



Innovative Models & Strategies for Data-Driven Governance

Fairfax Innovation Challenge – May 18, 2017

Sallie Keller

Director & Professor of Statistics

 VirginiaTech.
Biocomplexity Institute

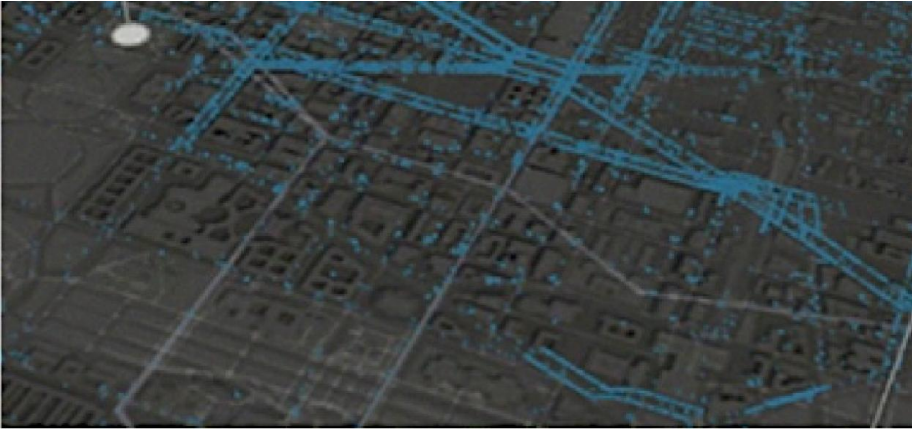


SDAL SOCIAL &
DECISION ANALYTICS
LABORATORY

Biocomplexity Institute of Virginia Tech

- The study of life and environment as a **complex system**
- Understanding biology **in the context of** ecosystems and human-created systems
- **Transdisciplinary** team science

“From molecules to policy”



Problem-Driven Science

Our information biology approach is putting research to work in the real world, breaking down barriers between science and policy.

Social and Decision Analytics Lab

The Social and Decision Analytics Laboratory brings together statisticians and social and behavioral scientists to embrace today's data revolution, developing evidence-based research and quantitative methods to inform policy decision-making.

- **Science of *ALL* Data**
- **Community Learning Data-Driven Discovery**
 - Education and Labor Force Analytics
 - Health and Well Being Analytics
 - Emergency Management Analytics
 - Industrial Innovation Analytics
- **Information Diffusion Analytics**

ALL Data-The new lens for social observing

Infrastructure



- Condition
- Operations
- Resilience
- Sustainability

Environment



- Climate
- Pollution
- Noise
- Flora/Fauna

People



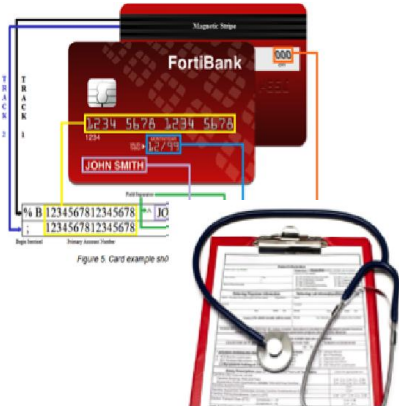
- Relationships
- Location
- Economic Condition
- Communication
- Activities
- Health

Bring the **ALL** data revolution to communities

Designed Data



Administrative Data

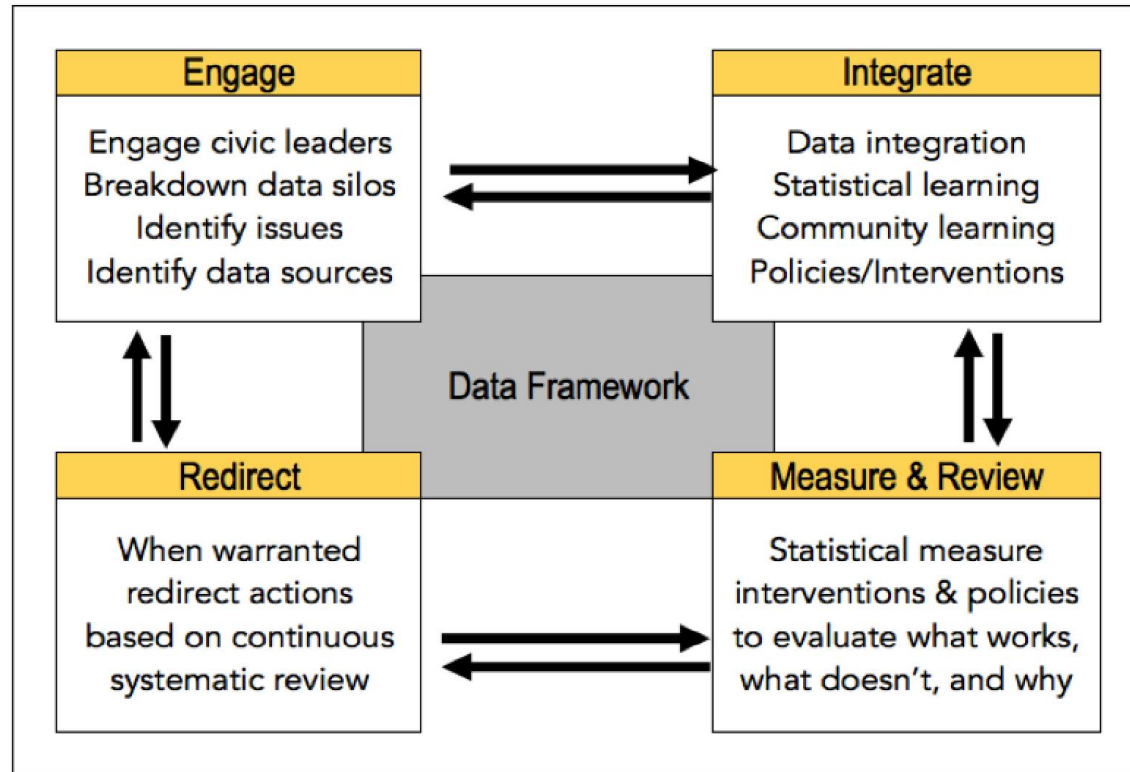


Opportunity Data

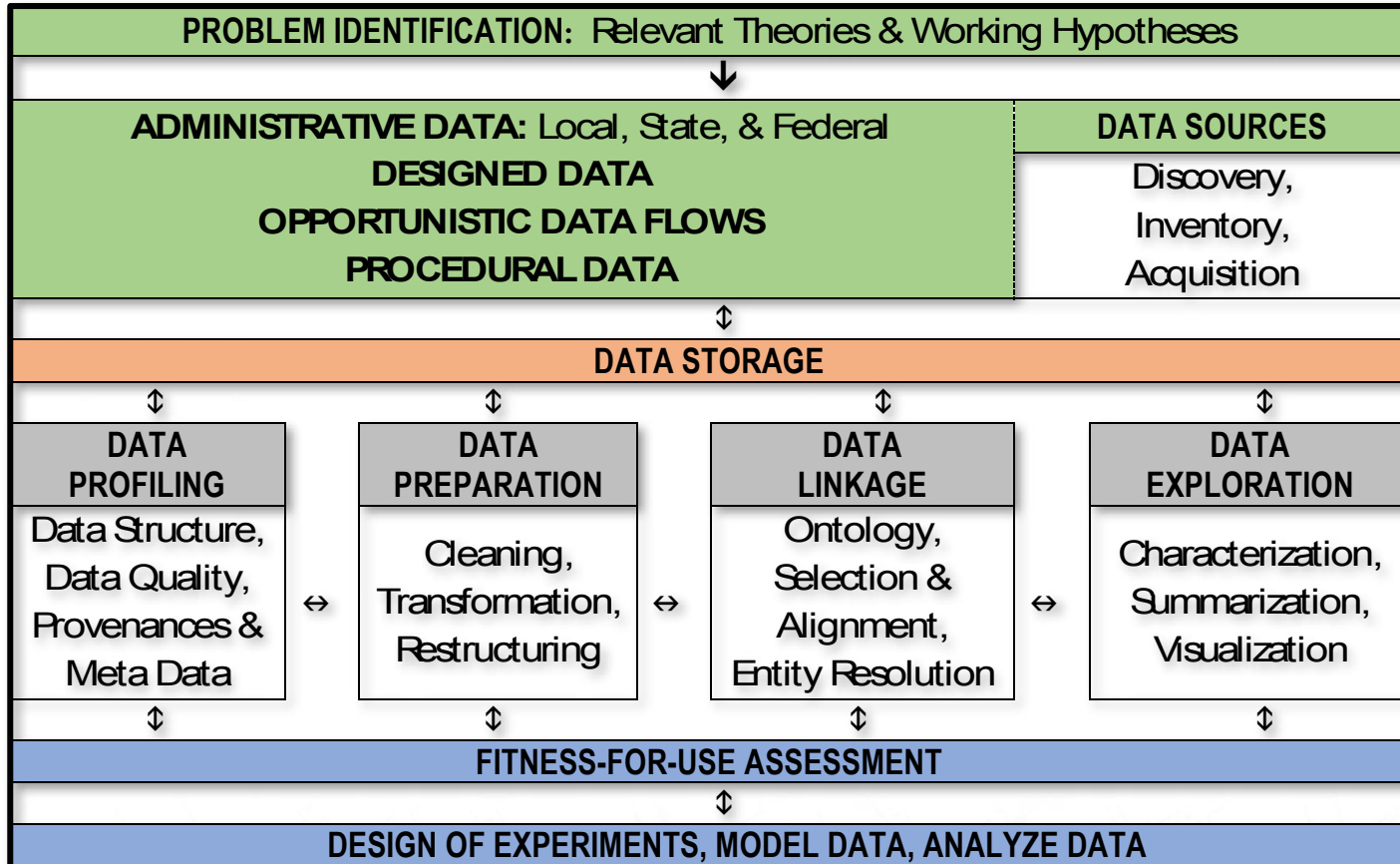


Procedural Data

This can be realized through a community



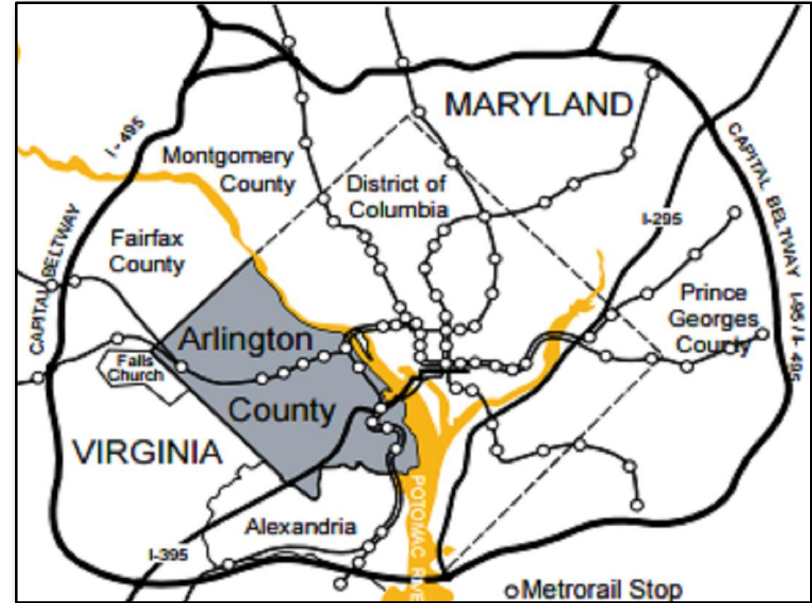
Data Framework



Partnership with Arlington County, VA - a successful demonstration



Improve quality of life and services while accelerating county's efficiency and resiliency



Develop situational awareness

Emergency Services Data

Internal Data

Flow of Events

External Data

CAD Call Summary

Unit Movement Summary

Incident Summary

NFIRS, Med Report

Incident Occurs 

Call Arrives @ E OC

Unit(s) Dispatched

Unit(s) Arrive

Incident Closed

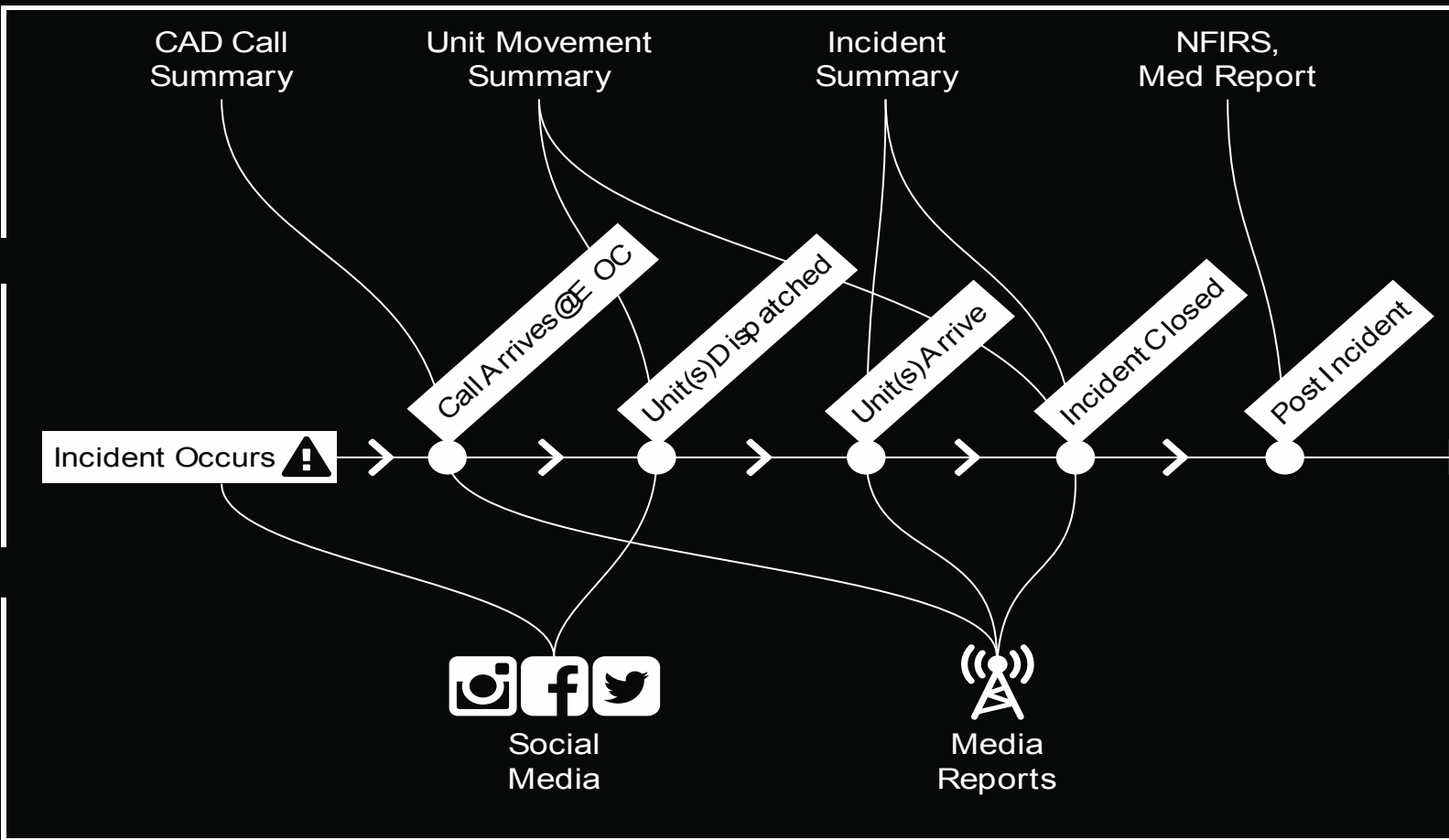
Post Incident



Social Media

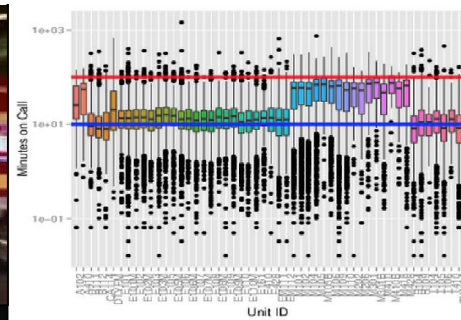
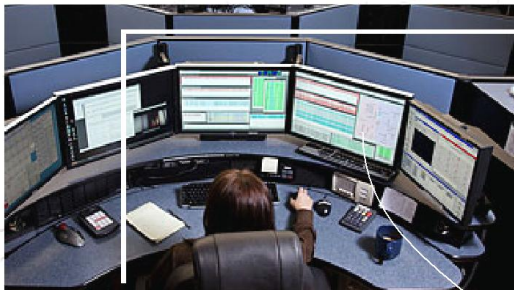


Media Reports



Flow of Events

External Data



NAME	STATUS
UNIT 1	ON CALL
UNIT 2	ON CALL
UNIT 3	ON CALL
UNIT 4	ON CALL
UNIT 5	ON CALL
UNIT 6	ON CALL
UNIT 7	ON CALL
UNIT 8	ON CALL
UNIT 9	ON CALL
UNIT 10	ON CALL
UNIT 11	ON CALL
UNIT 12	ON CALL
UNIT 13	ON CALL
UNIT 14	ON CALL
UNIT 15	ON CALL
UNIT 16	ON CALL
UNIT 17	ON CALL
UNIT 18	ON CALL
UNIT 19	ON CALL
UNIT 20	ON CALL

Incident Occurs

Call Arrives @ EOC

Unit(s) Dispatched

Unit(s) Arrive

Incident Closed

Post Incident

Just called 911 and held a woman's bloody hands after watching her get hit by a fast moving car. Crazy experience every day is a blessing

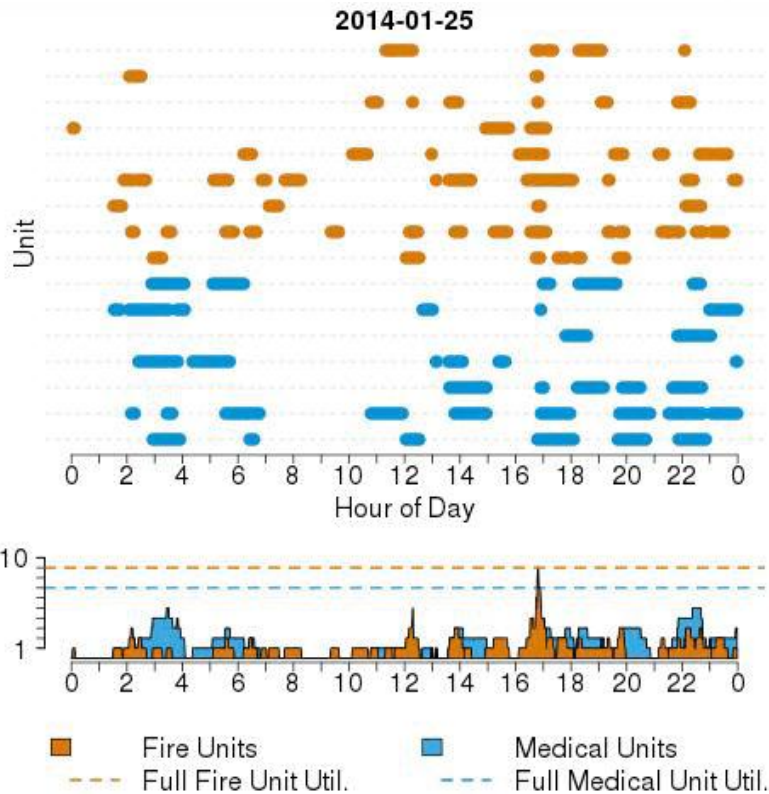
RETWEETS 1 FAVORITES 28

accidentally called 911 so i just set my neighbor Terrence's house on fire so i don't look dumb when the police arrive

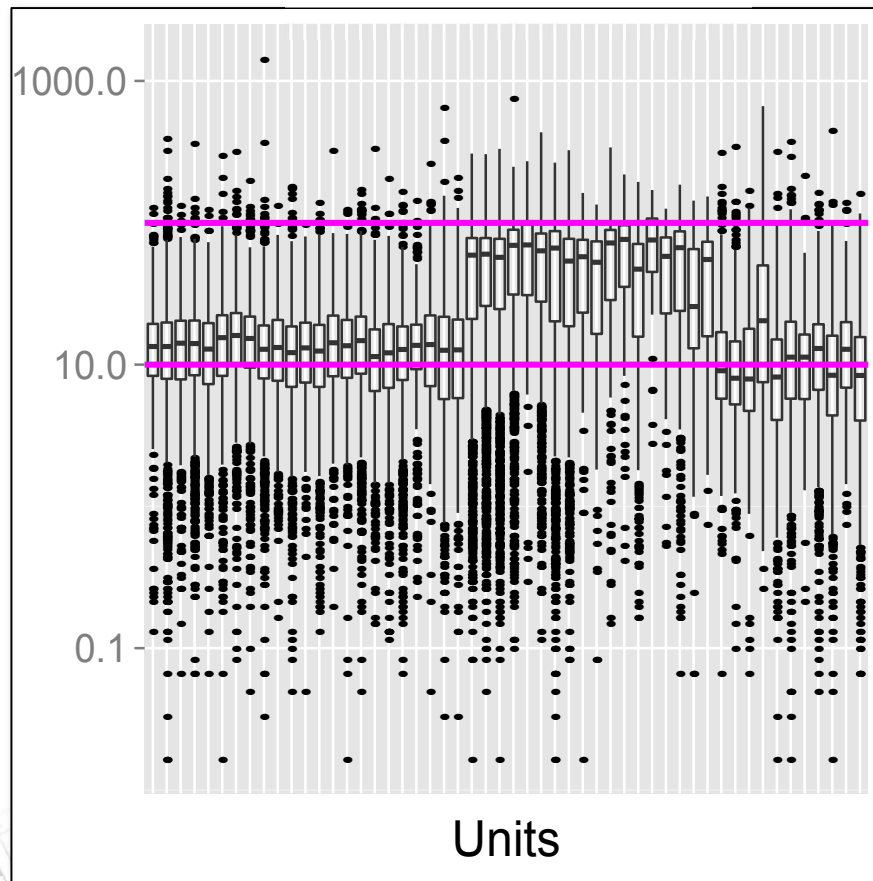
RETWEETS 82 FAVORITES 212

Unit Utilization

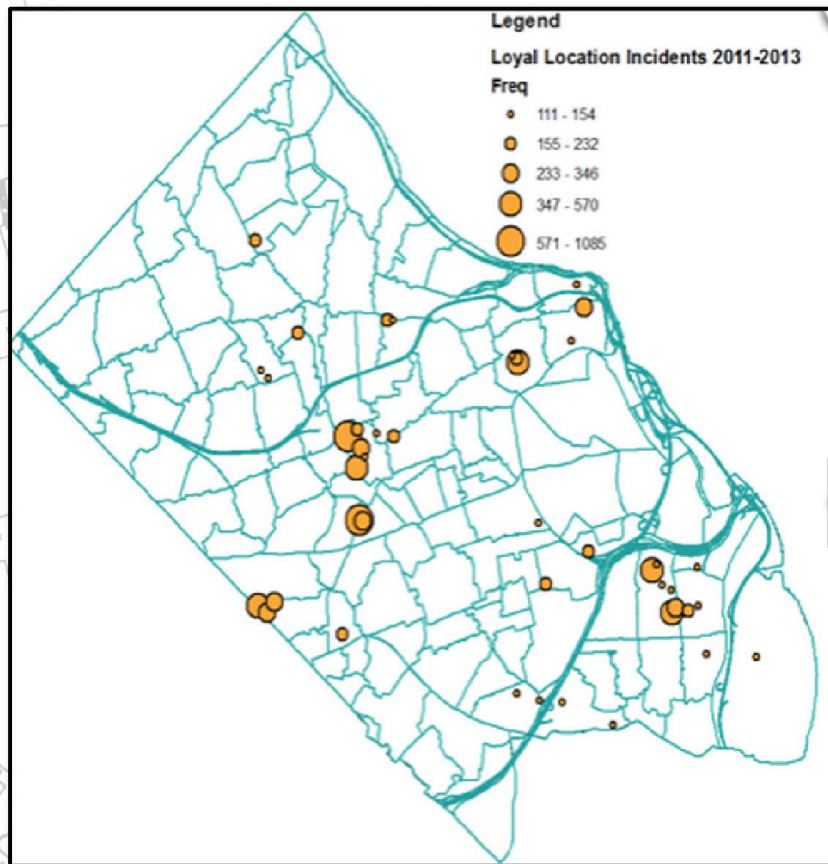
Utilization by Hour of the Day



Minutes on Call

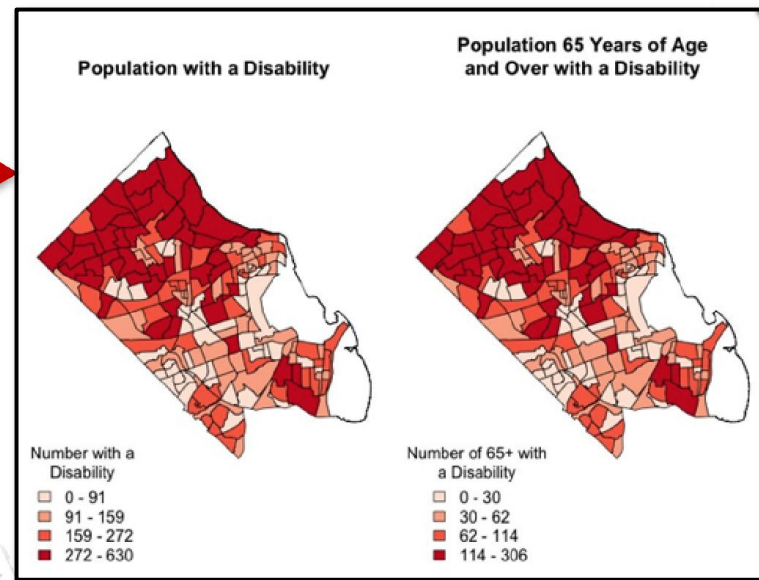


Loyal “Customers” or Loyal Locations



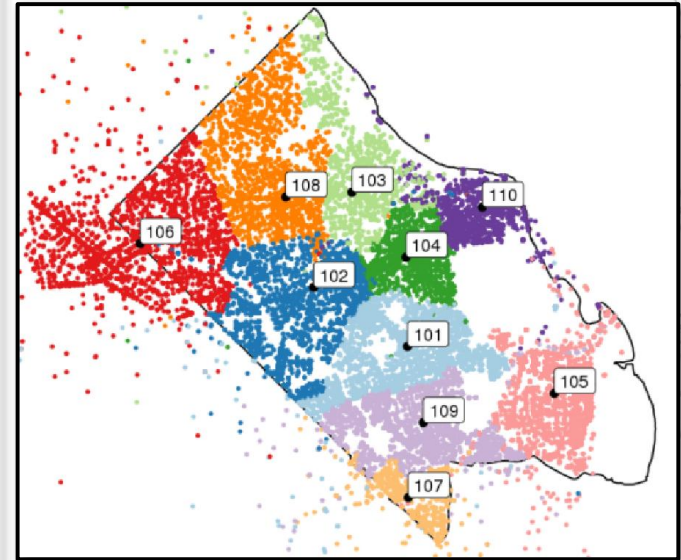
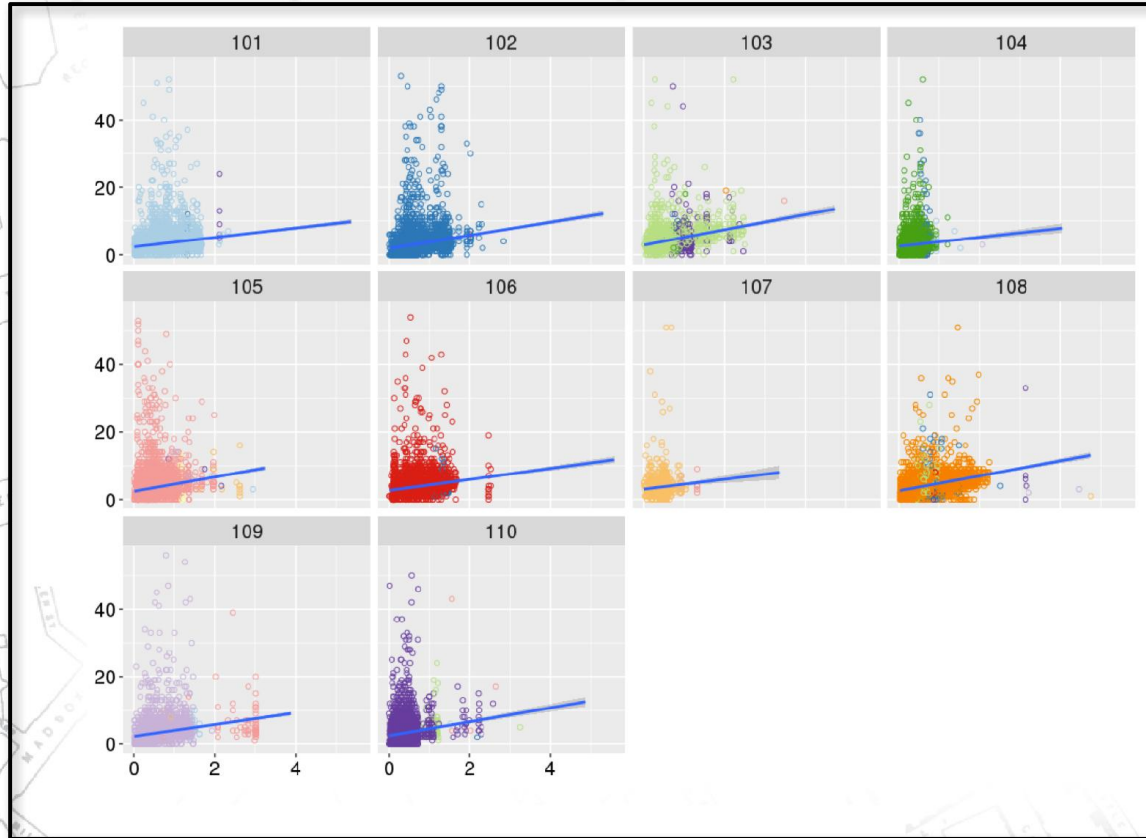
Mash up data with geography

- Identifies hotspots
- Deeper look into data and find these align with populations at risk in community

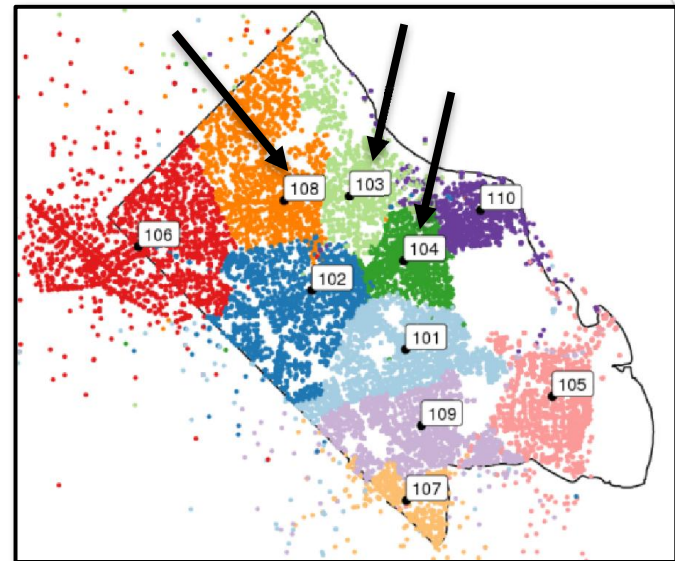
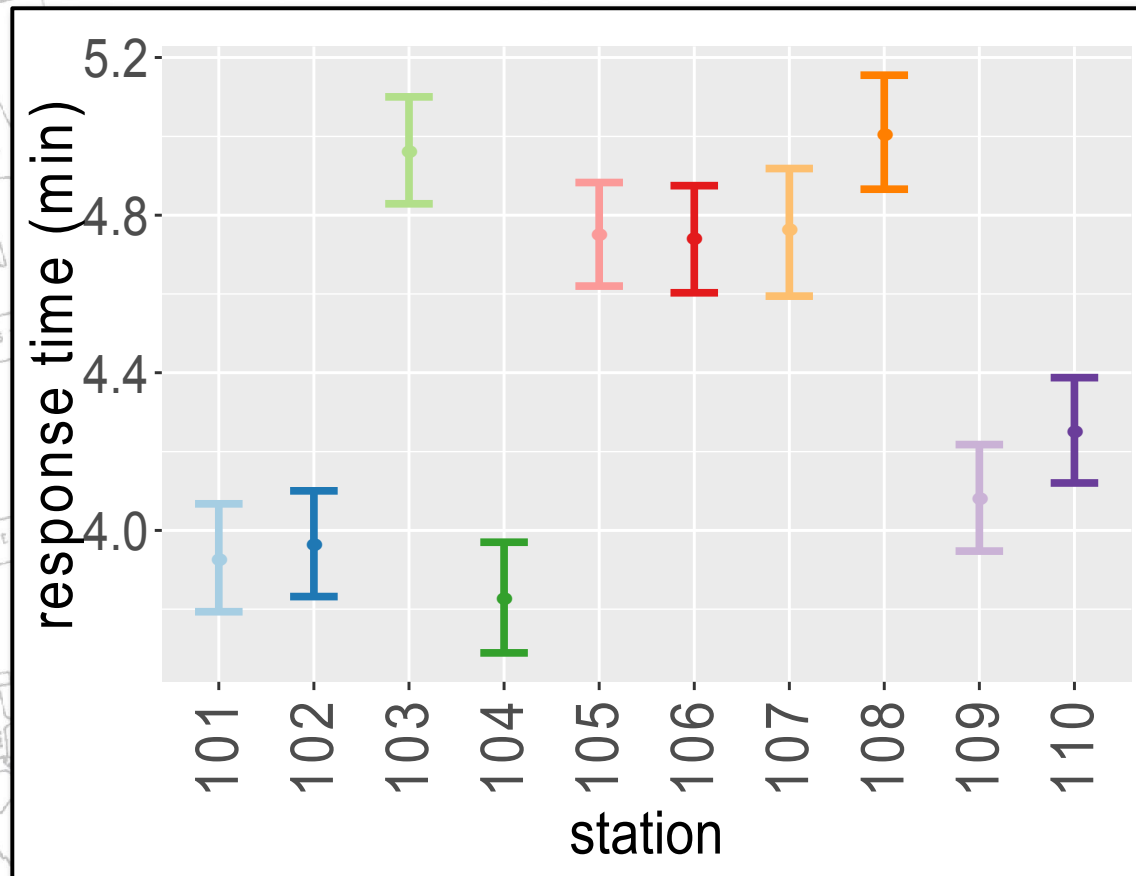


Emergency Response Time

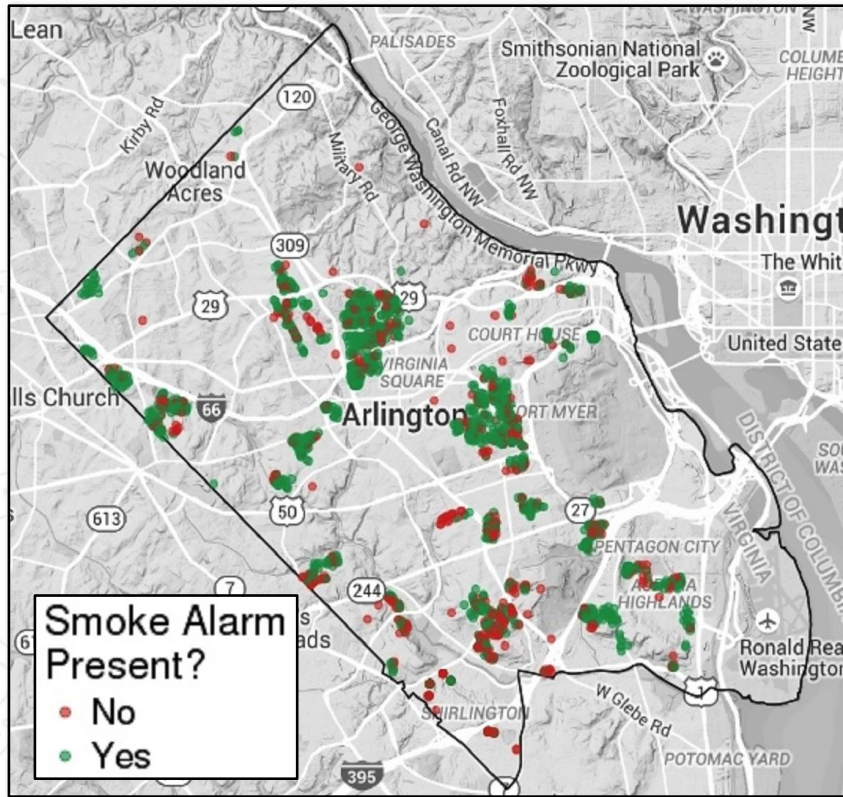
Response Time by Distance



Response Time by Fire Station



Arlington Operation FireSafe



Issue: The Fire Department wants to improve the efficiency of their Operation FireSafe program; out of 5,623 visits to single family homes only 1,799 smoke alarms were installed

Goal: Construct a model to estimate for each single family home the probability it has a smoke alarm

The DATA...

Household Level (Administrative Data): Operation FireSafe data on the location of the 5,623 single family homes visited, time and date of the visit and the outcome

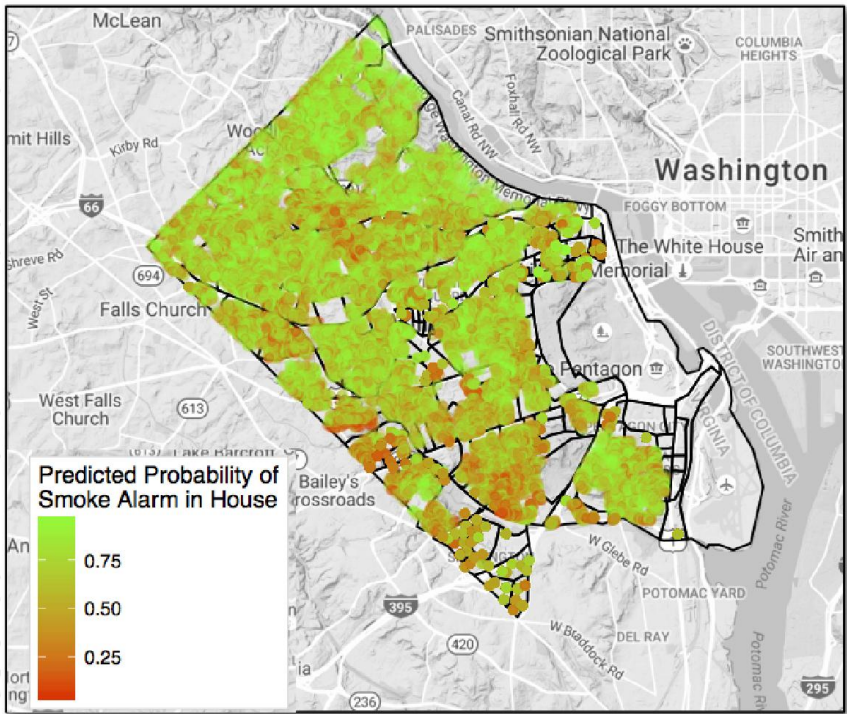
Household Level (Administrative Data): real estate tax assessments for 60,343 single family homes which includes tenure, home age, value, size and number of bedrooms

Household Level (Opportunity Data): geocoded the single family home locations

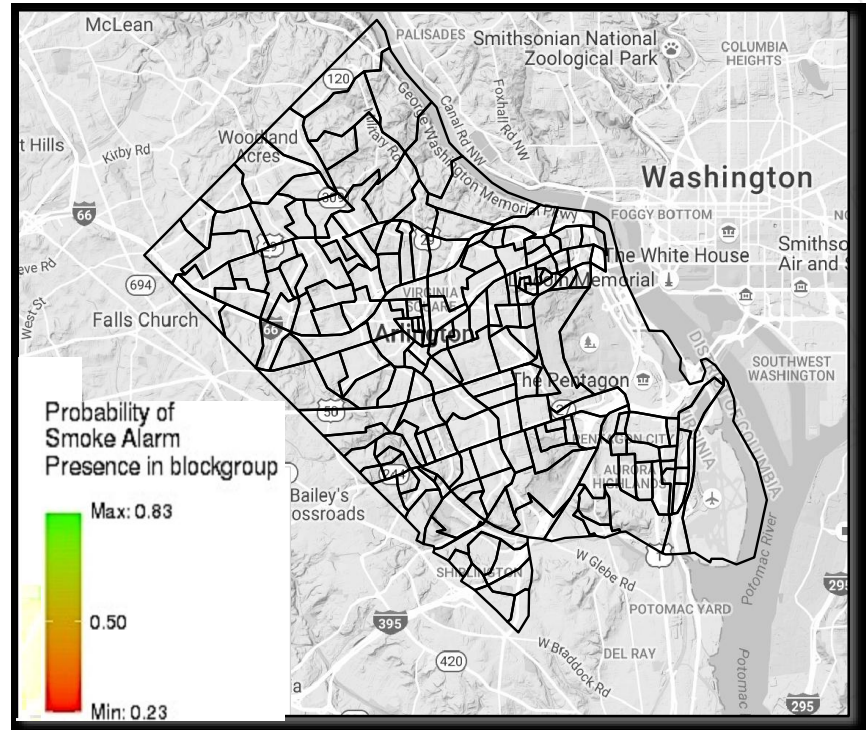
Census Tract Block Group Level (Designed Data): 5-year 2015 American Community Survey – household level demographic/socioeconomic data aggregated by block group

Probability of having a smoke detector

Bayesian logistic regression model with conditionally autoregressive spatial effects



Housing Unit Level Predictions



Census Block Level Predictions

Parks & Recreation Participation



Issue: The Department of Parks & Recreation (DPR) policy provides fee reductions to households that meet certain economic criteria

Goal: DPR wants to ensure that fees are not a barrier to program participation by targeting outreach

The DATA...

Household Level (Administrative Data): DPR 2016 enrollment data for all programs; BUT how a resident or household qualified for a fee reduction was not available?

Household Level (Opportunity Data): geocoded the household locations

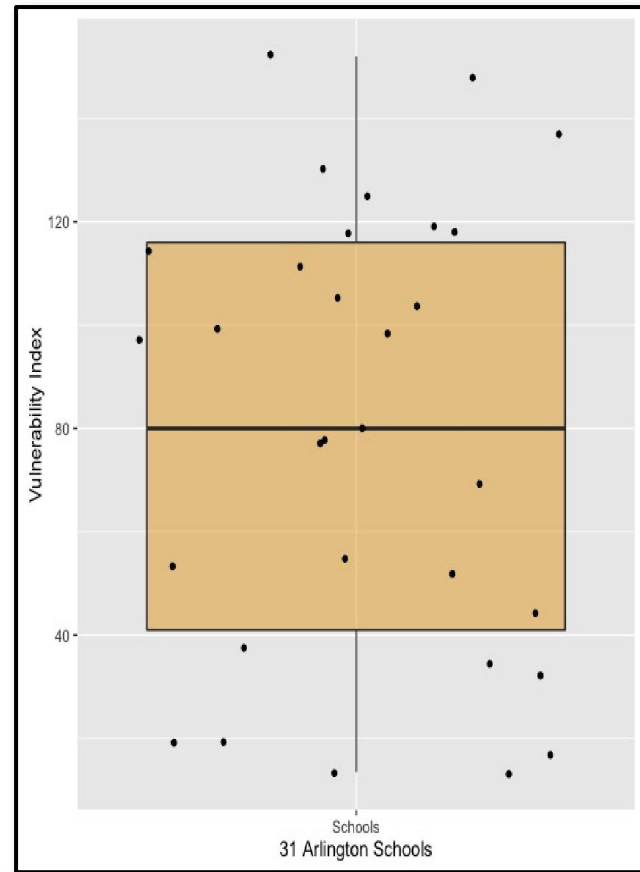
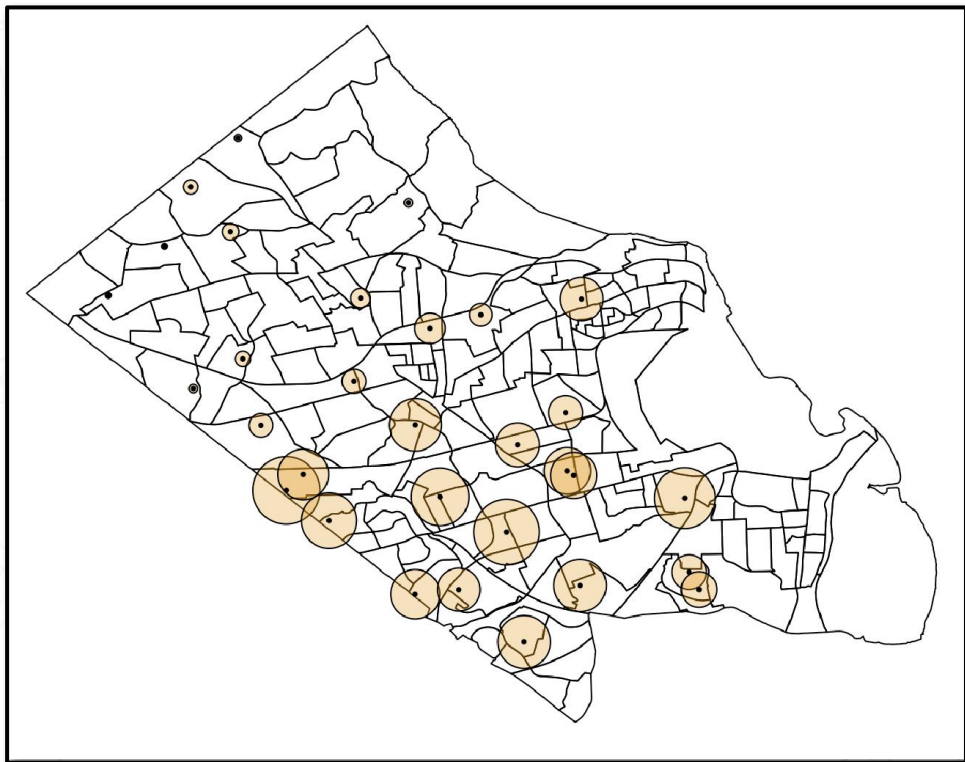
Census Tract Block Group Level (Designed Data): 5-year 2015 American Community Survey – household level measures of economic vulnerability aggregated by block group

School Level (Administrative Data): NCES CCD 2014-15 and the Virginia Dept. of Education – student level measures of economic vulnerability aggregated by school

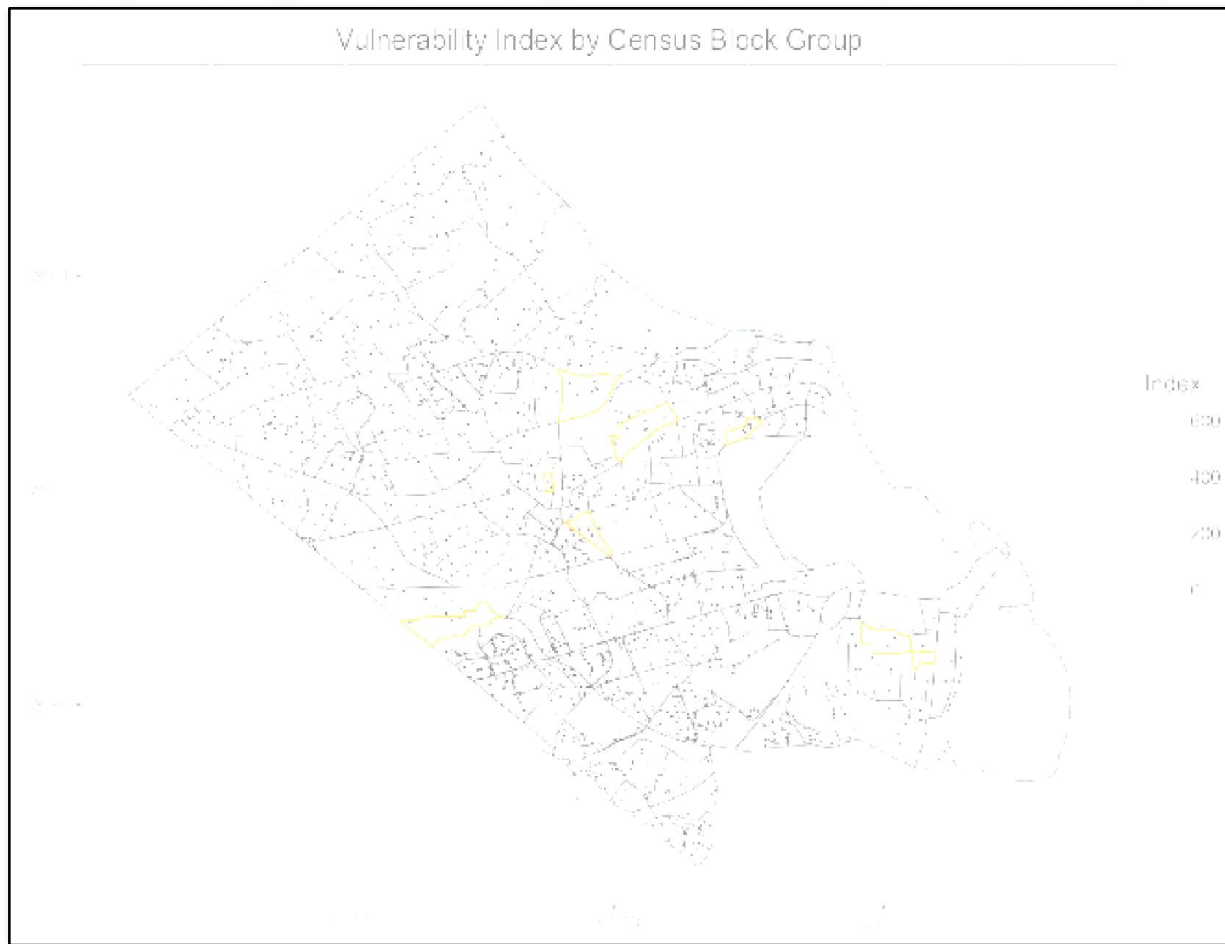
School Economic Vulnerability Index

Location of Arlington Schools

Symbol size is function of school index



Neighborhood economic vulnerability index



Democratization of Data

- Bring data in service of the public good to create resilient communities
- Build capacity for data-driven governance
- Provide the pathway to enable a community to become a ***data driven learning community***

