MAKING FRIENDS WITH COHORT ANALYSIS

5 Ways to Deep-Dive into Customer Data

Who are your best customers? Where did they come from? How can you retain them? From gaming apps to retail stores to software-as-a-service, countless businesses are asking these questions. And they’re using cohort analysis to find the answers.
What’s a cohort to you?

A cohort is a group of people who have something in common (for example, the channel that brought them to your website) over a certain period of time. Cohort analysis allows us to dig into groups of like users and model how the business is changing. For example, we might see positive trends in Months 1-3, with revenue flattening in Months 4-6, and maybe even a downward trend in later months. In one scenario, a shift in customer mix (perhaps from different sources or markets) could be driving more one-time customers; in another scenario, the downturn could be caused by a diminishing long-term user experience.

By analyzing a cohort, you can discern patterns across the customer journey—and answer any number of questions:

- **Channel attribution:** Which marketing channels delivered the highest-value customers?
- **Trends in user quality:** Do new customers spend as much as old ones?
- **Engagement points:** What steps, in what order, precede conversion?
- **Churn rates:** Do people drop off at a consistent rate and/or in the same places?
- **Product/Process issues:** At what point do users stop engaging?
More than a pretty picture.

The most common way to begin with cohort analysis is to define the customer’s starting point—the first time they clicked a marketing email or online ad, the first time they registered on your website, their first download, or their first booking or purchase—then build out over time, typically over days or weeks.

Here’s a cohort chart (the proverbial layer cake) that Buffer, a social media engagement company, uses to map signup month to revenue contribution.

When you produce a cohort analysis chart, you often end up with some sweet eye candy. (The “layer cake” never fails to delight.) But what business value are you gaining? How can you take action based on the results? Let’s take a look at several examples of companies who are using cohort analysis to make smarter marketing decisions, correct issues on their websites, improve product quality, and influence retention in a very controlled way.
How should you spend your online marketing budget?

The modern marketing organization takes advantage of a wide array of tactics across both online and offline channels. Which programs are succeeding? Which traffic sources are returning more in revenue than they’re costing you? Of course, each vendor will show you data that claims their service returns 5X on your spend. How can you get the upper hand on the data so you can compare sources for yourself?

To solve the “where and when to spend money” problem for themselves, Live Out There generated activation curves for paid traffic cohorts. They developed a method to “repatriate” cumulative ad spend to acquisition cohorts, which yielded a running cost for each cohort over time. They then compared the repatriated cohort to the contribution margin dollars for the cohort, so they could calculate a payback figure.

How are different traffic sources performing over time? This chart shows example activation curves for cohorts of paid traffic sources. The X axis shows the age of a cohort in weeks. The Y axis shows the percentage of the cohort who have become buyers each week. The X axis is normalized so cohorts starting in different weeks can be compared more easily to one another.

Using this method, the marketing team at Live Out There can see which campaigns generate new buyers who add more revenue dollars than the cost of acquiring them—and can allocate their budget to those traffic sources. And by setting a maximum duration in which they could expect a cohort to “pay back,” they could tune, tweak, and turn on and off traffic sources accordingly.
Is radio a viable marketing channel?

Even as opportunities for online marketing seem to multiply overnight, offline marketing channels continue to be excellent drivers of new customer engagement. Radio advertising, for example, remains an affordable choice for both brand-building and direct response campaigns—even for the most innovative Internet businesses.

Curious how radio advertising would perform, particularly side-by-side with digital channels, Simply Business ran an experiment. They designed a regionalized campaign that would enable them—along with a sophisticated data pipeline and datamining tools—to identify which website visitors had probably arrived based on a recent radio ad. The company then could establish a cohort of those visitors and track them all the way through the purchase funnel with tremendous clarity, calculating a reasonable attribution of new revenue to the radio initiative.

This visualization shows the impact of the radio ads in each region. Looking at website visitors by marketing channel, Simply Business could see that their direct channel (visitors who accessed the site by directly typing in their brand name) had spiked in the days immediately following each advertising flight—exactly the desired result.

Unlike other channels, such as paid search, which typically acquire customers much further along in their purchase consideration, the company expected radio to be more about driving awareness and therefore to have a longer payback period. However, by observing the behavior of the radio audience segments over time, they were able to validate that brand nurturing does eventually drive positive ROI directly—an insight which spells out the business case for further investment in radio advertising.
How can you understand user retention?

No matter what kind of online business you’re running—an ecommerce site, a mobile game, a media site, or a software subscription—you now have the ability to collect enormous quantities of user-level data. What can you learn from it? Cohort analysis provides a way to delve into all that data, returning valuable insights into who’s engaging with your brand, how they’re finding their way through the sales funnel, and how they’re using your product—and ultimately, how good is your product.

A growth analytics team at Yahoo! wanted to examine engagement metrics and user retention for several of the company’s mobile products. Out of a given set of users, how do they use the product? How long do they stick around? To answer these questions, the team set up cohort dimensions—attributes such as day of installation and device type.

<table>
<thead>
<tr>
<th>EVENT DATE</th>
<th>PRODUCT</th>
<th>INSTALL DATE</th>
<th>OS_NAME</th>
<th>ACTIVE USERS</th>
<th>COHORT SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mon</td>
<td>Mail</td>
<td>Mon</td>
<td>Android</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Tue</td>
<td>Mail</td>
<td>Mon</td>
<td>Android</td>
<td>83</td>
<td>100</td>
</tr>
<tr>
<td>Mon</td>
<td>Mail</td>
<td>Mon</td>
<td>iOS</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Tue</td>
<td>Mail</td>
<td>Mon</td>
<td>iOS</td>
<td>62</td>
<td>75</td>
</tr>
</tbody>
</table>

This simplified view shows that of 100 Mail installs on Android devices on a Monday, 83 devices were running the product on Tuesday. Installs on iOS devices were 75, dropping to 62 active user devices the next day.

By extending the cohort analysis, adding new dimensions and measures, the team wanted to understand what cohorts resulted in high LTV (lifetime value) customers. For each possible combination of cohort dimensions, for every possible event date, they looked at: How many devices belong to that cohort? How many devices from that cohort were active on that day? What is the retention rate for that cohort? And—most importantly—what are the characteristics of customers who stick around for a long, long time?
How can you use loyalty signals to drive your business forward?

Driving growth doesn’t stop at customer acquisition. Who are your most loyal consumers? How can you derive more revenue from those who already know and love your product or service? Cohort analysis can help you understand the behavior of your high-value customers and enable you to find smarter ways to engage with them, building a natural loyalty that will keep you friends for a long, long time.

The team at ThredUp realizes that every customer looks different and that they each spend differently over time. To understand their audience and their business, the company used cohort analysis as a way to capture patterns. Their goal? To find the inflection point in customer loyalty. When does a customer become a much better customer? What does it take to make that happen?

For example, the team noticed that purchasers of both kids’ and women’s clothing tend to stick around longer. How can ThredUp push mixed purchases? Customers who purchase from mobile devices had obviously downloaded the ThredUp app, indicating a deeper integration into their lives. How can ThredUp improve their mobile platform and drive customers to it?

ThredUp created a view that quantified the engagement value of different cohorts, based on first purchase. By classifying a customer within a cohort, they could assess the importance of retention efforts at particular inflection points.

By tracking all sorts of dimensions—acquisition costs, orders, order frequency and size, returns, profit margins, and more—they gained insight into how their best customers are acquired, what motivates them to purchase, and what to expect in terms of transactional volume and revenue contribution. They could also quantify the probability that a customer will return, not just in the next few months, but in the years.
How can you use engagement metrics to improve your game?

Gaming companies that produce game apps like Candy Crush, FarmVille, and Angry Birds love to track Daily Active Users (DAU) and Monthly Active Users (MAU), and data collectors regularly publish rankings of top apps based on these metrics. But DAU and MAU can easily mask more important metrics related to retention and growth. If you can learn where users are dropping off, you might gain the insights you need to build a better gaming experience.

Fortunately, it’s easy for a mobile gaming company to quickly iterate and update their app. By completely retuning the three levels, the team was able to eliminate the roadblocks that users were experiencing—and even to encourage churned players to return.

To ensure attention is paid to ongoing user engagement, producers of a mobile gaming app, much like Candy Crush, developed a quick cohort analysis and turned it into a custom metric that could track for each level of the game. In the example below, they saw that early, less-engaged users were struggling to get past Level 1, but loyalty quickly jumped and increased steadily over time, with some mild variance. More importantly, though, they also noticed that there were three levels that seemed to be driving users away.

This level-completion metric tells them how many users who start a given level complete that level.

The green columns show a smoothed retention curve, increasing user engagement, and a potential for more revenue.

The green columns show a smoothed retention curve, increasing user engagement, and a potential for more revenue.
Why does all this matter?

Once you group customers into cohorts—whether by the month they joined, by the day they downloaded, by the operating system they used to purchase—you get a much smoother view of your audience, even over very large sets of data. It’s as if a cohort is one giant customer, eliminating the noise of all those individuals. Cohort analysis works for any type of business, no matter what type of interactions, giving you the insights you need to execute smarter marketing, to engage with customers effectively, to redesign your product or service, and to focus on your most valuable customers.
Making Friends with Cohort Analysis

About Looker
Looker is an inventive software company that’s pioneering the next generation of business intelligence (BI). We believe in bringing better insights and data-driven decision-making to businesses of all sizes. The company has fast become the catalyst that is creating data-driven cultures at hundreds of industry-leading companies such as Yahoo!, Gilt, Warby Parker and Sony.

Looker is based in Santa Cruz, CA
Follow @lookerdata or visit www.looker.com

© 2015 Looker. All rights reserved. Looker and the Looker logo are trademarks of Looker Data Sciences, registered in the United States. Other trademarks are trademarks of their respective companies. All services are subject to change or discontinuance without notice.