Evaluation, data science, and the causal revolution

January 15, 2020

PMAF 8521: Program Evaluation for Public Service
Andrew Young School of Policy Studies · Georgia State University
Spring 2020
Plan for today

Data science and public service

Evidence, evaluation, and causation

Class details

Getting started!
Data science and public service
Policymakers at all levels of government are struggling to thoughtfully harness data in the service of public values. Many public servants grew up in an era of firmly separate disciplines: You were either an engineer or an economist, either a programmer or a social worker, but never both. In an era in which data is everything, the risks to core democratic principles—equity, fairness, support for the most vulnerable, delivery of effective government services—caused by technological illiteracy in policymakers, and policy illiteracy in computer scientists, are staggering.

...
Data and government

“To responsibly unleash the power of data to benefit all Americans”
FACT SHEET: Launching the Data-Driven Justice Initiative: Disrupting the Cycle of Incarceration

"[O]ur criminal justice system isn’t as smart as it should be. It’s not keeping us as safe as it should be. It is not as fair as it should be. Mass incarceration makes our country worse off, and we need to do something about it."

- President Barack Obama, July 14, 2015
## U.S. City Open Data Census

**Powered By Open Data Census**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cincinnati, OH</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>94%</td>
</tr>
<tr>
<td>2</td>
<td>Las Vegas, NV</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>89%</td>
</tr>
<tr>
<td>3</td>
<td>San Francisco, CA</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>85%</td>
</tr>
<tr>
<td>4</td>
<td>Virginia Beach, VA</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>79%</td>
</tr>
<tr>
<td>5</td>
<td>Baton Rouge/E.B.R.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>78%</td>
</tr>
</tbody>
</table>
## Breakdown

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Breakdown</th>
<th>Year</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budget</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business Listings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code Violations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction Permits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crime Reports</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Calls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee Salaries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lobbyist Activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parcels</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Police Use-of-Force</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procurement Contracts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Property Assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Property Transfers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public Facilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restaurant Inspections</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Service Requests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spending</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How do you use all this data to make the world better?
What is “statistics”?

Collecting and analyzing data from a representative sample in order to make inferences about a whole population.
What is “data science”?

- Big data
- Machine learning
- Artificial intelligence
- Data mining
- Cloud computing
- Neural networks
- Algorithms
- Artificial intelligence
- PR-speak for “statistics”
What is “data science”?

Turning raw data into understanding, insight, and knowledge

Collect
Analyze
Communicate
What’s the difference?

What is “program evaluation”?

Measuring the effect of social programs on society

Data and statistics  Communication  Causal inference (econometrics)
Evidence, evaluation, and causation
What is the relationship between social science research and public policy & administration?
Evidence-based medicine
Modern evidence-based medicine

Apply evidence to clinical treatment decisions

Move away from clinical judgment and “craft knowledge”

Is this good?
Can we find and measure evidence for policies and programs?
Evidence-based policy

- RAND health insurance study
- Oregon Medicaid expansion
- HUD’s Moving to Opportunity
- Tennessee STAR
Policy evidence industry

Jameel Poverty Action Lab (J-PAL)
Campbell Collaboration
Should we have evidence for every policy or program?

No!

Science vs. art/craft/intuition
IF U DONT SMOKE, U ALREADY BELIEVE IN CAUSAL INFERENCE WITHOUT RANDOMIZED TRIALS

(#HistorianSignBunny #Epidemiology)

10:13 PM - 12 Jul 2018

200 Retweets 612 Likes
Where does program evaluation fit with all this?

It’s a method for collecting evidence for policies and programs.
Types of evaluation

- Needs assessment
- Design and theory assessment
- Process evaluation and monitoring
- Impact evaluation
- Efficiency evaluation (CBA)
Theories of change

Three phases of truancy intervention

No truancy

Increased commitment to school

Better grades

Reduced risk factors

Impact evaluation!
Theory → impact

Grades

Outcome change

Pre-program grades

Post-program grades

Program activities

Program outcomes

Before Program

During Program

After Program

Grades with program

Program effect

Grades without program

Outcome change

Program activities

Program outcomes
Weeks before/after truancy intervention

Average number of absences

-10 -5 0 5

Lines
- Actual
- Predicted

Colors
- 80% Confidence
- 95% Confidence
- Truancy intervention
Godwin's law

From Wikipedia, the free encyclopedia

Godwin's law (or Godwin's rule of Hitler analogies)\(^1\)\(^2\) is an Internet adage asserting that "As an online discussion grows longer, the probability of a comparison involving Nazis or Hitler approaches 1";\(^2\)\(^3\) that is, if an online discussion (regardless of topic or scope) goes on long enough, sooner or later someone will compare someone or something to Adolf Hitler or his deeds, the point at which effectively the discussion or thread often ends. Promulgated by the American attorney and author Mike Godwin in 1990,\(^2\) Godwin's law
Godwin’s Law for statistics

Correlation does not imply causation

Except when it does

Even if it doesn’t, this phrase is useless and kills discussion
Not everyone found the news believable. “Facepalm. Correlation doesn’t imply causation,” wrote one unhappy Internet user. “That’s pretty much how I read this too... correlation is NOT causation,” agreed a Huffington Post superuser, seemingly distraught. “I was surprised not to find a discussion of correlation vs. causation,” cried someone at Hacker News. “Correlation does not mean causation,” a reader moaned at Slashdot. “There are so many variables here that it isn’t funny.”
Correlation vs. causation

How do we figure out correlation? Math and statistics

How do we figure out causation? Philosophy. No math.
John B. Holbein @JohnHolbein1 · Apr 7
Causality isn't achieved; it's approached.

Show this thread

John B. Holbein @JohnHolbein1 · Apr 7
Causality isn't binary; it's a continuum.

Show this thread
How do we know if X causes Y?

X causes Y if...

...we intervene and change X without changing anything else...

...and Y changes
Y “listens to” X

X isn’t the only thing that causes Y

A light switch causes a light to go on, but not if bulb is burned out (no Y despite X) or if the light was already on (Y without X)
Causal relationships?

- Lighting fireworks causes noise
- Rooster crows are followed by sunrise
- Getting an MPA increases your earnings
- Colds go away a few days after you take vitamin C
Causation

Causation = Correlation + time order + all other factors ruled out

How do you know if you have it right?

You need a philosophical model

That’s what this class is for!
The causal revolution

JUDEA PEARL
WINNER OF THE TURING AWARD
AND DANA MACKENZIE

THE BOOK OF WHY

THE NEW SCIENCE OF CAUSE AND EFFECT

α β
Causal diagrams

Directed acyclic graphs (DAGs)

Graphical model of the process that generates the data
Maps your philosophical model
Fancy math ("do-calculus") tells you what to control for to find causation
Minimal sufficient adjustment sets containing Age, Mental health, Risky behaviour, SES, Stress, Unplanned pregnancy for estimating the total effect of Smoking on FASD classification:

- Age, Mental health, Risky behaviour, SES, Stress, Unplanned pregnancy
Want to live longer? Try going to the opera. Researchers in Britain have found that people who reported going to a museum or concert even once a year lived longer than those who didn’t.

Another Benefit to Going to Museums? You May Live Longer
Researchers in Britain found that people who go to museums, the theater and the opera were less likely to die in the study period than those who didn’t. nyti.ms/2Q9AmZV
Break

Set up an RStudio.cloud account if you haven’t

Go to https://andhs.co/rstudio to join the class workspace
Ask me anything!
Class details
Evaluation and causation
- Program theories
- Logic models
- Measurement
- DAGs
- Potential outcomes

Tools and methods
- Randomization
- Matching
- Difference-in-differences
- Regression discontinuity
- Instrumental variables

Applied evaluation
- Preregistration
- Ethics
- Communication
- Other evaluations

R and the tidyverse
- Data manipulation
- Modeling
- R Markdown
- Visualization

Program Evaluation for Public Service
Stage 1: Regress each column of $X$ on $Z$, $(\hat{X} = Z\hat{\delta} + \text{errors})$:

$$\hat{\delta} = (Z^TZ)^{-1}Z^TX,$$

and save the predicted values:

$$\hat{X} = Z\hat{\delta} = Z(Z^TZ)^{-1}Z^TX = P_ZX.$$

Stage 2: Regress $Y$ on the predicted values from the first stage:

$$Y = \hat{X}\beta + \text{noise},$$

which gives

$$\beta_{2SLS} = (X^TP_ZX)^{-1}X^TP_ZY.$$

\[\text{model}_\text{2sls} \leftarrow \text{iv\_robust} (\text{health} \sim \text{bed\_net} \mid \text{treatment}, \text{data} = \text{bed\_nets})\]
Class technology

R Studio
The tidyverse
strike_damages_month <- bird_strikes %>%
group_by(Month) %>%
summarize(total_damages = sum(Cost, na.rm = TRUE),
average_damages = mean(Cost, na.rm = TRUE))

ggplot(data = strike_damages_month,
       mapping = aes(x = Month, y = total_damages)) +
geom_col() +
scale_y_continuous(labels = dollar) +
labs(x = "Month",
y = "Total damages",
title = "Really expensive collisions happen in the fall?",
subtitle = "Don't fly in August or October?",
source = "Source: FAA Wildlife Strike Database")
There is no way to go from knowing nothing about a subject to knowing something about a subject without going through a period of much frustration and suckiness.

Push through. You'll suck less.

Hadley Wickham, author of ggplot2 and the tidyverse
Dude, sucking at something is the first step towards being sort of good at something.

(It’s Great to) Suck at Something

By Karen Rinaldi

April 28, 2017
Am I making you computer scientists?

No!

You don’t need to be a mechanic to drive a car safely

You don’t need to be a computer scientist or developer to use R safely
My #rstats learning path:

1. Install R
2. Install RStudio
3. Google "How do I [THING I WANT TO DO] in R?"

Repeat step 3 ad infinitum.

7:19 AM - 18 Aug 2017
You can do this.
Goals for the class

- Become an expert with R
- Speak and do causation
- Design rigorous evaluations
- Change the world with data
Prerequisites

Math skills
Basic algebra

Computer science skills
None

Statistical skills
Regression and differences in means
(ideally; you can survive without it, though)
Miscellanea
Class expectations

- Late work
- Technology
- Participation
- Other?
Getting staRted!