

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

**Centro de Estatística Aplicada**

**Relatório de Análise Estatística**

RAE-CEA–22P08

**RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:**

**“Mapa da Desigualdade de Gênero e Interseccionalidades do Município de São Paulo”**

**Eduardo Janotti Cavalcante**

**Lúcia Pereira Barroso**

**São Paulo, julho de 2022**

**CENTRO DE ESTATÍSTICA APLICADA - CEA – USP**

**TÍTULO:** Relatório de Análise Estatística sobre o Projeto: “Mapa da Desigualdade de Gênero e Interseccionalidades do Município de São Paulo”.

**PESQUISADORA:** Luciana Fukimoto Itikawa

**INSTITUIÇÃO:** Instituto de Estudos Avançados da USP

**FINALIDADE DO PROJETO:** Pesquisa

**RESPONSÁVEIS PELA ANÁLISE:** Eduardo Janotti Cavalcante  
Lúcia Pereira Barroso

**REFERÊNCIA DESTE TRABALHO:** CAVALCANTE, E.J.; BARROSO, L.P. **Relatório de análise estatística sobre o projeto: “Mapa da Desigualdade de Gênero e Interseccionalidades do Município de São Paulo”.** São Paulo, IME-USP, 2022. (RAE–CEA-22P08)

## FICHA TÉCNICA

### REFERÊNCIAS BIBLIOGRÁFICAS:

Barrozo, L. V.; Fornaciali, M.; de André C. D. S.; Moraes, G. A. Z.; Mansur, G.; Cabral-Miranda, W.; de Mirandam M. J.; Sato, J. R.; Junior, E. A. (2020). GeoSES: A socioeconomic index for health and social research in Brazil. **PLoS ONE**, (15)4: e0232074. Disponível em: <<https://doi.org/10.1371/journal.pone.0232074>>. Acesso em: 26 de junho de 2022.

Buuren, S. V. (2018). **Flexible Imputation of Missing Data**. 2.ed. Nova York: Chapman & Hall/CRC Press. Disponível em: <<https://stefvanbuuren.name/fimd/>>. Acesso em: 28 de maio de 2022.

CENTRO DE ESTUDO DA METRÓPOLE (CEM). **Mapa da Vulnerabilidade Social do Município de São Paulo**. São Paulo: CEM/PMSP, 2004.

EQUAL MEASURES 2030 (EM2030). **Harnessing the Power of Data for Gender Equality. Introducing the 2019 EM 2030 SDG Gender Index**. Surrey, United Kingdom, 2019. Disponível em: <[https://data.em2030.org/wpcontent/ploads/2019/05/EM2030\\_2019\\_Global\\_Report\\_ENG.pdf](https://data.em2030.org/wpcontent/ploads/2019/05/EM2030_2019_Global_Report_ENG.pdf)>. Acesso em: outubro de 2020.

FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS (SEADE). **Índice Paulista de Vulnerabilidade Social versão 2010**. São Paulo: SEADE, 2013 Disponível em: <<https://www.seade.gov.br/ipvs>>. Acesso em: março de 2020.

Johnson, R. A.; Wichern, D. W. (2007). **Applied Multivariate Statistical Analysis**. 6.ed. Upper Saddle River: Pearson Prentice Hall.

NAÇÕES UNIDAS BRASIL. **Sobre o nosso trabalho para alcançar os Objetivos de Desenvolvimento Sustentável no Brasil**. Disponível em: <<https://brasil.un.org/pt-br/sdgs>> Acesso em: 12 de abril de 2022.

PUCSP/INPE/CEDEST. **Mapa da Inclusão e Exclusão da Cidade de São Paulo 2010**. V Fórum Social Sul. Disponível em: <[https://ceapg.fgv.br/sites/ceapg.fgv.br/files/u60/relatorio\\_\\_mapa\\_da\\_exclusao\\_social\\_-\\_sposati.pdf](https://ceapg.fgv.br/sites/ceapg.fgv.br/files/u60/relatorio__mapa_da_exclusao_social_-_sposati.pdf)> Acesso em: março de 2020.

REDE NOSSA SÃO PAULO (RNSP). **Mapa da Desigualdade do Município de São Paulo 2019 e 2020**. Disponível em: < <https://www.nossasaopaulo.org.br/campanhas/#13>>. Acesso em: novembro de 2020.

YAMAMOTO, J. K; LANDIM, P. M. B. (2013). **Geoestatística: conceitos e aplicações**. São Paulo: Oficina de Textos.

## **PROGRAMAS COMPUTACIONAIS UTILIZADOS:**

LibreOffice Writer versão 7.2.6.2

Google Sheets

Python versão 3.9.7

Spyder IDE versão 5.1.5

R versão 4.1.2

Rstudio versão 2022.2.3.492

QGIS versão 3.26.0-Buenos Aires

## **TÉCNICAS ESTATÍSTICAS UTILIZADAS**

Análise Descritiva Unidimensional (03:010)

Análise de Componentes Principais (06:070)

## **ÁREA DE APLICAÇÃO**

Sociometria (14:100)

Ciências Físicas e Geoestatística (14:010)

## **Resumo**

Não existe um índice síntese exclusivo para a desigualdade de gênero entre setores censitários que mostre as diversas diferenças dentro dos municípios. O estudo analisado neste relatório estatístico tem como escopo encontrar uma síntese de indicadores socioeconômicos, urbanos, laborais, criminalísticos, ambientais, sanitários etc. Com este intuito, foram construídos dois indicadores para o município de São Paulo, que apresentam características similares. Os resultados foram ilustrados no mapa do município.

## **Sumário**

<b>1. Introdução</b>	<b>8</b>
<b>2. Objetivo</b>	<b>9</b>
<b>3. Descrição do estudo</b>	<b>9</b>
<b>4. Descrição das variáveis</b>	<b>9</b>
<b>5. Análise descritiva</b>	<b>12</b>
<b>6. Construção do índice</b>	<b>14</b>
<b>7. Conclusão</b>	<b>18</b>
<b>APÊNDICE</b>	<b>19</b>
<b>ANEXO</b>	<b>69</b>

## 1. Introdução

Os Objetivos de Desenvolvimento Sustentável (ODSs) são os propósitos para os quais a Organização das Nações Unidas (ONU) está contribuindo a fim de garantir que pessoas de todos os lugares possam desfrutar de paz e de prosperidade, sendo divididos em 17 ODSs.

Dentre eles, o ODS 5 tem como objetivo alcançar a igualdade de gênero e empoderar todas as mulheres e meninas através das seguintes ações: acabar com todas as formas de discriminação, as formas de violência e as práticas nocivas; reconhecer e valorizar o trabalho de assistência e doméstico não remunerado, bem como a promoção de responsabilidade compartilhada dentro do lar e da família; garantir a participação plena e efetiva e a igualdade de oportunidades para a liderança, assim como o acesso universal à saúde sexual e reprodutiva e aos direitos reprodutivos; realizar reformas para dar direitos iguais aos recursos econômicos; aumentar o uso de tecnologias de base para promover o empoderamento e adotar e fortalecer políticas e legislação para a promoção da igualdade de gênero e o empoderamento em todos os níveis (NAÇÕES UNIDAS BRASIL).

Em relação ao ODS 5, o Brasil está classificado em septuagésimo sétimo dentre os 129 países considerados (EM2030, 2019). Os ODSs são apresentados pela ONU através de escala em nível nacional, mas a gestão de equipamentos e recursos no Brasil é delegada pela constituição aos municípios. Portanto, se torna importante entender as desigualdades dentro dos municípios a fim de subsidiar as políticas públicas para melhorar a posição do Brasil.

Apesar de existirem alguns poucos mapas que apresentam indicadores específicos de gênero como o Mapa da Vulnerabilidade Social do Município de São Paulo (CEM, 2004), o Índice Paulista de Vulnerabilidade Social (SEADE, 2013), o Mapa da Inclusão e Exclusão da Cidade de São Paulo (PUCSP/INPE/CEDEST, 2010) e o Mapa da Desigualdade do Município de São Paulo (RNSP, 2019; 2020), esses indicadores estão misturados com outros que não referenciam a questão do gênero.



Portanto, não temos ainda um índice síntese exclusivo para a desigualdade de gênero entre setores censitários que mostre as diversas diferenças dentro dos municípios.

## **2. Objetivo**

O objetivo do estudo é encontrar uma síntese de indicadores socioeconômicos, urbanos, laborais, criminalísticos, ambientais, sanitários etc, e avaliar o comportamento desse indicador-resumo nos setores censitários do município de São Paulo, a fim de elaborar mapas temáticos para dar subsídios a políticas públicas.

## **3. Descrição do estudo**

O estudo foi dividido em duas etapas. A primeira etapa consistiu no levantamento de dados referentes ao município de São Paulo por volta do ano de 2010, na criação de indicadores síntese a partir desses dados e na elaboração de mapas temáticos usando os indicadores como base. A segunda etapa consistirá em repetir a primeira, mas com dados por volta do ano de 2020, e, após a obtenção dos dados dos dois períodos, construir uma plataforma interativa a partir dos mapas. Este relatório aborda a criação dos indicadores síntese para os dados de 2010.

## **4. Descrição das variáveis**

Variável que caracteriza os dados:

- Setor censitário: identificação do setor censitário, tendo sido observados os 18.953 setores censitários do município de São Paulo no Censo de 2010.

Variáveis que caracterizam o setor censitário:

- As variáveis observadas são agrupadas nos seguintes fatores

- o Ambiental:

- Quantidade de áreas de riscos de deslizamento

- Quantidade de áreas contaminadas

Concentração de monóxido de carbono (ppm)

Concentração de dióxido de enxofre (mg/m<sup>3</sup>)

Temperatura superficial diurna média (°C)

o Habitacional:

Quantidade de domicílios adensados de mulheres chefes de família

Quantidade de domicílios sem segurança na posse

Quantidade de domicílios sem ligação de esgoto na rede

Quantidade de domicílios sem abastecimento de água na rede

Quantidade de favelas

Quantidade de cortiços

Quantidade de loteamentos irregulares

Quantidade de núcleos precários

Quantidade de assentamentos precários

o Saúde da mulher:

Quantidade de consultas pré-natal

Proporção de consultas pré-natal por mulher em idade gestacional

Quantidade de partos cesárea

Quantidade total de partos

Proporção de partos cesárea quanto ao total de partos

o Mobilidade (considerando o deslocamento habitual para o trabalho):

Quantidade de viagens feitas por mulheres

Quantidade de pessoas cujo tempo de deslocamento:

- é menor do que 5 minutos
- está entre 5 e 30 minutos
- está entre 30 e 60 minutos
- está entre 60 e 120 minutos
- é maior do que 120 minutos

o Violência:

Proporção de ocorrências de violência contra mulher por população total do setor

Quantidade de casos de feminicídio

Quantidade de casos de racismo

Quantidade de casos de homofobia/transfobia

Letalidade violenta intencional: quantidade de homicídios intencionais de mulheres

o População:

Quantidade de mulheres pretas e pardas chefes de família

Quantidade de mulheres chefes de família com até 3 salários mínimos

o Trabalho:

Quantidade de trabalhadores formais

Quantidade de trabalhadores informais

o Equipamentos sociais:

Proporção de vagas de creche por demanda

- o Representatividade política:

Quantidade de votos em vereadoras eleitas na eleição de 2012

- o Acesso a sinal de antena

Quantidade de antenas por habitante

O ano em que as variáveis foram medidas e suas respectivas fontes estão no anexo.

Variáveis que caracterizam o setor censitário utilizadas para padronização (obtidas em 2010 pelo IBGE):

- Tamanho do setor (Km<sup>2</sup>)
- Quantidade de domicílios
- Quantidade de pessoas
- Quantidade de mulheres
- Quantidade de mulheres chefes de família

## 5. Análise descritiva

Foram construídos histogramas ou gráficos de barras, e tabelas de medidas-resumo de todas as variáveis pelo programa Python (Python versão 3.9.7). Os gráficos e tabelas criados estão disponíveis no apêndice.

A partir desses gráficos e tabelas, foi possível observar determinados padrões de comportamento que se repetem para diferentes variáveis.

Os fatores ambiental, habitacional, mobilidade e acesso a sinal de antena estão associados a algumas variáveis com um número expressivo de observações iguais a zero, seguido de uma queda brusca nas frequências das demais observações, resultando em curvas que rapidamente decrescem, mas apresentam valores máximos

altos com baixa densidade, como visto na Tabela 1 e Figura 1 (quantidade de áreas de riscos de deslizamentos), Tabela 2 e Figura 2 (quantidade de áreas contaminadas), Tabela 8 e Figura 8 (quantidade de domicílios sem ligação de esgoto na rede), Tabela 9 e Figura 9 (quantidade de domicílios sem abastecimento de água na rede), Tabela 10 e Figura 10 (quantidade de favelas), Tabela 11 e Figura 11 (quantidade de cortiços), Tabela 12 e Figura 12 (quantidade de loteamentos irregulares), Tabela 13 e Figura 13 (quantidade de núcleos precários), Tabela 14 e Figura 14 (quantidade de assentamentos precários), Tabela 21 e Figura 21 (quantidade de pessoas cujo tempo de deslocamento é menor do que 5 minutos) e Tabela 37 e Figura 37 (quantidade de antenas por habitante).

Outro comportamento observado nos dados foi o de alguns valores apresentarem frequências altas de observações destacados dos demais, sem a observação de qualquer valor próximo. Esse comportamento foi observado nos fatores saúde da mulher, mobilidade, violência, trabalho, equipamentos sociais e representatividade política por meio dos valores das variáveis representadas na Tabela 17 e Figura 17 (quantidade de partos cesárea), Tabela 18 e Figura 18 (quantidade total de partos), Tabela 19 e Figura 19 (proporção de partos cesárea quanto ao total de partos), Tabela 20 e Figura 20 (quantidade de viagens feitas por mulheres), Tabela 22 e Figura 22 (quantidade de pessoas cujo tempo de deslocamento está entre 5 e 30 minutos), Tabela 23 e Figura 23 (quantidade de pessoas cujo tempo de deslocamento está entre 30 e 60 minutos), Tabela 24 e Figura 24 (quantidade de pessoas cujo tempo de deslocamento está entre 60 e 120 minutos), Tabela 25 e Figura 25 (quantidade de pessoas cujo tempo de deslocamento é maior do que 120 minutos), Tabela 28 e Figura 28 (quantidade de casos de racismo), Tabela 30 e Figura 30 (letalidade violenta intencional), Tabela 33 e Figura 33 (quantidade de trabalhadores formais), Tabela 34 e Figura 34 (quantidade de trabalhadores informais), Tabela 35 e Figura 35 (proporção de vagas de creche por demanda) e Tabela 36 e Figura 36 (quantidade de votos em vereadoras eleitas). Esse comportamento pode ter surgido a partir da redução da escala das observações, pois alguns dados foram obtidos para áreas maiores contendo vários setores, para depois serem adaptados à unidade de interesse do estudo pela atribuição da média aos setores censitários da área.

Algumas variáveis apresentaram destaque em seu valor máximo, com altas frequências. São elas as representadas nas Tabela 17 e Figura 17 (quantidade de partos cesárea), Tabela 18 e Figura 18 (quantidade total de partos), Tabela 19 e Figura 19 (proporção de partos cesárea quanto ao total de partos), Tabela 20 e Figura 20 (quantidade de viagens feitas por mulheres) e Tabela 36 e Figura 36 (quantidade de votos em vereadoras eleitas).

Além disso, foi possível observar uma alta quantidade de dados faltantes, de tal forma que apenas 86% dos setores apresentaram todas as observações, sendo o fator trabalho o único no qual todas as variáveis apresentaram dados completos.

## 6. Construção do índice

Antes da criação do índice resumo, as variáveis foram padronizadas porque os setores censitários têm tamanhos, quantidades de domicílios e características populacionais diferentes. A padronização foi realizada da seguinte forma:

- As variáveis quantidade de áreas de riscos de deslizamento, quantidade de áreas contaminadas, quantidade de favelas, quantidade de cortiços, quantidade de loteamentos irregulares e quantidade de núcleos precários tiveram seus valores divididos pelos respectivos valores da variável tamanho do setor.
- As variáveis quantidade de domicílios adensados de mulheres chefes de família, quantidade de domicílios sem segurança na posse, quantidade de domicílios sem ligação de esgoto na rede e quantidade de domicílios sem abastecimento de água na rede tiveram seus valores divididos pelos respectivos valores da variável quantidade de domicílios no setor.
- As variáveis quantidade de pessoas com certo tempo de deslocamento, quantidade de trabalhadores formais, quantidade de casos de racismo, quantidade de casos de homofobia/transfobia, quantidade de trabalhadores informais e quantidade de votos em vereadoras eleitas tiveram seus valores divididos pelos respectivos valores da variável quantidade de pessoas no setor.

- As variáveis quantidade de viagens feitas por mulheres, quantidade de casos de feminicídio e letalidade violenta intencional tiveram seus valores divididos pelos respectivos valores da variável quantidade de mulheres no setor.
- As variáveis quantidade de mulheres pretas e pardas chefes de família e quantidade de mulheres chefes de família com até 3 salários mínimos tiveram seus valores divididos pelos respectivos valores da variável quantidade de mulheres chefes de família no setor.
- As variáveis concentração de monóxido de carbono, concentração de dióxido de enxofre, temperatura superficial diurna média, quantidade de consultas pré-natal por mulher em idade gestacional, proporção de partos cesárea quanto ao total de partos, proporção de ocorrências de violência contra mulher por população total do setor, proporção de vagas de creche por demanda e quantidade de antenas por habitante não sofreram alterações em seus valores.
- As variáveis quantidade de assentamentos precários, quantidade de consultas pré-natal, quantidade de partos cesárea e quantidade total de partos não foram incluídas nos modelos, pois outras variáveis apresentam as mesmas informações e teriam sido redundante.

Após a divisão, as variáveis quantidade de domicílios adensados de mulheres chefes de família, quantidade de domicílios sem segurança na posse, quantidade de domicílios sem ligação de esgoto na rede, quantidade de domicílios sem abastecimento de água na rede, quantidade de pessoas com determinado tempo de deslocamento, quantidade de mulheres pretas e pardas chefes de família, quantidade de mulheres chefes de família com até 3 salários mínimos, quantidade de trabalhadores formais, quantidade de trabalhadores informais e quantidade de votos em vereadoras eleitas na eleição de 2012 apresentaram alguns valores maiores do que 1, o que não seria possível pois, no caso dessas variáveis, a divisão realizada deveria ser uma proporção,

o que significa que houve algum erro de coleta dos dados. Logo, nesses casos, os valores foram considerados iguais a 1.

Em seguida, todas as variáveis, tiveram seus valores subtraídos pelas respectivas médias e, depois, divididos pelos respectivos desvios padrões. Um problema que apareceu na padronização foi que 8 setores censitários (355030859000095, 355030846000158, 355030812000007, 355030892000174, 355030877000155, 355030816000097, 355030885000123 e 355030833000045) apresentaram número de mulheres maiores do que o número de pessoas. Para esses casos, a padronização foi feita com o número de mulheres.

Como as variáveis originais apresentavam dados faltantes, foi necessário o uso de técnicas de imputação de dados para não remover setores da análise. As duas técnicas consideradas foram a de imputação múltipla via correspondência de média preditiva (Buuren, 2018) e a interpolação por krigagem (Yamamoto e Landim, 2013). Como o método de krigagem se baseia na suposição de distribuição normal dos dados, as predições precisaram ser truncadas em zero. Assim, as estimativas provenientes de ambos os métodos foram similares. O método por imputação múltipla foi o escolhido.

Após a imputação, dois métodos foram considerados para a criação dos índices, análise de componentes principais e modelo de equações estruturais (Johnson e Wichern, 2007). Contudo, o último método apresentou problemas durante a convergência do software utilizado e 15 variáveis precisaram ser retiradas da análise. Os coeficientes apresentaram os mesmos sinais nos dois métodos, mostrando consistência na análise. Assim, o método escolhido foi o de componentes principais e os coeficientes obtidos para a construção da primeira componente principal, que seria utilizada como a base para o índice, estão na Tabela 38. As fórmulas para obter essa componente e o índice são as seguintes:

$$PC = \text{Variável Padronizada}_1 * \text{Coeficiente}_1 +$$

$$\text{Variável Padronizada}_2 * \text{Coeficiente}_2 + \dots + \text{Variável Padronizada}_{33} * \text{Coeficiente}_{33}$$

$$\text{Índice Resumo} = [PC - \text{mín}(PC)] / [\text{máx}(PC) - \text{mín}(PC)],$$



onde “ $\min(PC)$ ” é o menor valor obtido de PC para todos os setores censitários e “ $\max(PC)$ ” é o maior valor obtido.

A primeira componente principal explica 14% da variância total dos dados, enquanto que a adição de uma segunda componente aumentaria essa porcentagem para 24%. No entanto, devido à alta quantidade de variáveis, 14% pode ser considerado um valor substancial. A segunda componente também poderia ser utilizada. Para efeito de interpretação, consideramos as variáveis com coeficiente em valor absoluto maior que 0,1 (Tabela 38, primeira coluna). Note que entre esses, todos têm sinal positivo, exceto o que correspondente à proporção de partos cesárea em relação ao total de partos. As variáveis consideradas mais importantes, ou seja, com coeficiente maior que 0,1, indicam que quanto maior seu valor, pior a situação do setor censitário. Por essa perspectiva, era de se esperar que o coeficiente da quantidade de votos em vereadoras eleitas na eleição de 2012 fosse negativo, entretanto, como se observa na Figura 38, a correlação dessa variável com as demais é positiva, coerente com o sinal encontrado. Assim, quanto maior o valor do índice resultante, pior é a situação do setor censitário.

Outra estratégia de análise foi aplicar primeiro uma análise de componentes principais para cada fator e depois a análise de componentes principais sobre os fatores. Os fatores equipamentos sociais, representatividade política e acesso a sinal de antena participaram da análise por meio das próprias variáveis que os compõem, visto que têm uma única variável. Esse procedimento é similar ao que foi aplicado em Barrozo et al. (2020) e foi denominado neste relatório como análise de componentes principais de segunda ordem. Os coeficientes resultantes estão na Tabela 38, segunda coluna. Nessa tabela podemos perceber que as variáveis mais importantes são praticamente as mesmas encontradas na primeira análise e a porcentagem da variância total explicada pela componente resultante é 12%, próxima da anterior. Contudo, nesse novo modelo cada fator equivale a uma variável na construção do índice, enquanto que no anterior, fatores com mais variáveis poderiam a princípio parecer mais representados. No resultado final, os pesos não são tão diferentes.

Para as componentes principais de cada fator, as porcentagens de explicação da variância total são: Ambiental (42%), Habitacional (23%), Saúde da Mulher (52%), Mobilidade (35%), Violência (61%), População (90%), Trabalho (81%), Equipamentos Sociais (100%), Representação Política (100%) e Acesso a Sinal de Antena (100%). Os coeficientes dos índices intermediários estão nas Tabelas 39 a 45. Os coeficientes dos fatores equipamentos sociais, representatividade política e acesso a sinal de antena não estão nas tabelas, pois, como esses fatores contêm apenas uma variável, a primeira componente principal seria a própria variável padronizada.

A partir da Figura 39, Tabela 38 e Tabela 46, podemos observar que os dois índices obtidos têm comportamento bastante similar. Pelos mapas das Figuras 40 e 41, que ilustram os resultados para o município de São Paulo, observamos que as estimativas para cada setor são parecidas.

## **7. Conclusão**

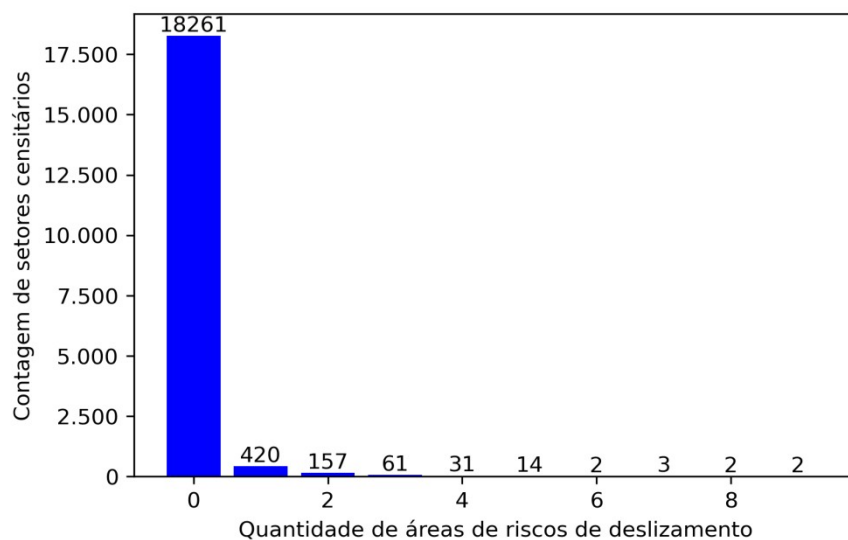
Dois índices diferentes foram apresentados neste relatório a fim de servir como medida resumo para quantificar a desigualdade entre os setores censitários, com o primeiro atribuindo a mesma importância para cada variável e o segundo, para cada fator. Contudo, devido à baixa correlação entre as variáveis dentro de um mesmo fator e as inconsistências encontradas durante a padronização, os índices resultantes podem ser melhorados. Neste relatório, o índice final foi padronizado linearmente para apresentar resultado entre zero e um. Outras padronizações poderiam ser consideradas, inclusive padronizações não-lineares.

Outro ponto a considerar é que quanto maior o valor do índice construído, pior é a situação do setor censitário. Isso poderia ser facilmente invertido, subtraindo o valor do índice do valor 1.

# APÊNDICE

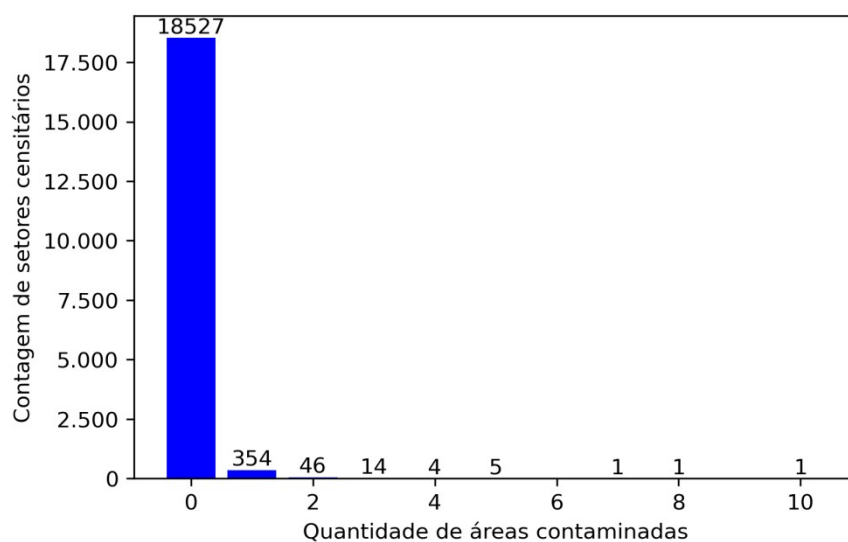
**Tabela 1** Medidas descritivas da quantidade de áreas de riscos de deslizamento

Medida	Valor
N	18953
Quantidade de zeros	18261
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	9
Média	0,062
Desvio-padrão	0,390

**Figura 1** Gráfico de barras da quantidade de áreas de riscos de deslizamento

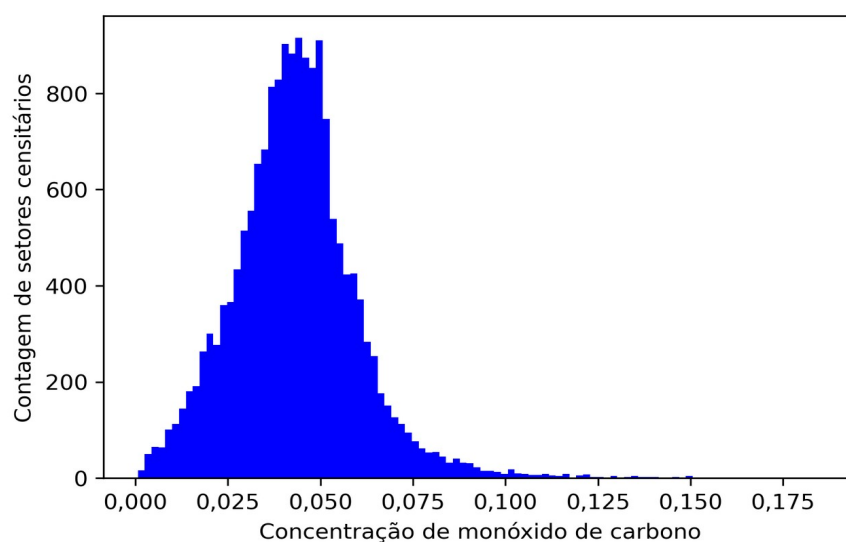
**Tabela 2** Medidas descritivas da quantidade de áreas contaminadas

Medida	Valor
N	18953
Quantidade de zeros	18527
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	10
Média	0,029
Desvio-padrão	0,235

**Figura 2** Gráfico de barras da quantidade de áreas contaminadas

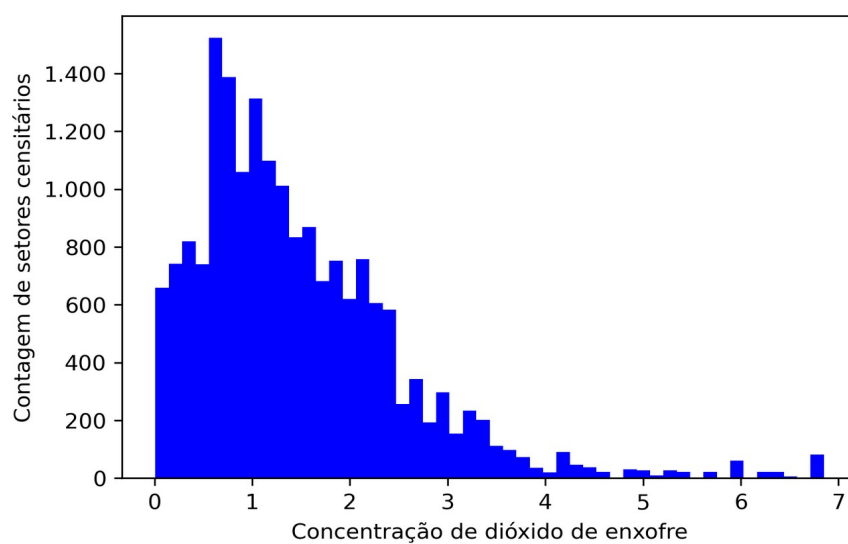
**Tabela 3** Medidas descritivas da concentração de monóxido de carbono (ppm)

Medida	Valor
N	17130
Quantidade de zeros	0
Mínimo	0,001
1º quartil	0,033
Mediana	0,043
3º quartil	0,052
Máximo	0,186
Média	0,043
Desvio-padrão	0,017

**Figura 3** Histograma da concentração de monóxido de carbono (ppm)

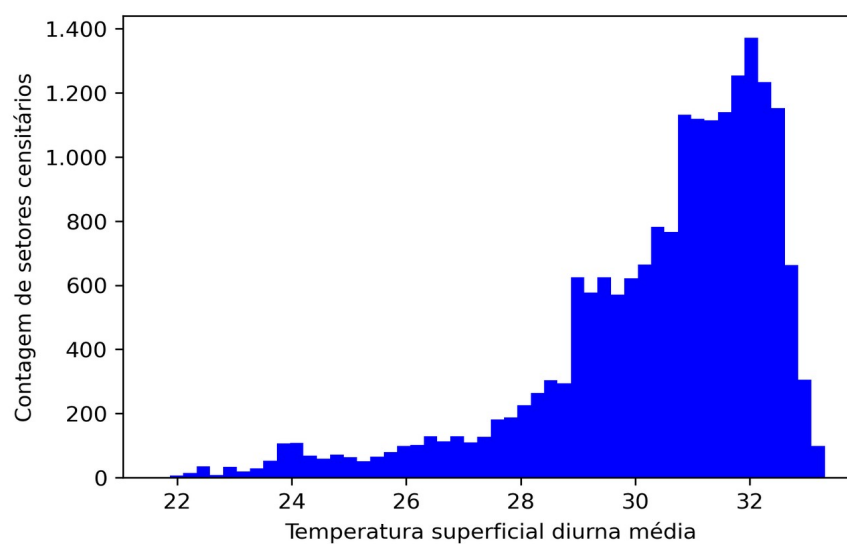
**Tabela 4** Medidas descritivas da concentração de dióxido de enxofre ( $\text{mg}/\text{m}^3$ )

Medida	Valor
N	18590
Quantidade de zeros	0
Mínimo	0,01
1º quartil	0,71
Mediana	1,24
3º quartil	2,03
Máximo	6,85
Média	1,477
Desvio-padrão	1,081

**Figura 4** Histograma da concentração de dióxido de enxofre ( $\text{mg}/\text{m}^3$ )

**Tabela 5** Medidas descritivas da temperatura superficial diurna média (°C)

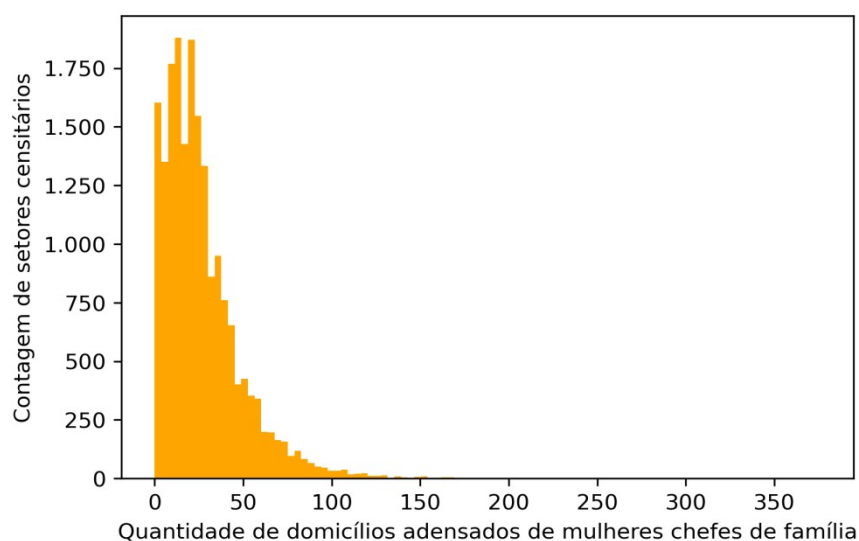
Medida	Valor
N	18953
Quantidade de zeros	0
Mínimo	21,637
1º quartil	19,49
Mediana	30,937
3º quartil	31,922
Máximo	33,313
Média	30,434
Desvio-padrão	2,048

**Figura 5** Gráfico de barras da temperatura superficial diurna média (°C)



**Tabela 6** Medidas descritivas da quantidade de domicílios adensados de mulheres chefes de família

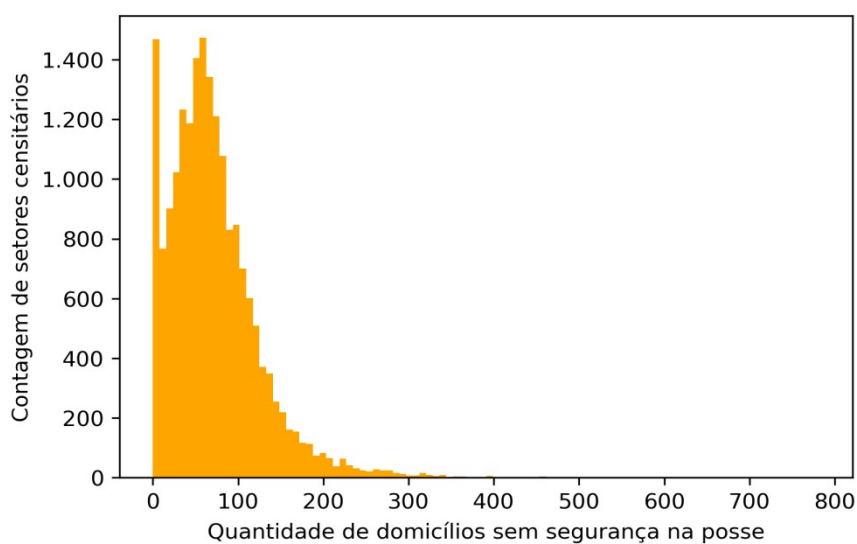
Medida	Valor
N	18953
Quantidade de zeros	889
Mínimo	0
1º quartil	12
Mediana	22
3º quartil	36
Máximo	376
Média	26,77
Desvio-padrão	22,502



**Figura 6** Histograma da quantidade de domicílios adensados de mulheres chefes de família

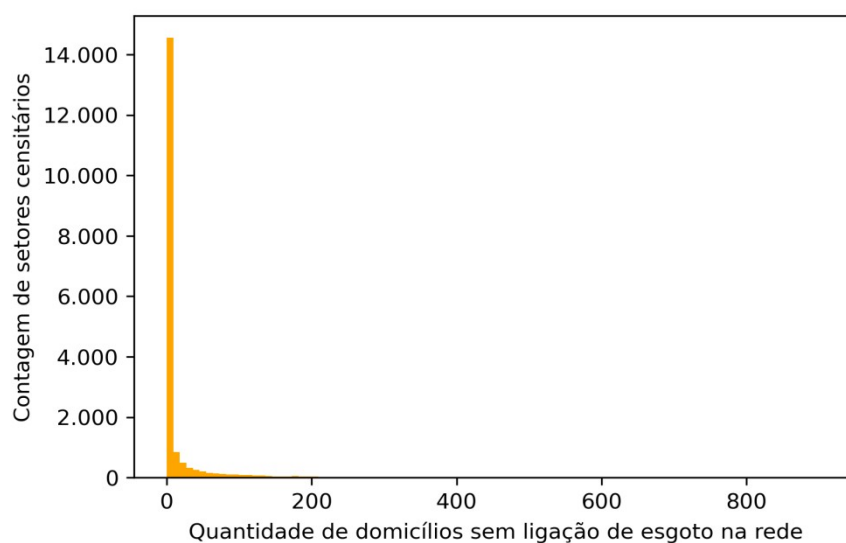
**Tabela 7** Medidas descritivas da quantidade de domicílios sem segurança na posse

Medida	Valor
N	18953
Quantidade de zeros	820
Mínimo	0
1º quartil	35
Mediana	63
3º quartil	96
Máximo	752
Média	71,496
Desvio-padrão	54,058

**Figura 7** Histograma da quantidade de domicílios sem segurança na posse

**Tabela 8** Medidas descritivas da quantidade de domicílios sem ligação de esgoto na rede

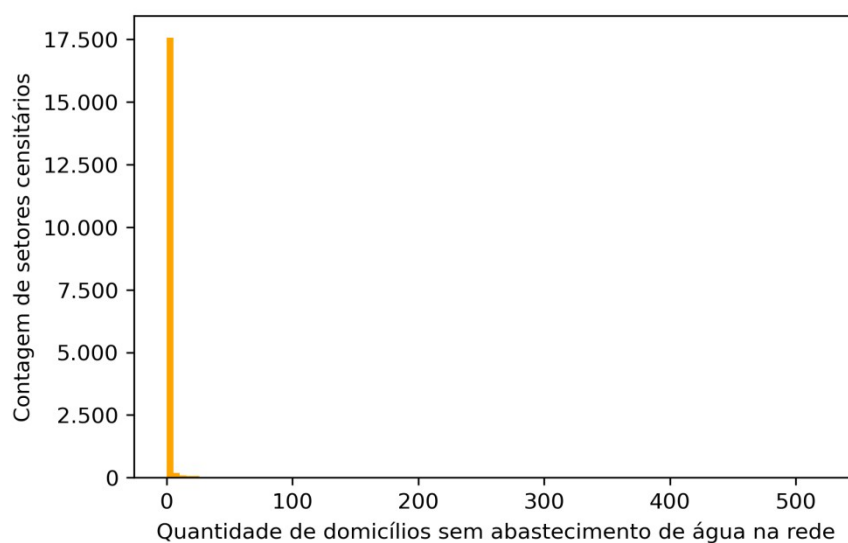
Medida	Valor
N	18206
Quantidade de zeros	9198
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	5
Máximo	908
Média	15,886
Desvio-padrão	48,16



**Figura 8** Histograma da quantidade de domicílios sem ligação de esgoto na rede

**Tabela 9** Medidas descritivas da quantidade de domicílios sem abastecimento de água na rede

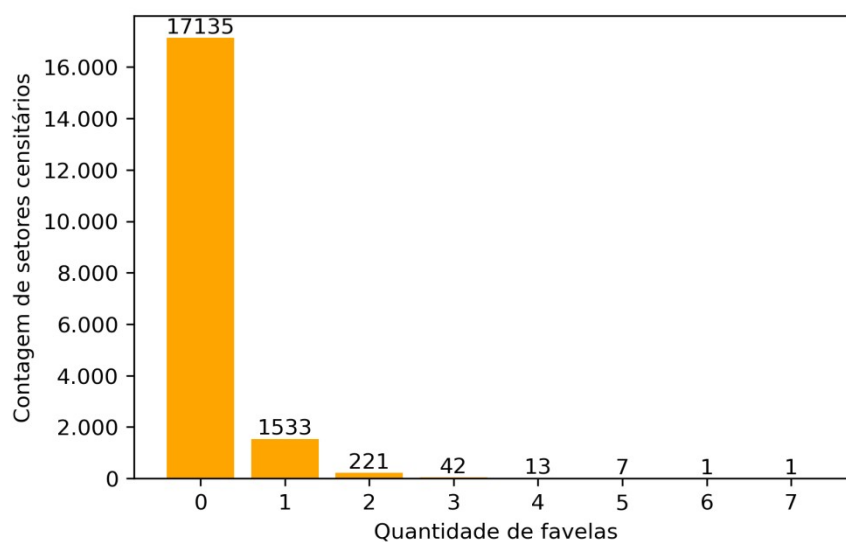
Medida	Valor
N	18206
Quantidade de zeros	15113
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	523
Média	1,779
Desvio-padrão	14,922



**Figura 9** Histograma da quantidade de domicílios sem abastecimento de água na rede

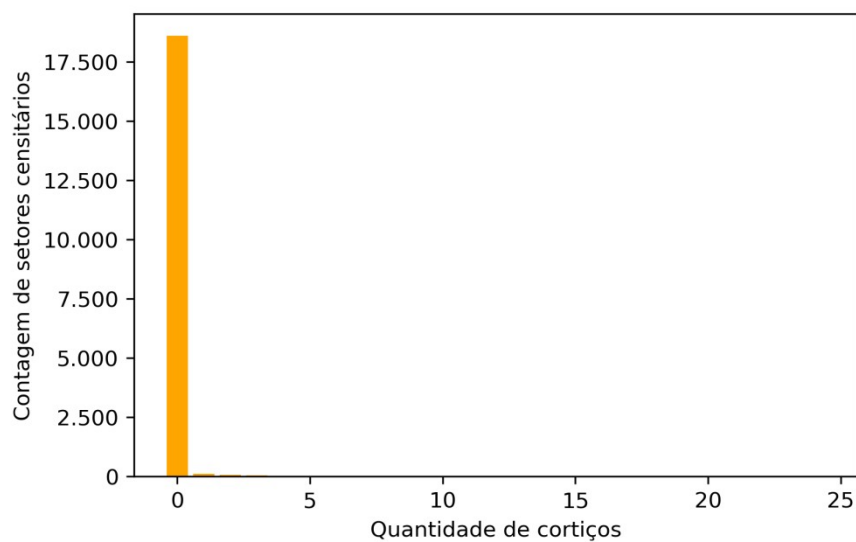
**Tabela 10** Medidas descritivas da quantidade de favelas

Medida	Valor
N	18953
Quantidade de zeros	17135
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	7
Média	0,116
Desvio-padrão	0,398

**Figura 10** Gráfico de barras da quantidade de favelas

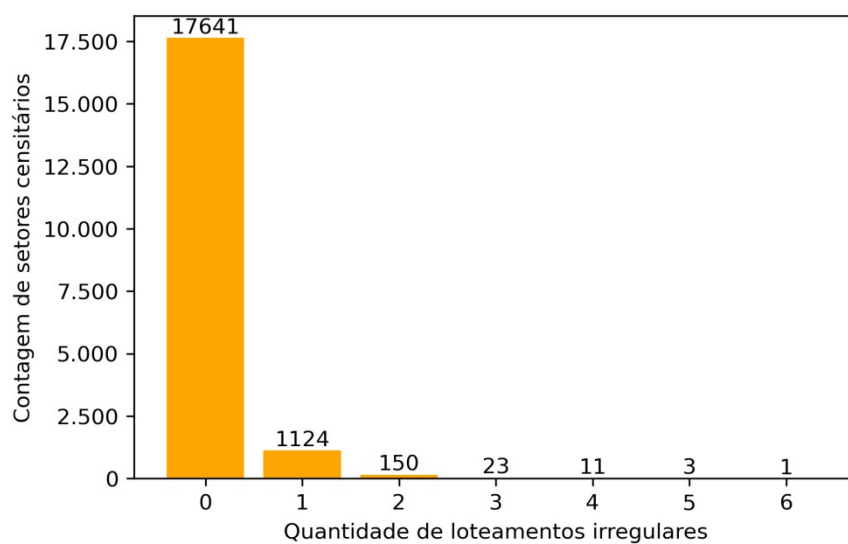
**Tabela 11** Medidas descritivas da quantidade de cortiços

Medida	Valor
N	18953
Quantidade de zeros	18957
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	24
Média	0,078
Desvio-padrão	0,818

**Figura 11** Gráfico de barras da quantidade de cortiços

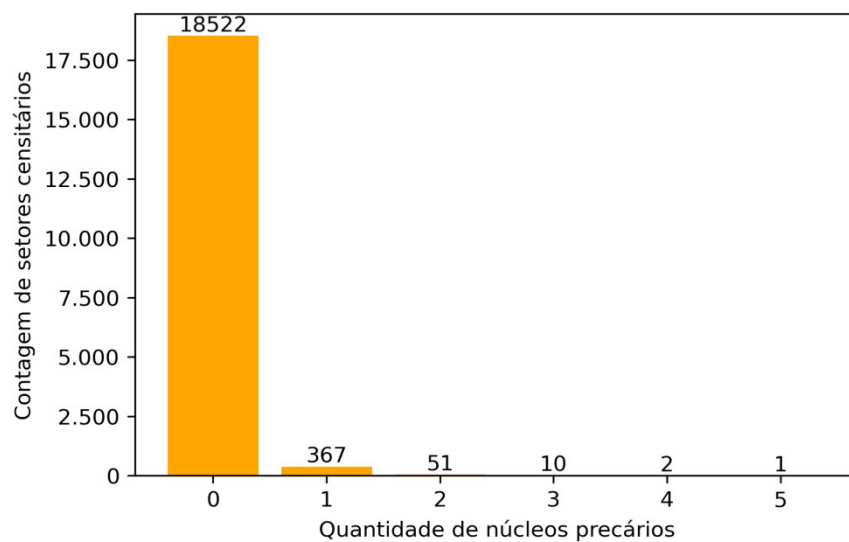
**Tabela 12** Medidas descritivas da quantidade de loteamentos irregulares

Medida	Valor
N	18953
Quantidade de zeros	17641
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	6
Média	0,082
Desvio-padrão	0,332

**Figura 12** Gráfico de barras da quantidade de loteamentos irregulares

**Tabela 13** Medidas descritivas da quantidade de núcleos precários

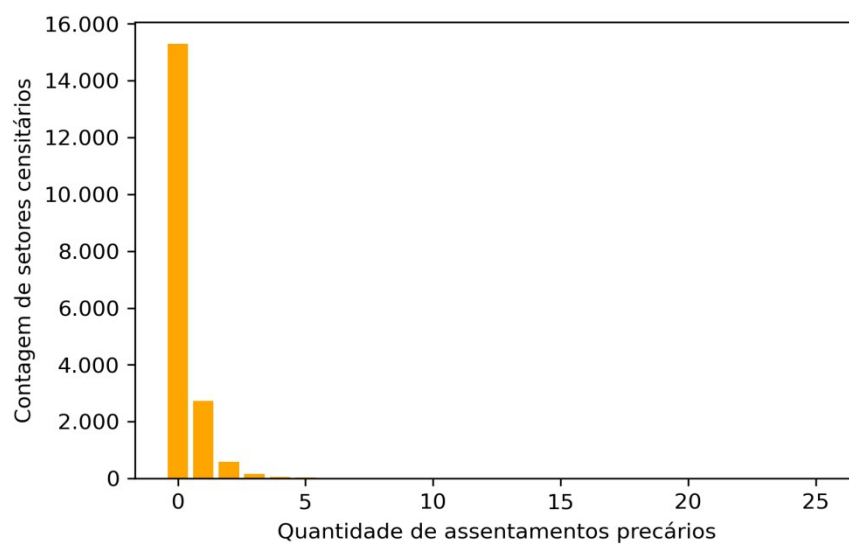
Medida	Valor
N	18953
Quantidade de zeros	18522
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	5
Média	0,027
Desvio-padrão	0,193

**Figura 13** Gráfico de barras da quantidade de núcleos precários



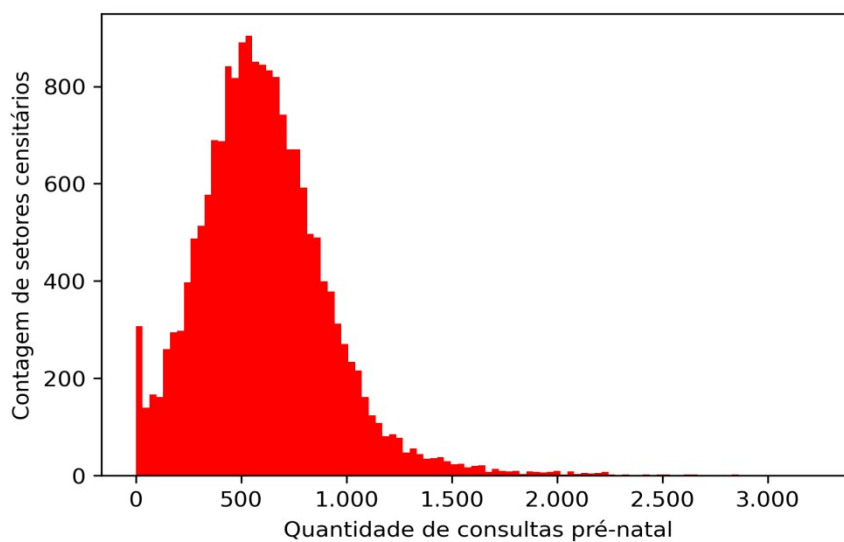
**Tabela 14** Medidas descritivas da quantidade de assentamentos precários

Medida	Valor
N	18953
Quantidade de zeros	15292
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	25
Média	0,303
Desvio-padrão	0,983

**Figura 14** Gráfico de barras da quantidade de assentamentos precários

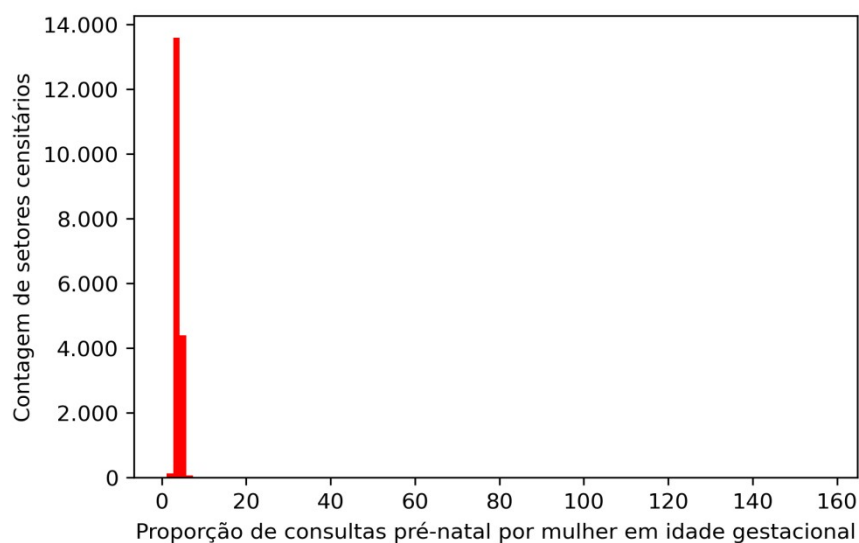
**Tabela 15** Medidas descritivas da quantidade de consultas pré-natal

Medida	Valor
N	18435
Quantidade de zeros	72
Mínimo	0
1º quartil	405
Mediana	583
3º quartil	778
Máximo	3252
Média	610,442
Desvio-padrão	316,121

**Figura 15** Histograma da quantidade de consultas pré-natal

**Tabela 16** Medidas descritivas da proporção de consultas pré-natal por mulher em idade gestacional

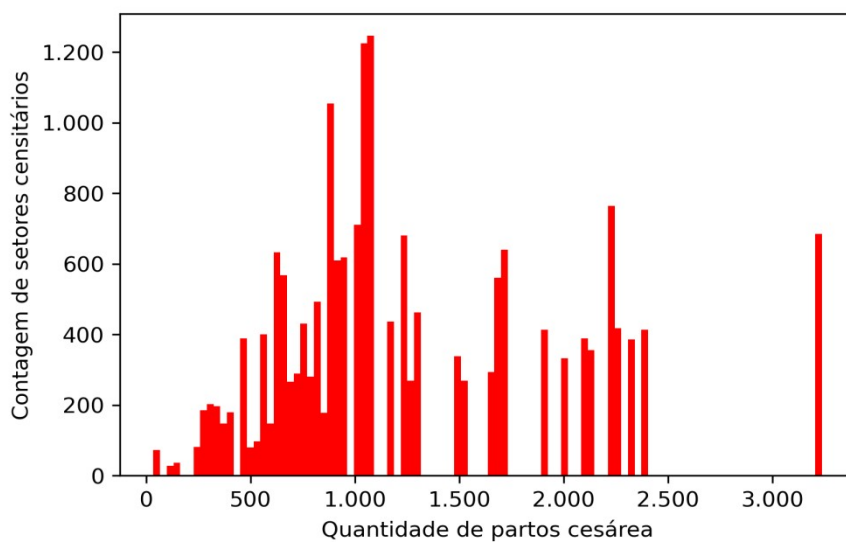
Medida	Valor
N	18193
Quantidade de zeros	0
Mínimo	1,139
1º quartil	3,71
Mediana	9,957
3º quartil	4,25
Máximo	157
Média	4,023
Desvio-padrão	1,399



**Figura 16** Histograma da proporção de consultas pré-natal por mulher em idade gestacional

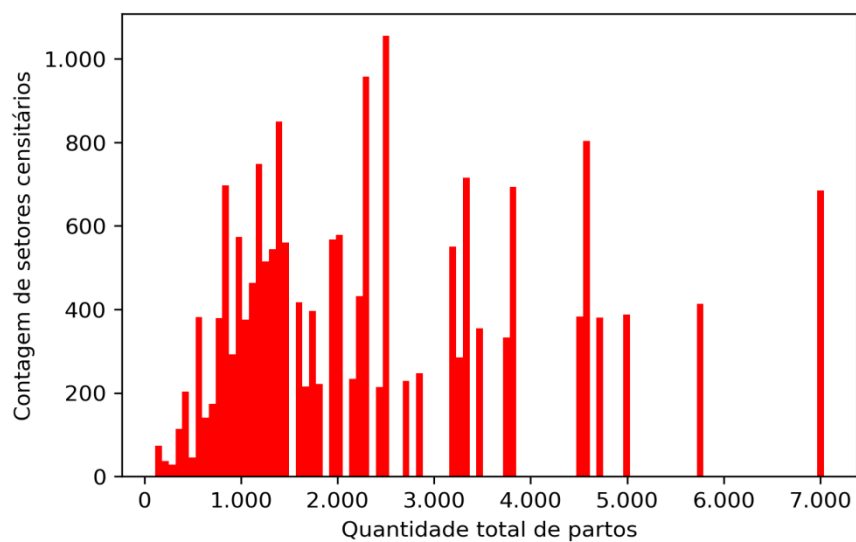
**Tabela 17** Medidas descritivas da quantidade de partos cesárea

Medida	Valor
N	18950
Quantidade de zeros	0
Mínimo	35
1º quartil	810
Mediana	1058
3º quartil	1709
Máximo	3236
Média	1272,238
Desvio-padrão	687,032

**Figura 17** Histograma da quantidade de partos cesárea

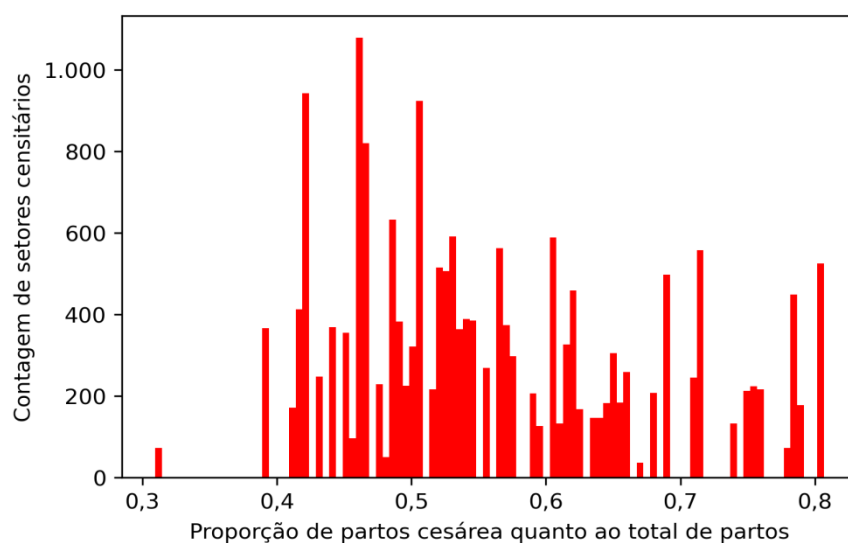
**Tabela 18** Medidas descritivas da quantidade total de partos

Medida	Valor
N	18950
Quantidade de zeros	0
Mínimo	113
1º quartil	1248
Mediana	2044
3º quartil	3351
Máximo	7037
Média	2460,192
Desvio-padrão	1592,205

**Figura 18** Histograma da quantidade total de partos

**Tabela 19** Medidas descritivas da proporção de partos cesárea quanto ao total de partos

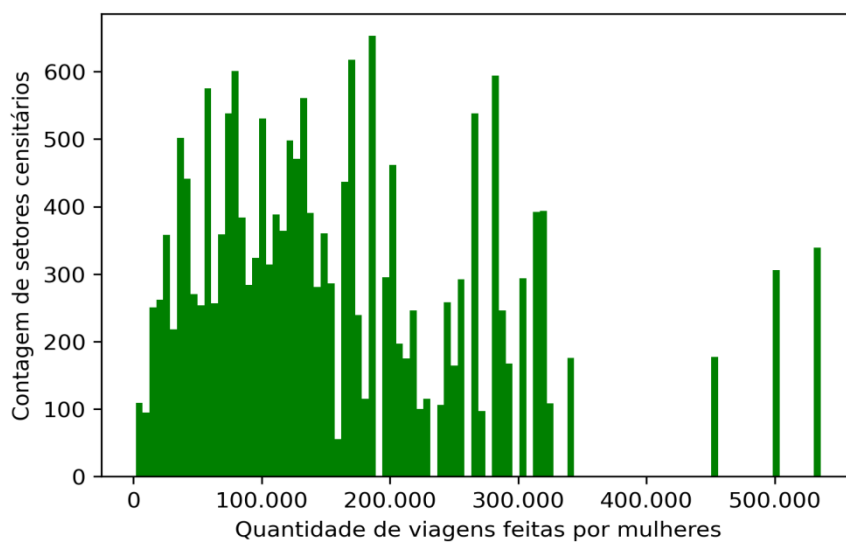
Medida	Valor
N	18950
Quantidade de zeros	0
Mínimo	0,31
1º quartil	0,468
Mediana	0,533
3º quartil	0,636
Máximo	0,807
Média	0,56
Desvio-padrão	0,111



**Figura 19** Histograma da proporção de partos cesárea quanto ao total de partos

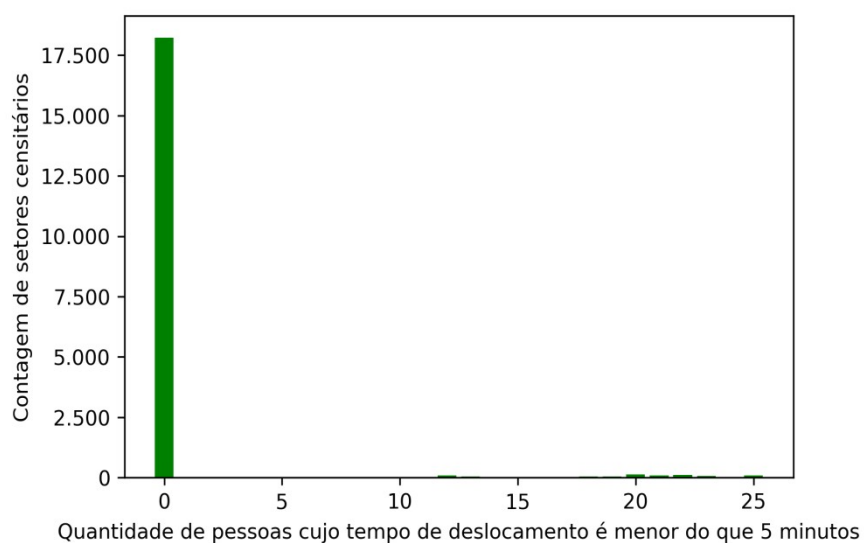
**Tabela 20** Medidas descritivas da quantidade de viagens feitas por mulheres

Medida	Valor
N	18882
Quantidade de zeros	0
Mínimo	1974,556
1º quartil	77769,812
Mediana	137421,505
3º quartil	231354,614
Máximo	535857,422
Média	164882,211
Desvio-padrão	113565,979

**Figura 20** Histograma da quantidade de viagens feitas por mulheres

**Tabela 21** Medidas descritivas da quantidade de pessoas cujo tempo de deslocamento é menor do que 5 minutos

Medida	Valor
N	18952
Quantidade de zeros	18222
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	25
Média	0,756
Desvio-padrão	3,857

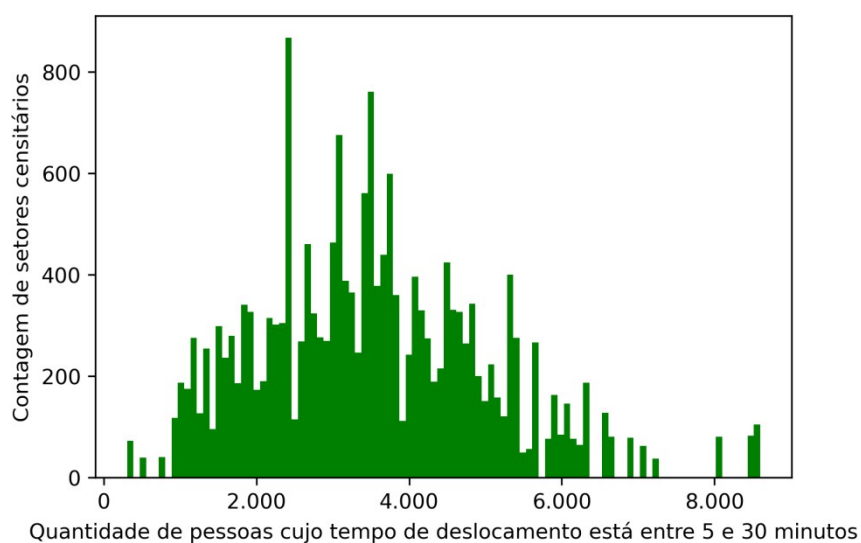


**Figura 21** Gráfico de barras da quantidade de pessoas cujo tempo de deslocamento é menor do que 5 minutos



**Tabela 22** Medidas descritivas da quantidade de pessoas cujo tempo de deslocamento está entre 5 e 30 minutos

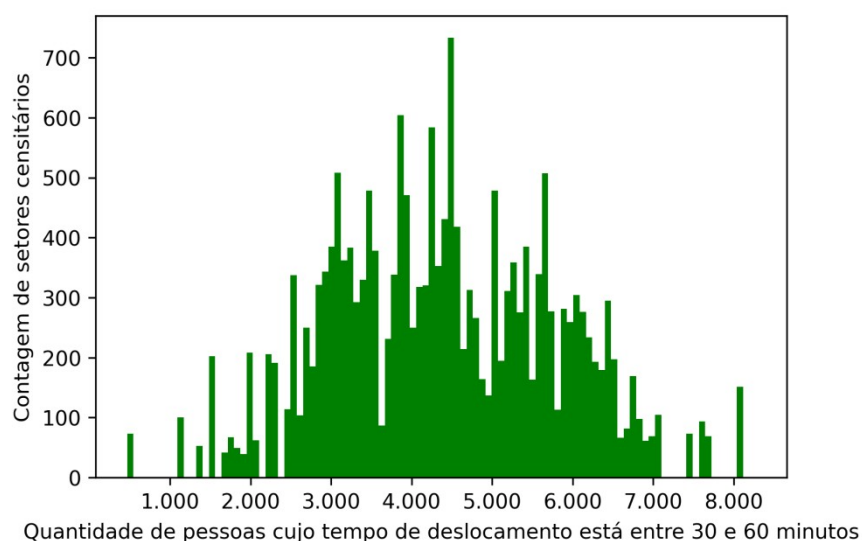
Medida	Valor
N	18952
Quantidade de zeros	2
Mínimo	307
1º quartil	2418
Mediana	3446
3º quartil	4516
Máximo	8599
Média	3525,158
Desvio-padrão	1507,94



**Figura 22** Histograma da quantidade de pessoas cujo tempo de deslocamento está entre 5 e 30 minutos

**Tabela 23** Medidas descritivas da quantidade de pessoas cujo tempo de deslocamento está entre 30 e 60 minutos

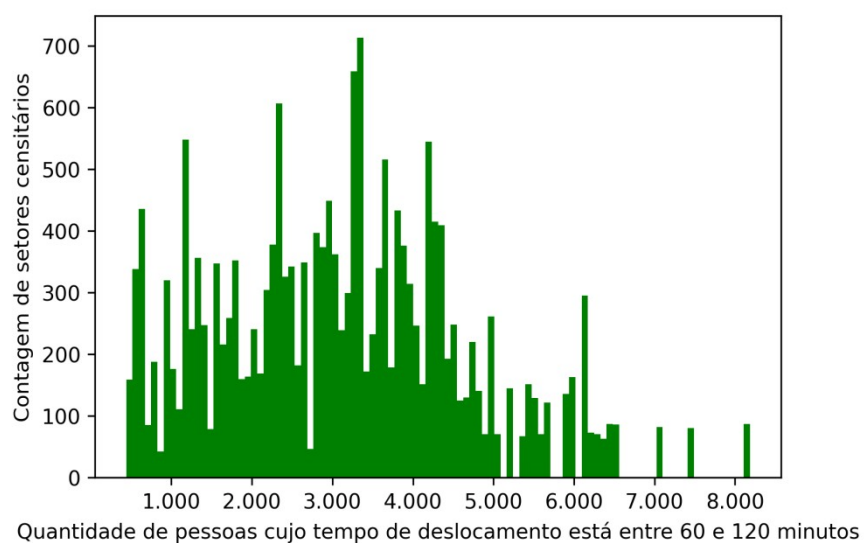
Medida	Valor
N	18952
Quantidade de zeros	0
Mínimo	468
1º quartil	3334
Mediana	4327
3º quartil	5447
Máximo	8270
Média	4387,467
Desvio-padrão	1408,955



**Figura 23** Histograma da quantidade de pessoas cujo tempo de deslocamento está entre 30 e 60 minutos

**Tabela 24** Medidas descritivas da quantidade de pessoas cujo tempo de deslocamento está entre 60 e 120 minutos

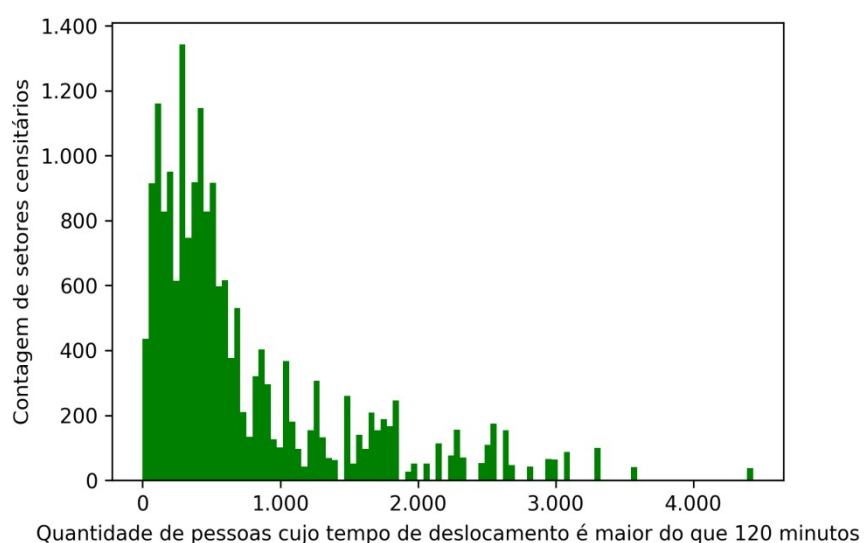
Medida	Valor
N	18952
Quantidade de zeros	0
Mínimo	441
1º quartil	1947
Mediana	3122
3º quartil	4155
Máximo	8186
Média	3120,795
Desvio-padrão	1542,13



**Figura 24** Histograma da quantidade de pessoas cujo tempo de deslocamento está entre 60 e 120 minutos

**Tabela 25** Medidas descritivas da quantidade de pessoas cujo tempo de deslocamento é maior do que 120 minutos

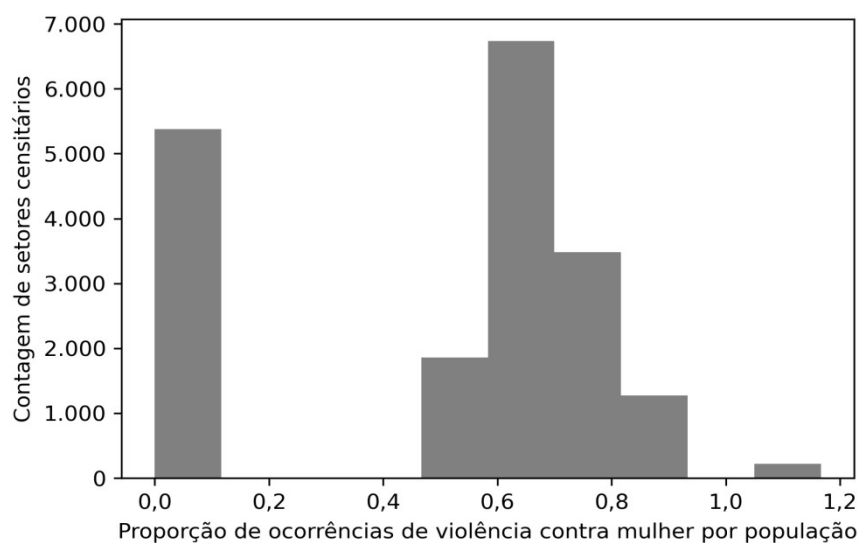
Medida	Valor
N	18952
Quantidade de zeros	73
Mínimo	0
1º quartil	258
Mediana	457
3º quartil	912
Máximo	4433
Média	730,574
Desvio-padrão	727,04



**Figura 25** Histograma da quantidade de pessoas cujo tempo de deslocamento é maior do que 120 minutos

**Tabela 26** Medidas descritivas da proporção de ocorrências de violência contra mulher por população

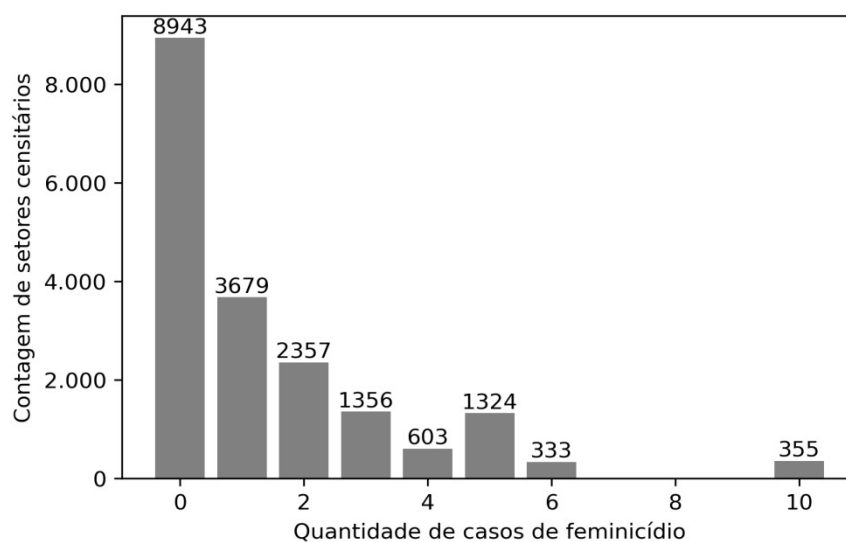
Medida	Valor
N	18950
Quantidade de zeros	5378
Mínimo	0
1º quartil	0
Mediana	0,591
3º quartil	0,706
Máximo	1,167
Média	0,479
Desvio-padrão	0,317



**Figura 26** Histograma da proporção de ocorrências de violência contra mulher por população

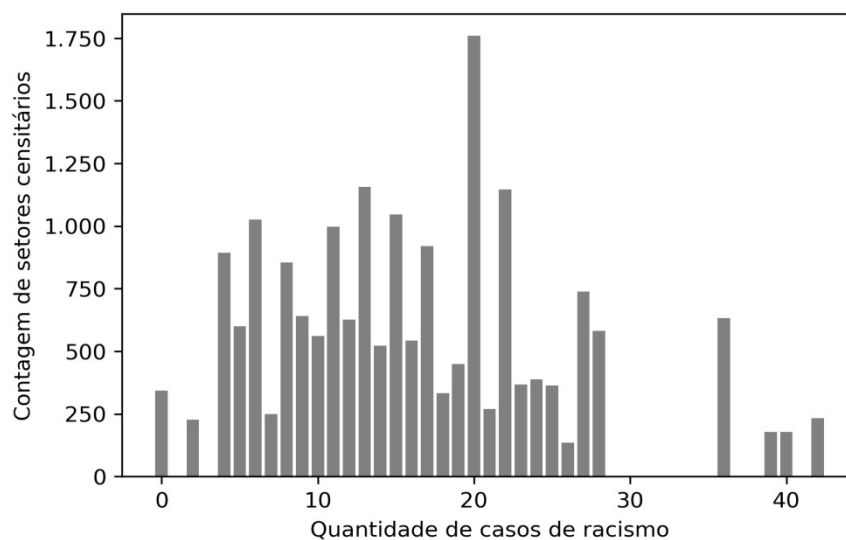
**Tabela 27** Medidas descritivas da quantidade de casos de feminicídio

Medida	Valor
N	18950
Quantidade de zeros	8943
Mínimo	0
1º quartil	0
Mediana	1
3º quartil	2
Máximo	10
Média	1,427
Desvio-padrão	2,015

**Figura 27** Gráfico de barras da quantidade de casos de feminicídio

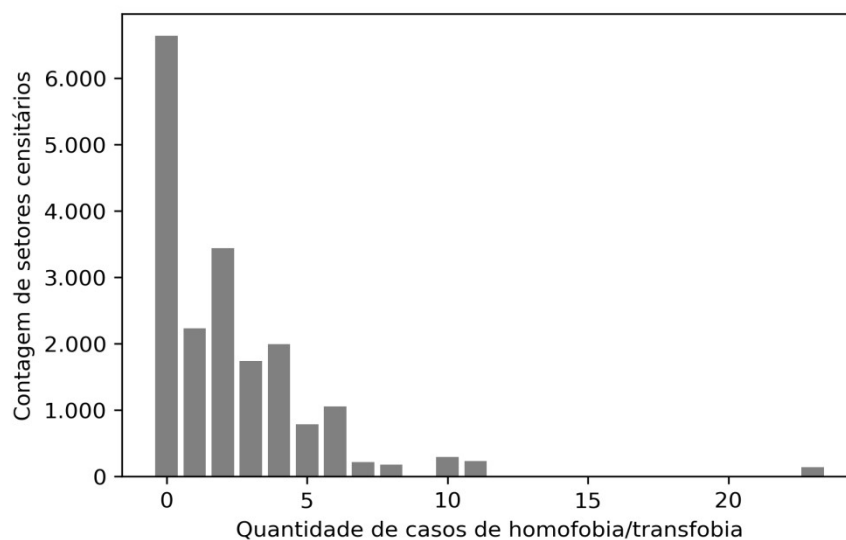
**Tabela 28** Medidas descritivas da quantidade de casos de racismo

Medida	Valor
N	18950
Quantidade de zeros	343
Mínimo	0
1º quartil	9
Mediana	15
3º quartil	22
Máximo	42
Média	16,231
Desvio-padrão	9,063

**Figura 28** Gráfico de barras da quantidade de casos de racismo

**Tabela 29** Medidas descritivas da quantidade de casos de homofobia/transfobia

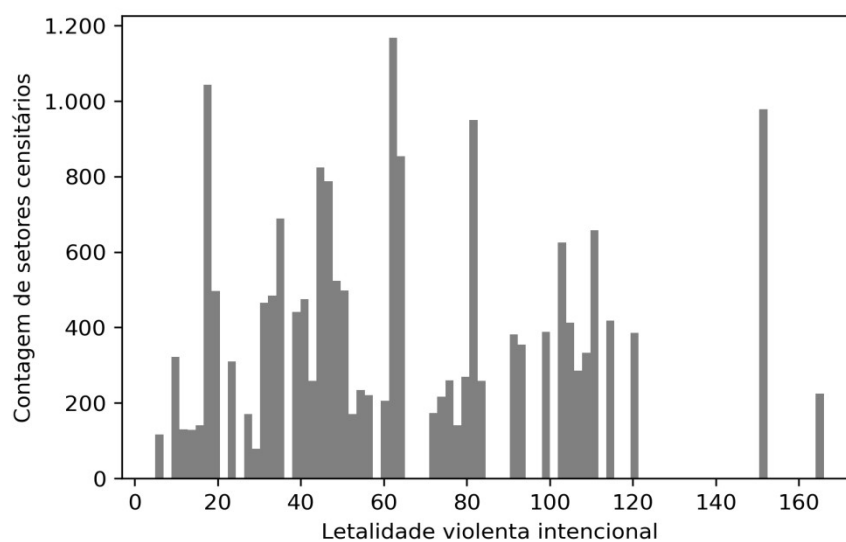
Medida	Valor
N	18950
Quantidade de zeros	6637
Mínimo	0
1º quartil	0
Mediana	2
3º quartil	4
Máximo	23
Média	2,336
Desvio-padrão	3,012

**Figura 29** Gráfico de barras da quantidade de casos de homofobia/transfobia



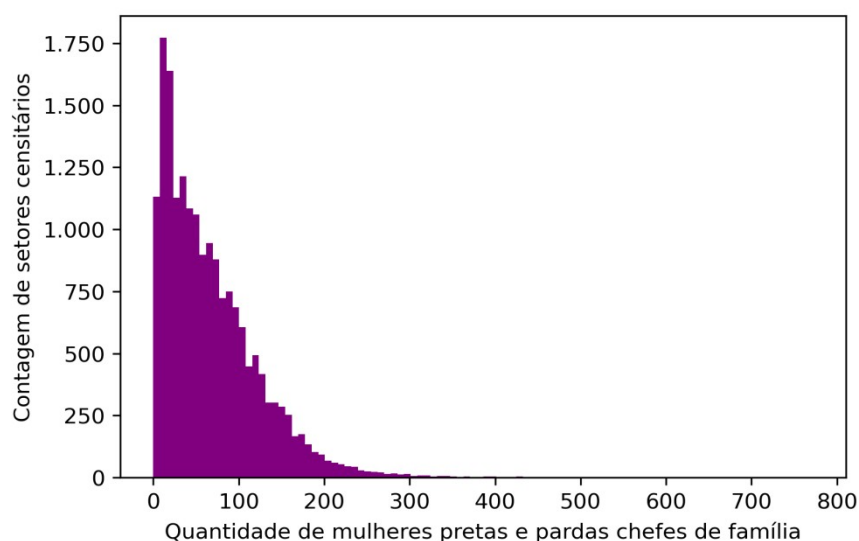
**Tabela 30** Medidas descritivas da letalidade violenta intencional

Medida	Valor
N	18950
Quantidade de zeros	0
Mínimo	5
1º quartil	38
Mediana	62
3º quartil	94
Máximo	166
Média	66,83
Desvio-padrão	37,937

**Figura 30** Histograma da letalidade violenta intencional

**Tabela 31** Medidas descritivas da quantidade de mulheres pretas e pardas chefes de família

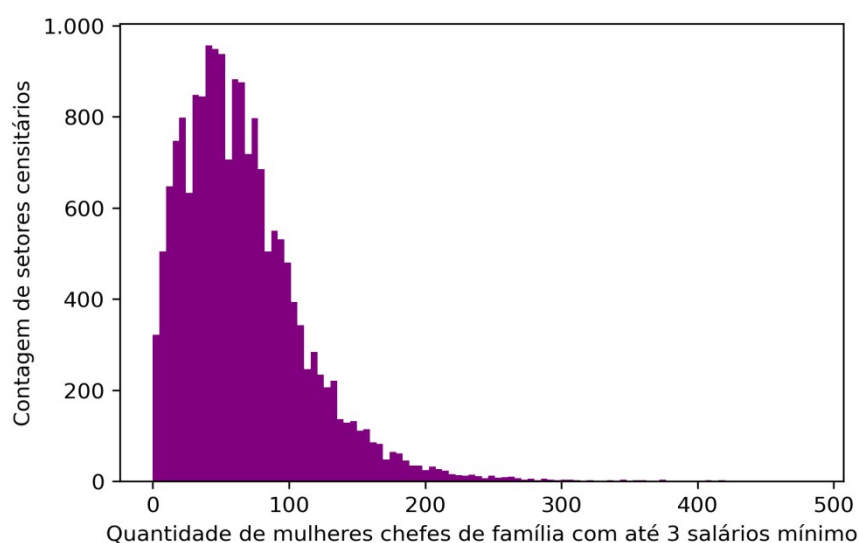
Medida	Valor
N	18206
Quantidade de zeros	55
Mínimo	0
1º quartil	2
Mediana	55
3º quartil	97
Máximo	772
Média	68,415
Desvio-padrão	58,01



**Figura 31** Histograma da quantidade de mulheres pretas e pardas chefes de família

**Tabela 32** Medidas descritivas da quantidade de mulheres chefes de família com até 3 salários mínimos

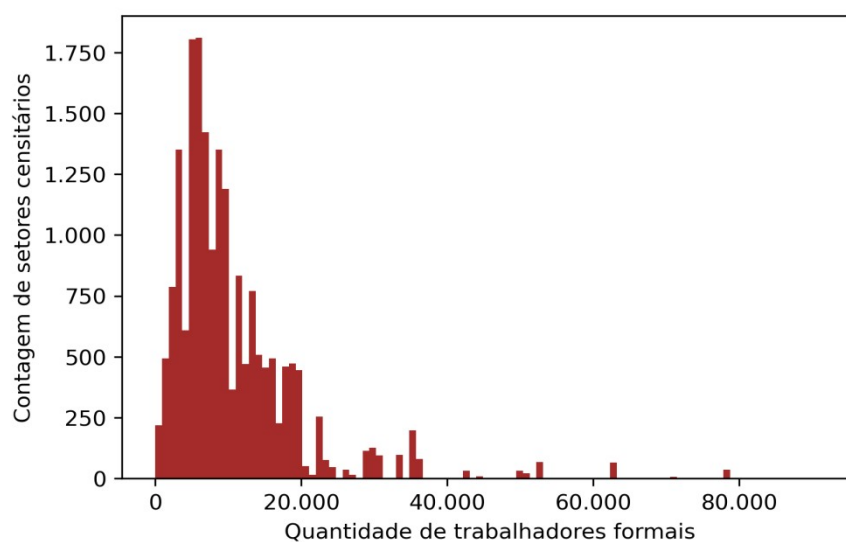
Medida	Valor
N	18206
Quantidade de zeros	50
Mínimo	0
1º quartil	34
Mediana	59
3º quartil	89
Máximo	483
Média	66,742
Desvio-padrão	46,098



**Figura 32** Histograma da quantidade de mulheres chefes de família com até 3 salários mínimos

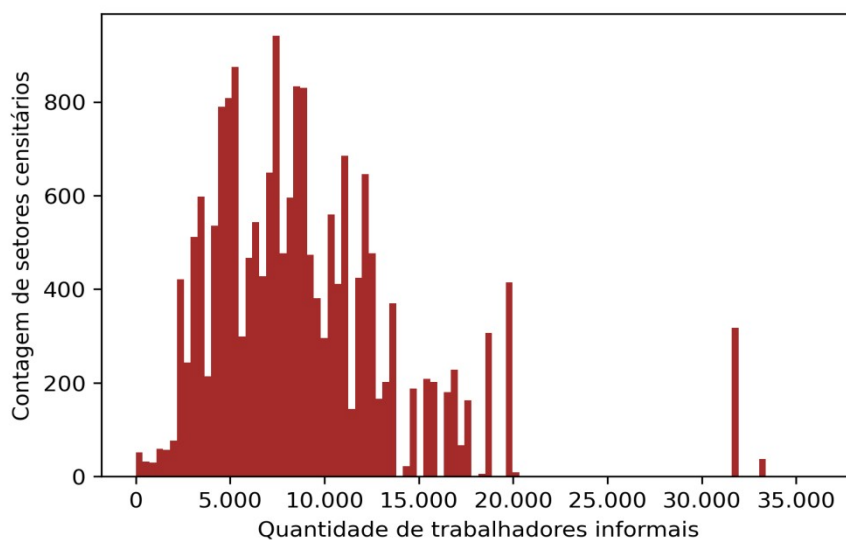
**Tabela 33** Medidas descritivas da quantidade de trabalhadores formais

Medida	Valor
N	18953
Quantidade de zeros	18
Mínimo	0
1º quartil	5253,07
Mediana	8269,04
3º quartil	13284,7
Máximo	91587,84
Média	10587,364
Desvio-padrão	8968,625

**Figura 33** Histograma da quantidade de trabalhadores formais

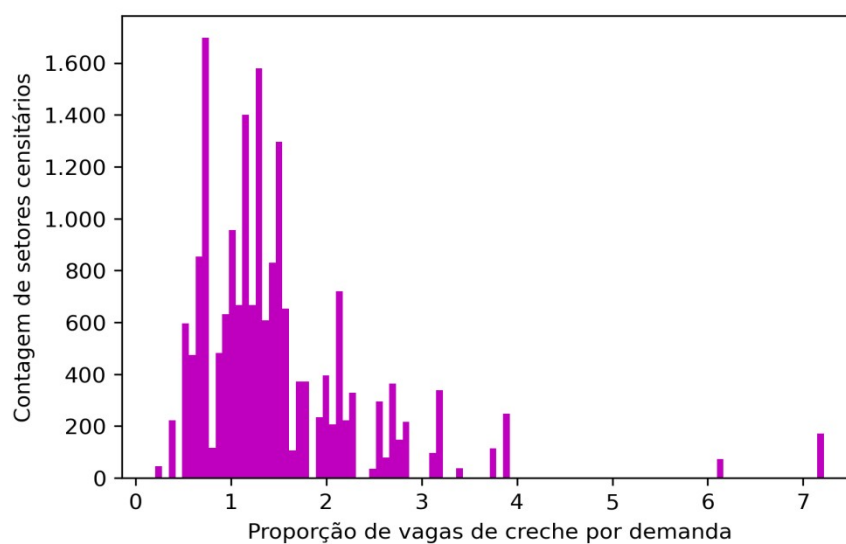
**Tabela 34** Medidas descritivas da quantidade de trabalhadores informais

Medida	Valor
N	18953
Quantidade de zeros	41
Mínimo	0
1º quartil	5214,91
Mediana	8303,44
3º quartil	11363,01
Máximo	36304,31
Média	9039,922
Desvio-padrão	5282,957

**Figura 34** Histograma da quantidade de trabalhadores informais

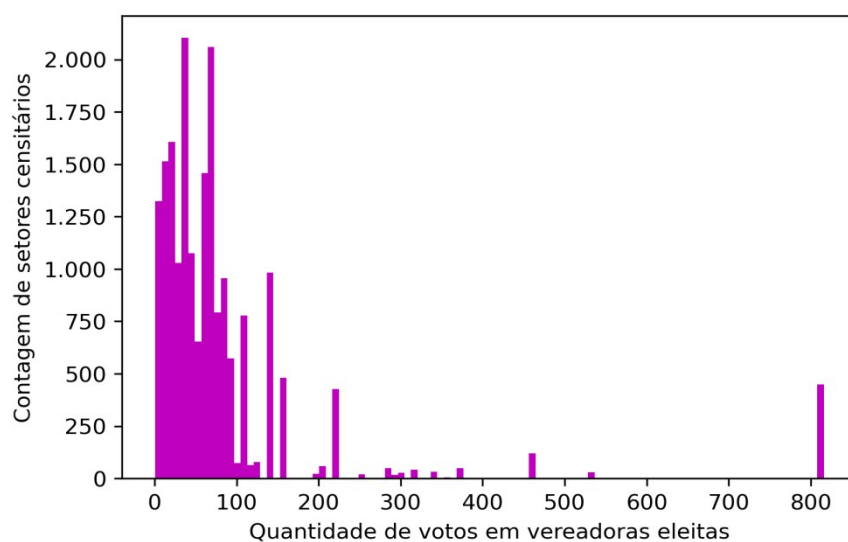
**Tabela 35** Medidas descritivas da proporção de vagas de creche por demanda

Medida	Valor
N	18950
Quantidade de zeros	0
Mínimo	0,208
1º quartil	0,929
Mediana	1,268
3º quartil	1,745
Máximo	7,217
Média	1,493
Desvio-padrão	0,945

**Figura 35** Histograma da proporção de vagas de creche por demanda

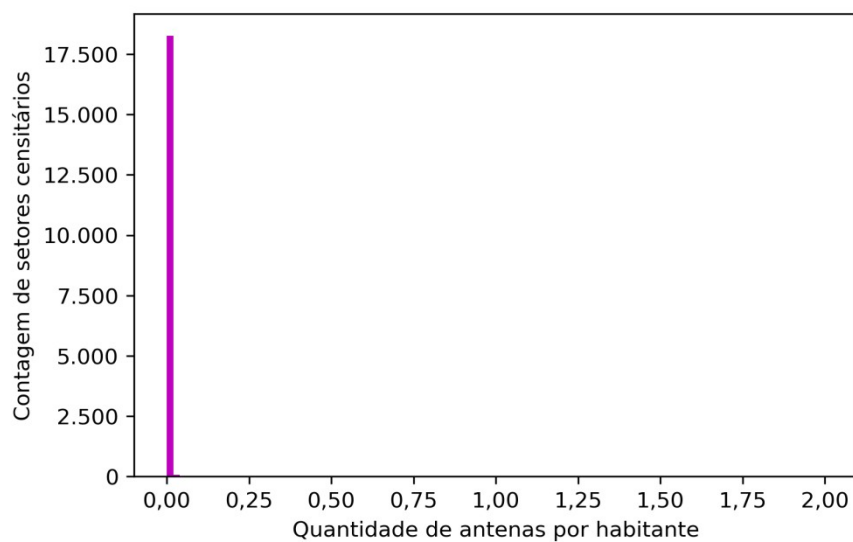
**Tabela 36** Medidas descritivas da quantidade de votos em vereadoras eleitas

Medida	Valor
N	18950
Quantidade de zeros	0
Mínimo	1
1º quartil	27
Mediana	58
3º quartil	87
Máximo	816
Média	85,604
Desvio-padrão	131,183

**Figura 36** Histograma da quantidade de votos em vereadoras eleitas

**Tabela 37** Medidas descritivas da quantidade de antenas por habitante

Medida	Valor
N	18354
Quantidade de zeros	15467
Mínimo	0
1º quartil	0
Mediana	0
3º quartil	0
Máximo	2
Média	0,001
Desvio-padrão	0,018

**Figura 37** Histograma da quantidade de antenas por habitante

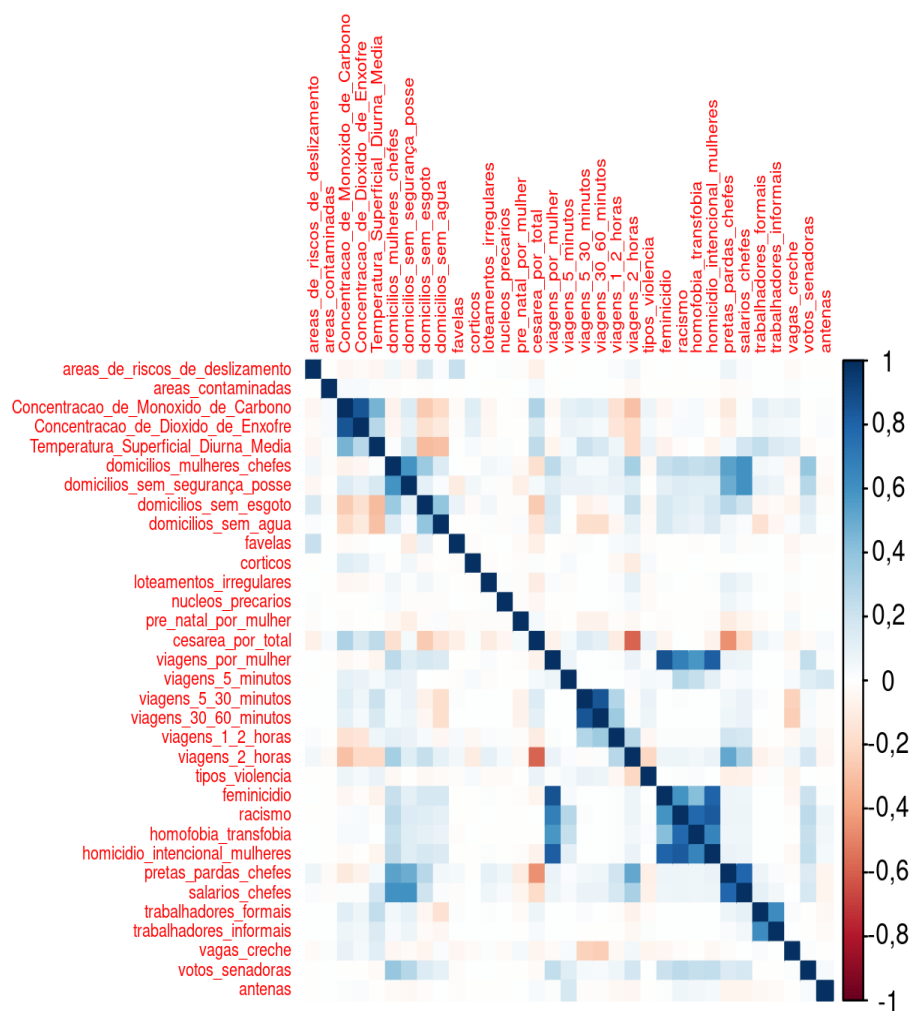


**Tabela 38** Coeficientes provenientes da análise de componentes principais sobre todos os dados

Variável	Coeficiente da análise de componentes principais	
	básica	de segunda ordem
quantidade de áreas de riscos de deslizamento	0,027	0,005
quantidade de áreas contaminadas	-0,011	-0,004
concentração de monóxido de carbono	-0,079	-0,058
concentração de dióxido de enxofre	-0,054	-0,055
temperatura superficial diurna média	-0,055	-0,038
quantidade de domicílios adensados de mulheres chefes de família	0,302	0,324
quantidade de domicílios sem segurança na posse	0,204	0,263
quantidade de domicílios sem ligação de esgoto na rede	0,199	0,255
quantidade de domicílios sem abastecimento de água na rede	0,131	0,182
quantidade de favelas	0,006	-0,007
quantidade de cortiços	-0,012	0,003
quantidade de loteamentos irregulares	0,039	0,047
quantidade de núcleos precários	0,010	-0,0002
quantidade de consultas pré-natal por mulher em idade gestacional	-0,027	-0,087

Variável	Coeficiente da análise de componentes principais	
	básica	de segunda ordem
proporção de partos cesárea quanto ao total de partos	-0,137	-0,087
quantidade de viagens feitas por mulheres	0,354	0,003
quantidade de pessoas cujo tempo de deslocamento é menor do que 5 minutos	0,088	0,002
quantidade de pessoas cujo tempo de deslocamento está entre 5 e 30 minutos	0,010	0,085
quantidade de pessoas cujo tempo de deslocamento está entre 30 e 60 minutos	0,014	0,087
quantidade de pessoas cujo tempo de deslocamento está entre 60 e 120 minutos	0,076	0,056
quantidade de pessoas cujo tempo de deslocamento é maior do que 2 horas	0,194	0,015
proporção de ocorrências de violência contra mulher por população total do setor	-0,028	0,007
quantidade de casos de feminicídio	0,327	0,311
quantidade de casos de racismo	0,345	0,358
quantidade de casos de homofobia/transfobia	0,301	0,318
letalidade violenta intencional	0,366	0,367
quantidade de mulheres pretas e pardas chefes de família	0,238	0,287

Variável	Coeficiente da análise de componentes principais	
	básica	de segunda ordem
quantidade de mulheres chefes de família com até 3 salários mínimos	0,219	0,287
quantidade de trabalhadores formais	0,003	0,014
quantidade de trabalhadores informais	0,003	0,014
proporção de vagas de creche por demanda	-0,020	-0,040
quantidade de votos em vereadoras eleitas	0,208	0,238
quantidade de antenas por habitante	0,006	0,001



**Figura 38** Gráfico da correlação entre as variáveis

**Tabela 39** Coeficientes provenientes da análise de componentes principais para o fator ambiental

Variável	Coeficiente
quantidade de áreas de riscos de deslizamento	-0,057
quantidade de áreas contaminadas	0,048
concentração de monóxido de carbono	0,656
concentração de dióxido de enxofre	0,614
temperatura superficial diurna média	0,432

**Tabela 40** Coeficientes provenientes da análise de componentes principais para o fator habitacional

Variável	Coeficiente
quantidade de domicílios adensados de mulheres chefes de família	0,618
quantidade de domicílios sem segurança na posse	0,502
quantidade de domicílios sem ligação de esgoto na rede	0,486
quantidade de domicílios sem abastecimento de água na rede	0,348
quantidade de favelas	-0,013
quantidade de cortiços	0,006
quantidade de loteamentos irregulares	0,089
quantidade de núcleos precários	-0,0004

**Tabela 41** Coeficientes provenientes da análise de componentes principais para o fator saúde

Variável	Coeficiente
quantidade de consultas pré-natal por mulher em idade gestacional	-0,707
proporção de partos cesárea quanto ao total de partos	-0,707

**Tabela 42** Coeficientes provenientes da análise de componentes principais para o fator mobilidade

Variável	Coeficiente
quantidade de viagens feitas por mulheres	0,023
quantidade de pessoas cujo tempo de deslocamento é menor do que 5 minutos	0,012
quantidade de pessoas cujo tempo de deslocamento está entre 5 e 30 minutos	0,629
quantidade de pessoas cujo tempo de deslocamento está entre 30 e 60 minutos	0,645
quantidade de pessoas cujo tempo de deslocamento está entre 60 e 120 minutos	0,418
quantidade de pessoas cujo tempo de deslocamento é maior do que 2 horas	0,110

**Tabela 43** Coeficientes provenientes da análise de componentes principais para o fator violência

Variável	Coeficiente
proporção de ocorrências de violência contra mulher por população total do setor	0,011
quantidade de casos de feminicídio	0,458
quantidade de casos de racismo	0,527
quantidade de casos de homofobia/transfobia	0,469
letalidade violenta intencional	0,541

**Tabela 44** Coeficientes provenientes da análise de componentes principais para o fator população

Variável	Coeficiente
quantidade de mulheres pretas e pardas chefes de família	0,707
quantidade de mulheres chefes de família com até 3 salários mínimos	0,707

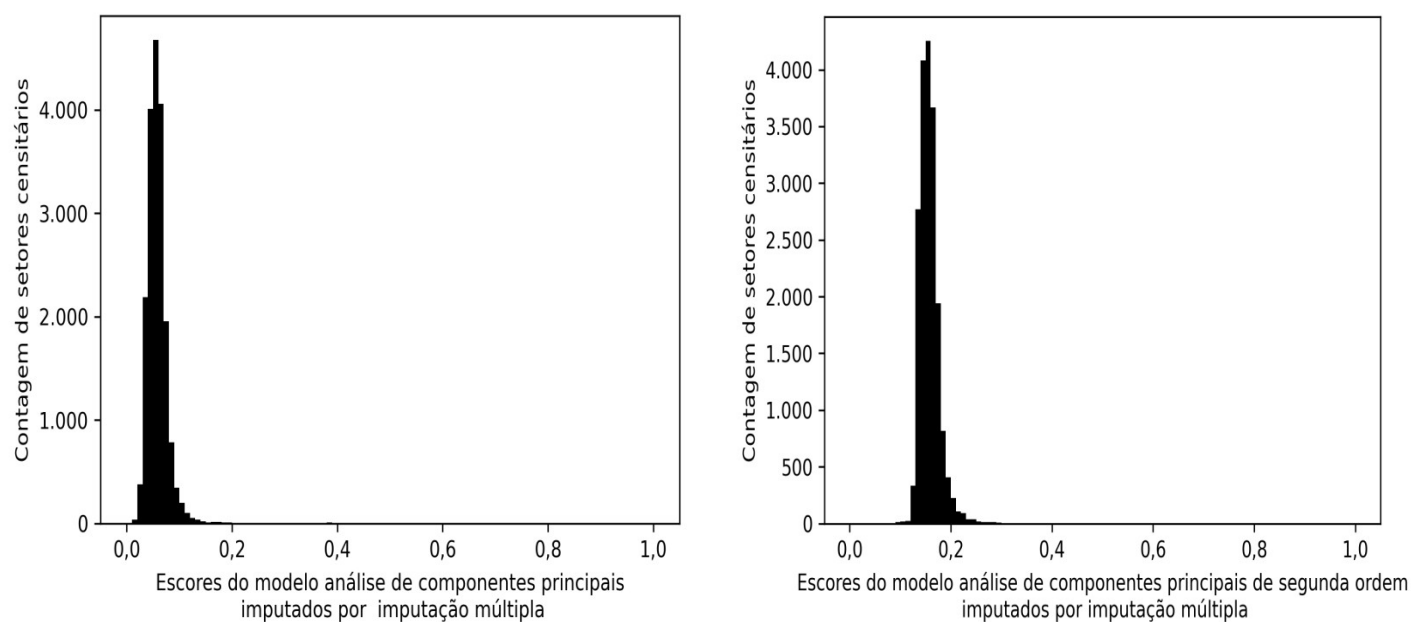


**Tabela 45** Coeficientes provenientes da análise de componentes principais para o fator trabalho

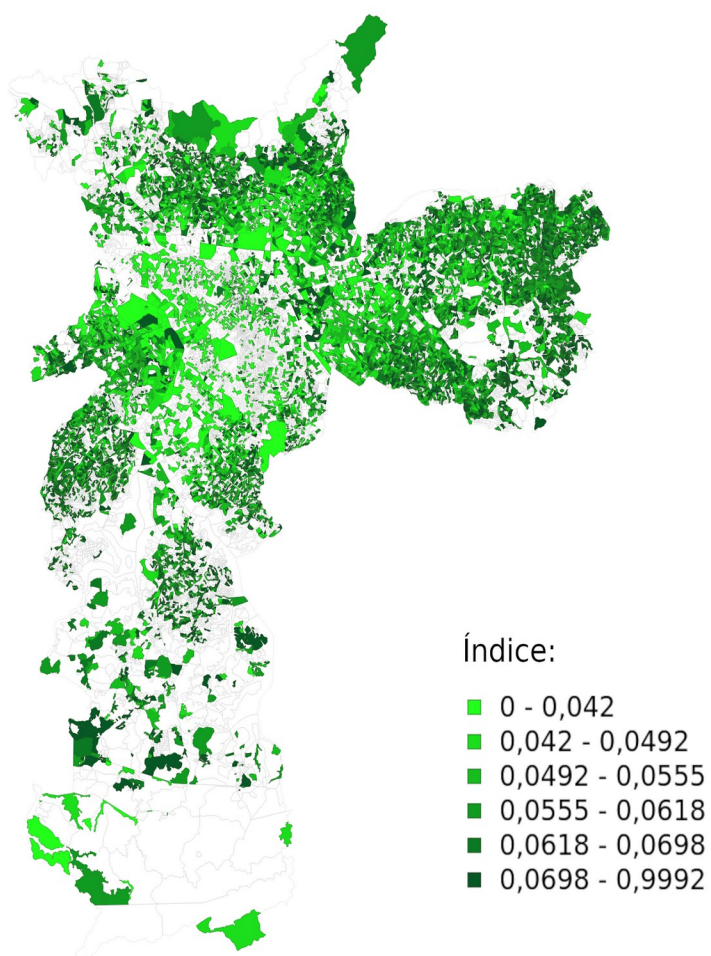
Variável	Coeficiente
quantidade de trabalhadores formais	-0,707
quantidade de trabalhadores informais	-0,707

**Tabela 46** Medidas descritivas dos índices obtidos

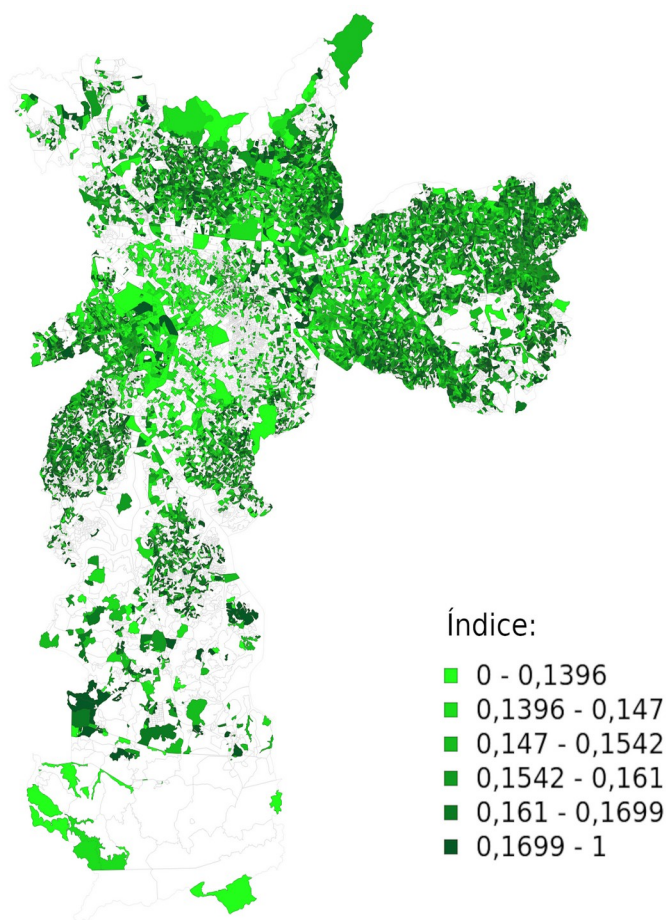
Medida	Índice análise de componentes principais básica	Índice análise de componentes principais de segunda ordem
Mínimo	0,0	0,0
1º quartil	0,046	0,144
Mediana	0,056	0,155
3º quartil	0,067	0,167
Máximo	1,0	1,0
Média	0,059	0,158
Desvio-padrão	0,028	0,027



**Figura 39** Histogramas dos índices-resumo



**Figura 40** Mapa dos setores censitários com o índice gerado pela análise de componentes principais básica



**Figura 41** Mapa dos setores censitários com o índice gerado pela análise de componentes principais de segunda ordem

# ANEXO

### Fonte e ano de coleta de cada variável do estudo

Variável	Fonte	Ano
Quantidade de áreas de riscos de deslizamento	IPT/geosampa	2010
Quantidade de áreas contaminadas	geosampa	2010
Concentração de monóxido de carbono (ppm)	instituto de energia e meio ambiente	2015
Concentração de dióxido de enxofre (mg/m3)	instituto de energia e meio ambiente	2015
Temperatura superficial diurna média (°C)	Schwandner, 2019	2010
Quantidade de domicílios adensados de mulheres chefes de família	IBGE	2010
Quantidade de domicílios sem segurança na posse	IBGE	2010
Quantidade de domicílios sem ligação de esgoto na rede	IBGE	2010
Quantidade de domicílios sem abastecimento de água na rede	IBGE	2010
Quantidade de favelas	geosampa	2020
Quantidade de cortiços	geosampa	2020
Quantidade de loteamentos irregulares	geosampa	2020
Quantidade de núcleos precários	geosampa	2020
Quantidade de assentamentos precários	geosampa	2020
Quantidade de consultas pré-natal	tabnet datasus PMSP	2010
Proporção de consultas pré-natal por mulher em idade gestacional	tabnet datasus PMSP	2010
Quantidade de partos cesárea	tabnet datasus PMSP	2010
Quantidade total de partos	tabnet datasus PMSP	2010
Proporção de partos cesárea quanto ao total de partos	tabnet datasus PMSP	2010
Quantidade de viagens feitas por mulheres	OD 2007	2007
Quantidade de pessoas cujo tempo de deslocamento é menor do que 5 minutos	IBGE Censo Amostra	2010
Quantidade de pessoas cujo tempo de deslocamento está entre 5 e 30 minutos	IBGE Censo Amostra	2010
Quantidade de pessoas cujo tempo de deslocamento está entre 30 e 60 minutos	IBGE Censo Amostra	2010
Quantidade de pessoas cujo tempo de deslocamento está entre 60 e 120 minutos	IBGE Censo Amostra	2010
Quantidade de pessoas cujo tempo de deslocamento é maior do que 120 minutos	IBGE Censo Amostra	2010
Proporção de ocorrências de violência contra mulher por população total do setor	SSP-SP	2016
Quantidade de casos de feminicídio	SSP-SP	2016
Quantidade de casos de racismo	SSP-SP	2016
Quantidade de casos de homofobia/transfobia	SSP-SP	2016
Letalidade violenta intencional	SSP-SP	2016
Quantidade de mulheres pretas e pardas chefes de família	IBGE	2010
Quantidade de mulheres chefes de família com até 3 salários mínimos	IBGE	2010
Quantidade de trabalhadores formais	OD 2007	2007
Quantidade de trabalhadores informais	OD 2007	2007
Proporção de vagas de creche por demanda	PMSP	2010
Quantidade de votos em vereadoras eleitas na eleição de 2012	TSE/ TRESP	2012
Quantidade de antenas por habitante	anatel	2021