[Become a member](https://medium.com/membership?source=upgrade_membership---nav_full------------------------)

[Sign in](https://medium.com/m/signin?operation=login&redirect=https%3A%2F%2Fmedium.com%2F%40josh.temple%2Fhow-to-estimate-the-value-of-your-customers-the-right-way-57c63fad093&source=post_page-----57c63fad093---------------------nav_reg-)

Get started

Top highlight

c:\users\syue003\onedrive\clv - how to estimate the value of your customers the right way (lifetimes) .docxc:\users\syue003\onedrive\clv - how to estimate the value of your customers the right way (lifetimes) .docxImplementing a probabilistic model for customer lifetime value

[Josh Temple](https://medium.com/@josh.temple?source=post_page-----57c63fad093----------------------)

Follow

[Mar 31](https://medium.com/@josh.temple/how-to-estimate-the-value-of-your-customers-the-right-way-57c63fad093?source=post_page-----57c63fad093----------------------) · 10 min read

When it comes to customer lifetime value (CLV), **most people are doing it wrong**, [according to Wharton marketing professor Peter Fader](https://www.youtube.com/watch?v=guj2gVEEx4s). At face value, CLV is an easy concept to understand —it’s a measurement of how much a business’s customers are worth over their lifetime. In practice, it’s deceptively hard to implement in a way that accurately captures the variation in customer behavior. CLV is so valuable to every business that it’s worth putting in the time and study to estimate it properly.

To help you estimate CLV the right way, we’ll walk through the formal definition, examine the pitfalls that plague the most popular approaches, learn some theory behind the buy-til-you-die model for CLV, and see how to implement a model with the lifetimes package in Python.

**Defining customer lifetime value**

Customer lifetime value is a prediction of how much value (in most cases, monetary value) a customer will bring to our company over their lifetime. The formal definition of customer lifetime value is as follows:

*The present value of the expected sum of discounted cash flows of an individual customer.*

We should intuitively grasp that the value of a customer over their lifetime is based on the sum of all their purchases (i.e. *cash flows*). We’ll come back to the *present value* and *discounted* parts of this definition. Determining how much a customer will spend in the future is the hardest part, and that’s what we’ll spend most of our time on.

Once we have an understanding of how much each customer is worth, we can…

* Determine the characteristics of our most valuable customer relationships and seek out customers with similar traits
* Understand how much we should spend to acquire a particular type of customer
* Push the channels that bring us our most valuable customers
* Reach out to our best customers for market research and product feedback

**Defining the business context**

Before we get started, we need to make an important distinction between two types of customers (**contractual** and **non-contractual**) and their purchase opportunity (**discrete** or **continuous**). The combination of customer type and purchase opportunity defines the**business context***,*which will affect our choice of CLV estimation approach.

Contractual customers are subscription customers — they churn on a defined date if they choose not to renew. Non-contractual customers exist in e-commerce or retail businesses, where a customer can purchase with a varying frequency and amount. We can’t confidently define churn in the non-contractual setting, as it’s impossible to be sure if a customer is “churned” or just purchasing infrequently.

A typical e-commerce company has a continuous purchase opportunity — the customer can buy at any time. However, some businesses have discrete purchase opportunities, where the customer can only buy at a specific time (e.g. prescription refills or concert tickets).

We must choose a CLV estimation approach that is appropriate for the type of customer we’re evaluating. For the purpose of this guide, I’ll be considering the **non-contractual, continuous setting**.

**What’s wrong with the popular approach to CLV?**

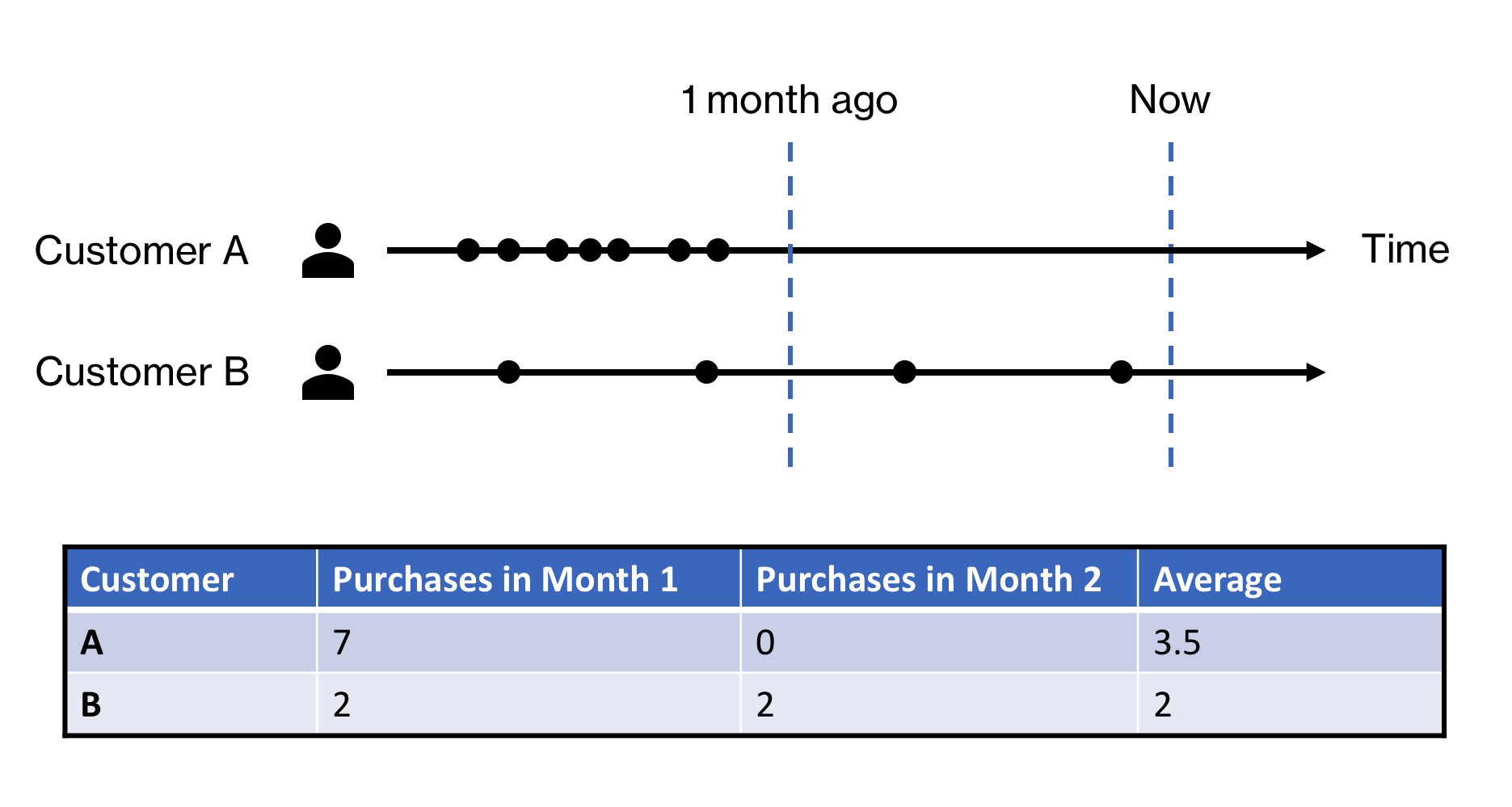
Google “*how to calculate customer lifetime value”* and you’ll find pages of tutorials from seemingly credible marketing brands like [HubSpot](https://blog.hubspot.com/service/how-to-calculate-customer-lifetime-value), [Kissmetrics](https://neilpatel.com/blog/how-to-calculate-lifetime-value/" \t "_blank), and [Optimizely](https://www.optimizely.com/optimization-glossary/lifetime-value/). Unfortunately, the popular approach on the front page of Google is oversimplified and sometimes is just plain wrong. Here’s an example of how CLV is [often described](https://www.optimizely.com/optimization-glossary/lifetime-value/):

*[CLV] is calculated by multiplying the Average Order Value by number of Expected Purchases and Time of Engagement.*

Why is this a misleading definition? First of all, we really shouldn’t be thinking about CLV as a calculation (in the same way that we calculate metrics like average order value or click-through rate). **CLV is fundamentally an *estimation of an unknown*, not a calculation of a known.** Sure, we can look at the past revenue total of a customer, but if we want to estimate the value of a customer over their lifetime, we have to introduce a predictive component that describes their future purchasing behavior. In statistics terms, we’re curious about the *expected* *value* of a customer. There will always be uncertainty within our final answer for CLV.

Secondly, this popular approach **relies too heavily on averages over the aggregate** of all your customers. The actual values for order frequency and order value can vary significantly from customer to customer. Averaging is not descriptive enough to capture the distribution of these parameters as they exist in reality.

For example, imagine two customers — Customer A placed a bunch of orders about a month ago, but then lost interest in our company and disappeared. Customer B buys less frequently, but has been a regular customer. If we base our CLV estimate on average purchase frequency and average order value, we would be misled into thinking that Customer A will be more valuable than Customer B over time. We’ve failed to account for the likelihood of the “death” of Customer A.



Purchasing behavior for two different customers over a two-month period.

If we want to estimate CLV correctly, we need to start thinking about how to estimate lifetime value for each customer and not only for the aggregate. Of course, that estimation is much more challenging than multiplying a few averages together, but our end result should be much more accurate.

**The buy-til-you-die model**

In 1987, a group of researchers from Wharton and Columbia [described a model](https://www.jstor.org/stable/2631608) called the Pareto/NBD for estimating the number of future purchases a customer will make. The Pareto/NBD model came to be part of a of a larger family of variations known as buy-til-you-die (BTYD) models.

BTYD models estimate the purchasing behavior of a customer through two stochastic processes — the probability that a customer places a repeat purchase and the probability that a customer “dies,” or permanently stops purchasing. You can think of this as two coins that are flipped continuously. The first coin dictates whether or not the customer places a purchase. The second coin dictates whether or not the customer stays alive. Each coin has its own probability of heads drawn from its distribution.

We don’t need that much information to estimate the parameters that govern these distributions. The driving factors for this model are **recency** and **frequency**, where recency refers to the time since a customer’s last purchase, and frequency refers to the number of repeat purchases placed by that customer in the given time period. These factors make intuitive sense — a customer who bought frequently in the past, but hasn’t bought recently is probably dead. A customer who bought infrequently in the past may not be dead but not yet ready to buy again.

Our goal with the BTYD model is to estimate the expected number of future purchases each customer will place over a time period given their recency and frequency.

**E[X(t)]** = expected number of transactions in time of length t  
 *(given a customer's recency and frequency)*

Once we have that expected number of future purchases for each customer, we can multiply it by that customer’s average order value to get their **residual lifetime value**(RLV). There are a number of ways we can estimate a customer’s average order value — we’ll go over one called the [Gamma Gamma model](http://www.brucehardie.com/notes/025/) later.

**RLV** = expected future purchases \* expected average order value

Residual lifetime value is the amount of additional value we expect to collect from a customer over a given time period. **From here, CLV is easy to get to — just add the sum of each customer’s past purchases to their RLV.** Though we may not be able to model an individual customer’s behavior that closely, the BTYD model performs very well for the aggregate because we’re able to capture variation in each customer’s purchasing behavior.

In 2003, Peter Fader and Bruce Hardie published [their seminal paper](http://brucehardie.com/papers/018/fader_et_al_mksc_05.pdf) on a somewhat simplified version of the Pareto/NBD called the Beta-Geometric/NBD (BG/NBD) that was much easier to implement. They dedicated a significant amount of time to making this model as accessible as possible — building an [R package](https://cran.r-project.org/web/packages/BTYD/index.html), [an Excel template](http://brucehardie.com/notes/004/), and [writing extensively](http://brucehardie.com/notes/) on how to think about and model CLV. Their work on this topic helped to establish the BTYD model as the preferred baseline for estimating CLV in the non-contractual business context.

In the spirit of keeping this extra practical, I won’t go into the math or distributions behind these models, but I encourage you to read the papers linked at the end of this guide to develop a deeper conceptual understanding before you implement this.

**Implementing a BTYD model with**lifetimes

When it comes to implementing a BG/NBD model, most of the work has already been done for us. [Cameron Davidson-Pilon](https://github.com/CamDavidsonPilon), former head of Data Science at Shopify, built a great [Python package](https://github.com/CamDavidsonPilon/lifetimes) called lifetimes that does most of the heavy lifting (if R is more your thing, check out [BTYD](https://cran.r-project.org/web/packages/BTYD/index.html)). We’ll be using the GammaGammaFitter which uses a [Gamma Gamma model](http://brucehardie.com/notes/025/) to estimate average order value and a BG/NBD model to estimate future purchase frequency.**The package is fairly**[**well-documented**](https://lifetimes.readthedocs.io/en/latest/)**, so I’ll refer you to the documentation for specific code examples.**

To use lifetimes, we’ll need a transaction event log with a customer ID, date of order, and order amount. If a customer ordered multiple times on the same day, sum their orders on that day together.

Next, we’ll use summary\_data\_from\_transaction\_data to transform our event log into recency, frequency, and monetary value ([RFM](https://en.wikipedia.org/wiki/RFM_(customer_value))) for each customer. These columns contain the following attributes for each customer:

* T is the length of the period of observation for the customer based on the value used for freq (the default is frequency is day). This can also be thought of the length of time between a customer’s first purchase and the end of the period, measured in increments defined by freq.
* recency is the point (relative to their own period of observation) at which the customer made their most recent *repeat* purchase. This can also be thought of the length of time between a customer’s first and last purchase for the period, measured in increments defined by freq.
* frequency is the number of *repeat* purchases the customer made during the observation period.
* monetary\_value is the average monetary amount of each *repeat* purchase made by a customer.

Before we fit our model, we need to check an important assumption. The model we’ll be using assumes that monetary value and frequency are linearly independent — that is, it assumes that a customer does not start to spend meaningfully more or less as they place additional purchases. [Following the documentation](https://lifetimes.readthedocs.io/en/latest/Quickstart.html#the-gamma-gamma-model-and-the-independence-assumption), we calculate the Pearson correlation between frequency and monetary\_value for customers with repeat purchases (frequency > 0) to validate this assumption.

Next, we’ll use a GammaGammaFitter to fit our model, using our RFM dataframe (filtered so it only includes customers with repeat purchases). After fitting the model, we can use the model’s method customer\_lifetime\_value to calculate how much more we predict each customer will spend over a given period of time (this is the customer’s residual lifetime value). Note that at the time of writing, there are some[open issues](https://github.com/CamDavidsonPilon/lifetimes/issues/233) with [using frequencies other than day](https://github.com/CamDavidsonPilon/lifetimes/issues/215) for the customer\_lifetime\_value method. To understand the discount\_rate argument, we need to revisit our original definition of CLV.

*The present value of the expected sum of discounted cash flows of an individual customer.*

So what does *present value*and *discounted*mean? These terms refer to the [time value of money](https://en.wikipedia.org/wiki/Time_value_of_money), which says that cash flows that happen in the future are less valuable today than if they had occurred in the present. Essentially, the lifetime value of a customer today is equal to the sum of their future purchases, adjusted for the fact that we’ll be receiving those cash flows at a later date.

To make that adjustment, we have to input a monthly adjusted value for discount\_rate. Choosing a discount rate is beyond the scope of this discussion, but you’ll want to input one that is appropriate for the growth of your business. The default monthly discount\_rate is 1%, which is a good place to start.

Once we’ve estimated each customer’s predicted residual lifetime value with the customer\_lifetime\_value method, we can add it to the sum of each customer’s historical order values to get an estimated CLV for every customer! This is very powerful information. Eventually, you’ll want to put this model in production so it updates regularly as your transaction data grows.

The lifetimes package has a ton of other interesting methods (check out plot\_probability\_alive\_matrix), so I encourage you to explore the documentation in more detail. I’ve also linked a bunch of other helpful papers, talks, and blog posts below for reference. Implementing a predictive CLV model isn’t trivial, but it will help you gain a more nuanced and accurate understanding of the people who buy your product. You can do it!

*If you enjoyed this article, please follow me*[*here*](https://medium.com/@josh.temple)*for more content about data engineering and analytics in small to mid-size companies.*

**Further reading**

1. [Peter Fader](https://marketing.wharton.upenn.edu/profile/faderp/) & [Bruce Hardie](http://brucehardie.com/)’s paper, [Counting Your Customers the Easy Way](http://brucehardie.com/papers/bgnbd_2004-04-20.pdf)
2. [Peter Fader](https://marketing.wharton.upenn.edu/profile/faderp/) & [Bruce Hardie](http://brucehardie.com/)’s paper, [What’s Wrong With This CLV Formula?](http://brucehardie.com/notes/033/)
3. [Jean-Rene Gauthier](https://www.linkedin.com/in/jr-gauthier/) & [Ben Van Dyke](https://www.linkedin.com/in/btvandyke/)’s 2017 [talk on CLV models](https://www.youtube.com/watch?v=gx6oHqpRgpY) at PyData
4. [Ben Keen](https://www.linkedin.com/in/benalexkeen/)’s [post](http://benalexkeen.com/bg-nbd-model-for-customer-base-analysis-in-python/) on implementing a BG/NBD model in Python from scratch.
5. [Roberto Medri](https://www.linkedin.com/in/robertomedri/)’s 2012 [presentation](https://conferences.oreilly.com/strata/stratany2012/public/schedule/detail/25680) on CLV and BG/NBD model at Etsy.

c:\users\syue003\onedrive\clv - how to estimate the value of your customers the right way (lifetimes) .docx