

Análise de Dados das Ocorrências no Trânsito da Cidade do Recife para Prevenção de Acidentes

Traffic Occurrences in the City of Recife: Use of Technology to Prevent Accidents

Anicely M. C. Santos¹

 orcid.org/0009-0003-6719-4495

João Vitor da Silva Gomes¹

 orcid.org/0000-0001-5727-2427

Paulo Cataldi¹

 orcid.org/0009-0006-6243-6143

Rodrigo Bessone¹

 orcid.org/0009-0004-3276-4002

Ygor Alcantara¹

 orcid.org/0009-0001-1001-2545

¹Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil. E-mail: amcs3@ecomp.poli.br

DOI: [10.25286/rep.v9i1.2766](https://doi.org/10.25286/rep.v9i1.2766)

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: Anicely M. C. Santos; João Vitor da Silva Gomes; Paulo Cataldi; Rodrigo Bessone; Ygor Alcantara. Análise de Dados das Ocorrências no Trânsito da Cidade do Recife para Prevenção de Acidentes. Revista de Engenharia e Pesquisa Aplicada, v.9, n. 1, p. 1-10, 2024. DOI: [10.25286/rep.v9i1.2766](https://doi.org/10.25286/rep.v9i1.2766)

RESUMO

Mortes por acidentes de trânsito ainda é a segunda maior causa de mortes no mundo, podendo ser derivado por três fatores: humano, veicular e vias. De acordo com o Departamento Nacional de Infraestrutura (DNIT), os tipos mais frequentes são: colisão no mesmo sentido, saída da pista, choque com objeto fixo, capotagem e colisão frontal. Já em termos de gravidade, temos: colisão frontal, atropelamento, saída da pista.

Este artigo analisou os acidentes de trânsito que aconteceram na cidade do Recife e região metropolitana com veículos de duas rodas, nos anos 2016 a 2021, utilizando dois modelos lineares: Autoregressive Moving Average (ARMA) e Autoregressive Integrated Moving Average (ARIMA), para encontrar padrões de tendências mensais. Os modelos não conseguiram superar a persistência.

PALAVRAS-CHAVE: Mineração de Dados; Trânsito; Séries Temporais; Modelos Estatísticos.

ABSTRACT

Deaths from traffic accidents are still the second biggest cause of death in the world, and can be caused by three factors: human, vehicle and road. According to the National Infrastructure Department (DNIT), the most frequent types are: collision in the same direction, leaving the roadway, collision with a fixed object, overturning and frontal collision. In terms of severity, we have: head-on collision, hit-and-run, lane departure. This article analyzed traffic accidents in the city of Recife and the metropolitan region involving two-wheeled vehicles, from 2016 to 2021, using two linear models: Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA), to find monthly trend patterns. The models failed to overcome persistence.

KEY-WORDS: Data Mining; Traffic; Time Series; Statistical Models.

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

Muito se tem discutido, recentemente, acerca dos avanços no campo da tecnologia, para a prevenção de acidentes de trânsito em todo país. Nesse sentido, destacamos como principais aspectos desse fenômeno [1], a redução dos custos, segurança, flexibilidade, agilidade, escalabilidade e o acesso de tecnologia de ponta em todo o processo; isso traz um impacto considerável na produtividade das instituições e em geral facilita o acesso às informações gerenciais, dando mais transparência a todas as etapas do processo.

Segundo a OMS [2] (Organização Mundial da Saúde) mais de 3,5 mil pessoas morrem por dia no trânsito em todo o mundo, isso equivale a quase 1,3 milhão de mortes que poderiam ter sido evitadas a cada ano. Além disso, entre 20 milhões e 50 milhões de pessoas ficam lesionadas por conta desses acidentes.

Constata-se a abaixo alguns fatores de riscos que causam os acidentes segundo Safe System:

1. Excesso de velocidade;
2. Dirigir sob a influência de álcool ou drogas;
3. Não utilização de capacetes para motociclistas, cinto de segurança e cadeirinha para crianças;
4. Direção distraída;

1.2 DESCRIÇÃO DO PROBLEMA

Atualmente, a Autarquia de Trânsito e Transporte Urbano do Recife - CTTU possui, no portal de dados abertos, as informações relativas às ocorrências de trânsito que ocorreram nos anos de 2016 a 2021. O intuito do projeto é combinar e tratar essas bases de dados para realizarmos sua aplicação para a diminuição das ocorrências, contribuindo assim com a queda nos números de vítimas fatais dos acidentes.

1.3 OBJETIVO

O objetivo do presente projeto é utilizar os *Datasets* do portal de dados abertos do Recife [3] para criar um modelo de predição do volume mensal de acidentes envolvendo veículos de duas rodas a fim de auxiliar na prevenção dessas ocorrências de trânsito, reduzindo o número de vítimas e diminuindo os gastos

hospitalares. Iremos analisar os dados históricos de acidentes e debater indicadores relevantes; analisar padrões nas ocorrências; e comparar modelos de previsão de séries temporais ARMA e ARIMA no contexto dos acidentes de trânsito envolvendo veículos de duas rodas.

1.4 JUSTIFICATIVA

Segundo estudo realizado pelo site UOL, a divisão administrativa de Recife é organizada em 94 bairros, agrupada em seis regiões administrativas (Centro, Norte, Nordeste, Oeste, Sudeste e Sul). Esse número considera os bairros oficiais da cidade. Porém, o município apresenta inúmeros assentamentos irregulares, principalmente nas áreas periféricas e em zonas de ocupação desaconselhada, como nas margens de rios e mangues. A ocupação irregular dificulta a mobilidade urbana e a oferta de serviços públicos para a população.

Outra pesquisa importante [4], realizada pelo Conselho Federal de Medicina, Pernambuco é o segundo estado do Brasil que mais registrou internações nos hospitais em consequência de acidentes de trânsito na última década, com um crescimento de 725% no número de vítimas. O estudo apontou que houve um média de 5,6 mil internados por acidentes de trânsito. Dados do banco de dados do estado de Pernambuco (2022) mostra que a cidade do Recife possui uma frota de 724.051 veículos, conforme mostra a Tabela 1.

Tabela 1 - Frota de veículos da cidade do Recife.

Tipo de veículos	Quantidade
Automóvel	406.065
Caminhão	20.408
Caminhonete	38.242
Camioneta	40.445
Ônibus	4.489
Micro-ônibus	2.283
Motocicletas	155.211
Outros*	56908

Fonte: Base de dados do Estado de Pernambuco – BDE.

A principal motivação para sustentar o presente projeto, reside na importância que o tema possui para a sociedade atual. Sendo considerado um

tema relevante. Podemos destacar a quantidade de ocorrências registradas pela prefeitura conforme mostra a Tabela 2.

Tabela 2 - Ocorrências de registradas na cidade do Recife (2016 a 2021).

Anos	Ocorrências	Número de Vítimas
2016	11256	2589
2017	11736	2658
2018	11394	2419
2019	12044	2612
2020	4092	2109
2021	2380	2148

Fonte: Portal de Dados abertos do Recife.

Observação: Nos anos de 2020 e 2021 tiveram uma queda considerável no número de ocorrências devido à crise sanitária sem precedentes “pandemia”, onde a interação social teve que ser diminuída, apesar disso tivemos um número de vítimas muito elevado. Trazer a tecnologia para o processo de prevenção de ocorrências de trânsito é essencial para a preservação da vida.

1.5 ESCOPO NEGATIVO

Não está proposto para este artigo predição de acidentes envolvendo outros tipos de veículos que não os de duas rodas. Não faremos *deploy* das soluções observadas, nem serão propostos novos modelos teóricos, matemáticos ou estatísticos, para tratamento ou modelagem dos dados. Aplicaremos alguns modelos já bem consolidados na literatura.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 ÁREA DE NEGÓCIO

O trânsito brasileiro é o quarto mais violento do continente americano, segundo dados divulgados pela Organização Mundial da Saúde (OMS). O país possui uma taxa de mortalidade de 23,4 mortes no trânsito para cada cem mil habitantes.

Segundo estudo realizado pelo Instituto de Pesquisa Econômica Aplicada - IPEA, em junho de 2020, os acidentes de trânsito no Brasil matam cerca de 45 mil pessoas por ano e deixam mais de 300 mil pessoas com lesões graves. Em uma estimativa conservadora, observou-se que os

acidentes em rodovias custam à sociedade brasileira cerca de R\$40 bilhões por ano, enquanto os acidentes nas áreas urbanas, em torno de R\$10 bilhões; sendo boa parte desse valor custos hospitalares [5].

A aplicação de tecnologia já é uma realidade na mitigação do trânsito em diversos centros urbanos, temos como exemplos os aplicativos de mobilidade, os dispositivos de segurança (semáforos com programação de tempo de funcionamento, câmeras inteligentes para avaliação de imagens) e etc.

2.2 TRABALHOS RELACIONADOS

A seguir serão descritos sete trabalhos que possuem semelhanças temáticas com o projeto proposto.

Em *Nonparametric Regression for the Short-term Traffic Flow Forecasting* (Regressão não paramétrica para a previsão de fluxo de tráfego de curto prazo), os autores usaram a base de tráfego, do ano de 2006, cedida pela *Universidade Duluth*, nos Estados Unidos, como teste foi utilizado o mês de janeiro de 2007. O intuito da pesquisa foi prever o fluxo de tráfego de curto prazo, no caso o mês seguinte. Com o uso de *knn* baseado em regressão não paramétrica, conseguem um resultado melhor que com o uso de redes neurais.

No artigo *A Time Series Model for Assessing the Trend and Forecasting the Road Traffic Accident Mortality* [6], os autores usam como base de treinamento 7 anos de trânsito da cidade de Zanzan, no Irã, entre 2007 e 2013, a fim de preverem os acidentes dos próximos 4 anos. Para isso usam séries temporais, e o melhor resultado é alcançado por meio do uso do modelo SARIMA.

Em *The prediction of traffic flow with regression analysis* [7], os autores usam dados de tráfego da cidade do Porto, Portugal, entre os anos de 2013 a 2015, sempre usando as três últimas semanas como treino, a fim de prever os dados da próxima. Foram usados cinco modelos: *Linear Regression*, *Sequential Minimal Optimization (SMO) Regression*, *Multilayer Perceptron*, *M5P model tree* e *Random Forest*.

No trabalho *Early Warning of Traffic Accident in Shanghai Based on Large Data set Mining* [8], são usados dados de trânsito da cidade de Xangai, China, entre julho de 2014 e abril de 2015, como entradas de algoritmos de classificação e análise de regressão.

Por fim, em *Mineração de Dados na Construção de Modelo de Predição de Acidentes com Vítimas*

em Recife [9], os autores têm como objetivo criação de estratégias que ajudem a mitigar os crescentes números de acidentes de trânsito com vítimas, e para isso utilizaram uma base de do portal de dados abertos da prefeitura do Recife do período de entre junho de 2015 até fevereiro de 2020.

3 MATERIAIS E MÉTODOS

3.1 DESCRIÇÃO DA BASE DOS DADOS

Utilizamos a metodologia exploratória do banco de dados da Autarquia de Trânsito e Transportes Urbanos do Recife - CTTU; trouxemos os dados quantitativos gerais e o filtro dos 10 bairros com maiores ocorrências, dados ênfase a alguns números que destoam dos resultados. A Tabela 3 mostra o dicionário de dados analisado neste trabalho.

Tabela 3 - Dicionário de dados Dataset

Coluna	Tipo	Descrição
dia	<i>timestamp</i>	descrição da data da ocorrência
mês_desc	<i>text</i>	descrição do mês da ocorrência
ano	<i>timestamp</i>	descrição do ano da ocorrência
hora	<i>text</i>	descrição da hora/minuto da ocorrência
natureza_acidente	<i>text</i>	descrição se o acidente possuiu ou não vítima
vítimas fatais	<i>numeric</i>	descrição o quantitativo de vítimas fatais
situação	<i>text</i>	descrição do status da ocorrência (andamento, finalizada, cancelada)
bairro	<i>text</i>	descrição do bairro da ocorrência
endereco	<i>text</i>	descrição do endereço da ocorrência
numero	<i>numeric</i>	descrição número do endereço da ocorrência
detalhe_endereco_acidente	<i>text</i>	descrição detalhada do endereço da ocorrência

complemento	<i>text</i>	descrição de um ponto de referência próximo da ocorrência
bairro_cruzamento	<i>text</i>	?
tipo	<i>text</i>	descrição do dito de acidente da ocorrência ex.: colisão, atropelamento, etc.
auto	<i>numeric</i>	quantidade de automóveis envolvidos na ocorrência
moto	<i>numeric</i>	quantidade de motos envolvidos na ocorrência
ciclom	<i>numeric</i>	quantidade de quadriciclos envolvidos na ocorrência
ciclista	<i>numeric</i>	quantidade de ciclista envolvidos na ocorrência
pedestre	<i>numeric</i>	quantidade de pedestre envolvidos na ocorrência
onibus	<i>numeric</i>	quantidade de ônibus envolvidos na ocorrência
caminhao	<i>numeric</i>	quantidade de caminhões envolvidos na ocorrência
viatura	<i>numeric</i>	quantidade de viaturas envolvidos na ocorrência
outros	<i>numeric</i>	Outros tipos de meios de transportes
vítimas	<i>numeric</i>	descrição o quantitativo de vítimas

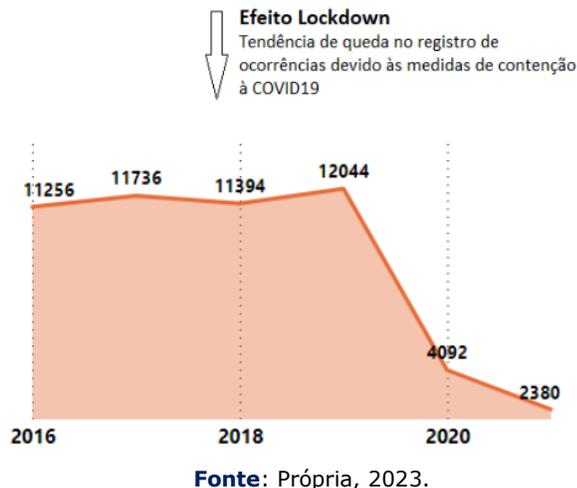
Fonte: Portal de Dados abertos do Recife.

3.2 ANÁLISE DESCRITIVA DOS DADOS

Realizamos o tratamento do dataset fazendo o processo de concatenação e tratamento das bases de dados dos anos de 2016 a 2021. Em relação aos dados faltantes, os registros com atributos nulos foram removidos. Os atributos faltantes foram preenchidos com os valores que mais ocorriam no dataset. Na seleção de atributos, novos atributos foram gerados a partir do conjunto de atributos fornecidos para ajudar no processo de mineração.

Após o tratamento verificamos que o número total de ocorrência registrado pela Autarquia de trânsito e transporte do Recife - CTTU foi de 52902. As figuras 1 e 2 detalhando a distribuição por ano, mês e dia da semana. Todas as informações foram retiradas do portal de dados aberto da cidade do Recife.

Figura 1 - Número de Ocorrências de Trânsito registradas no Recife de 2016 a 2021.

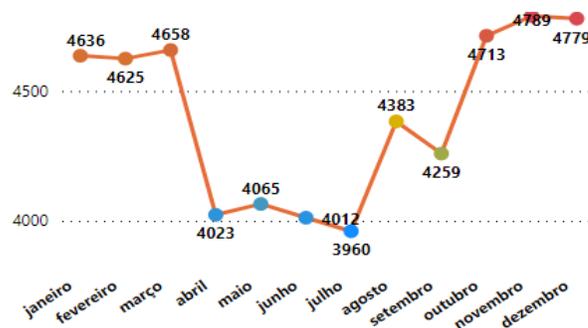


Observando a Figura 1 verificamos que no período de 2016 a 2019, o número de ocorrências foi bem semelhante em quantidade, já nos anos de 2020 e 2021, houve uma queda abrupta na quantidade de chamados, devido a proibição de locomoção, restrições impostas pela pandemia Covid-19.

Não ficou claro, durante o levantamento dos dados, se a queda na quantidade de chamados aconteceu devido a subnotificação, ou demais outros fatores não relacionados com a pandemia de Covid-19.

Outra informação importante é o número de vítimas fatais do período estudado (2016 a 2021) foram 175, na Figura 3 destacamos o número de vítimas fatais por meio de transporte.

Figura 2 - Número de ocorrências de trânsito agrupadas por mês entre os anos de 2016 e 2021.



Fonte: Própria, 2023.

Na Figura 02 vê-se que há uma maior concentração de registros de acidentes durante o período de outubro a março (verão), em comparação ao período de abril a setembro (inverno). A diferença entre os períodos é de em média cerca de 600 registros de acidentes.

Na Tabela 04, percebe-se uma distribuição uniforme dos acidentes durante os dias úteis (aproximadamente 16% cada) e uma diminuição nos finais de semana (8,32% no sábado e 11,9% no domingo).

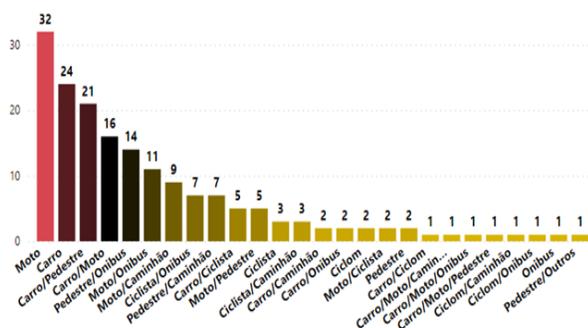
Tabela 4 - Número de Ocorrências por dia.

Dias da semana	Ocorrências	%
domingo	4430	8.32%
segunda-feira	8390	15.86%
terça-feira	8418	15.91%
quarta-feira	8371	15.82%
quinta-feira	8440	15.95%
sexta-feira	8583	16.22%
sábado	6297	11.90%
Total	52902	100%

Fonte: Própria, 2023.

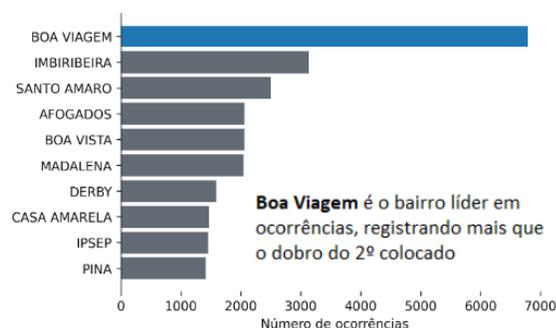
Destacamos os bairros onde ocorre o maior número de chamados de acidentes junto com o número de vítimas fatais, para isso realizamos um filtro para obter os 10 bairros com maior incidência de ocorrências. A Figura 4 mostra os detalhes.

Figura 3 – Número de Vítimas Fatais por Meio de Transporte.



Fonte: Própria, 2023.

Figura 4 - Top 10 Bairros do Recife com Maior Número de Ocorrências.



Fonte: Própria, 2023.

Além disso, é oportuno mencionar que o número de vítimas fatais envolvendo motos (Tabela 05) é bem significativo, fizemos também uma comparação com as vítimas fatais envolvendo carros (Tabela 06).

Tabela 5 - Número de vítimas fatais envolvendo Motos.

Moto	Vítimas fatais	%
Moto	32	41.03%
Carro/moto	16	20.51%
Moto/Ônibus	11	14.10%
Moto/Caminhão	9	11.54%
Moto/Pedestre	5	6.41%
Moto/Ciclista	2	2.56%
Carro/Moto/Caminhão	1	1.28%
Carro/Moto/Ônibus	1	1.28%
Carro/Moto/Pedestre	1	1.28%
Total	78	100%

Fonte: Própria, 2023.

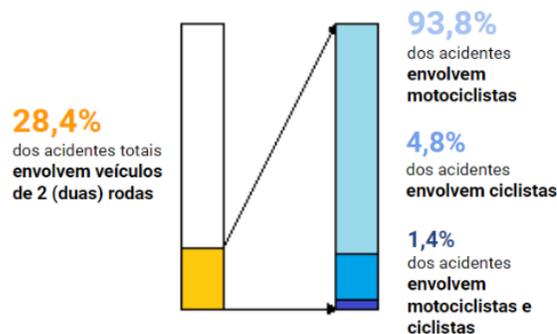
Tabela 6 - Número de vítimas fatais envolvendo carros.

Carro	Vítimas fatais	%
Carro	24	32.43%
Carro/Caminhão	2	2.70%
Carro/Ciclista	5	6.76%
Carro/Ciclomotor	1	1.35%
Carro/Moto	16	21.62%
Carro/Moto/Caminhão	1	1.35%
Carro/Moto/Ônibus	1	1.35%
Carro/Moto/Pedestre	1	1.35%
Carro/Ônibus	2	2.70%
Carro/Pedestre	21	28.38%
Total	74	100%

Fonte: Própria, 2023.

Segundo a Abraciclo (Associação Brasileira dos Fabricantes de Motocicletas, Ciclomotores, Motonetas, Bicicletas e Similares), no primeiro semestre de 2022, o número de motocicletas licenciadas no país (636.565) aumentou 32% em relação ao mesmo período de 2021. Baseado na informação da Abraciclo, achamos relevante realizar uma análise dos acidentes que envolveram veículos de duas rodas nas ruas da capital pernambucana. A Figura 5 mostra os detalhes.

Figura 5 - Relação dos Acidentes de Trânsito Envolvendo Veículos de Duas Rodas no Recife.

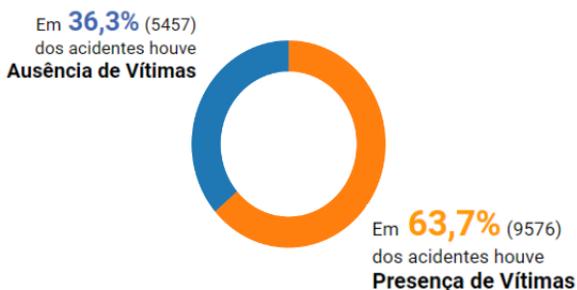


Fonte: Própria, 2023.

Das 52902 ocorrências de acidentes registradas pela CTTU entre 2016 e 2021, 28,4% envolveram veículos de duas rodas. Dos quais mais de 90% são motocicletas. Ainda conforme o gráfico 05, vemos que os ciclistas têm uma participação de menos de 5% nesses acidentes.

Embora representem menos de um terço dos registros totais, os condutores de veículos de duas rodas são os mais propensos a ter sua saúde debilitada nesses cenários. A ausência da proteção de uma carroceria, os expõe às possíveis sequelas decorrentes dos acidentes. Este fato reflete diretamente no que revela a Figura 6.

Figura 6 - Presença de Vítimas nos Acidentes de Trânsito Envolvendo Veículos de Duas Rodas.

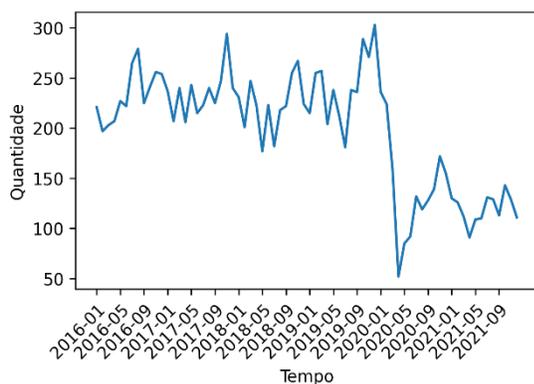


Fonte: Própria, 2023.

3.2 MÉTODO

Os dados originalmente estruturados conforme o dicionário de dados da seção 3.2, foram reduzidos para considerar somente os atributos de interesse: mes_dec, ano, moto e ciclista. Com base nestes atributos foi extraída uma série temporal através da contagem de registros de acidentes envolvendo veículos de duas rodas mês a mês ao longo dos anos de 2016 a 2021. A Figura 6 mostra a distribuição da série temporal.

Figura 6 - Série mensal da quantidade de acidentes envolvendo veículos de duas rodas.



Fonte: Própria, 2023.

Para analisar a série temporal, exibir os gráficos de autocorrelação e construir os modelos ARMA (*Autoregressive Moving Average*) e ARIMA (*Autoregressive Integrated Moving Average*) foi utilizada a biblioteca *statsmodels* disponível na linguagem Python. Já para dividir a série em trechos

de treinamento e teste, bem como a avaliação por meio da métrica Mean Squared Error (MSE), usamos a biblioteca *scikit-learn*.

O período que consta de jan/2016 a dez/2020 foi separado para a etapa de treinamento dos modelos, enquanto que o período que compreende os meses de jan/2021 a dez/2021 serviu como dados de teste.

A fim de estabelecer uma base de comparação durante a análise dos resultados, incluímos a persistência como modelo a ser superado. A persistência funciona conforme a seguinte premissa: "A previsão no instante futuro de interesse será igual ao que se observa no presente". Este é o modelo linear mais básico.

Os modelos foram desenvolvidos para que através da informação observável no instante atual t , e nos instantes passados $t-1$, $t-2$, ..., $t-k$ fosse possível prever a quantidade de acidentes envolvendo veículos de duas rodas no instante futuro $t+1$, considerando uma escala mensal. Ou seja, $t-1$, se refere a um mês no passado, e $t+1$, a um mês a frente.

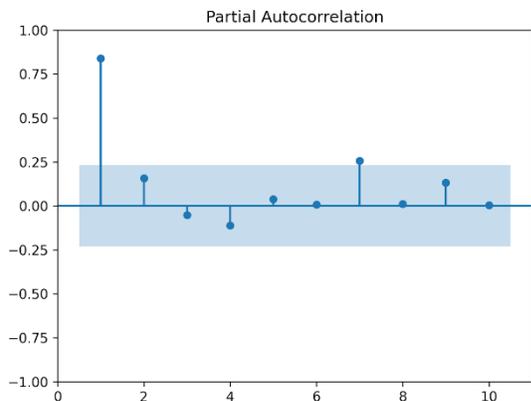
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

4.1 RESULTADOS

Iniciando pelo Figura 7 que exibe a autocorrelação parcial da série, verificamos que somente a informação do instante atual t é estatisticamente relevante para a previsão do instante $t+1$. Esse é o único lag que sobrepassa significativamente a região de confiança (região em azul). Esse foi um forte indicativo de que a componente auto-regressiva dos modelos ARMA e ARIMA deveria ter ordem 1.

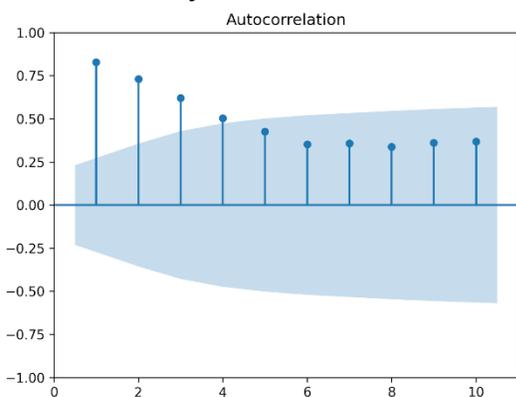
Por outro lado, olhando para o Figura 8 que contém a autocorrelação da série, vimos um pico estatisticamente relevante até o lag 4. Sendo esse um forte indicativo de que a componente de média móvel dos modelos ARMA e ARIMA ter ordem 4.

Figura 7 - Autocorrelação parcial da série.



Fonte: Própria, 2023.

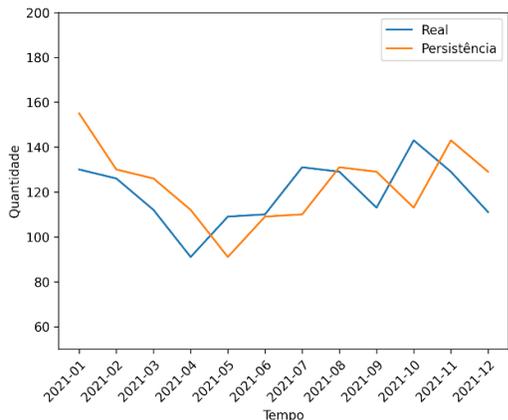
Figura 8 - Autocorrelação da série.



Fonte: Própria, 2023.

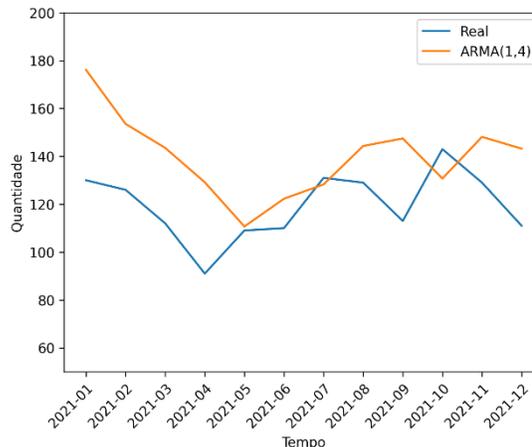
Os resultados das previsões dos modelos da persistência, ARMA(1,4) e ARIMA(1,1,4) são apresentados a seguir nos figuras 9, 10 e 11, respectivamente. A linha em azul representa a observação real da quantidade de acidentes envolvendo veículos de duas rodas, enquanto a linha em laranja representa as previsões.

Figura 9 - Previsão do modelo persistência.



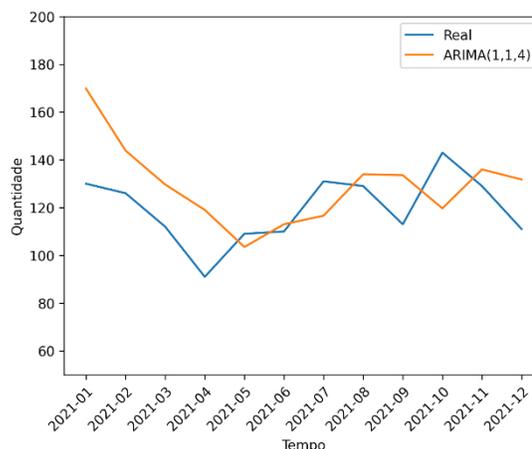
Fonte: Própria, 2023.

Figura 10 - Previsão do modelo ARMA(1,4).



Fonte: Própria, 2023.

Figura 11 - Previsão do modelo ARIMA(1,1,4).



Fonte: Própria, 2023.

As quantificações do erro entre as previsões dos modelos e as observações reais da série de teste estão dispostas na Tabela 07.

Tabela 7 - Erro quadrático médio das previsões obtidas para cada modelo.

Modelo	MSE
Persistência	310.3
ARMA(1,4)	706.17
ARIMA(1,1,4)	393.13

Fonte: Própria, 2023.

4.2 DISCUSSÃO

Os resultados dispostos revelaram que os modelos ARMA(1,4) e ARIMA(1,1,4) foram

incapazes de superar a persistência, o modelo base.

Através dos figuras 9, 10, 11 é possível notar que as previsões, no geral, se comportam como um mero deslocamento da série real no tempo. Até certo ponto, esta característica era esperada, uma vez que os modelos utilizados são lineares auto-regressivos de ordem 1.

Entretanto, o que chama atenção é a disparidade da medição do erro quadrático médio entre os modelos. O ARMA(1,4) apresentou mais que o dobro do erro da persistência. Por outro lado, a inclusão da diferenciação de primeira ordem no ARIMA(1,1,4) melhorou o desempenho em relação ao ARMA(1,4) em cerca de 44%.

Vale salientar que a métrica MSE penaliza os modelos considerando o quadrado dos erros e, portanto, mesmo que parte das previsões se equiparem a série real, a medição do erro médio será enviesada caso alguns poucos pontos divergirem muito da série real. É o que acontece com alguns pontos do ARMA(1,4) especialmente nos primeiros meses do ano, justificando o MSE elevado.

5 CONCLUSÃO E TRABALHOS FUTUROS

Este estudo se dispôs a investigar a possibilidade da previsão da quantidade de acidentes envolvendo veículos de duas rodas na cidade do Recife. Foram empregados dois modelos lineares, ARMA e ARIMA, a fim de que fossem encontrados padrões de tendência mensal nas ocorrências. Dados disponibilizados pela CTTU sujeitos a modelagem resultaram em um desempenho de previsão que não superou a persistência.

Entretanto, ficou evidente que o pré-processamento da série causou impacto direto na capacidade de aprendizagem dos modelos, especialmente no que tange a diferenciação da série temporal.

Este estudo pode ser ampliado no sentido de que se utilize modelos lineares com análise sazonal como o SARIMA (Seasonal ARIMA), ou de que se aplique outras técnicas de pré-processamento como a suavização exponencial.

5 REFERÊNCIAS

[1] 4 principais causas de mortes no trânsito. **Zul Digital**, São Paulo, 25 de maio de 2022. Disponível em: <https://www.zuldigital.com.br/blog/principais-causas-mortes-transito/>. Acesso em: 20 de setembro de 2022.

[2] OMS: Brasil é o quarto país com mais mortes no trânsito na América. **Gazeta do Povo**, Curitiba, 19 de maio de 2016. Disponível em: <https://www.gazetadopovo.com.br/vida-e-cidadania/oms-brasil-e-o-quarto-pais-com-mais-mortes-no-transito-na-america-7d9z0vpb4r3gotpl8u436lxms/>. Acesso em: 25 de outubro de 2022.

[3] Portal de Dados abertos da prefeitura do Recife. **Acidentes de Trânsito com e sem vítimas 2015 a 2022**. Disponível em: <http://dados.recife.pe.gov.br/>. Acesso em: 20 de setembro de 2022.

[4] SOUZA, Alice de. Internações por acidentes de trânsito crescem 725% em uma década em Pernambuco. **Diário de Pernambuco**, Pernambuco, 09 de junho de 2019. Disponível em: <https://www.diariodepernambuco.com.br/noticia/vida-urbana/2019/06/internacoes-por-acidentes-de-transito-crescem-725-em-uma-decada-em-pe.html>. Acesso em: 23 de outubro de 2022.

[5] CARVALHO, Carlos Henrique Ribeiro de. Custos dos acidentes de trânsito no Brasil: estimativa simplificada com base na atualização das pesquisas do IPEA sobre custos de acidentes nos aglomerados urbanos e rodovias. **Texto para discussão – Instituto de Pesquisa Econômica Aplicada (IPEA)**. Brasília: Rio de Janeiro, pág 7 – 11, 2020. Disponível em https://repositorio.ipea.gov.br/bitstream/11058/10075/1/td_2565.pdf.

[6] ZHANG, Tao et al. Nonparametric regression for their short-term traffic flow forecasting. In: **2010 International conference on mechanic automation and control engineering**. IEEE, 2010. p. 2850-2853.

[7] YOUSEFZADEH-CHABOK, Shahrokh et al. A Time Series Model for Assessing the Trend and Forecasting the Road Traffic Accident Mortality. **Archives of trauma research**. v.5, n.3, 2016.

[8] ALAM, Ishteaque; FARID, Dewan Md; ROSETTI, Rosaldo JF. The Prediction of traffic flow with regression analysis. In: **Emergin Technologies in Data Mining and Information Security**. Springer, Singapore, 2019. p. 661-671.

[9] YANBIN, Yang et al. Early Warning of Traffic Accident in Shanghai Based on Large Data Set

Mining. In: **2016 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)**. IEEE, 2016. p. 18-21.

[10] COSTA, Adriano de Melo; FREITAS, Arthur Guilherme Oliveira de; PINHEIRO, Ricardo Paranhos. Mineração de Dados na Construção de Modelo de Predição de Acidentes com Vítimas em Recife. **Revista de Engenharia e Pesquisa Aplicada (REPA)**. v.6, n.3, p. 70-80, 2021. DOI: 10.25286/rep.v6i3.1707. Acesso em: 20 de setembro de 2022.