ANDREI SOUZA DE OLIVEIRA

GABRIELA OHASHI DE SOUZA

MARINA OHASHI DE SOUZA

MATHEUS FERREIRA FRANÇA

MIGUEL MAURÍCIO TADEU PITALI DA SILVA

**PROJETO APLICADO I**

**Poluição atmosférica no estado de São Paulo e efeitos nocivos na saúde da população, no período de 2015 a 2021**

São Paulo

2023

**SUMÁRIO**

1. INTRODUÇÃO ......................................................................................... 3
2. PREMISSAS ............................................................................................ 4
3. OBJETIVOS E METAS ............................................................................ 5
4. GITHUB DO PROJETO ........................................................................... 6
5. CONTEXTO DO ESTUDO ....................................................................... 7
6. REFERÊNCIAS DE AQUISIÇÃO DO DATASET ..................................... 8
7. DESCRIÇÃO DA ORIGEM ...................................................................... 9
8. DESCRIÇÃO DO DATASET ................................................................... 10
9. ANÁLISE EXPLORATÓRIA DOS DADOS...............................................11
10. LISTA DE FIGURAS ............................................................................... 12
11. LISTA DE TABELAS ............................................................................... 13
12. CRONOGRAMA DO PROJETO ............................................................. 14

**INTRODUÇÃO**

A poluição ambiental vem causando impactos, nos últimos anos, e dados preocupantes sobre eles nos levam a refletir sobre o efeito dos mesmos em nossas vidas. O Brasil, que passou por um grande processo de urbanização, tende a sofrer ainda mais com as consequências desses impactos, e em uma cidade como São Paulo essa premissa não seria diferente. Os problemas de saúde são apenas uma pequena parcela dessas consequências.

O atual trabalho tem como foco expor a forma em que essa poluição afeta a saúde dos moradores de São Paulo, no período de 2015 a 2021.

**PREMISSAS**

Avaliar a poluição atmosférica no estado de São Paulo, no período de 2015 a 2021, e relacioná-la aos efeitos nocivos na saúde da população, com base no índice brasileiro da qualidade do ar definido pelo IEMA (Instituto de Energia e Meio Ambiente).

Utilização de base de dados disponível no Kaggle: <https://www.kaggle.com/datasets/samirnunesdasilva/sao-paulo-pollution-data>

**OBJETIVOS E METAS**

Identificar:

1 - Poluente mais crítico;

2 - Localidade mais afetada pela poluição;

3 - Período do ano com elevação no nível de concentração de CO; e

4 - Horário de maior concentração de CO.

Correlacionando esses indicadores com os efeitos na saúde da população.

Tabela 1 – Estrutura do índice brasileiro de qualidade do ar e efeitos à saúde

Tabela

Descrição gerada automaticamente

**GITHUB DO PROJETO**

[https://github.com/OhashiMarina/Projeto-Aplicado-I.git](https://bra01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fgithub.com%2FOhashiMarina%2FProjeto-Aplicado-I.git&data=05%7C01%7C%7Ca159a47157094da4aa8808db22ff7cec%7C84df9e7fe9f640afb435aaaaaaaaaaaa%7C1%7C0%7C638142252929122044%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=ddUPTaWCIXU%2BMwoq8OggR21JD%2BVL1BKK6QvFdLzdnuE%3D&reserved=0)

**CONTEXTO DO ESTUDO**

Avaliar a poluição atmosférica no estado de São Paulo, no período de 2015 a 2021, e relacioná-la aos efeitos nocivos na saúde da população, com base no índice brasileiro da qualidade do ar definido pelo IEMA (Instituto de Energia e Meio Ambiente).

**REFERÊNCIAS DE AQUISIÇÃO DO DATASET**

Utilização de base de dados disponível no Kaggle - [https://www.kaggle.com/datasets/samirnunesdasilva/sao-paulo-pollution-data](https://bra01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.kaggle.com%2Fdatasets%2Fsamirnunesdasilva%2Fsao-paulo-pollution-data&data=05%7C01%7C%7C8745d2d1f14c4ab6787008db22f3c52b%7C84df9e7fe9f640afb435aaaaaaaaaaaa%7C1%7C0%7C638142202605758390%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=twlb4MXFmjhcUAs5%2Ft8TtM2WDPAX6A8DX%2FqVO7UEvLA%3D&reserved=0)

Período: 2015 a 2021

Conjunto de dados com mais de 10 milhões de linhas englobando 87 estações de medição distintas distribuídas ao longo do estado de São Paulo que podem medir a concentração de 9 poluentes: MP10, O3, NO2, MP2.5, CO, SO2, NO, FMC e PTS.

**DESCRIÇÃO DA ORIGEM**

Fonte: Plataforma Qualidade do Ar, do Instituto de Energia e Meio Ambiente (IEMA);  
Site: <https://energiaeambiente.org.br/qualidadedoar#secao-14>.

**DESCRIÇÃO DO DATASET**

Colunas possuem as seguintes descrições:

**ID:** identificação, via índice inteiro, de cada registro do dataframe (chave primária);

**Data:** data em que foi feita a medição da concentração do poluente;

**Hora:** hora em que foi feita a medição;

**Estação:** local em que foi feita a medição;

**Código:** código associado à estação em que foi realizada a medição;

**Poluente:** poluente cuja concentração foi medida;

**Valor:** valor, na unidade especificada, da concentração do poluente;

**Unidade:** unidade de concentração utilizada; e

**Tipo:** como foi realizada a medição (de forma automática ou manual).

**ANÁLISE EXPLORATÓRIA DOS DADOS**

Realizamos os seguintes procedimentos em Python para a exploração dos dados:

**Avaliação primária dos dados:**

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

**Condicional dos Aferíveis**

Considerando a necessidade de correlacionar os indicadores estabelecidos no objetivo do projeto com a Tabela 1, identificou-se a necessidade de exclusão dos poluentes PTS, FMC e NO da base de dados.

Texto

Descrição gerada automaticamente

**Criação de novo *dataframe***

Novo *dataframe* considerando apenas os poluentes a serem analisados, conforme Tabela 1.

Interface gráfica do usuário, Tabela

Descrição gerada automaticamente

**Respondendo as perguntas do Objetivo do Projeto**

**1 - Poluente mais crítico**

Foi necessário definir a regra de classificação considerando a Tabela 1

Calendário

Descrição gerada automaticamente com confiança média

Como resultado, aparece um erro no notebook

Foi Atribuído um índice de nocividade de 0 a 5, sendo 0 o menos nocivo e 5 o mais nocivo, podemos notar que o poluente CO apresenta um maior risco à saúde que os demais poluentes.

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

**2 - Localidades mais afetadas pela poluição**

Considerando que o CO é o poluente de maior risco à saúde da população, foi verificado as estações de medição que tiveram a concentração desse poluente acima da média geral.

Tabela

Descrição gerada automaticamente com confiança média

**3 - Período do ano com maior nível de concentração de CO**

**4 - Horário de maior concentração de CO**

**LISTA DE FIGURAS**

**LISTA DE TABELAS**

Tabela 1 – Estrutura do índice brasileiro de qualidade do ar e efeitos à saúde

Tabela

Descrição gerada automaticamente

**CRONOGRAMA DO PROJETO**

Tabela

Descrição gerada automaticamente

Arquivo Excel disponível no Github.

Observação: Dias restantes contados da data da elaboração desse documento, 10/05/2023.