

# **Qualidade do Ar e Poluição da Água**

Uma análise técnica de perfis ambientais globais através de mineração de dados

Gabriel Vinicios Nanetti e Nathan Scremen



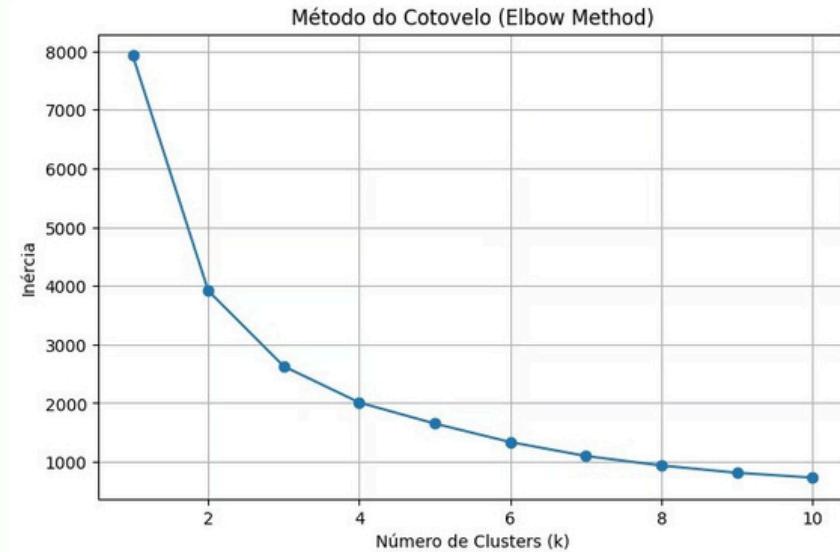
# Metodologia: Padronização e Definição de Cluster

Nossa abordagem utilizou técnicas matemáticas rigorosas para identificar padrões significativos nos dados ambientais globais, garantindo robustez científica na análise.

## Padronização

head() com padronização dos valores.

	City	AirQuality_Padronizado	WaterPollution_Padronizado
0	New York City	-0.498933	0.189768
1	Washington, D.C.	0.125258	0.174266
2	San Francisco	-0.056218	-0.063731
3	Berlin	0.003577	-0.624405
4	Los Angeles	-0.828414	0.649401



## Método do Cotovelo

Determinação matemática de K3 clusters ótimos

### Padronização Z-Score

Transformação dos dados para escala comum (Média=0, Desvio=1), permitindo comparação justa entre variáveis

### Algoritmo K-Means

Técnica de agrupamento não supervisionado para identificação de padrões naturais nos dados

### Validação Estatística

Ponto de inflexão identificado confirma a existência de três perfis ambientais distintos

# Segmentação Global: Os 3 Perfil Ambiental

A análise revelou três perfis ambientais distintos com características próprias de qualidade do ar e da água.



## Sustentabilidade Alta - 1156 cidades (29.2%)

**Cluster 0 (Roxo):** Ar Limpo + Água Limpa. Encontrado em cidades organizadas como Suíça e centros urbanos menores com infraestrutura exemplar.



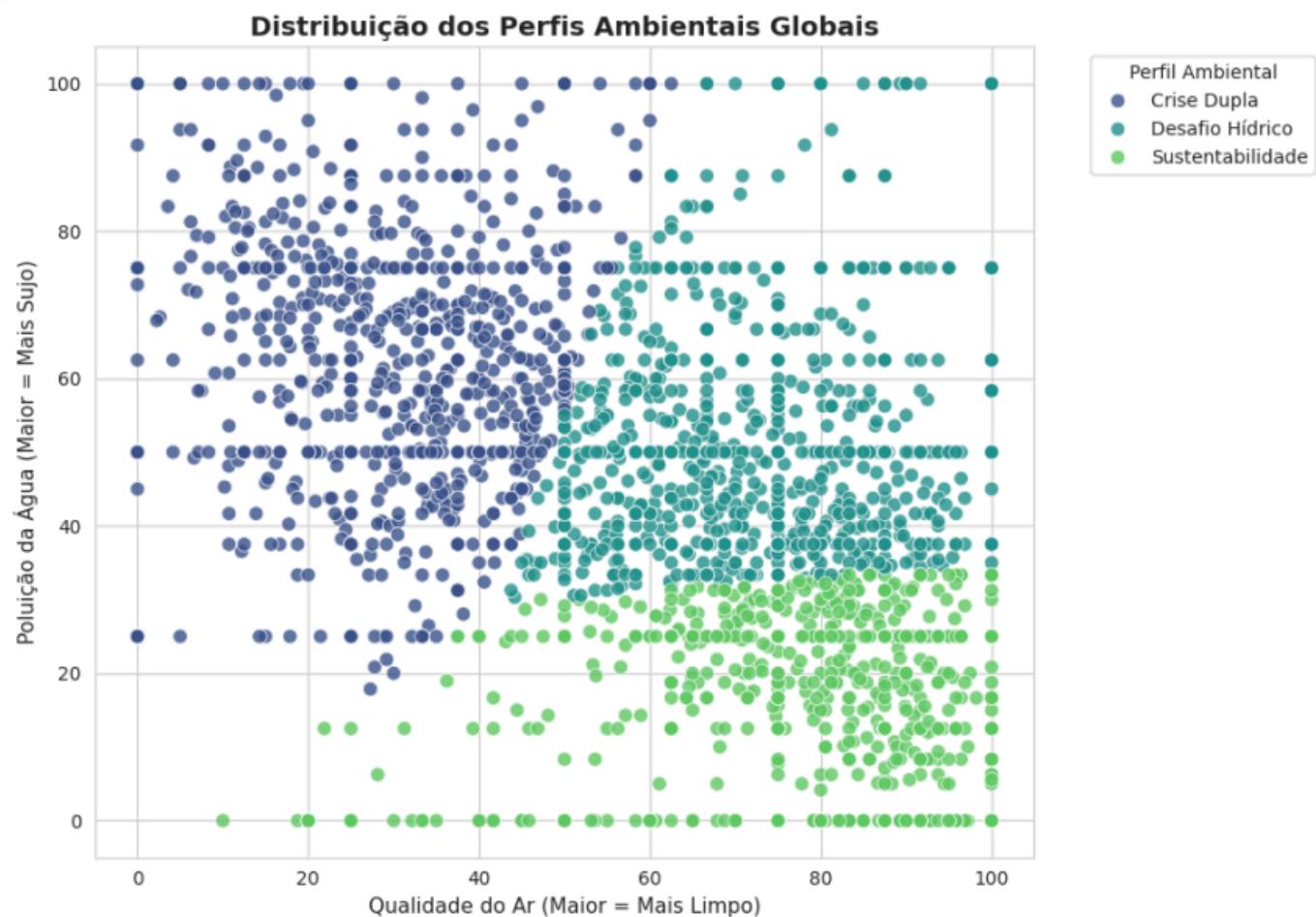
## Desafio Hídrico - 1565 cidades (39.5%)

**Cluster 2 (Amarelo):** Ar Bom + Água Ruim. Típico de cidades costeiras com boa dispersão atmosférica mas saneamento deficiente.



## Crise Dupla - 1242 cidades (31.3%)

**Cluster 1 (Ciano):** Ar Ruim + Água Ruim. Megalópoles e polos industriais enfrentando degradação ambiental em múltiplas frentes.



# Característica do Grupo

Cada perfil ambiental apresenta características únicas que demandam estratégias específicas de intervenção.



## Sustentabilidade Alta

O benchmark de excelência ambiental. Cidades que conseguem equilibrar desenvolvimento urbano com gestão eficiente de saneamento e controle de emissões atmosféricas.



## Desafio Hídrico

O grupo mais intrigante da análise. Demonstra que **ar limpo não garante água limpa**. Comum em locais com boa ventilação natural mas infraestrutura de esgoto precária.



## Crise Dupla

Falha sistemática nas duas métricas ambientais. Representa o alvo prioritário para implementação dos Objetivos de Desenvolvimento Sustentável 6 e 11.

# Validação Estatística: Prova de Robustez

Para garantir a solidade da nossa segmentação, empregamos métodos de validação estatística rigorosos, confirmando que os clusters identificados são significativamente distintos e refletem realidades ambientais específicas.

## → Análise com Boxplot

Geramos Boxplots para visualizar a distribuição dos dados e confirmar a separação clara entre os clusters em relação às métricas de qualidade do ar e da água.

## → Ar: Sustentabilidade Alta

Os Boxplots revelaram que o grupo de "Sustentabilidade Alta" possui uma qualidade do ar drasticamente superior, com baixa variabilidade, demonstrando seu desempenho excepcional.

## → Água: Crise Dupla

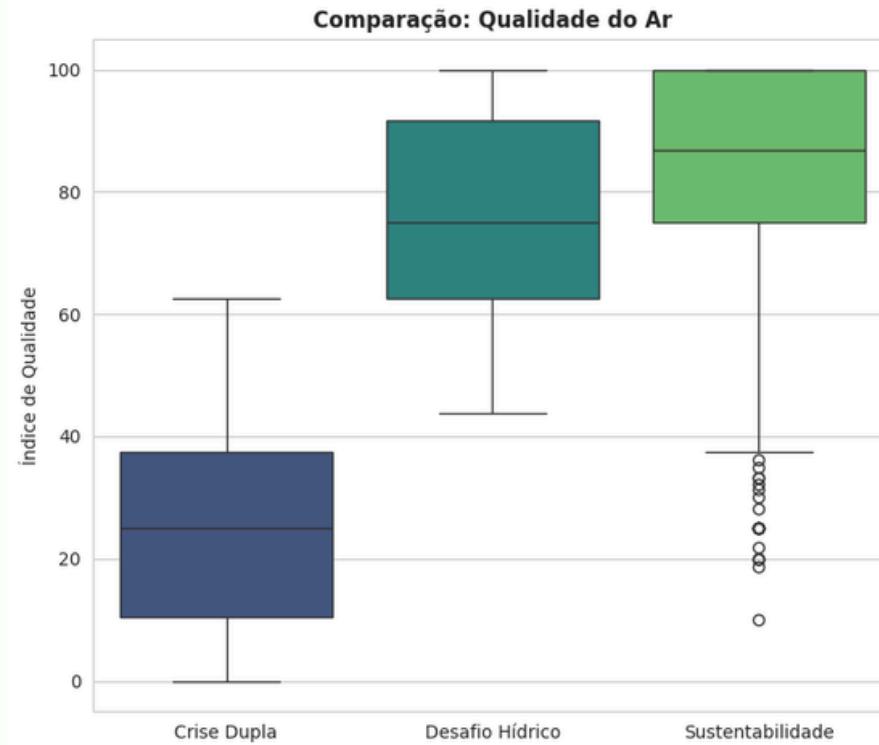
Para o grupo de "Crise Dupla", a análise mostrou níveis consistentemente altos de poluição da água, reforçando a gravidade de seu desafio ambiental.

## → Grupo Distintos e Reais

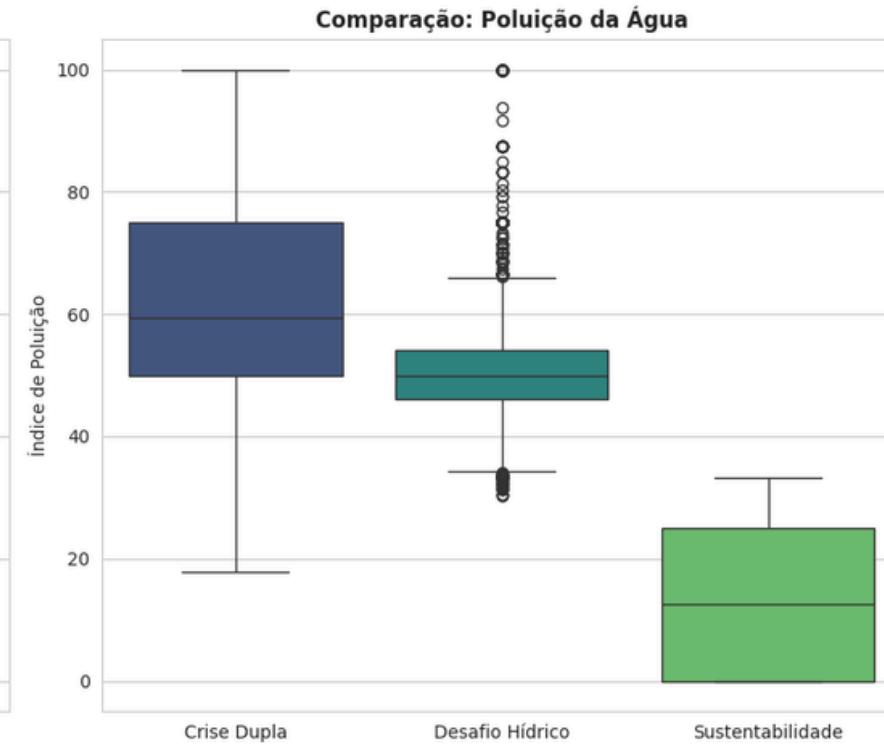
A clareza dos resultados dos Boxplots valida matematicamente a distinção entre os três perfis, atestando que os agrupamentos são robustos e baseados em dados reais.

# Diferenças de Comportamento entre Grupos

Baixa variação = Excelência Consistente



Mediana alta = Poluição Crônica



# O Mito do "Primeiro Mundo Limpo"

A análise quebra paradigmas sobre a relação entre desenvolvimento econômico e qualidade ambiental.

## Megalópole e Crise

Nova York, Londres e Los Angeles estão classificadas no **Cluster da Crise Dupla**, desafiando a percepção comum.



## O Modelo Suíço

O grupo de "Sustentabilidade Alta" é dominado por cidades médias e organizadas como Bern e Zurique.



### Conclusão Chave

Urbanização extrema e densidade populacional elevada geram degradação ambiental universal, independentemente da riqueza econômica do país. O tamanho e a gestão urbana importam mais que o PIB per capita.



# Brasil: Reflexo da Realidade Local

## Visão Ambiental: Realidade Brasileira

O cenário brasileiro reflete os padrões globais identificados, com particularidades regionais significativas.

City	Perfil	AirQuality	WaterPollution
Sao Paulo	Crise Dupla (Ar Ruim / Água Ruim)	24.1	73.7
Sao Luis	Crise Dupla (Ar Ruim / Água Ruim)	50.0	100.0
Manaus	Crise Dupla (Ar Ruim / Água Ruim)	31.2	46.4
Brasilia	Desafio Hídrico (Ar Bom / Água Poluída)	76.5	38.0
Curitiba	Desafio Hídrico (Ar Bom / Água Poluída)	73.1	48.9
Salvador	Desafio Hídrico (Ar Bom / Água Poluída)	83.0	51.3
Porto Alegre	Desafio Hídrico (Ar Bom / Água Poluída)	54.7	59.3
Osasco	Crise Dupla (Ar Ruim / Água Ruim)	35.0	66.7
Rio de Janeiro	Crise Dupla (Ar Ruim / Água Ruim)	46.8	77.2
Olinda	Crise Dupla (Ar Ruim / Água Ruim)	50.0	75.0
Belo Horizonte	Desafio Hídrico (Ar Bom / Água Poluída)	51.2	47.5
Fortaleza	Desafio Hídrico (Ar Bom / Água Poluída)	66.7	56.9

1

### Manaus

**Crise Dupla:** Polo industrial amazônico enfrenta poluição atmosférica de queimadas e desafios de saneamento básico.

2

### São Luís

**Crise Dupla:** Atividades portuárias e industriais comprometem simultaneamente qualidade do ar e dos recursos hídricos.

3

### Salvador

**Desafio Hídrico:** Boa dispersão atmosférica costeira, mas problemas críticos de balneabilidade e saneamento.

4

### Porto Alegre

**Desafio Hídrico:** Qualidade do ar relativamente preservada, porém infraestrutura hídrica demanda investimentos urgentes.

# Regras de Associação (Apriori)

Nesta fase, aprofundamos nossa análise. Saímos do agrupamento para buscar as relações de "causa e efeito" entre os fatores ambientais, revelando conexões que um simples agrupamento não poderia.

## Metodologia Detalhada:

1. **Categorização de Dados:** Primeiro, transformamos todos os valores numéricos de qualidade do ar e da água em 5 categorias distintas: Muito Baixo, Baixo, Médio, Alto e Muito Alto. Essa etapa é fundamental para o algoritmo de associação.

2. **Algoritmo Apriori:** Em seguida, rodamos o algoritmo Apriori, uma técnica robusta de mineração de dados, para identificar padrões frequentes e regras de associação entre as categorias. Isso nos permite ver, por exemplo, se "qualidade do ar muito baixa" frequentemente ocorre com "qualidade da água muito baixa".

	SE (Causa)	ENTAO (Efeito)	confidence
9	Água: Muito Baixo	Ar: Muito Alto	60.9%
3	Ar: Médio	Água: Médio	43.5%
0	Ar: Baixo	Água: Médio	41.6%
11	Água: Baixo	Ar: Muito Alto	38.5%
6	Água: Baixo	Ar: Alto	36.6%
7	Ar: Alto	Água: Médio	36.1%
12	Ar: Muito Alto	Água: Médio	33.9%
2	Ar: Baixo	Água: Alto	33.0%
5	Ar: Alto	Água: Baixo	31.4%
1	Água: Alto	Ar: Baixo	30.3%

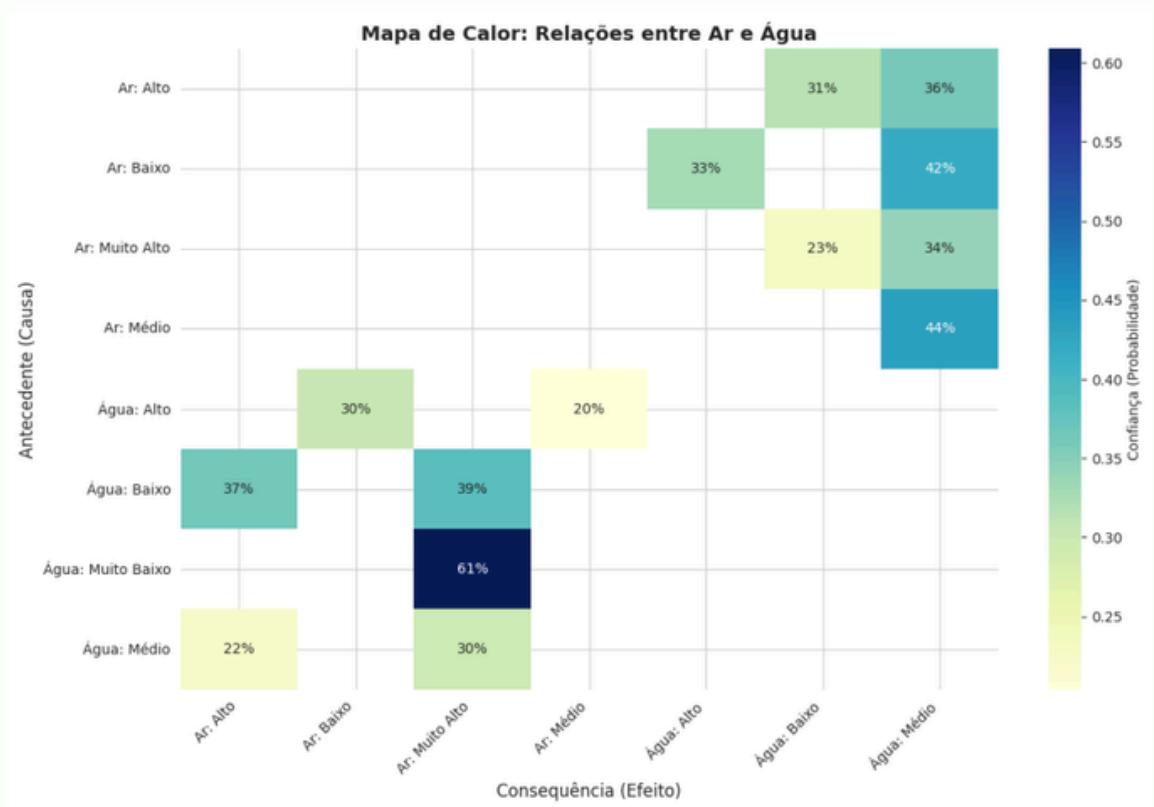
# Mapa de Calor das Regras de Associação

## Como Ler este Gráfico?

- 1. O Eixo Vertical (Causa):** Representa a condição inicial (ex: "Água Muito Baixa").
- 2. O Eixo Horizontal (Efeito):** Representa a consequência provável no Ar.
- 3. A Cor (Confiança):** Quanto mais Azul Escuro, maior a certeza estatística da regra.

## Regra Descoberta:

**Principal Descoberta:** As áreas em azul escuro mostram as conexões mais fortes. O destaque é a regra com **61% de confiança**, que prova estatisticamente: cidades que resolvem a poluição da água (nível 'Muito Baixo') tendem a ter, consequentemente, uma qualidade do ar excelente ('Muito Alta').



# Principais Descobertas: Regras de Associação

O algoritmo Apriori revelou correlações importantes entre a qualidade do ar e da água, destacando padrões na gestão ambiental.

## Excelência Ambiental Integrada

A regra mais forte (Confiança de 60%) indica que cidades com **Poluição da Água "Muito Baixa"** tendem a ter Qualidade do Ar **"Muito Alta"**. Isso sugere que a excelência ambiental é um fenômeno integrado e holístico.

## Sinergia a Sustentabilidade

Este achado reforça que é raro resolver problemas de poluição da água sem que haja uma gestão eficaz da qualidade do ar, e vice-versa. Ambientes verdadeiramente limpos demonstram um esforço coordenado em múltiplas frentes.

## A Inércia da Poluição

Observamos que cidades com ar **"Médio" ou "Ruim"** frequentemente apresentam poluição da água em níveis **similares**. Isso revela uma "inércia" na degradação ambiental, onde um problema tende a arrastar o outro, criando um ciclo vicioso difícil de quebrar.

Essas regras de associação fornecem uma base para estratégias mais direcionadas, enfatizando a interconexão dos desafios ambientais.

# A Explicação Científica: Deposição Atmosférica

## → O Ciclo da Poluição

Poluentes atmosféricos, provenientes de emissões industriais e veiculares, não desaparecem. Eles são reincorporados ao ciclo da água, precipitando-se de volta à superfície.

## → Dois Mecanismo Principais

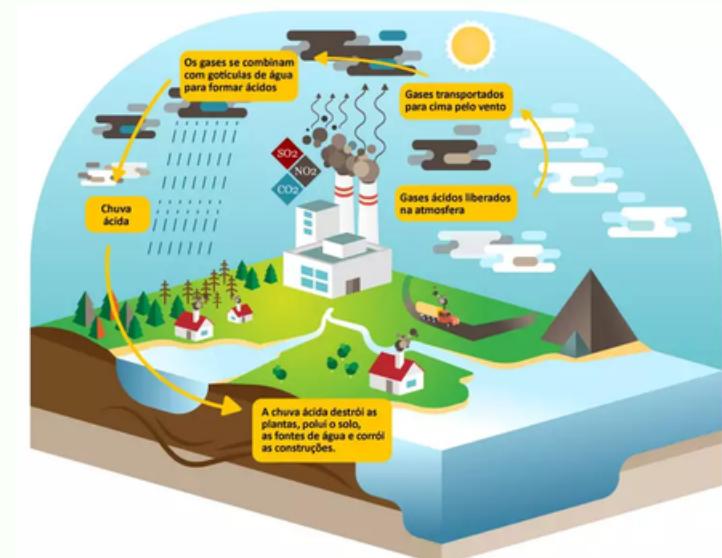
- **Deposição Seca:** Partículas de poeira e gases tóxicos se depositam diretamente em rios, solos e superfícies aquáticas.
- **Deposição Úmida:** Chuva ácida, neve ou neblina "lavam" a atmosfera, transportando poluentes dissolvidos para corpos d'água.

## → Poluição Difusão e Escoamento Urbano

Em áreas urbanas com alta impermeabilização do solo, a sujeira acumulada no ar e nas superfícies (ruas, telhados) é carregada diretamente para rios e mananciais durante eventos de chuva, configurando um processo de poluição difusa.

## □ A Prova do 60%

Nossa regra de associação (Apriori) confirmou estatisticamente este fenômeno físico. Onde o ar é sujo, a chuva transporta essa poluição para a água, validando a forte correlação encontrada entre os dois indicadores.



# Estudo de Caso: Shenzhen, China

Para evidenciar o impacto direto da deposição atmosférica, pesquisadores realizaram um experimento revelador em uma das maiores metrópoles da China.

01

## O Experimento

Água de escoamento foi coletada de telhados isolados em Shenzhen, longe de qualquer fonte direta de lixo humano ou esgoto, para testar a pureza da chuva.

02

## A Descoberta

Mesmo em superfícies sem contato humano direto, a água coletada apresentava altos níveis de **Nitratos** e **Carbono Orgânico**, indicando contaminação inesperada.

03

## O Veredito

A conclusão foi clara: a poluição não veio do chão, mas [diretamente do ar](#). A chuva "lavou" as emissões de veículos e fábricas, transferindo toxinas atmosféricas para a água por meio da deposição úmida.



# Conclusões e Implicações Globais

As análises realizadas revelaram insights cruciais, desmistificando percepções comuns e validando abordagens estratégicas para desafios ambientais globais.



## Quebra de Mitos Urbanos

Megacidades ricas como Nova York e Londres enfrentam "Crise Dupla" de poluição, alinhando-se a cidades em desenvolvimento. A densidade urbana, e não o PIB, emerge como o fator crítico para a qualidade ambiental.



## Validação da ODS

A forte associação estatística entre a qualidade do ar e da água reforça a necessidade de políticas públicas integradas, essenciais para o avanço dos Objetivos de Desenvolvimento Sustentável 6 (Água Limpa e Saneamento) e 11 (Cidades e Comunidades Sustentáveis).



## Impacto da Inteligência Artificial

O uso de algoritmos como K-Means e Apriori foi fundamental para diagnosticar perfis ambientais globais complexos, revelando padrões e interconexões que seriam imperceptíveis através de análises tradicionais.

Estes resultados apontam para a necessidade de uma visão holística e baseada em dados para a gestão ambiental, especialmente em contextos urbanos.

# Referências Bibliográficas

A base científica que sustenta nossa análise e metodologia.

**Kavipriya & Jothilakshmi (2020)** "K-Means Algorithm to Evaluate Air and Water Pollution in Tamilnadu." 2020 4th International Conference on Intelligent Computing and Control Systems. DOI: 10.1109/ICICCS48265.2020.9297493

**Andia et al. (2025)** "Cluster Based Classification of River Water Pollution Using K-Means for Policy Intervention and Environmental Justice in Central Java, Indonesia." Journal of Law, Environmental and Justice, 3(2), 5063541.

**Mohamad & Usman (2013)** "Standardization and Its Effects on K-Means Clustering Algorithm." Research Journal of Applied Sciences, Engineering and Technology, 6(17), 3299-3303.

# Obrigado!

Perguntas?

Esperamos que esta análise inspire ações e novas perspectivas para a qualidade ambiental global.

Gabriel Vinicios Nanetti & Nathan Scremen