

Big data visual analytics systems

CS524: Big Data Visualization & Analytics

Fabio Miranda

<https://fmiranda.me>

Visual analytics

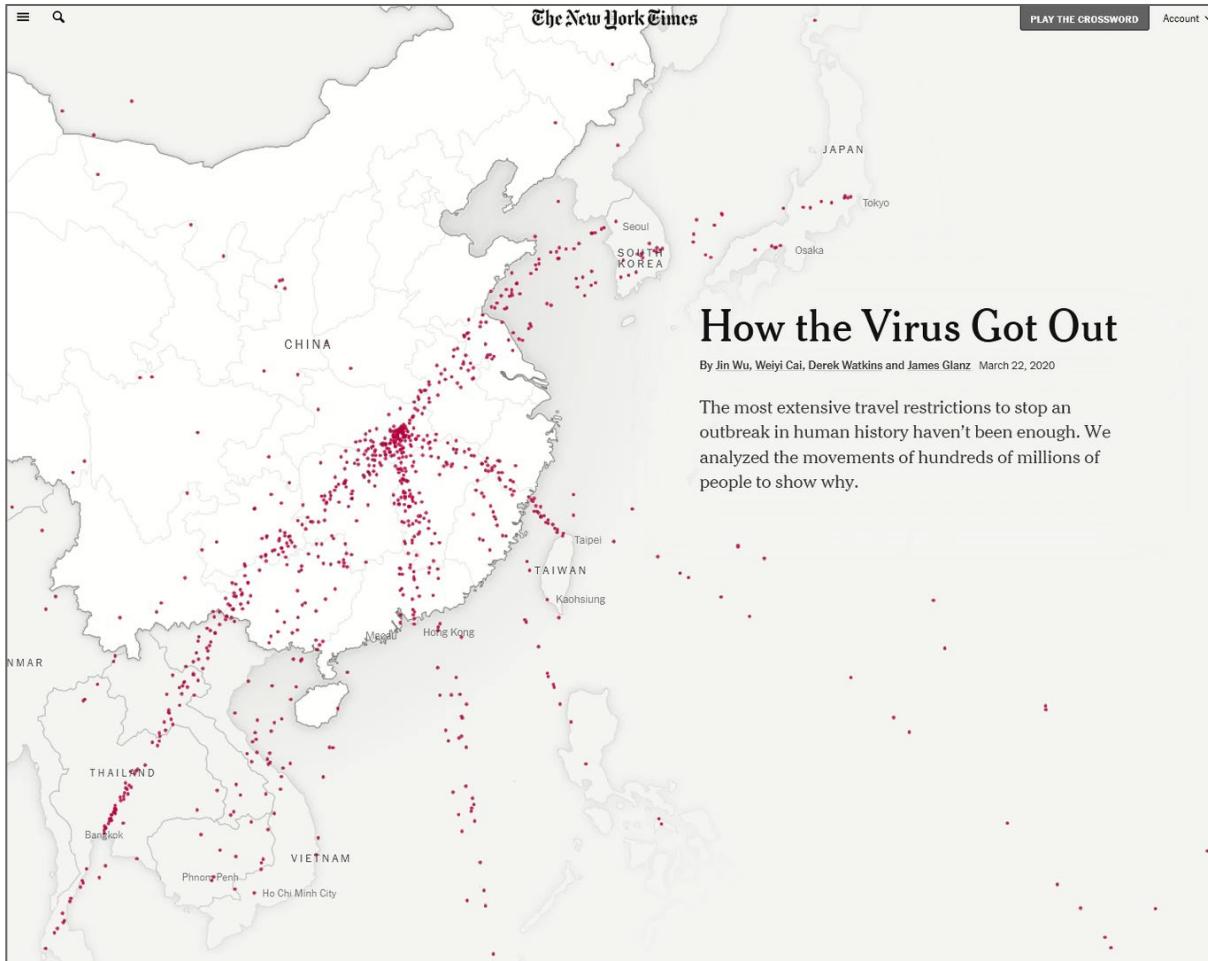
“Visual analytics is the formation of abstract visual metaphors in combination with a human information interaction that enables detection of the expected and discovery of the unexpected within massive, dynamically changing information spaces.”

[Wong and Thomas, 2004]

Visual analytics, infovis, scivis

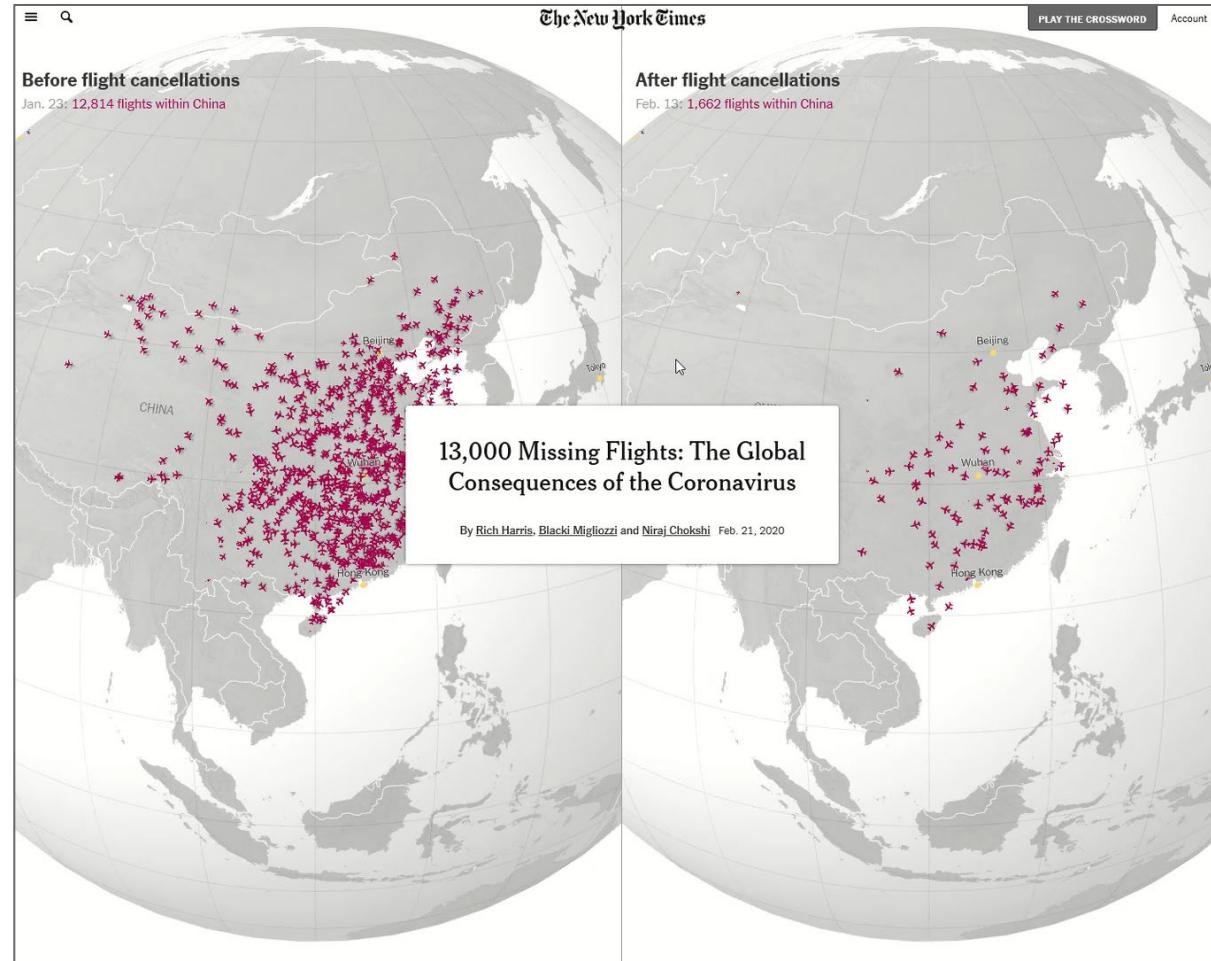
- Scientific visualization: data that has a natural geometric structure (wind flows, MRI).
- Information visualization: abstract representations (trees, graphs).
- Visual analytics: interactive visual representations and underlying analytical processes (data mining, stats, etc.)

Example: communicating findings



<https://www.nytimes.com/interactive/2020/03/22/world/coronavirus-spread.html>

Example: communicating findings



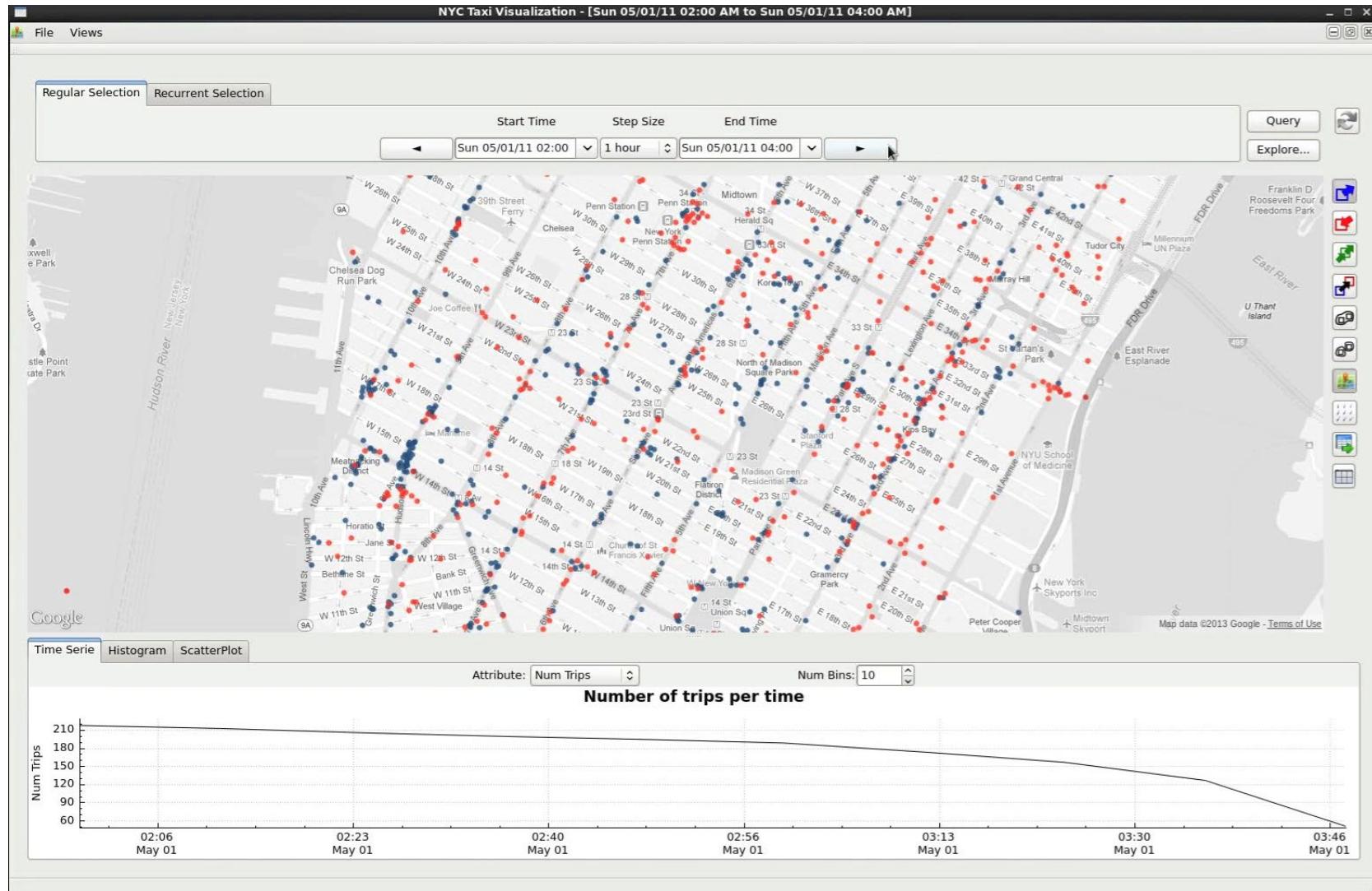
<https://www.nytimes.com/interactive/2020/02/21/business/coronavirus-airline-travel.html>

Example: exploration + communicating findings

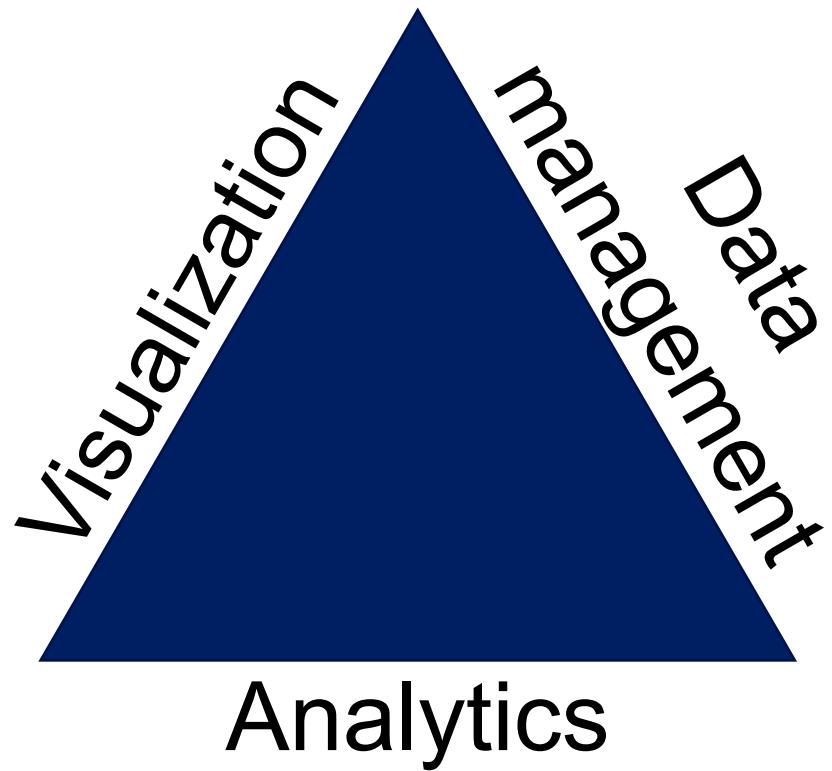


<https://www.nytimes.com/interactive/2016/12/21/upshot/Mapping-the-Shadows-of-New-York-City.html>

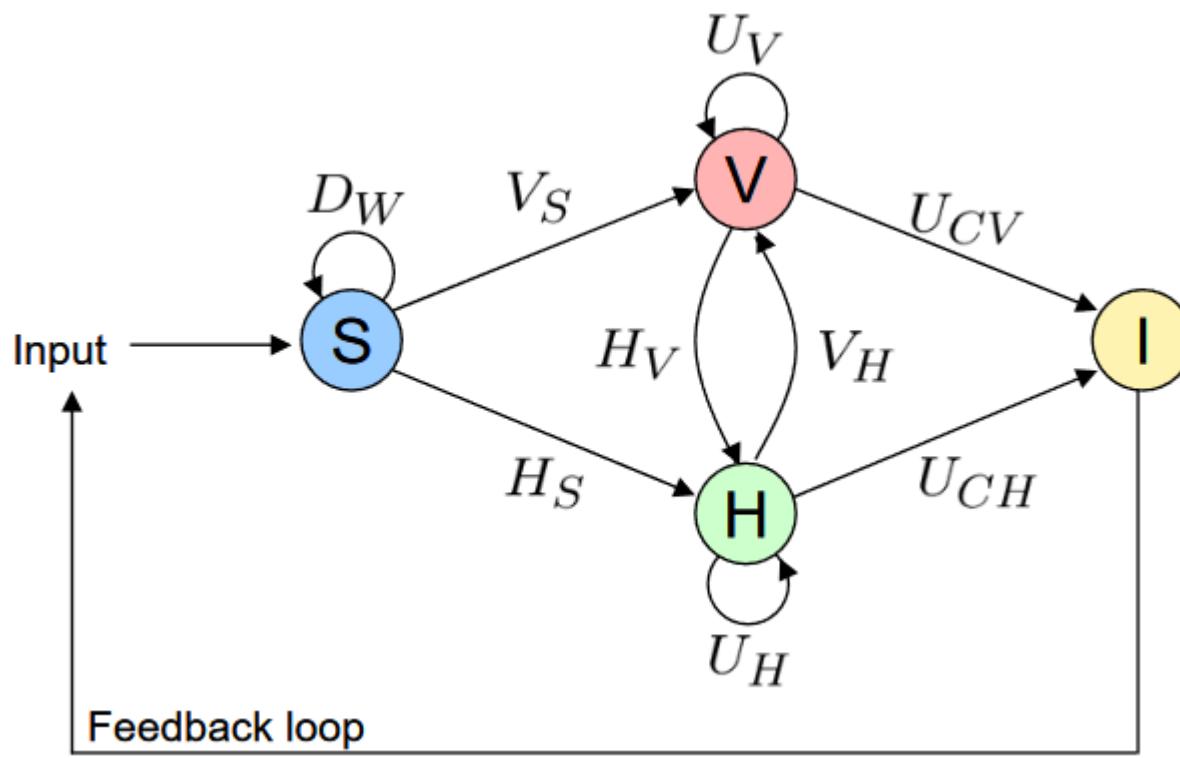
Example: visual analysis of data



Visual analytics



Visual analytics process



[“Visual Analytics: Scope and Challenges”, Kleim et al., 2008]

$$\begin{aligned} F: S &\rightarrow I \\ f \in \{D_w, V_x, H_y, U_z\} \\ S: \text{data}, I: \text{insight} \end{aligned}$$

D_w : data pre-processing functionalities (data transformation, data cleaning, data selection, data integration)

$V_x, x \in \{S, H\}$: visualization functions

V_S : functions visualizing data ($S \rightarrow V$)

V_H : functions visualizing hypotheses ($H \rightarrow V$)

$H_y, y \in \{S, V\}$: hypotheses generation process

H_S : hypotheses from automatic process ($S \rightarrow H$)

H_v : hypotheses from visualization ($V \rightarrow H$)

$U_z, z \in \{V, H, CV, CH\}$: user interactions

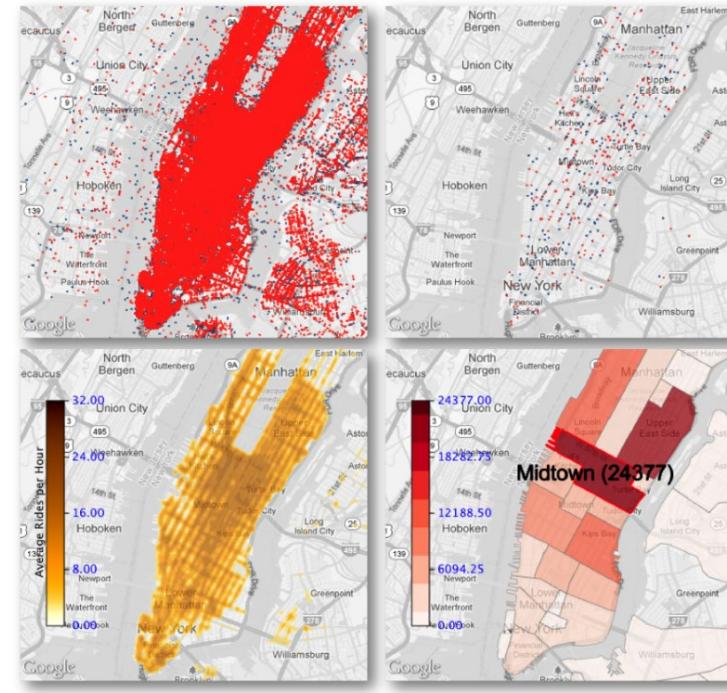
U_V : visualization interaction ($V \rightarrow V$)

U_H : hypotheses interaction ($H \rightarrow H$)

U_{cv} : insight concluded from visualization

U_{ch} : insight concluded from hypothesis

Visual analysis process: taxi data

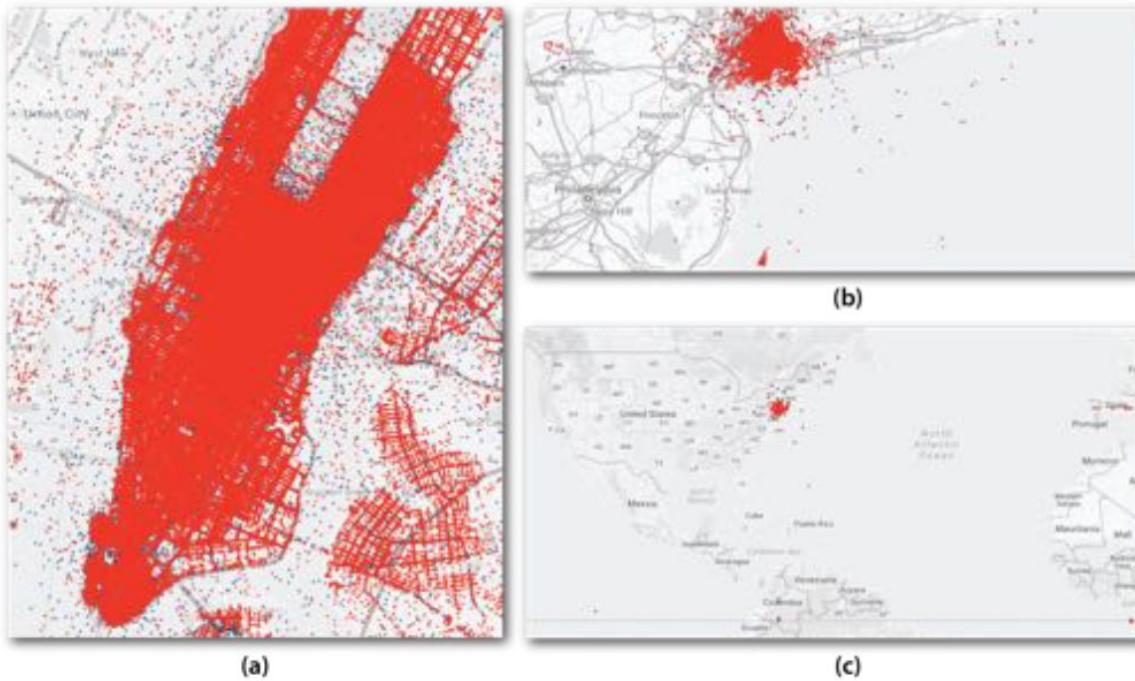


Distribution of NYC Taxi Pickups
and Dropoffs in Manhattan

Visual analysis process: taxi data

- Data: ~500,000 trips / day; ~1,000,000,000 trips in 5 years
 - Spatiotemporal: pickup + dropoff
 - Trip attributes: distance traveled, fare, tip, etc.
- Government, policy makers and scientists are usually unable to *interactive* explore the *whole* data.
 - Too many slices to examine.

Visual analysis process: taxi data

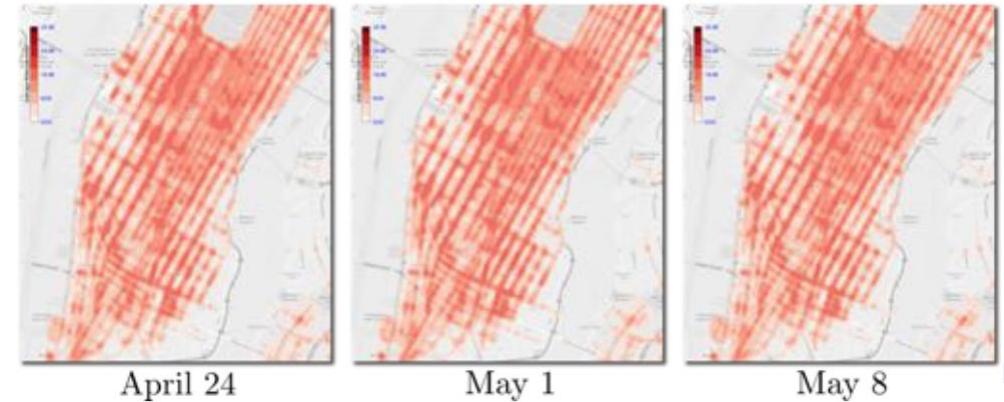
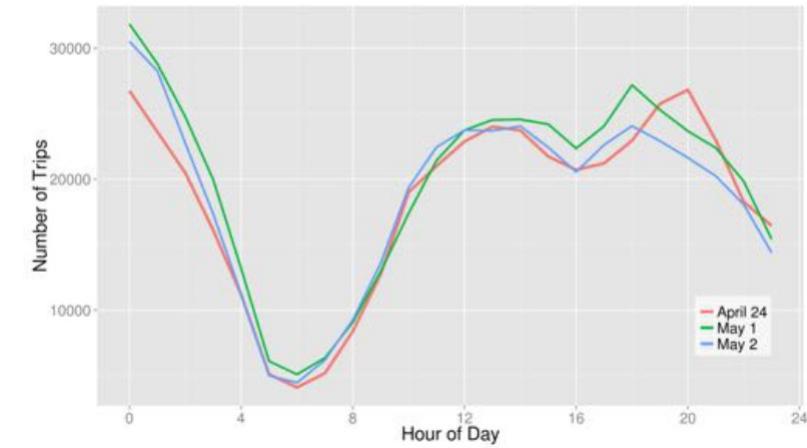


- Taxi pickups and drop-offs: data sets S_1, S_2, \dots, S_n
- Pre-processing
 - Data cleaning (e.g., remove points outside NYC): $d_c(S_1, \dots, S_n)$
 - Data integration (e.g., add weather column): $d_i(S_1, \dots, S_n)$
 - Data transformation (e.g., raw csv's to binary): $d_t(S_1, \dots, S_n)$
 - Data selection (e.g., only Manhattan trips): $d_s(S_1, \dots, S_n)$

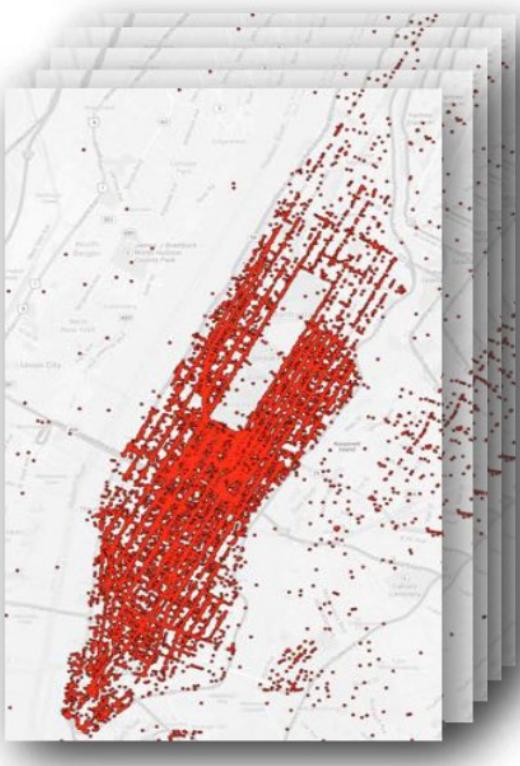
Visual analysis process: taxi data



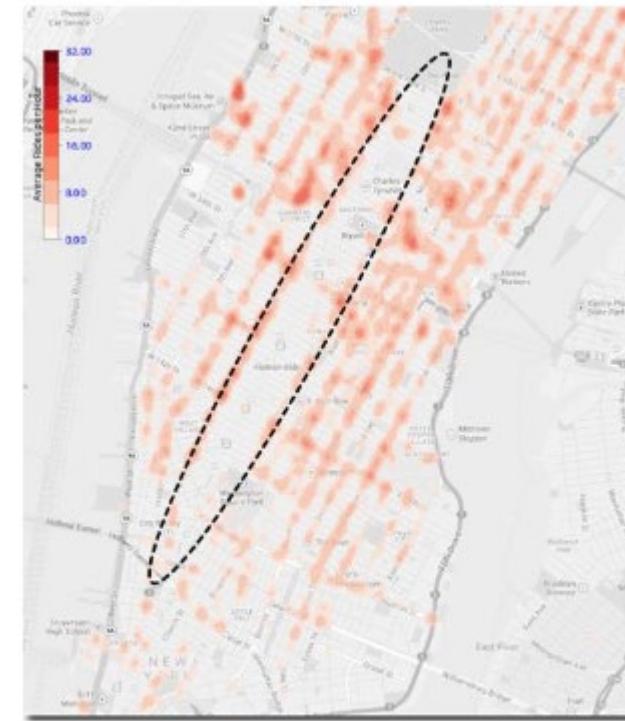
- Too many slices
- 365×24 1-hour slices in just one year
- Which slices are interesting?



Visual analysis process: taxi data

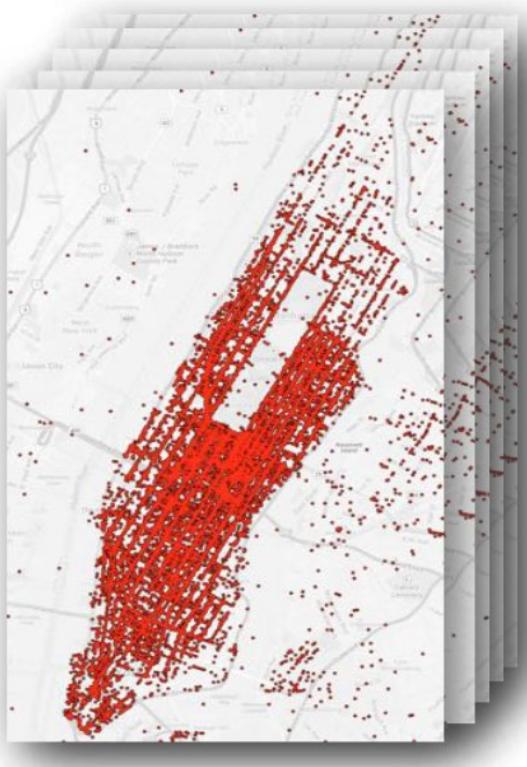


- Too many slices
- 365×24 1-hour slices in just one year
- Which slices are interesting?



May 1 (8-9am)

Visual analysis process: taxi data



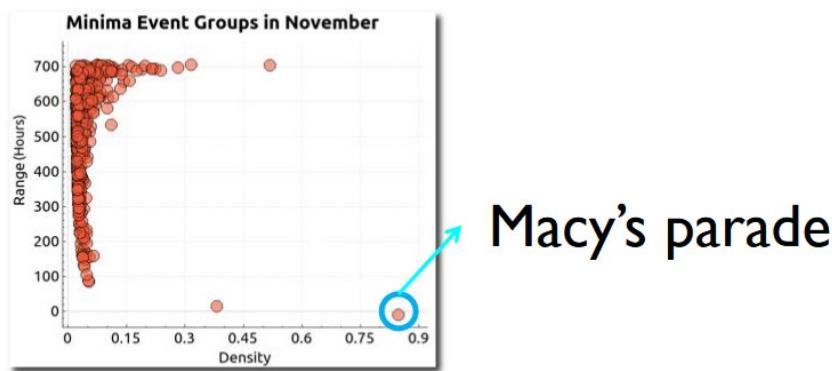
- Too many slices
- 365×24 1-hour slices in just one year
- Which slices are interesting?
- Guide users towards *interesting* data slices
- Data mining:
 - Identifying events with arbitrary spatial structure at multiple temporal scales:

$$h' = h_s(S')$$

with $S' = d_s(d_t(d_i(d_c(S_1, \dots, S_n))))$

Visual analysis process: taxi data

- Visualization:
 - Visualizing the previous hypothesis: $v' = v_h(h')$
 - Adjusting parameters: $v'' = u_v(v')$
 - Hypotheses from visualization: $h'' = h_v(v'')$, leading to an insight (unusual taxi activity around Macy's parade).



Visual analysis process: taxi data

- Taxi pickups and drop-offs: data sets S_1, S_2, \dots, S_n
- Pre-processing
 - Data cleaning (e.g., removing): $d_c(S_1, \dots, S_n)$
 - Data integration (e.g., adding extra attributes): $d_i(S_1, \dots, S_n)$
 - Data transformation (e.g., format suitable for analysis): $d_t(S_1, \dots, S_n)$
 - Data selection (e.g., subset of sensors): $d_s(S_1, \dots, S_n)$
- Data mining:
 - Identifying events with arbitrary spatial structure at multiple temporal scales:
$$h' = h_s(S'), \text{with } S' = d_s\left(d_t\left(d_i\left(d_c(S_1, \dots, S_n)\right)\right)\right)$$
- Visualization:
 - Visualizing the previous hypothesis: $v' = v_h(h')$
 - Adjusting parameters: $v'' = u_v(v')$
 - Hypotheses from visualization:
$$h'' = h_v(v'')$$

$$h_v(u_v(v_h(h_s(d_s(d_t(d_i(d_c(S_1, \dots, S_n))))))))$$

Visual analytics mantra

- Visualization mantra:

“Overview first, zoom / filter, details on demand”

[Shneiderman, 1996]

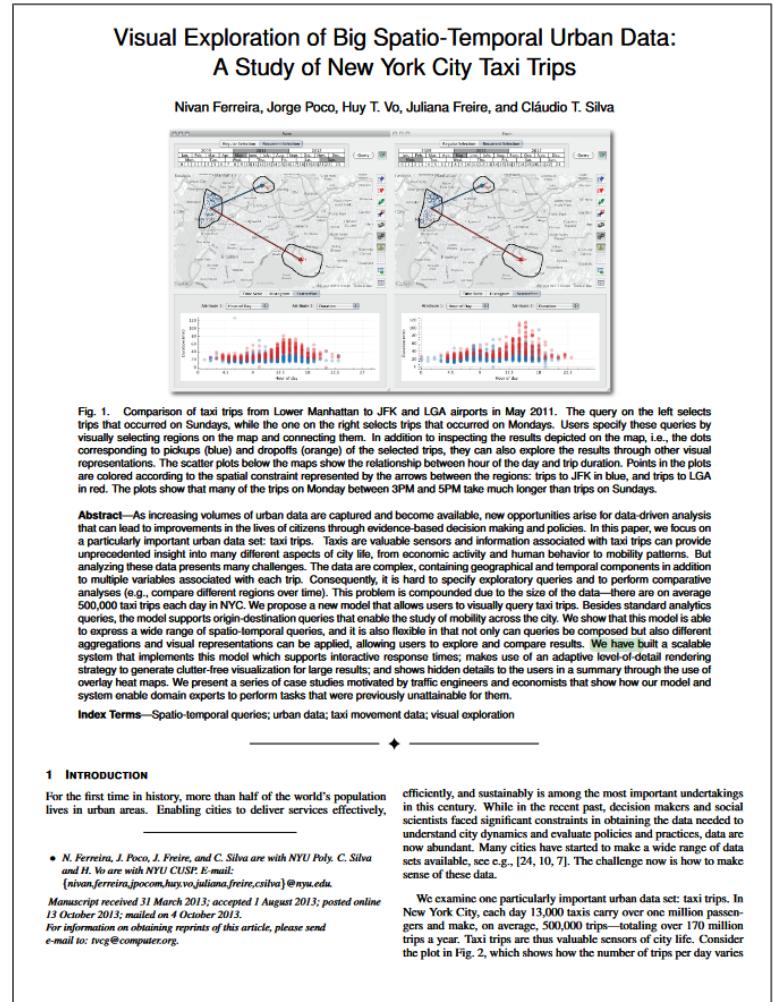
- Visual analytics mantra:

“Analyze first - show the important - zoom, filter and analyze further - details on demand”

[Klein et al., 2008]

Visual analytics + big data

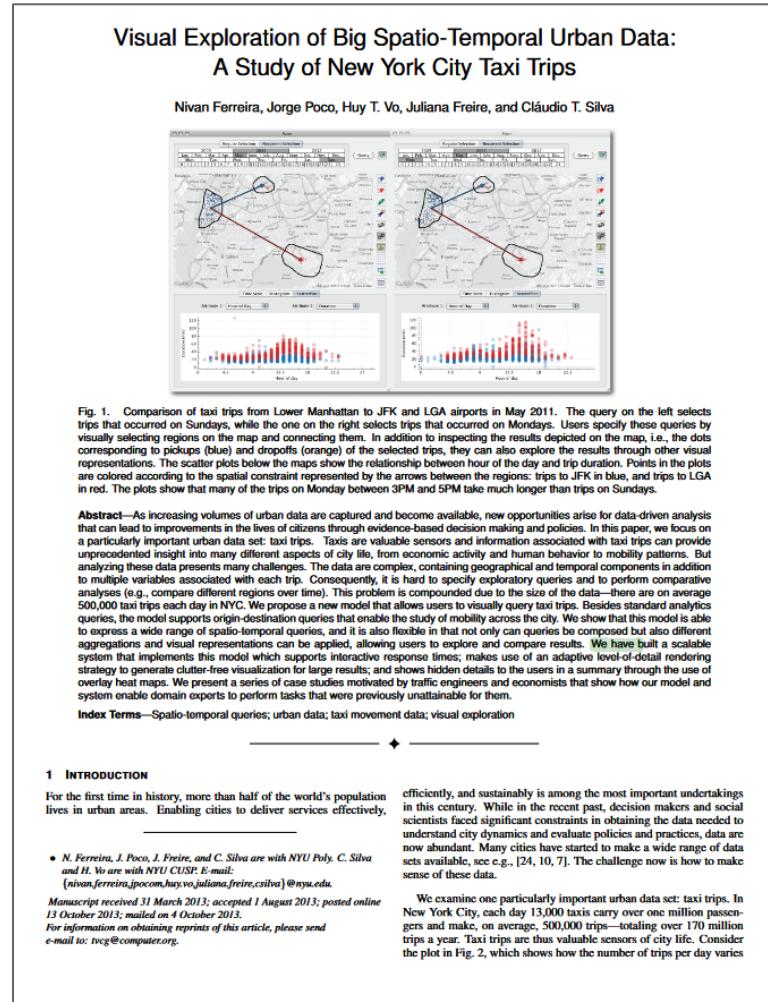
“We propose a new model that allows users to visually query taxi trips. Besides standard analytics queries, the model supports origin-destination queries that enable the study of mobility across the city. We show that this model is able to express a wide range of spatiotemporal queries, and it is also flexible in that not only can queries be composed but also different aggregations and visual representations can be applied, allowing users to explore and compare results.”



[Ferreira et al., 2013]

TaxiVis

- Interaction capabilities that enable users to pose queries over all the dimensions of the data and flexibly explore the associated attributes.
- Compare spatiotemporal slices through multiple coordinated views.
- Interactively compose and refine queries, and generalize them by performing parameter sweeps.
- Efficient data stored, and adaptive level-of-detail rendering to provide clutter-free visualizations.

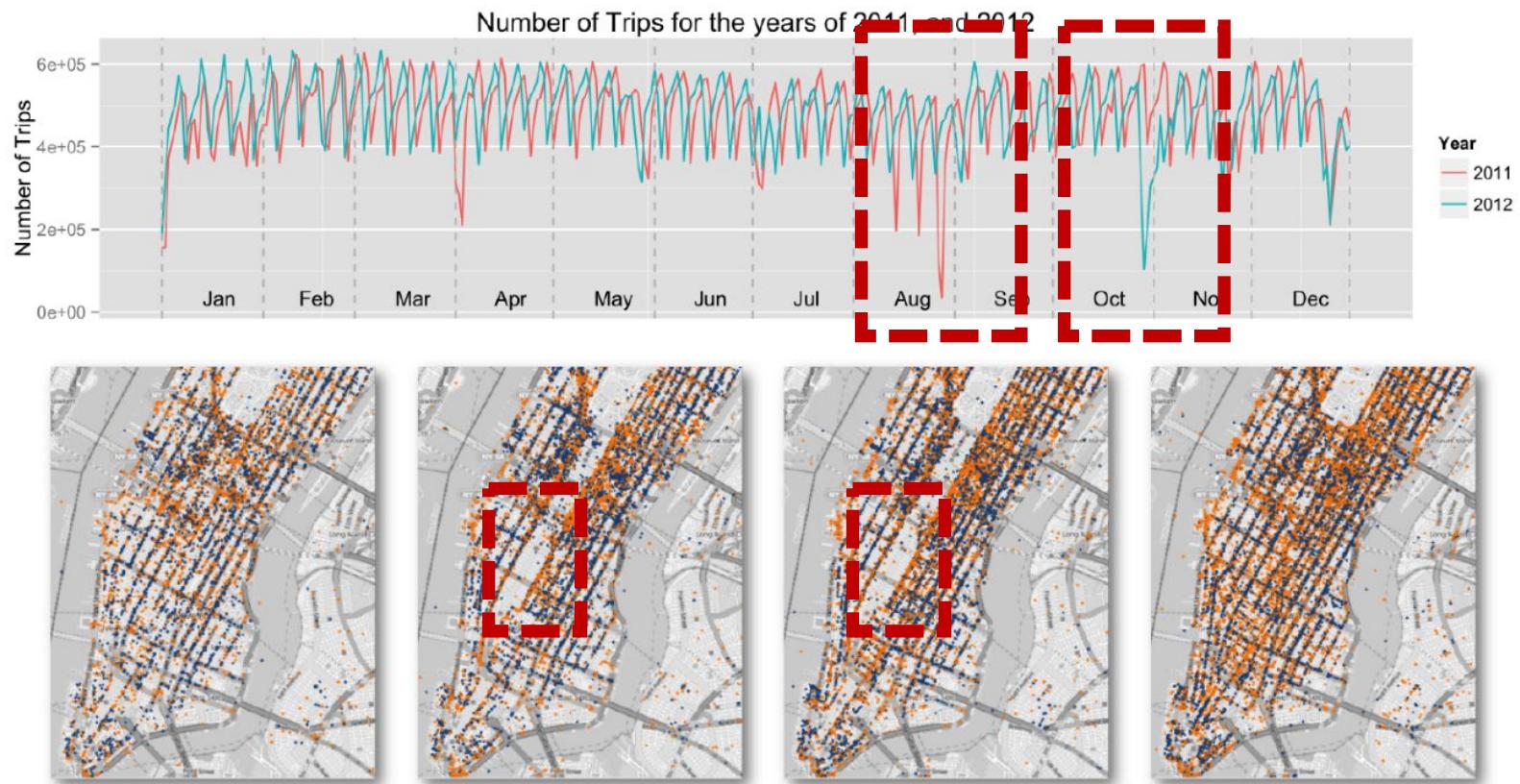


[Ferreira et al., 2013]

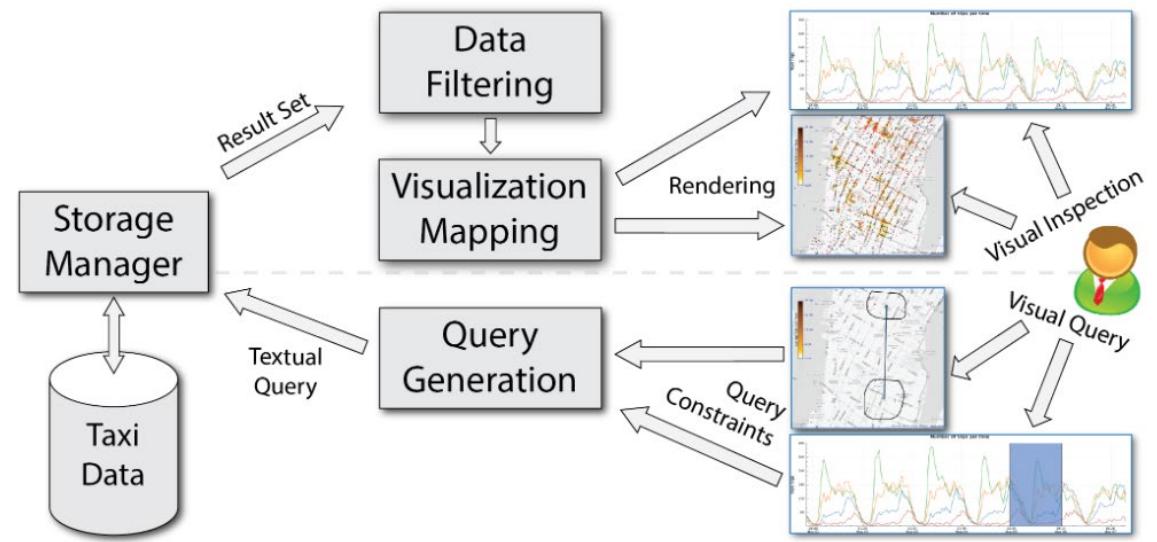
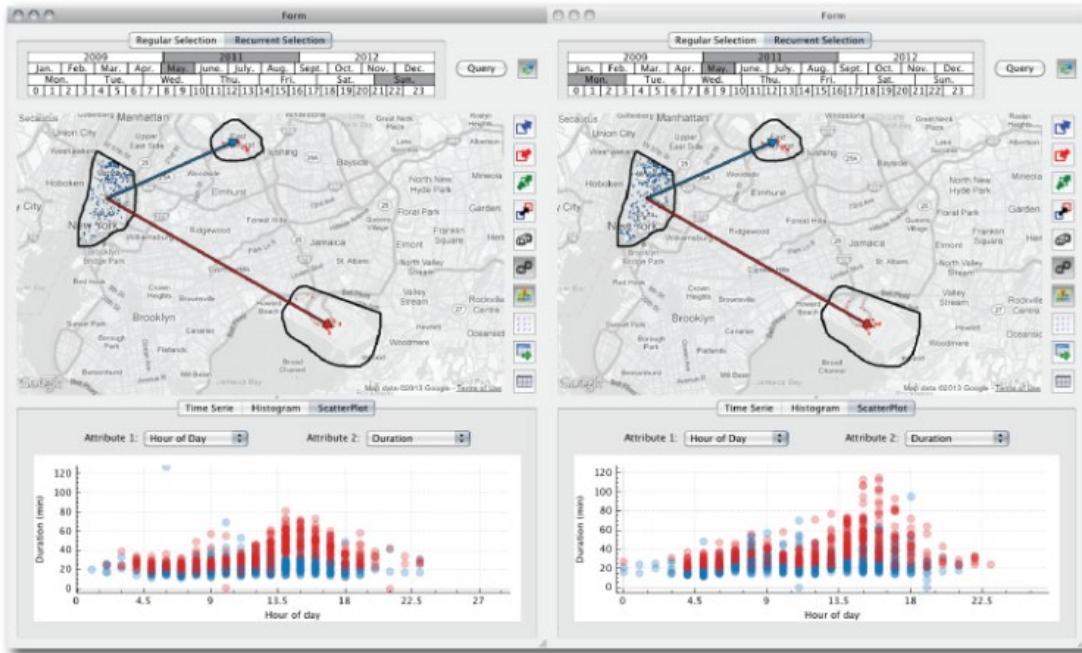
Taxi as sensors of city life

- Understanding the dynamics of the city, how different aspects of the data vary over space and time:
 - *“What is the average trip time from Midtown to the airports during weekdays?”*
 - *“How does the movement changes between Midtown and JFK throughout the day, over different days of the week?”*
 - *“How does the taxi fleet activity vary during weekdays?”*
- Ability to quickly test hypotheses: starting with one query about a specific place (*“movement patterns between Midtown and JFK”*) and generalize to all neighborhoods.

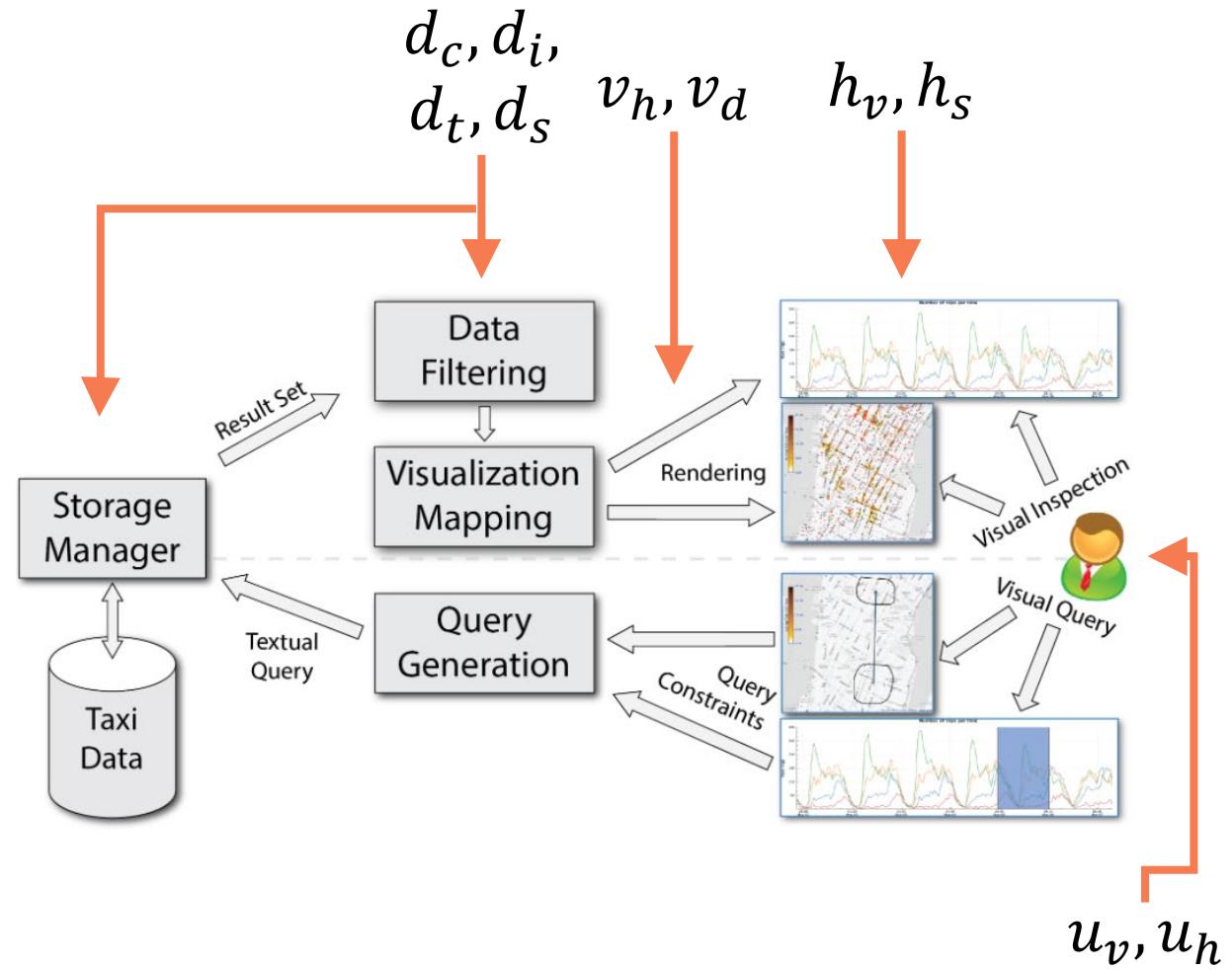
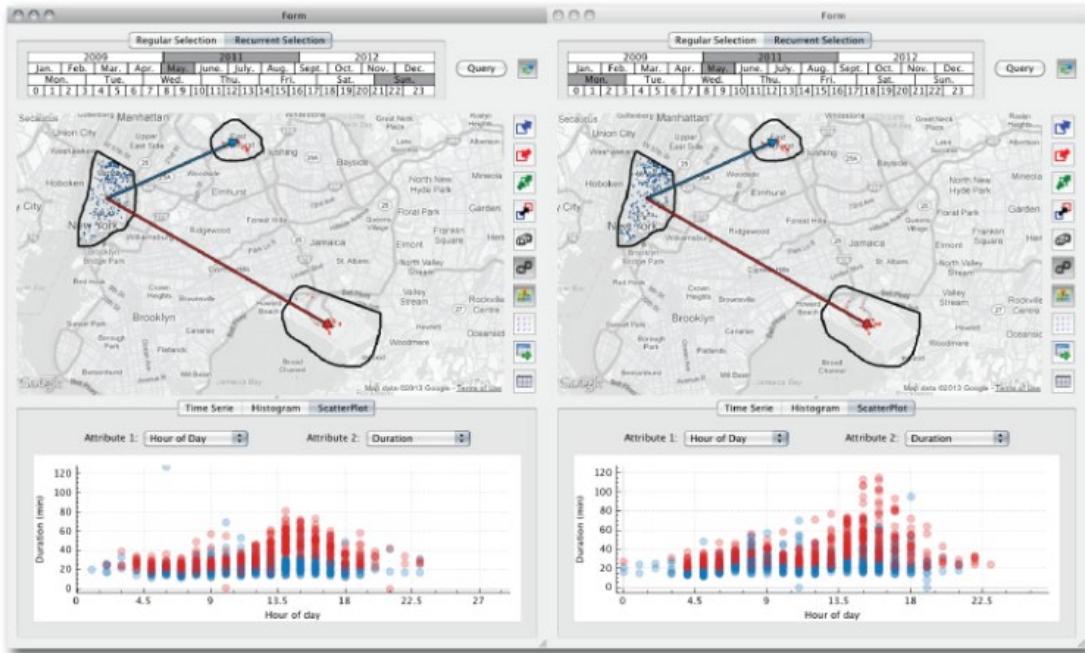
Taxis as sensors of city life



TaxiVis



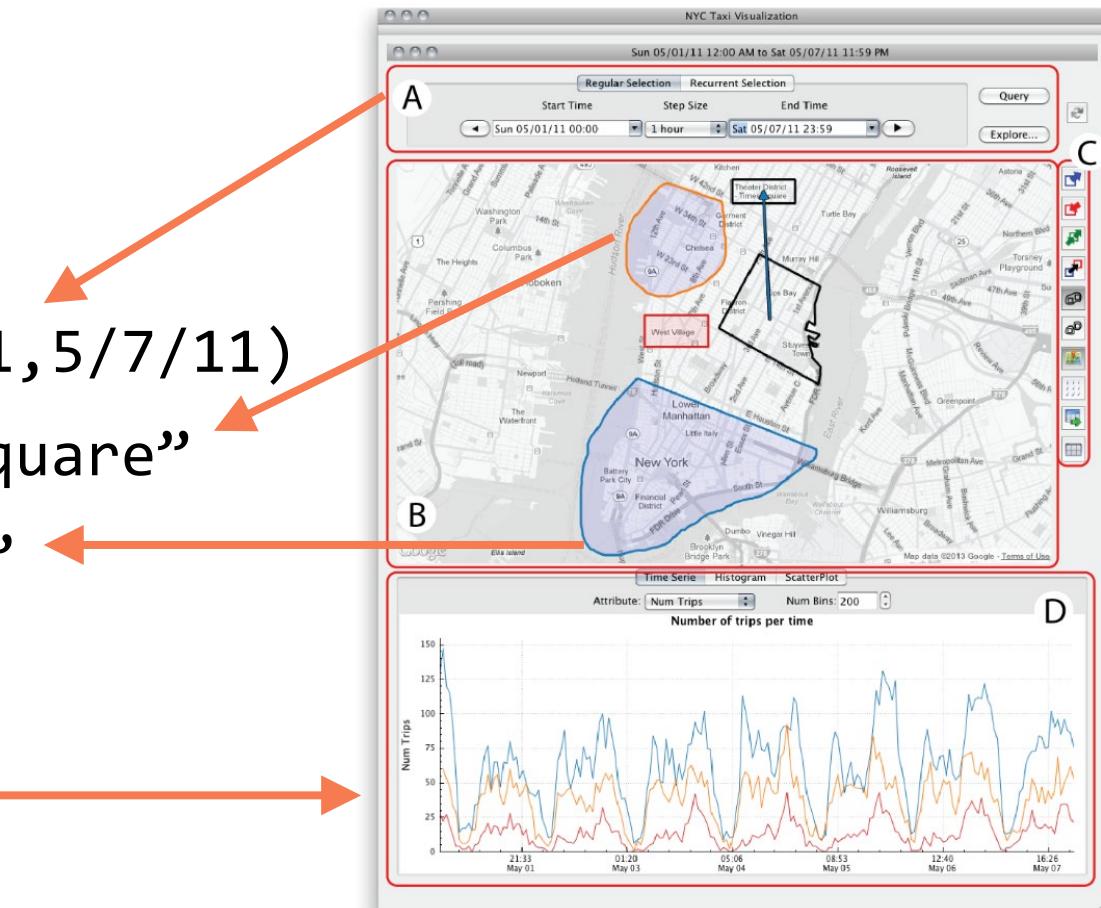
TaxiVis



Usability through visual operations

```
SELECT *  
FROM trips  
WHERE pickup_time in (5/1/11,5/7/11)  
AND dropoff_loc in “Times Square”  
AND pickup_loc in “Gramercy”
```

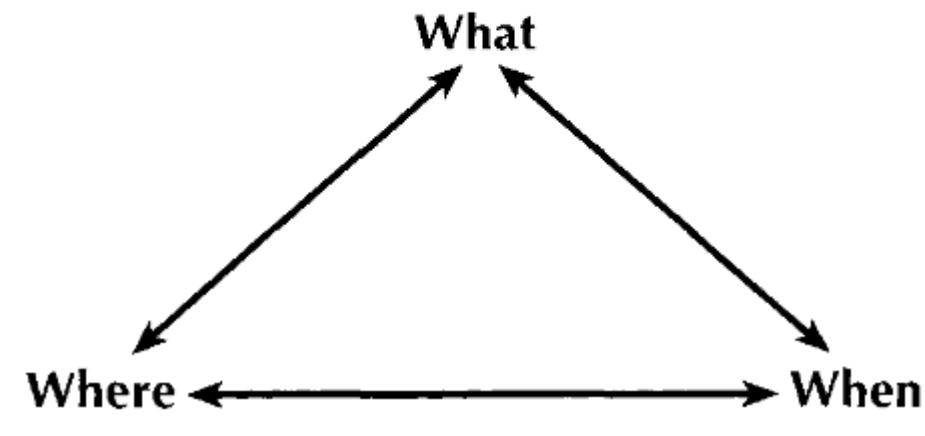
Data selection and result exploration are unified



Visual query model

- Expressive Triad framework:
 - *Where + when → what*: describe the objects (what) that are present at certain locations and times.
 - *When + what → where*: describe the locations occupied by objects at a given time.
 - *Where + what → when*: describe the times that objects occupied a given location.

THE BASIC VIEW COMPONENTS OF THE TRIAD FRAMEWORK



[Peuquet, 1994]

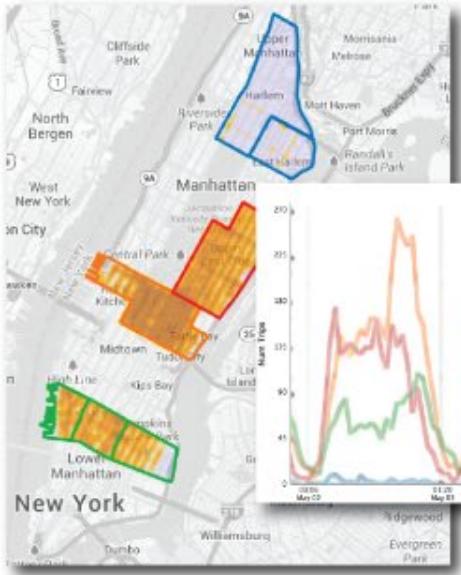
Visual query model

- Expressive Triad framework:
 - *Where + when → what*: “What is the average trip time from Midtown to the airports during weekdays?”
 - *When + what → where*: “Where are the hot spots in Manhattan in the weekends?”
 - *Where + what → when*: “When were activities restored in Lower Manhattan after the Sandy hurricane?”

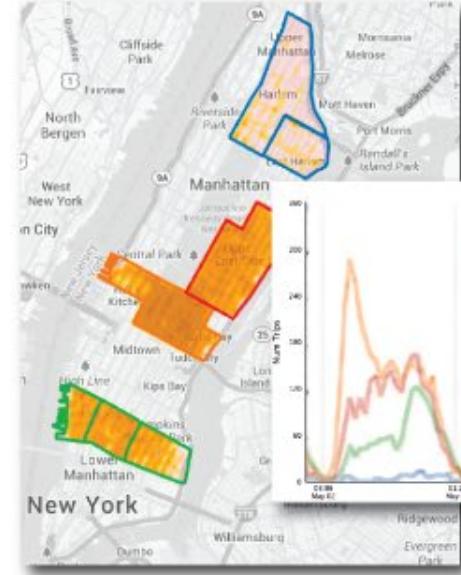
What? Where? When?



Comparison of taxi trips: characterizing neighborhoods



pickups



dropoffs

- Midtown and Upper East are the most active areas.
- Midtown is a popular destination in the morning and downtown in the night.
- Over the weekend, increase number of drop-offs in Downtown.
- Harlem is underserved by taxis.

Ability to quickly test hypotheses: starting with one query about a specific place and generalize to all neighborhoods.

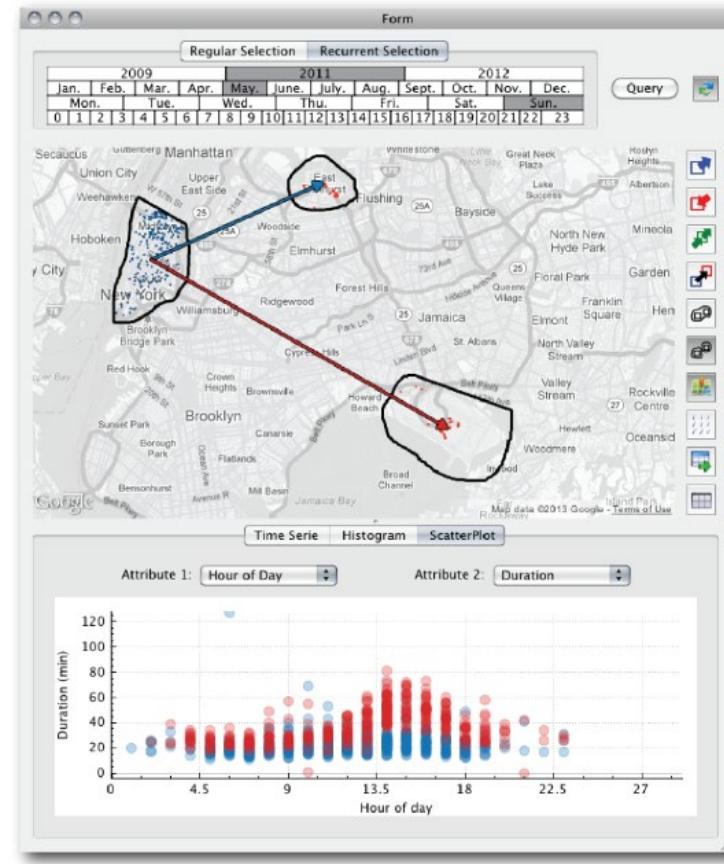
Challenge: interactive query evaluation

Typical query: find all trips that occurred between **lower Manhattan** and the two airports, **JFK** and **LGA** during **all Sundays in May 2011**.

Query time (sec)	PostgreSQL	ComDB
	503.9	20.6

“Increased latency reduces the rate at which users make observations, draw generalizations and generate hypotheses”

[Liu and Heer, 2014]

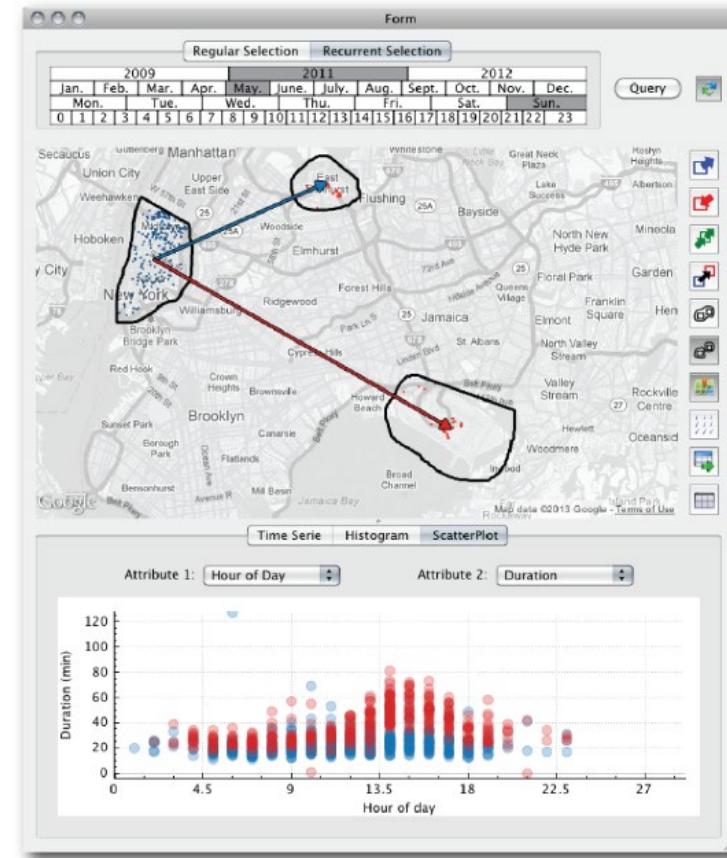


Challenge: interactive query evaluation

Typical query: find all trips that occurred between **lower Manhattan** and the two airports, **JFK** and **LGA** during **all Sundays in May 2011**.

Query time (sec)	PostgreSQL	ComDB
	503.9	20.6

Goal: support interactive queries

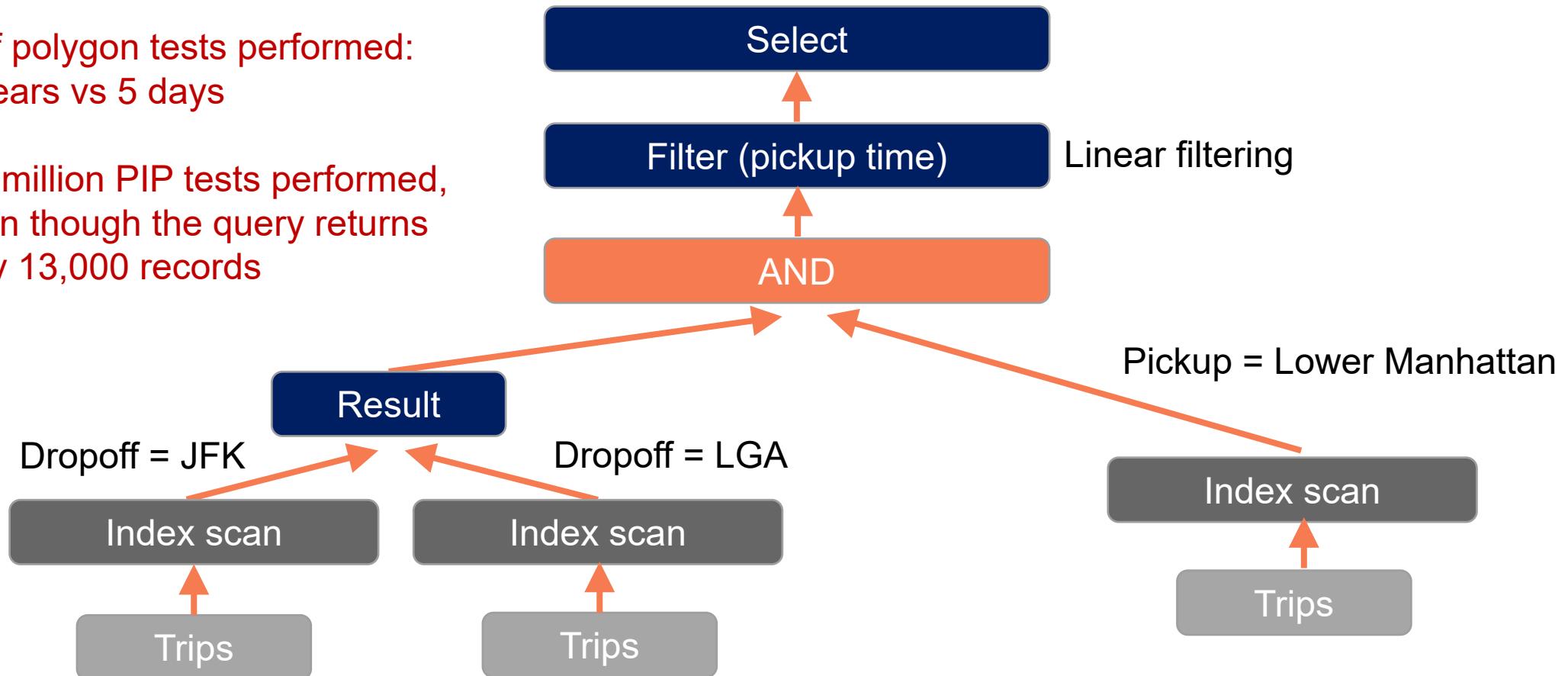


Challenge: interactive query evaluation

Typical query: find all trips that occurred between **lower Manhattan** and the two airports, **JFK** and **LGA** during **all Sundays in May 2011**.

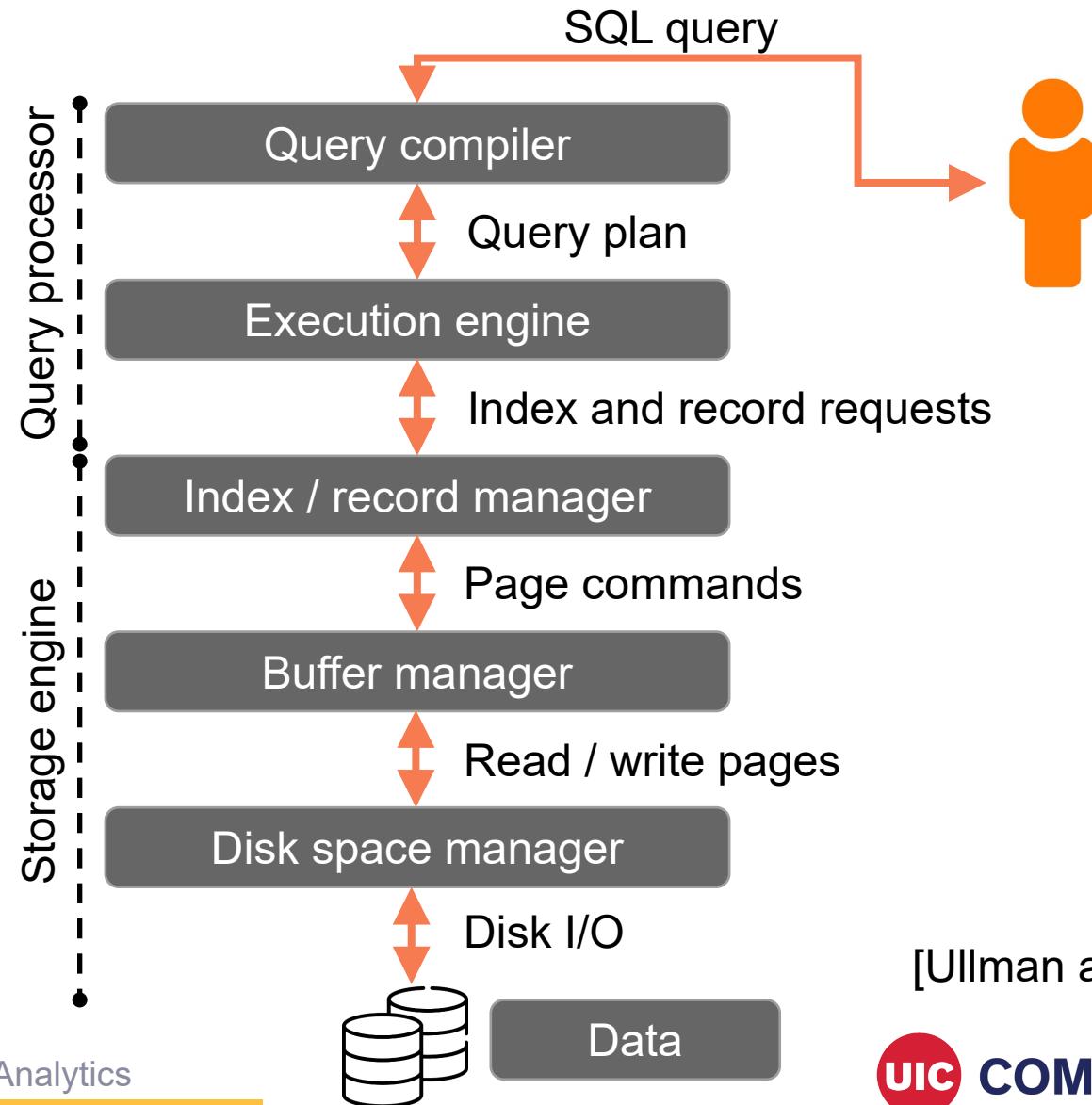
of polygon tests performed:
5 years vs 5 days

6.5 million PIP tests performed,
even though the query returns
only 13,000 records



Challenge: interactive query evaluation

Query processor:



[Ullman and Widow]

Challenge: interactive query evaluation

The diagram illustrates a challenge in interactive query evaluation. It features two orange human icons representing users. One user is interacting with a database table on the left, while the other is interacting with a database table on the right. Both interactions are triggered by an "SQL query".

Company	Units	Value
10	2	1.34
13	5	2.00
18	9	1.90
21	1	0.37
18	5	1.46
5	3	7.30
18	2	5.30

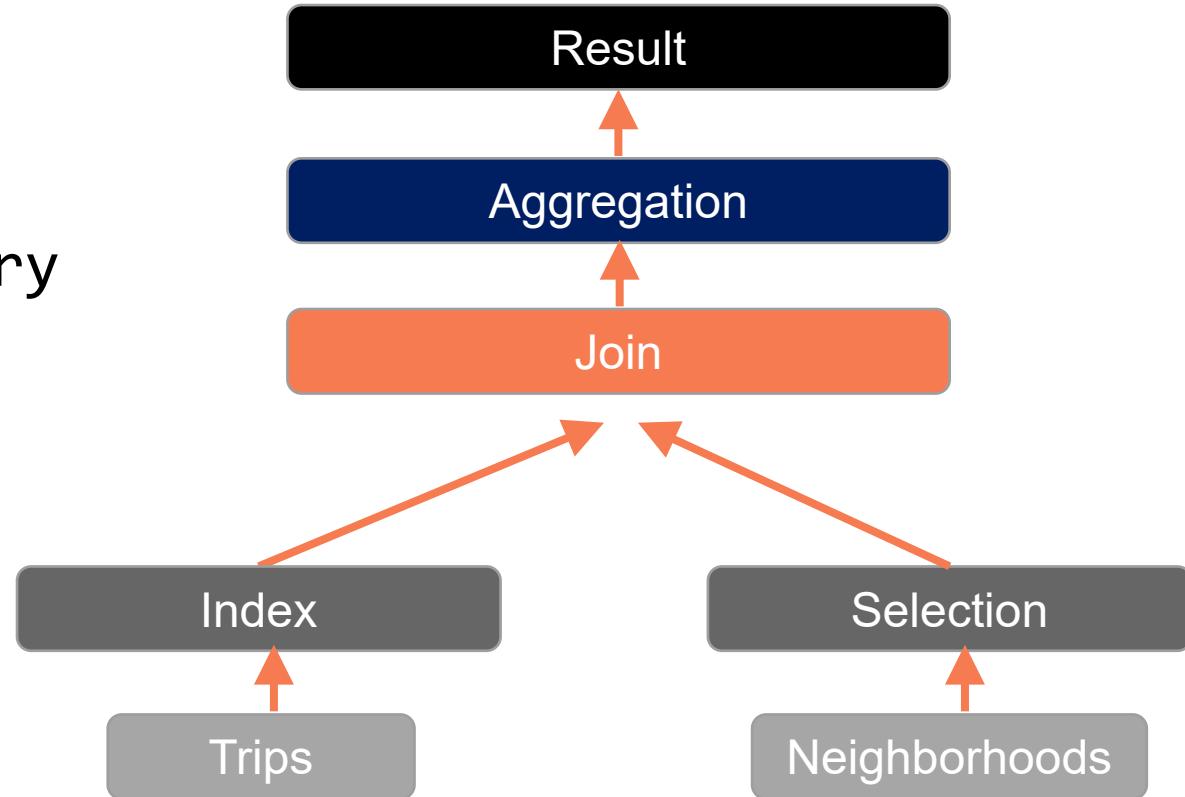
Company	Pointer
5	5
10	0
13	1
18	2
18	4
18	6
21	3

Pointer	Company	Units	Value
0	10	2	1.34
1	13	5	2.00
2	18	9	1.90
3	21	1	0.37
4	18	5	1.46
5	5	3	7.30
6	18	2	5.30

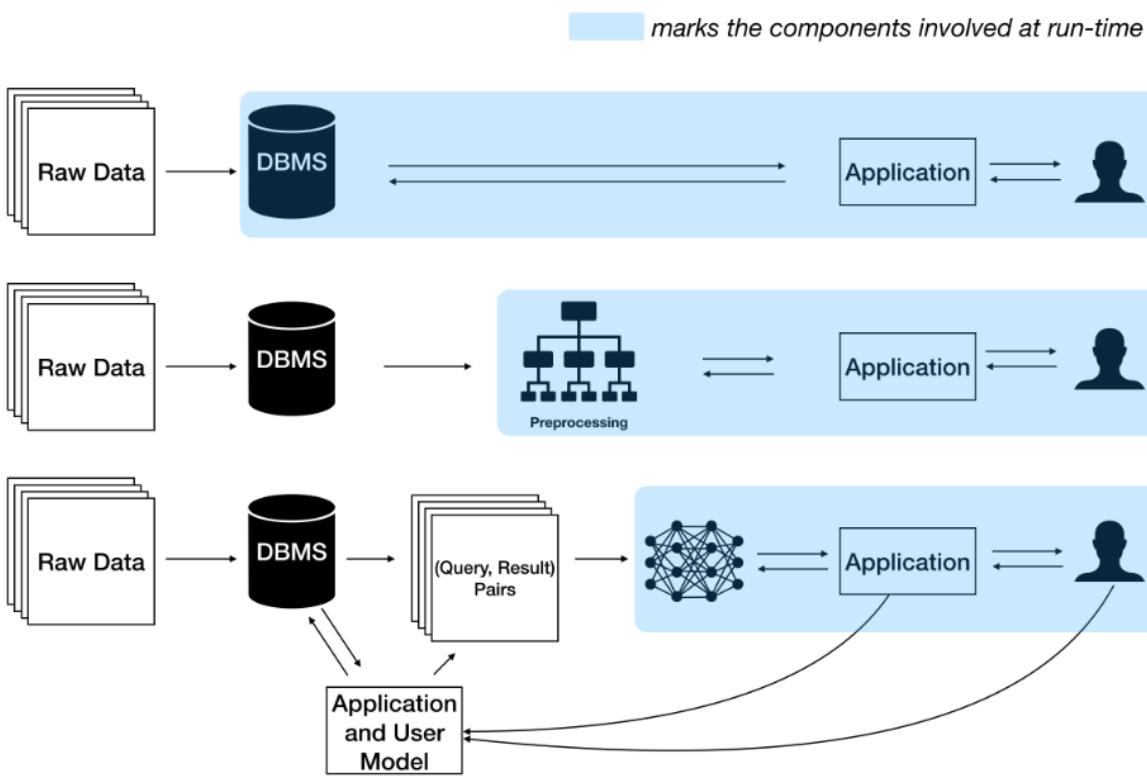
Challenge: interactive query evaluation

```
SELECT COUNT(*)  
FROM taxi T, neighborhoods N  
WHERE T.pickup INSIDE N.geometry  
GROUP BY N.id
```

- Solution in existing spatial databases
- Slow: several seconds / minutes



Challenge: interactive query evaluation

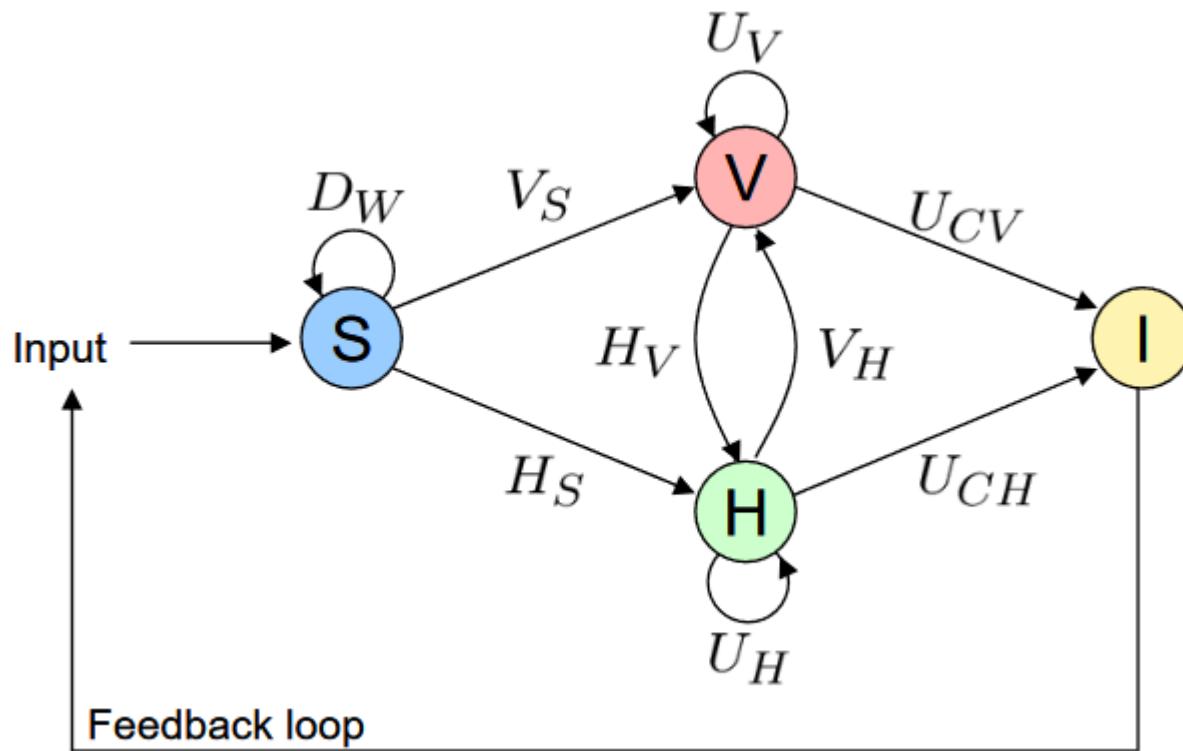


- **Solutions:**

- Pre-compute aggregations
- Use massively parallel architectures
- Approximate queries
- (all of the above)

[“NeuralCubes”, Wang et al., 2019]

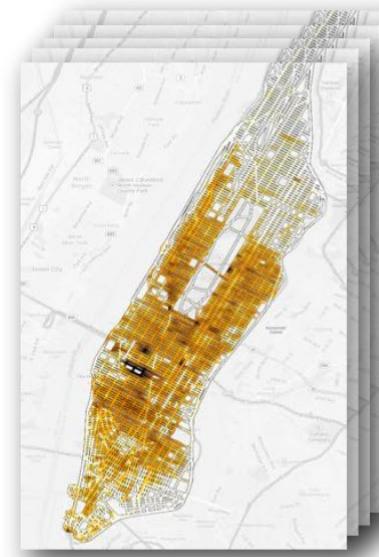
Finding interesting features



- How to tackle h_s (hypotheses from automatic methods)?
- Goal: guide users towards *interesting* data slices.

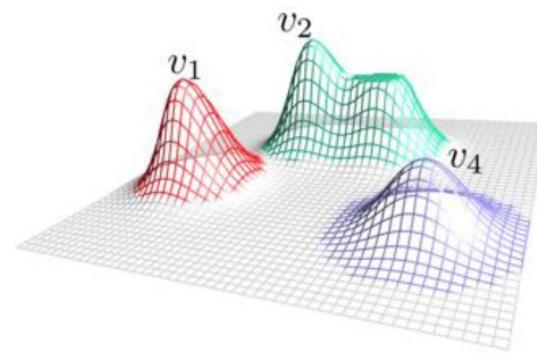
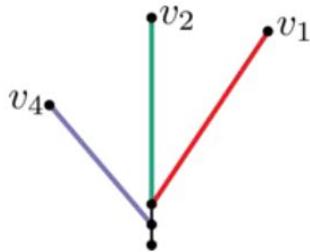
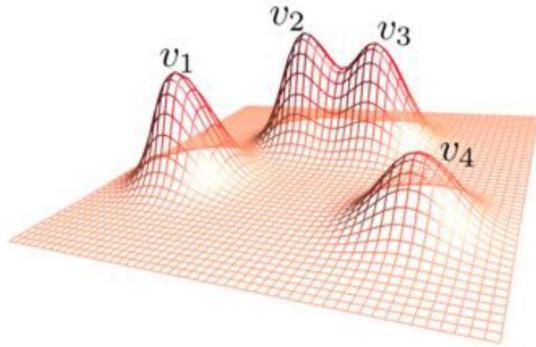
Identifying potential events

- Model data as a time-varying scalar function defined on a graph.
 - Taxi data: graph = road network; function = density of taxis
 - Subway data: graph = track network; function = delay of trains



Taxi data: potential events

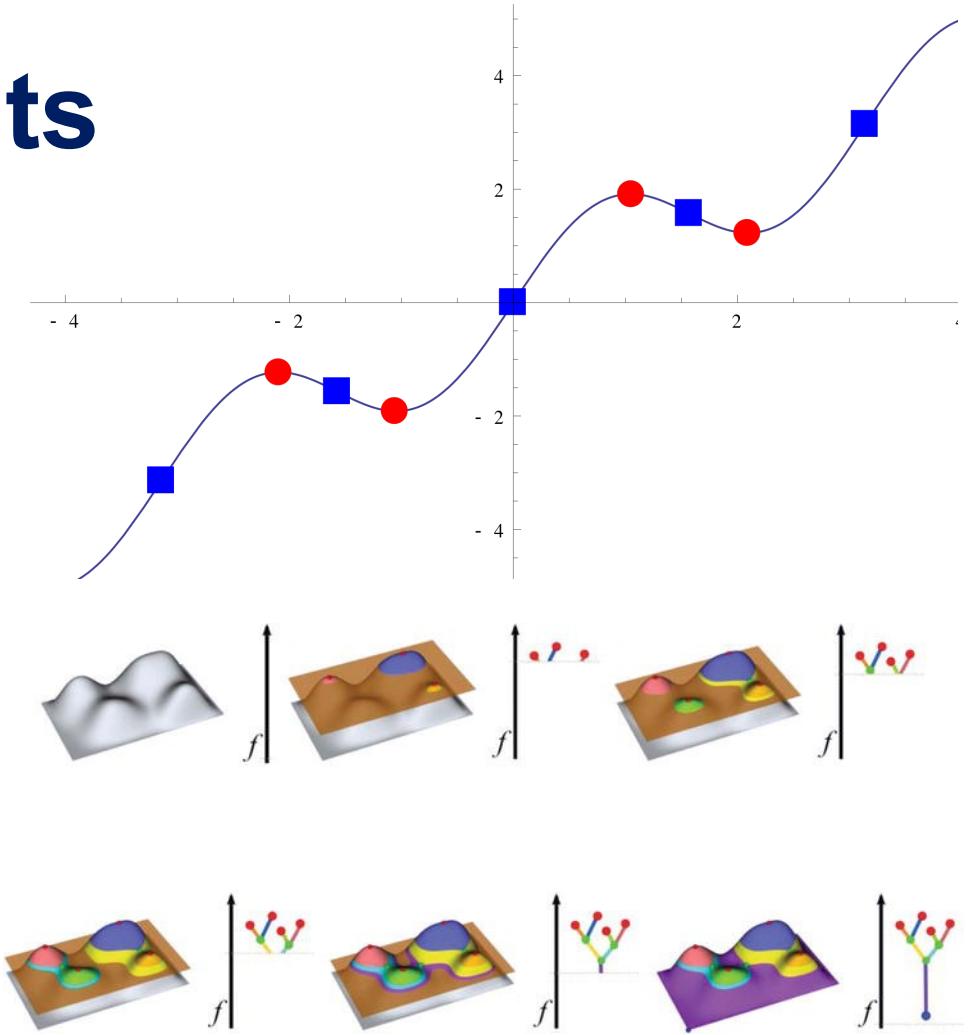
- Merge trees to efficiently identify events in each time step.
- Set of potential events: regions corresponding to maxima and minima.
 - A region is interesting if its behavior differs from its neighborhood.
 - Unimportant events can be simplified.



[Doraiswamy et al., 2014]

Taxi data: potential events

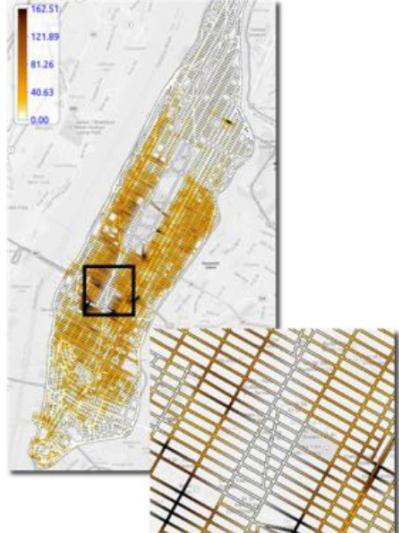
- Merge trees can be used to efficiently represent regions
- Topological changes occur at critical points (i.e., where the function is not differentiable or derivative equal to zero).
- Trees can be simplified to remove noise.
- Merge trees track level sets and capture local maxima as function height is reduced.



[Thomas et al., 2018]

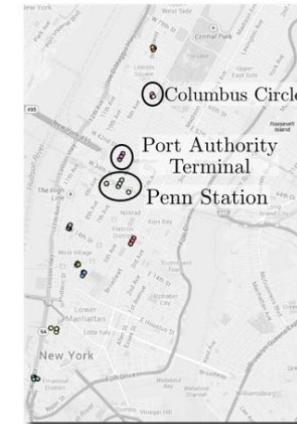
Taxi data: potential events

- Minima: lack of taxis
 - Density lower than local neighborhood.
 - Can denote road blocks.

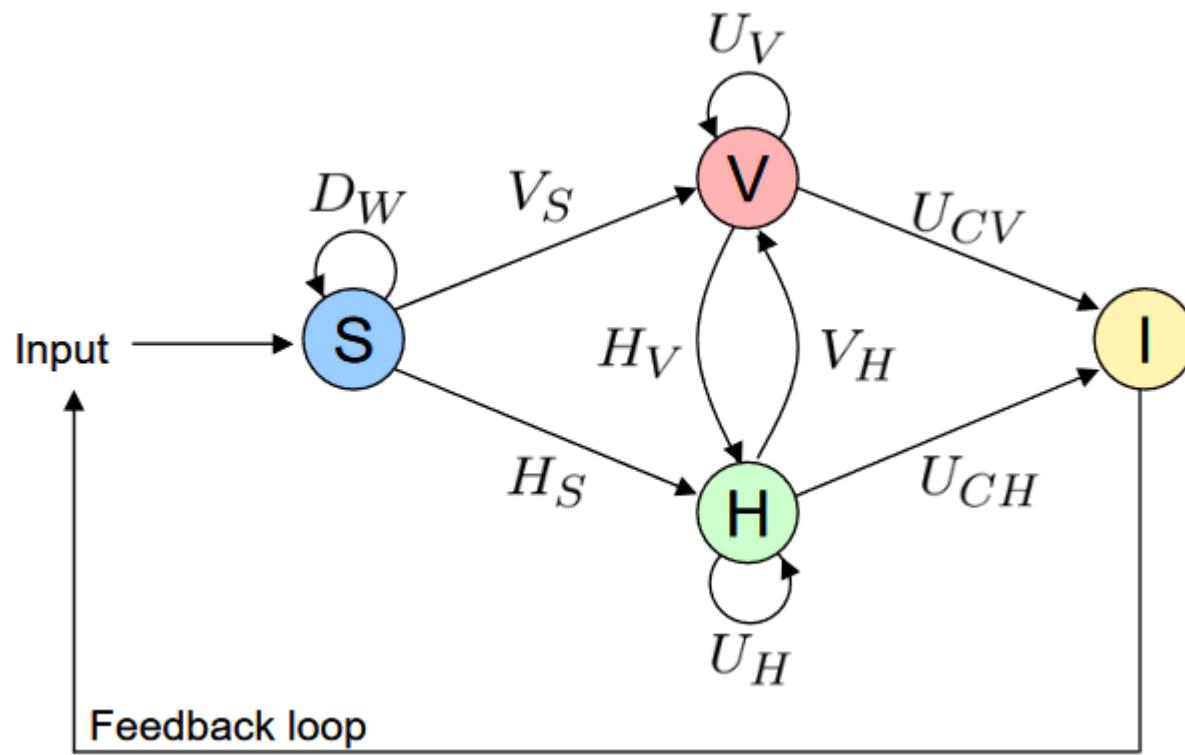


Maxima: popular taxi locations

- Density higher than local neighborhood.
- Can denote tourist locations, train stations.



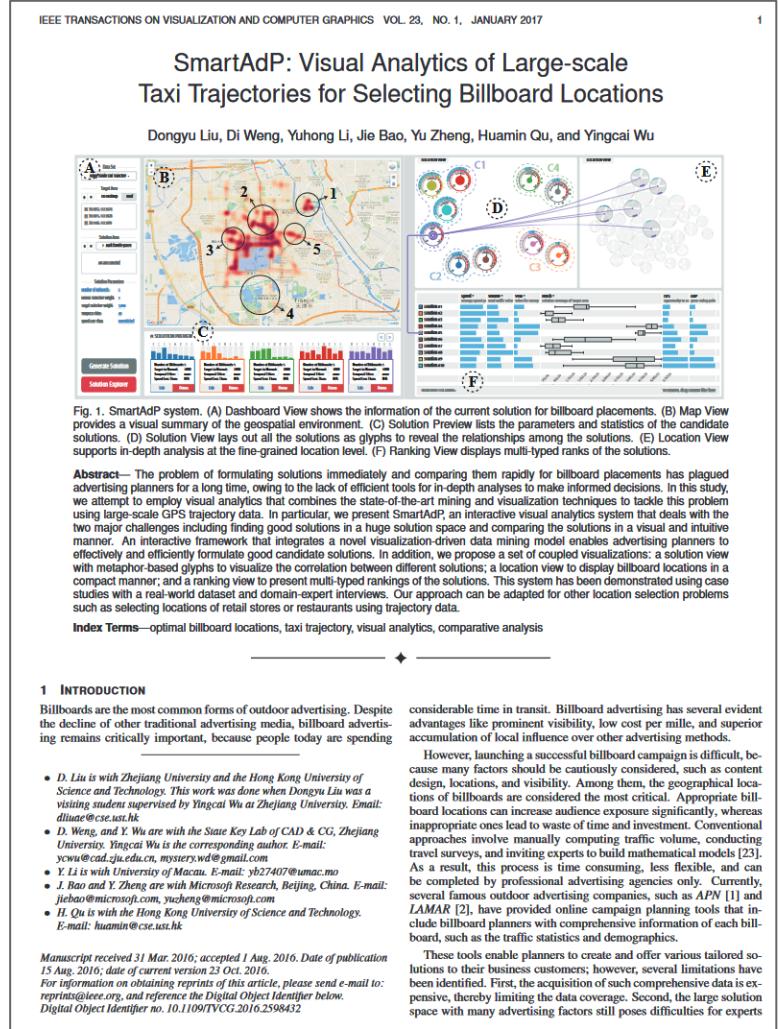
TaxiVis



- Visual analytics tool that includes:
 - h_s : hypotheses using computational topology
 - h_v : hypotheses from visualization
 - v_s : visualizing data
 - v_h : visualizing hypotheses
 - u : interaction

SmartAdP: taxi trajectories

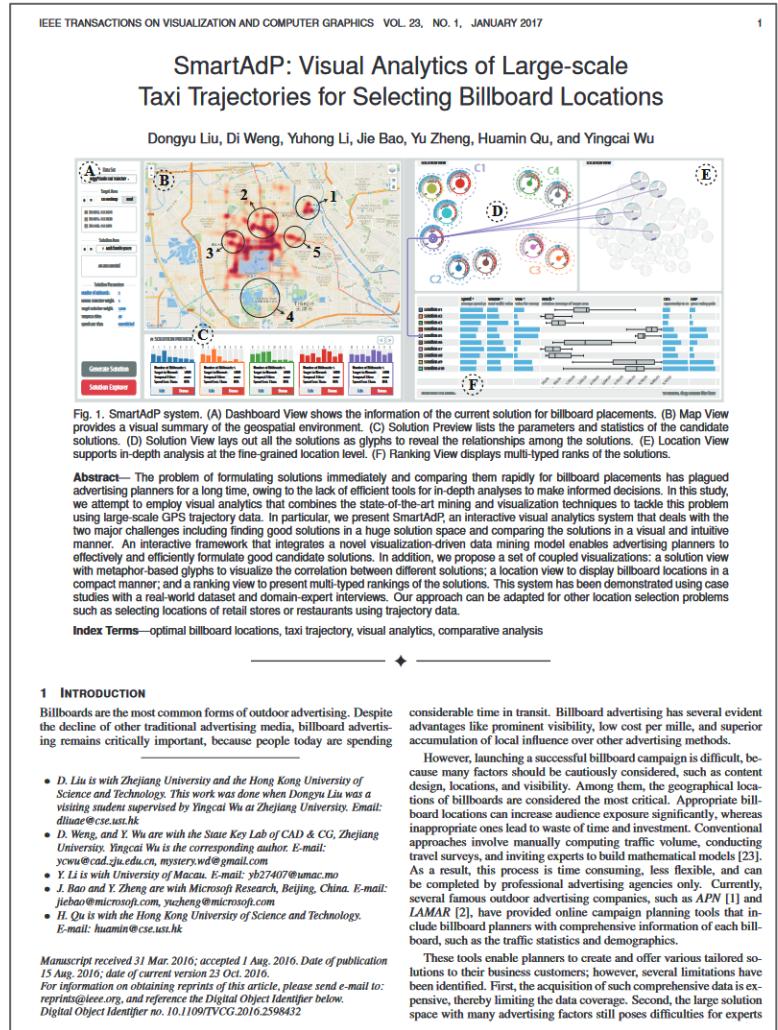
“ ... an interactive visual analytics system that deals with the two major challenges including finding good solutions in a huge solution space and comparing the solutions in a visual and intuitive manner. An interactive framework that integrates a novel visualization-driven data mining model enables advertising planners to effectively and efficiently formulate good candidate solutions. (...) This system has been demonstrated using case studies with a real-world dataset and domain-expert interviews.”



[Liu et al., 2017]

SmartAdP: taxi trajectories

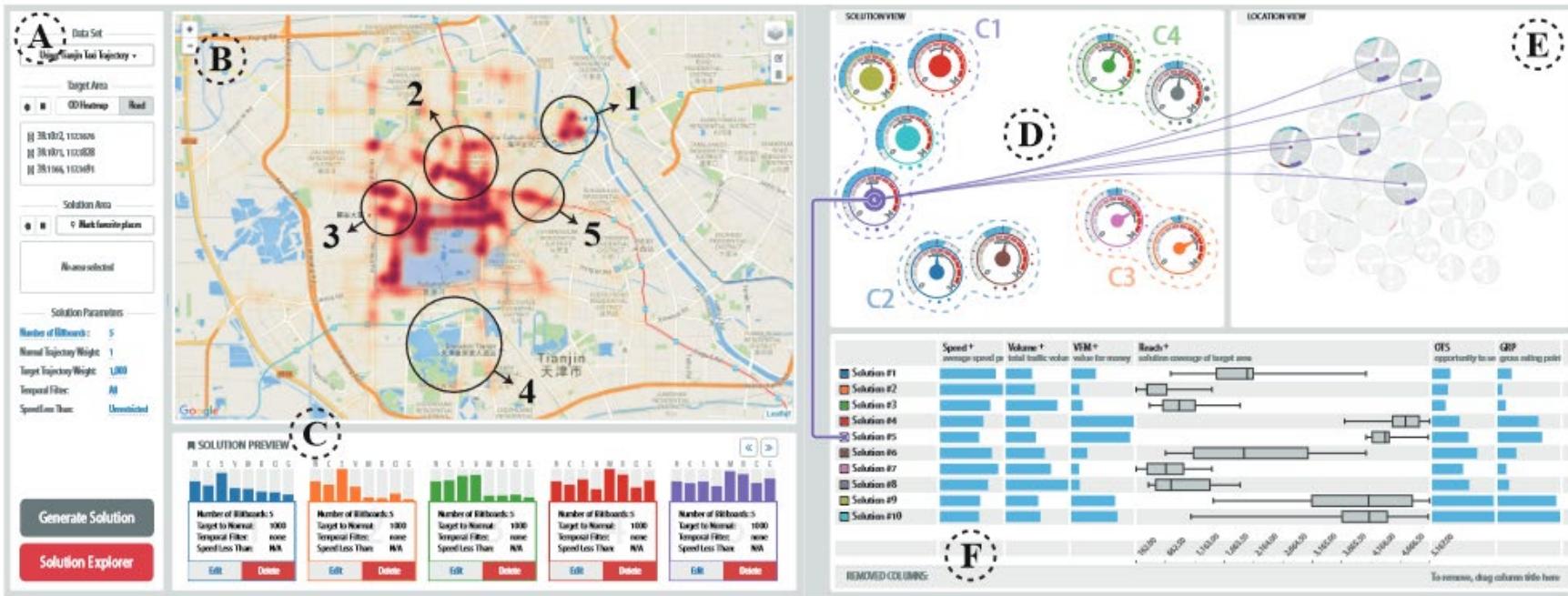
- Characterization of the problem of billboard location selection using taxi trajectory data.
- Interactive framework to generate billboard solutions with novel visualization-driven data mining model and application-specific data index mechanism.



[Liu et al., 2017]

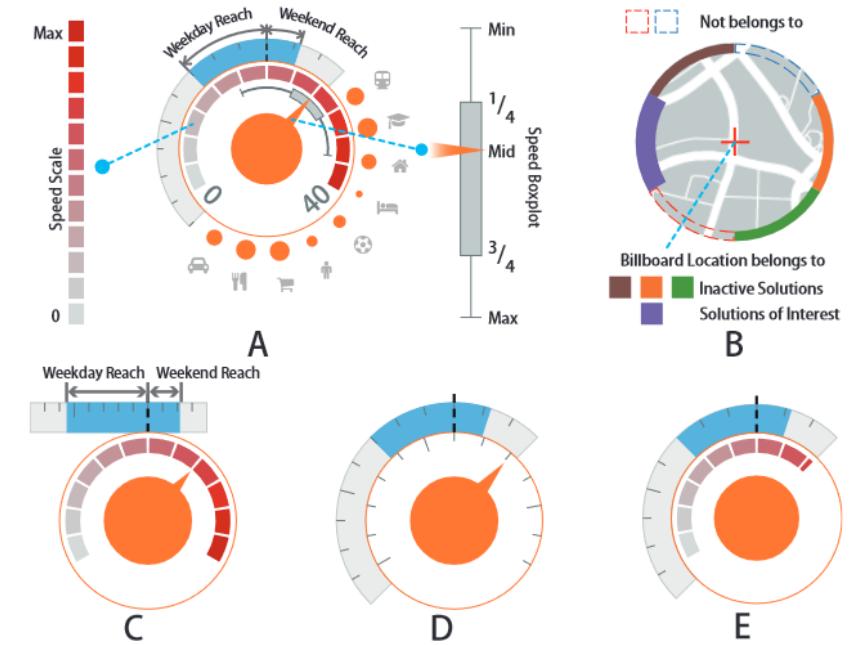
Billboard location

- Billboard location is a decisive factor for a billboard campaign. How to select the best locations (i.e., best intersections)?



Billboard location

- Data:
 - Traffic volume, traffic speed, traffic OD, environment, cost of billboard.
- Indicators:
 - *Coverage / reach*: percentage of covered trajectories among all trajectories.
 - *Opportunities to see*: average number of billboard contacts.
 - *Gross rating points*: average number of contacts that 100 target trajectories produce.
 - *Value for money*: value of the cost.



SmartAdP data structure

- K-location query:
 - Given a set of candidate vertices and target OD, select k locations.
 - Step 1: select trajectories with origin and destination within target OD.
 - All trajectories going from A to B.
 - Step 2: for each trajectory, update vertex coverage if it goes through vertex.
 - Does current trajectory go through candidate vertex? Update its coverage.

SmartAdP data structure

- How to support these operations?
 - We need to find trajectories within target OD (i.e., $vertex \rightarrow trajectory$).
 - We need to find where a trajectory goes through (i.e., $trajectory \rightarrow vertex$).
- MongoDB indices:
 - Trajectory-edge index: trajectory spatial points mapped to road network.
 $Trajectory \rightarrow road\ segment$
 - Trajectory-vertex index: covered vertices of trajectories.
 $Trajectory \rightarrow vertices$
 - Vertex-trajectory index: covered trajectories of vertices.
 $Vertices \rightarrow trajectories$

Urban Mosaic: street-level images

“ ... a tool for exploring the urban fabric through a spatially and temporally dense data set of 7.7 million street-level images from New York City, captured over the period of a year. Working in collaboration with professional practitioners, we use Urban Mosaic to investigate questions of accessibility and mobility, and preservation and retrofitting.”

Urban Mosaic: Visual Exploration of Streetscapes Using Large-Scale Image Data

Fabio Miranda¹, Maryam Hosseini², Marcos Lage³, Harish Doraiswamy¹, Graham Dove¹, Cláudio T. Silva¹

¹New York University; ²Rutgers University; ³Universidade Federal Fluminense
¹{fmiranda,harishd,graham dove,csilva}@nyu.edu; ²mary.hosseini@rutgers.edu; ³mlage@ic.uff.br

ABSTRACT

Urban planning is increasingly data driven, yet the challenge of designing with data at a city scale and remaining sensitive to the impact at a human scale is as important today as it was for Jane Jacobs. We address this challenge with Urban Mosaic, a tool for exploring the urban fabric through a spatially and temporally dense data set of 7.7 million street-level images from New York City, captured over the period of a year. Working in collaboration with professional practitioners, we use Urban Mosaic to investigate questions of accessibility and mobility, and preservation and retrofitting. In doing so, we demonstrate how tools such as this might provide a bridge between the city and the street, by supporting activities such as visual comparison of geographically distant neighborhoods, and temporal analysis of unfolding urban development.

Author Keywords

Urban planning; Interactive visualization; Data analysis; Urban data

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); Visual analytics; Visualization toolkits;

“A sense of place is built up, in the end, from many little things too, some so small people take them for granted, and yet the lack of them takes the flavor out of the city.”

(Jane Jacobs, *Downtown is for People*)

INTRODUCTION

For those of us living in or visiting the world’s major cities, their dynamism and complexity are immediately apparent. Yet urban planners and designers must work in a context where any single intervention, perhaps aimed at altering just one aspect, can have a wide ranging impact on a variety of interrelated components [12], affecting things at both a macro *city scale* and a micro *human scale*. Examples of changes at a city scale

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for this work (without other components) is held by the author(s). Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Requests for permission from permissions@acm.org.
CHI ’20, April 25–30, 2020, Honolulu, HI, USA
© 2020, copyright is held by the author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-6708-0/20/04...\$15.00.
<http://dx.doi.org/10.1145/3313831.3376399>

include, suburbanization, economic deconcentration, modification to transport infrastructure, rezoning and/or gentrification of neighborhoods, and major renewal projects. Examples of changes at a human scale, on the other hand, are reflected in the city’s urban fabric, and include aspects that make a city livable, encourage walking, and contribute to the perception of safety; in other words affect the day to day lives of its inhabitants. This might be manifested in lighting, shadow, sky exposure, open-front shops, the details of building facades, etc. [32].

Responding to these challenges at the city scale, planners and designers have been aided by a rapid growth in data from urban environments, and so are able to turn to computational methods and large-scale data analysis, which increase understanding by quantifying different aspects of the city (e.g., [36, 13, 14, 29, 70, 67, 48]). However, while sensitivity to the impact of change at a human scale remains as important today as it was for Jane Jacobs and others in the 1950s and 1960s [47], analyses of suitable data, which emphasize qualitative, visual details, are often difficult and time-consuming to perform. Although there is an increasing availability of street-view images, which support a degree of virtual assessment and auditing of the built environment, their distribution is often temporally sparse and so analysis is limited.

This paper introduces Urban Mosaic, a tool for visually exploring the urban fabric. It responds to the challenges practitioners face by employing a newly available spatially and temporally dense data set of street-level images from New York City (NYC). Urban Mosaic is a visual exploration system designed to help practitioners in urban planning and design gain insight into the human scale impact of changes in the urban fabric. It utilizes state-of-the-art computer vision techniques for image similarity search and clustering, together with efficient spatio-temporal selection and aggregation over the image metadata to visually explore and map this image data set. It further allows the analysis to be augmented using spatio-temporal urban data from a variety of other sources (e.g., census, transport, crime, weather, housing market, zoning, noise complaints). The image data set used in this work contains 7.7 million images captured in the Manhattan and Brooklyn boroughs between April 2016 and April 2017 using car-mounted cameras. Urban Mosaic has been developed as a collaboration between researchers in urban planning, visual analytics, and HCI. We include a detailed

[Miranda et al., 2020]



UIC COMPUTER SCIENCE

Urban Mosaic: street-level images

- Demonstrates the potential for images from large-scale street-view data sets to help bridge data-driven urban planning and design at city scale and at the human level.
- Exploration of the urban fabric using large spatiotemporal collection of images.
- Visual comparison of geographically distant areas to analyze unfolding urban developments.

Urban Mosaic: Visual Exploration of Streetscapes Using Large-Scale Image Data

Fabio Miranda¹, Maryam Hosseini², Marcos Lage³, Harish Doraiswamy¹, Graham Dove¹, Cláudio T. Silva¹

¹New York University; ²Rutgers University; ³Universidade Federal Fluminense

¹{fmiranda,harishd,graham dove,csilva}@nyu.edu; ²mary.hosseini@rutgers.edu; ³mlage@ic.uff.br

ABSTRACT

Urban planning is increasingly data driven, yet the challenge of designing with data at a city scale and remaining sensitive to the impact at a human scale is as important today as it was for Jane Jacobs. We address this challenge with Urban Mosaic, a tool for exploring the urban fabric through a spatially and temporally dense data set of 7.7 million street-level images from New York City, captured over the period of a year. Working in collaboration with professional practitioners, we use Urban Mosaic to investigate questions of accessibility and mobility, and preservation and retrofitting. In doing so, we demonstrate how tools such as this might provide a bridge between the city and the street, by supporting activities such as visual comparison of geographically distant neighborhoods, and temporal analysis of unfolding urban development.

Author Keywords

Urban planning; Interactive visualization; Data analysis; Urban data

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); Visual analytics; Visualization toolkits;

"A sense of place is built up, in the end, from many little things too, some so small people take them for granted, and yet the lack of them takes the flavor out of the city."

(Jane Jacobs, *Downtown is for People*)

INTRODUCTION

For those of us living in or visiting the world's major cities, their dynamism and complexity are immediately apparent. Yet urban planners and designers must work in a context where any single intervention, perhaps aimed at altering just one aspect, can have a wide ranging impact on a variety of interrelated components [12], affecting things at both a macro *city scale* and a micro *human scale*. Examples of changes at a city scale

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for this work (without others) is held by the author(s). It is illegal to abstract with credit not permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Requests for permission from permissions@acm.org.
CHI '20, April 25–30, 2020, Honolulu, HI, USA
© 2020, copyright is held by the author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-6708-0/20/04 ...\$15.00.
<http://dx.doi.org/10.1145/3313831.3376399>

include, suburbanization, economic deconcentration, modification to transport infrastructure, rezoning and/or gentrification of neighborhoods, and major renewal projects. Examples of changes at a human scale, on the other hand, are reflected in the city's urban fabric, and include aspects that make a city livable, encourage walking, and contribute to the perception of safety; in other words affect the day to day lives of its inhabitants. This might be manifested in lighting, shadow, sky exposure, open-front shops, the details of building facades, etc. [32].

Responding to these challenges at the city scale, planners and designers have been aided by a rapid growth in data from urban environments, and so are able to turn to computational methods and large-scale data analysis, which increase understanding by quantifying different aspects of the city (e.g., [36, 13, 14, 29, 70, 67, 48]). However, while sensitivity to the impact of change at a human scale remains as important today as it was for Jane Jacobs and others in the 1950s and 1960s [47], analyses of suitable data, which emphasize qualitative, visual details, are often difficult and time-consuming to perform. Although there is an increasing availability of street-view images, which support a degree of virtual assessment and auditing of the built environment, their distribution is often temporally sparse and so analysis is limited.

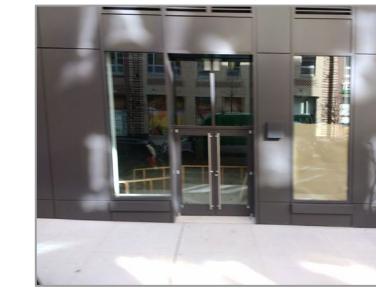
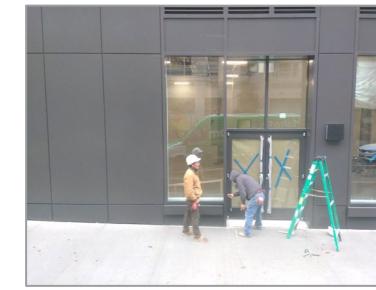
This paper introduces Urban Mosaic, a tool for visually exploring the urban fabric. It responds to the challenges practitioners face by employing a newly available spatially and temporally dense data set of street-level images from New York City (NYC). Urban Mosaic is a visual exploration system designed to help practitioners in urban planning and design gain insight into the human scale impact of changes in the urban fabric. It utilizes state-of-the-art computer vision techniques for image similarity search and clustering, together with efficient spatio-temporal selection and aggregation over the image metadata to visually explore and map this image data set. It further allows the analysis to be augmented using spatio-temporal urban data from a variety of other sources (e.g., census, transport, crime, weather, housing market, zoning, noise complaints). The image data set used in this work contains 7.7 million images captured in the Manhattan and Brooklyn boroughs between April 2016 and April 2017 using car-mounted cameras. Urban Mosaic has been developed as a collaboration between researchers in urban planning, visual analytics, and HCI. We include a detailed

[Miranda et al., 2020]



COMPUTER SCIENCE

Temporally-dense street-level images



Objectives

Support the **interactive** analysis of the city at the micro scale, over geographically distant regions.



Comparison of the urban fabric in different regions



Assessment of features in the built environment

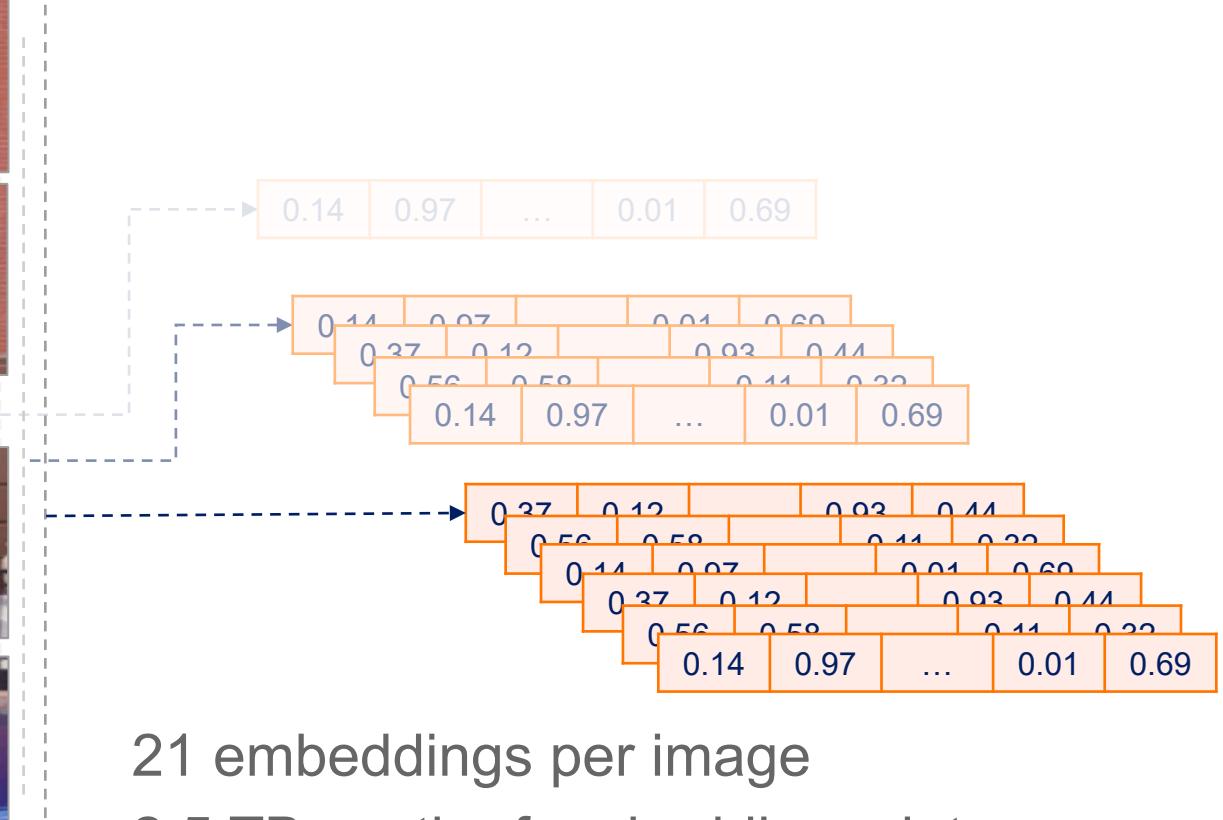


Assessment of walkability and accessibility

Image query composition

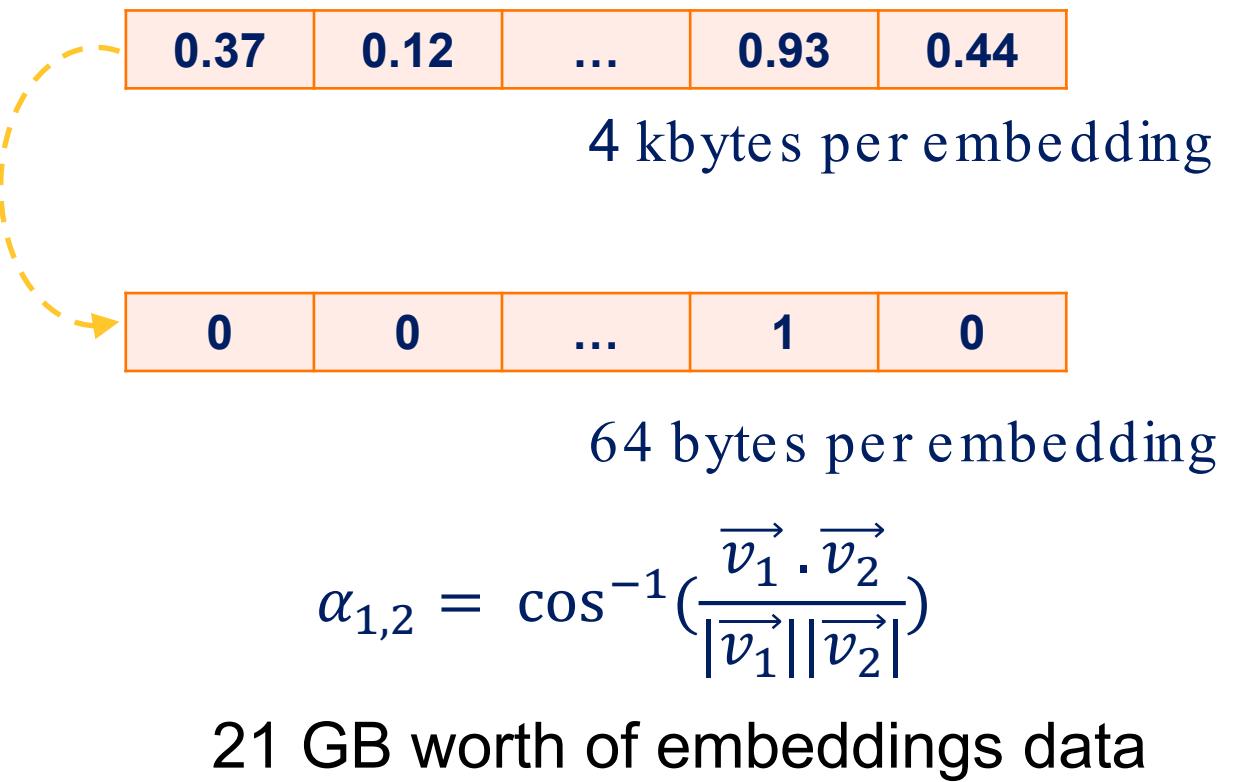
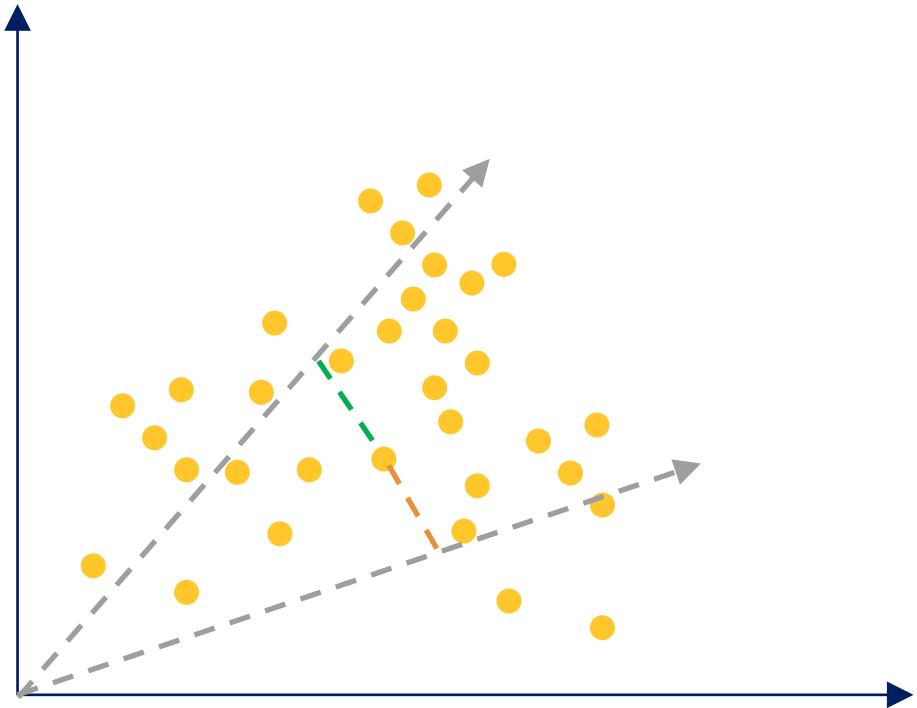


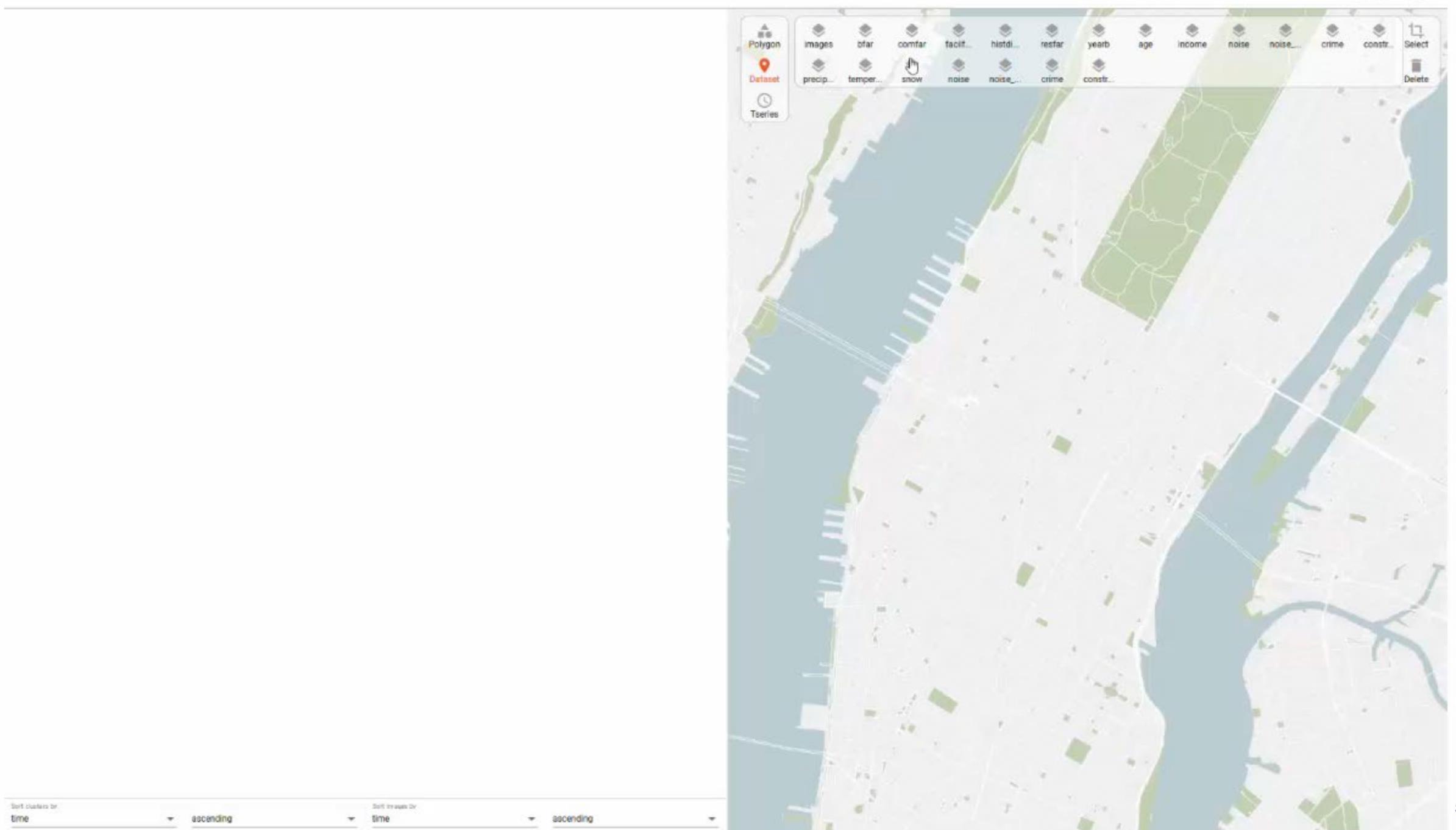
Image embeddings



21 embeddings per image
2.5 TB worth of embeddings data

Locality sensitive hashing





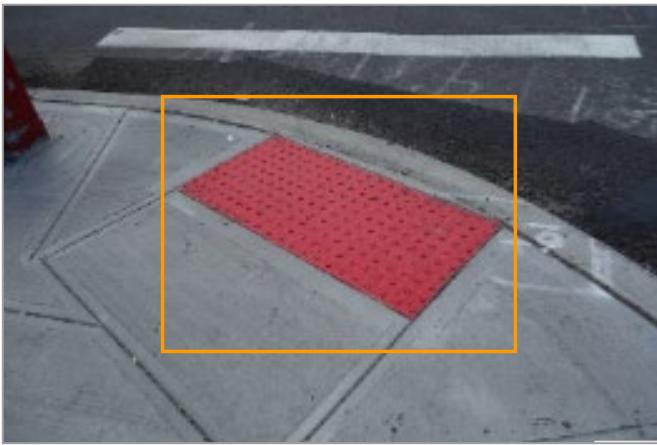
Accessibility for older adults



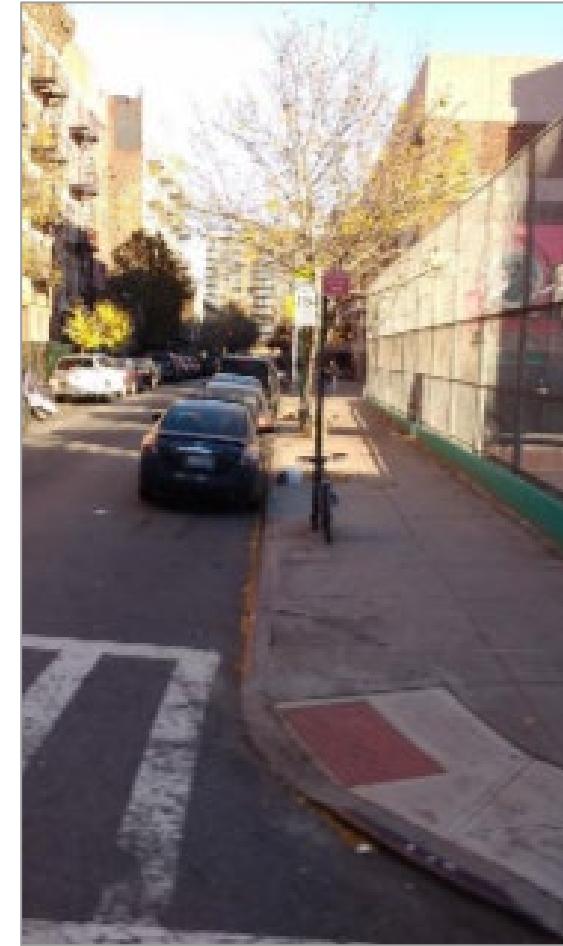
Focuses on interventions that enable older adults to “Age in Place” and prevent outdoor fall.

Impact of the neighborhood environment on falls.

Accessibility: installation of tactile pavings



June

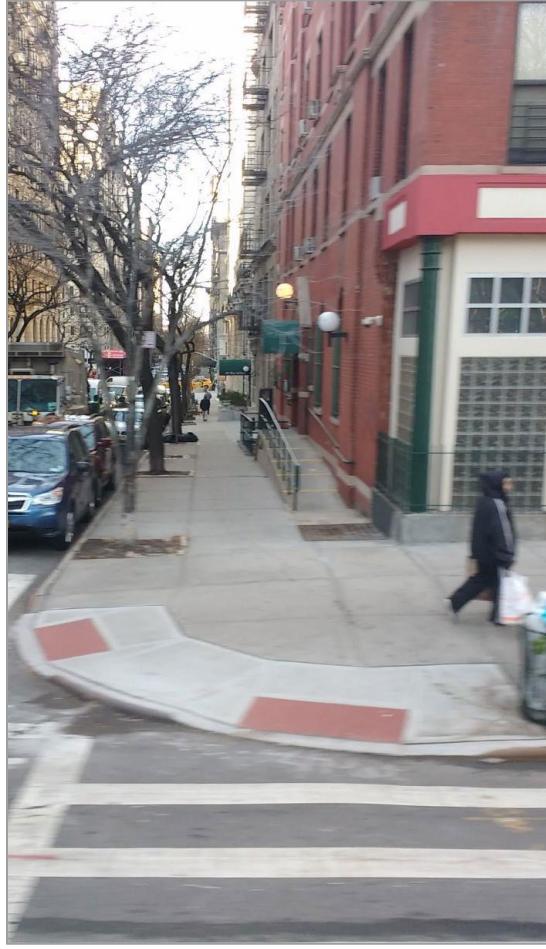
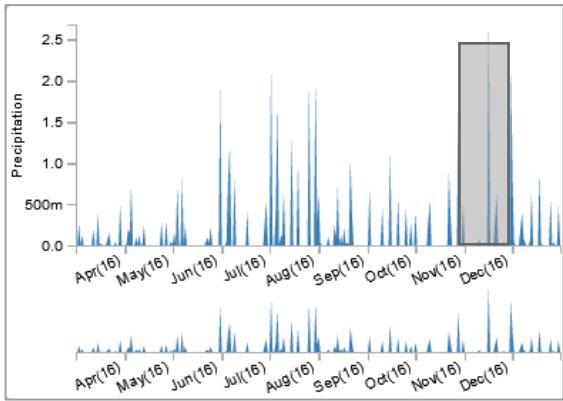
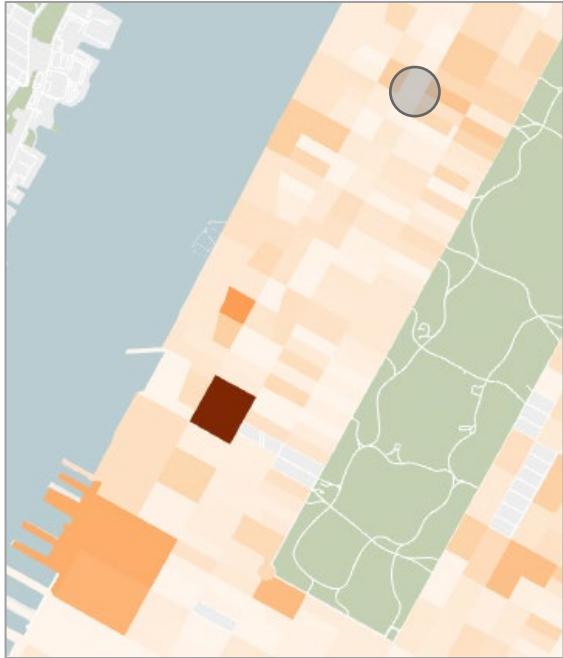


August



October

Accessibility: assessing hazards for older adults



Main takeaways

- General approaches might not scale.
- Big data visual analytics might require structures specifically designed for a given task.
- Understand the data, and how it can be transformed to best fit different computational resources and architectures.

Main takeaways

