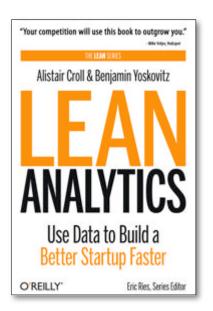
Lean Analytics

Use data to build a better business faster.



@byosko | @acroll

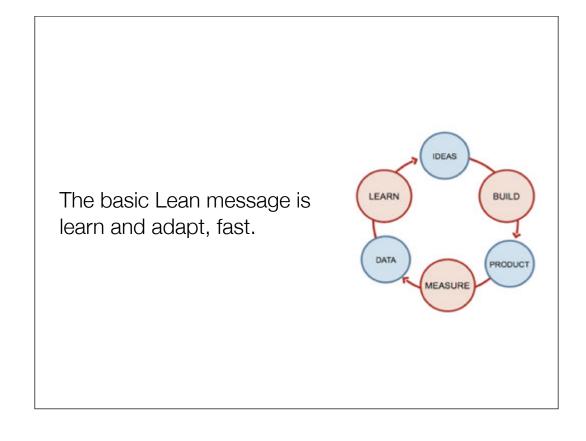
www.leananalyticsbook.com @leananalytics



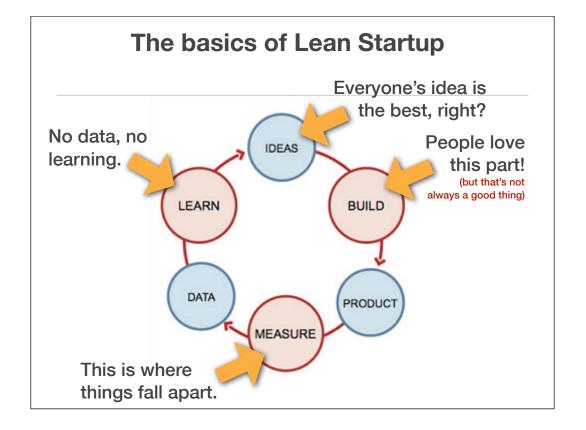
Don't sell what you can make.

Make what you can sell.

You've probably heard plenty about Lean today. So I'll summarize it as this: Don's sell what you can make, make what you can sell. It's a reverse field-of-dreams: If they come, you will build it.



We're all liars.





And you may think you know what analytics is, but you're probably thinking of Web analytics. That's only part of it.

In a startup, the purpose of analytics is to iterate to a product/market fit before the money runs out.

So here's a definition of analytics for startups—what we call Lean Analytics.

Hits	A metric from the early, foolish days of the Web. Count people instead.		
Page views	Marginally better than hits. Unless you're displaying ad inventory, count people.		
Visits	Is this one person visiting a hundred times, or are a hundred people visiting once? Fail.		
Unique visitors	This tells you nothing about what they did, why they stuck around, or if they left.		
Followers/ friends/likes	Count actions instead. Find out how many followers will do your bidding.		
Time on site, or pages/visit	Poor version of engagement. Lots of time spent on support pages is actually a bad sign.		
Emails collected	How many recipients will act on what's in them?		
Number of downloads	Outside app stores, downloads alone don't lead to lifetime value. Measure activations/active accounts.		

- 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.
- Number of page views: 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.
- Number of visits: Is this one person visiting a hundred times, or are a hundred people visiting once? Fail.
- Number of unique visitors: The only thing 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.
- Number of followers/friends/likes: Business can be a popularity contest if you can get those friends to do something, but counting followers rather than actions is a bad idea, particularly since many people follow back automatically. Once you know how many followers will do your bidding when asked, you've got something.
- 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 could be a bad thing.
- Emails collected: A big mailing list of people excited about your new startup is nice, but until you know how many will open your mails (and act on what's inside them) this isn't useful. Test some of them and see.

Number of downloads: While it sometimes affects your place in app stores and rankings, downloads alone don't lead to lifetime value. Measure activations, account creations, or something else.

How to think about metrics.

Leading mber today th

Number today that shows metric tomorrow—makes the news.

Lagging

Historical metric that shows how you're doing—reports the news.

• 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.

Small business example: Solare watches the numbers



Stage: RevenueModel: Retailer

• Solare is an Italian fine-dining restaurant under new management. The new team is trying to identify the key metrics and leading indicators

Here's an example from Solare, a restaurant in San Diego, that underscores just how important analytics are to businesses of any size and shape.



Solare watches the numbers

- A line in the sand: Gross Revenue to Labor Cost
 - Under 30% is good
 - Below 24% is great
 - Lower than 20% and you may be under-staffing, leading to dissatisfied customers
- A leading indicator: Total covers is 5x reservations at 5PM
 - If you have 50 reservations at 5, you'll have 250 covers that night.
 - This ratio varies by restaurant.

What mode of e-commerce are you?				
How many of your customers buy a second time in 90 days?	Then you are in this mode	Your customers will buy from you	You are just like	Focus on
1-15%	Acquisition	Once	70% of retailers	Low CAC, high checkout
15-30%	Hybrid	2-2.5 per year	20% of retailers	Increasing returns
>30%	Loyalty	>2.5 per year	10% of retailers (Thanks to Ke	Loyalty, inventory expansion evin Hilstrom for this.)

Here's a good example of how a leading indicator affects your behavior.

Keen filstorm of Minn That Data works with a number of e-commerce companies. He says it's essential for online retailers to know what kind of relationship they have with their buyers, because this drives everything from marketing strategy to shopping cart size. To understand this, he calculates the Annual Repurchase Rate: what percentage of people who bought something from you last year's lung will do so this year?

• Acquisition mode: If less than 40% of last year's buyers will buy this year, then the focus of the business is on new customer acquisition. Loyalty programs aren't good long-term investments for this kind of business. Kewin says that 70% of e-commerce businesses fall into this category when they're mature. Most catalogers (like Frontgate) are in this mode.

• Hydid and the strategies of the companies of the c

The Annual Repurchase Rate is an early indicator of how an e-commerce startup will succeed in the long term. Even before a year has elapsed, an e-commerce company can look at 90-day repurchase rates and get a sense of which model they're in.

18 to 15% 90-day Repurchase Rate means you're in Hydrind Mode.

18 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

18 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

18 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

18 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

18 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

19 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

19 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

10 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

10 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

10 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

10 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

10 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

11 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

12 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

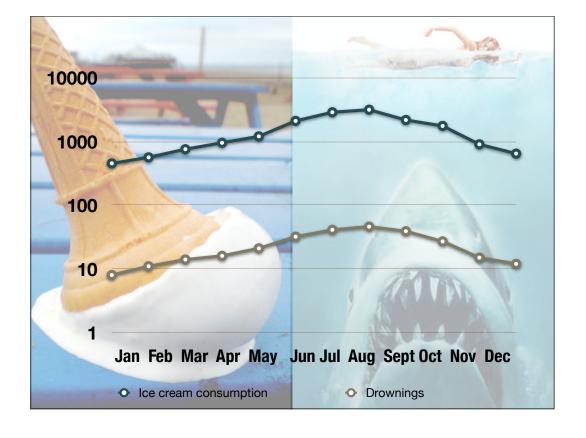
13 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

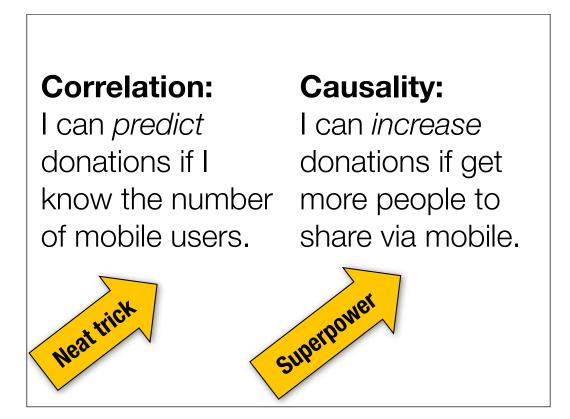
14 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

15 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

16 to 15% 90-day Repurchase Rate means you're in Acquisition Mode.

17 to 15% 90-day Repurchase Rate means you're in Acquisition Mode success) will be a support and the proposed acquisition Mode success and proposed acquisition Mode success) will be a support and a support a





One of the key concepts in Growth Hacking is the leading indicator—something that shows you how something you have right now links to something later on. Here are some examples.

Why causality beats correlation

Correlation lets you predict the future

"I will have 420 engaged users and 75 paying customers next month."

Causality lets you **change** the future

"If I can make more firsttime visitors stay on for 17 minutes I will increase sales in 90 days."

Find correlation

Test causality

Optimize the causal factor

- A Facebook user reaching 7 friends within 10 days of signing up (Chamath Palihapitiya)
- If someone comes back to **Zynga** a day after signing up for a game, they'll probably become an engaged, paying user (Nabeel Hyatt)
- A **Dropbox** user who puts at least one file in one folder on one device (ChenLi Wang)
- Twitter user following a certain number of people, and a certain percentage of those people following the user back (Josh Elman)
- A LinkedIn user getting to X connections in Y days (Elliot Schmukler)

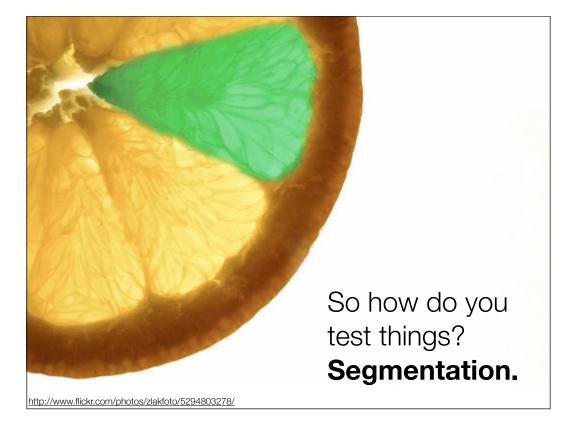
(These are also great segments to analyze.)

(from the 2012 Growth Hacking conference)

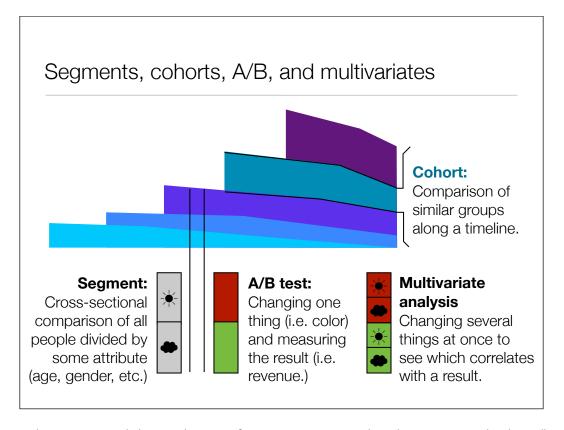
One of the key concepts in Growth Hacking is the leading indicator—something that shows you how something you have right now links to something later on. Here are some examples.



THis is a simple test for whether a metric is real or bullshit.



To measure a change, you need to run an experiment. All experiments consist of a control group and a test group. And you measure the experiment's result by comparing them. This is called segmentation.



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

On websites, we can often segment visitors according to a range of technical and demographic information. Then, by comparing one segment to another, we can learn things. If, for example, visitors using the Firefox browser have significantly fewer purchases, we may want to perform additional testing to find out why. If we see that a disproportionate number of buyers are coming from Australia, we might survey them to discover why, then try to replicate that success in other markets.

Segmentation works for any industry and any form of marketing, not just for websites. As we'll see, segmenting lists is an essential part of improving e-mail performance; but direct mail marketers have been segmenting for decades with great success. Even if you're selling to enterprise customers, dividing your messages by target audience, industry, or stage in the sales cycle—and then testing the results of those messages—is crucial.

Cohort analysis

A second kind of test is cohort analysis. As you build and test your product, you'll be iterating regularly. Users that join you in the first week will have a different experience from those that 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. As a result, the users who experienced the trial in month one may have a significantly different onboarding experience from those who experienced it in month five. How did that affect their churn? To find this 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.

Why use cohorts? Here's an example.						
Is this company		January	February	March	April	May
growing or stagnating?	Rev/customer	CA\$5.00	CA\$4.50	CA\$4.33	CA\$4.25	CA\$4.50
	Cohort	1	2	3	4	5
	January	CA\$5	CA\$3	CA\$2	CA\$1	CA\$0.5
How about	February		CA\$6	CA\$4	CA\$2	CA\$1
now?	March			CA\$7	CA\$6	CA\$5
	April				CA\$8	CA\$7
	May					CA\$9

Using cohorts is really, really important.

Why use cohorts? Here's an example.

Look at the same data in cohorts

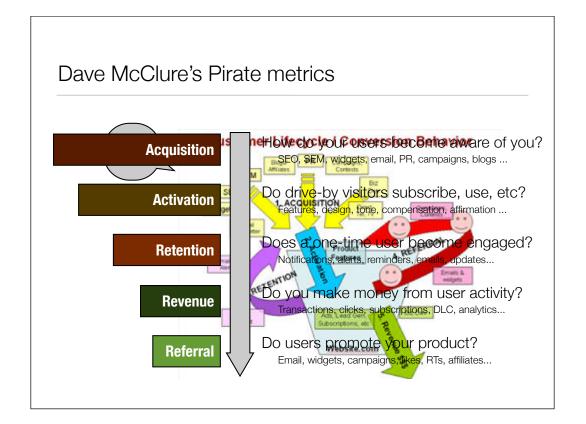
Cohort	1	2	3	4	5
January	CA\$5	CA\$3	CA\$2	CA\$1	CA\$0.5
February	CA\$6	CA\$4	CA\$2	CA\$1	
March	CA\$7	CA\$6	CA\$5		
April	CA\$8	CA\$7			
May	CA\$9				
Averages	CA\$7	CA\$5	CA\$3	CA\$1	CA\$0.5

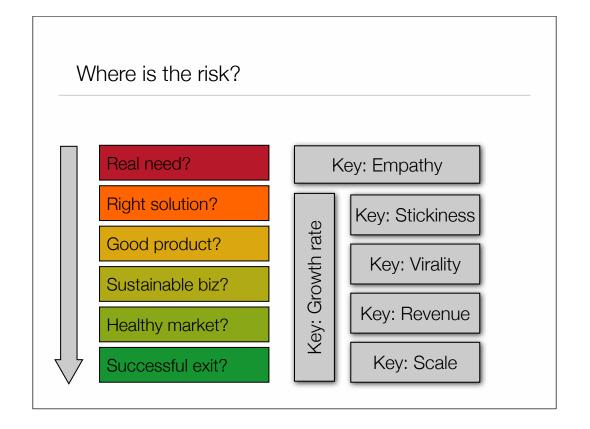
Using cohorts is really, really important.

The Stages of Lean Analytics.



Eric Ries talks about three engines of growth. Our framework suggests that all companies should go through these engines as three distinct stages.







Empathy stage: Localmind hacks Twitter

- Stage: EmpathyModel: UGC/mobile
- Real-time question and answer platform tied to locations.
- Needed to find out if a core behavior—answering questions about a place—happened enough to make the business real

Here's a concrete example of the empathy stage

localmind

Localmind hacks Twitter

- Before writing a line of code, Localmind was concerned that people would never answer questions.
 - This was their biggest risk: if questions went unanswered users would have a terrible experience and stop using Localmind.
- Ran an experiment on Twitter
 - Tracked geolocated tweets in Times Square
 - Sent @ messages to people who had just tweeted, asking questions about the area: how busy is it; is the subway running on time; is something open; etc.
- The response rate to their tweeted questions was very high.
 - Good enough proxy to de-risk the solution, and convince the team and investors that it was worth building Localmind.



Stickiness stage: qidiq streamlines invites

- Stage: Virality
- Model: SaaS
- Tool to poll small groups, built in the Year One Labs accelerator

Here's a concrete example of the virality stage

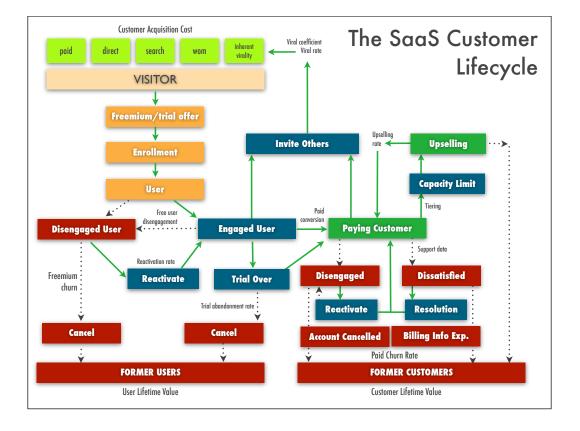
	Initial design		Redesigned workflow
	Survey owner adds recipient to group	<u></u>	Survey owner adds recipient to group
	Survey owner asks question	70-90% RESPONSE RATE	Survey owner asks question
	Recipient gets invite		Recipient reads survey question
E RATE	Recipient installs mobile app		Recipient responds to question
ESPONS	Recipient creates account, profile		Recipient sees survey results
10-25% RESPONSE RATE	Recipient can edit profile, view past questions, etc.		(Later, if needed)
-			Recipient visits website
	Recipient reads survey question		Recipient has no password!
	Recipient responds to question		Recipient does password recovery
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	Recipient sees survey results		One-time link sent to email
			Recipient creates password
			Recipient can edit profile, view past questions, etc.

Business model archetypes.

Six business model archetypes

- Transaction
- UGC
- Media
- SaaS
- Mobile app
- 2-sided mkt





SaaS model: WP Engine discovers the 2% cancellation rate



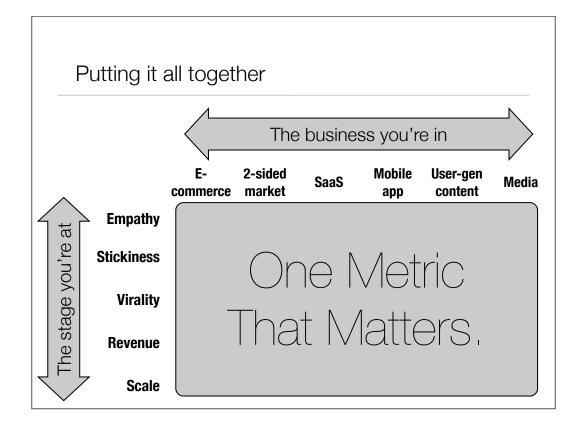
- Stage: Stickiness
- Model: SaaS
- Wordpress hosting company founded in July 2010, it raised \$1.2M in November 2011

Here's a concrete example of the stickiness stage.

WP-Engine discovers the 2% cancellation rate



- All companies have cancellations, but founder Jason Cohen was alarmed that he was losing a quarter of customers every year.
- Jason called customers himself. "Not everyone wanted to speak with me, but enough people were willing to talk, even after they had left, that I learned a lot about why they were leaving."
- Asked around. Turns out 2% is best case for most hosting companies.
- Without this, the company would have been getting diminishing returns overoptimizing churn; instead, they could focus on maximizing revenues or lowering acquisition costs.

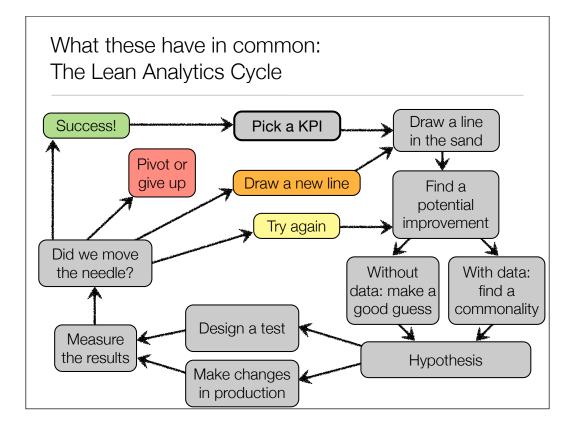


We've built on Eric's model, because we believe that companies naturally go through five stages of growth.

First, the need to get inside their customers' heads Then they need to make sure a small group of users keeps using the product

Then they need to make sure that word can spread Then they need to feed some of their revenues back into customer acquisition, accelerated by the virality they've built. Then they need to scale.

A cycle of learning.



What all these things share is a constant cycle of hypothesis and experimentation. We call this the Lean Analytics Cycle. It goes something like this:

- 1. Choose a KPI you want to improve, such as retention, revenue, or engagement, preferably one that represents the most fundamental business risk of your startup for this stage of growth and this business model
- 2. Draw a line in the sand for that KPI based on your business model
- 3. Figure out how to improve the KPI
- 4a. If you have no quantitative data, which will be the case early on, have a good idea you want to try as a result of Brainstorming, Stealing from others, User feedback, Etc.
- 4b. If you have quantitative data about what makes "good" users different from "bad" ones—meaning users, visitors, or customers who are doing what your business model wants them to (i.e. Buying, Signing up, Sharing, Contributing, Bringing in more revenue) then figure out what they have in common. Find a metric the "good" users share that's correlated with the KPI you want to change
- 5. Form a hypothesis ("if I do A then the KPI will improve.")
- 6a. If you're cautious, design A/B or multivariate experiments and implement the winning solution
- 6b. If you're not cautious, make changes to the business (product, market, pricing, etc.) that targets this commonality.
- 7. Measure the effect the change had on the KPI. Analyze a cohort of users that experienced the change.
- 8. Compare to the line in the sand. Did the KPI move past it?
- 9a. If it did, great. Fix the next thing that matters.
- 9b. If it didn't and you're out of time, pivot or quit.
- 9c. If it didn't, but you think the line was wrong, talk to customers and draw a new line.
- 9d. If it didn't but you have other ideas, try another potential improvement.





Gut instinct

Professional photography helps Airbnb's business



20 photographers in the field



Test results

Two to three times more bookings!



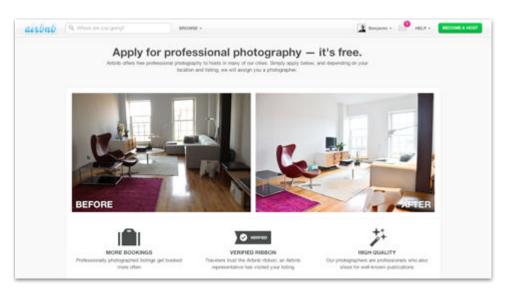
Back to the beginning

Use additional data to keep experimenting



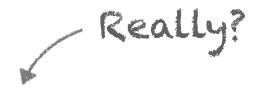
5,000 shoots / month in Feb. 2012







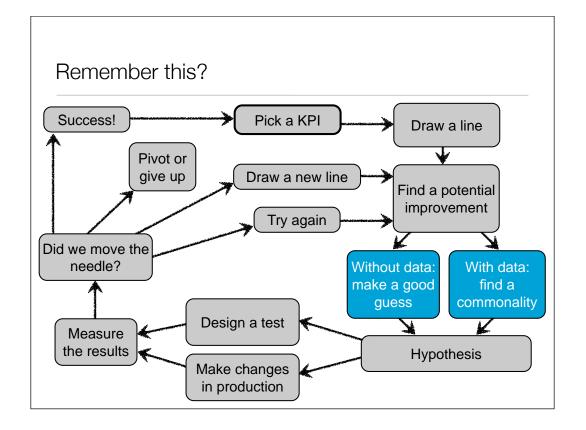
Hang on a second...



Gut instinct

Professional photography helps Airbnb's business







"Gee, those houses that do well look really nice."

Maybe it's the camera.

With data: find a commonality

"Computer: What do all the highly rented houses have in common?"

Camera model.

Mining the existing data: Circle of Moms finds an engaged market



- Stage: Stickiness
- Model: UGC
- Launched as Circle of Friends in 2007, it was a way for small groups to interact atop Facebook's platform; but when engagement wasn't good enough, the founders decided to dig deeper.

Mike Greenfield and his co-founders started Circle of Friends in September 2007, shortly after Facebook launched their 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 (Facebook had about 50 million users at the time) that was so open to reaching them. Circle of Friends was a simple idea: a Facebook application that allowed you to organize your friends into circles for targeted content sharing. Mike notes now that it was basically, "Google+ for Facebook" (before Google+ existed.)

The problem: Not enough engagement



- Too few people were actually using the product
- Less than 20% of any circles had any activity after their initial creation
- A few million monthly uniques from 10M registered users, but no sustained traction

By mid-2008, Circle of Friends had 10 million users. Greenfield focused on growth above everything else. "It was a land grab," he said. 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.

What Circle of Moms found



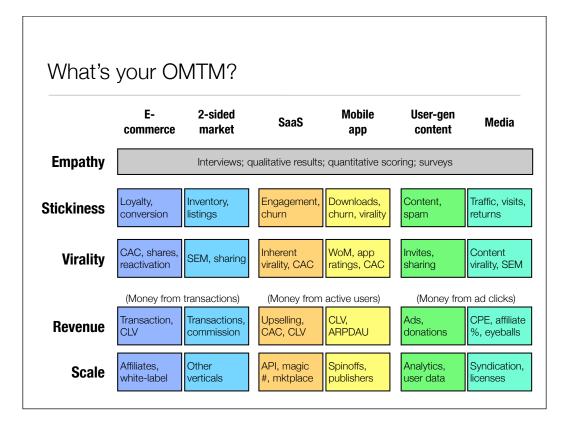
- They found moms were far more engaged
 - Messages to one another were on average 50% longer
 - 115% more likely to attach a picture to a post they wrote
 - 110% more likely to engage in a threaded (i.e. deep) conversation
 - Circle owners' friends were **50%** more likely to engage with the circle
 - 75% more likely to click on Facebook notifications
 - 180% more likely to click on Facebook news feed items
 - 60% more likely to accept invitations to the app
- Pivoted to the new market, including a name change
- By late 2009, 4.5M users and strong engagement
- Sold to Sugar, inc. in early 2012

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.

Circle of Moms proved to be extremely popular. By late 2009, Circle of Moms had 4.5 million users and engagement continued to be strong. 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.

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 bet—drop the generalized Circle of Friends to focus on a specific niche—a scary, gutsy bet, but one that was based on data.

Finding the One Metric That Matters.



As you move through the stages of Lean Analytics, the KPI you're trying to fix will shift. Here are some examples.

Choose only one metric.

Yes, one metric.



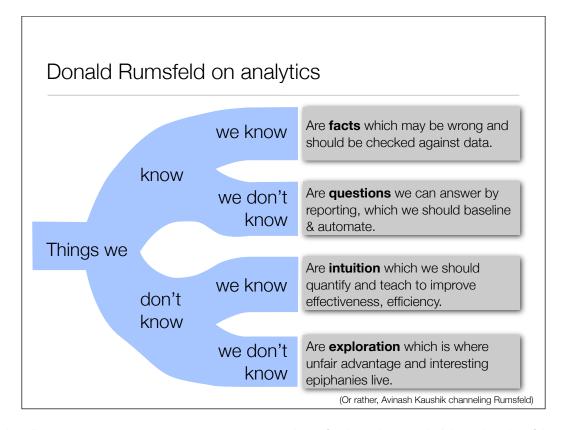
It will soon change.

In a startup, focus is hard to achieve.

Having only one metric addresses this problem.



Watch out for local maxima.



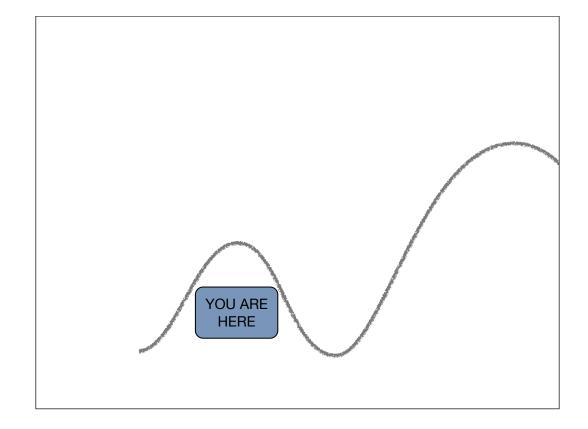
We work with data in two "postures." "Things we know" is a reporting posture—counting money, or users, or lines of code. We know we don't know the value of the metric, so we go find out. But we know the metric is needed, and big companies spend a lot of time doing it. It's the work of Business Analysts who work with Business Intelligence software and Data Warehouses.

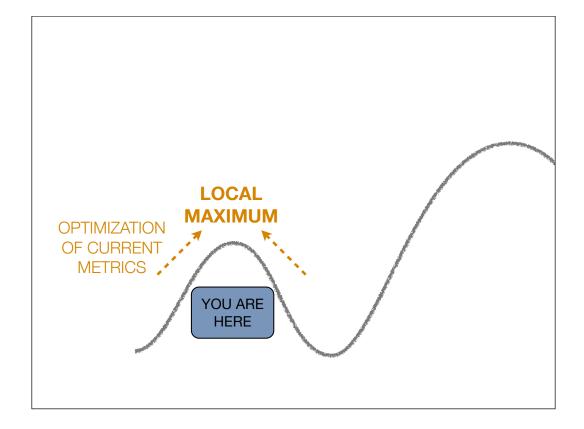
Reporting metrics track progress along an assumed path. A reporting metric says, "How many widgets did we sell today?" and tells our accountant how much tax to pay; an optimization metric says, "Did the green or the red widget sell more?" and chooses the official color of our widgets.

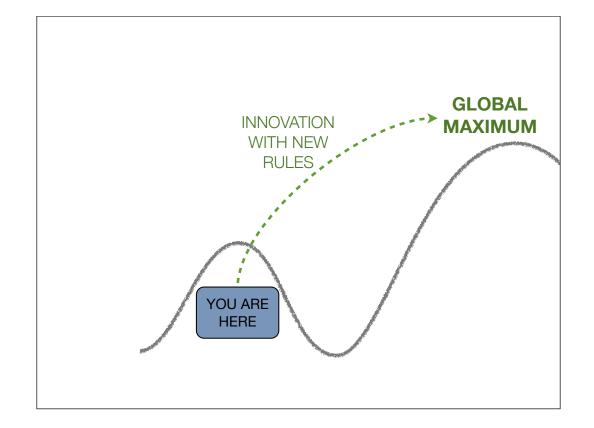
The other posture, "Things we don't know," is much more relevant to startups: exploring, to find out something new with which we can disrupt a market. It's repetitive, exploratory, and imprecise. It leads down plenty of wrong paths, and hopefully towards some kind of eureka moment when the idea falls into place. This fits what Steve Blank says a startup should spend its time doing: searching for a scalable, repeatable business model.

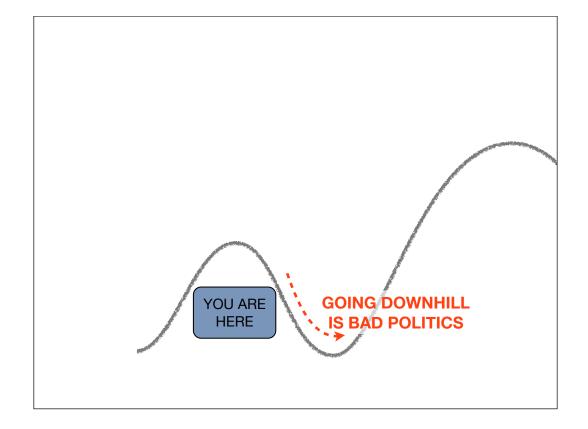
Exploratory metrics are "what if" metrics, the scores of our experiments. We expect them to change; we're testing one thing against another. In some cases, we may find that a completely unexpected cause produced the change we wanted. That cause may even become our secret advantage.

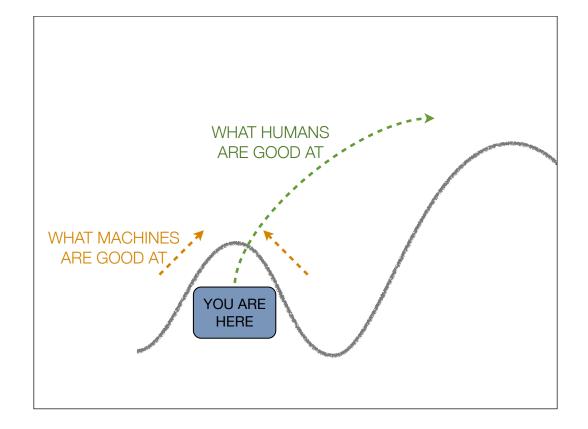
Analytics has a role to play in all four of Rumsfeld's quadrants. It can check our facts and assumptions, such as open rates or conversion rates, to be sure we're not kidding ourselves, and check that our business plans are accurate. It can test our intuitions, turning hypotheses into evidence. It can provide the data for our spreadsheets, waterfall charts, and board meetings. And it can help us find the nugget of opportunity on which to build a business. In the early stages of your startup, the latter matters most.







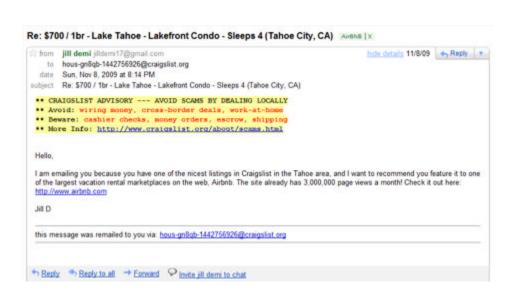




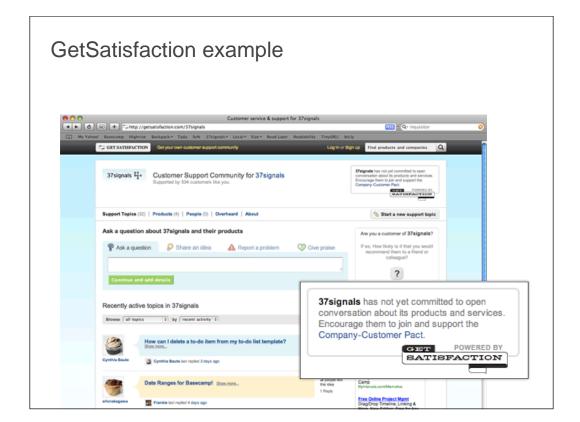
Dr. Daniel Goroff pointed me at this puzzler Martin Gardiner explained in the New York Times:

In the end, be **subversive**.

Airbnb and Craigslist







Coveting Gmail invites

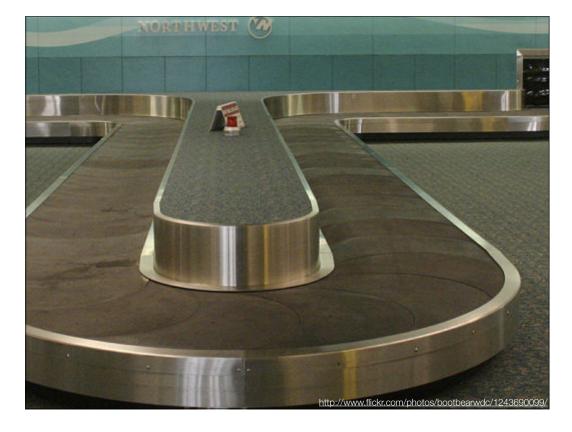




Farmville and wall posts







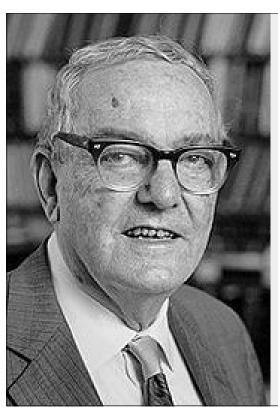
Here's an example of finding a cause.

SOME years ago, executives at a Houston airport faced a troubling customer-relations issue. Passengers were lodging an inordinate number of complaints about the long waits at baggage claim. In response, the executives increased the number of baggage handlers working that shift. The plan worked: the average wait fell to eight minutes, well within industry benchmarks. But the complaints persisted.

Puzzled, the airport executives undertook a more careful, on-site analysis. They found that it took passengers a minute to walk from their arrival gates to baggage claim and seven more minutes to get their bags. Roughly 88 percent of their time, in other words, was spent standing around waiting for their bags.

So the airport decided on a new approach: instead of reducing wait times, it moved the arrival gates away from the main terminal and routed bags to the outermost carousel. Passengers now had to walk six times longer to get their bags. Complaints dropped to near zero.

Turns out that long delivery wasn't causing dissatisfaction. Time spent waiting was.



"What information consumes is rather obvious: it consumes the attention of its recipients.

Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it."

(Computers, Communications and the Public Interest, pages 40-41, Martin Greenberger, ed., The Johns Hopkins Press, 1971.)

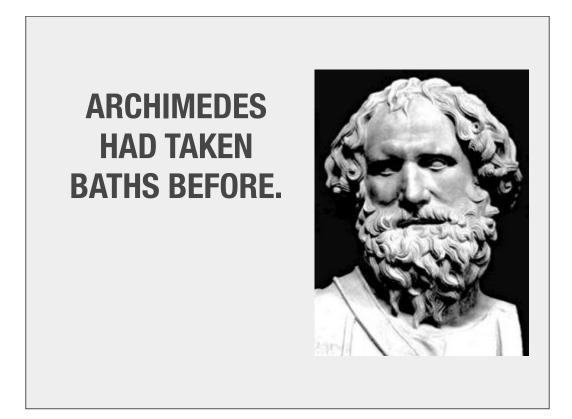
"The most important figures that one needs for management are unknown or unknowable, but successful management must nevertheless take account of them."

Lloyd S. Nelson



I want to tell you the story of Archimedes.

According to Vitruvius, a votive crown for a temple had been made for King Hiero II, who had supplied the pure gold to be used, and Archimedes was asked to determine whether some silver had been substituted by the dishonest goldsmith.[14] Archimedes had to solve the problem without damaging the crown, so he could not melt it down into a regularly shaped body in order to calculate its density. While taking a bath, he noticed that the level of the water in the tub rose as he got in, and realized that this effect could be used to determine the volume of the crown. For practical purposes water is incompressible, [15] so the submerged crown would displace an amount of water equal to its own volume. By dividing the mass of the crown by the volume of water displaced, the density of the crown could be obtained. This density would be lower than that of gold if cheaper and less dense metals had been added. Archimedes then took to the streets naked, so excited by his discovery that he had forgotten to dress, crying "Eureka!" (Greek: "εὕρηκα!," meaning "I have found it!"). The test was conducted successfully, proving that silver had indeed been mixed in.[16]



What's most important about this story is that Archimedes had taken baths before. It was the act of being asked that caused him to find an answer that had been there all along.

Once, a leader convinced others in the absence of data.

Once, a leader was someone who could convince people to do things in the absence of data. Now, a leader knows what questions to ask.

Now, the leader is the one who knows what questions to ask



Ben Yoskovitz byosko@gmail.com @byosko



Alistair Croll acroll@gmail.com @acroll

