Big Data for Cities

Week 3

Curt Savoie
Connor McKay

Agenda

- Recap on last week
- Effective Data-Driven Policy Discussion
- More R Demo
- Q&A

Recap

- First R homework, good/bad/problems?
- Reading overview, questions?
- How ridiculous were the things I was saying?
- Other Questions on last week?

Effective Data-Driven Policy Overview

- What makes a policy data-driven?
- What makes such a policy effective?
- Examples!

Definitions of Data-Driven Policy

- “Determined by or dependent on the collection or analysis of data.”
- “Development of quantifiable measures to assess policy performance and draw comparisons across similar circumstances or peer groups so that “best practices” can be identified and expanded” - (https://www.americanprogress.org/wp-content/uploads/issues/2007/04/pdf/data_driven_policy_report.pdf)

Keys To Effective Data-Driven Policy

- Clearly Communicated Methodology
  - Rigorous statistics
  - Reproducible Results
- Data Availability
- Engaged Stakeholders
- Small / Targeted Scope
- Achievable Recommendations
- Measurable Effects and definitions of success

Types of Data-Driven Policy

- Pay For Success / Outcomes based contracting
- Performance Management
- Predictive Models (Operations)
- “Negative Feedback” enforcement models
    - Boston “Problem Properties” Task Force
- “Positive Feedback” incentive models
    - Early Childhood Education incentives for more educated teachers

Data-Driven Policy Example #1

- Commonwealth of Massachusetts - Chapter 55 Opioid Overdose Study
  - Problem Statement: there is an epidemic of opioid overdoses in Massachusetts and little understanding of the trends.
  - Initially a 1 year project bringing together 7 datasets in order to present a report to the MA Legislature
  - Key findings that resulted in policy
    - 18-24 y.o. Released prisoners were 10X more likely to overdose
    - Most addictions stem from a legal prescription
    - People with multiple health providers are at increased risk

Data-Driven Policy Example #2

- Chicago Rodent Predictor
  - Problem Statement: rodents represent a public health hazard and can be difficult to track and monitor
  - Collaboration with outside partner, Carnegie Mellon with 311 data
  - Key Findings:
    - a 311 call or online request related to garbage produces a 7-day window in which an increased number of rodent calls will occur in the same area. Thus, rates of garbage-categorized 311 calls serve as a measurable indicator whose direction can signal changes in rat trends
    - Better precision operations for rat-baiting (yeah, that's a thing)

Data-Driven Policy Example #3

- NYC Risk-based Fire Inspections
  - Problem Statement: unsafe building conditions have higher risk for fires that cause serious injury and/or deaths. NYFD lacks the resources to inspect all buildings as timely as needed.
  - Used data from building permits, inspections and assessing, along with expert insights
  - Key findings:
    - Lack of legal activity with building conditions was a sign of illegal conversion
    - Age and type of building were variables
    - Was able to prioritize inspections to reduce fires

For Next Week

- **Reading on theory and practice**
  - http://www.mass.gov/chapter55/
  - http://donellameadows.org/archives/leverage-points-places-to-intervene-in-a-system/
  - http://datasmart.ash.harvard.edu/news/article/using-predictive-analytics-to-combat-rodents-in-chicago-271
  - http://www1.nyc.gov/office-of-the-mayor/news/163-13/mayor-bloomberg-fire-commissioner-cassano-new-risk-based-fire-inspections-citywide#/1
  - http://www.thirdsectorcap.org/portfolio/massachusetts-juvenile-justice-pfs-initiative/
  - https://www.aisp.upenn.edu/aisp-network/aisp-network-site-case-studies/
- In R
  - homework!