Geographic Data Science - Lecture IX

Points

Dani Arribas-Bel

Today

- The *point* of points
- Point patterns
- Visualization of point patterns

The *point* of points

Points like polygons

Points *can* represent **"fixed" entities**

Points like polygons

Points *can* represent **"fixed" entities**

In this case, points are **qualitatively** similar to **polygons/lines**

Points like polygons

Points *can* represent **"fixed" entities**

In this case, points are **qualitatively** similar to **polygons/lines**

The **goal** here is, taking location fixed, to model other aspects of the data

Points like polygons

Examples:

Points like polygons

Examples:

- Cities (in most cases)

Points like polygons

Examples:

- Cities (in most cases)
- Buildings

Points like polygons

Examples:

- Cities (in most cases)
- Buildings
- Polygons represented as their centroid

Points like polygons

Examples:

- Cities (in most cases)
- Buildings
- Polygons represented as their centroid
- ...

When points are not polygons

Point data are not only a different geometry than polygons or lines...

When points are not polygons

Point data are not only a different geometry than polygons or lines...

... Points can also represent a fundamentally different way to approach spatial analysis

Points unlike polygons

Points unlike polygons

- Rather than exhausting the entire space, points can be **events** subject to **occur anywhere**

Points unlike polygons

- The **location** of the event is **part** of what we are trying to understand/**model**

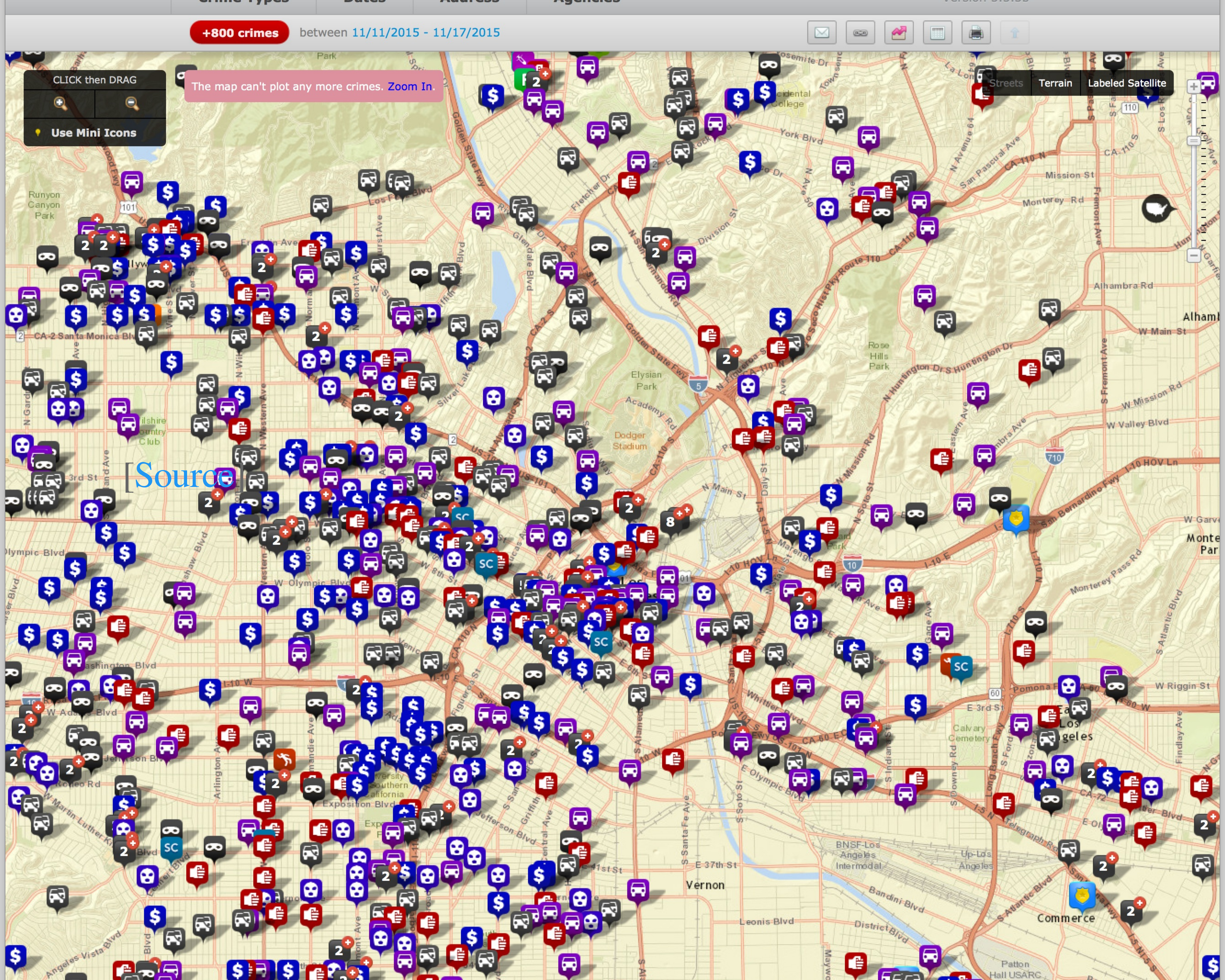
Points unlike polygons

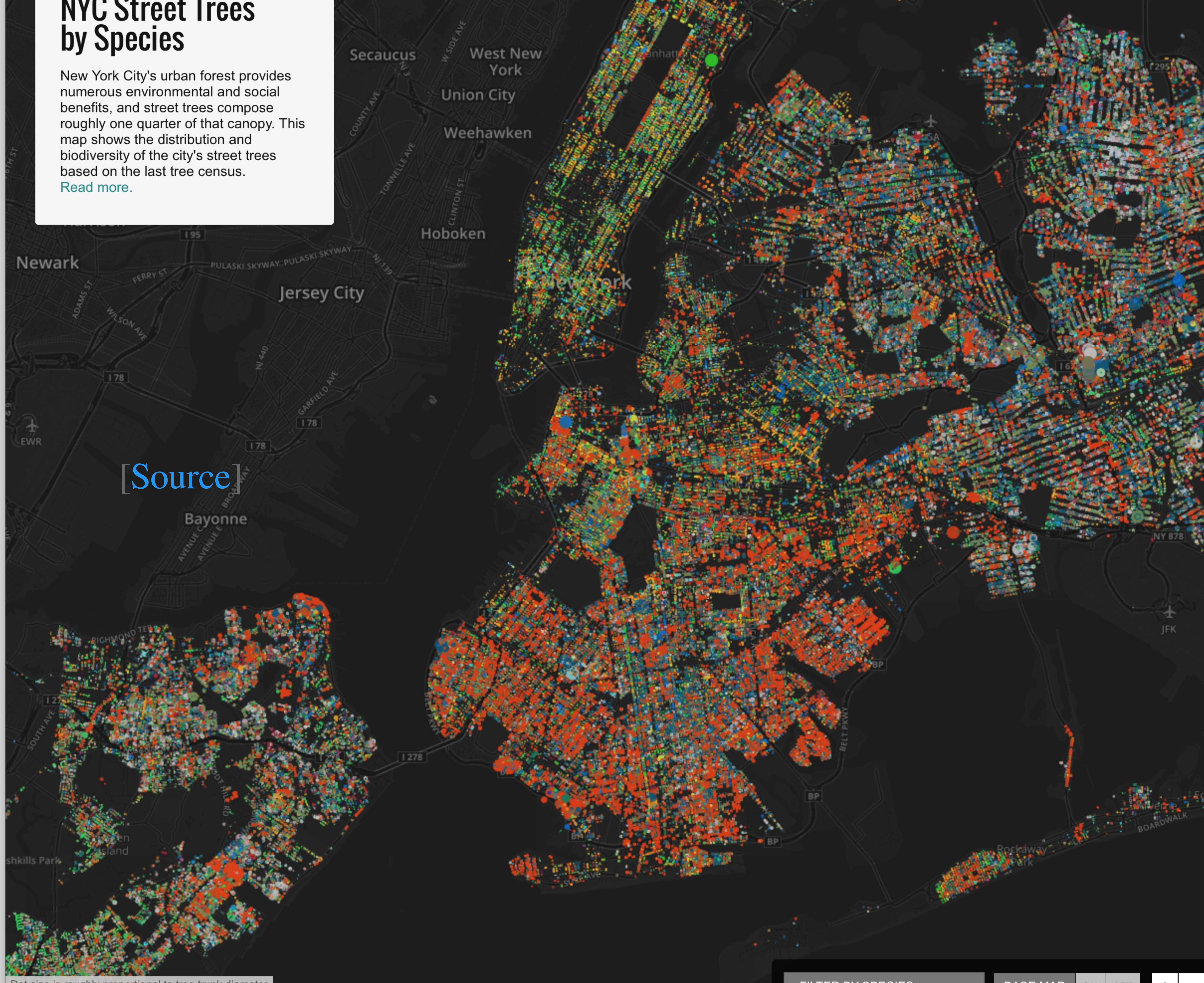
- The interest focuses on **characterizing** the **pattern** that the points follow **over space**

Points unlike polygons

- Rather than exhausting the entire space, points can be **events** subject to **occur anywhere**
- The **location** of the event is **part** of what we are trying to understand/**model**
- The interest focuses on **characterizing** the **pattern** that the points follow **over space**

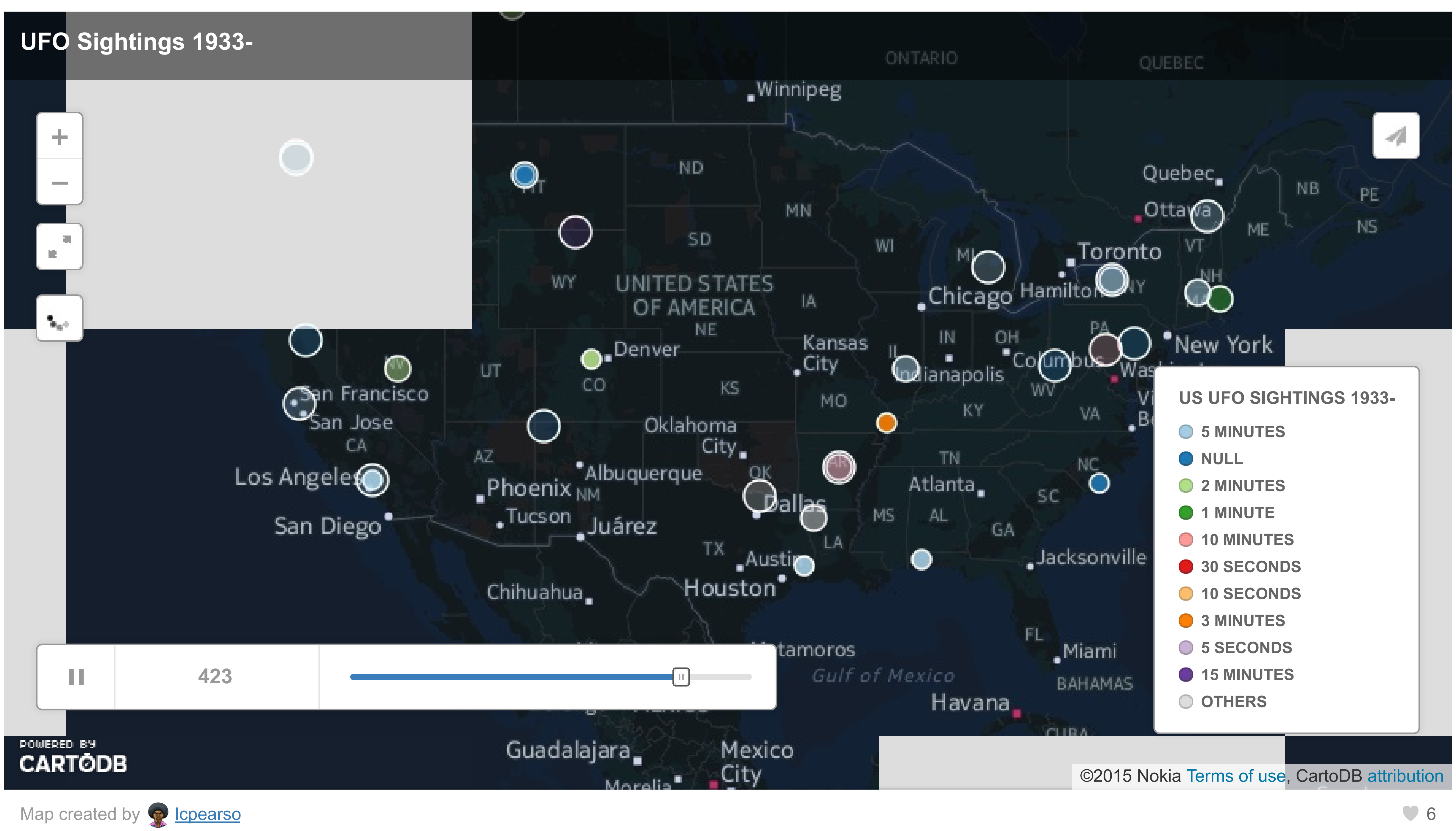
A few examples...


Crime Types
Dates
Address
Agencies
+800 crimes
between 11/11/2015 - 11/17/2015
CLICK then DRAG
Use Mini Icons
The map can't plot any more crimes. Zoom In.
Streets
Terrain
Labeled Satellite
[Source
Runyon Canyon Park
Elysian Park
Dodger Stadium
Rose Hills Park
Vernon
Commerce
East Los Angeles
Alhambra Rd
Mission St
Monterey Rd
York Blvd
W Main St
W Mission Rd
W Valley Blvd
W Riggin St
E 3rd St
E 37th St
Leonis Blvd
Bandini Blvd
District Blvd
Washington Blvd
Olympic Blvd
W Olympic Blvd
Exposition Blvd
Jefferson Blvd
Martin Luther King Jr Blvd
BNSF-Los Angeles Intermodal
Calvary Cemetery

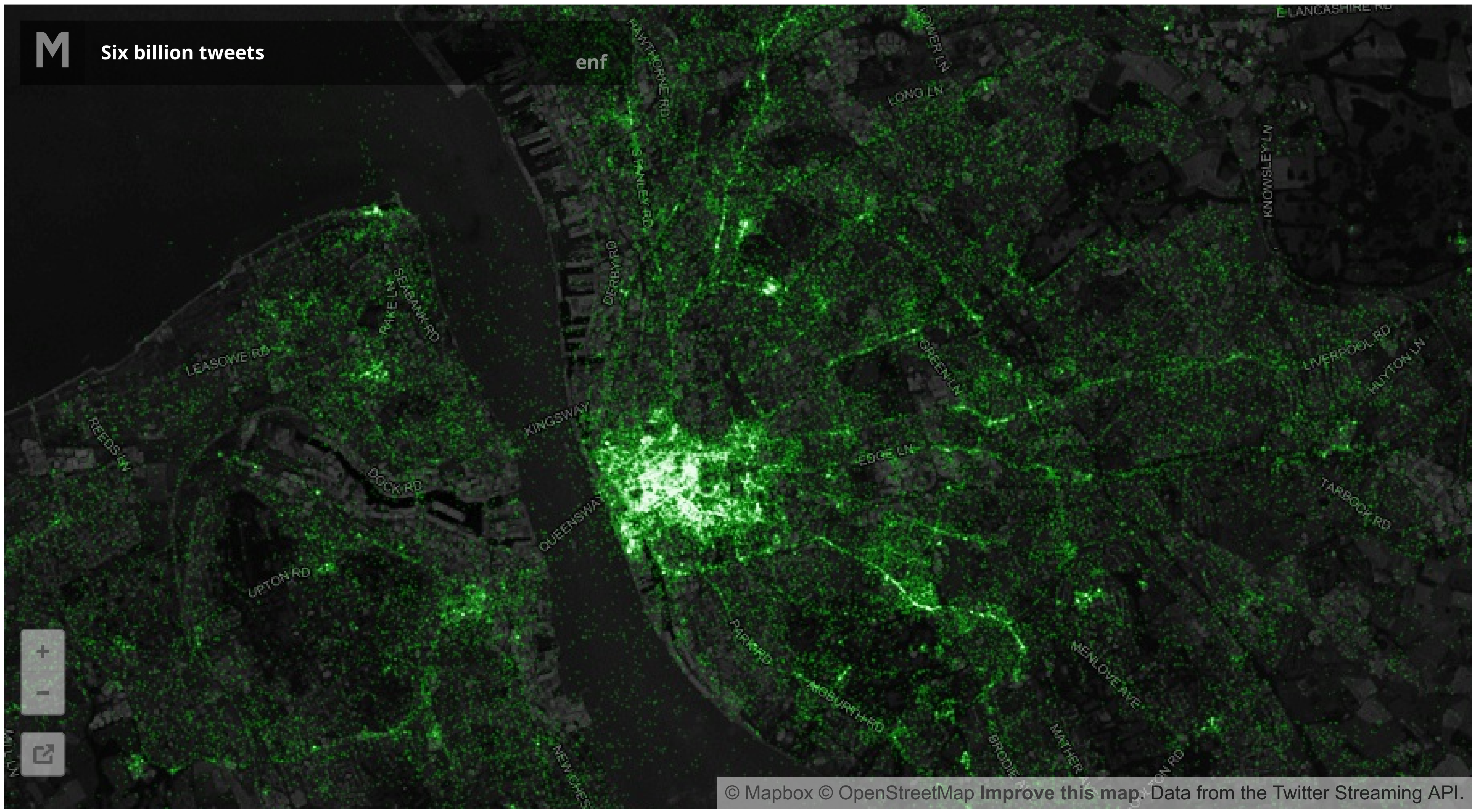

NYC Street Trees by Species
New York City's urban forest provides numerous environmental and social benefits, and street trees compose roughly one quarter of that canopy. This map shows the distribution and biodiversity of the city's street trees based on the last tree census.
Read more.
Secaucus
West New York
Union City
Weehawken
Hoboken
Newark
Jersey City
Bayonne
EWR
JFK
PULASKI SKYWAY
I 78
I 95
I 278
NY 878
BP
BOARDWALK

[Source]

UFO Sightings (1933-)



Geo-tagged tweets



Point patterns

Point patterns

Distribution of **points over** a portion of **space**

Assumption is a point can happen anywhere on that space, but only happens in specific locations

Point patterns

Distribution of **points over** a portion of **space**

Assumption is a point can happen anywhere on that space, but only happens in specific locations

- **Unmarked**: locations only
- **Marked**: values attached to each point

Point Pattern Analysis

Describe, characterize, and explain point patterns, focusing on their **generating process**

- Visual exploration
- Clustering properties
- Statistical modeling of the underlying processes

Visualization of PPs

Visualization of PPs

Two routes (today):

- *Aggregate*
- *Smooth*

Visualization of PPs

Two routes (today):

- *Aggregate* $\leftrightarrow$ "Histogram"
- *Smooth*

Visualization of PPs

Two routes (today):

- *Aggregate* $\leftrightarrow$ "Histogram"
- *Smooth* $\leftrightarrow$ KDE

Aggregation

Points meet polygons

Points meet polygons

Use **polygon** boundaries and **count** points per area

Points meet polygons

Use **polygon** boundaries and **count** points per area

[Insert your skills for **choropleth mapping** here!!!]

Points meet polygons

Use **polygon** boundaries and **count** points per area

[Insert your skills for **choropleth mapping** here!!!]

But,

Points meet polygons

Use **polygon** boundaries and **count** points per area

[Insert your skills for **choropleth mapping** here!!!]

But, the polygons need to *"make sense"* (their delineation needs to relate to the point generating process)

Aggregation

[Figures w/ points in polygons --> Choropleths of counts]

Hex-binning

If **no** polygon boundary seems like a **good candidate** for aggregation...

Hex-binning

If **no** polygon boundary seems like a **good candidate** for aggregation...

...draw a **hexagonal** (or squared) **tesselation**!!!

Hex-binning

If **no** polygon boundary seems like a **good candidate** for aggregation...

...draw a **hexagonal** (or squared) **tesselation**!!!

Hexagons...

- Are **regular**
- **Exhaust** the space (Unlike circles)
- Have **many sides** (minimize boundary problems)

Hex-binning

[Figure]

But…

But...

(Arbitrary) aggregation may induce **MAUP** (see Lecture 4)

But...

(Arbitrary) aggregation may induce **MAUP** (see Lecture 4)

+

But...

(Arbitrary) aggregation may induce **MAUP** (see Lecture 4)

+

Points usually represent events that affect to only **part** of the population and hence are best considered as **rates** (see Lecture 4)

Kernel Density Estimation

Kernel Density Estimation

*Estimate the (**continuous**) observed **distribution** of a variable*

Kernel Density Estimation

*Estimate the (**continuous**) observed **distribution** of a variable*

- Probability of finding an observation at a given point
- "Continuous histogram"
- Solves (much of) the MAUP problem, but not the underlying population issue

[Figure explaining the intuition of a KDE]

Bivariate (spatial) KDE

Probability of finding observations at a given point in space

- **Bivariate** version: distribution of **pairs of values**
- In **space**: values are coordinates (XY), locations
- Continuous "version" of a choropleth

Spatial KDE

[Figure with points over the kernel]

Recapitulation

- **Points** can be understood as a **fixed** or **random** process over space
- If seen as a random, *where* points are located is part of the interest in the (**point pattern**) analysis
- **Visualization** of point patterns can be done through **aggregation** or **smoothing** (but issues relating to the MAUP and underlying populations need to be kept in mind!)