

## **Especificação de um experimento de ciência de dados para análise de áreas contaminadas**

**Rosa Virginia Encinas Quille  
Gabriela dos Santos Luchetti Vieira  
Leandro Gomes de Freitas  
Pedro Luiz Pizzigatti Corrêa,  
Solange Nice Alves de Souza  
Alexandre Muselli Barbosa  
Felipe Valencia de Almeida**

*Palestra apresentado no  
WORKSHOP DE COMPUTAÇÃO APLICADA À  
GESTÃO DO MEIO AMBIENTE E RECURSOS  
NATURAIS, 16., 2025, Maceió. Palestra... 21 slides.*

A série “Comunicação Técnica” compreende trabalhos elaborados por técnicos do IPT, apresentados em eventos, publicados em revistas especializadas ou quando seu conteúdo apresentar relevância pública. **PROIBIDO REPRODUÇÃO**

Instituto de Pesquisas Tecnológicas do Estado de São Paulo  
S/A - IPT  
Av. Prof. Almeida Prado, 532 | Cidade Universitária ou  
Caixa Postal 0141 | CEP 01064-970  
São Paulo | SP | Brasil | CEP 05508-901  
Tel 11 3767 4374/4000 | Fax 11 3767-4099

[www.ipt.br](http://www.ipt.br)

# Especificação de um Experimento de Ciência de Dados para Análise de Áreas Contaminadas

Rosa Quille, Gabriela Vieira, **Leandro Gomes de Freitas**, Pedro Corrêa,  
Solange Alves-Souza, Alexandre Barbosa e Felipe Almeida



---

# Sumário

- Introdução
- Estudo de Caso USP-Leste
- Metodologia
- Resultados e Discussões
- Conclusões

# Importância do Gerenciamento de Áreas Contaminadas e o desafio da gestão de dados



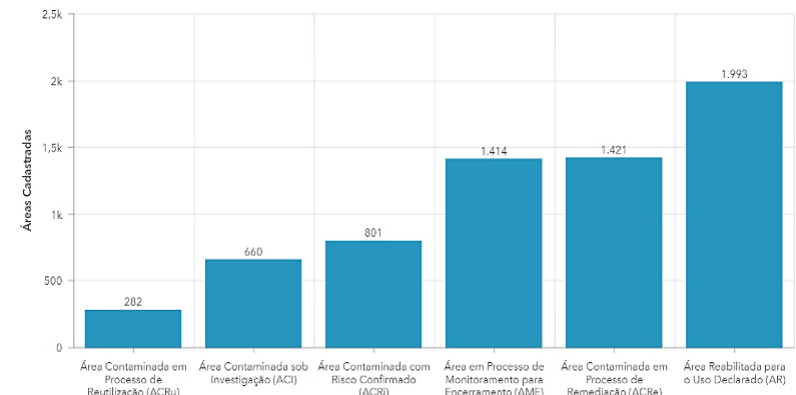
6.984



Áreas Contaminadas e Reabilitadas

Última Atualização: há 4 minutos

Áreas aptas para o uso ou com medidas de intervenção aprovadas:

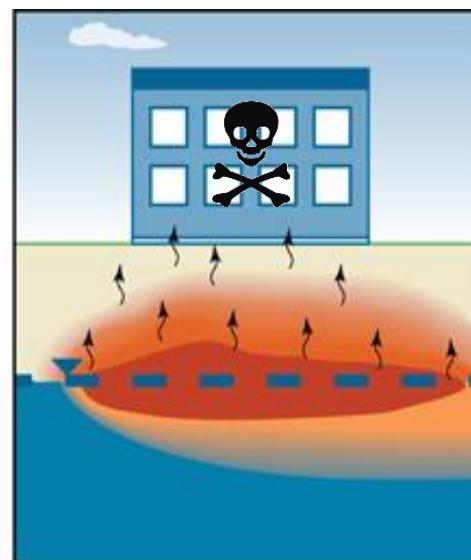
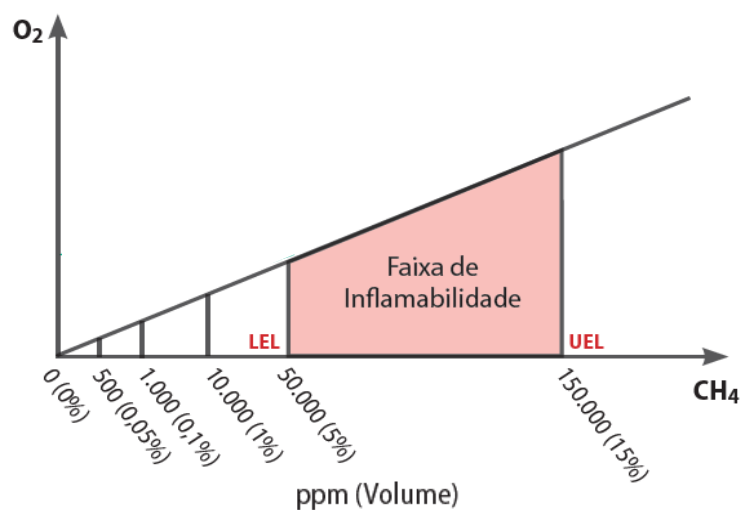
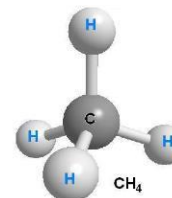


| building_id          | well_id | depth_level | sampling_date       | lower_chl_concentration | chl_peak_concentration | chl_concentration | nitrate_concentration | h2s_concentration | res_bacteria_percent | well_pressure_bar |
|----------------------|---------|-------------|---------------------|-------------------------|------------------------|-------------------|-----------------------|-------------------|----------------------|-------------------|
| Edifício 1 - Parte 1 | PMG-16  | A           | 2020-01-06 00:00:00 | 0%                      | 0%                     | 0%                | 17%                   | 0%                | 78%                  | 0,34              |
| Edifício 1 - Parte 1 | PMG-16  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,044             | 0,137                 | 0                 | 0,799                | 0,05              |
| Edifício 1 - Parte 1 | PMG-17  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,03              | 0,177                 | 0                 | 0,791                | 0,05              |
| Edifício 1 - Parte 1 | PMG-17  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,038             | 0,174                 | 0                 | 0,798                | 0,21              |
| Edifício 1 - Parte 1 | PMG-18  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,033             | 0,201                 | 0                 | 0,795                | 0,06              |
| Edifício 1 - Parte 1 | PMG-18  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,033             | 0,201                 | 0                 | 0,795                | 0,08              |
| Edifício 1 - Parte 1 | PMG-20  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,009             | 0,196                 | 0                 | 0,795                | 0,07              |
| Edifício 1 - Parte 1 | PMG-20  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,008             | 0,184                 | 0                 | 0,798                | 0,17              |
| Edifício 1 - Parte 1 | PMG-22  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,005             | 0,197                 | 0                 | 0,798                | 0,17              |
| Edifício 1 - Parte 1 | PMG-22  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,009             | 0,182                 | 0                 | 0,807                | 0,21              |
| Edifício 1 - Parte 1 | PMG-23  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,027             | 0,174                 | 0                 | 0,799                | 0,12              |
| Edifício 1 - Parte 1 | PMG-23  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,027             | 0,188                 | 0                 | 0,795                | 0,16              |
| Edifício 1 - Parte 1 | PMG-113 | A           | 2020-01-06 00:00:00 | 0                       | 0,002                  | 0,205             | 0                     | 1                 | 0,791                | 0,18              |
| Edifício 1 - Parte 1 | PMG-113 | B           | 2020-01-06 00:00:00 | 0                       | 0,019                  | 0,193             | 0                     | 1                 | 0,788                | 0,17              |
| Edifício 1 - Parte 1 | PMG-114 | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,01              | 0,197                 | 0                 | 0,791                | 0,17              |
| Edifício 1 - Parte 1 | PMG-114 | B           | 2020-01-06 00:00:00 | 0,001                   | 0,001                  | 0,006             | 0,195                 | 0                 | 0,798                | 0,14              |
| Edifício 1 - Parte 1 | PMG-115 | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,009             | 0,195                 | 0                 | 0,798                | 0,19              |
| Edifício 1 - Parte 1 | PMG-115 | B           | 2020-01-06 00:00:00 | 0                       | 0,012                  | 0,184             | 0                     | 0                 | 0,794                | 0,01              |
| Edifício 1 - Parte 2 | PMG-11  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,014             | 0,191                 | 0                 | 0,795                | 0,11              |
| Edifício 1 - Parte 2 | PMG-11  | B           | 2020-01-06 00:00:00 | 0,7%                    | 0,7%                   | 4,0%              | 14,6%                 | 0,0%              | 80,1%                | 0,15              |
| Edifício 1 - Parte 2 | PMG-12  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,005             | 0,203                 | 0                 | 0,791                | 0,19              |
| Edifício 1 - Parte 2 | PMG-12  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,006             | 0,203                 | 0                 | 0,791                | 0,11              |
| Edifício 1 - Parte 2 | PMG-13  | A           | 2020-01-06 00:00:00 | 0                       | 0,04                   | 0,191             | 0                     | 0                 | 0,789                | 0,21              |
| Edifício 1 - Parte 2 | PMG-13  | B           | 2020-01-06 00:00:00 | 0                       | 0,027                  | 0,173             | 0                     | 0                 | 0,791                | 0,29              |
| Edifício 1 - Parte 2 | PMG-14  | A           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,019             | 0,179                 | 0                 | 0,792                | 0,12              |
| Edifício 1 - Parte 2 | PMG-14  | B           | 2020-01-06 00:00:00 | 0                       | 0,042                  | 0,153             | 0                     | 0                 | 0,805                | 0,09              |
| Edifício 1 - Parte 2 | PMG-15  | A           | 2020-01-06 00:00:00 | 0                       | 0,01                   | 0,191             | 0                     | 0                 | 0,799                | 0,19              |
| Edifício 1 - Parte 2 | PMG-15  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,005             | 0,2                   | 0                 | 0,795                | 0,14              |
| Edifício 1 - Parte 2 | PMG-19  | A           | 2020-01-06 00:00:00 | 0                       | 0,01                   | 0,197             | 0                     | 0                 | 0,791                | 0,19              |
| Edifício 1 - Parte 2 | PMG-19  | B           | 2020-01-06 00:00:00 | 0                       | 0,012                  | 0,193             | 0                     | 0                 | 0,795                | 0,11              |
| Edifício 1 - Parte 2 | PMG-21  | A           | 2020-01-06 00:00:00 | 0                       | 0,01                   | 0,179             | 0                     | 0                 | 0,791                | 0,11              |
| Edifício 1 - Parte 2 | PMG-21  | B           | 2020-01-06 00:00:00 | 0                       | 0                      | 0,028             | 0,178                 | 0                 | 0,795                | 0,29              |

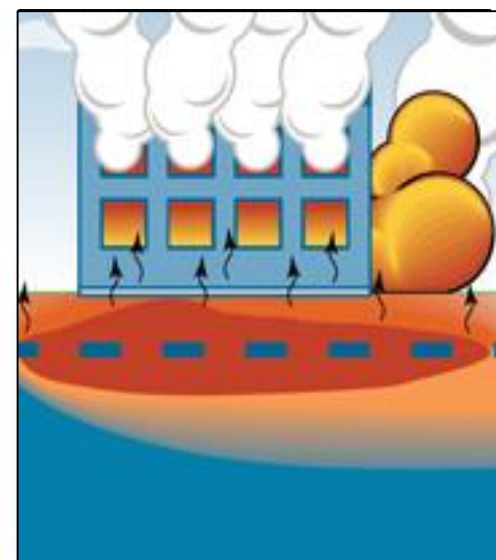


# Riscos relacionados ao gás metano ( $\text{CH}_4$ ) no subsolo

- Intrusão e acúmulo em ambientes fechados;
- Toxicidade, asfixia e explosividade;



Complete or Potentially Complete PVI Pathway



Emergency Situation

Fonte: ITRC, 2009

# O caso da USP-Leste

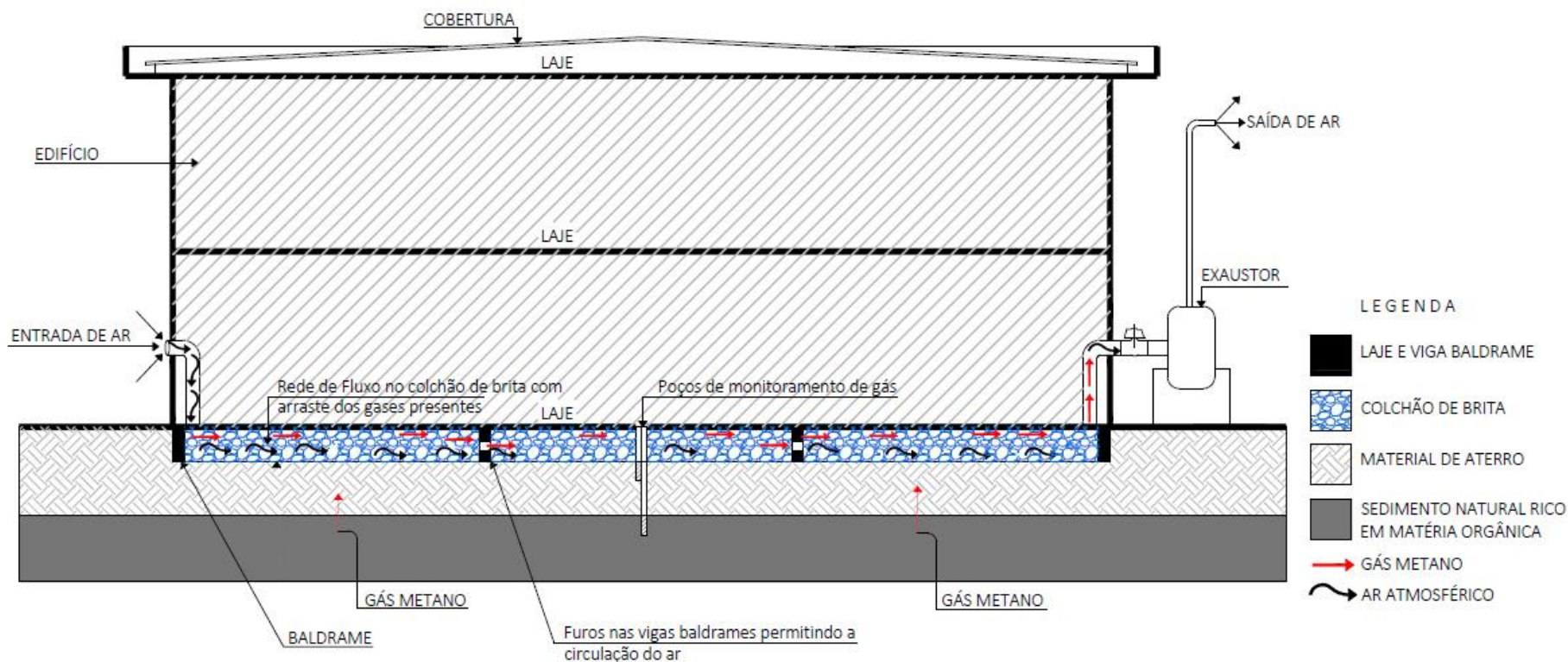


# Histórico dos estudos ambientais

- Entre as décadas de 1980 e 1990 a área foi bota-fora de materiais dragados no desassoreamento e retificação do rio Tietê;
- Relatório Ambiental Preliminar (2004) identificou a presença de gás metano no subsolo;
- Projeto e construção do Campus (entre 2005 e 2014) incluiu estruturas de proteção dos edifícios para evitar acúmulo de gases sob as edificações;
- Sistemas de drenagem dos gases foram instalados entre 2013 e 2014 e vêm sendo monitorados mensalmente;

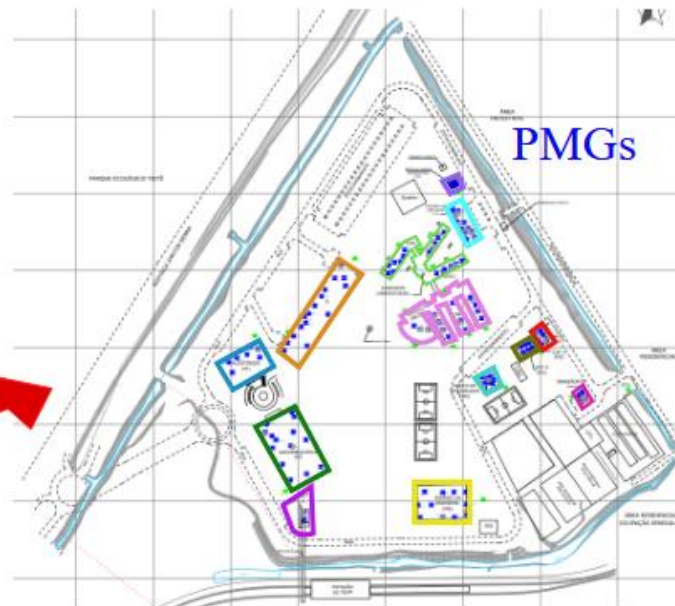
# O caso da USP-Leste

## Modelo esquemático dos sistemas de ventilação



# Malha de amostragem de PMGs

- 125 conjuntos de Poços de Monitoramento de Gases do solo (PMGs), com 2 níveis de profundidade cada, totalizando 250 PMGs. Monitorados sistematicamente;
- Conjunto de dados (dataset) considerado no período 2014 e 2022: 100.570 linhas de resultados: <https://wds.poli.usp.br/sga/discovery/>
- Fonte original dos dados (PDFs não estruturados): <https://www.sef.usp.br/usp-leste/acoes-ambientais/>



## Edifícios

- Edifício I-1 parte 1 e 2 (E5 e E6)
- Edifício I-3 Biblioteca (E3)
- Edifício I-3 Auditórios I-5 (E8)
- Edifício I-4 (E4)
- Conjunto Laboratorial (E1)
- Ginásio Poliesportivo (E10)
- CAT (E9)
- Incubadora (E11)
- Guarda Universitária (E12)
- Portaria P3 (E14)
- Transporte (E13)
- Bloco Inicial (Conjunto Didático) (E2)
- Enfermaria (E7)

# Metodologia – Ciclo de análise de dados

- 1. Estabelecer Questionamentos**
- 2. Aquisição de Dados**
- 3. Limpeza dos Dados**
- 4. Modelagem**
- 5. Análise de Dados**
- 6. Apresentação**
- 7. Análise de Resultados**
- 8. Refinar o Estudo**

# Ciclo de análise de dados

1. **Estabelecer Questionamentos:** Definir com clareza os objetivos da análise, guiando-se por perguntas específicas sobre o problema investigado.
2. **Aquisição de Dados:** Coletar dados relevantes com base nos questionamentos definidos, utilizando instrumentos e fontes adequadas ao contexto do estudo.
3. **Limpeza dos Dados:** Garantir a integridade dos dados por meio da exclusão de registros irrelevantes, tratamento de erros e padronização de formatos.

# Resultados

- Distribuição das frequências e limpeza dos dados medições na USP-Leste

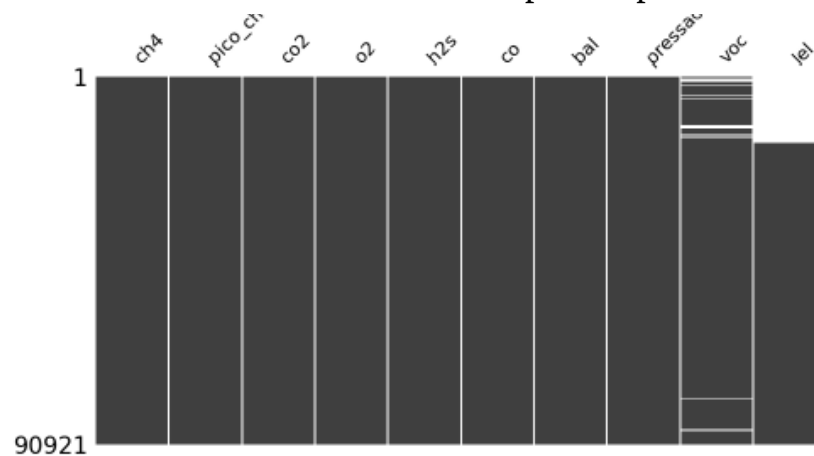
| ID  | Edifício                          | AF     | RF    | RF %    | CF      | CF %     |
|-----|-----------------------------------|--------|-------|---------|---------|----------|
| E1  | Conjunto Laboratorial             | 14.586 | 0,145 | 14,50 % | 14.586  | 14,50 %  |
| E2  | Bloco Inicial (conjunto didático) | 11.996 | 0,119 | 11,93 % | 26.582  | 26,43 %  |
| E3  | I-3 Biblioteca                    | 11.986 | 0,119 | 11,92 % | 38.568  | 38,35 %  |
| E4  | Edifício I-4                      | 10.272 | 0,102 | 10,21 % | 48.840  | 48,56 %  |
| E5  | Edifício I-1 Parte 1              | 7.706  | 0,077 | 7,66 %  | 56.546  | 56,23 %  |
| E6  | Edifício I-1 Parte 2              | 6.850  | 0,068 | 6,81 %  | 63.396  | 63,04 %  |
| E7  | Enfermaria                        | 5.982  | 0,059 | 5,95 %  | 69.378  | 68,98 %  |
| E8  | I-3 Auditórios                    | 5.978  | 0,059 | 5,94 %  | 75.356  | 74,93 %  |
| E9  | CAT                               | 5.978  | 0,059 | 5,94 %  | 81.334  | 80,87 %  |
| E10 | Ginásio Poliesportivo             | 5.978  | 0,059 | 5,94 %  | 87.312  | 86,82 %  |
| E11 | Incubadora                        | 5.124  | 0,051 | 5,09 %  | 92.436  | 91,91 %  |
| E12 | Guarda Universitária              | 3.096  | 0,031 | 3,08 %  | 95.532  | 94,99 %  |
| E13 | Transportes                       | 2.880  | 0,029 | 2,86 %  | 98.412  | 97,85 %  |
| E14 | Portaria P3                       | 2.158  | 0,022 | 2,15 %  | 100.570 | 100,00 % |

AF = Absolute Frequency; RF = Relative Frequency; CF = Cumulative Frequency.

Matriz de dados ausentes

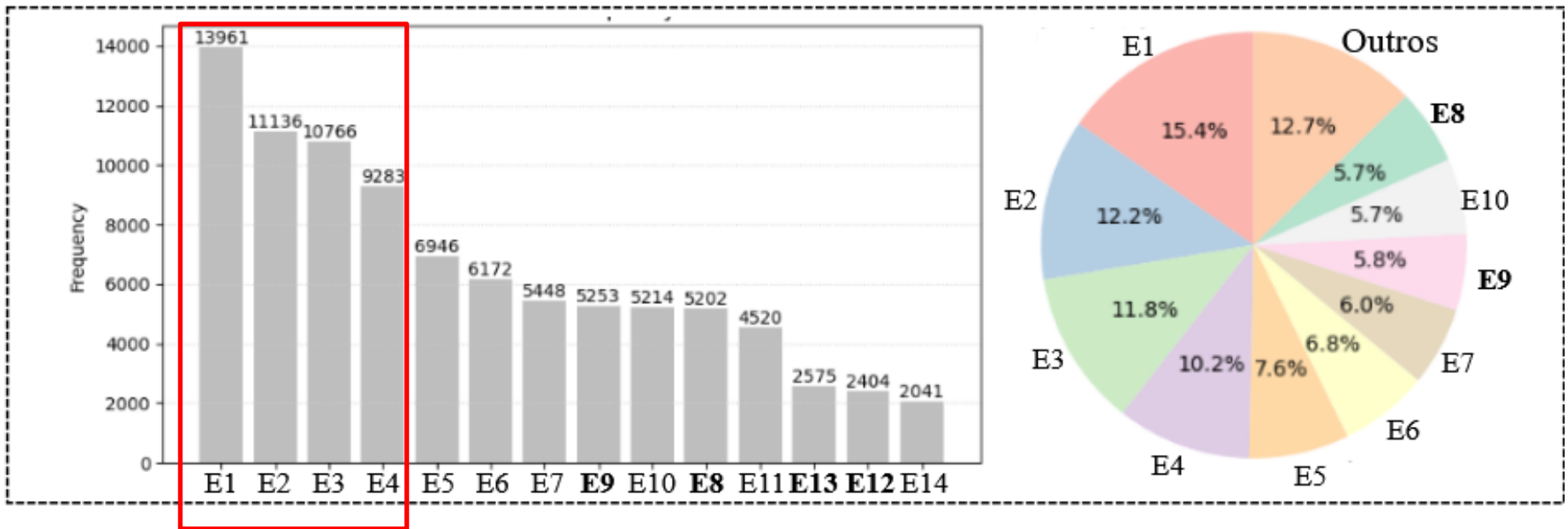


Matriz de dados após limpeza



# Resultados

- Distribuição das frequências de medições na USP-Leste



# Ciclo de análise de dados

4. **Modelagem:** Aplicar modelos estatísticos para explorar relações, padrões e realizar previsões.
5. **Análise de Dados:** Utilizar métodos analíticos e visualizações para interpretar os dados e gerar insights baseados em evidências.
6. **Apresentação:** Comunicar os achados de forma clara e visual, para facilitar o entendimento.
7. **Análise de Resultados:** Interpretar criticamente os resultados, considerando limitações e utilizando ferramentas visuais para apoiar decisões estratégicas.
8. **Refinar o Problema:** Reavaliar as perguntas iniciais com base nos resultados, ajustando o foco da análise e impulsionando ciclos mais eficazes.

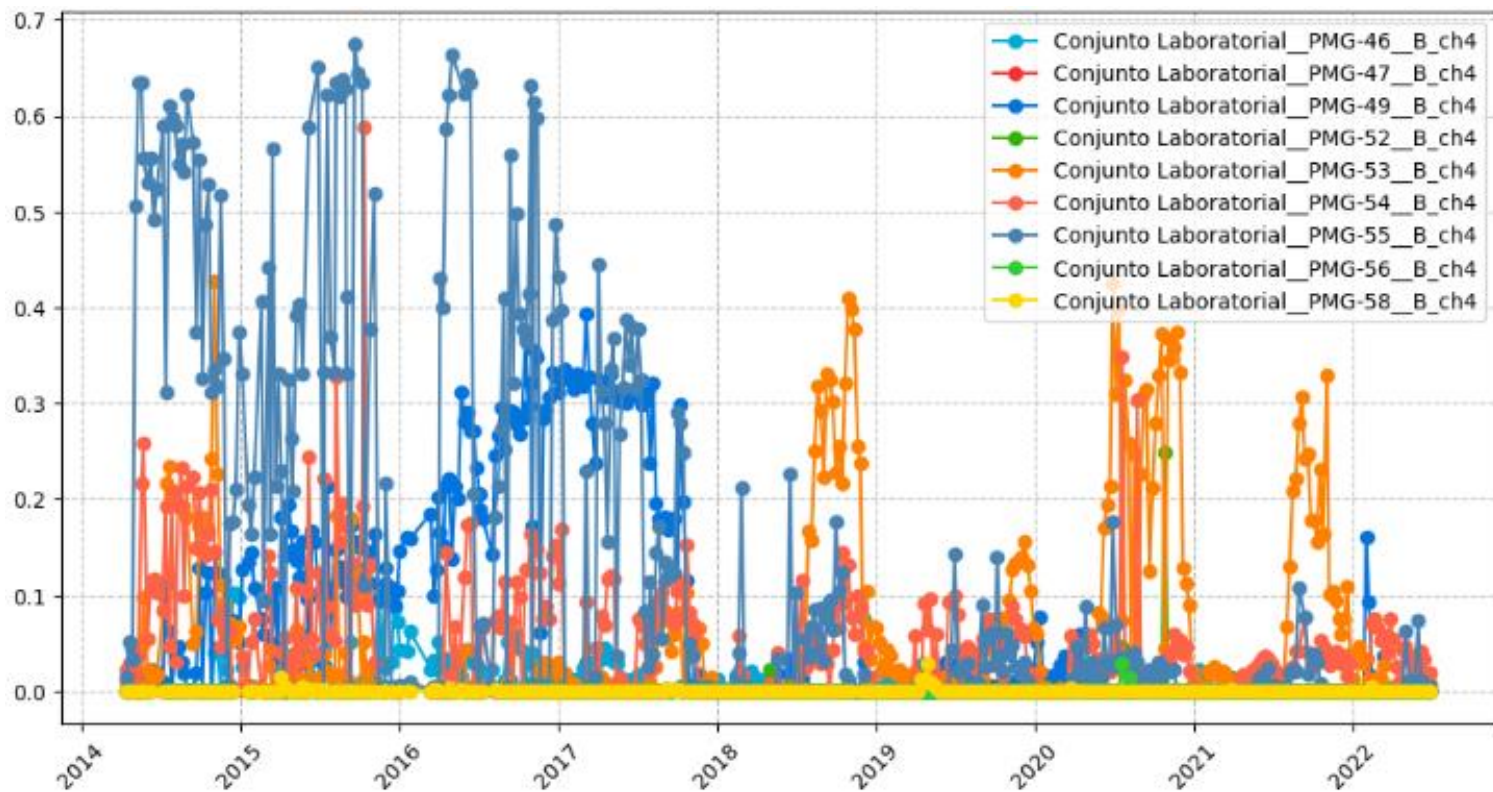
# Resultados e Discussões

- Matriz de Correlação (Pearson) entre variáveis

|         |         |        |        |         |         |          |         |          |         |
|---------|---------|--------|--------|---------|---------|----------|---------|----------|---------|
| ch4     | 1       | 0.28   | -0.42  | -0.0071 | 0.018   | -0.0056  | 0.15    | 0.025    | 0.7     |
| co2     | 0.28    | 1      | -0.79  | -0.033  | 0.025   | 0.013    | 0.02    | 0.035    | 0.34    |
| o2      | -0.42   | -0.79  | 1      | 0.019   | -0.02   | -0.021   | -0.042  | -0.035   | -0.49   |
| h2s     | -0.0071 | -0.033 | 0.019  | 1       | -0.029  | -0.0094  | -0.027  | 0.0052   | -0.0086 |
| co      | 0.018   | 0.025  | -0.02  | -0.029  | 1       | -0.0065  | 0.0092  | -0.0023  | 0.027   |
| bal     | -0.0056 | 0.013  | -0.021 | -0.0094 | -0.0065 | 1        | -0.0028 | -0.00092 | 0.0073  |
| pressao | 0.15    | 0.02   | -0.042 | -0.027  | 0.0092  | -0.0028  | 1       | 0.0016   | 0.08    |
| voc     | 0.025   | 0.035  | -0.035 | 0.0052  | -0.0023 | -0.00092 | 0.0016  | 1        | -0.019  |
| lel     | 0.7     | 0.34   | -0.49  | -0.0086 | 0.027   | 0.0073   | 0.08    | -0.019   | 1       |
|         | ch4     | co2    | o2     | h2s     | co      | bal      | pressao | voc      | lel     |

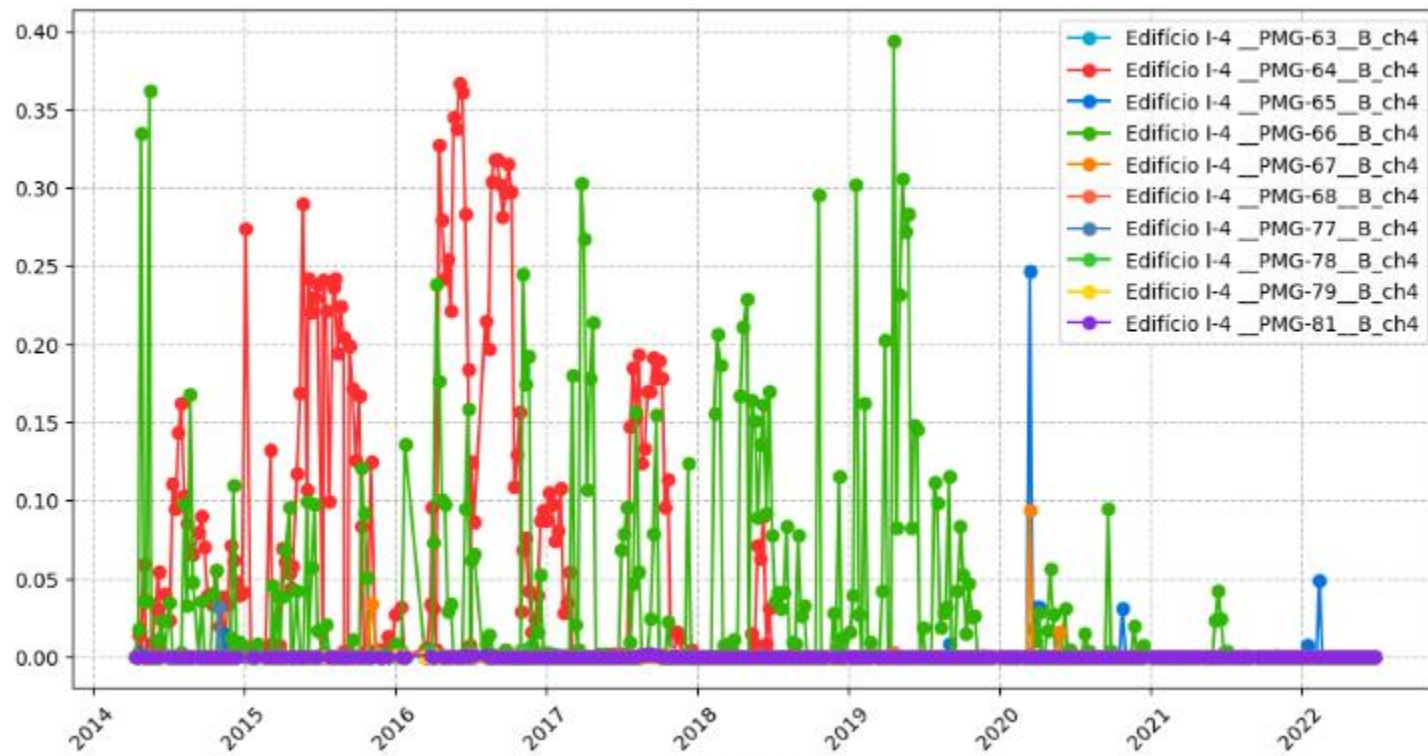
# Resultados e Discussões

- Evolução temporal da concentração de metano (Conjunto Laboratorial – E1)



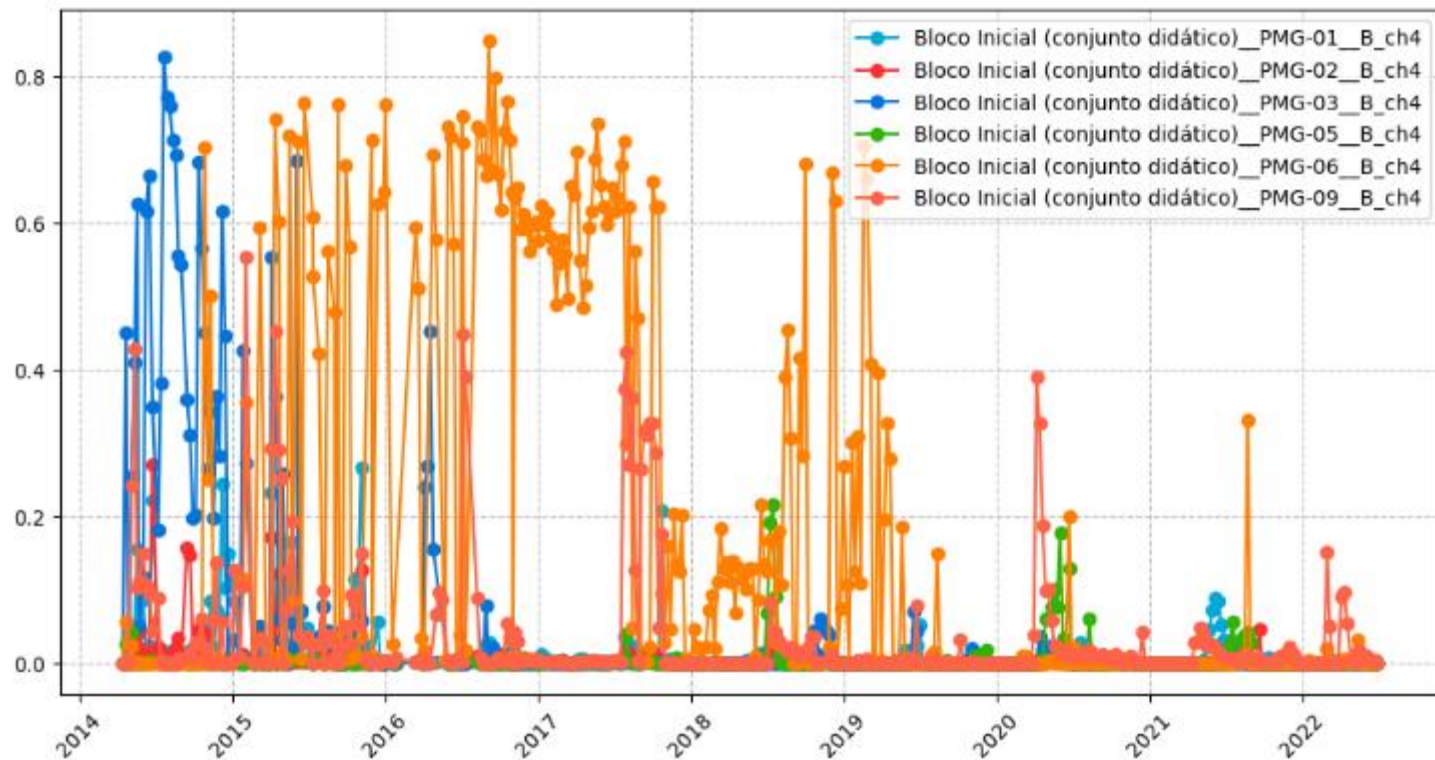
# Resultados e Discussões

## ■ Evolução temporal da concentração de metano (Edifício I-4 – E4)



# Resultados e Discussões

- Evolução temporal da concentração de metano (Conjunto Didático – E2)



# Conclusões

- Ciclo de análise de dados mostrou-se fundamental para a realização do experimento com um grande volume de dados de ACs;
- A identificação de correlações entre as variáveis coletadas fornece subsídios para o refinamento do Modelo Conceitual da área (MCA);
- A recorrência de medições com  $\text{CH}_4 > 5\%$  reforça a necessidade da continuidade dos monitoramentos periódicos e da operação dos sistemas de ventilação para mitigar potenciais cenários riscos no campus;
- Subsídios para otimizar o monitoramento ambiental e aperfeiçoar a gestão de riscos;
- Subsídios para novas pesquisas, correlacionando variáveis ambientais, com dados atmosféricos e dados geológicos do site.



<http://www5.each.usp.br/>



GOVERNO DO ESTADO DE SÃO PAULO  
SECRETARIA DE ESTADO DO MEIO AMBIENTE  
COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

## LICENÇA AMBIENTAL DE OPERAÇÃO RENOVAÇÃO

A Companhia Ambiental do Estado de São Paulo - CETESB, no uso das atribuições conferidas pela Lei Federal 6938, de 31 de agosto de 1981, que dispõe sobre a Política Nacional do Meio Ambiente, regulamentada pelo Decreto Federal 99.274, de 06 de junho de 1991, e pela Lei 13.542, de 08 de maio de 2009, e demais normas pertinentes, emite a presente Licença Ambiental de Operação, com base no Parecer Técnico 10/18/IE, para:

### IDENTIFICAÇÃO DO EMPREENDEDOR

RAZÃO SOCIAL: Universidade de São Paulo - USP  
CNPJ: 63.025.530/0001-04  
LOGRADOURO: Rua da Reitoria, 374  
BAIRRO: Butantã  
MUNICÍPIO: São Paulo

### IDENTIFICAÇÃO DO EMPREENDIMENTO

NOME: Campus Área Capital-Leste (USP Leste)  
LOGRADOURO: Rua Arlindo Bettio, 1000, Bairro Vila Guaraciba  
MUNICÍPIO: São Paulo

### CARACTERIZAÇÃO DO EMPREENDIMENTO

Campus Universitário em terreno de 1.240.578 m<sup>2</sup>, dos quais 211.136,96 m<sup>2</sup> são utilizados atualmente. As edificações ocupam 25.705,72 m<sup>2</sup> e somam 44.611,30 m<sup>2</sup> de área construída, enquanto 185.670,4 m<sup>2</sup>, incluindo áreas sem pavimentação, calçadas, jardins, gramados, áreas de vias, etc.

### OBSERVAÇÕES

- A presente Licença Ambiental de Operação deverá permanecer no local do empreendimento, sob pena de nulidade condicionada ao cumprimento das exigências relacionadas neste documento.
- A presente Licença Ambiental de Operação não dispensa nem substitui quaisquer outras autorizações ou certidões de qualquer natureza, exigidas pela legislação federal, estadual e municipal, bem como não significa reconhecimento de qualquer direito de propriedade.
- Integra(m) a presente Licença 1 anexo(s).
- O prazo de validade desta Licença Ambiental de Operação é de 7 (SETE) ano(s), a contar da data de emissão.
- A renovação da Licença Ambiental de Operação deverá ser requerida com antecedência de 90 (noventa) dias da data de vencimento de seu prazo de validade.

O presente documento foi emitido sem rasura e/ou colagem

Data: 15/10/11

*Geraldo do Amaral Filho*  
GERALDO DO AMARAL FILHO (Diretor de Avaliação de Impacto Ambiental)

CONTROLE Nº 1679729

# Agradecimentos



# Obrigado!

**Leandro Gomes de Freitas – IPT**

**lfreitas@ipt.br**

