportes bidimensionais de uma folha de papel ou uma tela de dispositivo eletrônico: para os eventos produzidos pelo Instituto Brasileiro de Geografia e Estatística, uma proposta interessante envolveria o uso de um conjunto de colunas de líquido iluminadas por LEDs, conforme a Figura 4.6, com a altura do líquido dentro da respectiva coluna de cada elemento sendo controlada dinamicamente com base nos valores dos dados estatísticos apresentados para aquele elemento.



Figura 4.6: Colunas de líquido iluminadas por LED. Extraído de <http://www.fsbaojiali.cn/>.

4.2 Carto-step

A proposta anterior traz o benefício de representar todas as unidades por meio de símbolos de forma e tamanho uniformes, mas ainda é possível pensar em soluções que limitem as diferenças dimensionais apenas a uma escala restrita de tamanhos, a fim de tornar todas as unidades visíveis simultaneamente e ainda proteger algum caráter proporcional e a forma geral do território representado.

Assim, a segunda proposta sugere a representação de cada entidade geográfica por meio de não apenas uma, mas de duas formas básicas concêntricas semelhantes: uma interna, de preenchimento sólido, possuindo área única e idêntica para todas as entidades; e outra externa, apresentada em contorno, cuja área é proporcionalmente relacionada à área geográfica real, porém restrita a uma escala de tamanhos válidos arbitrariamente escolhida. O elemento composto resulta semelhante a uma pirâmide truncada vista de cima, com a forma externa sendo associada à base da pirâmide e a forma interna correspondendo ao plano de corte, proporcionando assim um "degrau"imaginário entre as duas formas.

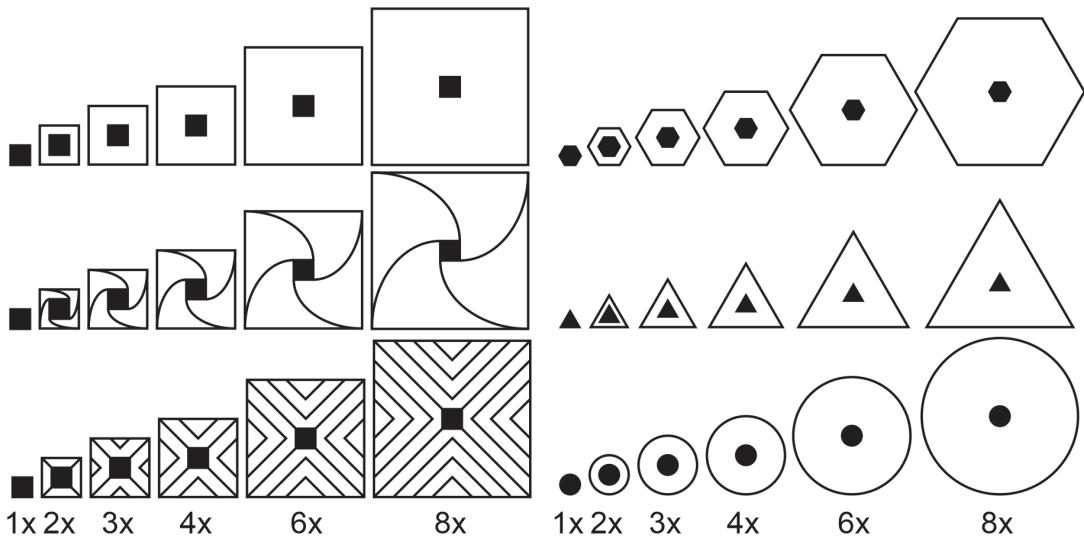


Figura 4.7: Sugestões de escalas para formas concêntricas básicas.

Nesta proposta, buscamos inicialmente identificar, para cada entidade geográfica, os seguintes aspectos-chave:

- seu centroide, a fim de definir seu posicionamento global no cartograma.
- sua área, a fim de definir, conjuntamente às áreas de todas as outras unidades, a escala global e o número de tamanhos discretos a serem aplicados na representação das distintas unidades no cartograma.

Após a coleta e estabelecimento desses aspectos, realiza-se então a distribuição espacial dos elementos no cartograma, que pode seguir a posição original dos centroides, ou aproxima-los de uma grade bidimensional regular. Não deve haver sobreposição das formas internas sólidas: é permitida, contudo, a sobreposição e o entrelaçamento das formas externas de contorno dos elementos.

Da mesma forma, para fins computacionais, algumas variáveis gerais precisam ser estabelecidas; estas são, nesta proposta, a forma básica e a dimensão interna de cada elemento, além da escala que compõe o conjunto estabelecido de valores discretos para as formas externas de contorno, em relação às formas sólidas internas.

Para representar computacionalmente os aspectos relacionados a cada elemento, foi utilizado um novo conjunto de tuplas; o conjunto de tuplas gerado para o exemplo do território brasileiro, seguindo o procedimento acima descrito, é apresentado na Figura 4.8. Cada uma das tuplas do conjunto é composta pelos seguintes valores numéricos:

- um identificador único (o qual pode ser, em alternativa a um número inteiro, um caractere ou uma string de texto);
- a posição central do elemento, a ser compartilhada pela forma interna sólida e pela forma externa em contorno;
- o valor discreto que multiplica a dimensão da forma interna para obter a dimensão externa do elemento combinado; e, opcionalmente,
- os dados a serem representados pelo cartograma.

```
const ufs = [
  [11, "RO", 15.5, 10, 4],
  [12, "AC", 11.5, 9, 4],
  [13, "AM", 15, 5.5, 8],
  [14, "RR", 16.5, 1.5, 4],
  [15, "PA", 22, 5.5, 6],
  [16, "AP", 22.5, 2, 3],
  [17, "TO", 25, 9, 4],
  [21, "MA", 27, 6.5, 4],
  [22, "PI", 29.5, 7, 3],
  [23, "CE", 31, 7.5, 2],
  [24, "RN", 32.5, 7, 1],
  [25, "PB", 33.5, 8, 1],
  [26, "PE", 32.5, 9, 1],
  [27, "AL", 33.5, 10, 1],
  [28, "SE", 32.5, 11, 1],
  [29, "BA", 29, 11.5, 6],
  [31, "MG", 27, 14.5, 4],
  [32, "ES", 30, 15.5, 2],
  [33, "RJ", 28, 17.5, 2],
  [34, "SP", 25, 17, 4],
  [41, "PR", 22, 18.5, 2],
  [42, "SC", 23.5, 20, 2],
  [43, "RS", 21.5, 22, 3],
  [50, "MS", 20.5, 16, 4],
  [51, "MT", 20.5, 11.5, 6],
  [52, "GO", 23.5, 13, 4],
  [53, "DF", 25, 13, 1]
];
```

Figura 4.8: Conjunto de tuplas para o cartograma de degraus referente às Unidades da Federação.

A partir deste ponto, denominaremos esta segunda proposta, apresentada na Figura 4.9, como “carto-step”, tendo em vista sua característica fundamental de combinar duas formas semelhantes e concêntricas, que podem ser visualmente compreendidas como um “degrau” originado por uma pirâmide regular truncada, ou seja, “cortada” em seu topo por um plano paralelo à base; favorecendo, assim, a

representação de mais uma variável pelo cartograma, através da aplicação da tridimensionalidade sugerida pelo sólido construído.

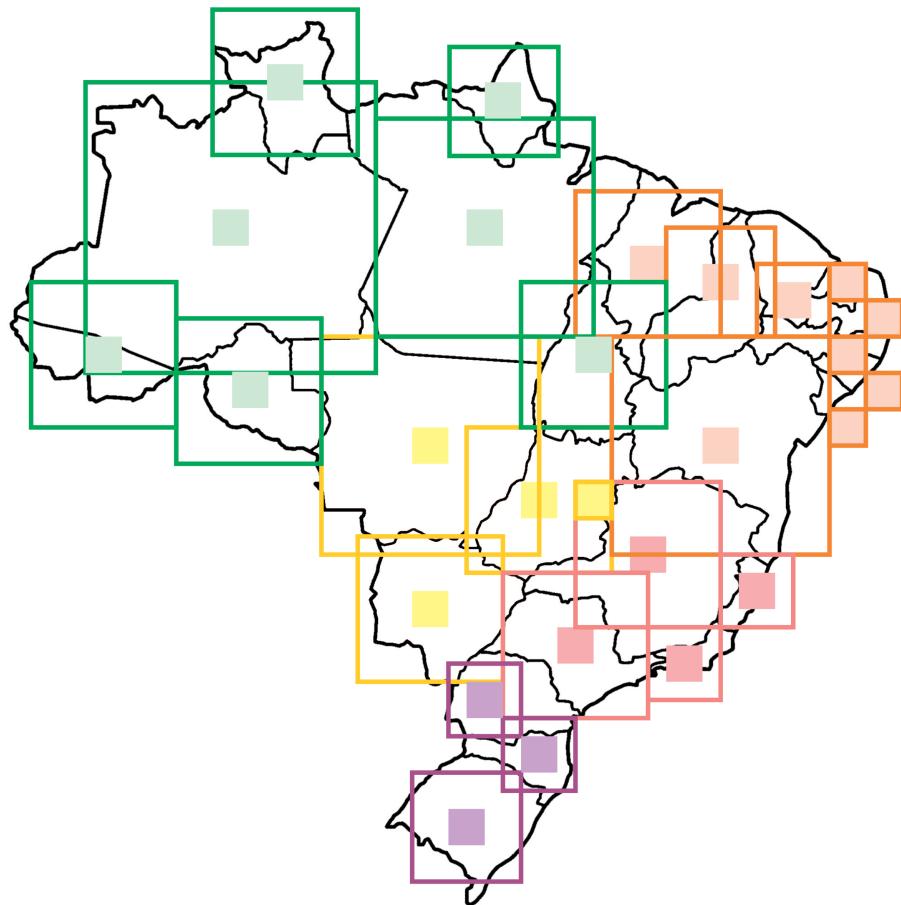


Figura 4.9: Cartograma de degrau (carto-step) representando o Brasil e as Unidades da Federação.

Ainda seguindo o procedimento acima descrito, observa-se que, na representação gerada para as Unidades da Federação divididas pelas respectivas Grandes Regiões e apresentada na Figura 4.10, o formato geral das Grandes Regiões é mantido, conforme a divisão oficialmente proposta.

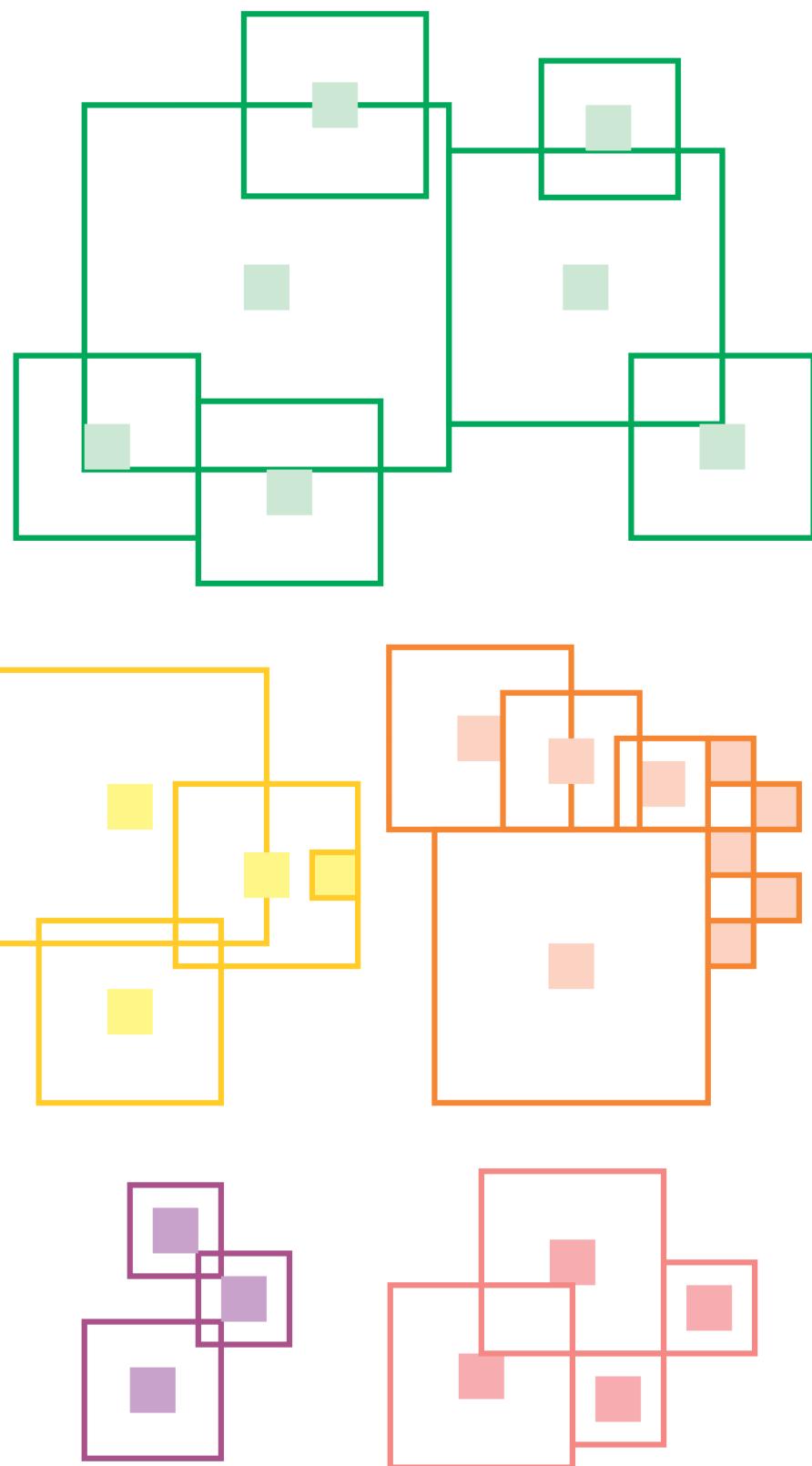


Figura 4.10: Cartograma de degrau (carto-step) aplicado separadamente às Grandes Regiões brasileiras.

Para verificar sua validade em cartogramas referentes a outras "áreas geográficas", o mesmo conceito de composição de formas foi experimentado com o mapa do

Estado do Rio de Janeiro (Figura 4.11). Aqui, a mesma sugestão de escala de tamanhos discretos, contendo valores de 1x a 8x, utilizada para o território brasileiro pode ser aplicada preservando a visualização de todas as unidades; outras escalas também podem ser adotadas, dependendo das características de cada território representado.

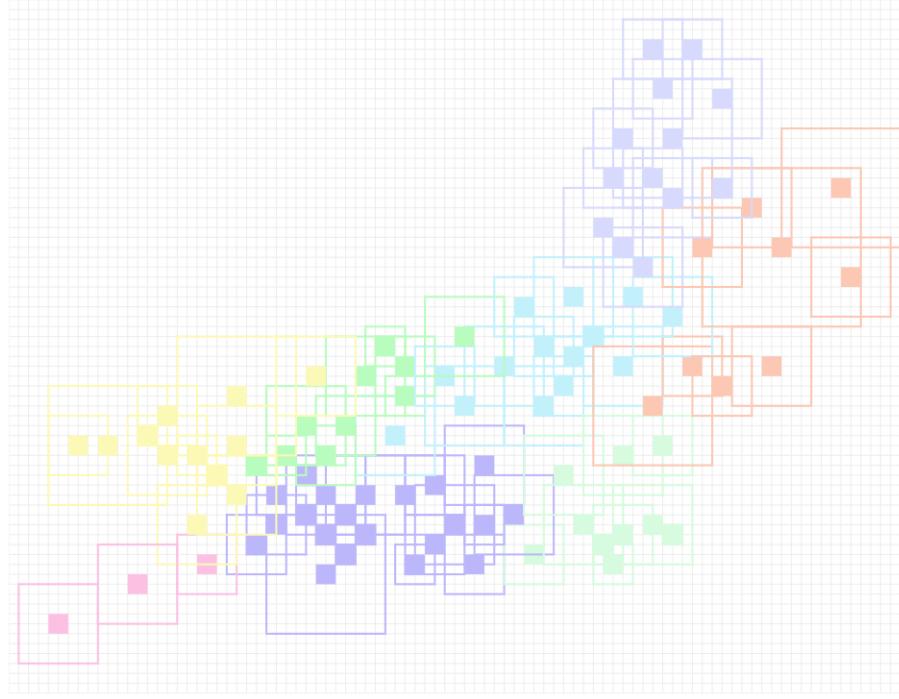


Figura 4.11: Cartograma de degrau (carto-step) representando os municípios do estado do Rio de Janeiro, com as regiões de governo identificadas por cores distintas.

É possível também vislumbrar a aplicação do conceito de elementos em degrau em termos de meios de visualização menos convencionais do que os suportes bidimensionais de uma folha de papel ou uma tela de dispositivo eletrônico: para os eventos produzidos pelo Instituto Brasileiro de Geografia e Estatística, uma proposta interessante envolveria o uso de um conjunto de colunas tridimensionais, revestidas com uma malha elástica desde o topo de dimensão fixa até a base de dimensão específica ao elemento, com a altura da coluna de cada elemento sendo controlada dinamicamente com base nos valores dos dados estatísticos apresentados para aquele elemento.

4.3 Step-centric

Vale ressaltar que as duas propostas aqui apresentadas não são excludentes; portanto, seus conceitos básicos característicos, concentricidade e degraus, poderiam ser combinados em um cartograma que apresentasse ambos os aspectos simultaneamente, produzindo o que poderia ser denominado um cartograma 'step-centric'.

```

const ufs =[[11,"RO",3,180,4],
            [12,"AC",4,180,4],
            [13,"AM",3,150,8],
            [14,"RR",4,150,4],
            [15,"PA",2,150,6],
            [16,"AP",3,120,3],
            [17,"TO",1,120,4],
            [21,"MA",2,120,4],
            [22,"PI",2,100,3],
            [23,"CE",2,80,2],
            [24,"RN",2,60,1],
            [25,"PB",2,40,1],
            [26,"PE",2,20,1],
            [27,"AL",2,0,1],
            [28,"SE",2,340,1],
            [29,"BA",2,320,6],
            [31,"MG",1,240,4],
            [32,"ES",2,290,2],
            [33,"RJ",2,265,2],
            [34,"SP",2,240,4],
            [41,"PR",3,240,2],
            [42,"SC",4,240,2],
            [43,"RS",5,240,3],
            [50,"MS",2,210,4],
            [51,"MT",2,180,6],
            [52,"GO",1,180,4],
            [53,"DF",0,0,1]
];

```

Figura 4.12: Conjunto de tuplas para o cartograma 'step-centric' referente às Unidades da Federação.

Ressalta-se que a simples aplicação de uma tupla contendo os mesmos valores utilizados em alguma das duas propostas separadas (Figura 4.12) pode não oferecer a solução mais adequada visualmente (Figura 4.13): os valores devem ser submetidos a revisão e ajustes, a fim de proporcionar a esta representação uma melhor adaptação espacial.

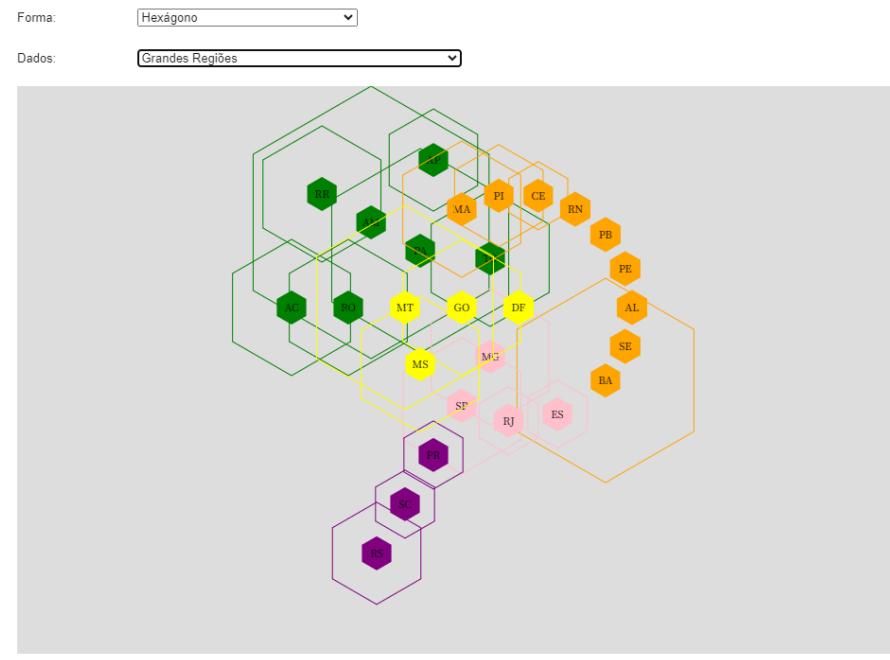


Figura 4.13: Cartograma 'step-centric' representando o Brasil e as Unidades da Federação, utilizando os valores de posicionamento produzidos para a tupla correspondente ao cartograma de concentricidade (carto-centric) e os valores de dimensionamento produzidos para a tupla correspondente ao cartograma de degrau (carto-step).

Capítulo 5

Implementações iniciais

As propostas apresentadas acima foram inicialmente implementadas no OBSERVABLE (2024), uma plataforma online de visualização de dados baseada em tecnologias para renderização, exibição e interação de páginas web, incluindo Javascript (WIKIPEDIA, 2025b) e bibliotecas baseadas em Javascript como a D3.js (BOSTOCK e OBSERVABLE, INC., 2025) para programação do lado do cliente, além de HyperText Markup Language (HTML) em sua versão 5 (W3C, 2025b) e Cascading Styles Sheets (CSS) (W3C, 2025a) para estruturação e exibição; na interface desenvolvida dentro da plataforma Observable, e acessível para interação por meio da Universal Resource Locator (URL) <https://observablehq.com/@lp>, é possível alternar diretamente entre os conjuntos de dados por meio de uma caixa de seleção.

As Figuras 5.1, 5.2 e 5.3 a seguir apresentam cartogramas gerados com base em dados obtidos do Censo Demográfico de 2022 realizado pelo IBGE, e utilizando o conceito carto-centric. Foram geradas novas tuplas contendo os valores das variáveis estatísticas associadas a cada unidade georreferenciada, e o conjunto de valores foi aplicado a uma escala monocromática de tonalidades, partindo do branco como valor mínimo em direção a uma cor selecionada arbitrariamente como valor máximo, para aplicação no preenchimento de uma das formas vetoriais simples sugeridas (triângulo, quadrado, hexágono, octógono ou círculo) e criadas para exibição por meio do recurso "canvas" da linguagem de marcação HTML 5 (WIKIPEDIA, 2025a) em modo bidimensional.

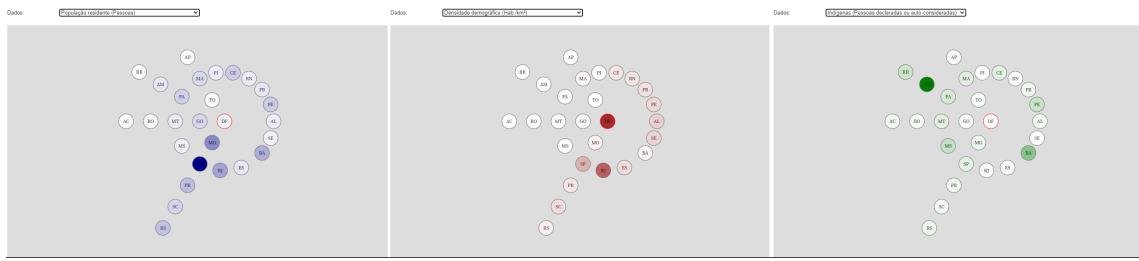


Figura 5.1: Carto-centric apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos às Unidades da Federação.

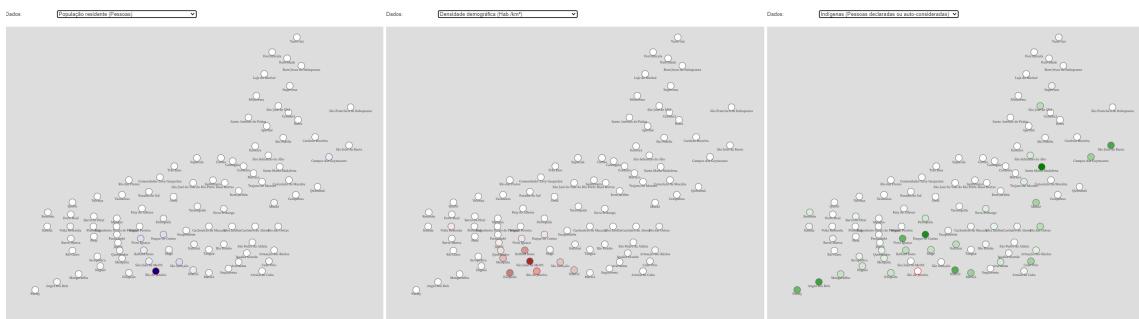


Figura 5.2: Carto-centric apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos aos municípios do Estado do Rio de Janeiro, com centro do cartograma sobre o município de Cachoeiras de Macacu.

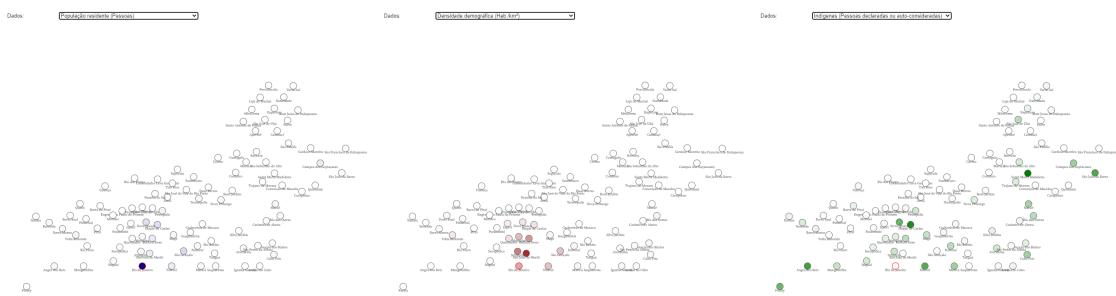


Figura 5.3: Carto-centric apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos aos municípios do Estado do Rio de Janeiro, com centro do cartograma sobre a capital do Estado.

As Figuras 5.4 e 5.5 a seguir apresentam os cartogramas gerados com base em dados obtidos do Censo Demográfico de 2022 realizado pelo IBGE, e utilizando o conceito carto-step proposto nesta tese. Foram geradas novas tuplas contendo os

valores das variáveis estatísticas associadas a cada unidade georreferenciada, e o conjunto de valores foi aplicado a uma escala monocromática de tonalidades, partindo do branco como valor mínimo em direção a uma cor selecionada arbitrariamente como valor máximo, para aplicação em uma das formas vetoriais simples sugeridas (triângulo, quadrado, hexágono, octógono ou círculo) como cor de preenchimento para a forma interna e como cor de contorno para a forma externa, e criadas para exibição por meio do recurso "canvas" da linguagem de marcação HTML 5 em modo bidimensional.

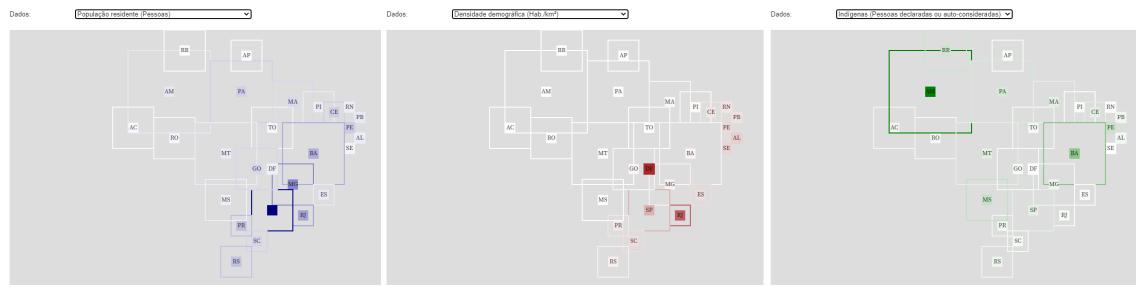


Figura 5.4: Carto-step apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos às Unidades da Federação.

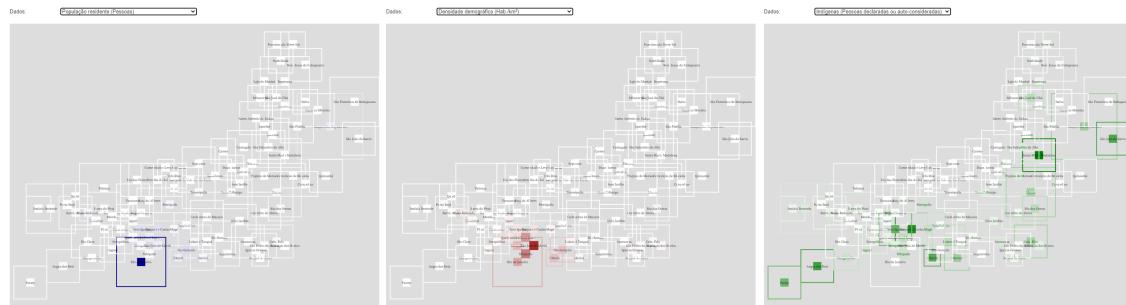


Figura 5.5: Carto-step apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos aos municípios do Estado do Rio de Janeiro.

A Figura 5.6 a seguir apresenta um cartograma gerado com base em dados obtidos do Censo Demográfico de 2022 realizado pelo IBGE, e utilizando também o conceito carto-step proposto nesta tese. Foram geradas novas tuplas contendo os valores das variáveis estatísticas associadas a cada unidade georreferenciada, e o conjunto de valores foi aplicado a uma escala monocromática de tonalidades, partindo do branco como valor mínimo em direção a uma cor selecionada arbitrariamente como valor máximo, para aplicação em um dos troncos de pirâmide baseados em formas simples sugeridas (triângulo, quadrado, hexágono, octógono ou círculo) como

cor de preenchimento; o mesmo conjunto de valores foi também aplicado à altura dos troncos, e tais sólidos foram criados por meio da biblioteca three.js (THREE.JS, 2025), utilizando a Application Programming Interface (API) WebGL (KHRONOS, 2025), para exibição por meio do recurso "canvas" da linguagem de marcação HTML 5, agora em modo tridimensional.

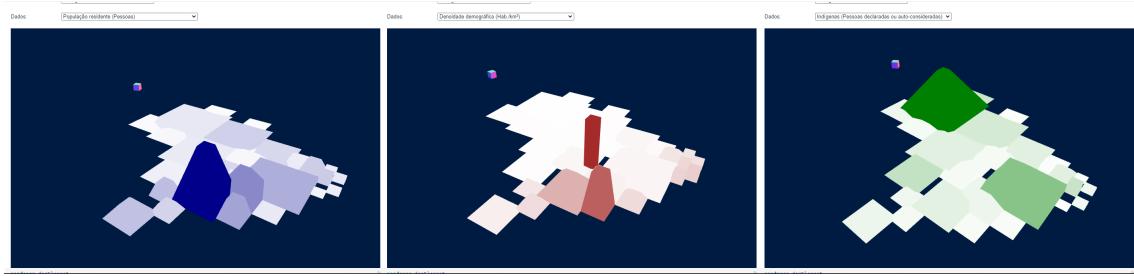


Figura 5.6: Carto-step tridimensional apresentando dados obtidos do Censo Demográfico de 2022 do Brasil, relativos às Unidades da Federação.

Por meio das apresentações geradas, é possível confirmar a validade das visualizações para auxiliar na compreensão de conjuntos de dados relativos às unidades geográficas representadas: obviamente, nem toda visualização proposta atenderá a todos os territórios, ficando a escolha da opção mais adequada a critério do produtor da representação; no entanto, ampliar as opções disponíveis será sempre desejável, a fim de favorecer a interpretação dos dados por meio de cartogramas.

Como apontado, as escolhas dos valores utilizados nas tuplas para o dimensionamento e posicionamento das unidades foram arbitradas individualmente, de acordo com critérios essencialmente visuais de adequação, como evitar sobreposição de unidades e manter a maior proximidade possível de suas posições geográficas relativas: para territórios com pequeno número de unidades, essa definição no modo "manual", como o implementado nas versões da plataforma Observable aqui apresentadas, ainda é viável; no entanto, ao aplicar essa tarefa a Unidades da Federação maiores (por exemplo, o estado com maior número de municípios é Minas Gerais, com 853 unidades para representar), a carga de trabalho para a definição de valores torna-se inadequada. Deve-se considerar, portanto, meios de automatizar em um maior grau a escolha dos valores aplicáveis para o dimensionamento e posicionamento das unidades georreferenciadas.

Nessa perspectiva, uma vez que a plataforma Observable é baseada em tecnologias para renderização, exibição e interação de páginas web, sua utilização ainda se justifica pela possibilidade do desenvolvimento de uma interface gráfica online, permitindo o carregamento de novos conjuntos de dados para os cartogramas, acessando bases de dados públicas e abertas, por meio da utilização de navegadores web

de uso amplo e independentes de sistemas operacionais. Para tal interface online, se vislumbra também a possibilidade de carregar novos mapas a partir de arquivos em formatos amplamente utilizados, como SHP, KML ou KMZ, para serem convertidos nos cartogramas especiais aqui apresentados.

Capítulo 6

Implementações para renderização automatizada de cartogramas

Dando continuidade à proposta de implementação sugerida no capítulo anterior, buscamos pesquisar o desenvolvimento da interface desejada utilizando aplicações que oferecessem opções simplificadas e interativas para geração de código, em oposição à codificação bruta, a qual não constitui o cerne deste trabalho; a plataforma Observable possui recursos valiosos para a manipulação visual de dados espaciais georreferenciados como, por exemplo, módulos que permitem a importação e exportação de arquivos em formato shapefile porém, para a presente proposta, tal recurso se mostrou limitado logo durante as etapas iniciais de implementação.

Além disso, para importação e exportação adequadas dos arquivos-fonte de dados georreferenciados, é necessária uma razoável compreensão e edição de dados a partir das informações e estruturas contidas nos formatos de mapas vetoriais utilizados atualmente; por este motivo, a consulta e o entendimento dos processos automatizados realizados por aplicações de Sistemas de Informação Geográfica (SIG), como QGIS e ArcGIS, se colocam como elementos importantes para essa compreensão, a fim de atingir o objetivo final de oferecer uma aplicação simplificada e de fácil utilização.

Considerando tais aspectos, a alternativa que se demonstrou interessante para desenvolvimento dos algoritmos sem uso de codificação bruta foi o KNIME (2024), uma plataforma de análise de dados, de desenvolvimento livre e aberto, baseada na IDE Eclipse de uso amplo e geral; o KNIME permite a criação de fluxos de trabalho visuais, compostos por módulos conectados por entradas e saídas que dispõem de modos de configuração simplificada de parâmetros.

As primeiras versões produzidas com o KNIME exigiram uma pequena readaptação das estruturas de dados, visto que o paradigma explorado para armazenar e manipular valores na plataforma Observable se baseia em matrizes multidimensionais como utilizadas no Javascript, enquanto na abordagem essencialmente visual

adotada pelo KNIME são utilizadas tabelas bidimensionais para tais funções, conforme a Figura 6.1.

Concatenated table - 3:114 - Concatenate																
Table "default" - Rows: 54 Spec - Columns: 21 Properties Flow Variables																
Row ID	I	extShape	D newX	D newY	D xceter	D yceter	D x	D y	S sides	D start	D scale	S shape	geometry(#1)			
Row3	5	3	-52	6	-69.909	-4.752	-52	6	4	0.785	1.414	POLYGON ((-51.0 7.0,-53.0 7.0,-53.0 5.0,-51.0 5.0,-51.0 7.0))	POLYGON - EPSG:4326			
Row4	1	4	-47	-8	-69.909	-4.752	-47	-8	4	0.785	1.414	POLYGON (-46.0 -7.0,-48.0 -7.0,-48.0 -9.0,-46.0 -9.0,-46.0 -7.0))	POLYGON - EPSG:4326			
Row5	5	4	-43	-3	-69.909	-4.752	-43	-3	4	0.785	1.414	POLYGON (-42.0 -1.999999999999999,-44.0 -1.999999999999999,...)	POLYGON - EPSG:4326			
Row6	4	5	-38	4	-69.909	-4.752	-38	4	4	0.785	1.414	POLYGON ((-37.0 -3.0,-39.0 -3.0,-39.0 -5.0,-37.0 5.0,-37.0 -3.0))	POLYGON - EPSG:4326			
Row7	7	2	-35	5	-69.909	-4.752	-35	5	4	0.785	1.414	POLYGON ((-34.0 -2.0,-36.0 -2.0,-36.0 -4.0,-34.0 -4.0,-34.0 -2.0))	POLYGON - EPSG:4326			
Row8	1	1	-32	-4	-69.909	-4.752	-32	-4	4	0.785	1.414	POLYGON ((-31.0 -3.0,-33.0 -3.0,-33.0 -5.0,-31.0 -5.0,-31.0 -3.0))	POLYGON - EPSG:4326			
Row9	1	1	-30	-6	-69.909	-4.752	-30	-6	4	0.785	1.414	POLYGON ((-29.0 -5.0,-31.0 -5.0,-31.0 -7.0,-29.0 7.0,-29.0 5.0))	POLYGON - EPSG:4326			
Row10	1	1	-32	-8	-69.909	-4.752	-32	-8	4	0.785	1.414	POLYGON ((-31.0 -7.0,-33.0 -7.0,-33.0 -9.0,-31.0 -9.0,-31.0 -7.0))	POLYGON - EPSG:4326			
Row11	5	1	-30	-10	-69.909	-4.752	-30	-10	4	0.785	1.414	POLYGON ((-29.0 -9.0,-31.0 -9.0,-31.0 -11.0,-29.0 11.0,-29.0 9.0))	POLYGON - EPSG:4326			
Row12	5	1	-32	-12	-69.909	-4.752	-32	-12	4	0.785	1.414	POLYGON ((-31.0 -11.0,-33.0 -11.0,-33.0 -13.0,-31.0 -13.0,-31.0 -11.0))	POLYGON - EPSG:4326			
Row13	2	6	-39	-13	-69.909	-4.752	-39	-13	4	0.785	1.414	POLYGON ((-38.0 -12.0,-40.0 -12.0,-40.0 -14.0,-38.0 -14.0,-38.0 -12.0))	POLYGON - EPSG:4326			
Row14	5	4	-44	-19	-69.909	-4.752	-44	-19	4	0.785	1.414	POLYGON ((-43.0 -18.0,-45.0 -18.0,-45.0 -20.0,-43.0 -20.0,-43.0 -18.0))	POLYGON - EPSG:4326			
Row15	1	2	-38	-21	-69.909	-4.752	-38	-21	4	0.785	1.414	POLYGON ((-37.0 -20.0,-39.0 -20.0,-39.0 -22.0,-37.0 -22.0,-37.0 -20.0))	POLYGON - EPSG:4326			
Row16	2	2	-42	-25	-69.909	-4.752	-42	-25	4	0.785	1.414	POLYGON ((-41.0 -24.0,-43.0 -24.0,-43.0 -26.0,-41.0 -26.0,-41.0 -24.0))	POLYGON - EPSG:4326			
Row17	2	4	-48	-24	-69.909	-4.752	-48	-24	4	0.785	1.414	POLYGON ((-47.0 -23.0,-49.0 -23.0,-49.0 -25.0,-47.0 -25.0,-47.0 -23.0))	POLYGON - EPSG:4326			
Row18	5	2	-54	-27	-69.909	-4.752	-54	-27	4	0.785	1.414	POLYGON ((-53.0 -26.0,-55.0 -26.0,-55.0 -28.0,-53.0 -28.0,-53.0 -26.0))	POLYGON - EPSG:4326			
Row19	3	2	-51	-30	-69.909	-4.752	-51	-30	4	0.785	1.414	POLYGON ((-50.0 -29.0,-52.0 -29.0,-52.0 -31.0,-50.0 -31.0,-50.0 -29.0))	POLYGON - EPSG:4326			
Row20	1	3	-55	-34	-69.909	-4.752	-55	-34	4	0.785	1.414	POLYGON ((-54.0 -33.0,-56.0 -33.0,-56.0 -35.0,-54.0 -35.0,-54.0 -33.0))	POLYGON - EPSG:4326			
Row21	3	4	-58	-22	-69.909	-4.752	-58	-22	4	0.785	1.414	POLYGON ((-57.0 -21.0,-59.0 -21.0,-59.0 -23.0,-57.0 -23.0,-57.0 -21.0))	POLYGON - EPSG:4326			
Row22	5	6	-56	-13	-69.909	-4.752	-56	-13	4	0.785	1.414	POLYGON ((-56.0 -12.0,-57.0 -12.0,-57.0 -14.0,-55.0 -14.0,-55.0 -12.0))	POLYGON - EPSG:4326			
Row23	4	50	-16	-16	-69.909	-4.752	-50	-16	4	0.785	1.414	POLYGON ((-49.0 -15.0,-51.0 -15.0,-51.0 -17.0,-49.0 -17.0,-49.0 -15.0))	POLYGON - EPSG:4326			
Row24	1	-47	-16	-69.909	-4.752	-47	-16	4	0.785	1.414	POLYGON ((-46.0 -15.0,-48.0 -15.0,-48.0 -17.0,-46.0 -17.0,-46.0 -15.0))	POLYGON - EPSG:4326				
Row25	3	4	-66	-10	-69.909	-4.752	-66	-10	4	0.785	1.414	POLYGON ((-45.0 -9.0,-47.0 -9.0,-47.0 -11.0,-45.0 -11.0,-45.0 -9.0))	POLYGON - EPSG:4326			
Row26	2	4	-64	7	-69.909	-4.752	-64	7	4	0.785	1.414	POLYGON ((-43.0 -8.0,-45.0 -8.0,-45.0 -6.0,-43.0 -6.0,-43.0 -8.0))	POLYGON - EPSG:4326			
Row27	4	4	-74	-8	-69.909	-4.752	-74	-8	4	0.785	1.414	MULTILINESTRING ((-44.0 -7.0,-45.0 -7.0,-45.0 -9.0,-44.0 -9.0,-44.0 -7.0))	MULTILINESTRING - EPSG:4326			
Row28	5	6	-67	-1	-69.909	-4.752	-67	-1	4	0.785	1.414	MULTILINESTRING ((-45.0 -7.0,-46.0 -7.0,-46.0 -9.0,-45.0 -9.0,-45.0 -7.0))	MULTILINESTRING - EPSG:4326			
Row29_dup	3	6	-53	-1	-69.909	-4.752	-53	-1	4	0.785	1.414	MULTILINESTRING ((-47.0 -5.0,-49.0 -5.0,-49.0 -7.0,-47.0 -7.0,-47.0 -5.0))	MULTILINESTRING - EPSG:4326			
Row30_dup	5	3	-52	6	-69.909	-4.752	-52	6	4	0.785	1.414	MULTILINESTRING ((-49.0 -9.0,-55.0 -9.0,-55.0 -11.0,-49.0 -11.0,-49.0 -9.0))	MULTILINESTRING - EPSG:4326			
Row31_dup	4	4	-47	-8	-69.909	-4.752	-47	-8	4	0.785	1.414	MULTILINESTRING ((-43.0 -3.0,-49.0 -3.0,-49.0 -5.0,-43.0 -5.0,-43.0 -3.0))	MULTILINESTRING - EPSG:4326			
Row32_dup	3	4	-43	-3	-69.909	-4.752	-43	-3	4	0.785	1.414	MULTILINESTRING ((-39.0 -1.0,-40.0 -1.0,-40.0 -0.0,-39.0 -0.0,-39.0 -1.0))	MULTILINESTRING - EPSG:4326			
Row33_dup	4	3	-38	-4	-69.909	-4.752	-38	-4	4	0.785	1.414	MULTILINESTRING ((-35.0 -0.0,-39.0 -0.0,-39.0 -1.0,-35.0 -1.0,-35.0 -0.0))	MULTILINESTRING - EPSG:4326			
Row34_dup	7	2	-35	-5	-69.909	-4.752	-35	-5	4	0.785	1.414	MULTILINESTRING ((-33.0 -2.0,-39.0 -2.0,-39.0 -4.0,-33.0 -4.0,-33.0 -2.0))	MULTILINESTRING - EPSG:4326			
Row35_dup	1	1	-32	-4	-69.909	-4.752	-32	-4	4	0.785	1.414	MULTILINESTRING ((-31.0 -3.0,-33.0 -3.0,-33.0 -5.0,-31.0 -5.0,-31.0 -3.0))	MULTILINESTRING - EPSG:4326			
Row36_dup	1	1	-30	-6	-69.909	-4.752	-30	-6	4	0.785	1.414	MULTILINESTRING ((-29.0 -5.0,-31.0 -5.0,-31.0 -7.0,-29.0 -7.0,-29.0 -5.0))	MULTILINESTRING - EPSG:4326			
Row37_dup	1	1	-32	-8	-69.909	-4.752	-32	-8	4	0.785	1.414	MULTILINESTRING ((-31.0 -7.0,-33.0 -7.0,-33.0 -9.0,-31.0 -9.0,-31.0 -7.0))	MULTILINESTRING - EPSG:4326			
Row38_dup	4	1	-30	-10	-69.909	-4.752	-30	-10	4	0.785	1.414	MULTILINESTRING ((-29.0 -9.0,-31.0 -9.0,-31.0 -11.0,-29.0 -11.0,-29.0 -9.0))	MULTILINESTRING - EPSG:4326			
Row39_dup	5	1	-32	-12	-69.909	-4.752	-32	-12	4	0.785	1.414	MULTILINESTRING ((-31.0 -11.0,-33.0 -11.0,-33.0 -13.0,-31.0 -13.0,-31.0 -11.0))	MULTILINESTRING - EPSG:4326			
Row40_dup	2	6	-39	-13	-69.909	-4.752	-39	-13	4	0.785	1.414	MULTILINESTRING ((-33.0 -6.0,-39.0 -6.0,-39.0 -9.0,-33.0 -9.0,-33.0 -6.0))	MULTILINESTRING - EPSG:4326			
Row41_dup	5	4	-44	-19	-69.909	-4.752	-44	-19	4	0.785	1.414	MULTILINESTRING ((-40.0 -15.0,-48.0 -15.0,-48.0 -23.0,-40.0 -23.0,-40.0 -15.0))	MULTILINESTRING - EPSG:4326			

Figura 6.1: Formato de tabela bidimensional utilizada pelo KNIME para intercâmbio de dados entre os módulos do fluxograma. Elaborado pelo autor.

Os dados de saída também foram redirecionados para outro formato: ao invés de formas vetoriais desenhadas em um elemento canvas do HMTL 5, foram gerados dados no formato shapefile, desenvolvido pela Esri e comumente usado por diferentes sistemas de informação geográfica (ESRI, 1998); os shapefiles podem armazenar não apenas informações geométricas, mas também atributos não geométricos adicionais (ou metadados, relativos por exemplo aos valores de dados estatísticos associados a cada unidade geográfica, como no caso do presente estudo).

Quanto às formas geométricas geradas, o formato legível por humanos mais comumente usado para representar objetos de geometria vetorial é o Well-Known Text Format (WKT); este formato foi estabelecido pelo Open Geospatial Consortium e especificado em seu padrão Simple Feature Access (HERRING et al., 2011). Para a proposta carto-centric, foi produzida uma camada (layer) contendo formas do tipo POLYGON correspondentes a cada unidade geográfica representada; para a proposta carto-step, além da camada do tipo POLYGON relativa às formas internas preenchidas, houve a necessidade de criar uma nova camada para as formas externas de contorno, utilizando aqui o tipo MULTILINESTRING.

Para a importação, exportação e manipulação de dados em formato georreferenciado oriundos de arquivos shapefile nas extensões .shp e .prj, também foi necessário baixar e instalar a extensão gratuita Geospatial Analytics, desenvolvida em conjunto

pela equipe de pesquisa do Centro de Análise Geográfica da Universidade de Harvard e incluída na plataforma de análise KNIME em dezembro de 2022 (LIU et al., 2024). Esta extensão, por ser baseada na extensão Python para KNIME, oferece a possibilidade de migração de código para aplicações em Python tanto em modo local quanto em modo online, o que atende ao objetivo inicialmente almejado nesta tese.

As Figuras 6.2 e 6.3 exibem os fluxogramas de módulos desenvolvidos para as propostas carto-centric e carto-step apresentadas nesta tese que, uma vez convertidos com base nos recursos oferecidos pela plataforma KNIME, podem dar origem a aplicações interativas com interfaces semelhantes às das capturas de tela correspondentes, apresentadas nas Figuras 6.4 e 6.5.

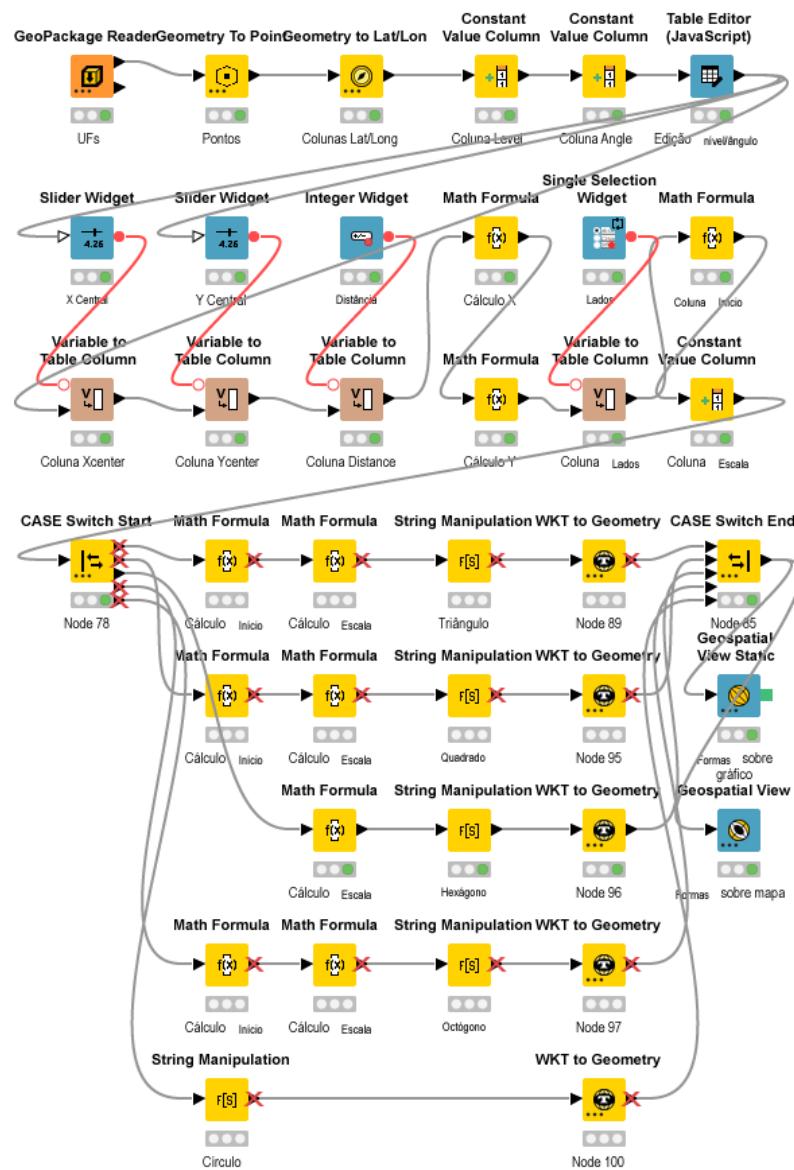


Figura 6.2: Fluxograma do KNIME para a proposta carto-centric. Elaborado pelo autor.

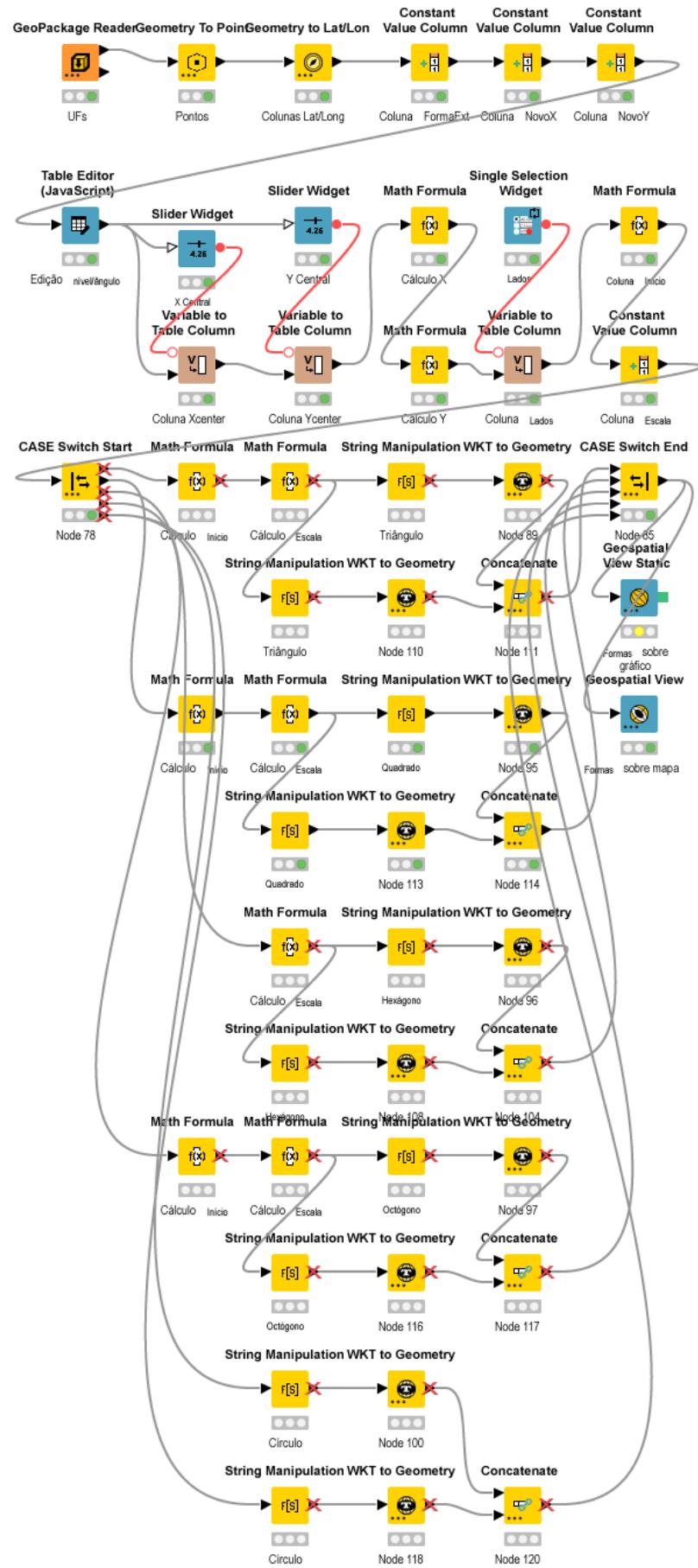


Figura 6.3: Fluxograma do KNIME para a proposta carto-step. Elaborado pelo autor.

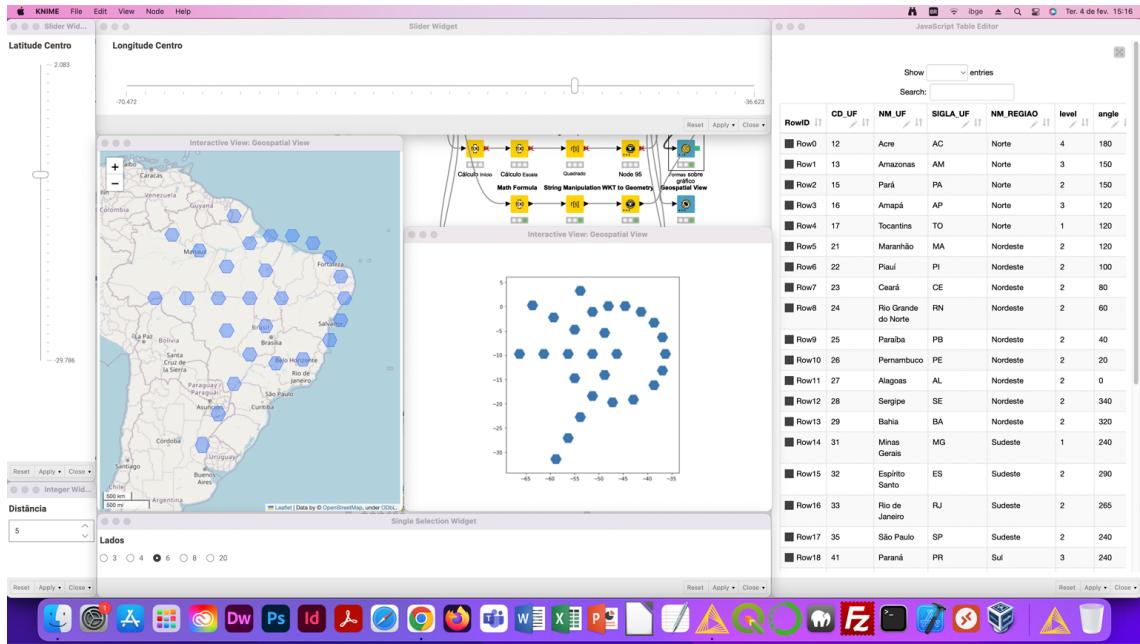


Figura 6.4: Tela capturada do KNIME para a proposta carto-centric. Elaborado pelo autor.

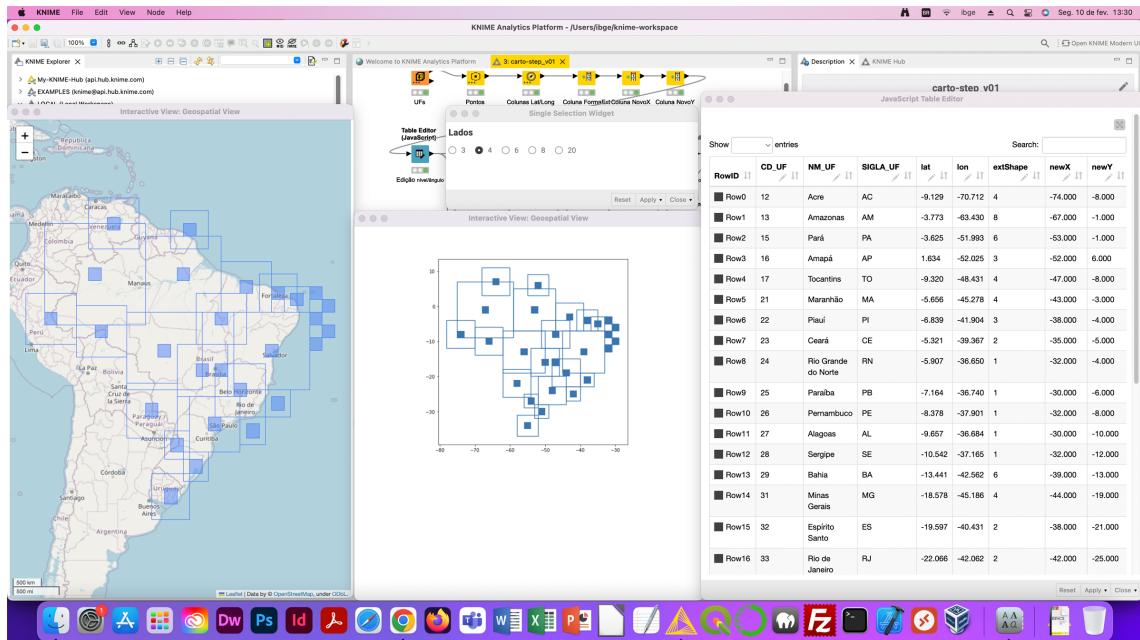


Figura 6.5: Tela capturada do KNIME para a proposta carto-step. Elaborado pelo autor.

Nota-se que a extensão Geospatial Analytics possibilita a saída dos dados tanto em formatos de imagem estática como Portable Network Graphics (PNG) ou Scalable Vector Graphics (SVG), por meio do módulo Geospatial View Static cuja visu-

alização é exibida mais ao centro nas telas capturadas, quanto em formato shapefile dinâmico e interativo por meio do módulo Geospatial View, cuja visualização está posicionada à esquerda nas telas: neste último, como efeito da projeção cartográfica aplicada ao mapa subjacente, as formas geométricas tendem a sofrer alongamento em sua dimensão vertical quando em latitudes mais elevadas e próximas aos polos; no caso das unidades geográficas pertencentes ao território brasileiro, tal alongamento, embora perceptível, não oferece grande impacto negativo visual na representação do cartograma.

As propostas aqui apresentadas foram publicadas em abril de 2025 na revista InfoDesign, sob o título "Alternatives for automatic rendering of Brazil and other geographic regions' territory thematic cartograms"(NASCIMENTO *et al.*, 2025).

Capítulo 7

Trabalhos Futuros e Considerações Finais

A presente tese e suas propostas oferecem inúmeras possibilidades de avanços de pesquisa: a primeira envolveria uma pesquisa comparativa entre os cartogramas aqui apresentados e outros tipos de cartogramas mais tradicionalmente utilizados, coletando vantagens, desvantagens e impressões junto a uma ampla gama de usuários, a fim de identificar as aplicações mais apropriadas a cada tipo de representação, bem como pontos e aspectos a serem aprimorados em versões futuras de cada proposta. Para tornar mais sólido tal estudo comparativo, deve também ser enriquecida a base produzida de cartogramas, incluindo, por exemplo, a representação das unidades da federação em sua totalidade, e podendo abranger outros recortes e níveis geográficos, como diferentes países e continentes.

Deve-se observar que não houve aqui análises de desempenho dos algoritmos propostos em termos de eficiência de tempo e recursos computacionais: um estudo voltado a este aspecto poderia indicar a implementação dos algoritmos de geração de cartogramas em tempo real utilizando linguagens para programação de Unidades de Processamento Gráfico (GPUs) como a API CUDA, visando a incorporação em dispositivos que disponham de Unidades de Processamento Gráfico de Propósito Geral (GPGPU) da fabricante nVidia (NVIDIA e FITZEK, 2006).

O objetivo de disponibilização dos algoritmos de geração automatizada de cartogramas por meio de uma interface web também deve ser viabilizado com a produção de novos estudos. A perspectiva mais promissora envolveria a migração dos fluxogramas desenvolvidos na plataforma KNIME para códigos na linguagem Python e suas bibliotecas de geoprocessamento, sobre as quais os módulos da extensão Geospatial Analytics já se baseiam: uma vez produzidos os códigos em Python, sua disponibilização como serviço web seria implementada de maneira natural. Alternativas a tal abordagem poderiam ser estudadas, a partir da criação de novos recursos de manipulação de dados georreferenciados para a plataforma Observable (OBSER-

VABLE, 2024), ou pela adoção da plataforma de visualização Processing em uma de suas versões para web como a biblioteca p5.js (MCCARTHY *et al.*, 2015).

Outro aspecto a ser considerado em um estudo posterior envolveria a geração das tuplas ou tabelas bidimensionais contendo os valores para posicionamento e dimensionamento dos elementos representacionais relativos às unidades: a definição arbitrária de tais valores tende a ser uma tarefa exaustiva para cartogramas com centenas ou milhares de unidades. Assim, algoritmos automatizados para a produção de tais dados, com o uso de recursos de aprendizado de máquina, constituem um tema futuro de pesquisa para as propostas apresentadas; nesse sentido, uma implementação em linguagem Python seria favorecida graças às diversas e já amplamente utilizadas bibliotecas para inteligência artificial e machine learning disponíveis atualmente.

Ainda assim, as propostas de cartogramas apresentadas na presente tese oferecem um potencial de utilização promissor: tanto para mídias estáticas e impressas em dimensões físicas reduzidas ou baixa qualidade de impressão, quanto para mídias eletrônicas dinâmicas, em dispositivos de telas reduzidas e/ou capacidade de processamento e armazenamento limitados, e também para mídias alternativas, como as sugestões de dispositivos tridimensionais para eventos apresentadas anteriormente. Tais implementações em si também constituem temas de pesquisa adicionais e complementares ao objeto de estudo aqui apresentado; qual seja, a geração de cartogramas especiais voltados aos problemas específicos de representação do território brasileiro. E, adicionalmente, fornecem recursos para atender de forma mais eficaz às atividades públicas da instituição à qual o autor da tese está vinculado, e que estão expressas em sua missão institucional: "Retratar o Brasil com informações necessárias ao conhecimento de sua realidade e ao exercício da cidadania."(IBGE, 2025b)

Referências Bibliográficas

- BARKOWSKY, T., LATECKI, L. J., RICHTER, K. F., 2000, “Schematizing maps: Simplification of geographic shape by discrete curve evolution”. In: *Spatial cognition ii: Integrating abstract theories, empirical studies, formal methods, and practical applications*, pp. 41–53. Springer, November.
- BARRETO, M. C., KOSMINSKY, D., ESPERANÇA, C., 2018, “Hexagonal hierarchical cartogram: Towards a thematic map of Brazil”, *InfoDesign-Revista Brasileira de Design da Informação*, v. 15, n. 1, pp. 45–62.
- BERTIN, J., 1983, *Semiology of graphics*. University of Wisconsin press.
- BOSTOCK, M., OBSERVABLE, INC., 2025. “D3.js”. Disponível em: <<https://d3js.org/>>. Acessado em: 15 Mai 2025.
- BREWER, C., CAMPBELL, A. J., 1998, “Beyond graduated circles: Varied point symbols for representing quantitative data on maps”, *Cartographic Perspectives*, v. 1, n. 29, pp. 6–25.
- BREWER, C. A., 1989, “Color chart use in map design”, *Cartographic Perspectives*, v. 1, n. 04, pp. 3–10.
- BREWER, C. A., MACEACHREN, A. M., PICKLE, L. W., et al., 1997, “Mapping mortality: Evaluating color schemes for choropleth maps”, *Annals of the Association of American Geographers*, v. 87, n. 3, pp. 411–438.
- BUENO, M., 2014, *Grade estatística: uma abordagem para ampliar o potencial analítico de dados censitários*. Tese de doutorado, Universidade Estadual de Campinas – Unicamp, Campinas. Disponível em: https://www.academia.edu/download/50854463/Bueno2014_homolog2.pdf.
- DEICHMANN, U., 1996, *A review of spatial population database design and modeling*. National Center for Geographic Information and Analysis.
- DENT, B. D., 1999, *Cartography: Thematic map design*. New York, McGraw-Hill Higher Education.

- EADES, P., 1984, “A heuristic for graph drawing”, *Congressus numerantium*, v. 42, n. 11, pp. 149–160.
- ELMER, M. E., 2012, *Symbol considerations for bivariate thematic mapping*. Tese de Mestrado, University of Wisconsin–Madison.
- ESRI, 1998, “ESRI shapefile technical description”, *Comput. Stat*, v. 16, pp. 370–371.
- FISHER, H. T., 1982, *Mapping information: The graphic display of quantitative information*. Cambridge, MA, Abt Books.
- FLANNERY, J. J., 1971, “The relative effectiveness of some common graduated point symbols in the presentation of quantitative data”, *Cartographica: The International Journal for Geographic Information and Geovisualization*, v. 8, n. 2, pp. 96–109.
- GOODCHILD, M. F., 1992, *Research Initiative 1, Accuracy of Spatial Databases, Final Report*. National Center for Geographic Information and Analysis.
- GRASLAND, C., MADELIN, M., MATHIAN, H., et al., 2006, *The modifiable areas unit problem*. Tese de Doutorado, ESPON| Inspire Policy Making with Territorial Evidence.
- GUZMÁN, J. M., SCHENSUL, D., ZHANG, S., 2013, “Understanding vulnerability and adaptation using census data”, *The demography of adaptation to climate change*. New York, London and Mexico City: UNFPA, IIED e El Colegio de México.
- HAGERSTRAND, T., 1968, “Innovation diffusion as a spatial process.” (No Title).
- HAKE, G., GRÜNREICH, D., 1994. “Kartographie. 7., völlig neubearb. Und erweiterte Aufl”. .
- HERRING, J., OTHERS, 2011, “Opengis® implementation standard for geographic information-simple feature access-part 1: Common architecture [corrigendum]”, (No Title).
- HOGAN, D. J., 1992, “Crescimento populacional, padrões de assentamento e o ambiente físico”, *Ciências sociais hoje*. Rio de Janeiro: ANPOCS, pp. 285–298.
- IBGE, 2025a. “IBGE | Portal do IBGE | IBGE”. a. Disponível em: <<https://www.ibge.gov.br/>>. Acessado em: 26 Abr 2025.

- IBGE, 2025b. “O IBGE | IBGE”. b. Disponível em: <<https://www.ibge.gov.br/acesso-informacao/institucional/o-ibge.html>>. Acessado em: 18 Mai 2025.
- INSPIRE, 2023. “Data Specification on Geographical Grid Systems - Technical Guidelines, Text, INSPIRE Maintenance and Implementation Group (MIG), European Union”.
- JENKS, G. F., CASPALL, F. C., 1971, “Error on choroplethic maps: definition, measurement, reduction”, *Annals of the Association of American Geographers*, v. 61, n. 2, pp. 217–244.
- JONES, C. B., 2014, *Geographical information systems and computer cartography*. Routledge.
- KHRONOS, 2025. “WebGL”. Disponível em: <<https://www.khronos.org/webgl/>>. Acessado em: 15 Mai 2025.
- KNIME, 2024. “Open for Innovation | KNIME”. <https://www.knime.com/>. Acessado em: 17 Dez 2024.
- KRYGIER, J., WOOD, D., 2024, *Making maps*. New York, Guilford Publications.
- LAGRANGE, J. P., WEIBEL, R., MULLER, J.-C., 1995, *GIS and Generalisation: Methodology and Practice*. CRC Press.
- LATECKI, L. J., LAKÄMPER, R., 1999a, “Convexity rule for shape decomposition based on discrete contour evolution”, *Computer Vision and Image Understanding*, v. 73, n. 3, pp. 441–454.
- LATECKI, L. J., LAKÄMPER, R., 1999b, “Polygon evolution by vertex deletion”. In: *Scale-Space Theories in Computer Vision: Second International Conference, Scale-Space'99 Corfu, Greece, September 26–27, 1999 Proceedings* 2, pp. 398–409. Springer, b.
- LINARD, C., TATEM, A. J., 2012, “Large-scale spatial population databases in infectious disease research”, *International journal of health geographics*, v. 11, pp. 1–13.
- LINARD, C., ALEGANA, V. A., NOOR, A. M., et al., 2010, “A high resolution spatial population database of Somalia for disease risk mapping”, *International Journal of Health Geographics*, v. 9, pp. 1–13.
- LIU, L., FU, X., KÖTTER, T., et al., 2024, “Geospatial analytics extension for knime”, *SoftwareX*, v. 25, pp. 101627.

- MACEACHREN, A. M., DIBIASE, D., 1991, “Animated maps of aggregate data: Conceptual and practical problems”, *Cartography and Geographic Information Systems*, v. 18, n. 4, pp. 221–229.
- MARTIN, D., 2000, “Census 2001: making the best of zonal geographies”. In: *CONFERENCE AT THE CATHIE MARSH CENTRE.[np].*, .
- MARTINE, G., SCHENSUL, D., 2013. “The Demography of Adaptation to Climate Change. New York, London and Mexico City”..
- MCCARTHY, L., REAS, C., FRY, B., 2015, *Getting started with P5.js: Making interactive graphics in JavaScript and processing*. Maker Media, Inc.
- MCMASTER, R., MCMASTER, S., 2002, “A history of twentieth-century American academic cartography”, *Cartography and Geographic Information Science*, v. 29, n. 3, pp. 305–321.
- MONMONIER, M., 1989, *Maps with the news: The development of American journalistic cartography*. Chicago, University of Chicago Press.
- MONTEIRO, A. M. V., CÂMARA, G., CARVALHO, M., et al., 2004, “Análise espacial de dados geográficos”, *Brasília: Embrapa*.
- MONTELLO, D. R., 2002, “Cognitive map-design research in the twentieth century: Theoretical and empirical approaches”, *Cartography and Geographic Information Science*, v. 29, n. 3, pp. 283–304.
- MUNZNER, T., 2014, *Visualization analysis and design*. CRC press.
- NASCIMENTO, L., REZNIK, R., ESPERANÇA, C., et al., 2025, “Alternatives for automatic rendering of Brazil and other geographic regions’ territory thematic cartograms”, *InfoDesign*, v. 22, n. 1.
- NORMAN, P., REES, P., BOYLE, P., 2003, “Achieving data compatibility over space and time: creating consistent geographical zones”, *International journal of population geography*, v. 9, n. 5, pp. 365–386.
- NVIDIA, VINGELMANN, P., FITZEK, F. H., 2006, “Cuda”, *Online http://www.nvidia.com/object/cuda_home_new*, v. 15.
- OBSERVABLE, 2024. “Observable: Build expressive charts and dashboards with code”. <https://observablehq.com/>. Acessado em: 26 Fev 2024.
- OJIMA, R., MARTINE, G., 2012, “Resgates sobre população e ambiente: breve análise da dinâmica demográfica e a urbanização nos biomas brasileiros”, *Ideias*, v. 3, n. 2, pp. 55–70.

- OLSON, J. M., BREWER, C. A., 1997, "An evaluation of color selections to accommodate map users with color-vision impairments", *Annals of the Association of American Geographers*, v. 87, n. 1, pp. 103–134.
- OPENSHAW, S., 1977, "A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modelling", *Transactions of the institute of british geographers*, pp. 459–472.
- OPENSHAW, S., RAO, L., 1995, "Algorithms for reengineering 1991 Census geography", *Environment and planning A*, v. 27, n. 3, pp. 425–446.
- PINHEIRO, T. M., ESPERANÇA, C., 2018, "Interactive Visualization of Trivariate Georeferenced Data", *Journal of Information and Data Management*, v. 9, n. 3. Disponível em: <<https://sol.sbc.org.br/journals/index.php/jidm/article/view/2043>>.
- ROBINSON, A. H., KIMERLING, A., OTHERS, 1969, *Elements of cartography*. New York, New York John Wiley and Sons.
- RUSANEN, J., MUILU, T., COPAERT, A., et al., 2001, "Finnish Socio-Economic Grid Data, GIS and the Hidden Geography of Unemployment", *Tijdschrift voor economische en sociale geografie*, v. 92, n. 2, pp. 139–147.
- SCHUURMAN, N., LESZCZYNSKI, A., FIEDLER, R., et al., 2006, "Building an integrated cadastral fabric for higher resolution socioeconomic spatial data analysis". In: *Progress in Spatial Data Handling: 12th International Symposium on Spatial Data Handling*, pp. 897–920. Springer, September.
- SLOCUM, T. A., MCMASTER, R. M., KESSLER, F. C., et al., 2008, *Thematic cartography and geographic visualization*. London, Prentice Hall.
- SOUZA, G., TORRES, H., 2003, "O estudo da metrópole e o uso de informações georreferenciadas", *São Paulo em perspectiva*, v. 17, pp. 35–44.
- SPECKMANN, B., VERBEEK, K., 2010, "Necklace maps", *IEEE Trans. Vis. Comput. Graph.*, v. 16, n. 6, pp. 881–889.
- SPECKMANN, B., VERBEEK, K., 2015, "Algorithms for necklace maps", *International Journal of Computational Geometry & Applications*, v. 25, n. 01, pp. 15–36.
- STRAND, G. H., BLOCH, V. V. H., 2009, "Statistical grids for Norway. Documentation of national grids for analysis and visualization of spatial data in Norway", *Statistics Norway*, v. 9.

- TAMMILEHTO-LUODE, M., 2011, “Opportunities and challenges of grid-based statistics”. In: *World Statistics Congress of the International Statistical Institute*, p. 209, August.
- TAMMILEHTO-LUODE, M., BACKER, L., ROGSTAD, L., 2000, “Grid data and area delimitation by definition towards a better European territorial statistical system”, *Statistical Journal of the United Nations Economic Commission for Europe*, v. 17, n. 2, pp. 109–117.
- THREE.JS, 2025. “three.js”. Disponível em: <<https://threejs.org/>>. Acessado em: 15 Mai 2025.
- TOBLER, W., DEICHMANN, U., GOTTSSEGEN, J., et al., 1997, “World population in a grid of spherical quadrilaterals”, *International Journal of Population Geography*, v. 3, n. 3, pp. 203–225.
- TUFTE, E. R., 2001, *The visual display of quantitative information*, v. 2. Graphics press Cheshire, CT.
- TVERSKY, B., 1995, “Cognitive origins of graphic productions”. In: *Understanding images: Finding meaning in digital imagery*, Springer-Verlag, pp. 29–53.
- TVERSKY, B., 2000, “Some ways that maps and diagrams communicate”. In: *Spatial cognition II: Integrating abstract theories, empirical studies, formal methods, and practical applications*, pp. 72–79. Springer, November.
- TYNER, J. A., 2014, *Principles of map design*. New York, Guilford Publications.
- UNSD, 2000, *Handbook on geographic information systems and digital mapping*. N. 79, F. United Nations Publications.
- UNSD, 2009, *Handbook on geospatial infrastructure in support of census activities*, v. 103. United Nations Publications.
- W3C, 2025a. “Cascading Style Sheets”. a. Disponível em: <<https://www.w3.org/Style/CSS/Overview.en.html>>. Acessado em: 15 Mai 2025.
- W3C, 2025b. “HTML 5”. b. Disponível em: <<https://www.w3.org/TR/2014/REC-html5-20141028/>>. Acessado em: 15 Mai 2025.
- WIKIPEDIA, 2025a. “Canvas (HTML5)”. a. Disponível em: <[https://pt.wikipedia.org/wiki/Canvas_\(HTML5\)](https://pt.wikipedia.org/wiki/Canvas_(HTML5))>. Acessado em: 15 Mai 2025.
- WIKIPEDIA, 2025b. “Javascript”. b. Disponível em: <<https://pt.wikipedia.org/wiki/JavaScript>>. Acessado em: 15 Mai 2025.

WRIGLEY, N., 1995, "Revisiting the modifiable areal unit problem and the ecological fallacy", *Diffusing Geography: Essays for Peter Haggett*, v. 49, pp. 71.

YAU, N., 2013, *Data points: Visualization that means something*. John Wiley & Sons.