**Nome dos alunos do grupo**

**Questão 1 (60 pontos)**

**Tema: Queimadas – Soluções Tecnológicas e Sustentabilidade**

**Enunciado:** Seu time faz parte de uma startup focada em tecnologia ambiental, com o objetivo de desenvolver soluções para monitoramento, prevenção e conscientização sobre queimadas no Brasil.

**Desafio:**

Aplicando as metodologias ágeis (Scrum ou Kanban), desenvolva uma solução mínima viável (MVP) que utilize programação em Python para:

Coletar dados reais ou simulados sobre focos de queimadas (fontes como INPE, IBGE ou datasets públicos).

Gerar, no mínimo, dois gráficos estatísticos ou geográficos (exemplos: mapa de calor, séries temporais, gráficos de barras comparativos) que ajudem na tomada de decisão ou conscientização sobre o problema.

Implementar uma análise exploratória dos dados, destacando padrões, tendências ou insights relevantes.

**Entregáveis:**

Documentação Agile:

Apresente o planejamento ágil usado:

Kanban ou Scrum Board (Sprint Backlog, Product Backlog etc.).

Lista de User Stories (Ex.: "Como cidadão, desejo visualizar os locais com maior número de queimadas para entender os riscos ambientais na minha região.").

Código em Python

Apresentação (Vídeo Pitch de até 3 minutos):

Explique o problema, a solução proposta, o uso de Agile no desenvolvimento e os insights gerados pelos dados e pelos gráficos.

**Relatório Final (PDF):**

Descreva:

O problema ambiental abordado.

A metodologia ágil aplicada.

O processo de desenvolvimento em Python.

Os gráficos e suas interpretações.

Conclusões e possíveis melhorias.

**Questão 2**

**Contexto:** As queimadas podem ser causadas por fatores naturais, como raios, ou por atividades humanas, como queimadas descontroladas, fogueiras mal apagadas ou atos criminosos. Com o aumento das temperaturas globais e da ocorrência de secas prolongadas, a frequência e a intensidade das queimadas têm crescido significativamente, impactando gravemente o meio ambiente e a saúde pública. Esses incêndios destroem vegetação nativa, prejudicam fauna e flora, e emitem grandes quantidades de fumaça tóxica, ameaçando a saúde de comunidades próximas e distantes, dependendo das condições atmosféricas e correntes de ar.

**Objetivos:** Analisar dados históricos de incêndios florestais para identificar padrões relacionados às causas, épocas do ano, localização e tamanho dos incêndios, utilizando tecnologia para sugerir métodos de prevenção e combate mais eficazes.

**Banco de Dados:** Para esta atividade, você deve utilizar o banco de dados ‘wildfires.csv’, composto de ocorrências de incêndios nos Estados Unidos.

**Variáveis:**

* **FOD\_ID:** Identificador único numérico do incêndio.
* **FIRE\_NAME:** Nome dado ao incêndio.
* **FIRE\_YEAR:** Ano em que o incêndio foi descoberto.
* **DISCOVERY\_DATE:** Data completa da descoberta do incêndio.
* **DISCOVERY\_DOY:** Dia do ano em que o incêndio foi descoberto.
* **NWCG\_CAUSE\_CLASSIFICATION:** Classificação ampla da causa do incêndio (Humana, Natural, Não especificada).
* **NWCG\_GENERAL\_CAUSE:** Causa geral do incêndio (ex.: Raio, Queima de resíduos, Incêndio criminoso, etc.).
* **CONT\_DATE:** Data em que o incêndio foi controlado.
* **CONT\_DOY:** Dia do ano em que o incêndio foi controlado.
* **FIRE\_SIZE:** Área afetada pelo incêndio (em acres).
* **FIRE\_SIZE\_CLASS:** Categoria do incêndio conforme o tamanho da área afetada.
* **LATITUDE:** Latitude da localização do incêndio.
* **LONGITUDE:** Longitude da localização do incêndio.
* **STATE:** Código do estado em que o incêndio ocorreu.

**Amostra:** O conjunto de dados contém mais de 2 milhões de registros de medições entre 1992 e 2020. Utilize uma semente aleatória (somatório do último número da matrícula de todos os integrantes do grupo) para selecionar uma amostra de 500 mil observações no dataset (sem reposição), que será utilizada na análise final.

**Entrega:**

* **Valor Total da Entrega (40 pontos)**
  + **Documentação (30 pontos)**
    - Análise, limpeza e formatação dos dados: Remoção de duplicatas, padronização de campos de datas ou categorias, identificação e tratamento de valores nulos e outliers;
    - Análise descritiva: Estatísticas descritivas das variáveis, tendências temporais para a ocorrência de incêndios e suas causas principais e estados onde ocorreram com maior e menor frequência;
    - Gráficos: pelo menos 4 gráficos, focando na distribuição temporal dos incêndios, principais causas, relação entre causas (humanas/naturais) e tamanho do incêndio ou localização geográfica;
    - Organização: Códigos bem estruturados e comentados e separação clara entre etapas de análise.
  + **Apresentação (10 pontos)**
    - Resumo do Projeto: Objetivo, contexto da análise e variáveis da base de dados;
    - Visualizações e Estatísticas-Chave: Comentários claros sobre os gráficos e insights extraídos;
    - Conclusões e Recomendações: Discuta como os resultados da análise dos dados poderiam ser usados por drones e sistemas de fiscalização para prevenir incêndios provocados por desmatamento ou ações humanas ilegais. Sintetize os principais padrões identificados sobre a ocorrência de incêndios florestais no dataset analisado. Apresente recomendações práticas para gestores públicos sobre como prevenir e combater incêndios florestais utilizando análise de dados e tecnologia.

**Entrega:** O arquivo com a documentação deverá ser entregue em formato ‘.ipynb’ (Notebook), contendo a análise completa, código comentado e estruturado. O arquivo da apresentação deverá ser entregue em formato ‘.pdf’ (Portable Document Format), que será utilizado para uma apresentação executiva do projeto. A não entrega de algum destes documentos ocasiona na diminuição da pontuação. A entrega deverá ser feita exclusivamente pelo Teams.

Observação: O arquivo para análise está nesta página do Kaggle, sobre incêndios florestais nos EUA, com mais de 2 milhões de ocorrências.

<https://www.kaggle.com/datasets/behroozsohrabi/us-wildfire-records-6th-edition?select=data.csv>

Foi aplicado um filtro nas principais colunas e o arquivo inicial, que era de 1GB, agora é de 200MB. Este arquivo será fornecido para os alunos pelo Teams.