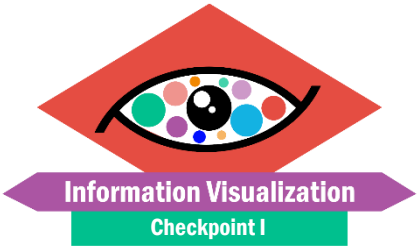

Checkpoint I: Project Proposal

Group: G43
Date: 2025/09/12

Problem Domain

This project addresses environmental hazards, public safety, and prevention by analyzing rural fires in Portugal. These fires threaten lives, infrastructure, and ecosystems, while also causing economic losses and community disruption. Portugal's Mediterranean climate, extensive forests, rural depopulation, and human activity increase vulnerability, making it vital to understand fire patterns, causes, and sizes. Prevention depends on both fire dynamics and resource allocation, with firefighter distribution playing a key role. By examining data on fires by year, region, cause, size, and personnel, this study seeks to identify trends, high-risk regions, and links between fire occurrence and resources, supporting proactive planning and prevention strategies.

Task Abstraction

This project analyzes rural fires in Portugal across temporal, spatial, categorical, and resource dimensions, examining trends, causes, sizes, and their relation to firefighting personnel, with emphasis on prevention strategies.

1. Temporal Trend Analysis

- **Abstract form:** How does X change over time?
- **Specific question:** How has the total number of fires in Alentejo Central evolved over the past fourteen years?
- **Task type:** Trends / Ordinal time

2. Regional Impact Comparison

- **Abstract form:** How does X differ across Y?
- **Specific question:** Which regions have the highest number of fires in 2024, and how does this compare with the number of firefighters in each region?
- **Task type:** Comparison / Geolocation / Relation

3. Cause Distribution Analysis

- **Abstract form:** What is the distribution of X?
- **Specific question:** What is the distribution of fire causes within Algarve in 2024, and does it relate with the number of firefighters available?
- **Task type:** Relation / Proportion / Relation

4. Multi-dimensional Relationships

- **Abstract form:** How does X relate to Y across multiple dimensions?
- **Specific question:** How do fire causes and dimensions relate with the number of firefighters in Alto Minho in 2023?

- **Task type:** Relation / Relationships / Multi-dimensional

5. Resource Proportional Analysis

- **Abstract form:** How does X proportionally compare across Y?
- **Specific question:** Which regions have the most or fewest firefighters relative to the number of fires, and does this reveal patterns of high fire risk?
- **Task type:** Proportion / Geolocation / Comparison

6. Multi-dimensional Relationships

- **Abstract form:** How does X relate to Y across multiple regions?
- **Specific question:** Is there a relationship between the number of fires and the number of firefighters in each region in 2020?
- **Task type:** Correlation / Relationship / Geolocation

7. Resource Proportional Analysis

- **Abstract form:** How does X distribute across Y?
- **Specific question:** How does the amount of burned area per firefighter vary across regions, and what does this suggest about prevention and response capacity?
- **Task type:** Distribution / Ratio / Multi-dimensional

These tasks enable exploration of trends, spatial differences, distributions, relations, and proportions, offering a comprehensive view of rural fire patterns and their influencing factors.

Data Abstraction

The data for this project comes from **multiple sources provided by INE (Instituto Nacional de Estatística)**, covering different aspects of rural fires in Portugal:

1. **Fires by Cause (Source)** – Number of fires in each region and year, broken down by cause (natural, negligence, intentional, etc.).
2. **Fires by Dimension (Source)** – Number of fires per size class for each region and year (< 1 ha, 1 - < 10 ha, …, >= 1 000 ha).
3. **Burned Area (Source)** – Percentage of area burned per region and year.
4. **Area (Source)** – Area per region in km²
5. **Firefighters (Source)** – Number of firefighting personnel in each region and year.

All datasets are **static** and table-based, downloadable in CSV. Each dataset contains a regional (NUTS-2024) and temporal (year) component that allows integration.

Data Processing and Cleaning

1. **Filtering and selection of relevant variables**
    - All variables in the datasets were deemed relevant to our study, so none were removed during preprocessing.
2. **Handling missing values**

- Missing fire counts or dimension counts → treated as zero.
- Missing causes → labeled as “Unknown.”
- Missing firefighters → labeled as “NULL.”
- Missing total area → labeled as “NULL.”

3. **Converting the CSV files into JSON files**
   - Convert the CSV files for the causes, dimensions, burned area percentage and firefighters into JSON files.
4. **Merging datasets**
   - Cause, dimension, burned area and firefighters’ data merged by Year and Region into a single unified dataset.
5. **Final export format**
   - Single JSON suitable for D3.js to power all dashboard components.

Attribute	Type	Scale	Description	Derived Measure?
Ano	Ordinal	Linear	Year of the fire occurrence	No
Região	Nominal	Categorical	NUTS-2024 region of the fire	No
Percentagem	Quantitative	Ratio	Percentage of burned area in the region	No
Sapadores	Quantitative	Ratio	Number of firefighters in the region for the year	No
Total	Quantitative	Ratio	Total number of fires in a region and year	No
Causas	Nominal	Categorical	List of causes of the fires	No
Dimensões	Ordinal	Categorical	List of the numbers of fires in each size class (<1 ha -> >= 1000 ha)	No
Area	Quantitative	Ratio	Total area of sub regions	No
Eficacia_Index	Quantitative	Ratio	Efficiency index per region and year (sapadores ÷ total fires). Higher values = better efficiency.	Yes
Prevencao_Index	Quantitative	Ratio	Preparedness index per region and year	Yes

			(sapadores ÷ area). Higher values = better preparedness.	

Mapping

- **How has the total number of fires in Alentejo Central evolved over the past fourteen years?**
  - **Attributes:** Região, Total
  - Varied between 142 and 369 annual occurrences. Peaks in 2012 (369) and 2019 (312); minimums in 2023 (142) and 2018 (159). We can see a slight downward trend overall.
- **Which regions have the highest number of fires in 2024, and how does this compare with the number of firefighters in each region?**
  - **Attributes:** Eficacia_Index
  - Highest number of fires: Alto Minho (0.230), Alto Tâmega e Barroso (0.640), Região de Coimbra (1.190).
  - Relation: The most affected regions don't always have the most firefighters (e.g., Alto Minho has far more fires than Coimbra but fewer firefighters).
- **What is the distribution of fire causes within Algarve in 2024, and does it relate with the number of firefighters available?**
  - **Attributes:** Causas
  - 345 fires; 209 negligence, 14 intentional, 115 undetermined and 40 Firefighters.
  - Conclusion: fires mostly due to negligence, no clear relation with the relatively low number of firefighters.
- **How do fire causes and dimensions relate with the number of firefighters in Alto Minho in 2023?**
  - **Attributes:** Causas, Dimensões and Sapadores
  - 1163 fires. Causes: 441 negligence, 481 intentional, 214 re-ignitions and 120 Firefighters.
  - Dimensions: 857 <1ha, 234 between 1–10ha, very few large fires.
  - Conclusion: despite many firefighters, most fires are human-caused and small.
- **Which regions have the most or fewest firefighters relative to the number of fires, and does this reveal patterns of high fire risk?**
  - **Attributes:** Eficacia_Index
  - Most firefighters /fire: In 2024, Região de Coimbra (194 sapadores/163 fires ≈ 1.19).
  - Fewest firefighters /fire: In 2024, Área Metropolitana do Porto (63 sapadores / 814 fires ≈ 0.07).
  - Pattern: Northern regions face more fires but proportionally fewer firefighters → higher risk.
- **Is there a relationship between the number of fires and the number of firefighters in each region in 2020?**
  - **Attributes:** Prevenção_Index
  - Overall, there is a linear relationship: regions with more fires tend to have more firefighters, reflected in high Eficacia_Index values.
- **How does the amount of burned area per firefighter vary across regions, and what does this suggest about prevention and response capacity?**
  - **Atributes:** Prevenção_Index
  - In Alentejo Central, there are only 16 sapadores for 7,393 km² (≈0.02 per km²). Large territories per firefighter are common across Portugal, highlighting coverage challenges and their impact on fire prevention.