

Datathon de Mobilidade Urbana 2024

Secretaria Nacional de Mobilidade Urbana (SEMOB) - Ministério das Cidades (MCID)

Plataforma de Ciência de Dados aplicada à Saúde (PCDaS)

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict) - Fundação Oswaldo Cruz (Fiocruz)

Vinícius da Fonseca Vieira

<https://github.com/vfvieira/datathon2024>

vfvieira@gmail.com

Junho de 2024

Contexto

Datathon de Mobilidade Urbana 2024, promovido pela Secretaria Nacional de Mobilidade Urbana (SEMOB) do Ministério das Cidades (MCID), em parceria com a Plataforma de Ciência de Dados aplicada à Saúde (PCDaS), iniciativa do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict) da Fundação Oswaldo Cruz (Fiocruz).

Desafio 1: Análise dos empreendimentos em mobilidade urbana da Carteira de Empreendimentos da MCID/SEMOB e sua correlação no território com os indicadores de mortos e feridos no trânsito mantidos pelo Atlas da Violência, considerando o tipo de programa do empreendimento.

Motivação da proposta

Tomando como base a **Carteira de Empreendimentos da MCID/SEMOB** e os **indicadores de mortos e feridos no trânsito mantidos pelo Atlas da Violência**.

Hipótese: Os indicadores de mortos e feridos nos municípios são diretamente impactados pelos empreendimentos de mobilidade urbana realizados pelo MCID.

Questões principais:

Questão 1: Como se caracteriza, ao longo do tempo, a relação entre os empreendimentos em mobilidade urbana realizados do MCID e os principais indicadores de mortos e feridos no trânsito?

Questão 2: Como essa relação varia em razão do programa ao qual os empreendimentos estão associados?

Questão 3: Como os dados sobre os empreendimentos de mobilidade podem ser utilizados para propor a redução em indicadores de mortos e feridos no trânsito?

Considerações iniciais

Foi utilizada a terminologia dos documentos oficiais.

Tentei me limitar aos dados das bases de dados descritas na página do Datathon 2024

Busquei uma solução mais generalista possível (sem filtros de escala, regiões ou parâmetros).

Foi utilizada a linguagem Python 3 e as bibliotecas:

- numpy
- scipy.stats
- pandas
- matplotlib
- synchrony

Ambiente computacional:

- PC com Ubuntu 22.04
- 32 Gb de memória
- Intel Core i9

Toda a solução leva cerca de 20 segundos para ser executada.

Uma breve investigação sobre os dados

Carteira de Empreendimentos da MCID/SEMOB:

Cada linha representa um empreendimento de mobilidade urbana distinto.

Foram selecionadas as variáveis:

- **programa** (nome do programa relacionado à obra, texto);
- **Código IBGE** (código do IBGE dos municípios onde a obra está localizada, valor inteiro);
- **pop_beneficiada** (população beneficiada pela obra, valor inteiro);
- **vlr_investimento** (valor total do investimento, valor real);
- **ano_fim_obra** (ano previsto de conclusão da obra, valor inteiro).

Mortos e feridos (Atlas de Violência):

Cada linha apresenta os indicadores de mortos e feridos em cada município Brasileiro em cada ano.

Foram selecionadas as variáveis:

- **Código IBGE** (código do IBGE dos municípios onde a obra está localizada, valor inteiro);
- **ano** (ano de registro dos indicadores no município, valor inteiro);
- **total_mortes** (total de mortes no município em um determinado ano);
- **total_feridos** (total de feridos no município em um determinado ano).

Sistema IBGE de Recuperação Automática (SIDRA):

Alto índice de valores faltantes na coluna **População** (Carteira de Empreendimentos) e um alto índice de valores com erros nas colunas **taxa_mun_mortes** e **taxa_mun_feridos** na base de acidentes.

Informações sobre a população dos municípios foram obtidas a partir de estimativas populacionais descritas no SIDRA.

Assim é possível encontrar valores proporcionais às populações das cidades para as propriedades estudadas.

Uma breve investigação sobre os dados

Escopo considerado

| Ano | AVANÇAR PÚBLICO | EMENDAS | PAC FIN | PAC OGU | PRÓ COMUNIDADE | SETOR PRIVADO |
|------|-----------------|---------|---------|---------|----------------|---------------|
| 2000 | | | | | 2 | |
| 2002 | | | | | 1 | |
| 2003 | | | | | 1 | |
| 2008 | | | | | | 1 |
| 2012 | | | 5 | | | 1 |
| 2013 | | | 13 | | | |
| 2014 | | | 32 | | | |
| 2015 | | 22916 | 37 | 2 | | |
| 2016 | | 1111 | 50 | 1 | | |
| 2017 | | 1748 | 66 | 4 | 1 | 1 |
| 2018 | | 2398 | 53 | 4 | | 51 |
| 2019 | | 3046 | 103 | 4 | | 42 |
| 2020 | 6 | 3467 | 88 | 6 | | 7 |
| 2021 | 41 | 3533 | 46 | 5 | | |
| 2022 | 43 | 2943 | 36 | 3 | | 10 |
| 2023 | 13 | 958 | 3 | 9 | | |

Mortos e feridos (Atlas de Violência):

Dados entre 2010 e 2019.

Carteira de empreendimentos:

Linhas com valor ausente nas colunas **ano_fimobra** ou **programa** foram desconsideradas.

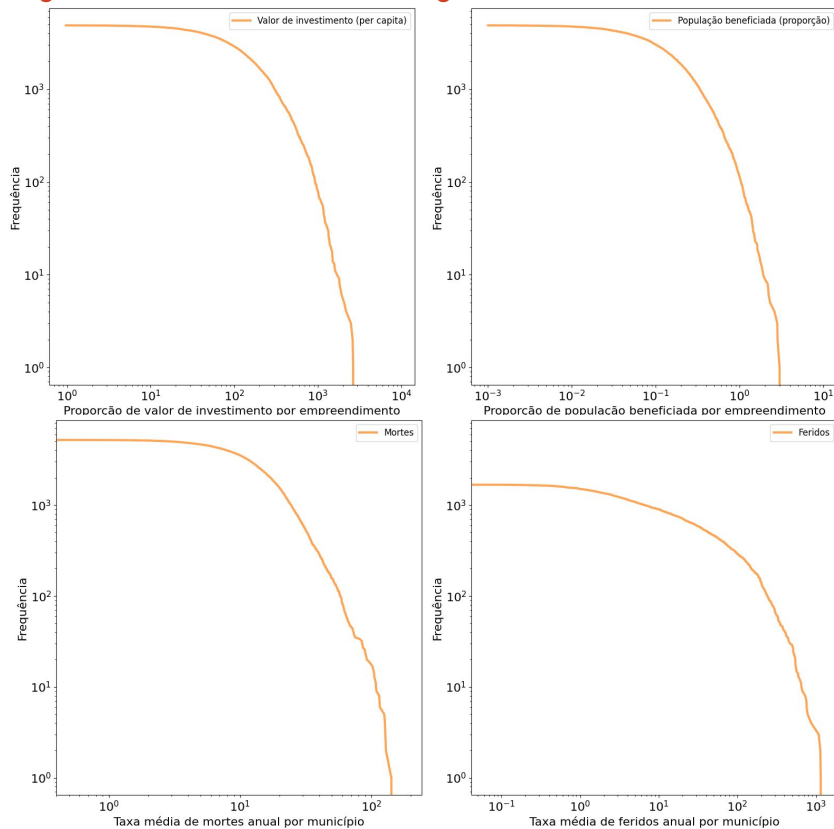
Anos com poucos registros foram desconsiderados.

Assim, ficamos com os registros referentes aos programas **EMENDAS**, **PAC FIN** e **SETOR PRIVADO** no intervalo entre os anos de **2015 a 2019**.

A análise da tabela revela um volume significativamente maior de programas do tipo EMENDA em todos os anos, mas especialmente no ano de 2015.

Uma breve investigação sobre os dados

Função de Distribuição Cumulativa Complementar



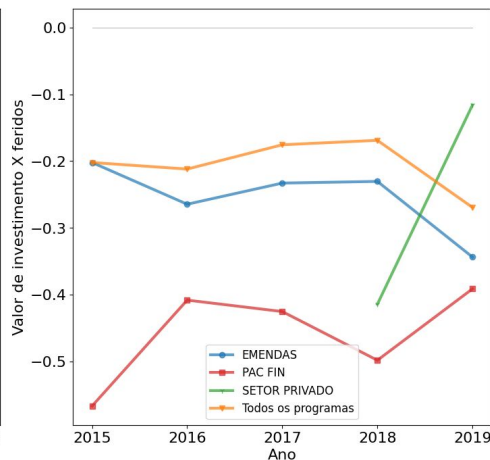
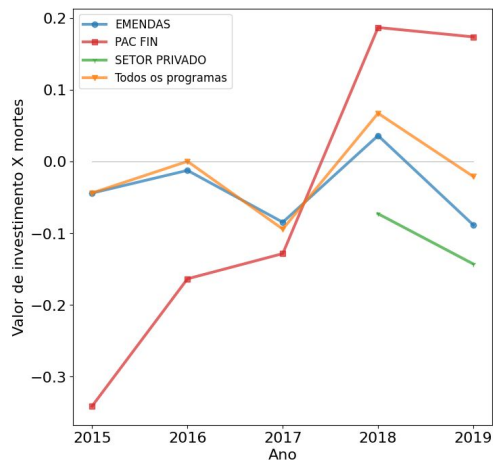
Investigação visual das funções de distribuição das variáveis sugeriu uma alta redundância entre as propriedades **vlr_investimento** e **pop_beneficiada**.

Isso foi confirmado pela alta correlação de Spearman entre as variáveis ($=0.97$).

Optou-se por abandonar a propriedade **pop_beneficiada** das análises (não traz informação que já não esteja contida em **vlr_investimento**).

Correlação estática

Relacionando propriedades em uma mesma época



Objetivo: correlacionar o **valor de investimento** (calculado *per capita*) em cada cidade com os indicadores de **mortes** e **feridos** em cada ano.

Método: Correlação entre as valores de propriedades, estratificados por programas, utilizando correlação de **Spearman**.

O que o resultado nos mostra:

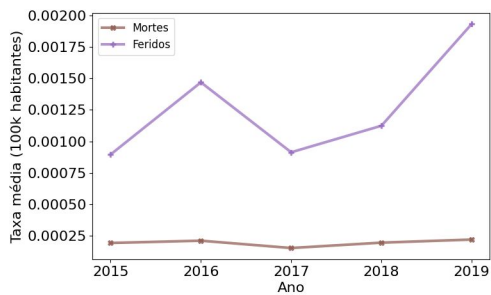
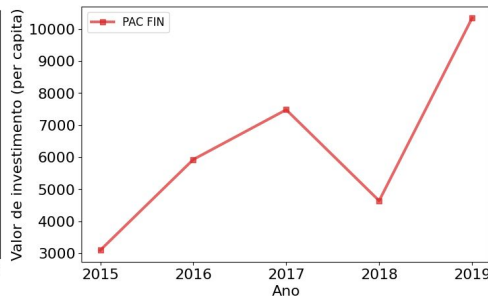
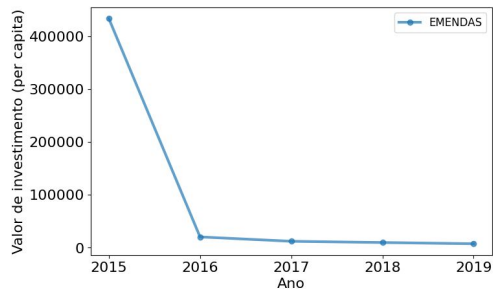
Não é possível observar uma correlação significativa entre o valor de investimento e o número de mortes e feridos quando consideramos uma correlação de seus valores em cada um dos anos.

Há uma correlação moderada entre o valor de investimento e o número de feridos no ano de 2015 para o programa PAC FIN (mesmo com uma frequência significativamente inferior a empreendimentos provenientes de EMENDAS).

Necessidade de visões alternativas sobre os dados!

Correlação dinâmica

Relacionando propriedades em épocas diferentes



Objetivo: verificar se o investimento em empreendimentos de mobilidade impacta os indicadores de mortos e feridos no futuro (e quando).

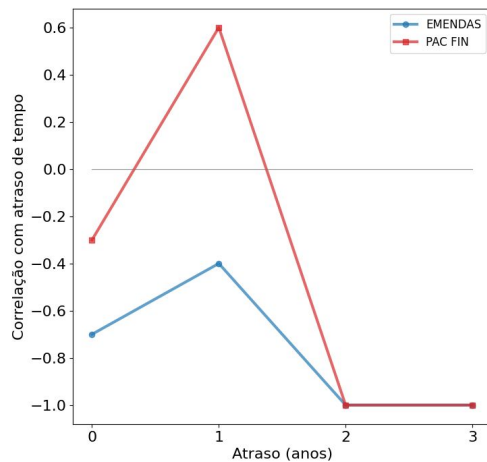
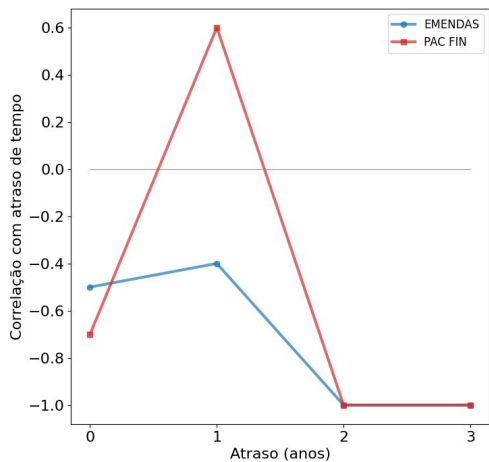
Passo preliminar: criação de séries temporais para as propriedades de interesse (vlr_investimento, mortes e feridos).

Cada série temporal armazena, por ano, a média dos valores da respectiva propriedade.

Método: Correlação Cruzada com Atraso de Tempo (*Time-Lagged Cross Correlation* - TLCC) entre as séries temporais utilizando correlação de **Spearman**.

Correlação dinâmica

Relacionando propriedades em épocas diferentes



O que o resultado nos mostra:

Mesmo que a correlação entre as propriedades seja apenas moderada quando analisamos as propriedades nos municípios ano-a-ano, há uma correlação entre as propriedades quando são analisadas de forma agregada ao longo do tempo.

O resultado sugere que um empreendimento de mobilidade urbana **não tem impacto imediato** nos indicadores de mortos e feridos, mas isso ocorre após um intervalo de **2 a 3 anos**.

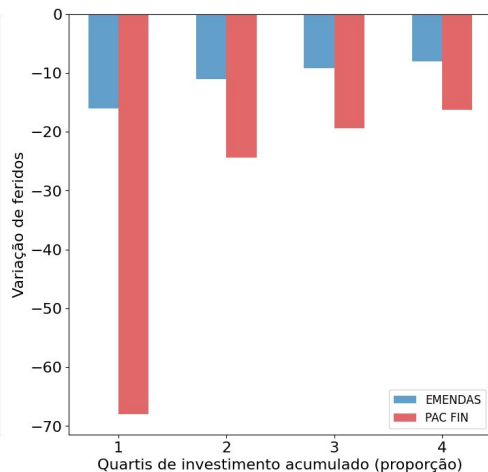
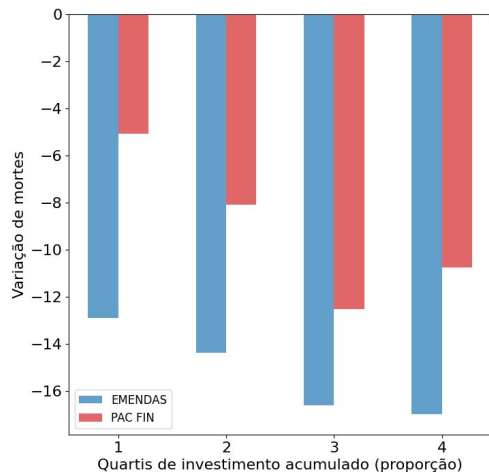
Isso reforça o caráter estratégico e de mais longo prazo da aplicação de recursos públicos para a melhoria da qualidade de vida de uma população.

Esse impacto não tem distinção quanto ao programa ao qual está relacionado o empreendimento.

Obs.: Para este estudo e os próximos, abandonamos também o programa SETOR PRIVADO, que apenas apresenta registros em dois anos, impedindo a avaliação dinâmica de seus empreendimentos.

Uma propriedade sob a perspectiva de outra

Avaliação sob a perspectiva de valor de investimento

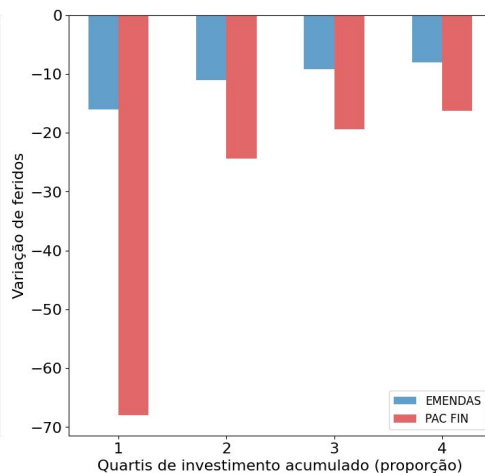
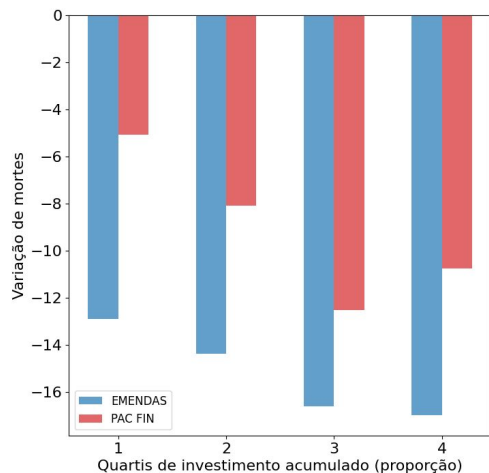


Objetivo: Avaliar como as propriedades **mortes** e **feridos** estão distribuídas por municípios com diferentes escalas da propriedade **valor de investimento**.

Método: Foi calculado, por município, o valor de investimento acumulado, estratificado por programa. Em seguida, foram calculados os quartis (Q1, Q2, Q3 e Q4) dos valores. Para os municípios em cada um desses quartis (eixo-x das figuras) foram calculadas as variações de mortes e feridos no intervalo entre 2015 e 2019 (eixo-y) para os municípios com redução de mortes e feridos.

Uma propriedade sob a perspectiva de outra

Avaliação sob a perspectiva de valor de investimento



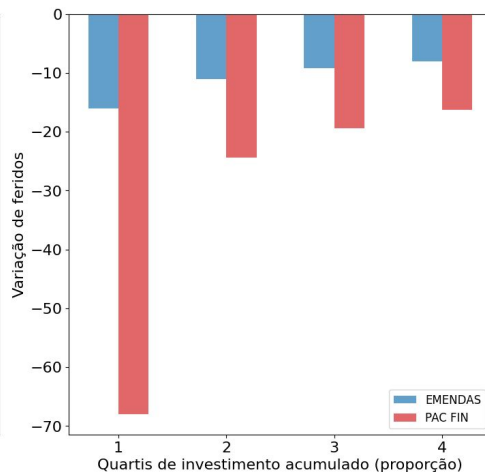
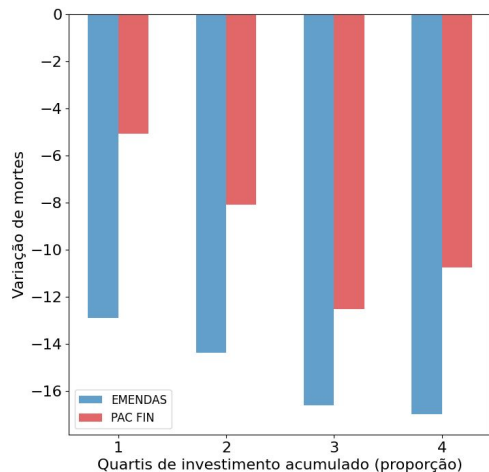
O que o resultado sobre mortes nos mostra (esquerda):

Os quartis superiores apresentaram variações significativamente mais importantes do que os quartis inferiores. Isso reforça a verificação da hipótese de que um maior investimento *per capita* em empreendimentos de mobilidade está relacionado à redução de mortes em uma janela de tempo de médio e longo prazo.

Mais ainda, é possível perceber que os programas do tipo EMENDA, que possuem um número muito superior de registros, trazem impacto significativamente maior do que os programas do tipo PAC FIN, especialmente em quartis inferiores.

Uma propriedade sob a perspectiva de outra

Avaliação sob a perspectiva de valor de investimento



O que o resultado sobre feridos nos mostra (direita):

Um resultado muito diferente pode ser visto para os resultados sobre feridos.

Os quartis inferiores apresentam variações significativamente mais importantes do que os quartis da direita. Além disso, investimentos do tipo PAC FIN trazem um impacto muito superior aos do tipo EMENDAS.

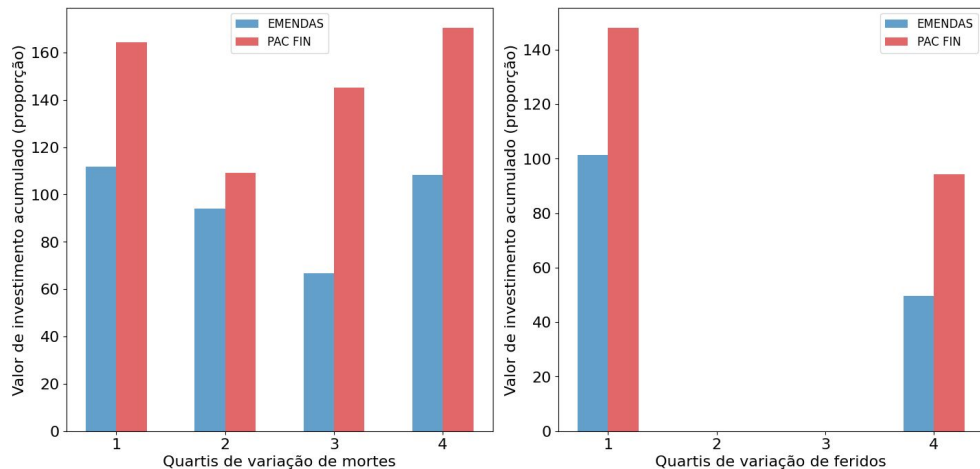
Isso pode ocorrer por alguns motivos.

Um deles é o alto número de valores ausentes de registros de feridos na base de dados.

Outro motivo é a possibilidade de que obras de menor impacto, portanto, de menor valor de investimento, sejam, de fato mais eficientes na redução de acidentes menos graves, que não acarretam em vítimas fatais.

Uma propriedade sob a perspectiva de outra

Avaliação sob a perspectiva de mortes e feridos

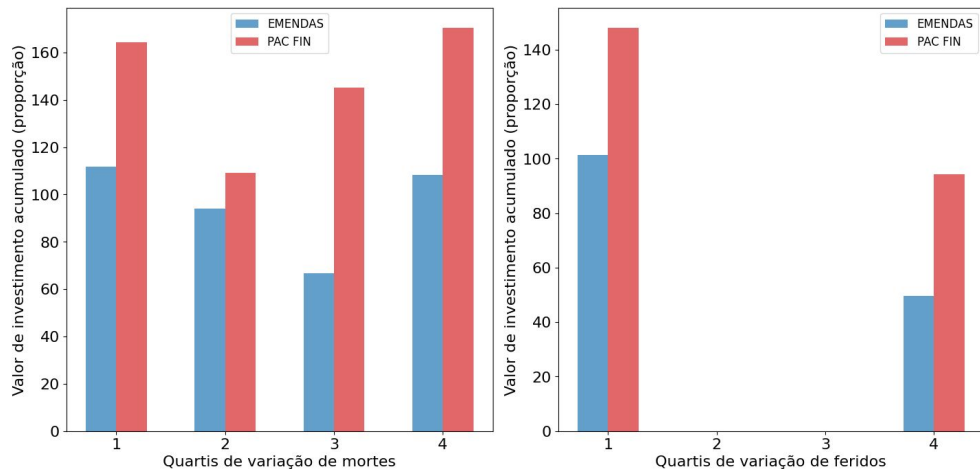


Objetivo: Avaliar como a propriedade **valor de investimento** está distribuída por municípios com diferentes escalas das propriedades **mortes e feridos**.

Método: Foi calculado, por município, a variação de mortes e feridos acumulados, estratificados por programa. Em seguida, foram calculados os quartis (Q1, Q2, Q3 e Q4) dos valores. Para os municípios em cada um desses quartis (eixo-x das figuras) foram calculadas os valores de investimento acumulados no intervalo entre 2015 e 2019 (eixo-y).

Uma propriedade sob a perspectiva de outra

Avaliação sob a perspectiva de mortes e feridos



O que o resultado sobre mortes nos mostra (esquerda):

Quando olhamos sob a perspectiva da variação de mortes, os municípios no quartil 1 e 4 (ou seja, os municípios com maior e menor redução de mortes, respectivamente) são aqueles com maior valor de investimento acumulado.

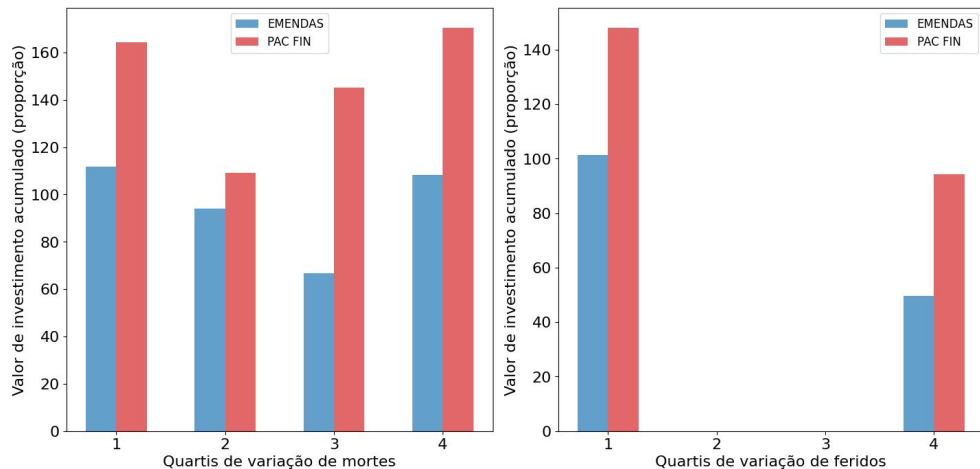
É possível que as maiores reduções de mortes, observadas em municípios do quartil 1, estejam relacionadas às maiores taxas de investimento em empreendimentos nesses municípios.

É possível também que municípios em outros quartis, com menores reduções em mortes, sejam aqueles que já possuem valores baixos de mortes no início do intervalo, eliminando as relações entre as variáveis nesse recorte.

Assim, embora seja possível elaborar algumas conjecturas sobre esses resultados, não é possível extrair conclusões definitivas.

Uma propriedade sob a perspectiva de outra

Avaliação sob a perspectiva de mortes e feridos



O que o resultado sobre feridos nos mostra (direita):

Sob a perspectiva dos municípios organizados pelos números de feridos, chama a atenção a ausência de valores nos quartis Q2 e Q3.

Isso ocorre devido a valores nulos de feridos registrados em um elevado número de municípios, que pode ser ocasionado por uma subnotificação de vítimas de acidentes não-fatais.

Mesmo assim, quando observamos os quartis Q1 e Q4, notamos que a hipótese é verificada e que municípios com maiores reduções em feridos apresentam maiores valores de investimento.

“E se...”

Explorando cenários projetados

Após analisar os indicativos de mortes e feridos nos municípios que receberam empreendimentos de mobilidade urbana, este estudo propõe uma visão sobre os municípios que **não foram contemplados** com tais empreendimentos.

Objetivo: Investigar a projeção de mortes e feridos em municípios que não foram contemplados com empreendimentos de mobilidade urbana com base em informações sobre municípios que foram contemplados.

Método:

- 1) Listar todos os municípios da base do Atlas de Violência que **não** estão contemplados na carteira de empreendimentos.
- 2) Para cada um desses municípios, descobrir, na carteira de empreendimentos, aquele que tem a taxa de mortes (ou feridos) mais parecida com a sua em 2015 (ano do início do intervalo). Esse será considerado seu **município mais próximo**.
- 3) Observar a taxa de variação de mortes (ou feridos) do seu município mais próximo entre 2015 e 2019.
- 4) Aplicar ao município não contemplado a taxa de variação de mortes (ou feridos) de seu município mais próximo.
- 5) Observar o número absoluto de mortes (ou feridos) que o município não contemplado teria em 2019 nesse cenário projetado.
- 6) Comparar o número de mortes (ou feridos) projetado com o valor real registrado na base de indicadores. Essa será a diferença entre o cenário projetado e o cenário real.

“E se...”

Explorando cenários projetados

Exemplo:

- O município de código **3303807** tem 40478 habitantes e não foi contemplado na carteira de empreendimentos.
- A taxa de mortes de **3303807** em 2015 é **93.8781** e em 2019 é **108.8845** por 100k habitantes.
- O município com taxa mais próxima é o de código **5300108**, com taxa 99.6970 por 100k habitantes em 2015.
- O município mais próximo teve uma **redução de 12.3085 mortes por 100k habitantes** em 2019, passando para 87.3886 mortes por 100k habitantes.
- Quando aplicamos essa redução ao município não contemplado, sua taxa projetada em 2019 passa a ser 96.5760.
- Ou seja, nesse cenário projetado, **4.9822 mortes teriam sido evitadas apenas em 2019 no município 3303807.**

“E se...”

Explorando cenários projetados

A base de indicadores de acidentes do Atlas de Violência apresenta **5570** municípios.

Desses, **4898 estão contemplados** na carteira de empreendimentos da SEMOB.

Ou seja, há **672 municípios que não estão contemplados**.

Desses, **433** apresentam registros de morte em 2015 e **126** apresentam registros de feridos em 2015.

Quando aplicamos o método de projeção a esses municípios, temos:

Mortes evitadas

134 mortes seriam evitadas nos **433 municípios apenas em 2019**
(a um custo de R\$41,58* por pessoa em empreendimentos de mobilidade urbana)

Feridos evitados

57 feridos seriam evitados nos **126 municípios apenas em 2019**
(a um custo de R\$57,31* por pessoa em empreendimentos de mobilidade urbana)

*Tomando como base o valor de investimentos dos empreendimentos dos municípios mais próximos no intervalo entre 2015 e 2019.

Considerações finais

Cidades são sistemas complexos e indicadores de violência no trânsito, como mortes e feridos, não estão relacionados apenas a empreendimentos de mobilidade em um intervalo de 5 anos, mas, pelo menos, em toda uma política pública de construção de infraestrutura de mobilidade por um longo prazo.

Mesmo assim, foi possível perceber que quando consideramos um atraso de 2 a 3 anos, os empreendimentos, representados pelo seu valor de investimento, têm impacto significativo sobre acidentes no trânsito.

O estudo sobre os municípios que não foram contemplados com empreendimentos de mobilidade urbana no intervalo estudado, nos ajuda a ter uma noção mais concreta das vidas que poderiam ser salvas, caso tais empreendimentos tivessem sido implementados.

Assim, é possível argumentar que o investimento em empreendimentos de mobilidade urbana é fundamental, não só para a melhoria da qualidade de vida humana nos centros urbanos do o ponto de vista de locomoção e garantia do direito à cidade, mas do ponto de vista de indicadores de saúde, elevando a dignidade das pessoas na nossa sociedade.

Datathon de Mobilidade Urbana 2024

Secretaria Nacional de Mobilidade Urbana (SEMOB) - Ministério das Cidades (MCID)

Plataforma de Ciência de Dados aplicada à Saúde (PCDaS)

Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (Icict) - Fundação Oswaldo Cruz (Fiocruz)

Obrigado! :)

Vinícius da Fonseca Vieira

<https://github.com/vfvieira/datathon2024>

vfvieira@gmail.com

Junho de 2024