**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS  
NÚCLEO DE EDUCAÇÃO A DISTÂNCIA**

**Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data**

**João Ricardo Côre Dutra**

**ANÁLISE DO FLUXO DE MOTOCICLETAS QUANTO ÀS CONDIÇÕES CLIMÁTICAS AO TIPO DE DIA E HORÁRIO NA AV. AFONSO PENA ESQUINA COM A RUA MARANHÃO NO MUNICÍPIO DE BELO HORIZONTE**

Belo Horizonte

2023

**João Ricardo Côre Dutra**

**ANÁLISE DO FLUXO DE MOTOCICLETAS QUANTO ÀS CONDIÇÕES CLIMÁTICAS AO TIPO DE DIA E HORÁRIO NA AV. AFONSO PENA ESQUINA COM A RUA MARANHÃO NO MUNICÍPIO DE BELO HORIZONTE**

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em Ciência de Dados e Big Data como requisito parcial à obtenção do título de especialista.

Belo Horizonte

2023**SUMÁRIO**

Contents

[1. Introdução 4](#_Toc119947443)

[1.1. Contextualização 4](#_Toc119947444)

[1.2. O problema proposto 5](#_Toc119947445)

[1.3. Objetivos 6](#_Toc119947446)

[3. Processamento/Tratamento de Dados 8](#_Toc119947447)

[4. Análise e Exploração dos Dados 9](#_Toc119947448)

[5. Criação de Modelos de Machine Learning 10](#_Toc119947449)

[6. Interpretação dos Resultados 11](#_Toc119947450)

[7. Apresentação dos Resultados 12](#_Toc119947451)

[8. Links 13](#_Toc119947452)

[REFERÊNCIAS 14](#_Toc119947453)

[REFERÊNCIAS 15](#_Toc119947454)

[APÊNDICE 16](#_Toc119947455)

# Introdução

## Contextualização

Viver em sociedade é uma característica de muitos animais, assim como os demais primatas organizam seus grupos a fim de potencializar a sua capacidade de sobrevivência e melhorar sua qualidade de vida, o ser humano desde muito tempo vive em grupos de diversos tipos e tamanhos com o mesmo objetivo.

O homem moderno já surgiu vivendo em sociedade, herança dos seus ancestrais. Viver em células organizadas onde cada indivíduo desempenha um papel especifico para contribuir de forma benéfica para os outros membros do grupo faz o ser humano potencializar a sua capacidade de dominar a natureza em benefício próprio.

As cidades é um dos exemplos de células sociais mais complexos em que o homem se organiza e vive. Dentro delas os cidadãos se organizam de forma a contribuir de diferentes maneiras em benefício próprio e para o coletivo, melhorando assim a qualidade de vida de todos. Entretanto, essa forma de se organizar pode gerar alguns malefícios para a vida humana.

Durante o século XX, pode-se observar um alto crescimento das cidades. Dois dos fatores que mais contribuíram para esse crescimento foi a queda de mortalidade proporcionada pelas inovações tecnológicas e também o êxodo rural principalmente nos países em desenvolvimento. Esse alto crescimento fez aparecer diversas metrópoles pelo mundo e no Brasil não foi diferente.

Durante a história do mundo pode-se observar diversas mazelas oriundas da formação e crescimento das cidades e formação de metrópoles. A pandemia da peste negra por exemplo se fundamenta na falta de higiene das cidades do século XIV. Alguns vírus que vieram acometer os humanos oriundos de grandes criações de aves e suínos também é outro exemplo já que grandes criações animais são necessárias para alimentar muitas pessoas que não necessariamente trabalham na produção de alimentos.

Apesar de grandes metrópoles terem um papel fundamental no desenvolvimento de um país, viver em cidades muito grandes sem um planejamento urbano adequado pode causar diversos problemas, como o aumento da poluição do ar, problemas quanto ao transito de pessoas e cargas por haver altos fluxos de veículos, fazendo gastar mais tempo que o necessário para se deslocar-se de um ponto a outro, também o aumento da violência por conta dos bolsões de pobrezas oriundos de uma má gestão social e muitos outros.

## O problema proposto

Muitas das vezes, o crescimento de uma cidade traz problemas para a sociedade urbana por falta de planejamento. Estudar e entender a intensidade de fluxo de veículos em uma cidade pode trazer diversos benefícios para o município que o faz, possibilitando uma melhor estruturação das vias públicas e do trânsito, gerando economia de tempo, de recursos e cuidado com o meio ambiente, portanto melhorando a vida dos cidadãos que ali residem.

Intuitivamente, pressupõe-se que o fluxo de alguns tipos de veículos podem variar em função de alguns fatores, como por exemplo o fluxo de caminhões pode ser maior em dias úteis quando comparados à finais de semana e feriados, assim como o fluxo de motocicletas também pode variar de acordo com o dia, horário e possivelmente condições climáticas. Prever e entender o comportamento do fluxo de veículos em um determinado ponto de uma cidade se mostra de extrema importância para melhor gestão de recursos, melhoria da qualidade de vida de uma população, tomadas de decisões quanto ao planejamento urbano, cuidado com o meio ambiente e muitos outros motivos.

Atualmente, a prefeitura da cidade de Belo Horizonte, disponibiliza através do portal BHTrans (<https://prefeitura.pbh.gov.br/bhtrans>) diversos dados relacionados ao transito na cidade, dentre esses dados existe a contagem volumétrica de radares (<https://dados.pbh.gov.br/dataset/contagens-volumetricas-de-radares>) onde estão presentes os registros de cada veículo que passou em frente aos radares de trânsitos espalhados pela cidade. Nessa massa de dados estão presentes informações de data, horário, tipo de veículo e localização.

O fluxo de motocicletas em uma cidade impacta diretamente na qualidade e na segurança do transito como um todo e também esse fluxo pode ser bastante volátil em função de muitos fatores externos, logo, optou-se por medir o fluxo de motocicletas na Av. Afonso Pena, equina com a Rua Maranhão nos meses de maio e junho de 2022 a fim de conseguir entender como alguns fatores externos o influenciam. A Av. Afonso pena foi escolhida por ser conhecidamente uma avenida bastante movimentada na cidade de Belo Horizonte, portanto, proporcionando uma quantidade significativa de dados a serem analisados. Os fatores escolhidos para serem estudados como influenciadores na intensidade de fluxo de motocicletas foram os tipos de dia (dia útil, final de semana e feriado), horários e condições climáticas (temperatura e precipitação).

# **Objetivos**

Este trabalho tem como objetivo propor e avaliar a influência fatores externos que podem afetar a intensidade de fluxo de motocicletas na Av. Afonso Pena equina com a rua Maranhão em Belo Horizonte.

O estudo consistiu em calcular a frequência de cruzamento de motocicletas em cada horário de cada dia no intervalo do dia 01 de maio de 2022 até o dia 30 de Junho de 2022 e depois classificar cada registro, se estava acima da média ou abaixo da média da frequência com relação a todos os registros.

Os fatores escolhidos como possíveis influenciadores na frequência de cruzamentos de motocicletas no ponto estudado foram:

1. Tipo de dia. Onde pode ser sábado, domingo, feriado ou dia útil. Esse fator foi escolhido pois nos dias úteis as atividades econômicas e sociais ocorrem com maior intensidade, podendo influenciar na frequência da quantidade de motocicletas que cruzam por hora o ponto de estudo.
2. Hora. A hora em que a dada frequência calculada é um possível fator de influência, pois existem horários ao longo das 24h de um dia em que as pessoas saem de casa para trabalhar ou para lazer. Também há os horários de funcionamento das atividades comerciais e econômicas, o que pode influenciar na maior quantidade de pessoas e transporte de cargas utilizando as vias públicas para se locomoverem.
3. Temperatura Ambiente: A temperatura pode ser um fator que influencia na decisão de uma pessoa utilizar ou não um motocicleta, onde em dias muito quentes ou muito frios, pode acontecer de optarem por outro tipo de transporte.
4. Precipitação: Assim como a temperatura, a precipitação também pode ser um fator que influencia na decisão da utilização de uma motocicleta para transporte ou outro tipo de veículo.

Os quatro fatores apresentados foram classificados em dois grupos: o primeiro é o de fatores climáticos, compreendendo dados de precipitação e temperatura ambiente, o segundo grupo com os dados de tipos de dias e horários, compreendendo o tipo de dia em que a medição da frequência está sendo feita e o respectivo horário.

Por fim, este trabalho procurou verificar o quanto e como esses grupos aumento ou diminuiu do fluxo de motocicletas no ponto estudado.

2. Coleta de Dados

Nessa seção você deve deixar claro onde obteve os dados, o formato e estrutura dos datasets, o relacionamento entre os datasets utilizados, etc. Caso os dados tenham sido obtidos na internet, informe a data e o link em que os dados foram obtidos. Sugere-se que você crie uma tabela com a descrição de cada campo/coluna de cada dataset conforme o exemplo a seguir:

|  |  |  |
| --- | --- | --- |
| Nome da coluna/campo | Descrição | Tipo |

# 3. Processamento/Tratamento de Dados

Nessa seção você deve deixar registrado todo o processamento e tratamento feitos sobre os dados obtidos. É importante que você informe a quantidade de registros obtidos, a quantidade de registros duplicados ou com informações ausentes, que tratamento você deu para cada problema encontrado em seus datasets, etc. Você deve descrever cada passo de forma minuciosa, de forma que outra pessoa consiga reproduzir o seu processamento/tratamento de forma precisa. Justifique as decisões tomadas no tratamento dos dados. Por exemplo: para os valores ausentes para o campo X eu decidi preenchê-los utilizando o cálculo da média aritmética pelo motivo ..., e então justifique sua decisão.

# 4. Análise e Exploração dos Dados

Nessa seção você deve mostrar como foi realizada a análise e exploração dos dados do seu trabalho. Mostre as hipóteses levantadas durante essa etapa e os padrões e *insights* identificados.

# 5. Criação de Modelos de Machine Learning

Conforme o documento de instruções para o TCC, essa etapa é obrigatória. Nessa seção você irá descrever as ferramentas e algoritmos utilizados. Se você utilizou o Knime, coloque aqui um print dos seus modelos e a descrição detalhada do workflow de cada modelo. Caso você tenha escrito scripts em Python ou R, por exemplo, coloque aqui apenas os trechos do código que você considera extremamente importantes para entendimento do seu trabalho. Explique as *features* utilizadas, faça a comparação entre diferentes algoritmos/modelos, justifique a escolha por determinado modelo, os parâmetros utilizados, etc. Por fim, salienta-se que embora você possa utilizar o KNIME para testar protótipos do seu modelo de ML, encorajamos você a fazer seus modelos em Python ou R.

# 6. Interpretação dos Resultados

Nessa seção você deve interpretar os resultados obtidos na análise e exploração de dados e também interpretar os resultados da aplicação dos algoritmos de Machine Learning, descobrindo insights importantes para responder o problema proposto.

# 7. Apresentação dos Resultados

Nessa seção você deve apresentar os resultados obtidos. Apresente gráficos, *dashboards*, conte a sua história de forma bastante criativa. Aqui você pode utilizar os modelos de Canvas propostos por Dourard (clique [aqui](https://www.louisdorard.com/machine-learning-canvas)) ou por Vasandani (clique [aqui](https://towardsdatascience.com/a-data-science-workflow-canvas-to-kickstart-your-projects-db62556be4d0)).



# 8. Links

Aqui você deve disponibilizar os links para o vídeo com sua apresentação de 5 minutos e para o repositório contendo os dados utilizados no projeto, scripts criados, etc.

Link para o vídeo: youtube.com/...

Link para o repositório: github.com/...

# 

# REFERÊNCIAS

BHTrans, **Empresa de Transportes e Trânsito de Belo Horizonte**, 2022. Disponível em: https://prefeitura.pbh.gov.br/bhtrans. Acesso em: 21 nov. 2022.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

# APÊNDICE

**Programação/Scripts**

Cole aqui seus scripts em Python e/ou R.

**Gráficos**

Cole aqui workflows (KNIME), gráficos e figuras que você tenha gerado e não colocou no texto principal.

**Tabelas**

Cole aqui tabelas de dados que você tenha gerado e não colocou no texto principal.