Rhombus Challenge – First Round

Objective:

Each candidate/team is to:

1. **Ingest a data stream of historical and current fires in California last 20 years (2006-2025)**
2. **Analyze the data** to detect high-risk fire events
3. **Build a basic interactive dashboard** (web or notebook) to display events and summary statistics

Input:

Use the CA Fire API Endpoint(s)

- https://www.fire.ca.gov/api/sitecore/Incident/GetFiresByYear?year={yr}
*where yr is in a range of [2006-2025]

- https://www.fire.ca.gov/api/sitecore/Incident/GetSingleFire?IncidentUrl={url}
*where url is for each fire from the data above (e.g., /incidents/2025/8/9/richbar-fire/)

Example response from the API (1)

```
[
  {
    "Name": "Sierra Fire",
    "Updated": "2006-02-12",
    "Started": "2006-02-06",
    "AdminUnit": "USFS/Orange County Fire",
    "County": "Orange,Riverside",
    "Location": "Corona/Orange",
    "AcresBurned": 10584,
    "PercentContained": 100,
    "Longitude": 0,
    "Latitude": 0,
    "Type": "",
    "UniqueId": "7bc3e973-7c13-4e62-926a-c6d25d4fa426",
    "Url": "/incidents/2006/2/6/sierra-fire/",
    "StartedDateOnly": null,
    "IsActive": false,
    "ExternalUrl": ""
  },
  {
    "Name": "Arrastre Fire",
    "Updated": "2006-06-09",
    "Started": "2006-06-09",
```

Example response from the API (2)

```
{
  "Name": "Richbar Fire",
  "Final": true,
  "Updated": "/Date(1755421199557)/",
  "Started": "/Date(1754761157000)/",
  "AdminUnit": "Sequioa National Forest",
  "AdminUnitUrl": "",
  "County": null,
  "Counties": [
    "Kern"
  ],
  "Location": "Highway 178 and Upper Richbar, Lake Isabella",
  "AcresBurned": 158.3,
  "PercentContained": 100,
  "ControlStatement": "",
  "AgencyNames": "",
  "Longitude": -118.722809,
  "Latitude": 35.468395,
  "MapLongitude": 0,
  "MapLatitude": 0,
  "Type": "Wildfire",
  "UniqueId": "ae18777a-ca05-418d-965c-b34355a324d1",
  "Url": "/incidents/2025/8/9/richbar-fire/",
  "ExternalUrl": "",
  "ExtinguishedDate": "/Date(1755381600000)/",
```

Requirements:

Setup

You must do your own library/ module installs and everything on your own. This is part of the drill too. Even though this is a fairly simple problem, however, this will ensure that you are comfortable in doing such setups as and when required.

Algorithmic/Data Component (For All)

- Parse the stream and aggregate:
    - Total number of events and Acres burned
    - Identify high-risk fires (e.g., Acres Burned > 50000 or active more than 30 days: extinguished date - started date)
    - Time series of fire events and Acres burned per month
- **Bonus 1** (if time allows): Cluster fires within a 100 km radius that overlap in their active periods

- **Bonus 2** (if time allows): Calculate the average assessed property value by county and identify the top five fires with the greatest property losses over the past 20 years

    - Total Land Assess Value by County:
    https://www.boe.ca.gov/DataPortal/api/odata/County_Assessed_Property_Values_by_Property_Class_and_County
    - Total Land Area by County (Square Miles):
    https://gis.data.ca.gov/datasets/California::california-county-boundaries-and-identifiers

Data Science Track

- Use **Streamlit or Jupyter Notebook** to:
  - Display charts (e.g., matplotlib/seaborn/Plotly)
  - Show alert dashboard with filtering
  - Interactive map using `folium` or `plotly`
  - Use cross filtering option where when you filter something in the table, then the map will filter out those points in a similar way.

- Find (1) seasonality and (2) annual trend of fire events and severity in CA