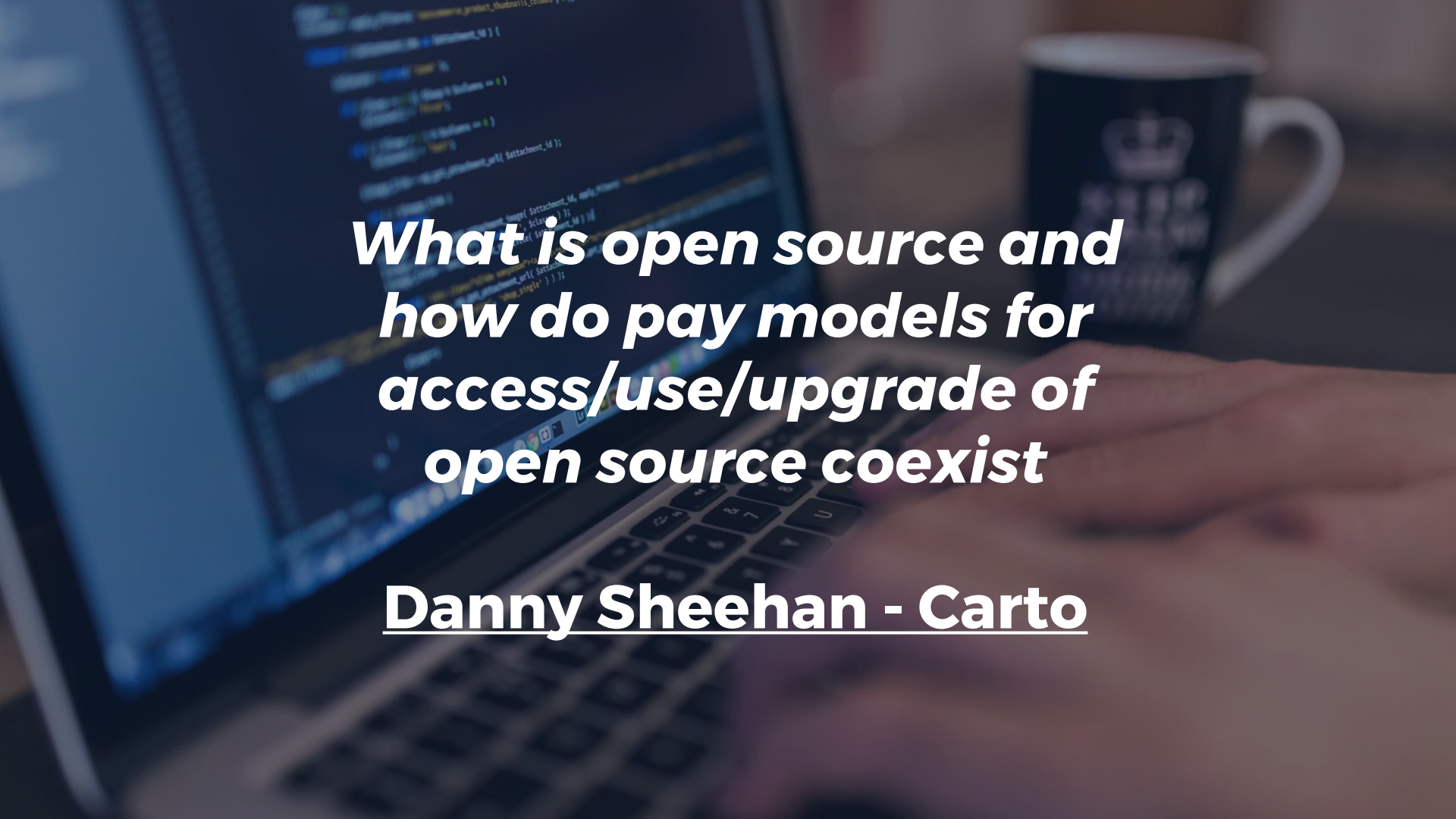

International
Cartographic
Association



What is open source and how do pay models for access/use/upgrade of open source coexist

Danny Sheehan - Carto

ABOUT

- **Solutions Engineer at Carto**
  - user of FOSS4G tools/code/platforms
  - 14+ years other Desktop GIS Software product(s), geoprocessing and spatial analysis.

- **Email:** danny@carto.com

ABOUT

- Filling in for
  - MAMATA AKELLA - SENIOR CARTOGRAPHER at CARTO

ABOUT

- **Education**
  - **Columbia University**
    - Certificate in Professional Achievement in Data Sciences
  - **University at Buffalo**
    - Masters in Geography/Specialization in GIS
  - **Geneseo**
    - Bachelors in Geography

ABOUT

- **Prior experience**
  - **Columbia University** - 6.5 years Senior Research Scientist/GIS Analyst at Department of Epidemiology Mailman School of Public Health - Built Environment & Health Project (beh.columbia.edu),
    - Affiliations with Institute of Social, Economic Research & Policy, Fu School of Engineering
    - Teaching at Dept. of Epidemiology, Quantitative Methods for Social Sciences, Barnard College
  - **AECOM** - 3.5 years lead GIS Analyst for FEIS/FOEIS for US Navy projects, Dahlgren, USWTR, Guam, I-287/Tappan Zee Bridge

ABOUT

- **Journal publications**
  - American Journal of Preventive Medicine
  - American Journal of Public Health
  - European Journal of Gastroenterology & Hepatology
  - Journal of Exposure Science and Environmental Epidemiology
  - Cancer Causes & Control
  - AIDS and Behavior
  - Journal of Maps
  - Etc,...

- **Google School Link**

[https://scholar.google.com/citations?user=K6iTYsUAAAAJ&hl=en]

ABOUT

- Current Role as Solutions Engineer at Carto:
  - On-Prem Install
  - Leveraging our APIs and the Carto stack
  - Setting up Database Connectors
  - Building middleware!!!
  - ...(continue)

ABOUT

- Current Role as Solutions Engineer at Carto:
  - Bug Reporting and Fixes
  - Geoprocessing with PostGIS
  - Strategizing analytical methods
  - Coordinating between Carto's Data Science teams and our clients' Data Science needs
  - Getting the Feature Request you ask for into the Carto Platform

What is Open Source?

What is Open Source?

What is Open Source?

Open-source software (OSS) is computer software with its source code made available with a license in which the copyright holder provides the rights to study, change, and distribute the software to anyone and for any purpose. [1]

What is Open Source?

Open-source software (OSS) is computer software with its source code made available with a license in which the copyright holder provides the rights to study, change, and distribute the software to anyone and for any purpose. [1]

- Wikipedia:

- 1. St. Laurent, Andrew M. (2008). Understanding Open Source and Free Software Licensing. O'Reilly Media. p. 4. ISBN 9780596553951.

What is Open Data?

What is Open Data?

Open data is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control. [1]

What is Open Data?

Open data is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control. [1]

- Wikipedia:

- 1. Auer, S. R.; Bizer, C.; Kobilarov, G.; Lehmann, J.; Cyganiak, R.; Ives, Z. (2007). "DBpedia: A Nucleus for a Web of Open Data". The Semantic Web. Lecture Notes in Computer Science. 4825. p. 722. ISBN 978-3-540-76297-3. doi:10.1007/978-3-540-76298-0_52.

OPEN SOURCE

TYPES OF OPEN SOURCE

1. **Desktop vs. Server architecture**

OPEN SOURCE

TYPES OF OPEN SOURCE

1. Desktop vs. Server architecture

2. **Installer scripts/UI vs. command line approach/dependencies are up to the installer.**

OPEN SOURCE

TYPES OF OPEN SOURCE

1. Desktop vs. Server architecture
2. Installer scripts/UI vs. command line approach/dependencies are up to the installer.
3. **Add-on features/plugins only available at paid level vs. everything is open.**

RESEARCH

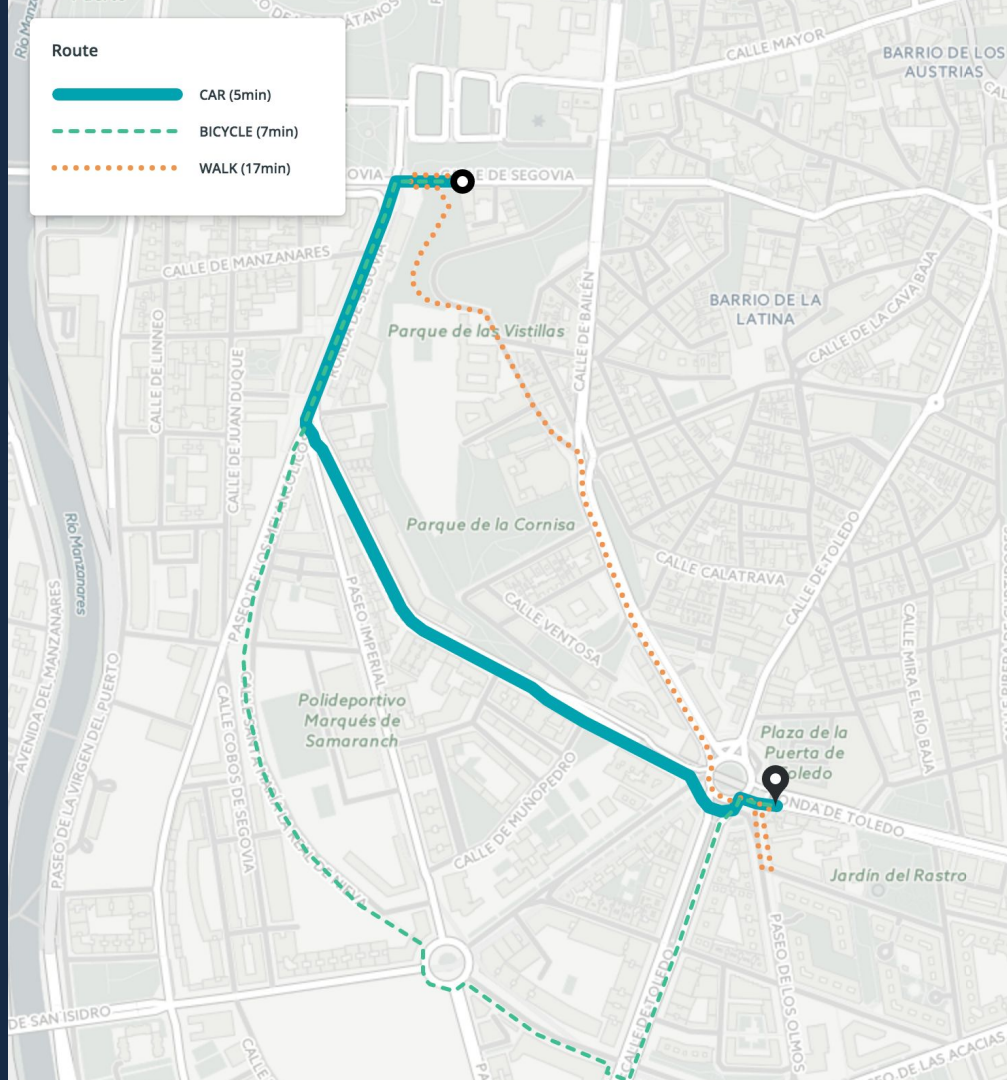
Spatial Machine Learning

Leverage features in your own data and location context to train gradient boosted random forests and predict new markets

LOCATION DATA SERVICES

Exceptional data, technology, and cartography to power your applications.

A set of Location Data Services including vector basemaps, geocoding, routing, and demographic data augmentation services to help extract the full potential of location intelligence.



RESEARCH

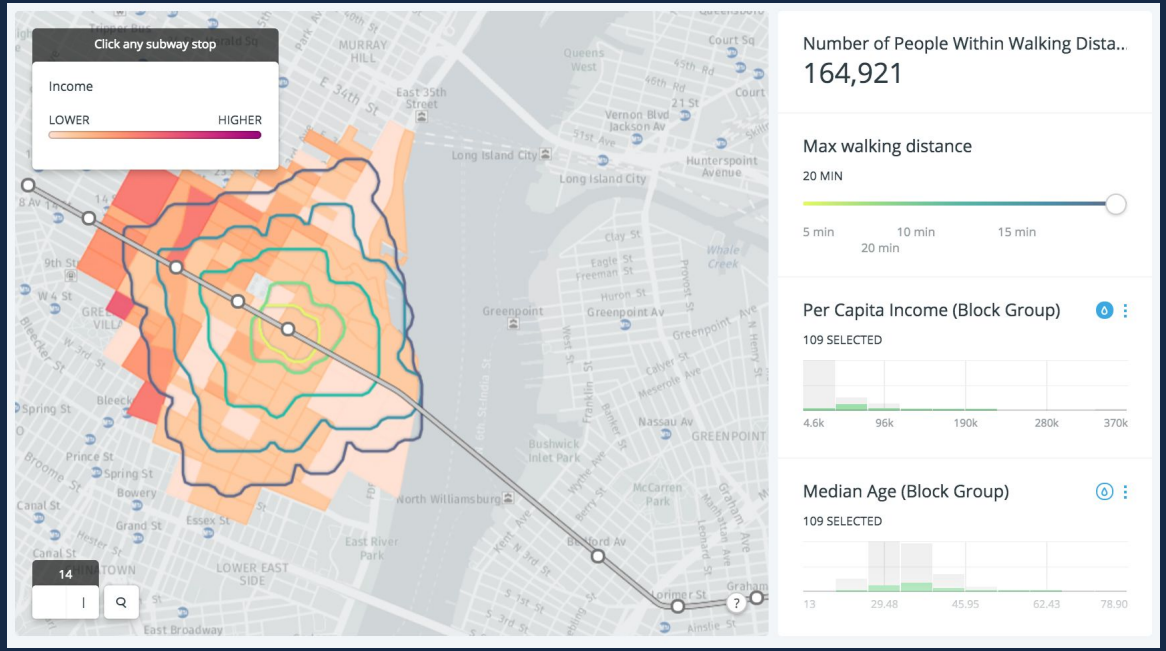
Data Observatory

Measurements

Access demographic, economic, and both numeric and categorical measurements at places of interest.

Segmentation

Classifications are built up from leading population data and provide a comprehensive view of where your potential customers, voters, or supporters exist.

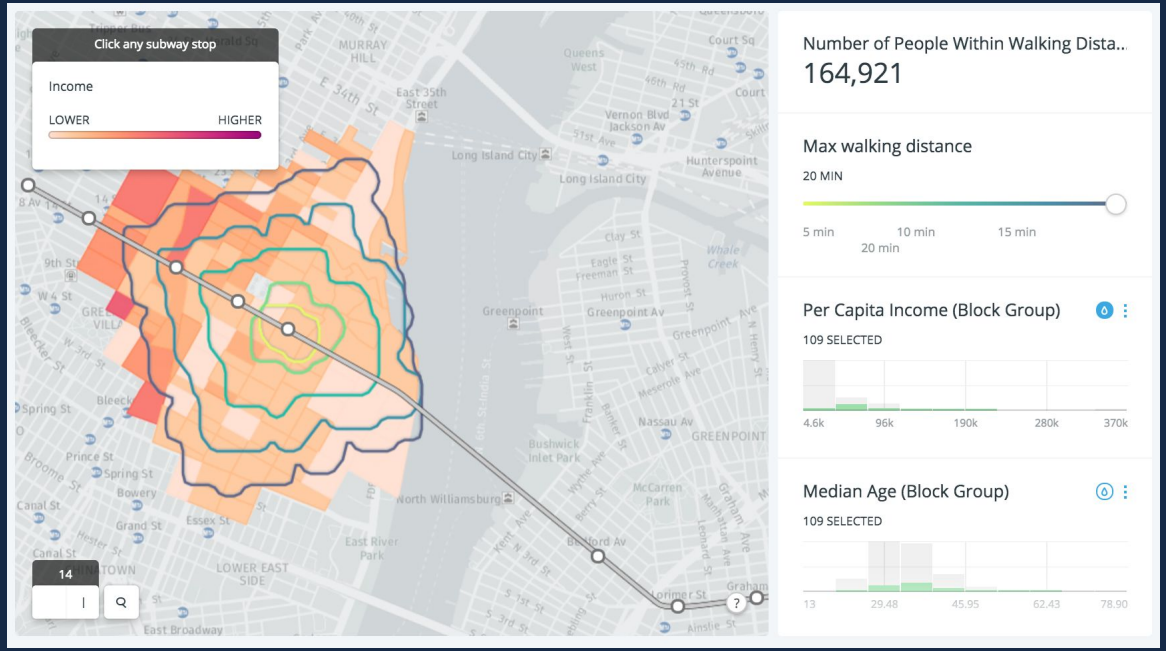


RESEARCH

Data Observatory

[Big Metadata]

https://cartodb.github.io/bigmetadata/



OPEN SOURCE

TYPES OF OPEN SOURCE

1. Desktop vs. Server architecture
2. Installer scripts/UI vs. command line approach/dependencies are up to the installer.
3. Add-on features/plugins only available at paid level vs. everything is open.

4. License types of Open Source

[https://opensource.org/licenses]

OPEN SOURCE

OPEN SOURCE ADVANTAGES

1. **Quickly fixable... code often in public repo.**

OPEN SOURCE

OPEN SOURCE ADVANTAGES

1. Quickly fixable... code often in public repo.

2. More secure... more eyes on the code that can identify security holes/flaws.

OPEN SOURCE

OPEN SOURCE ADVANTAGES

1. Quickly fixable... more eyes on the code.
2. More secure... more eyes on the code that can identify security holes/flaws.
3. Active collaboration and community. Which breeds innovation, health debates.


How do
pay models for
access/use/upgrade
of Open Source
coexist?

Open Source Models Examples:

From where else than StackOverflow:

How do open source developers make money?

https://stackoverflow.com/questions/8159049/how-do-open-source-developers-make-money

Sorry for the duplicate question. My bad. – jamesfzhang Nov 16 '11 at 21:37

1 There are many ways... donations, offering support... I did not make a response since i c too vague. I would like to view a detailled list of most common and effective ways to mak Nov 16 '11 at 21:37

Many options, charge for support, dual license, donations, charge for the development o Nov 16 '11 at 21:38

add a comment

1 Answer

activ

24

I can think of four ways:

1) The open source project is a side project and have a job as their source of inc

2) They have a corporate sponsor that pays the devs a salary

3) They develop their OSS project and then sell services based on that project (MySQL for free, and do consulting installing and maintaining MySQL for compar

4) They develop a base version of the OSS project, and then develop and sell c for it

share improve this answer

answered N

Gab

3,58

3 I would like to add this line too. " They become famous and after that they can introduce and services " – UnKnown Aug 26 '16 at 12:52

OPEN SOURCE MODELS

1. **The open source project is a side project and have developers have other jobs as their source of income**

OPEN SOURCE

OPEN SOURCE MODELS

1. The open source project is a side project and have developers have other jobs as their source of income

2. **Corporate Sponsorship pays developers and company**

OPEN SOURCE

OPEN SOURCE MODELS

1. The open source project is a side project and have developers have other jobs as their source of income

2. Corporate Sponsorship pays developers and company

3. **OSS project is developed by company and sold services based on that project**
    - (ie, I give away MySQL for free, and do consulting installing and maintaining MySQL for companies)

OPEN SOURCE

OPEN SOURCE MODELS

1. **The open source project is a side project and have developers have other jobs as their source of income**

2. **Corporate Sponsorship pays developers and company**

3. **OSS project is developed by company and sold services based on that project**
    - (ie, I give away MySQL for free, and do consulting installing and maintaining MySQL for companies)

4. **They develop a base version of the OSS project, and then develop and sell commercial add-ons for it**

At Carto seems like we have a combo of all 4

Sponsorship for features from companies, selling of additional or hosted features/storage, we consult, etc. and some contributions are made by our dev. community/our dependencies also have a mix of all four (4).

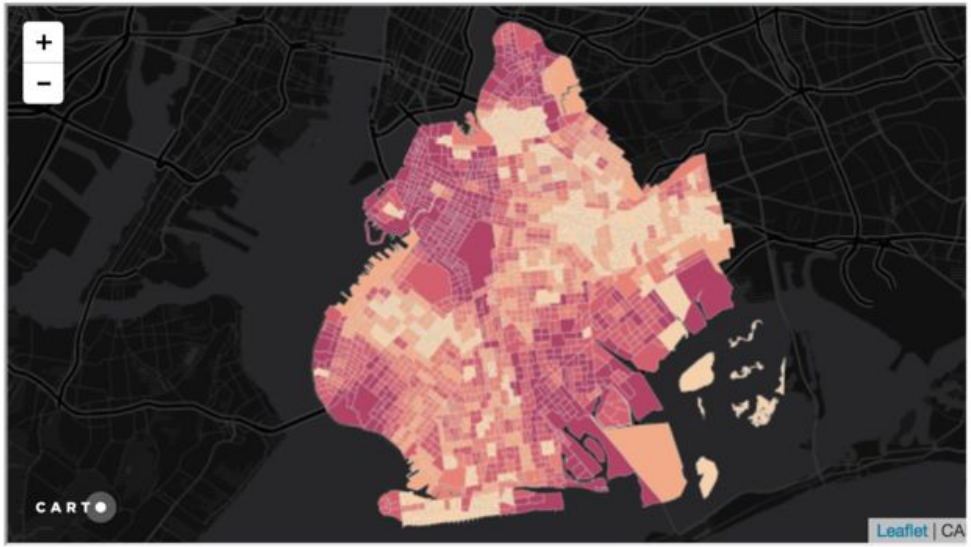
```
df = pd.read_carto(username=USERNAME,
                   api_key=APIKEY,
                   tablename='brooklyn_poverty_notnull_geoms')
df.head()
```

Out[1]:

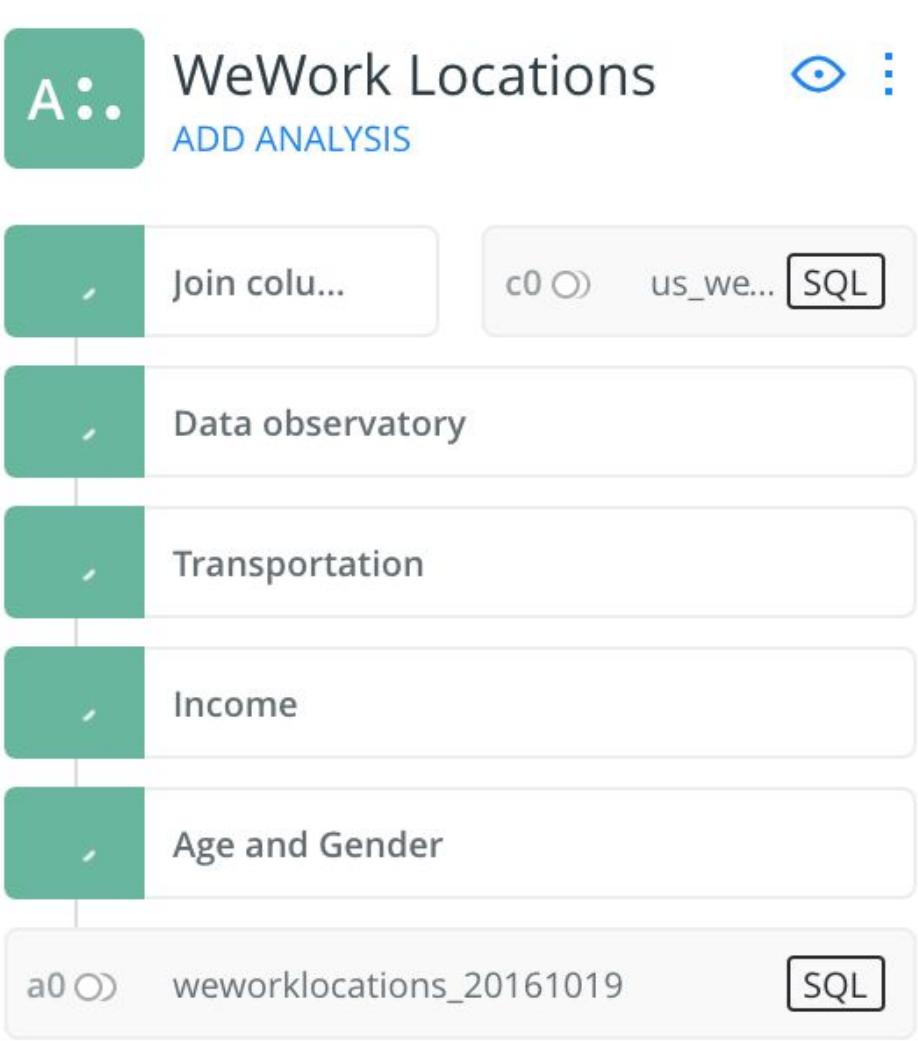
	geoid	high_school_diploma_2011_2015	median_income_2011_2015	p
cartodb_id				
153	360470018001	0.219213	36414.226756	0.
946	360470352001	0.258464	17156.148349	0.
15	360470514002	0.264986	45746.135380	0.
37	360470137001	0.105692	100039.579057	0.
1	360470043001	0.059511	112115.160834	0.

In [2]:
```
df.carto_map(interactive=True, stylecol='median_income_2011_2015')
```

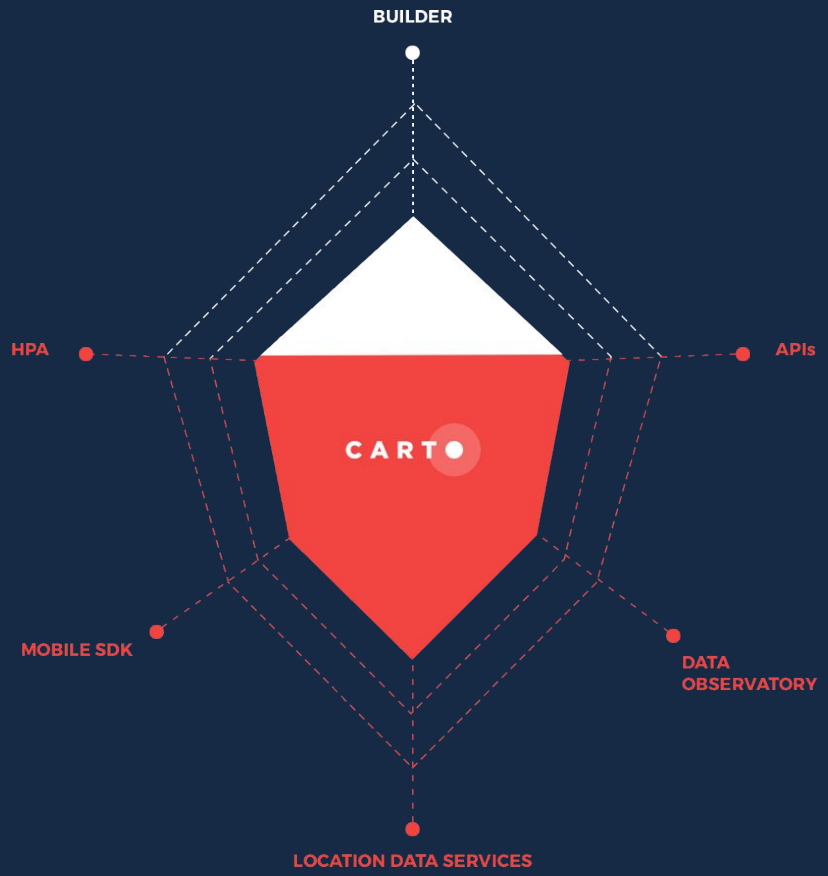
Out[2]:



We're Open Source built on top of other Open Source



Access/Use,
&
Upgrade



OPEN SOURCE

Access/Use

1. **Open Source version, self installed.**

OPEN SOURCE

Access/Use

1. Open Source version, self installed.
2. **Hosted Cloud version (SaaS)**

OPEN SOURCE

Access/Use

1. Open Source version, self installed.
2. Hosted Cloud version (SaaS)
3. On-premise installation licensed and sold by Carto.
    - Could be either in clients cloud environment, local/virtual on physical server.

OPEN SOURCE

Upgrade

1. **Storage - clients pay for additional storage.**

OPEN SOURCE

Upgrade

1. **Storage - clients pay for additional storage.**

2. **Speed/Processing Power**
    - Handle Larger Data Volumes
    - Handle Complex Processes
        - NP Hard, etc.
    - Handle Processes faster
    - Upgrades to core technology/dependencies, i.e. Postgres

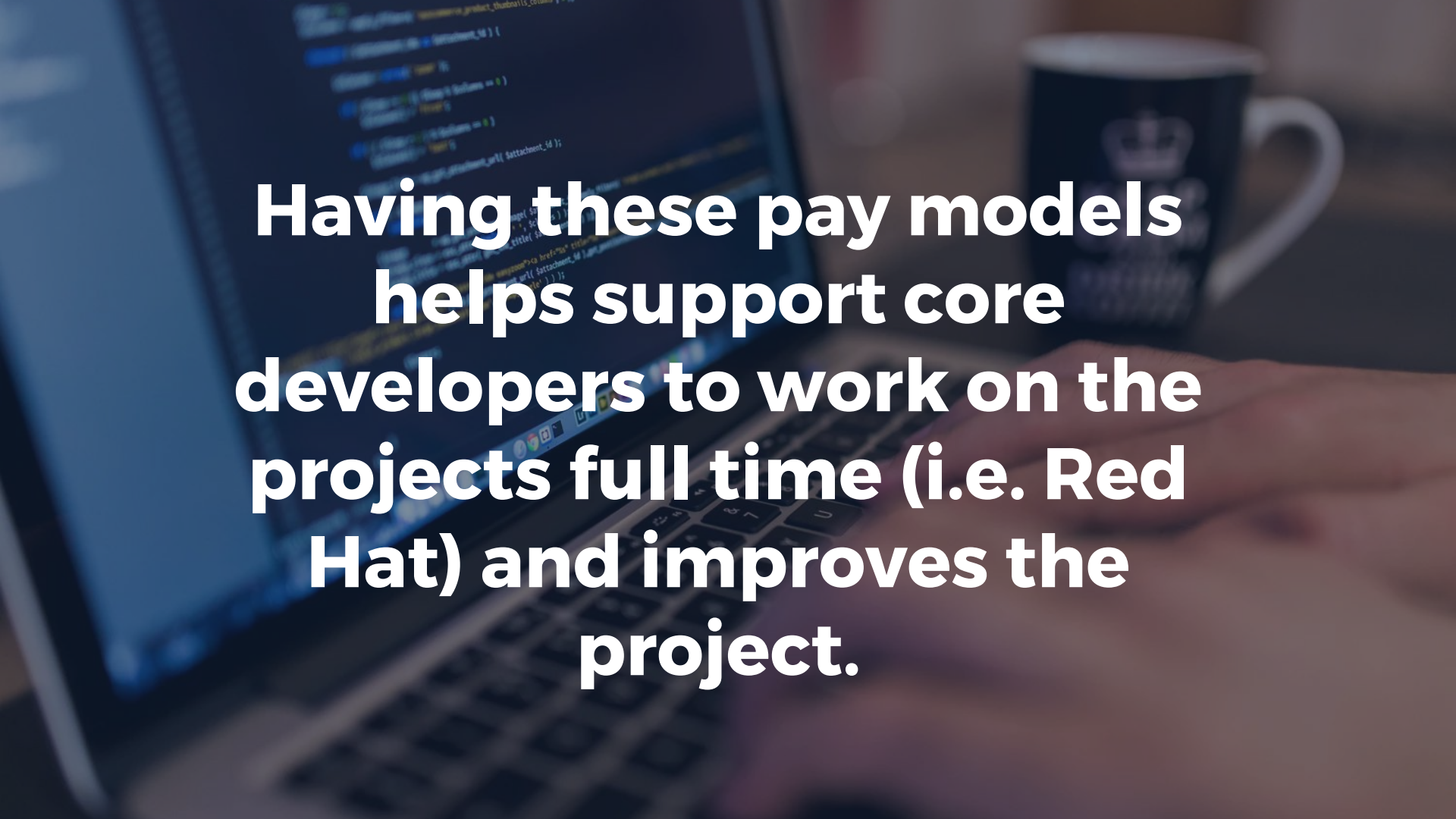
OPEN SOURCE

Upgrade

1. **Storage - clients pay for additional storage.**

2. **Speed/Processing Power**

3. **Additional Features**
    - Added analysis methods
    - Custom tools
    - Sponsored features


Having these pay models helps support core developers to work on the projects full time (i.e. Red Hat) and improves the project.

At Carto:

Our core development and Product team, Data Science/Research and Solutions team(s) are all individuals that firmly believe in Open Source as a philosophy.

Get a list of all the current import jobs

```
from carto.file_import import FileImportJobManager

file_import_manager = FileImportJobManager(auth_client)
file_imports = file_import_manager.all()
```

Get all the datasets

```
from carto.datasets import DatasetManager

dataset_manager = DatasetManager(auth_client)
datasets = dataset_manager.all()
```

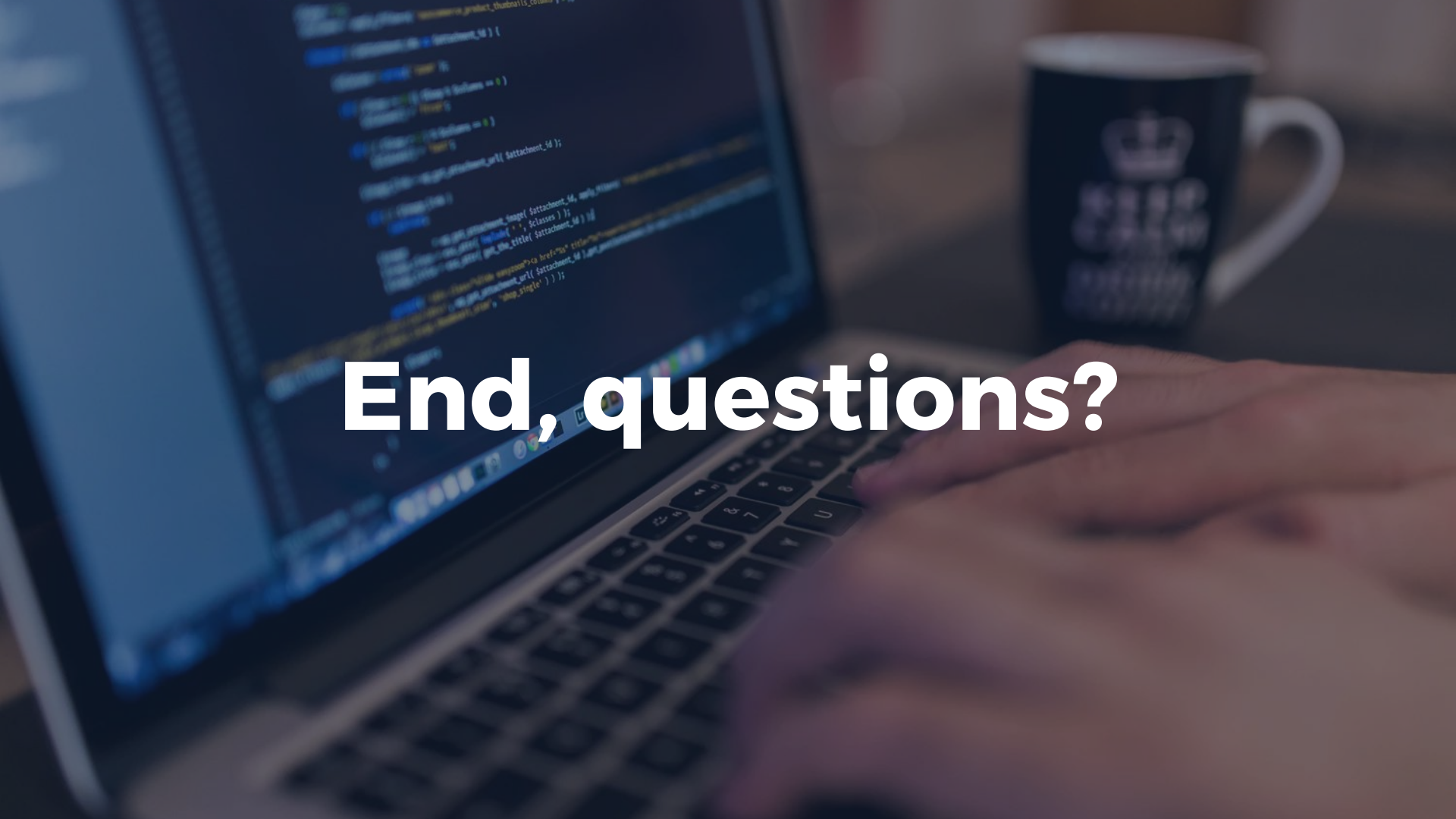
Get a specific dataset

```
from carto.datasets import DatasetManager

# write here the ID of the dataset to retrieve
DATASET_ID = ""

dataset_manager = DatasetManager(auth_client)
dataset = dataset_manager.get(DATASET_ID)
```

Carto Install:

Install:

https://github.com/CartoDB/cartodb/blob/master/doc/manual/source/install.rst


End, questions?

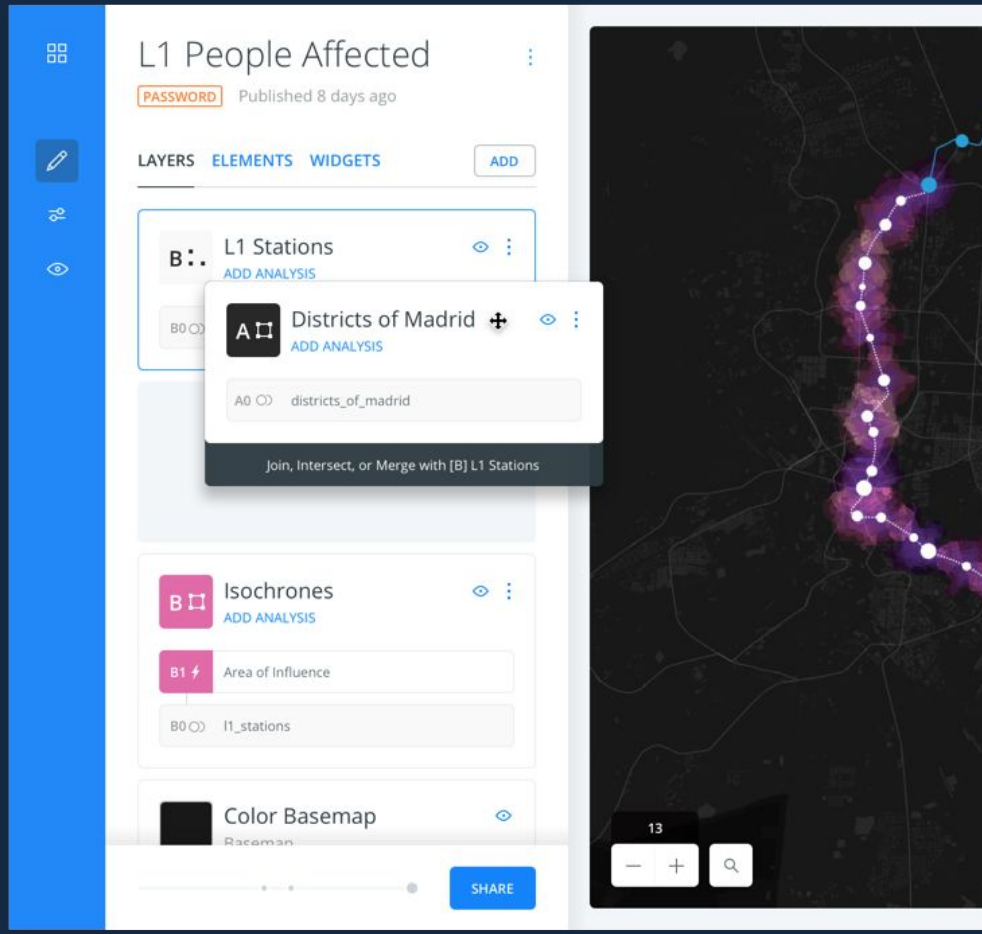

Contact:
@nygeog
danny@carto.com

BUILDER

LOCATION INTELLIGENCE FINALLY INTUITIVE

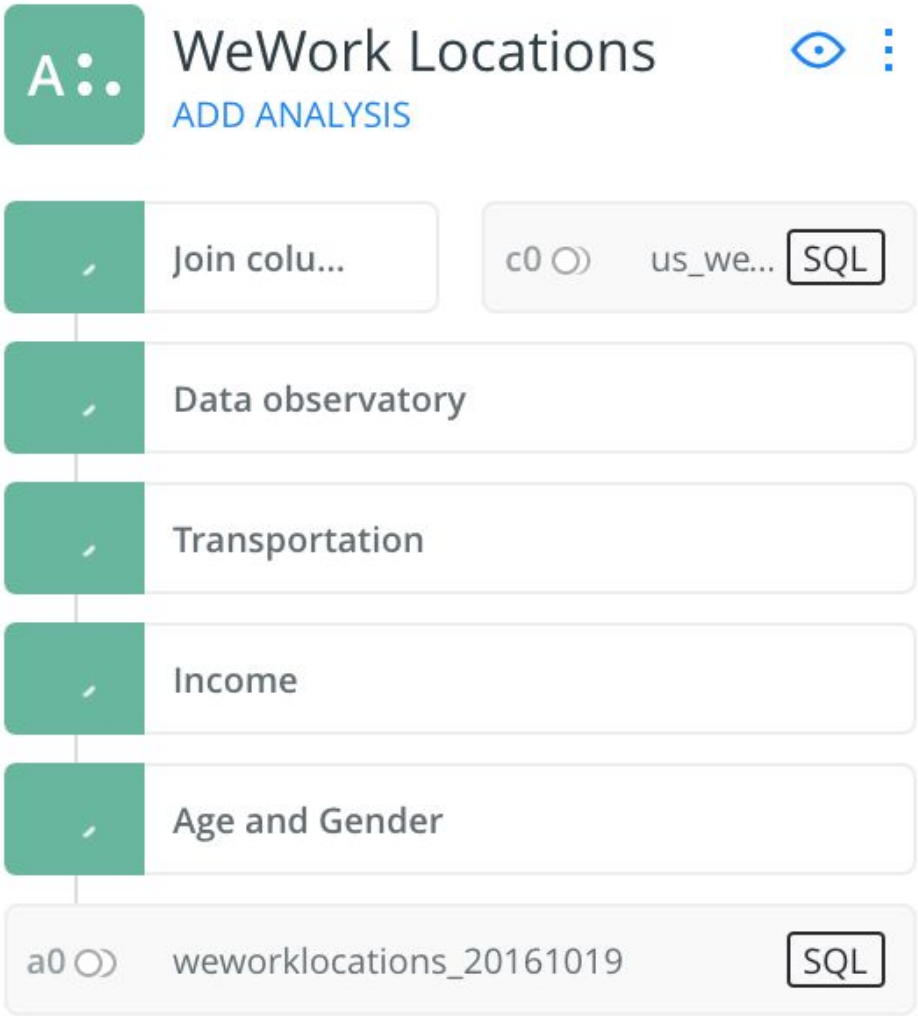
A web-based drag and drop analysis tool for business users and analysts to discover and predict key insights from location data.

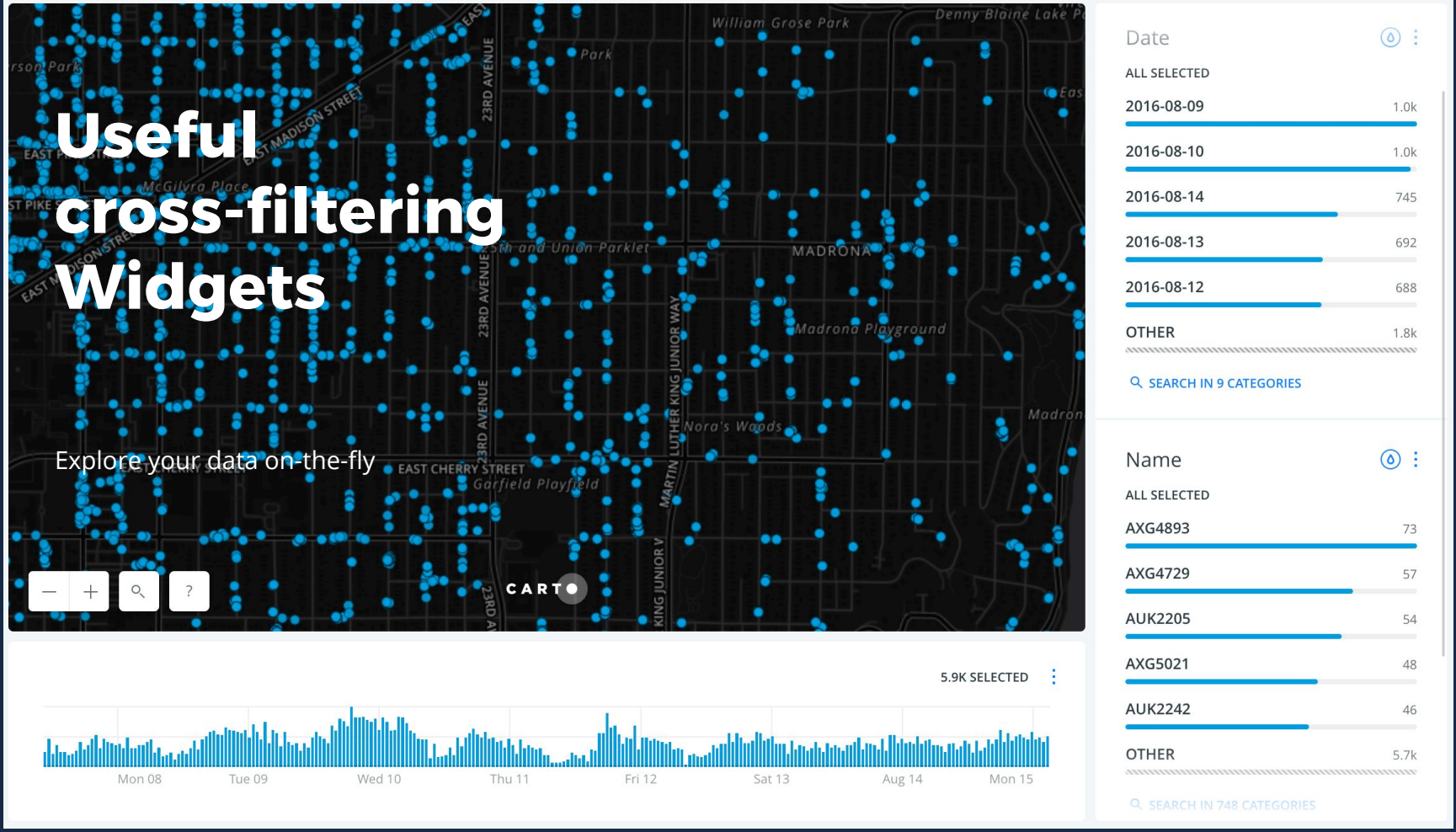
CARTO Builder unleashes the power of location intelligence with self-service, actionable dashboards you can share across your whole organization.



Intuitive Analysis pipelines

Easily generate powerful analysis pipelines that combine GIS and Data Science workflows




Useful cross-filtering Widgets
Explore your data on-the-fly
William Grose Park
Denny Blaine Lake P
Park
23RD AVENUE
EAST MADISON STREET
McGilvra Place
MADRONA
Madrona Playground
Nora's Woods
EAST CHERRY STREET
Garfield Playfield
MARTIN LUTHER KING JUNIOR WAY
CARTO
Date
ALL SELECTED
2016-08-09 1.0k
2016-08-10 1.0k
2016-08-14 745
2016-08-13 692
2016-08-12 688
OTHER 1.8k
SEARCH IN 9 CATEGORIES
Name
ALL SELECTED
AXG4893 73
AXG4729 57
AUK2205 54
AXG5021 48
AUK2242 46
OTHER 5.7k
SEARCH IN 748 CATEGORIES
5.9K SELECTED
Mon 08
Tue 09
Wed 10
Thu 11
Fri 12
Sat 13
Aug 14
Mon 15

Work in your own Data Science world

CARTO offers **CartoFrames** a convenient and intuitive Python module that fits into your existing Data Science Workflows and our newly reconfigured **Carto Python SDK** (new version 1.0.0) that allows you to easily interact with the Carto Engine in your Python environment.

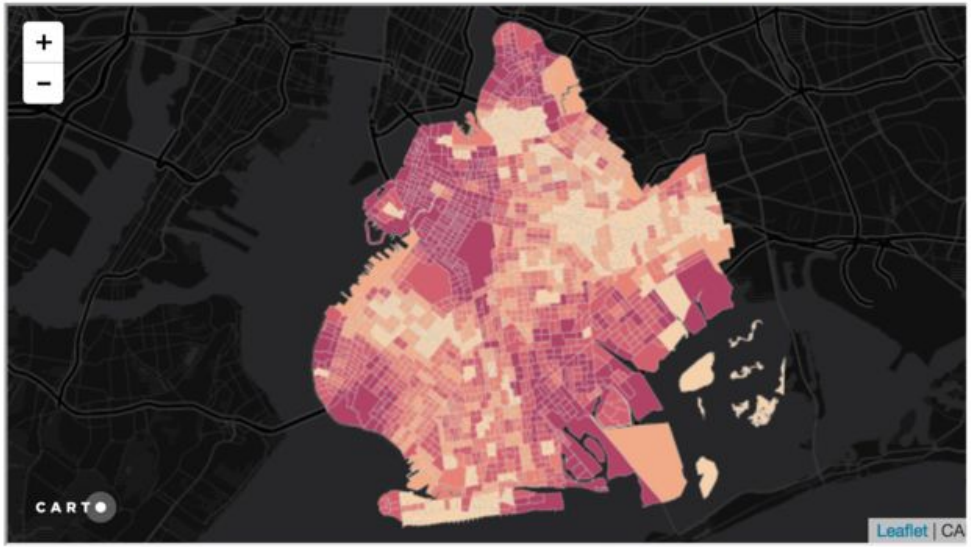
```
df = pd.read_carto(username=USERNAME,
                   api_key=APIKEY,
                   tablename='brooklyn_poverty_notnull_geoms')
df.head()
```

Out[1]:

	geoid	high_school_diploma_2011_2015	median_income_2011_2015	p
cartodb_id				
153	360470018001	0.219213	36414.226756	0.
946	360470352001	0.258464	17156.148349	0.
15	360470514002	0.264986	45746.135380	0.
37	360470137001	0.105692	100039.579057	0.
1	360470043001	0.059511	112115.160834	0.

In [2]:
```
df.carto_map(interactive=True, stylecol='median_income_2011_2015')
```

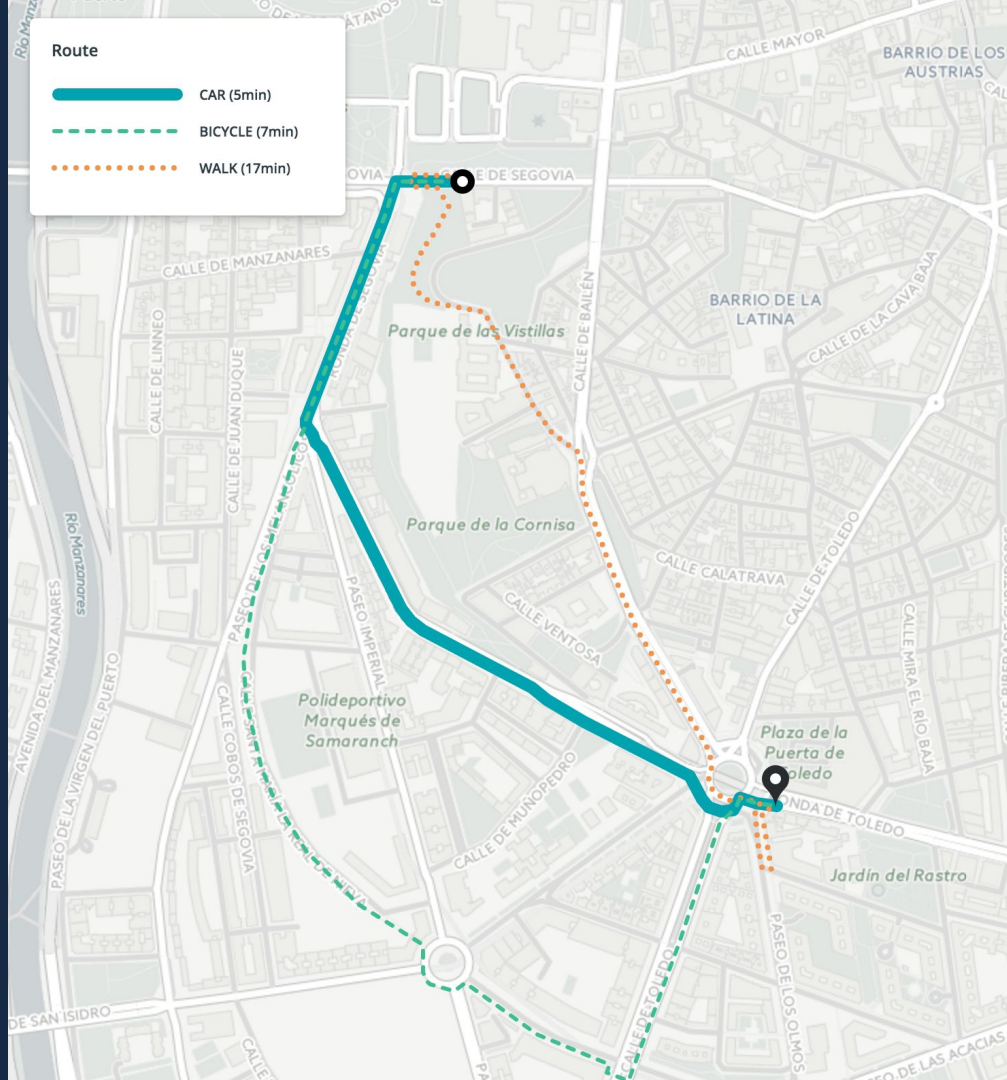
Out[2]:



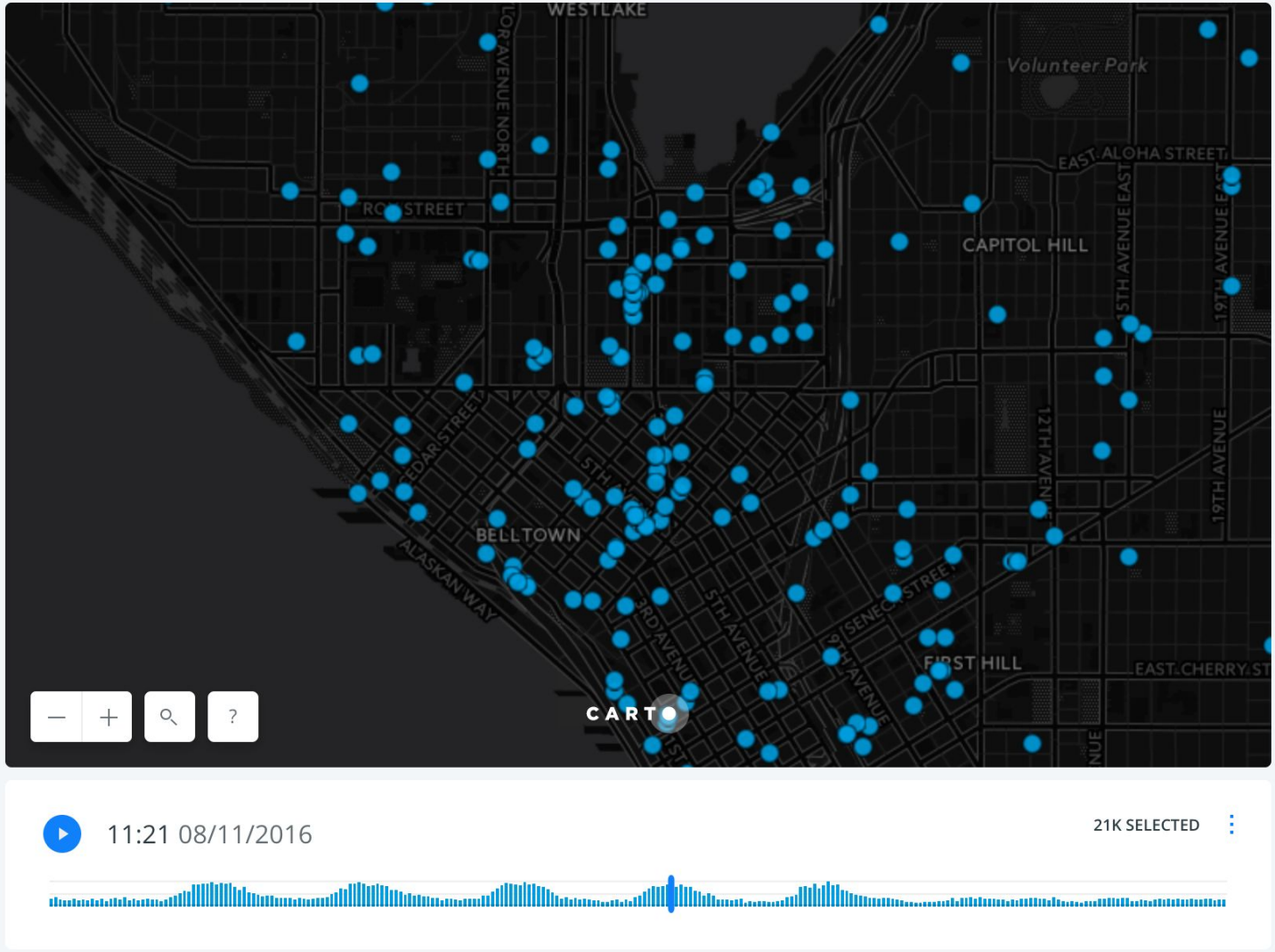
LOCATION DATA SERVICES

Exceptional data, technology, and cartography to power your applications.

A set of Location Data Services including vector basemaps, geocoding, routing, and demographic data augmentation services to help extract the full potential of location intelligence.



Car2Go Examples: Animated Time-Series



Car2Go Examples: Car2Go Locations Seattle

5.4K SELECTED

Date

ALL SELECTED

2016-08-09	933
2016-08-10	908
2016-08-14	686
2016-08-12	638
2016-08-08	609
OTHER	1.6k

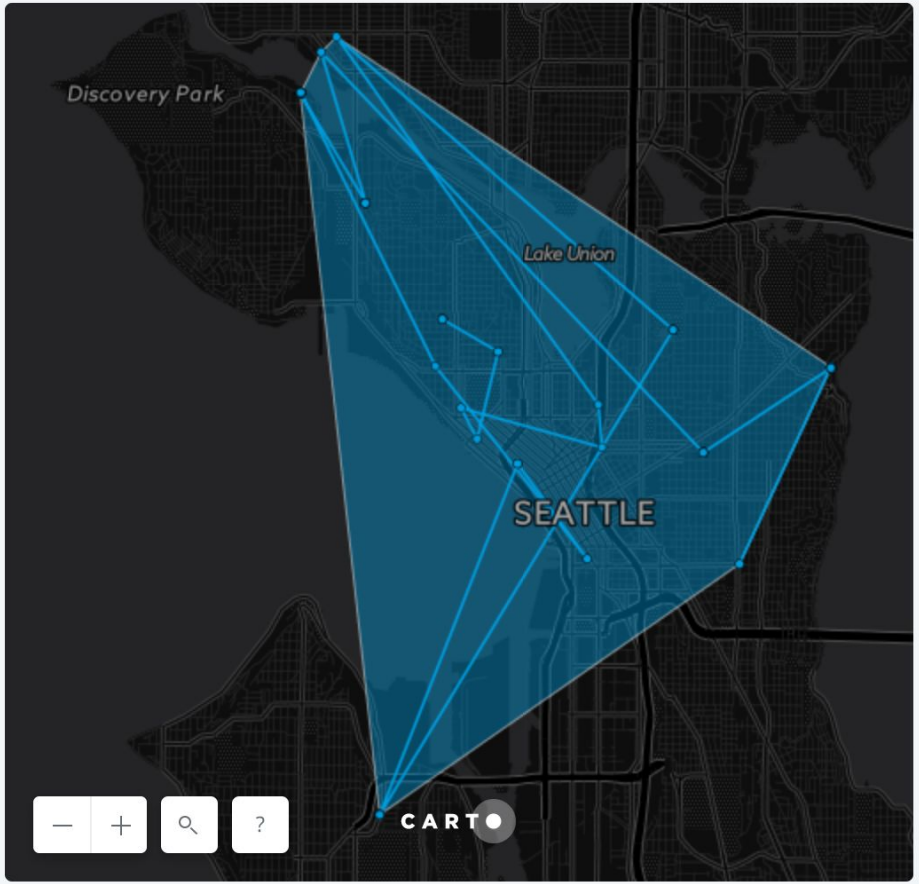
SEARCH IN 9 CATEGORIES

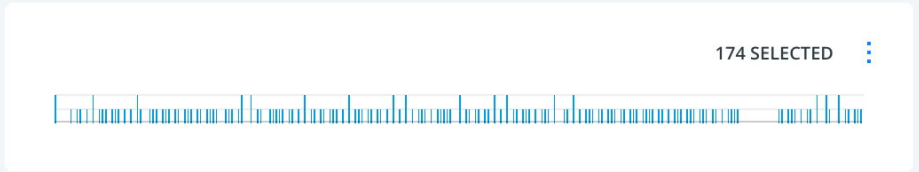
Name

ALL SELECTED

AXG4893	73
AXG4729	57
AXG5021	48
AXG5744	46
AUK2242	46

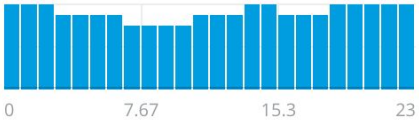
Car2Go Examples: Car2Go Locations Seattle with Convex Hulls and Connected Lines





Hour

174 SELECTED



Count

174

Date

ALL SELECTED

2016-08-10	24
2016-08-08	24
2016-08-13	24
2016-08-11	24
2016-08-12	24
OTHER	54

SEARCH IN 9 CATEGORIES

Carto | Python

Python is super powerful. Thus, we have our Carto SDK for Carto APIs, new version 1.0.0!

Also, for fans of Python Pandas (R-like operations and data structures) we have newly released CartoFrames.

Get a list of all the current import jobs

```
from carto.file_import import FileImportJobManager

file_import_manager = FileImportJobManager(auth_client)
file_imports = file_import_manager.all()
```

Get all the datasets

```
from carto.datasets import DatasetManager

dataset_manager = DatasetManager(auth_client)
datasets = dataset_manager.all()
```

Get a specific dataset

```
from carto.datasets import DatasetManager

# write here the ID of the dataset to retrieve
DATASET_ID = ""

dataset_manager = DatasetManager(auth_client)
dataset = dataset_manager.get(DATASET_ID)
```

Carto | Python SDK

Carto-Python is a full, backwards incompatible rewrite of the deprecated cartodb-python SDK. Since the initial rewrite, carto-python has been loaded with a lot of new features, not present in old cartodb-python.

Get a list of all the current import jobs

```
from carto.file_import import FileImportJobManager

file_import_manager = FileImportJobManager(auth_client)
file_imports = file_import_manager.all()
```

Get all the datasets

```
from carto.datasets import DatasetManager

dataset_manager = DatasetManager(auth_client)
datasets = dataset_manager.all()
```

Get a specific dataset

```
from carto.datasets import DatasetManager

# write here the ID of the dataset to retrieve
DATASET_ID = ""

dataset_manager = DatasetManager(auth_client)
dataset = dataset_manager.get(DATASET_ID)
```

Carto | CartoFrames

A Python package for integrating CARTO maps and services into data science workflows.

Examples: https://github.com/CartoDB/cartoframes/tree/master/examples

- https://github.com/CartoDB/cartoframes/blob/master/examples/Basic%20Usage.ipynb
- https://github.com/CartoDB/cartoframes/blob/master/examples/cartoframes-and-dask.ipynb
- https://github.com/CartoDB/cartoframes/blob/master/examples/Shapely%2C%20Geopandas%2C%20and%20Cartoframes.ipynb

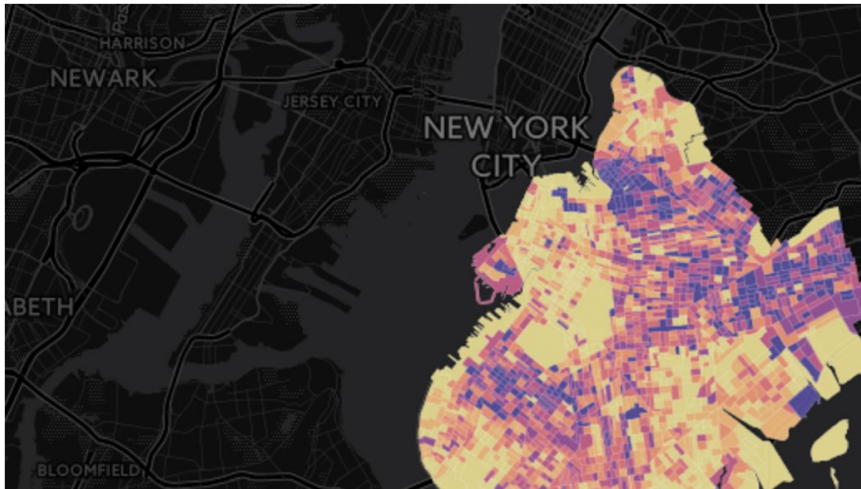
In [2]:
```
# Get a CARTO table as a pandas DataFrame
df = cc.read('brooklyn_poverty')
df.head()
```

Out[2]:

	the_geom	the_g
cartodb_id		
2	0103000020E6100000010000000B0000006D3A02B85982...	0103
3	0103000020E610000001000000090000006AA164726A80...	0103
4	0103000020E6100000010000000E000000591822A7AF7F...	0103
6	0103000020E610000001000000080000000F5BD86E0B87E...	0103
7	0103000020E6100000010000000C000000E6CDE15AED7E...	0103

In [3]:
```
from cartoframes import Layer, styling
l = Layer('brooklyn_poverty', color={'column': 'poverty_per_p
cc.map(layers=l, interactive=False)
```

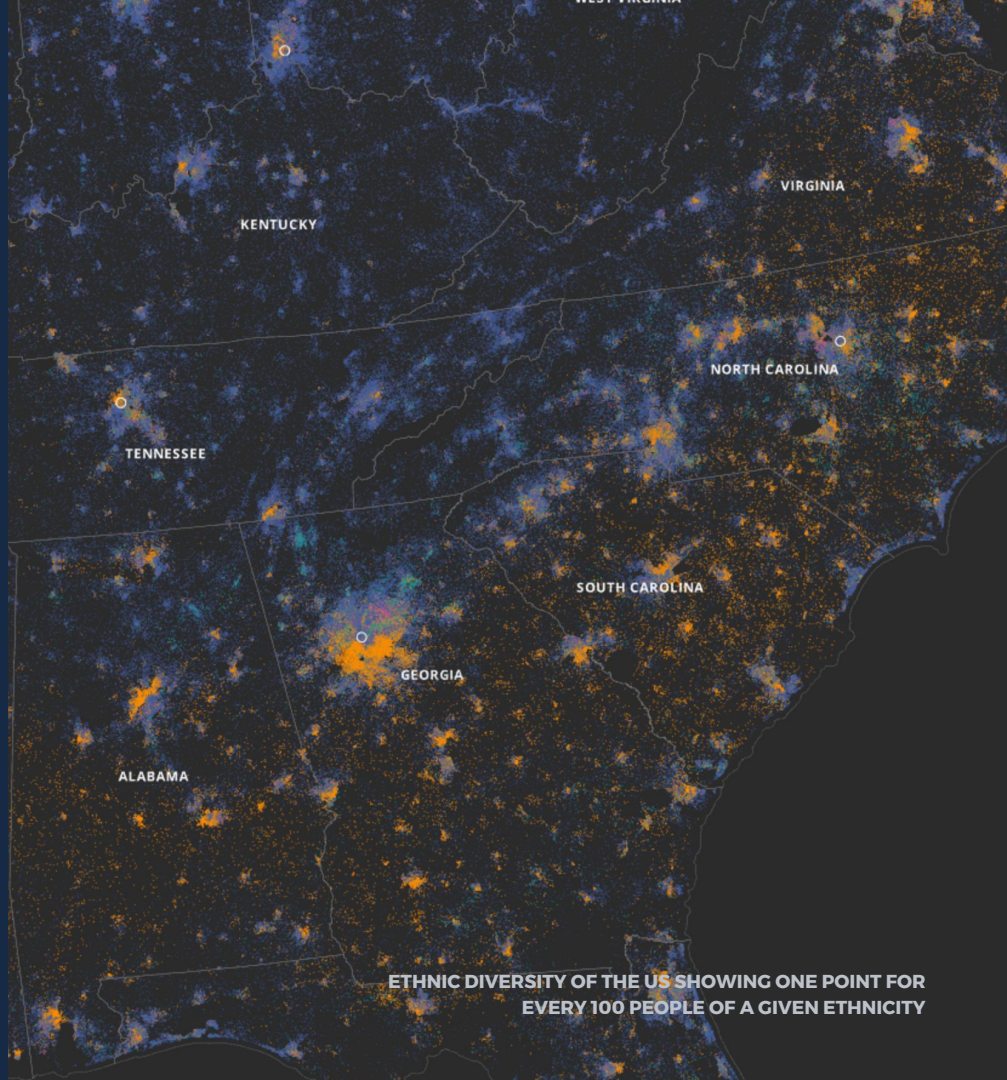
Out[3]:



Spatial Analysis for insights and prediction

It's not just your data that is spatial, the underlying processes that produce it and drive your business are as well.

CARTO offers a suite of analysis methods to understand, quantify and predict those spatial relationships to generate fresh insight and opportunity from your data.



ETHNIC DIVERSITY OF THE US SHOWING ONE POINT FOR EVERY 100 PEOPLE OF A GIVEN ETHNICITY

Analysis

Including spatial analysis in your data science workflow can uncover new insights and opportunities

Optimization

Analyze your operational infrastructure for inefficiencies and find new optimizations. See how a change will affect your network.

Inference

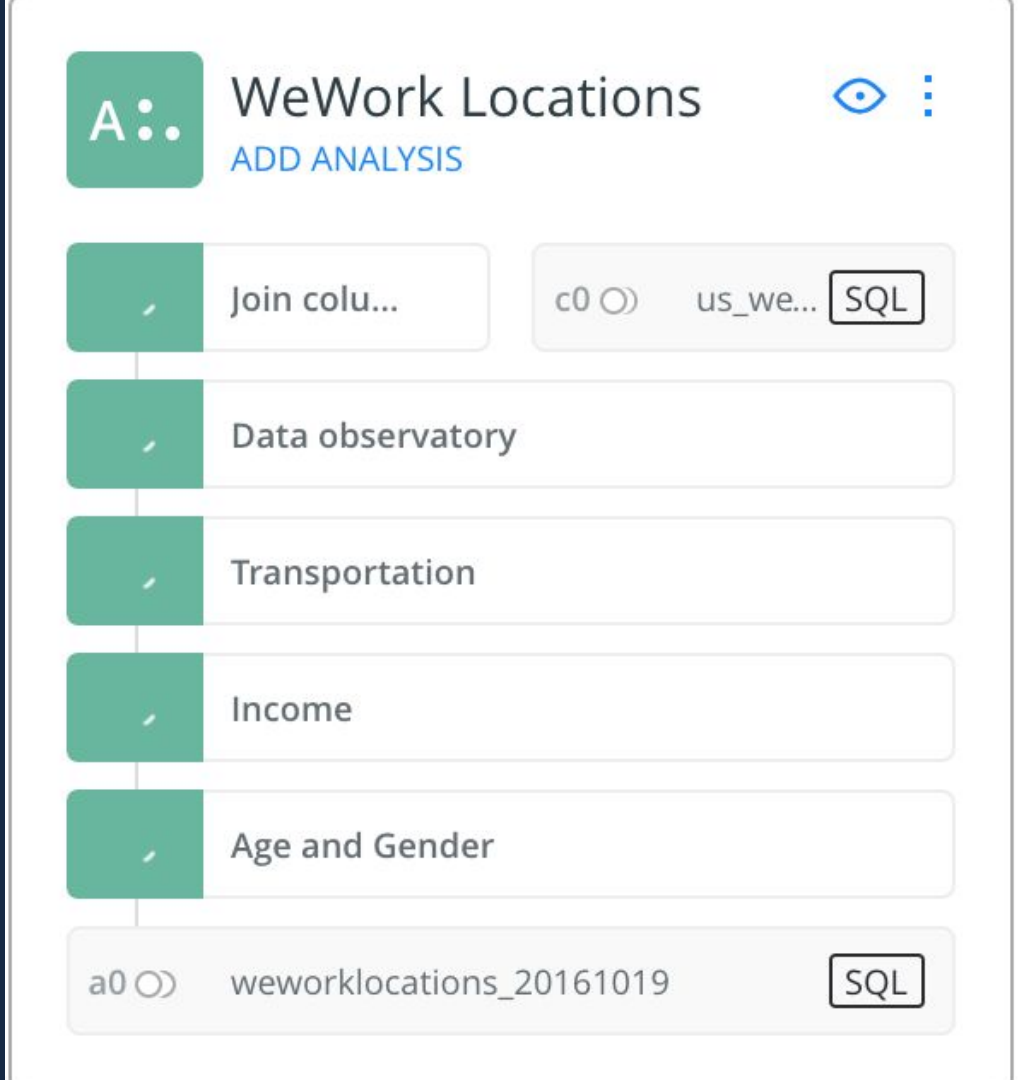
Uncover the hidden relationships in your data to inform decision making, planning and understanding customers behavior.

Prediction

Use the spatial and temporal patterns in your data to predict performance in new markets and identify new opportunities for growth.

Intuitive Analysis pipelines

Easily generate powerful analysis pipelines that combine GIS and Data Science workflows



Work in your own Data Science world

CARTO offers **CartoFrames** a convenient and intuitive Python module that fits into your existing Data Science Workflows

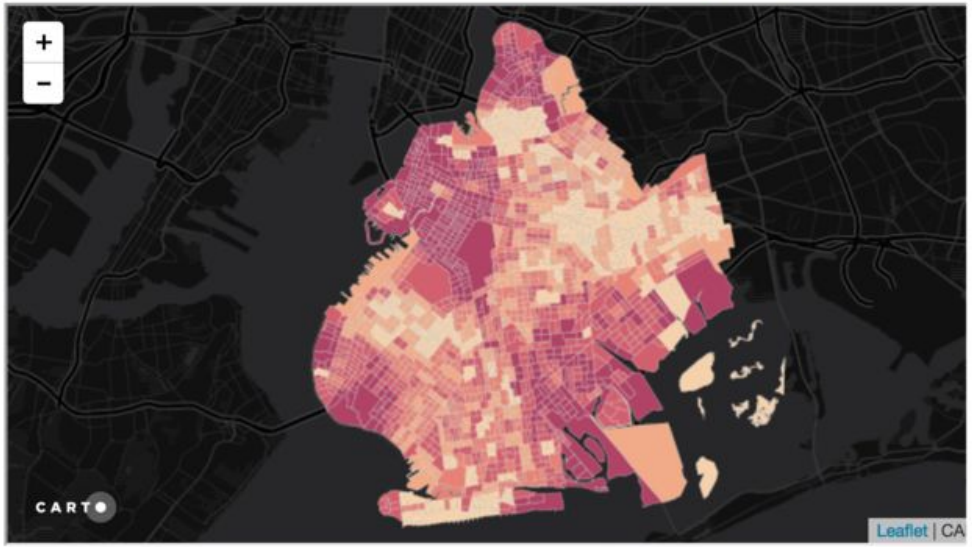
```
df = pd.read_carto(username=USERNAME,
                   api_key=APIKEY,
                   tablename='brooklyn_poverty_notnull_geoms')
df.head()
```

Out[1]:

	geoid	high_school_diploma_2011_2015	median_income_2011_2015	p
cartodb_id				
153	360470018001	0.219213	36414.226756	0.
946	360470352001	0.258464	17156.148349	0.
15	360470514002	0.264986	45746.135380	0.
37	360470137001	0.105692	100039.579057	0.
1	360470043001	0.059511	112115.160834	0.

In [2]:
```
df.carto_map(interactive=True, stylecol='median_income_2011_2015')
```

Out[2]:



Inference

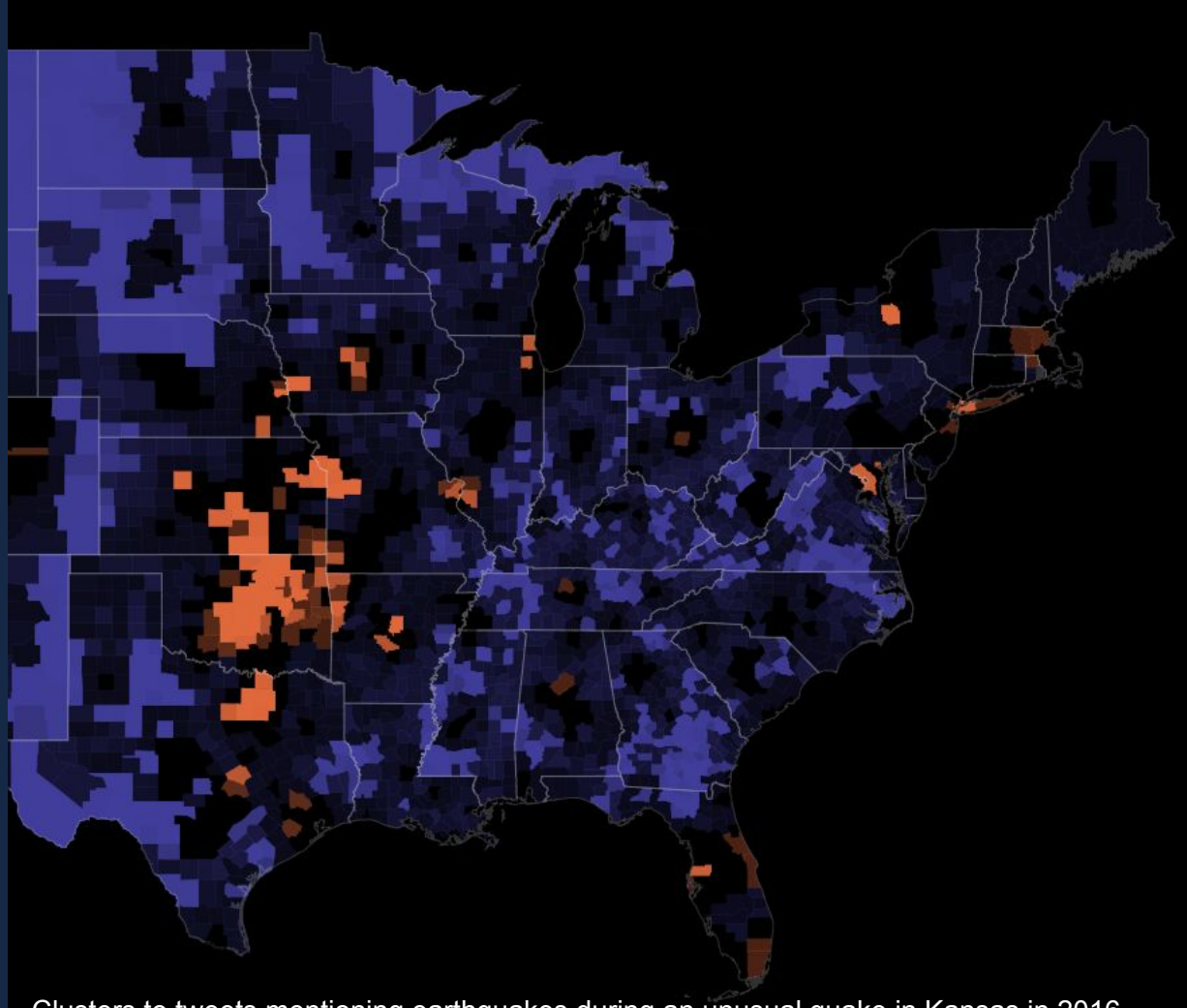
Your data contains underlying patterns and relationships that vary from place to place.

Use CARTO's inference tools to uncover them and use them to make better decisions.

ANALYSIS

Outliers and Clusters

Find clusters of high or low values and also outliers by measuring Spatial Autocorrelation.

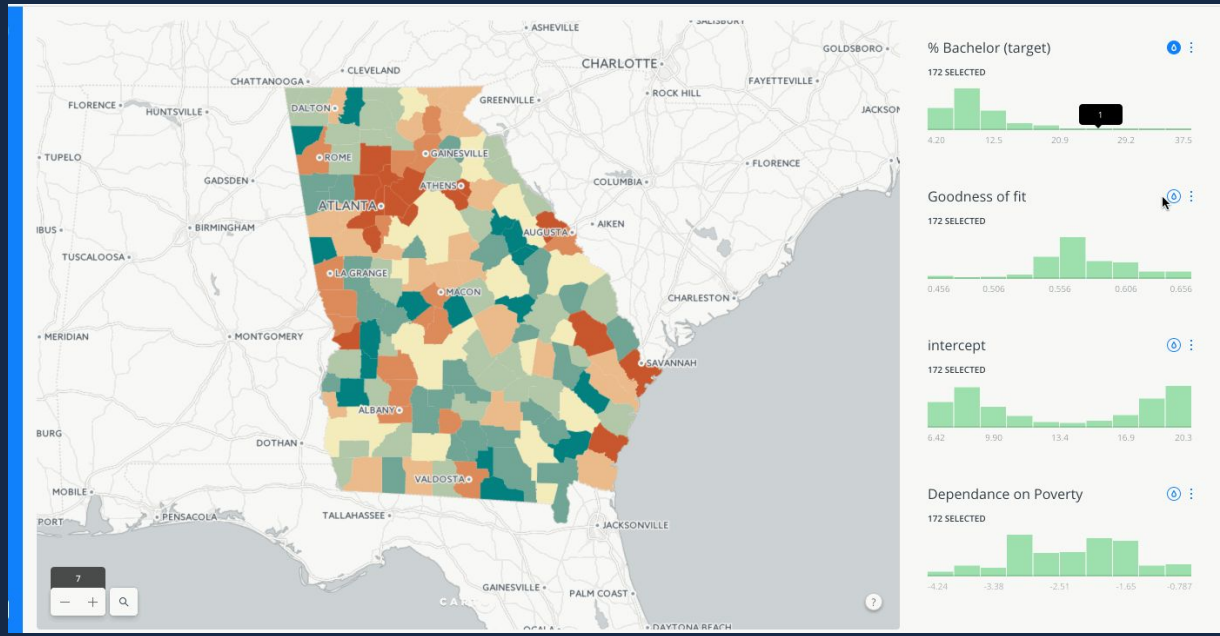


Clusters to tweets mentioning earthquakes during an unusual quake in Kansas in 2016

RESEARCH

Geographically Weighted Regression

Uncover the Relationships In your data and how they vary spatially.



Relationships between educational attainment, poverty and diversity in rural and urban environments.

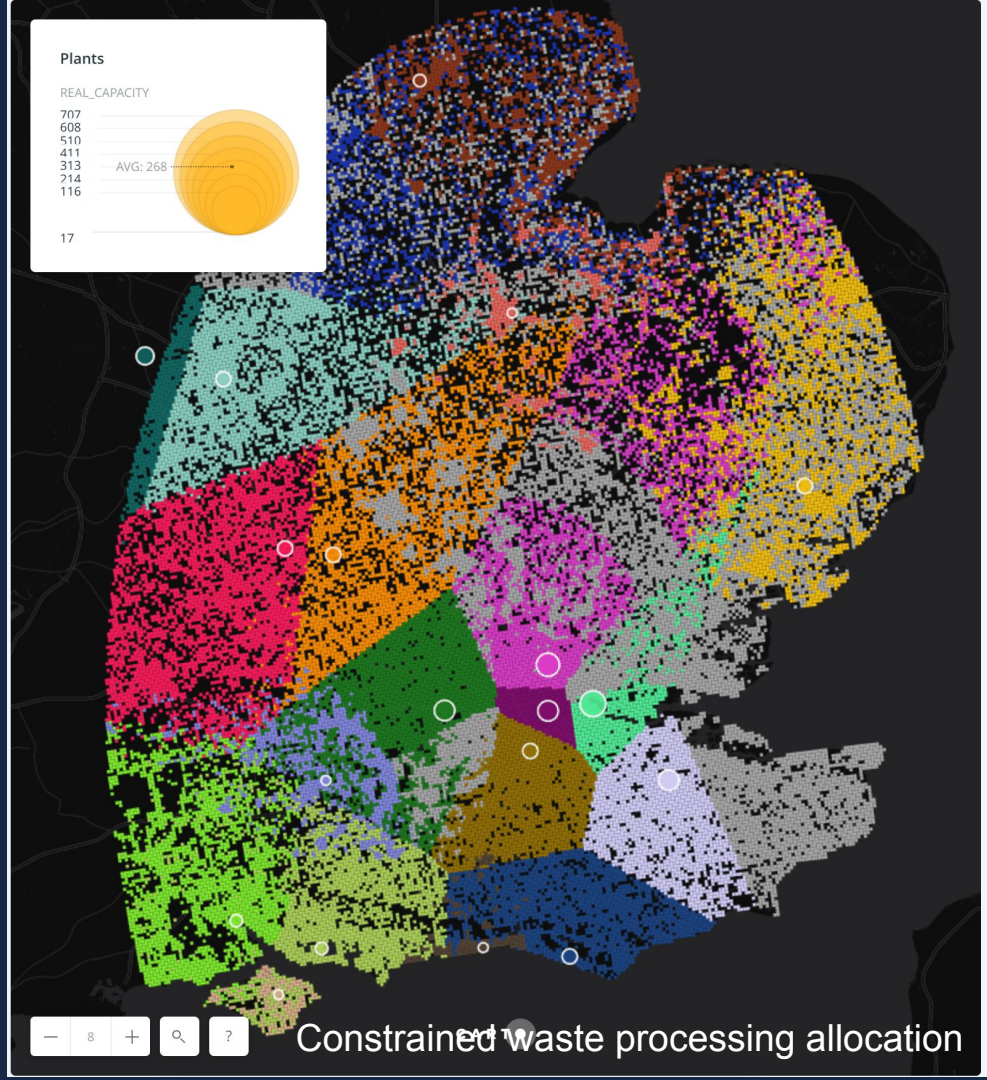
Optimization

Most infrastructure is spatially distributed and incurs costs. Use CARTO to profile and configure your operation to minimize costs and identify key high nodes within your operation .

ANALYSIS

Linear programing

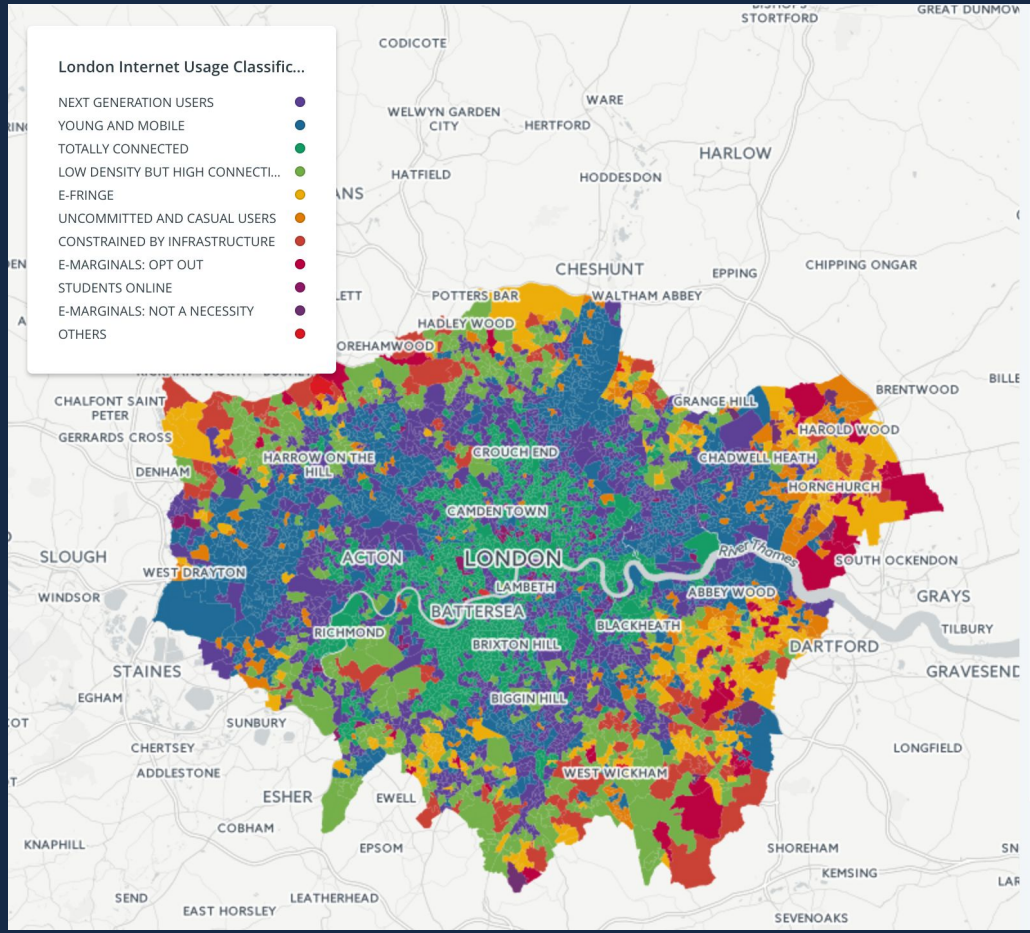
Use our linear programing module in conjunction with Data Locations Services to optimize and fine tune your decision making.



ANALYSIS

Region Definition

Use our Max-P and spatial Agglomerate Analysis Methods to define balanced sales regions, distinct markets and more.



Internet useage clusters in London

Prediction

Use your existing data to predict how your business will perform in new markets and how existing markets will evolve

ANALYSIS

Spatial Markov Chains

Use the past performance of your locations and their neighbors to predict trends in the market