

**Universidade de São Paulo
Instituto de Matemática e Estatística**

Centro de Estatística Aplicada

Relatório de Análise Estatística

RAE-CEA-24P03

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:

“Avaliação da distribuição espacial em alta resolução do perfil de internações hospitalares no SUS no município de São Paulo”

Bruno Groper Morbin

Fabio Kon

São Paulo, junho de 2024

CENTRO DE ESTATÍSTICA APLICADA - CEA – USP

TÍTULO: Relatório de Análise Estatística sobre o Projeto: “Avaliação da distribuição espacial em alta resolução do perfil de internações hospitalares no SUS no município de São Paulo”.

PESQUISADOR: Prof. Dr. Fabio Kon

ORIENTADOR: Prof. Dr. Fabio Kon

INSTITUIÇÃO: Instituto de Matemática e Estatística da Universidade de São Paulo -
IME USP

FINALIDADE DO PROJETO: Publicação de artigo científico

RESPONSÁVEIS PELA ANÁLISE: Bruno Groper Morbin
Fabio Kon

REFERÊNCIA DESTE TRABALHO: MORBIN, B.G; KON, F. **Relatório de Análise Estatística sobre o Projeto: “Avaliação da distribuição espacial em alta resolução do perfil de internações hospitalares no SUS no município de São Paulo”**. São Paulo, IME-USP, 2024. (RAE-CEA-24P03).

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

ANDERSON, M. J. (2001). **A new method for non-parametric multivariate analysis of variance**. *Austral Ecology*, 26, 32-46. doi:10.1111/j.1442-9993.2001.01070.pp.x.

ARTES, R; BARROSO, L.P. (2023). **Métodos multivariados de análise estatística**. 1.ed. São Paulo: Edgard Blucher. ISBN:9786555067026.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (2012). **Censo Brasileiro de 2010**. Rio de Janeiro: IBGE.

MARTINELLI, G. (2023). **Análise da distribuição geográfica de doenças causadoras de óbitos na cidade de São Paulo utilizando aprendizado de máquina**. São Paulo. Dissertação (Mestrado em Ciência da Computação). Instituto de Matemática e Estatística, Universidade de São Paulo. doi:10.11606/D.45.2023.tde-24112023-144503.

OMS - ORGANIZAÇÃO MUNDIAL DA SAÚDE (2017). **CID-10: Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde**. São Paulo: EDUSP. ISBN:9788531401930.

SILVA, L.J.D. (2000). O conceito de espaço na epidemiologia das doenças infecciosas. **Cadernos de Saúde Pública**, 595-617.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

RStudio (versão 4.4.0)

TÉCNICAS ESTATÍSTICAS UTILIZADAS

Análise Descritiva Unidimensional (03:010)

Análise de Conglomerados (06:120)

Análise de Variância Multivariada Permutacional (08:990)

ÁREA DE APLICAÇÃO

Medicina Epidemiologia (14:040)

Resumo

Este estudo analisa a distribuição geográfica das internações hospitalares no SUS em São Paulo para os anos de 2015, 2018 e 2019. Utilizando dados espaciais e técnicas estatísticas como análise descritiva unidimensional e análise de conglomerados, mapeamos a incidência de doenças conforme a CID-10 e relacionamos essas informações com variáveis demográficas dos setores censitários. Os resultados identificam áreas com diferentes taxas de internação, o que pode ajudar a orientar políticas públicas e melhorar a gestão dos recursos de saúde.

Sumário

1. Introdução.....	8
2. Objetivo.....	8
3. Descrição do estudo.....	8
4. Descrição das variáveis.....	9
5. Análise descritiva.....	11
6. Análise inferencial.....	21
7. Conclusões.....	29
APÊNDICE A.....	31
APÊNDICE B.....	34
ANEXO.....	49

1. Introdução

A epidemiologia pode ser estabelecida como a área da ciência que estuda a distribuição de doenças na população humana e o espaço geográfico está totalmente atrelado a esse conceito (Martinelli, 2023). De acordo com Silva (2000), utilizar o espaço geográfico para o estudo da ocorrência e distribuição das doenças é uma prática que se difundiu antes mesmo da existência da epidemiologia como disciplina científica.

Segundo Martinelli (2023), a disponibilidade de informações com maior qualidade nos dias de hoje e o avanço tecnológico permitem realizar novos estudos com ajuda da Ciência de Dados, possibilitando compreender melhor os fenômenos relevantes para a sociedade.

2. Objetivo

O objetivo do estudo é identificar a distribuição geográfica de doenças de internações hospitalares do SUS agrupadas pelos capítulos da CID-10, catalogados pela OMS (2017). A análise será limitada ao município de São Paulo, observando informações coletadas por setores censitários.

Ao final do estudo, deseja-se concluir quais setores apresentam as maiores taxas de ocorrência das doenças catalogadas e quais possuem as menores, além de visualizar a localização espacial dos mesmos para dar suporte a políticas públicas.

3. Descrição do estudo

O estudo conta com dados disponibilizados pela Secretaria Municipal de Saúde da Prefeitura de São Paulo (SMS-SP) contendo informações de internações

hospitalares no SUS (Sistema Único de Saúde) nos anos de 2015, 2018 e 2019. Esses registros apresentam informações como idade, sexo, setor censitário da residência do paciente, bem como diagnóstico registrado na AIH (Autorização para Internação Hospitalar) e estabelecimento de saúde em que foi internado.

Dessa forma, internações registradas em estabelecimentos de saúde privados, e portanto que não atendem o SUS, não estão contempladas na amostra do estudo.

Também são utilizados dados do censo de 2010 do IBGE (2012) para coleta de informações demográficas dos setores censitários, além de tabela de catalogação das doenças da CID-10 pela OMS (2017) para agrupamento das doenças diagnosticadas.

Para a localização dos estabelecimentos de saúde, são utilizados dados do Cadastro Nacional de Estabelecimento de Saúde (CNES) de 2019.

Com isso, será realizada técnica de agrupamento entre os setores similares em relação à taxa de ocorrência de cada doença.

Há também a divisão de Coordenadorias Regionais de Saúde (CRS) entre os distritos administrativos do município que servirão apenas para análise exploratória, sendo que as técnicas de agrupamento serão realizadas sobre os setores censitários. A observação da localização dos centros hospitalares do SUS que atendem as internações registradas no conjunto de dados da pesquisa também servirá exclusivamente para a exploração dos dados.

4. Descrição das variáveis

Variáveis que caracterizam a amostra dos pacientes internados:

- Idade do paciente (em anos);
- Gênero: masculino ou feminino;
- Setor censitário: código do IBGE, composto por 2 dígitos para o Estado, 5 dígitos para o Município, 2 dígitos para o Distrito Administrativo, 2 dígitos para o Subdistrito e por fim 4 dígitos para o setor censitário;

- Latitude e longitude do setor censitário;
- CNES (Cadastro Nacional de Estabelecimentos de Saúde) em que o paciente fora internado;
- Latitude e longitude do CNES;
- Ano de internação do paciente;
- Diagnóstico (primário, secundário 1 e secundário 2) da internação: em código da CID-10;

Variáveis referentes às características demográficas coletadas pelos dados agregados por setores censitários do resultado do universo do Censo de 2010 do IBGE (2012):

- Setor censitário (código do IBGE);
- Tipo do setor: urbano ou rural;
- Nome do distrito administrativo;
- Área do setor censitário (em km²);
- População no setor censitário;
- População por gênero (masculino e feminino);
- População por faixa etária: crianças de 0 a 14 anos, jovens de 15 a 29 anos, adultos de 30 a 59 anos e idosos de 60 anos ou mais;
- População por gênero e faixa etária;
- Número de alfabetizados: pessoas alfabetizadas com 5 anos ou mais de idade;
- Renda média (em R\$): valor do rendimento nominal médio mensal das pessoas de 10 anos ou mais de idade (com e sem rendimento);
- Coordenadoria Regional de Saúde: Centro, Leste, Norte, Oeste, Sudeste, Sul;
- Número de domicílios com e sem arborização na quadra;
- Número de domicílios com e sem coleta de lixo na quadra;

Variáveis que foram criadas a partir de transformações das originais:

- Proporção de idosos no setor: quantidade de residentes com 60 anos ou mais dividida pela quantidade total de residentes em cada setor censitário;

- Proporção de alfabetizados no setor: quantidade de residentes alfabetizados dividida pela quantidade de pessoas residentes em cada setor censitário;
- Proporção de domicílios em quadras arborizadas: quantidade de domicílios que responderam estarem localizados em quadra com alguma arborização próxima dividida pela quantidade de domicílios entrevistados em cada setor censitário;
- Taxa de internação/população ou internação/população: dado um capítulo da CID-10, calcula-se quantos diagnósticos foram realizados somados os anos de 2015, 2018 e 2019 e divide-se pela quantidade de pessoas residentes no setor censitário;
- Taxa de internação padronizada: calcula-se, para cada capítulo da CID-10, a média e o desvio padrão das taxas de internação/população. Posteriormente a isso, transforma-se a taxa internação/população de cada setor subtraindo a média e dividindo pelo desvio padrão. Essa transformação foi utilizada no trabalho realizado por Martinelli (2023) e é utilizada no final do estudo;

5. Análise descritiva

Para o conjunto de dados trabalhado tem-se ao todo 2.431.045 diagnósticos de internações, contemplando 17.897 setores censitários no município de São Paulo referentes à residência dos pacientes. Entre esses setores, 312 (1,7%) não apresentaram informações demográficas no Censo de 2010 do IBGE, como é possível visualizar na Figura 1.

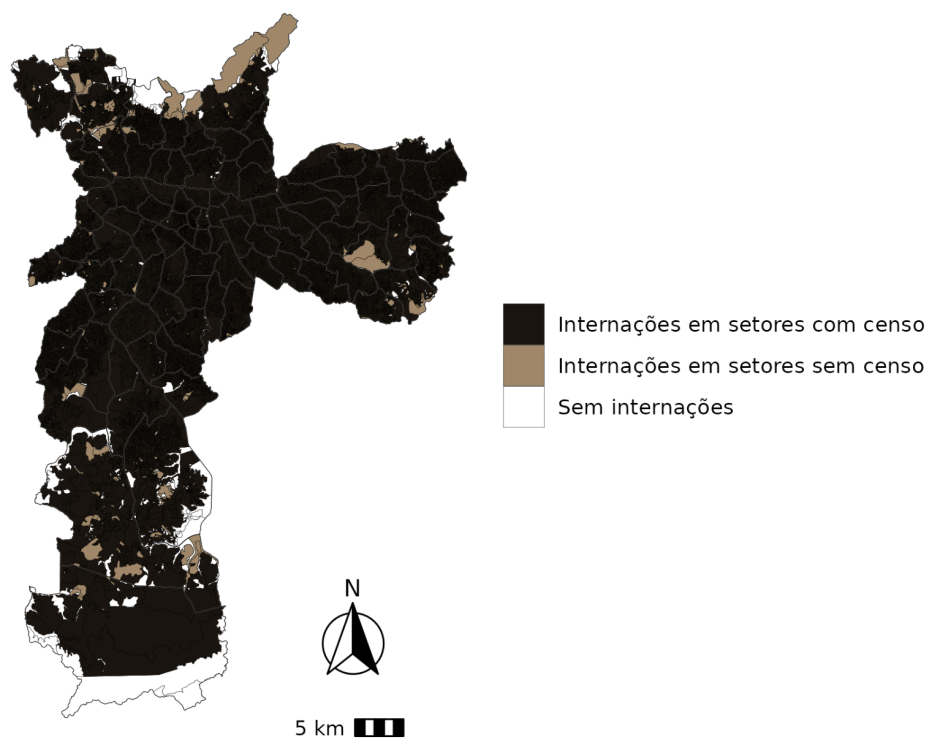


Figura 1: Mapa dos setores censitários e distritos administrativos do município de São Paulo que compõem os registros de internações mapeados na pesquisa.

O motivo pelo qual setores censitários não apresentaram informações demográficas do Censo de 2010 do IBGE pode ser explicado tendo em vista que são áreas de parques ou ZOE (Zonas de Ocupação Especial), como pode ser observado na Figura I do Anexo.

Os capítulos que compõem a CID-10 foram listados na Tabela 1 e serviram como base para análise da taxa de ocorrência das doenças por espaço geográfico. Nota-se que os nomes dos capítulos dessas doenças apresentados durante a pesquisa foram encurtados dos originais, podendo esses serem consultados na Tabela A.2.

Tabela 1: Capítulos da CID-10 (nomes encurtados).

Capítulo	Descrição	Intervalo de códigos
I	Doenças infecciosas e parasitárias	A00-B99

II	Neoplasmas (tumores)	C00-D48
III	Transtornos sanguíneos e imunitários	D50-D89
IV	Doenças endócrinas	E00-E90
V	Transtornos mentais e comportamentais	F00-F99
VI	Doenças do sistema nervoso	G00-G99
VII	Doenças do olho e anexos	H00-H59
VIII	Doenças do ouvido e da apófise mastóide	H60-H95
IX	Doenças do aparelho circulatório	I00-I99
X	Doenças do aparelho respiratório	J00-J99
XI	Doenças do aparelho digestivo	K00-K93
XII	Doenças da pele e do tecido subcutâneo	L00-L99
XIII	Sistema osteomuscular e tecido conjuntivo	M00-M99
XIV	Doenças do aparelho geniturinário	N00-N99
XV	Gravidez	O00-O99
XVI	Afecções perinatais	P00-P96
XVII	Malformações e anomalias cromossômicas	Q00-Q99
XVIII	Achados anormais	R00-R99
XIX	Lesões	S00-T98
XX	Causas externas de morbidade e de mortalidade	V01-Y98
XXI	Fatores influenciadores de saúde	Z00-Z99
**	Não preenchido	U04-U99

Também podem ser visualizados na Figura 2 os estabelecimentos de saúde do SUS que apresentaram ao menos uma internação nos registros dos pacientes mapeados, bem como a divisão das Coordenadorias Regionais de Saúde (CRS) do município por Distrito Administrativo na Figura 3. A lista de distritos por coordenadoria regional pode ser vista na Tabela A.1.

Nota-se que há poucos hospitais na porção sul do mapa, devido ao tipo da região ser predominantemente rural. Além disso, é possível observar que na Zona

Leste não há Hospitais Especializados para atendimento das internações vindas dessa região.

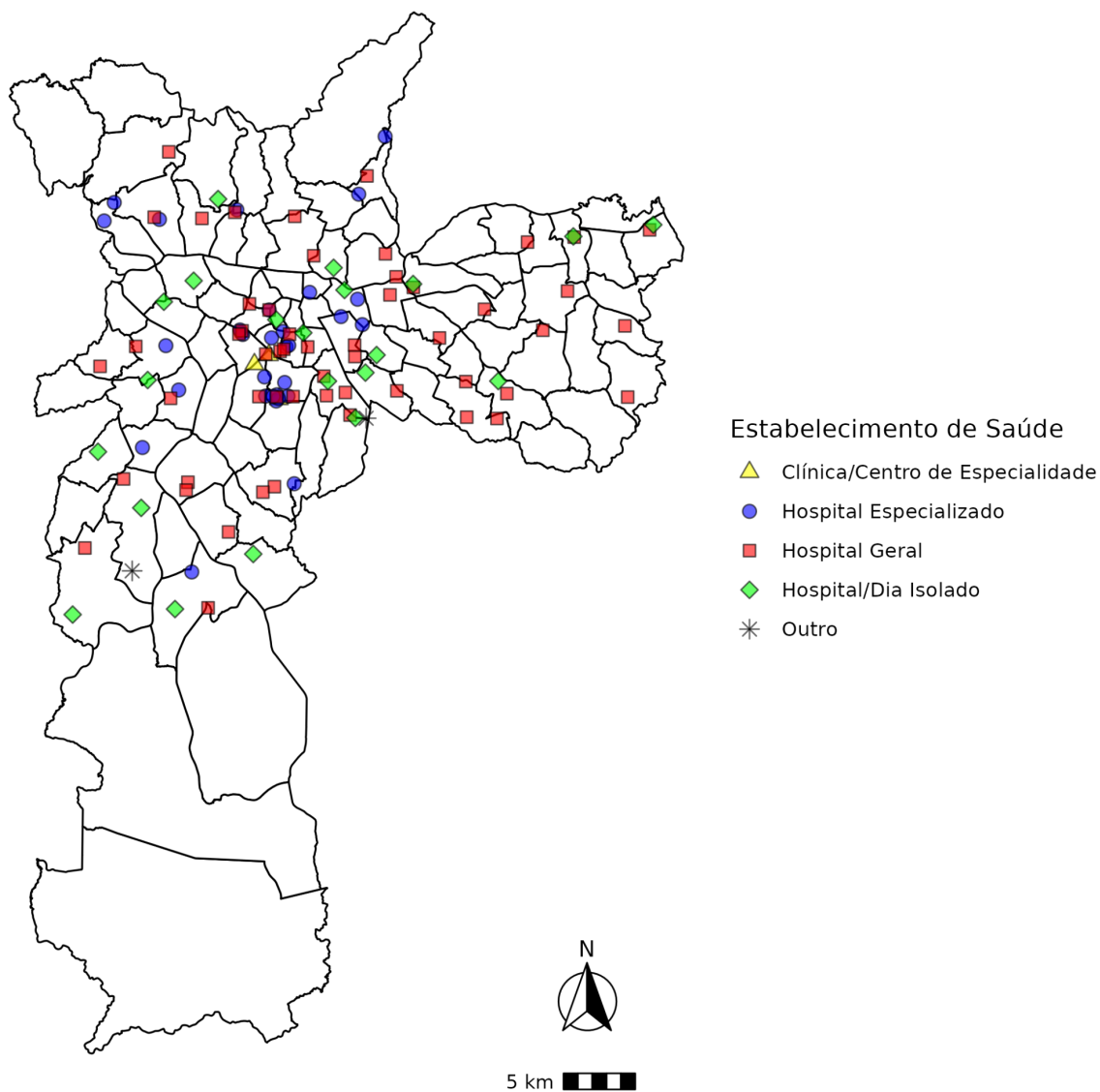


Figura 2: Estabelecimentos de saúde do SUS que atenderam ao menos uma interação dentro da base de dados utilizada na pesquisa.

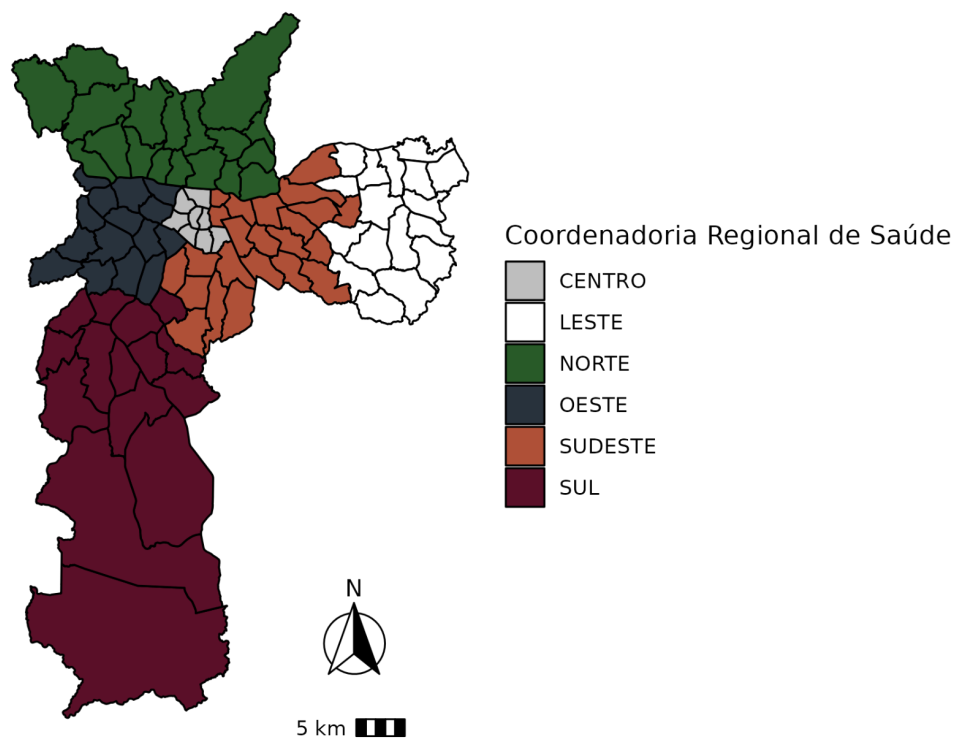


Figura 3: Coordenadorias Regionais de Saúde no município de São Paulo.

A Figura B.1 indica que a população média de cada setor censitário não ultrapassa 1.000 habitantes, contudo a densidade demográfica apresenta de forma mais clara a proporção de habitantes por área dentro de cada setor. Com essa última medida, é possível observar que a região Sul em sua porção mais extrema concentra o menor número de habitantes por quilômetro quadrado. Tal fato pode ser explicado pela abrangência territorial nessa região de áreas rurais, como verifica-se na Figura B.2.

5.1 Perfil demográfico dos setores

De acordo com a Figura 4 pode-se observar que há uma concentração maior do público mais velho (na faixa etária de idosos de 60 anos ou mais) na região central do mapa. Vale destacar que há uma baixa concentração de idosos na região ao Leste, por mais que a mesma seja bastante populosa.

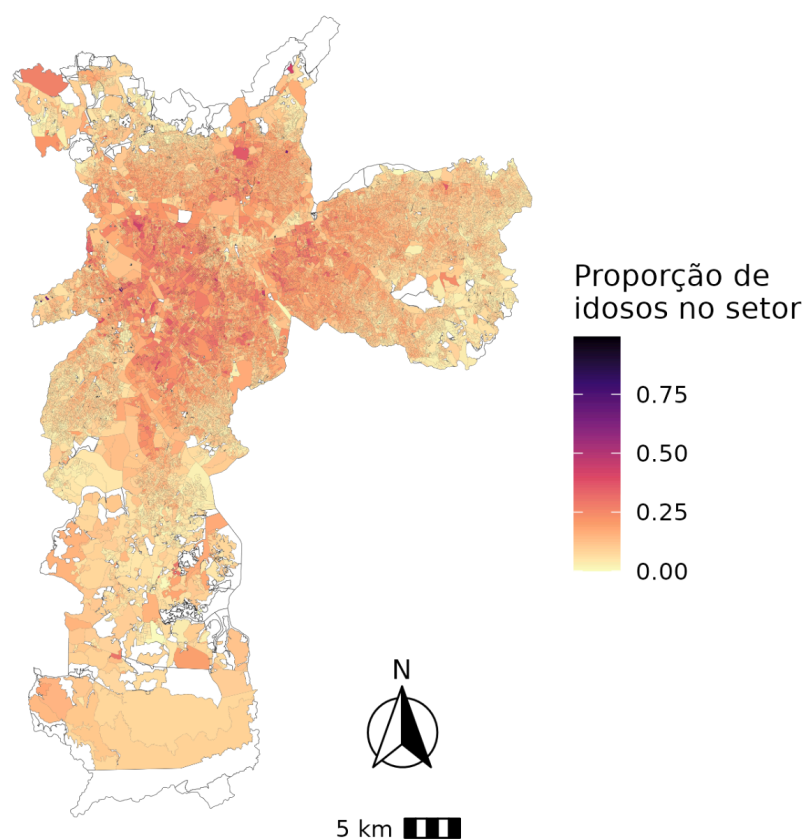


Figura 4: Proporção de idosos de 60 anos ou mais nos setores censitários do município.

Pela Figura 5 é possível observar que a taxa de alfabetização de pessoas com 5 anos ou mais de idade em relação à população total do setor censitário situa-se em torno de 89% para a grande maioria dos setores.

Já com a Figura 6 é possível observar a distribuição da renda média por pessoa com 10 anos ou mais e verifica-se que as maiores rendas se concentram na região Central do município englobando porção norte da região Sul, além de boa parte da região Oeste, Norte e Sudeste. Vale mencionar que na época da coleta dos dados do Censo de 2010 pelo IBGE, o salário mínimo situava-se em torno de R\$600,00 e, portanto, mais de 25% dos setores apresentaram renda média de até 1 salário mínimo da época, como é possível verificar pelo gráfico da distribuição acumulada da renda média nessa mesma figura mencionada anteriormente.

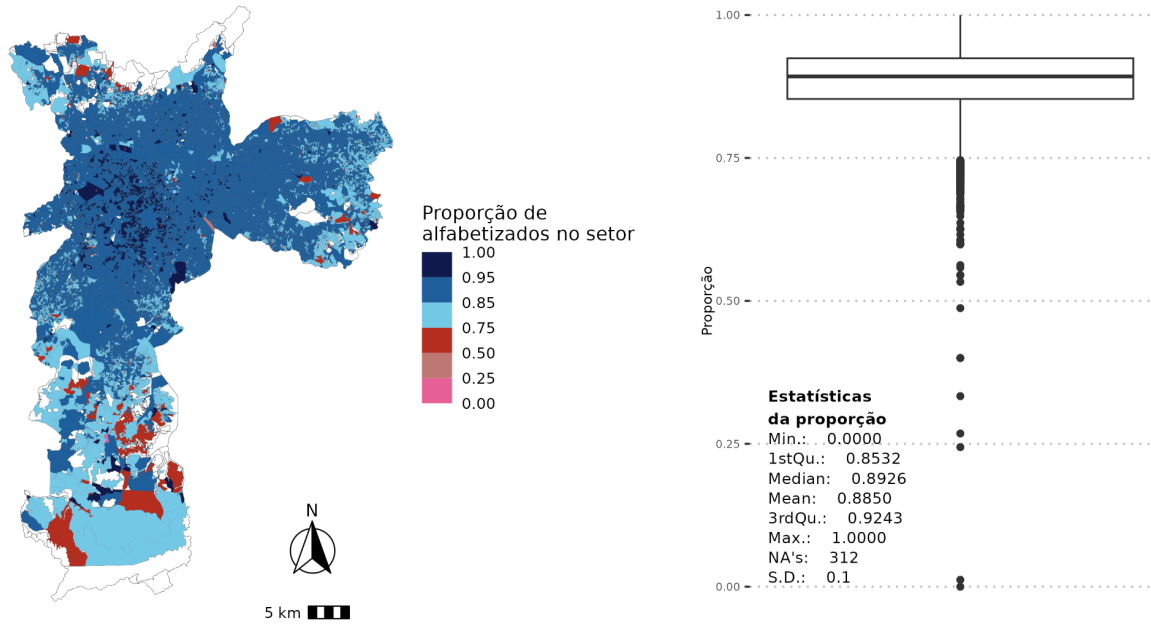


Figura 5: Proporção de alfabetizados no setor.

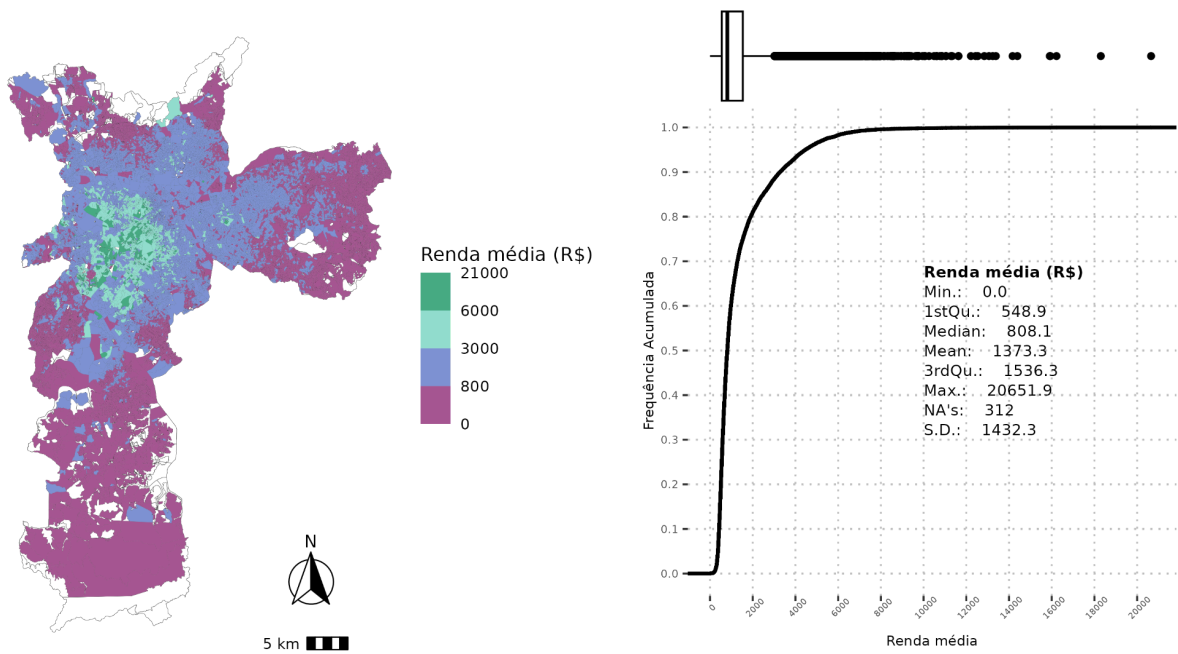


Figura 6: Renda média no setor.

5.2 Entorno dos domicílios nos setores

A Figura 7 mostra os setores com maiores e menores proporções de domicílios consultados que haviam arborização presente. Essa informação foi coletada pelo IBGE (2012) da seguinte forma: “Foi pesquisado se na face ou na sua face confrontante ou no canteiro central, existia arborização, ou seja, existia árvore ao longo da calçada/passeio e/ou em canteiro que divide pistas de um mesmo logradouro, mesmo que apenas em parte. Considerou-se também a arborização quando existente em logradouros sem pavimentação e/ou sem calçada/passeio”. Destaca-se a região da Zona Leste com baixas taxas de arborização na área.

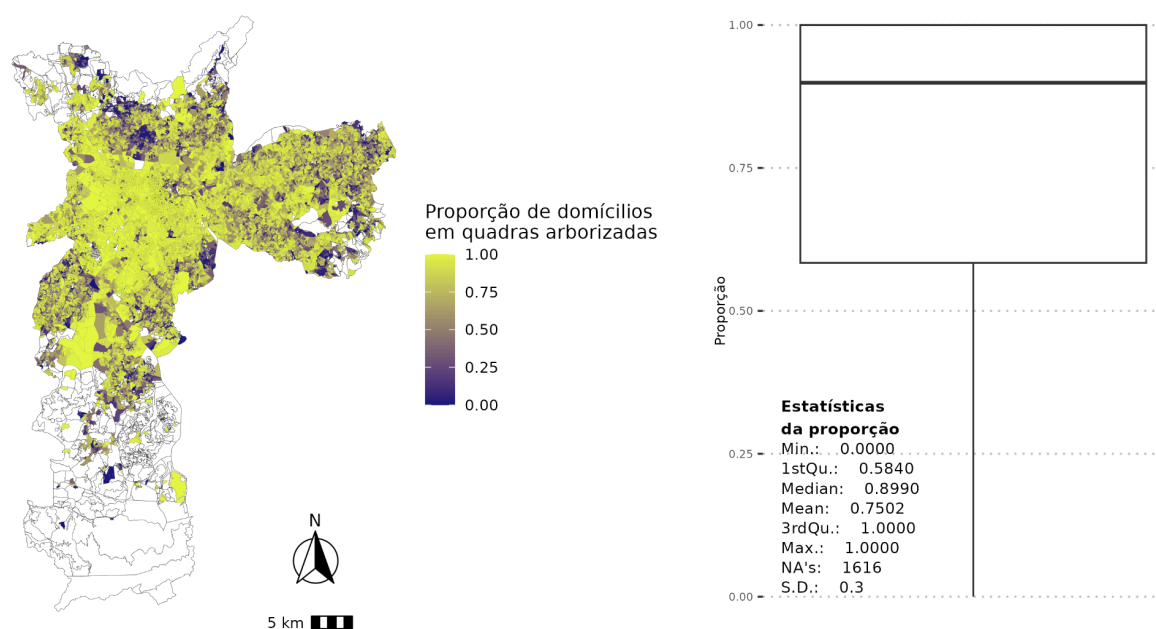


Figura 7: Proporção de domicílios no setor que responderam haver arborização próxima.

A Figura B.3 é parecida com a anterior, porém representa a proporção de domicílios consultados que indicaram haver coleta de lixo no entorno da residência. De acordo com o IBGE (2012), “Considerou-se quando, na face ou na sua confrontante,

existia local de depósito e acúmulo de lixo. A existência de caçamba de serviço de limpeza não foi considerada como lixo acumulado em via pública”.

Nota-se que essa última medida citada não possui uma discrepância significativa nos diferentes setores censitários, enquanto que a primeira relacionada à arborização demonstra uma dispersão clara.

5.3 Características dos capítulos das doenças

Com base na Figura 8 é possível observar a proporção de internações por capítulo da CID-10 nos anos de 2015, 2018 e 2019. Nota-se que doenças relacionadas a gravidez, parto e puerpério (Capítulo XV da CID-10) se apresentam com maior quantidade de diagnósticos representando 17,3%. Em seguida, doenças do aparelho circulatório, respiratório e digestivo compõem, nesta ordem, 11,3%, 8,6% e 8,1% do conjunto de registros dos diagnósticos coletados dos pacientes. Alguns capítulos de destaque são doenças relacionadas ao câncer (Neoplasmas) abrangendo 6,8% dos diagnósticos obtidos, além de algumas doenças infecciosas e parasitárias (Capítulo I da CID-10) representando 4% do universo dos dados diagnosticados.

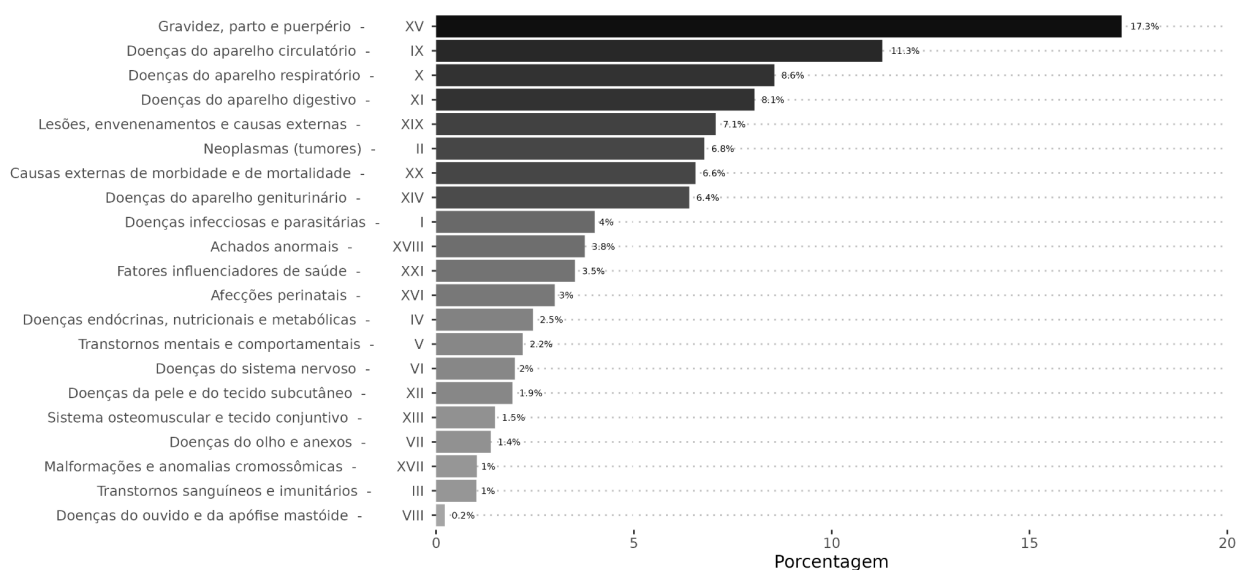


Figura 8: Porcentagem de diagnósticos coletados para cada Capítulo da CID-10.

Pela Figura 9 tem-se uma visão das faixas etárias mais afetadas por capítulo da CID-10, sendo possível observar que a faixa etária predominante para as internações do Capítulo XV (Gravidez, parto e puerpério) é de jovens (em torno de 25 anos). Além disso, doenças do aparelho circulatório afetam mais indivíduos acima de 50 anos, enquanto que malformações congênitas, deformidades e anomalias cromossômicas (Capítulo XVII da CID-10) abrange crianças de até 15 anos majoritariamente.

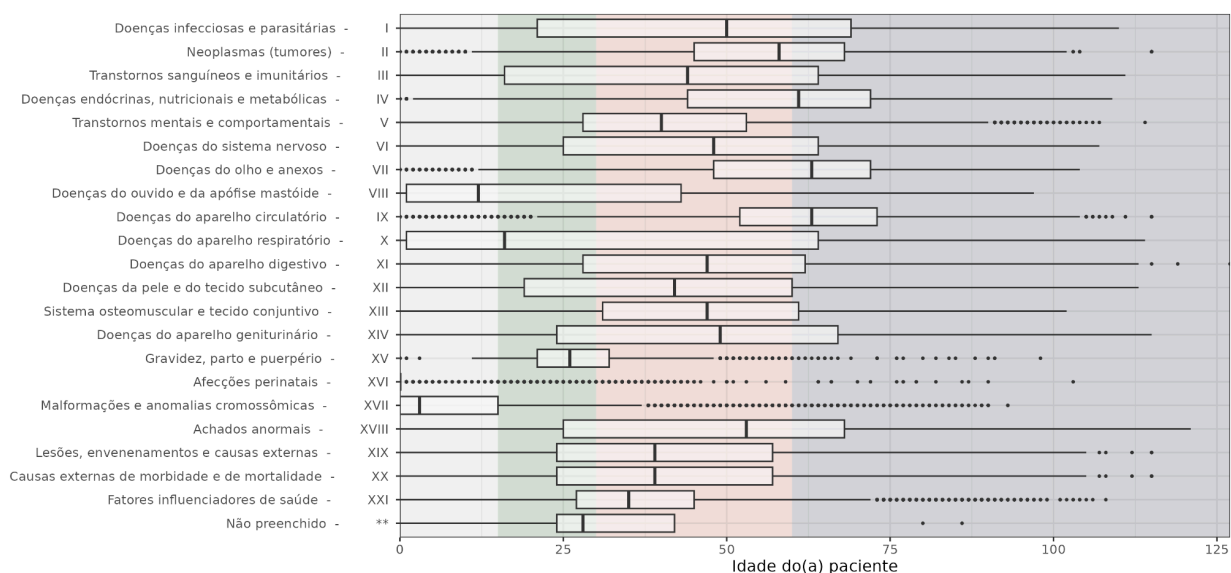


Figura 9: Box plots da idade dos pacientes internados por capítulo da CID-10. As faixas coloridas representam crianças, jovens, adultos e idosos respectivamente.

A Figura B.4 permite visualizar a quantidade de ocorrências por capítulo da CID-10 em cada ano da amostra. Vale ressaltar no entanto que essa variação anual não será levada em consideração no estudo, pois serão agregadas todas as internações dentro desse período. Além disso, o último capítulo da CID-10, foi suprimido de algumas análises pelo fato de apresentar apenas 9 diagnósticos nos três anos coletados.

5.4 Distribuição das doenças por setor

O mapa da distribuição de cada capítulo de doenças da CID-10 pelos setores censitários do município pode ser visualizado na Figura B.5. Nele, é possível observar que há setores em que a população residente registrada no Censo do IBGE de 2010 é menor que o número de diagnósticos de internações somados dos anos de 2015, 2018 e 2019 e tal comportamento pode ser explicado pelo fato de serem áreas com população muito pequena. Além disso, é possível verificar que todos os capítulos da CID-10 apresentam a maioria dos setores com a taxa internação/população menor que 30%. Isso norteará os estudos inferenciais que serão descritos a seguir.

6. Análise inferencial

Para iniciar a análise inferencial, resolveu-se truncar os valores da taxa de internações/população acima de 30% em 30%, uma vez que mais de 95% dos valores dessa eram inferiores a 30% e os demais 5% eram muito superiores a esse valor. Esse corte foi realizado, por capítulo da CID-10. O número de setores afetados pelo truncamento em cada capítulo varia de 1 a 138, mas as proporções de diagnósticos (número de diagnósticos nos setores afetados sobre o total de diagnósticos) são muito pequenas (valor máximo igual a 2,16%), como mostra a Tabela 2.

Tabela 2: Informações sobre os setores censitários que sofreram truncamento no valor da Taxa internações/população por capítulo da CID-10.

Capítulo	Nome encurtado	Quantidade de setores afetados	Total de diagnósticos	Quantidade de diagnósticos nos setores afetados e proporção sobre o total
I	Doenças infecciosas e parasitárias	23	97.017	352 (0,36%)
II	Neoplasmas (tumores)	29	164.199	856 (0,52%)
III	Transtornos sanguíneos e imunitários	7	24.728	16 (0,06%)
IV	Doenças endócrinas	14	59.202	136 (0,23%)
V	Transtornos mentais e comportamentais	17	53.017	1.147 (2,16%)

VI	Doenças do sistema nervoso	10	48.168	120 (0,25%)
VII	Doenças do olho e anexos	6	33.353	35 (0,10%)
VIII	Doenças do ouvido e da apófise mastóide	1	5.413	3 (0,06%)
IX	Doenças do aparelho circulatório	59	272.773	1.838 (0,67%)
X	Doenças do aparelho respiratório	51	206.533	1.527 (0,74%)
XI	Doenças do aparelho digestivo	51	194.666	1.087 (0,56%)
XII	Doenças da pele e do tecido subcutâneo	12	46.622	131 (0,28%)
XIII	Sistema osteomuscular e tecido conjuntivo	7	35.907	39 (0,11%)
XIV	Doenças do aparelho geniturinário	33	154.666	489 (0,32%)
XV	Gravidez	138	418.418	8.552 (2,04%)
XVI	Afecções perinatais	20	72.329	150 (0,21%)
XVII	Malformações e anomalias cromossômicas	8	24.875	54 (0,22%)
XVIII	Achados anormais	33	90.826	445 (0,49%)
XIX	Lesões	47	170.730	2.347 (1,37%)
XX	Causas externas de morbidade e de mortalidade	41	158.460	1.732 (1,09%)
XXI	Fatores influenciadores de saúde	29	84.672	536 (0,63%)

6.1 Metodologia

Foi utilizado o seguinte processo:

- Foram separados os setores com maior população (acima de 750 residentes – representando 5.265 setores censitários) dos com menor população;
- Com base nas taxas de internação/população dos diferentes capítulos da CID-10 dos setores mais populosos, aplicou-se o método de agrupamento hierárquico usando a matriz de distâncias euclidianas e a ligação de Ward como medida de dissimilaridade.
- Definidos os clusters pelo método hierárquico, foram calculados os centróides das taxas de internação/população de cada cluster.

- Utilizando o algoritmo de K-médias os setores menos populosos foram alocados aos clusters obtidos pelo método hierárquico.

Vale ressaltar que o algoritmo hierárquico demanda grande espaço de armazenamento. Assim, seria inviável utilizar todos os 17.585 setores censitários de uma única vez para rodar o agrupamento hierárquico.

O método de Ward foi escolhido por carregar a métrica de Soma de Quadrados entre os grupos e dentro dos grupos para definir as ligações de dissimilaridades.

6.2 Definição do número de clusters

Para definição do número de clusters, usou-se o gráfico *Elbow Plot*, apresentado na Figura 10. A partir dessa figura, optou-se por utilizar 5 clusters para agrupamento dos setores.

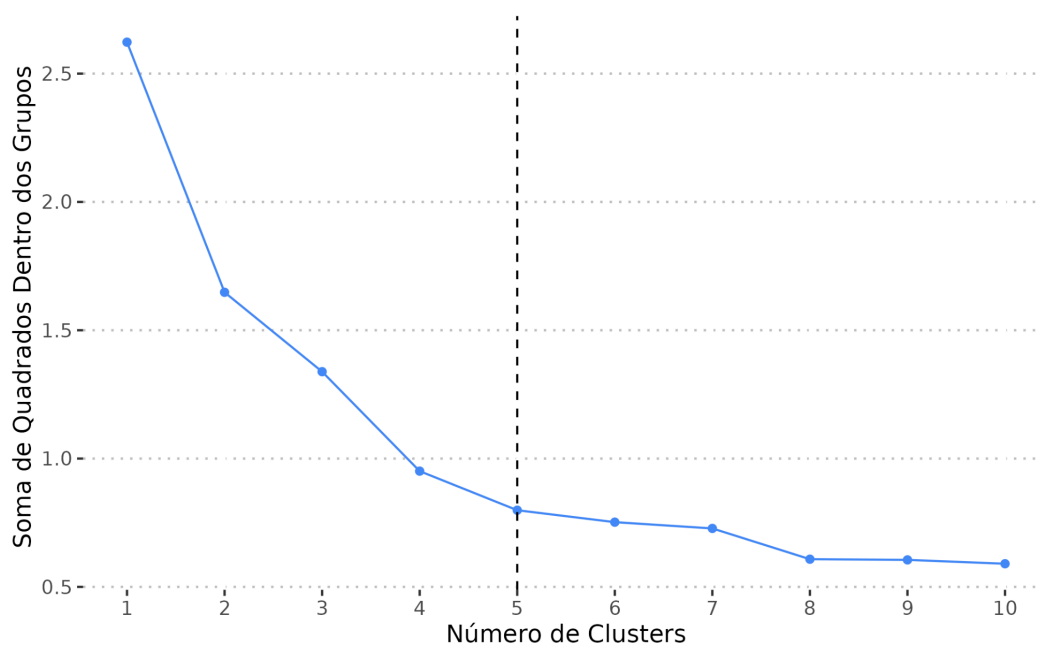


Figura 10: *Elbow plot* para definição do número de clusters.

6.3 Visualização da distribuição geográfica dos clusters

Finalizado o agrupamento hierárquico com os setores mais populosos e posteriormente o encaixe dos setores menos populosos sobre os clusters definidos, chegou-se à seguinte visualização da Figura 11.

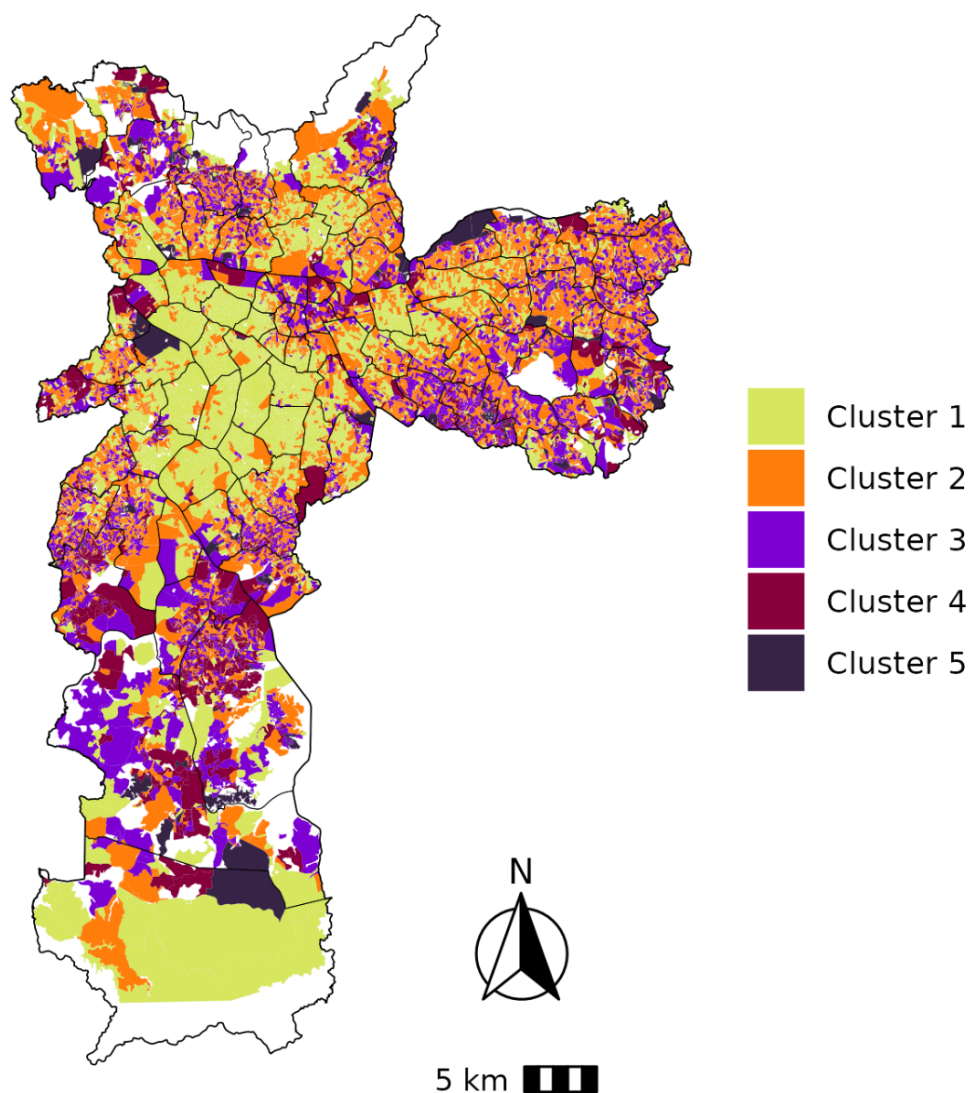


Figura 11: Distribuição geográfica dos clusters baseados nas Taxas de internação/população de cada capítulo da CID-10.

A partir dos clusters definidos, é possível visualizar a distribuição das taxas de internação/população para cada capítulo de doenças da CID-10 nos gráficos *box plot* apresentados na Figura 12.

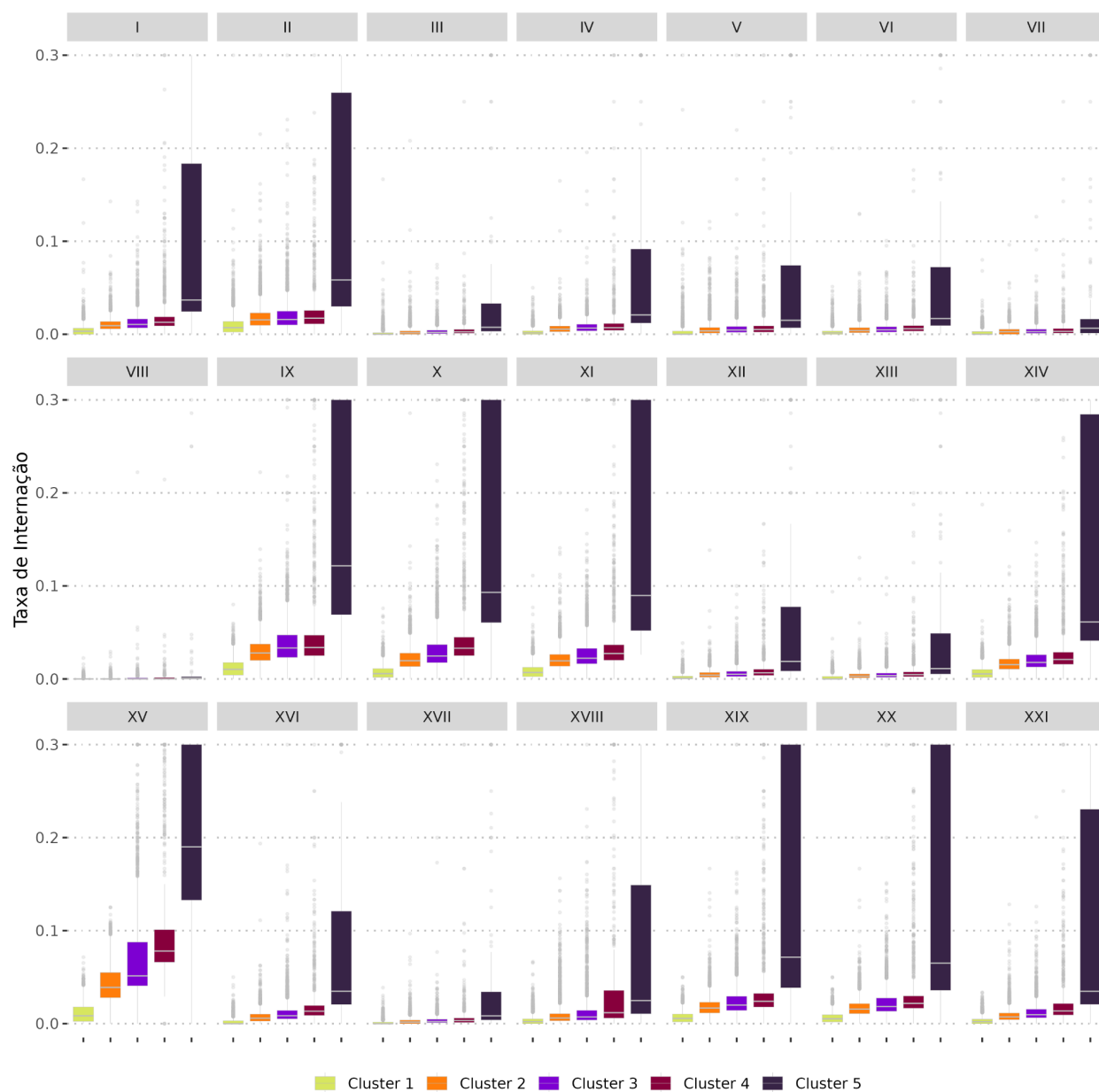


Figura 12: *Box plots* da Taxa de internações/população por capítulo da CID-10 em cada cluster.

Observa-se que o cluster 1 até o 5 apresenta, nessa ordem, as menores para as maiores taxas de internação das doenças nos setores censitários. Vale ressaltar porém que as análises de quantidade e taxa de internações estão baseadas exclusivamente em pacientes que foram atendidos pelo SUS. Logo, os setores censitários do cluster 1, por exemplo, não necessariamente são aqueles em que ocorrem menos internações na população residente, pois podem estar ocorrendo internações em hospitais privados que não foram contemplados nessa pesquisa.

Nota-se então que o Cluster 1 abrange boa parte da região Central do município, possuindo territórios da Zona Oeste, e da parte superior da Zona Sul.

Na Figura II do Anexo, é possível identificar a localização do distrito administrativo do Morumbi, e pela Figura III observam-se algumas partes desse distrito ocupadas por moradores de uma comunidade. Essas partes pertencem ao cluster 3, enquanto o restante do distrito pertence ao cluster 1. Tal característica explica a discrepância de estilo de vida dentro do mesmo distrito.

O cluster 5 mostra as maiores taxas, sendo que os setores censitário que sofreram truncamento nas taxas pertencem a ele. Portanto, pode ser considerado um cluster com valores mais anômalos quando comparado aos demais.

6.4 Analisando diferença significativa entre os clusters

Para testar a hipótese de que os 5 clusters são significativamente distintos, utiliza-se o método de Análise de Variância Multivariada Permutacional (ANDERSON, 2001) que retorna a distribuição da estatística F em um cenário não paramétrico.

O valor da estatística F observada foi de 342,5, resultando em um valor-p menor que 0,001%, levando a rejeitar a hipótese de que os clusters não são diferentes. Na Figura 13, o valor 342,5 localiza-se no extremo da cauda direita da distribuição F sob a hipótese nula.

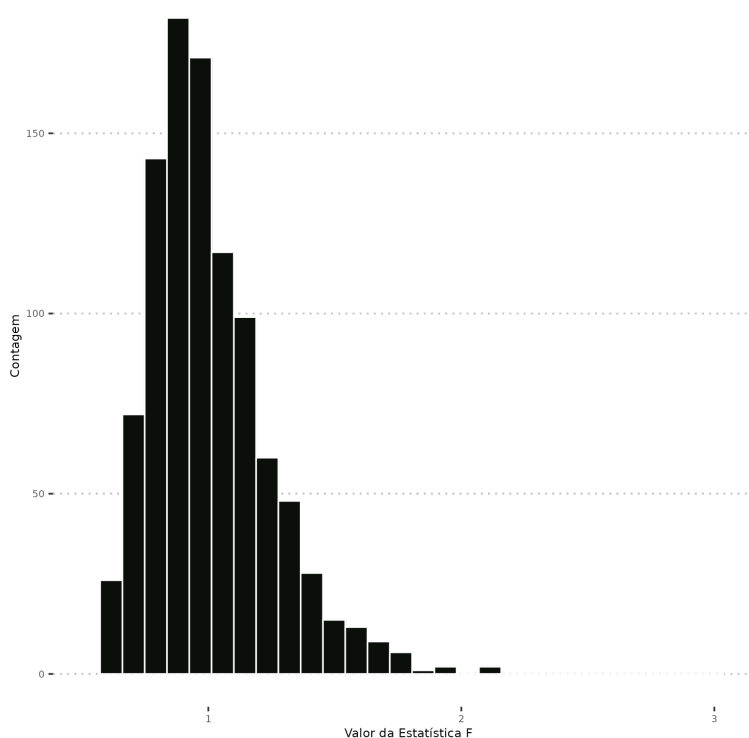


Figura 13: Distribuição empírica da estatística F a partir do método de PERMANOVA sob hipótese nula de que os clusters não são distintos.

6.5 Aplicando a metodologia usando taxa padronizada

Pode-se padronizar a taxa de internações/população de todos os capítulos de doenças por cada setor censitário, tornando a taxa passível de ser interpretada de uma forma que demonstre a presença mais ou menos acentuada de determinado capítulo de doenças dentro do setor. Essa padronização elimina o viés devido à ordem de grandeza das taxas, assim como de suas variabilidades, por capítulo da CID-10.

Aplicando-se o mesmo procedimento descrito anteriormente nas taxas padronizadas, chega-se nos resultados observados nas Figuras 14 e 15, agora com 4 clusters. A definição dos 4 clusters nesse caso é justificada pelo *Elbow Plot* da Figura B.6, bem como para comparar com os resultados do estudo de Martinelli (2023), destacados nas Figuras IV e V do Anexo.

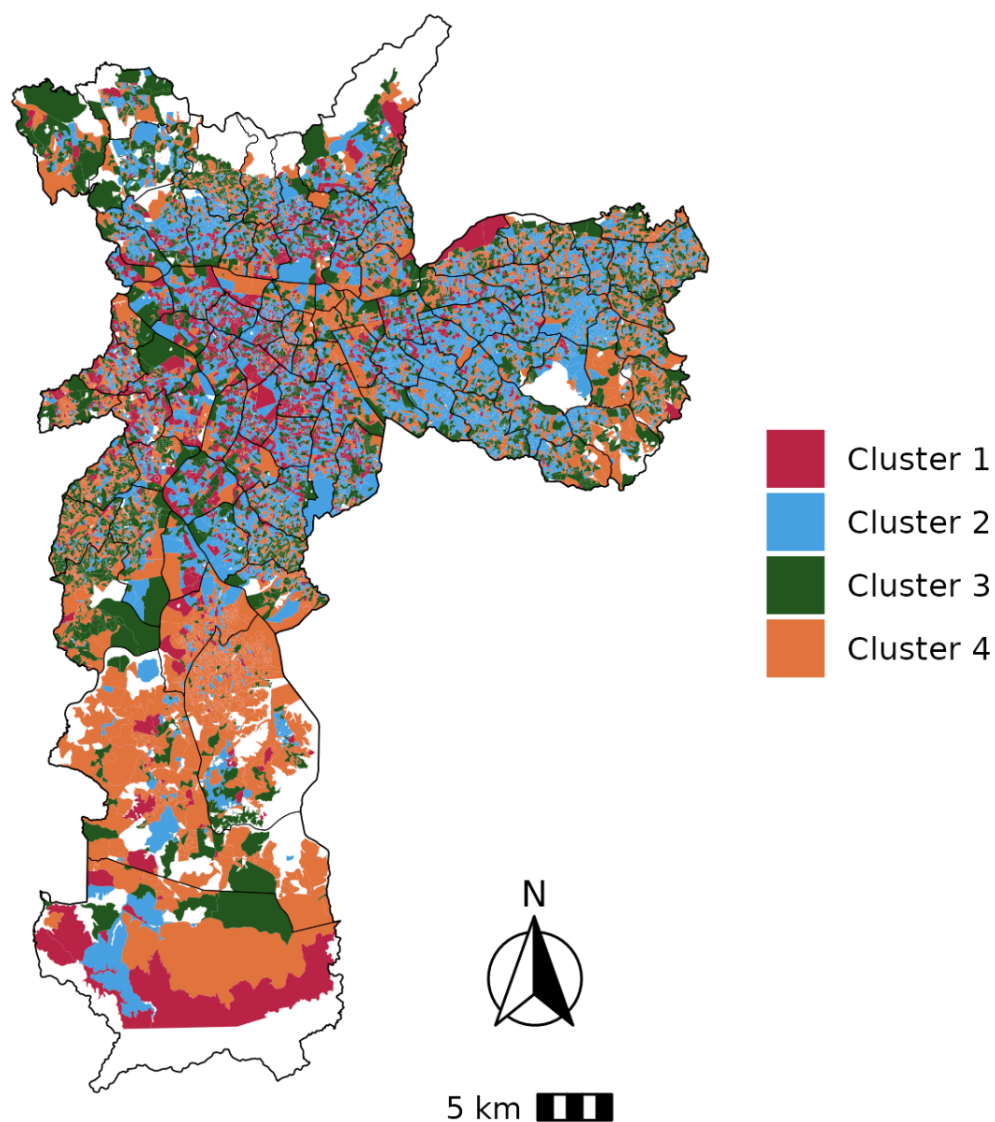


Figura 13: Clusters criados a partir da Taxa de internação padronizada.

Pode-se, portanto, fazer um comparativo desse agrupamento obtido com o resultado do trabalho de Martinelli (2023) que analisa o perfil de mortalidade a partir das doenças catalogadas na CID-10, por distrito administrativo. Esse gráfico pode ser checado na Figura IV do Anexo, junto com o perfil desses clusters baseados na taxa de mortalidade apresentado na Figura V.



Figura 14: Perfil dos clusters criados a partir da média da Taxa de internação padronizada.

O perfil dos clusters, utilizando as médias das taxas de internação padronizadas (Figura 14) destaca algumas doenças que caracterizam bem o cluster como doenças do aparelho circulatório, que se sobressaem no cluster 2, e doenças relacionadas a Neoplasmas e tumores ligadas com o cluster 1. Além disso, percebe-se as internações por doenças relacionadas à gravidez, parto e puerpério sendo mais presentes no cluster 4 – caracterizado por regiões mais afastadas do centro do município e dos estabelecimentos de saúde do SUS.

7. Conclusões

Com base nas análises realizadas, chega-se às seguintes conclusões:

- As regiões com maior renda média no município de São Paulo apresentam as menores taxas de internação/população levando em consideração que foram analisadas dentro do Sistema Único de Saúde (SUS); tal fato leva a suspeitar que os pacientes residentes nessas regiões tendem a utilizar mais serviços de saúde privados.
- Internações causadas por doenças relacionadas ao câncer (Neoplasmas) estão mais concentradas na área central, mais nobre, do município, enquanto as relacionadas à doenças do aparelho circulatório estão mais presentes na Zona Leste;
- Os setores com presença de arborização apresentam menores taxas de internação padronizada de doenças do aparelho respiratório (Capítulo X da CID-10);
- É possível identificar dentro do mesmo distrito administrativo diferentes perfis de internação por doenças da CID-10. Exemplo, seria o caso do Morumbi que possui setores nos 4 clusters;
- As comunidades espalhadas pelo município apresentam as maiores taxas de internação por doenças relacionadas à causas externas (Capítulos XIX e XX da CID-10).

APÊNDICE A

Tabela

Tabela A.1: Distritos Administrativos por Coordenadoria Regional de Saúde no município de São Paulo

Coordenadoria Regional de Saúde	Distritos Administrativos
Centro	Bom Retiro, Santa Cecília, Consolação, República, Sé, Bela Vista, Liberdade, Cambuci
Leste	Ermelino Matarazzo, Vila Jacuí, São Miguel, Jardim Helena, Vila Curuçá, Ponte Rasa, Itaquera, Lajeado, Itaim Paulista, Cidade Líder, Parque do Carmo, José Bonifácio, Guaianases, Cidade Tiradentes, São Mateus, Iguatemi, São Rafael
Norte	Perus, Anhanguera, Jaraguá, Pirituba, São Domingos, Brasilândia, Freguesia do Ó, Limão, Cachoeirinha, Mandaqui, Casa Verde, Santana, Tucuruvi, Vila Guilherme, Vila Maria, Vila Medeiros, Jaçanã, Tremembé
Oeste	Jaguara, Vila Leopoldina, Jaguaré, Lapa, Barra Funda, Perdizes, Alto de Pinheiros, Pinheiros, Jardim Paulista, Butantã, Rio Pequeno, Raposo Tavares, Vila Sônia, Morumbi, Itaim Bibi
Sudeste	Cangaíba, Penha, Vila Matilde, Carrão, Tatuapé, Belém, Pari, Brás, Mooca, Água Rasa, Vila Formosa, Artur Alvim, Aricanduva, Vila Prudente, São Lucas, Sapopemba, Ipiranga, Sacomã, Cursino, Vila Mariana, Saúde, Jabaquara, Moema
Sul	Campo Limpo, Vila Andrade, Capão Redondo, Jardim São Luís, Jardim Ângela, Socorro, Santo Amaro, Campo Belo, Campo Grande, Cidade Ademar, Pedreira, Cidade Dutra, Grajaú, Parelheiros, Marsilac

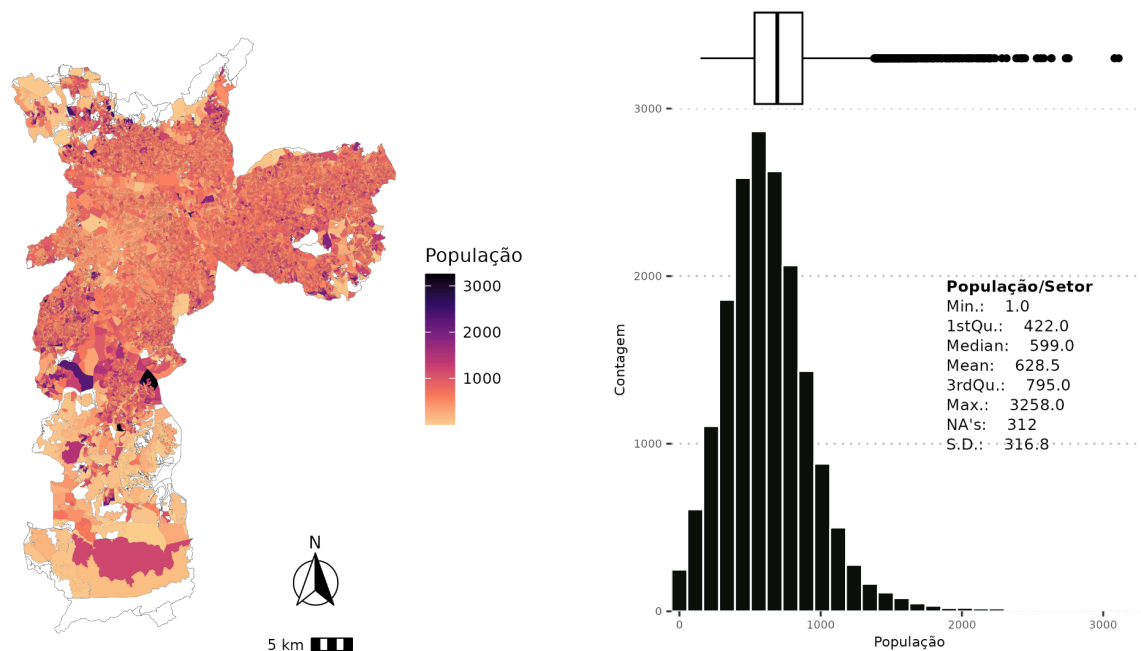
Tabela A.2: Capítulos da CID-10 (nomes originais).

Capítulo	Descrição	Intervalo de códigos
I	Algumas doenças infecciosas e parasitárias	A00-B99
II	Neoplasmas [tumores]	C00-D48
III	Doenças do sangue e dos órgãos hematopoéticos e alguns transtornos imunitários	D50-D89
IV	Doenças endócrinas, nutricionais e metabólicas	E00-E90
V	Transtornos mentais e comportamentais	F00-F99
VI	Doenças do sistema nervoso	G00-G99
VII	Doenças do olho e anexos	H00-H59
VIII	Doenças do ouvido e da apófise mastóide	H60-H95
IX	Doenças do aparelho circulatório	I00-I99
X	Doenças do aparelho respiratório	J00-J99
XI	Doenças do aparelho digestivo	K00-K93
XII	Doenças da pele e do tecido subcutâneo	L00-L99

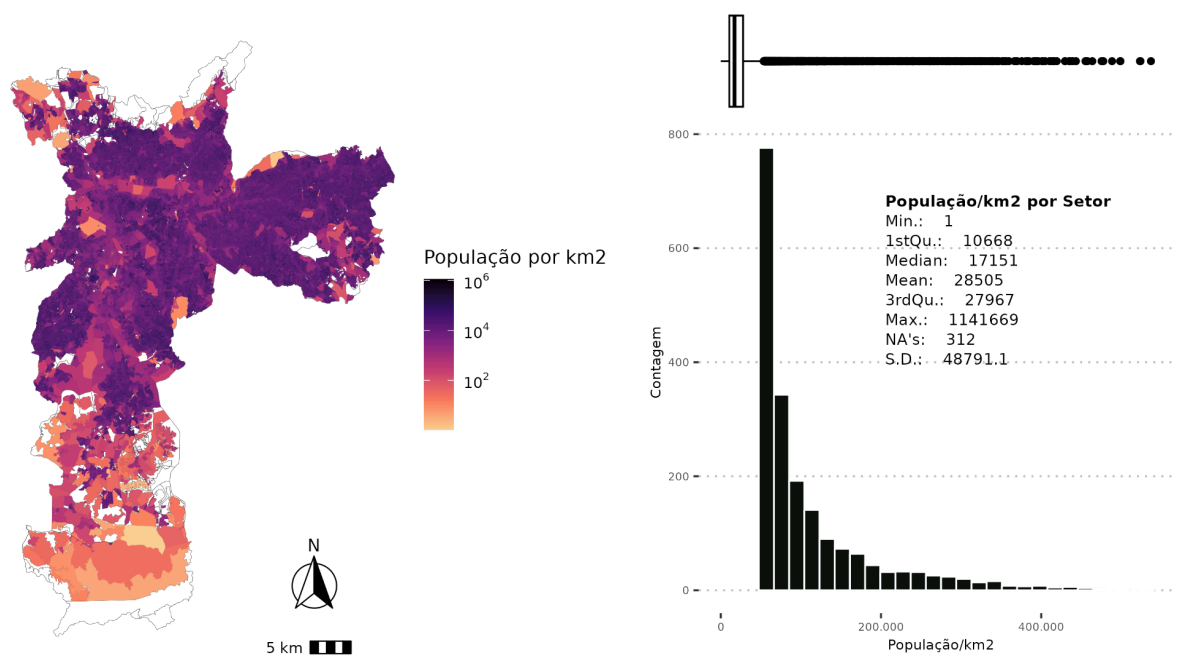
XIII	Doenças do sistema osteomuscular e do tecido conjuntivo	M00-M99
XIV	Doenças do aparelho geniturinário	N00-N99
XV	Gravidez, parto e puerpério	O00-O99
XVI	Algumas afecções originadas no período perinatal	P00-P96
XVII	Malformações congênitas, deformidades e anomalias cromossômicas	Q00-Q99
XVIII	Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte	R00-R99
XIX	Lesões, envenenamentos e algumas outras conseqüências de causas externas	S00-T98
XX	Causas externas de morbidade e de mortalidade	V01-Y98
XXI	Fatores que influenciam o estado de saúde e o contato com os serviços de saúde	Z00-Z99
**	CID 10ª Revisão não disponível ou não preenchido ou inválido	U04-U99

APÊNDICE B

Figuras



(a)



(b)

Figura B.1: (a) Mapa e distribuição da população absoluta por setor censitário; (b) Mapa e distribuição da população por área em quilômetro quadrado de cada setor censitário.

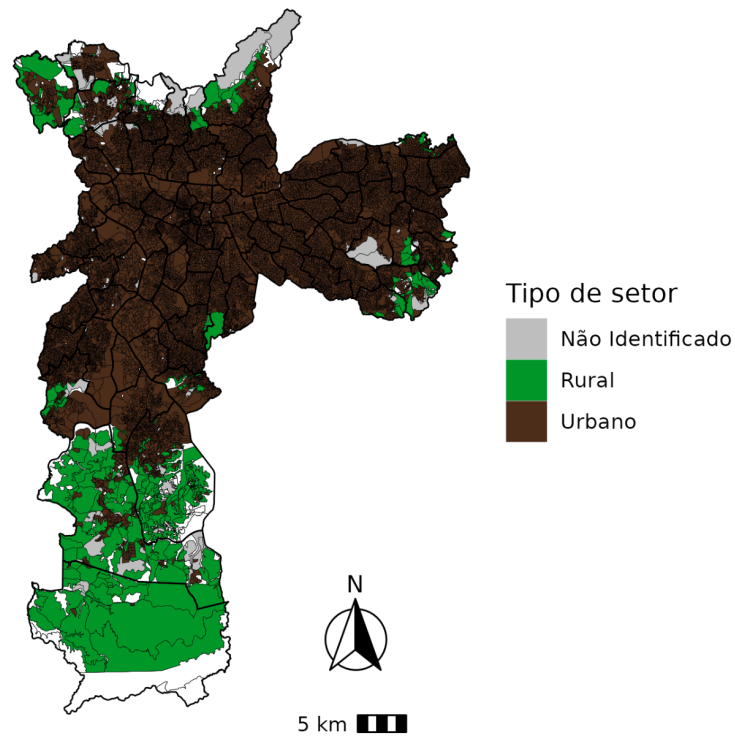


Figura B.2: Tipo de setor por distrito administrativo.

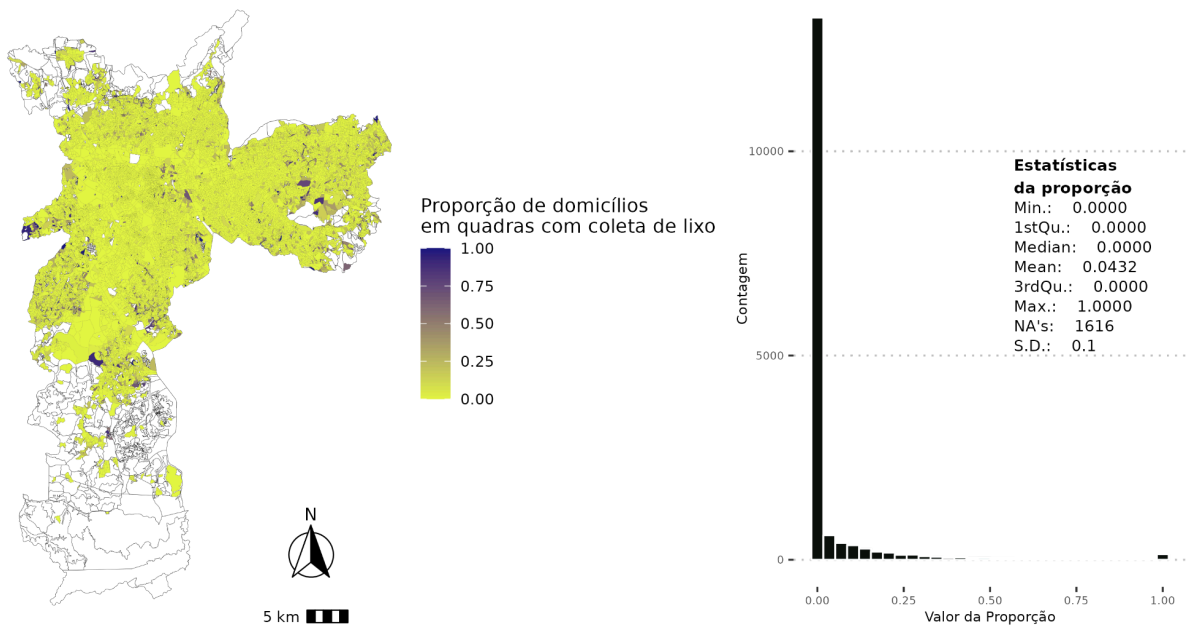
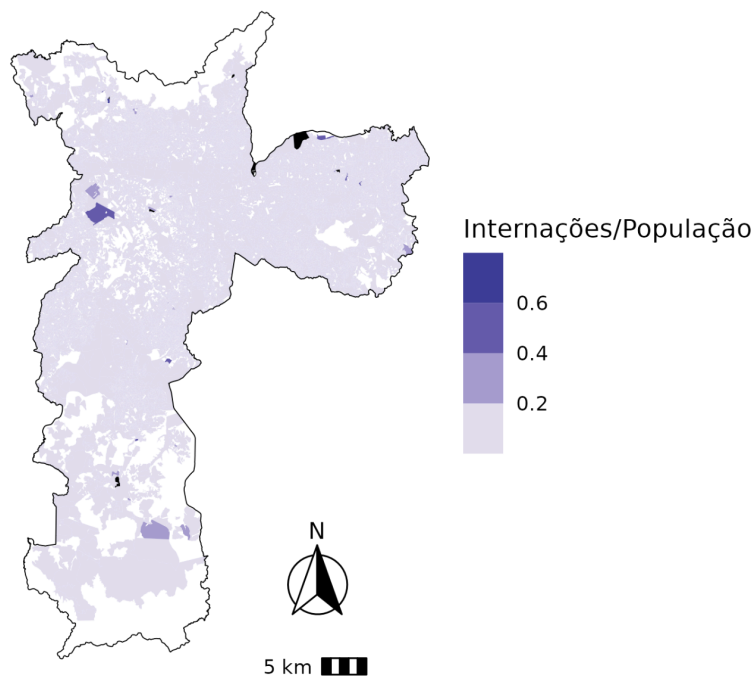


Figura B.3: Mapa e distribuição da proporção de domicílios no setor censitário que responderam haver coleta de lixo.



Figura B.4: Número de ocorrências de internações por capítulo da CID-10 para cada ano.

Doenças infecciosas e parasitárias - I



Neoplasmas (tumores) - II

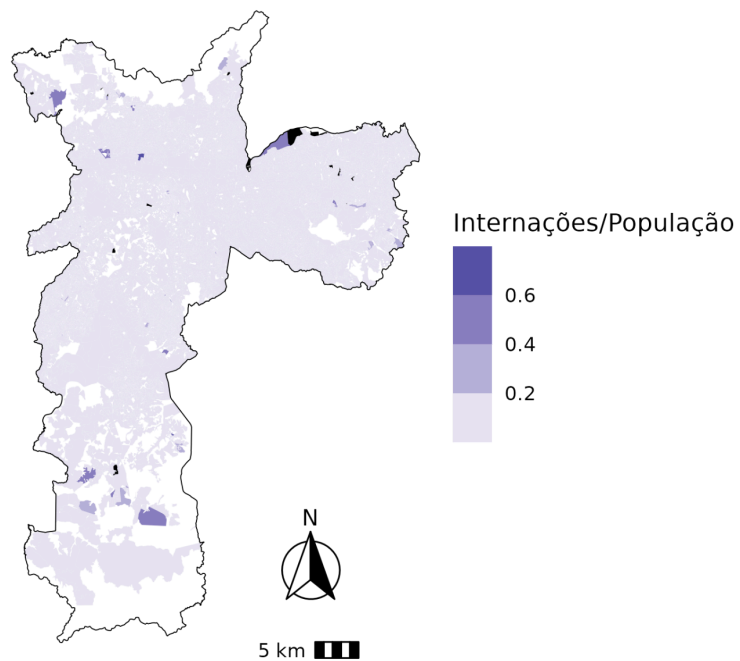
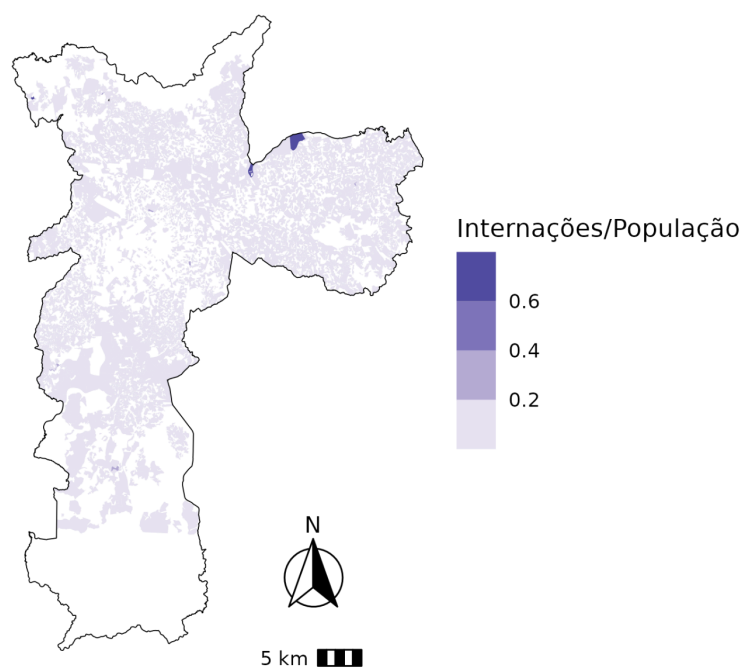


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 1/11)

Transtornos sanguíneos e imunitários - III



Doenças endócrinas, nutricionais e metabólicas - IV

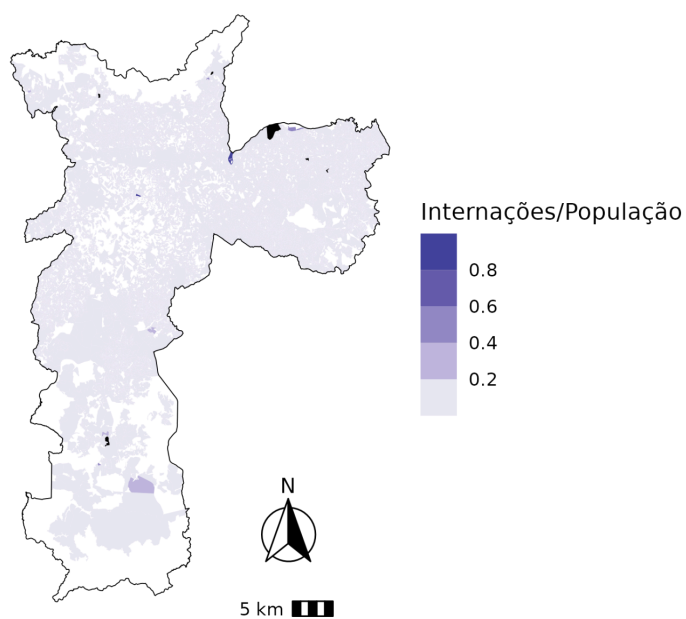
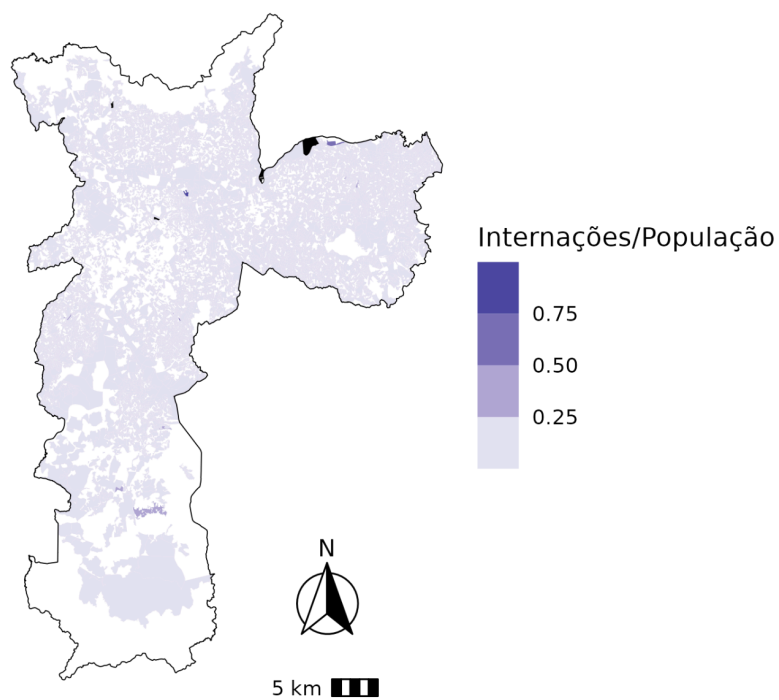


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 2/11)

Transtornos mentais e comportamentais - V



Doenças do sistema nervoso - VI

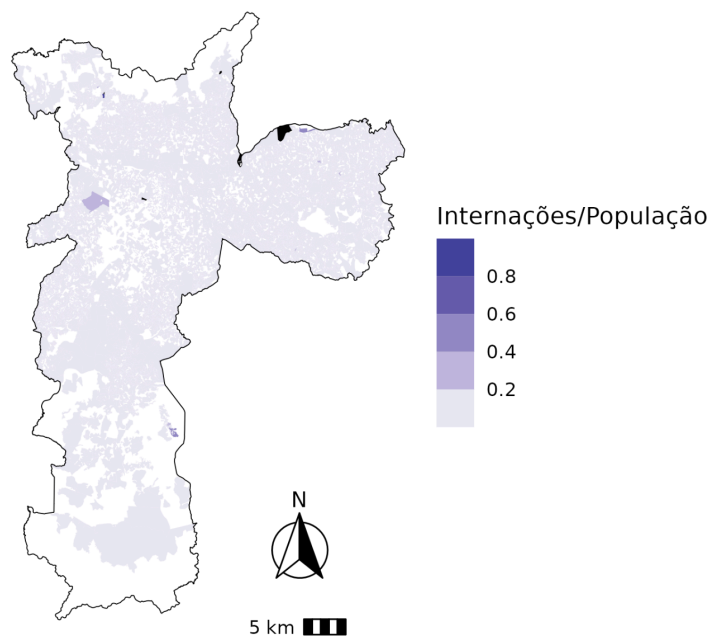
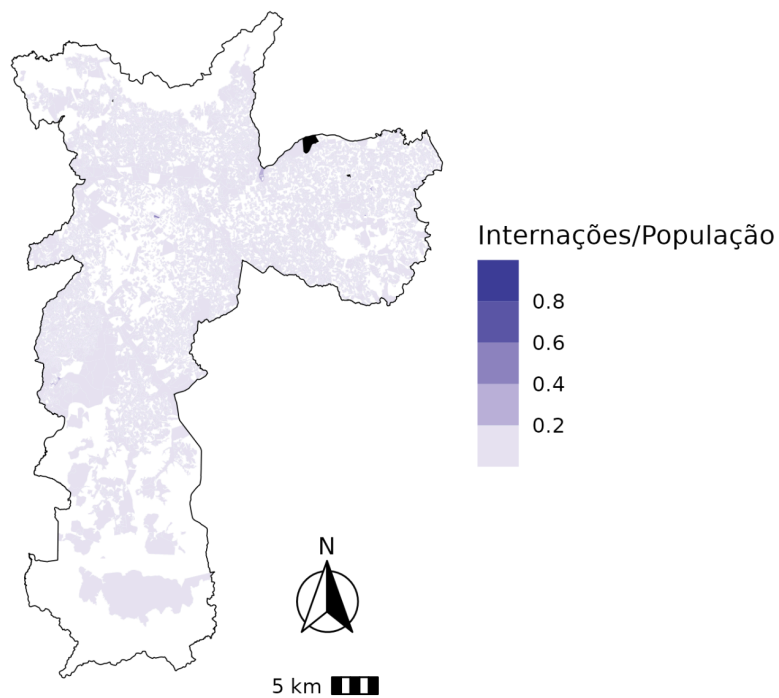


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 3/11)

Doenças do olho e anexos - VII



Doenças do ouvido e da apófise mastóide - VIII

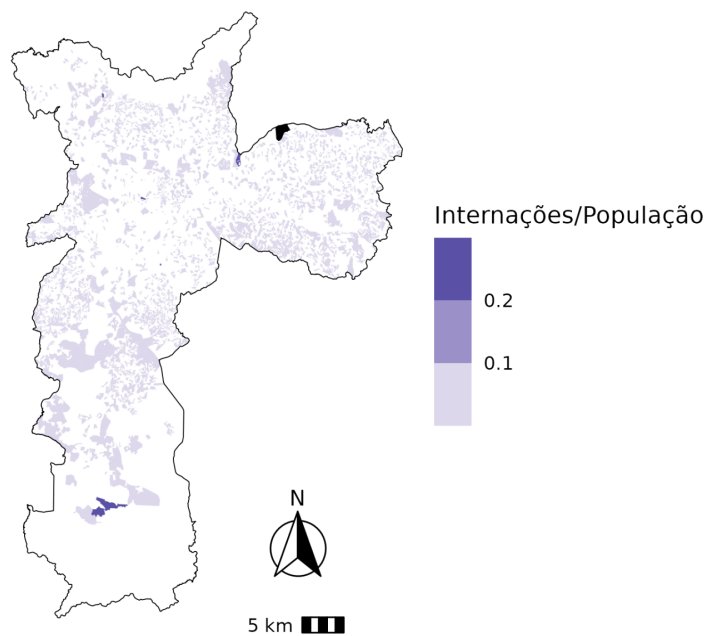
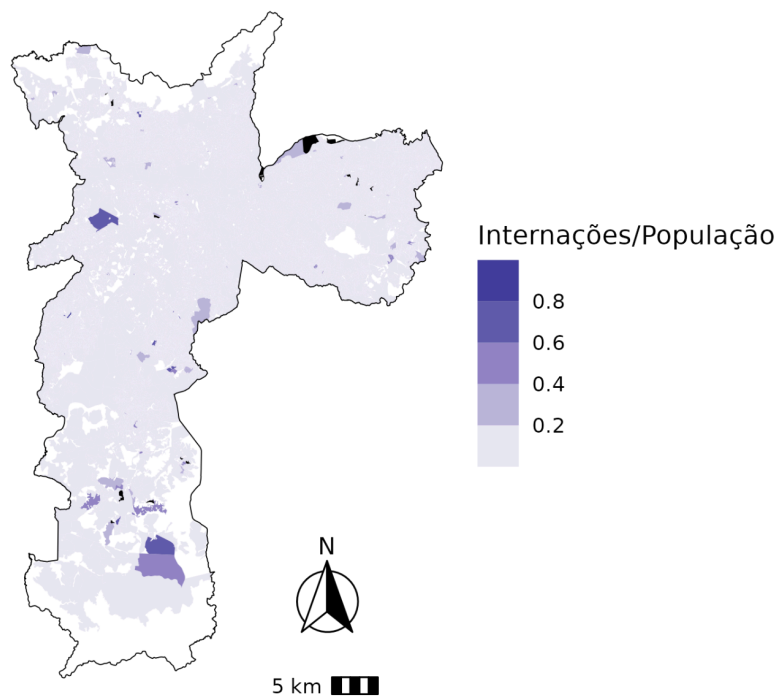


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 4/11)

Doenças do aparelho circulatório - IX



Doenças do aparelho respiratório - X

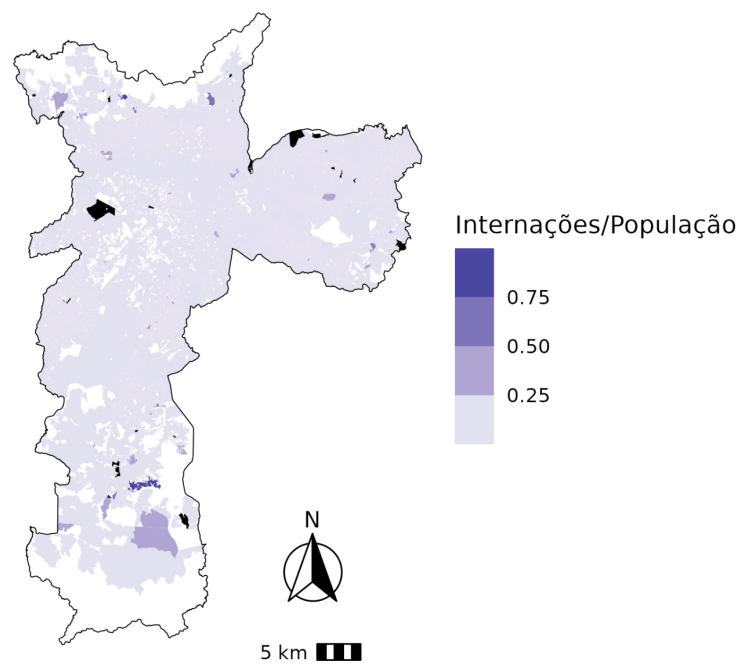
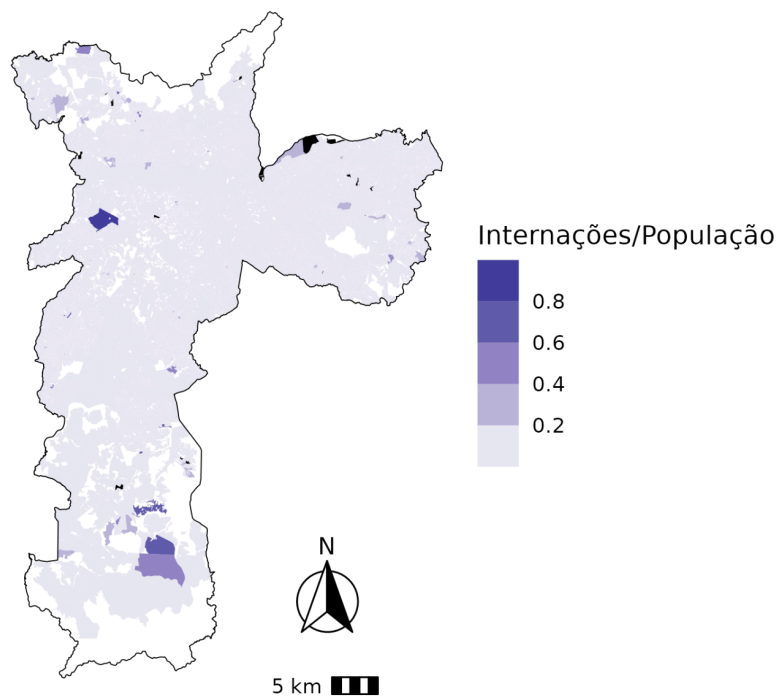


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 5/11)

Doenças do aparelho digestivo - XI



Doenças da pele e do tecido subcutâneo - XII

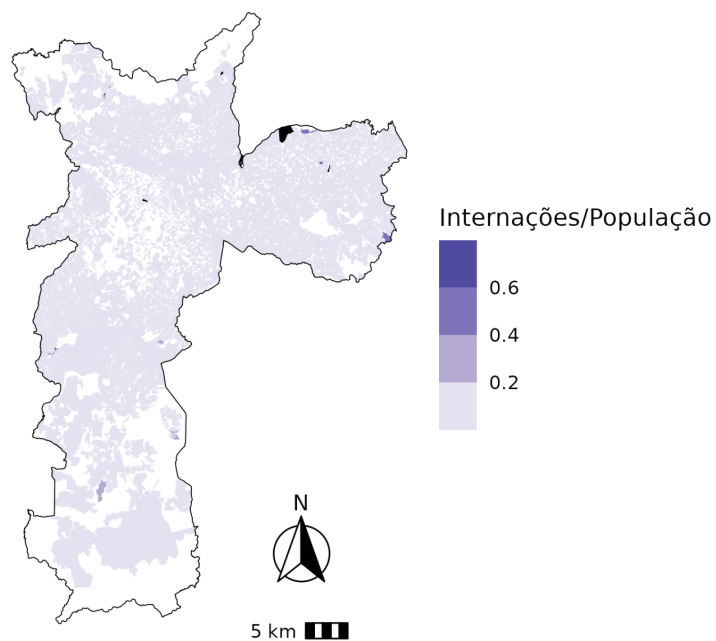
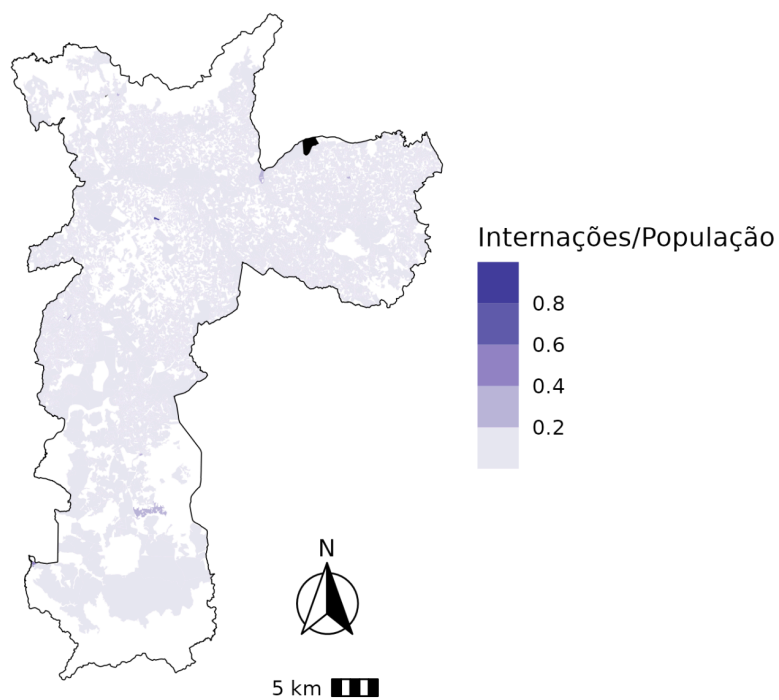


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 6/11)

Sistema osteomuscular e tecido conjuntivo - XIII



Doenças do aparelho geniturinário - XIV

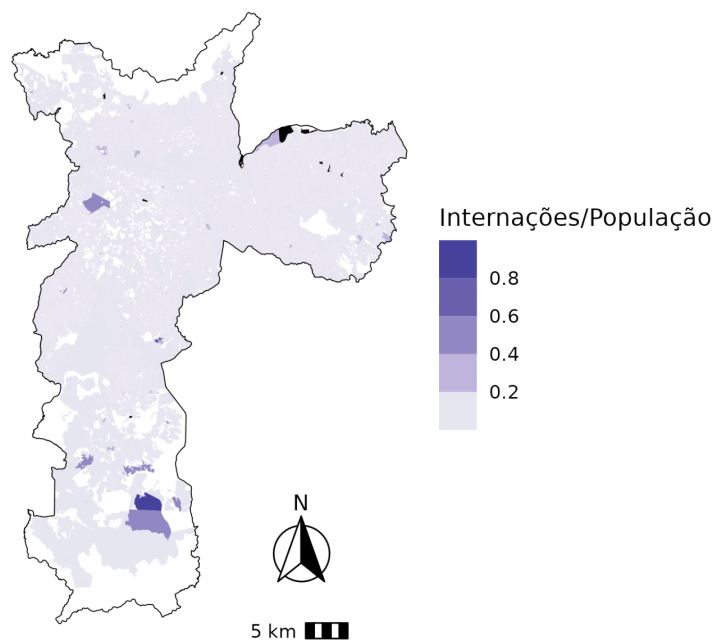
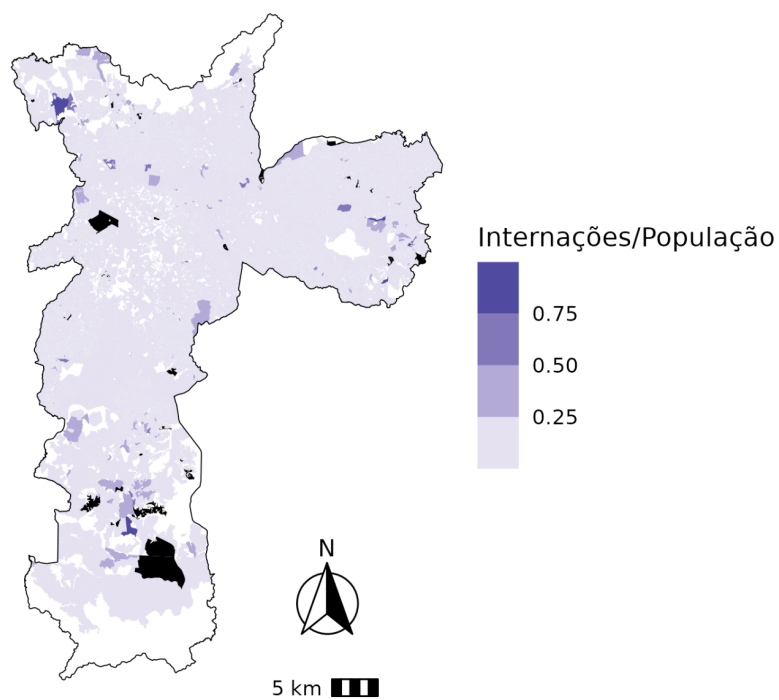


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 7/11)

Gravidez, parto e puerpério - XV



Afecções perinatais - XVI

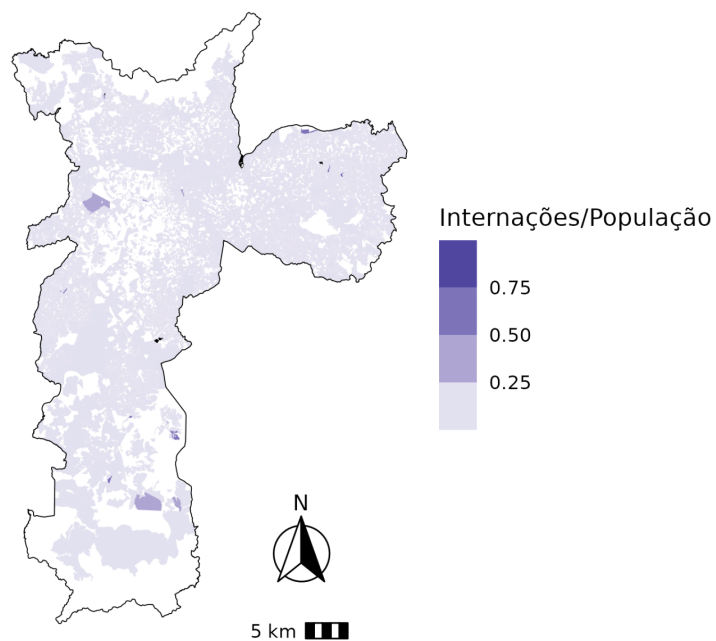
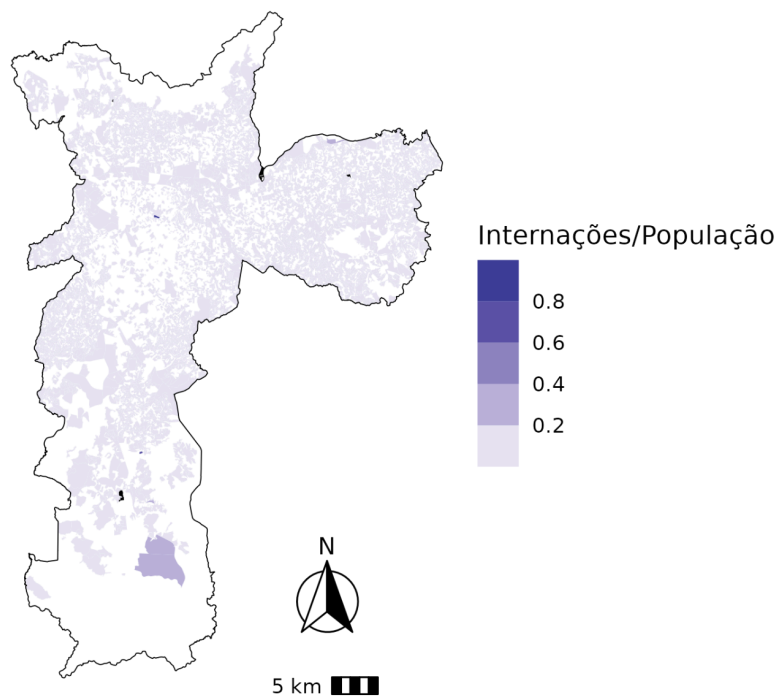


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 8/11)

Malformações e anomalias cromossômicas - XVII



Achados anormais - XVIII

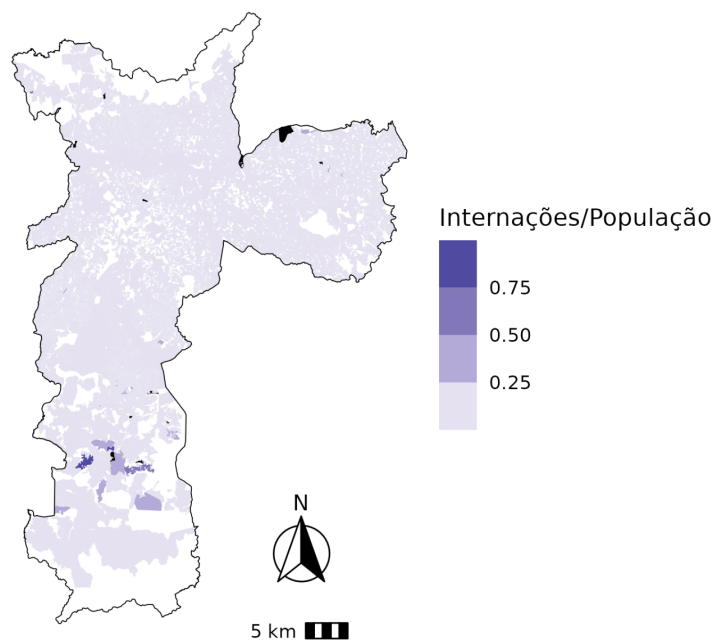
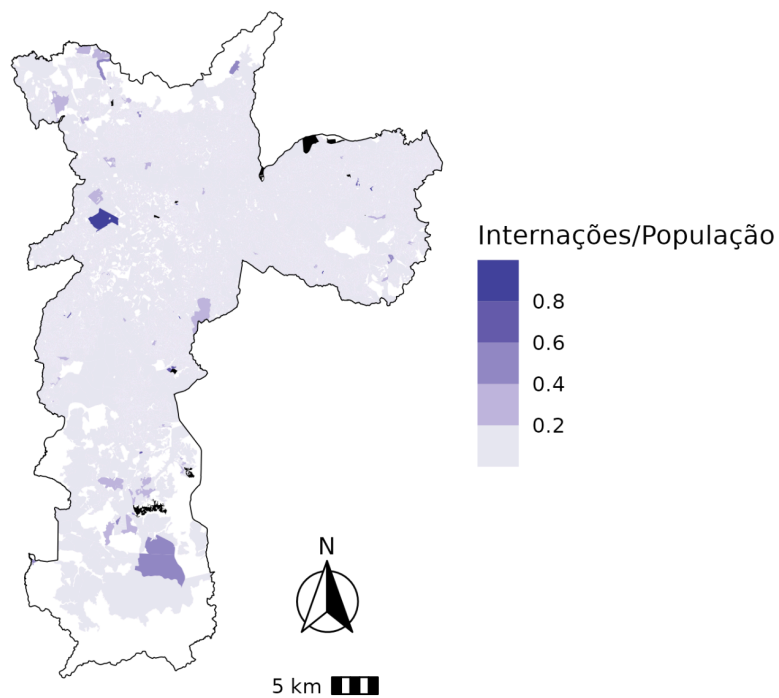


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 9/11)

Lesões, envenenamentos e causas externas - XIX



Causas externas de morbidade e de mortalidade - XX

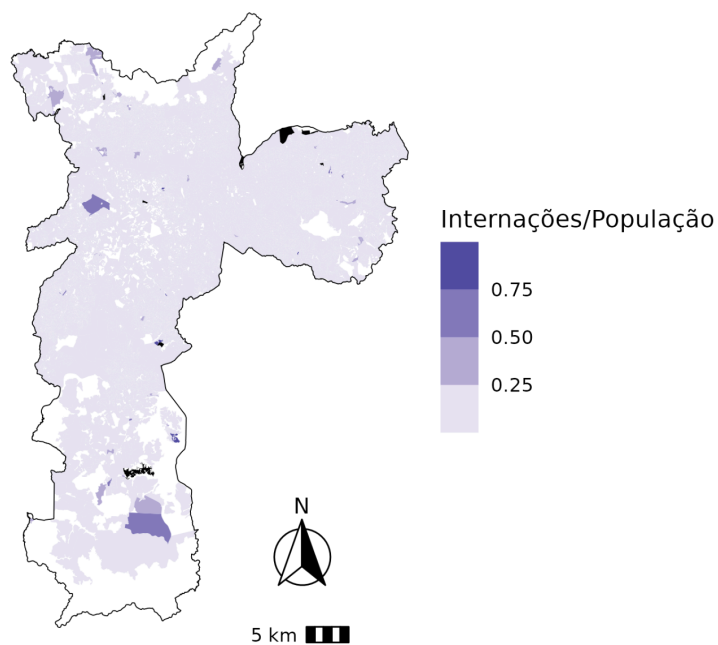


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 10/11)

Fatores influenciadores de saúde - XXI

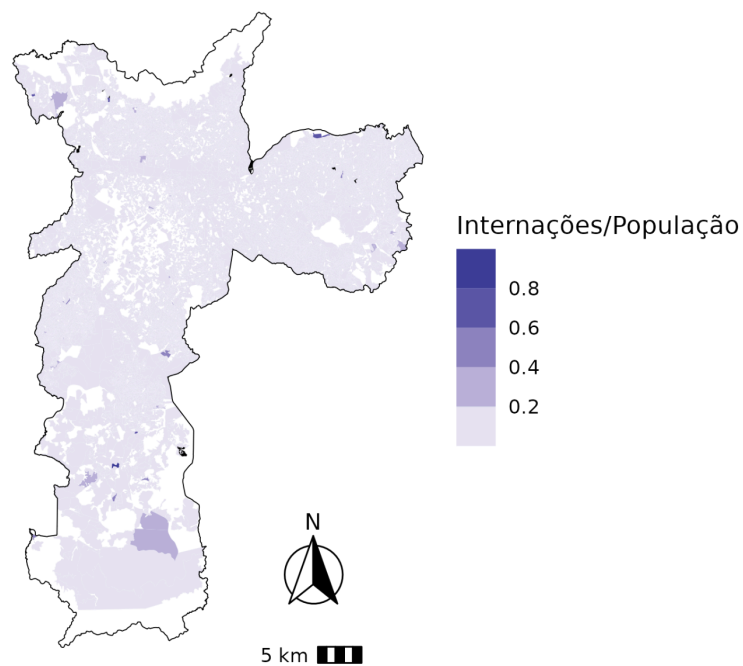


Figura B.5: Taxa internação/população por capítulo da CID-10. Regiões em preto apresentaram taxas acima de 1. (Página 11/11)

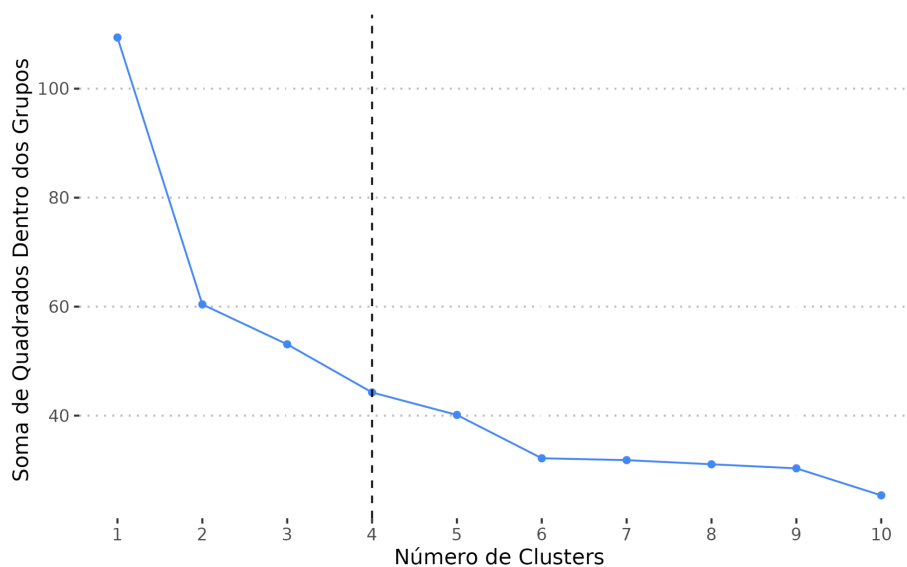


Figura B.6: *Elbow plot* para definição do número de clusters no agrupamento dos setores com base na Taxa de internações padronizadas por setor.

ANEXO

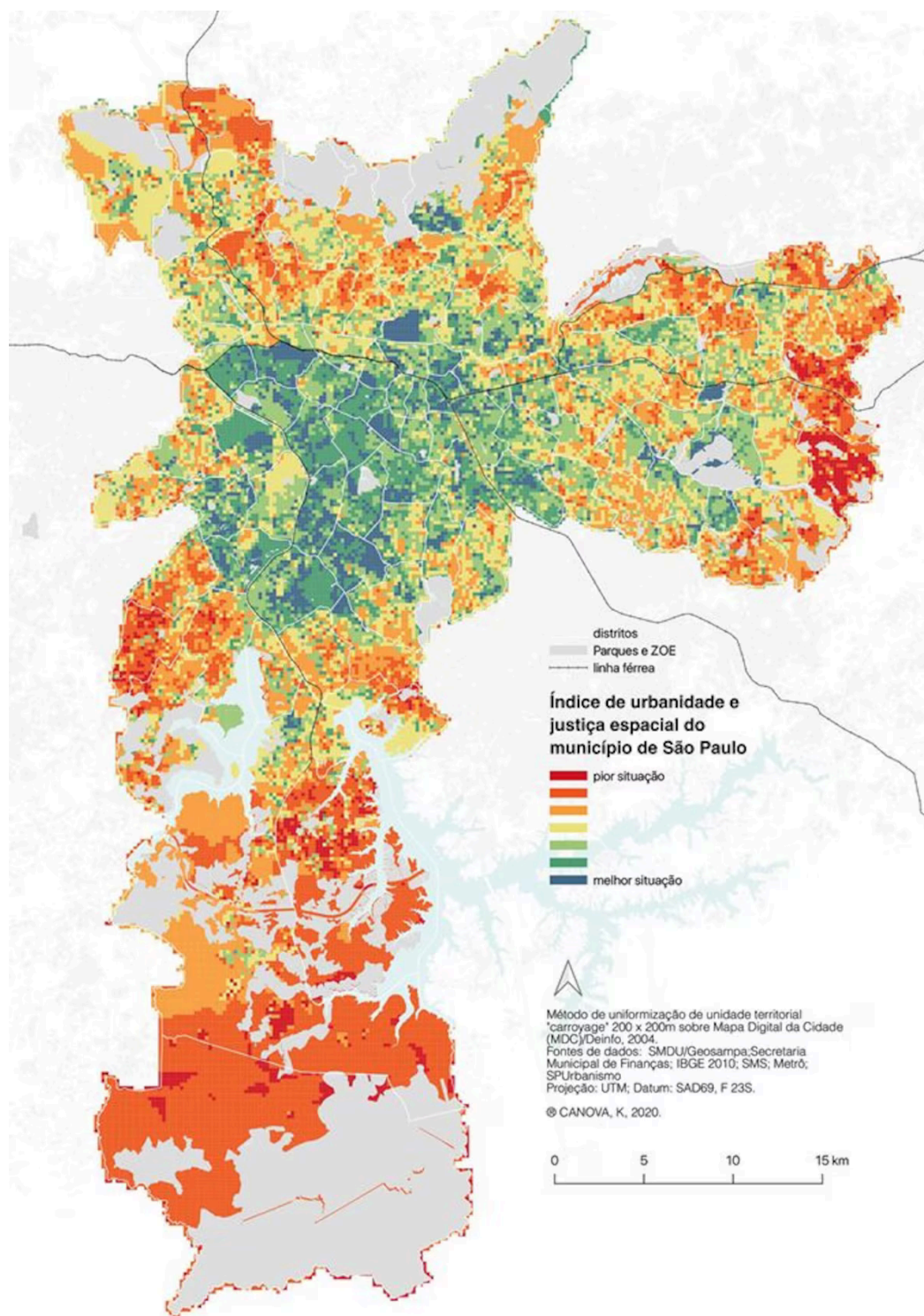


Figura I: Visualização de Zonas de Ocupação Especial (ZOE) e Parques no município de São Paulo. Fonte:

<https://www.estadao.com.br/sao-paulo/caminhadas-urbanas/o-mapa-da-desigualdade-em-sao-paulo/>

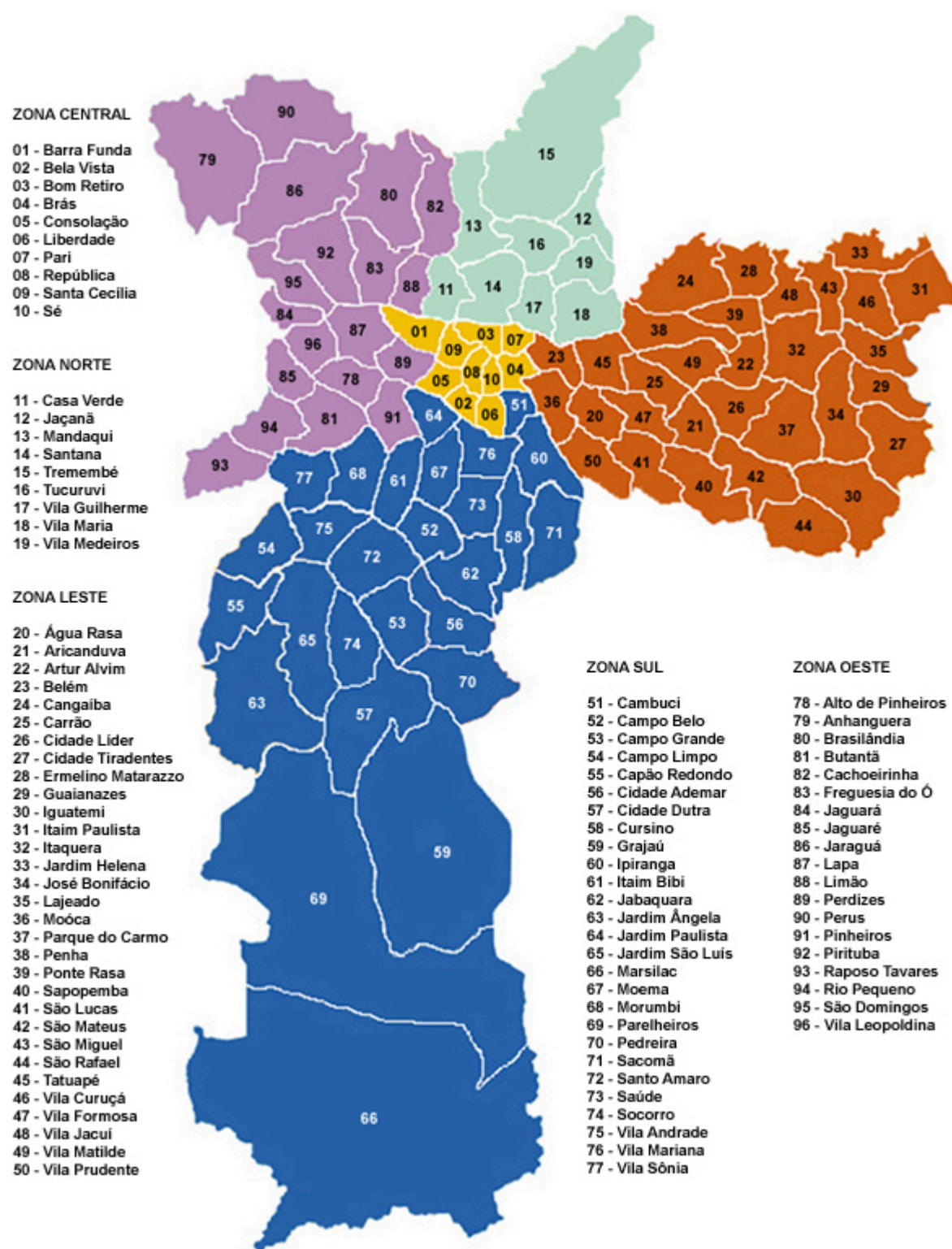


Figura II: Distritos Administrativos no município de São Paulo. Fonte: <https://jaraquasp.blogspot.com/2016/08/fronteiras-bairro-jaragua.html>

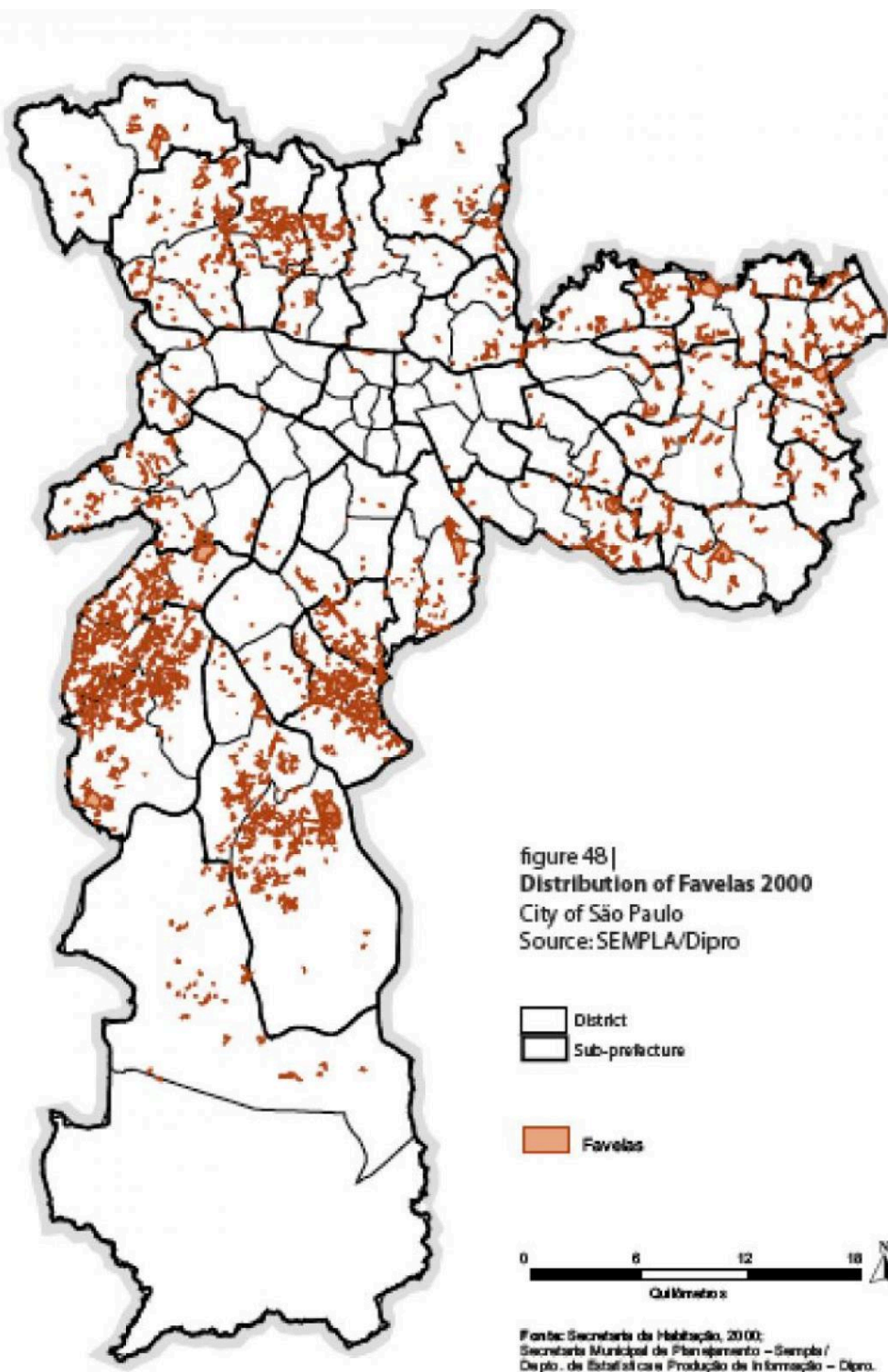


Figura III: Localização das comunidades no município de São Paulo. Fonte: <<https://pt.map-of-sao-paulo.com/outros-mapas/são-paulo-favelas-mapa>>

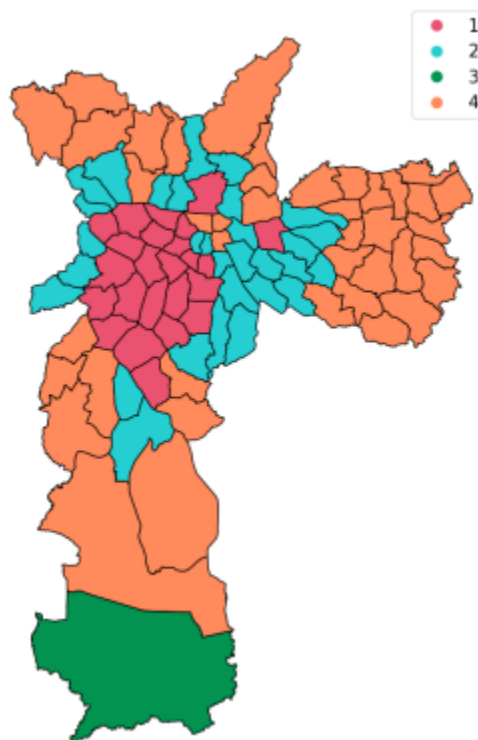


Figura IV: Distribuição geográfica dos clusters realizados no estudo de Martinelli (2023) considerando o agrupamento de taxas de mortalidade padronizadas por distrito administrativo no município de São Paulo.

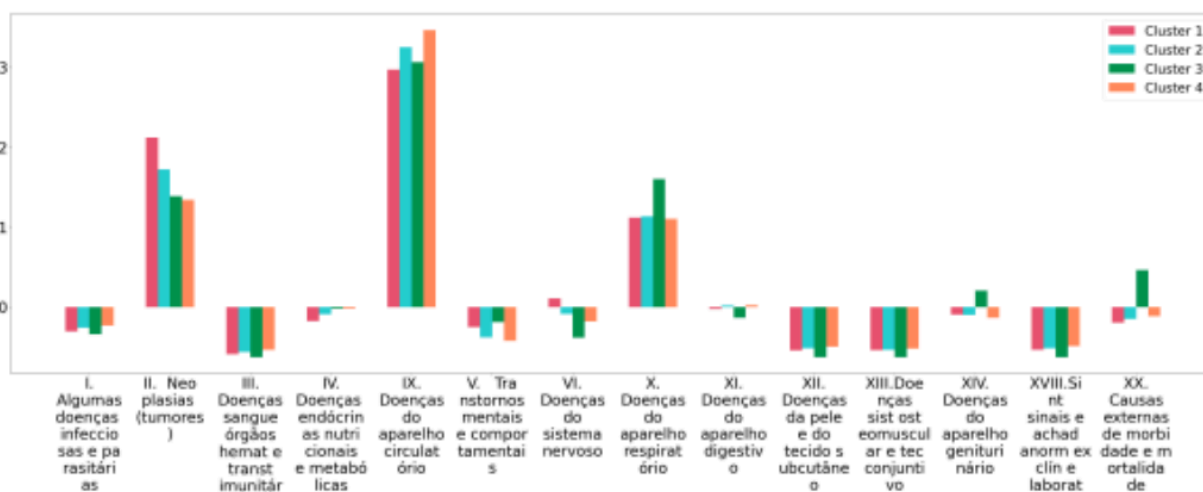


Figura V: Taxas médias de mortalidade padronizadas por distrito administrativo, para cada cluster realizado no estudo de Martinelli (2023).