

Spatial Data Science Conference

Integrating CARTOframes into Spatial Data Science workflows

Introductions



Miguel Álvarez

**Data Scientist
at CARTO**



Giulia Carella, PhD

**Data Scientist
at CARTO**

Data Science Workflows using CARTO



HOW IT WORKS

CARTO turns your Location Data Into Business Outcomes

Whether it's more efficient delivery routes, strategic store placements or targeted geomarketing campaigns - CARTO makes it simple in 5 key steps:



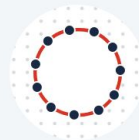
**Data
Ingestion**



**Data
Enrichment**



Analysis



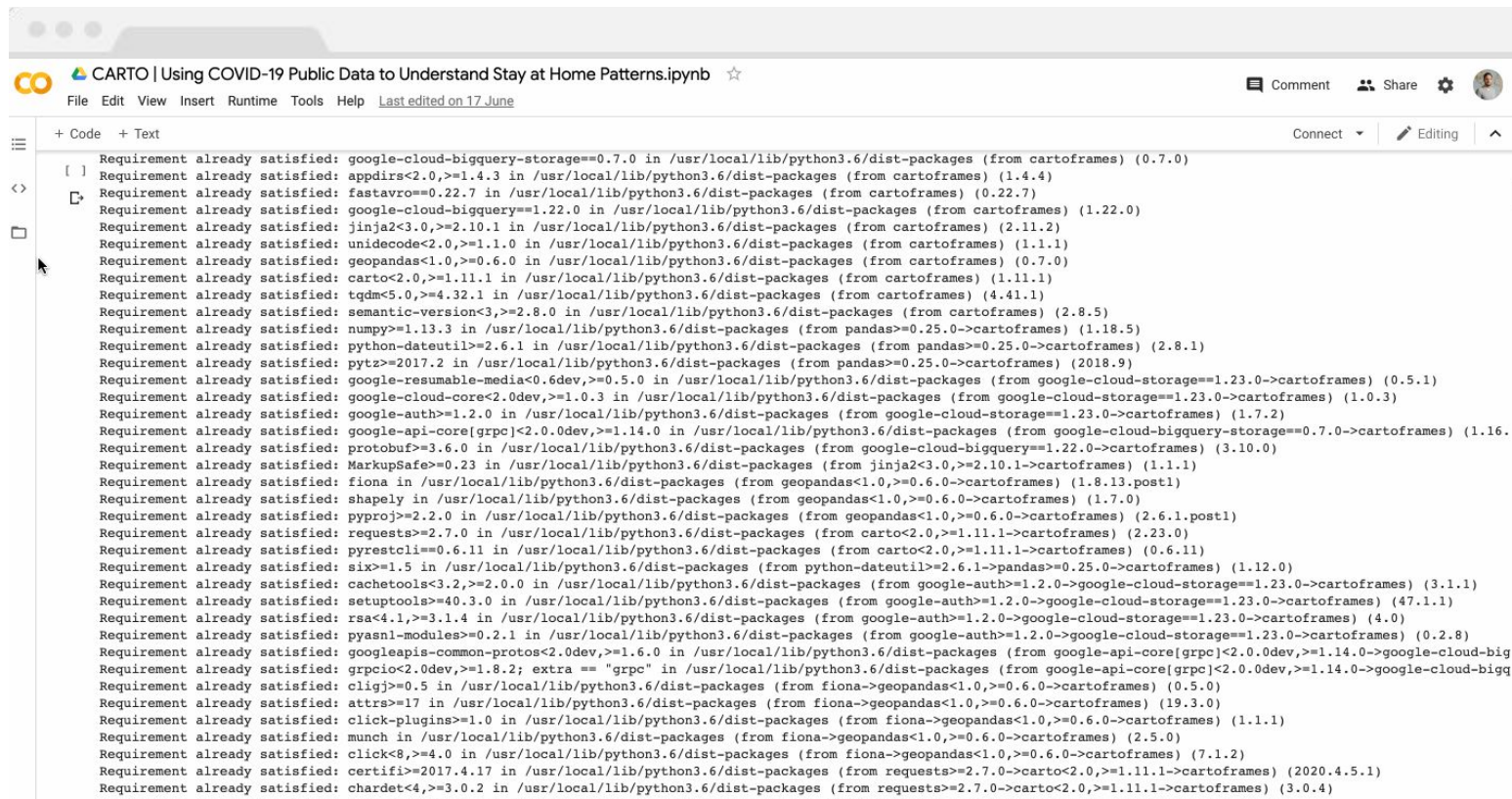
**Solutions &
Visualization**



Integrations

CARTOframes

A Python Library to Facilitate your Spatial Data Analysis Workflow




The screenshot shows a Jupyter Notebook interface with a toolbar at the top containing icons for file operations, runtime, tools, help, and a star icon. The notebook title is "CARTO | Using COVID-19 Public Data to Understand Stay at Home Patterns.ipynb". The interface includes a menu bar with "File", "Edit", "View", "Insert", "Runtime", "Tools", "Help", and "Last edited on 17 June". On the right side, there are buttons for "Comment", "Share", and a settings gear icon. Below the menu bar, there are tabs for "Code" and "Text", and a "Connect" button. The main area displays a list of installed Python packages and their versions, along with the paths where they are installed. The list includes packages like google-cloud-bigquery-storage, appdirs, fastavro, google-cloud-bigquery, Jinja2, unicode, geopandas, carto, tqdm, semantic-version, numpy, python-dateutil, pytz, google-resumable-media, google-cloud-core, google-auth, google-api-core, protobuf, MarkupSafe, Fiona, shapely, pyproj, requests, pyrestcli, six, cachetools, setuptools, rsa, pyasn1-modules, googleapis-common-protos, grpcio, cligj, attrs, click-plugins, munch, click, certifi, and chardet.

```
Requirement already satisfied: google-cloud-bigquery-storage==0.7.0 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (0.7.0)
Requirement already satisfied: appdirs<2.0,>=1.4.3 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (1.4.4)
Requirement already satisfied: fastavro==0.22.7 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (0.22.7)
Requirement already satisfied: google-cloud-bigquery==1.22.0 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (1.22.0)
Requirement already satisfied: Jinja2<3.0,>=2.10.1 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (2.11.2)
Requirement already satisfied: unicode<2.0,>=1.1.0 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (1.1.1)
Requirement already satisfied: geopandas<1.0,>=0.6.0 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (0.7.0)
Requirement already satisfied: carto<2.0,>=1.11.1 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (1.11.1)
Requirement already satisfied: tqdm<5.0,>=4.32.1 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (4.41.1)
Requirement already satisfied: semantic-version<3,>=2.8.0 in /usr/local/lib/python3.6/dist-packages (from cartoframes) (2.8.5)
Requirement already satisfied: numpy>=1.13.3 in /usr/local/lib/python3.6/dist-packages (from pandas>=0.25.0->cartoframes) (1.18.5)
Requirement already satisfied: python-dateutil>=2.6.1 in /usr/local/lib/python3.6/dist-packages (from pandas>=0.25.0->cartoframes) (2.8.1)
Requirement already satisfied: pytz>=2017.2 in /usr/local/lib/python3.6/dist-packages (from pandas>=0.25.0->cartoframes) (2018.9)
Requirement already satisfied: google-resumable-media<0.6dev,>=0.5.0 in /usr/local/lib/python3.6/dist-packages (from google-cloud-storage==1.23.0->cartoframes) (0.5.1)
Requirement already satisfied: google-cloud-core<2.0dev,>=1.0.3 in /usr/local/lib/python3.6/dist-packages (from google-cloud-storage==1.23.0->cartoframes) (1.0.3)
Requirement already satisfied: google-auth==1.2.0 in /usr/local/lib/python3.6/dist-packages (from google-cloud-storage==1.23.0->cartoframes) (1.7.2)
Requirement already satisfied: google-api-core[grpc]<2.0.0dev,>=1.14.0 in /usr/local/lib/python3.6/dist-packages (from google-cloud-bigquery-storage==0.7.0->cartoframes) (1.16.0)
Requirement already satisfied: protobuf>=3.6.0 in /usr/local/lib/python3.6/dist-packages (from google-cloud-bigquery==1.22.0->cartoframes) (3.10.0)
Requirement already satisfied: MarkupSafe>=0.23 in /usr/local/lib/python3.6/dist-packages (from Jinja2<3.0,>=2.10.1->cartoframes) (1.1.1)
Requirement already satisfied: Fiona in /usr/local/lib/python3.6/dist-packages (from geopandas<1.0,>=0.6.0->cartoframes) (1.8.13.post1)
Requirement already satisfied: shapely in /usr/local/lib/python3.6/dist-packages (from geopandas<1.0,>=0.6.0->cartoframes) (1.7.0)
Requirement already satisfied: pyproj>=2.2.0 in /usr/local/lib/python3.6/dist-packages (from geopandas<1.0,>=0.6.0->cartoframes) (2.6.1.post1)
Requirement already satisfied: requests>=2.7.0 in /usr/local/lib/python3.6/dist-packages (from carto<2.0,>=1.11.1->cartoframes) (2.23.0)
Requirement already satisfied: pyrestcli==0.6.11 in /usr/local/lib/python3.6/dist-packages (from carto<2.0,>=1.11.1->cartoframes) (0.6.11)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.6/dist-packages (from python-dateutil>=2.6.1->pandas>=0.25.0->cartoframes) (1.12.0)
Requirement already satisfied: cachetools<3.2,>=2.0.0 in /usr/local/lib/python3.6/dist-packages (from google-auth==1.2.0->google-cloud-storage==1.23.0->cartoframes) (3.1.1)
Requirement already satisfied: setuptools>=40.3.0 in /usr/local/lib/python3.6/dist-packages (from google-auth==1.2.0->google-cloud-storage==1.23.0->cartoframes) (47.1.1)
Requirement already satisfied: rsa<4.1,>=3.1.4 in /usr/local/lib/python3.6/dist-packages (from google-auth==1.2.0->google-cloud-storage==1.23.0->cartoframes) (4.0)
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.6/dist-packages (from google-auth==1.2.0->google-cloud-storage==1.23.0->cartoframes) (0.2.8)
Requirement already satisfied: googleapis-common-protos<2.0dev,>=1.6.0 in /usr/local/lib/python3.6/dist-packages (from google-api-core[grpc]<2.0.0dev,>=1.14.0->google-cloud-bigquery==1.22.0->cartoframes) (1.5.1)
Requirement already satisfied: grpcio<2.0dev,>=1.8.2; extra == "grpc" in /usr/local/lib/python3.6/dist-packages (from google-api-core[grpc]<2.0.0dev,>=1.14.0->google-cloud-bigquery==1.22.0->cartoframes) (1.29.0)
Requirement already satisfied: cligj>=0.5 in /usr/local/lib/python3.6/dist-packages (from Fiona>=0.6.0->geopandas<1.0,>=0.6.0->cartoframes) (0.5.0)
Requirement already satisfied: attrs>=17 in /usr/local/lib/python3.6/dist-packages (from Fiona>=0.6.0->geopandas<1.0,>=0.6.0->cartoframes) (19.3.0)
Requirement already satisfied: click-plugins>=1.0 in /usr/local/lib/python3.6/dist-packages (from Fiona>=0.6.0->geopandas<1.0,>=0.6.0->cartoframes) (1.1.1)
Requirement already satisfied: munch in /usr/local/lib/python3.6/dist-packages (from Fiona>=0.6.0->geopandas<1.0,>=0.6.0->cartoframes) (2.5.0)
Requirement already satisfied: click>=8.0 in /usr/local/lib/python3.6/dist-packages (from Fiona>=0.6.0->geopandas<1.0,>=0.6.0->cartoframes) (7.1.2)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.6/dist-packages (from requests>=2.7.0->carto<2.0,>=1.11.1->cartoframes) (2020.4.5.1)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.6/dist-packages (from requests>=2.7.0->carto<2.0,>=1.11.1->cartoframes) (3.0.4)
```

Why CARTOframes?

Reduce context switching



Powering end-to-end data science workflows

Explore

Clean, geocode, and visualize your data straight out of Jupyter notebooks.



Enrich

Access a wide range of datasets - all on standardized spatial aggregations to reduce your time to insight.



Analyze

Get insights from your data using our API and your own libraries, functions, and workflows.



Share

Once your analysis is done, add widgets and share your results.

1. Explore

→ Manage your data

Load a CSV file

Load data from a CSV file

Load a JSON file

Load data from a JSON file

Load a GeoJSON file

Load data from a GeoJSON file

Load a shapefile

Load data from a shapefile

Load a CARTO table

Load data from a CARTO table

Load a CARTO SQL query

Load data from a CARTO table using a SQL Query

Upload to CARTO

Upload data to CARTO

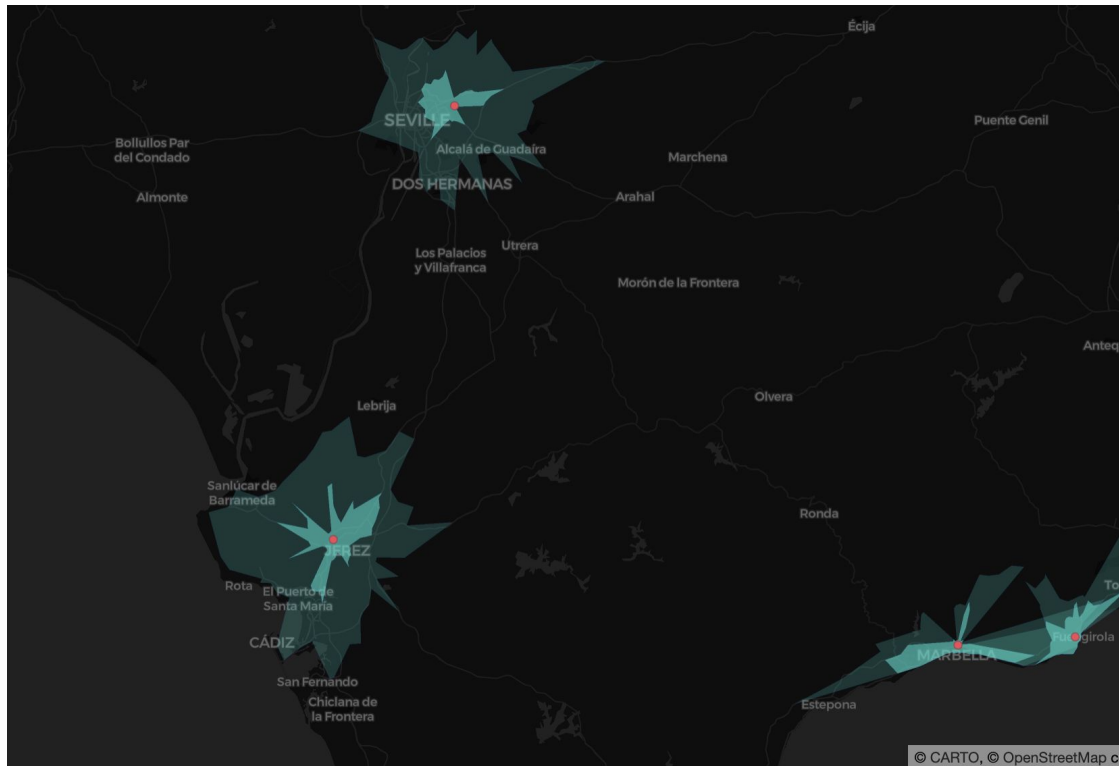
Change CARTO table privacy

Change the privacy of a CARTO table

1. Explore

→ Get your data ready

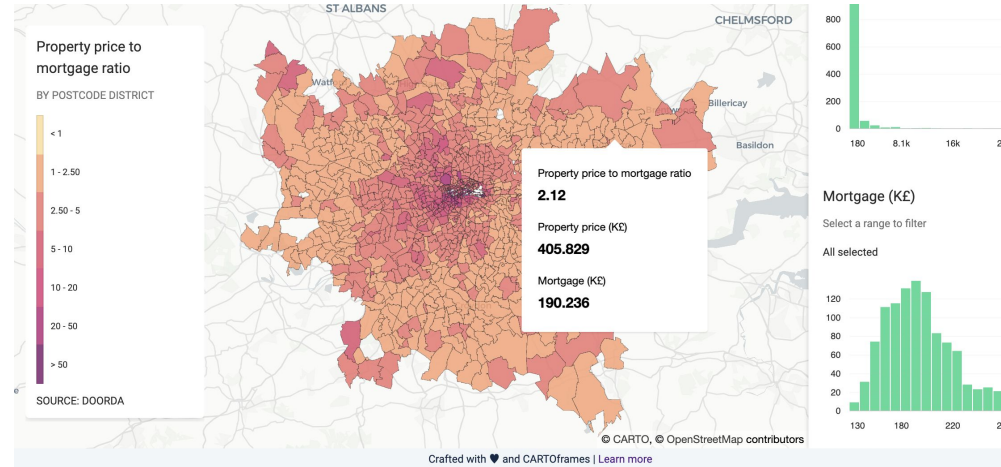
- Geocode large datasets in just one request
- Create isochrones for your points



1. Explore

→ Visualize








- Local data and hosted datasets
- Maps with multiple layers
- Styling for numerical and categorical variables
- Custom basemaps
- Legends, pop-ups, and widgets
- Layouts



2. Enrich

→ CARTO Data Observatory

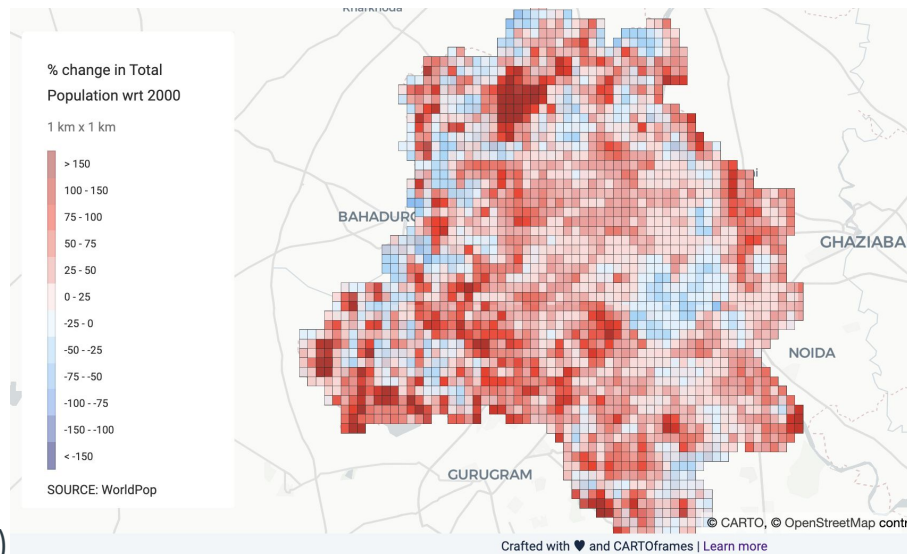
- Access to different location data streams on common geometries. Working with market-leaders, we bring together high-quality curated datasets to reduce the time to insight.
<https://carto.com/platform/spatial-data-catalog/>

 Financial	 Housing
 Human Mobility	 POI's
 Road Traffic	 Environmental
 Demographics	 Global Boundaries
 GeoSocial	 COVID-19

2. Enrich

→ Discover and enrich your data

- Direct access to our Data Observatory
- Open and premium datasets
- Discover the data you need
 - By category
 - By country
 - By geography
 - By provider
- Check stats about available datasets
- Request a dataset
- Enrich your dataframe (points or polygons)



We are working to help you *productize* your workflows and reduce your data preparation time, enabling you to focus on your analysis.

- ```
[390]: Map(Layer(nyc_geoms_enriched.reset_index(),
 style=color_category_style('cluster'),
 popup_hover=[popup_element('geoid'),
 popup_element('index'),
 popup_element('cluster')]))
```
- 
- cluster
- 15
  - 01
  - 03
  - 04
  - 17
  - 16
  - 08
  - 07
  - 10
  - 12
  - 05
  - Others
- PATERSON
- Clifton
- Bloomfield
- East Orange
- Union City
- NEWARK
- IRVINGTON
- Union
- ELIZABETH
- Bayonne
- Jersey City
- NEW YORK
- Mount Vernon
- Glenn Cove
- Minerva
- Garden City
- Hempstead
- Valley Stream
- Freeport
- East Rockaway
- Long Beach
- Perth Amboy
- Woodbridge
- Piscataway
- Plainfield
- CARTO, OpenStreetMap

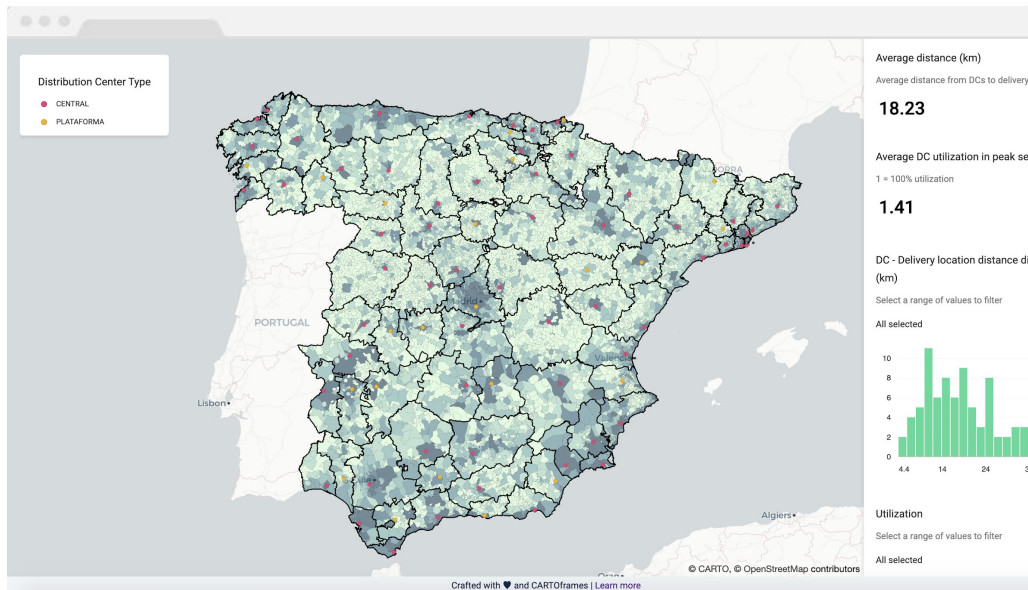
## 4. Share

### → Prepare your analysis output

- Make your analysis easy to consume by others in your organization by adding widgets: histogram, category, animation control, time series.

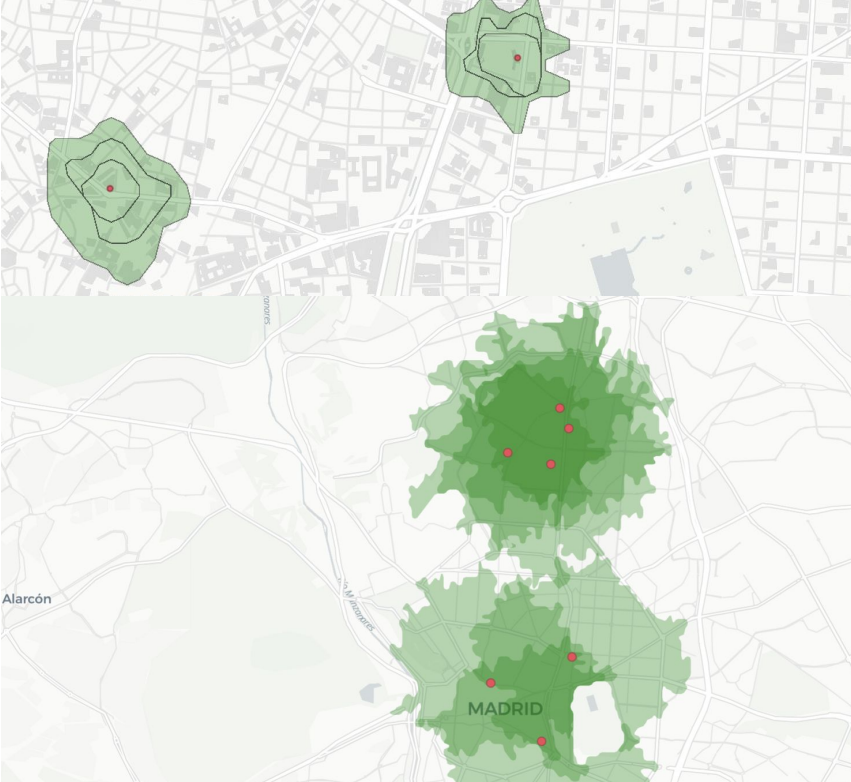
### → Publish

- Publish your map to CARTO and get the shareable link as response.



# Hands-on with CARTO

## Practical Spatial Data Science in Python



|   | the_geom                  | address     |   |
|---|---------------------------|-------------|---|
| 0 | POINT (-3.70588 40.42049) | Gran Vía 46 | M |
| 1 | POINT (-5.98312 37.35547) | Ebro 1      | S |



# Hands-on with CARTO

## Practical Spatial Data Science in Python

*Using CARTO tech stack, in this session we will go through a step by step demo using Jupyter notebooks, from data exploration, to external data discovery and augmentation, to model formulation and results.*

- **Site selection:** where should Starbucks open new coffee shops in Long Island, NY? In this demo we will go through a typical site selection use case, from modelling the revenues of the existing stores as a function of socioeconomic covariates, to predicting the potential revenues in new locations.
- **Logistic spatial optimization:** where should a parcel delivery company locate their distribution and fulfilment centers? What areas should they service? In this demo we will go through a supply chain network optimization use case, from analysing past data to identify spatio-temporal patterns to building an optimization model to analyze and quantify the impact of changes in the current network.

# Thank you for listening!

Request a demo at [CARTO.com](https://CARTO.com)

**Giulia Carella**

Data Scientist | [giulia@carto.com](mailto:giulia@carto.com)

**Miguel Álvarez**

Data Scientist | [marvarez@carto.com](mailto:marvarez@carto.com)

