Spatial data analysis

1. Show-cases of projects on spatial data
2. Spatial networks and mobility

Your projects: choose network to visualise

Spatial data analysis

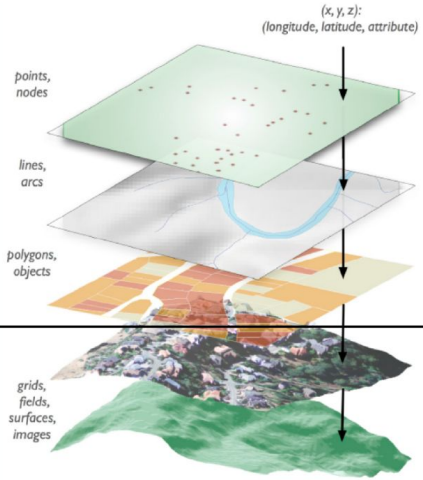
1. packages to keep in mind: shapely, geopandas, osmnx, moving pandas

2. new data types (like shape files), polygons

3. standard methods (e.g. correlation analysis, statistical tests) need to be adapted

What are “shape files”?

Vector: Geometric objects
.shp, .svg



Raster: Grid of pixels
.tif, .jpg, .png, .bmp

Shape files main objects

- Point
- LineString
- Polygon

which are the basic ingredients when working with spatial data in vector format.

Shapefiles - they are usually groups of files containing geospatial vector data, with the same name and extensions like .shp, .shx, .dbf. Here we only care about the .shp file which contains the geometric features.

See also https://github.com/mszell/geospatialdatascience
Image Maggi Kelly

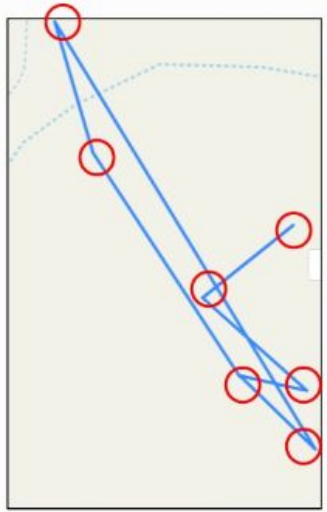
Examples of data sets

GIS programs vs. python vs. specific applications

Movement of one User during an Earthquake

Typhoons

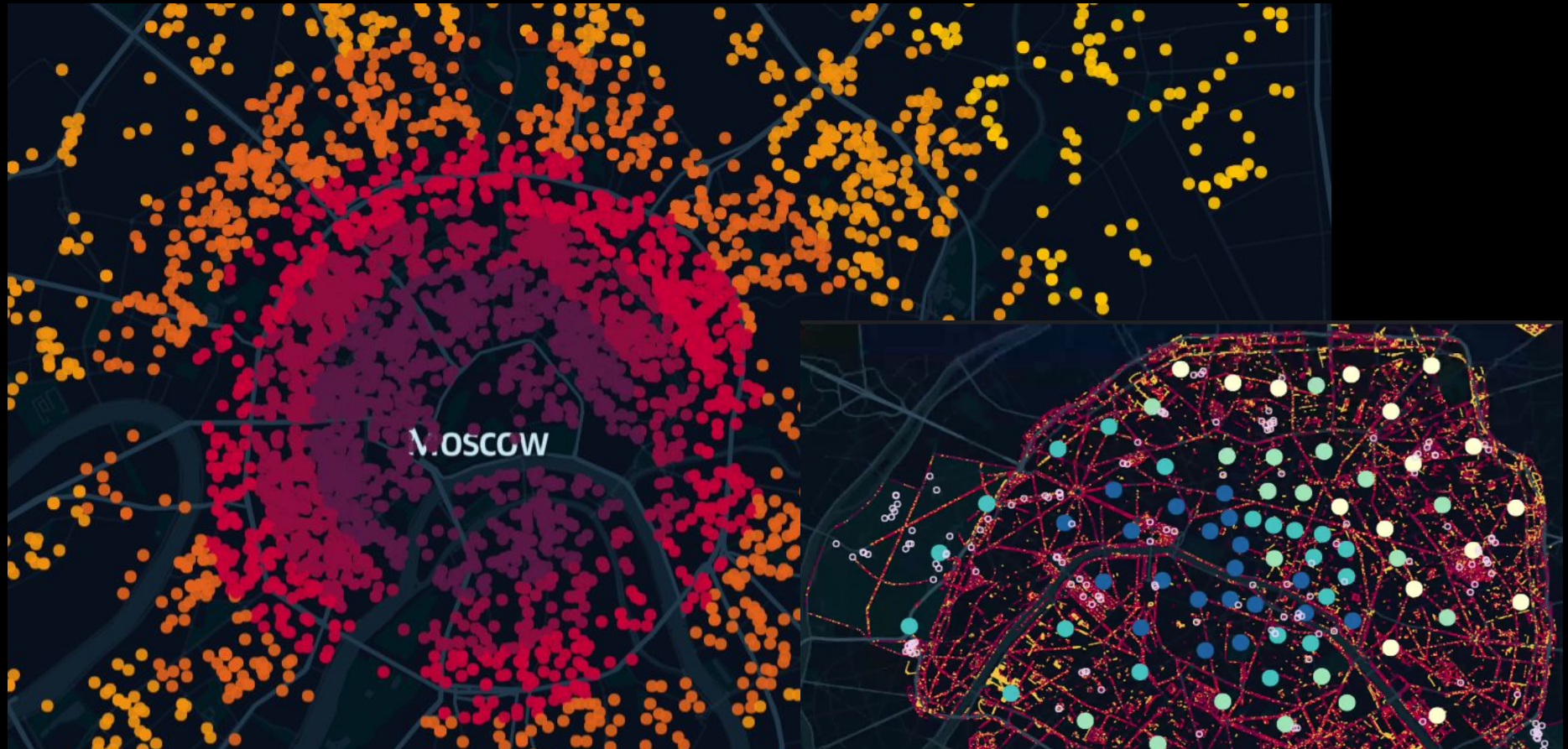
Earthquake in Iquique Chile

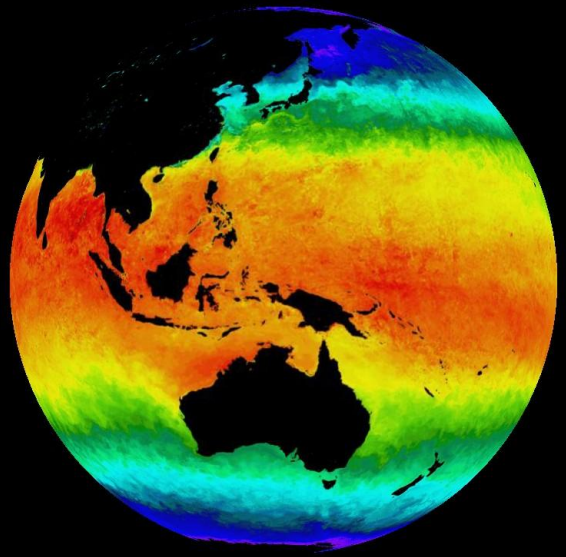


GeoJson File

```
[-70.144336, -17.975391, 0, 1395904540],
[-70.146159, -17.976769, 0, 1395996037],
[-70.144062, -17.978553, 0, 1396482110],
[-70.145409, -17.978258, 0, 1396653147],
[-70.143879, -17.979667, 0, 1396864856],
[-70.149145, -17.97148, 0, 1397655993],
[-70.148378, -17.974008, 0, 1397656479],
[-70.145405, -17.978278, 0, 1395655322],
[-70.145405, -17.978278, 0, 1395655428]
```

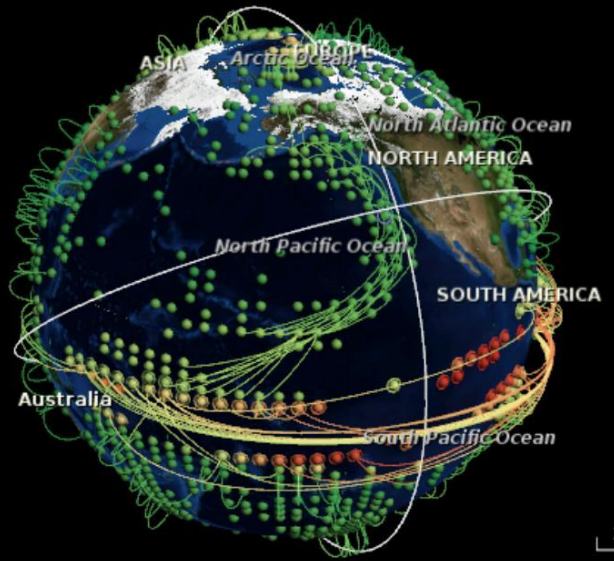
Examples of data from Memorial

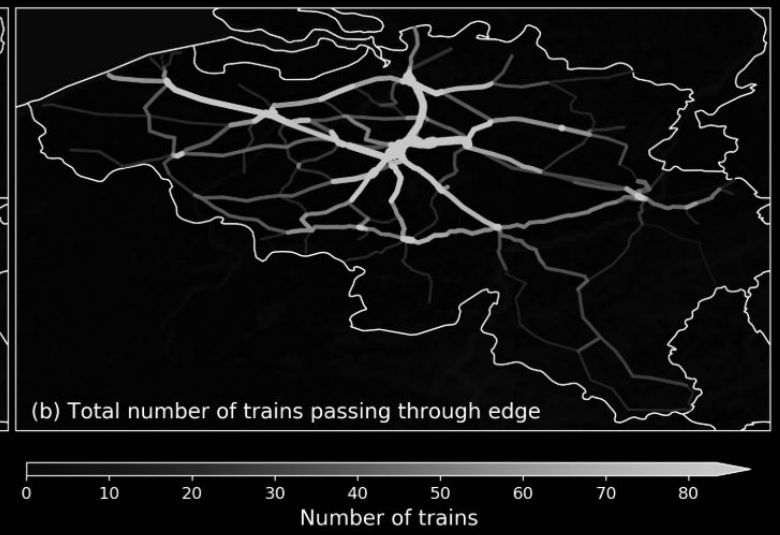




Spatial data projects

Analysis of climate data:
Pyunicorn python package
https://github.com/pik-copan/pyunicorn





Transport in networks Nature Comm. (2020)

Modelling railway delay https://arxiv.org/abs/2105.061

City evolution in time
C.Lagesse, L.Tupikina, P.Bauman

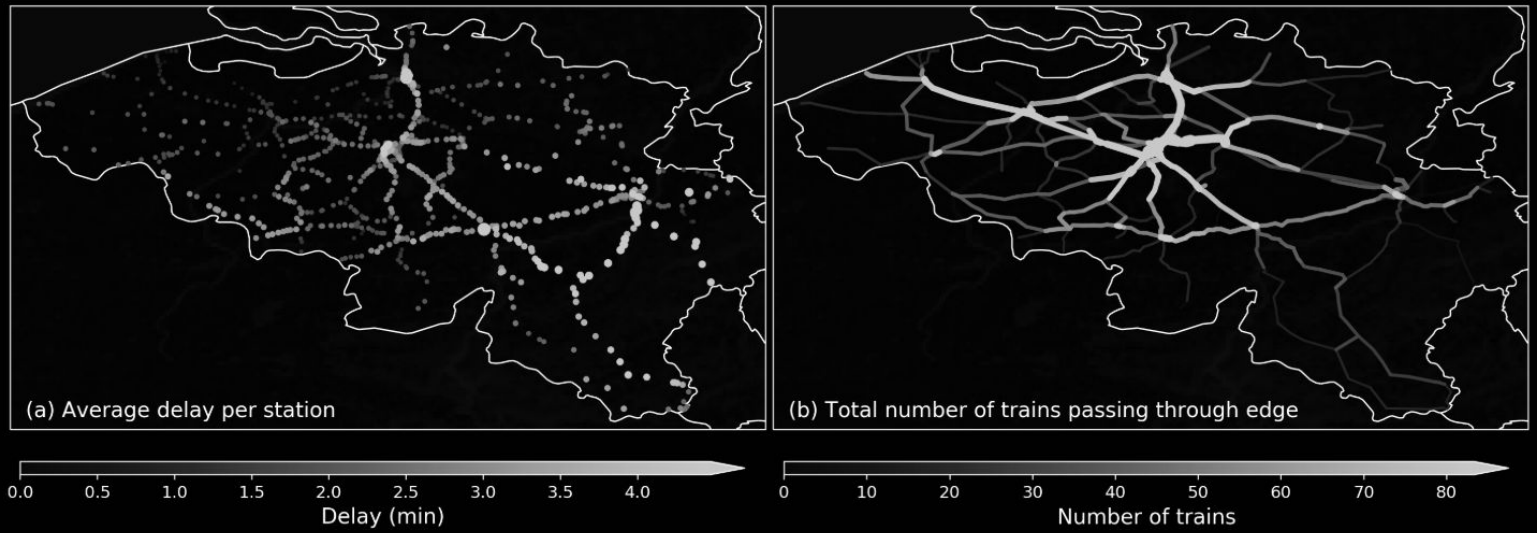
Shared bike systems
https://github.com/Liyubov/bike-sharing

Citizen science city analysis
https://github.com/correlaid-paris/citizen_science_ina
ralist

Analysis of innovations in science
cri-paris.org

Closed cities ZATOs
https://theterraforming.strelka.com/

Examples of projects



Modelling railway delay
https://arxiv.org/abs/2105.06111
About me
https://sites.google.com/view/liubovkmatematike/

NEWS FEATURE | 23 October 2018

No PhDs needed: how citizen science is transforming research

Projects that recruit the public are getting more ambitious and diverse, but the field faces some growing pains.

Aisling Irwin

ZOOniverse - 1 mln users volunteers, 2014

Openstreetmaps > 7 mln users 2021

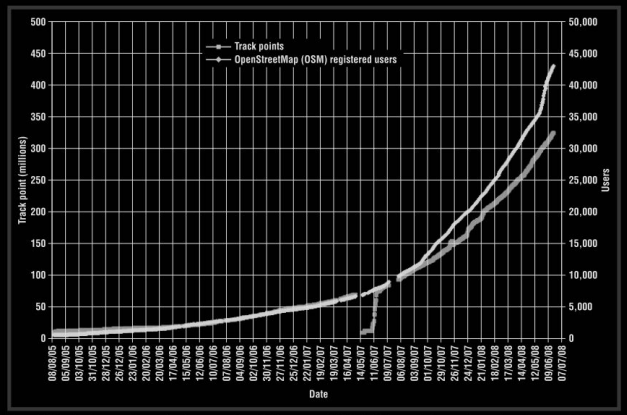


Figure 1. Graph of user and contributions growth to OSM on a monthly basis. The graph shows the accelerating growth in number of users and the rapid increase in data entry measured in track points (source: http://wiki.openstreetmap.org).

NEWS FEATURE | 23 October 2018

No PhDs needed: how citizen science is transforming research

ZOOniverse - 1 mln users volunteers, 2014

Openstreetmaps > 7 mln users 2021

Sacramento

Stockton

Modesto

YOSEMITE NATIONAL PARK

SIERRA NATIONAL FOREST

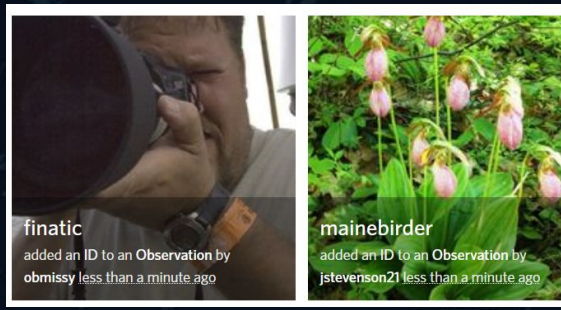
San José

Gilroy

Santa Cruz

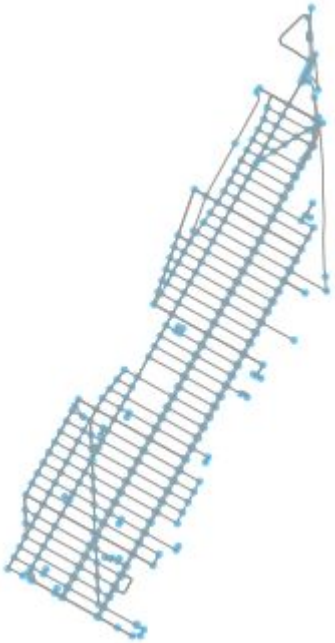
Showcase of some projects

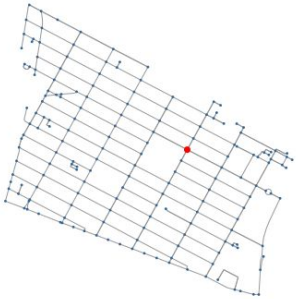
iNaturalist citizen science project

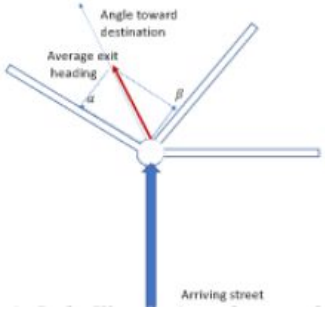


Data, code, conference
https://github.com/correlaid-paris/citizen_science_inaturali

CitSciVirtual
2021
LOCAL, GLOBAL, CONNECTED
Coming N
Registration Op



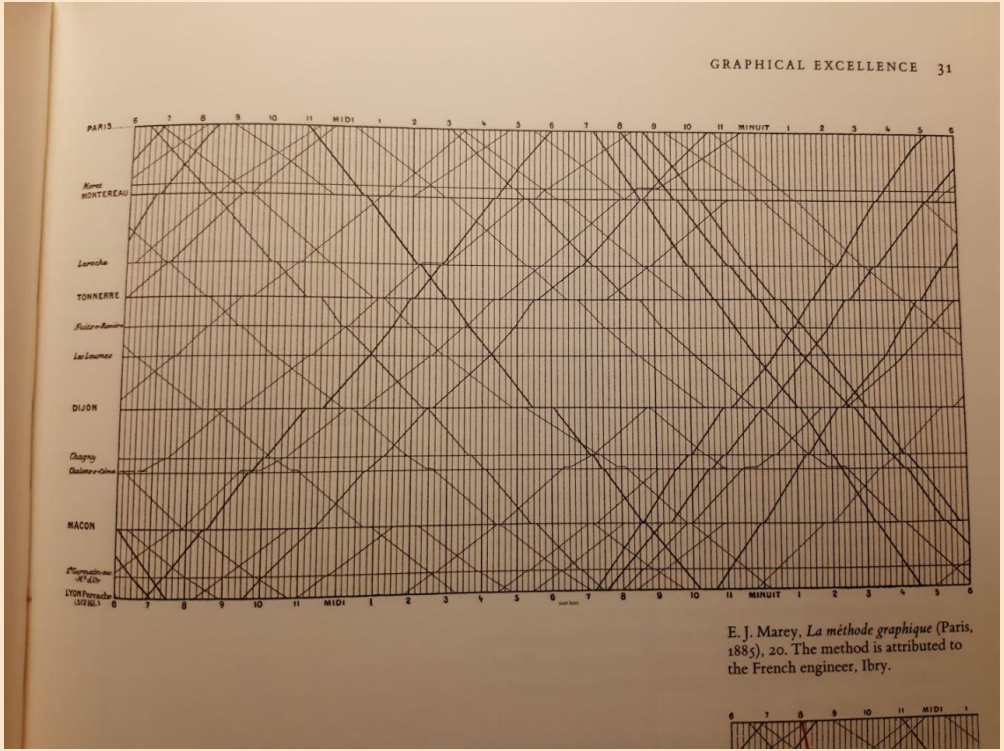



Angle toward destination
Average exit heading
α
β
Arriving street

Openstreetmaps analysis

Links to github
https://github.com/cityinteractionlab/openstreetmaps_osmnx_workshop
https://github.com/gboeing/osmnx-examples
https://github.com/cityinteractionlab/openstreetmaps_osmnx_workshop
Google colab
(geo packages required)

Data analysis from 19th century



Tufte book: spatial data visualisation

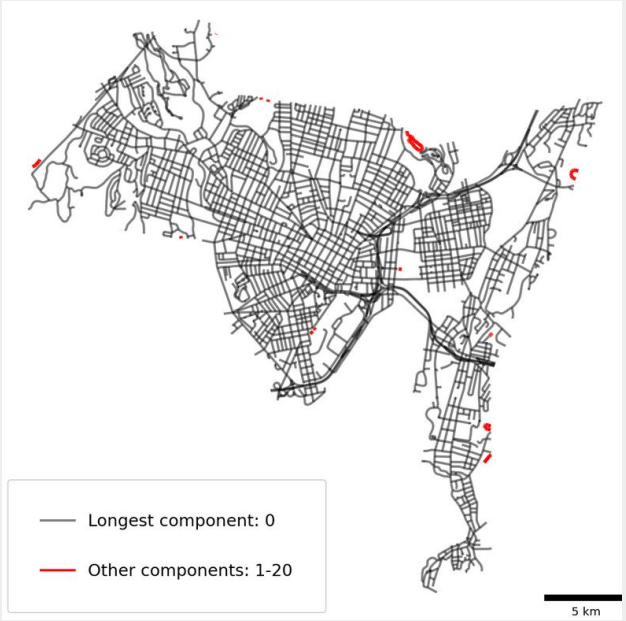
Resources pages

www.worldpop.org spatial population analysis
hdx platform

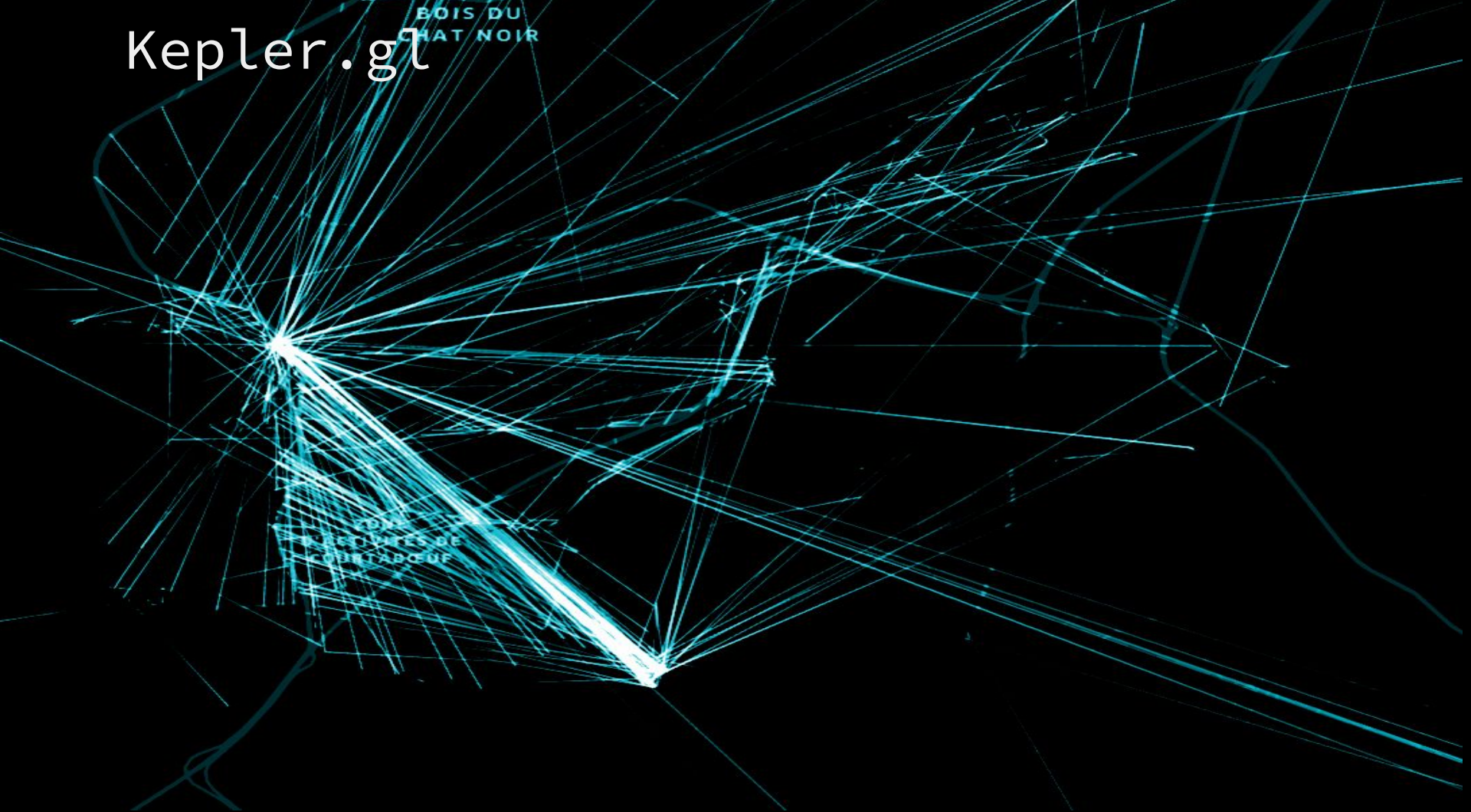
www.kepler.gl for mobility data visualisation (browser)
Mapbox python integration

Open source tools, python and R
https://www.python-graph-gallery.com/

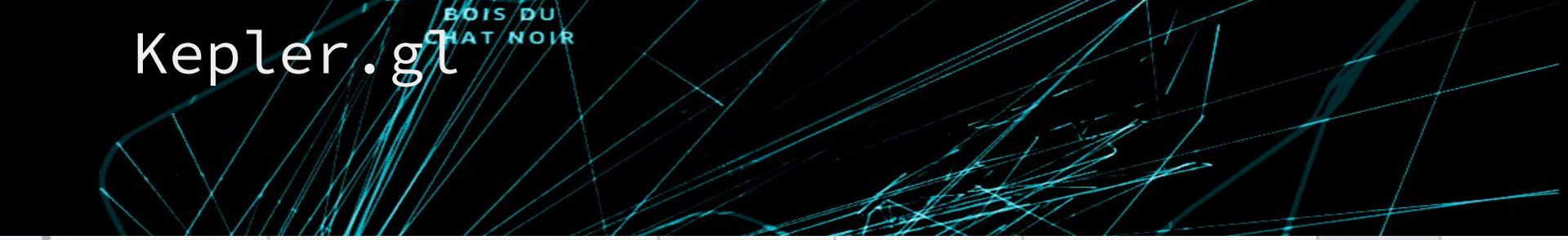
Packages:
Scikit https://github.com/scikit-mobility/scikit-mobility
Matplotlib, cartopy - simple plotting,
Folium - online plotting,
Geopandas - python package,
libpysal - spatial distribution,
Osmnx - python package for analysis of openstreetmaps
Spagetti https://github.com/pysal/spaghetti



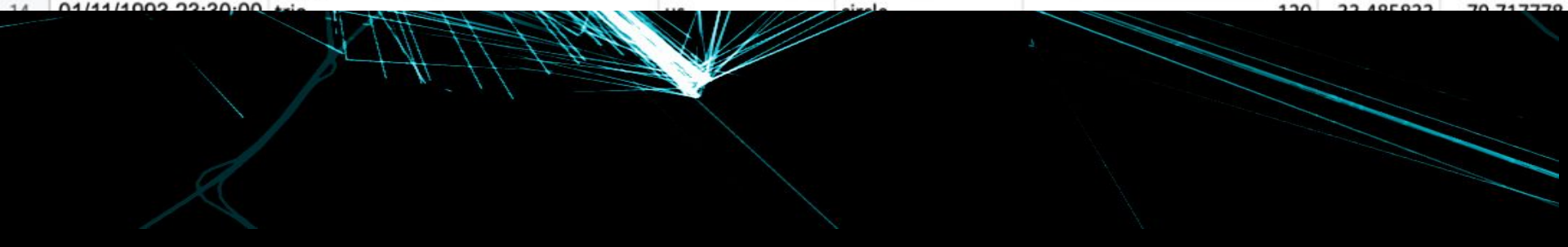
Kepler.gl



Kepler.gl



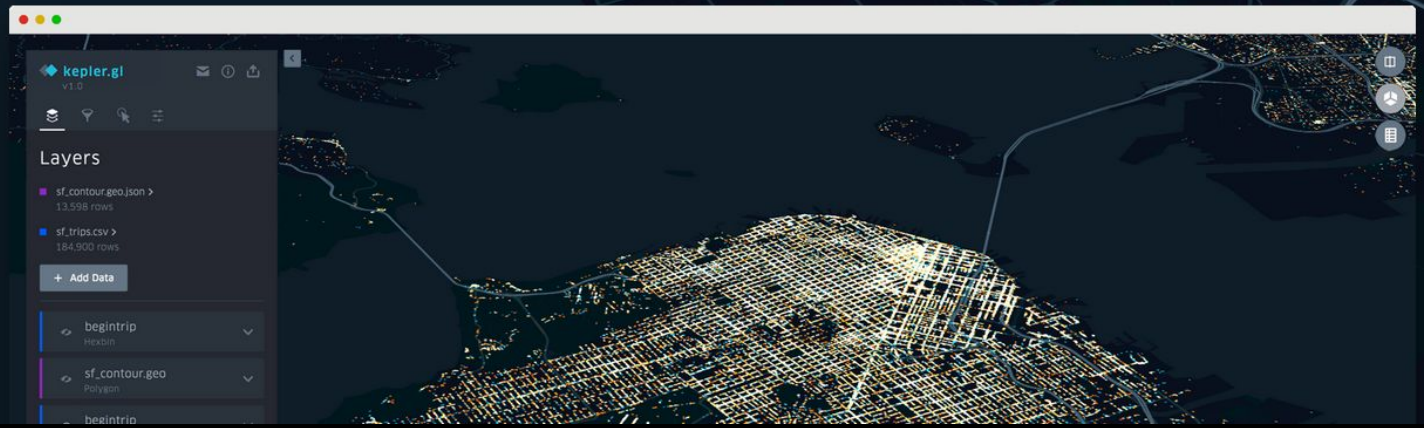
1	Date_time	city	country	UFO_shape	length_of_encounter_seconds	latitude	longitude
2	01/03/1993 12:00:00	prescott	us	diamond	900	34.540000	-112.467778
3	01/03/1993 13:00:00	chattanooga	us	sphere	300	35.045556	-85.309722
4	01/06/1993 14:00:00	pittsburgh	us	sphere	15	40.440556	-79.996111
5	01/06/1993 21:00:00	san jose (snell rd / blossom hill rd)	us	circle	300	37.339444	-121.893889
6	01/06/1993 22:00:00	billings	us	light	20	45.783333	-108.500000
7	01/07/1993 00:00:00	phoenix	us	oval	60	33.448333	-112.073333
8	01/07/1993 03:30:00	katy	us	sphere	15	29.785556	-95.824167
9	01/08/1993 17:00:00	warrenton	us	circle	300	38.713333	-77.795556
10	01/08/1993 21:30:00	tillamook(lees camp)	us	light	900	45.456389	-123.842778
11	01/09/1993 22:00:00	bethel (albany township)	us	light	120	44.404167	-70.791111
12	01/10/1993 20:00:00	delaware	us	light	7200	40.298611	-83.068056
13	01/10/1993 20:00:00	pryor	us	light	1800	45.429722	-108.532500
14	01/11/1993 22:30:00	[illegible]	[illegible]	[illegible]	120	33.485833	-70.717778



Kepler.gl is a powerful **open source** geospatial analysis tool for **large-scale** data sets.

GET STARTED

GITHUB

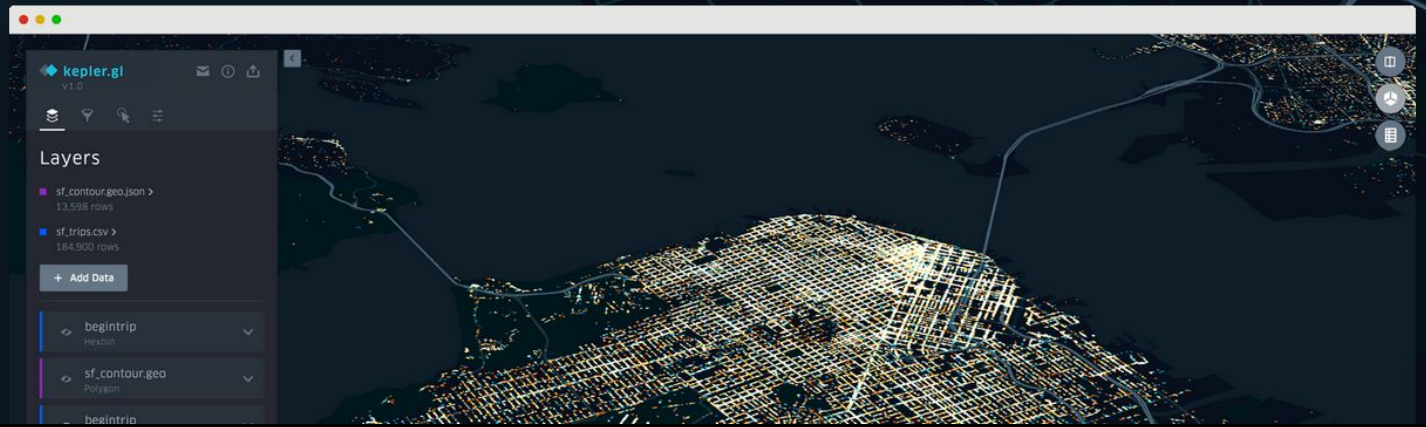


Kepler.gl open source program for visualising spatial large scale data, client based web application

Kepler.gl is a powerful **open source** geospatial analysis tool for **large-scale** data sets.

GET STARTED

GITHUB

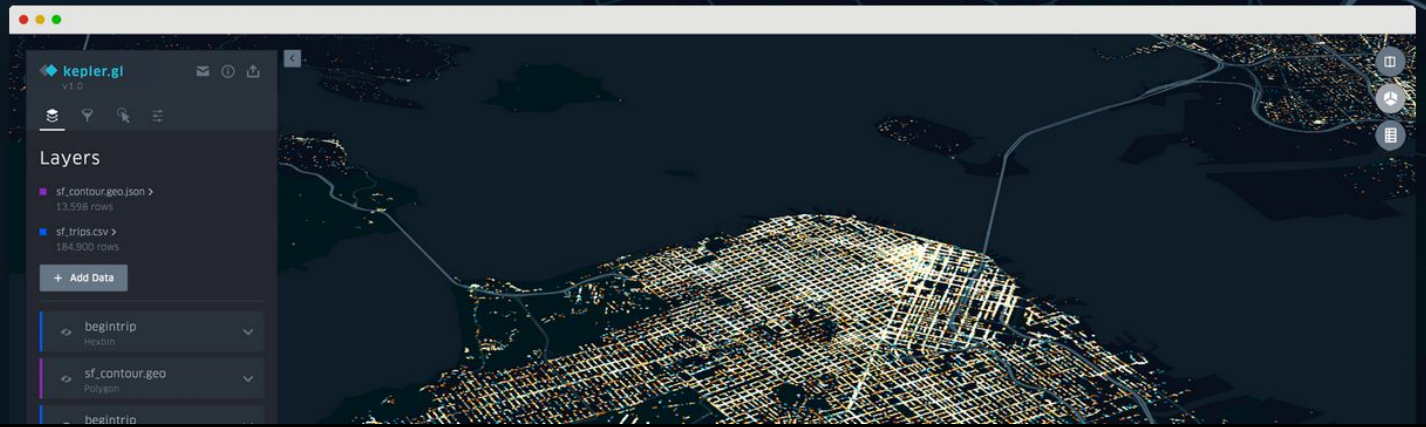


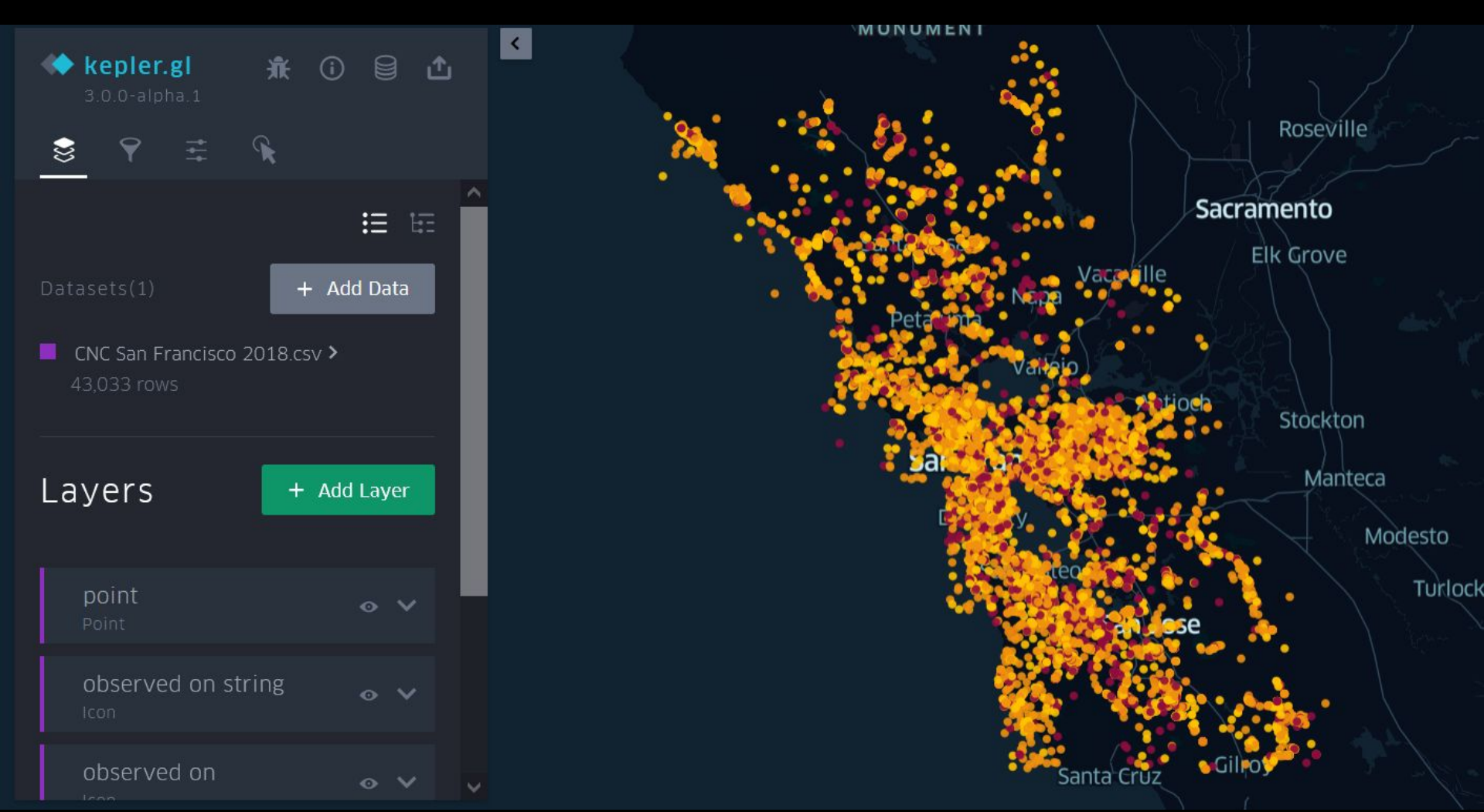
Kepler.gl open source program for visualising spatial large scale data, client based web application: all data is on your machine, export project to work on it further

Kepler.gl is a powerful **open source** geospatial analysis tool for **large scale** data sets.

GET STARTED

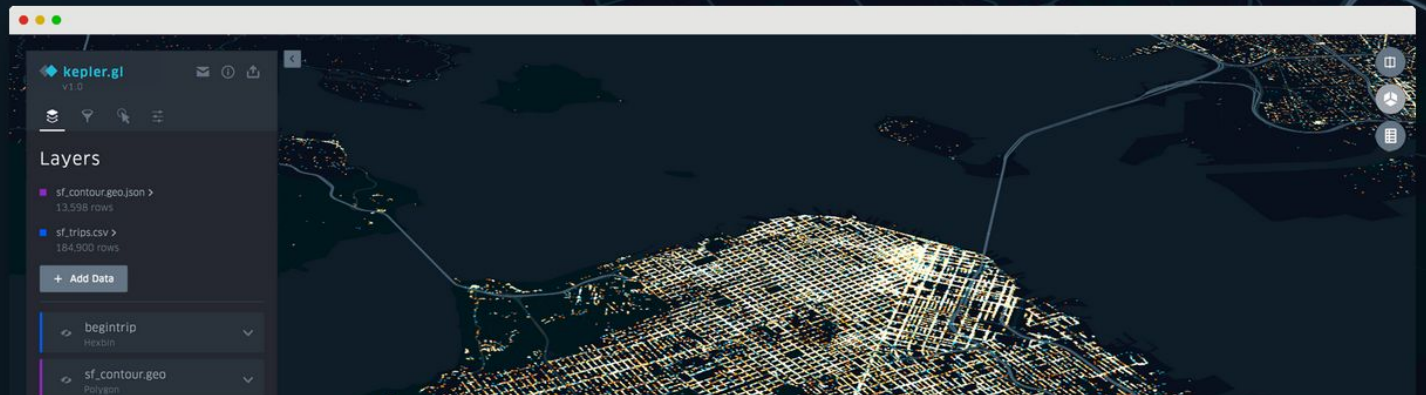
GITHUB



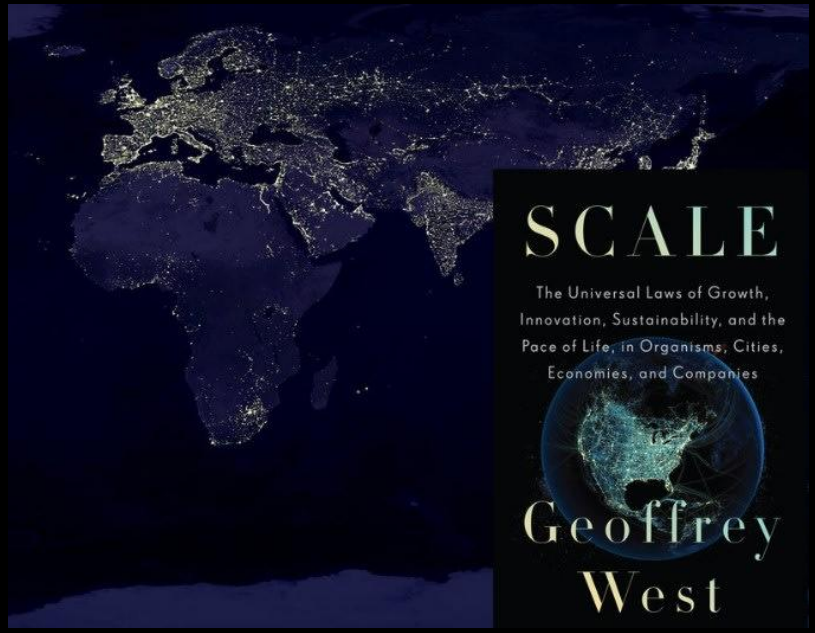

kepler.gl
3.0.0-alpha.1
Datasets(1)
+ Add Data
CNC San Francisco 2018.csv
43,033 rows
Layers
+ Add Layer
point
Point
observed on string
Icon
observed on
Roseville
Sacramento
Elk Grove
Vacaville
Stockton
Manteca
Modesto
Turlock
Santa Cruz
Gilroy

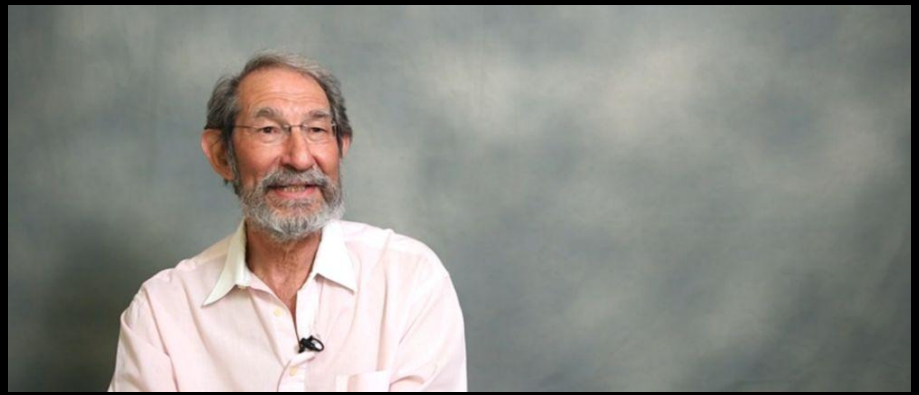
Try your own dataset, example one from data.world, set it from github

Name the geo information as latitude, longitude



Annex





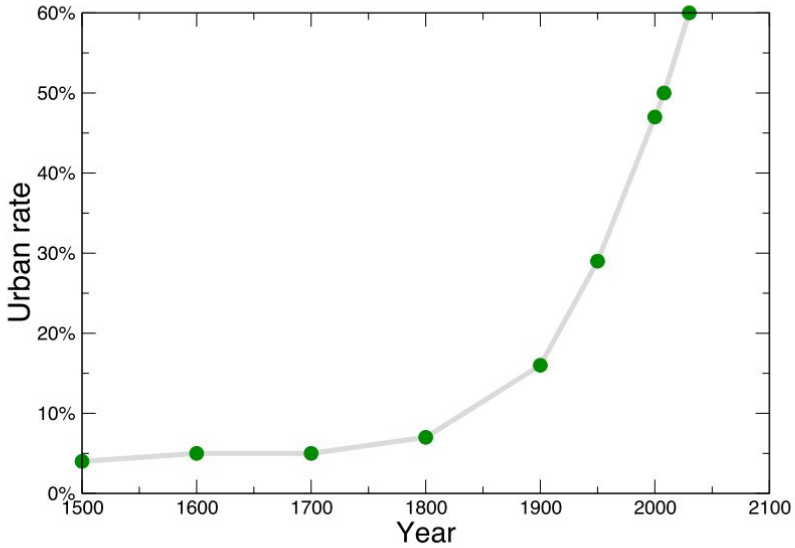
City science

Non-equilibrium, open system

L.Bettencourt, G. West (Santa Fe) “A unified theory of urban living”, Nature (2010)

Why city science?

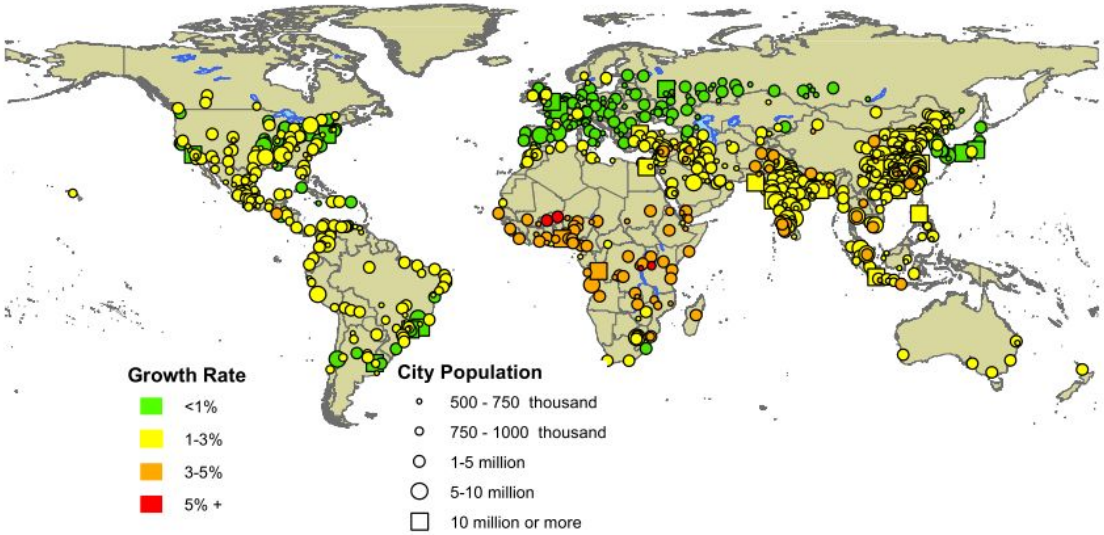
Importance of cities: urbanization rate



Projection: in 2050: 70% of the world population lives in cities

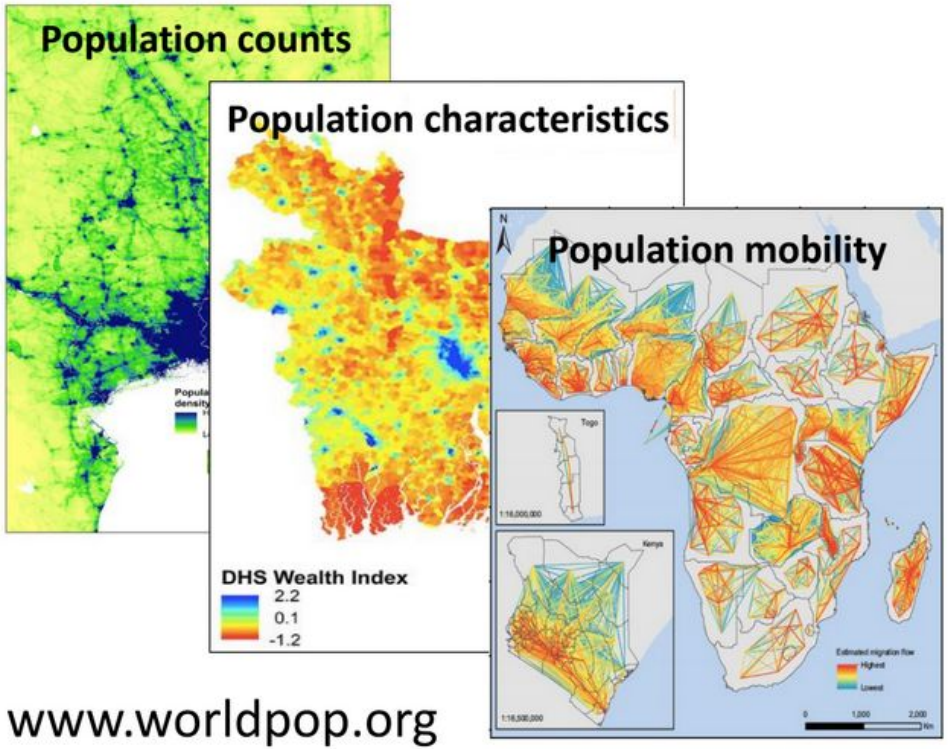
Ourworld in data

Why city science?



Ourworld in data
https://ourworldindata.org/

Why spatial analysis?



www.worldpop.org

Applied research and implementation group

Mapping small area population distributions, demographics and dynamics

Open data, open peer-reviewed statistical methods, user engagement

Application in epidemiology, maternal/newborn health, childhood vaccination

Some papers for inspiration

Paper on human mobility review https://arxiv.org/abs/1710.00004

Paper on spatial planning and bicycle infrastructure Natera, L.G., Battiston, F., Iñiguez, G., and Szell, M. *"Data-driven strategies for optimal bicycle network growth"*, 2019.

Paper on spatial networks https://arxiv.org/abs/1010.0302

M.Barthelemy "Spatial networks" https://arxiv.org/pdf/1010.0302.pdf

M.C. Gonzalez, C.A. Hidalgo, and A.-L. Barabasi "Understanding individual human mobility patterns" Na-ture, 453:779 - 782, 2009.
https://www.ncbi.nlm.nih.gov/pubmed/18528393

Human mobility

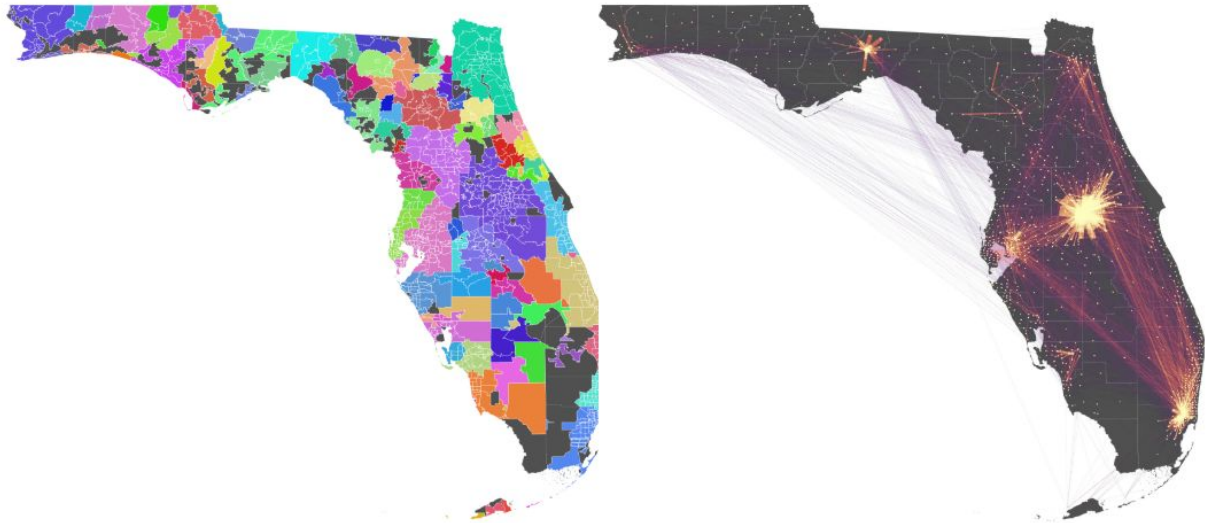


Fig. 3: Commuting flows compiled from census data. Left panel: The state of Florida partitioned according to its counties. Right panel: Commuting flows between counties, where thickness of lines correspond to volume of flow. Data compiled from the United States Census Bureau.

How do people move in real life? Human Mobility: Models and Applications - Hugo Barbosa-Filho et al.

Human mobility

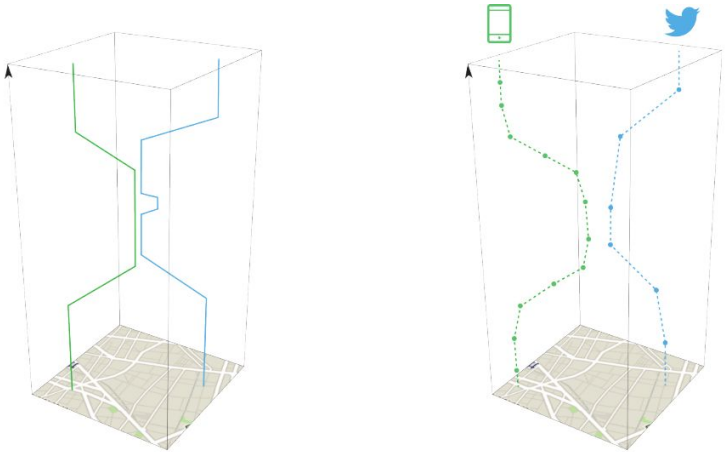


Fig. 1: The cubes of time geography, as first proposed by Torsten Hägerstrand in [36]. The geographical space is represented by the 2D plan, while time is figured by the vertical axis. (Left) The two curves represent the daily space-time trajectories of two individuals living in the same neighborhood and working in the same place. (Right) The geographical footprints continuously and passively produced by individuals through the use of their ICT devices allow to approximate their trajectories. While these re-constructed trajectories are partial and contain errors that might mislead the understanding of underlying trajectories, they are nonetheless more precise nowadays than they were 10 years ago, and produced by a constantly growing number of individuals worldwide.

How do people move in real life? Human Mobility: Models and Applications - Hugo Barbosa-Filho et al.