



Creating new Features

Answer the questions in teams of 2 to 4 students

The Forest Fires dataset (<https://archive.ics.uci.edu/dataset/162/forest+fires>) contains information about forest fires in the northeast region of Portugal. It includes various attributes related to the meteorological conditions and fire details:

- **X** X-axis spatial coordinate within the Montesinho park map (integer value ranging from 1 to 9). Represents the east-west position of the fire within the park.
- **Y** Y-axis spatial coordinate within the Montesinho park map (integer value ranging from 2 to 9). Represents the north-south position of the fire within the park.
- **month** Month of the year (string value like 'jan', 'feb', etc.). Indicates the month when the fire occurred, useful for understanding seasonal patterns.
- **day** Day of the week (string value like 'mon', 'tue', etc.). Indicates the day of the week when the fire occurred, useful for weekly trends analysis.
- **FFMC** Fine Fuel Moisture Code (numeric value). Represents the moisture content of surface litter and other fine fuels. High values indicate dry conditions conducive to fire ignition.
- **DMC** Duff Moisture Code (numeric value). Indicates the moisture content of loosely compacted organic layers of moderate depth. High values suggest dry duff and increased fire potential.
- **DC** Drought Code (numeric value). Reflects the moisture content of deep, compact organic layers. High values indicate long-term dryness and potential for deep-burning fires.
- **ISI** Initial Spread Index (numeric value). Represents the rate of fire spread immediately after ignition, influenced by wind and fine fuel moisture.
- **temp** Temperature in Celsius (numeric value). The ambient temperature, which affects fire behaviour and intensity.
- **RH** Relative Humidity in percentage (numeric value). The amount of moisture in the air relative to the maximum moisture the air can hold at the same temperature. Lower values indicate drier air, which can contribute to fire spread.
- **wind** Wind speed in km/h (numeric value). Wind speed at the fire location, a critical factor in fire spread and intensity.
- **rain** Rainfall in mm/m² (numeric value). The amount of rainfall, which can help in fire suppression by increasing fuel moisture.
- **area** Burned area of the forest (in hectares) (numeric value). The target variable representing the size of the area affected by the fire.

Exercise

Create **three** new features that might be interesting for a Visualization project explaining the causes of forest fires in Montesinho park. For each of your new features, explain the **idea** behind the feature and **how you will create it** exactly.