

Chapter 4. Data-Driven Versus Data-Informed

Data is a powerful thing. It can be addictive, making you overanalyze everything. But much of what we actually do is unconscious, based on past experience and pragmatism. And with good reason: relying on wisdom and experience, rather than rigid analysis, helps us get through our day. After all, you don't run A/B testing before deciding what pants to put on in the morning; if you did, you'd never get out the door.

One of the criticisms of Lean Startup is that it's too data-driven. Rather than be a slave to the data, these critics say, we should use it as a tool. We should be data-informed, not data-driven. Mostly, they're just being lazy, and looking for reasons not to do the hard work. But sometimes, they have a point: using data to optimize one part of your business, without stepping back and looking at the big picture, can be dangerous—even fatal.

Consider travel agency Orbitz and its discovery that Mac users were willing to reserve a more expensive hotel room. CTO Roger Liew told the *Wall Street Journal*, “We had the intuition [that Mac users are 40% more likely to book a four-or five-star hotel than PC users and to stay in more expensive rooms], and we were able to confirm it based on the data.”^[14]

On the one hand, an algorithm that ignores seemingly unrelated customer data (in this case, whether visitors were using a Mac) wouldn't have found this opportunity to increase revenues. On the other hand, an algorithm that blindly optimizes based on customer data, regardless of its relationship to the sale, may have unintended consequences—like bad PR. Data-driven machine optimization, when not moderated by human judgment, can cause problems.

Years ago, Gail Ennis, then CMO of analytics giant Omniture, told one of us that users of the company's content optimization tools had to temper machine optimization with human judgment. Left to its own devices, the software quickly

learned that scantily clad women generated a far higher click-through rate on web pages than other forms of content. But that click-through rate was a short-term gain, offset by damage to the brand of the company that relied on it. So Omniture's software works alongside curators who understand the bigger picture and provide suitable imagery for the machine to test. *Humans do inspiration; machines do validation.*

In mathematics, a local maximum is the largest value of a function within a given neighborhood.^[15] That doesn't mean it's the largest *possible* value, just the largest one in a particular range. As an analogy, consider a lake on a mountainside. The water isn't at its lowest possible level—that would be sea level—but it's at the lowest possible level in the area surrounding the lake.

Optimization is all about finding the lowest or highest values of a particular function. A machine can find the optimal settings for something, but only within the constraints and problem space of which it's aware, in much the same way that the water in a mountainside lake can't find the lowest possible value, just the lowest value within the constraints provided.

To understand the problem with constrained optimization, imagine that you're given three wheels and asked to evolve the best, most stable vehicle. After many iterations of pitting different wheel layouts against one another, you come up with a tricycle-like configuration. It's the optimal three-wheeled configuration.

Data-driven optimization can perform this kind of iterative improvement. What it can't do, however, is say, "You know what? Four wheels would be way better!" Math is good at optimizing a known system; humans are good at finding a new one. Put another way, *change favors local maxima; innovation favors global disruption.*

In his book *River Out Of Eden* (Basic Books), Richard Dawkins uses the analogy of a flowing river to describe evolution. Evolution, he explains, can create the eye. In fact, it can create dozens of versions of it, for wasps, octopods, humans, eagles, and whales. What it can't do well is go *backward*: once you have an eye that's useful, slight mutations don't usually yield improvements. A human won't evolve an eagle's eye, because the intermediate steps all result in bad eyesight.

Machine-only optimization suffers from similar limitations as evolution. If you're optimizing for local maxima, you might be missing a bigger, more important opportunity. It's your job to be the intelligent designer to data's

evolution.

Many of the startup founders with whom we've spoken have a fundamental mistrust of leaving their businesses to numbers alone. They want to trust their guts. They're uneasy with their companies being optimized without a soul, and see the need to look at the bigger picture of the market, the problem they're solving, and their fundamental business models.

Ultimately, quantitative data is great for testing hypotheses, but it's lousy for generating new ones unless combined with human introspection.

How to Think Like a Data Scientist

Monica Rogati, a data scientist at LinkedIn, gave us the following 10 common pitfalls that entrepreneurs should avoid as they dig into the data their startups capture.

1. **Assuming the data is clean.** Cleaning the data you capture is often most of the work, and the simple act of cleaning it up can often reveal important patterns. "Is an instrumentation bug causing 30% of your numbers to be null?" asks Monica. "Do you really have that many users in the 90210 zip code?" Check your data at the door to be sure it's valid and useful.
2. **Not normalizing.** Let's say you're making a list of popular wedding destinations. You could count the number of people flying in for a wedding, but unless you consider the total number of air travellers coming to that city as well, you'll just get a list of cities with busy airports.
3. **Excluding outliers.** Those 21 people using your product more than a thousand times a day are either your biggest fans, or bots crawling your site for content. Whichever they are, ignoring them would be a mistake.
4. **Including outliers.** While those 21 people using your product a thousand times a day are interesting from a *qualitative* perspective, because they can show you things you didn't expect, they're not good for building a general model. "You probably want to exclude them when building data products," cautions Monica. "Otherwise, the 'you may also like' feature on your site will have the same items everywhere—the ones your hardcore fans

wanted.”

5. **Ignoring seasonality.** “Whoa, is ‘intern’ the fastest-growing job of the year? Oh, wait, it’s June.” Failure to consider time of day, day of week, and monthly changes when looking at patterns leads to bad decision making.
6. **Ignoring size when reporting growth.** Context is critical. Or, as Monica puts it, “When you’ve just started, technically, your dad signing up does count as doubling your user base.”
7. **Data vomit.** A dashboard isn’t much use if you don’t know where to look.
8. **Metrics that cry wolf.** You want to be responsive, so you set up alerts to let you know when something is awry in order to fix it quickly. But if your thresholds are too sensitive, they get “whiny”—and you’ll start to ignore them.
9. **The “Not Collected Here” syndrome.** “Mashing up your data with data from other sources can lead to valuable insights,” says Monica. “Do your best customers come from zip codes with a high concentration of sushi restaurants?” This might give you a few great ideas about what experiments to run next—or even influence your growth strategy.
10. **Focusing on noise.** “We’re hardwired (and then programmed) to see patterns where there are none,” Monica warns. “It helps to set aside the vanity metrics, step back, and look at the bigger picture.”

Lean Startup and Big Vision

Some entrepreneurs are maniacally, almost compulsively, data-obsessed, but tend to get mired in analysis paralysis. Others are casual, shoot-from-the-hip intuitionists who ignore data unless it suits them, and pivot lazily from idea to idea without discipline. At the root of this divide is the fundamental challenge that Lean Startup advocates face: how do you have a minimum viable product and a hugely compelling vision at the same time?

Plenty of founders use Lean Startup as an excuse to start a company without a vision. “It’s so easy to start a company these days.” They reason, “the barriers are so low that everyone can do it, right?” Yet having a big vision is important: starting a company without one makes you susceptible to outside influences, be they from customers, investors, competition, press, or anything else. Without a big vision, you’ll lack purpose, and over time you’ll find yourself wandering aimlessly.

So if a big, hairy, audacious vision is important—one with a changing-the-world type goal—how does that reconcile with the step-by-step, always-questioning approach of Lean Startup?

The answer is actually pretty simple. You need to think of Lean Startup as the process you use to move toward and achieve your vision.

We sometimes remind early-stage founders that, in many ways, they aren’t building a product. *They’re building a tool to learn what product to build.* This helps separate the task at hand—finding a sustainable business model—from the screens, lines of code, and mailing lists they’ve carefully built along the way.

Lean Startup is focused on learning above everything else, and encourages broad thinking, exploration, and experimentation. It’s not about mindlessly going through the motions of *build>measure>learn*—it’s about really understanding what’s going on and being open to new possibilities.

Be Lean. Don’t be small. We’ve talked to founders who want to be the leading provider in their state or province. Why not the world? Even the Allies had to pick a beachhead, but landing in Normandy didn’t mean they lacked a big vision. They just found a good place to start.

Some people believe Lean Startup encourages that smallness, but in fact, used properly, Lean Startup helps *expand* your vision, because you’re encouraged to question everything. As you dig deeper and peel away more layers of what you’re doing—whether you’re looking at problems, solutions, customers, revenue, or anything else—you’re likely to find a lot more than you expected. If you’re opportunistic about it, you can expand your vision and understand how to get there faster, all at the same time.

[14] <http://online.wsj.com/article/SB10001424052702304458604577488822667325882.html>