

Mastering Big Data Visualization & Storytelling Workshop II: Assessing the damages of La Palma Volcano

Since 11 September 2021, a swarm of seismic activity has been ongoing in the southern part of the Spanish Canary Island of La Palma (Cumbre Vieja region). The increasing frequency, magnitude, and shallowness of the seismic events were an indication of a pending volcanic eruption; which occurred on 16th September, leading to evacuation of people living in the vicinity.

In this tutorial we are going to assess the number of buildings and population that may get affected by the lava flow. We'll also estimate the value of damaged residential properties affected by the volcano eruption. You can access the map [here](#).

Your CARTO account

For this session, you'll need a CARTO account!

If you don't have one, you can set up a free 14-day trial at app.carto.com. This should only take a couple of minutes to do, but we do recommend setting this up before coming to the workshops so you can dive right in!

! There is a maximum of one CARTO account per email address. If you have previously set up a free trial with your email, we recommend using an alternative email address for this session. If you run into any issues setting up an account, please contact support@carto.com.

Access third-party data using CARTO

Before starting processing and visualizing the data, we will see how to access third-party data from CARTO. There are public and premium data sources available that can be useful for our analysis, and some of them can be subscribed for free!

Specifically, we will be using the [Spatial Features](#) dataset. This unique data product provides derived variables across a wide range of themes including demographics, points of interest, and climatology data with global coverage.

1. In the [CARTO workspace](#), head to the **Data Observatory** and type in 'Spatial Features Spain' in the search tool bar
2. Select the H3 dataset (interested in Spatial Indexes? Click [here](#) for more details)
3. Then click on **Subscribe for free**. You should now have access to the full dataset!

The screenshot shows the Data Observatory interface within the CARTO workspace. The search bar at the top contains the query "Spatial Features Spain". Below the search bar, there are two sections: "Popular data products" and "Top providers". Under "Popular data products", several datasets are listed: "Spatial Features - Spain (H3 Resolution 8)", "Spatial Features - Spain (Quadgrid 18)", "Spatial Features - Spain (Quadgrid 15)", "Sociodemographics - Spain (Grid 100m)", and "Spatial Features - United States of America (Quadgrid 18)". Under "Top providers", providers like Experian, Michael Bauer International, Overture Maps, The Data Appeal Company, and Applied Geographic Solutions are listed with their respective data product counts. At the bottom of the interface, there is a search bar with the placeholder "Search for *Spatial Features Spain*" and a progress bar indicating the total number of data products found: 4659, 2279, 1743, and 1364.

4. To verify that the data was correctly subscribed, go to the **Data Explorer > Data Observatory > CARTO > Subscriptions > Spatial Features - Spain (H3 Resolution 8) [v3]** and take a look at the data preview

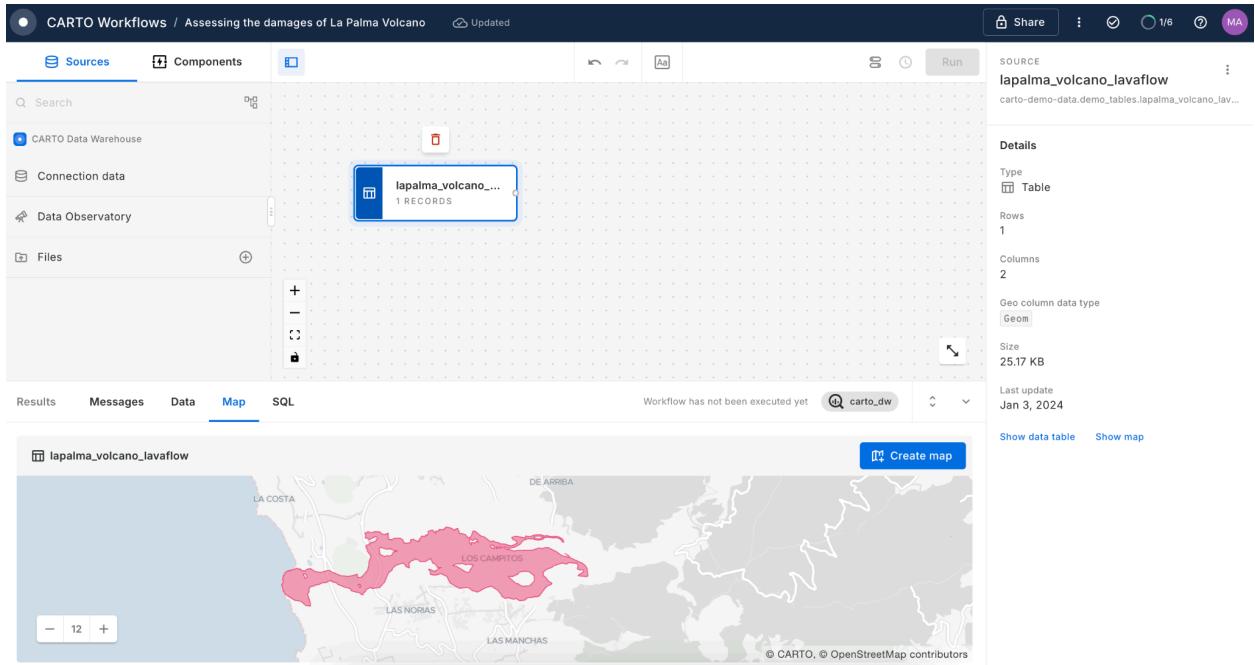
Column name	Type	Description	Slug
male_25_to_29	Float	Male population between 25 and 29 years old	male_25_to_29_b7ce9817
geoid	String	Unique cell identifier	geoid_6eda440d
do_date	Date	First day of the year of the corresponding data	do_date_36be358f
country_iso	String	Country name, following the ISO 3166 standard	country_iso_e7deda83
country_iso_a3	String	Three-letter country code, following the ISO 3166 standard	country_iso_a3_a258fc17
female_under_1	Float	Female population below 1 year	female_under_1_12624c93
female_1_to_4	Float	Female population between 1 and 4 years old	female_1_to_4_75835a74
female_5_to_9	Float	Female population between 5 and 9 years old	female_5_to_9_90a364df
female_10_to_14	Float	Female population between 10 and 14 years old	female_10_to_14_d5d12369
female_15_to_19	Float	Female population between 15 and 19 years old	female_15_to_19_f9587073
female_20_to_24	Float	Female population between 20 and 24 years old	female_20_to_24_70737749
female_25_to_29	Float	Female population between 25 and 29 years old	female_25_to_29_5cfa2453
female_30_to_34	Float	Female population between 30 and 34 years old	female_30_to_34_a5c24696

Assessing the damages using Workflows

Lava flow

The first thing that we are going to do is to take a look at the lava flow from the Volcano eruption in La Palma, Spain as measured by the Copernicus satellite on 10/04/2021.

1. In the [CARTO workspace](#), head to the **Data Explorer > CARTO Data Warehouse > demo data > demo_tables > lapalma_volcano_lavaflow**
2. Click on **Create workflow**. This will load your data into a new **CARTO Workflow** canvas so that you can start working with it!
3. **Rename the workflow** to “Assessing the damages of La Palma Volcano”
4. Click on the **Map** tab in the bottom section to see a preview or the lava flow



Buildings and population

We will now estimate the property loss and affected population by the lava flow and its deposits:

1. Go to the **Sources** section on the left-side menu, then to **Connection data > demo data > demo_tables > lapalma_buildings**
2. Drag and drop the source to the canvas. Navigate through the **Data** tab in the bottom section to see the information provided for each building
3. Then, go to the **Sources** section on the left-side menu, then to **Data Observatory > CARTO > Subscriptions > Spatial Features - Spain (H3 Resolution 8) [v3]**

To process these data, then go to the **Components** section, and drag and drop the **Enrich Polygons** component to obtain the total of estimated property value of those residential properties affected by the lava flow as well as the total number of parcels affected:

1. Connect `lapalma_volcano_lavaflow` as the **target** polygon and `lapalma_buildings` as the **source**
2. In the Variables section, in the node, add `SUM` for `estimated_prop_value` column and `COUNT` aggregation for `numberOfBuildingUnits` column
3. The output result is the lava flow source with the addition of the two new properties

SOURCE
lapalma_buildings
carto-demo-data.demo_tables.lapalma_buildings

Details
Type: Table
Rows: 35,605
Columns: 27
Geo column data type: Geom
Size: 30 MB
Last update: Apr 13, 2023

DATA ENRICHMENT
Enrich Polygons
SUCCESS

This uses a data source (either a table or a DO subscription) to enrich another target table.

Enrich Polygons documentation

Settings
Target polygons geo column: geom
Source geo column: geom
Variables
estimated_prop_value: SUM
numberOfBuildingUnits: COUNT
population_93405ad7, sum, populat...
Remove all

Use another **Enrich Polygons** component to obtain the total population affected by the lava flow extent:

1. Get the geometries of the H3 cells by using **H3 Boundary**
2. **Create Column** to obtain the total vulnerable population (young and elderly population). Call your new variable **vulnerable_population**:

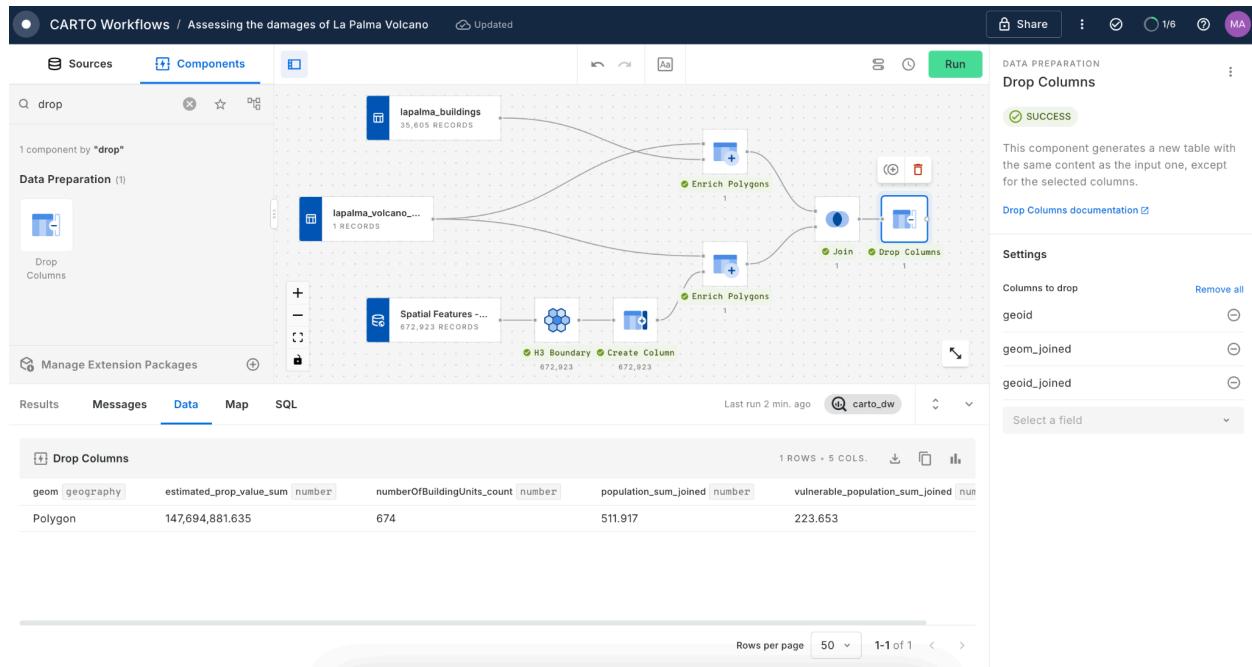
```

female_under_1 + female_1_to_4 + female_5_to_9 + female_5_to_9 +
female_10_to_14 + female_15_to_19 + female_70_to_74 + female_75_to_79 +
female_80_and_over + male_under_1 + male_1_to_4 + male_5_to_9 + male_5_to_9 +
male_10_to_14 + male_15_to_19 + male_70_to_74 + male_75_to_79 +
male_80_and_over

```

3. Link `lapalma_volcano_lavaflow` as the **target** and your processed Spatial Features data (the **Create Column** component) as the **source**
4. Then, in the Variables section add **SUM** of population and **vulnerable_population** columns
5. The output result is the lava flow source with the addition of the two new properties

Lastly, **Join** the two outputs and **Drop Columns** that are unnecessary:



Now you have two options:

- Use the **Save as Table** component to save your results in a specific destination (i.e. `lapalma_volcano_lavaflow_enriched`) and later show them in a map. Click on **Rename and select destination**, then navigate to **organization data > shared_us** and type in the table name.

⚠️ Be careful not to overwrite any existing table!

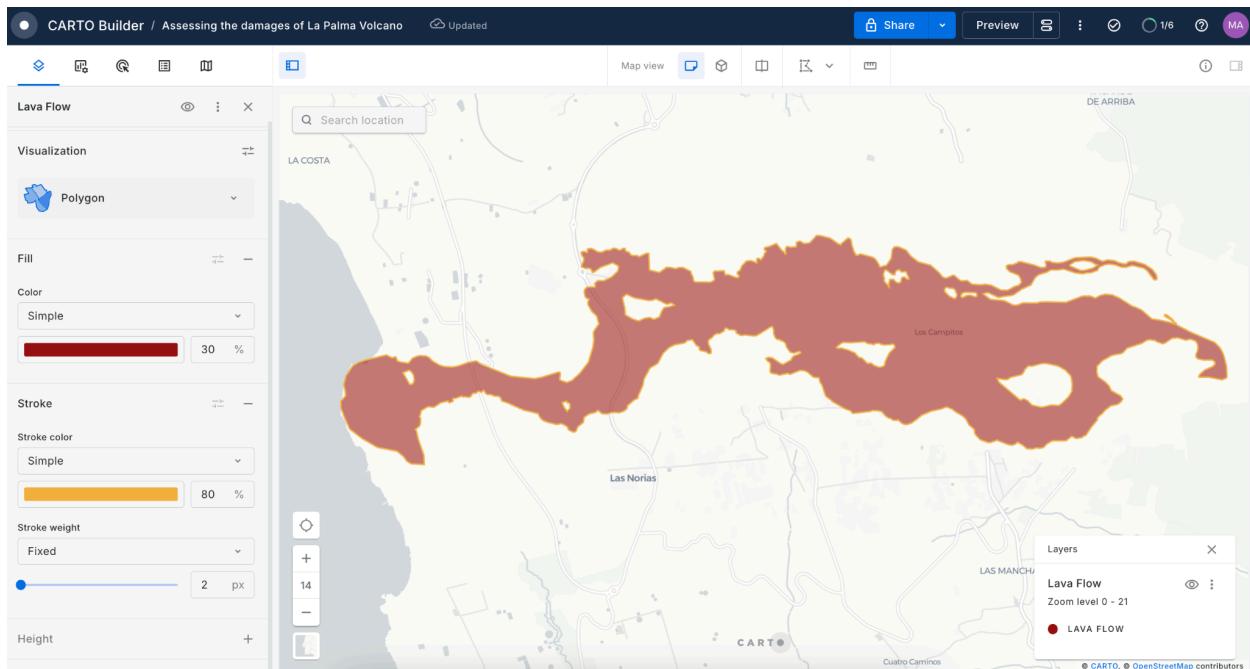
- Click on the **Map** tab in the bottom section to see a preview and directly create a map by clicking on **Create map**
-

Style your map!

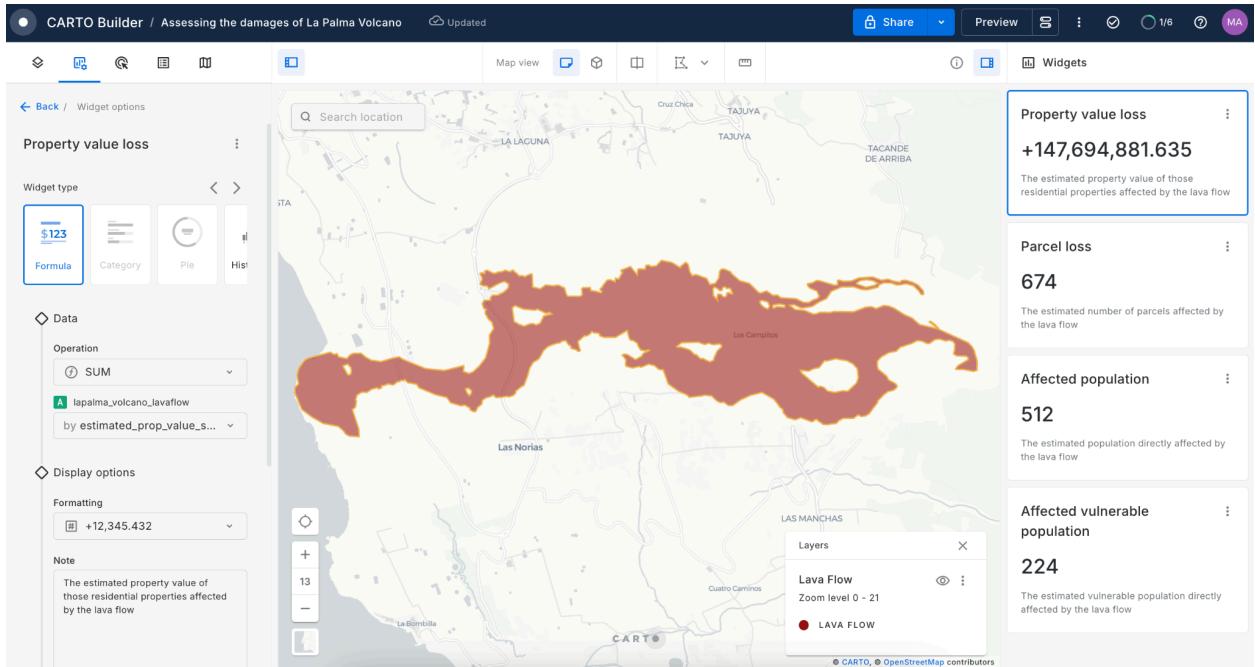
Lava flow

The first thing that we are going to do is to take a look at the lava flow from the Volcano eruption in La Palma, Spain as measured by the Copernicus satellite on 10/04/2021.

1. In the [CARTO workspace](#), head to the **Data Explorer > CARTO Data Warehouse > demo data > demo_tables > lapalma_volcano_lavaflow_enriched**
2. Click on **Create map**. This will load your data into a new **CARTO Builder** map so that you can customize it!
3. **Rename the map** to “Assessing the damages of La Palma Volcano”
4. **Rename the layer** to “Lava Flow”



5. Add some **Widgets** to show the metrics we just computed. Use the **Formula** widget to show the details.



Buildings

Before importing the data to Builder, we will need to do some processing in advance using **Workflows**. We are interested in the geography of each building, the current use of each building, the year of construction and other details like the number of building units or the number of dwellings. We will create a pipeline to extract this information:

1. **Select** the relevant features: `geom, beginning, currentUse, numberOfWorkingUnits, numberOfWorkingDwellings, numberOfWorkingFloorsAboveGround` from the `lapalma_buildings` dataset
2. **Create Column** to extract the year of construction from the attribute `beginning`: `LEFT(beginning, 4)` & set your column name to `construction_year` → this just takes the four first characters of the string
3. Filter the results using **Where** and remove those buildings whose construction year has not been well-informed: `construction_year NOT LIKE '%-%'` → this just takes out all those rows whose construction year contains the character -
4. **Cast** the new column `construction_year` to `INT64`
5. **Drop Columns** `beginning`
6. **Save as Table** as `lapalma_buildings_processed`

⚠️ Be careful not to overwrite any existing table!

CARTO Workflows / Assessing the damages of La Palma Volcano

Components

Input/Output (2)

Save as Table **Send by Email**

lapalma_buildings 35,605 RECORDS

lapalma_volcano...

lapalma_buildings_processed carto-dw-ac-c12zs8sn.shared_us.lapalma_bu...

Append to existing table

Advanced options

Save as Table

35,368 ROWS + 6 COLS.

geom (geography)	currentUse (string)	numberBuildingUnits (number)	numberDwellings (number)	numberFloorsAboveGround (number)	construction_year (number)
Polygon	_agriculture	2	0	1	1,960
Polygon	4_2_retail	1	0	1	1,960
MultiPolygon	_agriculture	1	0	1	1,960
MultiPolygon	_agriculture	1	0	1	1,960

Rows per page: 200 | 2201-2400 of 35368

7. Go to our Lava flow map and click on **Add source from...** > **Data Explorer** > **organization data** > **shared_us** > **lapalma_buildings_processed** below the sources section
8. **Rename the layer** to "La Palma Buildings" and style it by construction year. We have set a custom color scale that divides the buildings into 4 groups, from older to newer. Make sure you also get your **Legend** ready!

CARTO Builder / Assessing the damages of La Palma Volcano

La Palma Buildings

Simple

Color scheme by construction_year

Stroke

Stroke color Simple

Stroke weight Fixed

Property value loss +147,694,881.635

The estimated property value of those residential properties affected by the lava flow

Parcel loss 674

The estimated number of parcels affected by the lava flow

Affected population 512

The estimated population directly affected by the lava flow

Affected vulnerable population 224

The estimated vulnerable population directly affected by the lava flow

Population

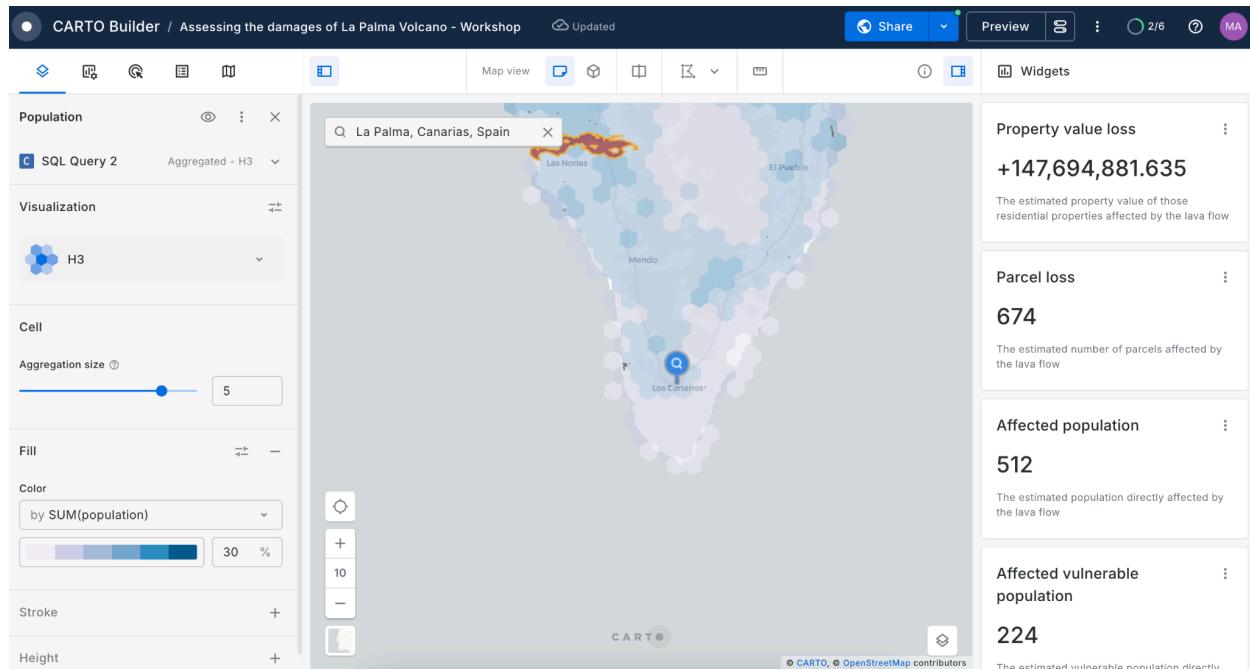
Lastly, let's add a population layer by clicking on **Add source from... > From Data Observatory > CARTO > Subscriptions > Spatial Features - Spain (H3 Resolution 8)**

[v3]. You'll see that a SQL Editor opens up, that's just the query that we are using to load the data. Let's tweak it a bit to add the vulnerable population variable that we are interested in, just copy the following and click on **Run**:

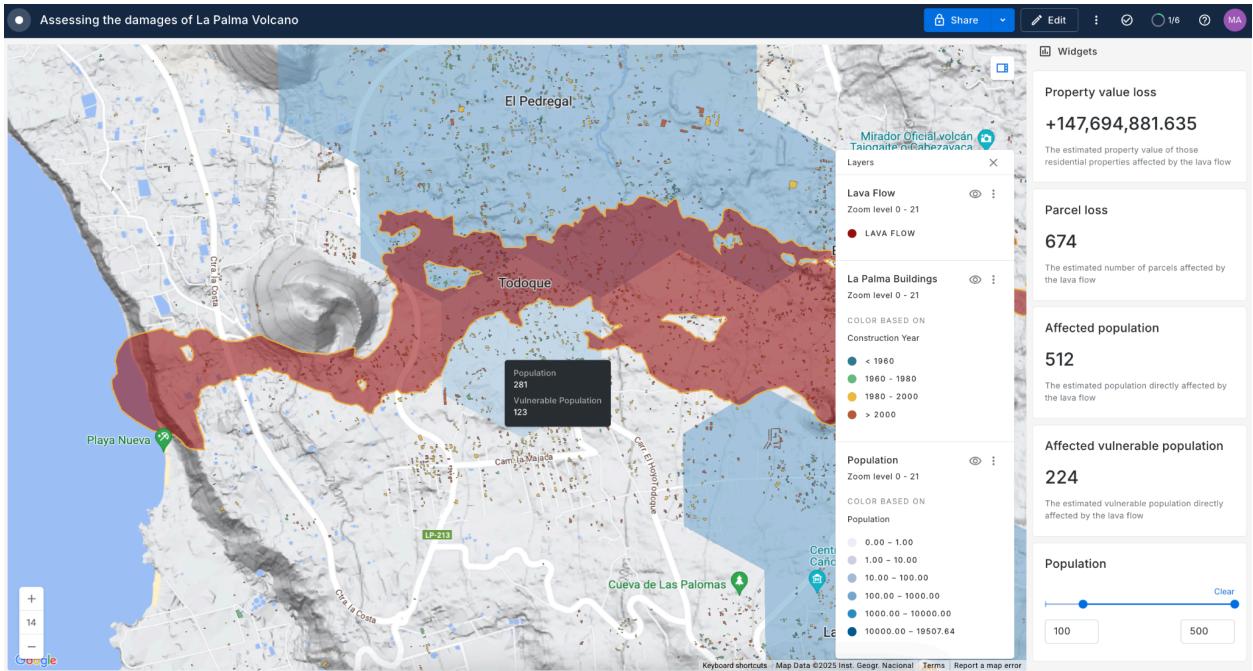
```
SELECT geoid AS h3, *EXCEPT(geoid), female_under_1 + female_1_to_4 + female_5_to_9 + female_5_to_9 + female_10_to_14 + female_15_to_19 + female_70_to_74 + female_75_to_79 + female_80_and_over + male_under_1 + male_1_to_4 + male_5_to_9 + male_5_to_9 + male_10_to_14 + male_15_to_19 + male_70_to_74 + male_75_to_79 + male_80_and_over as vulnerable_population FROM `carto-data.ac_c12zs8sn.sub_carto_derived_spatialfeatures_esp_h3res8_v1_yearly_v3`
```

Now, **rename your layer** to "Population", style it accordingly and add some widgets/tooltips. Be creative! Notice that, since this is aggregated data, we need to specify how the data should be aggregated as we navigate through the map. For instance, population is to be summed.

#ProTip: you can type in the name of the location (La Palma, Canarias, Spain) you are interested in the search tool bar, and the map will be automatically centered in that area!



After some styling, you can end up with a map looking like this:



Congratulations - **you've completed the workshop!**