

# GOVERNING

THE STATES AND LOCALITIES

## The Next Big Thing in Data Analytics

As the amount of data that governments accumulate grows, so does the need to disaggregate it.

BY KATHERINE BARRETT & RICHARD GREENE | JULY 2015

Drive through most Connecticut communities on trash pickup day, and you'll discover two containers in front of many homes. One is for run-of-the-mill garbage. The other is for recycling.

Obviously, most communities would prefer that their citizens recycle as vigorously as possible. In Stamford, Conn., for example, leaders know that citywide about 28 percent of the trash is recycled. That may be useful information, but far more helpful is to know how much trash is recycled neighborhood by neighborhood. "Certain communities will recycle over 60 percent," says Jay Fountain, chair of the Fiscal Committee of the Stamford Board of Representatives. "Others will recycle from 5 to 10 percent."

That breakdown -- or disaggregation as it's known in data circles -- provides the kind of information that allows the city to "find the areas in which we need to encourage recycling," Fountain says.

An emphasis on disaggregating information has been growing in importance as states and cities tap into huge quantities of information for more sophisticated analyses. It's at the "heart of the new focus on data analytics," says Harry Hatry, director of the public management program at the Urban Institute, who has been an advocate of using disaggregated data for decades.

Not only does disaggregating data make it more useful for policymakers, it's also key to engaging public interest. Consider schools in New York City, our hometown. Any data related to the quality of the public schools citywide is only minimally useful to parents and school administrators.

Schools in Staten Island are very different from those in the South Bronx. But school-by-school breakouts capture the public's attention and are of use in determining which schools may need the most help. "People will pay attention to personalized information," says Diane Lim, vice president of economic research at the Committee for Economic Development. "If you're looking at colleges, you want to see about your alma mater."

Sometimes, seemingly disaggregated data needs yet more fine tuning. Just a few weeks ago, the Centers for Disease Control and Prevention (CDC) released a study about Hispanic health risks in America. One of its top findings was that "fewer Hispanics than whites die from the 10

leading causes of death, but Hispanics had higher death rates than whites from diabetes and chronic liver disease and cirrhosis.”

Think about this for a moment. Hispanics have many different nations of origin. Lumping them all together can be misleading. Consider smoking rates. According to more disaggregated CDC data, only about 14 percent of Hispanics smoke, compared with 24 percent of whites. But if you look at Puerto Rican males, the picture changes, with 26 percent of that population smoking.

Or split things up another way. It turns out that Hispanics are as likely as whites to have high blood pressure, but Hispanic women are “twice as likely as Hispanic men” to get it under control, reported the CDC.

When it comes to economic information, disaggregation is particularly crucial. “The economy is the sum of the parts and all the parts are moving in different ways,” says Lim. “So, you can’t see the direction of the trend from the aggregate.”

Of course, she’s right. This is kind of like the old joke that features three men shooting at a target. The first fellow shoots 10 feet to the right of the target. The second fellow’s arrow goes 10 feet to the left. Observing this, the third gentleman, an economist who leans on aggregated numbers, says, “Hey, the two of you hit a bull’s-eye.”

Disaggregation is necessary for effective use of performance information. Research has found that performance data is best used in individual agencies -- as opposed to governmentwide -- because it’s far easier for agencies to make use of the separate pieces of information. “This is important for managers to be able to manage both their money and their people,” says Katherine Willoughby, professor of public policy and management in the Andrew Young School of Policy Studies at Georgia State University.

All of this is not to say that providing policymakers and citizens with only disaggregated numbers is enough. It’s the combination of both the aggregated and the disaggregated that gives a more complete story. Yet there are many instances in which there is no central source to put all the individual numbers from cities, counties or states together in a way that would provide an aggregate benchmark as a means of measure.

In criminal justice, for instance, one area of current interest focuses on local governments that are charging fees and fines. In some places, men and women are going to jail for not paying parking fees. “But, we don’t have any kind of aggregated national source for understanding the extent of this or how it’s changed over time,” says Michael Leachman, director of state fiscal research for the Center on Budget and Policy Priorities.

Of course, even the proper combination of aggregated and disaggregated information isn’t a silver bullet. No single general rule is a panacea. “Having more good data doesn’t necessarily make for better policy decisions,” Leachman says. “But at least it makes it possible to make good policy decisions.”