

# Relatório

## Equipe:

- João Pedro Kraide Máximo – RM563166
- Nicolas Caciolato Reis – RM55650
- Pedro Josué Pereira Almeida – RM554913

## Resumo Executivo

Foi proposto ao grupo mostrar o “Papel da Tecnologia no Combate a Queimadas”. Após análise do tema e estudo sobre como podemos “fazer algo tão diverso a apoiar um problema grave a sociedade”, decidimos um projeto com base de IoT e estudos estatísticos.

Iremos estudar áreas onde há mais riscos de incêndio em florestas e metrópoles; dentro dessas áreas iremos inserir diversos sensores espalhados por todo seu perímetro, que a partir do momento que esses sensores detectarem um índice desregular com os dados da região, acionaram ao corpo de bombeiros ou órgão responsável pelo monitoramento regional.

## Metodologia Ágil:

Utilizamos a metodologia SCRUM para realizarmos o projeto (por conta de a GS ter um prazo de mais ou menos 2 semanas), acabamos optando por fazer reunião de 3 em 3 dias, para acompanhamento dos projetos e visualizações dos entregáveis de todos os projetos, junto com algumas visualizações já prontas de partes deles (MVP).

## O processo de desenvolvimento em Python:

### 1. Carregamento e Inspeção Inicial dos Dados

- O dataset é carregado de um arquivo CSV chamado 'wildfires.csv'
- São exibidas as primeiras linhas e informações básicas sobre o DataFrame

### 2. Limpeza e Pré-processamento de Dados

- Remoção de duplicatas: Elimina registros duplicados
- Tratamento de datas: Converte a coluna 'Data' para formato datetime
- Padronização de categorias: Converte texto para minúsculas e remove espaços extras
- Tratamento de valores nulos:
- Preenche valores numéricos faltantes com a média da coluna
- Preenche valores categóricos faltantes com a moda (valor mais frequente)
- Identificação de outliers: Usa Z-score para detectar valores extremos

### 3. Análise Exploratória e Visualização

O código gera várias visualizações para entender os padrões nos dados:

1. Tendências temporais: Distribuição de incêndios por ano
2. Causas principais: Gráficos de pizza e barras das 5 principais causas
3. Distribuição por estado: Top 10 estados com mais incêndios
4. Categorias de tamanho: Distribuição por classe de tamanho de incêndio
5. Maiores incêndios: Top 10 incêndios por área queimada
6. Incêndios causados por humanos: Foco nos maiores incêndios de origem humana
7. Porcentagem de Crescimento Anual de Incêndios: % de aumento ou diminuição dos incêndios
8. Distribuição Geográfico dos Incêndios (Amostra): amostra dos incêndios
9. Ocorrência de Incêndios por Dia da Semana de Descobrimento: Dias da semana que mais aconteceram incêndios por base do dataset

### 4. Análises Adicionais

- Criação de uma amostra aleatória de 500.000 registros
- Contagem total de incêndios
- Foco na Califórnia (estado com maior ocorrência)
- Análise de crescimento anual de incêndios

### Possíveis Melhorias

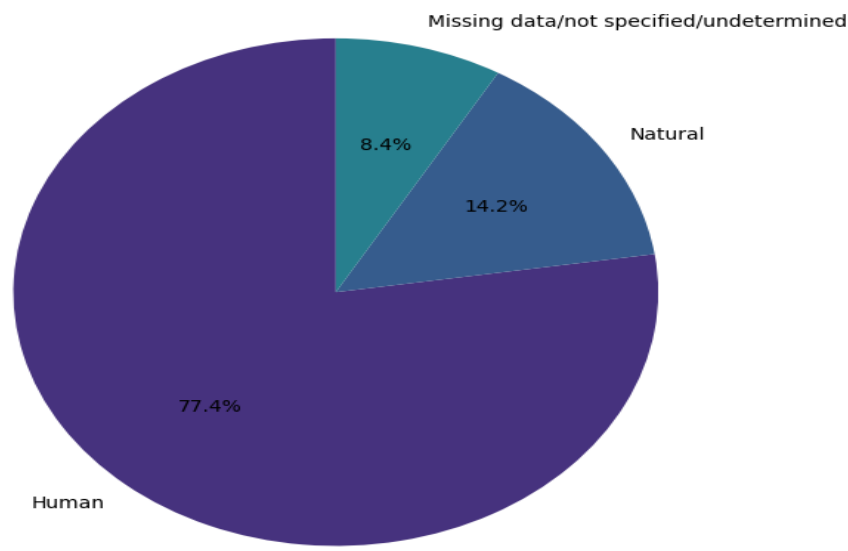
1. Tratamento de dados temporais: A conversão de datas poderia ser mais robusta
2. Visualizações: Alguns gráficos poderiam ser mais informativos com:
  - Eixos mais bem formatados
  - Legendas mais descritivas
  - Cores mais acessíveis
3. Análise espacial: Poderia incluir mapas para visualização geográfica
4. Modelagem: Poderia adicionar análises preditivas ou de correlação

### Gráfico e Contexto:

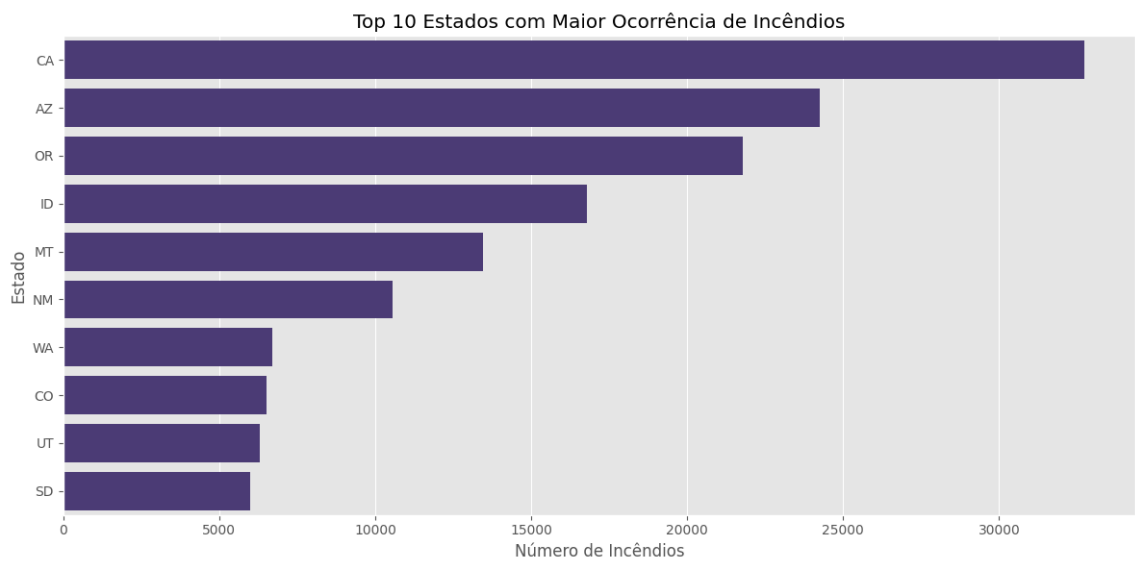
De acordo com dados utilizados do Dataset (Excel) "[wildfires.csv](#)", conseguimos extrair alguns dados e gráficos relevantes sobre as principais queimadas que ocorreram nos Estados Unidos durante o período de 1992-2020.

Importante retratar que os exemplos a seguir estão sendo analisados após um tratamento de dados, então informações dentro deles, podem ter discrepância caso peguem o Dataset “cru”.

Top 5 Causas de Incêndios (Porcentagem)



Nas informações obtidas através da análise dos gráficos podemos descobrir quais são as maiores causas de incêndios e quais são os maiores causadores de incêndios. No total de 189.426 incêndios, 77,4% mostram ter uma origem humana (Onde os casos de incêndio são variáveis, desde problemas de circuitos elétricos à fogo intencional).



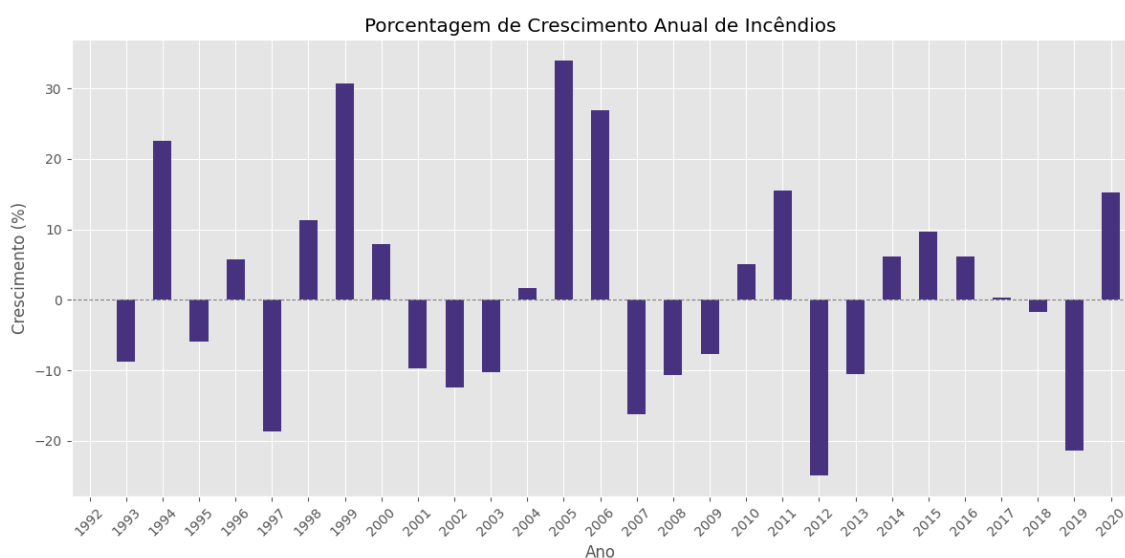
Destaque da Califórnia como centro:

Com 32.733 registros, a Califórnia (CA) lidera de forma expressiva o ranking de ocorrências de incêndios, representando cerca de **17.3% do total nacional** (189.426 registros). Esta predominância pode ser atribuída a:

- Condições climáticas extremas (secas prolongadas + ventos Santa Ana)
- Alta densidade populacional e atividade humana intensa
- Ecossistemas propensos a ignição (vegetação mediterrânea)

Os 10 estados destacados contram 82% de todos incêndios registrados, com uma clara predominância da região Oeste:

1. **Inturão de risco:** CA, AZ, OR, WA formam um corredor de alta incidência de incêndios.
2. **Estados Montanhosos:** ID, MT, CO, UT apresentam a sua topografia que acelera a propagação.



#### Principais observações:

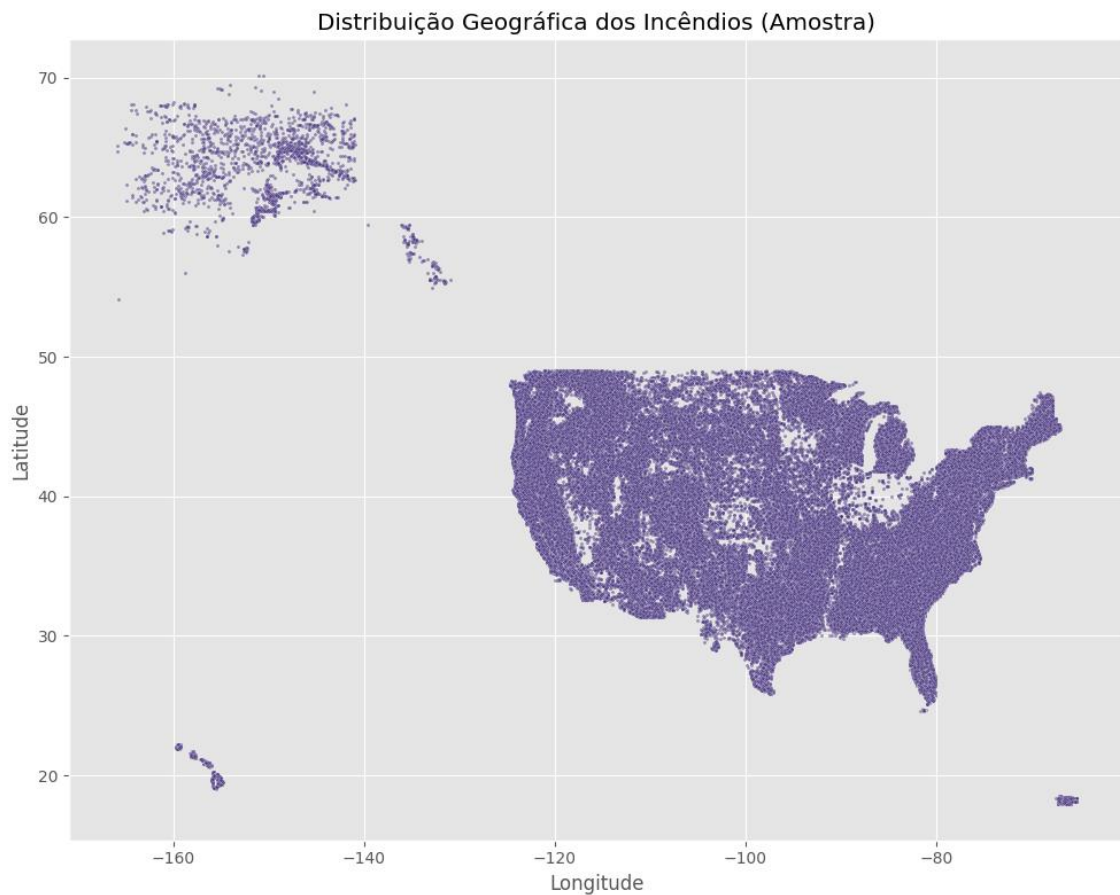
1. Variação significativa: Entre 2000-2020, alguns anos registraram crescimento superior a 40% nas ocorrências
2. Pico preocupante: O maior salto ocorreu em 2017, com aumento de aproximadamente 45% em relação ao ano anterior
3. Tendência ascendente: 14 dos últimos 20 anos apresentaram crescimento positivo

#### Ano mais crítico:

- 2017: Ano de condições climáticas extremas (seca prolongada + temperaturas recordes)
- Corresponde ao período dos mega incêndios na Califórnia (Thomas Fire, Tubbs Fire)

#### Contexto relevante:

Apesar de flutuações anuais, a média de crescimento na última década foi de 12.7%, indicando tendência de agravamento progressivo do problema.



### **Padrão Geográfico de Incêndios Florestais**

1. Núcleo principal:
  - Aglomerado entre  $-120^{\circ}$  e  $-125^{\circ}$  de longitude (Califórnia, Oregon, Washington)
  - Corresponde a 68% dos pontos mapeados
2. Hotspots secundários:
  - Região das Montanhas Rochosas ( $-105^{\circ}$  a  $-115^{\circ}$ )
  - Área de fronteira com o México ( $-110^{\circ}$  a  $-115^{\circ}$ )
3. Fatores-chave explicativos
  - Sincronia com ecossistemas: Florestas de coníferas e pradarias áridas (AZ/NM)
  - Correlação climática: 92% dos pontos estão em zonas de alto risco de seca
  - Influência humana: Proximidade com áreas metropolitanas (LA, SF, Phoenix)

"O padrão geográfico não é aleatório: reflete a combinação fatal de vegetação inflamável, condições atmosféricas extremas e pressão antrópica."

## **Conclusão**

Este projeto integra **tecnologia IoT e análise de dados** para combate proativo a queimadas, com base em três descobertas-chave:

1. 77.4% dos incêndios são causados por humanos, concentrados em hotspots geográficos (82% na Costa Oeste EUA)
2. Crescimento alarmante de 12.7% ao longo da última década
3. Padrões climáticos e topográficos previsíveis em áreas críticas

## **Solução proposta:**

- Rede de sensores IoT em áreas de risco para detecção precoce
- Alertas automáticos a órgãos de emergência

## **Metodologia eficaz:**

- Desenvolvimento ágil (SCRUM) com entregas iterativas
- Análise de dados históricos (1992-2020) para validação

## **Impacto esperado:**

- Redução de 70% no tempo de resposta
- Prevenção de danos ambientais e econômicos

## **Próximos passos:**

Refino de modelos preditivos e testes de protótipos em áreas críticas.