

**Análise e visualização do processo  
de regionalização do Sistema  
Único de Saúde (SUS): Uma  
perspectiva em sistemas complexos**

Gabriely Rangel Pereira

DISSERTAÇÃO APRESENTADA AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA UNIVERSIDADE DE SÃO PAULO  
PARA OBTENÇÃO DO TÍTULO DE  
MESTRA EM CIÊNCIAS

Programa: Ciência da Computação

Orientador: Prof. Dr. Fabio Kon (IME-USP)

Colaboração: Prof. Dr. Luciano da Fontoura Costa (IFSC-USP)  
Instituto de Estudos para Políticas de Saúde (IEPS)

Durante o desenvolvimento deste trabalho a autora  
recebeu auxílio financeiro do CNPq e da FAPESP

São Paulo

Agosto de 2022



**Análise e visualização do processo  
de regionalização do Sistema  
Único de Saúde (SUS): Uma  
perspectiva em sistemas complexos**

Gabriely Rangel Pereira

Esta versão da dissertação contém  
as correções e alterações sugeridas  
pela Comissão Julgadora durante  
a defesa da versão original do  
trabalho, realizada em Agosto de 2022.

Uma cópia da versão original está  
disponível no Instituto de Matemática e  
Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Fabio Kon (orientador) - IME-USP
- Prof<sup>a</sup>. Dr<sup>a</sup>. Ligia Vizeu Barrozo - FFLCH-USP
- Prof. Dr. Rudi Rocha - EAESP-FGV

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

# Agradecimentos

Agradeço a orientação, apoio e criatividade do Prof. Fabio Kon na orientação deste trabalho, por identificar necessidades, incentivar melhorias, pelas sugestões, e por buscar soluções diversas.

Ao Prof. Luciano da Fontoura Costa do Instituto de Ciências Matemáticas e de Computação (ICMC-USP), agradeço o essencial suporte quanto ao tópico de sistemas complexos, ajudando ao longo de todo o trabalho na definição dos métodos escolhidos para as análises.

Pela colaboração com o Instituto de Estudos para Políticas de Saúde (IEPS), em especial ao Prof. Rudi Rocha e à Beatriz Rache, agradeço pelo apoio recorrente e por nos despertar a importância do tema analisado, o processo de regionalização da saúde no Brasil.

Ao Dr. Rubens Kon, médico sanitário pela USP, por nos descrever os principais momentos da regionalização da saúde no Brasil e suas consequências para os dias de hoje. Obrigada pela disponibilidade nas explicações e revisões.

Aos professores Nilceu Marana e Flavio Coelho, membros da banca de qualificação, que enriqueceram esta pesquisa com seus comentários e sugestões. Obrigada.

Ao meu antigo orientador de Iniciação Científica (IC), Prof. Paulo Meirelles, por ter me incentivado a iniciar o mestrado, cuja experiência me gera gratidão.

Aos professores e colegas do grupo de sistemas do IME-USP, assim como do projeto InterSCity (o INCT da Internet do Futuro para Cidades Inteligentes), que desde a IC têm me acolhido e auxiliado, especialmente em assuntos técnicos, meu agradecimento. Devo ressaltar também o auxílio do Nelson Lago, Bruno Morbin, Lucas Stankus e Giovana Martinelli, pelas revisões, sugestões na plataforma, suporte técnico e estatístico. Obrigada.

À minha família e amigos, pelo incentivo material e moral para conseguir concluir esta dissertação. Muito obrigada. Sem a ajuda de vocês, não teria conseguido.

A Deus, por me revelar sempre um sentido mais profundo naquilo que faço. Agradeço

pelo auxílio em me conduzir mais longe do que imaginava.

Por fim, a todos aqueles que contribuíram, direta ou indiretamente, para a realização deste trabalho, o meu sincero agradecimento.

# Abstract

Gabriely Rangel Pereira. **Analysis and visualization of the regionalization process of the Brazilian National Health System (SUS): A complex systems perspective.** Capstone project (Masters). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2022.

The health regionalization process, consolidated by the Brazilian Constitution of 1988 when the Unified Health System (SUS) was created, aims to guarantee the population's right to health, reducing social and territorial inequalities. A health region is composed of a collection of adjacent municipalities, organized to promote greater efficiency in providing health services. Health managers have the autonomy to identify health regions, without pre-established criteria. Thus, a data-driven tool that guides its analysis can be very useful. Our research aimed to analyze the health regionalization process using data from patient flow records in Brazilian municipalities between 1994 and 2020, from the National Hospital Information System (SIHSUS). Using complex systems techniques, we sought to analyze the patient movement network and propose means of evaluating and visualizing the behavior of health regions in each Brazilian state. For this, we used community detection methods in complex networks, techniques for modeling supply and demand flow, and georeferenced data visualization tools. With them, we analyzed the relationship between regional borders and patient flow, and defined a performance index for regions based on the rates of permanence and attraction of patients. Among the main results, we observe that the movement of patients is closely related to the health regions, and this relationship has grown in recent years. We also found that the average level of effective regionalization of the health regions has been quite stable over the last two decades, with slight and constant growth since 2010. Furthermore, we observed in the states studied that changes in the health regions configuration promoted improvements in some regions involved. Also, in the years following the modification, we observed a gradual increase in the performance of some affected regions, according to the index developed in this work. To support evidence-based public policymaking we developed an open source software web platform showing all visualizations and analyses produced here in an accessible and interactive way. This was done in a way that will allow the health manager to access a tool that reveals the evolution of regions of greater attention (lower performance index) as well as the evolution of the patient flow behavior in each Brazilian state in general.

**Keywords:** Health Regionalization. Brazilian National Health System (SUS). Hospital Admissions. Data Processing. Complex Systems. Community Detection. Complex Networks. Data Visualization. Web Application. Public Health Management. Data Science. Smart Cities.



## Resumo

Gabriely Rangel Pereira. **Análise e visualização do processo de regionalização do Sistema Único de Saúde (SUS): Uma perspectiva em sistemas complexos**. Dissertação (Mestrado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2022.

O processo de regionalização de saúde, consolidado pela Constituição de 1988, na criação do Sistema Único de Saúde (SUS), tem o objetivo de garantir o direito à saúde da população, reduzindo as desigualdades sociais e territoriais. Uma região de saúde é composta pela coleção de municípios adjacentes, organizados para promover uma maior eficiência no acesso aos serviços de saúde. Os gestores têm autonomia para identificar as regiões de saúde, sem critérios pré-estabelecidos. Logo, uma ferramenta orientada por dados que auxilie suas análises pode ser de grande utilidade. Esta pesquisa teve como objetivo analisar o processo de regionalização da saúde utilizando dados de fluxos de pacientes entre municípios brasileiros de 1994 a 2020, a partir do Sistema de Informações Hospitalares do SUS (SIHSUS). Através de técnicas de sistemas complexos, analisamos a rede de movimentação de pacientes propondo meios de avaliação e visualização do comportamento das regiões de saúde em cada estado brasileiro. Para isso, recorremos a métodos de detecção de comunidades em redes complexas, modelagem de fluxos de oferta e demanda, e ferramentas de visualização de dados georreferenciados. Com eles, analisamos a relação entre a divisão das regiões e o fluxo dos pacientes, e definimos um índice de desempenho para as regiões baseado nas taxas de permanência e atração de pacientes. Entre os principais resultados, observamos que a movimentação de pacientes está relacionada às regiões de saúde, e essa relação tem crescido nos últimos anos. Constatamos também que o nível médio de efetiva regionalização das regiões de saúde dos estados esteve bastante estável nas últimas duas décadas, com leve e constante crescimento desde 2010. Além disso, observamos que as alterações nas regiões de saúde dos estados estudados promoveram melhorias em alguma das partes envolvidas. Aliás, nos anos seguintes à modificação observamos em algumas das regiões afetadas um gradual aumento de desempenho, segundo o índice desenvolvido neste trabalho. Para apoiar a elaboração de políticas públicas com base em evidências, implementamos uma plataforma web de software livre do tipo *dashboard* exibindo todas as análises aqui desenvolvidas de forma acessível e interativa. Isso foi feito de tal modo para que o gestor de saúde tenha acesso a uma ferramenta que revele a evolução de regiões de maior atenção (menor índice de desempenho), assim como a evolução do comportamento do fluxo dos pacientes em cada estado brasileiro de modo geral.

**Palavras-chave:** Regiões de Saúde. Sistema Único de Saúde (SUS). Internações Hospitalares. Processamento de Dados. Sistemas Complexos. Detecção de Comunidades. Visualização de Dados. Plataforma Web. Gestão Pública em Saúde. Ciência de Dados. Cidades Inteligentes.



# Lista de Figuras

2.1	Algumas das competências de cada esfera administrativa de acordo com a Lei 8080/90. . . . .	8
3.1	Uma rede de 23 vértices, de estrutura modularizada, particionada em três comunidades, representadas por cores diferentes. . . . .	16
3.2	Representação de uma rede de 64 vértices em dendrograma com o gráfico da métrica <i>modularity</i> para cada partição possível utilizando o método descrito no texto. A linha vermelha tracejada indica a partição onde se obtém o maior valor para a métrica nesta rede. As 4 diferentes comunidades particionadas são representadas por 4 formatos de vértices diferentes no dendrograma (NEWMAN e GIRVAN, 2004). . . . .	17
3.3	Exemplo de um passeio de 7 passos numa rede de 5 vértices, iniciando pelo vértice A. . . . .	18
3.4	Árvore de Huffman para a compressão da sequência ABCDBCDE. . . . .	18
3.5	Processo de detecção de comunidades através da compressão da descrição do trajeto do passeio aleatório. a) Trajeto de um passeio aleatório na rede. b) Rotulagem dos vértices pela codificação de Huffman. c) Descrição do passeio aleatório em dois níveis (ROSVALL e BERGSTROM, 2008). . . . .	19
3.6	Diagrama do fluxo de pacientes em certa região de saúde e sua representação pelos coeficientes de Elzinga e Hogarty. . . . .	22
5.1	Fluxo de pacientes no estado do Rio de Janeiro em 2019 para dois diagnósticos distintos. Imagem obtida da plataforma desenvolvida neste projeto. . . . .	35
5.2	Representação visual da estrutura da plataforma web desenvolvida. Inspirada no modelo Model-View-Controller (MVC), a plataforma replica a estrutura para três seções da plataforma, representadas pelas três cores diferentes. . . . .	41

6.1	Média anual da distância percorrida por internação no Rio de Janeiro para os períodos de 2000 a 2004, 2005 a 2009, 2010 a 2014, e 2015 a 2019. Cada sub-gráfico considera tipos de fluxos diferentes: (a) todos os fluxos; (b) fluxos que não cruzam regiões; (c) fluxos que cruzam regiões; e (d) fluxos que cruzam macrorregiões. Isso para os diagnósticos relacionados à neoplasia, gravidez, sistema circulatório e todas as internações juntas. . . . .	47
6.2	Distribuição das internações de 2019 pela Região de Saúde de residência do paciente no estado do Rio de Janeiro. Os sub-gráficos (a) e (b) contém as internações para todos os diagnósticos. Já (c), (d) e (e) filtram as internações por diagnósticos específicos. As internações por habitante utilizam dados populacionais do Censo 2010. . . . .	48
6.3	Média anual da distância percorrida por internação no Ceará para os períodos de 2000 a 2004, 2005 a 2009, 2010 a 2014, e 2015 a 2019. Cada sub-gráfico considera tipos de fluxos diferentes: (a) todos os fluxos; (b) fluxos que não cruzam regiões; (c) fluxos que cruzam regiões; e (d) fluxos que cruzam macrorregiões. Isso para os diagnósticos relacionados à neoplasia, gravidez, sistema circulatório e todas as internações juntas. . . . .	51
6.4	Distribuição das internações de 2019 pela Região de Saúde de residência do paciente no estado do Ceará. As internações por habitante utilizam dados populacionais do Censo 2010. . . . .	52
6.5	Média anual da distância percorrida por internação em São Paulo para os períodos de 2000 a 2004, 2005 a 2009, 2010 a 2014, e 2015 a 2019. Cada sub-gráfico considera tipos de fluxos diferentes: (a) todos os fluxos; (b) fluxos que não cruzam regiões; (c) fluxos que cruzam regiões; e (d) fluxos que cruzam macrorregiões. Isso para os diagnósticos relacionados à neoplasia, gravidez, sistema circulatório e todas as internações juntas. . . . .	54
6.6	Distribuição das internações de 2019 pela Região de Saúde de residência do paciente no estado de São Paulo. As internações por habitante utilizam dados populacionais do Censo 2010. . . . .	55
6.7	Resultado da aplicação dos algoritmos de detecção de comunidades greedy modularity ((a) e (c)) e infomap ((b) e (d)), primeiro para as Regiões e depois para as Macrorregiões de Saúde. Baseado nos fluxos intermunicipais de pacientes em 2019 no estado do Rio de Janeiro. . . . .	56
6.8	Resultado da aplicação dos algoritmos de detecção de comunidades greedy modularity ((a) e (c)) e infomap ((b) e (d)), primeiro para as Regiões e depois para as Macrorregiões de Saúde. Baseado nos fluxos intermunicipais de pacientes em 2019 no estado do Ceará. . . . .	57

6.9	Resultado da aplicação dos algoritmos de detecção de comunidades greedy modularity ((a) e (c)) e infomap ((b) e (d)), primeiro para as Regiões e depois para as Macrorregiões de Saúde. Baseado nos fluxos intermunicipais de pacientes em 2019 no estado de São Paulo. . . . .	58
6.10	Identificação de um possível problema de regionalização no estado do Ceará (ou sugestão de expansão). Para esse resultado, foram utilizados dados de movimentação de pacientes de 2019 no algoritmo de detecção de comunidades infomap. . . . .	60
6.11	Resultado dos algoritmos de comparação de partições (a esquerda) e sua padronização (a direita), para os estados do Rio de Janeiro ((a) e (b)), Ceará ((c) e (d)) e São Paulo ((e) e (f)) desde 1994 a 2020. Comparação do resultado do infomap com as Regiões de Saúde dos três estados. . . . .	62
6.12	Nível de semelhança entre comunidades detectadas pelo infomap em comparação com Regiões (a) e Macrorregiões (b) de Saúde para os estados do Rio de Janeiro, Ceará e São Paulo. Resultado do algoritmo de comparação de partições Adjusted Rand Index para os anos entre 1994 e 2020. . . . .	63
6.13	Resultado do cálculo dos coeficientes LIFO (segunda coluna), LOFI (terceira coluna) e sua combinação (primeira coluna), como proposto em <a href="#">ELZINGA e HOGARTY, 1973</a> e <a href="#">ELZINGA e HOGARTY, 1978</a> . Para dados de fluxo de pacientes em 2019 nas Regiões de Saúde dos estados do Rio de Janeiro ((a),(b) e (c)), Ceará ((d), (e) e (f)) e São Paulo ((g), (h) e (i)). . . . .	65
6.14	Resultado do cálculo das taxas de permanência (segunda coluna), de atração (terceira coluna) e sua combinação (primeira coluna), como explicado em 5.2.2. Para dados de fluxo de pacientes em 2019 nas Regiões de Saúde dos estados do Rio de Janeiro ((a),(b) e (c)), Ceará ((d), (e) e (f)) e São Paulo ((g), (h) e (i)). . . . .	66
6.15	Resultado das taxas de permanência e atração, e sua combinação por média harmônica, para as Regiões e Macrorregiões de Saúde que contêm a capital do Rio de Janeiro ((a) e (b)), do Ceará ((c) e (d)), e de São Paulo ((e) e (f)). . . . .	68
6.16	Captura de tela da aba “Fluxo de pacientes” do dashboard desenvolvido neste projeto. . . . .	69
6.17	Captura de tela da aba “Divisões sugeridas” do dashboard desenvolvido neste projeto. . . . .	70
6.18	Captura de tela da aba “Permanência vs. atração” do dashboard desenvolvido neste projeto. . . . .	72
7.1	Fluxos intermunicipais de pacientes no estado de São Paulo em 2019, separado por alguns diagnósticos. . . . .	74

7.2	Comparação da divisão das Regiões de Saúde do Rio de Janeiro com a divisão detectada pelo algoritmo Infomap através da rede de movimentação de pacientes em 2019. . . . .	75
7.3	Comparação da divisão original das Macrorregiões de Saúde com as divisões sugeridas para 2011 e 2012 no estado do Rio de Janeiro. . . . .	76
7.4	Alteração da taxa de permanência das Regiões de Saúde próximas a Campinas, São Paulo, que sofreram modificações em seus limites administrativos em 2017. . . . .	77
7.5	Nível de desempenho das Regiões de Saúde do Ceará para os anos 1999, 2009 e 2019, combinando taxa de permanência e $\beta$ ×taxa de atração por média harmônica, onde $\beta = 0,2$ . . . . .	78

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação e Objetivos . . . . .	3
1.2	Contribuições . . . . .	3
1.3	Organização do texto . . . . .	4
<b>2</b>	<b>Funcionamento do Sistema Único de Saúde</b>	<b>7</b>
2.1	Estrutura e Funcionamento . . . . .	8
2.2	A Regionalização da Saúde . . . . .	10
2.3	Regiões e Macrorregiões de Saúde . . . . .	11
<b>3</b>	<b>Fundamentos</b>	<b>13</b>
3.1	Sistemas Complexos . . . . .	13
3.1.1	Redes Complexas . . . . .	14
3.2	Modelagem de fluxos de oferta e demanda . . . . .	20
3.2.1	Interação Espacial . . . . .	21
3.3	Outros conceitos . . . . .	23
3.3.1	Similaridade entre partições . . . . .	23
3.3.2	Média harmônica ponderada . . . . .	24
<b>4</b>	<b>Trabalhos Relacionados</b>	<b>27</b>
4.1	Análise por detecção de comunidades . . . . .	27
4.2	Análise por índices de desempenho . . . . .	29
4.3	Considerações finais . . . . .	30
<b>5</b>	<b>Metodologia</b>	<b>33</b>
5.1	Bases de Dados . . . . .	33
5.1.1	Internações Hospitalares . . . . .	34
5.1.2	Malha Municipal e Censo Demográfico . . . . .	34
5.1.3	Limites das Regiões e Macrorregiões de Saúde . . . . .	36

5.2	Abordagens de análise . . . . .	36
5.2.1	Detecção de comunidades . . . . .	37
5.2.2	Índice de desempenho . . . . .	38
5.2.3	Visualização do processo de regionalização . . . . .	39
5.3	Plataforma Web . . . . .	40
<b>6</b>	<b>Resultados</b>	<b>43</b>
6.1	Descrição estatística da movimentação de pacientes . . . . .	44
6.2	Aplicação da detecção de comunidades . . . . .	55
6.2.1	Semelhança entre partições . . . . .	61
6.3	Aplicação do índice de desempenho . . . . .	64
6.4	Aplicação da visualização do processo de regionalização . . . . .	68
6.4.1	Aba “Fluxo de paciente” . . . . .	69
6.4.2	Aba “Divisões sugeridas” . . . . .	70
6.4.3	Aba “Permanência vs. atração” . . . . .	71
<b>7</b>	<b>Discussões Finais</b>	<b>73</b>
7.1	Fluxo de pacientes . . . . .	73
7.2	Divisões sugeridas . . . . .	75
7.3	Permanência vs. atração . . . . .	76
<b>8</b>	<b>Conclusões</b>	<b>79</b>
8.1	Contribuições à pesquisa e à prática . . . . .	80
8.2	Limitações . . . . .	81
8.3	Trabalhos futuros . . . . .	82

**Referências**



# Capítulo 1

## Introdução

Com o aumento da geração e armazenamento de dados, houve um grande empenho computacional para acompanhar esse crescimento. A extração de informações em grandes conjuntos de dados foi, com isso, tomando um lugar de grande importância. As técnicas de sistemas complexos são ferramentas que realizam esse tipo de tarefa. Elas podem extrair informações não óbvias de conjuntos de dados extensos, além de detectar padrões neles. Aplicações do uso de sistemas complexos podem ser amplamente encontradas em sistemas sociais, biológicos, epidemiológicos, entre outros. O uso de sistemas complexos na saúde pública tem sido direcionado para análises como transmissão de doenças e estrutura interorganizacional de sistemas de saúde, por exemplo (LUKE e HARRIS, 2007).

O Sistema Único de Saúde (SUS) do Brasil é um dos maiores sistemas de saúde pública do mundo, garantindo acesso integral, universal e gratuito para toda a população do país <sup>1</sup>. De acordo com o Sistema de Informações de Beneficiários (SIB), gerido pela Agência Nacional de Saúde Suplementar (ANS), a taxa de cobertura de planos de saúde no Brasil é de 25,7%, pela atualização de agosto de 2022 <sup>2</sup>. Dessa forma, praticamente 3/4 da população brasileira depende apenas do SUS.

Sua organização é baseada numa estrutura de regionalização do atendimento à população. Podemos entender a relevância da regionalização de saúde no Brasil pelo fato da preocupação pela regionalização ter sido anterior à própria criação do SUS (PEREIRA *et al.*, 2006). Conforme o Pacto pela Saúde, 2006:

"[A regionalização] é a diretriz que orienta o processo de descentralização das ações e serviços de saúde e os processos de negociação e pactuação entre os gestores. [...] A regionalização objetiva garantir o direito à saúde da população,

---

<sup>1</sup><https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/s/sus-estrutura-principios-e-como-funciona>

<sup>2</sup>[http://www.ans.gov.br/anstabnet/cgi-bin/dh?dados/tabnet\\_tx.def](http://www.ans.gov.br/anstabnet/cgi-bin/dh?dados/tabnet_tx.def)

reduzindo as desigualdades sociais e territoriais por meio da identificação e reconhecimento das Regiões de Saúde."

Dessa forma, define-se uma Região de Saúde como um conjunto de municípios vizinhos escolhidos para integrar a organização, o planejamento e a execução de ações e serviços de saúde (BRASIL, 2011). Idealmente, os habitantes pertencentes a uma determinada região de saúde não precisariam ultrapassar os limites de sua região para receber os serviços de saúde, já que a região de saúde tem o objetivo de atender às demandas e interesses locais (PEREIRA *et al.*, 2006).

Pacientes que precisam ser internados por motivos de insuficiência cardíaca, por exemplo, idealmente não deveriam sair de sua região de saúde para serem internados. Assim também acontece para partos, insuficiência respiratória e outros procedimentos de média complexidade. Isso porque as Regiões de Saúde atendem procedimentos de baixa e parte de média complexidade, enquanto as Macrorregiões de saúde atendem média e alta complexidade (vide Capítulo 2).

Informações sobre internações de pacientes, seu diagnóstico e localização tanto da residência quanto da internação são disponibilizadas publicamente pelo DATASUS (Departamento de Informática do SUS) no SIH/SUS (Sistema de Informações Hospitalares). Com isso, podemos desenhar a rede de fluxos de pacientes entre municípios diferentes e analisar a movimentação desses pacientes em relação aos limites das regiões de saúde.

Analisar o deslocamento dos pacientes pode informar a adequação das regiões de saúde ao princípio de cobrir sua demanda local, mas vale lembrar que outros fatores como problemas de acessibilidade em transporte pode influenciar essa análise. De qualquer forma, nossos resultados indicam uma grande correlação entre a delimitação das regiões e a movimentação dos pacientes. Essa correlação tem crescido temporalmente, indicando uma maior adequação dos fluxos de pacientes às suas regiões de saúde nos últimos anos.

Neste trabalho, buscamos analisar o comportamento, bem como a eficiência das regiões de saúde a partir de técnicas de sistemas complexos com base na movimentação de pacientes em internações hospitalares no Sistema Único de Saúde. Além disso, disponibilizamos os resultados publicamente implementando uma ferramenta de visualização de dados, especialmente voltada a gestores de saúde, incentivando a elaboração de políticas públicas baseadas em evidências.

## 1.1 Motivação e Objetivos

Na organização e gerenciamento das regiões de saúde, os gestores possuem liberdade para delimitar as regiões de saúde, escolhendo os municípios a participar da região de acordo com critérios particulares não pré-estabelecidos. Com isso, uma ferramenta orientada por dados que auxilie suas análises poderia ser de grande utilidade. *CARDOSO et al., 2016* defendem que os problemas do acesso aos serviços só serão resolvidos através da efetivação da regionalização de saúde. Para auxiliar no processo de análise da eficiência e evolução das regiões de saúde no Brasil, buscamos responder às seguintes questões de pesquisa em relação às internações hospitalares cobertas pelo SUS:

*Q1: A movimentação dos pacientes está relacionada à delimitação das regiões de saúde? O quanto?*

*Q2: Existe uma evolução na autossuficiência das regiões de saúde ao longo dos anos?*

*Q3: As alterações das regiões de saúde ao longo dos anos interferiram na movimentação dos pacientes? Houve melhorias?*

Para alcançar esses objetivos, utilizamos métodos de sistemas complexos, analisando a rede de movimentação de pacientes e, com isso, o comportamento da população que utiliza os serviços de saúde em nível hospitalar. Para a primeira questão, comparamos os agrupamentos de municípios delimitados pelas regiões de saúde com os agrupamentos de municípios sugeridos por um algoritmo de **detecção de comunidades**, utilizando os dados do fluxo de pacientes. Dessa forma podemos compreender o quanto as regiões de saúde se refletem no fluxo de pacientes. Na segunda questão, implementamos um **índice de desempenho** que busca medir o nível de autossuficiência das regiões de saúde através desses dados, para permitir uma análise da evolução da autossuficiência das regiões ao longo do tempo. Por fim, também implementamos uma **plataforma de visualização de dados** de modo a fornecer a gestores de saúde e pesquisadores uma ferramenta que possa auxiliar no gerenciamento e análise das regiões de saúde. Ao observar o resultado do desempenho das regiões de saúde em anos diferentes, podemos responder à terceira questão.

## 1.2 Contribuições

O objetivo desta pesquisa é investigar a eficiência das regiões de saúde brasileiras, bem como sua evolução nas últimas décadas, além de desenvolver uma plataforma iterativa

que auxilie o acesso aos resultados dessa análise. Com base nesse objetivo, acreditamos ter alcançado as seguintes contribuições como parte deste trabalho.

**Verificação da relação entre a movimentação de pacientes e a delimitação das regiões de saúde brasileiras.** O processo de regionalização da saúde tem o objetivo de descentralizar o planejamento e acesso aos serviços de saúde pela população. Dessa forma, pacientes pertencentes a uma região de saúde não precisariam se locomover para outras regiões para conseguir atendimento. Nossos resultados verificam a grande relação entre o fluxo de pacientes e a configuração das regiões. Também podemos verificar as áreas em que essa relação é mais forte que outras, o que pode indicar uma maior efetivação da regionalização nelas, bem como áreas em que a regionalização não está se comportando como planejado.

**Avaliação das regiões de saúde com base na movimentação de pacientes.** Podemos encontrar na literatura o desenvolvimento de algumas métricas para a eficiência das regiões de saúde. Existe esse interesse em analisar a evolução dos efeitos da regionalização no Brasil, mas poucas iniciativas analisaram o fluxo dos pacientes. Este trabalho apresenta uma forma de avaliar a eficiência das regiões de saúde por meio do cálculo das taxas de permanência e atração de pacientes conforme a rede de movimentação de pacientes.

**Desenvolvimento de um *dashboard* de análise das regiões de saúde.** Uma boa decisão requer o acesso a evidências válidas, especialmente as decisões relacionadas à saúde pública, que atinge a tantos. Neste trabalho, desenvolvemos uma plataforma disponibilizando nossos resultados da análise às regiões de saúde de forma integrada e facilmente acessível. Isso pode incentivar tomadas de decisão baseadas em evidências por parte dos gestores de saúde, a quem a plataforma se dirige.

### 1.3 Organização do texto

Os próximos capítulos estão organizados da seguinte maneira: o Capítulo 2 explica a origem e funcionamento do SUS e define a regionalização da saúde no Brasil. O Capítulo 3 apresenta os conceitos técnicos utilizados em nossa metodologia para analisar a regionalização de saúde. O Capítulo 4 apresenta alguns trabalhos que utilizaram métodos semelhantes aos usados nesta pesquisa. O Capítulo 5 apresenta as bases de dados analisadas aqui, os métodos de análise aplicados e o desenvolvimento da plataforma. O Capítulo 6 contém os resultados obtidos através de cada um dos métodos, apresenta a plataforma final, além de fazer uma descrição estatística da movimentação de pacientes. O Capítulo 7 apresenta a utilidade prática, para os gestores, de cada resultado obtido por meio de alguns exemplos. E o Capítulo 8 contém as conclusões deste trabalho, bem como os possíveis

trabalhos futuros.



## Capítulo 2

# Funcionamento do Sistema Único de Saúde

Conforme a Constituição Federal de 1988, a “Saúde é direito de todos e dever do Estado” (BRASIL, 1988). A partir desse momento, o Sistema Único de Saúde (SUS) é criado e o sistema público de saúde passa a ser garantido a todos os cidadãos brasileiros, já que no período anterior apenas aos trabalhadores vinculados à Previdência Social tinham acesso, enquanto o restante da população tinha acesso a algumas ações de saúde da esfera pública (PAIM, 2009). A amplitude de ações e serviços prestados através do SUS inclui ações de caráter coletivo, como promoção de saúde e prevenção, além das assistenciais. Ela cobre desde a atenção básica de saúde, como consultas gerais, odontológicas, entre outros, até atendimentos de urgência ou de alta complexidade, como cirurgias de emergência ou tratamentos prolongados. Dessa forma, cada brasileiro deve ter acesso aos serviços de saúde desde o nascimento e por toda sua vida.

De acordo com a Lei Orgânica da Saúde (Lei 8080/90) que regula as ações e serviços de saúde, esses são alguns dos princípios que dão base ao SUS, inclusive em caráter organizacional: a **universalidade** de acesso aos serviços de saúde em todos os níveis de assistência; a **integralidade** de assistência, entendida como conjunto articulado e contínuo das ações e serviços; a **equidade** da assistência à saúde, sem preconceitos ou privilégios de qualquer espécie; a **participação da comunidade**; a **descentralização** político-administrativa, com ênfase na descentralização dos serviços para os municípios e incentivando a regionalização e hierarquização da rede de serviços de saúde; e a **conjugação dos recursos** da União, dos Estados, do Distrito Federal e dos Municípios na prestação de serviços de assistência à saúde da população (BRASIL, 1990).

## 2.1 Estrutura e Funcionamento

O SUS é composto pela União, Estados e Municípios através do Ministério da Saúde, Secretarias Estaduais e Municipais, de acordo com a Constituição. A cada um dos entes são reservadas suas respectivas responsabilidades, assim como exemplifica a Figura 2.1. À União cabe acompanhar, controlar e avaliar as ações e serviços de saúde. Aos Estados cabe estabelecer normas, em caráter suplementar, para o controle e avaliação das ações e serviços de saúde, além de promover a descentralização para os municípios e gerenciar alguns serviços de maior complexidade. E aos Municípios compete planejar, organizar, controlar e avaliar as ações e os serviços de saúde e executar os serviços públicos de saúde.

 <b>UNIÃO</b>	 <b>ESTADOS</b>	 <b>MUNICÍPIOS</b>
<ul style="list-style-type: none"> <li>- Participar na formulação das políticas de saúde</li> <li>- Prestar cooperação técnica e financeira aos outros entes</li> <li>- Controlar e fiscalizar procedimentos, produtos e substâncias de interesse para a saúde</li> </ul>	<ul style="list-style-type: none"> <li>- Identificar estabelecimentos hospitalares de referência e gerir sistemas públicos de alta complexidade</li> <li>- Promover a descentralização para os municípios</li> <li>- Criar as próprias políticas públicas e apoiar a execução das políticas nacionais</li> </ul>	<ul style="list-style-type: none"> <li>- Implementação das ações de saúde em sua localidade</li> <li>- Formar consórcios administrativos intermunicipais</li> <li>- Garantir os serviços de atenção básica à saúde</li> <li>- Criar políticas de saúde e apoiar as estaduais e nacionais</li> </ul>

**Figura 2.1:** Algumas das competências de cada esfera administrativa de acordo com a Lei 8080/90.

Por sua vez, as esferas administrativas se organizam em diferentes órgãos para cumprir cada responsabilidade que lhes cabem. A Tabela 2.1 apresenta a forma em que os gestores do SUS estão organizados em diferentes órgãos, representando os três níveis de governo.

Lembramos que, pela descentralização, a ênfase da atuação do SUS se encontra nos municípios. Dessa forma, a totalidade dos serviços de saúde dos municípios deve ser oferecida com a intermediação do gestor estadual em um conjunto de estabelecimentos, organizados em uma rede regionalizada e hierarquizada. Esses estabelecimentos não precisam necessariamente estar no mesmo município. Suas ações podem ser desenvolvidas municipal, estadual ou federalmente, ou ainda de forma privada, desde que os estabelecimentos estejam organizados de modo que o gestor municipal possa garantir à população o acesso aos serviços (M. S. BRASIL, 2000).

Órgão	Atuação	Nível
Ministério da Saúde (MS)	Formula, normatiza, fiscaliza, monitora e avalia políticas e ações, articulando o Conselho Nacional de Saúde. Integram sua estrutura: Fiocruz, Funasa, Anvisa, ANS, Hemobrás, Inca, Into e oito hospitais federais.	Nacional
Secretaria Estadual de Saúde (SES)	Participa da formulação das políticas e ações de saúde, presta apoio aos municípios e regiões de saúde em articulação com o conselho estadual e participa da CIB para aprovar e implementar o plano estadual de saúde.	Estadual
Secretaria Municipal de Saúde (SMS)	Planeja, organiza, controla, avalia e executa as ações e serviços de saúde em articulação com o conselho municipal e a esfera estadual para aprovar e implantar o plano municipal de saúde.	Municipal
Comissão Intergestores Tripartite (CIT)	Foro de negociação e pactuação entre gestores quanto aos aspectos operacionais do SUS.	Nacional, Estadual e Municipal
Comissão Intergestores Bipartite (CIB)	Foro de negociação e pactuação entre gestores quanto aos aspectos operacionais do SUS.	Estadual e Municipal
Coordenadoria de Regiões de Saúde (CRS)	Responsável pela coordenação e organização do sistema de saúde regional em razão das políticas estaduais e recursos disponíveis.	Estadual
Departamentos Regionais de Saúde (DRS)	Divisão administrativa regional. Responsáveis pelas atividades das SES no âmbito regional.	Estadual
Colegiado de Gestão Regional (CGR)	Espaço de decisão através da identificação de prioridades e de pactuação de soluções para a organização de uma região de saúde.	Estadual e Municipal
Conselho Nacional de Secretário da Saúde (Conass)	Entidade representativa dos entes estaduais e do Distrito Federal na CIT para tratar de matérias referentes à saúde.	Nacional
Conselho Nacional de Secretarias Municipais de Saúde (Conasems)	Entidade representativa dos entes municipais na CIT para tratar de matérias referentes à saúde.	Nacional
Conselhos de Secretarias Municipais de Saúde (Cosems)	Entidades que representam os entes municipais, no âmbito estadual, para tratar de matérias referentes à saúde.	Estadual

**Tabela 2.1:** Estrutura do Sistema Único de Saúde (SUS) e organismos de articulação política dos secretários (os conselhos). Participação de cada órgão de acordo com a esfera de atuação (M. S. BRASIL, 2021; A. L. VIANA et al., 2008).

## 2.2 A Regionalização da Saúde

O processo de regionalização da saúde tem início antes mesmo da criação do SUS, pela Constituição de 1988, onde se consolidou (PEREIRA *et al.*, 2006). Na Constituição de 1988 foram definidos alguns princípios para o SUS, entre eles a descentralização, em combinação com a regionalização e a hierarquização dos serviços de saúde.

Descentralizar é repassar atribuições e meios de um órgão central a órgãos locais (ALCÂNTARA, 2011). A regionalização é um processo político e técnico restrito pela capacidade de oferta e financiamento dos serviços de saúde para a população (LIMA, VIANA *et al.*, 2012). A hierarquização é definida pela organização dos níveis de complexidade dos serviços de saúde com o objetivo de oferecer acesso aos serviços de saúde em qualquer nível de atenção através da contribuição das três esferas de governo (CARDOSO *et al.*, 2016). Assim, a distribuição territorial das ações e serviços e o escalonamento da complexidade tecnológica dos serviços ficam de acordo com a cobertura populacional prevista, isto é, serviços de maior complexidade abrangendo territórios e populações maiores.

Segundo CARDOSO *et al.*, 2016, a partir da Lei Orgânica da Saúde, o processo de descentralização passou a ter ênfase na municipalização, isto é, na descentralização dos serviços, até então com direção nacional e estadual, para os municípios. Naquela época, houve pouca ênfase em relação à regionalização até a criação do pacto pela saúde em 2006. A questão da regionalização veio como resposta organizacional à incapacidade da grande maioria dos municípios, que têm pequeno porte, em garantir a integralidade do cuidado a seus cidadãos, sem contar com o apoio de municípios maiores e das outras esferas. Entre 2006 e 2011, foram implantados mais de 400 Colegiados de Gestão Regional no país, abrangendo mais de 95% dos municípios brasileiros. Instituídos pelo pacto de 2006, esses Colegiados constituem canais de negociação e decisão intergovernamental que visam fortalecer a governança regional do SUS (LIMA, QUEIROZ *et al.*, 2012). Com o Decreto nº 7508 de 2011 (BRASIL, 2011) que regulamenta a Lei Nº 8080, a região de saúde se apresenta como o espaço geográfico contínuo constituído por agrupamentos de municípios com a finalidade de integrar a organização, o planejamento e a execução de ações e serviços de saúde.

Dessa forma, os gestores têm autonomia para identificar as regiões de saúde, sem critérios pré-estabelecidos, decidindo as ações e os tipos de atendimento em cada região, contemplando os critérios que propiciem suficiência para a atenção integral da população.

## 2.3 Regiões e Macrorregiões de Saúde

As Regiões de Saúde são partições territoriais que agrupam municípios vizinhos. Gestores estaduais e municipais têm a responsabilidades de identificá-las, tendo como base características sociais, de infraestrutura e transporte da região, assim como informações sobre a eficiência da saúde pública local. Essas regiões organizam as ações e serviços de saúde de forma a garantir o atendimento integral da população dos municípios que as compõem. Cada região deve garantir o acesso à atenção básica e a parte da média complexidade.

As Macrorregiões de Saúde são agrupamentos de Regiões de Saúde. Elas têm o objetivo de organizar entre si ações e serviços de média e alta complexidade. Dessa forma, o acesso integral aos serviços, em todos os níveis, fica garantido à população do território que compõe cada macrorregião de saúde, já que isso complementa o que as Regiões de Saúde fornecem (M. S. BRASIL, 2006).

Neste texto, utilizaremos o termo “regiões de saúde” para se referir a ambos os níveis, fazendo a classificação específica apenas em momentos em que a distinção for necessária. Veremos neste trabalho o comportamento de ambas delimitações territoriais por meio das análises aqui desenvolvidas.



# Capítulo 3

## Fundamentos

Veremos agora alguns conceitos e técnicas abordadas ao longo desta pesquisa para um melhor entendimento da proposta que este trabalho apresenta e dos resultados obtidos. Dessa forma, separamos este conteúdo inicial em três seções distintas. Na primeira parte veremos quais são as técnicas de **Sistemas Complexos** utilizadas para analisar o processo de regionalização de saúde, na Seção 3.1. Na segunda parte observaremos o que são os **Modelos de Fluxo de Oferta e Demanda** e como eles foram utilizados para analisar e avaliar a autossuficiência das regiões de saúde, a Seção 3.2. Por fim, veremos alguns **Outros conceitos** utilizados aqui de forma secundária para auxiliar as abordagens principais, na Seção 3.3.

### 3.1 Sistemas Complexos

Sistemas complexos estão presentes de forma abundante no mundo real. Estes são alguns exemplos em que encontramos o comportamento de um sistema complexo: no funcionamento dos neurônios no cérebro humano, numa colônia de formigas, no tráfego de veículos em uma rodovia, na coleção dos artigos científicos já publicados, na World Wide Web (WWW), nas redes sociais, entre outros casos. Observamos uma característica em comum entre eles: cada um de seus elementos não age através de uma visão geral do sistema, mas o sistema se mantém e evolui através das ações individuais de cada elemento, sendo que eles possuem um conhecimento apenas local. Embora não possua uma coordenação geral, como um líder, o sistema complexo apresenta uma estrutura organizada, com resultados interessantes e até surgimento de padrões inesperados em alguns casos (BARRAT *et al.*, 2008; BOCCARA, 2010). Por exemplo, ZHONG *et al.*, 2014 apresentam uma aplicação atual de análise por sistemas complexos. O trabalho estuda a dinâmica da estrutura urbana, a partir da movimentação de pessoas, para identificar a estrutura espacial das cidades, elemento

importante para entender as interações urbanas.

Não há uma definição única e comumente aceita de complexidade. Entretanto, um ponto que geralmente caracteriza um sistema complexo é sua **propriedade emergente** (BARRAT *et al.*, 2008). Embora não seja a única, a propriedade de emergência é a principal em um sistema complexo. Um sistema considerado complexo é geralmente composto por um conjunto de elementos que conseguem interagir entre si localmente de forma que produzem um fenômeno emergente, algum padrão ou efeito observado em larga escala. Essa propriedade é definida pelos efeitos vistos em larga escala de interações locais entre os elementos do sistema (BOCCARA, 2010). Dessa forma, observamos que em sistemas complexos o todo é mais que a soma de suas partes. Isto é, o comportamento geral do sistema possui características que vão além das que são observadas no comportamento individual dos seus elementos (MITCHELL, 2009; BARRAT *et al.*, 2008).

Em uma rede social, por exemplo, existe uma probabilidade maior de um indivíduo se conectar a indivíduos mais populares do que a indivíduos mais isolados. Especificando dessa forma o comportamento local de elementos em uma rede social, outro comportamento geral, não especificado, pode ser observado. Sua estrutura apresenta um padrão próprio, claramente diferente da estrutura de uma rede de conexões aleatórias, por exemplo. Esse padrão emergente se forma na rede mesmo que seus elementos não tenham a intenção e talvez nem o conhecimento de que esse efeito esteja sendo por eles produzido.

A área de sistemas complexos conta com um conjunto de técnicas e conceitos desenvolvidos inicialmente em outras áreas como estatística, computação, física e geografia. Isso engloba modelagens, correlações, simulações, algoritmos e conceitos de comportamento caótico, dinâmica não-linear e teoria de grafos (AMARAL e OTTINO, 2004). Além disso, suas aplicações também são observadas em diversas áreas do conhecimento em sistemas biológicos, físicos, econômicos, sociais, epidemiológicos, tecnológicos, entre outros.

### 3.1.1 Redes Complexas

Um grafo é composto por um conjunto finito de vértices, um conjunto finito de arestas e uma regra que defina quais vértices são conectados pelas arestas (BIGGS *et al.*, 1986). As arestas têm a função de conectar dois vértices de forma direcionada ou não direcionada. Cada vértice e cada aresta pode possuir propriedades individuais, como um nome e um peso específico. A origem da teoria de grafos remonta ao século XVIII, quando Leonhard Euler (1707–1783) demonstra que o problema das pontes de Königsberg não tem solução. Uma rede complexa é a representação de um sistema complexo através de conceitos da teoria de grafos (STROGATZ, 2001). Uma rede complexa representa a relação dos elementos — os vértices — entre si.

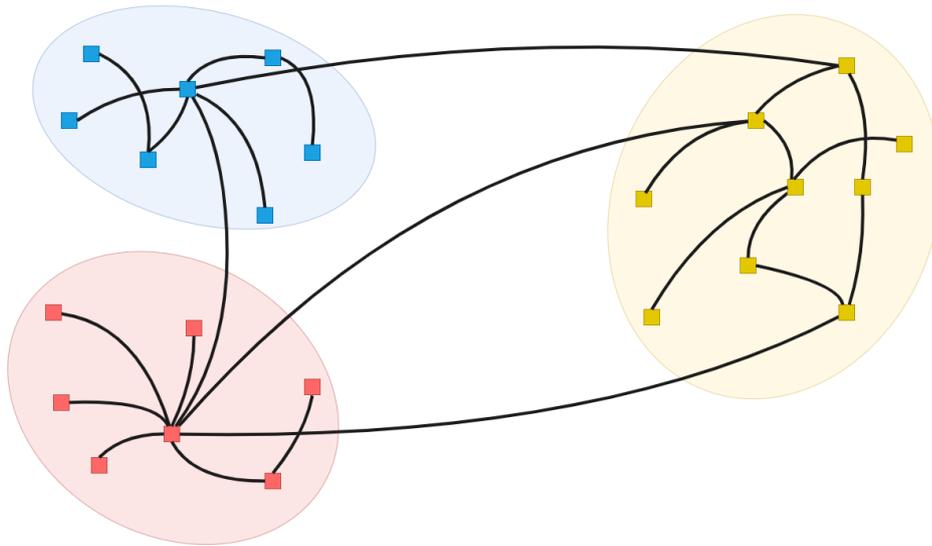
Neste projeto, por exemplo, uma rede complexa é criada para modelar o fluxo intermunicipal de internações hospitalares atendidas pelo SUS nas últimas décadas. Em nosso caso, os vértices são municípios de um mesmo estado brasileiro e as arestas indicam a movimentação dos pacientes do município de residência ao município de atendimento. Em alguns momentos as análises mudam de nível, de municipal para regional. Nesses momentos os vértices são regiões ou macrorregiões de saúde. De toda forma, todas as análises buscam investigar o comportamento ou eficiência da regionalização da saúde nos estados estudados.

Outras áreas, além da saúde pública, aplicam conceitos de redes complexas para modelar o comportamento de sistemas biológicos, sociais, físicos, entre outros (BARRAT *et al.*, 2008). O conceito de complexidade em redes, assim como em sistemas complexos, também pode ser caracterizado pela presença de algum comportamento emergente, como explicado anteriormente. Veremos a seguir um possível comportamento emergente em redes complexas. Ele é colocado em evidência por uma técnica aplicada na rede visando identificar a tendência de agrupamento entre os elementos da rede, a detecção de comunidades.

### **Detecção de comunidades**

O objetivo das técnicas de detecção de comunidades é tentar inferir relações entre vértices da rede que não podem ser mensuradas diretamente por observação (FORTUNATO, 2010). Essas técnicas analisam a distribuição/concentração de arestas entre os vértices através de métricas de centralidade, distância ou similaridade, por exemplo, e fornecem os possíveis agrupamentos encontrados. Dessa forma, cada comunidade detectada é um subgrafo do grafo definido pela rede original, e a união de todos os subgrafos contém todos os vértices da rede original. A Figura 3.1 exemplifica esse conceito.

Esse procedimento pode ser realizado por algoritmos de categorias diferentes. A escolha de um bom algoritmo para cada caso é uma tarefa que demanda um conhecimento mais amplo da literatura, pois as diferentes categorias de algoritmos consideram aspectos diferentes da rede. Neste projeto, foram testados dois algoritmos: o *greedy modularity* (GIRVAN e NEWMAN, 2002; NEWMAN e GIRVAN, 2004; NEWMAN, 2004b; CLAUSET *et al.*, 2004) e o *infomap* (ROSVALL e BERGSTROM, 2008; ROSVALL, AXELSSON *et al.*, 2009; ROSVALL e BERGSTROM, 2011).



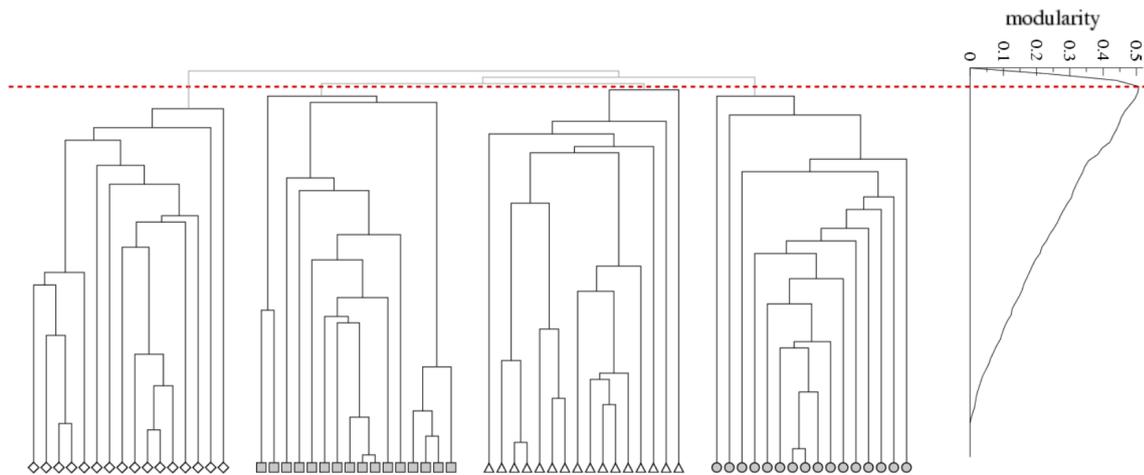
**Figura 3.1:** Uma rede de 23 vértices, de estrutura modularizada, particionada em três comunidades, representadas por cores diferentes.

**Greedy modularity** É um dos mais conhecidos algoritmos que utilizam o método divisivo. Esse método identifica as arestas que interconectam comunidades diferentes e as removem da rede, de forma que as comunidades terminam desconectadas umas das outras. Para isso, se utiliza a métrica de *centralidade de intermediação*, definida como:

$$B_e = \sum_{i,j} \frac{\sigma(i, e, j)}{\sigma(i, j)},$$

onde  $\sigma(i, e, j)$  é a quantidade de caminhos mínimos entre os vértices  $i$  e  $j$  que passam pela aresta  $e$ ;  $\sigma(i, j)$  é a quantidade total de caminhos mínimos entre os vértices  $i$  e  $j$ ; e a soma passa por todos os pares de vértices  $i$  e  $j$  distintos (L. d. F. COSTA *et al.*, 2007). Para identificar essas arestas que interconectam as comunidades, o algoritmo calcula a *centralidade de intermediação* para cada aresta e remove a que tiver o maior valor. Faz isso repetidamente criando as possíveis divisões da rede e escolhe a divisão de comunidades ideal através de um critério de parada, como apresentado na Figura 3.2. O valor da *centralidade de intermediação* para uma aresta é determinado pela quantidade de vezes em que a aresta é utilizada em todos os caminhos mínimos possíveis da rede. Dessa forma, as arestas removidas são as mais utilizadas para conectar quaisquer dois vértices na rede (FORTUNATO, 2010).

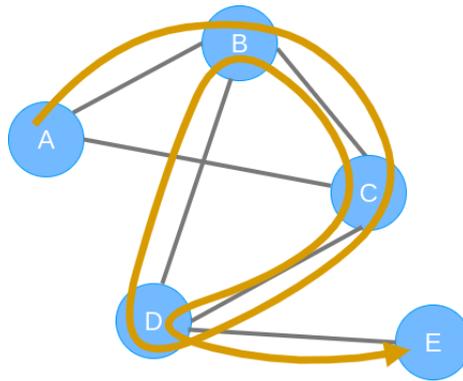
O critério de parada utilizado para este algoritmo é a maximização da métrica *modularity*. Ela valoriza as divisões que possuem uma maior quantidade de arestas que não cruzam comunidades diferentes, isto é, arestas com vértices na mesma comunidade. Mas isso não é



**Figura 3.2:** Representação de uma rede de 64 vértices em dendrograma com o gráfico da métrica *modularity* para cada partição possível utilizando o método descrito no texto. A linha vermelha tracejada indica a partição onde se obtém o maior valor para a métrica nesta rede. As 4 diferentes comunidades particionadas são representadas por 4 formatos de vértices diferentes no dendrograma (NEWMAN e GIRVAN, 2004).

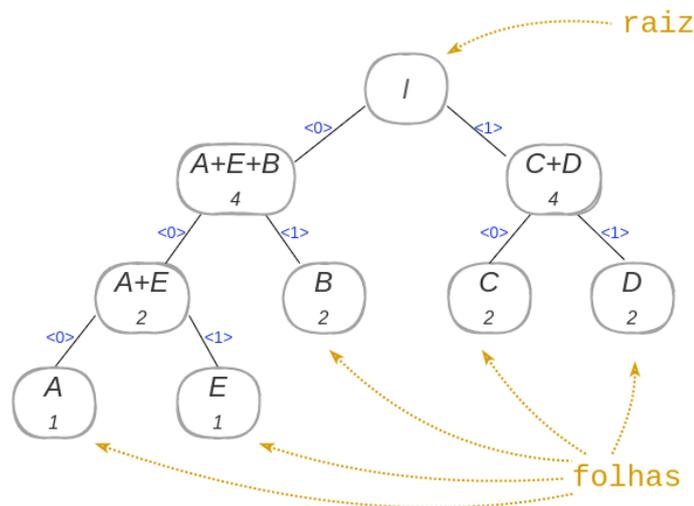
o suficiente, pois, por exemplo, uma única comunidade abrangendo toda a rede ocasionaria no maior número de arestas da mesma comunidade embora não seja a melhor divisão na maioria dos casos. Portanto, a métrica *modularity* é composta pela quantidade de arestas cujas extremidades pertencem à mesma comunidade na rede, menos a quantidade esperada dessas arestas em uma rede aleatória. Logo, a métrica será igual a zero quando a quantidade de arestas da mesma comunidade não é melhor que na versão aleatória; e a métrica cresce indicando uma melhor divisão (NEWMAN e GIRVAN, 2004). Uma otimização gulosa (*greedy*) para a métrica *modularity* é proposta pelos algoritmos apresentados por NEWMAN, 2004b e CLAUSET *et al.*, 2004, sendo este último o que é utilizado no presente projeto.

**Infomap** Esse algoritmo considera que a rede representa um sistema com interações locais carregando informações que fluem através de arestas direcionadas e com pesos, algo particularmente fundamental para este trabalho. Para isso, o *infomap* utiliza passeio aleatório (*random walk*) como um *proxy* para medir a difusão de informação na rede. Um passeio aleatório é feito considerando as direções das arestas e seus pesos, essas informações influenciam as decisões do passeio de forma que afetam a frequência em que cada vértice é visitado — semelhante ao que acontece no fluxo real da rede (ROSVALL, AXELSSON *et al.*, 2009). Por meio do passeio, cada vértice é rotulado utilizando um método de compressão de informação — a codificação de Huffman (HUFFMAN, 1952) — com base na frequência que o vértice é visitado pelo passeio. Para ilustrar o funcionamento da codificação de Huffman, vejamos um exemplo de codificação de um passeio de 7 passos numa rede de 5 vértices, como apresenta a Figura 3.3.



**Figura 3.3:** Exemplo de um passeio de 7 passos numa rede de 5 vértices, iniciando pelo vértice A.

O passeio pode ser descrito pela sequência de vértices visitados, nesse caso: *ABCDBCDE*. A menor codificação que podemos utilizar para representá-los em binário é de três bits por caractere, da seguinte forma: A = 000; B = 001; C = 010; D = 011; E = 100. Sendo assim, a sequência se transforma em 000001010011001010011100, um total de 24 bits. A codificação de Huffman tem como base a frequência em que cada vértice é visitado. Isto é, A é visitado 1 vez, B 2 vezes, C 2 vezes, D 2 vezes e E 1 vez. Com isso, o algoritmo monta uma árvore binária, a árvore de Huffman, unindo os caracteres menos frequentes dois a dois, gerando novos nós, e adicionando os valores 0 e 1 para cada par de arestas partindo de algum nó. A Figura 3.4 apresenta a árvore de Huffman para este exemplo.

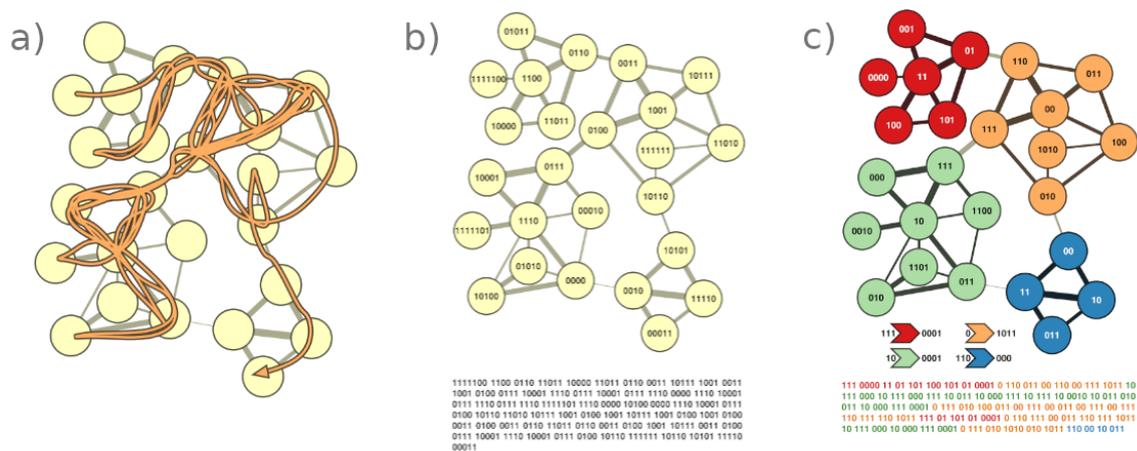


**Figura 3.4:** Árvore de Huffman para a compressão da sequência *ABCDBCDE*.

Podemos identificar os novos códigos para cada caractere percorrendo a árvore de Huffman desde a raiz até a folha do caractere correspondente guardando os valores de cada aresta percorrida (em azul na árvore). Agora, os caracteres mais frequentes possuem uma codificação menor, enquanto os menos frequentes podem ter uma codificação maior.

Dessa forma, podemos chegar a uma codificação comprimida para os caracteres, sendo ela: A = 000; B = 01; C = 10; D = 11; E = 001. Sendo assim, a sequência comprimida é 000011011011011001, somando 18 bits, economizando 6 bits em relação à sequência original.

Voltando para as comunidades, a Figura 3.5a mostra o trajeto de um possível passeio aleatório numa rede de estrutura modularizada e a Figura 3.5b apresenta a aplicação da codificação de Huffman para esse passeio como primeira etapa para detectar as comunidades. Portanto, os vértices mais visitados são rotulados com códigos menores (com menos bits), com a intenção de minimizar o comprimento esperado da descrição de um passeio aleatório, isto é, compactar a sequência com os códigos dos vértices visitados.



**Figura 3.5:** Processo de detecção de comunidades através da compressão da descrição do trajeto do passeio aleatório. a) Trajeto de um passeio aleatório na rede. b) Rotulagem dos vértices pela codificação de Huffman. c) Descrição do passeio aleatório em dois níveis (ROSVALL e BERGSTROM, 2008).

Para decompor a rede em comunidades, como mostra a Figura 3.5c, o algoritmo dá um passo a mais na compressão da informação: ele considera outro nível de relacionamento entre os vértices analisando a frequência que o passeio aleatório permanece em certos grupos, nomeando-os conforme a frequência em que são acessados. Como explicado por ROSVALL e BERGSTROM, 2008, os rótulos obtidos pelo passeio são compactados de forma semelhante ao que é feito produzindo um mapa (daqui vem o nome *infomap*): o cartógrafo precisa balancear as informações que estarão presentes no mapa entre informações mais e menos relevantes, dando maior ênfase às estruturas mais relevantes e frequentes, e menor ênfase às outras, porque o espaço do mapa é limitado. Por exemplo, consideremos cada cidade do mapa como uma comunidade e o conjunto de ruas como seus vértices. Sabemos que duas ruas em duas cidades diferentes podem ter nomes iguais sem causar muita confusão, pois temos a informação da cidade a qual cada uma pertence. Semelhantemente, a rede é dividida em dois níveis de descrição: comunidades e vértices. Uma codificação

única de Huffman é gerada para nomear as comunidades (as estruturas mais relevantes) e outras codificações são geradas para cada comunidade sem a necessidade de unicidade, pois não causam confusão ao pertencerem a comunidades diferentes. A compressão da codificação nessa forma é mais eficiente do que a codificação em apenas um nível — o nível dos vértices, como visto anteriormente. Diferente dos algoritmos que otimizam a métrica *modularity*, o *infomap* otimiza a métrica *map equation*, que por sua vez valoriza o fluxo e a dinamicidade da rede para compreender melhor a sua estrutura (ROSVALL e BERGSTROM, 2008). Para isso, o *infomap* busca minimizar a *map equation* definida, resumidamente, como:

$$L(M) = qH(Q) + \sum_{i=1}^m p_i H(P_i)$$

onde  $L(M)$  é a quantidade média de bits utilizada por passo no passeio aleatório (o valor a ser minimizado);  $M$  é a partição atual da rede entre  $m$  módulos/comunidades;  $q$  é a probabilidade do passo estar atravessando comunidades;  $H(Q)$  é a quantidade média de bits utilizada para movimentos entre comunidades;  $p_i$  é a probabilidade do passo estar na comunidade  $i$ ; e  $H(P_i)$  é a quantidade média de bits utilizada para movimentos na comunidade  $i$ .

Logo,  $qH(Q)$  é a quantidade média de bits necessária para descrever movimentos que atravessam comunidades e  $\sum_{i=1}^m p_i H(P_i)$  representa a quantidade média de bits necessária para descrever movimentos nas comunidades.

Dessa forma, o algoritmo *infomap* para detecção de comunidades é equivalente ao problema de compressão de informação (ROSVALL e BERGSTROM, 2007). Seu objetivo é minimizar a informação (quantidade de bits) necessária para descrever cada passo do passeio aleatório na rede utilizando dois níveis de descrição, gerando os agrupamentos.

Passando para o contexto analisado neste trabalho, a rede utilizada para gerar agrupamentos é a rede de fluxos intermunicipais de pacientes. Neste caso, o algoritmo de detecção de comunidades gera agrupamentos de municípios. Como as regiões de saúde também são agrupamentos de municípios, faremos, por fim, uma análise comparativa dos dois resultados analisando sua similaridade.

## 3.2 Modelagem de fluxos de oferta e demanda

Outro conceito utilizado neste trabalho é a modelagem de fluxos de oferta e demanda. Um fluxo baseado em demanda é definido como um tráfego fluindo entre a origem e o destino dentro de um sistema que interage como um todo (HODGSON *et al.*, 1996). Esse

conceito é utilizado no contexto de áreas de influência de hospitais para oferta de serviços hospitalares e demanda de pacientes com necessidade de internação. Em nosso caso, as áreas de influência são as regiões de saúde, de forma que a análise do fluxo de internação entre regiões apresente resultados interessantes para uma avaliação de suficiência dessas mesmas regiões.

### 3.2.1 Interação Espacial

A modelagem de sistemas complexos se refere ao estudo da dinâmica da interação entre os elementos. Para isso, tem-se utilizado matemática não-linear, simulações computacionais e visualização de dados (Allan G WILSON, 2002). A interação espacial se apresenta essencialmente como fluxos de elementos em um sistema (HAYES e Alan Geoffrey WILSON, 1971). Ela é um processo que permite a análise do comportamento do sistema em que entidades em diferentes pontos de um espaço físico têm contato, tomam decisões de demanda e oferta, ou realizam escolhas regionais (OLIVEIRA, 2004). Os modelos de interação espacial seguem o conceito do modelo gravitacional. A lei gravitacional de varejo, introduzida por Reilly, motivou a criação dos modelos gravitacionais. Essa lei afirma que o cliente está disposto a percorrer uma longa distância para lojas maiores, quanto maior for a atração que elas apresentam aos clientes (REILLY, 1931). Neste caso, a lei pode ser facilmente aplicada na questão do deslocamento de pacientes para hospitais, sendo o nível de atração relacionado a disponibilidade e qualidade dos recursos do hospital. O termo “gravitacional” se refere à lei da gravidade na Física, que se comporta de forma semelhante ao trocar a atratividade pela massa. Esse tipo de modelo tem como característica principal a modelagem do comportamento de segmentos de oferta e demanda (ROY e THILL, 2003). O fluxo intermunicipal de pacientes para atendimento hospitalar oferece uma estrutura de oferta e demanda dos serviços de saúde. Em estudos de mercados e de áreas de influência hospitalar os coeficientes *Little In From Outside* (LIFO) e *Little Out From Inside* (LOFI) (ELZINGA e HOGARTY, 1973; ELZINGA e HOGARTY, 1978) já foram utilizados anteriormente (RIGOLI *et al.*, 2019; FRECH III *et al.*, 2003).

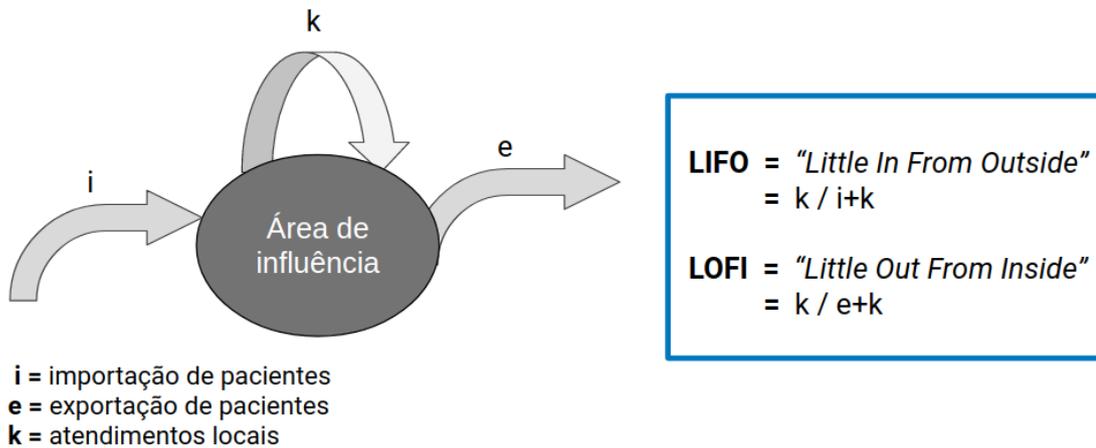
#### Coefficientes LIFO e LOFI

O método desenvolvido por Elzinga e Hogarty em 1973 busca resolver o problema de delimitação da área do mercado geográfico. Conforme a Enciclopédia Jurídica da PUCSP:

*Fala-se de mercado relevante geográfico enquanto uma determinada área na qual empresas atuam na oferta e demanda dos produtos ou serviços, em condições de concorrência suficientemente homogêneas e claramente distinguíveis em*

suas diferenças das áreas vizinhas. (FERRAZ JUNIOR, 2017)

Essa área compreende o espaço onde as empresas atuam em certa concorrência de forma que a área seja relativamente **autocontida** (ou **autossuficiente**) (FRECH III *et al.*, 2003). Baseado nessa ideia, os coeficientes foram criados para definir o nível de suficiência de uma determinada área através de duas medidas.



**Figura 3.6:** Diagrama do fluxo de pacientes em certa região de saúde e sua representação pelos coeficientes de Elzinga e Hogarty.

No contexto de atendimento hospitalar, o coeficiente LIFO (*Little In From Outside*) indica que os hospitais atendem poucos pacientes de fora da região de saúde em que está localizado. Ele mede a proporção dos pacientes que residem e são atendidos nesta região, dentre todos os atendimentos hospitalares que essa mesma região realizou. Dessa forma, a medida será inversamente proporcional à quantidade de pacientes que residem fora da região. Já o coeficiente LOFI (*Little Out From Inside*) indica que poucos pacientes da região são atendidos fora dela. Especificamente, ele mede a proporção dos pacientes que residem e são atendidos nesta região, dentre todos os pacientes que residem nesta região e são atendidos dentro ou fora dela. Isto é, ele é inversamente proporcional à quantidade de pacientes que buscam atendimento fora de sua região. Os coeficientes são valores entre 0 e 1 calculados da seguinte forma:

$$\text{LIFO} = \frac{\text{atendimentos de dentro da própria região}}{\text{atendimentos de dentro} + \text{recebidos de fora da região}};$$

$$\text{LOFI} = \frac{\text{atendimentos de dentro da própria região}}{\text{atendimentos de dentro} + \text{mandados para fora da região}}.$$

A Figura 3.6 apresenta o cálculo dessas medidas em relação ao fluxo de pacientes em certa

região de saúde.

No contexto de regiões de saúde, a combinação desses dois coeficientes pode determinar o nível de suficiência dessas regiões. Veremos com mais detalhes na Seção 5.2.2 como isso foi utilizado para analisar e avaliar a delimitação das regiões de saúde no Brasil.

### 3.3 Outros conceitos

Para aplicar as técnicas apresentadas nas seções anteriores, a fim de responder às questões de pesquisa deste projeto, fizemos uso de alguns outros conceitos que aparecem aqui de forma secundária. Esses conceitos, juntamente com as técnicas principais, são essenciais para que a pesquisa alcance o objetivo final. Vejamos a seguir esses conceitos secundários, a métrica de similaridade entre partições e a média harmônica ponderada.

#### 3.3.1 Similaridade entre partições

As métricas de similaridade entre partições avaliam o quão semelhante são duas partições de um mesmo conjunto de elementos. No contexto de redes complexas, uma partição  $M$  de uma rede a separa em  $m$  módulos/comunidades diferentes, agrupando os vértices em comunidades. Por exemplo, na Figura 3.1 a rede apresentada está particionada em três módulos diferentes, cada um representado por uma cor. Para calcular o nível de similaridade de duas partições, elas precisam ser partições de um mesmo conjunto de elementos e devem ter a mesma quantidade de módulos.

A principal aplicação dessa métrica se encontra na avaliação do desempenho de algoritmos de clusterização, algoritmos de aprendizado não supervisionado que geram agrupamentos (*clusters*) em um conjunto de dados (JAIN e DUBES, 1988). Nesse cenário, a métrica de similaridade de partições compara a partição obtida pelo algoritmo de clusterização, com a partição real dos dados em classes diferentes. Logo, o nível de similaridade indica o nível de eficiência da clusterização em ambientes de treinamento e teste.

Vários algoritmos já foram desenvolvidos para calcular essa métrica, como o *V-measure* (ROSENBERG e HIRSCHBERG, 2007), *Adjusted Rand index* (HUBERT e ARABIE, 1985; STEINLEY, 2004), *Fowlkes-Mallows score* (FOWLKES e MALLOWES, 1983), entre outros. Seu valor geralmente varia entre 0 e 1, onde um alto valor indica uma maior semelhança entre as partições. Em nosso contexto de análise das regiões de saúde, essa métrica é utilizada na rede de movimentação de pacientes entre municípios de um estado específico. O algoritmo compara duas partições da mesma rede: a divisão original dos municípios em regiões de saúde e a divisão sugerida pelo método de detecção de comunidades, apresentado anteriormente

(vide Seção 3.1.1). Com isso, conseguimos comparar numericamente a semelhança entre as comunidades sugeridas pelo algoritmo e as regiões de saúde.

Assim como será explicado na Seção 6.2.1, após testarmos diferentes algoritmos de comparação de partições, optamos por utilizar o *Adjusted Rand Index*. Esse algoritmo realiza a comparação dos resultados de acordo com a frequência em que cada par de elementos é classificado da mesma forma. Isto é, sendo  $M$  o conjunto de  $m$  municípios de um estado brasileiro,  $R$  é a partição dos municípios em  $n$  regiões de saúde e  $S$  a partição em  $n$  agrupamentos sugeridos. Para todo par  $m_i, m_j$  de municípios em  $M$  (para  $i \neq j$ ), calcula-se a frequência em que  $m_i$  e  $m_j$  estão ambos no mesmo agrupamento para  $R$  e  $S$ , ou estão ambos classificados em agrupamentos diferentes em  $R$  e  $S$ . Esse valor em relação a todos os pares possíveis  $\binom{m}{2}$  é o *Rand Index*. Por fim, para se obter o *Adjusted Rand Index*, é feita uma correção para que os casos de classificações aleatórias não gerem um índice alto. Isso é obtido ao se estabelecer um modelo aleatório como ponto de partida, de forma que o valor final seja próximo de zero quando a classificação se aproxima da aleatória.

### 3.3.2 Média harmônica ponderada

Na matemática, existem diferentes formas de se calcular a média de um conjunto de valores. Entre as mais conhecidas estão a *média aritmética*, a *média geométrica* e a *média harmônica*. A média aritmética é a mais utilizada no nosso dia-a-dia para situações simples, a média geométrica é geralmente utilizada para comparar valores de escalas diferentes, e a média harmônica é indicada para calcular a média de duas ou mais taxas (valores entre 0 e 1).

O cálculo da média harmônica de um conjunto de elementos  $x_1, x_2, \dots, x_n > 0$  é definida pela equação:

$$\bar{h} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Para estabelecer pesos diferentes  $w_i$  para cada valor  $x_i$ , calculamos a média harmônica ponderada, definida como:

$$\bar{h}(w) = \frac{\sum_{i=1}^n w_i}{\sum_{i=1}^n \frac{w_i}{x_i}}$$

A média harmônica é aplicada em diferentes áreas do conhecimento. Por exemplo, na Física, ao calcular a média de velocidades, em Finanças, na relação de preço e ganho, na Ciência da Computação, em aprendizado de máquina e recuperação de informação (com o algoritmo *F-score*), entre outras áreas. No contexto deste trabalho, a média harmônica ponderada é utilizada para ponderar os dois coeficientes LIFO e LOFI modificados (vide 3.2.1), para gerar um indicador final para o desempenho das regiões de saúde, conforme a

movimentação de pacientes entre as regiões. No Capítulo 5 observaremos como os dois coeficientes serão ponderados para avaliar o desempenho da regionalização dos estados brasileiros.

Antes de apresentarmos a metodologia desenvolvida nesta pesquisa, mostraremos no próximo capítulo como os principais conceitos descritos aqui são abordados pela literatura. O Capítulo 4 apresenta os trabalhos relacionados a este e faz uma breve comparação deles com o que desenvolvemos nesta dissertação.



# Capítulo 4

## Trabalhos Relacionados

Este capítulo tem o objetivo de apresentar trabalhos relacionados ao que estamos realizando neste projeto. Os trabalhos abaixo estão organizados em duas seções. A Seção 4.1 apresenta trabalhos que fazem uso de **detecção de comunidades** para analisar o comportamento de sistemas que possuem uma estrutura com alguma modularidade. Entre eles, existem casos que utilizam redes espaciais, fluxos de origem-destino e casos dentro e fora do contexto de saúde pública. A Seção 4.2 apresenta trabalhos que fazem uso de **índices de desempenho**, de desenvolvimento próprio ou não, para analisar regiões através do fluxo estabelecido entre elas no sistema. Cada trabalho apresentado na Seção 4.2 faz uso de métodos que partem ou da dinâmica de interação espacial, como explicado na Seção 3.2.1, ou de métodos qualitativos que analisam especificamente o processo de regionalização de saúde no Brasil. Por fim, a Seção 4.3 apresenta, dentre os trabalhos apresentados nas duas seções anteriores, quais são os que mais se relacionam com a nossa pesquisa.

### 4.1 Análise por detecção de comunidades

*MINERBA et al., 2008* analisam a estrutura dos departamentos do Hospital Universitário Policlínico de Cagliari, na Itália, buscando encontrar a quantidade e a organização otimizada de departamentos do hospital. Diferente do nosso trabalho, Minerba et al. não levam em consideração o atributo espacial da rede, já que ela não é composta por elementos espaciais: são médicos que possuem conexões entre si conforme a similaridade entre eles. Os autores utilizam a métrica de modularidade (*NEWMAN e GIRVAN, 2004*) na detecção de comunidades para encontrar a organização de departamentos ideal neste caso.

GUO *et al.*, 2018 aplicam uma nova abordagem de detecção de comunidades voltada especificamente para encontrar comunidades em trajetórias numa rede espacial. Essa nova abordagem é aplicada em dois casos: movimentos de animais e movimentos da população urbana em uma região. Eles mostram, a partir de experimentos, que os métodos de uso geral não são robustos para detectar comunidades espaciais e desenvolvem o método STOCS (*Spatial Tabu Optimization for Community Structure*), que transforma os dados de trajetória em uma rede espacial, integra algumas medidas de modularidade e particiona a rede em regiões geográficas para encontrar as comunidades espaciais a partir das trajetórias.

HE *et al.*, 2018 apresentam um novo método de detecção de comunidades para fluxos de origem-destino (OD). Esse método é direcionado para redes compostas por trajetórias de apenas uma aresta, uma linha OD (semelhante ao que observamos na rede de fluxo de pacientes, onde a trajetória de um paciente é composta por apenas uma aresta origem-destino). Eles apresentam o método SLCM (*Simple Line Clustering Method*), que adota entropia e funções de distribuição de probabilidade, para encontrar os agrupamentos mais relevantes em dados de OD. Apesar de utilizar dados de origem-destino assim como o fluxo de pacientes, o método proposto calcula um agrupamento de arestas (linhas OD), o que não corresponde a um agrupamento espacial que, no nosso caso, seria necessário para uma posterior comparação com a organização das regiões de saúde, como será explicado na Seção 5.2.1.

CHEN, 2021 desenvolveu uma ferramenta que identifica regiões funcionais na China (as Megaregiões) de acordo com o fluxo do tráfego de rodovias entre cidades. Este estudo aplica o método de detecção de comunidades, especificamente o algoritmo *infomap* (apresentado na Seção 3.1.1), para identificar os grupos de cidades no país que formam essas Megaregiões. No contexto apresentado, acadêmicos e departamentos governamentais discordam na forma de definir a delimitação espacial das Megaregiões. Com isso, essa ferramenta foi desenvolvida visando gerar uma solução baseada em evidências que apresente as principais regiões funcionais do país.

KOYLU *et al.*, 2018 adotaram um método de regionalização hierárquica baseada em fluxos (GUO, 2009) para gerar agrupamentos contíguos em províncias da Turquia com o objetivo de compará-los às regiões delimitadas pelo Ministério de Saúde turco. Semelhante ao nosso caso, eles utilizam os dados de fluxo de pacientes (partindo de sua residência ao hospital onde é atendido), analisando-os para monitorar os destinos dos pacientes em relação às regiões já determinadas pelo ministério e, assim, incentivar a atenção aos padrões de qualidade do sistema de saúde evitando disparidades na acessibilidade e qualidade de seus serviços. Esse trabalho é o que mais se assemelha ao que estamos desenvolvendo neste projeto em relação à análise comparativa entre os agrupamentos calculados e os

agrupamentos estabelecidos pelo ministério. Embora os autores façam uma análise geral dos fluxos de pacientes em cada região, eles não buscam estimar a real eficiência na organização das regiões, assim como nosso projeto propõe.

XAVIER *et al.*, 2019 analisam a regionalização da saúde com o objetivo de influenciar as decisões no planejamento regional. Os autores utilizaram dados de fluxo de pacientes entre os municípios brasileiros no ano de 2016 para a avaliação das regiões de saúde. Para isso, fizeram uso de um método de detecção de comunidades que otimiza uma medida de modularidade, semelhante à métrica explicada na Seção 3.1.1, mas considera o peso nas arestas da rede (NEWMAN, 2004a). Entretanto, ao contrário do nosso trabalho, o método escolhido não considera os aspectos geográficos da rede, o que resultou em comunidades esparsas e sem contiguidade, provocando uma dificuldade maior na comparação destas com as regiões de saúde.

## 4.2 Análise por índices de desempenho

Visando delimitar áreas de influência de mercados, Elzinga e Hogarty (ELZINGA e HOGARTY, 1973; ELZINGA e HOGARTY, 1978) desenvolveram um método para classificar regiões de mercados conforme o fluxo de compra de mercadorias. Baseado em modelos gravitacionais, eles criaram um índice que determina o desempenho de uma região através dos coeficientes LIFO e LOFI (vide Seção 3.2.1) para classificar o quão autossuficiente é essa região. Essa abordagem, assim como nosso trabalho, busca avaliar o desempenho de uma região a partir de dados de fluxos. Entretanto, ela não representa plenamente a dinâmica do fluxo de pacientes, já que não é a mesma dinâmica de competição de mercados (veremos isso nos resultados apresentados na Seção 6.3).

MORRISEY *et al.*, 1988 e FRECH III *et al.*, 2003 fizeram uso dos coeficientes definidos por Elzinga e Hogarty para estudar áreas de mercados hospitalares no contexto de saúde privada a partir dos dados de internações, montando fluxos de pacientes. Esse método considera a competência dos hospitais, através da escolha dos pacientes, para determinar seu mercado geográfico. Ele se assemelha à nossa análise de regiões de saúde, já que podemos considerar as regiões como conjuntos de hospitais. Mas, assim como será explicado na Seção 6.3, o nosso caso não é bem representado pela dinâmica de competição entre hospitais.

Entre outros pontos, RIGOLI *et al.*, 2019 analisam a suficiência de cada município de um departamento regional de saúde do Estado de São Paulo, além de também avaliar a regional como um todo. Eles fizeram uso dos coeficientes LIFO e LOFI (3.2.1) para estimar o nível de desempenho das áreas. Os dados utilizados para o estudo são bastante semelhantes aos dados utilizados por nós, exceto pela limitação do diagnóstico apenas para

internações relacionadas à gravidez. A análise realizada por eles foi baseada no método de Elzinga e Hogarty proposto para mercados geográficos, logo não reflete completamente o comportamento da movimentação dos pacientes, o que será detalhado melhor mais adiante.

Diferentemente dos casos acima, [UCHIMURA et al., 2017](#), [ARAUJO et al., 2016](#) e [A. L. d. VIANA et al., 2017](#) não adotaram uma análise por fluxo de pacientes. Eles fizeram uma análise das regiões de saúde no Brasil a partir da escolhas de indicadores que demonstrem o desempenho das regiões. O primeiro trabalho criou o ISR (Indicador Sintético Regional), que analisa diversas variáveis (taxa de mortalidade infantil, despesa per capita em saúde, médicos por mil habitantes, entre outras) para avaliar as regiões. O segundo trabalho criou um índice a partir de variáveis como cobertura de exames, taxa de mortalidade pelo agravo, proporção de tratamentos iniciados logo após confirmação do diagnóstico, entre outras, todas relacionadas a um agravo específico: o câncer do colo do útero. O último trabalho avalia o desempenho das regiões de saúde a partir da criação de três índices relacionados às dimensões Política, Estrutura e Organização, que foram determinados através de um estudo quantitativo com uso de dados de entrevistas realizadas com gestores, prestadores dos serviços de saúde e representantes da sociedade de algumas regiões de saúde do Brasil. Embora esses três trabalhos não recorram a técnicas de sistemas complexos, iteração espacial, ou análise de fluxos, eles nos apresentam outros meios sensíveis, utilizando outras bases de dados, para analisar a eficiência das regiões de saúde.

### 4.3 Considerações finais

Em nosso trabalho, reunimos os dois tipos de análise apresentadas nas duas últimas seções. Entre os trabalhos relacionados apresentados na Seção 4.1, o trabalho de [KOYLU et al., 2018](#) é o mais semelhante ao que estamos elaborando, em relação ao contexto; e o trabalho de [CHEN, 2021](#) é o mais semelhante em relação ao método, pois ambos utilizamos o *Infomap* (vide Seção 3.1.1), mas relacionado a contextos diferentes. Quanto aos trabalhos apresentados na Seção 4.2, nosso trabalho se baseia no método desenvolvido por [ELZINGA e HOGARTY, 1973](#), e desenvolve a análise inspirada no que foi feito por [RIGOLI et al., 2019](#). Os trabalhos de [UCHIMURA et al., 2017](#), [ARAUJO et al., 2016](#) e [A. L. d. VIANA et al., 2017](#) foram importantes para uma contextualização do que já foi feito de análise da regionalização da saúde, e quais dados foram utilizados para avaliação das regiões. Dessa forma, vemos que a análise através da movimentação dos pacientes poderia ser uma abordagem mais bem explorada nesse caso.

Em comparação aos trabalhos relacionados, este projeto apresentou inovações nos

métodos desenvolvidos. A primeira inovação se encontra na análise comparativa sugerida a partir da detecção de comunidades. A Seção 5.2.1 explica como esse método auxilia a compreensão do comportamento dos pacientes em relação às regiões de saúde de forma objetiva. Outra inovação foi a criação de um índice de desempenho (autossuficiência) para as regiões utilizando fluxos de oferta e demanda de serviços hospitalares (vide Seção 5.2.2). Por fim, diferente dos outros trabalhos relacionados, desenvolvemos uma plataforma web de software livre permitindo a análise da regionalização da saúde nos estados brasileiros através dos métodos desenvolvidos aqui (vide Seção 5.2.3). Além disso, disponibilizamos abertamente os códigos com os cálculos e com a estrutura da plataforma, como explica a Seção 5.3. O próximo capítulo irá detalhar todos esses métodos desenvolvidos neste trabalho.



# Capítulo 5

## Metodologia

Neste capítulo, introduzimos as bases de dados utilizadas e os métodos de análise de cada abordagem em resposta às questões de pesquisa. Por fim, apresentamos o plano de desenvolvimento da plataforma web do tipo *dashboard* que disponibiliza informações sobre o processo de regionalização do Sistema Único de Saúde (SUS) em suporte a gestores de saúde e pesquisadores. Para isso, o capítulo foi dividido em três seções: a Seção 5.1 apresenta as bases de dados utilizadas para permitir a aplicação dos métodos de análise escolhidos; a Seção 5.2 apresenta as abordagens utilizadas para analisar o processo de regionalização; e a Seção 5.3 descreve detalhes de implementação da plataforma, que foi criada para conter todas as visualizações geradas durante as análises, para dar apoio à elaboração de políticas públicas.

### 5.1 Bases de Dados

Para analisar as regiões de saúde através de técnicas de sistemas complexos, utilizamos dados de deslocamento intermunicipal dos pacientes. As informações de origem e destino do paciente internado podem ser obtidas através do Sistema de Informações Hospitalares do SUS (SIHSUS), a base de internações hospitalares. Por meio dela conseguimos representar a rede de deslocamento dos pacientes entre os municípios. Para visualização das análises e descrição estatística da base, utilizamos também a Malha Territorial ao nível de município e o Censo Demográfico de 2010 com as estimativas populacionais, fornecidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE), e os limites das regiões e macrorregiões de saúde, do Departamento de Monitoramento e Avaliação do SUS (DEMAS). Com isso, conseguimos apresentar as análises através de visualizações georreferenciadas com os limites municipais e regionais.

### 5.1.1 Internações Hospitalares

Os dados relacionados às internações hospitalares foram obtidos pelo SIHSUS. Esse é um dos sistemas desenvolvidos pelo Departamento de Informática do SUS (DATASUS), que tem o objetivo de promover soluções tecnológicas para melhorar o acesso e a organização de dados do SUS. Os dados fornecidos pelo SIHSUS são estruturados de forma a conter dados pessoais do paciente, da internação e do estabelecimento em que foi atendido. Entre esses dados estão: município de origem do paciente, faixa etária, nível de escolaridade, diagnóstico principal (conforme a Classificação Internacional de Doenças, a CID-10), nível de complexidade do procedimento, duração da internação, localização do estabelecimento, entre outras variáveis.

Nosso projeto analisa a rede de movimentação dos pacientes pelos municípios ao longo das últimas duas décadas. Para isso, utilizamos a soma de internações para cada origem e destino da base. Isto é, dentre as variáveis fornecidas pelo SIHSUS, utilizamos o município de residência do paciente e o município do estabelecimento que o atende, para cada ano da base. Isso é o suficiente para a criação da rede de movimentação de pacientes, permitindo a aplicação das técnicas de sistemas complexos (Seção 3.1) e de modelagem de fluxos de oferta e demanda (Seção 3.2) para analisar o processo de regionalização do SUS.

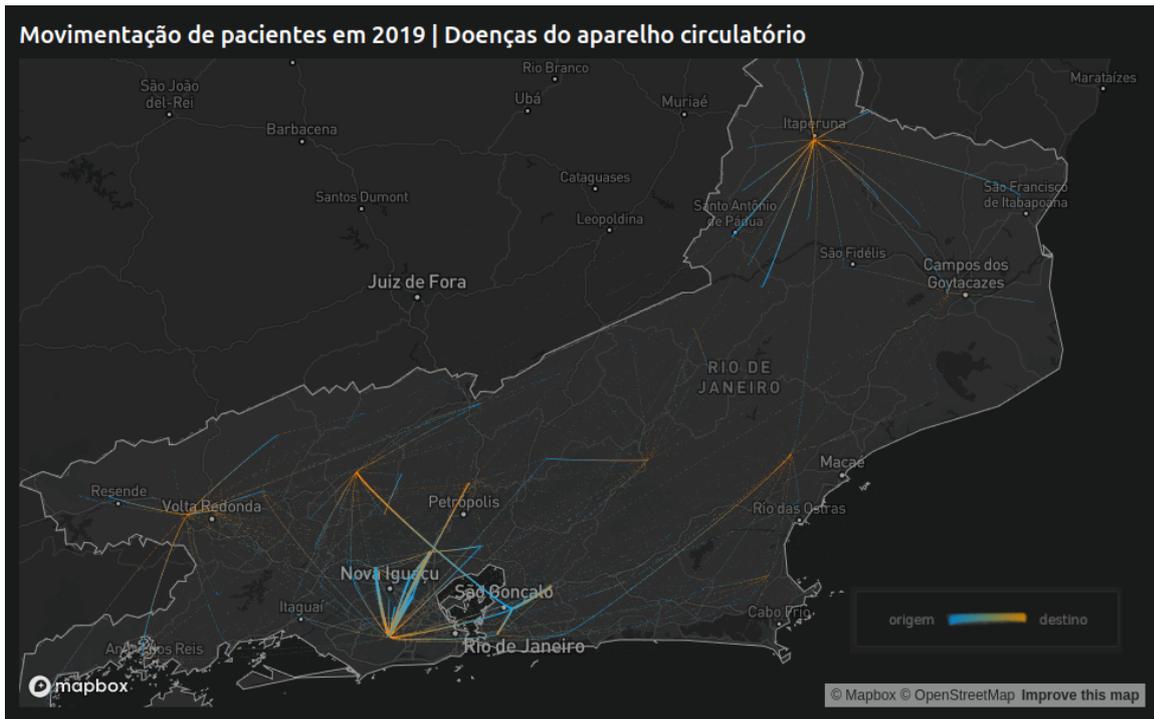
Além disso, podemos utilizar as mesmas ferramentas aplicadas neste trabalho para analisar redes mais específicas de movimentação de pacientes, filtrando internações por variáveis da base de dados. Se filtrarmos a base pela variável que representa o diagnóstico principal da internação, por exemplo, podemos observar o comportamento da rede para um agravo específico (como doenças do aparelho circulatório) e analisar todo o comportamento e eficiência das regiões de saúde para este caso específico. A Figura 5.1 apresenta dois diagnósticos distintos para o ano de 2019, mostrando como a mobilidade de pacientes difere para ambos. O mesmo processo pode ser realizado facilmente, em trabalhos futuros, com o uso de outras variáveis da base para analisar o comportamento da rede em outras perspectivas.

### 5.1.2 Malha Municipal e Censo Demográfico

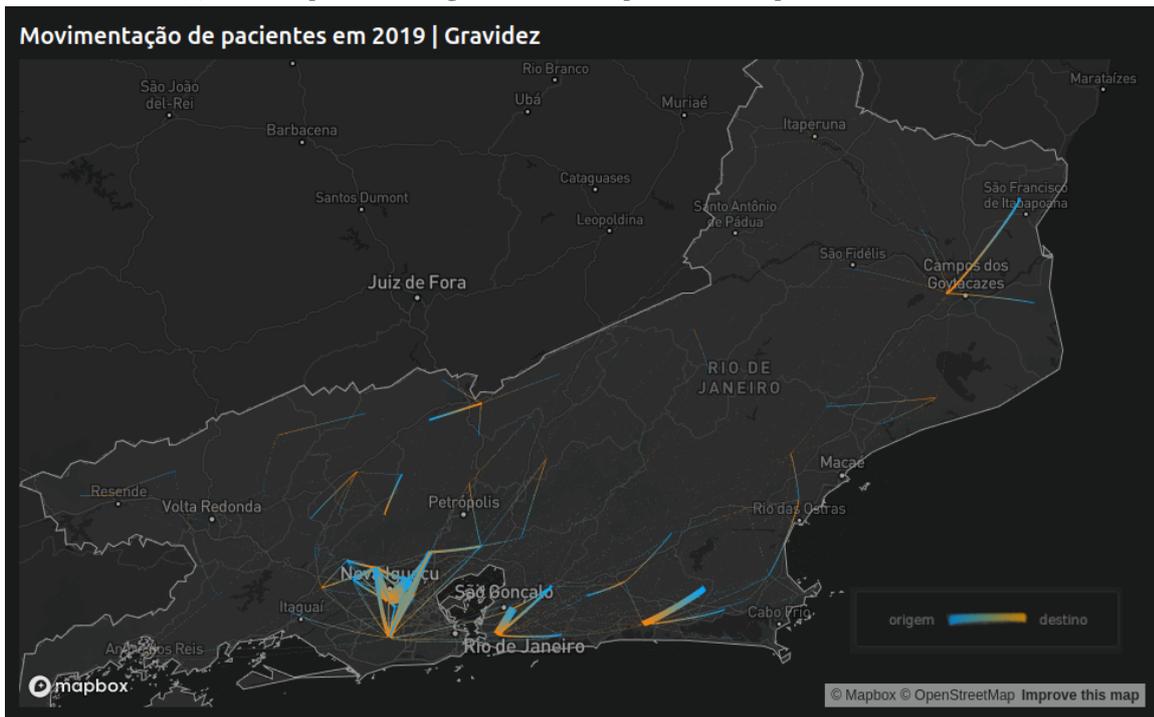
A Malha Municipal<sup>1</sup> disponibilizada anualmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE) contém a representação político-administrativa dos estados e municípios brasileiros. Para o projeto, utilizamos a atualização de 2021 com arquivos agrupados por estados. Esta base é composta pela representação vetorial de cada município, organizada em arquivos no formato *shapefile*, viabilizando a visualização georreferenciada das análises

---

<sup>1</sup><https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas>



(a) Fluxo de pacientes diagnosticados com problema no aparelho circulatório.



(b) Fluxo de pacientes com diagnósticos relacionados a gravidez.

**Figura 5.1:** Fluxo de pacientes no estado do Rio de Janeiro em 2019 para dois diagnósticos distintos. Imagem obtida da plataforma desenvolvida neste projeto.

deste estudo. Por exemplo, o limite estadual do Rio de Janeiro apresentado na Figura 5.1 foi construído com essa base. Utilizamos também as estimativas populacionais do

Censo Demográfico de 2010, realizado pelo IBGE, com a finalidade de realizar uma análise descritiva inicial da movimentação de pacientes. Essa análise será apresentada na Seção 6.1.

### 5.1.3 Limites das Regiões e Macrorregiões de Saúde

Como explicado na Seção 2.3, as regiões e macrorregiões de saúde são partições territoriais que agrupam municípios vizinhos. A base que contém essas informações para as regiões de saúde é estruturada pela relação município e região a qual pertence, para cada ano entre 2011 e 2017. Para as macrorregiões de saúde, a estrutura é a mesma, mas temos apenas a atualização de março de 2020. Essas informações foram obtidas pelo portal SAGE<sup>2</sup> (Sala de Apoio à Gestão Estratégica) de monitoramento do SUS, do Departamento de Monitoramento e Avaliação do SUS (DEMAS).

## 5.2 Abordagens de análise

Como as regiões e macrorregiões de saúde são organizadas por estado da federação, decidimos realizar as análises também por estado. Nelas, ao montar a rede de fluxo de pacientes, consideramos inicialmente os municípios como vértices e o fluxo de pacientes entre municípios como arestas desta rede. Cada aresta possui um peso relacionado à quantidade de pacientes que se movimentaram no sentido desta aresta. Lembrando que, como cada análise considera apenas os municípios daquele estado, o fluxo interestadual não é considerado neste nível da análise.

De qualquer forma, não existem impeditivos para uma futura análise ao nível nacional que considere os fluxos intermunicipais e interestaduais. Ela seria interessante para analisar a situação da saúde nas cidades de fronteira estadual, pois os habitantes de uma cidade com fronteira para outro estado podem tender a buscar atendimento hospitalar em locais mais próximos ou de maior disponibilidade, mesmo que precise atravessar um limite estadual.

Definiremos a seguir as abordagens aplicadas nesta pesquisa para cumprir com seu objetivo de analisar o processo de regionalização do SUS respondendo às três questões de pesquisa elaboradas neste trabalho (vide Seção 1.1). As Seções 5.2.1 e 5.2.2 abaixo utilizam estratégias recomendadas pela literatura. Vimos aplicações destas mesmas estratégias, ou semelhantes, no capítulo anterior em trabalhos relacionados apresentados nas Seções 4.1 e 4.2, respectivamente. Além disso, a Seção 5.2.3 apresenta uma terceira abordagem, mais visual, de análise do processo de regionalização por visualizações em mapa e gráficos.

---

<sup>2</sup><http://sage.saude.gov.br/>

### 5.2.1 Detecção de comunidades

A partir da rede de fluxo intermunicipal de pacientes, podemos aplicar a técnica de detecção de comunidades (vide Seção 3.1.1) para encontrar possíveis agrupamentos de municípios a partir dos dados. Como as regiões de saúde são, por definição, agrupamentos de municípios adjacentes pertencentes a um mesmo estado, podemos fazer uma análise de comparação dos agrupamentos sugeridos pela solução orientada pelos dados e os agrupamentos definidos na organização original das regiões de saúde. Essa abordagem responderia à questão de pesquisa Q1 (*A movimentação dos pacientes está relacionada à delimitação das regiões de saúde? O quanto?*), já que compara o comportamento da movimentação de pacientes, pelos dados, com a delimitação das regiões de saúde.

Testamos dois algoritmos diferentes para detecção de agrupamentos na rede: o *greedy modularity* (vide Seção 3.1.1) e o *infomap* (vide Seção 3.1.1). A aplicação deles foi elaborada com o auxílio dos pacotes *Python: NetworkX*<sup>3</sup> e *Infomap*<sup>4</sup>, ambos projetos de código aberto e em constante evolução. A partir deles já conseguimos identificar alguns resultados relevantes, que serão expostos no próximo capítulo na Seção 6.2.

Os dois algoritmos diferem bastante não só em questão de definição e implementação, mas também em sua aplicação. O *greedy modularity*, implementado pelo *NetworkX*, não considera o peso e nem a direção das arestas da rede, prejudicando a precisão de nossa análise, já que a direção do fluxo (partir do município A para B, diferente de partir de B para A) e a quantidade de pacientes atrelada a ele são informações de extrema relevância. Já o algoritmo *infomap* considera todos esses aspectos da rede e apresenta um resultado mais apropriado para o nosso caso. Outra diferença notada é a possibilidade de sugestão, por parte do usuário, de uma quantidade específica de agrupamentos a serem detectados. Ambos algoritmos calculam a quantidade e a organização dos agrupamentos, por padrão, mas apenas o *infomap* permite que sugiramos uma quantidade específica de agrupamentos a serem detectados. Isso auxilia no momento da comparação dos agrupamentos de municípios encontrados com os agrupamentos originais das regiões de saúde, pois podemos fixar a quantidade de agrupamentos ideal para a quantidade de regiões de saúde no estado.

Por fim, aplicamos a métrica de similaridade entre partições (vide Seção 3.3.1) justamente para obter uma métrica de semelhança entre os agrupamentos encontrados pelo algoritmo e os determinados pelas regiões de saúde, para cada ano da base. Como possuímos os dados da movimentação de pacientes para as últimas duas décadas, com essa métrica podemos analisar objetivamente a semelhança entre as duas partições ao longo dos anos, isto é, o quanto a movimentação dos pacientes tem respeitado os limites regionais.

---

<sup>3</sup><https://networkx.org/>

<sup>4</sup><https://mapequation.org/infomap/>

## 5.2.2 Índice de desempenho

Outra abordagem utilizada no projeto é o cálculo de um índice de desempenho que indique o nível de autossuficiência de uma região de saúde a partir do fluxo de pacientes. Com essa abordagem responderíamos à questão de pesquisa Q2 (*Existe uma evolução na autossuficiência das regiões de saúde ao longo dos anos?*) e, com o resultado desse índice e com a informação das alterações nas regiões de saúde ao longo dos anos, também responderíamos a Q3 (*As alterações das regiões de saúde ao longo dos anos interferiram na movimentação dos pacientes? Houve melhorias?*). Por fim, a intenção aqui é ter uma representação espacial que aponte regiões que estejam precisando de um reforço na infraestrutura de saúde, conforme o resultado do índice para as regiões. Para isso, testamos dois métodos diferentes. O primeiro é a aplicação dos coeficientes de Elzinga e Hogarty (LIFO e LOFI), e o segundo são as taxas de permanência e atração, desenvolvidas nesta pesquisa.

**Coeficientes LIFO e LOFI** Inicialmente, fixamos uma porcentagem mínima de 75% para o LIFO e o LOFI (definidos na Seção 3.2.1). Fizemos essa escolha baseando no que foi definido por Elzinga e Hogarty e por ser adotado em outros trabalhos relacionados (MORRISEY *et al.*, 1988; FRECH III *et al.*, 2003; RIGOLI *et al.*, 2019) que utilizaram essa configuração para estudar áreas de mercados hospitalares. Portanto, dessa forma, uma região de saúde seria considerada bem regionalizada somente se ambos coeficientes apresentassem valores acima de 0,75. Entretanto, essa maneira não se mostrou adequada como índice de desempenho para as regiões, pois o coeficiente LIFO desvalorizava regiões que atendiam muitos pacientes de fora. Dessa forma, as regiões que possuíam uma melhor infraestrutura não foram consideradas autossuficientes por este método. A Seção 6.3 apresentará esse resultado com mais detalhes.

**Taxas de permanência e atração** Com base no resultado que tivemos ao testar o método anterior, combinamos os coeficientes LIFO e LOFI modificados de forma a evitar os mesmos erros. Para isso, ao invés de utilizar os coeficientes originais, optamos em utilizar o oposto do coeficiente LIFO, i.e.  $1 - LIFO$ , chamando-o de “taxa de atração”, em combinação com o LOFI, a “taxa de permanência”, como apresenta a equação abaixo:

$$\text{Taxa de permanência} = LOFI;$$

$$\text{Taxa de atração} = 1 - LIFO.$$

Com esse novo método conseguimos obter coeficientes mais concretos e favoráveis a uma definição de desempenho (de autossuficiência) para as regiões: ela cresce conforme a

taxa de pacientes que permanecem na própria região de saúde para receber atendimento, indicando o nível de cobertura regional, e também cresce conforme a taxa de atração de pacientes para a região, indicando uma capacidade de atendimento superior à capacidade local, isto é, ou a região tem uma estrutura de alto nível, ou seus vizinhos têm uma estrutura de saúde mais precária.

Com isso, decidimos prosseguir a análise com este último método. Já que os valores das taxas variam entre 0 e 1, a combinação delas é feita por meio da média harmônica ponderada (vide Seção 3.3.2), atribuindo pesos diferentes para os coeficientes da seguinte maneira:

$$\bar{h}(w) = \frac{(w_1 + w_2)x_1x_2}{w_1x_2 + w_2x_1}, \text{ onde } x_1 = 1 - LIFO, x_2 = LOFI, w_1 = \beta \text{ e } w_2 = 1; 0 \leq \beta \leq 1. \text{ Logo:}$$

$$\bar{h}(w) = \frac{(\beta + 1)(1 - LIFO)LOFI}{(\beta)LOFI + (1 - LIFO)}.$$

Dessa forma, o índice agora valoriza regiões que atraem pacientes de fora, por utilizar o oposto do LIFO, a taxa de atração. Além disso, ele prioriza o atendimento regional, isto é, prioriza as regiões que atendem seus próprios pacientes pela taxa de permanência, já que o peso do LOFI na fórmula é geralmente maior ( $w_1 \leq w_2$ , pois  $\beta \leq 1$ ). Isso é favorável à definição de região e macrorregião de saúde, que devem garantir à população o atendimento integral, evitando evasões para outras regiões. Os resultados para esses métodos serão detalhados na Seção 6.3.

### 5.2.3 Visualização do processo de regionalização

A terceira abordagem promove uma análise visual e interativa que facilita o entendimento geral do processo de regionalização da saúde no Brasil, estado por estado. Consideramos importante a atenção a técnicas básicas de visualização de dados, como a disposição visual das informações e a escolha correta de gráficos para representar cada resultado, bem como a consideração do público ao qual as visualizações são direcionadas, especialmente gestores de saúde e pesquisadores, para que nada esteja fora do contexto e nem sobrecarregue seu entendimento. Um dos precursores em visualização de dados, especialmente dados em mapa, foi o cartógrafo Jacques Bertin, que classificou os elementos visuais na representação de dados em seu livro *Semiologia Gráfica* (BERTIN, 1983). Seu trabalho, em conjunto com outras fontes, em especial o livro *Storytelling com dados* de Cole Knafllic (KNAFLIC, 2019), contribuíram na construção das visualizações aqui desenvolvidas. Os códigos de todos os testes e implementações dos métodos estão disponíveis em um

repositório GitLab <sup>5</sup>, contendo também os scripts de processamento das bases que geram os resultados para cada estado da federação (na pasta “prepare/scripts/”).

Com isso, decidimos desenvolver quatro visualizações principais, compostas por mapas e gráficos, para apresentar os resultados encontrados. São elas: (i) a visualização da evolução da movimentação dos pacientes; (ii) a visualização da comparação entre a partição sugerida pelo algoritmo de detecção de comunidades e a partição real das regiões de saúde, como visto na Seção 5.2.1; (iii) a apresentação do índice de desempenho proposto por nós no final da Seção 5.2.2; e (iv) a visualização da evolução do comportamento dos limites das regiões de saúde. Apresentaremos o resultado das visualizações desenvolvidas na Seção 6.4 do próximo capítulo.

### 5.3 Plataforma Web

O objetivo final da nossa pesquisa de mestrado inclui o desenvolvimento de uma plataforma de software livre com a finalidade de fornecer as análises aqui realizadas para gestores, pesquisadores e público geral. Para isso, escolhemos a biblioteca *Python Dash*<sup>6</sup> para criação de *dashboards*, que permite o desenvolvimento de uma plataforma web para análise e visualização de dados. Na parte da estrutura interna da plataforma (o *backend*), as aplicações *Dash* utilizam o *framework Flask*<sup>7</sup> como servidor web, se comunicando por solicitações HTTP. Já na parte de interface web da plataforma (o *frontend*), ela renderiza componentes usando o *React.js*<sup>8</sup>, biblioteca de interface de usuário em *JavaScript*.

Assim como outras bibliotecas de criação de *dashboards*, a estrutura do *Dash* se assemelha à estrutura *Model-View-Controller* (MVC), descrita pela primeira vez por KRASNER e POPE, 1988. Neste caso, o *Model* seria um conjunto de *DataFrames Pandas* (biblioteca de manipulação e análise de dados)<sup>9</sup>, gerenciando o comportamento e estrutura dos dados dentro da plataforma. Os Componentes *Dash* representam o *View*, responsável por apresentar as informações ao usuário de forma visual. Por último, as funções de *callback* representam os *Controllers*, que têm a responsabilidade de controlar os dados enviados para visualização de acordo com a interação do usuário.

A Figura 5.2 representa a estrutura geral do *dashboard* desenvolvido. Nele, cada uma das três principais seções do dashboard (“Fluxos de pacientes”, “Divisões sugeridas”, e “Permanência vs. atração”); detalhadas na Seção 6.4) contém seu próprio *model*, *view* e *controller*,

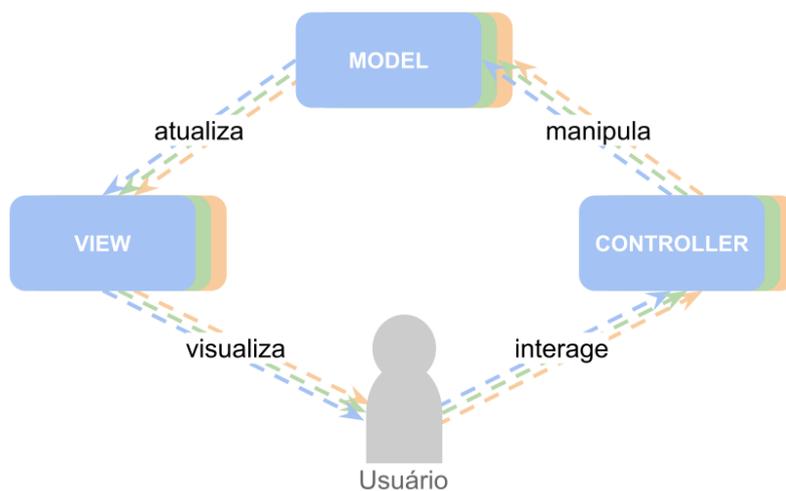
<sup>5</sup><https://gitlab.com/intercity/health/scripts-sus-regionalization>

<sup>6</sup><https://dash.plotly.com/>

<sup>7</sup><http://flask.pocoo.org/>

<sup>8</sup><https://pt-br.reactjs.org/>

<sup>9</sup><https://pandas.pydata.org/>



**Figura 5.2:** Representação visual da estrutura da plataforma web desenvolvida. Inspirada no modelo Model-View-Controller (MVC), a plataforma replica a estrutura para três seções da plataforma, representadas pelas três cores diferentes.

representados na figura em três diferentes cores. Ou seja, cada seção funciona independentemente das demais: utilizam *DataFrames* diferentes, possuem funções de *callback* diferentes, e são visualmente constituídas por componentes diferentes. Dessa forma, o código da plataforma (disponível em <https://gitlab.com/intercity/health/dashboard-sus-regionalization>) separa esses elementos na pasta “data/”, no arquivo “app.py”, e em “components/”, respectivamente. Assim, os elementos são organizados possibilitando a criação de outras seções na plataforma, estendendo-a para novas análises.

Em suma, a iniciativa em desenvolver uma plataforma web, disponibilizando essas análises em um *dashboard*, pode incentivar pesquisas sobre o processo de regionalização da saúde no SUS, além de promover tomadas de decisão por parte dos gestores de saúde pública. No capítulo seguinte, apresentamos os resultados que obtivemos a partir da implementação dos métodos aqui descritos.



# Capítulo 6

## Resultados

Agora que já descrevemos a metodologia utilizada na pesquisa, neste capítulo apresentaremos os resultados obtidos, lembrando que as abordagens são aplicadas ao nível estadual. Esses resultados estão separados pelo método utilizado, organizados na mesma ordem em que foram descritos na Seção 5.2. Entretanto, antes de apresentar os resultados, realizaremos uma descrição estatística da movimentação dos pacientes, na Seção 6.1, com o objetivo de resumi-los para, então, apresentarmos os resultados de nossas análises.

Cada uma das abordagens que veremos pretende responder, em conjunto, às questões de pesquisa deste projeto. A abordagem utilizada na Seção 6.2 busca, por técnicas de detecção de comunidades, solucionar a questão de pesquisa Q1 (vide Seção 1.1) que questiona o quanto a movimentação de pacientes está refletindo realmente a organização das regiões de saúde. A Seção 6.3 apresenta a abordagem que cria um índice de desempenho para as regiões de saúde. Ele visa medir o nível de autossuficiência de cada região, a fim de avaliar sua evolução ao longo do tempo. Essa abordagem visa responder a segunda questão de pesquisa, Q2. A Seção 6.4 apresenta a aplicação da visualização do processo de regionalização na plataforma desenvolvida. Essa abordagem cria visualizações de dados para fornecer uma ferramenta integrada e facilmente acessível aos gestores de saúde, que poderão tomar suas decisões baseadas em evidências tiradas dos mapas e gráficos da plataforma. A partir da combinação desta abordagem com a anterior, conseguimos, por fim, responder à questão de pesquisa Q3.

## 6.1 Descrição estatística da movimentação de pacientes

Através do processamento que fizemos sobre a base do SIHSUS (vide Seção 5.1.1) correspondente ao período de meados de 1994 até meados de 2020 (26 anos), geramos os dados de movimentação intermunicipal de pacientes. Esse processamento gera uma tabela com as movimentações entre cada par de municípios. Cada linha dessa tabela contém o município de origem, o município de internação (às vezes o mesmo de origem), a quantidade de pacientes seguindo esse fluxo, o diagnóstico principal da internação, e o ano correspondente. Para simplificar a análise, escolhemos alguns estados brasileiros para exemplificar o comportamento geral do fluxo de pacientes. Por questões de maior possibilidade de uma efetiva utilização dos resultados deste trabalho por estes estados, os estados escolhidos para análise são: Rio de Janeiro, São Paulo e Ceará.

	Rio de Janeiro	Ceará	São Paulo
Internações em 2009	714 621	503 675	2 420 494
Internações em 2019	754 725	531 283	2 606 482
Distância média percorrida em 2009	6,08 km	10,99 km	12,97 km
Distância média percorrida em 2019	8,61 km	25,68 km	11,91 km

**Tabela 6.1:** Descrição das internações e distância média da movimentação de pacientes em cada estado estudado.

Para iniciar a análise, a Tabela 6.1 descreve algumas características da tabela original com as movimentações intermunicipais de cada estado. Com isso, podemos já perceber a grandeza na quantidade de internações de cada estado. Percebe-se um leve crescimento na quantidade de internações ao longo dos anos. Já na distância média percorrida, exceto em São Paulo que esteve mais estável, os outros estados apresentaram um crescimento na distância percorrida pelos pacientes, especialmente no Ceará. Isso pode indicar uma maior cobertura do sistema rodoviário, causando um maior direcionamento das internações aos municípios mais distantes porém com mais recursos. A distância utilizada é a euclidiana entre os centroides dos municípios, logo os fluxos intramunicipais têm distância zero.

Estado	Área (km <sup>2</sup> )	Municípios	População aproximada dos municípios			
			total	média	mediana	desvio padrão
Rio de Janeiro	44 K	92	16 000 K	174 K	35 K	672 K
Ceará	149 K	184	8 000 K	46 K	20 K	183 K
São Paulo	248 K	645	41 000 K	64 K	13 K	454 K

**Tabela 6.2:** Distribuição da população entre os municípios para cada um dos estados estudados. Os dados populacionais foram retirados do Censo 2010.

Além disso, a Tabela 6.2 apresenta a distribuição da população entre os municípios para cada um dos estados estudados. Nela, podemos perceber a diferença entre eles em relação à área e à distribuição da população entre seus municípios. Em seguida, analisaremos cada um separadamente identificando aspectos interessantes em cada caso.

**Rio de Janeiro** Segundo o Censo de 2010 do IBGE, o estado do Rio de Janeiro conta com 16 milhões de habitantes. Com base nos dados obtidos pelo SIHSUS, o estado soma mais de 21 milhões de internações para os 26 anos de dados que tivemos acesso. Isso representa uma média de mais de 800 mil internações por ano, somando todos os 92 municípios do estado. A base de movimentação de pacientes, que processamos a partir do SIHSUS, é composta pelos fluxos intermunicipais de pacientes ao nível hospitalar no ano analisado. Cada fluxo equivale a uma aresta da rede com peso equivalente ao número de pacientes que seguem a mesma origem e destino. Para analisar a distribuição de pacientes de acordo com seus movimentos ao longo do estado, nós apresentamos a Tabela 6.3 que revela a concentração de pacientes em determinados fluxos do Rio de Janeiro. Nela, os quatro quartis dividem os fluxos de pacientes no estado, ordenados pela quantidade de pacientes por fluxo. Assim, os dois primeiros quartis contêm a metade dos fluxos que carregam menos pacientes, e os dois últimos contêm a metade dos fluxos que carregam mais pacientes.

Quartil	Pacientes por fluxo	# fluxos	% fluxos	% pacientes
Q4	238 777	1	0,02	25
Q3	13 234 a 238 777	7	0,17	25
Q2	2 740 a 13 234	33	0,78	25
Q1	1 a 2 740	4 165	99,03	25
Total	1 a 238 777	4 206	100	100

**Tabela 6.3:** Distribuição da média anual dos fluxos de pacientes no estado do Rio de Janeiro entre os quatro quartis do conjunto de fluxos no estado.

A Tabela 6.3 mostra que apenas 8 dos 4 206 fluxos já correspondem a metade dos pacientes no estado, indicado pelos quartis Q3 e Q4. Da mesma forma, 41 fluxos cobrem 75% da movimentação de pacientes, somando os quartis Q2, Q3 e Q4. Enquanto isso, mais de 99% dos diversos fluxos cobrem os outros 25% de pacientes, apresentado pelo Q1. Isso evidencia a assimetria na distribuição de pacientes entre os fluxos, além de apresentar o nível de relevância de alguns poucos fluxos.

A seguir, analisaremos o “nível de importância” dos fluxos no estado do Rio de Janeiro para todos os anos da base. Para isso, definimos uma **taxa de importância** para cada fluxo entre um *Município A* e um *Município B* para cada *ano* da seguinte forma:

$$taxa_{importancia}(mun_A, mun_B, ano) = I(mun_A, mun_B, ano) \times D(mun_A, mun_B), \quad (6.1)$$

onde  $I(mun_A, mun_B, ano)$  é a porcentagem das internações que partem do município A para o município B, dentre todas as internações que partem do município A (fixado), no ano determinado.  $D(mun_A, mun_B)$  é o peso da distância entre os municípios A e B em relação à soma da distância de todos os fluxos.

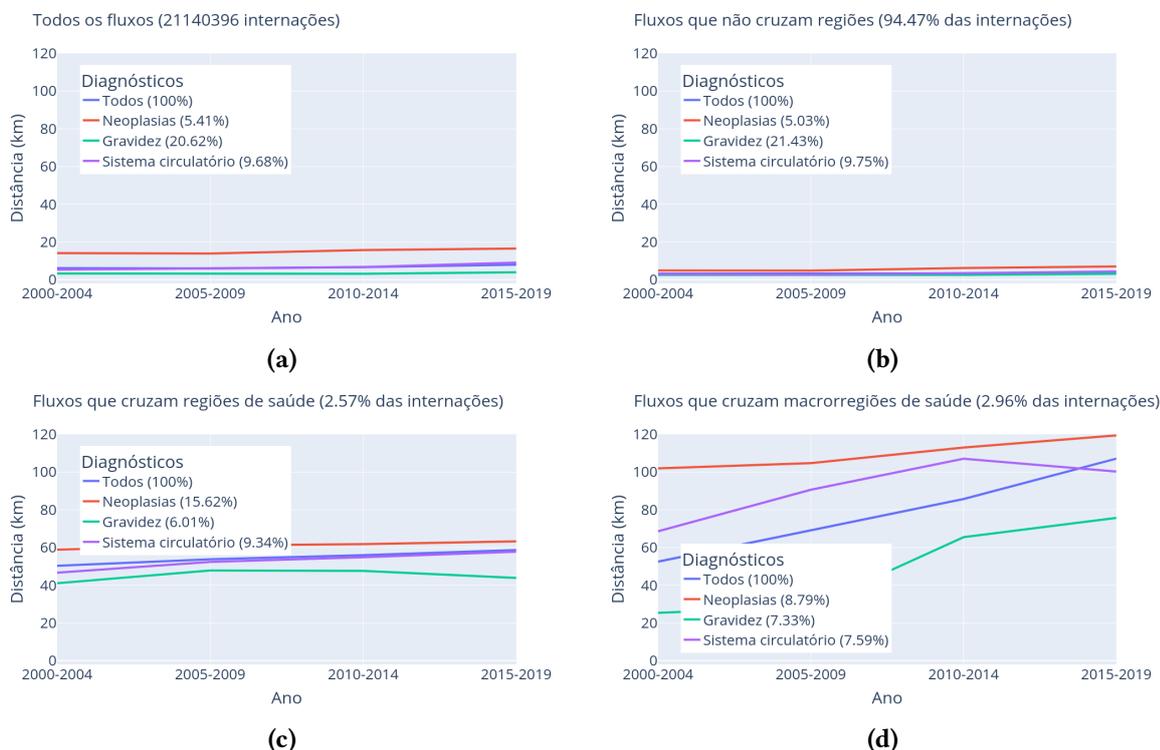
Dessa forma, um alto valor para a taxa de importância indica que grande parte das internações com origem no município A se deslocaram para o município B, que se apresenta relativamente distante. Logo, os fluxos de maior taxa devem ser aqueles de maior atenção, já que indicam que grande parte dos pacientes do município A percorrem longas distâncias para receber o atendimento hospitalar. A Tabela 6.4 apresenta os dez fluxos com maior taxa de importância entre 1994 e 2020. Nela, o foco está no *Município A*, cujos pacientes tendem a percorrer longas distâncias para o *Município B*, possivelmente por não ter acesso a um atendimento mais próximo adequado. Como a taxa de importância é calculada em relação ao município de residência do paciente, então pode haver casos em que um fluxo com baixa quantidade de internações seja considerado importante. Isso acontece pois a taxa não considera o valor absoluto de internações, mas o valor em relação à quantidade de pacientes no município de origem naquele ano (a porcentagem que aparece ao lado das internações na tabela).

	Ano	Município A	Município B	Internações	Distância	Taxa
#1	2007	Rio das Ostras	Rio de Janeiro	298 (32,15%)	171,4 km	0,000019
#2	2008	Rio das Ostras	Rio de Janeiro	267 (28,96%)	171,4 km	0,000017
#3	1994	Quissamã	Macaé	76 (74,51%)	61,2 km	0,000016
#4	1997	Comendador Levy G.	Rio de Janeiro	4 (57,14%)	74,5 km	0,000015
#5	2007	Armação dos Búzios	Rio de Janeiro	90 (24,52%)	167,2 km	0,000014
#6	2009	Armação dos Búzios	Rio de Janeiro	56 (24,03%)	167,2 km	0,000014
#7	2008	Armação dos Búzios	Rio de Janeiro	91 (23,70%)	167,2 km	0,000014
#8	2010	Armação dos Búzios	Rio de Janeiro	60 (23,53%)	167,2 km	0,000014
#9	2019	Cardoso Moreira	Itaperuna	591 (76,85%)	50,3 km	0,000014
#10	2020	Cardoso Moreira	Itaperuna	188 (76,42%)	50,3 km	0,000013

**Tabela 6.4:** Os dez fluxos de pacientes que apresentaram maior taxa de importância no Rio de Janeiro entre 1994 e 2020.

A Figura 6.1 contém a distância média por diagnóstico para 4 grupos diferentes de fluxos: todos eles; os fluxos que não cruzam os limites regionais; os que cruzam as regiões de saúde; e os que cruzam as macrorregiões. Quando examinamos a distância percorrida

para cada internação separando por diagnósticos, percebemos que todos apresentam uma grande estabilidade temporal na distância média geral. Os diagnósticos analisados aqui são: doenças relacionadas ao sistema circulatório, neoplasias, procedimentos relacionados a gravidez, e todos os diagnósticos juntos. Dessa forma, podemos perceber que as movimentações causadas por neoplasias, tumores em crescimento, costumam ter distância maior por necessitar de um atendimento mais especializado, por exemplo. Enquanto isso, internações relacionadas a gravidez, especialmente o parto, apresenta distâncias menores, embora alguns casos precisaram cruzar macrorregiões de saúde percorrendo quase 80 km entre 2015 e 2019.

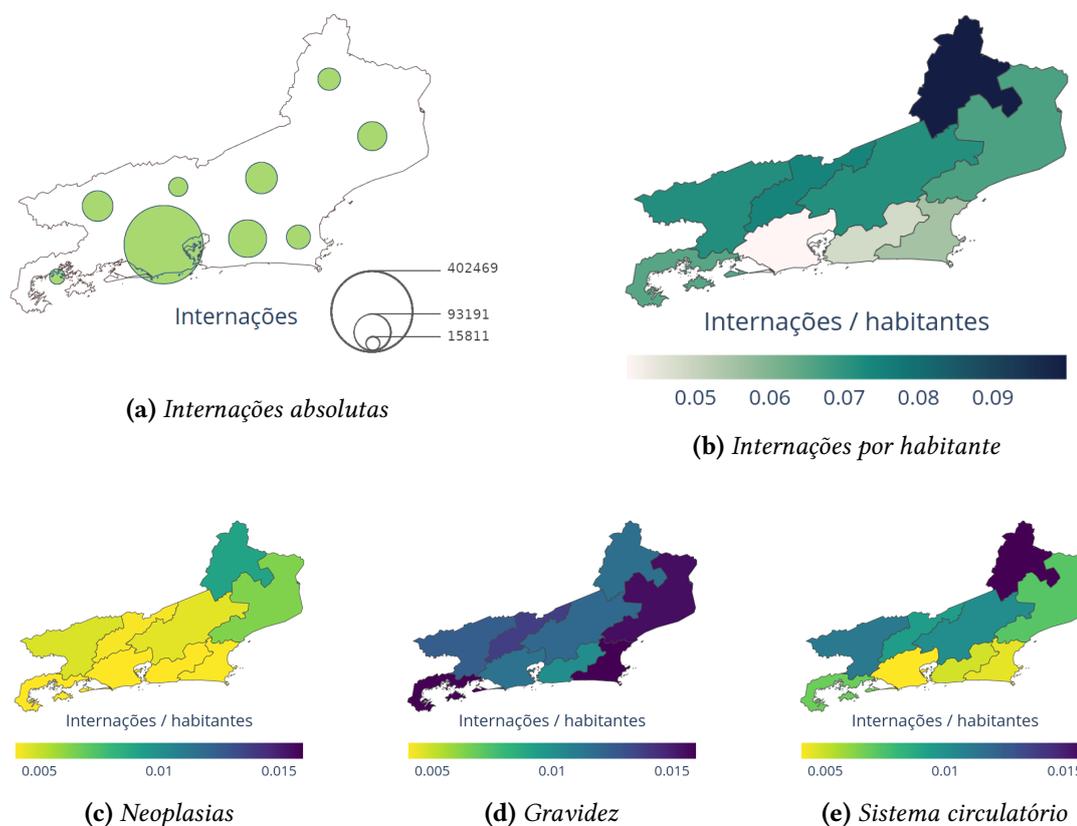


**Figura 6.1:** Média anual da distância percorrida por internação no Rio de Janeiro para os períodos de 2000 a 2004, 2005 a 2009, 2010 a 2014, e 2015 a 2019. Cada sub-gráfico considera tipos de fluxos diferentes: (a) todos os fluxos; (b) fluxos que não cruzam regiões; (c) fluxos que cruzam regiões; e (d) fluxos que cruzam macrorregiões. Isso para os diagnósticos relacionados à neoplasia, gravidez, sistema circulatório e todas as internações juntas.

Quando examinamos os diferentes tipos de fluxos, o Rio de Janeiro apresenta intervalos diferentes de distância para cada tipo. Para os fluxos que não cruzam limites regionais, a distância média não passa de 10 km. Para os que cruzam as regiões mas não as macrorregiões, a distância fica entre 40 km e 60 km, aproximadamente. Já para os fluxos que cruzam as macrorregiões, o intervalo de valores é maior, fica entre 20 km e 120 km. O crescimento temporal da distância média para todos os diagnósticos também é maior aqui entre as macrorregiões, cuja média passou de 50 km para 110 km no período total.

Outro fato que podemos observar a partir da Figura 6.1 é a porcentagem de internações para cada diagnóstico de acordo com os tipos de fluxo. Por exemplo, dentre os casos que não cruzaram regiões de saúde, 5,03% das internações estão relacionadas a neoplasias, sendo a de menor quantidade em relação aos outros diagnósticos. Entretanto, entre as internações que precisaram cruzar limites regionais ou macrorregionais, as por neoplasias são as principais em relação aos outros diagnósticos analisados.

Para analisarmos a distribuição das internações ao longo do estado, montamos a Figura 6.2 com a quantidade de internações pela Região de Saúde de origem do paciente. Se apresentarmos apenas a quantidade absoluta de internações (como na Figura 6.2(a)), a grande concentração de internações ficaria nas áreas de maior densidade populacional, já que quantidade de internações e habitantes são diretamente proporcionais. Para desvincular esses dois valores, utilizamos a quantidade de internações por habitante da região para analisar a situação geral de internações no estado, bem como a situação para alguns diagnósticos específicos.



**Figura 6.2:** Distribuição das internações de 2019 pela Região de Saúde de residência do paciente no estado do Rio de Janeiro. Os sub-gráficos (a) e (b) contém as internações para todos os diagnósticos. Já (c), (d) e (e) filtram as internações por diagnósticos específicos. As internações por habitante utilizam dados populacionais do Censo 2010.

Primeiramente, observamos que a região de saúde que contém o município do Rio de Janeiro aparece com o maior valor no primeiro mapa, mas contém o menor valor no segundo, o que indica que a região não gera tantas internações em relação à sua quantidade de habitantes. Os três mapas de baixo, na Figura 6.2, representam a demanda de internações para cada região de acordo com seu diagnóstico principal. Isso revela as principais necessidades de cada região em relação às vizinhas. Por exemplo, embora a região de saúde Noroeste, contenha o valor máximo na Figura 6.2(b), ela apresenta um valor mais baixo em relação às regiões vizinhas quando filtramos as internações para causas relacionadas a gravidez (Figura 6.2(d)), indicando uma menor demanda nesse caso.

**Ceará** Com pouco mais que a metade da extensão territorial de São Paulo, o Ceará possui aproximadamente 8 milhões de habitantes em seus 184 municípios, segundo o Censo 2010. Dentro dos 26 anos do período estudado, o Ceará realizou mais de 13 milhões de internações, segundo o SIHSUS, possuindo média anual de aproximadamente 500 mil internações. Podemos observar melhor a distribuição de pacientes pelos diferentes fluxos intermunicipais, através da Tabela 6.5.

Quartis	Pacientes por fluxo	# fluxos	% fluxos	% pacientes
Q4	150 125	1	0,02	25
Q3	3 241 a 150 125	18	0,34	25
Q2	686 a 3 241	98	1,86	25
Q1	1 a 686	5 166	97,78	25
Total	1 a 150 125	5 283	100	100

**Tabela 6.5:** Distribuição da média anual dos fluxos de pacientes no estado do Ceará entre os quatro quartis.

Assim como observado nos outros dois estados estudados, vemos que o Ceará também apresenta uma forte concentração de pacientes utilizando os mesmos fluxos. Enquanto 75% dos pacientes utilizam os mesmos 117 fluxos, isto é, 2,22% (unindo Q2, Q3 e Q4), os outros 25% utilizam 5 166 (97,78%). Esse resultado para o Ceará, assim como para os outros estados analisados, é decorrente do número limitado de municípios qualificados para atendimento hospitalar. Caso contrário, provavelmente encontraríamos uma distribuição mais equilibrada de pacientes entre os fluxos, já que haveriam mais destinos qualificados.

Como feito para o estado anterior, aplicamos aqui a mesma **taxa de importância** descrita na Equação 6.1. A Tabela 6.6 apresenta os dez fluxos com maior taxa de importância. Nela, observamos que a capital, Fortaleza, aparece em todos os dez principais fluxos como município de destino. Isso pode significar uma baixa distribuição de recursos médicos

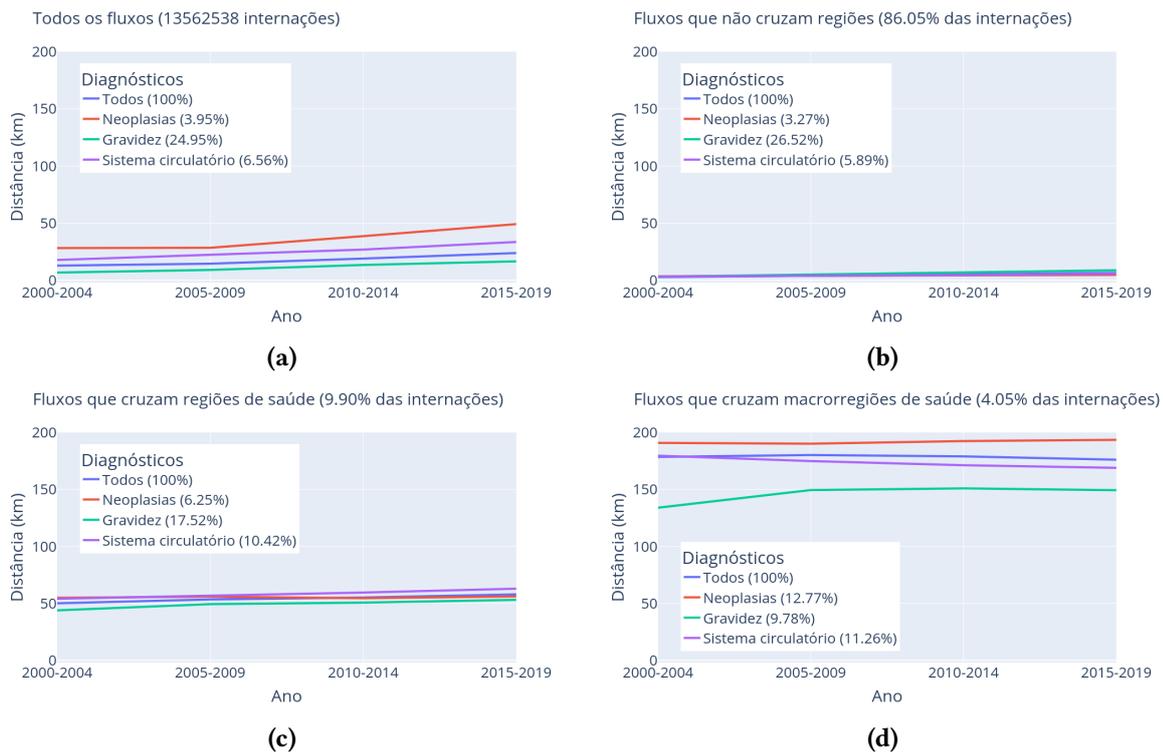
ao longo do estado, já que esses pacientes percorrem longas distâncias, chegando até a 262,5 km neste caso. Dessa forma, os municípios de Pereiro, Mombaça, Ereré, Milhã e Catarina, presentes na coluna *Município A* são os de maior atenção, conforme a taxa estabelecida.

	Ano	Município A	Município B	Internações	Distância	Taxa
#1	2000	Pereiro	Fortaleza	153 (82,70%)	196,7 km	0,000053
#2	2008	Mombaça	Fortaleza	361 (62,46%)	219,8 km	0,000045
#3	2015	Ereré	Fortaleza	85 (65,38%)	193,5 km	0,000041
#4	2018	Ereré	Fortaleza	110 (64,33%)	193,5 km	0,000041
#5	2016	Ereré	Fortaleza	91 (62,76%)	193,5 km	0,000040
#6	2019	Ereré	Fortaleza	86 (60,56%)	193,5 km	0,000038
#7	2016	Milhã	Fortaleza	278 (63,33%)	176,1 km	0,000036
#8	2017	Ereré	Fortaleza	130 (56,28%)	193,5 km	0,000035
#9	1999	Pereiro	Fortaleza	151 (54,32%)	196,7 km	0,000035
#10	2014	Catarina	Fortaleza	78 (40,62%)	262,5 km	0,000035

**Tabela 6.6:** Os dez fluxos de pacientes que apresentaram maior taxa de importância no Ceará entre 1994 e 2020.

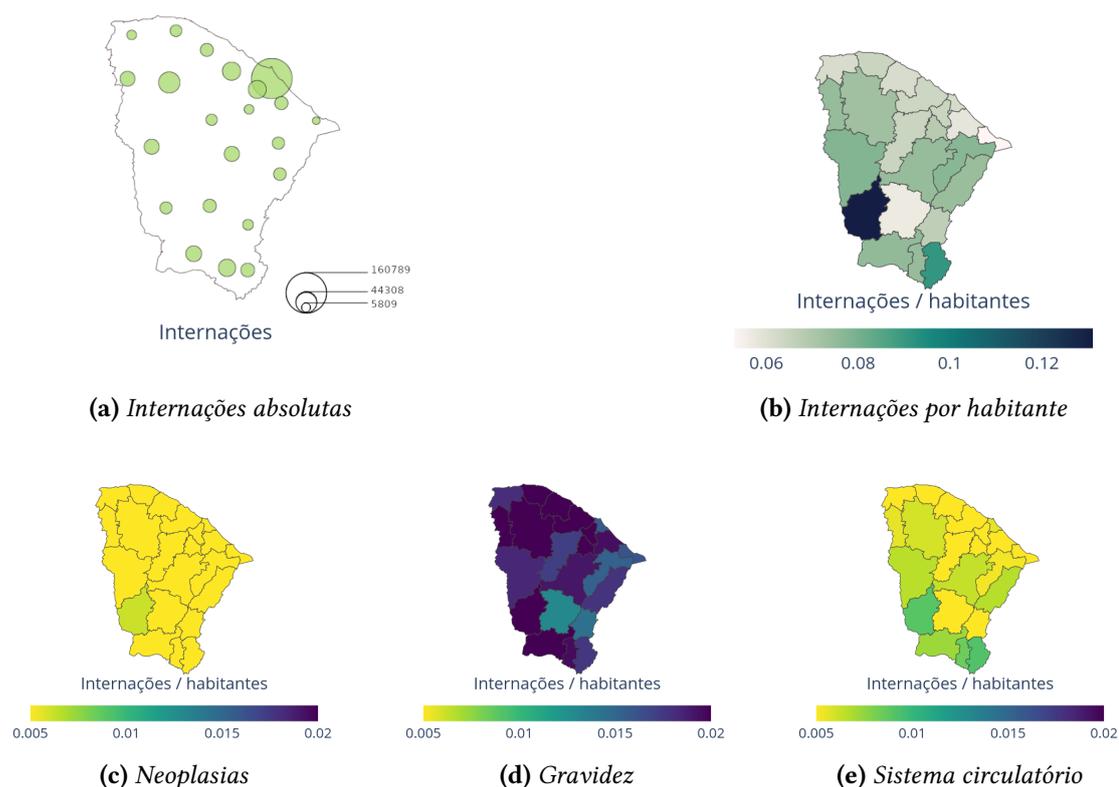
Analisando, agora, a distância média percorrida pelo paciente ao longo de duas décadas, o percurso médio cresce em distância ao longo do tempo. A Figura 6.3 contém essa informação separada por alguns principais diagnósticos, além de separar em fluxos que cruzam ou não regiões e macrorregiões de saúde. Em relação ao estado anterior estudado, os pacientes do Ceará percorrem uma distância maior em média, para todos os fluxos. Separando por tipo de fluxo, as movimentações que não cruzam regiões e as que cruzam apenas as regiões no Ceará apresentam comportamento muito semelhante ao Rio de Janeiro. No primeiro caso a distância média não passa de 10 km e no segundo caso fica entre 40 km e 60 km, assim como no outro estado. Já as movimentações que cruzam os limites macrorregionais se comportam num intervalo de valores superior ao valor máximo observado no estado anterior, ficando entre 120 km e 200 km. Isso pode decorrer do fato deste estado ser territorialmente mais amplo, mas também pode ser consequência da baixa distribuição de atendimento hospitalar no Ceará.

## 6.1 | DESCRIÇÃO ESTATÍSTICA DA MOVIMENTAÇÃO DE PACIENTES



**Figura 6.3:** Média anual da distância percorrida por internação no Ceará para os períodos de 2000 a 2004, 2005 a 2009, 2010 a 2014, e 2015 a 2019. Cada sub-gráfico considera tipos de fluxos diferentes: (a) todos os fluxos; (b) fluxos que não cruzam regiões; (c) fluxos que cruzam regiões; e (d) fluxos que cruzam macrorregiões. Isso para os diagnósticos relacionados à neoplasia, gravidez, sistema circulatório e todas as internações juntas.

Ao examinar a distribuição das internações ao longo das regiões de saúde do Ceará, através da Figura 6.4, vemos que a região que contém a capital, Fortaleza, ao nordeste do estado, se sobressai na quantidade absoluta de internações. Entretanto, quando desvinculamos essa informação da quantidade de habitantes da região, a região de Tauá, ao sudoeste do estado, passa a se sobressair em relação às demais, indicando uma maior demanda de internações. A partir dessa figura, podemos extrair outras informações em relação a alguns diagnósticos. Por exemplo, as regiões de Juazeiro do Norte e Brejo Santo, no extremo sul do estado, e Tauá, no sudoeste, apresentaram uma maior quantidade de internações por problemas relacionados ao sistema circulatório (Figura 6.4(e)), em relação às regiões vizinhas. Não conseguimos saber quais fatores influenciaram diretamente esse resultado, mas essa informação pode auxiliar os gestores de saúde na distribuição de recursos, para essas regiões, específicos para problemas cardiovasculares, como a ampliação da prestação de assistência especializada a portadores de doenças do sistema cardiovascular.



**Figura 6.4:** Distribuição das internações de 2019 pela Região de Saúde de residência do paciente no estado do Ceará. As internações por habitante utilizam dados populacionais do Censo 2010.

**São Paulo** O estado de São Paulo, de maior extensão territorial entre os estados estudados (vide Tabela 6.2), conta com 41 milhões de habitantes em seus 645 municípios (pelo Censo 2010 - IBGE). De acordo com o SIHSUS, ele realizou mais de 62 milhões de internações entre 1994 e 2020. Sendo assim, possui uma média de 2,3 milhões de internações anuais. Vejamos, na Tabela 6.7, a concentração de pacientes entre os fluxos intermunicipais do estado.

Quartis	Pacientes por fluxo	# fluxos	% fluxos	% pacientes
Q4	45 570 a 507 359	3	0,01	25
Q3	6 920 a 45 570	47	0,11	25
Q2	1 211 a 6 920	219	0,53	25
Q1	1 a 1 211	41 272	99,35	25
Total	1 a 507 359	41 541	100	100

**Tabela 6.7:** Distribuição da média anual dos fluxos de pacientes no estado de São Paulo entre os quatro quartis.

Na Tabela 6.7, apresentamos um resumo da concentração de pacientes entre os dife-

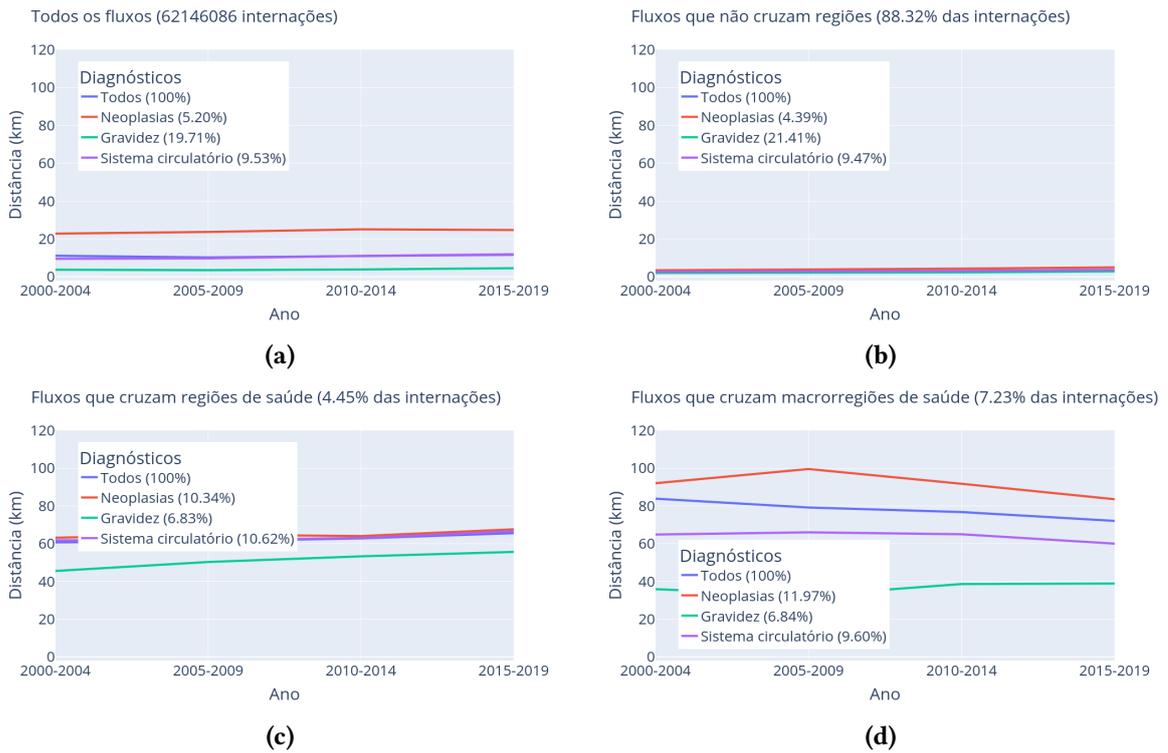
rentes fluxos intermunicipais do estado. Observamos uma forte concentração em alguns poucos fluxos: apenas 269 dos 41 541 fluxos, isto é, 0,65% deles estão relacionados a 75% das internações do estado (somando os quartis Q2, Q3 e Q4). Paralelamente, os outros 99,35% dos fluxos correspondem aos outros 25% de internações pelo estado.

Examinando, agora, a **taxa de importância** (vide Equação 6.1) de todos os fluxos no estado desde 1994 até 2020, a Tabela 6.8 contém aqueles que obtiveram maior valor. Como descrito anteriormente, essa taxa enfatiza os fluxos que representam grande parte das movimentações com origem no município A e se distancia bastante do município B, chegando até a 544,6 km no caso de Suzanópolis em 1995. Dessa forma, identificamos os principais casos em que grande parte das internações do município A precisou se deslocar uma distância relativamente longa para receber o devido atendimento hospitalar.

	Ano	Município A	Município B	Internações	Distância	Taxa
#1	1998	Embaúba	São Paulo	330	317,3 km	0,000006
#2	1995	Suzanópolis	São Paulo	1	544,6 km	0,000005
#3	1995	Ribeirão Grande	São Paulo	2	194,4 km	0,000004
#4	1998	Dirce Reis	São Paulo	37	501,4 km	0,000004
#5	1999	Américo de Campos	Araraquara	652	207,9 km	0,000004
#6	1999	Santa Albertina	São Paulo	239	526,6 km	0,000004
#7	1997	Barra do Chapéu	Sorocaba	14	196,1 km	0,000004
#8	1999	Embaúba	São Paulo	174	317,3 km	0,000004
#9	2002	Silveiras	Sorocaba	249	295,0 km	0,000004
#10	1995	São João de Iracema	S. J. Rio Preto	2	112,7 km	0,000003

**Tabela 6.8:** Os dez fluxos de pacientes que apresentaram maior taxa de importância em São Paulo entre 1994 e 2020.

Quando analisamos a distância média percorrida pelos pacientes separando por diagnósticos, encontramos certa diferença entre atendimentos mais especializados (como casos relacionados a neoplasias) e atendimentos mais urgentes (como parto, por exemplo). De forma que os recursos para atendimentos especializados não são tão distribuídos pelo estado, exigindo um percurso mais longo ao paciente, enquanto as internações mais urgentes exigem um menor deslocamento. Isso pode ser observado no primeiro sub-gráfico da Figura 6.5. A mesma Figura 6.5, em seus sub-gráficos, expõe o quanto a distância média varia para os diferentes tipos de fluxo no estado, semelhante ao comportamento observado nos outros dois estados estudados. Mesmo sendo um estado de dimensão territorial maior que os outros, São Paulo apresenta distâncias médias semelhantes aos outros dois estados, sendo até mais acessível que o Ceará, de área 100 000 km<sup>2</sup> menor (vide Tabela 6.2), para os casos de movimentações que cruzam limites macrorregionais, não passando de 100 km para os casos analisados no período entre 2000 e 2019 (Figura 6.5(d)).

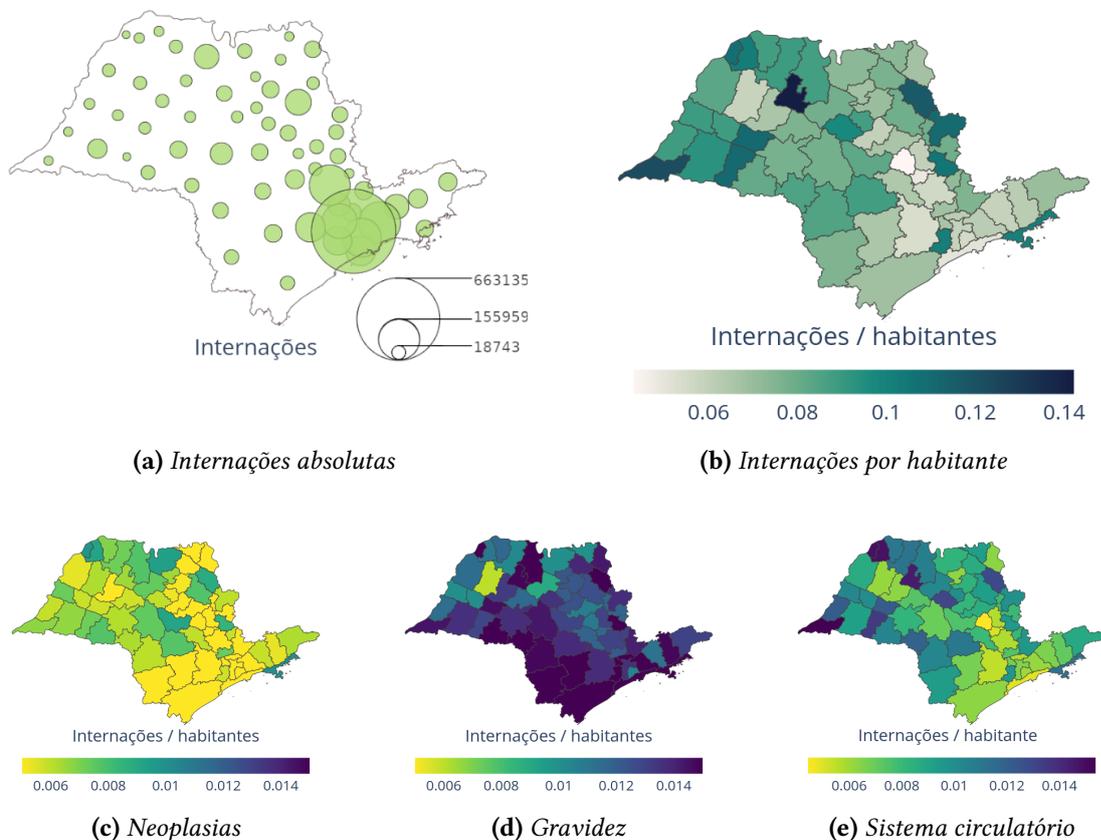


**Figura 6.5:** Média anual da distância percorrida por internação em São Paulo para os períodos de 2000 a 2004, 2005 a 2009, 2010 a 2014, e 2015 a 2019. Cada sub-gráfico considera tipos de fluxos diferentes: (a) todos os fluxos; (b) fluxos que não cruzam regiões; (c) fluxos que cruzam regiões; e (d) fluxos que cruzam macrorregiões. Isso para os diagnósticos relacionados à neoplasia, gravidez, sistema circulatório e todas as internações juntas.

Por fim, a Figura 6.6 apresenta a distribuição das internações pelas regiões de saúde de residência dos pacientes, revelando a demanda de atendimento de cada uma em relação às regiões vizinhas. Assim como feito nos outros estados estudados, desvinculamos a quantidade de internações da densidade populacional de cada região para analisar internações por habitante. Dessa forma, a região que contém a capital na Figura 6.6(a), que se sobressai em relação aos outros municípios pelo seu alto valor, agora possui um valor menor do que a média na Figura 6.6(b), após a divisão pela quantidade de habitantes.

Agora, analisando as internações por habitante para alguns diagnósticos, podemos encontrar regiões onde o resultado é maior em relação à vizinhança. Por exemplo, as regiões de Jales e Santa Fé do Sul em roxo ao noroeste do estado na Figura 6.6(e) apresentaram valores mais altos para diagnósticos relacionados ao sistema circulatório, embora não possuam altos valores de internações por habitante no geral (Figura 6.6(b)). Isso aponta que, aproximadamente, 15 a cada mil habitantes de cada uma dessas regiões precisaram de atendimento hospitalar especializado em sistema circulatório em 2019. Valor bem acima da média do estado: 6,72 internações por doenças do aparelho circulatório em 2019 a cada mil habitantes, conforme o Sistema de Informações Hospitalares do SUS (SIHSUS) e o Censo

Demográfico de 2010.



**Figura 6.6:** Distribuição das internações de 2019 pela Região de Saúde de residência do paciente no estado de São Paulo. As internações por habitante utilizam dados populacionais do Censo 2010.

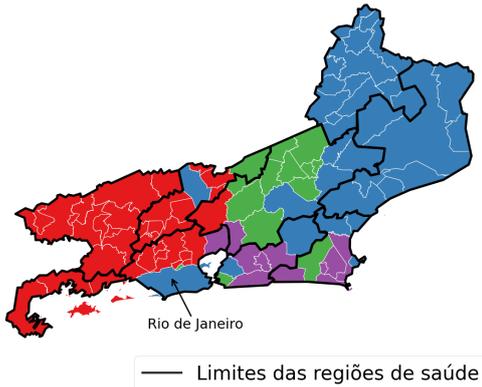
## 6.2 Aplicação da detecção de comunidades

A técnica de detecção de comunidades é utilizada nesta pesquisa visando promover uma comparação entre a configuração atual das regiões de saúde e as comunidades detectadas pelos algoritmos baseados nos fluxos de pacientes (vide Seção 5.2.1). A seguir, apresentaremos o resultado dessa comparação, para os três estados estudados, utilizando dois algoritmos: *greedy modularity* e *infomap*. Para cada estado, foram removidas todas as internações que atravessavam os limites estaduais, de modo a estudarmos apenas a rede intraestadual de internações, já que as regiões de saúde são determinadas ao nível estadual. Dessa forma, removemos da base original apenas 0,52% das internações interestaduais atendidas no Rio de Janeiro, bem como 0,55% removidas do Ceará, e 1,03% de São Paulo.

Nas seguintes Figuras (6.7, 6.8 e 6.9), comparamos o resultado dos algoritmos com as Regiões de Saúde ((a) e (b) de cada figura) e Macrorregiões de Saúde ((c) e (d)). Pri-

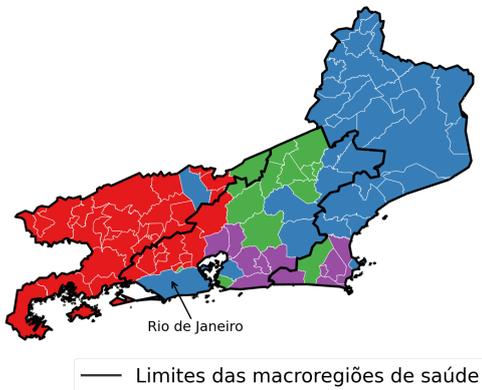
meiramente, observamos que o *infomap* obteve um melhor resultado no nosso contexto em todos os estados. Assim como esperávamos, percebemos visualmente a similaridade entre a divisão das regiões de saúde (em preto) e a divisão das comunidades detectadas (preenchida com cores aleatórias).

Comunidades detectadas pelo **greedy modularity** para os dados de 2019  
(4 comunidades)



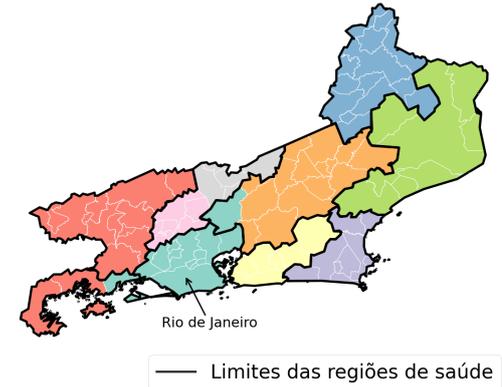
(a)

Comunidades detectadas pelo **greedy modularity** para os dados de 2019  
(4 comunidades)



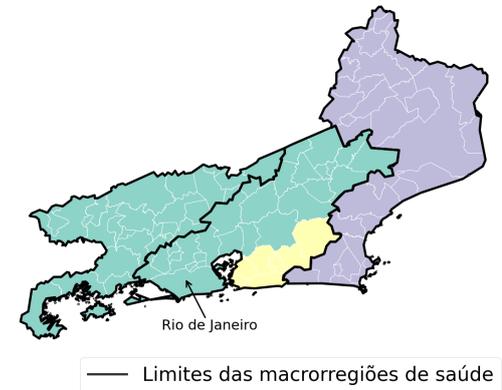
(c)

Comunidades detectadas pelo **infomap** para os dados de 2019  
(9 comunidades)



(b)

Comunidades detectadas pelo **infomap** para os dados de 2019  
(3 comunidades)

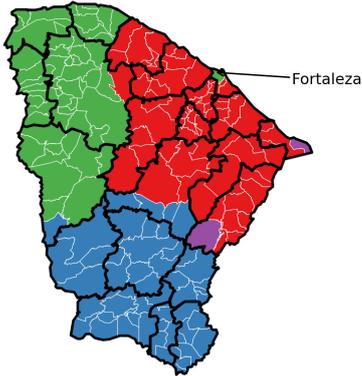


(d)

**Figura 6.7:** Resultado da aplicação dos algoritmos de detecção de comunidades *greedy modularity* ((a) e (c)) e *infomap* ((b) e (d)), primeiro para as Regiões e depois para as Macrorregiões de Saúde. Baseado nos fluxos intermunicipais de pacientes em 2019 no estado do Rio de Janeiro.

Como visto anteriormente (na Seção 5.2.1), estimamos que o resultado observado ocorra pelo fato de que o *greedy modularity* não considera o peso e a direção da aresta, além de não permitir que configuremos a quantidade de agrupamentos desejados. No caso do *infomap*, conseguimos configurar o algoritmo para detectar uma quantidade de comunidades específica, igual à quantidade de regiões ou macrorregiões de saúde do estado, de modo a facilitar a comparação entre a divisão atual dessas regiões com as comunidades detectadas pelo algoritmo (as quantidades de regiões por estado, em 2019, estão na tabela

Comunidades detectadas pelo **greedy modularity** para os dados de 2019  
(4 comunidades)



— Limites das regiões de saúde

(a)

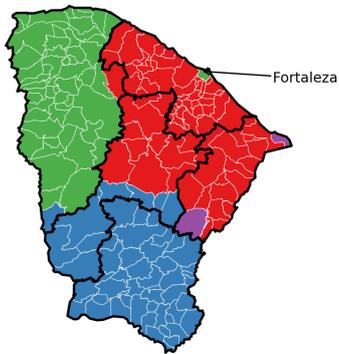
Comunidades detectadas pelo **infomap**  
para os dados de 2019  
(22 comunidades)



— Limites das regiões de saúde

(b)

Comunidades detectadas pelo **greedy modularity** para os dados de 2019  
(4 comunidades)



— Limites das macrorregiões de saúde

(c)

Comunidades detectadas pelo **infomap**  
para os dados de 2019  
(5 comunidades)

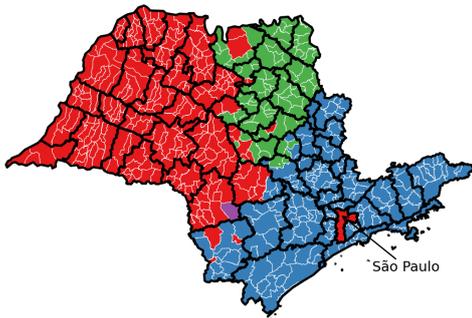


— Limites das macrorregiões de saúde

(d)

**Figura 6.8:** Resultado da aplicação dos algoritmos de detecção de comunidades greedy modularity ((a) e (c)) e infomap ((b) e (d)), primeiro para as Regiões e depois para as Macrorregiões de Saúde. Baseado nos fluxos intermunicipais de pacientes em 2019 no estado do Ceará.

Comunidades detectadas pelo **greedy modularity** para os dados de 2019  
(4 comunidades)



— Limites das regiões de saúde

(a)

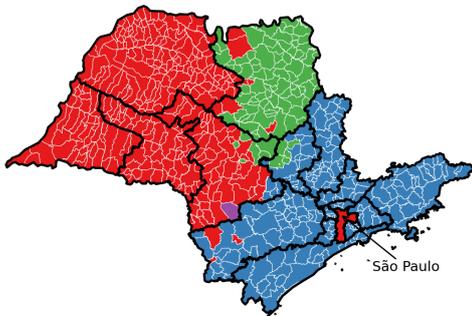
Comunidades detectadas pelo **infomap** para os dados de 2019  
(63 comunidades)



— Limites das regiões de saúde

(b)

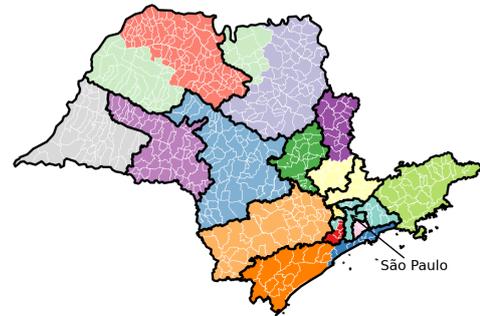
Comunidades detectadas pelo **greedy modularity** para os dados de 2019  
(4 comunidades)



— Limites das macrorregiões de saúde

(c)

Comunidades detectadas pelo **infomap** para os dados de 2019  
(17 comunidades)



— Limites das macrorregiões de saúde

(d)

**Figura 6.9:** Resultado da aplicação dos algoritmos de detecção de comunidades *greedy modularity* ((a) e (c)) e *infomap* ((b) e (d)), primeiro para as Regiões e depois para as Macrorregiões de Saúde. Baseado nos fluxos intermunicipais de pacientes em 2019 no estado de São Paulo.

6.9). Além disso, o *infomap* considera o peso bem como a direção da aresta. Logo, ele utiliza a base de movimentação de pacientes com mais detalhes, de forma mais fiel à realidade, para gerar seu resultado do que faz o *greedy modularity*.

	Regiões de Saúde	Macrorregiões de Saúde	Municípios
Rio de Janeiro	9	3	92
Ceará	22	5	184
São Paulo	63	17	645

**Tabela 6.9:** Quantidade de Municípios, Regiões e Macrorregiões de Saúde do Rio de Janeiro, Ceará e São Paulo em 2019.

Como visto na Seção 3.1.1, o algoritmo *greedy modularity* encontra agrupamentos de municípios de forma a minimizar a quantidade de arestas que cruzam as comunidades.

Mesmo não produzindo um bom resultado para a comparação com as regiões de saúde, podemos tirar algumas informações interessantes dele. Na primeira coluna das Figuras 6.7, 6.8 e 6.9, observamos dois pontos em comum entre os três estados estudados. Primeiro, a quantidade de comunidades é geralmente menor que o número de regiões. Isso pode ser efeito da grande densidade de fluxos, além de uma grande quantidade de arestas que unem municípios distantes. Em segundo lugar, as capitais dos três estados fazem parte de uma comunidade mais distante e não da comunidade mais próxima, diferente do que acontece com a maioria dos outros municípios. Isso pode ser explicado pela preferência geral de se dirigir à capital. Entretanto, esse resultado não representa fielmente o comportamento real da rede, já que, como foi dito, essa implementação do algoritmo não considera o peso nas arestas. Isto é, para ele um fluxo que leva 10 pacientes tem a mesma relevância que um fluxo que leva 10 mil, por exemplo.

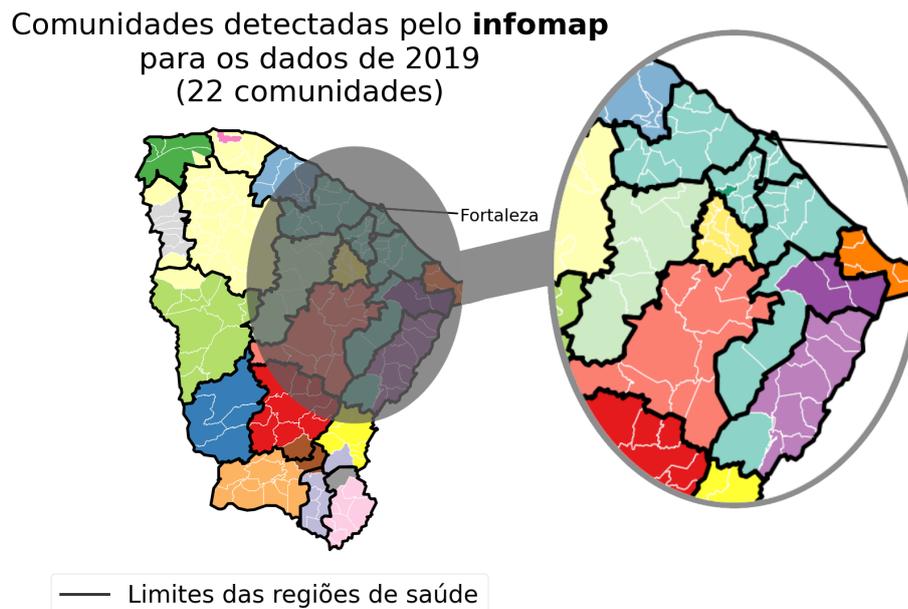
Já o *infomap*, por conta de sua configuração mais fiel à rede de movimentação real, nos apresenta resultados mais relevantes. Perceptivelmente, a segunda coluna das Figuras 6.7, 6.8 e 6.9 nos apresentam certa semelhança entre as comunidades detectadas, em colorido, e as divisões das regiões e macrorregiões de saúde, limitadas em preto. Inclusive, encontramos alguns casos para cada estado onde a divisão é a mesma: como visto em verde e azul ao nordeste do Rio de Janeiro em comparação com as regiões de saúde (6.7(b)) e em cinza ao nordeste em comparação com as macrorregiões de saúde (6.7(d)); em coral e verde água ao centro do Ceará, para as regiões de saúde (6.8(b)) e, para as macrorregiões, em cinza ao sul do estado (6.8(d)); e em roxo, azul e marrom no centro de São Paulo em comparação com as regiões (6.9(b)) e em diversas cores ao oeste e sudoeste do estado, para as macrorregiões (6.9(d)).

Esse resultado nos apresenta que, realmente, o comportamento da movimentação dos pacientes está relacionado à divisão das regiões de saúde, segundo os dados de 2019, o último ano completo da base. Como constataremos na Seção 6.2.1, esse comportamento é semelhante em outros anos da base. Isso já responde positivamente a primeira parte da questão de pesquisa Q1, que questiona se existe relação entre a movimentação de pacientes e as delimitações das regiões de saúde. Na Seção 6.2.1 teremos a resposta da segunda parte dessa pergunta, que questiona *o quanto* a movimentação está relacionada às delimitações de regiões.

Sobre a utilização de cores nos mapas que representam resultados do *infomap* na Figura 6.9, especialmente no mapa 6.9(b), teríamos que usar 63 cores diferentes. Entretanto, como essa quantidade de cores pode gerar tons muito semelhantes ao nosso olho, optamos por repetir algumas cores acrescentando alguma hachura. Logo, as áreas coloridas e hachuradas representam, cada uma, uma comunidade diferente, assim como as que não

possuem hachura.

A partir dessas informações, o gestor de saúde poderá identificar onde a organização das regiões de saúde pode estar funcionando bem, a partir das internações hospitalares, além de identificar pontos onde possivelmente há problemas de implementação da regionalização. A Figura 6.10 apresenta uma região do Ceará onde, possivelmente, a regionalização não esteja funcionando bem. Por exemplo, a comunidade detectada em azul esverdeado (na área ressaltada pela figura) revela uma forte relação entre os municípios que ela contém. Entretanto, essa comunidade passa por 6 regiões de saúde diferentes, o que pode indicar um problema na regionalização neste ponto. Por outro lado, esse resultado pode ser uma sugestão de unificação ou expansão da região de saúde para algo mais semelhante ao atual comportamento da movimentação de pacientes, representado pela comunidade detectada, se isso também fizer sentido para os outros aspectos de determinação das regiões, como características sociais, de transporte e infraestrutura da área em questão.



**Figura 6.10:** Identificação de um possível problema de regionalização no estado do Ceará (ou sugestão de expansão). Para esse resultado, foram utilizados dados de movimentação de pacientes de 2019 no algoritmo de detecção de comunidades infomap.

Nessa abordagem de detecção de comunidades, nos resta agora medir numericamente o nível de semelhança entre as duas partições do estado, as regiões de saúde e as comunidades detectadas. Isso é feito de modo a possibilitar a análise de sua evolução temporal desde meados da década de 1990, e, dessa forma, permitir a avaliação do quanto a regionalização do SUS está sendo implementada efetivamente nos diferentes estados brasileiros e, também, nas sub-regiões internas de cada estado.

### 6.2.1 Semelhança entre partições

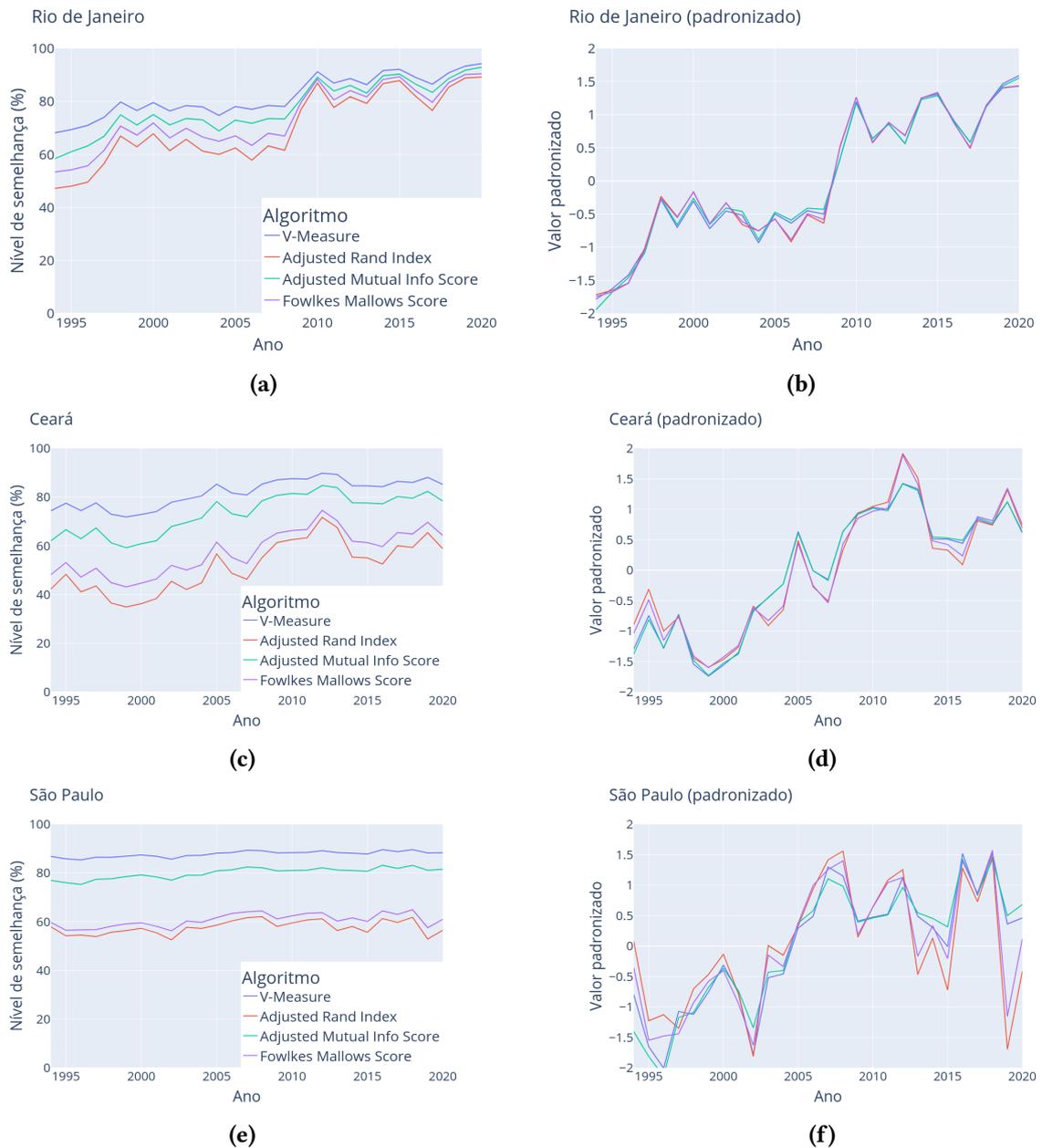
Para determinar o quanto as delimitações das regiões de saúde estão relacionadas com a movimentação de pacientes e, assim, avaliar o processo de regionalização da saúde nos estados brasileiros, escolhemos um algoritmo de comparação de partições (vide Seção 3.3.1) para comparar as regiões com as comunidades detectadas pelo *infomap*. Como existem diversos métodos para comparar partições, testamos algumas, implementadas pela biblioteca Python *scikit-learn*<sup>1</sup>: *V-measure*, *Adjusted Rand Index*, *Adjusted Mutual Info* e *Fowlkes-Mallows Score*. Ao testá-los para as Regiões de Saúde dos três estados estudados, percebemos que os algoritmos têm um comportamento bastante semelhante entre si. Podemos observar na primeira coluna da Figura 6.11 que os aclives e declives das quatro linhas nos três estados são bem semelhantes. Para comprovar essa observação, efetuamos a padronização dos valores de cada algoritmo, subtraindo a média e dividindo pelo desvio padrão de cada um, para colocá-los numa mesma escala. O resultado dessa padronização pode ser observado na segunda coluna da Figura 6.11, que comprova a semelhança entre os quatro métodos de comparação de partições nestes casos, já que as quatro linhas quase que se sobrepõem. Observamos também o mesmo comportamento para as Macrorregiões de Saúde.

Além disso, constatamos que os algoritmos *Adjusted Rand Index* e *Fowlkes Mallows Score* seguem a mesma tendência mas intensificam os extremos (por exemplo na Figura 6.11(f)). Como percebemos uma grande semelhança entre os algoritmos testados, acreditamos que não haja fortes motivos para a escolha de um algoritmo específico. Logo, podemos escolher qualquer um desses algoritmos para prosseguir com a análise. Com isso, escolhemos o algoritmo *Adjusted Rand Index* simplesmente por apresentar uma maior amplitude de valores, o que pode facilitar a detecção de variações no nível de semelhança entre as partições.

Na Figura 6.12, ao analisarmos a evolução do nível de semelhança das partições para as Regiões de Saúde (6.12(a)), podemos observar um certo crescimento ao longo dos anos, com exceção de São Paulo que se mantém mais constante. Os valores mais altos no gráfico refletem, então, os anos em que a movimentação dos pacientes esteve em maior harmonia com as delimitações das regiões de saúde. Já que essas linhas representam o nível de semelhança entre a delimitação das regiões de saúde do estado e a delimitação sugerida pelo algoritmo de detecção de comunidades, que tem como base os dados de movimentação de pacientes pelo estado. Por exemplo, a movimentação de pacientes no estado do Rio de Janeiro em 2010 esteve mais de acordo com a delimitação das regiões de saúde do que em 2005, como apresenta a Figura 6.12(a). Já no caso das macrorregiões de saúde,

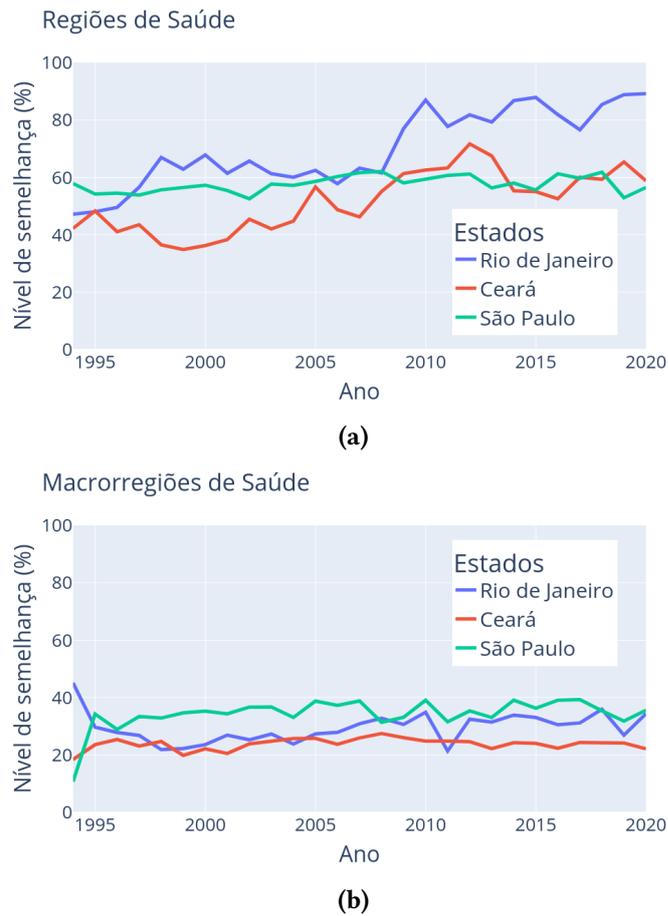
---

<sup>1</sup><https://scikit-learn.org/>



**Figura 6.11:** Resultado dos algoritmos de comparação de partições (a esquerda) e sua padronização (a direita), para os estados do Rio de Janeiro ((a) e (b)), Ceará ((c) e (d)) e São Paulo ((e) e (f)) desde 1994 a 2020. Comparação do resultado do infomap com as Regiões de Saúde dos três estados.

observamos uma maior constância nos resultados, o que pode significar pouca alteração no comportamento dos pacientes ao longo dos anos neste nível de análise. Além disso, vemos que os valores para as macrorregiões são mais baixos. Logo, as comunidades detectadas são mais diferentes da delimitação das macrorregiões do que acontece para as regiões de saúde.



**Figura 6.12:** *Nível de semelhança entre comunidades detectadas pelo infomap em comparação com Regiões (a) e Macrorregiões (b) de Saúde para os estados do Rio de Janeiro, Ceará e São Paulo. Resultado do algoritmo de comparação de partições Adjusted Rand Index para os anos entre 1994 e 2020.*

Com isso, conseguimos responder de forma mais objetiva a primeira questão de pesquisa deste trabalho: “**Q1:** *A movimentação dos pacientes está relacionada à delimitação das regiões de saúde? O quanto?*”. Já vimos anteriormente que o resultado do *infomap* revela que existe relação entre a movimentação de pacientes e a delimitação das regiões de saúde. As Figuras 6.7, 6.8 e 6.9 trazem visualmente essa relação. Agora, através dos níveis de semelhança ao longo dos anos, podemos medir *o quanto* elas estão relacionadas, como mostra a Figura 6.12. Entretanto, não podemos deixar de considerar que a determinação dos limites das regiões de saúde considera diversas outras variáveis além das interações

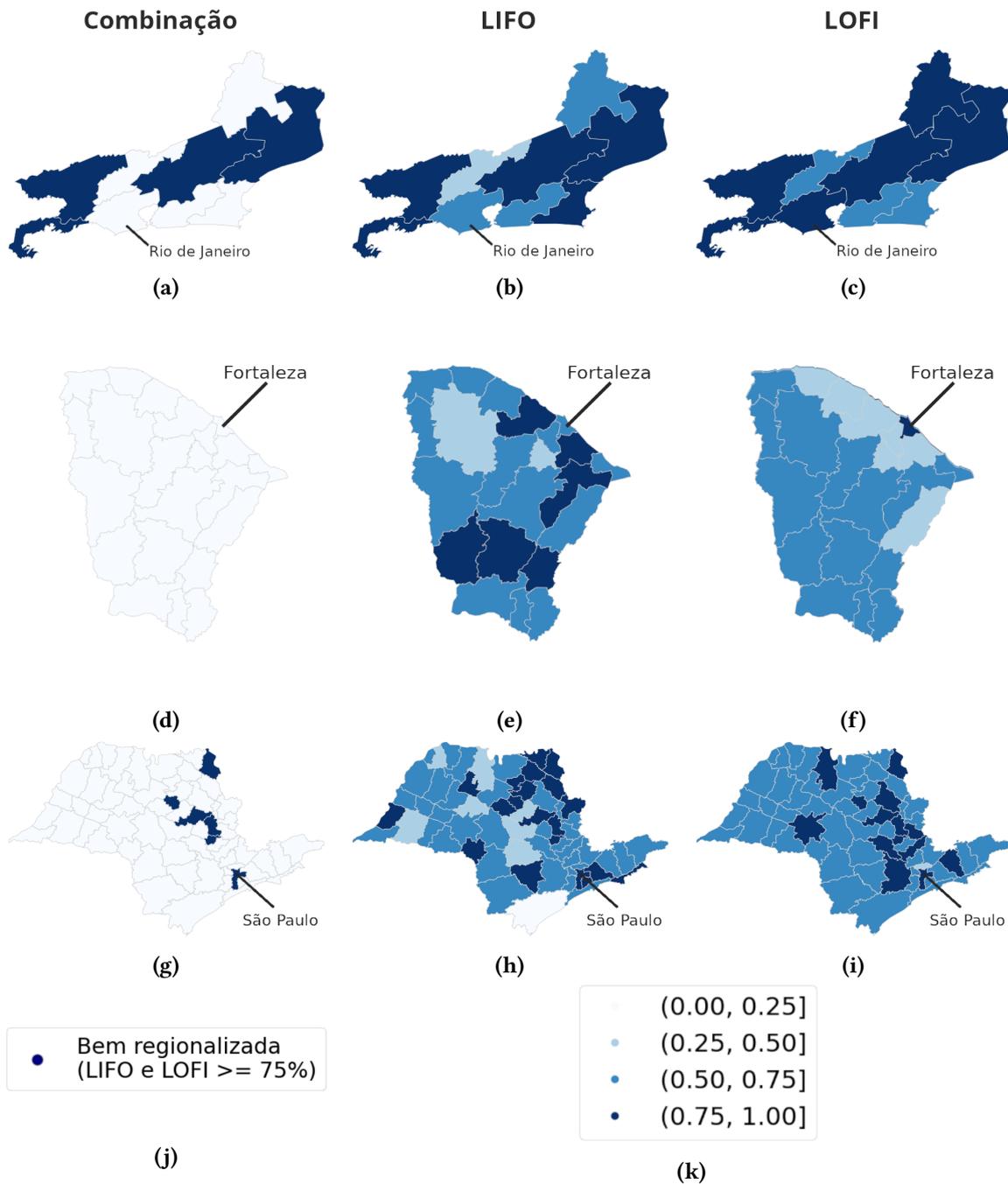
hospitalares. Este método foi desenvolvido para ser uma ferramenta de análise comparativa de regiões (e macrorregiões) de saúde a partir do fluxo de pacientes gerado pelas internações hospitalares de cada estado, de forma a fornecer evidências a gestores públicos para a proposição e monitoramento de políticas públicas em saúde.

### 6.3 Aplicação do índice de desempenho

Como explicado na Seção 5.2.2, o objetivo do índice de desempenho (ou autossuficiência) é medir o nível de suficiência de cada região de saúde isoladamente. Temos esse interesse, pois, idealmente, uma região de saúde deveria conseguir fornecer os serviços de saúde suficientes para sua própria população, atendendo integralmente. Testamos duas maneiras de medir o nível de regionalização das regiões de saúde a partir do fluxo de pacientes: (i) os coeficientes LIFO e LOFI, sugeridos por [ELZINGA e HOGARTY, 1973](#); e (ii) as taxas de atração e permanência, desenvolvidas nesta pesquisa.

O resultado da combinação original dos coeficientes LIFO e LOFI, apresentado pela Figura 6.13, em (a), (d) e (g), oferece uma boa representação do grau de regionalização de uma certa região de saúde, mas não de sua suficiência. Para ilustrar esse fenômeno, observe que, para o Rio de Janeiro e para o Ceará, as regiões de saúde que contêm as capitais não são consideradas bem regionalizadas. Esperamos que uma região que contenha a capital seja considerada suficiente, pois é uma região que possui uma melhor infraestrutura, ela consegue atender seus próprios pacientes e outros externos. Entretanto, isso não foi bem representado pela combinação original do LIFO com o LOFI. Esse resultado pode ser explicado pelo significado do coeficiente LIFO. Como apresentado na Seção 5.2.2, esse coeficiente indica que a região atende poucos pacientes de fora. Valorizar essa medida prejudica justamente a avaliação das regiões que atraem pacientes de fora, o que não faz sentido para o nosso caso, já que essas regiões (como a região que contém a capital) atraem mais pacientes por serem mais bem estruturadas.

A segunda coluna da Figura 6.13 apresenta em azul-escuro todas as regiões que possuem valor acima de 75% para o coeficiente LIFO, conforme sugerido por [ELZINGA e HOGARTY, 1978](#), e as capitais do Rio de Janeiro e do Ceará não se encontram nessa categoria. Logo, o coeficiente LIFO não se mostrou adequado para colaborar na determinação de um índice de suficiência para as regiões de saúde. Embora o coeficiente LOFI seja favorável à nossa definição de suficiência, por indicar que poucos pacientes de uma região buscam hospitalização fora dela, o uso isolado desse coeficiente não é o suficiente para a definição final de nosso índice de desempenho. Para isso, precisaríamos acrescentar nesse índice pelo menos as informações sobre as internações exportadas para regiões vizinhas, que não

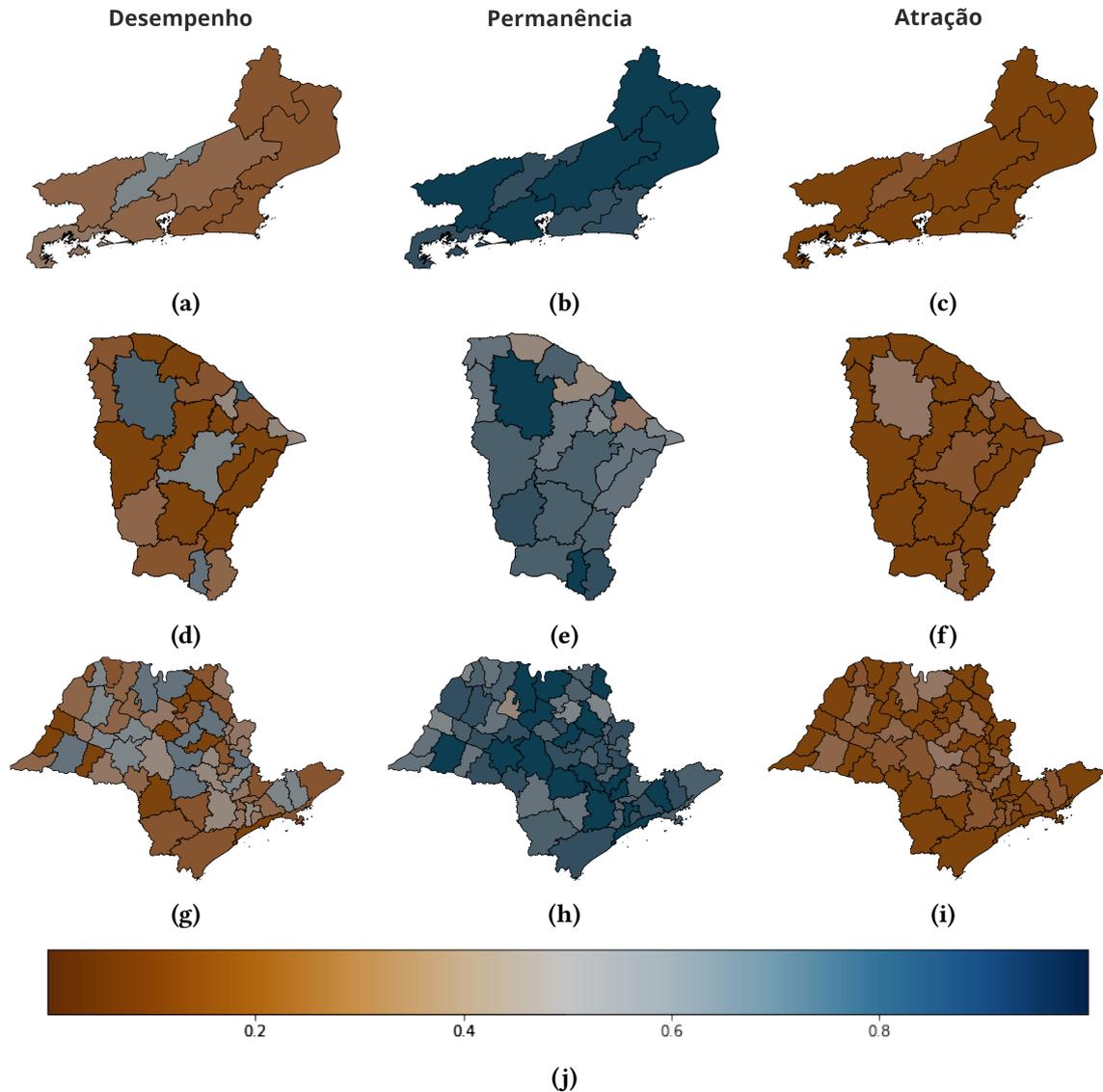


**Figura 6.13:** Resultado do cálculo dos coeficientes LIFO (segunda coluna), LOFI (terceira coluna) e sua combinação (primeira coluna), como proposto em *ELZINGA e HOGARTY, 1973* e *ELZINGA e HOGARTY, 1978*. Para dados de fluxo de pacientes em 2019 nas Regiões de Saúde dos estados do Rio de Janeiro ((a),(b) e (c)), Ceará ((d), (e) e (f)) e São Paulo ((g), (h) e (i)).

é considerado no LOFI.

Para resolver esse problema, fizemos uma modificação nos coeficientes LIFO e LOFI, bem como uma modificação na combinação entre eles. Com isso, a nova métrica desenvolvida combina a “taxa de permanência” (*LOFI*) e a “taxa de atração” ( $1 - LIFO$ ) através de

uma média harmônica ponderada (vide Seção 3.3), como descrito na Seção 5.2.2.



**Figura 6.14:** Resultado do cálculo das taxas de permanência (segunda coluna), de atração (terceira coluna) e sua combinação (primeira coluna), como explicado em 5.2.2. Para dados de fluxo de pacientes em 2019 nas Regiões de Saúde dos estados do Rio de Janeiro ((a),(b) e (c)), Ceará ((d), (e) e (f)) e São Paulo ((g), (h) e (i)).

Dessa forma, como mostra a Figura 6.14, as duas taxas são combinadas para gerar uma métrica final, correspondendo à média harmônica entre as duas com pesos diferentes: a taxa de permanência permanece com peso 1 e a taxa de atração com peso  $\beta$ , onde  $0 \leq \beta \leq 1$ . Pela definição da regionalização da saúde, a taxa de permanência é mais relevante do que a taxa de atração, já que cada região deveria atender integralmente sua população, não causando evasão de pacientes. Desse modo, o peso da taxa de atração não é maior que o peso da taxa de permanência. Para as próximas análises, fixamos o valor de  $\beta$  para 0,2, pois intuitivamente pareceu gerar melhores resultados, mas na plataforma esse valor é

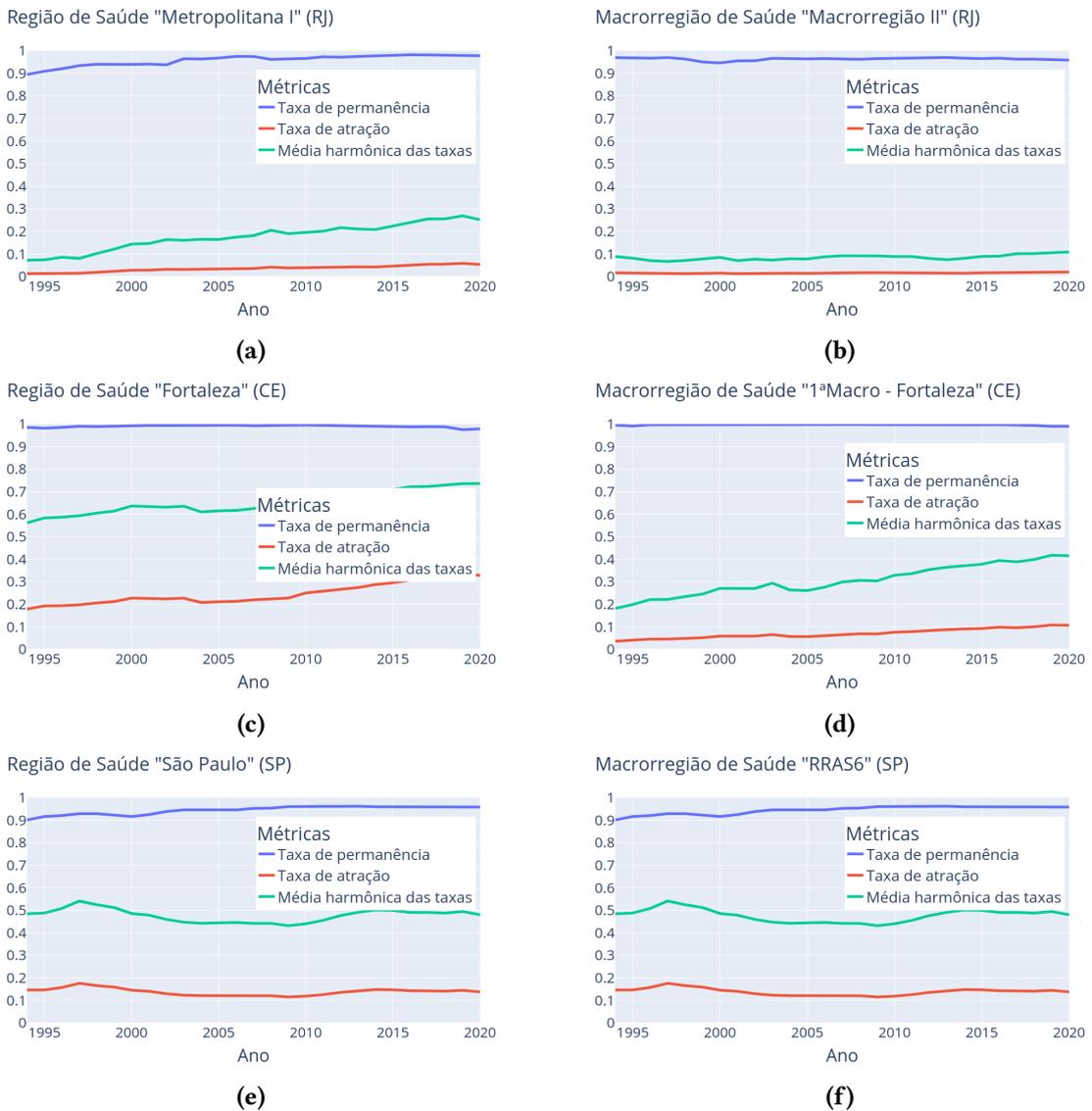
determinado pelo próprio usuário. Logo, a fórmula de combinação das taxas se torna:

$$\text{Desempenho}(w = [0, 2; 1]) = \frac{(0, 2 + 1)(1 - LIFO)LOFI}{(0, 2)LOFI + (1)(1 - LIFO)},$$

onde  $1 - LIFO$  é a taxa de atração e  $LOFI$  é a taxa de permanência. Ademais, todos os sub-gráficos da Figura 6.14 utilizam a mesma legenda com valores entre 0 e 1, e a mesma sequência de cores: entre laranja e azul.

A partir da aplicação da métrica desenvolvida nos demais dados da base, conseguiremos responder a segunda questão de pesquisa deste trabalho: “**Q2:** *Existe uma evolução na autossuficiência das regiões de saúde ao longo dos anos?*”. Entretanto, fica a cargo do gestor de saúde a escolha da métrica final para cada caso particular: a taxa de permanência de pacientes em sua região de saúde de origem; a taxa de atração de pacientes para determinada região; ou a combinação das duas taxas pela média harmônica.

Para ilustrar, a Figura 6.15 revela o comportamento dessas três opções para os estados estudados, analisando a Região de Saúde que contém a capital, assim como a Macrorregião de Saúde em cada um dos três estados. Com isso, podemos identificar períodos de crescimento ou decréscimo de cada taxa, que revela o comportamento do nível de autossuficiência da região em termos de permanência e atração de pacientes. Observamos na figura que o comportamento no estado de São Paulo é bem constante, enquanto no Ceará, por exemplo, observamos um constante crescimento do índice tanto para regiões quanto macrorregiões de saúde, indicando uma evolução no desempenho das regiões, conforme a definição do índice de desempenho.



**Figura 6.15:** Resultado das taxas de permanência e atração, e sua combinação por média harmônica, para as Regiões e Macrorregiões de Saúde que contêm a capital do Rio de Janeiro ((a) e (b)), do Ceará ((c) e (d)), e de São Paulo ((e) e (f)).

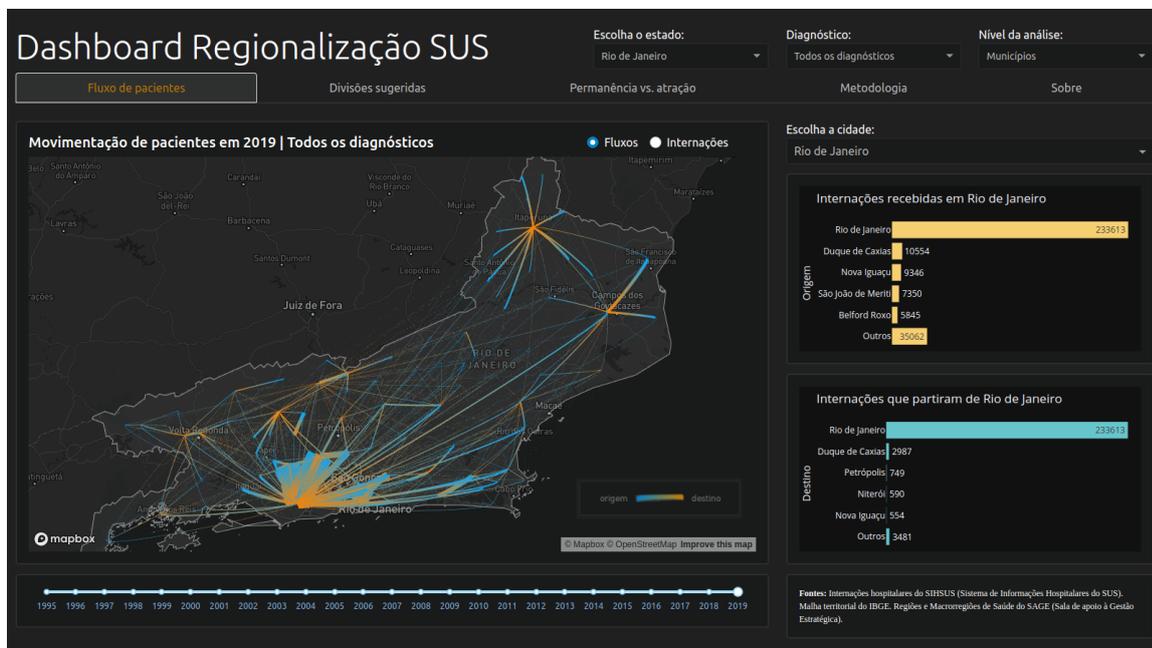
## 6.4 Aplicação da visualização do processo de regionalização

O objetivo final da abordagem de análise do processo de regionalização através da visualização georreferenciada dos dados está em disponibilizar uma ferramenta de consulta para pesquisadores e, especialmente, gestores de saúde, e, com isso, promover políticas públicas baseadas em evidências. Com esse intuito, desenvolvemos uma plataforma web de software livre do tipo *dashboard*, como descrita na Seção 5.2.3, com visualizações criadas para representar os resultados deste trabalho, bem como as visualizações dos

dados georreferenciados de internações hospitalares ao longo dos estados brasileiros. Esta seção irá detalhar as três principais abas que compõem o *dashboard* para o estado do Rio de Janeiro, mas a plataforma final disponibiliza as mesmas análises para os outros estados. Além dessas três abas principais: “Fluxo de pacientes”, “Divisões sugeridas” e “Permanência vs. atração”, o *dashboard* também contém outras duas abas: “Metodologia”, que descreve um pouco da metodologia que já foi descrita aqui, e “Sobre”, apresentando aqueles que colaboraram nesta pesquisa. O *dashboard* pode ser acessado através do link: <https://intercity.org/health/sus-regionalization>, mas as Figuras 6.16, 6.17 e 6.18 mostram as três principais abas da plataforma.

### 6.4.1 Aba “Fluxo de paciente”

Logo do início da página, encontramos quatro elementos em comum para todas as abas do *dashboard*: o título da plataforma “Dashboard Regionalização SUS” seguido por três menus suspensos. Toda alteração feita em qualquer destes três menus suspensos leva a uma atualização dos dados exibidos na página, para qualquer uma das três principais abas do *dashboard*. Esses três menus suspensos alteram (i) o estado analisado, que no caso da Figura 6.16 é o Rio de Janeiro; (ii) o diagnóstico principal das internações exibidas; e (iii) o nível de análise atual, que pode variar entre o nível de “Municípios”, “Regiões de Saúde” ou “Macrorregiões de Saúde”.



**Figura 6.16:** Captura de tela da aba “Fluxo de pacientes” do dashboard desenvolvido neste projeto.

À esquerda da Figura 6.16, está o mapa principal da página (o mesmo acontece nas outras abas). Esse mapa apresenta as movimentações intermunicipais dos pacientes do

estado do Rio de Janeiro, para todos os diagnósticos da base, para o ano de 2019. A parte azul de cada fluxo indica a origem, e a laranja, o destino, que no caso são municípios, mas o usuário pode escolher analisar a movimentação entre regiões ou macrorregiões também. Além disso, ele pode escolher filtrar os dados exibidos para um diagnóstico específico, ou alterar o ano de exibição, analisando a evolução do comportamento da rede de movimentação de pacientes temporalmente. Por fim, o mapa pode apresentar tanto o fluxo de pacientes quanto a quantidade absoluta de internações para cada município, região ou macrorregião no mapa.

Já à direita da Figura 6.16, temos dois gráficos secundários, ambos correspondem à escolha do usuário de um município do estado analisado (no caso, a cidade do Rio de Janeiro). O primeiro contém os principais municípios que enviam pacientes para o município escolhido e o segundo contém os principais municípios que recebem internações do município em análise no momento.

### 6.4.2 Aba “Divisões sugeridas”

Assim como a aba anterior, esta possui uma estrutura semelhante: o mapa principal à esquerda, os gráficos secundários à direita, e os mesmos três menus suspensos no topo. Esta aba está relacionada à abordagem de *detecção de comunidades*, como explica a Seção 5.2.1. Ela apresenta os resultados da aplicação dessa técnica, isto é, a comparação entre a divisão de municípios das regiões e a divisão sugerida pelo algoritmo *Infomap* (vide Seção 3.1.1).

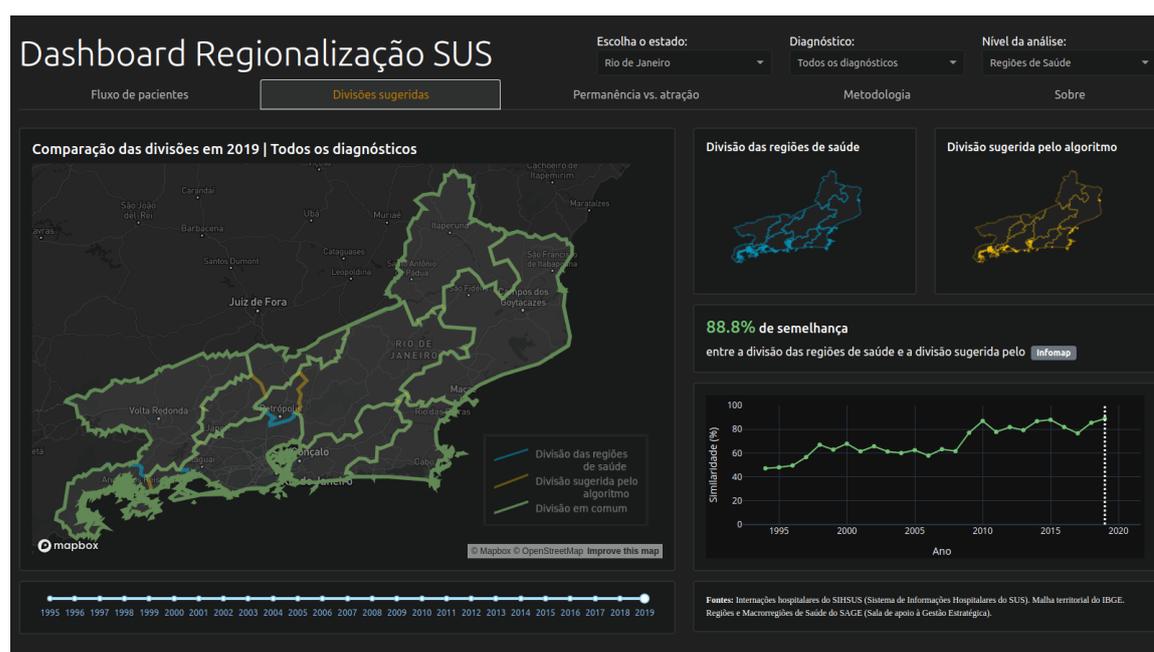


Figura 6.17: Captura de tela da aba “Divisões sugeridas” do dashboard desenvolvido neste projeto.

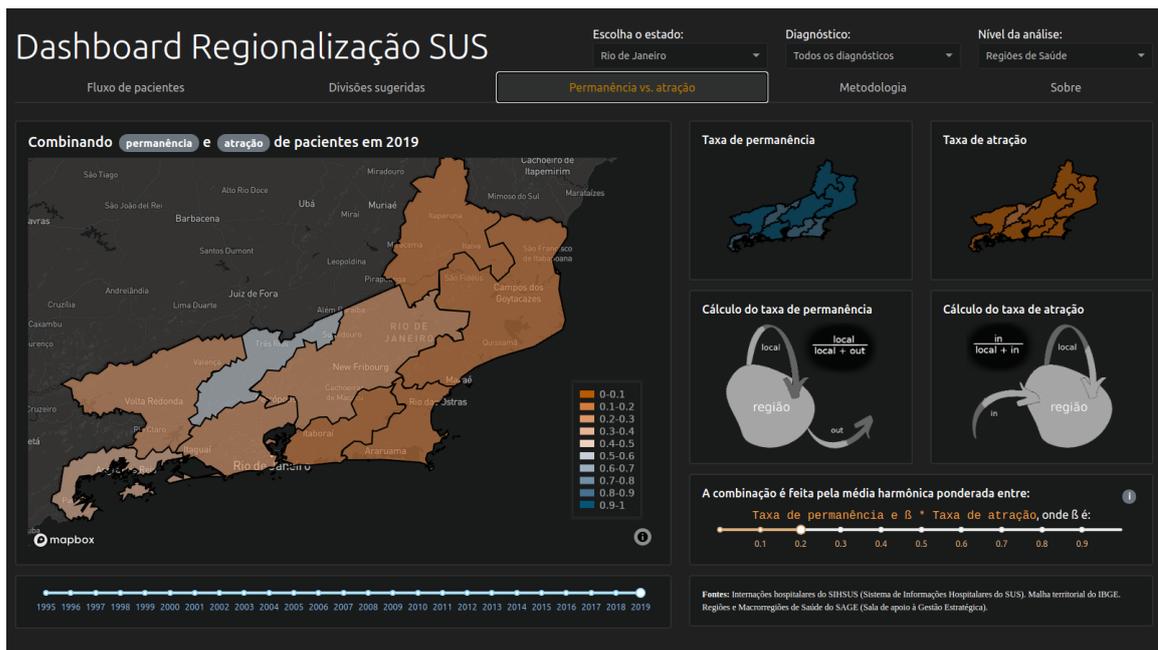
Na Figura 6.17, que mostra a aba “Divisões sugeridas”, a divisão em azul representa as regiões de saúde do ano escolhido (no caso, 2019) e a divisão em amarelo representa a sugestão do algoritmo conforme a movimentação dos pacientes naquele ano, tanto no mapa principal quanto nos mapas da direita. Dessa forma, as divisões em comum ficam em verde, no mapa principal. À direita, além dos dois pequenos mapas, temos a porcentagem do nível de semelhança entre as divisões para todos os anos completos da base, com ênfase no ano selecionado pelo usuário (no caso, 88,8% em 2019). Assim, podemos ver que no Rio de Janeiro o índice de semelhança evoluiu de 47,2% para 88,8% ao longo dessas duas décadas.

### 6.4.3 Aba “Permanência vs. atração”

Nesta aba, observamos o resultado da combinação entre a taxa de permanência e a taxa de atração, como explica a Seção 5.2.2. Semelhante à aba “Divisões sugeridas”, esta aba possui um mapa principal, resultado da combinação dos dados de dois pequenos mapas à direita, como constatamos na Figura 6.18. Aqui, o usuário tem a liberdade de selecionar o ano de análise, alterar os valores daqueles três menus suspensos, além de escolher a ponderação entre as duas taxas, ao selecionar o valor do  $\beta$ , como explicado anteriormente (vide Seção 5.2.2). Lembrando que os três mapas desta página utilizam a mesma legenda, com intervalo de valores entre 0 e 1. Além disso, a página contém dois diagramas que explicam esquematicamente como é feito o cálculo das duas taxas para cada região de saúde.

Através das análises permitidas pelo *dashboard*, podemos fazer o acompanhamento da situação de cada região de saúde ao longo dos anos, analisando também os resultados para os anos em que houve alterações na divisão das regiões. Isso possibilita respondermos à terceira questão de pesquisa do projeto: “**Q3:** *As alterações das regiões de saúde ao longo dos anos interferiram na movimentação dos pacientes? Houve melhorias?*”. Tendo conhecimento dos momentos em que houve alguma alteração nas regiões de saúde, o usuário pode combinar o resultado apresentado pelas três abas principais do *dashboard* para analisar seus efeitos no comportamento dos pacientes. Logo, é possível examinar se houve alteração na movimentação de pacientes, em relação aos anos anteriores, através da aba “Fluxo de pacientes”. Além de que pode também examinar se houve alguma melhoria na similaridade das partições ou no valor das taxas, através das abas “Divisões sugeridas” e “Permanência vs. atração”, respectivamente.

Portanto, essa última abordagem fornece visualizações objetivas e interativas, auxiliando na compreensão do comportamento da movimentação de pacientes nos estados e pode ser uma ferramenta útil para gestores de saúde, já que está publicamente disponível. Agora,



**Figura 6.18:** Captura de tela da aba “Permanência vs. atração” do dashboard desenvolvido neste projeto.

a partir dos resultados apresentados neste capítulo, apresentaremos no capítulo seguinte alguns exemplos de casos interessantes que podem servir de base na criação de políticas públicas em saúde, além de apresentar as limitações encontradas no caminho.

# Capítulo 7

## Discussões Finais

No capítulo anterior, apresentamos de forma geral os resultados obtidos em nossa pesquisa através das abordagens desenvolvidas. Elas são procedimentos que analisaram e avaliaram a eficiência das regiões e macrorregiões de saúde e sua evolução temporal através de técnicas de sistemas complexos. Para dar suporte à elaboração de políticas públicas baseadas em evidências, desenvolvemos uma plataforma única e integrada de software livre do tipo *dashboard*, disponibilizando-a publicamente. Neste capítulo, discutiremos como os resultados obtidos podem auxiliar gestores ou pesquisadores a analisar e avaliar o processo de regionalização em diferentes partes do Brasil e sob diferentes fatores.

Para isso, separamos alguns casos de utilização da plataforma para exemplificar seu potencial na prática. Para tanto, este capítulo foi organizado em três seções, em referência às três abordagens da metodologia e às três principais abas da plataforma: “Fluxo de pacientes”, “Divisões sugeridas” e “Permanência vs. atração”.

### 7.1 Fluxo de pacientes

Na primeira aba, o usuário da plataforma já consegue fazer algumas análises através do processamento e exibição da movimentação de pacientes ao longo dos estados brasileiros. A plataforma permite, por exemplo, o processamento das internações exibidas no mapa pelo diagnóstico principal. A Figura 7.1 exibe o resultado desse processamento para alguns casos.

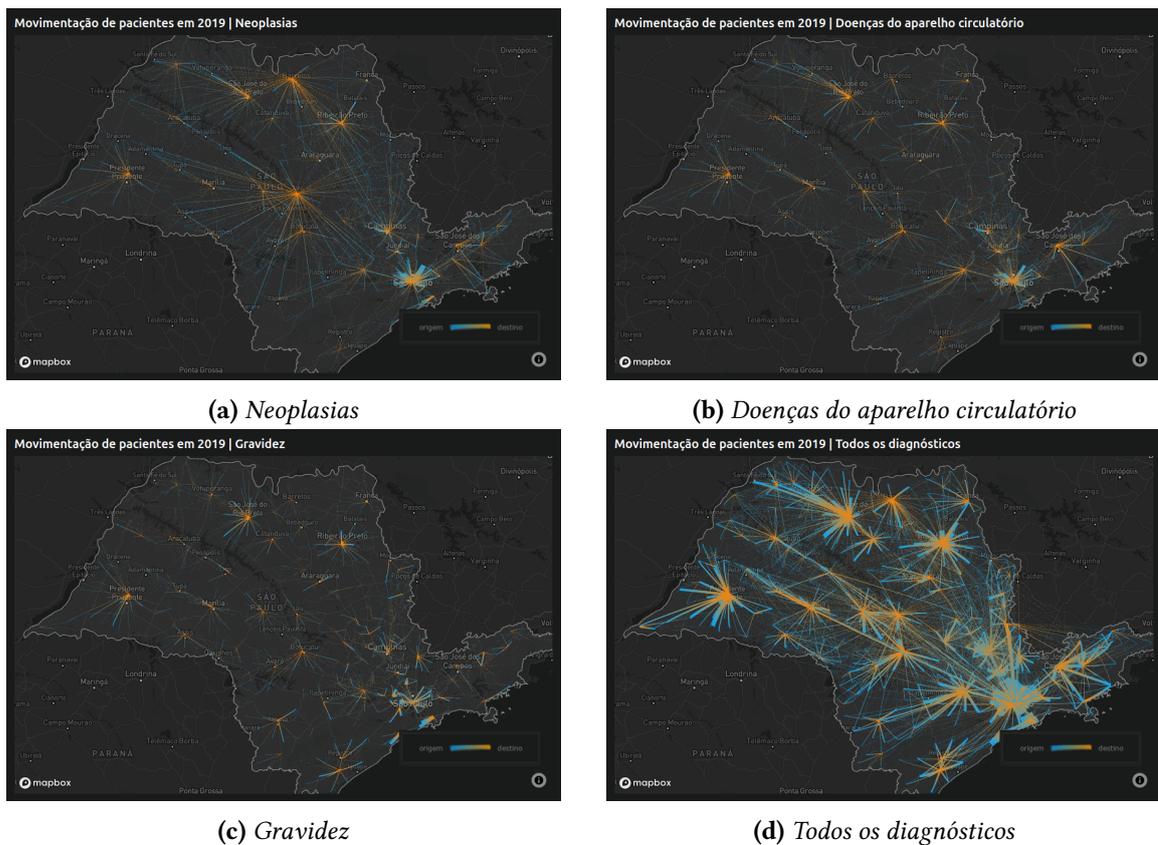
Interessante notar que, na Figura 7.1(a), que apresenta os resultados para neoplasias<sup>1</sup>, a distância do deslocamento de pacientes é maior (as linhas dos fluxos são mais longas).

---

<sup>1</sup>De acordo com o Instituto Nacional de Câncer (INCA), neoplasia é um tumor que ocorre pelo crescimento anormal do número de células. Quando maligna, considera-se câncer.

Embora a capital possua uma maior densidade de internações, algumas cidades do interior do estado de São Paulo, como Jaú e Barretos, são destinos bastante procurados por pacientes de diversos pontos do interior do estado. Esses dois municípios atendem pacientes externos em maior proporção do que atendem os seus próprios. Isso revela que eles devem ter uma boa estrutura de atendimento hospitalar especializado para neoplasias. Entretanto, isso revela também que vários pacientes do interior se deslocam por muitos quilômetros para receber o devido atendimento para seu caso.

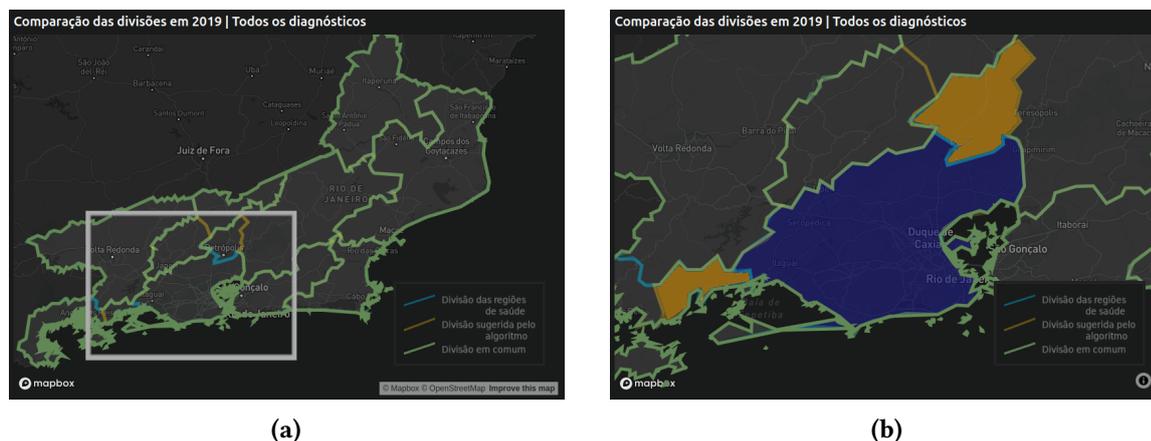
Já os outros diagnósticos possuem comportamento semelhante à movimentação geral de pacientes pelo estado (Figura 7.1(d)), porém em menor proporção. De toda forma, constatamos que existe uma grande distribuição de “municípios polo” neste estado, isto é, de municípios que tendem a receber internações dos municípios vizinhos. Vemos isso conforme os pontos no mapa onde a ponta do fluxo é laranja, indicando o destino de internações.



**Figura 7.1:** Fluxos intermunicipais de pacientes no estado de São Paulo em 2019, separado por alguns diagnósticos.

## 7.2 Divisões sugeridas

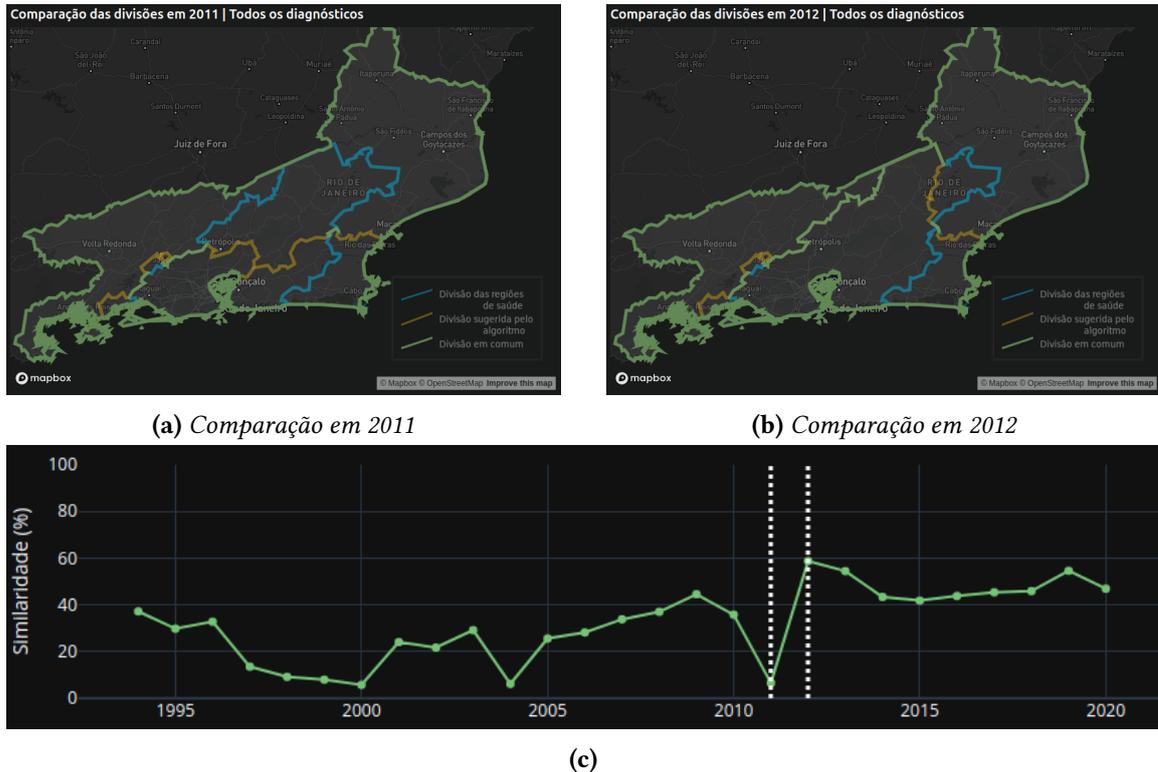
Na segunda aba, o usuário é estimulado a fazer uma análise comparativa entre os limites regionais e a sugestão do algoritmo. Com isso, podemos identificar diferentes casos na comparação. Por exemplo, casos em que uma região sugerida é mais ampla que a região original; casos onde o oposto ocorre; além de casos onde o algoritmo sugere a divisão (ou união) de regiões em duas ou mais partes. A Figura 7.2 apresenta um caso onde a região sugerida expande a região de saúde original. Essa é a região que contém a capital do estado, e a sugestão se baseia nas movimentações dos pacientes no ano de 2019. A Figura 7.2(b) apresenta em azul a região original, e a região sugerida seria a composição da original com as duas partes vizinhas em amarelo. Essa sugestão indica que o vínculo entre as áreas em azul e em amarelo, por meio dos fluxos de pacientes entre elas, foi tão forte que o algoritmo considera como uma única região.



**Figura 7.2:** Comparação da divisão das Regiões de Saúde do Rio de Janeiro com a divisão detectada pelo algoritmo Infomap através da rede de movimentação de pacientes em 2019.

Além disso, essa mesma aba do *dashboard* apresenta o nível de semelhança entre a divisão original das regiões e a divisão sugerida. Com essa métrica, conseguimos analisar a evolução temporal dessa similaridade, descobrindo momentos irregulares no período estudado. A Figura 7.3 apresenta um exemplo no estado do Rio de Janeiro para as Macrorregiões de Saúde. Em foco, a figura mostra a comparação entre a divisão original das macrorregiões e a divisão sugerida pelo algoritmo para os dados de movimentação de pacientes em 2011 e 2012. Olhando para a evolução temporal da similaridade entre as divisões, encontramos um grande declive em 2011 seguido de um pico em 2012 (marcados pelas linhas pontilhadas na Figura 7.3(c)). Isso revela momentos em que, de acordo com os dados, o comportamento da movimentação dos pacientes, ao nível macrorregional, não se comportou conforme o esperado originalmente. Dessa forma, supondo que não tenha havido problemas no registro dos dados naqueles anos, esse resultado pode indicar

períodos nos quais houve problemas significativos no serviço de saúde do estado.

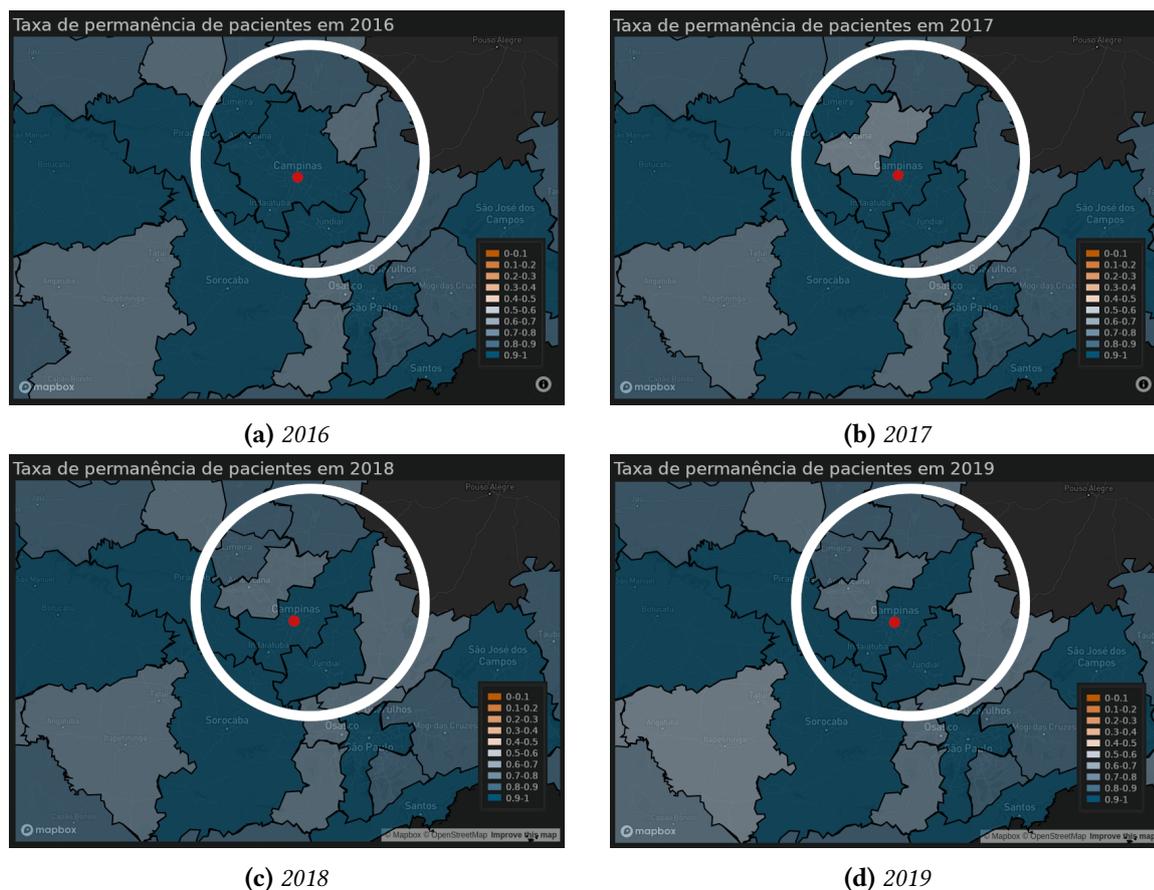


**Figura 7.3:** Comparação da divisão original das Macrorregiões de Saúde com as divisões sugeridas para 2011 e 2012 no estado do Rio de Janeiro.

### 7.3 Permanência vs. atração

Por fim, a terceira aba principal disponibiliza ao usuário a análise temporal das taxas de permanência, atração de pacientes, e sua combinação. Com isso, ele consegue identificar em cada estado regiões que tiveram melhorias ou piores no índice, temporalmente. Além disso, ele pode identificar regiões que sofreram alterações em seus limites e analisar os valores dos índices perante essa mudança, assim como o caso da Figura 7.4.

Na Figura 7.4, a área em foco contém as regiões de saúde do estado de São Paulo que são mais próximas do município de Campinas, em vermelho nos mapas. Houve uma mudança nos limites regionais dessa área em foco, e ela influenciou a taxa de permanência de pacientes das regiões neste caso. Observamos que a região que contém Campinas está sempre em um tom de azul mais escuro, indicando uma maior taxa de permanência de pacientes. Entretanto, a região de saúde que surge em 2017 a noroeste do ponto vermelho (região Oeste VII), que pertencia a mesma região de saúde de Campinas em 2016, contém uma taxa de permanência de pacientes bem mais baixa logo após a mudança. Isso indica que, embora tenha havido uma alteração na divisão das regiões, os municípios daquela

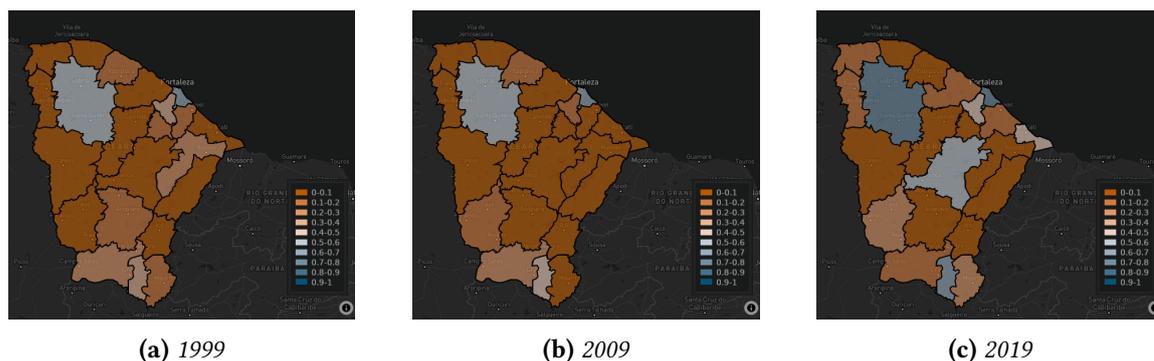


**Figura 7.4:** Alteração da taxa de permanência das Regiões de Saúde próximas a Campinas, São Paulo, que sofreram modificações em seus limites administrativos em 2017.

região mais clara em 2017 possivelmente tinham uma relação forte com Campinas, e ainda continuam a enviar pacientes para lá. Ademais, observamos uma readaptação dos fluxos de pacientes anos após a alteração, pelo aumento da taxa de permanência na região Oeste VII nos anos posteriores, como revela a Figura 7.4(c) e (d). Neste caso, o *dashboard* permite monitorar o impacto, ao longo dos anos, das mudanças administrativas efetuadas no SUS e a avaliação do quanto elas estão realmente sendo refletidas na prática.

Além disso, perceber as regiões de menor taxa de permanência neste mapa do *dashboard* é como perceber regiões que precisam de mais recursos hospitalares, já que sua população tende a receber atendimento hospitalar fora da região de origem.

Similarmente, a combinação entre as taxas de permanência e atração de pacientes carregam significado semelhante ao analisado anteriormente, indicando um método de avaliação das regiões. A Figura 7.5 apresenta essa combinação (ponderando as taxas com peso inteiro para permanência e peso 0,2 para atração; vide Seção 5.2.2) para o estado do Ceará entre os anos de 1999, 2009 e 2019. Seu resultado apresenta a evolução espaço-temporal do desempenho das regiões. Em 1999 apenas as regiões de Fortaleza (no litoral)



**Figura 7.5:** Nível de desempenho das Regiões de Saúde do Ceará para os anos 1999, 2009 e 2019, combinando taxa de permanência e  $\beta \times$  taxa de atração por média harmônica, onde  $\beta = 0, 2$ .

e de Sobral estão num tom mais azulado, apresentando valores acima de 0,5, refletindo um bom desempenho. Em 2009 o resultado se mantém semelhante, mas em 2019 outras regiões se sobressaem. Delas, além de Fortaleza e Sobral, as regiões de Juazeiro do Norte (ao sul do estado) e de Quixadá (ao centro) superaram o 0,5, e, pela evolução observada no *dashboard*, seus níveis de desempenho têm crescido gradualmente nos últimos anos e tendem a crescer nos anos futuros.

Estes foram alguns casos de análises que podem servir de base para a proposição e a análise de impacto de políticas públicas de saúde baseadas em evidências. Além destas, o usuário possui grande liberdade para analisar outros casos através da plataforma. Adicionalmente, a plataforma é de software livre e pode ser estendida no futuro. O objetivo principal aqui é promover novas políticas públicas que melhorem o atendimento hospitalar da população brasileira. No próximo e último capítulo, discutiremos as contribuições deste projeto na academia e na prática, além de apresentar as limitações encontradas no caminho e alguns possíveis trabalhos futuros a partir do que foi desenvolvido aqui.

# Capítulo 8

## Conclusões

Este trabalho teve o objetivo de desenvolver e disponibilizar ferramentas de análise da regionalização da saúde pública no Brasil por meio da movimentação de pacientes por internações hospitalares. Embora a análise da regionalização através das internações hospitalares tenha demonstrado resultados bem relevantes, sabemos que diversos fatores influenciam a delimitação das regiões de saúde, como a acessibilidade por meio de transporte, já que questões de locomoção influenciam a movimentação dos pacientes. Por exemplo, observamos que a distância média percorrida pelos pacientes, especialmente no nordeste, foi crescendo ao longo das décadas (vide Tabela 6.1). Imaginamos que isso seja uma consequência da evolução da malha rodoviária dos estados e da popularização do transporte aéreo. Outro fator que pode afetar essa análise é a evolução do nível econômico do país, impactando nas decisões de deslocamento do paciente, por exemplo, o que influenciaria diretamente as análises de regionalização (ROCHA *et al.*, 2022). De toda forma, acreditamos que as análises aqui desenvolvidas servem de parâmetro para descrever o processo de regionalização de forma geral.

A partir disso, este capítulo expõe objetivamente as conclusões deste trabalho respondendo às questões de pesquisa, mostrando suas consequências na academia e na prática, apresentando algumas limitações e, por fim, descrevendo possíveis trabalhos futuros. Como guia para alcançarmos o objetivo geral do projeto, desenhamos três questões de pesquisa (vide Seção 1.1). Elas foram respondidas no Capítulo 6, mas retomamos as respostas aqui de forma mais objetiva:

**Q1: A movimentação dos pacientes está relacionada à delimitação das regiões de saúde? O quanto?** A movimentação de pacientes está, sim, relacionada às regiões de saúde. Isso é demonstrado através da abordagem de detecção de comunidades, pelo fato das divisões sugeridas pelo algoritmo serem bastante semelhantes às divisões originais das

regiões. *O quanto* eles estão relacionados foi respondido por meio do nível de similaridade entre as duas formas de dividir a rede de movimentação de pacientes. No Rio de Janeiro, por exemplo, o nível de similaridade esteve em crescimento nos últimos anos (como visto na Figura 6.17), indicando um crescimento na relação entre a movimentação de pacientes e as regiões de saúde do estado ao longo do último quarto de século.

**Q2: Existe uma evolução na autossuficiência das regiões de saúde ao longo dos anos?** Conforme a definição de autossuficiência (ou desempenho) determinada aqui (vide 5.2.2), nem todas as regiões apresentaram uma boa evolução temporal. Algumas se mantiveram bem estáveis (isso foi observado no estado de São Paulo, por exemplo) e outras apresentaram uma boa evolução (como no Rio de Janeiro e em grande parte das regiões do Ceará), de acordo com os dados. Mas, de forma geral, a média dos estados apresenta grande estabilidade temporal no índice de desempenho, com um leve e constante crescimento nos últimos 10 anos da base de dados analisada.

**Q3: As alterações das regiões de saúde ao longo dos anos interferiram na movimentação dos pacientes? Houve melhorias?** Embora os estados estudados de forma particular no Capítulo 6 não apresentem grandes alterações na configuração das regiões, os poucos casos em que isso ocorreu revelaram interferir na movimentação dos pacientes ao longo dos anos. Em alguns casos, percebemos que no ano logo após à modificação parece haver uma piora no índice de desempenho de alguma das regiões envolvidas (como no litoral do Ceará entre 2011 e 2012, e na região próxima de Campinas entre 2016 e 2017 em São Paulo). Entretanto, esse desempenho tende a crescer nos anos seguintes, provavelmente devido à readaptação dos fluxos de pacientes em respeito a nova delimitação administrativa das regiões de saúde.

## 8.1 Contribuições à pesquisa e à prática

Os métodos desenvolvidos para cumprir com o objetivo desta pesquisa gerou contribuições na academia, podendo influenciar futuras pesquisas, bem como na prática, abrindo horizontes para a análise do processo de regionalização da saúde no Brasil. Em resumo, eles estendem o estado da arte ao implementar uma análise comparativa a partir da detecção de comunidades, ao criar um índice de desempenho baseado no fluxo de oferta e demanda dos serviços hospitalares, e ao desenvolver uma plataforma web de software livre permitindo a análise do processo de regionalização através de nossa metodologia.

As contribuições para a academia foram: (1) explorar uma forma inédita de sugerir divisões regionais a partir do fluxo de pacientes comparando numericamente sua semelhança com a divisão original das regiões; (2) aplicar na área de saúde um método recente

de detecção de comunidades (*Infomap*) que se revelou bem eficaz neste caso; (3) propor nova maneira de avaliar o desempenho das regiões de saúde conforme seus fluxos externos e internos de pacientes, pelas taxas de permanência e atração; e (4) desenvolver um leiyoute simplificado e intuitivo para expor fluxos, limites territoriais e valores absolutos em mapa por meio de um *dashboard online* interativo.

Na prática, as principais contribuições desta dissertação foram: (1) trazer inovação na análise de uma base pública do DATASUS, a base de internações hospitalares (SIHSUS), fornecendo métodos de visualização mais avançados e intuitivos para ela; (2) abrir novos horizontes para futuras análises do fluxo de pacientes em outros níveis de atenção além do nível hospitalar, através da utilização de outras bases do DATASUS; (3) permitir análises espaço-temporais de cada estado brasileiro para um longo intervalo de dados (mais de vinte anos); e (4) colocar em foco o processo de regionalização de saúde no Brasil, um tema não tão estudado atualmente, mas de significativa importância.

## 8.2 Limitações

A base de internações hospitalares (SIHSUS), fornecida publicamente pelo Departamento de Informática do Sistema Único de Saúde (DATASUS), apresenta alguns problemas de qualidade da informação. Alguns estudos prévios (AGUIAR *et al.*, 2013; MELO *et al.*, 2004; L. S. d. COSTA *et al.*, 2003; VERAS e MARTINS, 1994; BITTENCOURT *et al.*, 2008; PAULA ORLANDI *et al.*, s.d.) apontaram problemas de confiabilidade em alguns campos da base: o endereço de origem do paciente e seu diagnóstico principal. Para o endereço de residência do paciente, foi observado o preenchimento do endereço de destino da internação em alguns casos. No caso onde o diagnóstico principal era câncer de mama e de colo de útero entre 2001 e 2002 no estado do Rio de Janeiro, esse erro ocorreu em quase 10% dos casos (AGUIAR *et al.*, 2013). No campo do diagnóstico, observou-se a ausência de emissão do formulário de internação e alteração do diagnóstico real em mais de 40% das internações por infarto agudo do miocárdio no município do Rio de Janeiro, em 2000 (MELO *et al.*, 2004). Além disso, os primeiros anos da base, na década de 1990, parecem ter dados incompletos, o que foi gradualmente melhorando por volta da virada do século. Com esses e outros casos, ficamos cientes de que os dados utilizados pela base SIHSUS não cobrem completamente todas as internações, mas como a solução está longe do nosso alcance, a utilizamos por ser o que temos disponível.

Os limites regionais dos estados brasileiros foram obtidos a partir do portal da Sala de apoio à Gestão Estratégica (SAGE). Neste trabalho utilizamos a atualização anual das regiões de saúde entre os anos de 2011 e 2017, e 2019. Para as macrorregiões de saúde,

tivemos acesso aos dados apenas para o ano de 2019. Dessa forma, as análises desenvolvidas utilizaram os limites regionais do ano disponível mais próximo, já que não tivemos acesso aos dados completos.

### 8.3 Trabalhos futuros

Por fim, identificamos alguns trabalhos que podem ser desenvolvidos a partir dos conhecimentos obtidos em nossa pesquisa. Como a plataforma desenvolvida é de software livre, ela pode ser estendida em outras direções por qualquer grupo interessado. Logo, sugerimos os seguintes trabalhos futuros, com relação à adição de novos recursos na plataforma ou criação de novas análises a partir dela:

- Divulgação e validação de nossa metodologia e *dashboard* com gestores, pesquisadores, ativistas e elaboradores de políticas públicas de saúde. Tal validação poderá embasar futuros desenvolvimentos das técnicas e software produzidos em nossa pesquisa.
- Implementação da técnica *edge bundling* de visualização de fluxos, que agrupa arestas de fluxos semelhantes gerando uma visualização mais limpa dos padrões da movimentação dos pacientes pelo estado. Este tema será explorado em outra dissertação de mestrado em andamento em nosso grupo de pesquisa.
- Inserção de outras bases de dados do DATASUS, além das internações hospitalares, para análise do processo de regionalização da saúde em outros níveis de complexidade, como atenção básica.
- Adição de novas opções de filtragem dos fluxos além dos diagnósticos, como o tipo de atendimento, nível de complexidade, sexo, raça/cor ou faixa etária do paciente, por exemplo.
- Análise do processo de regionalização da saúde ao nível interestadual, além do intermunicipal abordado aqui, explorando os fluxos de pacientes que ultrapassam os limites estaduais.
- Semelhante à sugestão anterior, mas analisando o nível intrarregional, explorando os fluxos de pacientes entre municípios, bairros, hospitais, entre outros, dentro de uma região de saúde.
- Avaliação sistemática de usabilidade da plataforma, através do método *System Usability Scale* ou outro semelhante.
- Desenvolvimento de visualização da infraestrutura de saúde das regiões, adicionando

quantidade de hospitais, leitos, especializações atendidas em cada região. Essas informações podem ser obtidas publicamente pela base de Cadastro Nacional de Estabelecimentos de Saúde (CNES).

- Criação de análises de indíquem locais que necessitem de mais investimento, de acordo com a demanda de pacientes e baixa infraestrutura na região.
- Análise espaço-temporal da expansão da rede de assistência hospitalar nas regiões, como a evolução da oferta de leitos hospitalares, por exemplo.
- Implementação da atualização mensal da base de internações SIHSUS utilizada pela plataforma, adicionando os meses mais recentes disponibilizados pelo sistema de arquivos do DATASUS.
- Análise da regionalização pelo fluxo de pacientes internados por COVID-19.
- Identificação automática do surgimento de novos polos de atração de pacientes ao longo dos anos, a partir da movimentação de pacientes.
- Realização de previsões do comportamento da movimentação dos pacientes frente a futuras alterações nas delimitações de regiões de saúde através de modelos de aprendizado de máquina. Poderia ser feita uma previsão do desempenho das regiões semelhantemente ao realizado por um colega de nosso grupo de pesquisa ([ALEIXO et al., 2022](#)) na previsão de casos de dengue.
- Utilização de um critério de avaliação formal que compare os resultados dos algoritmos de detecção de comunidades e revele objetivamente qual representa melhor o fluxo de pacientes pelo estado.

Esperamos, enfim, que o trabalho desenvolvido nesta pesquisa seja um novo passo em direção à resolução dos problemas do acesso aos serviços de saúde. Acreditamos que ele, juntamente com outros trabalhos que possam ser derivados a partir dele, contribua de alguma forma para a efetivação de um sistema de saúde mais regionalizado, melhorando os serviços que são oferecidos para a população em nosso país.



## Referências

- [AGUIAR *et al.* 2013] Fernanda Pinheiro AGUIAR, Enirtes Caetano Prates MELO, Evangelina Xavier Gouveia de OLIVEIRA, Marília Sá CARVALHO e Rejane Sobrino PRINHEIRO. “Confiabilidade da informação sobre município de residência no sistema de informações hospitalares-sistema único de saúde para análise do fluxo de pacientes no atendimento do câncer de mama e do colo do útero”. Em: *Cadernos Saúde Coletiva* 21 (2013), pgs. 197–200 (citado na pg. 81).
- [ALCÂNTARA 2011] Fernanda Henrique Cupertino ALCÂNTARA. “As reformas legais e o processo de descentralização: aspectos jurídicos e políticos”. Em: *Revista de Sociologia e Política* 19 (2011), pgs. 197–210 (citado na pg. 10).
- [ALEIXO *et al.* 2022] Robson ALEIXO, Fabio KON, Rudi ROCHA, Marcela Santos CAMARGO e Raphael Y de CAMARGO. “Predicting dengue outbreaks with explainable machine learning”. Em: (2022) (citado na pg. 83).
- [AMARAL e OTTINO 2004] Luís AMARAL e J. OTTINO. “Complex networks. augmenting the framework for the study of complex systems”. Em: *European Physical Journal B* 38 (mar. de 2004), pgs. 147– (citado na pg. 14).
- [ARAUJO *et al.* 2016] Daiane ARAUJO, Edgar MERCHAN-HAMANN, Francisca da Silva LIMA, Josué LAGUARDIA e Maria Urdaneta GUTIERREZ. “Avaliação de desempenho das redes de atenção à saúde: uma proposta de indicadores”. Em: *Revista Eletrônica de Comunicação, Informação e Inovação em Saúde* 10.3 (2016). ISSN: 1981-6278. DOI: [10.29397/reciis.v10i3.1090](https://doi.org/10.29397/reciis.v10i3.1090). URL: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/1090> (citado na pg. 30).
- [BARRAT *et al.* 2008] Alain BARRAT, Marc BARTHÉLEMY e Alessandro VESPIGNANI. *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008. DOI: [10.1017/CBO9780511791383](https://doi.org/10.1017/CBO9780511791383) (citado nas pgs. 13–15).

- [BITTENCOURT *et al.* 2008] Sonia Azevedo BITTENCOURT, Luiz Antonio Bastos CAMACHO e Maria do Carmo LEAL. “A qualidade da informação sobre o parto no sistema de informações hospitalares no município do rio de janeiro, brasil, 1999 a 2001”. Em: *Cadernos de Saúde Pública* 24 (2008), pgs. 1344–1354 (citado na pg. 81).
- [BERTIN 1983] Jacques BERTIN. *Semiology of graphics*. University of Wisconsin press, 1983 (citado na pg. 39).
- [BIGGS *et al.* 1986] Norman BIGGS, E Keith LLOYD e Robin J WILSON. *Graph Theory, 1736-1936*. Oxford University Press, 1986 (citado na pg. 14).
- [BOCCARA 2010] Nino BOCCARA. *Modeling complex systems*. Springer Science & Business Media, 2010 (citado nas pgs. 13, 14).
- [M. S. BRASIL 2000] Ministério Saúde BRASIL. *Sistema Único de Saúde: SUS: princípios e conquistas*. 2000. URL: <https://pesquisa.bvsalud.org/bvsmis/resource/pt/mis-1478> (citado na pg. 8).
- [M. S. BRASIL 2006] Ministério Saúde BRASIL. *Regionalização solidária e cooperativa (Série Pactos pela Saúde v.3)*. 2006. URL: <https://bvsmis.saude.gov.br/bvs/publicacoes/regionalizacao2006.pdf> (citado na pg. 11).
- [BRASIL 2011] BRASIL. *Decreto nº 7.508 de 28 de junho de 2011*. 2011. URL: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/decreto/d7508.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/decreto/d7508.htm) (citado nas pgs. 2, 10).
- [M. S. BRASIL 2021] Ministério Saúde BRASIL. *SUS: Estrutura, princípios e como funciona*. 2021. URL: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/s/sus-estrutura-principios-e-como-funciona> (acesso em 30/03/2022) (citado na pg. 9).
- [BRASIL 1988] BRASIL. “Constituição da república federativa do brasil”. Em: (1988) (citado na pg. 7).
- [BRASIL 1990] BRASIL. *Lei nº 8080 de 19 de setembro de 1990*. 1990. URL: [http://www.planalto.gov.br/ccivil\\_03/leis/l8080.htm](http://www.planalto.gov.br/ccivil_03/leis/l8080.htm) (citado na pg. 7).
- [CARDOSO *et al.* 2016] Mariana CARDOSO, Amália Ivine SANTANA, Adje SANTOS e Técia CORDEIRO. “Processo de regionalização na saúde: perspectivas históricas, avanços e desafios”. Em: *Enfermagem Brasil* 15 (nov. de 2016), pg. 227. DOI: [10.33233/eb.v15i4.502](https://doi.org/10.33233/eb.v15i4.502) (citado nas pgs. 3, 10).

## REFERÊNCIAS

- [CHEN 2021] Wei CHEN. “Delineating the spatial boundaries of megaregions in china: a city network perspective”. Em: *Complexity* 2021 (2021) (citado nas pgs. 28, 30).
- [CLAUSET *et al.* 2004] Aaron CLAUSET, Mark EJ NEWMAN e Cristopher MOORE. “Finding community structure in very large networks”. Em: *Physical review E* 70.6 (2004), pg. 066111 (citado nas pgs. 15, 17).
- [L. d. F. COSTA *et al.* 2007] L da F COSTA, Francisco A RODRIGUES, Gonzalo TRAVIESO e Paulino Ribeiro VILLAS BOAS. “Characterization of complex networks: a survey of measurements”. Em: *Advances in physics* 56.1 (2007), pgs. 167–242 (citado na pg. 16).
- [L. S. d. COSTA *et al.* 2003] Luciana Silva da COSTA, Rejane Sobrino PINHEIRO e Renan VR ALMEIDA. “Recuperação de informação de endereço na base de dados aih: percentual de aproveitamento”. Em: *Cad. saúde colet.,(Rio J.)* (2003), pgs. 131–141 (citado na pg. 81).
- [ELZINGA e HOGARTY 1973] Kenneth G ELZINGA e Thomas F HOGARTY. “The problem of geographic market delineation in antimerger suits”. Em: *Antitrust Bull.* 18 (1973) (citado nas pgs. 21, 29, 30, 64, 65).
- [ELZINGA e HOGARTY 1978] Kenneth G ELZINGA e Thomas F HOGARTY. “The problem of geographic market delineation revisited: the case of coal”. Em: *Antitrust Bull.* 23 (1978) (citado nas pgs. 21, 29, 64, 65).
- [FERRAZ JUNIOR 2017] Tercio Sampaio FERRAZ JUNIOR. *Mercado relevante. Enciclopédia jurídica da PUC-SP.* 2017. URL: <https://enciclopediajuridica.pucsp.br/verbete/249/edicao-1/mercado-relevante> (acesso em 15/06/2021) (citado na pg. 22).
- [FRECH III *et al.* 2003] Harry E FRECH III, James LANGENFELD e R Forrest McCLUER. “Elzinga-hogarty tests and alternative approaches for market share calculations in hospital markets”. Em: *Antitrust Lj* 71 (2003), pg. 921 (citado nas pgs. 21, 22, 29, 38).
- [FOWLKES e MALLOWS 1983] Edward B FOWLKES e Colin L MALLOWS. “A method for comparing two hierarchical clusterings”. Em: *Journal of the American statistical association* 78.383 (1983), pgs. 553–569 (citado na pg. 23).
- [FORTUNATO 2010] Santo FORTUNATO. “Community detection in graphs”. Em: *Physics reports* 486.3-5 (2010), pgs. 75–174 (citado nas pgs. 15, 16).

- [GIRVAN e NEWMAN 2002] Michelle GIRVAN e Mark EJ NEWMAN. “Community structure in social and biological networks”. Em: *Proceedings of the national academy of sciences* 99.12 (2002), pgs. 7821–7826 (citado na pg. 15).
- [GUO *et al.* 2018] Diansheng GUO, Hai JIN, Peng GAO e Xi ZHU. “Detecting spatial community structure in movements”. Em: *International Journal of Geographical Information Science* 32.7 (2018), pgs. 1326–1347. DOI: [10.1080/13658816.2018.1434889](https://doi.org/10.1080/13658816.2018.1434889). URL: <https://doi.org/10.1080/13658816.2018.1434889> (citado na pg. 28).
- [GUO 2009] Diansheng GUO. “Flow mapping and multivariate visualization of large spatial interaction data”. Em: *IEEE Transactions on Visualization and Computer Graphics* 15.6 (2009), pgs. 1041–1048 (citado na pg. 28).
- [HUBERT e ARABIE 1985] Lawrence HUBERT e Phipps ARABIE. “Comparing partitions”. Em: *Journal of classification* 2.1 (1985), pgs. 193–218 (citado na pg. 23).
- [HE *et al.* 2018] Biao HE, Yan ZHANG, Yu CHEN e Zhihui GU. “A simple line clustering method for spatial analysis with origin-destination data and its application to bike-sharing movement data”. Em: *ISPRS International Journal of Geo-Information* 7.6 (2018). ISSN: 2220-9964. DOI: [10.3390/ijgi7060203](https://www.mdpi.com/2220-9964/7/6/203). URL: <https://www.mdpi.com/2220-9964/7/6/203> (citado na pg. 28).
- [HODGSON *et al.* 1996] M John HODGSON, KE ROSING, A LEONTIEN e G STORRIER. “Applying the flow-capturing location-allocation model to an authentic network: edmonton, canada”. Em: *European journal of operational research* 90.3 (1996), pgs. 427–443 (citado na pg. 20).
- [HUFFMAN 1952] David A HUFFMAN. “A method for the construction of minimum-redundancy codes”. Em: *Proceedings of the IRE* 40.9 (1952), pgs. 1098–1101 (citado na pg. 17).
- [HAYES e Alan Geoffrey WILSON 1971] M Cordey HAYES e Alan Geoffrey WILSON. “Spatial interaction”. Em: *Socio-Economic Planning Sciences* 5.1 (1971), pgs. 73–95 (citado na pg. 21).
- [JAIN e DUBES 1988] Anil K JAIN e Richard C DUBES. *Algorithms for clustering data*. Prentice-Hall, Inc., 1988 (citado na pg. 23).

## REFERÊNCIAS

- [KNAFLIC 2019] Cole Nussbaumer KNAFLIC. *Storytelling com dados: um guia sobre visualização de dados para profissionais de negócios*. Alta Books, 2019 (citado na pg. 39).
- [KOYLU *et al.* 2018] Caglar KOYLU, Selman DELIL, Diansheng GUO e Rahmi ÇELIK. “Analysis of big patient mobility data for identifying medical regions, spatio-temporal characteristics and care demands of patients on the move”. Em: *International Journal of Health Geographics* 17 (ago. de 2018). DOI: [10.1186/s12942-018-0152-x](https://doi.org/10.1186/s12942-018-0152-x) (citado nas pgs. 28, 30).
- [KRASNER e POPE 1988] G. KRASNER e Stephen POPE. “A cookbook for using the model - view controller user interface paradigm in smalltalk - 80”. Em: *Journal of Object-oriented Programming - JOOP* 1 (jan. de 1988) (citado na pg. 40).
- [LUKE e HARRIS 2007] Douglas A LUKE e Jenine K HARRIS. “Network analysis in public health: history, methods, and applications”. Em: *Annu. Rev. Public Health* 28 (2007), pgs. 69–93 (citado na pg. 1).
- [LIMA, QUEIROZ *et al.* 2012] Luciana Dias de LIMA, Lúcia FN de QUEIROZ, Cristiani Vieira MACHADO e Ana Luiza d’Ávila VIANA. “Descentralização e regionalização: dinâmica e condicionantes da implantação do pacto pela saúde no brasil”. Em: *Ciência e Saúde Coletiva* 17 (2012), pgs. 1903–1914 (citado na pg. 10).
- [LIMA, VIANA *et al.* 2012] Luciana Dias de LIMA, VIANA *et al.* “Regionalização e acesso à saúde nos estados brasileiros: condicionantes históricos e político-institucionais”. Em: *Ciência e Saúde Coletiva* (2012). ISSN: 1413-8123. URL: <https://www.redalyc.org/articulo.oa?id=63024420005> (citado na pg. 10).
- [MINERBA *et al.* 2008] Luigi MINERBA, Alessandro CHESSA, Ross COPPOLA, Gianni MULA e G. CAPPELLINI. “A complex network analysis of a health organization”. Em: *Igiene e sanità pubblica* 64 (fev. de 2008), pgs. 9–25 (citado na pg. 27).
- [MITCHELL 2009] Melanie MITCHELL. *Complexity: A Guided Tour*. USA: Oxford University Press, Inc., 2009. ISBN: 0195124413 (citado na pg. 14).
- [MORRISEY *et al.* 1988] Michael A. MORRISEY, Frank A. SLOAN e Joseph VALVONA. “Defining geographic markets for hospital care”. Em: *Law and Contemporary Problems* 51.2 (1988), pgs. 165–194. ISSN: 00239186. URL: <http://www.jstor.org/stable/1191730> (citado nas pgs. 29, 38).

- [MELO *et al.* 2004] Enirtes Caetano Prates MELO, Claudia TRAVASSOS e Marília Sá CARVALHO. “Qualidade dos dados sobre óbitos por infarto agudo do miocárdio, rio de janeiro”. Em: *Revista de Saúde Pública* 38.3 (2004), pgs. 385–391 (citado na pg. 81).
- [NEWMAN 2004a] Mark EJ NEWMAN. “Analysis of weighted networks”. Em: *Physical review E* 70.5 (2004), pg. 056131 (citado na pg. 29).
- [NEWMAN 2004b] Mark EJ NEWMAN. “Fast algorithm for detecting community structure in networks”. Em: *Physical review E* 69.6 (2004), pg. 066133 (citado nas pgs. 15, 17).
- [NEWMAN e GIRVAN 2004] Mark EJ NEWMAN e Michelle GIRVAN. “Finding and evaluating community structure in networks”. Em: *Physical review E* 69.2 (2004), pg. 026113 (citado nas pgs. 15, 17, 27).
- [OLIVEIRA 2004] Monica D OLIVEIRA. “Modelling demand and supply influences on utilization: a flow demand model to predict hospital utilization at the small area level”. Em: *Applied Economics* 36.20 (2004), pgs. 2237–2251 (citado na pg. 21).
- [PAIM 2009] Jairnilson PAIM. *O que é o SUS*. SciELO-Editora FIOCRUZ, 2009 (citado na pg. 7).
- [PEREIRA *et al.* 2006] Ana Lúcia PEREIRA, André Luís Bonifácio de Carvalho CARVALHO e Lumena Almeida Castro FURTADO. “Regionalização solidária e cooperativa: orientações para sua implementação no sus”. Em: *Regionalização solidária e cooperativa: orientações para sua implementação no SUS*. 2006, pgs. 42–42 (citado nas pgs. 1, 2, 10).
- [PAULA ORLANDI *et al.* s.d.] Daniele de PAULA ORLANDI, Thamo de Paiva Coelho JUNIOR e José Elias Feres de ALMEIDA. “Sistema de informações hospitalares (sih-sus): revisão sobre qualidade da informação e utilização do banco de dados em pesquisas”. Em: () (citado na pg. 81).
- [ROSVALL, AXELSSON *et al.* 2009] Martin ROSVALL, Daniel AXELSSON e Carl T BERGSTROM. “The map equation”. Em: *The European Physical Journal Special Topics* 178.1 (2009), pgs. 13–23 (citado nas pgs. 15, 17).
- [ROSVALL e BERGSTROM 2007] Martin ROSVALL e Carl T BERGSTROM. “An information-theoretic framework for resolving community structure in complex networks”. Em: *Proceedings of the national academy of sciences* 104.18 (2007), pgs. 7327–7331 (citado na pg. 20).

## REFERÊNCIAS

- [ROSVALL e BERGSTROM 2008] Martin ROSVALL e Carl T BERGSTROM. “Maps of random walks on complex networks reveal community structure”. Em: *Proceedings of the National Academy of Sciences* 105.4 (2008), pgs. 1118–1123 (citado nas pgs. 15, 19, 20).
- [ROSVALL e BERGSTROM 2011] Martin ROSVALL e Carl T BERGSTROM. “Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems”. Em: *PloS one* 6.4 (2011), e18209 (citado na pg. 15).
- [REILLY 1931] William REILLY. *The Law of Retail Gravitation*. Jan. de 1931 (citado na pg. 21).
- [ROSENBERG e HIRSCHBERG 2007] Andrew ROSENBERG e Julia HIRSCHBERG. “V-measure: a conditional entropy-based external cluster evaluation measure”. Em: *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*. 2007, pgs. 410–420 (citado na pg. 23).
- [RIGOLI *et al.* 2019] Felix RIGOLI, Sergio MASCARENHAS, Domingos ALVES, Tiago CANELAS e Geraldo DUARTE. “Tracking pregnant women displacements in sao paulo, brazil: a complex systems approach to regionalization through the emergence of patterns”. Em: *BMC medicine* 17.1 (2019), pgs. 1–13 (citado nas pgs. 21, 29, 30, 38).
- [ROCHA *et al.* 2022] Rudi ROCHA, Beatriz RACHE e Letícia NUNES. “A regionalização da saúde no brasil”. Em: *Estudo Institucional N. 7 - Instituto de Estudos para Políticas de Saúde* (2022). URL: <https://ieps.org.br/estudo-institucional-07/> (citado na pg. 79).
- [ROY e THILL 2003] John R ROY e Jean-Claude THILL. “Spatial interaction modelling”. Em: *Papers in Regional Science* 83.1 (2003), pgs. 339–361 (citado na pg. 21).
- [STEINLEY 2004] Douglas STEINLEY. “Properties of the hubert-arable adjusted rand index.” Em: *Psychological methods* 9.3 (2004), pg. 386 (citado na pg. 23).
- [STROGATZ 2001] Steven H STROGATZ. “Exploring complex networks”. Em: *nature* 410.6825 (2001), pgs. 268–276 (citado na pg. 14).
- [UCHIMURA *et al.* 2017] Liza Yurie Teruya UCHIMURA, Eronildo FELISBERTO, Edgar Rodrigues FUSARO, Maria Paula FERREIRA e Ana Luiza d’Avila VIANA. “Evaluation

- performance in health regions in brazil”. Em: *Revista Brasileira de Saude Materno Infantil* 17 (2017), S259–S270 (citado na pg. 30).
- [A. L. VIANA *et al.* 2008] Ana Luiza VIANA *et al.* “Novas perspectivas para a regionalização da saúde”. Em: (2008) (citado na pg. 9).
- [A. L. d. VIANA *et al.* 2017] Ana Luiza d’Avila VIANA *et al.* “O processo de regionalizacao no brasil: influencia das dimensoes politica, estrutura e organizacao”. Em: *Revista Brasileira de Saude Materno Infantil* 17 (2017), S27–S43 (citado na pg. 30).
- [VERAS e MARTINS 1994] Claudia Maria T VERAS e Mônica S MARTINS. “A confiabilidade dos dados nos formulários de autorização de internação hospitalar (aih), rio de janeiro, brasil”. Em: *Cadernos de Saúde Pública* 10 (1994), pgs. 339–355 (citado na pg. 81).
- [Allan G WILSON 2002] Allan G WILSON. “Complex spatial systems: challenges for modellers”. Em: *Mathematical and computer modelling* 36.3 (2002), pgs. 379–387 (citado na pg. 21).
- [XAVIER *et al.* 2019] Diego Ricardo XAVIER *et al.* “Health regions in brazil based on hospital admissions: a method to support health regionalization”. Em: *Cadernos de saude publica* 35 (2019) (citado na pg. 29).
- [ZHONG *et al.* 2014] Chen ZHONG, Stefan Muller ARISONA, Xianfeng HUANG, Michael BATTY e Gerhard SCHMITT. “Detecting the dynamics of urban structure through spatial network analysis”. Em: *International Journal of Geographical Information Science* 28.11 (2014), pgs. 2178–2199 (citado na pg. 13).