

Growth that compounds

What is RFM analysis?

RFM analysis ranks customers by how recently and how often they bought and what they spent per purchase. It describes the past; forecasts need a model.

Author	Published	Reading time	Words
Sinan Isoglu, MBA (Quantic)	18 October 2026	10 min	2,199
Commercial growth leader, lecturer and doctoral researcher			

Management summary

RFM analysis summarises each customer's purchase history in three measures, recency, frequency and monetary value, and groups customers by them. This concept guide defines the three measures (Fader et al.), shows how an RFM code is built and what it revealed about customer value in one published study, explains why RFM describes the past and how probability models turn it into a forecast (Fader and Hardie), separates the noncontractual settings where RFM analyses are usually run from subscription settings, which the sources value from retention, and gives a one-page record for an RFM analysis. The sources are one modelling study of a single cohort of an online retailer, a review and one contractual-setting study.

Contents	§ 1	What is RFM analysis?	2
	§ 2	How is an RFM code built?	2
	Figure 1	Average customer lifetime value by RFM tercile at CDNOW	3
	§ 3	Why is an RFM score not a forecast?	4
	§ 4	How do probability models turn RFM into a forecast?	5
	§ 5	Where does RFM fit, and where does it not?	6
	§ 6	What goes into an RFM analysis record?	7
	Figure 2	An RFM analysis record	7
	§ 7	How far have these sources tested RFM?	8
		References	9

What is RFM analysis?

§ 1 of 7

RFM analysis summarises each customer's purchase history in three measures and groups customers by them. Fader et al. (2005)³ describe the practice: in direct marketing it is common to summarise customers' prior behaviour in terms of their "recency (time of most recent purchase), frequency (number of prior purchases), and monetary value (average purchase amount per transaction), that is, their RFM characteristics".

³ Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research*, 42(4), 415-430.

The measures are a description of the past. Fader and Hardie (2009)¹ note that "Initial analyses of these transaction databases are typically descriptive in nature." and that such analyses "are descriptive in nature and focus on the past behavior of the firm's customer base". Customer lifetime value, by contrast, is defined by the future. At a purely conceptual level, Fader et al. (2005) write, "It is simply the present value of the future cash flows associated with a customer", and "measures should look to the future not the past".

¹ Fader, P. S., & Hardie, B. G. S. (2009). Probability models for customer-base analysis. *Journal of Interactive Marketing*, 23(1), 61-69.

In this guide, RFM analysis means recording, for each customer and a stated observation window, the time since the last purchase, the number of repeat purchases and the average purchase value, and grouping customers by those three numbers. It describes behaviour; it becomes a statement about future value only through a model that has been checked against later data.

How is an RFM code built?

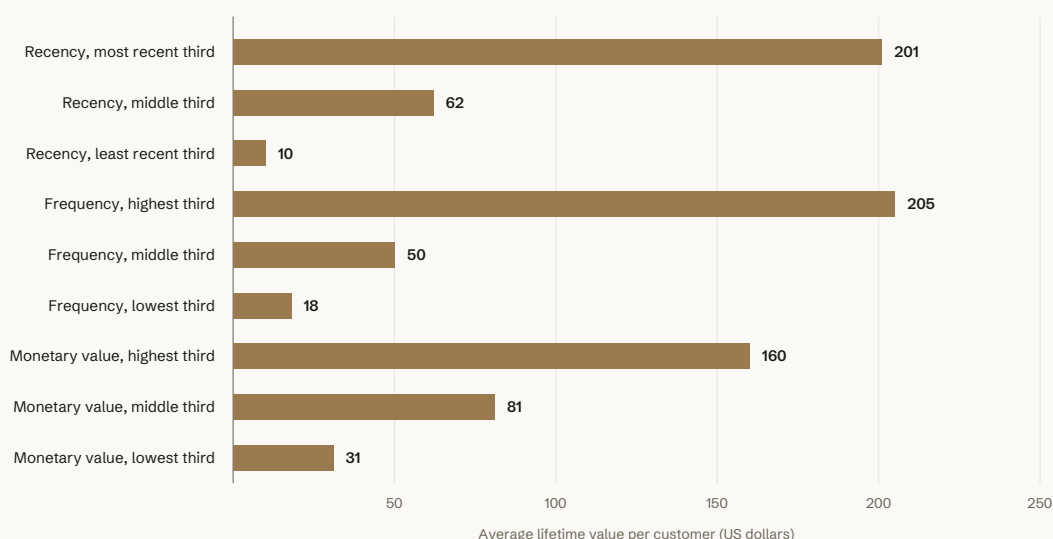
§ 2 of 7

By ranking customers on each measure and coding the ranks. Fader et al. (2005)³ show one way, using the 23,570 people who made their first purchase at the online music retailer CDNOW in the first quarter of 1997, observed for 78 weeks, less ten probable resellers they removed. They "set aside the 12,054 customers who made no repeat purchases over the 78-week observation period". For the rest, "we coded the customers in the top tercile (most recent) $R = 3$, in the second tercile $R = 2$, and in the third tercile (least recent) $R = 1$ ", and coded frequency and monetary value the same way. "We coded the customers who made no repeat purchases $R = F = M = 0$." That gives 27 cells plus a zero cell.

With a model of future purchasing, they then estimated the lifetime value of each group. Recency separated customers most: “We note that the greatest variability in CLV is on the recency dimension, followed closely by the frequency dimension. The least variation is on the monetary value dimension.” They add that “This is consistent with the widely held view that recency is usually a more powerful discriminator than frequency or monetary value”.

Figure 1 Average customer lifetime value by RFM tercile at CDNOW

Customers in the most recent third were estimated to be worth about twenty times those in the least recent third. Monetary value separated them least.



One cohort of one online retailer in 1997 and 1998. The values are model estimates of future value, not observed revenue.

Source Fader et al. (2005), Table 3, p. 426. Average estimated lifetime value per customer, in US dollars, for the 11,506 customers of one 1997 CDNOW cohort who made at least one repeat purchase, by tercile of each measure; about 3,835 customers per tercile. The authors assume a 30 per cent gross margin and a 15 per cent annual discount rate. The values are expected lifetime value beyond week 78, with no acquisition cost deducted. The authors' estimates, not a re-estimation.

The best cell and the zero cell both mattered. The cell with high values on all three measures “shows an average net present value of \$435 per customer”, and “This represents approximately 38% of the future value of the entire cohort”. But “Perhaps the most striking observation is the significant contribution of the” zero cell: customers with no repeat purchase in 78 weeks each had little value, yet “these very light

buyers collectively constitute approximately 5% of the total future value of the entire cohort, which is greater than most of the 27 other RFM cells”. On this guide’s reading, an RFM analysis that drops the customers who did not buy again drops part of the base’s value. Fader et al. draw the same lesson: “A thorough analysis of the customer base requires careful consideration of the zero class”.

Why is an RFM score not a forecast?

§ 3 of 7

Because the three measures are only signals of how a customer will behave. Fader et al. (2005)³ list the problems with the usual regression-type scoring models. “Scoring models attempt to predict behavior in the next period”, while lifetime value needs every later period. “Two periods of purchasing behavior are required”, one to define the measures and one for the outcome. And “the observed RFM variables are only imperfect indicators of underlying behavioral traits”, so a different slice of the data gives different RFM values and different model parameters. Fader and Hardie (2009)¹ put it more simply: “the past is not a perfect mirror for the future”.

The relationships are also not the straight lines a score implies. In the raw data, “It is clear that recency and frequency each has a direct and positive association with future purchasing”, but “There is a highly nonlinear relationship between recency/frequency and future transactions.” One pattern runs against intuition: “For low levels of recency, customers with greater frequency are likely to have lower future purchasing potential than customers with lower prior purchasing rates.” Fader et al. (2005)³ resolve it with two hypothetical customers whose last purchases fall in the same week: the frequent buyer’s long silence suggests that customer has stopped, while the light buyer, who buys rarely, is more likely still active, so in their model the light buyer’s discounted expected transactions come out higher, 4.6 against 1.9. Fader and Hardie (2009)¹ use a light buyer to make the same point about the chance of still being a customer: her long gap “could simply be a function of the fact she appears to be a light buyer”, and whether she will also buy more is, they write, “an open question”. Fader et al. (2005) also warn that in plots of raw data “the patterns observed will depend on the amount of time for which the customers are observed”.

How do probability models turn RFM into a forecast?

§ 4 of 7

By treating observed purchases as the result of each customer's unobserved buying rate and chance of having stopped. Fader and Hardie (2009)¹ review such probability models as tools for customer-base analysis. The original model of this kind, the Pareto/NBD, which they date to a seminal paper of 1987, assumes that customers buy at their own steady rates while they remain customers and stop at some unobserved point, the “buy till you die” story. The data it needs are modest: the only customer-level inputs required for its “individual-level predictions are recency and frequency”. Fader et al. (2005)³ add monetary value and show that, in their model, “RFM measures are sufficient statistics for an individual customer's purchasing history”, so that, as Fader and Hardie² summarise, “we only need to know three things about a customer's buying behavior in a given time period in order to compute their residual lifetime value: recency, frequency, and monetary value (i.e., RFM)”.

² Fader, P. S., & Hardie, B. G. S. (2010). Customer-base valuation in a contractual setting: The perils of ignoring heterogeneity. *Marketing Science*, 29(1), 85-93.

Such models also remove two of the scoring model's problems: “there is no need to split the observed transaction data into two periods to create a dependent variable”, and “we can predict behavior over future time periods of any length”. Fader and Hardie³²¹ report for the Pareto/NBD that “its predictive performance is impressive”, and describe a simpler variant whose “parameters can be obtained quite easily in Microsoft Excel”, noting that “the two models yield very similar results in a wide variety of purchasing environments”. The model is only as good as its assumptions: Fader et al. treat spend per transaction as independent of when and how often a customer buys, and write that “any researcher applying our model to a new data set must test the validity of the assumption of independence between recency/frequency and monetary value”. The [CLV guide](#) sets out what such a forecast is used for.

Where does RFM fit, and where does it not?

§ 5 of 7

In settings where the firm cannot see a customer leave. Fader and Hardie (2009)¹ classify customer bases on two questions. The first is whether the relationship is contractual. In a noncontractual setting, “a key characteristic of which is that the time at which a customer becomes inactive is unobserved by the firm”, and “The key challenge of noncontractual settings is how to differentiate those customers who have ended their relationship with the firm (without informing the firm) from those who are simply in the midst of a long hiatus between transactions.” This is why a fixed definition of an active customer, such as a purchase in the last twelve months, misleads: “this 12-month cut-off point is arbitrary”. The second question is whether purchases can happen at any time or only at set points: “are the opportunities for transactions restricted to discrete points in time or can they occur at any point in time?”

The boundary between the two settings is strict: “it is completely inappropriate to apply a model developed for a contractual setting in a noncontractual setting (and vice versa)”. Fader et al. (2005)³, whose own setting is noncontractual, note that “It is true that, in general, RFM analyses are conducted in such a domain, but the use of CLV goes well beyond it.” In a contractual or subscription setting, as Fader and Hardie (2010)² put it, following their 2009 review, “a defining characteristic of a contractual (or subscription) business setting is that the loss of a customer is observed”, and value is usually projected with retention rates. They warn that “The application of standard textbook discussions of CLV leads to calculations based on a single retention rate.”, although “at the cohort level, retention rates typically increase over time”, and “We show that failing to recognize these dynamics yields a downward-biased estimate of the residual value of the customer base”. The [cohort analysis guide](#) explains how retention is read by cohort. On this guide’s reading, a B2B firm with renewing contracts should value those accounts with a model of retention by cohort, as Fader and Hardie (2010) do, and keep RFM for purchases that can stop without the firm being told, such as orders from accounts that hold no contract. Fader and Hardie (2009)¹ place purchases and usage made under a contract, such as a credit card or mobile phone usage, on the contractual side, so a subscriber’s add-on orders are not a clean case for RFM.

What goes into an RFM analysis record?

§ 6 of 7

Six fields, in one record per analysis: the setting, the observation window, the definition of each measure, the grouping, the forecast and its check, and the decision each group feeds. The fields follow Fader and Hardie's³²¹ classification, Fader et al.'s definitions and coding, and their warnings about windows, nonlinearity and the zero cell.

Figure 2 An RFM analysis record

Write down the setting, the window, the three definitions, the grouping, the forecast check and the decision in separate fields before acting on a score.

Setting	Window	Measures	Grouping	Forecast	Decision
Can the firm see a customer leave? Any time or set points?	Which period? Would a longer one change the groups?	How are recency, repeat count and average value defined?	Terciles or quintiles? Where do non-repeaters go?	Which model, and checked against which later period?	Which action does each group get, and who owns it?

If the firm can see customers leave, value them from retention by cohort, not from recency.

One row per analysis. A score with no forecast check is a description, not a prediction.

Source Author's decision worksheet. The setting field follows Fader and Hardie (2009), p. 63, and Fader and Hardie (2010), p. 85; the window, measures and grouping fields follow Fader et al. (2005), pp. 415–417 and 425; the forecast field follows Fader and Hardie (2009), pp. 62 and 65, and Fader et al. (2005), p. 426. The fields are reader inputs; no company data is supplied.

An illustration, with no company data: a firm sells a software subscription and also sells training seats and add-on modules ordered as needed, some of them to accounts that hold no subscription. The record marks the subscriptions, and the add-on orders of subscribers, as contractual and values those accounts with a model of retention by cohort. For the orders of accounts without a subscription it fixes a two-year window, defines recency as the time since the last order and monetary value as the average order value, keeps the accounts with a single order as their own group, checks a model's forecast against the

following quarter before using it, and names the action for each group, for example a review by the account manager for frequent buyers who have gone quiet. Whether a group is worth acting on is the subject of the segmentation essay. These are the author's examples and have not been tested.

How far have these sources tested RFM?

§ 7 of 7

On one cohort, with models from one team of authors. Fader et al. (2005)³ estimate value for one cohort of one online retailer, and list the limits themselves: “We also recognize that our context of noncontractual purchasing is a limitation.” and “Another limitation is that we examined only one cohort for CDNOW.” Their values rest on an assumed margin, because “we have no information about the actual margins for CDNOW”. Fader and Hardie (2009)¹ review models and their use rather than test RFM groups against each other, and Fader and Hardie (2010)² work through a hypothetical contractual firm whose retention comes from a stylised two-segment example, then apply parameters fitted to two actual data sets in the authors' earlier work to the same hypothetical firm to size the bias. All three works share the same two authors, Fader and Hardie (the 2005 study adds Lee), and this guide cites no independent test of their models. None of the three studies B2B accounts. The B2B illustration in this guide is the author's reasoning.

 End of essay

References

- 1** Fader, P. S., & Hardie, B. G. S. (2009). Probability models for customer-base analysis. *Journal of Interactive Marketing*, 23(1), 61-69. <https://doi.org/10.1016/j.intmar.2008.11.003>
- 2** Fader, P. S., & Hardie, B. G. S. (2010). Customer-base valuation in a contractual setting: The perils of ignoring heterogeneity. *Marketing Science*, 29(1), 85-93. <https://doi.org/10.1287/mksc.1080.0482>
- 3** Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research*, 42(4), 415-430. <https://doi.org/10.1509/jmkr.2005.42.4.415>

How to cite this essay

Isoglu, S. (2026, October 18). What is RFM analysis? Sinan Isoglu.
<https://isoglu.com/insights/journal/what-is-rfm-analysis/>
Online version: <https://isoglu.com/insights/journal/what-is-rfm-analysis/>

Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher

© 2026 Sinan Isoglu · isoglu.com · All rights reserved