

Locational Big Data and Analytics: Implications for the Sharing Economy AMCIS 2017 SIGGIS Workshop

Brian N. Hilton, Ph.D.

Associate Professor

Director, Advanced GIS Lab

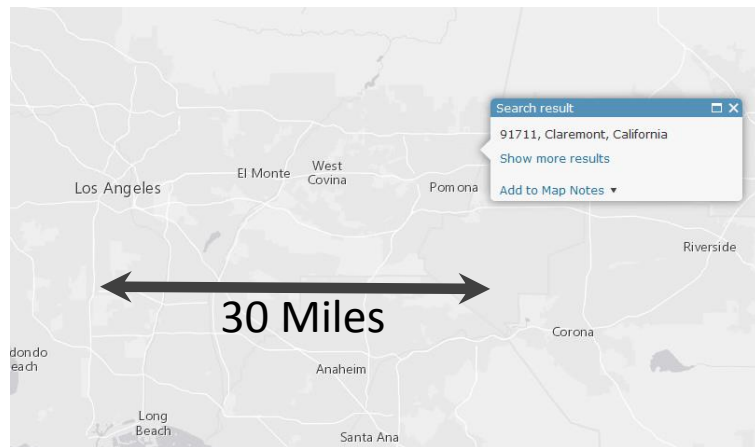
Center for Information Systems and Technology

Claremont Graduate University

Claremont, CA



 Claremont Graduate University



CENTER FOR INFORMATION SYSTEMS & TECHNOLOGY

Advanced GIS Lab Center for Information Systems and Technology Claremont Graduate University

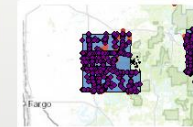
Featured Maps



LA's Urban Forest Since 1985



Los Angeles - Soils for Groundwater Recharge



MN Tribes, Crashes, and Road Segments



Monterey County: California Coastal Trail Association

The Advanced GIS Lab at the Center for Information Systems and Technology (CISAT) focuses on advanced spatial analysis and the research and development of advanced GIS solutions. Led by Clinical Associate Professor Dr. Brian Hilton, the Lab stresses a transdisciplinary approach to knowledge, research, and problem-solving using core ideas, methods, and concepts from several disciplines to critically examine a broad range of real-world problems. This transdisciplinary approach, a hallmark of Claremont Graduate University's (CGU) research philosophy, extends the scope of interdisciplinary or multidisciplinary scholarship by traversing the range of traditional disciplines for the advancement of knowledge and solutions to the world's most pressing issues.

Currently, the Lab and its associated students and faculty are examining the implementation of GIS technologies to improve community health, better understand road transportation safety, support humanitarian efforts, and examine ecosystem services.

In addition, the Lab is an Esri Development Center (EDC). The EDC program was created to confer special recognition and status to university departments worldwide that have exemplary programs focused on educating students in the design and development of GIS applications using Esri's geospatial technologies. As an inaugural EDC, CISAT is a unique resource for CGU that: provides students and faculty with the capabilities to teach and develop state-of-the-art applications in the Lab; provides Esri training focused on GIS and related technologies; and offers students special recognition through an annual achievement award.

Founded in 1925, Claremont Graduate University is an independent graduate-level university. CGU is located in the city of Claremont, California, 35 miles east of Los Angeles. The CGU community is characterized by its unusual diversity, collegiality, and environmental beauty.

[Esri.com](#) | [Help](#) | [Terms of Use](#) | [Privacy](#) | [Contact Esri](#) | [Report Abuse](#) | [Contact Us](#)



Agenda

- Review of Three Research Projects / Use Cases
 - Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement
 - Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement
 - Stop-and-Frisk Policy from a Quantitative and Spatial Perspective
- Hands-On Demonstration (Sharing Economy Examples)
 - Insights for ArcGIS

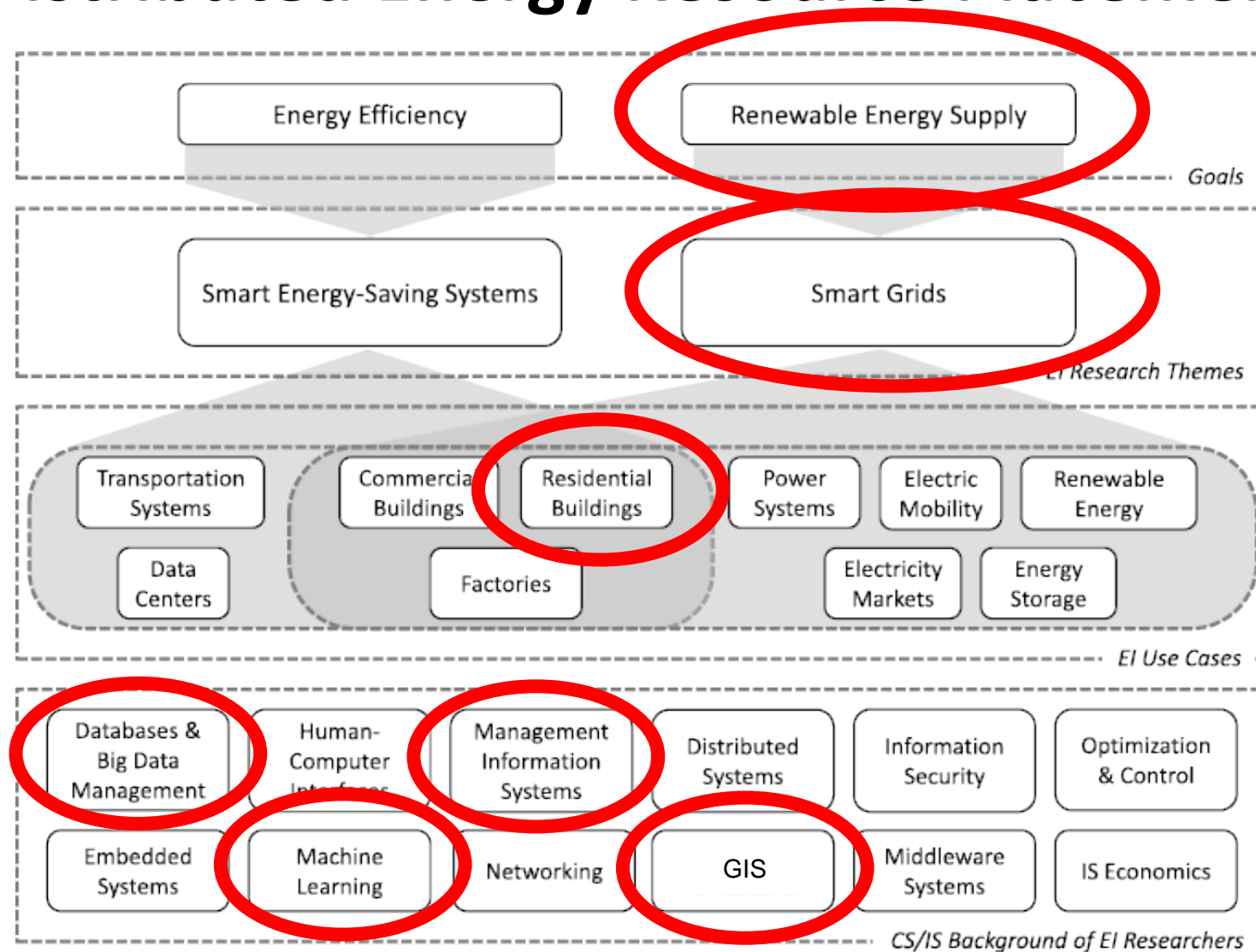
Agenda

- Locational Big Data and Analytics has created a need for the efficient manipulation and scalable analysis of spatial big data on disparate, and distributed, datasets. As a result, this has opened a number of research areas such as:
 - Developing capabilities for accessing, formatting, and combining spatial big data in ways that enable it to be easily consumed;
 - Developing methodologies to derive insight into spatial big data for inferential understanding and decision making;
 - Developing teaching resources to better understand the use of data manipulation techniques, spatial statistics, and spatial data-mining tasks related to spatial big data; and
 - Developing novel spatial and spatiotemporal methods that can take advantage of newly emerging data-intensive computational resources.

Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Domain – Energy Informatics
- Research Question – “Where are the optimal locations for the placement of Distributed Energy Resources, specifically, lithium-ion (Li-ion) batteries on the electricity grid?”

Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement



Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

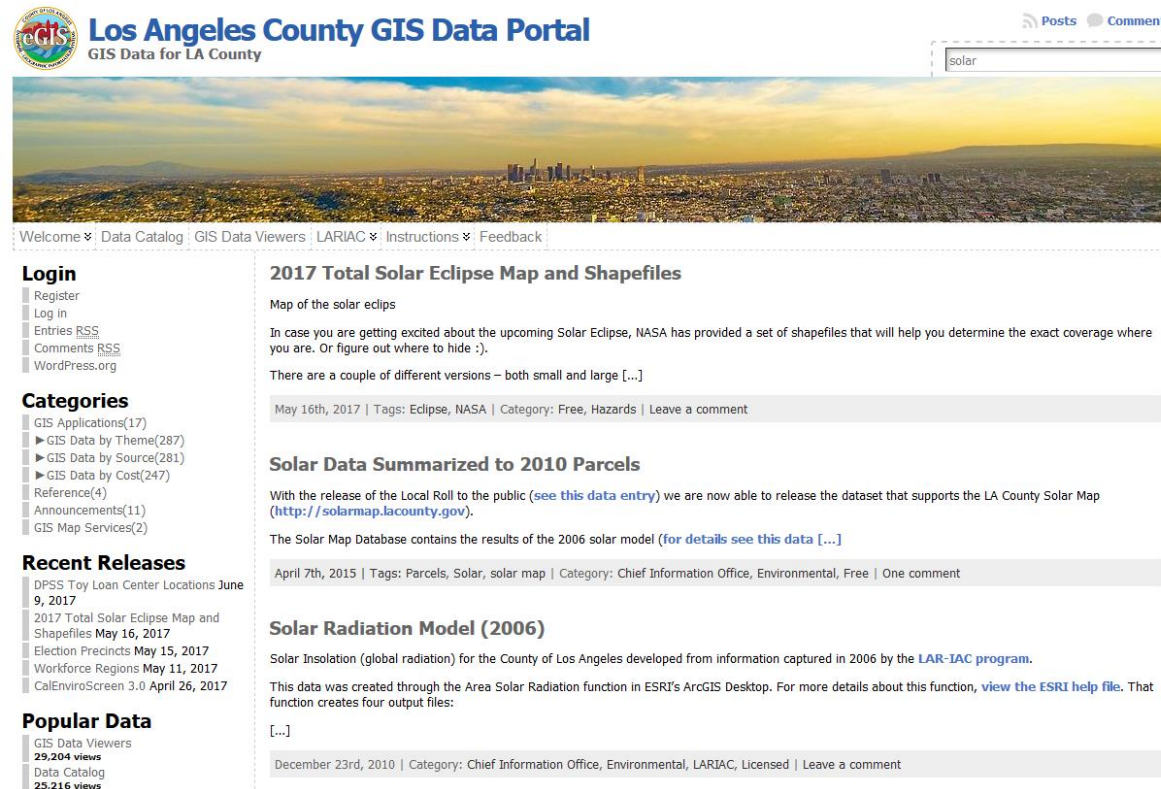
- Deploying **Distributed Energy Resources** in a widespread, efficient, and cost-effective manner requires complex integration with the existing electricity grid.
- The global scale-up of lithium-ion (**Li-ion**) **batteries** is enabling cost-effective energy storage systems for electric utility use.
- Policy incentives have increased **solar panel adoption** (grid-connected photovoltaic energy (PV) systems) – **California ranks first** among all states in number of solar PV systems installed.
- Research can identify and resolve the challenges of PV system integration, facilitating the transition to a smarter grid.

Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Research steps:
 1. Understanding **Solar Panel Adoption** across three main customer types: **Residential**, Commercial, and Industrial.
 2. Development of GIS-based planning algorithm(s) for the optimal placement of new DERs (Li-ion batteries) given the spatial constraints of the existing electricity grid.

Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Data Provider
 - “The LA County GIS Data Portal is the place to search for GIS data created, maintained, licensed, and stored by the County of Los Angeles.”



Los Angeles County GIS Data Portal
GIS Data for LA County

Posts Comments

solar

Welcome Data Catalog GIS Data Viewers LARIAC Instructions Feedback

Login
Register
Log in
Entries RSS
Comments RSS
WordPress.org

Categories
GIS Applications(17)
GIS Data by Theme(287)
GIS Data by Source(281)
GIS Data by Cost(247)
Reference(4)
Announcements(11)
GIS Map Services(2)

Recent Releases
DPSS Toy Loan Center Locations June 9, 2017
2017 Total Solar Eclipse Map and Shapefiles May 16, 2017
Election Precincts May 15, 2017
Workforce Regions May 11, 2017
CalEnviroScreen 3.0 April 26, 2017

Popular Data
GIS Data Viewers
29,204 views
Data Catalog
25,216 views

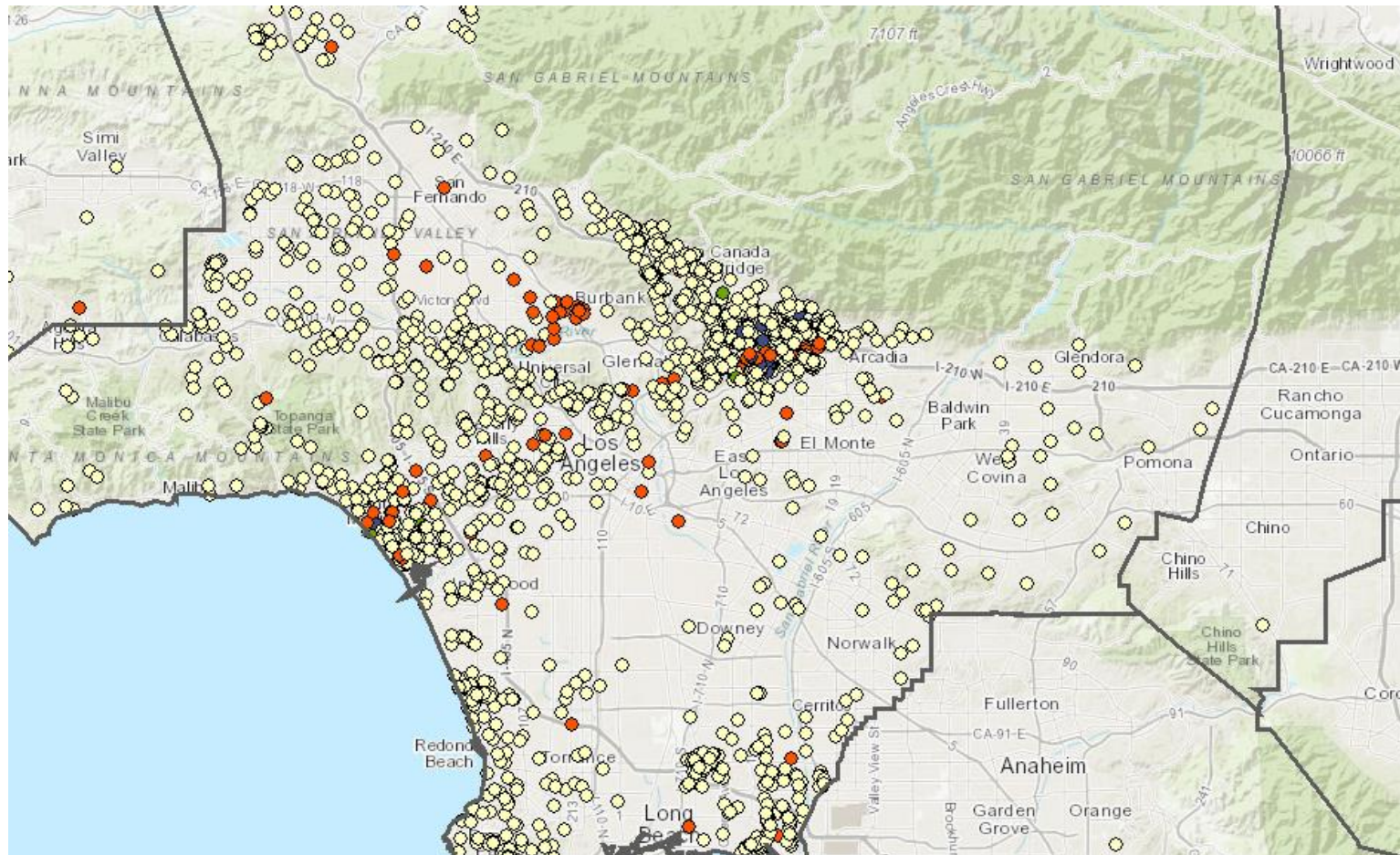
2017 Total Solar Eclipse Map and Shapefiles
Map of the solar eclips
In case you are getting excited about the upcoming Solar Eclipse, NASA has provided a set of shapefiles that will help you determine the exact coverage where you are. Or figure out where to hide :).
There are a couple of different versions – both small and large [...]
May 16th, 2017 | Tags: Eclipse, NASA | Category: Free, Hazards | Leave a comment

Solar Data Summarized to 2010 Parcels
With the release of the Local Roll to the public (see this data entry) we are now able to release the dataset that supports the LA County Solar Map (<http://solamap.lacounty.gov>).
The Solar Map Database contains the results of the 2006 solar model (for details see this data [...])
April 7th, 2015 | Tags: Parcels, Solar, solar map | Category: Chief Information Office, Environmental, Free | One comment

Solar Radiation Model (2006)
Solar Insolation (global radiation) for the County of Los Angeles developed from information captured in 2006 by the LAR-IAC program.
This data was created through the Area Solar Radiation function in ESRI's ArcGIS Desktop. For more details about this function, view the ESRI help file. That function creates four output files:
[...]
December 23rd, 2010 | Category: Chief Information Office, Environmental, LARIAC, Licensed | Leave a comment

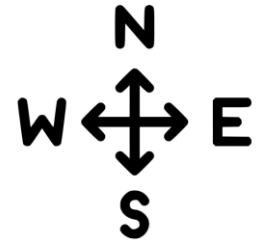
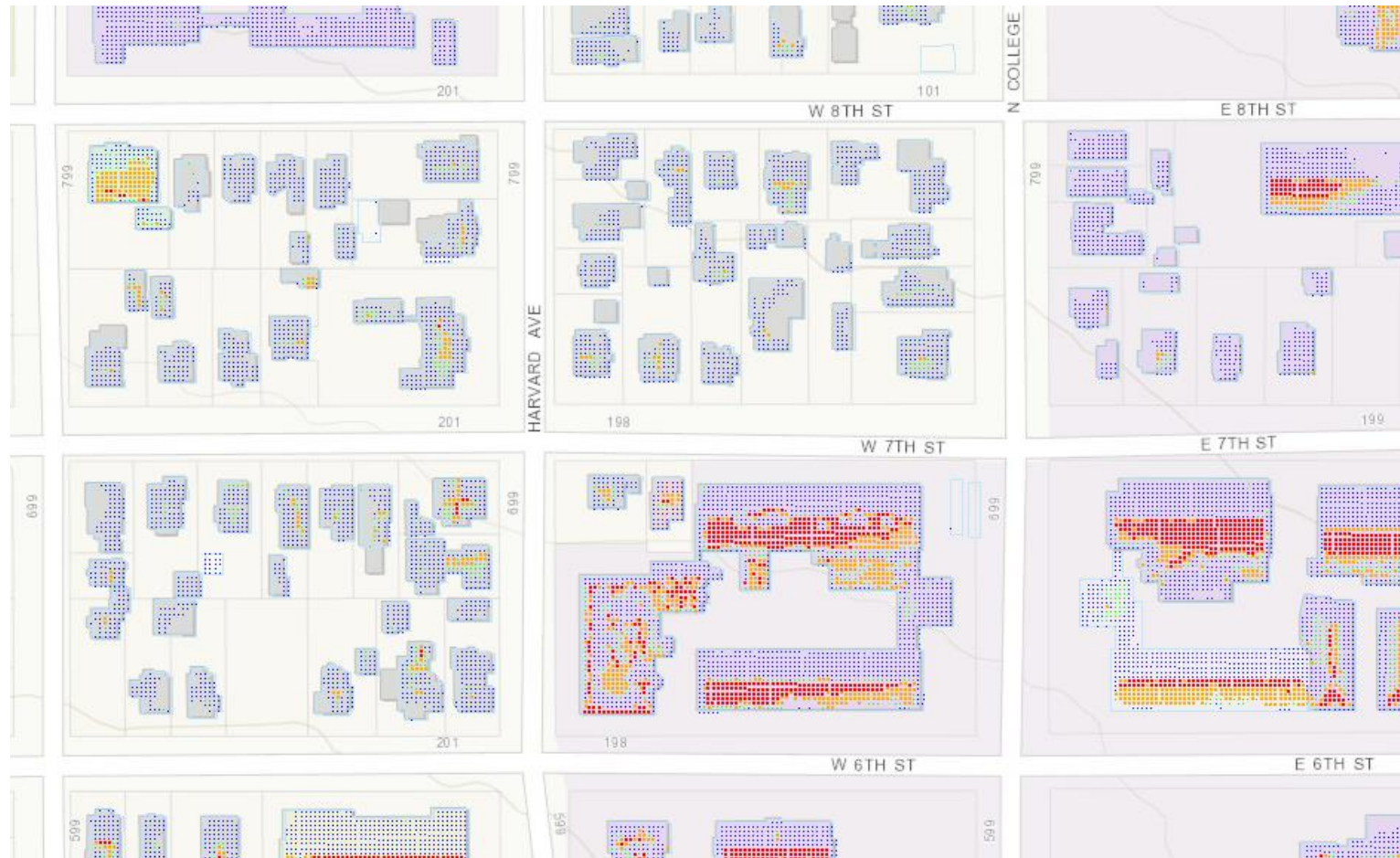
Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Solar Installation Data Description



Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Solar PV Potential Data Description



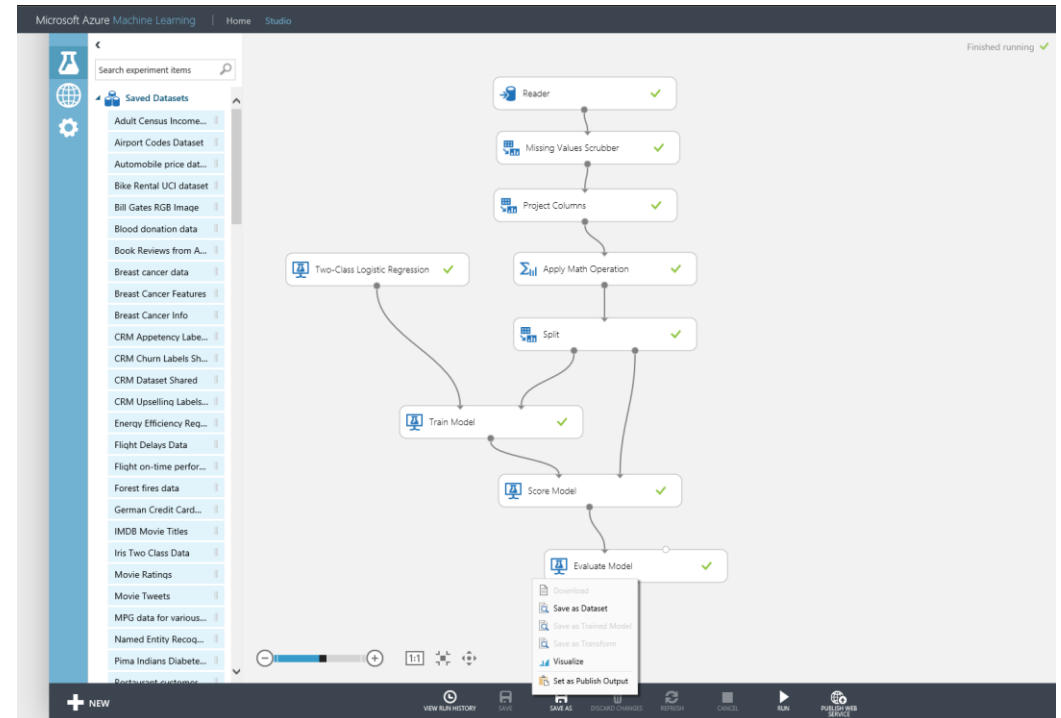
Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Research Step (1)
- Creating a Predictive Model for Residential Solar Panel Adoption
 - *Los Angeles County: Residential Parcels = 1,868,519 out of 2,392,100 (78%)*
 - Dependent variable:
Likelihood of a household adopting solar energy panels
 - Independent variables:
 - (1) parcel information, such as: parcel age, parcel value, etc.
 - (2) customer demographics, such as: household income, household size, etc.
 - (3) expenditure data, such as: electricity usage, mortgage value, etc.

Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Research Step (1)
- The results of the Two-Class Logistic Regression (Azure Machine Learning) indicated that six factors emerged as significant predictors of solar adoption:

- parcel age,
- average household size,
- total area suitable for solar roof top,
- total building area square feet,
- average household income, and
- average home value



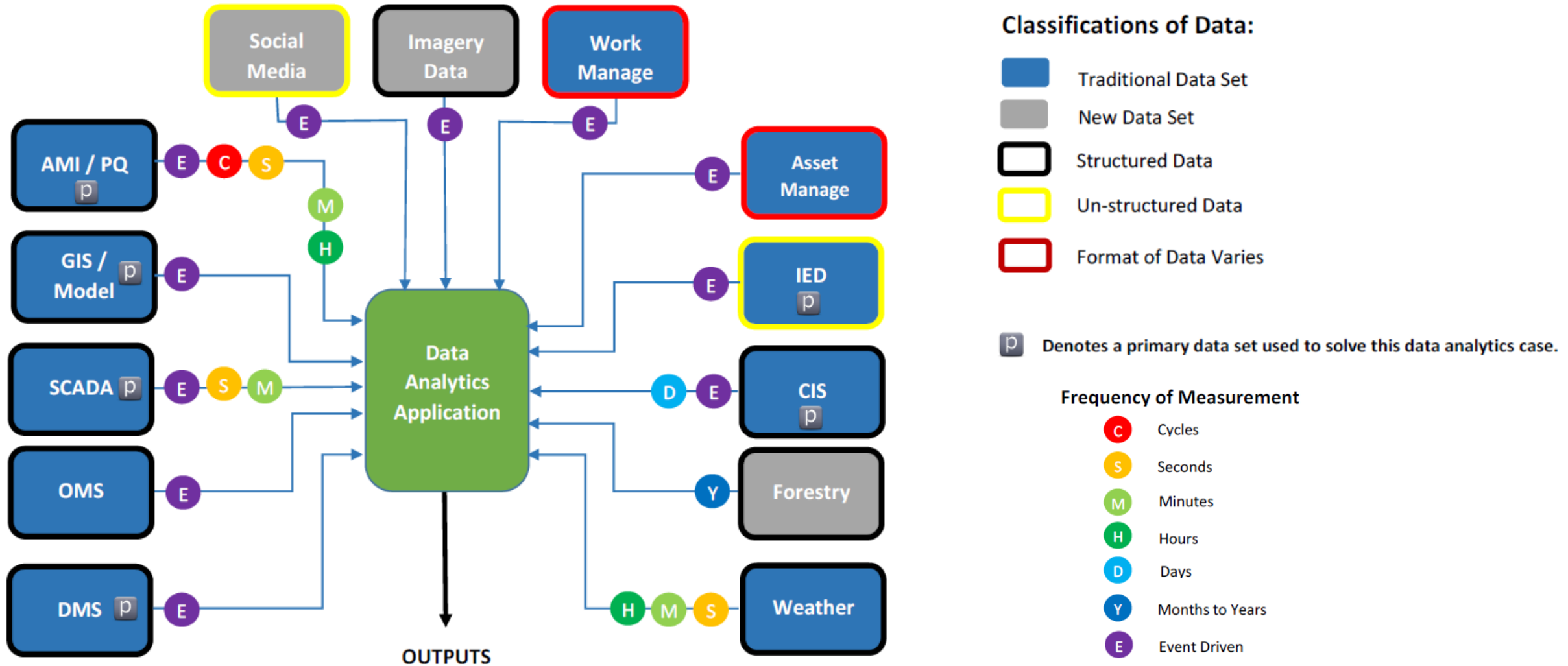
Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

- Future Work / Analysis
 - Research Step (2)
 - Development of GIS-based planning algorithm(s) for the optimal placement of new DERs
 - Data Provider
 - “The Electric Power Research Institute, or EPRI, conducts research on issues related to the electric power industry.”



- Devise a methodology for organizing our disparate datasets...

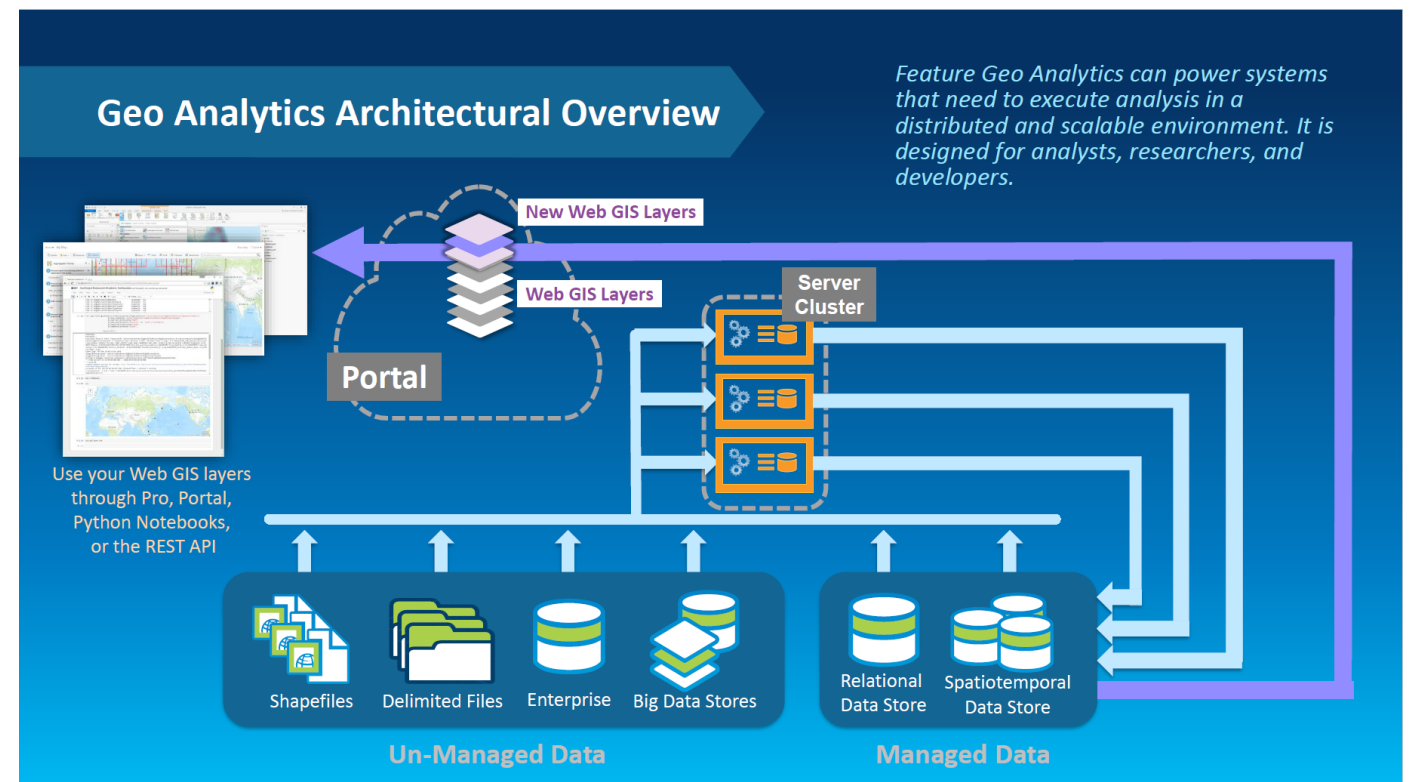
Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement



Optimization Algorithm for Spatially Constrained Distributed Energy Resource Placement

Developing capabilities for accessing, formatting, and combining spatial big data in ways that enable it to be easily consumed.

- ArcGIS GeoAnalytics Server



Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

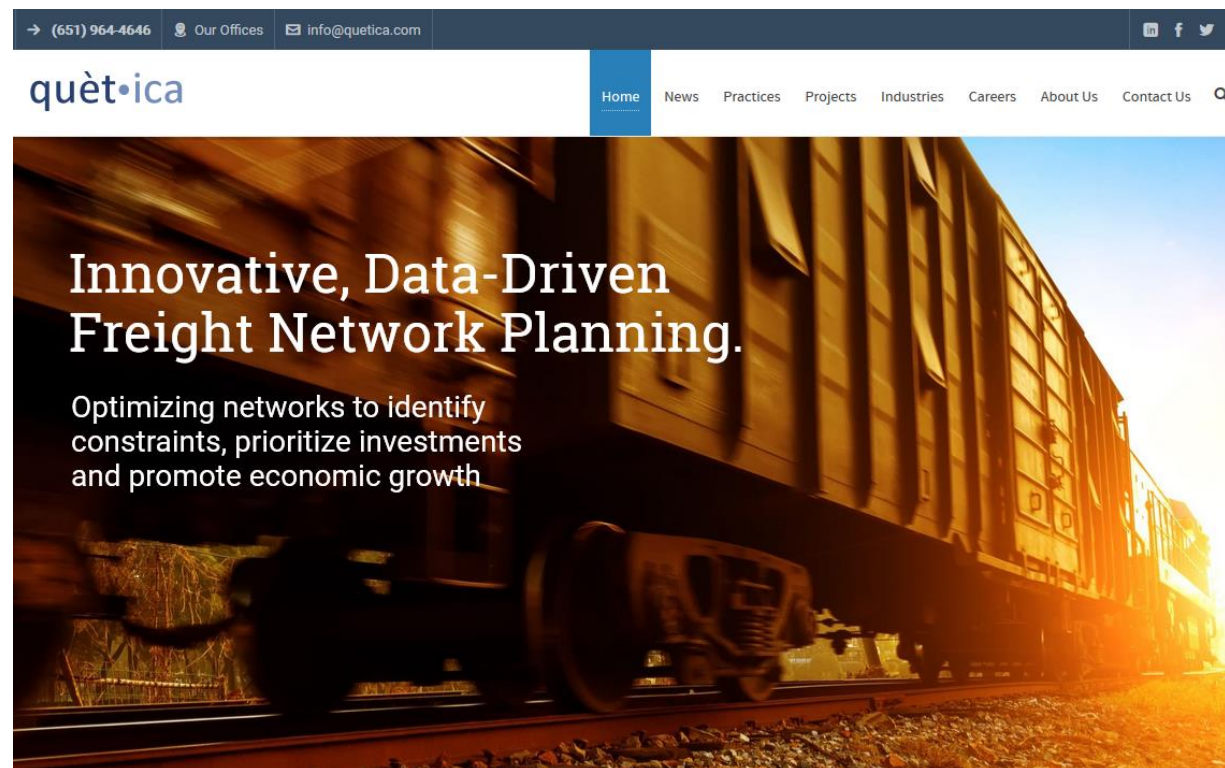
- Domain – Logistics / Freight Network Planning
- Research Question – “How can we better understand commodity flows for economic development, for freight policy analysis, and transportation infrastructure impacts?”

Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

- In Minnesota, technological and economic shifts in the grain supply chain have altered the way grain producers and sellers navigate their local freight network.
- In particular, many producers have been increasing their personal trucking capacity and taking longer trips to intermodal and domestic market options.
- This logistical reshaping of local grain supply chains pressure transportation officials to reconsider the consequences for road infrastructure and congested freight corridors.

Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

- Data Provider
 - Quetica, a Minnesota-based, supply chain management company that uses commodity flow analysis to optimize freight network planning.

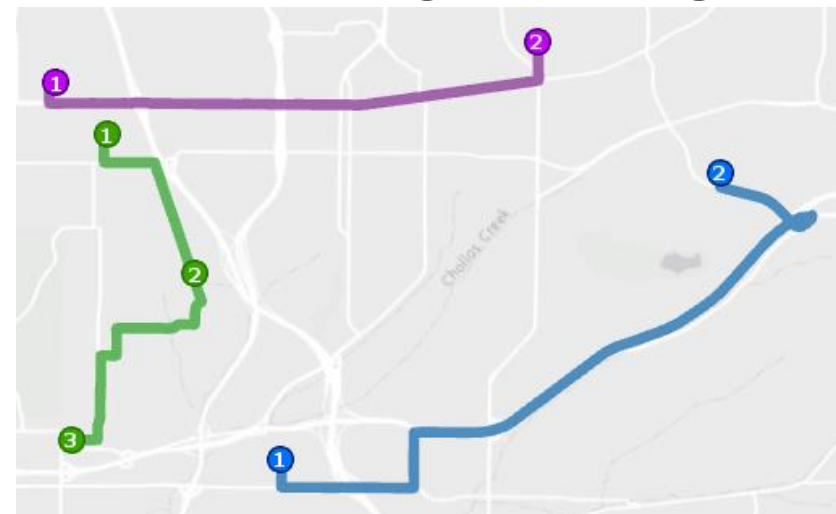


Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

- Freight Data Description
 - The Quetica sample dataset included **cereal grain shipments, via truck**, including **shipment weight**, for Midwest **U.S. counties** in 2014:
 - 257,006 - Midwest U.S. shipments - total tons 764,848,291
 - 15,920 - MN-related (internal/external) shipments - total tons 79,638,868
 - 4,489 - MN-only shipments (internal/internal) - total tons 66,789,589

Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

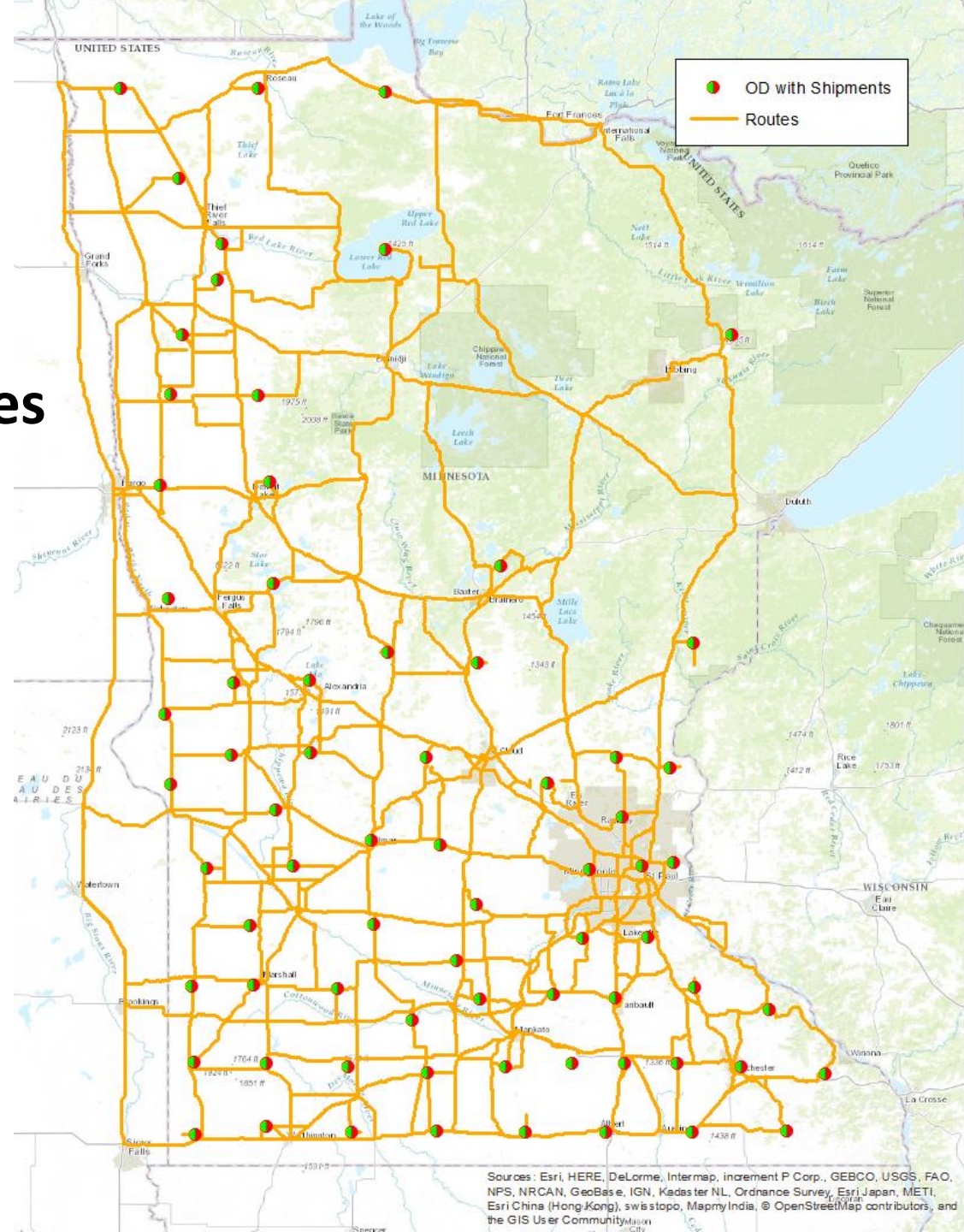
- Network Data Description
 - (87 MN counties) * (87 MN counties) = 7,569 total O-D routes
 - Appended shipment data to these O-D routes (4,489 routes)
 - Merged these O-D routes into one “flattened” dataset
 - Joined the merged O-D routes with 30,389 MN road segments
 - Resultant layer contains 30,389 road segments containing **road usage** and **shipment weight totals**



Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

Origins, Destinations, and O-D Routes

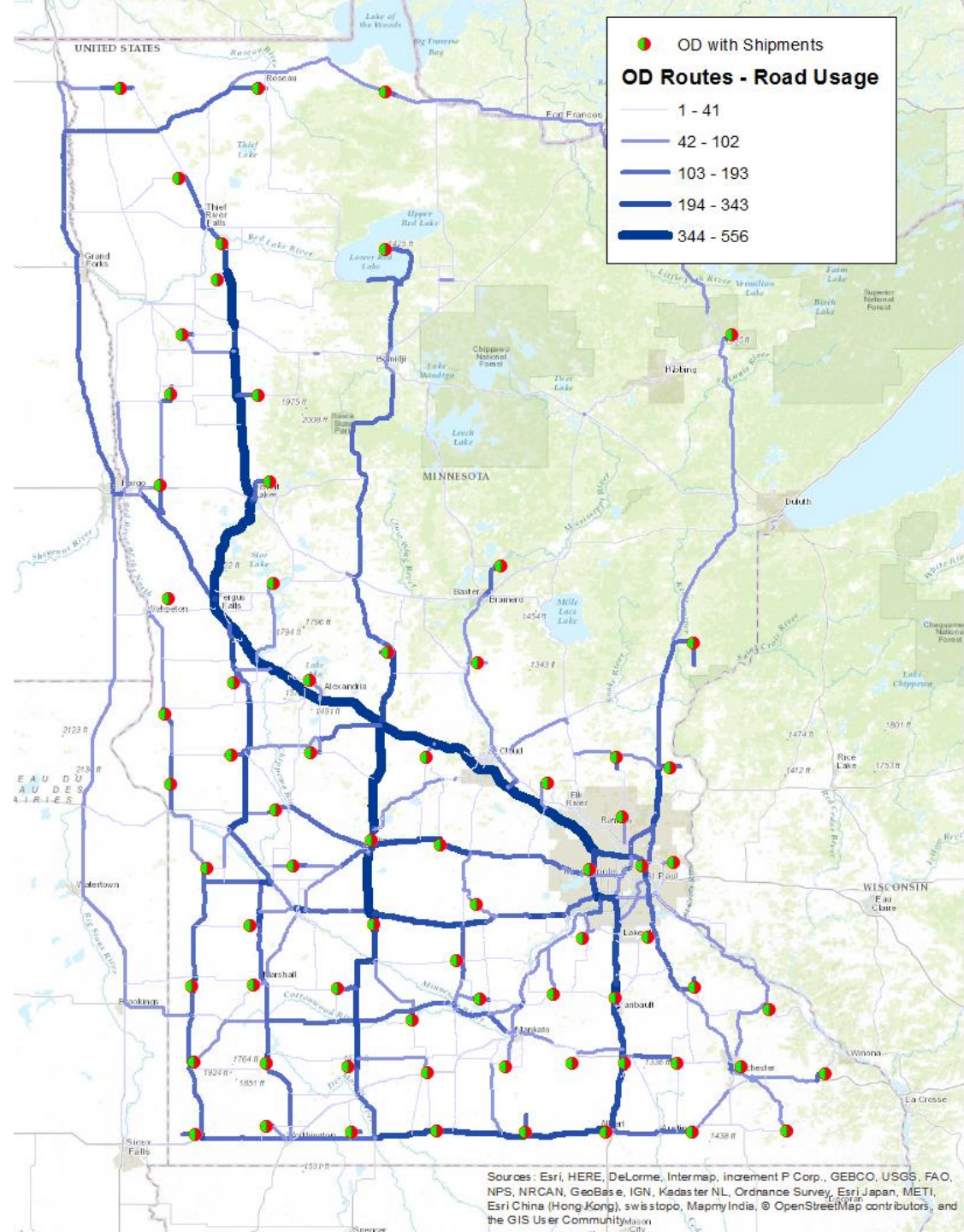
Origins (All MN Counties)
Destinations (All MN Counties)



Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

Road Usage

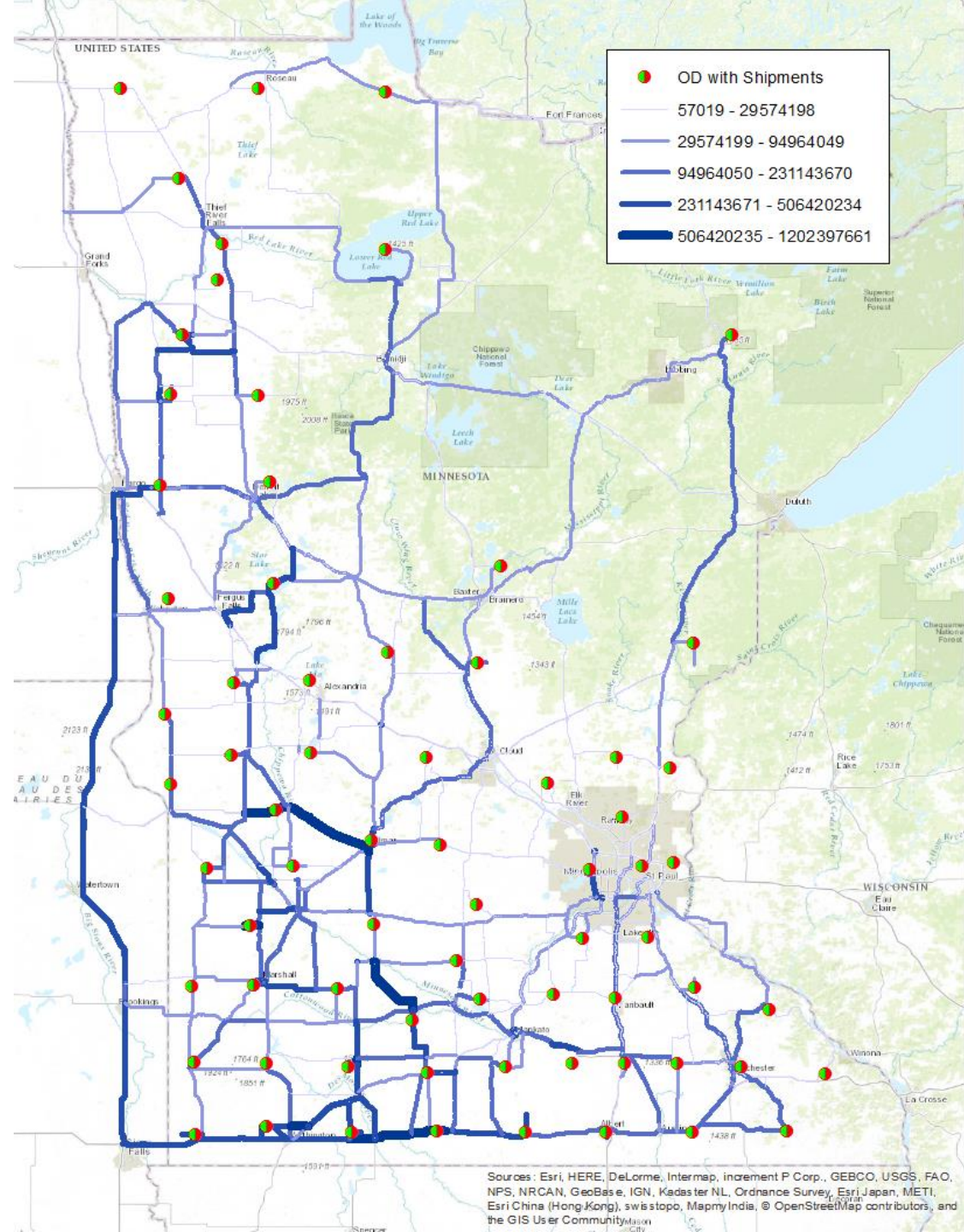
Origins (All MN Counties)
Destinations (All MN Counties)



Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

Shipping Weight

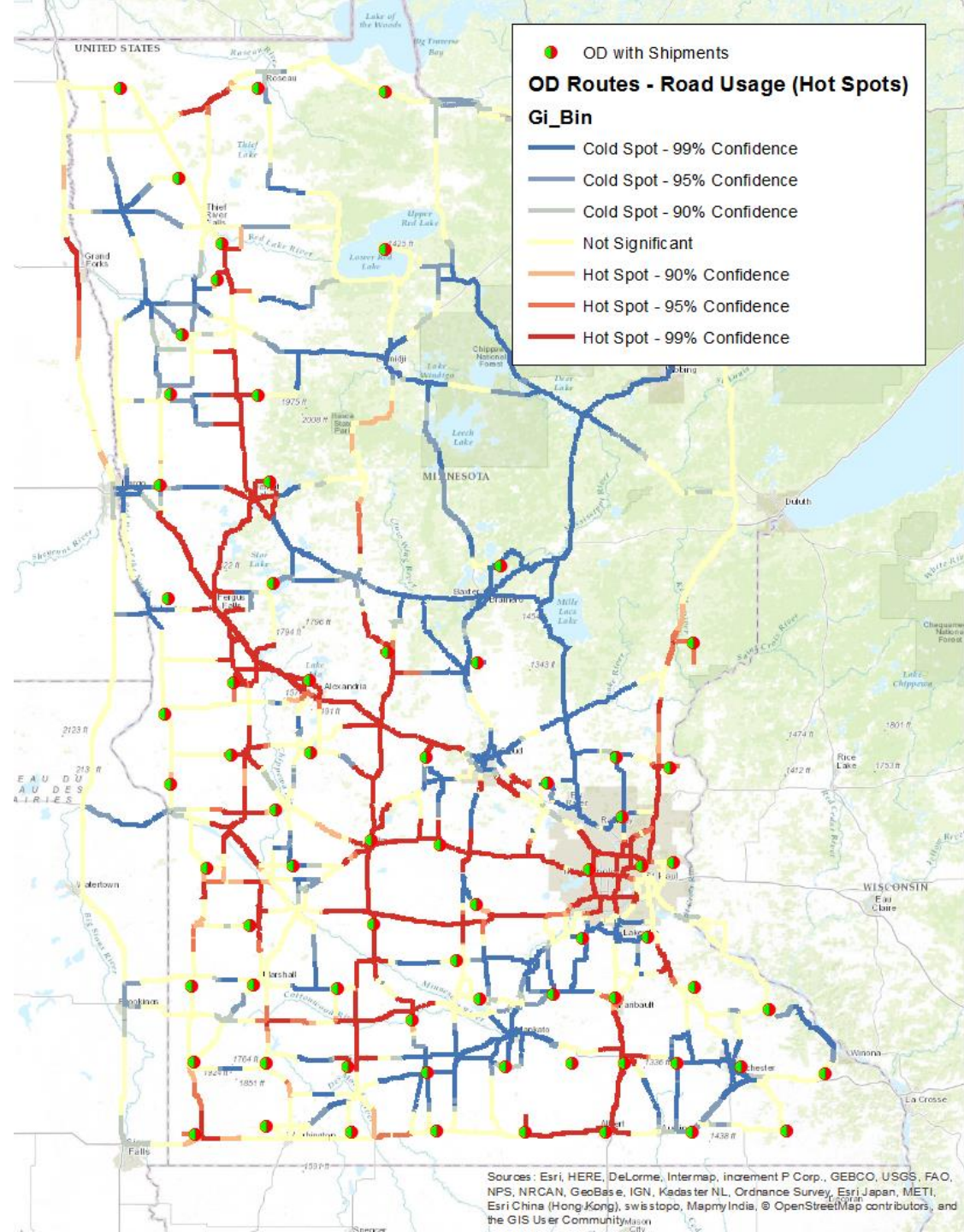
Origins (All MN Counties)
Destinations (All MN Counties)



Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

Road Usage (Hot Spots)

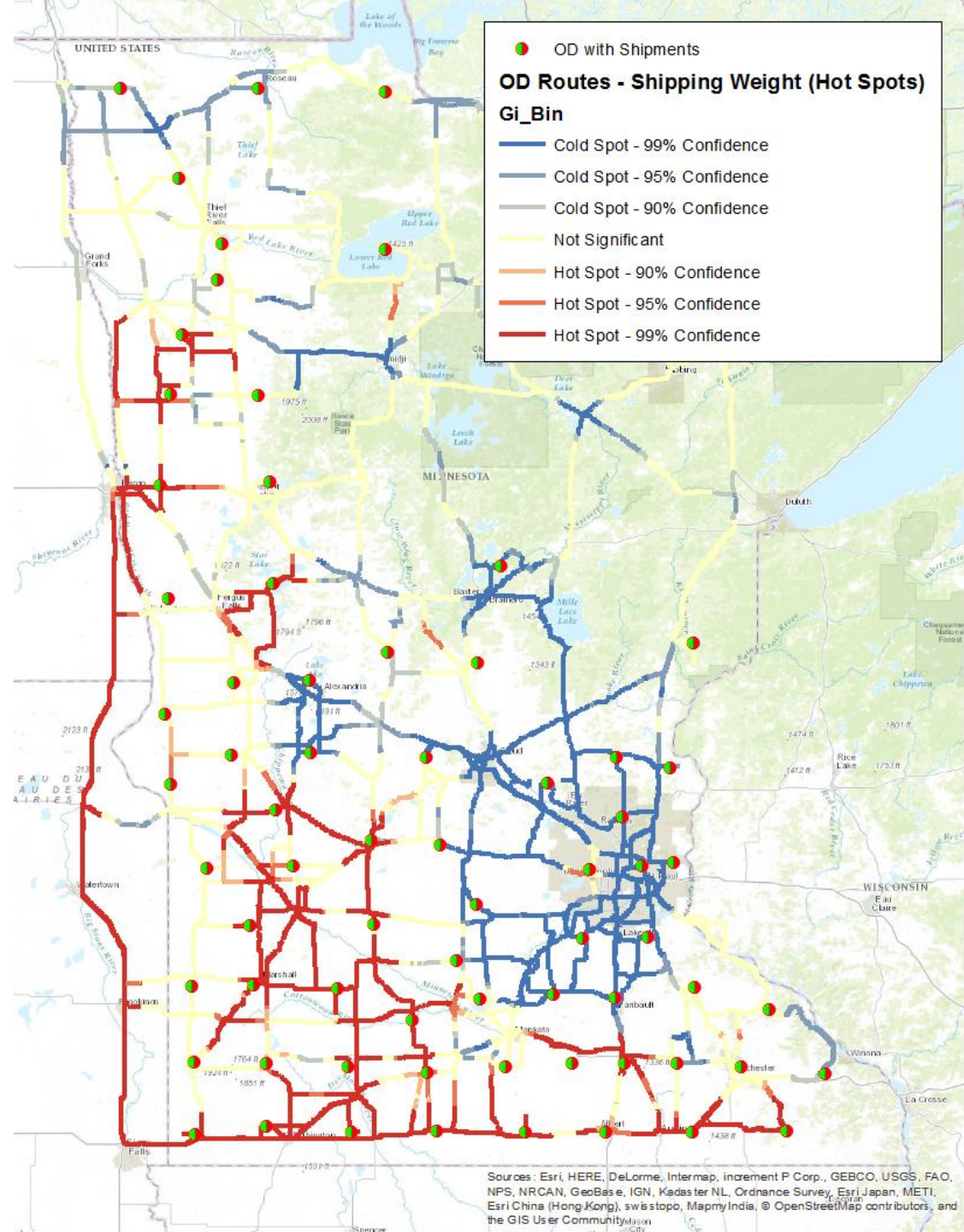
Origins (All MN Counties)
Destinations (All MN Counties)



Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

Shipping Weight (Hot Spots)

Origins (All MN Counties)
Destinations (All MN Counties)



Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

- Future Work / Analysis
 - Additional Industry Clusters (e.g., Agricultural Products, Animal Products, Mining Products, etc.) and Commodity Shipment Types (e.g., Processed Food Products, Dimension Stone, etc.)
 - More granular unit of analysis (1,031 MN zip codes) * (1,031 MN zip codes) = 1,062,961 total O-D routes
- Devise streamlined, and faster, data processing workflows...

Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

- Issue / Need

Run analytics:

- **against data that is too big for a single desktop machine**
 - hundreds of millions of 911 calls accumulated over years
 - billions of observations of ship movements ingested through GeoEvent
- **designed to gain insight into both spatial and temporal patterns**
- **against massive collections in a scalable manner**
- **and meet time constraints**



months



weeks



days



hours



minutes

Evolving Supply Chains and Local Freight Flows: GIS Analysis of Minnesota Cereal Grain Movement

Developing novel spatial and spatiotemporal methods that can take advantage of newly emerging data-intensive computational resources.

- ArcGIS Enterprise
- XSEDE and GIS (<http://www.gisandbox.org>)

SDSC SAN DIEGO
SUPERCOMPUTER CENTER



Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

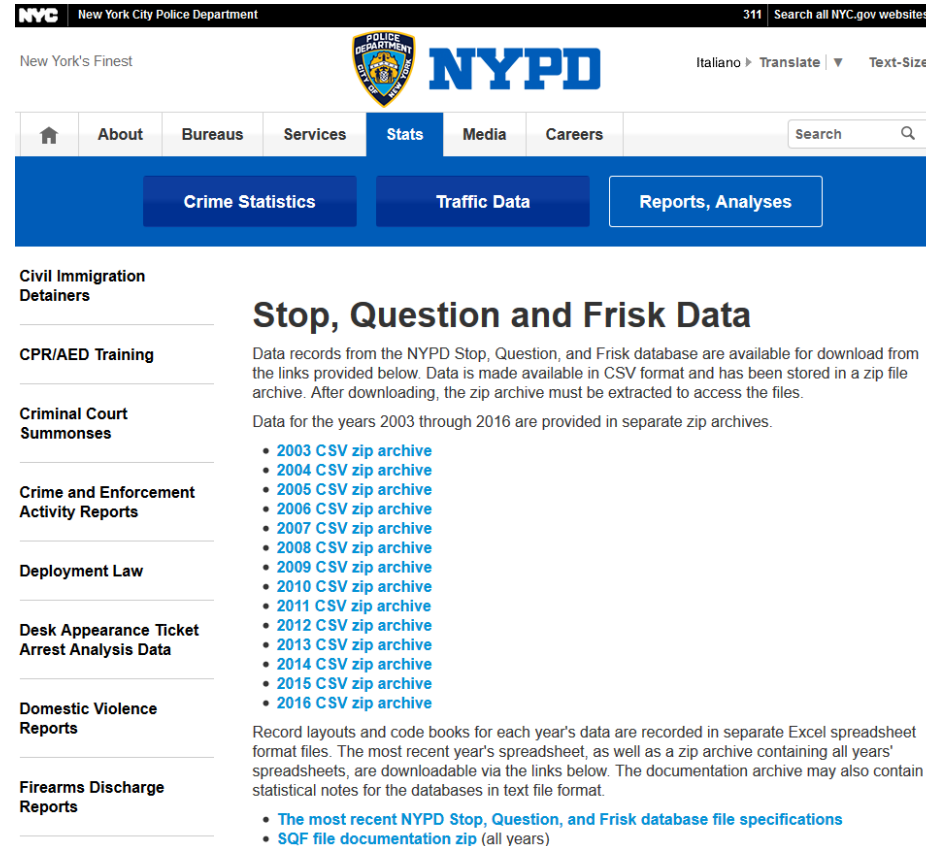
- Domain – Spatial Justice / Spatial Equality
- Research Question – “Does the race or ethnicity of an individual being stopped by a police officer have a significant role in an individual being frisked and by how much?”

Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- Policy encourages police officers to stop people they deem suspicious, question them, and to frisk them for drugs, contraband, or weapons if illegal activities are suspected.
 - Reasonable suspicion is the belief that someone poses a danger, has committed a crime, or is about to commit a crime.
- Race cannot be a factor for the frisk.
- The New York City Stop-and-Frisk Policy is an example of how a policy intended to keep the public safe, now has a negative public perspective.

Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- Data Provider
 - Stop-and-Frisk data records are available from the NYPD Stop, Question, and Frisk database.



NYC New York City Police Department 311 Search all NYC.gov websites

New York's Finest

Italliano Translate Text-Size

Home About Bureaus Services **Stats** Media Careers Search

Crime Statistics Traffic Data Reports, Analyses

Civil Immigration Detainers

CPR/AED Training

Criminal Court Summonses

Crime and Enforcement Activity Reports

Deployment Law

Desk Appearance Ticket Arrest Analysis Data

Domestic Violence Reports

Firearms Discharge Reports

Stop, Question and Frisk Data

Data records from the NYPD Stop, Question, and Frisk database are available for download from the links provided below. Data is made available in CSV format and has been stored in a zip file archive. After downloading, the zip archive must be extracted to access the files.

Data for the years 2003 through 2016 are provided in separate zip archives.

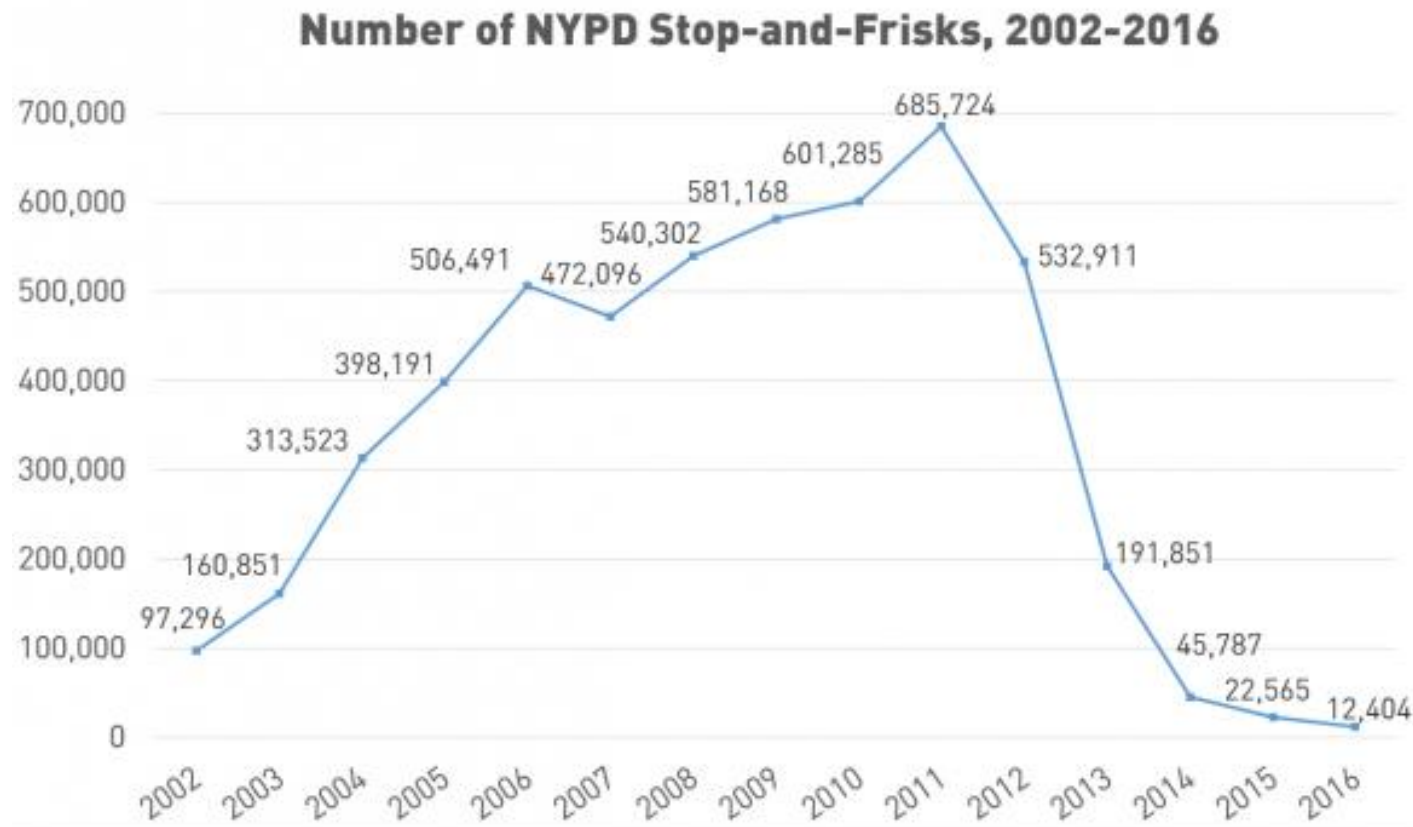
- [2003 CSV zip archive](#)
- [2004 CSV zip archive](#)
- [2005 CSV zip archive](#)
- [2006 CSV zip archive](#)
- [2007 CSV zip archive](#)
- [2008 CSV zip archive](#)
- [2009 CSV zip archive](#)
- [2010 CSV zip archive](#)
- [2011 CSV zip archive](#)
- [2012 CSV zip archive](#)
- [2013 CSV zip archive](#)
- [2014 CSV zip archive](#)
- [2015 CSV zip archive](#)
- [2016 CSV zip archive](#)

Record layouts and code books for each year's data are recorded in separate Excel spreadsheet format files. The most recent year's spreadsheet, as well as a zip archive containing all years' spreadsheets, are downloadable via the links below. The documentation archive may also contain statistical notes for the databases in text file format.

- [The most recent NYPD Stop, Question, and Frisk database file specifications](#)
- [SQF file documentation zip](#) (all years)

Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

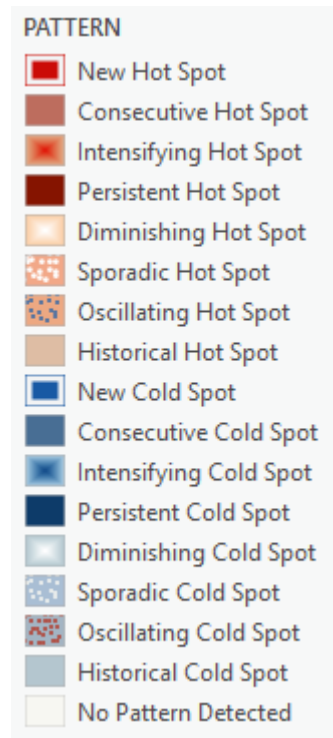
- Stop-and-Frisk Data Description
 - 5,162,445 New York City police stops and street interrogations (2002-2016)



Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- Asian
- Stop-and-Frisk
- 2014

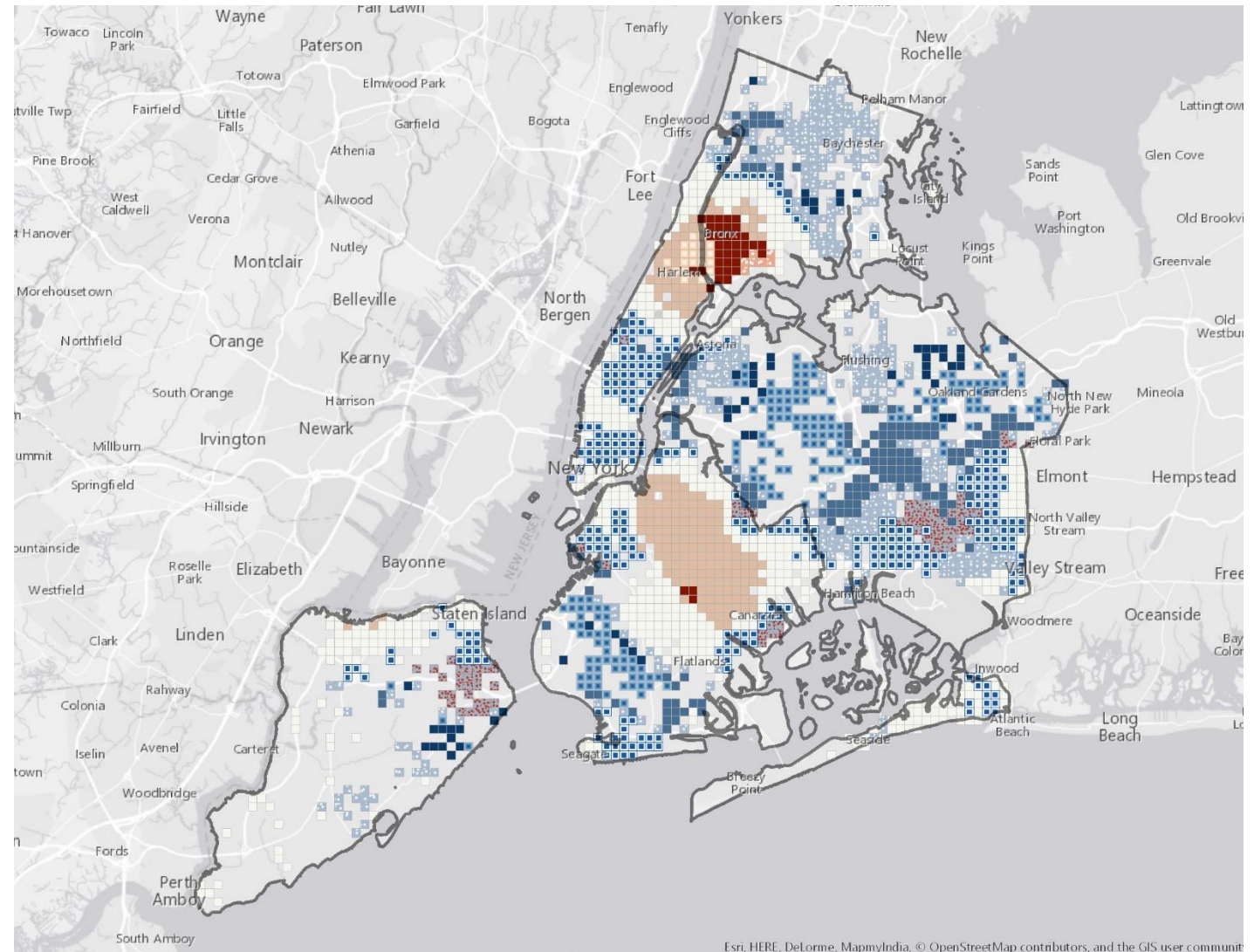
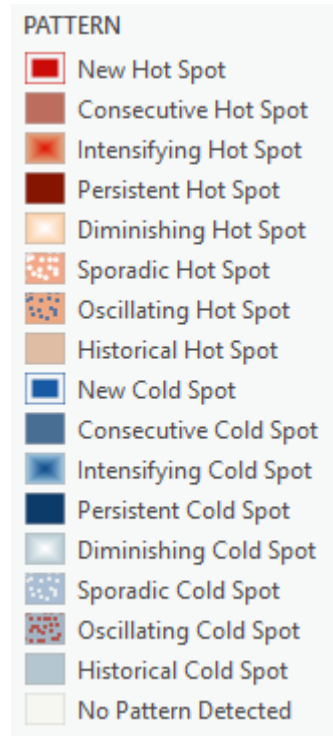
Emerging Hot Spot Analysis
Spatiotemporal Trends



Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- Black
- Stop-and-Frisk
- 2014

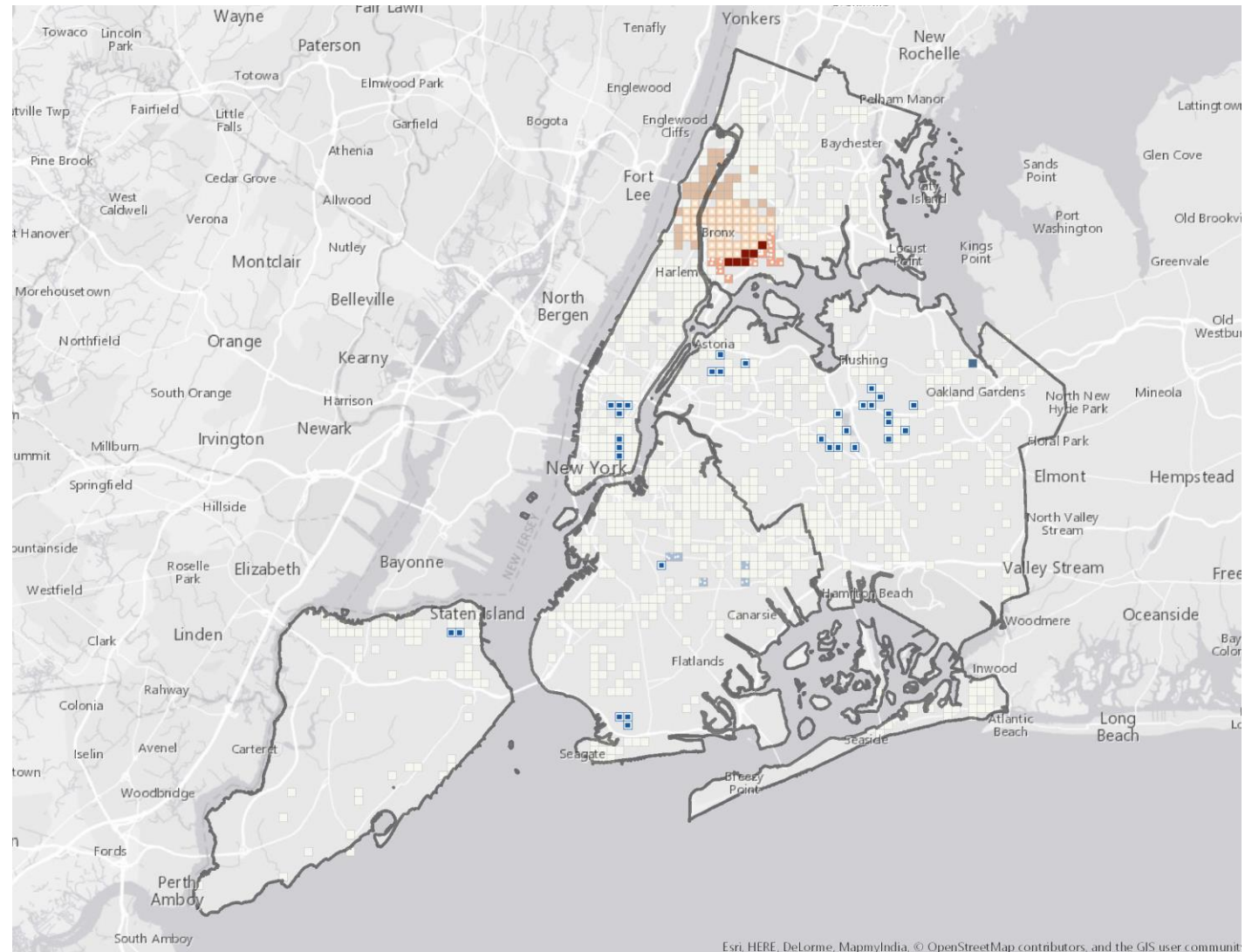
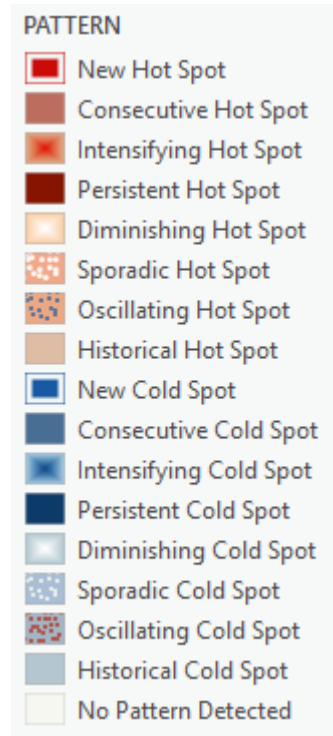
Emerging Hot Spot Analysis
Spatiotemporal Trends



Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- Black Hispanic
- Stop-and-Frisk
- 2014

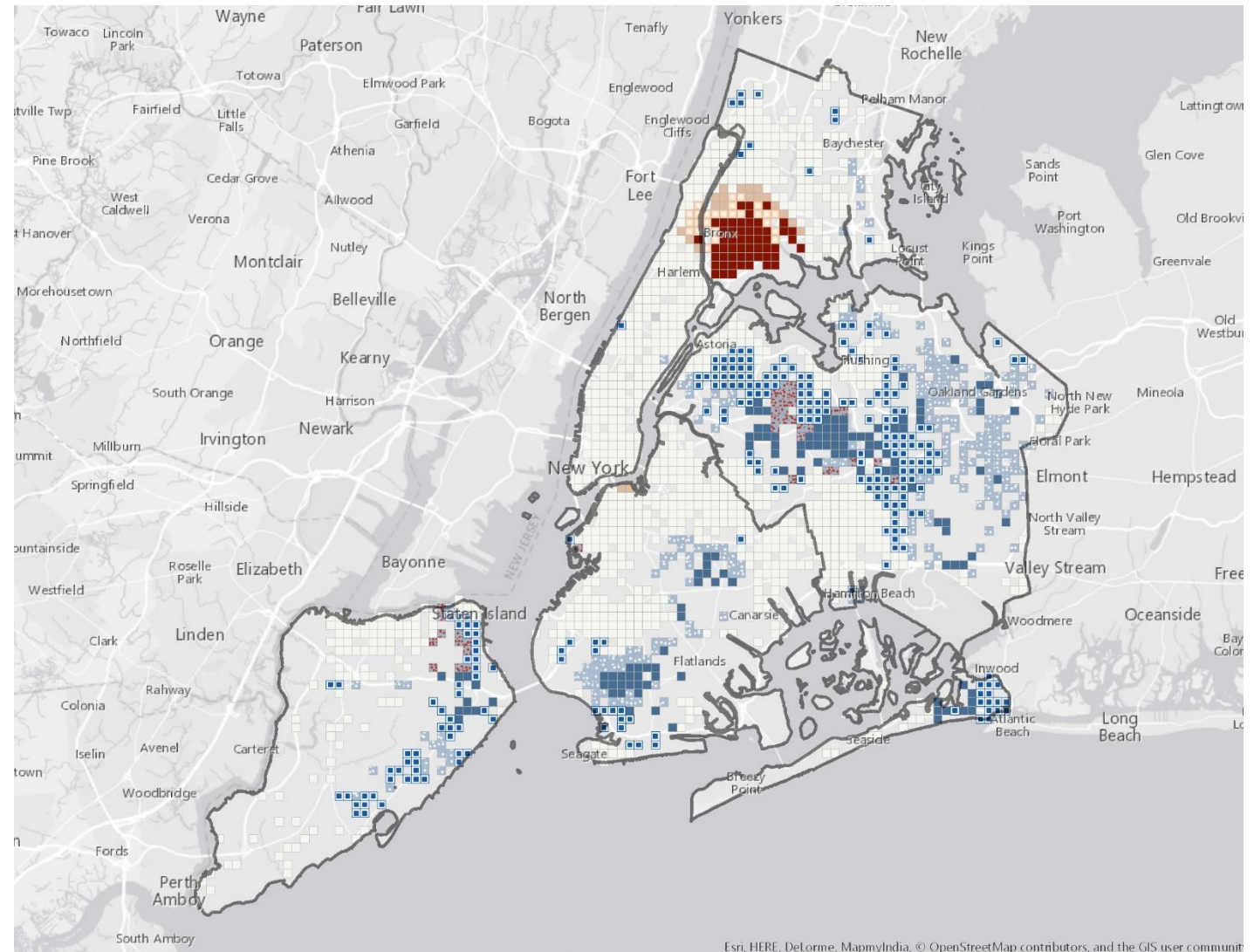
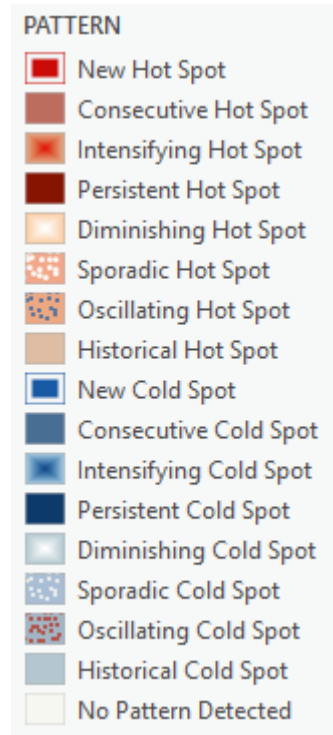
Emerging Hot Spot Analysis
Spatiotemporal Trends



Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- White Hispanic
- Stop-and-Frisk
- 2014

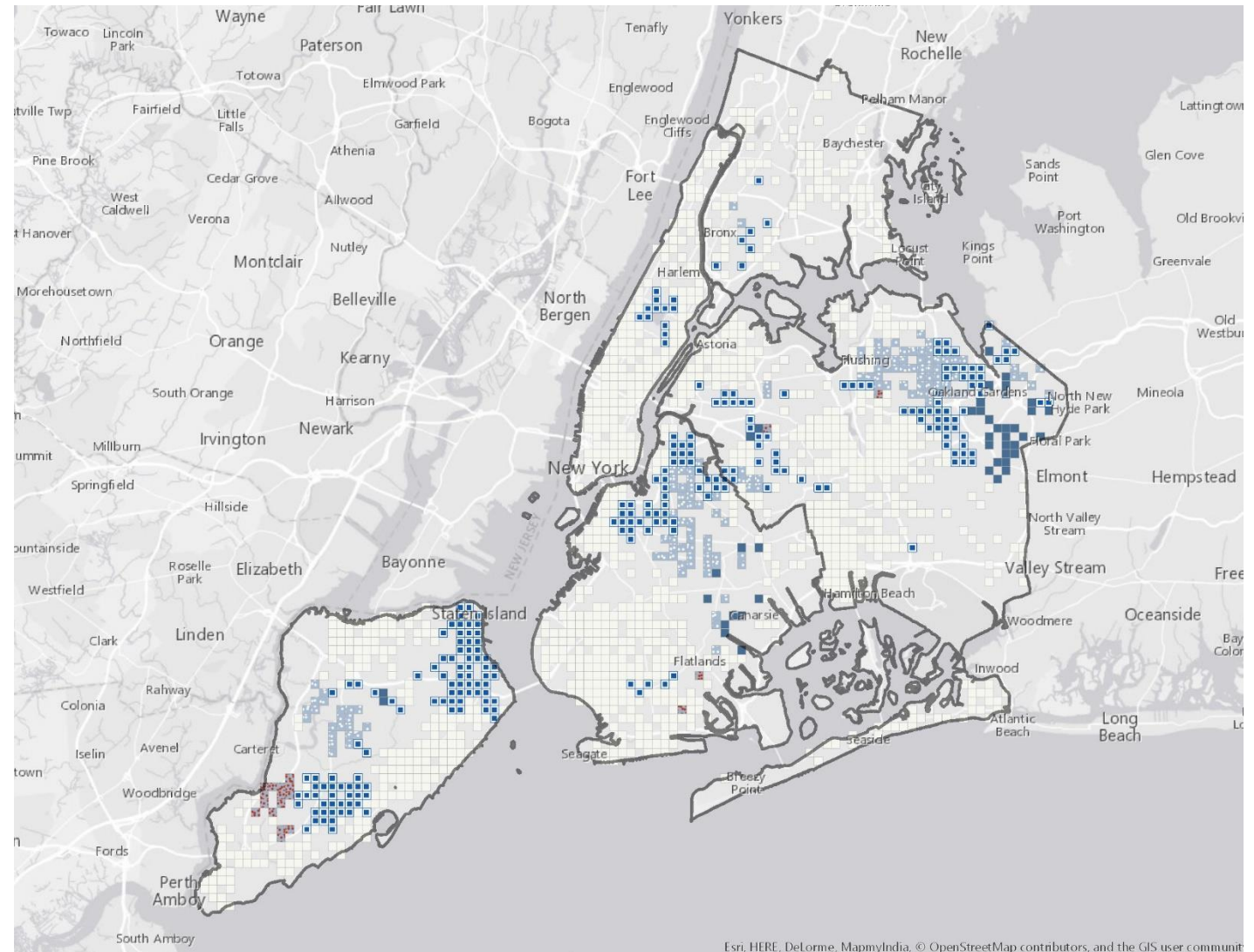
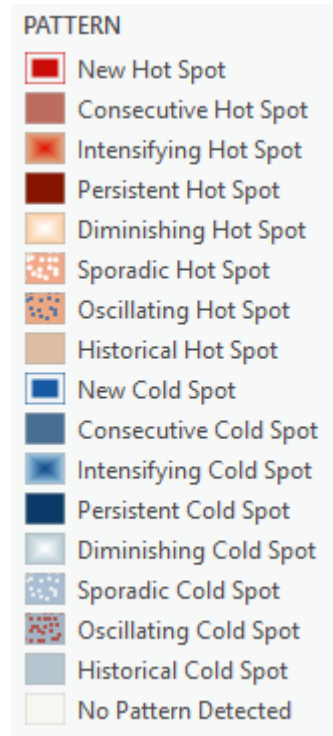
Emerging Hot Spot Analysis
Spatiotemporal Trends



Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

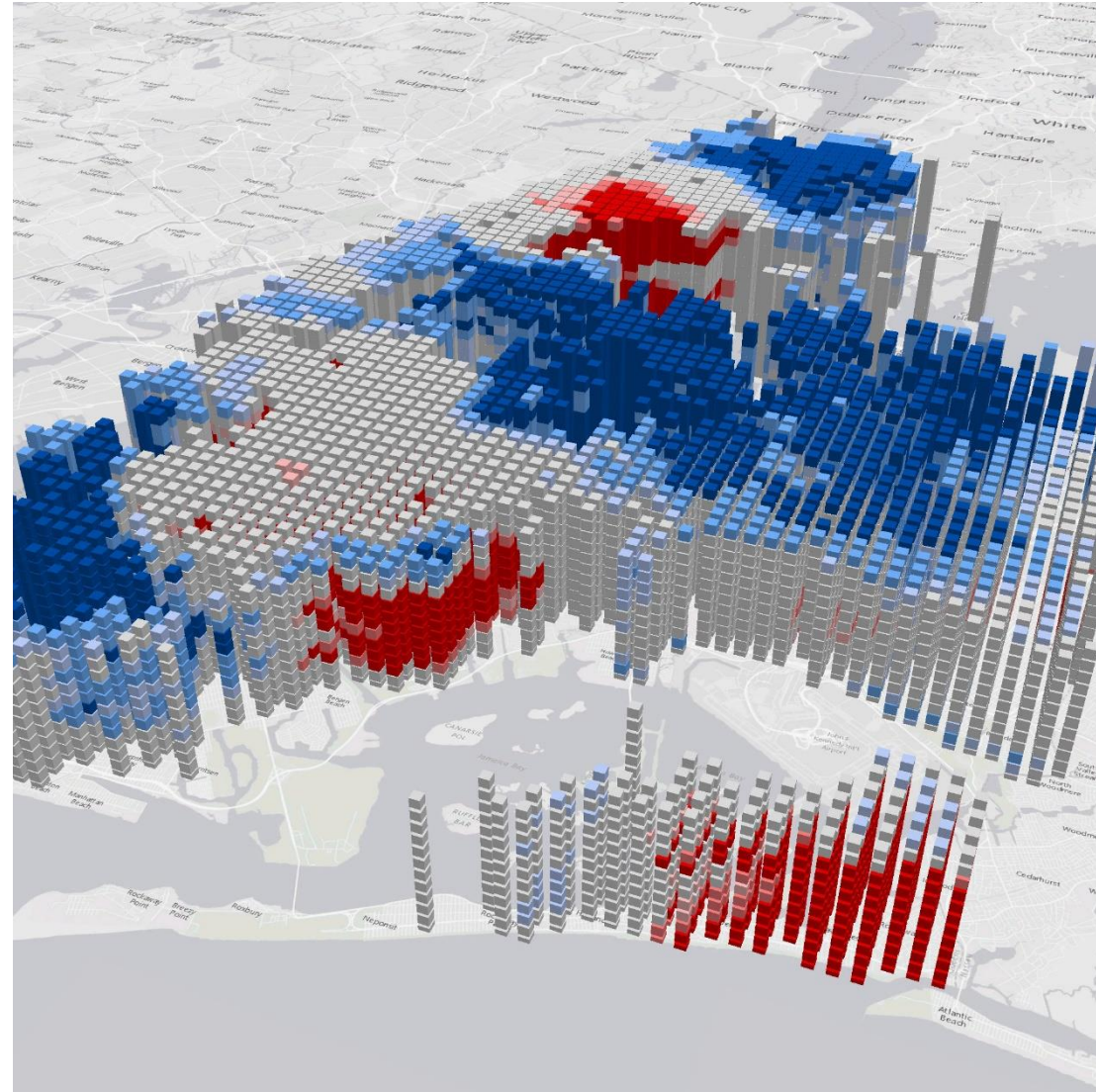
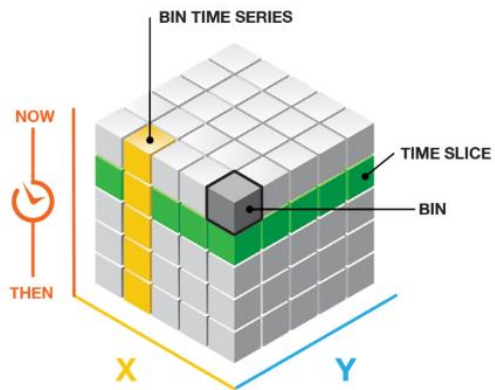
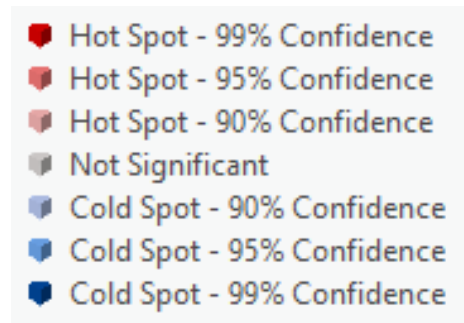
- White
- Stop-and-Frisk
- 2014

Emerging Hot Spot Analysis
Spatiotemporal Trends



Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- Black
- Stop-and-Frisk
- 2014



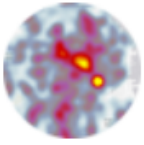







Emerging Hot Spot Analysis
Visualize Space-Time Cube

Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

- Future Work / Analysis
 - Conduct multiple, spatiotemporal analyses (e.g., across years, specific months, days, hours, specific attribute types, and combinations of these) for NYC.
 - Examine this issue in other cities (e.g., Philadelphia, Chicago, Los Angeles)
- Devise streamlined, shareable, analytic processing workflows...

Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

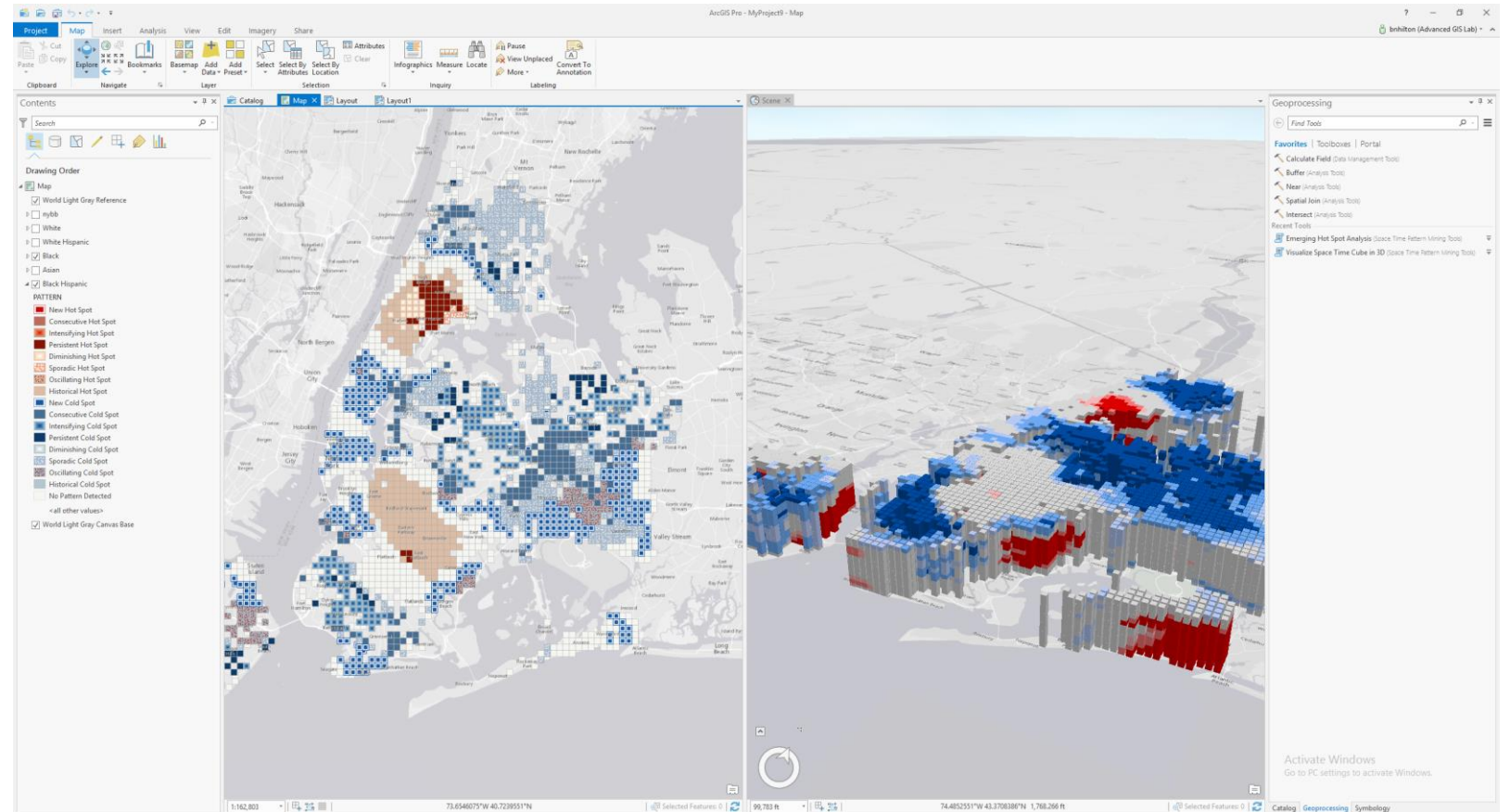
- Issue / Need
 - Web-based, data analytics and exploration tool for conducting multiple, spatiotemporal analyses, of spatial and non-spatial data, where you can:

 Explore, Analyze, Iterate	 Visualization and Analysis
 Drag-and-Drop Visualization	 Add Demographic Data
 Guided Workflows	 Share and Communicate
 Connect to Data Sources	 Record and Rerun Workflows

Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

Developing methodologies to derive insight into spatial big data for inferential understanding and decision making.

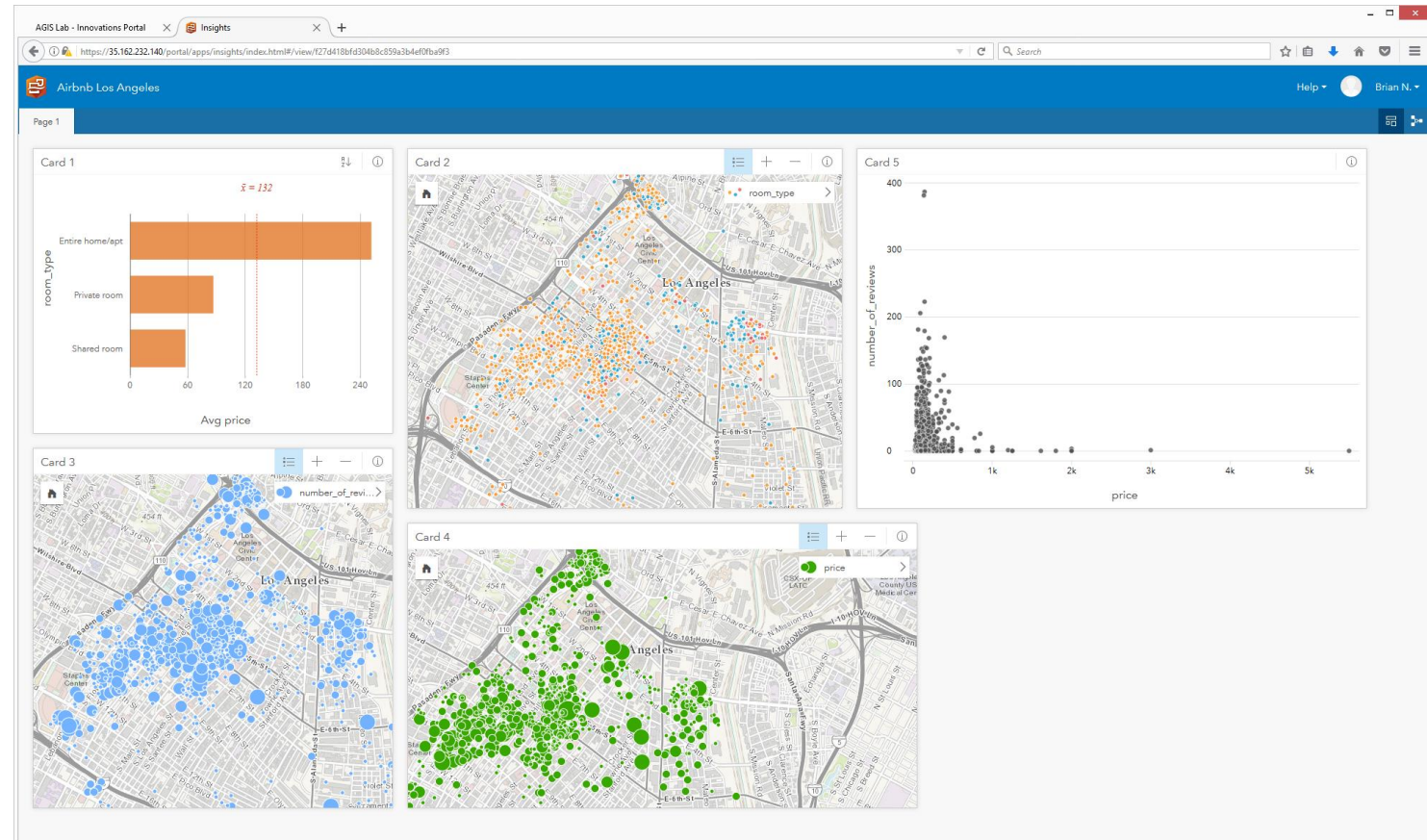
- ArcGIS Pro
- Emerging Hot Spot Analysis
- Space-Time Cube



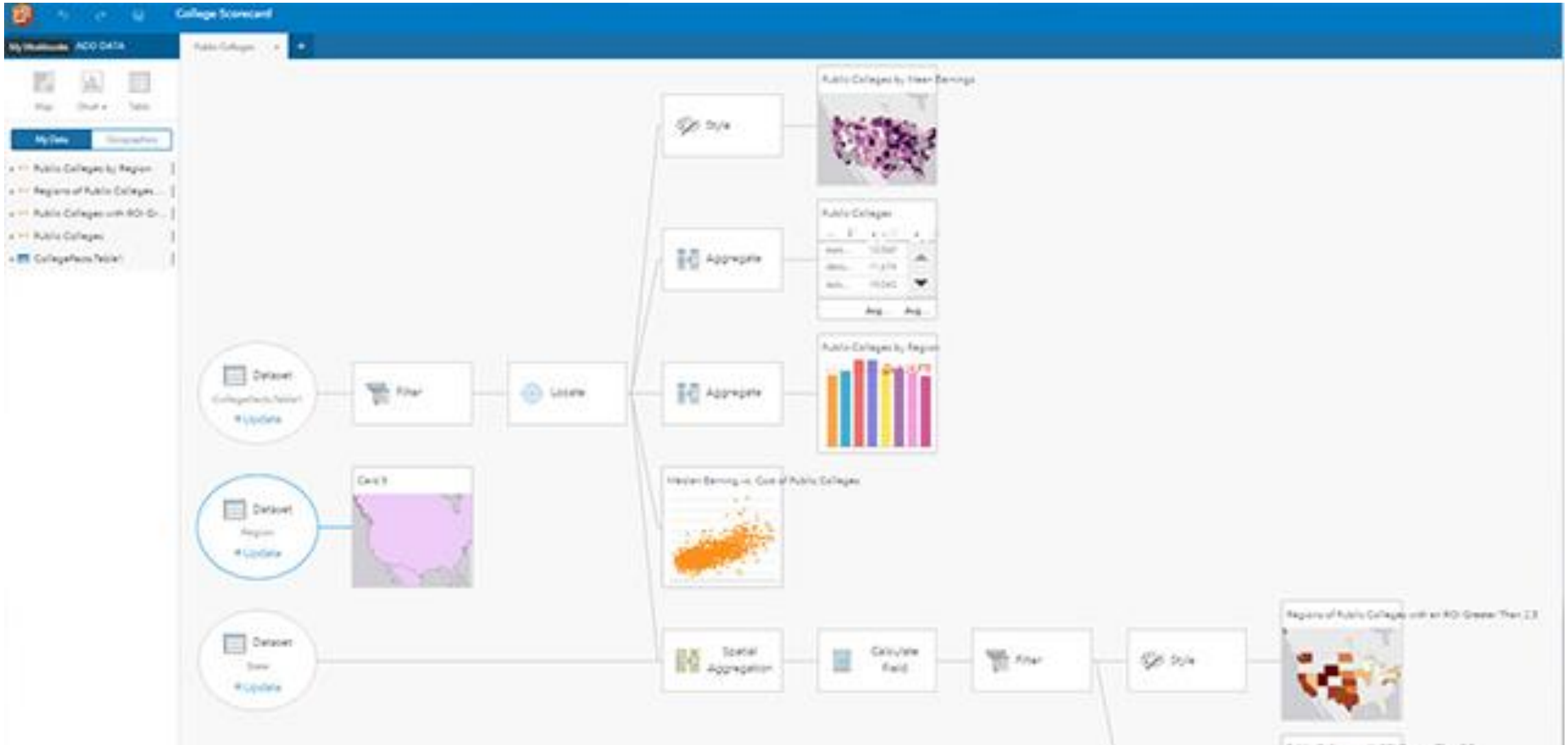
Stop-and-Frisk Policy from a Quantitative and Spatial Perspective

Developing methodologies to derive insight into spatial big data for inferential understanding and decision making.

- Insights for ArcGIS
- Web-based
- Data analytics tool

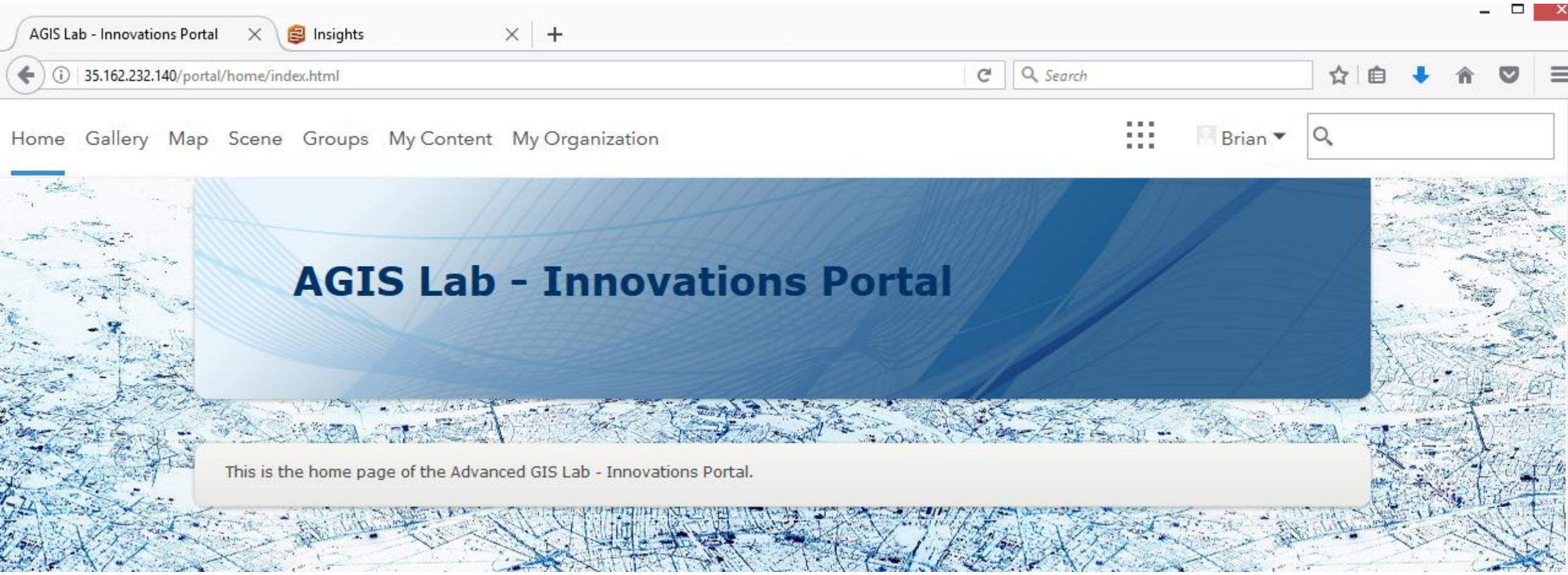


Stop-and-Frisk Policy from a Quantitative and Spatial Perspective



Insights for ArcGIS - Hands-on demo

<http://agislab.org>



The image shows a web browser window displaying the AGIS Lab - Innovations Portal. The browser's address bar shows the URL `35.162.232.140/portal/home/index.html`. The page features a navigation menu with links for Home, Gallery, Map, Scene, Groups, My Content, and My Organization. A user profile for 'Brian' is visible in the top right corner. The main content area has a blue header with the text 'AGIS Lab - Innovations Portal' and a white box below it stating 'This is the home page of the Advanced GIS Lab - Innovations Portal.' The background of the page is a satellite map.

AGIS Lab - Innovations Portal

This is the home page of the Advanced GIS Lab - Innovations Portal.

Insights for ArcGIS - Hands-on demo

Home Gallery Map Scene Groups My Content My Organization

Brian



AGIS Lab - Innovations Portal



Insights for ArcGIS

EDIT SETTINGS ADD MEMBERS VIEW STATUS MANAGE LICENSES

Members

Viewing:

Name	Username	Last Login	Level	Role	Action
amcis_user1 amcis_gis1	amcis_user1	Aug 3, 2017	2	Publisher	
amcis_user10 amcis_gis10	amcis_user10	Not yet	2	Publisher	
amcis_user11 amcis_gis11	amcis_user11	Not yet	2	Publisher	

Members

Members

- 1 0
- 2 20

Total Members

Find...

The most...

Appendix



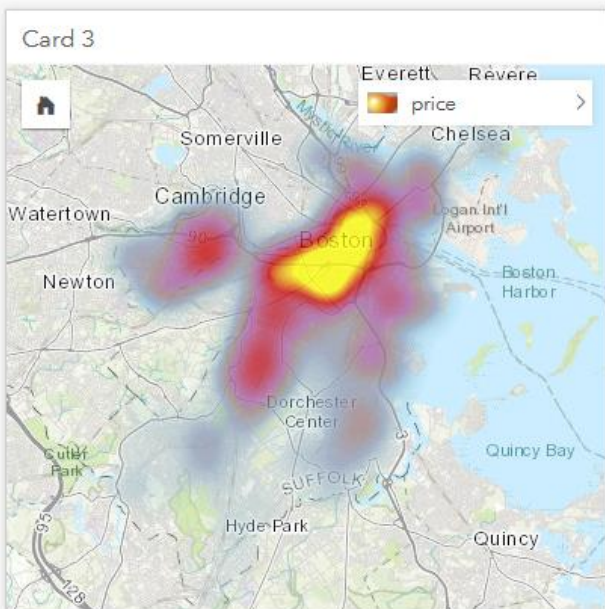
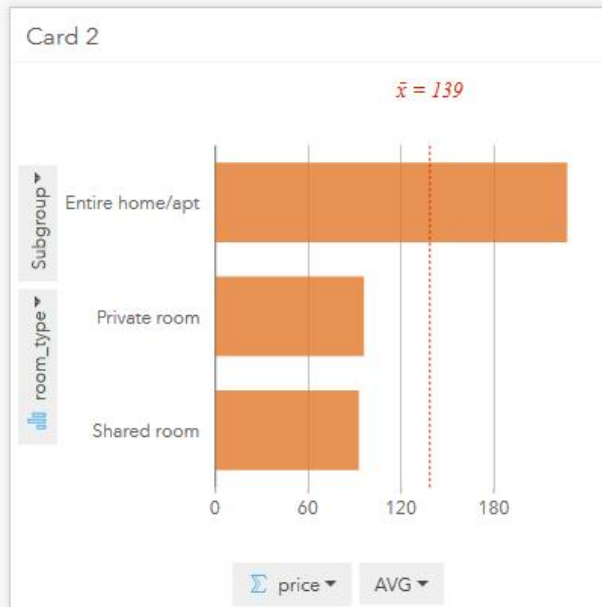
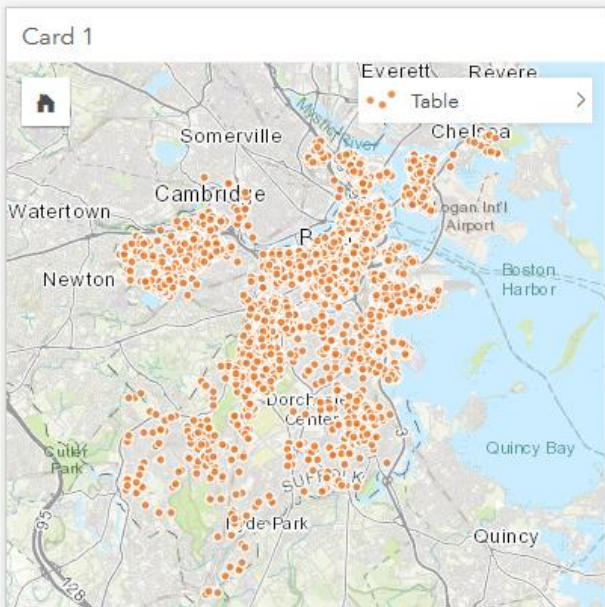
+ ADD DATA

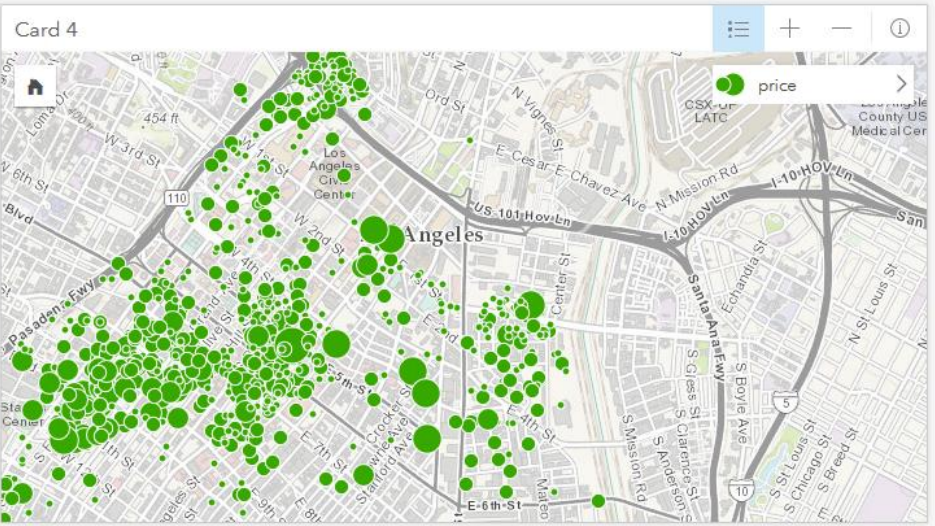
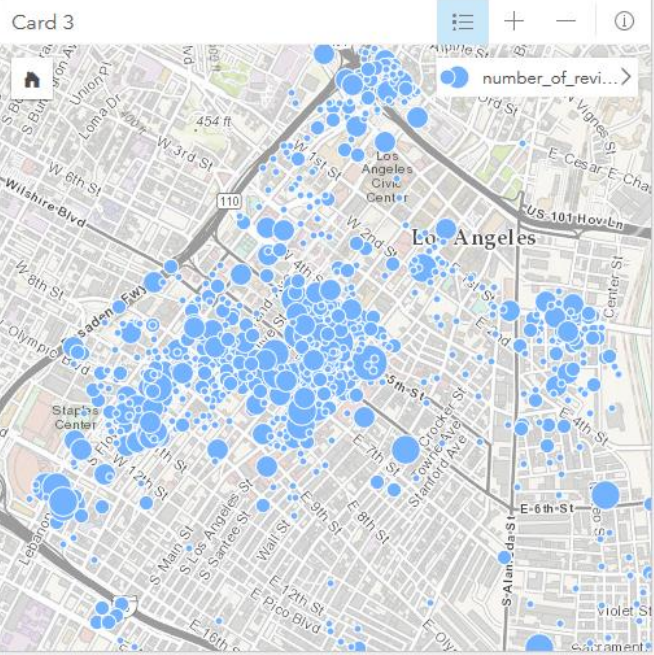
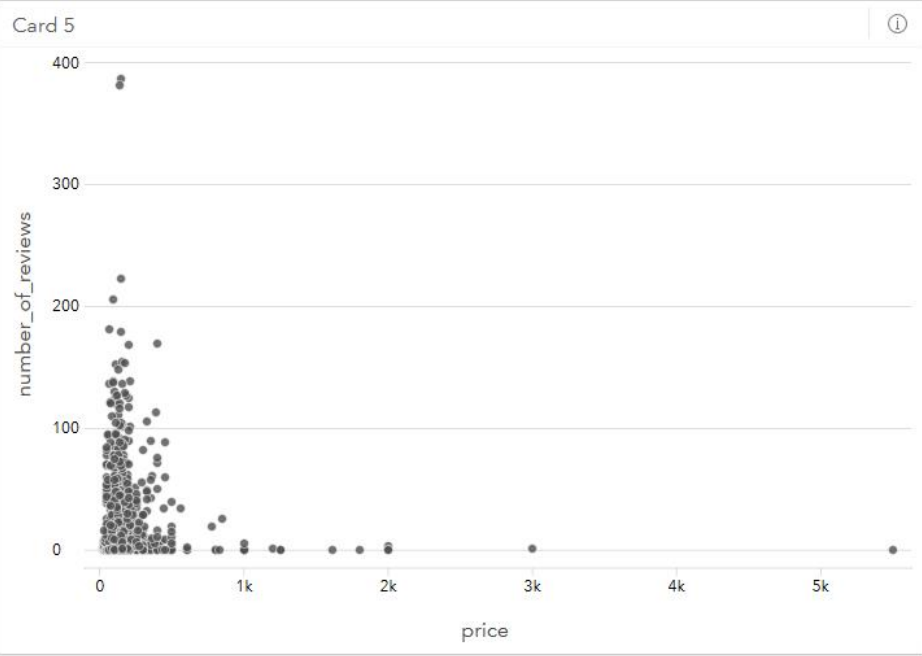
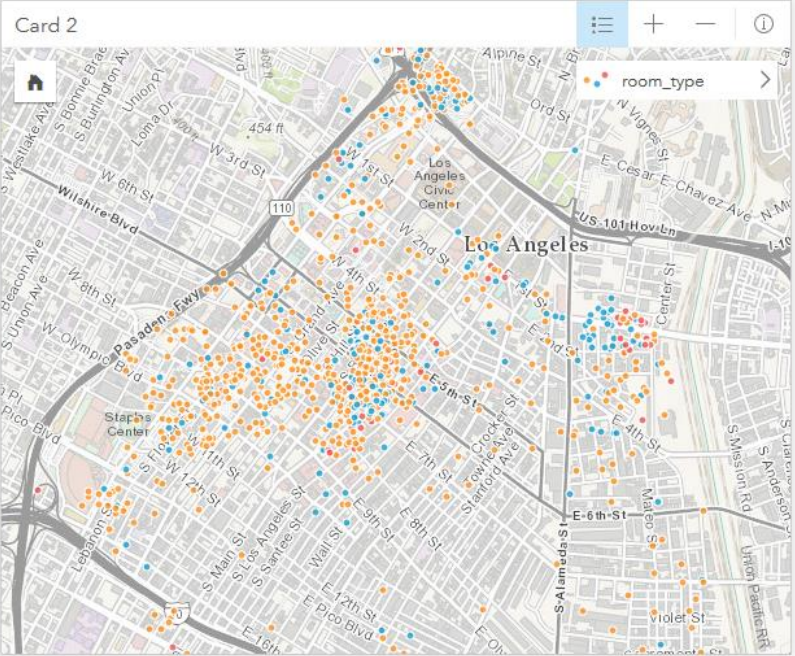
Page 1

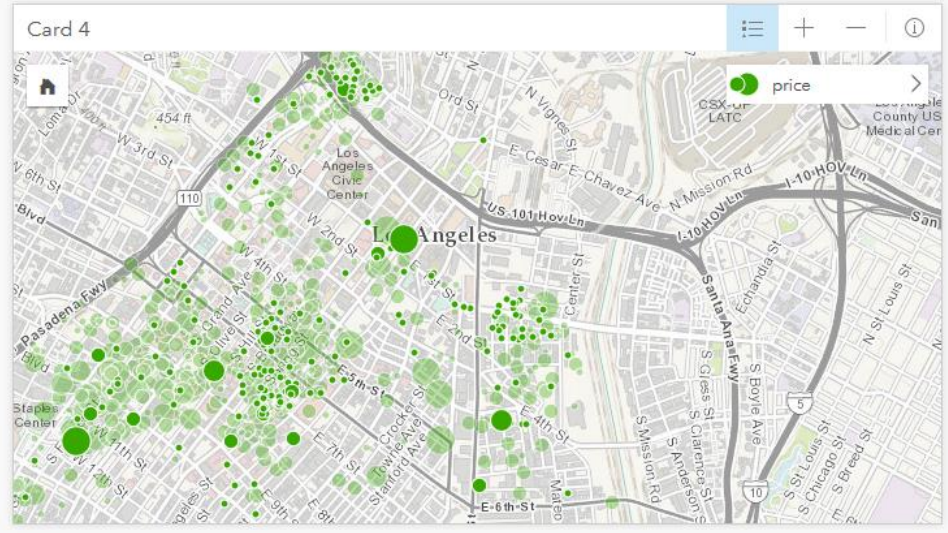
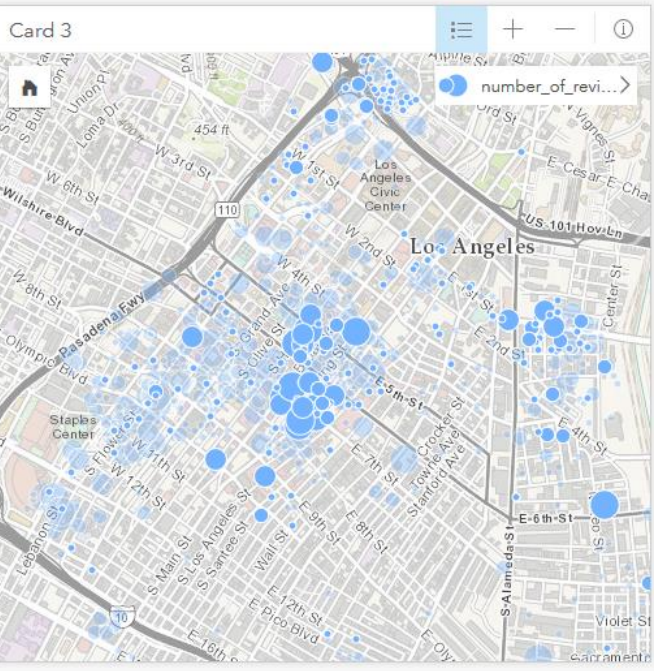
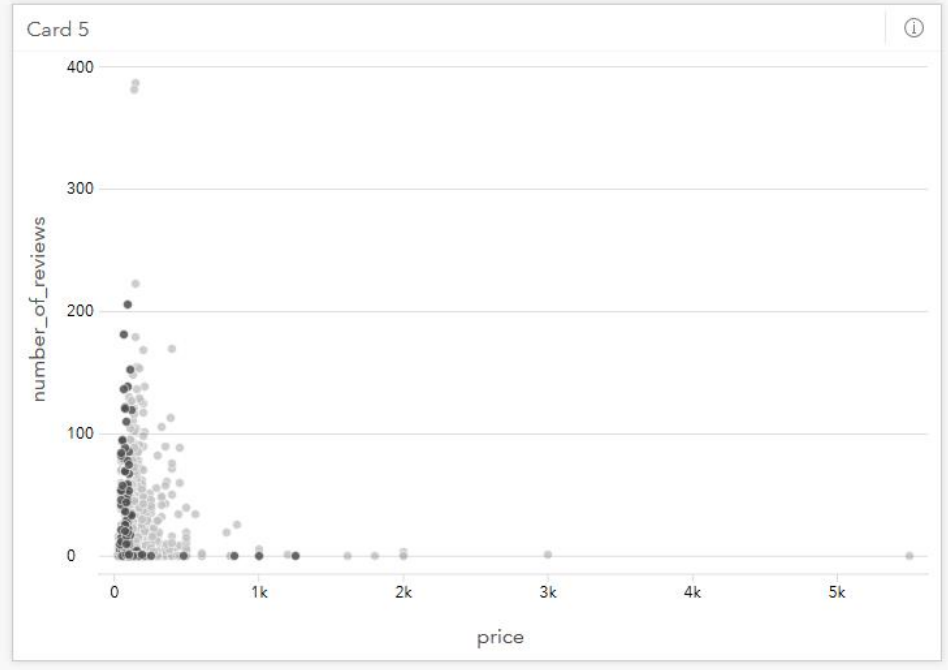
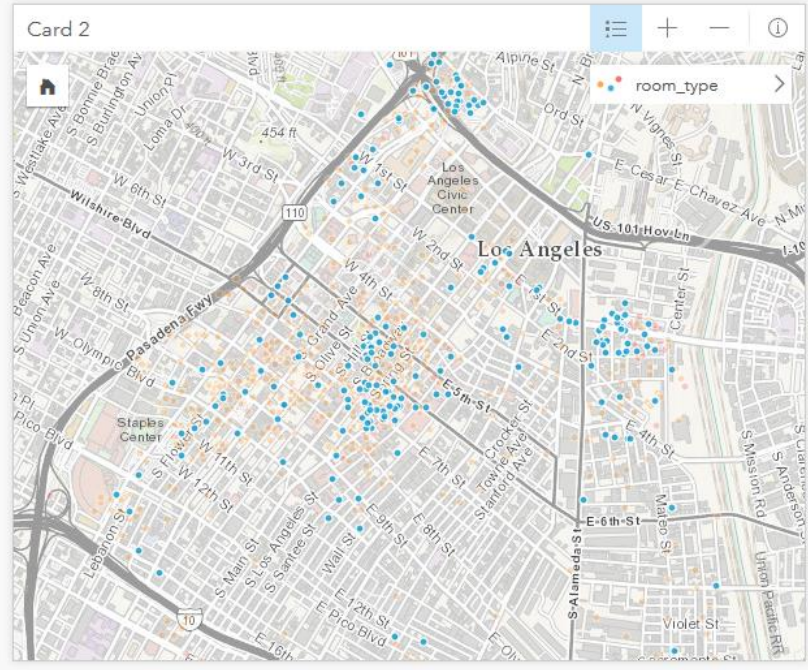
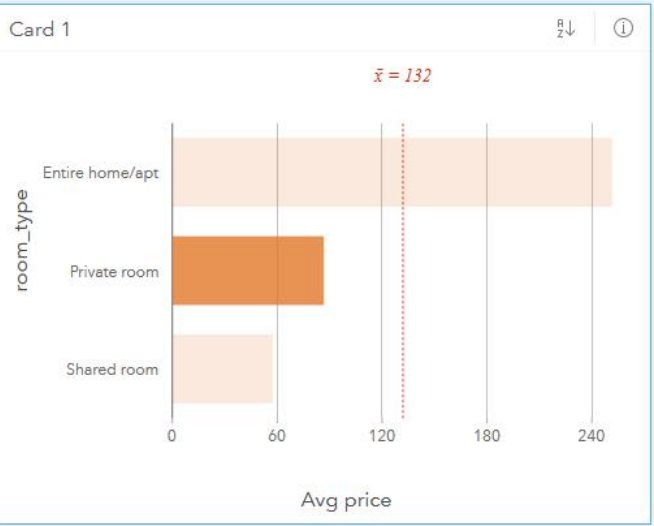


Map Chart Table

- Card 2
- Table







Additional Resources

Developing teaching resources to better understand the use of data manipulation techniques, spatial statistics, and spatial data-mining tasks related to spatial big data.

- Jupyter Notebooks
 - <https://developers.arcgis.com/python/>
 - <https://developers.arcgis.com/python/sample-notebooks/>
 - https://notebooks.esri.com/user/VFVGul1U7slf/notebooks/samples/04_gis_analysts_data_scientists/analyze_new_york_city_taxi_data.ipynb
- MapD: <https://www.mapd.com/>