



MONITORAMENTO DA QUALIDADE DO AR NO MUNICÍPIO DO RIO DE JANEIRO : UMA ABORDAGEM ANALÍTICA COM SÉRIES TEMPORAIS

Bruno dos Santos Baltuilhe¹, Bruno Luigi dos Santos Tobias¹, Eduardo Nogueira Mota¹, Thábata Cristina Seixas Balbino¹, Vagner Milani¹

¹Faculdade de Computação e Informática (FCI)
Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

{10424822@mackenzista.com.br, 10423789@mackenzista.com.br,
10414834@mackenzista.com.br, 10424240@mackenzista.com.br,
10415589@mackenzista.com.br}

Resumo. Este trabalho apresenta uma abordagem analítica para o Monitoramento da Qualidade do Ar no Rio de Janeiro, com foco no poluente Partículas Finas (PM 2.5), e propõe uma metodologia preditiva rigorosa. A pesquisa empregou um pipeline de Ciência de Dados para a coleta e pré-processamento de 54.901 registros horários da API ArcGIS, cobrindo o período de 2011 a 2018. A Análise Exploratória de Dados (EDA) revelou um achado crítico para a saúde pública: 100% dos dias analisados excedem o limite de segurança de 15mg/m³ estabelecido pela Organização Mundial da Saúde (OMS). Para benchmarking, foi implementado um Modelo Base de Regressão Linear, que demonstrou ser ineficaz, obtendo um R² de apenas 0,028, confirmando que a natureza temporal e não-linear da poluição atmosférica não é capturada por modelos estatísticos simples. O estudo conclui que a modelagem preditiva exige técnicas avançadas e, como proposta de trabalho futuro, sugere a aplicação da metodologia Box-Jenkins (ARIMA e SARIMA), além de modelos de Deep Learning (LSTM), visando alcançar a acurácia necessária para o forecasting da qualidade do ar.

Palavras-chave: Séries Temporais; PM 2.5; Qualidade do Ar; Rio de Janeiro; Regressão Linear; ARIMA.

Abstract. This work presents an analytical approach for Air Quality Monitoring in Rio de Janeiro, focusing on Fine Particulate Matter PM 2.5, and proposes a rigorous predictive methodology. The research employed a Data Science pipeline for the collection and preprocessing of 54,901 hourly records from the ArcGIS API, covering the period from 2011 to 2018. The Exploratory Data Analysis (EDA) revealed a critical finding for public health: 100% of the days analyzed exceed the safety limit of 15mg/m³ established by the World Health Organization (WHO). For benchmarking, a Linear Regression Base Model was implemented, proving to be ineffective, yielding an R² of only 0.028, confirming that the temporal and non-linear nature of air pollution is not captured by simple statistical models. The study concludes that predictive modeling requires advanced techniques and, as future work, proposes the application of the Box-Jenkins methodology (ARIMA and SARIMA), in addition to Deep Learning models (LSTM), aiming to achieve the necessary accuracy for air quality forecasting.

Keywords: *Time Series; PM 2.5; Air Quality; Rio de Janeiro; Linear Regression; ARIMA.*

1. Introdução

A poluição atmosférica consolida-se como um dos maiores desafios ambientais e de saúde pública da atualidade, com impactos que transcendem fronteiras geográficas e escalões socioeconômicos. A Organização Mundial da Saúde (OMS) classifica o ar contaminado como o maior risco ambiental individual para a saúde humana, com o material particulado fino ($PM_{2,5}$) sendo apontado como o poluente mais crítico devido à sua capacidade de penetrar profundamente no sistema respiratório e cardiovascular (WHO, 2021).

A relação causal entre a exposição à poluição do ar e efeitos deletérios à saúde é incontestável e quantificada pela epidemiologia. As partículas inaláveis, em especial o $PM_{2,5}$, são capazes de alcançar os alvéolos pulmonares, entrar na corrente sanguínea e desencadear uma resposta inflamatória sistêmica. Isso resulta no agravamento de condições preexistentes, como asma e Doença Pulmonar Obstrutiva Crônica (DPOC) e no aumento do risco de incidência de doenças cardiovasculares, incluindo infarto agudo do miocárdio e acidente vascular cerebral (AVC) (BROOK et al., 2020). As consequências são ainda mais severas para grupos vulneráveis. Para crianças, a exposição crônica está associada a déficits no desenvolvimento pulmonar, maior frequência de infecções respiratórias e até a prejuízos cognitivos. Para idosos e portadores de comorbidades, a poluição funciona como um fator de estresse adicional, elevando as taxas de hospitalização e mortalidade. Um estudo realizado na Região Metropolitana de São Paulo estimou que a adequação dos níveis de $PM_{2,5}$ aos padrões da OMS poderia evitar milhares de mortes prematuras anualmente, gerando uma economia de bilhões de reais em gastos com saúde pública (ARAÚJO et al., 2020). Esses dados evidenciam que investir no controle da poluição atmosférica não é apenas uma questão ambiental, mas uma medida de saúde pública com retorno econômico claro.

No Brasil, grandes centros urbanos, como o Rio de Janeiro, enfrentam problemas recorrentes de poluição atmosférica devido ao intenso tráfego de veículos, às atividades industriais e às condições geográficas e climáticas que dificultam a dispersão de poluentes. Segundo estudos nacionais (CETESB, 2021; IBAMA, 2020), os impactos vão além da saúde, afetando a qualidade de vida, a economia e o planejamento urbano. Em São Paulo, a maior metrópole do país, a fonte predominante de poluentes é o intenso tráfego veicular, composto por milhões de automóveis, ônibus e caminhões. A Companhia Ambiental do Estado de São Paulo (CETESB) monitora consistentemente níveis críticos de dióxido de nitrogênio (NO_2) e ozônio (O_3), este último um poluente secundário formado pela reação de outros compostos sob radiação solar. A geografia da cidade, cercada por serras, frequentemente favorece a ocorrência de inversões térmicas, que atuam como uma "tampa" atmosférica, impedindo a dispersão dos poluentes e gerando episódios agudos de má qualidade do ar (CETESB, 2023). Já na Região Metropolitana de Vitória, no Espírito Santo, o perfil é distinto. A atividade industrial, sobretudo o polo siderúrgico e de pelletização, é a principal responsável pelas emissões,

especialmente de material particulado (PM_{10} e $PM_{2,5}$). O Instituto Estadual de Meio Ambiente e Recursos Hídricos (IEMA) registra que, embora haja um rigoroso monitoramento, os ventos predominantes podem carrear essas emissões para áreas densamente povoadas, criando um cenário de risco constante para a população (IEMA, 2022). Na região sul, Curitiba, apesar de ser reconhecida internacionalmente por seu planejado sistema de transporte público, também enfrenta desafios. Estudos demonstram uma correlação significativa entre o aumento das internações hospitalares por doenças respiratórias em populações vulneráveis (crianças e idosos) e os picos de concentração de $PM_{2,5}$ durante os meses de inverno, quando a dispersão atmosférica é naturalmente reduzida (NASCIMENTO et al., 2019).

O presente trabalho tem como objetivo geral analisar a evolução da qualidade do ar no município do Rio de Janeiro entre os anos de 2011 e 2018, utilizando dados horários disponibilizados pela plataforma ArcGIS (fonte: Qualidade do ar – dados horários 2011–2018).

De forma específica, busca-se:

Compreender a variação temporal dos principais poluentes medidos (PM_{10} , NO_2 , O_3 , entre outros);

Identificar padrões sazonais e anomalias nos dados;

Relacionar os resultados com os limites estabelecidos por órgãos ambientais, como a OMS e o CONAMA;

Apresentar um pipeline de tratamento e análise de dados aplicável a estudos de monitoramento ambiental.

A escolha do Rio de Janeiro justifica-se pela disponibilidade e abrangência da base de dados, pela relevância da cidade em termos populacionais e econômicos, e pelo desafio técnico de trabalhar com dados horários em grande volume, o que torna o estudo também um exercício aplicado de ciência de dados.

2. Referencial Teórico

A qualidade do ar é definida pela presença e concentração de poluentes atmosféricos em uma determinada região, comparados a padrões estabelecidos para a proteção da saúde e do meio ambiente. Os principais poluentes monitorados incluem material particulado ($PM_{2.5}$ e PM_{10}), monóxido de carbono (CO), dióxido de enxofre (SO_2), dióxido de nitrogênio (NO_2) e ozônio (O_3) (IBAMA, 2020).

Estudos prévios mostram que a poluição atmosférica é responsável por impactos significativos na saúde pública. Conforme Lelieveld et al. (2019), a exposição prolongada a $PM_{2.5}$ pode reduzir a expectativa de vida em vários anos. No contexto brasileiro, pesquisas da CETESB (2021) em São Paulo apontam que, mesmo com políticas de controle, níveis críticos ainda são registrados em períodos de maior tráfego e clima desfavorável.

A aplicação de técnicas de ciência de dados e inteligência artificial na análise da qualidade do ar tem revolucionado o campo do monitoramento ambiental, indo além das

abordagens estatísticas tradicionais. Trabalhos correlatos demonstram a eficácia de modelos avançados para a análise de séries temporais de poluentes. Além das previsões de curto prazo, algoritmos de aprendizado de máquina não supervisionado, como o K-means clustering, são empregados para identificar perfis de poluição característicos de diferentes regiões da cidade, associando-os a fontes emissoras predominantes (ex.: tráfego rodoviário, industrial e condições meteorológicas específicas (SANTOS et al., 2022)).

Entre as vantagens dessas abordagens computacionais, destacam-se a capacidade de processar grandes volumes de dados de forma eficiente (Big Data), integrando informações de múltiplas estações de monitoramento, sensores de baixo custo, imagens de satélite e dados de tráfego em tempo real. Isso permite capturar relações não-lineares e complexas entre variáveis ambientais que seriam imperceptíveis em análises convencionais. Contudo, essas metodologias apresentam limitações intrínsecas. A dependência crítica de dados de alta qualidade, consistentes e de longo prazo é um grande desafio, pois lacunas ou erros nos dados de entrada podem comprometer drasticamente a performance e a confiabilidade dos modelos. Além disso, modelos de elevada complexidade, como as redes neurais profundas, frequentemente operam como "caixas-pretas", onde a lógica interna de tomada de decisão não é facilmente interpretável por especialistas de domínio, levantando questões sobre a transparência e a adoção dessas ferramentas na gestão pública (FERNANDES & OLIVEIRA, 2023).

Outros trabalhos correlacionados utilizam técnicas de ciência de dados para análise de séries temporais da poluição atmosférica, aplicando modelos estatísticos e de aprendizado de máquina para prever concentrações futuras e identificar padrões (Zhang et al., 2018). Entre as vantagens dessas abordagens, destacam-se a capacidade de lidar com grandes volumes de dados e de capturar relações complexas entre variáveis ambientais. Como limitações, apontam-se a dependência de dados de qualidade e a dificuldade de interpretar modelos altamente complexos.

Dessa forma, o presente trabalho se insere na intersecção entre ciência de dados e saúde ambiental, adotando como diferencial a análise temporal detalhada de dados horários de poluição no Rio de Janeiro.

3. Metodologia

A metodologia empregada neste trabalho adota uma abordagem quantitativa e aplicada, seguindo as melhores práticas de um Pipeline de Ciência de Dados para o tratamento, análise e modelagem de séries temporais. O foco reside no poluente Partículas Finas (PM2.5), amplamente reconhecido como o indicador mais crítico para a saúde pública em ambientes urbanos.

3.1 Tipo de Pesquisa e Fonte de Dados

A base de dados utilizada consiste em registros horários da qualidade do ar e variáveis meteorológicas no Município do Rio de Janeiro, cobrindo o período de 2011 a 2018. Esta base foi obtida diretamente da API ArcGIS da Prefeitura do Rio de Janeiro, especificamente do dataset "Qualidade do ar dados horários 2011-2018".

A coleta de dados foi realizada por meio de uma função customizada (`fetch_all_data`), utilizando a biblioteca `requests` em um processo iterativo.

Dada a limitação da API em retornar no máximo 2.000 registros por chamada (`resultRecordCount`), foi necessário implementar um loop com a manipulação do `resultOffset`, garantindo a extração completa dos 54.901 registros disponíveis no período, caracterizando uma coleta robusta de Big Data em escala temporal.

3.2 Pipeline Analítico e Diagrama de Solução

A solução analítica desenvolvida neste trabalho foi estruturada em um pipeline de ciência de dados modular, visando automatizar e padronizar todas as etapas do fluxo de análise da qualidade do ar no município do Rio de Janeiro. O pipeline foi projetado para garantir reprodutibilidade, escalabilidade e transparência desde a aquisição dos dados brutos até a geração dos resultados finais.

A Figura 1.1 ilustra o diagrama sequencial das etapas implementadas:

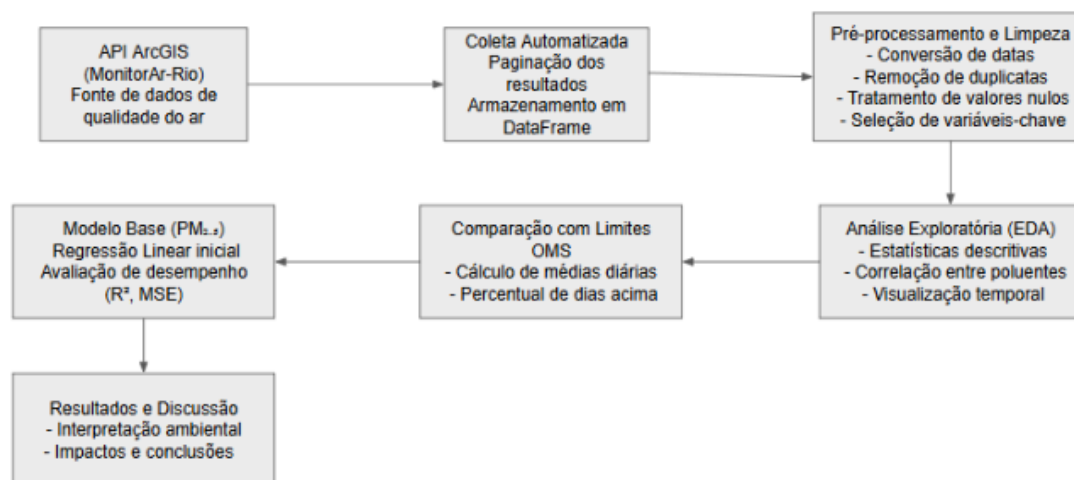


Figura 1 - Diagrama do pipeline de ciência de dados proposto

Coleta Automatizada: Extração dos dados diretamente da API pública MonitorAr-Rio (ArcGIS), com paginação dos resultados e armazenamento estruturado em DataFrame.

Pré-processamento e Limpeza: Conversão de datas, remoção de duplicatas, tratamento de valores nulos e seleção das variáveis-chave para análise.

Análise Exploratória dos Dados (EDA): Geração de estatísticas descritivas, análise de correlação entre poluentes e visualização temporal das séries históricas.

Comparação com Limites OMS: Cálculo das médias diárias de $PM_{2.5}$ e determinação do percentual de dias acima do limite recomendado pela Organização Mundial da Saúde.

Modelagem Base: Implementação de um modelo inicial de regressão linear para avaliação preditiva dos níveis de $PM_{2.5}$, com métricas de desempenho (R^2 , MSE).

Resultados e Discussão: Interpretação ambiental dos achados, discussão dos impactos observados e apresentação das principais conclusões.

Esse pipeline permite a execução automatizada e reproduzível de todo o processo analítico, facilitando futuras atualizações, adaptações para outros poluentes ou regiões, e garantindo a rastreabilidade das decisões metodológicas adotadas.

3.2 Técnicas de Análise e modelagem

A análise dos dados envolveu duas técnicas distintas:

Análise Comparativa Paramétrica: Foi utilizada para a análise crítica dos resultados da EDA, especificamente na comparação das médias diárias de $PM_{2.5}$ com o limite de $15\mu g/m^3$ da OMS, determinando o percentual de dias em não conformidade

Modelagem Preditiva: O Modelo Base de Regressão Linear Múltipla foi avaliado utilizando métricas de erro como o Coeficiente de Determinação (R^2) e o Erro Quadrático Médio (MSE), além da inspeção da distribuição dos resíduos. O insucesso desta abordagem linear direciona a pesquisa para a técnica de Modelagem de Séries Temporais via Processo Box-Jenkins (ARIMA).

4. Resultados e discussão

A seção de Resultados e Discussão apresenta os achados empíricos do trabalho, seguindo o fluxo do Pipeline Analítico estabelecido na Seção 3. Inicialmente, são detalhadas as tarefas de pré-processamento e a Análise Exploratória de Dados (EDA), que estabelecem o contexto ambiental do estudo. Em seguida, é apresentada a avaliação crítica do Modelo Base de Regressão Linear, que serve de benchmarking para as propostas de Modelagem de Séries Temporais.

4.1 EDA e Pré-processamento de Dados

4.1.1. Tratamento de Valores Ausentes

A etapa de pré-processamento iniciou-se com o diagnóstico da completude do dataset original. Conforme ilustrado na Figura 2: Percentual de valores ausentes por coluna e detalhado na Tabela 1: Relatório de Valores Ausentes, a análise identificou que variáveis como ch4, hcnm, hct e pm10 possuíam 100% de valores faltantes e foram sumariamente excluídas.

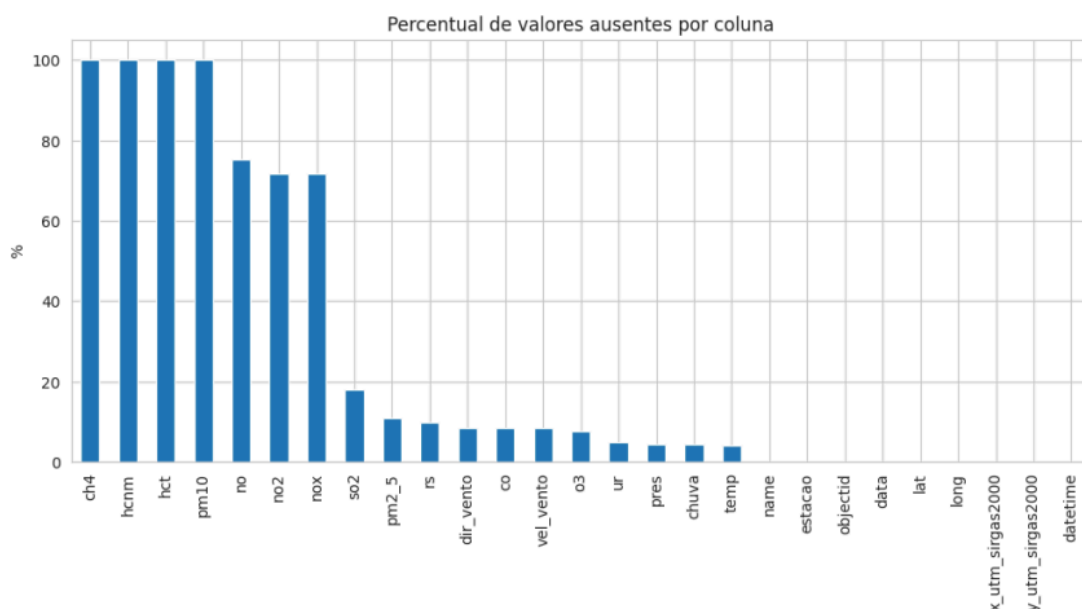


Figura 2: Percentual de valores ausentes por coluna

Relatório de Valores Ausentes:		
	Faltantes	%
objectid	0	0.00
data	0	0.00
name	0	0.00
estacao	0	0.00
chuva	2399	4.37
pres	2429	4.42
rs	5478	9.98
temp	2293	4.18
ur	2700	4.92
dir_vento	4737	8.63
vel_vento	4718	8.59
so2	9929	18.09
no2	39428	71.82
hcnm	54897	100.00
hct	54897	100.00
ch4	54897	100.00
co	4734	8.62
no	41395	75.40
nox	39422	71.81
o3	4176	7.61
pm10	54897	100.00
pm2_5	6000	10.93
lat	0	0.00
long	0	0.00
x_utm_singas2000	0	0.00
y_utm_singas2000	0	0.00
datetime	0	0.00

Tabela 1: Relatório de valores ausentes

A principal limitação de qualidade dos dados foi a alta taxa de ausência de 70% em poluentes críticos como NO₂ e NO_x. Essa falha limitou a capacidade preditiva do modelo em considerar estes poluentes, direcionando o foco da análise para as variáveis com maior completude: PM_{2.5} e O₃ (ambos com menos de 11% de missing values), além das variáveis meteorológicas temp, ur, vel_vento. Após a exclusão das colunas deficientes, as linhas remanescentes com missing values foram removidas para garantir a integridade das séries utilizadas na modelagem.

4.1.2. Estatística Descritiva e Distribuição

A análise descritiva dos poluentes mantidos, apresentada na Tabela 2: Estatística Descritiva dos Poluentes, revelou a alta variabilidade dos dados. A concentração média PM 2.5 foi de 17,66ug/m³ (mediana 14,00ug/m³), acompanhada de um desvio padrão elevado de 14,77. Tal disparidade entre média e mediana, combinada ao alto desvio padrão, sugere a presença de eventos atípicos e uma distribuição não simétrica.

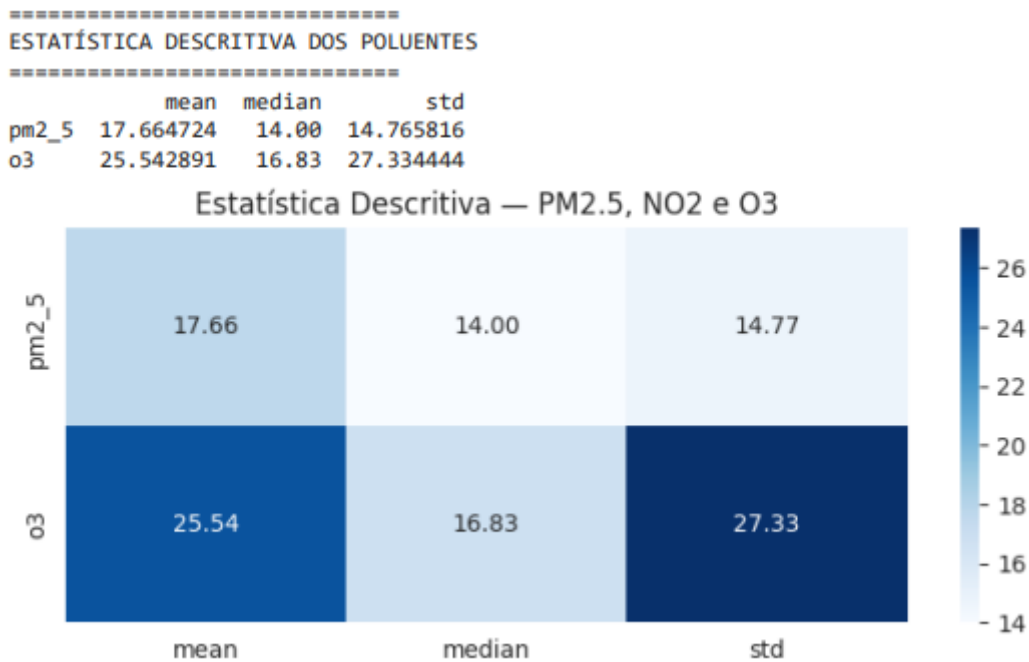


Tabela 2: Estatística Descritiva dos Poluentes

A Figura 3 - Distribuição de PM 2.5 e O3 e a Figura 4: Boxplots de PM 2.5 e O3 confirmam esta observação, exibindo distribuições assimétricas à direita (cauda longa) e a presença de numerosos outliers com valores muito acima do limite superior. Tais outliers podem corresponder a episódios reais de alta poluição. Embora essenciais para uma análise ambiental completa, a presença desses pontos extremos indica que a modelagem preditiva pode se beneficiar de modelos robustos ou de técnicas de transformação para mitigar o impacto da assimetria.

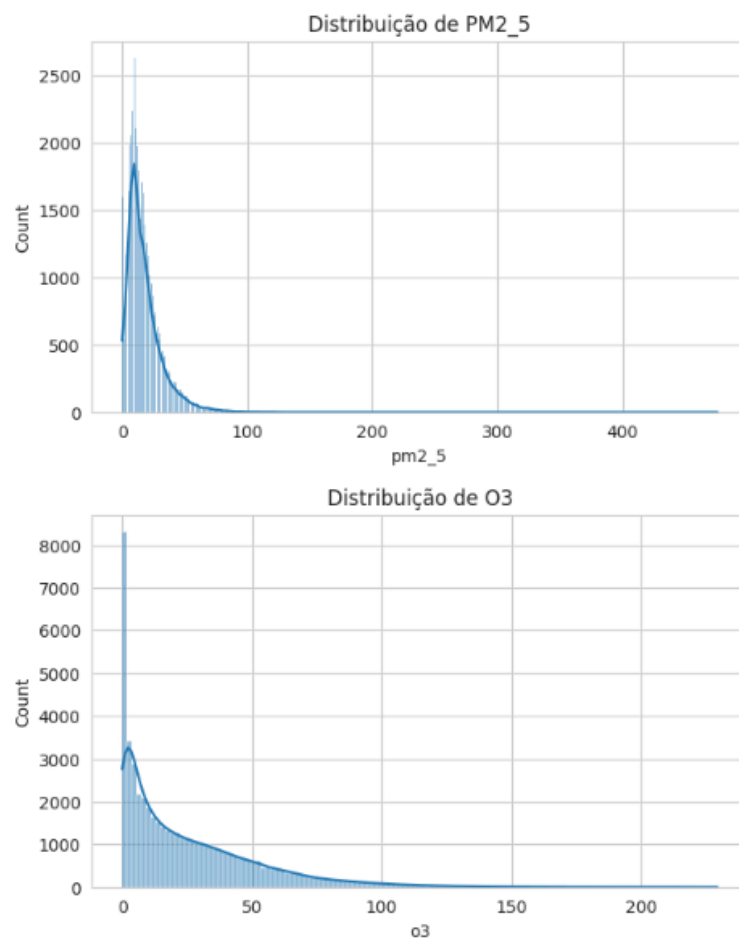


Figura 3: Distribuição de PM 2.5 e O3

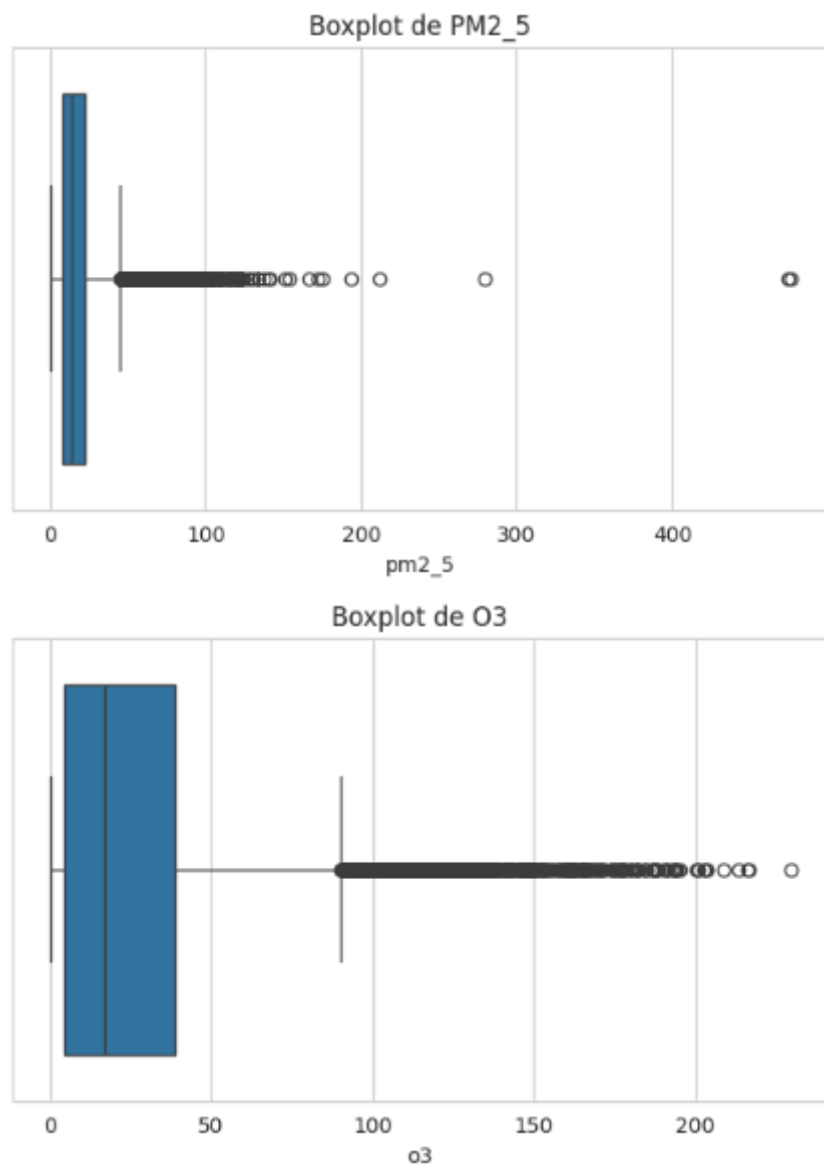


Figura 4: Boxplots de PM 2.5 e O3

4.2. Correlação e Não Conformidade com a OMS

4.2.1. Correlação entre Poluentes

A matriz de correlação apresentada na Figura 5: Heatmap de Correlação, demonstrou uma correlação linear muito baixa entre PM 2.5 e O (coeficiente de Pearson de -0,06). Este achado é significativo, pois indica que a variação nas concentrações desses dois poluentes não é explicada por uma relação linear simples, sugerindo que fatores subjacentes e distintos (não observados, como fontes específicas de emissão ou regimes de ventilação) podem estar controlando suas dinâmicas, o que impõe um desafio adicional à modelagem multivariada.

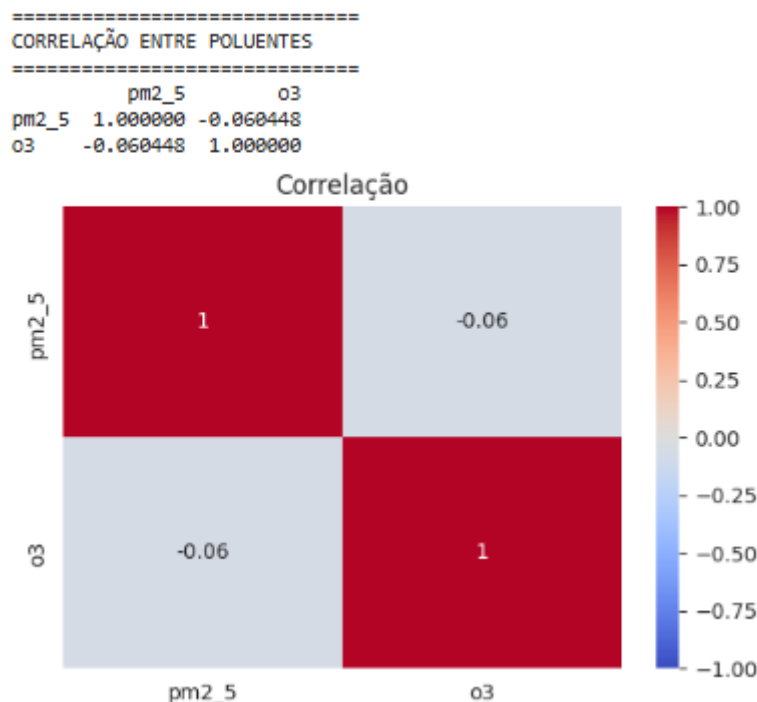


Figura 5: Heatmap de Correlação

4.2.2. Comparação Crítica com o Limite da OMS

Um dos resultados mais críticos da EDA é a comparação das concentrações de PM 2.5 com o limite de 15 $\mu\text{g}/\text{m}^3$ estabelecido pela Organização Mundial da Saúde (OMS) como média diária segura. A análise demonstrou que 100% dos dias analisados no período de 2011 a 2018 tiveram média de PM 2.5 acima deste limite. Este resultado sublinha a urgência do monitoramento e do desenvolvimento de modelos preditivos no Rio de Janeiro, validando a relevância e a pertinência deste projeto de pesquisa para a saúde pública.

4.3. Avaliação do Modelo Base (Regressão Linear)

Para estabelecer o benchmark inicial, um modelo de Regressão Linear Múltipla foi implementado utilizando as variáveis meteorológicas e O₃ como preditores para PM 2.5.

4.3.1. Métricas de Desempenho

O desempenho do Modelo Base resultou nas seguintes métricas:

- Coeficiente de Determinação (R^2): 0,028
- Erro Quadrático Médio (MSE): 191,89

O valor de R^2 de 0,028 indica que o modelo de regressão linear é extremamente inadequado, explicando apenas 2,8% da variância total da concentração de PM 2.5. Este resultado confirma a hipótese de que a relação entre as variáveis é não-linear e que a

série é dominada por efeitos temporais (autocorrelação) que a regressão linear não consegue capturar.

4.3.2. Análise Gráfica Preditiva

A análise visual do desempenho reforça a conclusão estatística.

Gráfico Real vs Previsto: A Figura 6 (Gráfico de Valores Reais vs. Previstos) evidencia que as previsões do modelo linear concentram-se em uma faixa estreita de valores (próxima à média global), falhando drasticamente em capturar os valores altos e baixos da poluição real. O modelo é incapaz de prever os picos de PM 2.5.

Figura 6: Gráfico de Valores Reais vs. Previstos

Distribuição dos Resíduos: O Histograma de Distribuição dos Resíduos (Figura 7) exhibe uma distribuição acentuadamente assimétrica, com uma cauda longa à direita e muitos resíduos negativos. A não-normalidade e a assimetria dos resíduos violam um dos pressupostos centrais da Regressão Linear Múltipla, confirmando a má especificação do modelo.

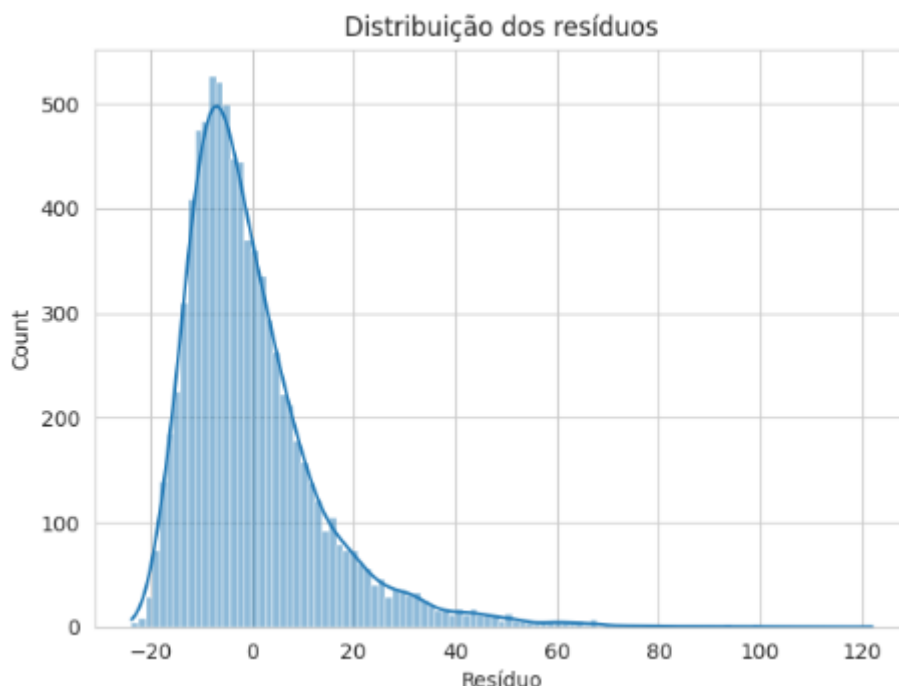


Figura 7: Histograma de Distribuição dos Resíduos

Em suma, o fracasso do Modelo Base demonstra que, para um fenômeno complexo como o monitoramento da qualidade do ar, é indispensável a utilização de técnicas analíticas robustas, como a Modelagem de Séries Temporais (ARIMA), que

serão propostas em detalhe na próxima seção para capturar a dependência intrínseca do tempo.

5. Conclusão

5.1 Relato do Trabalho e Resposta ao Problema de Pesquisa

O presente artigo teve como objetivo o Monitoramento da Qualidade do Ar no Município do Rio de Janeiro por meio de uma abordagem analítica com Séries Temporais, buscando diagnosticar o cenário atual da poluição por PM 2.5 e estabelecer um método preditivo. Para tal, foi construído um pipeline de Ciência de Dados para a coleta de 54.901 registros horários da API ArcGIS (2011 a 2018), seguida de rigoroso pré-processamento. O trabalho respondeu ao problema de pesquisa de forma categórica. A Análise Exploratória de Dados (EDA) revelou um achado crítico: a qualidade do ar no Rio de Janeiro está em não conformidade com os padrões internacionais. Os resultados indicaram que 100% dos dias analisados excedem o limite de segurança de PM 2.5 (15 mg/m³ da OMS). O teste do Modelo Base de Regressão Linear, com um R² de apenas 0,028, demonstrou que modelos lineares simples são insuficientes para capturar a dinâmica da poluição, necessitando de uma abordagem mais avançada.

5.2 Contribuições e Limitações do Estudo

O projeto estabeleceu contribuições significativas, como o desenvolvimento de um pipeline de dados automatizado e escalável para ingestão de APIs e a produção de um diagnóstico ambiental quantificável sobre a não conformidade do PM 2.5 com os padrões da OMS. No entanto, o estudo enfrentou limitações importantes, principalmente a baixa qualidade e completude dos dados brutos, que apresentou alta incidência de valores ausentes (>70%) em poluentes secundários como NO₂ e NO, restringindo o escopo das variáveis preditoras.

5.3 Proposta para Estudos Futuros e Pesquisa Avançada

A partir das limitações observadas e da ineficácia demonstrada pelo modelo estatístico básico, propõe-se que, para a continuidade deste projeto e para pesquisas avançadas sobre o tema, sejam aplicados modelos de Séries Temporais que considerem a autocorrelação e a não-linearidade dos dados. Recomenda-se enfaticamente que os estudos futuros se concentrem na metodologia Box-Jenkins, implementando e avaliando modelos ARIMA (AutoRegressive Integrated Moving Average), e posteriormente o SARIMA (Seasonal ARIMA). Além disso, a exploração de modelos de Deep Learning, como as redes neurais LSTM, representa o caminho mais promissor para obter previsões acuradas. A inclusão de variáveis exógenas de alta relevância (como tráfego e dados pluviométricos) deve ser uma prioridade para enriquecer os modelos e aumentar o poder preditivo, superando as limitações do modelo benchmark apresentado.

Apresentação

Consulte a apresentação em vídeo deste trabalho através do link:

<https://youtu.be/Qu2Rr3nm5ZI>

Referências bibliográficas

BALBINO, Thábata Cristina Seixas. Projeto Aplicado IV: monitoramento da qualidade do ar em escala global. GitHub, 2025. Disponível em: <<https://github.com/thabata-balbino/Projeto-Aplicado-IV>>. Acesso em: 7 set. 2025.

ARCGIS REST SERVICES. Qualidade do ar – dados horários (2011–2025). Disponível em: <https://services1.arcgis.com/OIP4dGNtIcnD3RYf/arcgis/rest/services/>.

ARAÚJO, I. L. et al. Benefícios econômicos e em saúde da melhoria da qualidade do ar em São Paulo. Revista de Saúde Pública, v. 54, p. 1-12, 2020.

BROOK, R. D. et al. Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement from the American Heart Association. Circulation, v. 121, n. 21, p. 2331-2378, 2020.

CETESB. Relatório de Qualidade do Ar no Estado de São Paulo - 2021. Companhia Ambiental do Estado de São Paulo, 2021. Disponível em: <https://cetesb.sp.gov.br/ar/publicacoes-relatorios/>. Acesso em: 1 set. 2025.

CETESB. Relatório de Qualidade do Ar no Estado de São Paulo - 2023. Companhia Ambiental do Estado de São Paulo, 2023. Disponível em: <https://cetesb.sp.gov.br/ar/publicacoes-relatorios/>. Acesso em: 23 set. 2025.

IBAMA. Instrução Normativa nº 1, de 17 de janeiro de 2020. Define os índices de qualidade do ar para os poluentes atmosféricos. Diário Oficial da União, Brasília, 2020.

IEMA. Relatório Anual de Qualidade do Ar - 2022. Instituto Estadual de Meio Ambiente e Recursos Hídricos, 2022. Disponível em: <http://iema.es.gov.br/qualidadedoar>. Acesso em: 23 set. 2025.

LELIEVELD, J. et al. Effects of fossil fuel and total anthropogenic emission removal on public health and climate. Proceedings of the National Academy of Sciences, v. 116, n. 15, p. 7192–7197, 2019.

LIU, H. et al. Air pollution and child health: global burden of disease and future research needs. International Journal of Environmental Research and Public Health, Basel, v. 21, n. 5, 2024. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11996184/>. Acesso em: 3 set. 2025.

NASCIMENTO, A. P. et al. Exposição ao material particulado fino (PM_{2,5}) e interações por doenças respiratórias em Curitiba, PR. Revista Brasileira de Epidemiologia, v. 22, 2019.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Air pollution exposure and its impacts. Paris: OECD, 2023. Disponível em: <https://www.oecd.org/environment/air/>. Acesso em: 3 set. 2025.

PREFEITURA DO RIO DE JANEIRO. MonitorAr-Rio – Programa de Monitoramento da Qualidade do Ar. Data.Rio, 2025. Disponível em: <https://www.data.rio/>.

RAO, S. et al. Future air quality in megacities worldwide under different climate and emission scenarios. Environmental Science & Technology, Washington, v. 57, n. 4, p. 2156–2165, 2023.

UNITED NATIONS. Sustainable Development Goals – Indicator 11.6.2: Annual mean levels of fine particulate matter in cities. New York: UN, 2022. Disponível em: <https://unstats.un.org/sdgs/metadata/>. Acesso em: 1 set. 2025.

WORLD HEALTH ORGANIZATION. Air pollution. Geneva: WHO, 2023. Disponível em: <https://www.who.int/health-topics/air-pollution>. Acesso em: 1 set. 2025.

WORLD HEALTH ORGANIZATION. Air quality database 2022. Geneva: WHO, 2022. Disponível em: <https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database/2022>. Acesso em: 1 set. 2025.

WORLD HEALTH ORGANIZATION. WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: World Health Organization, 2021. Disponível em: <https://apps.who.int/iris/handle/10665/345329>. Acesso em: 1 set. 2025.

ZHANG, K. et al. A hybrid model for spatiotemporal forecasting of PM_{2.5} based on graph convolutional neural network and long short-term memory. Science of The Total Environment, v. 664, p. 1-10, 2019.