Closing the big data analysis access gap between low- and middle-income countries

July 20 2021, by David Gunderman and Eric Vance

Ecologists, like the one Eric Vance met in western Africa, use statistics to learn about the animals and ecosystems they study. Credit: Eric Vance, CC BY-ND

The ability to collect and learn from large amounts of data has been a [major driver of innovation](#) over recent decades. Everything from [health care](#)—think patient analytics, wearable devices and the COVID-19 response—to [transportation](#)—Uber and Lyft—to [entertainment](#)—Netflix—is now driven by data and statistics.

Yet the ability to collect good data, the capacity to derive insights from it and the skills to turn those insights into change aren't spread evenly across the globe.

Taking a page from the way Doctors Without Borders sends medical personnel and expertise to developing countries, [some organizations](#) have begun to do the same with statistics. But overall, the need to improve local statistical capacity in developing nations [remains largely unmet](#).

We are two mathematicians at the [University of Colorado Boulder](#) and are part of a project called the [Laboratory for Interdisciplinary Statistical Analysis](#) that is working to develop statistical infrastructure across the world. The goal of the program is to help build data science infrastructure in developing nations. In 10 countries and counting, we have started "stat labs"—academic centers that train young statisticians to collaborate on important local statistics projects.

Where statistics matter

The benefit of a program like Doctors Without Borders is obvious—the group provides medical care. The benefit of improved statistical capacity is harder to see but can be just as important.

For example, during the great cholera outbreak in London in 1854, John Snow used statistical data collection and analysis to [identify and close off the contaminated water pump](#). Later that year, Florence Nightingale, the founder of modern nursing, used statistics to show that simple

hygiene measures could [drastically reduce infection and death in hospitals](link).

Every year, the [World Bank scores countries](link) on a scale of 1-100. One represents a complete lack of basic statistical data and analysis capacity, and 100 represents the statistical capacity of a developed nation like the U.S. According to [the 2020 report](link), the average statistical capacity of countries in sub-Saharan Africa, South Asia and Latin America are 57.1, 69.8 and 70.1, respectively.

This disparate statistical capacity has played an important role in the pandemic. Strong data collection and analysis of COVID-19 cases allowed some countries—like Nigeria and the U.S. – to [better respond to the initial outbreaks](link) and take an informed approach when [reopening sections of the economy](link).
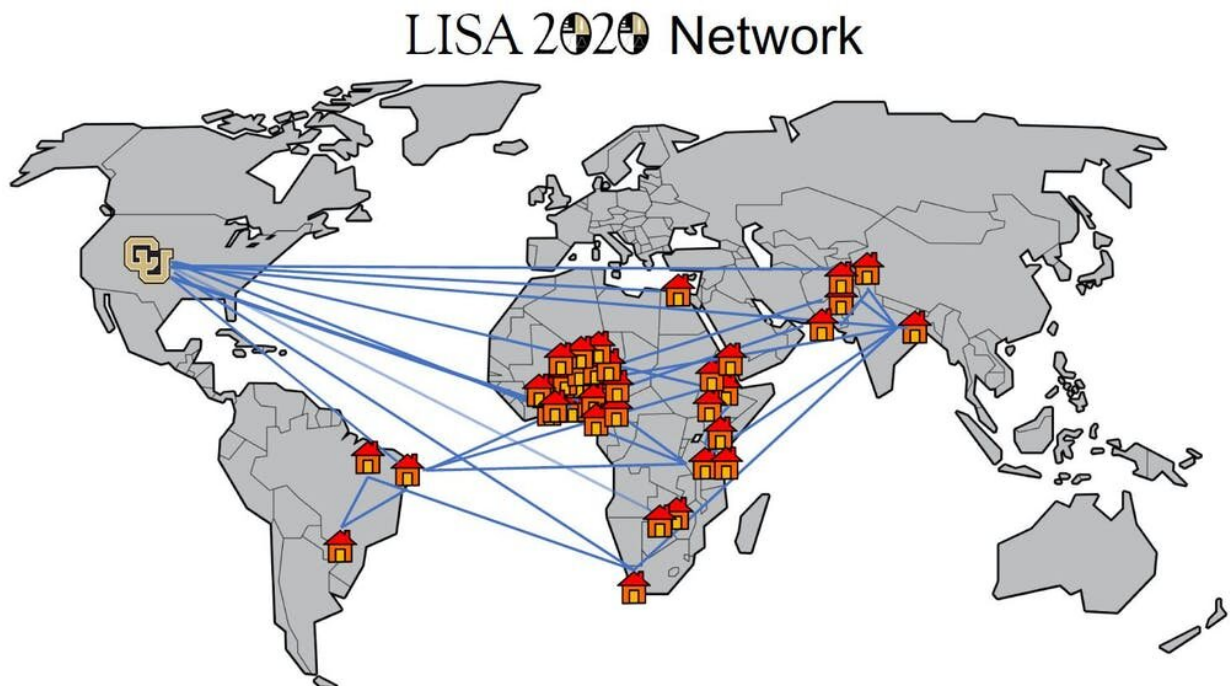
The LISA 2020 network has grown to include more than 30 statistical

Unfortunately, during the pandemic, fully [80% of national statistics offices](#) in lower- to [middle-income countries](#) indicated they needed additional support to perform important data collection and analysis.

Just as good data can lead to good decisions, lack of data can often lead to less effective decisions. For example, during the 2014 to 2016 Ebola epidemic in Liberia, the government [did not initially](#) have access to accurate, real-time mortality data or effective analysis tools. This lack prevented public health authorities from quickly and effectively responding to outbreaks. Once the government introduced a [phone-based data collection system](#), officials were better able to allocate doctors and nurses to where they were needed.

Statistics in ecology, health and politics

The idea for the Laboratory for Interdisciplinary Statistical Analysis started in northwestern Africa, on the border of Western Sahara and Mauritania. One of us, Eric Vance, was in the middle of a five-year stint traveling around the world before his Ph.D. At a border checkpoint in the middle of an old minefield, he coincidentally met a biologist who was studying the Saharan desert fox.

When the biologist found out Vance was studying statistics, his eyes lit up, and he said, "Oh, a statistician! I have questions for you." But before Vance could offer any help, he had to get on a bus and cross the mine-filled border. When Vance got back to the U.S., he realized the widespread need for statistics capacity and education [in developing countries](#). To address this gap, he launched the global [LISA 2020 Network](#) in 2012.

The goal of the program is to give local college students the [skills and tools](#) to do the statistics they need to drive development. We help local professors establish a statistics laboratory at universities where they work. These statistics labs are collaborative centers where local professors teach students to provide statistics consulting to other academics, businesses and policy makers. While the students are learning statistics they're also using their technical skills to drive real, local change.

One of our [partner labs](#) is working with Nigeria's [Independent National Electoral Commission](#). Together, they are assessing the accuracy, completeness, consistency and reliability of data within Nigeria's Continuous Voter Registration policy to explore ways to improve the electoral process for voters.

In Ethiopia, another local lab is helping the Ethiopian government improve the registry of births and deaths. Using surveys, effective database management and statistical training programs the goal is to improve health outcomes.

Since launching in 2012, our network of stat labs has grown substantially, with particularly strong roots in Africa, South Asia and Brazil. As of July 2021, it consists of [31 stat labs](#) located in 10 low- and middle-income countries.

As statistics continues to play an ever more important role in society, equal access to data resources in [developing countries is becoming more essential](#).

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).