What Leaders Get Wrong About Data-Driven Decisions

By Bart de Langhe and Stefano Puntoni

Data-driven decision-making has been hailed as an antidote to the biases of human intuition. Companies have more data than ever, but surveys cast doubt on the effectiveness of data-driven decisions in organizations. The majority of executives say their data analytics initiatives produce disappointing results,[i] and only about a quarter of executives say their data analytics projects produce actionable insights.[ii]

There is a clue to the problem in the name "data-driven decision-making." It's *data-driven*.

Data-driven **often means taking data at face value**

Data have the potential to increase our understanding of how the world works, but using data to make decisions is not by itself a safeguard against the influence of pre-existing beliefs and incentives.

Take the case of "Gwenn & Jenny's," an ice-cream vendor that wants to know how its online advertising impacts sales.

To answer this question, Twitter offers a three-step process.[iii] First, a data broker, like Datalogix, shares identifying information from a brand's customers with Twitter (like browser cookies, email addresses, and phone numbers). Next, Twitter searches for these customers in their records and, if there is a match, they add information about these customers' activities on Twitter (like whether they viewed or clicked on the brand's Tweets). Finally, analysts can compare purchase decisions of customers who engaged with the brand on Twitter with purchase decisions of customers who did not.

Twitter's approach invariably reveals stark differences: Customers that saw and engaged with a brand on Twitter visit its stores more often, and they spend more on each visit. Taking these data to suggest that social media advertising has a large impact on sales supports Twitter's business model. It's also consistent with clients' beliefs that social media advertising works and that its effectiveness can easily be measured.

Twitter has sold its three-step process to many companies. We use it in our teaching as an example of flawed data-driven decision-making.

Comparing customers who saw a brand's content with customers who did not see the content is like comparing apples and oranges. These customers differ in many other ways. Gwenn & Jenny's most loyal customers are more likely to engage with the brand on Twitter and they are also more likely to buy the brand. They don't buy because they saw the brand's Tweets. They buy because they like Gwenn & Jenny's, and because they like Gwenn & Jenny's, they follow the brand on Twitter. Twitter's approach dramatically exaggerates the impact of advertising on sales.

Data-driven decision making starts from the data that is provided by data providers and data scientists. This often leads decision-makers to accept data uncritically.

Data-driven often means answering the wrong question

Data-driven decision-making in practice often comes down to finding a purpose for data. Companies look for ways to extract insight from data they have collected, but the data that are available may not be the data that are needed to make a decision.

Take customer relationship management at "RollingBoulder", a media company with a subscription-based business model. RollingBoulder's customers have annual memberships. These can be renewed by completing a renewal letter, which the company sends to each customer when a membership is about to expire. To reduce customer attrition, the organization sometimes adds a "thank you" gift to renewal letters.[iv]

Over the years, RollingBoulder has developed a rich dataset that describes past and current customers along various dimensions (like zip code, when a customer first became a member, and how a customer behaves on the website). The company has developed a sophisticated predictive algorithm that quantifies the likelihood that an active member will churn. It then sends a gift to the customers most likely to churn.

This data-driven approach to churn management is considered best practice in the industry.[v] Unfortunately, it's another example of flawed data-driven decision-making.

RollingBoulder's predictive algorithm gives a precise answer to the question "How likely is a customer to churn?" This is valuable information. For example, it allows the firm to make projections about the value of its customer base. However, it does not address the question that is relevant here: "What is the effect of including a gift on a customer's likelihood to churn?" This question cannot be answered based on the available data. It requires collecting new data.

Data-driven decision-making anchors on data that are available. This often leads decision-makers to focus on the wrong question.

Decision-driven data analytics

Decision-making shouldn't be data-driven. Rather the opposite: Data analytics should be decision-driven.

Data-driven decision-making starts with data and the people who analyze them. Decision-driven data-analytics starts with decisions and the people who make them. Data-driven decision-making is about finding a purpose for data. Decision-driven data analytics is about finding data for a purpose.

Data-driven decision-making	**Decision-driven data analytics**
Starts from the data that are available	Starts from the decision to be made
Find a question for an answer	Find an answer to a question
Finding a purpose for data	Finding data for a purpose
Emphasis on what is known	Emphasis on what is unknown
Empowers data scientists	Empowers decision makers

When we tell executives about the decision-driven approach to data analytics, some are quick to point out a potential problem. They caution that decision-makers who utilize data to support a decision that has already been made may be falling prey to confirmation bias. But that is not decision-driven data analytics. That is "preference-driven data analytics," and it might be the worst way to make decisions—but unfortunately a very common one.

Decision-driven data analytics involves two important steps. In the first step, it is the responsibility of decision-makers to form a narrow consideration set of alternative courses of action. In the second step, it is the joint responsibility of decision-makers and data scientists to identify the data needed to figure out which course of action is best.

Step 1: What are the alternative courses of action?

In this step, decision-makers think "wide-then-narrow". Many decisions are made on autopilot, after considering only one course of action. This can harm the quality of decision-making. Thinking "wide" means that decision-makers generate many alternative courses of action. To illustrate, let's go back to RollingBoulder. If the business objective is to increase the value of its customer base, a "thank you" gift to reduce churn is just one of many possible courses of action. The company could also improve customer development (e.g., through up- or cross-selling), make investments to acquire new customers (e.g., through sales promotions), or improve editorial content (e.g., by hiring new writers).

However, too many alternatives can make the problem intractable from a managerial and data-analytic perspective. Thinking "narrow" means that decision-makers use their judgment to winnow courses of action, for instance because of anticipated difficulties in implementation, or lack of support from top management. For instance, the customer relationship manager at RollingBoulder may realize that improving editorial content falls outside her responsibilities.

By thinking "wide-then-narrow" decision-makers increase the likelihood that the final consideration set includes high-quality and feasible courses of action.

Step 2: What data do we need in order to rank alternative courses of action?

In the second step, decision-makers and data scientists develop criteria to discriminate and rank the alternative courses of action selected in the first step. The goal of data analytics is to turn unknowns into knowns, so that alternative courses of action can be ranked more objectively.

Starting from the decision draws attention to unknowns, and this has a major advantage. It makes it immediately obvious that there are limits to what can be known and that unknowns can be tackled in many different ways. For instance, if you tell people that Interbrand has determined that the Mastercard brand is worth \$9,430M, and ranks 62th in the world,[vi] most people will take this point estimate and rank at face value. That's one of the problems with data-driven decision-making, as we saw in the Gwenn and Jenny's example. Instead, if you ask people what they think the Mastercard brand is worth, most will realize that brand value is a complex multi-dimensional construct that can be quantified only imperfectly, and in different ways. They are correct. Kantar Millward Brown estimates the Mastercard brand to be worth \$91,929M (ranked 12th in the world)[vii] and Brand Finance estimates it at \$18,293M (ranked 94th in the world).[viii] Starting from what is unknown highlights that the world is complex and uncertain.

Decision-driven data analytics is not about collecting as much data as possible. It's crucial to consider the value of data. If one would take the same decision before turning an unknown into a known than after, then there is no benefit of engaging in data collection and analysis.

Oftentimes, data that are collected for the purpose of making a decision have more value than data that are already available. RollingBoulder had to decide whether to add a gift to a customer's renewal letter. To make this decision, the company needed to know how sending a gift influences a customer's likelihood to churn. This question cannot be addressed based on the available data. It requires running a randomized controlled trial (an RCT, or A/B test): Customers are randomly selected to receive the gift or not, and the company then observes which customers churn and which customers stay.

Step 3: What is the best course of action?

The final step should be easy. If the first two steps were executed well, data analytics will now reveal the best course of action.

Analyzing the data from the RCT, RollingBoulder learned an important lesson. The gift successfully reduced churn likelihood for some customers, but it backfired for others. The customers that were least likely to churn without the intervention became even less likely to churn when targeted with the intervention. Instead, the customers that were most likely to churn without the intervention became even more likely to churn when targeted with the intervention. In other words, the company had always been targeting precisely the wrong customers, along with many other businesses that blindly follow the "best-practice" of targeting high-risk customers.

Conclusion

Decision-driven data analytics emphasizes the importance of asking questions, and thus, the importance of managerial judgment. The job of decision-makers is not to accept or reject data and analyses offered by data scientists. Their job is to critically reflect on alternative courses of action, and the data and analyses that are needed to make a decision. All this is done *before* any analyses are performed. This encourages decision-makers to make their beliefs about the world explicit. It draws attention to unknowns and the value of additional data collection and analysis.

Pablo Picasso once said: “Computers are useless. They can only give you answers.”

---

[i] See here: https://www.thehackettgroup.com/2020-key-issues-it-1912/?leadSourceMostRecent=Social%20Media&leadSource=Social%20Media&campaignID=701330000012mjlAAA&leadSourceDescription=HCKT%20Download:%20Key%20Issues%2020Q1IT%20-%20SM

[ii] See here: https://www.accenture.com/_acnmedia/pdf-108/accenture-closing-data-value-gap-fixed.pdf

[iii] See here: https://help.twitter.com/en/safety-and-security/measuring-sales-impact

[iv] “RollingBoulder” is modeled after a company described in Ascarza (2018). See here: https://journals.sagepub.com/doi/full/10.1509/jmr.16.0163

[v] See for instance here: https://en.wikipedia.org/wiki/Customer_attrition

[vi] See here: https://www.interbrand.com/best-brands/best-global-brands/2019/ranking/mastercard/

[vii] See here: http://online.pubhtml5.com/bydd/ksdy/#p=33

[viii] See here: https://brandfinance.com/images/upload/global_500_2019_free.pdf