

Sign in

Publish AI, ML & data-science insights to a global community of data professionals.

Submit an Article

LATEST

EDITOR'S PICKS

DEEP DIVES

NEWSLETTER

WRITE FOR TDS



DATA SCIENCE

# The Power of Framework Dimensions: What Data Scientists Should Know

Practical guidance and a case study

Chinmay Kakatkar

Oct 26, 2025 15 min read





A [previous article](#) provided a holistic overview of conceptual frameworks – analytical structures for representing abstract concepts and organizing data. Data scientists use such

frameworks in a wide variety of contexts, from use case ideation and validation of machine learning models to productization and operation of user-facing solutions. The framework type (e.g., hierarchy, matrix, process flow, relational map) and framework dimensions (e.g., categorical, ordinal, continuous) largely determine the look and feel of a conceptual framework. While the previous article devoted more space to a discussion of framework types, this article will place the spotlight on framework dimensions. With the help of a real-life case study, we will see how modifying the framework dimensions can yield a perceptual shift that can unlock new insights. This deep dive aims to better equip readers to use and build conceptual frameworks more effectively.

**Note:** All figures in the following sections have been created by the author of this article.

## **A Primer on Framework Dimensions**

Whereas the framework type defines the structure of what you are trying to represent, the framework dimensions determine the content. The dimensions generally fall into three classes: categorical, ordinal, and continuous. The following sections

examine this classification of framework dimensions in more detail and go over some aspects that you should consider when including multiple dimensions in a framework.

## **The Big Three**

Let us start with the class of *categorical* dimensions, which is possibly the simplest class of dimensions. As the name suggests, the dimension consists of a finite set of discrete categories that need not be in any particular order. For instance, if the dimension represents a company's markets, it could be divided into geographic categories such as "USA," "Germany," and "China." Similarly, you could have a categorical dimension that breaks down the company's products into different product segments (e.g., by ingredients, relevance to customers, and so on). It is also a good idea to keep the MECE principle (*mutually exclusive* and *cumulatively exhaustive*) in mind whenever you are breaking down a dimension into smaller categories; after all, you want the categories to fully cover the scope of the dimension and avoid redundant categories.

*Ordinal* dimensions are similar to categorical ones, with the additional feature that the categories making up the dimension

are also ordered in some way. The ordering allows you to say that one category is “greater than,” “less than,” “equal,” or “unequal” to another. Suppose you took a company’s set of markets and ranked them by a criterion like profitability. The ranking would impose an ordering on the set of markets, thereby producing an ordinal dimension representing the profit-based (ascending or descending) ordering of markets. However, the rankings need not imply that the profitability values of countries are evenly spaced; the profitability gap between the top-ranked and second-ranked country could be different from the gap between second- and third-ranked countries. Ordinal dimensions are also often used to construct survey questions, taking the form of a Likert scale (e.g., “disagree,” “neutral,” “agree”). The ordering allows responses across the survey participants to be analyzed in terms of where they lie on the scale for each question.

Finally, a *continuous* dimension gives a quantitative measure of something. Unlike categorical and ordinal dimensions (which consist of discrete categories or values), continuous dimensions can potentially take on any value (however tiny) within a given range. For example, the probability, in percentage terms, of some event occurring can lie anywhere between 0% and 100%; values

such as 5%, 10% and 10.00123% would all be permissible. The values of a continuous dimension are also inherently ordered.

## **Choosing Dimensions Wisely**

It is important to consider the strengths and limitations of each dimension class before applying them to your framework. For instance, you could look at the information content of each dimension class. The presence of an ordering and the ability to take on increasingly fine-grained values within a given range contribute to the depth of the information content. Based on information content, ordinal dimensions should be favored over categorical ones, and continuous dimensions should be favored over the other two whenever they can be measured in a granular, quantitative manner. However, the information richness comes at the cost of the resources needed to procure and analyze the data underlying the dimensions. Also, presenting and explaining information-rich dimensions to an audience can be hard, since there is a lot of content that needs to be unpacked and digested. As such, even if you use continuous dimensions to perform the analysis, it may make sense to “bucket” the continuous data into ordinal or even categorical data to simplify what is shown to an audience.

Furthermore, since frameworks can involve multiple dimensions, it is important to achieve an optimal interplay between the dimensions. There are at least two basic decisions that you will need to make in this regard – how many dimensions, and what kinds, to include in the framework. Especially in the early stages of analyzing a problem, the tendency is to be generous with the number of dimensions considered, since the problem may not be well-understood at this point and there is a risk of eliminating potentially valuable dimensions prematurely. But as your analysis progresses, a handful of dimensions will typically stand out from the rest as being especially key; these dimensions may be the ones that explain the solution most completely and succinctly, or the ones that unlock novel insights. The number of dimensions may also depend on the framework type that you want to use. For example, whereas a two-by-two matrix can only handle two dimensions, a hierarchy can potentially handle many more.

When deciding on the kinds of dimensions to include in the framework, you can choose either dimensions of the same class or of different classes. Each class comes with a unique way of thinking about the underlying data. Using dimensions of the same class has the advantage of letting you transfer one way of thinking across the dimensions in the framework. For instance, if

you know that the framework only uses continuous dimensions, then you can potentially apply the same quantitative way of thinking – and the associated machinery, such as arithmetic operators and statistics – to all of them. You can thus also compare dimensions of the same class more easily (think “apples to apples” versus “apples to oranges”). However, using dimensions of different classes also has its merits. In a hierarchical framework, using different dimension classes for each level in the hierarchy can help distinguish the levels from one another more clearly. For example, the top-level concepts in a given hierarchy may be categorical, while the sub-concepts may be ordinal or continuous; in this case, going deeper into the hierarchical structure would also be paralleled by an increase in the information-richness of the dimensions involved, which may help your analytical thought process.

Ultimately, the choice of framework dimensions in terms of quantity and diversity will most likely be part of an iterative process. The dimensions that you start off with at the beginning of the framework-building process may not necessarily be the ones you end up including in the final framework. Also, as with most things, there is likely no “perfect” dimension, just dimensions that are more or less suitable for your framework



objective. Being aware of the strengths and limitations of the dimensions and seeing framework-building as an iterative process should help take the pressure off at the outset and allow you to focus on building a useful conceptual framework.

## **Case Study: Sales Performance at SoftCo**

The sheer variety of framework dimensions, and their strong coupling with the framework objective, means that hand-picking “the most important” dimensions (or selecting based on some other criteria) can be difficult. Yet, changing the dimensions while maintaining the same framework type can lead to very different interpretations of the framework. In the following anonymized case study, we will see how even slight modifications to the dimensions can make a big difference and yield new insights.

SoftCo is a mid-sized technology company that offers marketing-related software products and services to businesses. The company operates in the US and has about two dozen sales reps spread out nationally across different territories. The sales reps are responsible for growing the business in their territory, which includes everything from identifying prospective customers to interacting with them and closing the sale. At the end of every

month, Sally, SoftCo's veteran Head of Sales, reviews the performance across all territories and reports her findings to the CEO. She also gives feedback to the sales reps to recognize achievements and suggest ways to improve. Over the years, Sally has identified several factors that can influence the performance of individual sales reps, including the amount of customer interaction (typically phone calls, with a few field visits). Figure 1 shows a simple scatter plot (a matrix framework with two continuous dimensions) that compares sales performance to customer interactions for individual sales reps.

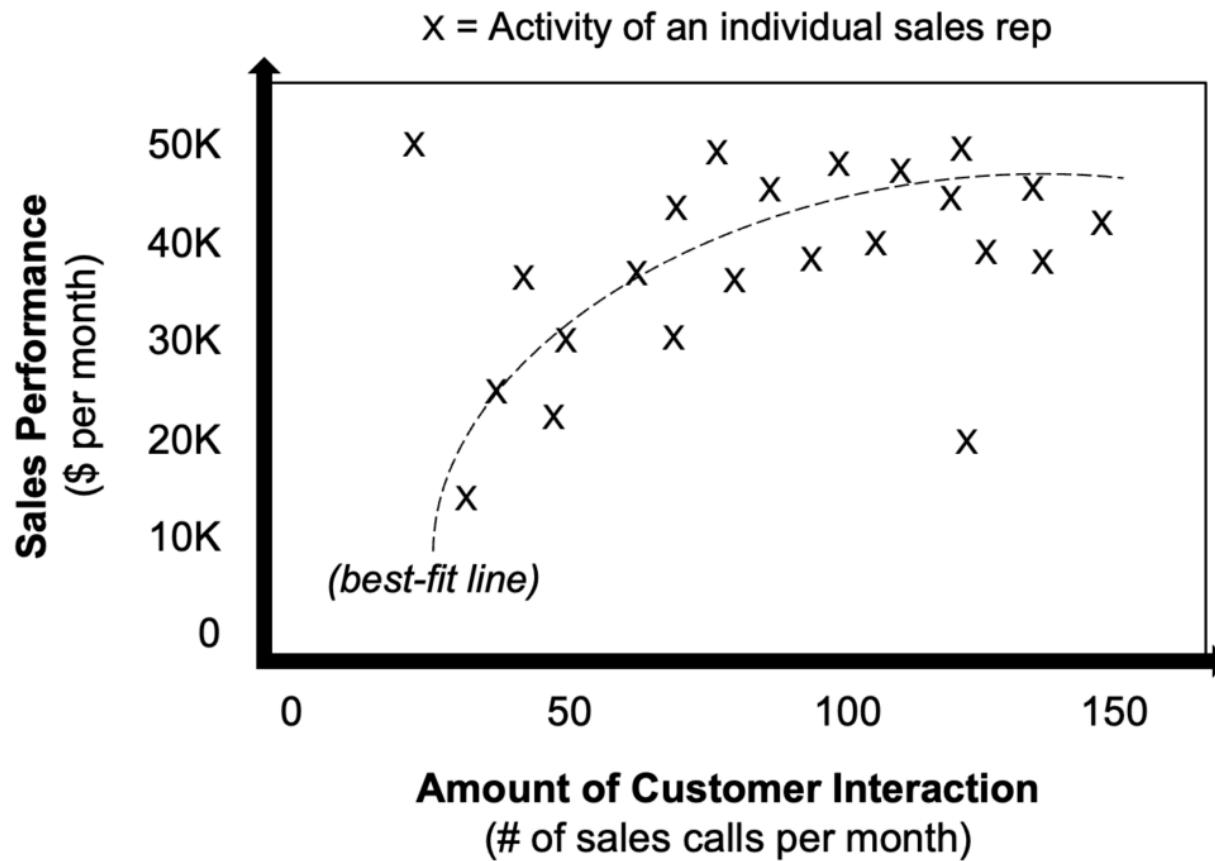


Figure 1: Scatterplot of Sales Performance at SoftCo

The choice of dimensions in Figure 1 guides the interpretation of the framework in many ways, beyond the fact that Sally has chosen specifically to examine customer interaction as a key predictor of sales performance. The use of continuous dimensions lends itself naturally to quantitative measurement. Sales performance is thus measured by the amount of money

each rep generates per month, while customer interaction is measured by the number of sales calls made per month. Of course, these measures alone are probably not sufficient to fully capture the two framework dimensions. For instance, the number of calls does not tell us anything about the quality and distribution of the calls across customers, and the dollar value of the deals a sales rep generates in a month does not tell us much about the strategic nature of the deals (e.g., whether the deals were about growing the business with existing customers, or “door openers” for a new stream of business with new customers). Nevertheless, by looking at the scatterplot in Figure 1, we can derive several interesting insights:

- There were 23 sales reps working for SoftCo during the observed month. In total, the sales team made about \$858,000 in this time period.
- On average, each sales rep made about \$37,300 worth of sales in the observed month. The highest and lowest individual sales were about \$50,000 and \$14,000, respectively.
- The most efficient and least efficient sales reps (in terms of \$/calls) made about \$2,000/call and \$160/call, respectively;

that is a roughly 12x difference in efficiency.

- There seems to be a non-linear relationship between customer interaction and sales performance. Up to about 75 calls, each additional call seems to be correlated with a big boost in sales performance. But beyond 75 calls the link with sales performance is less strong.

Figure 1 thus leads to a range of insights that are derived by looking at the performance of individual sales reps and the performance of the entire group. Some of the insights are fairly straightforward (e.g., the number of sales reps, average sales performance), giving us a general understanding of the scale of SoftCo's sales operation and the nature of the business. Other insights, such as the gap between the most and least efficient sales reps, and the non-linear relationship between sales performance and customer interaction, are potentially more thought-provoking; besides highlighting possible gaps between the abilities of different sales reps and diminishing returns from too many calls, the insights also suggest that other factors beyond customer interaction may also be good predictors of sales performance. The scatterplot representation also makes it easy to identify the outliers among the sales reps, which can be

useful for further analysis of what sets these outliers apart from the rest of the sales reps.

Now, to show how changing the class of the dimensions can lead to a different perspective, Figure 2 presents a two-by-two matrix that is based on the same information as the previous scatterplot. The two continuous dimensions of the scatterplot have been transformed into ordinal dimensions by splitting them along certain threshold values. Sales performance figures below \$25,000/month are considered “low,” while those above are “high.” Similarly, customer interaction figures below 75 calls/month are “low,” and those above are “high.” The choice of the threshold value is clearly important and should be based on reasonable argument. For example, the sales performance threshold may be based on a minimum sales target that each sales rep is required to hit, and the customer interaction threshold could be related to the point at which the curve in Figure 1 starts to flatten (indicating a shift in the marginal value of additional sales calls).

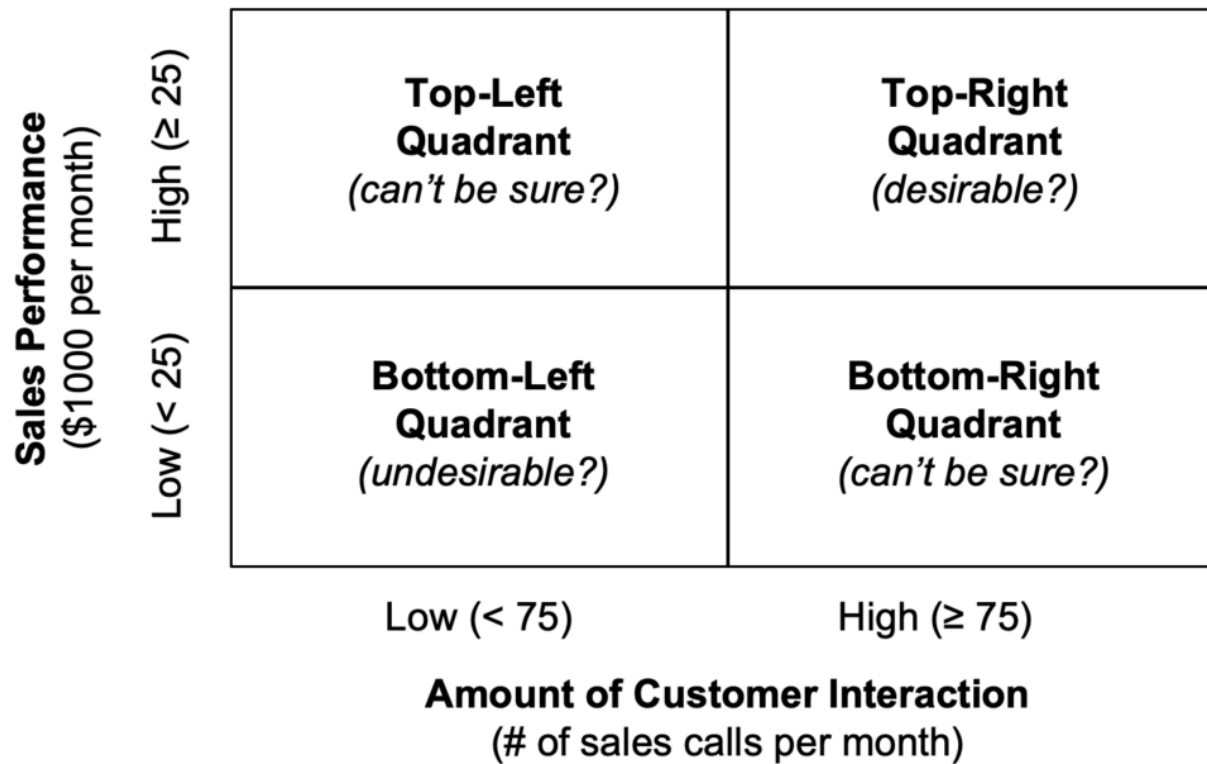


Figure 2: Simplified Matrix of Sales Performance at SoftCo

Whereas the scatterplot in Figure 1 drew our attention to the performances of individual sales reps and the overall trend in the relationship between sales performance and customer interaction, the two-by-two matrix in Figure 2 enables a more simplified view that lends itself to a segmentation of sales reps into different groups. In keeping with conventions, the bottom-left quadrant of the two-by-two matrix shows the group of sales

reps that may be in an undesirable position; these reps are making relatively few calls and generating few sales. The top-right quadrant contains “star performers” that evidently seem to interact extensively with customers and make sure that this hard work translates into actual sales. The dynamics in the other two quadrants seem less clear. The reps in the top-left quadrant seem to achieve high sales despite making relatively few calls – what is the secret behind their efficiency and is it sustainable? The reps in the bottom-right quadrant have the opposite dynamic, making a lot of calls that do not seem to pay off – if these reps are essentially working as hard as the star performers, why are they not achieving similarly high sales figures?

By drawing attention to different segments of the sales team, the two-by-two matrix can be used to develop tailored strategies that address the unique characteristics of each segment. For those in the bottom-left of the matrix, it is important to find out why both customer interaction and sales performance are relatively low. Do these sales reps have to deal with difficult customers, do the reps need more training, or are the reps allocating some of their time to other valuable activities that are not captured by this month’s sales performance (e.g., training other staff, strategic planning, and personal development)?



Armed with these additional insights, Sally can develop measures that better capture the true value that the sales reps in the bottom-left quadrant of Figure 2 create for SoftCo.

Similarly, for the bottom-right quadrant, a new strategy may be needed to increase efficiency by translating the relatively high level of customer interaction into actual sales; this may involve prioritizing certain leads over others, training the sales reps to be more tenacious in closing each sale, and motivating them to continue hustling. For the remaining two quadrants, achieving sustainability could be the key objective. It is worth understanding what makes the sales reps in the top-left quadrant so efficient and what the other sales reps can learn from them. At the same time the reps in the top-left also need a strategy for reducing the risk of slipping down if their customer interaction does not consistently pan out. Finally, a strategy is needed to keep the reps in the top-right quadrant motivated (e.g., by social recognition, monetary rewards, opportunities for promotion) to keep them performing consistently at a high level.

To close off, here is a helpful video by Mike Gastin that expands on some of the considerations discussed above when choosing dimensions for two-by-two matrices: