

Impacto dos Empreendimentos de Mobilidade Urbana nos Acidentes de Trânsito: Uma Análise Baseada em dados do SIMU

Richard Sousa Antunes | Ana Yoon Faria de Lima

Junho, 2024

Datathon de Mobilidade Urbana 2024

SIMU | SEMOB | MCID | PCDaS | Icict-Fiocruz

Sumário

- ➊ **Análise exploratória de dados**
- ➋ **Tratamento de dados**
- ➌ **Modelos**
- ➍ **Conclusões**

Sumário

- ➊ **Análise exploratória de dados**
- ➋ Tratamento de dados
- ➌ Modelos
- ➍ Conclusões

Base de acidentes

- ▶ O número total de mortos diminui gradativamente ao longo dos anos enquanto total de feridos aumenta.

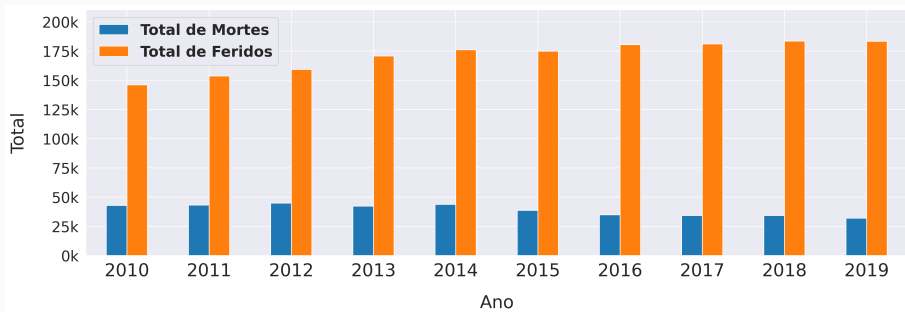


Figura 1: *Total de mortes e feridos por ano em todos os municípios brasileiros.*

- ▶ Regiões N, NE e CO possuem maior aumento no número de feridos.

Caso de exemplo: São Paulo

- **São Paulo fica em 1º lugar no número de mortes de 2010 a 2015 e em 2019, porém de 2016 a 2018 não fica nem no top 5.**
 - ▶ Investimentos em mobilidade em São Paulo tiveram um grande aumento em 2014 e podem estar correlacionados com essa melhora.
 - ▶ Além disso, esse período coincide com a diminuição do limite de velocidade de marginais na cidade e com maior fiscalização da Lei Seca.

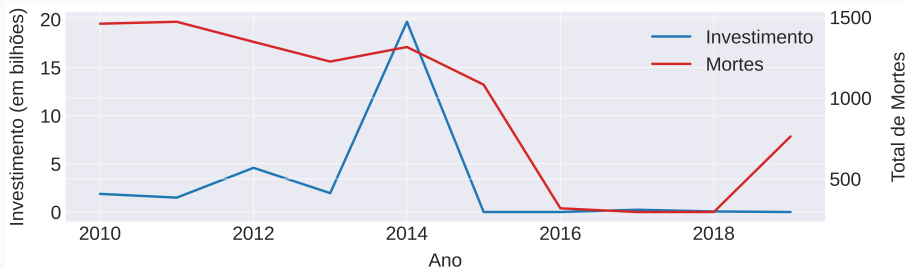


Figura 2: *Investimentos versus número de mortes em São Paulo.*

Proporção de acidentes por meio de transporte

- ▶ Motociclistas representam cerca de metade do total de feridos.

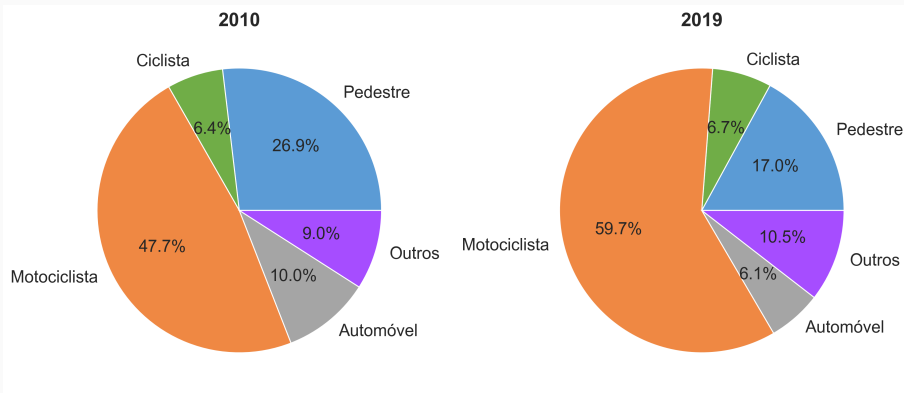


Figura 3: *Porcentagem de feridos por categoria de modo de transporte.*

- ▶ Proporção de motociclistas feridos vêm aumentando ao longo dos anos, enquanto ciclistas e pedestres feridos diminuíram.

Sumário

- ① Análise exploratória de dados
- ② Tratamento de dados**
- ③ Modelos
- ④ Conclusões

Tratamento de dados

- **Foi realizado agrupamentos a fim de analisar os dados de forma mais resumida e detectar relações.**
 - ▶ Por município e ano.
- **Tratamento de dados nulos.**
 - ▶ 47 % da base tinha a maioria das colunas de município nulas.
- **Tratamentos de empreendimentos de forma a categorizá-los para conseguir relacionar com mortos e feridos.**
 - ▶ Ex: "IMPLANTA«YO...", "PAVMTCAO E RECAP VIAS,AMPLIAC.."

Tratamento de dados - Empreendimentos

- **Técnicas avançadas: Códigos corretores de textos, vetorização de NLP e clusterização HDBSCAN, Técnicas fuzzys de proximidade de palavras.**
 - ▶ Mais que 5h de execução (GPU do Colab era > 1h) e após executadas não tinham um resultado satisfatório, pois traziam muitas categorias irrelevantes.
- **Técnica simples**
 - ▶ Substituição de palavras com dicionários, Regex e remoção de stopwords com pacote de NLP. Com execução < 1min, obteve-se um resultado mapeado e controlável:
 - » *METRO, PAVIMENTACAO, REFROTA, SINALIZACAO, CALCAMENTO, CICLOVIA, CORREDOR, DRENAGEM, VLT, etc.*

Sumário

- ① Análise exploratória de dados
- ② Tratamento de dados
- ③ Modelos**
- ④ Conclusões

- **Modelos de Regressão para Dados em Panel.**

- ▶ Especialmente útil por ser constituído de uma dimensão temporal e outra espacial.

- » *Dimensões utilizadas: Município e ano.*

- ▶ Controle de Variáveis Não Observáveis que variam apenas no tempo ou entre os municípios.

- » *Efeitos Fixos e Aleatórios.*

- ▶ Extensa literatura com testes estatísticos consolidados e amplamente utilizados pelas comunidades da saúde, econometria, estatística, etc.

- **Pooled: Modelo simples de dados empilhados.**

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \dots + \beta_n x_{nit} + \varepsilon_{it} \quad (1)$$

- **EF: Modelo de dados em painel com Efeitos Fixos α_i .**

$$y_{it} = \alpha_i + \beta_1 x_{1it} + \dots + \beta_n x_{nit} + \varepsilon_{it} \quad (2)$$

- **EA: Modelo de dados em painel com Efeitos Aleatórios γ_t .**

$$y_{it} = \gamma_t + \beta_1 x_{1it} + \dots + \beta_n x_{nit} + \varepsilon_{it} \quad (3)$$

» y_{it} : mortes ou feridos, x_{jit} : variáveis explicativas, ε : erro, β_j : coeficientes

- **Foram realizados 3 testes estatísticos para comparar os modelos:**
 - ▶ Teste de Chow: EF vs Pooled.
 - ▶ Teste de Breusch - Pagan: EA vs Pooled.
 - ▶ Teste de especificação de Hausman: EF vs EA.
- **Como resultado tem-se que o modelo de efeitos aleatórios (EA) é o mais adequado para esse trabalho, tendo sido o escolhido.**

Sumário

- ① Análise exploratória de dados
- ② Tratamento de dados
- ③ Modelos
- ④ Conclusões**

Principais conclusões

- Encontramos as categorias de empreendimentos que têm relação positiva ou negativa com a base de acidentes.
 - ▶ Por exemplo, o metrô tem uma relação negativa com o número de mortes.

Categoria de empreendimento	Coeficiente	P-valor
METRÔ	$-1.11 \cdot 10^{-8}$	0.0

Tabela 1: *Coeficiente e p-valor da categoria ‘metrô’, resultante do modelo para todas as cidades.*

- ▶ Vai ao encontro do caso de São Paulo.

Principais conclusões

- Itens que corroboram com o aumento de velocidade nas ruas possuem uma correlação positiva com o número de acidentes.

Categoria de empreendimento	Coeficiente	P-valor
PAVIMENTAÇÃO	$4.40 \cdot 10^{-7}$	0.0
TRANSPORTE RÁPIDO	$1.59 \cdot 10^{-7}$	0.0
CORREDORES/FAIXAS EXCLUSIVAS	$3.48 \cdot 10^{-8}$	0.003

Tabela 2: *Categorias com correlação positiva com o número de mortes, com os respectivos coeficientes e p-valores resultantes do modelo.*

Principais conclusões

- Investimentos em alternativas ao transporte rodoviário e aumento da qualidade do transporte possuem um coeficiente negativo em relação aos acidentes.

Categoria de empreendimento	Coeficiente	P-valor
CICLOVIA	$-4.82 \cdot 10^{-7}$	0.0
METRÔ	$-1.11 \cdot 10^{-8}$	0.0
REFROTA	$-3.34 \cdot 10^{-6}$	0.0
DRENAGEM	$-1.93 \cdot 10^{-7}$	0.0

Tabela 3: *Categorias com correlação negativa com o número de mortes, com os respectivos coeficientes e p-valores resultantes do modelo.*

Obrigado!