

## Data analysis can help us understand how to build healthy societies

Marina Kaminsky holds a Master's degree in public administration from the Evans School of Public Policy and Governance at the University of Washington and a Bachelor's degree in Economics from Reed College. She has experience in research on international rural and agricultural development, public health, and migration through positions at the Evans School Policy Analysis and Research Group, Innovations for Poverty Action, and the Center for Global Development.

She is currently part of the team of the [North Star Civic Foundation](#) in Portland, Oregon. She is passionate about working to identify solutions to economic and social inequities, both locally and in international contexts, and has pursued opportunities to do so through research and direct service.

**How did you end up working with data? Was that the plan from the very beginning, or did you start with a different idea in mind?**



*Source: <https://www.northstarcivic.org/>*

That's an interesting question! To answer how I got to work with data, I would probably have to go all the way back to my high school years. I took an Economics class, prior to which I thought Economics had something to do with money. Since I wasn't ever that motivated by money, I told myself "I probably won't like this, but I may as well try it out." However, the teacher was really passionate, and that class, together with History and Criminology, made me realize how much I could learn about people and trends from data, which I found intriguing. I thought of it as a really useful tool because we are often good at observing external phenomena, but when we are trying to reflect back on our own behavior as individuals or societies, we can get really myopic.

Thus, having data to inform our understanding of the world could be really useful in that context.

**Could you tell us more about why working with data is important? Can you give us an example?**

I think it is really important to pair quantitative analysis with qualitative data. We live in a world where more and more importance is attributed to quantitative information but sometimes it may

be misleading without the proper context. I would like to give an example of this. When I was in college I wrote a thesis looking at conditional cash transfers - those are basically used by a lot of governments in developing countries. They will give money as an incentive to families from lower economic backgrounds to do certain things like keep their kids in school or attend regular health checkups during pregnancy. These are activities that the government views as socially beneficial. So, I was examining a case in which I wanted to understand whether such a transfer in the sphere of education actually improves education outcomes for students whose families are recipients of the money. I was checking data such as enrollment rates, literacy rates, etc. After doing all that analysis it seemed like these transfers were indeed having a beneficial impact. However, years later, in Graduate School, I heard about some work that a researcher was doing on a similar topic. She was looking at a different program and there had been a lot of data that supported the idea that these conditional cash transfers worked. However, what she wanted to do was go into communities, and build trust over time so they didn't feel like someone from the government was coming in and asking them how they felt about the program.

She found out that, despite the benefits, there were a lot of struggles that people would go through when provided with transfers. The case of her study concerned pregnant mothers, who were receiving a certain amount of money if they would never miss their medical checkups. To access the healthcare facilities, some of the women had to walk several kilometers to reach them and sometimes the facilities were closed because they weren't well funded. In such a case the women had to come back the next day. That put a lot of burden on their bodies, especially as they were getting close to delivery, and was actually putting them at a higher risk.

I think that's a really good example of how in order to get a full picture of what's happening in the world, it's important to have quantitative data, but also pair it with information about individual and community experiences.

**What are the challenges that you have encountered while working with data? Can you share some of them?**

It is ideal when you can just get a data set that's really clean and work with it. Nevertheless, among the biggest challenges I've encountered is accessing data because it is either not open-access, or people are just not collecting it. Thus, there is no source of information and you know that this particular type of data would be really useful, but it has not been collected and you may not have the capacity to do it yourself. That's the first challenge that I would typically encounter. The second one is when the data is available, but in formats that take a lot of processing until you get to a point where it's clean, it's in one place, and it's ready to be analyzed. I would say that data collection and data cleaning are challenges most of the time. Once you have a complete and clean data set, you understand all of the components of the data. Then, the analysis tends to become more straightforward.

**In such a case what is the role of nongovernmental organizations when it comes to convincing public officials to make more data accessible or better organized? Is public data accessible in the US?**

There are a lot of public institutions that do produce data and make it available to people. You see that more in the case of institutions that are focused on research. The Washington State Department of Ecology, for example, started working as a research center with a lot of scientists and is an example of an institution that publishes a lot of data. These types of entities usually recognize and value making data accessible to other people. It is harder for smaller governmental organizations to produce data just because of limited capacity. You can still request data from the government if it is not confidential in the US (there is a number of reasons why it might be confidential). Usually, after such a request they will provide it to the public. So I think there's generally a good degree of data accessibility. Where I see it being less accessible is in agencies and organizations that may not typically have thought of themselves as scientists or data producers.

**As you mentioned, sometimes working with data seems distant to institutions, but also to the wider public. Can you give me an example of how your work with data has achieved a positive change?**

I think I can speak about a project from the last year, which was putting together a [Fact Book](#) that we published online. We read through academic articles trying to understand what researchers had been able to study on the topic of electoral reforms that were being proposed in the city of Portland OR. The aim was to share that research with the Community, specifically with journalists, but through them, with the broader public as well, so that people would get a sense of what impact this type of reform could have in Portland. Most of the research we were looking at was not theoretical but actually based on data. So these were researchers that had been looking at data from multiple election cycles, where this specific type of proportional rate choice voting was used. They were examining places either prior to this reform being implemented or they would look broadly at a lot of different countries to make comparisons and say what the results of one electoral system were (e.g., that it has led to more positive campaigning than other types of electoral systems).



When an organization like ours is able to collect all of these different research papers and say “This is what we're seeing broadly across the literature”, I think that's really powerful and it can allow people to understand which decisions are more appropriate for their communities.

I'm talking about this more from a public policy perspective because that's the sphere I am involved in, but there are a lot of other types of data-informed decisions related to the environment or marketing, for instance. However, in terms of public policy, I think having data to analyze allows us to understand which programs support our communities, how we should distribute resources, and how to create healthy societies. I just view data as a really powerful tool.

One more thing is worth mentioning here. Sometimes people do not understand how somebody sitting and working on an Excel sheet can contribute to society. Sometimes public administrators also feel burdened by collecting data, as it can take a lot of work to do the data collection, to clean the data, to get it to a point where researchers can analyze it, and just to get it to the point where it can be useful in that sense. All of the effort on data collection or data cleaning, while it may not seem like directly helping people, is in fact crucial. The whole research process is dependent on that work. I think those people should be recognized a lot more than they are, not just the people who actually use the data and write a paper, or the journalists who publish the information or try to make it available to communities. It just takes a lot of people and a lot of effort to get the data and then share it with the world.

**Working with data is indeed important but often requires specific knowledge. Do you think that nongovernmental organizations in the US as a whole, possess this required knowledge, and what should be done better, in your opinion?**

I think the situation is pretty much the same as with governmental entities. There is a really wide range across nongovernmental organizations in terms of their ability and capacity to do data analysis. For example, I used to consult for the Bill and Melinda Gates Foundation, which is a huge nongovernmental organization and has a lot of resources. They use data really frequently to make data-driven decisions and that's a priority for them. There are a lot of very data-savvy people in the organization. In contrast, a small community-based organization in Portland focused on a particular issue of a particular group might be under-resourced. Thus, they often may not have the capacity for that. Even bigger organizations sometimes haven't developed those skills on their team yet because they've just been slower to adopt data as a tool in the work they do.

I think more and more organizations are recognizing the value of being able to do data analysis. Even if it is a small entity, working on advocacy. Having information, all kinds of information from the community about the community is a very important part of decision-making and programming. If we think about the issue of homelessness, for instance, having data about what homeless folks need, what part of the city they're residing in, what they do when it's really cold outside, etc. would help the organizations that are trying to support the homeless community be able to make decisions on how to do that in a way that benefits more people.

In terms of how to address the lack of capacity, there are multiple ways to support improving it in non-governmental organizations. One of them is starting with funders who can provide opportunities for that. They can provide multi-year funds under the condition of building data analysis capacity. Another way is just to start teaching data analysis in schools. If there are

young people who are passionate about working to support their communities and understand the importance of data, they will bring those skills as they move out of school and start working in these places.

Additionally, even if organizations can't hire somebody full-time to do data analysis, they might try to identify ways to train their staff. At my previous job, we were discussing creating a program that would bring mid-level career people into public entities or nongovernmental organizations from different countries to the university. We would do a couple of month-long training programs that would teach them basic computer coding skills and data analysis skills. Such programs are a good example and they don't have to be international. I think the tougher part is convincing people that working with data is useful because if people don't see it as beneficial, then it's hard to make them invest in developing in that direction.