**Growth Analytics**

*Introduction*

-focus: B2C, internet-based companies

-resources: text (Lean Analytics, Amplitude’s Product Analytics Playbook)

*Growth*

-viral coefficient

-cycle time

-caveat: not all new users arrive as a result of contact with an existing user

*Churn*

-retention

-churn rate/probability

-customer lifetime

*Profitability*

-customer lifetime value

-customer acquisition cost

-payback period

*Introduction*

The core focus of most companies is growing and monetizing their user base. Doing this intelligently is the aim of a corner of data science often referred to as growth analytics. With the associated concepts (e.g., viral growth, customer retention, lifetime value), companies are able to perform the vital work of tracking and modeling changes to their total number of users and the associated cash flows.

Internet and businesses-to-consumer (“B2C”) companies, whose potential customer base includes the entire internet-connected world ([4.5B+ billion people](t.ly/1VZm) as of 2020) are particularly well-suited to benefit from such analysis. Their customers can easily sign onto a service, and their engagement with particular product features serves as an important measure of business health. These organizations include platforms hosting user-generated content alongside embedded advertisements (Facebook, Twitter, TikTok), logistics companies physically delivering items ordered online (Amazon, Uber, Instacart), and financial technology firms involved in transactions processing (PayPal, Square, Stripe).

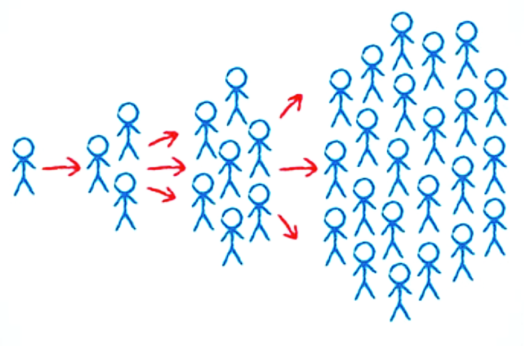
On the other hand, growth analytics tools may be of limited use to companies whose core products are not online do not have the benefit of easily monitoring user engagement. Similarly, business-to-business (“B2B”) companies typically court a small number of prospective customers and often customize their product and pricing to each one.

This article aims to summarize key concepts in growth data science taken from both text ([Lean Analytics](t.ly/JgJYm), [Amplitude](t.ly/P9qB), [Social Capital](https://medium.com/swlh/diligence-at-social-capital-part-1-accounting-for-user-growth-4a8a449fddfc#.bvu0we2z3), [Tribe Capital](https://tribecap.co/a-quantitative-approach-to-product-market-fit/)) and video ([Alex Schultz](t.ly/Q6G3), [Elliot Shmukler](t.ly/3pZZ5)) resources.

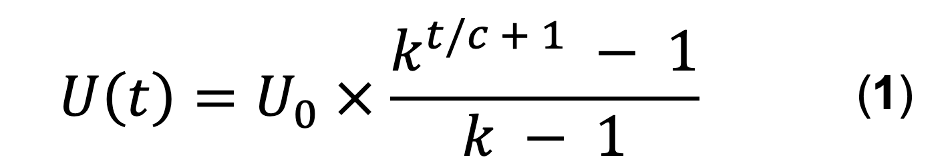
***I. Growth***

*1. Viral growth model*

Borrowed from epidemiology, the [viral growth model](https://readwrite.com/2015/10/09/virality-math-formula/#:~:text=To%20fix%20this%2C%20we%20simply,and%20the%20cycle%20time%20%E2%80%93%20Ct.) allows companies to project the size of their user base over time under the assumption that new users are brought in by existing users.

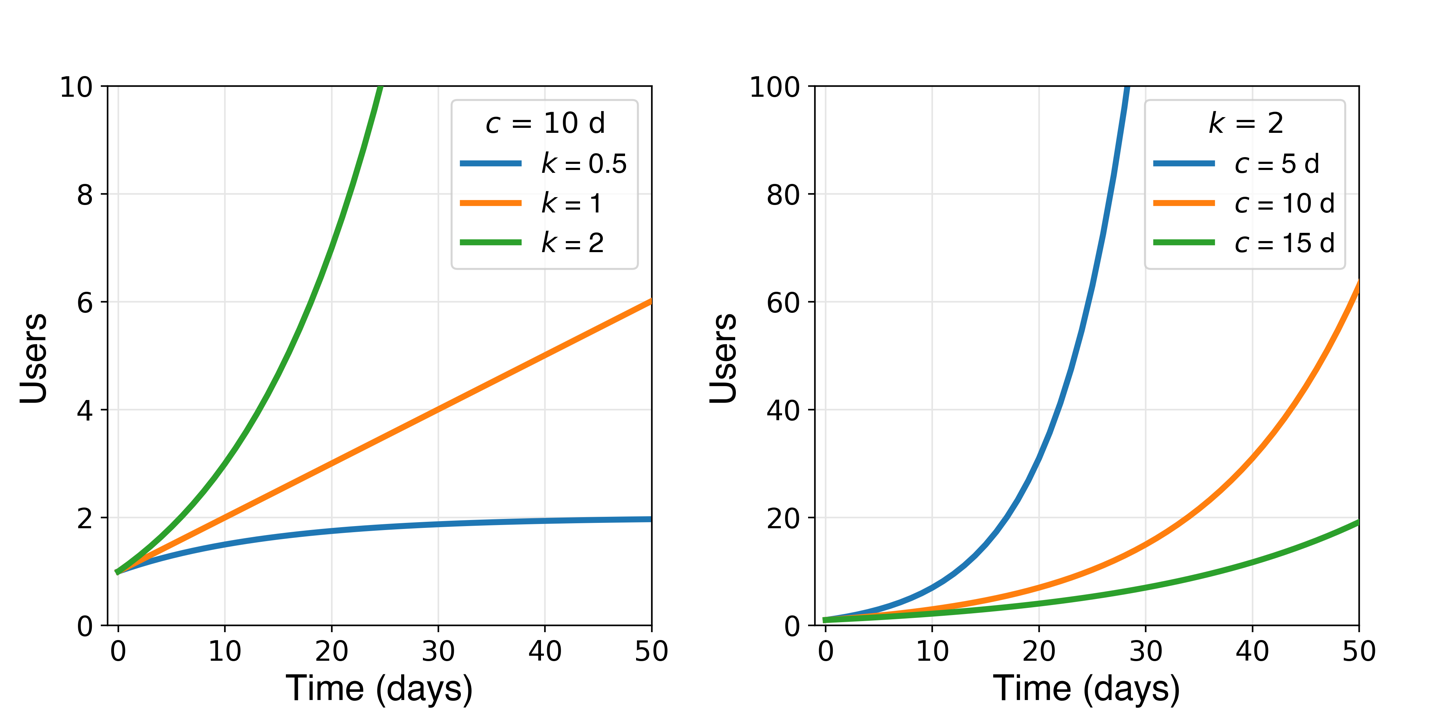


Specifically, the number of users at any particular time can be expressed in terms of current user base *U*0, the number of viral cycles *t*/*c* (*t* is time elapsed and *c* is cycle time), and the number of new users per existing user *k*, also known as the viral coefficient:



A. Viral coefficient, *k*

The key assumption of the viral growth model is that new users engage with your product as a result of interaction with current users. Practically speaking, a company can empirically determine its viral coefficient *k* by multiplying the invitation rate of a current user by the acceptance rate of an invited user. For example, if existing users send on average 5 invites, and 40% of invitees accept, the viral coefficient is 5 × 0.4 = 2. In this scenario, each current user will bring on 2 new users. A key goal for many young companies attaining a viral coefficient of one, since *k* = 1 defines the boundary between saturating and accelerating growth (Figure 1, left).

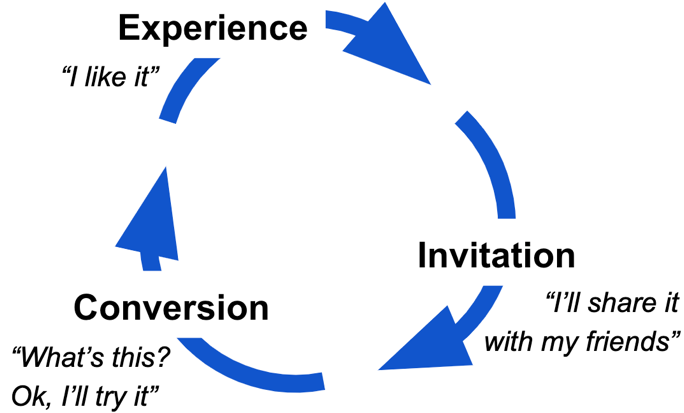


**Figure 1**. Effect of viral coefficient *k* and cycle time *c* on growth of user base according to Equation 1. (Left): linear growth results for *k* ≈ 1 (note: *U* is not defined at *k* = 1) while *k* > 1 and *k* < 1 yield exponential growth and saturation, respectively. (Right) shortening cycle time increases growth rate.

B. Cycle time, *c*

To bring time into the analysis, the viral coefficient *k* is raised to the power of *t*/*c* + 1, where *t/c* is the number of viral cycles elapsed. Starting with 1 user and a viral coefficient of 2, the number of users after 1, 2, and 3 cycles will be 3, 7, and 15, respectively. In the first cycle, the first user brought on 2 new users. In the second cycle, those 2 latest users each brought on 2 (4 total) new users. In the third cycle, those 4 latest users each brought on 2 (8 total) new users. Following this logic, for a given time *t* elapsed, reducing the cycle time *c* is equivalent to increasing the number of elapsed cycles (Figure 1, right). While the viral coefficient dictates the shape of the user growth curve, cycle time controls the compression of the growth curve along the x-direction.

The viral cycle is composed of three steps: (1) a positive user experience is delivered, (2) the user invites others, (3) some of those invitations are accepted (Scheme 1). The cycle time c is can then be calculated as the average time interval separating the conversion of one user and conversion of the user(s) they bring in.



**Scheme 1**. Viral cycle.

There are a couple important things to note about the viral growth model:

1. Viral coefficient and cycle time can change. Ideally, improvements to your product increase *k* and reduce *c* (see next section), but it’s also possible that competition, market saturation, or unfavorable product changes do the opposite*.*
2. Some new users may have discovered your product without any invitation. In this case, calculating future user count with measured viral coefficient and cycle time will yield an underestimate, since it does not take self-conversions into account.

As such, the viral growth expression outlined in Equation 1 offers a clear illustration of how two key levers (*k* and *c*) drive user growth, but an extrapolated future user count based on instantaneous measurements of viral coefficient and cycle time should be interpreted with caution.

*2. Increasing growth rate*

Even if Eq. 1 is an oversimplification, increasing the viral coefficient *k* and reducing cycle time *c* are two critical approaches to increasing user growth.

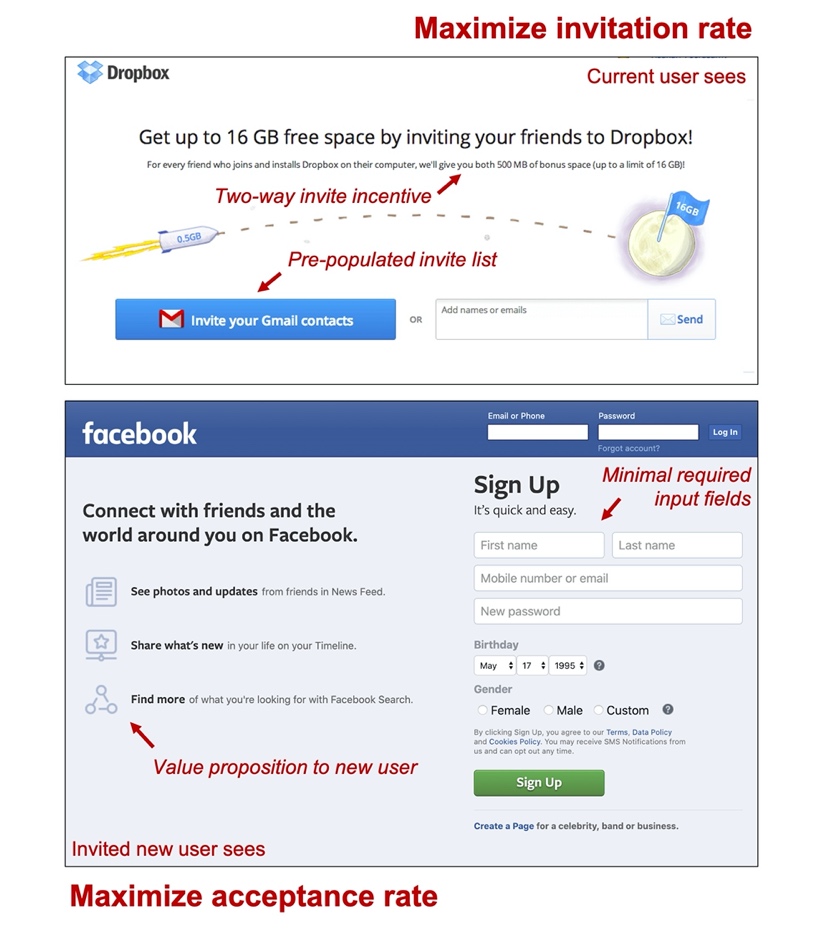
**Increasing viral coefficient, *k***

Since *k* is the product of invitation rate and acceptance rate, increasing either will increase viral coefficient. To increase ***invitation rate*** (number of invites sent out per existing user, see Figure 2 top), you might consider doing the following:

1. Present users with an option to invite a list of their contacts pulled from email, social media, etc. This makes it easier to for users to invite others.
2. Build in incentives for sharing. These can be one- or two-way, benefitting the inviter or both inviter and invitee, and can be monetary or otherwise. This makes sending invitations more appealing.
3. Work on extending the user lifetime to allow for more sharing. Once a user drops off, they won’t be inviting potential new users.

On the other hand, to increase ***acceptance rate*** (fraction of invitees who convert, see Figure 2 bottom), you might:

1. Add compelling content to landing page, such as a brief summary of the value proposition to new users.
2. Reduce barriers to new user signup. Requiring as few input fields as possible (e.g., name, email, password) will make it easier for interested prospective users to convert.



**Figure 2.** Increasing viral coefficient by optimizing invitation rate from current users (top: Dropbox example) and acceptance rate from invited prospective new users (bottom: Facebook example).

B. Reducing cycle time, *c*

Reducing cycle time is equivalent to fast-forwarding your growth trajectory, and might be accomplished by simply speeding up each stage of the viral cycle shown in Scheme 1. Consider doing the following:

1. Quickly delivering the key positive user experience (aka. critical event or [magic moment](https://medium.com/egyptian-startup-manual/how-to-conduct-aha-moment-aka-magic-moment-analysis-without-knowledge-of-statistics-or-data-10e59c38ee5)) that brings a new user from evaluation to loyalty mode. Optimize the sequence of navigation choices the user is presented with ([conversion funnel](https://www.hotjar.com/blog/funnel-analysis/)) to get them to this positive experience as fast as possible.
2. Incorporate a “share” feature alongside or shortly after the magic moment so that the satisfied user can invite friends while the positive experience is fresh in their mind. Consider introducing expiring incentives for invitations.
3. Reduce barriers to new user signup. In addition to increasing acceptance rate, requiring as few input fields as possible (e.g., name, email, password) will reduce the time spent on this step.

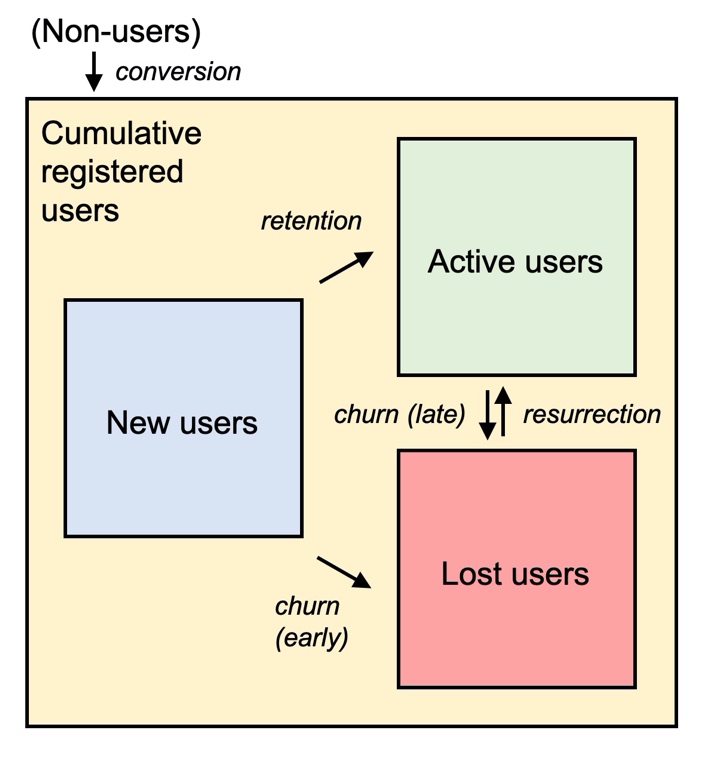
***II. Retention and churn***

While it’s adding new users is essential, it is ultimately *active users* who contribute to revenue by clicking ads or making a purchase. For this reason, it’s important to focus on these users, separating them from the rest of the user base. Indeed, active users are often a small fraction of cumulative registered users: the average mobile app, for instance, [loses 80% of its daily active users](https://andrewchen.co/new-data-shows-why-losing-80-of-your-mobile-users-is-normal-and-that-the-best-apps-do-much-better/) within just 3 days.

*1. Categorizing users*

Most companies can claim as registered users only a tiny fraction of the global internet-connected population. Within this group of registered users, there are three smaller subsets (Scheme 2) of users:

1. New users: those who have signed up within the specified time interval
2. Active users: those who have taken a particular action (see next section) within that time
3. Lost users: those who have not taken the action within that time

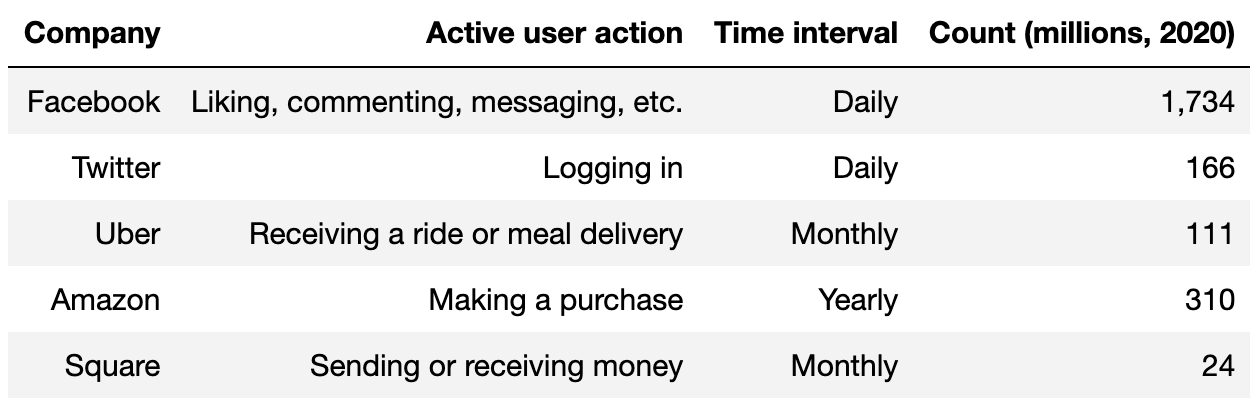


**Scheme 2**. Users within the cumulative registered user base include new, active, and lost (churned) users. New users are either retained as active users or churn due to inactivity. Active users can churn, and churned users can reactivate.

*2. Defining an active user*

An active user is one who has completed some action, ideally one tied to revenue generation, within a particular time interval. Publicly-traded tech companies typically share their active user count and definition of the action that distinguishes active- from inactive users in quarterly investor presentations (Table 1).

It’s also necessary to stipulate the time interval in which an action must be taken for the user to be considered active. This is often tied to the frequency with which users are expected to engage with the product. Social media offers up free, personalized content whenever the user has a moment to browse, and many users visit daily. On the other hand, ride-hailing, e-commerce, and mobile payments platforms address a specific need, typically cost money, and are thus used less frequently. Very few people are daily Uber riders, for example, and focusing on that population would amount to ignoring the core customer base.



**Table 1**. Active user definition and count for selected companies. See associated references: [Facebook](https://s21.q4cdn.com/399680738/files/doc_financials/2020/q1/Q1-2020-FB-Earnings-Presentation.pdf), [Twitter](https://s22.q4cdn.com/826641620/files/doc_financials/2020/q1/Q1-2020-Shareholder-Letter.pdf), [Amazon](https://etaileast.wbresearch.com/blog/amazons-engaged-buyers-drive-social-media-revenue), [Uber](https://s23.q4cdn.com/407969754/files/doc_financials/2019/sr/InvestorPresentation_2020_Feb13.pdf), [Square](https://s21.q4cdn.com/114365585/files/doc_financials/2019/q4/2019-Q4-Shareholder-Letter-Square.pdf).

*3. Visualizing user retention*

2. Retention

A. N-day retention

B.

Caveat:

-viral coefficient

-cycle time

-caveat: not all new users arrive as a result of contact with an existing user

*Churn*

-retention

-churn rate/probability

-customer lifetime

*Profitability*

-customer lifetime value

-customer acquisition cost

-payback period