

Matéria: Estatística Aplicada

Professor: Esdras Adriano

Integrantes: Edvaldo Dos Santos, Gabriel Augusto Sá Silva

# 1. Introdução

Este trabalho visa analisar a quantidade de feridos gravemente em acidentes ao longo de semanas divididas em duas metades. A relevância deste estudo reside na possibilidade de identificar padrões ou diferenças significativas entre as duas metades das semanas, o que pode contribuir para a compreensão dos fatores que influenciam a gravidade dos acidentes e, consequentemente, para o desenvolvimento de estratégias de prevenção mais eficazes.

## 2. Objetivo

O objetivo principal deste trabalho é investigar se existe uma diferença estatisticamente significativa na quantidade de feridos gravemente entre a primeira e a segunda metade das semanas. Além disso, busca-se modelar a relação entre os dias da semana e a quantidade de feridos, utilizando regressão linear, para prever ou explicar a variabilidade na quantidade de feridos com base no dia da semana.

## 3. Materiais e Métodos

Para a realização deste estudo, foram utilizadas técnicas de estatística descritiva, teste t de Student e regressão linear. Os dados foram coletados e organizados em um conjunto de dados específico, e as análises foram realizadas utilizando o software Python, com bibliotecas especializadas como statistics para cálculos estatísticos, scipy para o teste t de Student e sklearn para a regressão linear. O ambiente de desenvolvimento utilizado foi o Google Colab, que permitiu a execução do código Python e a documentação do processo e dos resultados através de Markdown.

## 4. Resultados

- A média da quantidade de feridos gravemente durante a primeira metade das semanas foi de 2044.0, enquanto na segunda metade foi de 2576.67.
- O desvio padrão da quantidade de feridos gravemente durante a primeira metade das semanas foi de 112.54, e na segunda metade foi de 423.37.
- O teste t de Student resultou em uma estatística de teste de -1.7195823253985905 e um p-valor de 0.16062854557289966, indicando que não há evidências suficientes para rejeitar a hipótese nula de que não existe diferença significativa entre as duas metades das semanas.

- Foi realizada uma regressão linear para modelar a relação entre os dias da semana e a quantidade de feridos, substituindo os nomes dos dias por valores numéricos para facilitar a análise.

## 5. Conclusões

Os resultados indicam que não há diferença estatisticamente significativa na quantidade de feridos gravemente entre a primeira e a segunda metade das semanas. Além disso, a aplicação da regressão linear permite uma análise mais aprofundada da relação entre os dias da semana e a quantidade de feridos, contribuindo para um melhor entendimento dos fatores que podem influenciar a gravidade dos acidentes. Este trabalho destaca a importância de utilizar métodos estatísticos e de modelagem para analisar e interpretar dados de acidentes, visando a prevenção e a melhoria da segurança.

## 6. Relato

Como alunos, achamos este trabalho desafiador e enriquecedor. Ele permitiu a aplicação prática de conceitos estatísticos e de análise de dados, além de proporcionar uma oportunidade de contribuir para um tema de grande importância social, que é a segurança no trânsito. A utilização do Google Colab facilitou o processo de análise e documentação do trabalho.

## 7. Referência

- Python Software Foundation. Python Language Reference, versão 3.x. Disponível em: <https://www.python.org>
- Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.
- Documentação do Google Colab. Disponível em: <https://colab.research.google.com/notebooks/intro.ipynb>
- link do arquivo.ipynb:  
<<https://drive.google.com/file/d/1cHzKMrEt0eTCKYM-jNFOFlqZRUeyP4LU/view?usp=drivesdk>>, basta abrir esse arquivo pelo google colab.
- link para a base de dados:  
<<https://drive.google.com/file/d/1gT6TSLWtTzqvY09qHfkN9vPP6dO9Gkk/view?usp=sharing>>