WHY DATA STORIES?

The leaders of DataSpark, a program at the University of Rhode Island, recognized long ago that rigorous data analysis is only one step toward achieving their mission of informing and inspiring innovative decisionmaking. Over the years, they have developed a collaborative process of producing data stories—a series of data visualizations focused on answering a specific question. From “What is the employment landscape for graduates from the public higher education system?” to “What value do volunteers bring to neighborhoods and to the state?” these visualizations and their accompanying text guide the audience through data points to illuminate the answers.

By inviting local stakeholders into the process of producing a data story, DataSpark breaks down institutional silos, builds participant technical expertise, and expands dissemination of the product. Cross-sector collaboration is central in their model. Regular meetings of the data-story team encourage more communication among government agencies, nonprofit, and philanthropic partners. Team member feedback helps to inform the research, providing information on potential data use and reliability concerns and guiding the team away from issues that may be overplayed in favor of ones that are fresh and timely. Diverse teams bring a range of policy perspectives and backgrounds to the table, improving the data story beyond what DataSpark might accomplish alone. Moreover, discussions surrounding the best presentation of data and potential pitfalls can help build technical expertise among team members. Once launched, stakeholders bring the story more visibility by linking to the content on their websites and using the story at events or briefings.

DataSpark’s contributions to policy conversations are particularly noteworthy because they often tap the existing integrated data system, the Rhode Island DataHUB. The data-story process results in sharing new insights from data across sources that are linked at the individual-level with the public in a way that preserves confidentiality.

This brief generalizes lessons from DataSpark’s experience to encourage other organizations to adopt the data-story model. It outlines the seven steps in the data-story process, presents important considerations for successful data stories, and concludes with spotlights on two DataSpark data stories.
DATA-STORY PROCESS

Step One: Determine a convening organization.

The convener will be a bridge between the data and the data-story team for the 6 to 12 months that the process generally takes. An appropriate convenor values and builds capacity for data usage and analysis across the data-story team. The staff must have (1) technical capacity to analyze trends and patterns in the data; (2) the ability to produce actionable information and visualizations that make data easy to understand; and (3) experience in communicating with nontechnical audiences. In addition, having a store of data sources with data-sharing agreements in place will reduce the extra time needed to access data once a topic is selected. The convening organization should also have a website to host and promote the final product. An appropriate convenor may have deep topical expertise on an issue or may have broad expertise across many issues, as is the case with DataSpark.

Step Two: Identify a topic and related data sources.

The funder (e.g., a state agency or foundation) or the convener typically identifies the topic for the data story. The topics can be broad, such as a policy area, like workforce development, postsecondary education, or health, or more specific, like the educational costs of unhealthy housing. The topic selection will depend on having relevant and accessible data to understand the policy or research questions. This may be from a single data source, multiple data sources analyzed separately, or multiple data sources linked at the record level, like the integrated data system used by DataSpark.

Step Three: Build a team.

The convener brings together a cross-sector team of organizations with experience in the topic area, typically involving 8 to 12 stakeholders. In addition to the convener, key team members include the data provider(s) and data stewards, who weigh in on technical questions and data limitations. Other potential members include content experts (including program staff and policy experts), who provide important insight on the issue and key audiences.

DataSpark recommends only bringing decisionmakers to the table at the end of the process, when the data story has actionable components. Some data stories will only be shared internally with the team or presented to an audience as a source of background information. In these latter instances, a decisionmaker would not be a necessary member of the team.

Teams can be as small as 3 to 4 stakeholders and as large as 15 to 20, in part determined by the level of interest. The number will also depend on the constraints and goals of the project. In general, the more stakeholders involved, the...
longer it takes to move through the process. If time or funding is an issue, the convener may prioritize the efficient completion of the final product, involving fewer stakeholders. If collaboration is one of the central goals, the convener may invest more time in the process and include more stakeholders. The size of the team may also change over time, naturally shrinking as the intensity of engagement in the process or number of meetings increases.

Ultimately, the convener needs to balance the logistics of scheduling meetings and maintaining active engagement against the desire to involve the largest number of stakeholders.

**Step Four: Convene the team and draft a question.**

Once the data-story team has been identified, the process begins with a kickoff meeting. For larger teams, DataSpark hires an independent consultant to moderate the meetings, finding that a neutral facilitator helps balance the interests of multiple stakeholders and promotes conversation. The purpose of the initial team meeting is to introduce the data-story process and draft a specific question related to the topic of interest. The relevance of the policy question to the stakeholders and the community should be primary considerations, as well as the urgency of the issue. Certain topics require a faster response than what this process typically allows. Practically, data limitations and availability will also determine which questions can be chosen.

The kickoff meeting may not end with agreement on a single question, but the convener should work to narrow the ideas down to two to three questions and have the team finalize the questions at the next meeting. Producing multiple data stories over time with the same team can also create opportunities to revisit ideas that were not selected initially.

Following the kickoff meeting, the team meets on a regular basis to explore the data and develop the data story. Meeting frequency depends on available funding and desired stakeholder engagement, as well as the project timeline. More meetings allow for greater stakeholder engagement and feedback, but they can extend the timeline of producing the final product.

**Step Five: Explore the data and develop a story.**

The objective of this step is to figure out what the data say and the importance of those findings, and to frame the story so that it is relevant to the audience. Frequently the longest step in the process, this phase can take up to about a third of the entire time required to make a data story. This step is more time intensive because it is the most iterative. An analyst for the convener (e.g., one of DataSpark’s data analysts) will explore the data and report back to the team through an online forum or in a regular meeting to describe their early analysis and share initial data visualizations. The discussion of this analysis will generate new questions, prompting the analyst
to return to the data. Occasionally, the analyst will return with findings of the data that are not interesting, and the team will have to explore a new direction.

To conclude this step, the convener drafts a series of final figures and charts that work best to tell that story.

**Step Six: Vet the story.**

The team shares the draft data story externally (beyond the team) to solicit feedback and expand the circle of stakeholders invested in the final project. Teams should identify for reviewers the aspects of the project that they are willing to change (for example, whether to address questions that are beyond the scope of the original issue area) when they share the story externally; this will focus stakeholder feedback to specific components of the data story and ensure the reviewers time is spent wisely.

**Step Seven: Publish and disseminate the story.**

The convener and team members determine how and where to publicize the data story. Depending on the stakeholders involved with the team and the focus of the project, the presentation of the data story could occur in a public meeting, an internal briefing for an agency or funder, or over social media. The team should also take advantage of their networks to determine the best strategies for promotion and dissemination.

**SPOTLIGHT: THE IMPACT OF NATIONAL SERVICE ON RI COMMUNITIES**

The data story “The Impact of National Service on RI Communities” studies the role of AmeriCorps volunteers working in Rhode Island over the last 10 years. DataSpark used their integrated data system, the RI DataHUB, to link data for each person from Serve RI (the state AmeriCorps agency), the Rhode Island Departments of Education and Labor and Training, the RI Voter Database, and the National Student Clearinghouse. The data-story team was composed of staff from several of these agencies.

Using a combination of different types of graphs, the data story explores the impact of AmeriCorps members on the state’s social and economic development. AmeriCorps volunteers tend to concentrate their services in areas with the highest poverty rates (as seen in the figure 1). Members frequently come from the communities they serve, and members tend to remain in Rhode Island for school and work after their service. AmeriCorps members are also more likely than not to later work in public service fields, including education services, health care, and social assistance. Voting, a measure of civic engagement, also tends to be higher among AmeriCorps members compared with the
Among the AmeriCorps members studied, voter turnout exceeded that of the general population in both 2012 and 2014 (by 21 percentage points and 16 percentage points, respectively).

Consisting of 14 individual visualizations, the data story helped ServeRI understand the impact of the AmeriCorps members it places in the state, lobby for additional match funding, and encourage RI colleges and universities to inspire high-intensity volunteerism.

**SPOTLIGHT: PREPARING WORKERS FOR RI’S HIGH-DEMAND OCCUPATIONS**

The data story “Preparing Workers for RI’s High-Demand Occupations” explores potential skill gaps or labor market mismatch in Rhode Island. Using data from the Rhode Island Department of Labor and Training’s Labor Market Information division and the RI Office of the Postsecondary Commissioner, a series of six interactive tables and graphs identify potential labor shortages (areas of high demand) across different occupations within the state. The data story identifies a shortage of graduates trained for finance and computer occupations. At the same time, the state may be overproducing graduates in logistics and, though to a lesser extent, graduates in marketing, pharmacy, and registered nursing. Wages for graduates in high-demand fields tend to be on par with (or exceed) entry-level wages in neighboring Massachussetts and the nation as a whole, suggesting that wages might not be the best explanation for labor shortages.

Figure 2 explores whether more years of education for nurses means higher wages. Using the linked data, the team found that registered nurses with
associate’s degrees could expect to earn as much as those with bachelor’s degrees, though they tended to be more likely to work in skilled nursing facilities than in hospitals.

Employing a combination of tables, bar charts, and pie charts, the data-story process involved about 20 people in a kickoff meeting including data providers, policy people, and data analysts. A smaller working group made up of the main data providers, the Department of Labor, and the Office of the Postsecondary Commissioner, continued to meet to accomplish the tasks laid out in step 5 above.

The data story was shared with decisionmakers at the two state agencies providing the data. Ultimately, a finding that was tangential to the labor shortage question (what types of training produced graduates that stayed in state for work), was most interesting to the data providers. The data story was also shared widely on social media by the data providers and DataSpark.

To view more data stories from DataSpark visit: http://ridatahub.org/datastories/.

Source: DataSpark
Amanda Gold is a research associate in the Metropolitan Housing and Communities Center at the Urban Institute.

Leah Hendey is a senior research associate in the Metropolitan Housing and Communities Center at the Urban Institute.

Thanks to Kimberly Pierson, director of DataSpark and Megan Swindall, data analyst, for sharing her insights about the data-story process.

Coordinated by the Urban Institute, the National Neighborhood Indicators Partnership consists of independent organizations, like DataSpark, in more than 30 cities that have a shared mission to help community stakeholders use neighborhood data for better decisionmaking, with a focus on assisting organizations and residents in low-income communities.

This paper was supported by the Annie E. Casey Foundation. The views expressed are those of the authors and do not necessarily represent those of Annie E. Casey Foundation or the Urban Institute, its trustees, or its funders.

Copyright © 2017. Urban Institute. Permission is granted for reproduction of this file, with attribution to the Urban Institute.