

Spatial Data in R

Why We Need Spatial Data



I WANT TO MAKE A DISASTER MOVIE
THAT JUST SHOWS SCIENTISTS RUSHING
TO UPDATE ALL THEIR DATA SETS.

What is Spatial Data?

Spatial data are either discrete locations (roads, countries, oceans) or continuous fields (elevation, air quality, land type). They are represented in 2 ways

1. Vector Data – points, lines and polygons composed by coordinate pairs
2. Raster Data – represents continuous variables by dividing the world into grids/pixels, with each pixel representing an average value for the area it covers

Coordinate Reference System

To represent the location and shape of an object you need 2 pieces of information:

1. Coordinates
2. Reference system for how the coordinates relate to the shape of the Earth.

Coordinate Reference system (CRS): a geometric model of the shape of the Earth, which identifies the origin and orientation of the coordinate axes on the ellipsoid and the units of measurement.

Vector Data in R

sp (spatial polygons)

defines *classes* to represent spatial data. It is an S4 object, where objects store spatial geometries separately from associated attribute data, matching by order

sf (simple features)

stored as a native R dataframe, where the geometries are in a list-column. It implements the “simple features standard” that is also used in PostGIS, GeoJSON, and ArcGIS, while sp predates industry standards

Objects in the spatial features package

sfg: geometry

- geometry of a single feature
- vector, matrix, or list of matrices of coordinates with defined dimension and type of geometry
- seven main geometry types

sfc: geospatial geometry

- list of sfg objects
- coordinate reference system through crs attribute
- seven subclasses based on geometries

sf: geospatial geometry with attributes

- data frame housing attributes
- geometry column of class sfc