**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 as of 2020 [t.ly/1VZm] ) 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).

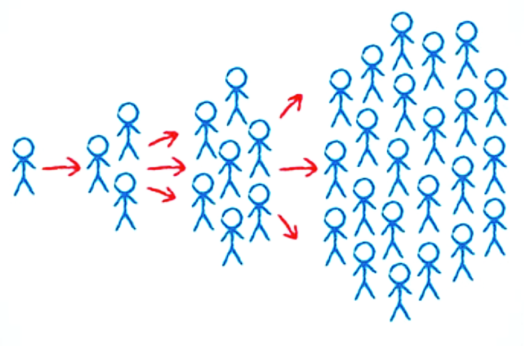
On the other hand, growth analytics tools may be of limited use to companies whose core products are not online do not have the benefit of easily monitoring user engagement. Similarly, business-to-business (“B2B”) companies typically court a small number of prospective customers and often customize their 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] chapters 16, 17, Amplitude [t.ly/P9qB] ) and video (Alex Schultz [t.ly/Q6G3], Elliot Shmukler [t.ly/3pZZ5]) resources.

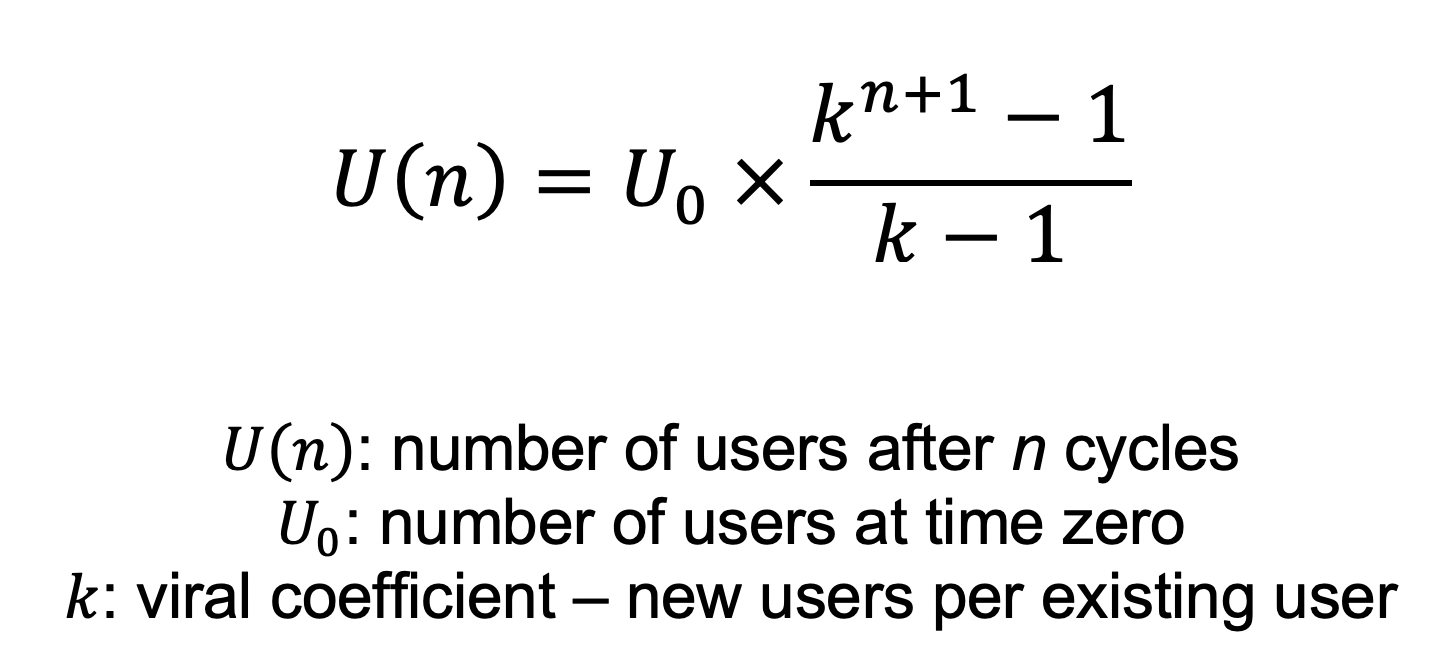
*Growth*

*1. Viral growth model*

Borrowed from epidemiology [t.ly/xdt0], the viral growth model allows companies to project the size of their user base over time.

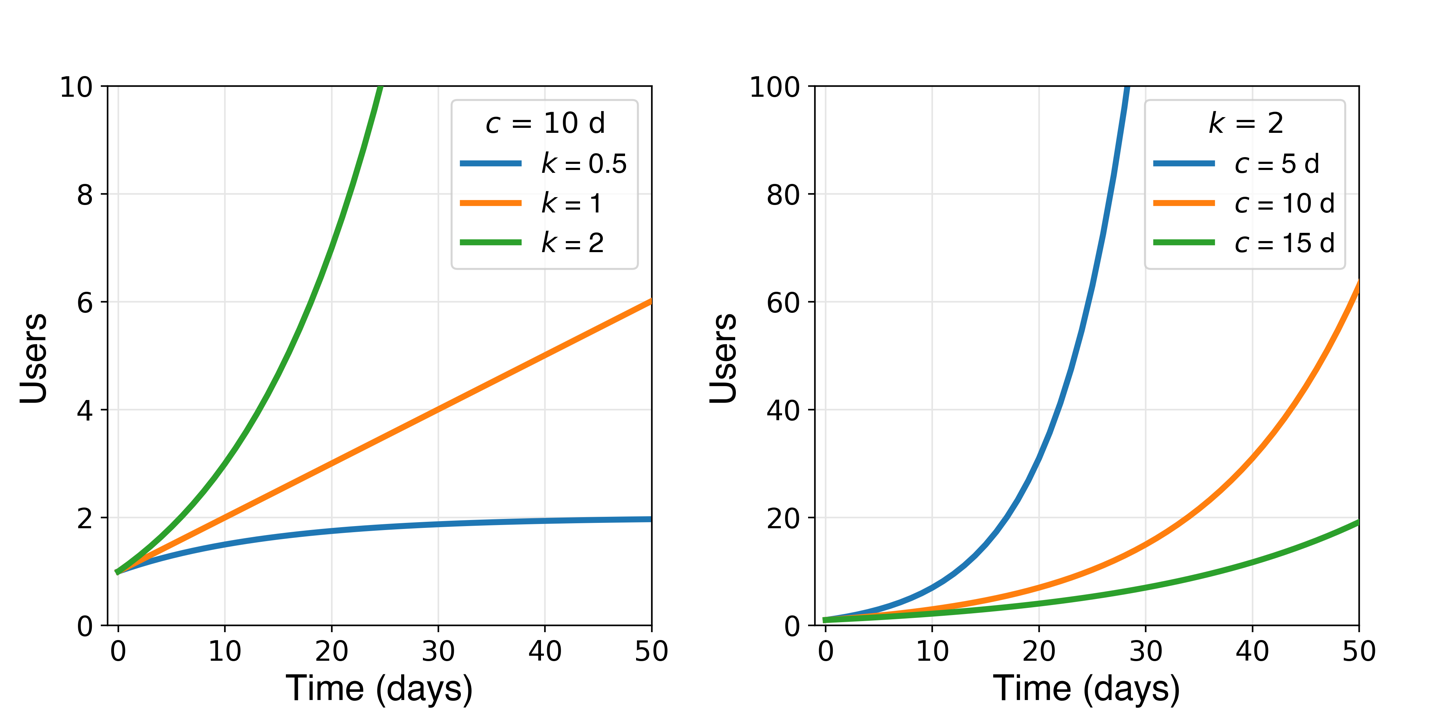


Specifically, the number of users at any particular time can be expressed in terms of current user base *U*0, the number of viral cycles *n*, and the number of new users per existing user *k*, also known as the viral coefficient:

  
Eq. 1

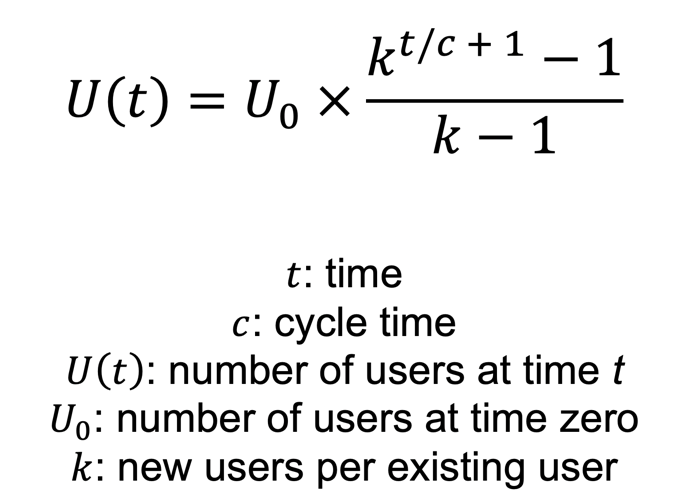
The key assumption is that new users engage with your product as a result of interaction with current users. Practically speaking, a company can empirically determine its viral coefficient *k* by multiplying the invitation rate of a current user by the acceptance rate of an invited user. For example, 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.

To bring time into the analysis, the viral coefficient *k* is raised to the power of *n* + 1, where *n* 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.

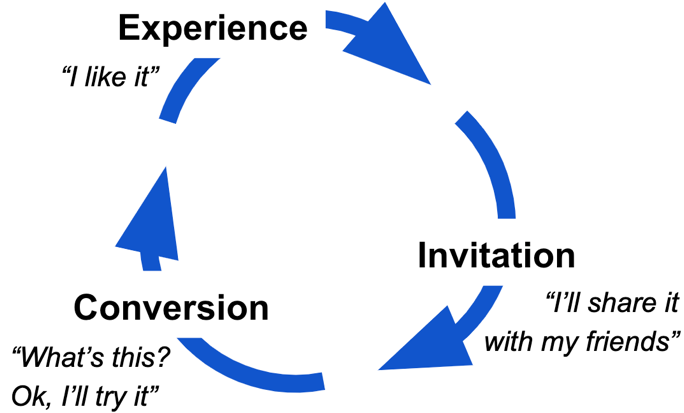


**Figure 1**. Effect of viral coefficient *k* and cycle time *c* on growth of user base. (Left): linear growth results for *k* ≈ 1 (note: *U* is not defined at *k* = 1) while *k* > 1 and *k* < 1 yield exponential growth and saturation, respectively. (Right)

Equation 1 can be expressed more generally by replacing the number of viral cycles *n* with *t*/*c*, where *t* refers to time elapsed and *c* is the cycle time:

  
Eq. 2

Accordingly, for a given time *t* elapsed, a reducing the cycle time *c* is equivalent to increasing the number of elapsed cycles. In practical terms, companies with a viral growth model can calculate cycle time as the average time interval separating the conversion of one user and conversion of the user(s) they bring in (Scheme 1).



**Scheme 1**. Viral cycle.

A few 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 reduce *k.*
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.

Given these shortcomings, one might argue that the viral growth model is more suited to quantify the two key levers (*k* and *c*) driving the rate of user growth than to accurately predict the number of users at some point in the future.

*2. Increasing viral growth rate*

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