<http://ire.org/events-and-training/event/2702/3021/>

Andy Lehren, The New York Times

Sharad Goel, assistant professor at Stanford University of Engineering. Mathematician.

Josh Hinkle, KXAN’s Executive Producer of In-depth & Investigative Content.

Cheryl W. Thompson, an IRE board member, writes investigative stories for The Washington Post and teaches investigative reporting at GW University

-----------------------------

Andy:

**The Disproportionate Risks of Driving While Black**

Data gathered: traffic stop data- are blacks stopped more than whites? Hard to say b/c don’t know race of everyone on the road. So instead, focus on collateral damage. You can get stopped and get a ticket. This is where the officer must make a decision. Types of searches- probable cause, consent searches (4th amend. Searches, officer just says “I’m searching your car” and it’s up to the driver to dispute). So instead of stoppage rates, look at who got probable cause searches.

In Greensboro, NC- officers did probable cause searches on blacks more than whites. Response was officers were looking for guns and drugs. But guns and drugs were not found in cars of white and black drivers at the same rates. In fact, searches of whites were more likely to turn up guns and drugs.

Then compared to other cities; nearby cities.

FOIA’d: every single police stop involving any single kind of traffic ticket. Wanted race, reason of the stop, duration of the stop, whether search was conducted and what kind. Some states better than others and diff agencies in diff states. Look at state police agencies. Got almost everything as a database.

Mapped crime and mapped ticketing to see if ticketing was different in low crime vs high areas and saw rates that ppl got pulled over and searched.

Look at ticketing patterns:

“High discretionary stops”- non-moving violations, not speeding tickets. These are stops done by police for minimal stuff-- shows cops might have been looking for a reason to pull them over. (tail light, seat belt, license plate not visible enough, etc).

Examined innocent bystanders shot by NYPD; almost always in minority areas. Then the city doesn’t want to pay for medical expenses.

**No Money, No Mercy: After a crime, the price of a second change**

Looked at who got Pretrial Diversion funding, which allows you to sign a slip of paper that says you admit you committed the crime and you will do some sort of classes, community service, make a donation, etc.

This didn’t work that way. Usually only people with money could afford this program. Prosecutors had almost total control of who gets this funding. Local gov really was just making money off this program.

There was no data on this. NYT had to build their own database.

Places to find information:

* Jstor
* University of Michigan has repository ICPSR
* Bureau of Justice Stats.
* John Jay university
* State statistical analysis centers
* National conference of state legislatures
* Bar associations
* Beware of purpose of data collectors.

Learn the math: learn formulas that matter. With ticketing, it’s the hit rate. Learn about how recidivism rates are calculated; it’s diff state to state and even on a county level.

Build your own database. The documents exist somewhere, just maybe not in the way you want them.

------------------

Sharad:

Algorithmic Fairness

Discrimination vs. disparate impact (stopping blacks at a higher rate to find drugs and guns is disparate impact. Stopping blacks at a higher rate when stops of whites yield more drugs and guns is discrimination)

Machine Bias by ProPublica

Are algorithms like COMPAS fair? - little evidence of bias (hard to prove) but are they fair? (hard to prove as well)

Algorithms in criminal justice: estimate defendant’s likelihood of committing a violent offence based on other info (not race). Then, set a detection threshold (i.e. detain defendants deemed more than 20 percent likely to commit a violent offense by that model).

**Black defendants are detailed much more often than whites, but the algorithm does not use race. Is this discrimination? Disparate impact? Why is this the case?**

* Black defendants on average have more serious criminal histories (according to measures the algorithm uses. A statement about the correlation of the algorithm’s factors and race). Defendants with more serious criminal histories are more likely to reoffend and thus more likely to be detained. So, the algorithm is working how it’s designed to….but there is a disparate impact.
* Error rates are also higher for blacks than whites. (detention rates differing)

Reported crime is a proxy for actual crime.

**What should we do?**

* Stop using algorithms - but judges are trying to do the same thing as algorithms but are typically worse at it.
* Set race-based detection thresholds. (detain whites who are at least 20% less likely to commit a violent crime and black who are at least 30% less likely or some other measure)

Cut short b/c propublica reporter Jeff said he was misrepresenting the story.

--------

Josh:

Racial Profiling Whitewash (Texas)

Lots of Dept of Public Safety records had ppl with traditionally spanish last names marked as white.

Tracked down drivers through citations issued- actual citation to avoid data entry errors. Found online court and jail records for mug shots, dash cam footage.

Used MySQL to filter the 16 million records. Took hours to load data and make queries.

Now a state law to ask drivers to provide their race rather than marking it on their own. Austin police was doing the same thing.

State and city traffic stop data

Texas is one of 33 states with racial profiling laws with reporting requirements.

City data can come from municipal courts.

Request actual citations to verify the data.

In NC you have to go to the court to request and look at dash cam video

-----------