

What is lead scoring?

Lead scoring ranks leads by how likely they are to reach a named outcome. Without that outcome, a time window and a population, a score says little.

Author	Published	Reading time	Words
Sinan Isoglu, MBA (Quantic) Commercial growth leader, lecturer and doctoral researcher	18 October 2026	16 min	3,557

Management summary

Lead scoring attaches to each lead an estimate of how likely it is to reach a named outcome, or of what that outcome would be worth, so that leads can be ranked. This concept guide defines it, separates scores set by rules from scores fitted to data, and separates scoring from qualification and routing. It reports a survey of sales reps in four B2B firms in which reps who rated their firm's lead screening higher spent a larger share of their time on marketing leads, by different amounts in each firm, and two models whose accuracy has to be read against the commonest outcome. It gives a record a score should carry before anyone acts on it. The sources are one survey, one model tested on simulated data and on one seller's data, and one firm's CRM records.

Contents	§ 1	What is lead scoring?	2
	§ 2	What must a score name before it means anything?	3
	§ 3	Who sets the weights: people or data?	4
	§ 4	Is a score the same as qualification or routing?	5
	§ 5	What happens to a lead after the score?	5
	Figure 1	How strongly reps' time on marketing leads rose with their rating of the lead screen	6
	§ 6	Why can an accurate-looking score still mislead?	8
	§ 7	What should travel with a score into the sales queue?	9
	Figure 2	A lead score record	10
	§ 8	What would a software team still have to test?	11
		References	12

What is lead scoring?

§ 1 of 8

Lead scoring attaches to each lead a number or a class that estimates how likely the lead is to reach a named outcome, such as a first meeting or a signed order, or what that outcome would be worth. The estimate is used to rank leads, so that the people who follow them up start where the payoff is most likely. None of the three studies this guide draws on uses the phrase. Each covers one part of it: how sales reps respond to the screening that decides which leads reach them, a model that estimates a buyer's stage from its behaviour, and a model that explains which leads were won, lost or canceled.

The screening step Sabnis et al. (2013)² studied is, on this guide's definition, where a score meets its threshold. In their words, "Prequalification is a process by which the marketing department screens the leads it has generated and discards those it deems unattractive." It is done inside the firm: "lead prequalification is typically executed internally by marketing personnel". The survey items they used describe what such a screen is meant to do from the sales rep's side, for instance "My company is very good at giving only hot leads to salespeople."

² Sabnis, G., Chatterjee, S. C., Grewal, R., & Lilien, G. L. (2013). The sales lead black hole: On sales reps' follow-up of marketing leads. *Journal of Marketing*, 77(1), 52-67.

Marvasti et al. (2021)¹ put the scoring question in their title, "Is this company a lead customer?", and answer it with a model: "the authors propose a machine learning approach to infer the stages of the B2B buying journey by observing the online browsing behavior of buyer companies". The stage itself is not observed, since "These buying stages are hidden from the sellers"; the browsing stands in for it. Virtanen et al. (2015)³ modelled the outcome of each of 4,194 sales leads in one IT solutions company. "The final dataset contains 4194 sales leads, and each lead is nested under a certain customer." A model of that kind, built on past leads and their outcomes, is what a model-based score rests on.

¹ Marvasti, N. B., Huhtala, J.-P., Yousefi, Z. R., Vaniala, I., Upreti, B., Malo, P., Kaski, S., & Tikkanen, H. (2021). Is this company a lead customer? Estimating stages of B2B buying journey. *Industrial Marketing Management*, 97, 126-133.

³ Virtanen, T., Parvinen, P., & Rollins, M. (2015). Complexity of sales situation and sales lead performance: An empirical study in business-to-business company. *Industrial Marketing Management*, 45, 49-58.

In this guide, a lead score is an estimate, attached to one lead, of how likely that lead is to reach a named outcome within a stated time window, or of what that outcome would be worth, made for a stated population of leads from information available when the score is set. Lead scoring is the practice of producing such estimates and ranking leads by them. A score ranks; it does not decide. Turning it into a decision takes a threshold, which is qualification, and an owner, which is routing.

Likelihood and value are different estimates, and not all sales reps weigh them the same way. Sabnis et al. note, citing earlier research, that “Sales reps with better previous sales performance are more likely to follow up on leads that offer a lower likelihood of closure but higher revenue potential, rather than the reverse”. On this guide’s reading, a score that ranks by likelihood alone can therefore order leads differently from the way a firm’s strongest reps would.

What must a score name before it means anything?

§ 2 of 8

Three things: the outcome it predicts, the window in which the outcome has to happen, and the population of leads it was built for. The sources show how easily each one shifts.

The outcome first. In Marvasti et al.¹’s real data, the stages to be predicted came from a seller’s opportunity records, and “there are 15 different opportunity types defined in this dataset mapped to previously described, four buying states with the help of a sales specialist”. In that mapping, “Closed Won” and “Closed Lost” both count as “Late Funnel” (Table 6, p. 131). A model that is right about the late stage is not, for that reason, right about who buys. Sabnis et al.²’s survey fixed the end of follow-up with three outcomes: “We count follow-up as starting when you contact a prospect in any way at all and ending when you know what the prospect will do (you get the order, somebody else gets the order, or you write them off as hopeless).” A score that predicts the first of those is a different score from one that predicts only that a lead will not be written off.

The window next. Marvasti et al.¹ checked each day’s estimate against the stage of an opportunity dated close by: “Note that we have considered an interval of 2 weeks before and after each opportunity date as a valid interval for each buying stage.” With a different interval the same model would face a different test, so the window belongs beside any accuracy figure.

Then the population. A browsing model can score only the companies that browse: in Marvasti et al.¹'s real data that was 648 companies that visited the seller's website in just over a year. Virtanen et al.³ narrowed their leads before modelling them: "Leads that did not include all necessary data points, or had low revenue, were ruled out of the study." A score holds for the kind of lead it was built on, and says nothing yet about the leads that were left out.

Who sets the weights: people or data?

§ 3 of 8

In a rule-based score, people decide what counts and by how much: points for a job title, a company size or a download, and a threshold above which marketing passes the lead on. Sabnis et al.²'s prequalification is a screen set by people in this sense, run by marketing staff, and the reps they surveyed judged it by whether it filtered: one item read "My company does a terrible job of filtering out cold leads", scored in reverse. In a model-based score, the weights come from data. In Marvasti et al.¹'s hidden Markov model the parameters are not set in advance; "it should initially be estimated from the data (observed buying behavior) itself". Virtanen et al.³ fitted a multinomial logistic regression to the recorded outcome of each lead.

The two need different things. A rule needs someone to decide what counts, and it stays a guess until it is checked against outcomes. A model needs data to learn from and outcomes to be checked against, and the outcomes are recorded by people too. Marvasti et al.¹'s model learned from browsing alone; the seller's opportunity records served only to check it, and the authors warn that "the provided ground truth may not be precise, and the seller may have missed existing opportunities on some dates". An accuracy figure measured against incomplete records measures the records as well as the model, and a model fitted to such records, as Virtanen et al.³'s was fitted to the outcomes salespeople recorded, learns the records as well as the buyers. No source compares a rule-based with a model-based score on the same leads, so this guide does not say which of the two ranks better.

Is a score the same as qualification or routing?

§ 4 of 8

No. Scoring estimates, qualification decides and routing assigns. Qualification is the yes-or-no decision that a lead is good enough to pass on, which is what Sabnis et al.² call prequalification. A score can feed it through a threshold, but the threshold is a separate choice with its own cost: every lead just below it is treated as not worth a call. Routing then gives an eligible lead an owner or a queue. The [lead routing guide](#) draws the same line from the other side: a score may rank or prioritise records, routing assigns them, and the assignment record does not show whether a lead was any good. The [product-qualified lead guide](#) covers one kind of qualification rule, a threshold on product use within a stated window.

What happens to a lead after the score?

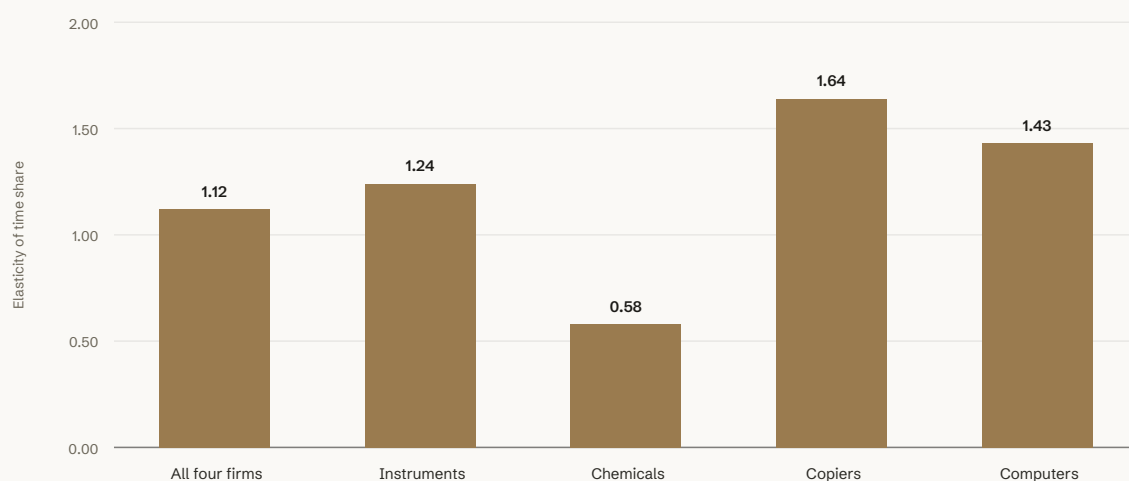
§ 5 of 8

It reaches a person who decides whether to spend time on it, and who judges it partly by what they think of the screen it came through. Sabnis et al.² surveyed sales reps in four B2B firms, one each in scientific instruments, chemicals, copiers and computers, and analysed the answers of 461 of them. Their starting point: “Sales reps cannot observe the objective quality of leads; they are motivated by their perceptions of lead quality or, more operationally, of the marketing department’s lead qualification process.” They measured that perception with survey items and related it to the share of each rep’s working time spent on marketing leads, as against leads the reps found themselves and all other work.

Reps who rated the screen higher spent a larger share of their time on marketing leads, and the authors report this as support for their first hypothesis, “which suggests that as the perceived quality of prequalification increases, the proportion of time spent on marketing leads increases”. Stated as an elasticity for the whole sample, “a 1% increase in perceived prequalification quality produces a 1.12% increase in marketing lead follow-up time allocation”, and “The elasticity is positive across all four firms but varies considerably (from .58 to 1.64).” The authors write “produces”; the data come from one survey at one point in time, so this guide reads the figure as an association. The link was stronger among more experienced reps: the authors found support for their argument that “experience enhances the positive effect of the perceived quality of prequalification on the proportion of time spent following up marketing leads”. Figure 1 shows the elasticity overall and by firm.

Figure 1 How strongly reps' time on marketing leads rose with their rating of the lead screen

In every firm, reps who rated the screen higher gave a larger share of their time to marketing leads, but in the authors' point estimates the strength of that link differed about threefold between firms.



An elasticity of 1.12 means that, in the authors' model, a 1 per cent higher rating of the screen went with a 1.12 per cent larger share of time on marketing leads. The rating is each rep's own answer to survey items, not a measure of the leads. The same table puts the elasticity for managerial tracking at $-.14$ and for lead volume at $-.07$ across the sample. Table 3 gives no interval or test for any elasticity. The firm values come from a model that pulls each firm toward the sample-wide estimate, Firm 2 rests on 30 reps, and the sample-wide value is the model's own estimate, not an average of the four firm bars.

Source Sabnis et al. (2013), p. 63, Table 3, row Quality of prequalification: elasticity of the share of a sales rep's time spent on marketing leads with respect to the rep's perceived quality of the firm's lead prequalification, in the authors' Dirichlet regression of survey data. Reps analysed: 51 in Firm 1 (scientific instruments), 30 in Firm 2 (chemicals), 100 in Firm 3 (copiers) and 280 in Firm 4 (computers), 461 in all (Table 1, p. 59; industries pp. 59 and 63).

Two limits keep this from being a finding about scores. The study measured shares of time, not leads: the authors say their implications “may not extend fully to related constructs such as actual time spent on marketing leads (in hours) or the number of leads followed”, and that “Lead-level data from firms’ customer relationship management systems might make such analyses possible.” And the screen was rated by the reps, not tested against outcomes: “the extent of prequalification quality is not an objective, firm-level measure but rather is based on individual sales reps’ perceptions”, and they count

such measures among their limits, since they “inherently suffer from measurement error and perhaps some psychological bias”. What the study does show is that the share of time reps give to marketing leads goes with how highly they rate the screen behind those leads, by different amounts in different firms. The study’s other two levers, managerial tracking of follow-up and the number of leads sent, are discussed in the essay on the funnel bottleneck nobody measures.

The reps who reported no time at all on marketing leads are outside those numbers. “Our preliminary examination of the data showed that 39 respondents reported spending no time pursuing marketing leads.” They rated their firm’s prequalification significantly lower than the others, and “37 of the 39 respondents who indicated no marketing lead follow-up time came from one firm (computer industry)”, which the authors read as “a potential for systematic bias” before removing them from the analysis. The phrase in the paper’s title, the sales lead black hole, names the leads that marketing generates and sales never pursues. The share the paper gives for it comes from two industry research reports it cites, not from its own respondents, and the authors add that the figure “reflects inadequacies of marketing as well as the sales function”, since reps can receive so many leads that “it is not possible for the sales reps to pursue all of them, even if they desired to do so”.

Marvasti et al.¹ expect that sellers can use stage estimates to “improve the hit rate of marketing and sales activities”, and for the link between matching marketing to the buying stage and a higher hit rate they cite earlier work. Their study estimates stages; it does not observe what sellers did with the estimates, and none of the three sources tests whether acting on a score changes how many leads convert.

Why can an accurate-looking score still mislead?

§ 6 of 8

Because accuracy depends on how much the score sees, on what it is checked against and on how common the outcome is. Marvasti et al.¹ first tested their model on simulated buyers, where the true stage is known; each simulated dataset held 300 companies. Accuracy was 83 per cent when browsing was observed every day and 71 per cent when it was observed on about a third of days, and the authors sum up that “the fewer browsing activities are observed, the lower the performance is achieved”. In the simulations, “Both parameter learning and buying state estimations are performed on the same dataset”, so the model was checked on the same companies whose browsing it had learned from, though without seeing their true stages. On one seller’s real data, where “in each month, at most, four days include some browsing behavior”, the reported accuracy was 62 per cent.

A share of correct answers means little until it is set against the base rate: how often each outcome occurs, and so how often a rule that always names the commonest one would be right. Virtanen et al.³ made two such comparisons for their own model. It classified 63.0 per cent of leads correctly, “which is 50% better than by chance alone”, where chance meant the 41.6 per cent that guessing in proportion to the three outcomes’ shares would get right. Against the commonest outcome, a rule “which predicts that every single lead is canceled”, the gain was smaller: “Our model failed to meet a 25% improvement over this test model, as the improvement rate of our model was 17%.” Marvasti et al.¹ report no such comparison. This guide’s arithmetic on their real-data confusion matrix (Table 7, p. 132) finds most of the scored days in one stage, Late Funnel: a rule that named it every day would have been right on 65 to 76 per cent of them, against the model’s 62 per cent. The range comes from the table itself. Its labels make the columns the actual stages, but its first column then holds 342 days of the No-funnel stage, which the authors say the opportunity data never record, so the table can be read either way.

Both models sort each case into one of several classes. A lead score can instead rank leads by their chance of one rare outcome, such as a paid contract, and then the commonest outcome is that nothing happens: a rule that always predicted it would be right most of the time and would still pick out no lead. On this guide’s reading, the base-rate check for such a score is whether the leads it ranks highest, among leads it was not fitted to, reach the outcome more often than all of those leads do.

A model built to explain outcomes is not yet a score. In Virtanen et al.³'s model, "The third control variable, lead-time, had a clear effect on sales closing performance", and "lead time was calculated in quarters from lead identification until the deal was won, lost, or canceled". That suits the authors' aim, which was to explain outcomes. But lead time is known only once a lead has closed, and a score set when the lead arrives can use only what is known at that moment. On this guide's reading, an input measured up to the outcome makes a model look better at explaining than a score set when the lead arrives could be at predicting.

A score can also confirm itself. If leads with low scores get less follow-up, fewer of them convert, and the next round of data says the low scores were right. Sabnis et al.²'s results make that plausible, since reps' time went with their rating of the screen, but no source here measures it. Recording who followed up which lead, and when, is what separates a weak lead from a neglected one. This is the guide's reasoning and has not been tested.

What should travel with a score into the sales queue?

§ 7 of 8

Six, in one record per score: the outcome it predicts and how that outcome is defined, the window, the population and what was excluded, the inputs and whether each was known when the score was set, the check against the base rate on leads the score was not fitted to, and the follow-up record. The fields follow the outcome mapping and validation window in Marvasti et al.¹, the exclusions and the chance and commonest-outcome comparisons in Virtanen et al.³, the ranking check this guide adds, and the follow-up findings in Sabnis et al.²

Figure 2 A lead score record

Write down the outcome, the window, the population, the inputs, the check and the follow-up before a score decides who gets a call.

Outcome	Window	Population	Inputs	Check	Follow-up
Which event, defined how?	By when after scoring?	Which leads, and which left out?	Each known when scored?	On unseen leads, better than the base rate?	Who acted on it, and when?

A score that does no better than the base rate on leads it has not seen should not decide who gets a call.

One row per score. Check the score on leads from a later period than those used to set it. Without the follow-up column, a low score and a neglected lead look the same.

Source Author's decision worksheet. The outcome and window fields follow Marvasti et al. (2021), p. 131; the population field follows Marvasti et al. (2021), p. 131, and Virtanen et al. (2015); the inputs and check fields follow Virtanen et al. (2015) and Marvasti et al. (2021), pp. 130 to 132, and the check's ranking form is this guide's; the follow-up field follows Sabnis et al. (2013), pp. 55, 59 and 63 to 64. The fields are reader inputs; no company data is supplied.

An illustration, with no company data: a B2B software firm scores the companies that start a free trial, on the fourteenth day of the trial. The record names the outcome (a paid contract, not a booked sales meeting), the window (to day 90 of the trial, 76 days after scoring), the population (trials from business email domains, leaving out existing customers, partners and the firm's own staff) and the inputs (company size and industry from the sign-up form and product use in the first 14 days, nothing recorded later). It checks the score on trials that began after the months used to set the weights, and compares the share of paid contracts among the top tenth of scores with the share among all trials. It records which trials a seller contacted and when, so that a low score is not confirmed by neglect. To learn whether calling the top-scored trials first changes conversion, rather than whether the score predicts it, the firm would draw lots among trials above the score threshold, pass half to a seller at once and half on the normal schedule,

and compare the share with a paid contract in each half at the end of the full 90 days. These are the author's examples and have not been tested.

What would a software team still have to test?

§ 8 of 8

How large any of this is in a software business. Sabnis et al.² surveyed reps in four firms that were “major players in their industries” for decades, about shares of their time, and warn that “our conclusions might not apply to smaller or younger firms whose sales processes are still evolving”. Marvasti et al.¹ tested their model on simulated companies and on one seller's sparse browsing data, in which “the browsing data is available only at an aggregated level, where it is not possible to separate browsing associated with different products or decision-makers”. Virtanen et al.³ studied one IT solutions company's CRM records, and “the total amount of leads won, lost, and canceled cannot be mentioned in this paper”.

On this guide's reading, two things carry over to a software team with some confidence, because they follow from what a score is rather than from any one sample: a score has to name its outcome, window and population, and what it achieves means little until it is set against the base rate on leads it was not fitted to. A third is a hypothesis to check, since the authors warn it might not apply to smaller or younger firms: that the time people give to the leads they receive goes with how highly they rate the screen those leads came through. What the sources do not say is whether rules or models rank leads better, how much acting on a score changes conversion, or whether reps trust a model's score more or less than a rule they can read.

 End of essay

References

- 1 Marvasti, N. B., Huhtala, J.-P., Yousefi, Z. R., Vaniala, I., Upreti, B., Malo, P., Kaski, S., & Tikkanen, H. (2021). Is this company a lead customer? Estimating stages of B2B buying journey. *Industrial Marketing Management*, 97, 126–133. <https://doi.org/10.1016/j.indmarman.2021.06.003>
- 2 Sabnis, G., Chatterjee, S. C., Grewal, R., & Lilien, G. L. (2013). The sales lead black hole: On sales reps' follow-up of marketing leads. *Journal of Marketing*, 77(1), 52–67. <https://doi.org/10.1509/jm.10.0047>
- 3 Virtanen, T., Parvinen, P., & Rollins, M. (2015). Complexity of sales situation and sales lead performance: An empirical study in business-to-business company. *Industrial Marketing Management*, 45, 49–58. <https://doi.org/10.1016/j.indmarman.2015.02.024>

How to cite this essay

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

Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher
© 2026 Sinan Isoglu · isoglu.com · All rights reserved