

## Chapter 2. How to Keep Score

---

Analytics is about tracking the metrics that are critical to your business. Usually, those metrics matter because they relate to your business model—where money comes from, how much things cost, how many customers you have, and the effectiveness of your customer acquisition strategies.

In a startup, you don't always know which metrics are key, because you're not entirely sure what business you're in. You're frequently changing the activity you analyze. You're still trying to find the right product, or the right target audience. In a startup, the purpose of analytics is *to find your way to the right product and market before the money runs out*.

### What Makes a Good Metric?

Here are some rules of thumb for what makes a good metric—a number that will drive the changes you're looking for.

**A good metric is comparative.** Being able to compare a metric to other time periods, groups of users, or competitors helps you understand which way things are moving. “Increased conversion from last week” is more meaningful than “2% conversion.”

**A good metric is understandable.** If people can't remember it and discuss it, it's much harder to turn a change in the data into a change in the culture.

**A good metric is a ratio or a rate.** Accountants and financial analysts have several ratios they look at to understand, at a glance, the fundamental health of a company.<sup>[5]</sup> You need some, too.

There are several reasons ratios tend to be the best metrics:

- **Ratios are easier to act on.** Think about driving a car. Distance travelled is informational. But speed—distance per hour—is something you can act on, because it tells you about your current state, and whether you need to go

faster or slower to get to your destination on time.

- **Ratios are inherently comparative.** If you compare a daily metric to the same metric over a month, you'll see whether you're looking at a sudden spike or a long-term trend. In a car, speed is one metric, but speed right now over average speed this hour shows you a lot about whether you're accelerating or slowing down.
- **Ratios are also good for comparing factors that are somehow opposed, or for which there's an inherent tension.** In a car, this might be distance covered divided by traffic tickets. The faster you drive, the more distance you cover—but the more tickets you get. This ratio might suggest whether or not you should be breaking the speed limit.

Leaving our car analogy for a moment, consider a startup with free and paid versions of its software. The company has a choice to make: offer a rich set of features for free to acquire new users, or reserve those features for paying customers, so they will spend money to unlock them. Having a full-featured free product might reduce sales, but having a crippled product might reduce new users. You need a metric that combines the two, so you can understand how changes affect overall health. Otherwise, you might do something that increases sales revenue at the expense of growth.

**A good metric changes the way you behave.** This is by far the most important criterion for a metric: what will you do differently based on changes in the metric?

- “Accounting” metrics like daily sales revenue, when entered into your spreadsheet, need to make your predictions more accurate. These metrics form the basis of Lean Startup’s *innovation accounting*, showing you how close you are to an ideal model and whether your actual results are converging on your business plan.
- “Experimental” metrics, like the results of a test, help you to optimize the product, pricing, or market. Changes in these metrics will significantly change your behavior. Agree on what that change will be before you collect the data: if the pink website generates more revenue than the alternative, you’re going pink; if more than half your respondents say they won’t pay for

a feature, don't build it; if your curated MVP doesn't increase order size by 30%, try something else.

Drawing a line in the sand is a great way to enforce a disciplined approach. A good metric changes the way you behave precisely *because* it's aligned to your goals of keeping users, encouraging word of mouth, acquiring customers efficiently, or generating revenue.

Unfortunately, that's not always how it happens.

Renowned author, entrepreneur, and public speaker Seth Godin cites several examples of this in a blog post entitled "Avoiding false metrics."<sup>[6]</sup> Funnily enough (or maybe not!), one of Seth's examples, which involves car salespeople, recently happened to Ben.

While finalizing the paperwork for his new car, the dealer said to Ben, "You'll get a call in the next week or so. They'll want to know about your experience at the dealership. It's a quick thing, won't take you more than a minute or two. It's on a scale from 1 to 5. You'll give us a 5, right? Nothing in the experience would warrant less, right? If so, I'm very, very sorry, but a 5 would be great."

Ben didn't give it a lot of thought (and strangely, no one ever did call). Seth would call this a *false metric*, because the car salesman spent more time asking for a good rating (which was clearly important to him) than he did providing a great experience, which was supposedly what the rating was for in the first place.

Misguided sales teams do this too. At one company, Alistair saw a sales executive tie quarterly compensation to the number of deals in the pipeline, rather than to the number of deals closed, or to margin on those sales.

Salespeople are coin-operated, so they did what they always do: they followed the money. In this case, that meant a glut of junk leads that took two quarters to clean out of the pipeline—time that would have been far better spent closing qualified prospects.

Of course, customer satisfaction or pipeline flow is vital to a successful business. But if you want to change behavior, your metric must be tied to the behavioral change you want. If you measure something and it's not attached to a goal, in turn changing your behavior, you're wasting your time. Worse, you may be lying to yourself and fooling yourself into believing that everything is OK.

That's no way to succeed.

One other thing you'll notice about metrics is that they often come in pairs.

*Conversion rate* (the percentage of people who buy something) is tied to *time-to-purchase* (how long it takes someone to buy something). Together, they tell you a lot about your cash flow. Similarly, *viral coefficient* (the number of people a user successfully invites to your service) and *viral cycle time* (how long it takes them to invite others) drive your adoption rate. As you start to explore the numbers that underpin your business, you'll notice these pairs. Behind them lurks a fundamental metric like revenue, cash flow, or user adoption.

If you want to choose the right metrics, you need to keep five things in mind:

### ***Qualitative versus quantitative metrics***

Qualitative metrics are unstructured, anecdotal, revealing, and hard to aggregate; quantitative metrics involve numbers and statistics, and provide hard numbers but less insight.

### ***Vanity versus actionable metrics***

Vanity metrics might make you feel good, but they don't change how you act. Actionable metrics change your behavior by helping you pick a course of action.

### ***Exploratory versus reporting metrics***

Exploratory metrics are speculative and try to find unknown insights to give you the upper hand, while reporting metrics keep you abreast of normal, managerial, day-to-day operations.

### ***Leading versus lagging metrics***

Leading metrics give you a predictive understanding of the future; lagging metrics explain the past. Leading metrics are better because you still have time to act on them—the horse hasn't left the barn yet.

### ***Correlated versus causal metrics***

If two metrics change together, they're correlated, but if one metric *causes*

another metric to change, they're causal. If you find a causal relationship between something you want (like revenue) and something you can control (like which ad you show), then you can change the future.

Analysts look at specific metrics that drive the business, called *key performance indicators* (KPIs). Every industry has KPIs—if you're a restaurant owner, it's the number of covers (tables) in a night; if you're an investor, it's the return on an investment; if you're a media website, it's ad clicks; and so on.

## Qualitative Versus Quantitative Metrics

Quantitative data is easy to understand. It's the numbers we track and measure—for example, sports scores and movie ratings. As soon as something is ranked, counted, or put on a scale, it's quantified. Quantitative data is nice and scientific, and (assuming you do the math right) you can aggregate it, extrapolate it, and put it into a spreadsheet. But it's seldom enough to get a business started. You can't walk up to people, ask them what problems they're facing, and get a quantitative answer. For that, you need qualitative input.

Qualitative data is messy, subjective, and imprecise. It's the stuff of interviews and debates. It's hard to quantify. You can't measure qualitative data easily. If quantitative data answers “what” and “how much,” qualitative data answers “why.” *Quantitative data abhors emotion; qualitative data marinates in it.*

Initially, you're looking for qualitative data. You're not measuring results numerically. Instead, you're speaking to people—specifically, to people you think are potential customers in the right target market. You're exploring. *You're getting out of the building.*

Collecting good qualitative data takes preparation. You need to ask specific questions without leading potential customers or skewing their answers. You have to avoid letting your enthusiasm and reality distortion rub off on your interview subjects. Unprepared interviews yield misleading or meaningless results.

## Vanity Versus Real Metrics

Many companies claim they're data-driven. Unfortunately, while they embrace the *data* part of that mantra, few focus on the second word: *driven*. If you have a piece of data on which you cannot act, it's a vanity metric. If all it does is stroke your ego, it won't help. You want your data to inform, to guide, to improve your business model, to help you decide on a course of action.

Whenever you look at a metric, ask yourself, "What will I do differently based on this information?" If you can't answer that question, you probably shouldn't worry about the metric too much. And if you don't know which metrics *would* change your organization's behavior, you aren't being data-driven. You're floundering in data quicksand.

Consider, for example, "total signups." This is a vanity metric. The number can only increase over time (a classic "up and to the right" graph). It tells us nothing about what those users are doing or whether they're valuable to us. They may have signed up for the application and vanished forever.

"Total active users" is a bit better—assuming that you've done a decent job of defining an active user—but it's still a vanity metric. It will gradually increase over time, too, unless you do something horribly wrong.

The real metric of interest—the *actionable* one—is "percent of users who are active." This is a critical metric because it tells us about the level of engagement your users have with your product. When you change something about the product, this metric should change, and if you change it in a good way, it should go up. That means you can experiment, learn, and iterate with it.

Another interesting metric to look at is "number of users acquired over a specific time period." Often, this will help you compare different marketing approaches—for example, a Facebook campaign in the first week, a reddit campaign in the second, a Google AdWords campaign in the third, and a LinkedIn campaign in the fourth. Segmenting experiments by time in this way isn't precise, but it's relatively easy.<sup>[2]</sup> And it's actionable: if Facebook works better than LinkedIn, you know where to spend your money.

Actionable metrics aren't magic. They won't tell you what to do—in the previous example, you could try changing your pricing, or your medium, or your wording. The point here is that you're doing *something* based on the data you collect.

## Eight Vanity Metrics to Watch Out For

It's easy to fall in love with numbers that go up and to the right. Here's a list of eight notorious vanity metrics you should avoid.

1. **Number of hits.** This is a metric from the early, foolish days of the Web. If you have a site with many objects on it, this will be a big number. Count people instead.
2. **Number of page views.** This is only slightly better than hits, since it counts the number of times someone requests a page. Unless your business model depends on page views (i.e., display advertising inventory), you should count people instead.
3. **Number of visits.** Is this one person who visits a hundred times, or are a hundred people visiting once? Fail.
4. **Number of unique visitors.** All this shows you is how many people saw your home page. It tells you nothing about what they did, why they stuck around, or if they left.
5. **Number of followers/friends/likes.** Counting followers and friends is nothing more than a popularity contest, unless you can get them to do something useful for you. Once you know how many followers will do your bidding when asked, you've got something.
6. **Time on site/number of pages.** These are a poor substitute for actual engagement or activity unless your business is tied to this behavior. If customers spend a lot of time on your support or complaints pages, that's probably a bad thing.
7. **Emails collected.** A big mailing list of people excited about your new startup is nice, but until you know how many will open your emails (and act on what's inside them), this isn't useful. Send test emails to some of your registered subscribers and see if they'll do what you tell them.
8. **Number of downloads.** While it sometimes affects your ranking in app stores, downloads alone don't lead to real value. Measure activations,



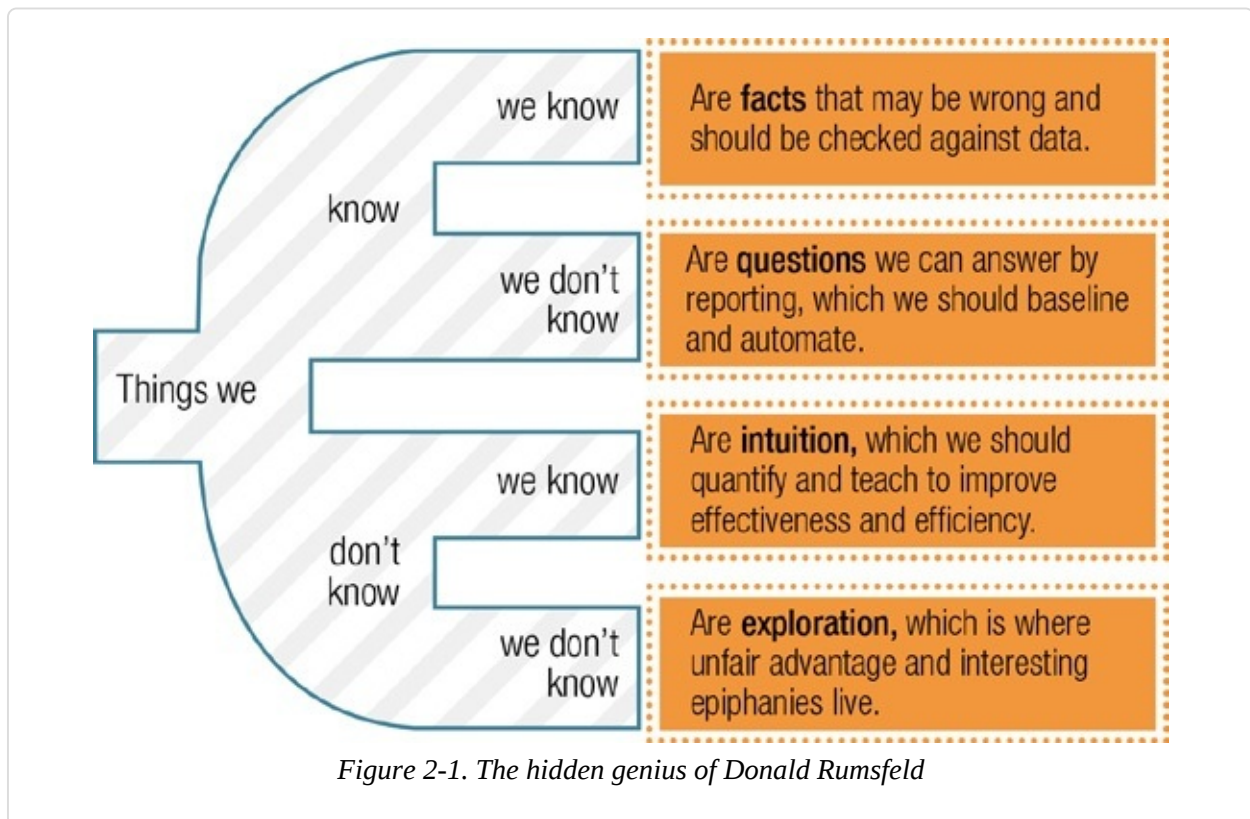
account creations, or something else.

## Exploratory Versus Reporting Metrics

Avinash Kaushik, author and Digital Marketing Evangelist at Google, says former US Secretary of Defense Donald Rumsfeld knew a thing or two about analytics. According to Rumsfeld:

*There are known knowns; there are things we know that we know. There are known unknowns; that is to say there are things that we now know we don't know. But there are also unknown unknowns—there are things we do not know, we don't know.*

Figure 2-1 shows these four kinds of information.



The “known unknowns” is a reporting posture—counting money, or users, or lines of code. We *know* we don’t know the value of the metric, so we go find out. We may use these metrics for accounting (“How many widgets did we sell today?”) or to measure the outcome of an experiment (“Did the green or the red



widget sell more?”), but in both cases, we know the metric is needed.

The “unknown unknowns” are most relevant to startups: exploring to discover something new that will help you disrupt a market. As we’ll see in the next case study, it’s how Circle of Friends found out that moms were its best users. These “unknown unknowns” are where the magic lives. They lead down plenty of wrong paths, and hopefully toward some kind of “eureka!” moment when the idea falls into place. This fits what Steve Blank says a startup should spend its time doing: searching for a scalable, repeatable business model.

Analytics has a role to play in all four of Rumsfeld’s quadrants:

- It can check our facts and assumptions—such as open rates or conversion rates—to be sure we’re not kidding ourselves, and check that our business plans are accurate.
- It can test our intuitions, turning hypotheses into evidence.
- It can provide the data for our spreadsheets, waterfall charts, and board meetings.
- It can help us find the nugget of opportunity on which to build a business.

In the early stages of your startup, the unknown unknowns matter most, because they can become your secret weapons.

## Circle of Moms Explores Its Way to Success

Circle of Friends was a simple idea: a Facebook application that allowed you to organize your friends into circles for targeted content sharing. Mike Greenfield and his co-founders started the company in September 2007, shortly after Facebook launched its developer platform. The timing was perfect: Facebook became an open, viral place to acquire users as quickly as possible and build a startup. There had never been a platform with so many users and that was so open (Facebook had about 50 million users at the time).

By mid-2008, Circle of Friends had 10 million users. Mike focused on growth above everything else. “It was a land grab,” he says, and Circle of Friends was clearly viral. But there was a problem. Too few people were actually *using* the

product.

According to Mike, less than 20% of circles had any activity whatsoever after their initial creation. “We had a few million monthly uniques from those 10 million users, but as a general social network we knew that wasn’t good enough and monetization would likely be poor.”

So Mike went digging.

He started looking through the database of users and what they were doing. The company didn’t have an in-depth analytical dashboard at the time, but Mike could still do some exploratory analysis. And he found a segment of users—moms, to be precise—that bucked the poor engagement trend of most users. Here’s what he found:

- Their messages to one another were on average 50% longer.
- They were 115% more likely to attach a picture to a post they wrote.
- They were 110% more likely to engage in a threaded (i.e., deep) conversation.
- They had friends who, once invited, were 50% more likely to become engaged users themselves.
- They were 75% more likely to click on Facebook notifications.
- They were 180% more likely to click on Facebook news feed items.
- They were 60% more likely to accept invitations to the app.

The numbers were so compelling that in June 2008, Mike and his team switched focus completely. They pivoted. And in October 2008, they launched Circle of Moms on Facebook.

Initially, numbers dropped as a result of the new focus, but by 2009, the team grew its community to 4.5 million users—and unlike the users who’d been lost in the change, these were actively engaged. The company went through some ups and downs after that, as Facebook limited applications’ abilities to spread virally. Ultimately, the company moved off Facebook, grew independently, and sold to Sugar Inc. in early 2012.

## Summary

- Circle of Friends was a social graph application in the right place at the right time—with the wrong market.
- By analyzing patterns of engagement and desirable behavior, then finding out what those users had in common, the company found the right market for its offering.
- Once the company had found its target, it focused—all the way to changing its name. Pivot hard or go home, and be prepared to burn some bridges.

## Analytics Lessons Learned

The key to Mike's success with Circle of Moms was his ability to dig into the data and look for meaningful patterns and opportunities. Mike discovered an “unknown unknown” that led to a big, scary, gutsy bet (drop the generalized Circle of Friends to focus on a specific niche) that was a gamble—but one that was based on data.

There's a “critical mass” of engagement necessary for any community to take off. Mild success may not give you escape velocity. As a result, it's better to have fervent engagement with a smaller, more easily addressable target market. Virality requires focus.

## Leading Versus Lagging Metrics

Both leading and lagging metrics are useful, but they serve different purposes.

A leading metric (sometimes called a *leading indicator*) tries to predict the future. For example, the current number of prospects in your sales funnel gives you a sense of how many new customers you'll acquire in the future. If the current number of prospects is very small, you're not likely to add many new customers. You can increase the number of prospects and expect an increase in new customers.

On the other hand, a lagging metric, such as *churn* (which is the number of customers who leave in a given time period) gives you an indication that there's a problem—but by the time you're able to collect the data and identify the

problem, it's too late. The customers who churned out aren't coming back. That doesn't mean you can't act on a lagging metric (i.e., work to improve churn and then measure it again), but it's akin to closing the barn door after the horses have left. New horses won't leave, but you've already lost a few.

In the early days of your startup, you won't have enough data to know how a current metric relates to one down the road, so measure lagging metrics at first. Lagging metrics are still useful and can provide a solid baseline of performance. For leading indicators to work, you need to be able to do cohort analysis and compare groups of customers over periods of time.

Consider, for example, the volume of customer complaints. You might track the number of support calls that happen in a day—once you've got a call volume to make that useful. Earlier on, you might track the number of customer complaints in a 90-day period. Both could be leading indicators of churn: if complaints are increasing, it's likely that more customers will stop using your product or service. As a leading indicator, customer complaints also give you ammunition to dig into what's going on, figure out why customers are complaining more, and address those issues.

Now consider account cancellation or product returns. Both are important metrics—but they measure after the fact. They pinpoint problems, but only after it's too late to avert the loss of a customer. Churn is important (and we discuss it at length throughout the book), but looking at it myopically won't let you iterate and adapt at the speed you need.

Indicators are everywhere. In an enterprise software company, quarterly new product bookings are a lagging metric of sales success. By contrast, new qualified leads are a leading indicator, because they let you predict sales success ahead of time. But as anyone who's ever worked in B2B (business-to-business) sales will tell you, in addition to qualified leads you need a good understanding of conversion rate and sales-cycle length. Only then can you make a realistic estimate of how much new business you'll book.

In some cases, a lagging metric for one group within a company is a leading metric for another. For example, we know that the number of quarterly bookings is a lagging metric for salespeople (the contracts are signed already), but for the finance department that's focused on collecting payment, they're a leading indicator of expected revenue (since the revenue hasn't yet been realized).

Ultimately, you need to decide whether the thing you're tracking helps you make better decisions sooner. As we've said, a real metric has to be actionable. Lagging and leading metrics can both be actionable, but leading indicators show you what *will* happen, reducing your cycle time and making you leaner.

## Correlated Versus Causal Metrics

In Canada, the use of winter tires is correlated with a decrease in accidents. People put softer winter tires on their cars in cold weather, and there are more accidents in the summer.<sup>[8]</sup> Does that mean we should make drivers use winter tires year-round? Almost certainly not—softer tires stop poorly on warm summer roads, and accidents would increase.

Other factors, such as the number of hours driven and summer vacations, are likely responsible for the increased accident rates. But looking at a simple correlation without demanding causality leads to some bad decisions. There's a correlation between ice cream consumption and drowning. Does that mean we should ban ice cream to avert drowning deaths? Or measure ice cream consumption to predict the fortunes of funeral home stock prices? No: ice cream and drowning rates both happen *because of* summer weather.

Finding a correlation between two metrics is a good thing. Correlations can help you predict what will happen. But finding the *cause* of something means you can change it. Usually, causations aren't simple one-to-one relationships. Many factors conspire to cause something. In the case of summertime car crashes, we have to consider alcohol consumption, the number of inexperienced drivers on the road, the greater number of daylight hours, summer vacations, and so on. So you'll seldom get a 100% causal relationship. You'll get several independent metrics, each of which "explains" a portion of the behavior of the dependent metric. But even a degree of causality is valuable.

You prove causality by finding a correlation, then running an experiment in which you control the other variables and measure the difference. This is hard to do because no two users are identical; it's often impossible to subject a statistically significant number of people to a properly controlled experiment in the real world.

If you have a big enough sample of users, you can run a reliable test without

controlling all the other variables, because eventually the impact of the other variables is relatively unimportant. That's why Google can test subtle factors like the color of a hyperlink,<sup>[9]</sup> and why Microsoft knows exactly what effect a slower page load time has on search rates.<sup>[10]</sup> But for the average startup, you'll need to run simpler tests that experiment with only a few things, and then compare how that changed the business.

We'll look at different kinds of testing and segmentation shortly, but for now, recognize this: correlation is good. Causality is great. Sometimes, you may have to settle for the former—but you should always be trying to discover the latter.

## Moving Targets

When picking a goal early on, you're drawing a line in the sand—not carving it in stone. You're chasing a moving target, because you really don't know how to define success.

Adjusting your goals and how you define your key metrics is acceptable, provided that you're being honest with yourself, recognizing the change this means for your business, and not just lowering expectations so that you can keep going in spite of the evidence.

When your initial offering—your minimum viable product—is in the market and you're acquiring early-adopter customers and testing their use of your product, you won't even know how they're going to use it (although you'll have assumptions). Sometimes there's a huge gulf between what you assume and what users actually do. You might think that people will play your multiplayer game, only to discover that they're using you as a photo upload service. Unlikely? That's how Flickr got started.

Sometimes, however, the differences are subtler. You might assume your product has to be used daily to succeed, only to find out that's not so. In these situations, it's reasonable to update your metrics accordingly, provided that you're able to prove the value created.

## HighScore House Defines an “Active User”

HighScore House started as a simple application that allowed parents to list

chores and challenges for their children with point values. Kids could complete the tasks, collect points, and redeem the points for rewards they wanted.

When HighScore House launched its MVP, the company had several hundred families ready to test it. The founders drew a line in the sand: in order for the MVP to be considered successful, parents and kids would have to each use the application four times per week. These families would be considered “active.” It was a high, but good, bar.

After a month or so, the percentage of active families was lower than this line in the sand. The founders were disappointed but determined to keep experimenting in an effort to improve engagement:

- They modified the sign-up flow (making it clearer and more educational to increase quality signups and to improve onboarding).
- They sent email notifications as daily reminders to parents.
- They sent transactional emails to parents based on actions their kids took in the system.

There was an incremental improvement each time, but nothing that moved the needle significantly enough to say that the MVP was a success.

Then co-founder and CEO Kyle Seaman did something critical: *he picked up the phone*. Kyle spoke with dozens of parents. He started calling parents who had signed up, but who weren’t active. First he reached out to those that had abandoned HighScore House completely (“churned out”). For many of them, the application wasn’t solving a big enough pain point. That’s fine. The founders never assumed the market was “all parents”—that’s just too broad a definition, particularly for a first version of a product. Kyle was looking for a smaller subset of families where HighScore House would resonate, to narrow the market segment and focus.

Kyle then called those families who *were* using HighScore House, but not using it enough to be defined as active. Many of these families responded positively: “We’re using HighScore House. It’s great. The kids are making their beds consistently for the first time ever!”

The response from parents was a surprise. Many of them were using HighScore



House only once or twice a week, but they were getting value out of the product. From this, Kyle learned about segmentation and which types of families were more or less interested in what the company was offering. He began to understand that the initial baseline of usage the team had set wasn't consistent with how engaged customers were using the product.

That doesn't mean the team shouldn't have taken a guess. Without that initial line in the sand, they would have had no benchmark for learning, and Kyle might not have picked up the phone. But now he really understood his customers. The combination of quantitative and qualitative data was key.

As a result of this learning, the team redefined the "active user" threshold to more accurately reflect existing users' behavior. It was okay for them to adjust a key metric because they truly understood why they were doing it and could justify the change.

### **Summary**

- HighScore House drew an early, audacious line in the sand—which it couldn't hit.
- The team experimented quickly to improve the number of active users but couldn't move the needle enough.
- They picked up the phone and spoke to customers, realizing that they were creating value for a segment of users with lower usage metrics.

### **Analytics Lessons Learned**

First, know your customer. There's no substitute for engaging with customers and users directly. All the numbers in the world can't explain why something is happening. Pick up the phone right now and call a customer, even one who's disengaged.

Second, make early assumptions and set targets for what you think success looks like, but don't experiment yourself into oblivion. Lower the bar if necessary, but not for the sake of getting over it: that's just cheating. Use qualitative data to understand what value you're creating and adjust only if the new line in the sand reflects how customers (in specific segments) are using your product.

# Segments, Cohorts, A/B Testing, and Multivariate Analysis

Testing is at the heart of Lean Analytics. Testing usually involves comparing two things against each other through segmentation, cohort analysis, or A/B testing. These are important concepts for anyone trying to perform the kind of scientific comparison needed to justify a change, so we'll explain them in some detail here.

## Segmentation

A segment is simply a group that shares some common characteristic. It might be users who run Firefox, or restaurant patrons who make reservations rather than walking in, or passengers who buy first-class tickets, or parents who drive minivans.

On websites, you segment visitors according to a range of technical and demographic information, then compare one segment to another. If visitors using the Firefox browser have significantly fewer purchases, do additional testing to find out why. If a disproportionate number of engaged users are coming from Australia, survey them to discover why, and then try to replicate that success in other markets.

Segmentation works for any industry and any form of marketing, not just for websites. Direct mail marketers have been segmenting for decades with great success.

## Cohort Analysis

A second kind of analysis, which compares similar groups over time, is cohort analysis. As you build and test your product, you'll iterate constantly. Users who join you in the first week will have a different experience from those who join later on. For example, all of your users might go through an initial free trial, usage, payment, and abandonment cycle. As this happens, you'll make changes to your business model. The users who experienced the trial in month one will have a different onboarding experience from those who experience it in month five. How did that affect their churn? To find out, we use cohort analysis.

Each group of users is a cohort—participants in an experiment across their

lifecycle. You can compare cohorts against one another to see if, on the whole, key metrics are getting better over time. Here's an example of why cohort analysis is critical for startups.

Imagine that you're running an online retailer. Each month, you acquire a thousand new customers, and they spend some money. **Table 2-1** shows your customers' average revenues from the first five months of the business.

*Table 2-1. Average revenues for five months*

	January	February	March	April	May
<b>Total customers</b>	1,000	2,000	3,000	4,000	5,000
<b>Average revenue per customer</b>	\$5.00	\$4.50	\$4.33	\$4.25	\$4.50

From this table, you can't learn much. Are things getting better or worse? Since you aren't comparing recent customers to older ones—and because you're commingling the purchases of a customer who's been around for five months with those of a brand new one—it's hard to tell. All this data shows is a slight drop in revenues, then a recovery. But average revenue is pretty static.

Now consider the same data, broken out by the month in which that customer group started using the site. As **Table 2-2** shows, something important is going on. Customers who arrived in month five are spending, on average, \$9 in their first month—nearly double that of those who arrived in month one. That's huge growth!

*Table 2-2. Comparing revenues by the month customers arrived*

	January	February	March	April	May
<b>New users</b>	1,000	1,000	1,000	1,000	1,000
<b>Total users</b>	1,000	2,000	3,000	4,000	5,000
<b>Month 1</b>	\$5.00	\$3.00	\$2.00	\$1.00	\$0.50
<b>Month 2</b>		\$6.00	\$4.00	\$2.00	\$1.00
<b>Month 3</b>			\$7.00	\$6.00	\$5.00

<b>Month 4</b>				\$8.00	\$7.00
<b>Month 5</b>					\$9.00

Another way to understand cohorts is to line up the data by the users' experience—in the case of **Table 2-3**, we've done this by the number of months they've used the system. This shows another critical metric: how quickly revenue declines after the first month.

*Table 2-3. Cohort analysis of revenue data*

	Month of use				
<b>Cohort</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>January</b>	\$5.00	\$3.00	\$2.00	\$1.00	\$0.50
<b>February</b>	\$6.00	\$4.00	\$2.00	\$1.00	
<b>March</b>	\$7.00	\$6.00	\$5.00		
<b>April</b>	\$8.00	\$7.00			
<b>May</b>	\$9.00				
<b>Averages</b>	\$7.00	\$5.00	\$3.00	\$1.00	\$0.50

A cohort analysis presents a *much* clearer perspective. In this example, poor monetization in early months was diluting the overall health of the metrics. The January cohort—the first row—spent \$5 in its first month, then tapered off to only \$0.50 in its fifth month. But first-month spending is growing dramatically, and the drop-off seems better, too: April's cohort spent \$8 in its first month and \$7 in its second month. A company that seemed stalled is in fact flourishing. And you know what metric to focus on: drop-off in sales after the first month.

This kind of reporting allows you to see patterns clearly against the lifecycle of a customer, rather than slicing across all customers blindly without accounting for the natural cycle a customer undergoes. Cohort analysis can be done for revenue, churn, viral word of mouth, support costs, or any other metric you care about.

## A/B and Multivariate Testing

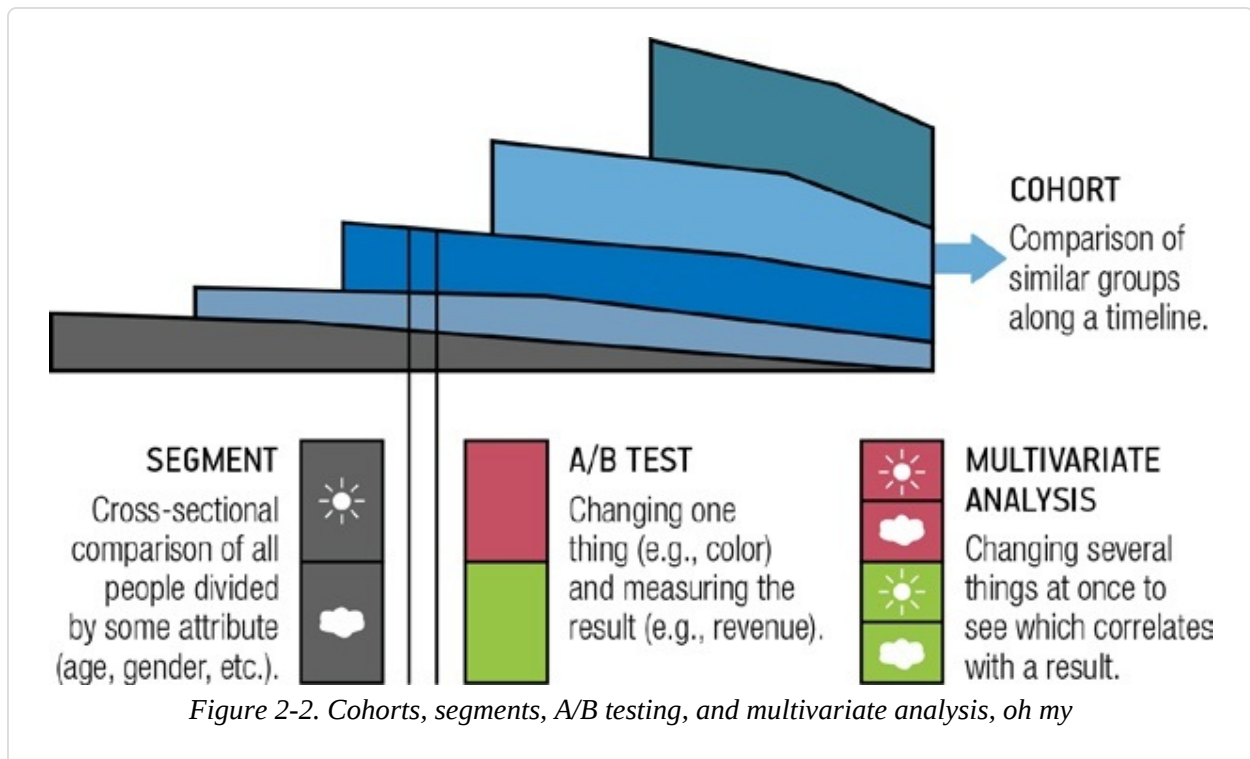
Cohort experiments that compare groups like the one in [Table 2-2](#) are called *longitudinal* studies, since the data is collected *along* the natural lifespan of a customer group. By contrast, studies in which different groups of test subjects are given different experiences at the same time are called *cross-sectional* studies. Showing half of the visitors a blue link and half of them a green link in order to see which group is more likely to click that link is a cross-sectional study. When we're comparing one attribute of a subject's experience, such as link color, and assuming everything else is equal, we're doing A/B testing.

You can test everything about your product, but it's best to focus on the critical steps and assumptions. The results can pay off dramatically: Jay Parmar, co-founder of crowdfunded ticketing site Picatic, told us that simply changing the company's call to action from "Get started free" to "Try it out free" increased the number of people who clicked on an offer—known as the *click-through rate*—by 376% for a 10-day period.

A/B tests seem relatively simple, but they have a problem. Unless you're a huge web property—like Bing or Google—with enough traffic to run a test on a single factor like link color or page speed and get an answer quickly, you'll have more things to test than you have traffic. You might want to test the color of a web page, the text in a call to action, and the picture you're showing to visitors.

Rather than running a series of separate tests one after the other—which will delay your learning cycle—you can analyze them all at once using a technique called *multivariate analysis*. This relies on statistical analysis of the results to see which of many factors correlates strongly with an improvement in a key metric.

[Figure 2-2](#) illustrates these four ways of slicing users into subgroups and analyzing or testing them.



## The Lean Analytics Cycle

Much of Lean Analytics is about finding a meaningful metric, then running experiments to improve it until that metric is good enough for you to move to the next problem or the next stage of your business, as shown in [Figure 2-3](#).

Eventually, you'll find a business model that is sustainable, repeatable, and growing, and learn how to scale it.

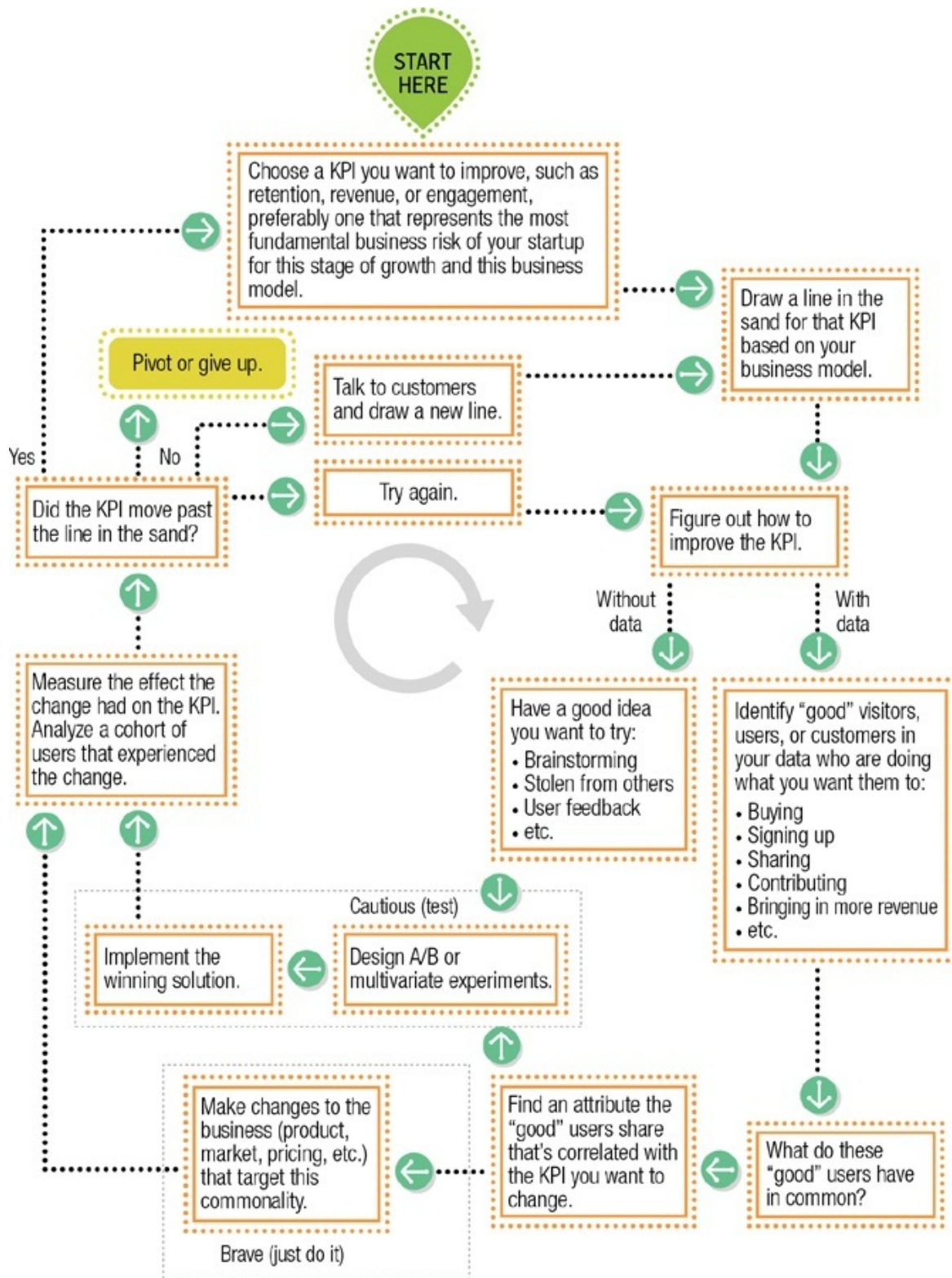


Figure 2-3. The circle of life for analytical startups



We've covered a lot of background on metrics and analytics in this chapter, and your head might be a bit full at this point. You've learned:

- What makes a good metric
- What vanity metrics are and how to avoid them
- The difference between qualitative and quantitative metrics, between exploratory and reporting metrics, between leading and lagging metrics, and between correlated and causal metrics
- What A/B testing is, and why multivariate testing is more common
- The difference between segments and cohorts

In the coming chapters, you'll put all of these dimensions to work on a variety of business models and stages of startup growth.

## Evaluating the Metrics You Track

Take a look at the top three to five metrics that you track religiously and review daily. Write them down. Now answer these questions about them:

- How many of those metrics are good metrics?
- How many do you use to make business decisions, and how many are just vanity metrics?
- Can you eliminate any that aren't adding value?
- Are there others that you're now thinking about that may be more meaningful?

Cross off the bad ones and add new ones to the bottom of your list, and let's keep going through the book.

---

[5] This includes fundamentals such as the price-to-earnings ratio, sales margins, the cost of sales, revenue per employee, and so on.

[6] [http://sethgodin.typepad.com/seths\\_blog/2012/05/avoiding-false-metrics.html](http://sethgodin.typepad.com/seths_blog/2012/05/avoiding-false-metrics.html)

[7] A better way is to run the four campaigns concurrently, using analytics to group the users you acquire into distinct *segments*. You'll get your answer in one week rather than four, and control for other variables like seasonal variation. We'll get into more detail about segmentation and cohort analysis later.

[8] <http://www.statcan.gc.ca/pub/82-003-x/2008003/article/10648/c-g/5202438-eng.htm>

[9] <http://gigaom.com/2009/07/09/when-it-comes-to-links-color-matters/>

[10] <http://velocityconf.com/velocity2009/public/schedule/detail/8523>