|  |
| --- |
| CURSO: Tecnologia em Ciência de Dados |
| SEMESTRE: 2° |
| COMPONENTE CURRICULAR / Projeto Aplicado 1 |
| NOMES COMPLETOS DOS ALUNOS:  Andrei Souza de Oliveira – TIA 1092250600  Gabriela Ohashi de Souza – TIA 10922521097  Marina Ohashi de Souza – TIA 109225209714  Matheus Ferreira França – TIA 10922519211  Miguel Maurício Tadeu Pitali da Silva – TIA 10922507310 |
| NOME DO PROFESSOR: Leonardo Massayuki Takuno |

**ETAPA 1 – 30 dias – Metas e milestones**

• **Premissas do projeto.**

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](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)

• **Objetivos e metas.**

Identificar:

- poluente mais crítico;

- localidade mais afetada pela poluição;

- período do ano com elevação no nível de concentração de poluentes; e

- horário de maior concentração de poluentes.

Correlacionando esses indicadores com os efeitos na saúde da população.

• **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)

• **Título do trabalho**

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

• **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 (origem dos dados, limitações de uso e período da coleta)**

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 (informações sobre a organização que gerou os dados e o contexto em que os dados foram gerados)**

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

**• Descrição do dataset (o que ele contém, qual é a proposta, quais problemas ou fenômenos foram registrados)**

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).

• **Pensamento computacional em contextos organizacionais**

O Pensamento Computacional é um processo de resolução de problemas que é essencial nas organizações das mais diversas áreas.