Harvard Business Review

ARTICLE
DATA

Help Your Team Understand What Data Is and Isn't Good For

by Joel Shapiro

DATA

Help Your Team Understand What Data Is and Isn't Good For

by Joel Shapiro

OCTOBER 12, 2018

JOHN LUND/GETTY IMAGES

Leaders today increasingly turn to big data and advanced analytics in hopes of solving their most pressing problems, whether it's a drop-off of repeat customers, a shift in consumption patterns, or an attempt to reach new markets. The prevailing thought is that more data is better, especially given advancements in tools and technologies such as artificial intelligence and predictive analytics.

But when it comes to uncovering the motivations and rationale behind individual behaviors within a social system, data can only do so much. It can guide the discovery of a problem, but it won't determine the solution. In other words, data analytics can tell you *what* is happening, but it will rarely tell you *why*. To effectively bring together the what and the why — a problem and its cause, in order to find a probable solution — leaders need to combine the advanced capabilities of big data and analytics with tried-and-true qualitative approaches such as interviewing groups of individuals, conducting focus groups, and in-depth observation.

In my conversations with business leaders about how they use data analytics, a primary focus is on technical, large-scale systems. This is where big data and analytics can really shine, in applications such as predictive maintenance. Industrial companies, from railroads to oilfields, use predictive analytics to ensure smooth operations; rather than wait for a mechanical breakdown to occur, predictive maintenance prevents problems and avoids downtime.

What works with locomotives and oil rigs, however, can be far less effective when it comes to influencing people's behaviors. With social systems and the behaviors generated by large groups of individuals — who does what and under what conditions — it is far harder to identify solutions to problems. This points to the shortcoming of using data analytics alone for solving problems that arise from individual behavior.

That's not to say big data and analytics don't play an important part. Rather, by understanding the strengths and limitations of using big data in this way, leaders can employ the most effective strategies for identifying the what and why of a problem, and how to solve it — and can help their teams learn to do the same. Here are five important considerations that everyone who works with big data needs to understand:

1. **Data can determine the "what" of a problem:** Data analysis is helpful in determining patterns of behavior, both positive and negative — for example, the success of an organization or enterprise in motivating people to engage in certain activities. Analyses may reveal, for instance, that a certain type of customer is more or less likely to buy a particular product or renew a subscription or membership. Sophisticated data analytics can reveal patterns among large groups and smaller subgroups.

2. **Data rarely reveal the "why":** In the aggregate, individual behaviors show up in the data, revealing patterns among certain demographics and groups. But just because data show, say, what the typical 33-year-old women making less than $100,000 a year who has children is likely to do or not do, that won't reveal the *why*. Data may prompt people to make assumptions; for example, that a price point was too high for a particular customer, or that a subscription service related to a leisure activity (e.g., a gym membership) no longer appeals to a consumer who has time constraints. Assumptions also can be made about root causes of behaviors, such as why millennials prefer companies that prioritize social impact or why particular subgroup of employees underperform. Assumptions, though, are only guesses about the rationale of others' behaviors, not a reliable basis for determining the best solution to address a problem.

3. **The "why" needs a qualitative approach:** Whether the social group involves current customers, potential customers, vendors, or any other population, the only way to discover the "why" is to engage with them in qualitative research such as interviews, focus groups, and observation. The result is an iterative process that starts with the "who" and the "what," which the data can reveal, and proceeds to the next step of diagnosing the "why," which the data cannot typically reveal. In the past, companies often hired experts in qualitative research to help determine how and why customers use particular products or gravitated to certain brands. Today, though, many business leaders try to use big data and analytics to automate the entire process. But the shortcomings of using data for diagnostics of social behaviors are quickly revealed. For example, social media analytics can identify influencers for well-defined customer segments. But the real challenge is knowing *why* customers are drawn to those influencers in order to craft effective strategies to entice customers to buy more or become brand advocates themselves.

4. **You need to consider temporal and other factors:** Other factors also influence behavior, making solutions more difficult to find and less likely to remain effective over time. For example, several years ago, an auto club discovered that motorists who had longer-than-average wait times by the side of the road were less likely to renew their subscriptions. Based on that data, the company emphasized the need to reduce wait times. Since then, the proliferation of smartphones and other devices have given people ways to occupy themselves, altering their perception of how long they're waiting. As a result, focusing on wait time alone today (as opposed to other factors such as pricing and quality) has proved to be less effective in reducing churn among auto club members.

5. **You need rigorous testing to find the right solution:** With big data analysis and smaller-scale qualitative research combined, organizations can gain deeper insights into both problems and their causes, which can then help inform solutions that are likely to produce a desired result. The best way to know the effectiveness of a solution is to conduct randomized testing using two similar groups: one that is offered the solution and one that is not. Data analysis from this experimentation will reveal whether the solution actually solves the problem. Although randomized experiments can be expensive and complex, the data analysis involved brings the process full circle, and often pays for itself in terms of the return on investment.

Data analytics are most effective as part of an overall process to identify, explore, and test, but are not the only tool for the task. Solving social behaviors still requires small-scale qualitative exploration to engage people and learn more about what's truly motivating the behaviors that show up in the data.

Joel Shapiro is clinical associate professor at the Kellogg School of Management at Northwestern University and academic director of Executive Education.