Data and slides

http://bit.ly/4ms5lci

Is my house at risk of flooding? A data-driven investigation to track buildings in flood-prone zones

Miguel Ros García
Data Journalist at Newtral.es

miguel.ros@newtral.es

Newtral

Index

1. Summary of our investigation
2. What is spatial data and why we use it
3. How did we find the data
4. Difficulties
5. How can be replicated (for any research)
6. The process
7. **Hands on**

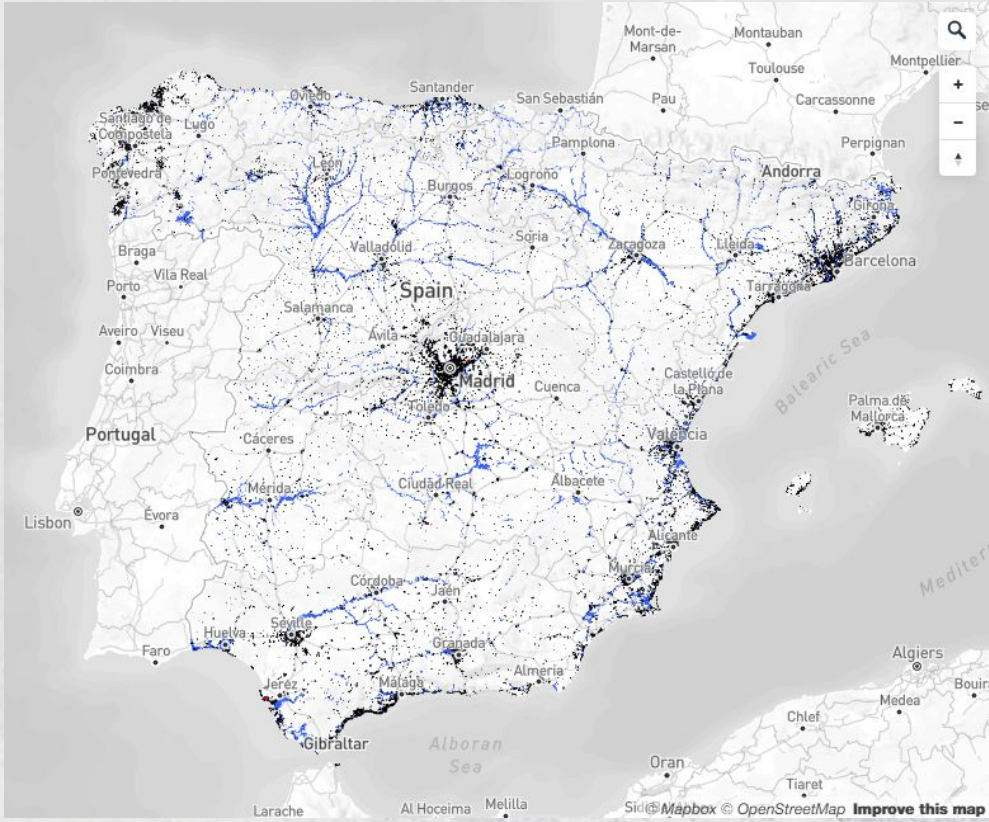
We found 200,000 buildings constructed in areas with a high risk of flooding in Spain

Why did we decide to do this project?

How did we carry out this data-driven investigation?

What kind of data did we use?

How did we find and analyze the data?



Leverage the data

Try to make the most of our data

With the same data we have
released a special with 8 reports

Why did we decide to do this project?

Nevtral

Why did we decide to do this project?

October 29, 2024

Isolated Depression at High Levels.

One of the greatest natural disasters in Spain.

227 deceased.

Thousands of homes affected.

Deaths in Spain by type of natural disaster

Data between 1995 and 2015

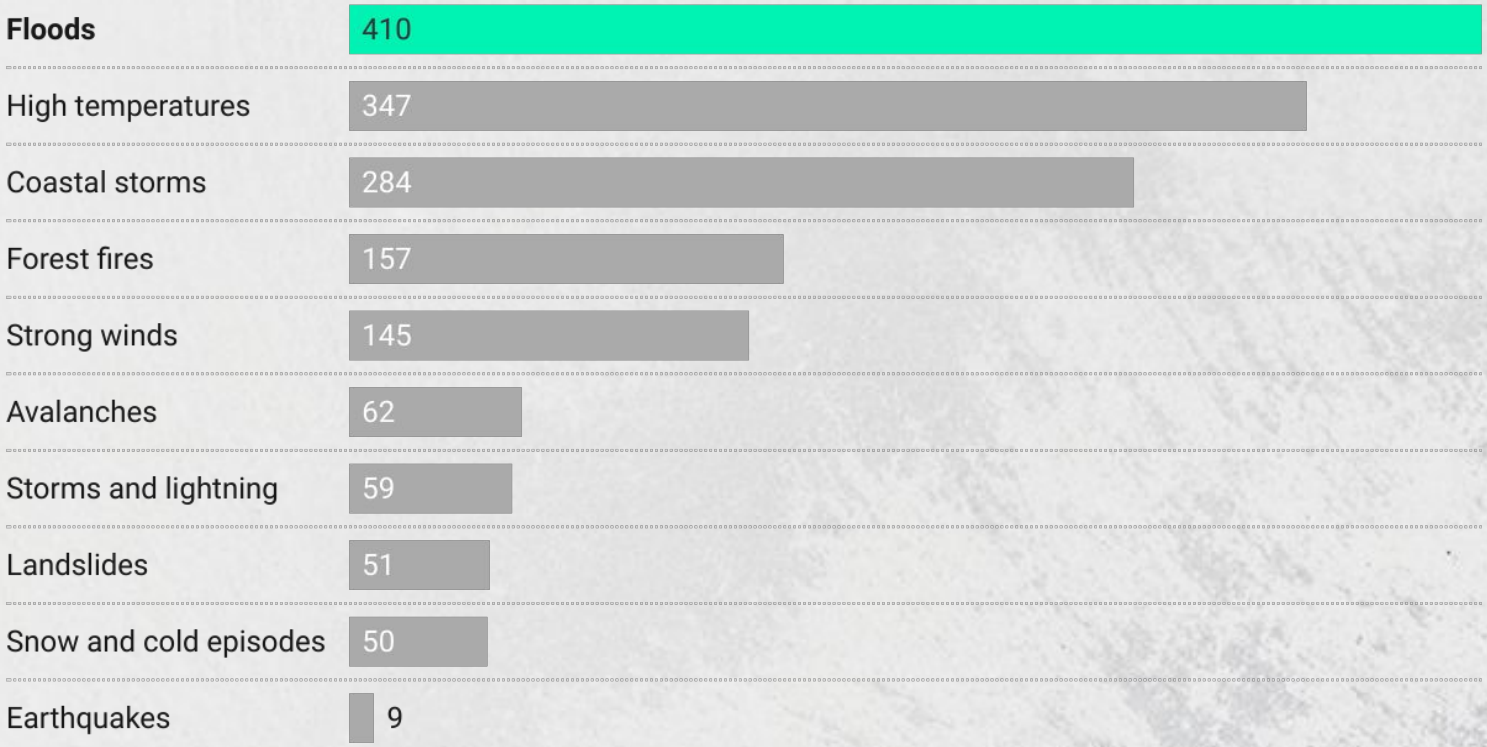


Gráfico: Newtral.es • Fuente: Compilation of data from Civil Protection • Creado con Datawrapper

Newtral

Our hypothesis

Floods in Spain are recurrent and are the natural disaster that affects the country the most.

Just like in Valencia, there must be many homes at risk of flooding throughout the country.

Our objective

Quantify the number of homes at risk of flooding

What is geospatial data?

Information which describes objects with a specific location on Earth

- Points
- Lines
- Polygons

Why use geo data?

It allows you to verify data, uncover relevant information and tell engaging stories about the environment.

How did we find the data?

Flood prone zones

Ministry for Ecological Transition and the Demographic Challenge

Cadastral data

General Directorate of Cadastre of Spain

Difficulties we face

Difficulties we face

1. Unify the data
2. Handle large volumes of data
3. Optimize data to be able to visualize it
4. Reduce costs

Geo data resolutions

Low: over 60m/pixel
Medium: 10 – 30m/pixel
High: 3m – 5m/pixel
Very high: less than 1m/pixel

How can it be replicated?

How can it be replicated?

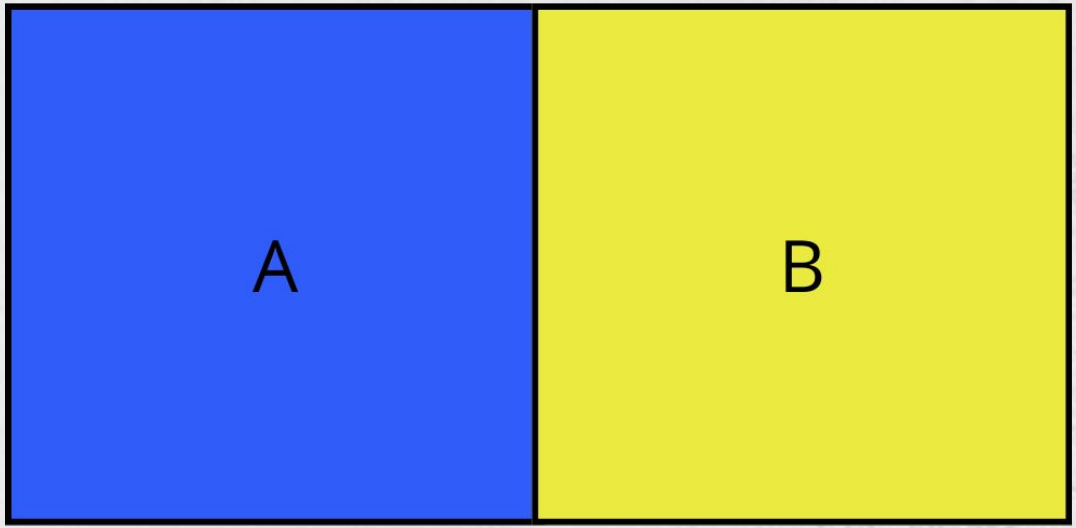
Spatial join

How can it be replicated? - SPATIAL JOIN

Different from traditional ‘joins’.

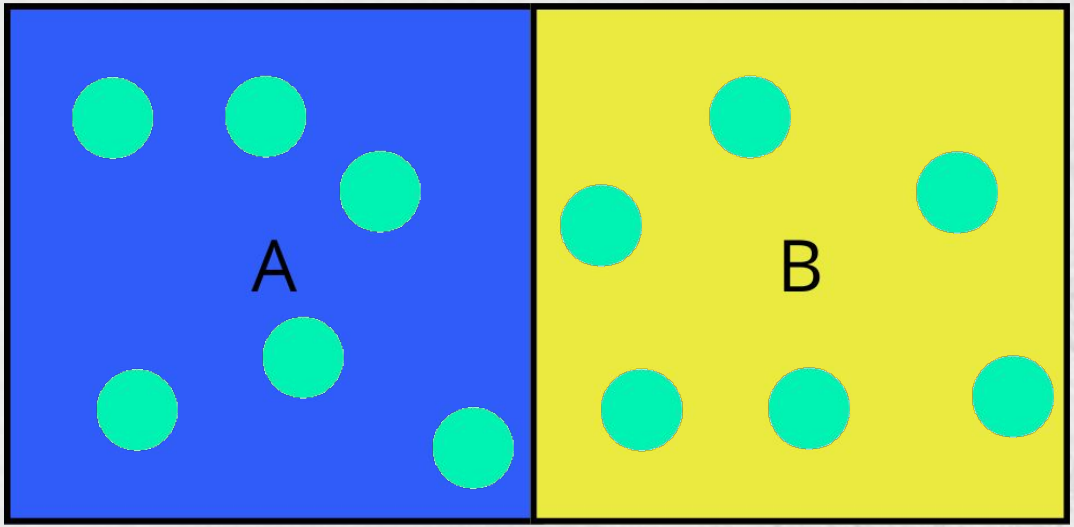
It compares two geospatial files based on the spatial relationship of their geometries.

How can it be replicated? - SPATIAL JOIN



How can it be replicated? - SPATIAL JOIN

How can it be replicated? - SPATIAL JOIN



How can it be replicated? - SPATIAL JOIN

1 2 3 4 5 7 8

```
parcels_flooded = gpd.sjoin(parcels, flood_zones, how='inner', predicate='intersects')
log_message(f"Parcels in flood zones found in {parcel_path}: {len(parcels_flooded)}")
```

What tools are needed?

What tools are needed?

python™

mapbox

The process

Transform

Transform the raw files downloaded from official sources, which are typically in Shapefile or GML formats, into more manageable formats such as GeoJSON. Python is used in this step.

Analyse and cleaning

Once we have the files in the desired format and organized into folders, we explore and clean the files (removing unwanted columns or omitting errors).

In this way, we can reduce the file size by half. In this step, we also use Python.

Spatial Join

With the clean and streamlined data, we proceed to perform a spatial join between two geospatial data files.

This allows us to identify and extract the matching geometries.

Python is also used in this step.

Optimize

The results are combined into one, two, or three files (depending on the size). Both the initial cadastral data and the geometries isolated from the spatial join need to be processed again.

This process is done with Tippecanoe, and the result is files in Mbtiles format.

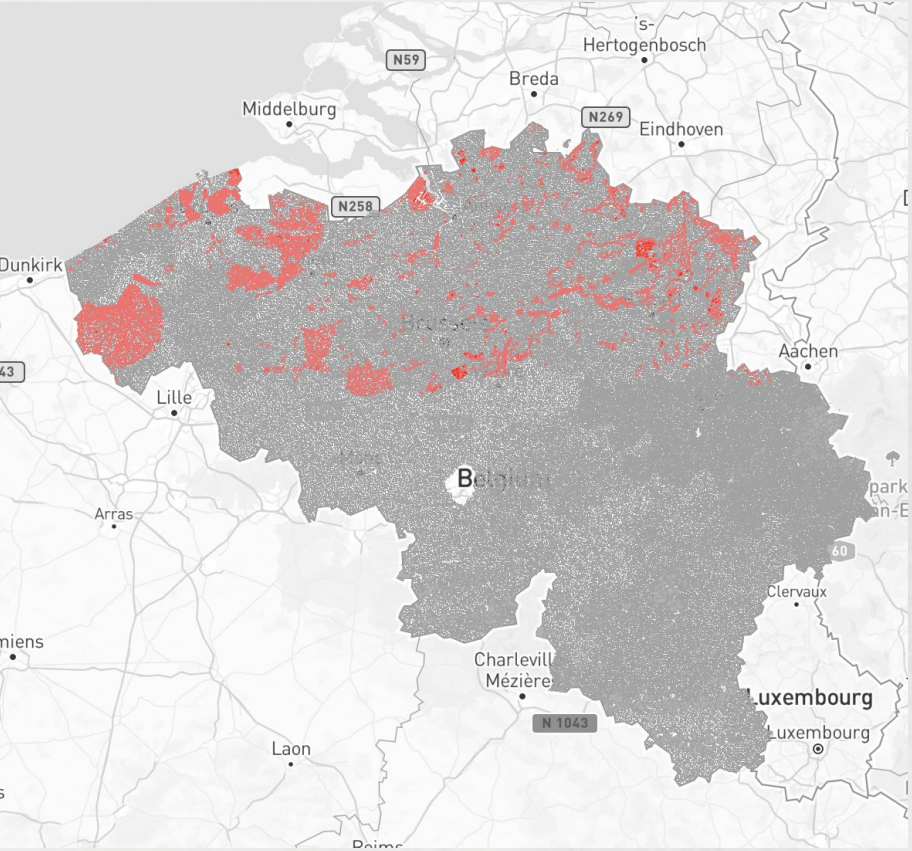
Visualize

These Mbtiles files are now ready to be visualized in an application like Mapbox. We upload them from Mapbox Studio, and that's it!

For a more professional touch, the map can be created using HTML and JavaScript code.

Our objective

https://goo.su/D2vOTb



Nevtral