Data Analysis Exam - Fall 2020

Data Documentation

The study from which the data originate considers forest fires in the Montesinho natural park from the Trás-os-Montes northeast region of Portugal (see Figure 1). This park contains high flora and fauna diversity. Inserted within a supra-Mediterranean climate, the average annual temperature is within the range 8 to 12 degrees Celsius. The data were collected from January 2000 to December 2003. On a daily basis, every time a forest fire occurred, several features were registered, including the spatial location within a 9 x 9 grid (see Figure 1), the date, several covariates relevant for predicting fire severity, and the total burned area. A description of the measured variables is included in the table below.

247 of the 517 measurements have a zero value for area. A zero value means that an area smaller than 0.01ha = $100m^2$ was burned. To reduce skewness, the logarithm function y=log(x + 1) was applied in the original analysis.

Variable	Description
X	horizontal coordinate (see Figure 1)
Y	vertical coordinate (see Figure 1)
month	month of year
day	day of the week
FFMC	Fine Fuel Moisture Code is the moisture content of *surface* organic material and affects ignition and fire spread
DMC	Duff Moisture Code is the moisture content of *shallow* depth organic layers and affects fire intensity
DC	Drought Code is the moisture content of *deep* organic layers and affects fire intensity
ISI	Initial Spread Index is a score that predicts the velocity of fire spread
temp	outside temperature (degrees Celsius)
RH	outside relative humidity in (%)
wind	outside wind speed in (km/hr)
rain	outside rain volume (mm/m^2)
area	total burned area (hectares)

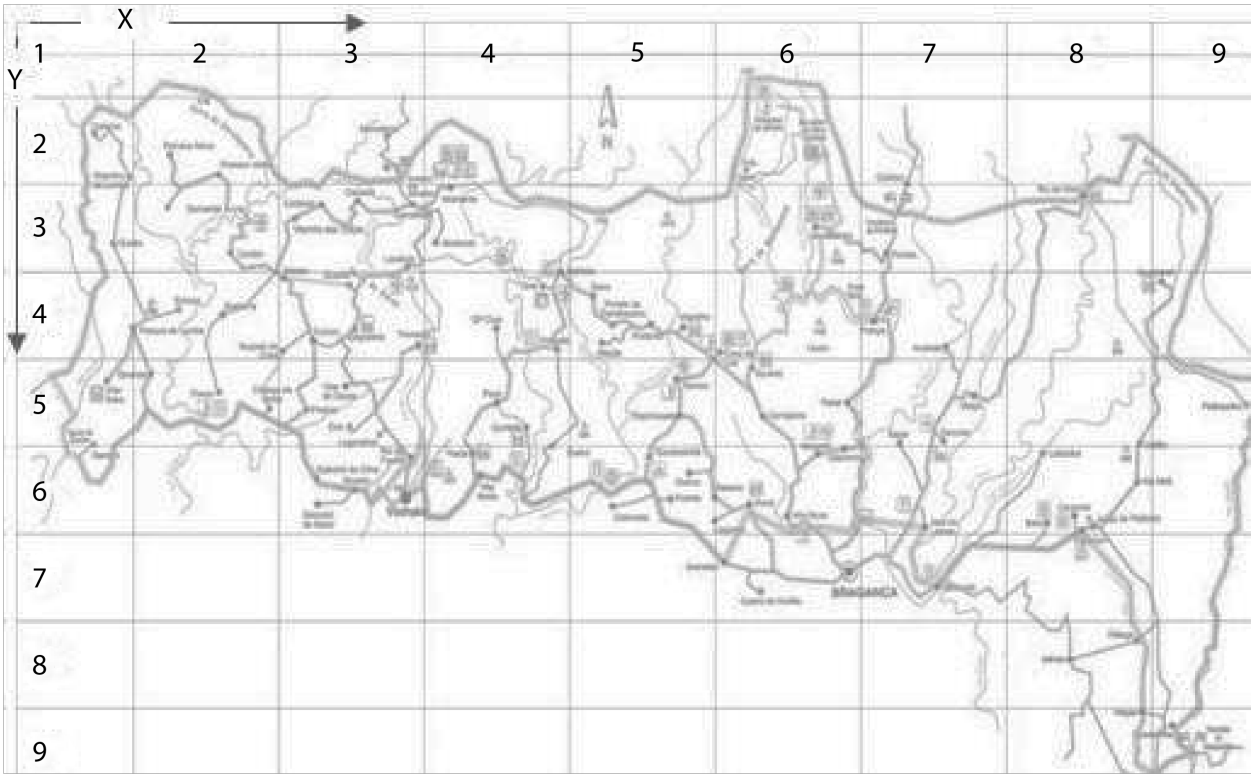

Figure 1: Montesinho natural park