A MODE ANALYTICS REPORT

The Essential Guide to Improving User Retention

MODE

Nail retention. Grow your user base.

Conventional wisdom says it costs a lot more to get a new user than to keep a current one. Exactly *how much more* expensive is up for debate, but the fundamental concept is pretty clear: you're wasting resources if the users you work hard to acquire don't stick around.

The measures of how long users "stick around" are commonly referred to as user retention.

Strong retention has positive effects on a wide variety of metrics that likely matter to your business. The longer users stick around, the more likely they are to invite or refer others, increasing virality. You'll also have more opportunities to monetize—whether that means showing more ads, increasing average subscription length, or boosting repeat purchases—and increase user lifetime value.

If you want to move the needle on your most important metrics, start by improving retention.

Getting started

In this guide, we'll start simple by showing you how to measure retention. Armed with a baseline, we'll dive into deeper questions to find opportunities to improve retention. By the end, you'll be able to answer the following questions:

- How long do users stick around?
- Who are our most valuable users?
- What actions do retained users take?

The best way to begin improving retention is to establish a baseline of **overall retention** for your product. Overall retention is calculated by finding out how many users who sign up in one period are still around in a subsequent time period—and then dividing the second number by the first.

Imagine that in the last week of April, 79 users sign up for your product. Of those users, 17 return the following week. To calculate the retention rate, simply divide the number of returning users (17) by the number original users (79). That means your retention rate is 22%. That's a massive drop.

And that was only one week! What about the weeks that follow?

Date	New Users	1	2	3
Apr 27, 2014	79	22%	19%	13%

The table above shows that 19% of the original group of users returned two weeks after signing up and only 13% returned on the third week. Looks like you might have a leaky bucket on your hands.

Before you sound the alarms, make sure the users who signed up during the last week of April weren't an anomaly. Are the users who signed up during other weeks exhibiting similar behavior?

By looking at additional data, we can see that users who signed up in subsequent weeks are dropping off at about the same rate as the signups from the last week in April. After one week, about 80% of users have abandoned the product. With the baseline established, we need to begin investigating why users drop off so quickly and identify areas that might help us fix the problem.

Date	New Users	1	2	3
Apr 27, 2014	79	22%	19%	13%
May 04, 2014	168	23%	21%	21%
May 11, 2014	188	19%	19%	13%
May 18, 2014	191	23%	21%	22%
May 25, 2014	191	21%	16%	20%

Of course, weekly numbers might not be right for your product. Instead, use time periods that roughly mirror how often people typically use your product. If your product is used daily, like a social app or messaging service, look at day to day changes in retention. By contrast, if you're building a service to help people pay credit card bills, monthly retention rates are likely more meaningful.

Exploring retention with SQL

There are plenty of tools that help you track retention. But if your goal is to *improve* retention, there's no better tool than working directly with raw data from your product. No two products are exactly the same, and one-size-fits-all retention reports don't provide the flexibility needed to analyze retention, not just measure it.

The explorations that follow can be used to build retention reporting from the ground up. Each exploration is accompanied by a report you can tailor to your product data, stored in a relational database like Redshift, Postgres, or MySQL.

The reports assume data is stored in schema that includes:

- **A user table** of people, accounts, visitors, or any other type of identity.
- **An events table** of page views, purchases, logins, or other actions taken by users.

To learn more about the expected schema, read this article.

HOW TO TAILOR RETENTION REPORTS TO YOUR DATA

Open a report from this guide in your browser. Click Clone to copy the report into your Mode account and modify the SQL query to reference your schema. Once you have the report working with your data, you can start to tailor the queries to fit the nuances of your product—something no out-of-the-box retention tool can achieve.

For more information on how to modify the queries, click here.

Exploration 1

How long do users stick around?

In the introduction, we calculated retention by grouping users based on the week they signed up. In this exploration, we'll be grouping users on a monthly basis. These groups are known as cohorts. Cohorts are commonly defined by the time period in which users sign up.

Acquisition cohorts allow you to monitor retention over time and are useful tools for understanding how users react to product changes. For instance, if a new product is unpopular, retention rates for new users might start to drop.

A slow decline indicates that users are satisfied with your product—it's natural for users to drift away over time. However, if the retention rate dips significantly from one period to another, you'll want to investigate further.

The chart below has become synonymous with retention. The chart has a triangular shape because the later rows—which represent recent signups—haven't been users for long enough to populate the right side of the table. Since we can't measure the 9-month retention rates for users who joined two months ago, those cells are left blank.

The first column shows the month users signed up, and the second column shows the number of new users. Reading from left to right, you can see the percentage of users who retained during each subsequent month. From top to bottom, you can see how retention rates are improving or worsening with each new cohort.

Monthly Retention Rates by Signup Date

Signup Date	New Users	1	2	3	4	5	6	7	8	9
Jan 01, 2014	78	96%	92%	90%	88%	86%	85%	85%	83%	82%
Feb 01, 2014	88	98%	93%	89%	88%	86%	83%	82%	80%	
Mar 01, 2014	103	100%	95%	91%	89%	87%	83%	80%		
Apr 01, 2014	107	99%	95%	93%	91%	86%	84%			
May 01, 2014	114	98%	92%	86%	85%	84%				
Jun 01, 2014	128	95%	93%	90%	86%					
Jul 01, 2014	136	95%	90%	87%						
Aug 01, 2014	149	97%	91%							
Sep 01, 2014	158	97%								

TAILORING TIPS

Start by opening the interactive version of this report. After cloning the report, check out the SQL query and update the column references to match your schema.

For this report, you'll need:

- **the date** each user activated their account
- **the timestamp** of the actions that the user has performed since signing up, such as homepage visits, messages sent, or downloads

The report is designed to allow you to switch between weekly and monthly retention, but you can easily modify the query to look at daily retention, too. And because you're working with raw data you can exclude events that might not be meaningful actions.

This report also includes two churn tables. Churn is the opposite of retention—it lets you know how many users are abandoning your product instead of sticking around.

To learn more about using this report with your data, read the documentation.

Exploration 2

Who are our most valuable users?

Although acquisition cohorts are great for measuring what's happening, they aren't very actionable. You can't get more users to sign up last week, so you'll have to look for other clues into what makes certain users more likely to retain than others.

Demographic information is a great place to start. By segmenting users into demographic cohorts—like the device they were using when they signed up, where they live, or where on the web they came from—you can begin to suss out patterns.

In some cases, the necessary actions are clear. If retention rates are higher for users acquired through LinkedIn ads than through Facebook ads, you can reduce Facebook spend and increase LinkedIn spend.

In other cases, the conclusions are more complex. Suppose you find that invited users retain at a higher rate than the baseline. Might this indicate that social connections make your product experience more enjoyable? Maybe. A few experiments to increase connections amongst non-invited users could uncover a retention goldmine.

The idea here is that demographic cohort patterns can inspire ideas for experiments. As you deploy experiments, watch for changes in the cohort's retention rate. If retention rate increases, it's an indicator that your experiment may have worked.

The chart below shows demographic cohorts based on user language. Look familiar? It's very similar to the acquisition cohort chart from Exploration 1. You can still read from left to right to see how each cohort retained over time. Reading from top to bottom allows you to compare retention across all cohorts of the same age. For instance, we can see that speakers of Chinese, Arabic, and German drop off sharply three weeks after signup.

Weekly Retention Rates by User Language

Language	New Users	1	2	3	4	5	6	7	8	9
arabic	133	18%	14%	11%	14%	11%	11%	9.8%	11%	6.0%
chinese	118	20%	14%	7.6%	14%	14%	14%	15%	8.5%	8.5%
english	1,858	19%	16%	16%	15%	15%	13%	12%	9.0%	6.6%
french	311	20%	17%	17%	18%	13%	10%	12%	6.1%	5.8%
german	207	19%	14%	12%	16%	15%	14%	14%	13%	8.2%
indian	113	20%	17%	18%	12%	16%	11%	9.7%	12%	8.8%

TAILORING TIPS

To use this report with your data, follow the same steps as in Exploration 1. Open the interactive version of this report, clone the report, and change the column references to match your schema.

For this report, you'll need:

- **the date** each user activated their account
- **the timestamp** of the actions that the user has performed since signing up, such as homepage visits, messages sent, or downloads
- **a demographic attribute** (in this case, language)

With a little tweaking in SQL, you can easily customize these tables to show any type of demographic attribute (as long as you're tracking it): gender, location, age, signup source, etc. For example, here's another report where users are divided into cohorts by device type.

As before, you can modify this report to show whatever time period makes the most sense for your business.

For more tips on customizing this chart, read the documentation.

Exploration 3

What actions do retained users take?

Exploration 2 focused acquisition efforts on user attributes. Now we'll dive into behavioral data to determine how to encourage long-term usage early on in the customer journey.

The goal here is to help users find value as soon as possible. To figure out what creates value, you need to examine how users interact with your product after they sign up and see if there's a correlation between their actions and long-term retention.

For instance, Facebook found that new users who add at least 7 friends in their 10 days on the platform are very likely to continue using Facebook. By adding friends, users are exposed to the core value of Facebook—connecting with friends and family. Users who don't reach this point are much less likely to return.

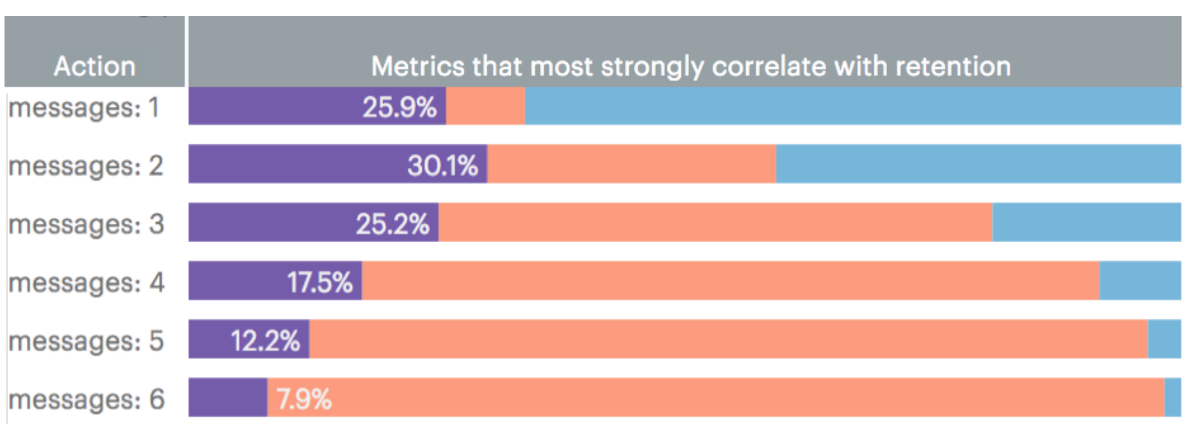
This simple metric of when a user becomes hooked on a product is popularly known as an "aha moment." "Aha moments" aren't about precision, but about defining a core goal. Once Facebook defined their "aha moment," they focused on features that encouraged users to add more friends, like making friending a step in the signup flow and adding a box that suggests people you may know.

The "aha moment" is a set of actions that separate users who find value in your product from those who don't. Users who find value come back. To identify which actions separate retained users from lost ones, you'll need to group your users into **behavioral cohorts**.

The chart below compares the number of times users take action in their first week to the likelihood that they'll return in their second week.

- **Purple bars** represent users who both retained and did the listed action.
- **Red-orange bars** represent users who retained, but didn't do the listed action in their first week.

Retention Rates by Action in Users' First Week



- **Blue bars** represent users who took the action but didn't retain.

The larger the purple bar (users retaining and taking the actions) relative to red-orange bar (users just retaining) and blue bar (users just taking the action), the more strongly this action correlates with retention. In this case, it looks like we should encouraging users to send two messages early on.

TAILORING TIPS

By now, you know the drill. Here's the interactive version of the report. Clone it and update the column references to match your schema.

If you want to examine how combinations of actions affect retention (e.g. one message sent + three homepage visits) check out this second report.

Both of these reports use the same data. You'll need:

- **the date** each user activated their account
- **the timestamp** of the actions that the user has performed since signing up, such as homepage visits, messages sent, or downloads
- **the names** of the actions that the user has performed since signing up

You can pick any action or set of actions, but we recommend starting out with what you think might drive retention, and go from there.

For more tips on customizing the first report, read this documentation. For the second report, read this documentation.

A word of caution

Correlation doesn't imply causation

All of the reports included here identify correlations, not causations. Just because retention and a variable appear related, doesn't mean they are. For instance, if retention dips when you release a new product feature, it might not be *because* you released that feature. Likewise, French speakers might stick around longer, but it doesn't mean they do this *because* they speak French. And although the users who send two Facebook messages retain better, it doesn't mean that they retain *because* they've sent two Facebook messages.

That being said, correlations are usually a good place to start experimenting. When you find a demographic attribute that appears to correlate with strong retention, try to target those users. When you find an action that appears to correlate with strong retention, push more users to take that action. If the relationship is causal, more people will retain and your correlation will hold. If it's not causal, people will take the action but won't retain at higher rates. As a result, the original correlations will break down until another demographic or behavior emerges as being more strongly correlated with retention. Rinse and repeat until your correlation holds.

To understand how this works, imagine that analysts at Facebook discovered that sending two messages correlates strongly with retention. It's possible that sending two messages shows users great chat features, which makes them hooked. But it's also possible that people like Facebook for other reasons and messages are just a side effect.

If Facebook wanted to test the idea that sending messages drives retention, they could put a big button on their homepage that said "CHAT WITH YOUR FRIENDS!" This would almost certainly increase the number of messages sent. If messaging drove retention, the experience would result in increased retention. However, if the correlation between messages and retention wasn't causal, it'd have very little effect on retention and the correlation would start to break down. Facebook could then check that idea off the list, and start to look at other features.

Retention, full circle

Together acquisition, demographic, and behavioral cohorts form a framework for iteration that sets you up to master retention and accelerate growth. When looking for how you're doing, turn to acquisition cohorts. When looking for what to do next, start segmenting with demographic and behavioral cohorts. Cycling through cohort analysis can lead to quicker insights to improve retention, and because you've built these reports with raw data, there's no end to the explorations you can do.

Questions?

If you have any questions about using these reports with your data, let us know. We're happy to help.

About Mode

Mode is a collaborative analytics platform built by analysts, for analysts. Sign up for a free trial at www.modeanalytics.com.

MODE