Jeff Interview

Tue, Apr 25, 2023 2:17PM • 15:36

**SUMMARY KEYWORDS**

data, estimation, estimates, fbi, crime, reported, agencies, county, quarterly estimates, records management system, inaccurate, collect, ucr, neighbors, new orleans, share, datasets, obvious, evaluate, released

00:00

estimation has been 21. And we, I will share some of our findings, when I asked you a question later, but I'm just want to start from when did you start noticing FBIs missing crime data and the issues within their estimations?

00:18

I mean, it's been obvious they they announced in 2015, that they were going to make the transition the neighbors. And so it's been obvious from that point that they were well short. I mean, they had 30 or 40% of agencies were reporting via nighters in 2015. So it hasn't been something where they've gotten close. It's it's been a slow ramp up, it's been an obvious ramp up.

00:44

Yeah. So have you noticed any improvements throughout this years from 2015? To now?

00:54

Improvements? Yeah, they've gone from under 50%. In 2020, I believe to I think they're almost 70% of the US population is covered. So there's been a substantial jump in

01:10

coverage. It's just It used to be 95 to 97%, was covered by an agency that was with fear. So they're still well short of where you'd want to be in order to be able to make assessments.

01:23

Yeah, exactly. And we, like doing our research, we found that the FBI did not make it very clear about his methodology, when we're coming up with his estimation. So I don't know how much do you know about their methodology and techniques to make estimations? And how would you evaluate them? I don't really know, under the hood, exactly what they're doing to make their estimations. Um, that's, that's coming from BJs, not FBI. So they, you know, they've got a bunch of statisticians and they do their best. But then you get examples like New York, where they estimated data from really complete data sources. And I think that they, they probably felt a little bit of pressure to do estimates where they otherwise may not have. And so I think that that's why you get such a funky and messed up estimate for New York State.

02:22

And then I started places where they just there was nowhere near enough data to even pretend to do and so

02:28

I see, um, and I was looking at your profile and find you was served as the analyst for the New Orleans Police Department. So I'm wondering, could you just share some of your experience working on the crime data in Louisiana? And what kind of challenges have you encountered when collecting and reporting crime data to the FBI? I mean, it's the same in Louisiana, you get some places that do it some places that don't New Orleans, where I was not doing it, not, not doing either. So it makes it very difficult to evaluate current data and think about it in a way that makes sense, because we just don't have access to good crime data. And for New Orleans, it's a technical problem, they don't have the records management system, and it's not clear when they're going to get it. So it's not something that sort of has an easy solution either.

03:22

Um, but how they started working on it, like how, how they approach this type technical problems, like, maybe to buy a new RMS and they haven't yet. So any plan to buy it? Yes. But it's, it's been for four years at least have plans to buy it. And they don't seem to be any closer. So I'm not going to hold my breath on that one. I see. Is that something like a lot of states share? Or that's something you only observe in Louisiana? Now, Fair enough, fair enough, just because nybridge is so in terms of what data to collect. And so if your records management system can't do that, then they need to go out and and find a new one. And there's, they can be expensive, they can be challenging.

04:14

And if you're not feeling intense pressure to do something, then you may not do it.

04:20

I see. And how about a staff shortage like that's collecting and reporting the neighbors data needs a lot of people on it and

04:31

does

04:33

like Louisiana has like shortage shortage and stuff to put people into on this job.

04:40

So some places do some don't and Orleans does. I don't know that it's necessarily a staffing issue for inputting data, but you have to have the technical know how, what exactly you need and then you need to train your staff to do it. So it's not necessarily a numbers game, but it is there's a component where you have to be because there's

05:00

So many more things that you have to be putting into each report to fill out members correctly, you have the people that are doing that the officers on the ground have to be able to do that. And and have the time to do that.

05:13

I see. Um, and I know like the the FBI estimation for 2022 just released recently. I don't know, have you had a chance to look at them?

05:27

The what their estimates for 2022? Yes. The

05:33

mean, the quarterly estimates? No, I think for for just the whole estimates for last year. No, they haven't done that. Yet. They, they released those in September, October. They've been doing quarterly estimates. But

05:54

but the quarterly estimates are have been flawed. And they don't have enough agencies, big agencies reporting to do them. So they're just, instead of doing actual quarterly estimates, they're just releasing data for a bunch of cities that's of dubious quality. I see. So why why did you Why do you say it's flawed, like in the same way as their estimates for the past years? Or is there any other problems? Well, so they just published the data, they don't audit it for this quarterly data. And so you get things like the first quarter of 2022, Chicago reported one murder,

06:32

an obvious data submission error, and then included in the target on the last quarter. So you get things and if you compare the data online to the data being reported quarterly to the FBI, there's probably a dozen, two dozen cities where there's obviously the the number is is nowhere near where it should be.

06:55

I see and have, you know, noticed any impacts or consequences of their inaccurate estimates? Because we were trying to find is there any politician or media has been using? FBI is inaccurate estimation? I don't know if do you have a few examples of like, the data? I don't specifically have examples. I think.

07:20

People use crime data to make their point. And so I don't know that the like accuracy of it. It usually matters to the politicians. I don't have any specific examples of politicians, citing inaccurate fibers data. Do you think the neighbor's data is using widely or decided widely?

07:40

I'm not really, I think that it because it's so complicated. It's hard for people like, I barely understand how to use divers. And I like to think that I have a certain level of expertise in it.

07:55

It's really difficult. There's so much data there and so much to go through that lay users are not able to use it in a way that they could other datasets or previous datasets. So it's good for researchers, it's not necessarily good for the public at large.

08:12

I see. Um,

08:15

yeah. And

08:17

so can I share my screen just to show you one of the data that we call this rule

08:25

before?

08:45

So right here, we merged

08:48

actual crime data with

08:51

ask the FBI estimation for New York State. And we found like there are four types of crime that the FBI did not give the estimation and I don't know if you have a chance to know, like, the reason why behind us and why why the FBI give estimation for certain types of crime that have like less counts, for example, rape, they only have like 6000s but like for for robbery and aggravated assault. They have like 15,000. But but the FBI didn't give estimations. Do you know, the reasons behind? I don't? They may not have gotten about submission that I'm not sure why. Okay, I see. No problem. Yeah. And, Travis, do want to ask your questions. I know what I would feel right. Yeah. Hi.

09:44

My name is Travis. And I'm curious about your chance that you can't get to know the accuracy level or bias of these estimations.

09:53

No, not really. I don't I mean, you'd have to ask BJs then I'd be willing to take you into the math of it.

10:00

But I don't

10:03

publish the code or anything like that.

10:06

Okay, I see. And or is there any way to compare the estimation amount of different states or counties? Because we know they are some factors that lead to the bias? So can we just collect these factors and use these factors to like evaluate the estimation?

10:25

Not that I know of, I'm not aware of any way to evaluate the estimations in any way. So it's kind of hard to say with much confidence what

10:37

when he 20 is, I see. And I'm also curious about the estimation procedure on library data. So I read an article about it in the sense that the non participation rate for nybridge is higher, because it includes not only those agencies who did not submit any crime data, but also those SRS agencies whose data cannot be converted to neighbors. So do you think they have done something on any of these two aspects to try to improve the rate in the past two years? Yeah, I mean, they've done they've done a lot to improve the rate, it's, it's gone from, like I said, almost 50 to 55%, to almost 70%, which is great, but it's still well, sort of where it needs to be. So until the the number is a 588 90%, somewhere in there, we're not going to have really close estimates. And so that makes the estimates kind of useless.

11:34

I see. And things, you know, different steps to statistical weights are created for like different types of clients of any of the estimations. So do you think it's correct to simply sum up the estimation of violent and property crimes to get the index total? I mean, would there be any interactions between these different types of crimes in a motto?

12:00

No, there shouldn't be. I mean, the only issue is that it, you know, it's going to create an enormous plus or minus of enormous comments in a row. But

12:13

yeah, the elements are bad. So you can use whatever you want with them.

12:17

So can I also these different types of crimes, they are independent.

12:25

When it comes to the estimation, they should be yes. Okay. I see.

12:32

That's all my question.

12:35

Okay, just one thing to close our conversation like, how do like in the near future? How long do you think we'll be able to see the FBI make substantial improvement in their estimation, and when they graduate,

12:51

but we're graduating, this may like, next?

12:56

I don't know you're well, on your way to your first job. I think by that.

13:01

I mean, hopefully,

13:05

you know, next years. So the 2024 data released in 2025 is close to that 80%. Mark.

13:14

It might be the end of the decade before we're at where we were with UCR or SRS, I should say.

13:23

I don't know if that's really pessimistic or well, but

13:27

each year, should it be getting better? I wouldn't want to guess when it's going to get good, but it should, in theory get better each year.

13:36

I see. Okay. Um, all right. Thank you. Thank you so much. Yeah, my pleasure. Thank you.

13:45

So you mean, we cannot use any methods to compare the estimation between different between different counties? If I just I have a county and it has higher population and higher crime, data's in 2020. And it reported fewer a month in 2021? Can I claim that these counties estimation would be worse than other counties, which has, like fewer population and fewer crimes is in 2020. Also, they're not estimated at the county level, they're only estimated in the state and national level. So if they're, it's sort of a binary, either the county or the agency reported data or didn't.

14:29

And so if it did, then you can just use the data that it recorded, you shouldn't have to estimate at that level. The UCR county data that the old system was was pretty bad. Ooh. So it's generally recommended to use county level and to rely on agency level data instead.

14:48

You mean, the estimation is based on the county level or state level? The county the estimation is based on the state level estimated on the local level. So the

15:00

So the data can be collected from each agency, which when it comes to the estimation, we should only look at the state level. Is that correct? Yeah. For an estimation, yes. Okay. Okay. Got it. Thank you. It really helps.

15:14

Thank you.

15:16

Thank you so much. Thank you. Have a good day. I'll probably follow you up with any questions we encounter. I hope you don't mind. Sure. Of course. Happy to help. Thank you. Thank you. Thank you. Thank you. Bye