

# *datastreamx*

A Practical Guide to  
Pricing Data Products

Everett Leonidas  
Whit Walker

## Welcome!

First and foremost, congratulations! Welcome to DataStreamX! We've worked hard to provide you with the tools and reference materials necessary for structuring and monetizing data products that can be sold globally.

Vendors on the DataStreamX platform have varying levels of experience in distributing and selling their proprietary data into the public realm. For first-time Vendors, we have prepared a comprehensive set of tools to guide the process. For experienced Vendors, we provide advice on how to structure and price data so that it can be most effectively sold on the DataStreamX marketplace.

Whichever Vendor personality you identify with, we want to offer our help and make your experience as delightful as possible. We believe this document and pricing framework will help you calibrate the best pricing structure for your products. Please keep in mind, that pricing is totally in your control. We want to provide the tools and insights that drive you the most revenue.

Congratulations, again!

-The DataStreamX Team

## Contents

<b>Welcome!</b> .....	1
<b>Section 1. Thinking about price</b> .....	3
<b>Section 2. Understanding customers</b> .....	4
<b>Known use cases</b> .....	4
<b>Willingness to pay</b> .....	4
<b>Section 3. Assessing pricing levers</b> .....	6
<b>Brand strength</b> .....	6
<b>Frequency and latency</b> .....	6
<b>Completeness</b> .....	7
<b>Volume</b> .....	7
<b>Competition and rarity</b> .....	7
<b>Organization and tidiness</b> .....	7
<b>Ease of understanding</b> .....	8
<b>Dependency</b> .....	9
<b>Section 4. Scoring a data product</b> .....	10
<b>Section 5. Setting the pricing strategy</b> .....	12
<b>Section 6. Conclusion</b> .....	13

## Section 1. Thinking about price

Setting the price of data products will be among the most important decisions to make as a data Vendor on the DataStreamX marketplace. Pricing—long understood as a blend of art and science—has a profound impact on revenue generation and the ability to build an effective business case for creating data products. Generally, there are two primary ways to construct prices for data products: cost-based and value-based.

### **Cost-based pricing**

Cost-based pricing is frequently referred to “cost-plus pricing”; the two terms can be used interchangeably without confusion. In brief, this pricing strategy determines a selling price by adding a markup to unit cost. Vendors need to assess the fixed and variable costs required to build, maintain and deliver data products on DataStreamX. These costs are classified as set-up, personnel, overhead and on-going IT costs.

The good news for Vendors is that the fixed costs associated with creating data products for sale on DataStreamX is typically very low relative to previous methods of delivery. The tech investment involved can be as simple as adding snippets of server-side code on existing infrastructure. Likewise, variable costs also could be nearly negligible.

This means that it is very economical for an organization to get started selling data, however, it also means that a cost-based pricing strategy may be ineffective. Total costs could be very small and thus the corresponding mark-up would be orders of magnitude greater than those in other industries that use cost-based pricing (such as manufacturing).

DataStreamX recommends that data Vendors complete this costing exercise for evaluating costs, but not for setting actual selling price. Vendors should use their cost structure as the absolute minimum amount to accept for their data products. If priced lower than this, the organization would lose money on every sale. Instead of a cost-based pricing approach, DataStreamX suggests using a value-based pricing strategy.

### **Value-based pricing**

In software and services, the value a customer receives from a product is many times the cost of producing it. With value-based pricing, Vendors attempt to capture the majority of what customers are willing to pay for their data products. By no means is this an easy exercise. Value-based pricing requires many considerations and a critical self-assessment.

As a marketplace for digital assets, DataStreamX endorses a value-based pricing philosophy, we will devote the remainder of the guide to walking through the value-based approach and highlight the key considerations throughout the pricing process.

Unfortunately, we can't say what the exact price is for a point of data itself. However, we can help Vendors consider the most important factors for determining the selling price of their data. We will cover these factors in the following sections as we detail customer identification, value levers, product scoring and pricing strategy.

## Section 2. Understanding customers

### Known use cases

Vendors should try to understand their customers as best as possible. A good first step to this understanding is to assess the various use cases for the data product. Vendors should consider many possible ways their data products could deliver value to a user. Organizations use data in a variety of ways, including, but not limited to:

- Enhancing existing products
- Creating new products
- Improving customer experience
- Lowering costs
- Mitigating risk
- Improving supply chain efficiency
- Pricing more effectively

Each use case will have a different impact on an organization. Each organization will have a different opinion on the value they receive for it, and thus a different interest in buying the data product at a given price.

Vendors that can clearly define their target customer and use cases will be more capable at pricing their data products. If not, they risk overpricing for their intended targets without much supporting evidence.

### Willingness to pay

A Vendor's ultimate objective on DataStreamX should be to maximize total revenue—by converting the maximum number of customers at the highest possible price point. A data product's ideal price would capture the customers' maximum willingness to pay, without deterring too many prospective buyers. The challenge is that all customers are different, each with a unique willingness to pay. Mastering the nuances of this dynamic requires precisely measured price elasticity of the data products.

In brief, the concept of price elasticity means that pricing low would attract highest number of customers, but total revenue is non-optimal because price is below their willingness to pay (also known as consumer surplus). Conversely, pricing too high would yield non-optimal total revenue because customers would be deterred by prices above their willingness to pay.

Vendors can reasonably estimate customers' maximum price they are willing to pay by examining who the target customers are and what behaviors define them. Most likely, commercial data products on DataStreamX will be most attractive to enterprise customers, also known as "B2B" customers.

For the sake of our pricing discussion, B2B organizations are made up of two relevant groups of people: **users** and **decision makers**. Users generate demand—they *want* the data product. Sometimes a **user** can make a purchase decision individually—a great situation when upper management approval is unnecessary, thus the sales cycle is short. Sometimes a user must obtain approval from **decision makers**—situations where the sales cycle is longer.

As stated above, a Vendor's primary objective should be to maximize revenue. The secondary objective should be to encourage customers to make a purchase decision as quickly as possible. Once a user has realized the need for a data product, the Vendor can accelerate the purchase decision with price. We will discuss two alternatives for 'fast-tracking' the purchase decision for B2B customers:

1) Setting price just below the threshold where **users** must request purchase approval

Each organization will have different thresholds. These can range anywhere from \$1,000 to \$5,000 to \$15,000. With a good understanding of the target customers, Vendors should have a good understanding of their users' threshold, before their purchases escalate to decision maker approval. If Vendors do not know their customers just yet, we encourage them set prices more conservatively so as not to discourage buyers based on unsupported assumptions.

2) Set the price at a point where **decision makers** would find it hard to object

Decision makers will choose whether or not to buy a data product. They will have a hunch on what an acceptable price is—they will look at how comparable products are priced in the open market. These are called a "reference prices."

For example, data used in IT consulting/research projects is typically provided by industry analyst firms such as Gartner, IDC, Forrester, ABI Research and Strategy Analytics. As a result, companies with IT sector data would want to use the prices these companies charge as reference points. Typically, these firms price each of their "premium" data sets at \$15,000 to \$40,000 per year. This range should provide a suitable reference for how much a robust, "premium" data set can command in the marketplace.

Likewise, any data sets that are priced higher than these "premium" data products would require a significant value add above and beyond what customers currently receive from the currently available, comparable data products. We generally recommend prices do not exceed such higher end reference prices.

Decision makers will also consider "make vs. buy" scenarios. "Make vs. buy" entail decision makers comparing the price of a Vendor's data product against the costs of building or assembling the data themselves. If the total "make" cost exceeds "buy" price of the Vendor's product, then the decision maker would find it easier to justify purchasing the data product.

After considering currently available reference prices, data Vendors should have a better understanding of the price the market is willing to accept for a given data product. With this understanding, it is now possible to adjust pricing according to the value represented by a new data product. In order to understand how Vendors can achieve the maximum willingness to pay, we will now dive deeper into the levers that drive value.

## Section 3. Assessing pricing levers

In this section, we will discuss the primary value-to-price tradeoffs elements that determine how a given data product will compare to a customer's maximum willingness to pay. Many of these value drivers will be influenced by the expectations of the customers, i.e. what are minimum requirements or **benchmarks** for data products of a given type.

We would like to emphasize again, that data Vendors should have target customers in mind when calibrating the pricing levers below in order to arrive at their data product's price. Customers will likely have pre-existing benchmarks in mind when evaluating any data product.

For illustrative purposes, we will create a fictional new data product for global weather forecasts ("**New Weather Product**"). An appropriate benchmark is the Global Forecast System that is published four (4) times per day by the US National Weather Service ("**NWS**") and produces forecasts for up to 16 days in advance.

### Brand strength

Brand strength refers to how recognizable and respected Vendors are with their target customers. The stronger association there is between a Vendor's name and valuable use cases, the more confidence buyers will have with the Vendor's products. Thus, a stronger brand allows a Vendor to charge a higher price relative to competition and alternatives.

It should be noted that a Vendor does not need to have a "household name" to have brand strength. Instead, the Vendor could have a passionate user community, strong reference cases or subject matter expertise.

For example, the Vendor of the New Weather Product may not be well known outside of the meteorologist community. However, the New Weather Product Vendor may be well-respected and supported by meteorologists enough to influence commercial buyers of the New Weather Product. In fact, meteorologists may base their analyses, research and references on New Weather Product, thereby strengthening the brand of New Weather Product even further.

### Frequency and latency

Frequency refers to the intervals between when data is captured, refreshed and distributed. For example, the NWS is published four times daily. If New Weather Product is published hourly, then it is delivering significant added value to users who are highly sensitive to weather patterns, such as maritime vessels, who need data at more frequent intervals. Thus higher frequency would command a higher price.

Latency refers to the promptness of when data product is distributed. For example, suppose New Weather Product published their forecasts and current weather conditions four (4) times daily, like the NWS forecasts. If current weather conditions are then published with a one hour delay, and if NWS has no delay, then New Weather Product is less valuable since it has no time advantage over the benchmark. New Weather Product would then be difficult to sell at a premium price point.

## Completeness

The completeness of a data product refers to the quantity of useful data fields it contains. Each user is different and will likely be searching for different signals or measurements. If a data product can deliver all the critical information to a customer, then other data products are not necessary. Thus, a more complete data product generally delivers more utility to the target customers.

For example, the NWS benchmark may not give granular detail on the weather forecast, such as time, precipitation amount, wind direction, humidity and pollen count. This may not be necessary for most customers, but it may be necessary for the target customers of New Weather Product. This would be a point of differentiation for New Weather Product. If this is the case, then the target customers would then value the New Weather Product at a premium.

## Volume

While completeness refers to the number of variable types, volume refers to the number of reported items or entries. Using the above example, New Weather Product's volume could significantly increase if it included hourly predictions of the weather for each data field. At the same time, volume could significantly increase if New Weather Product reported information across a broader geography, such as multiple locations within a city rather than the city as a whole.

More volume in a data product generally translates into increased value because it provides more detail and leads to more actionable insights.

## Competition and rarity

Competition refers to the substitutes and competing data products available, either on DataStreamX or elsewhere. As with most products, more competition typically decreases the price a Vendor is capable of charging. With the New Weather Product example, the NWS is a direct competitor. If all other factors held equal and there was no differentiation from NWS, then the ability to charge a premium price for New Weather Product would be compromised.

Rarity refers to a data product's uniqueness and ability to be replicated by another party. Proprietary data can be extremely rare, and if useful, can command a significant premium. For example, if the New Weather Product has exclusive rights to certain weather measurements, such as pollution levels, then it will be extremely rare. This rarity provides New Weather Product considerable market leverage and ability to charge a premium.

## Organization and tidiness

Organization and tidiness refers to how the data product is compiled and delivered to users. Buyers will use data products in various ways, such as inputting the information into models or joining with other data sources. If the data product is in a format that is readily consumed, then it requires less preparation and is thus more valuable. For example, if the New Weather Product is output as an image or PDF unreadable by machines, then it may not be of any use to buyers since the cost of using the data product outweighs its benefits.

Organization and tidiness also refers to the consistency of the format of the data product. Data Vendors should maintain a consistent structure to their data products so that buyers won't have to change the models, processes and analyses developed based upon prior data products. In the case of our New Weather Product, if city weather is reported using four components, then it should maintain this structure and not change to only report three components. This would frustrate and confuse buyers.

Additionally, numeric format should be consistent. For example, if one million is reported as "1.000" in the new economic data product, then ten thousand should be reported as "0.010".

Sloppy, shifting or difficult to manipulate data products may deter potential buyers, particularly if they require a significant investment of time to use or maintain. As such, buyers' willingness to pay for sloppy data products will fall.

### Ease of understanding

Ease of understanding refers to how self-explanatory a data product is to a buyer. Clear and detailed documentation vastly improves the ease of understanding. All data products must have documentation provided by the data Vendor. Clear documentation is important for two reasons: marketability and usability.

In terms of marketability, clearer documentation leads to better understanding and application. When buyers are able to clearly understand what a data product is and how to use it, then the more likely they will be to purchase it.

For usability, clear documentation is especially important if the data fields are not formatted to generally accepted industry standards. Likewise, all data should be clearly marked and labeled so that the users can understand the units of measurement.

Generally, the less clear a data product is, the less valuable it is to a user.

For example, New Weather Product may measure wind speed with a propriety unit of measure: NPS. Whereas, the NWS benchmark measures output in miles per hour. A user would have to know exactly how to convert NPS into miles or kilometers per hour. This should be detailed both in the documentation and clearly marked within the data set. So long as the user can understand the information contained in the new economic data product, then he or she won't discount the value relative to NWS.

### Accuracy

Accuracy refers to how correct the data product is. Many times, accuracy is gauged by the ability to indicate relative conditions such as with a tracking index. For other circumstances, accuracy is an absolute measure such as with temperature, GPS and stock prices.

Customers will determine their tolerance levels regarding data accuracy. Accuracy is easily evaluated by a data product's competition and how often the data is revised or restated. For example, New Weather Product may be more accurate than the NWS because its forecasts are rarely revised and have a higher rate of correctly predicting future weather events. If true, then it would be more reliable than the NWS forecasts.

## Dependency

Dependency refers to the critical nature of the data product for the buyers' use cases. If decisions cannot be made without the data product, then dependency is high, and the data producer can charge a higher price for it. However, if the data product can help make more accurate decisions, but is simply a "nice to have" feature, then the data still has value, but the buyer's willingness to pay for it will be lower.

For example, commercial fishing fleets need weather data to plan routes and manage risk. If New Weather Product provides the most up-to-date and accurate weather information, then there would be a high dependency (and value) for customers involved in commercial fishing.

## Section 4. Scoring a data product

In this step, we help data Vendors score data products based on the value they can deliver in relation to customer expectations. This step is very introspective—it requires Vendors to honestly assess how they believe customers will perceive the data products. After scoring their own product, Vendors would be wise to score the competition data products available. This will help Vendors visualize how the data product compares to reference products—for both price and value.

The DataStreamX scorecard is based on a scale of 1 to 10. '10' indicates the product ranks at the highest level of customer expectations. Conversely, '0' indicates the product ranks at the lowest level of customer expectations. '5' indicates the product is generally acceptable for customers. The table below defines these rankings in more detail.

### Brand

---

- 10 - Vendor has recognizable, trusted reputation
- 5 - Vendor has undifferentiated reputation
- 1 - Vendor is unknown or has suspicious reputation

### Latency & frequency

---

- 10 - Fastest of any benchmarks or competition
- 5 - Prompt, and timely; meets expectations
- 1 - Extremely delayed in distribution

### Completeness

---

- 10 - No more data needed is needed for use cases
- 5 - More data may be needed for non-core uses
- 1 - Data is incomplete and useful for few analyses

### Volume

---

- 10 - Data provides rich detail for thorough insights
- 5 - Enough data is provided to draw simple insights
- 1 - Data is very sparse

### Competition & rarity

---

- 10 - Data set is unique and irreplaceable
- 5 - Data set has few substitutes
- 1 - Data set is a commodity, no differentiation

### Organization & tidiness

---

- 10 - Data can be manipulated and used immediately
- 5 - Data requires some scrubbing and preparation
- 1 - Data is unusable for most situations

### Ease of understanding

---

- 10 - Concise, thorough documentation & use cases
- 5 - Detailed documentation
- 1 - Limited documentation, unclear labeling

### Accuracy

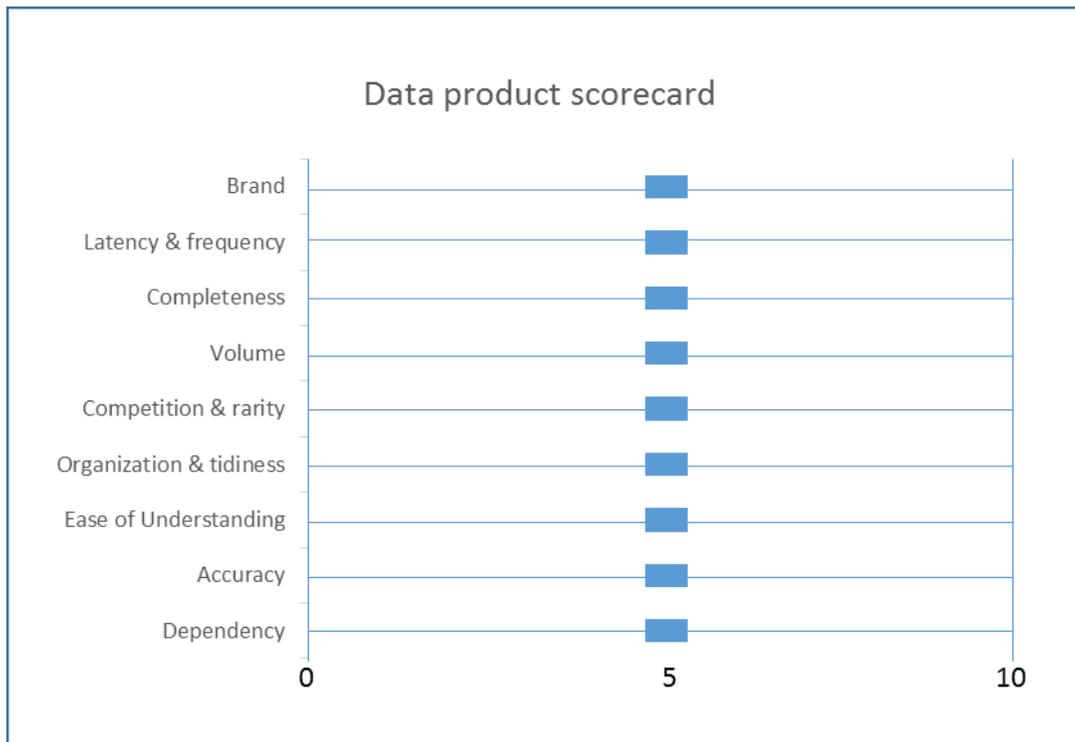
---

- 10 - Unquestionably accurate and exact
- 5 - As exact as competing products or sources
- 1 - Frequently incorrect, revised or restated

### Dependency

---

- 10 - Use cases would fail without this data
- 5 - Use cases would improve with this data
- 1 - Use cases do not benefit much from this data



A data product that scores '10' across all key metrics would be considered a "gold standard" and thus could capture the maximum willingness to pay (e.g. \$15,000). However, Vendors may not want to structure products that deliver all the features a customer wants for a "gold standard" data set, such as update frequency or intervals. Doing so might price the data set at too high of a value. Likewise, it might be smarter to divide up all of the relevant elements into smaller data products that can address specific use cases. Customers may lower their willingness to pay (e.g. \$2,500) and accept the feature tradeoffs that Vendors make with their data products.

It is worth noting again that all customers are different and have different price / feature expectations. The scorecard is for Vendors to complete to understand how a data product will be perceived in the marketplace. In setting their price, Vendors should consider all variables mentioned in order to maximize the revenues they can earn on DataStreamX.

## Section 5. Setting the pricing strategy

DataStreamX gives Vendors full control on how to set the price of their data products. Vendors can list and price data products in several ways:

- Monthly subscription with optional one-time fee
- Single purchase data product for one-time downloads of reference and historic data

Typically for streaming data or continuously updating data, a monthly subscription is most common. This data type holds most value when it is consumed close to the time of origin. When buyers derive a larger amount of value during the initial months consuming a data product than they do in subsequent months, Vendors should consider charging a one-time fee to the buyer. Since DataStreamX does not force long-term contracts, a one-time fee is best used to capture the total value of a data product, while also ensuring an ongoing revenue stream from the subscription.

For reference data and historic data, a clear understanding of the target buyer is needed in order to understand the inherent value that the data possesses. For example, real estate developers would want to have robust, analyzed historic data sets in order to forecast and manage risk before making large capital investments in new commercial properties. Conversely, historic and reference data for oil commodities is less valuable than real-time data since futures pricing can change in milliseconds. The ability to assign premium prices to reference and historic data products largely depends on who will primarily consume the data and the necessity for immediacy.

Finally, DataStreamX recommends pricing data competitively in order to gain wider adoption and lower the inhibitions of “fringe” customers wanting to test and sample data products. Because DataStreamX provides Vendors complete visibility into their customer base, Vendors can better understand who their target customers are. By setting a competitive price on the marketplace, Vendors have the opportunity to attract a broader set of users and be able to better serve these customers. With a broader set of users, Vendors can become smarter: building better data products to meet market needs or adjusting prices to capture the maximum willingness to pay.

It must be noted that all paying customers will have a six month price protection from the date of any price increases. So it is in Vendors’ best interests to complete the analyses recommended in this document to ascertain what they believe is the best price to charge on the DataStreamX marketplace.

Keep in mind, the primary goal is to maximize the revenue Vendors earn on DataStreamX.

## Section 6. Conclusion

Pricing is never an easy exercise, but with the right framework Vendors can quickly become more skilled at finding the ideal price. We designed this pricing guide to provide clear guidance and a sound framework to price data products that maximize revenue for a Vendor.

DataStreamX provides several free resources to help Vendors successfully build and market data products on our marketplace. We will soon have a companion whitepaper—“Structuring Valuable Data Products”— that helps Vendors build valuable data products for external consumption. Additionally, more resources are available on our community portal at: <http://community.datastreamx.com>.

## About the Authors

### Everett Leonidas



Everett Leonidas is a director of business development at DataStreamX. Everett has over a decade of experience in customer development in high technology and finance. Everett continuously aspires to delight clients worldwide with a consultative approach to unlock new business opportunities. He has worked in various international markets including Korea, the Philippines, Singapore and the United States. Currently, Everett lives in Singapore and enjoys cooking, golfing, traveling and discovering data.

For inquiries, contact [everett@datastreamx.com](mailto:everett@datastreamx.com)

### Whit Walker



Whit Walker is a director of business development at DataStreamX. Whit has nearly a decade of experience in business development for media and technology companies. Whit helps organizations worldwide to monetize their way in a meaningful and impactful way. Whit has lived and worked in various international markets including Korea, Argentina, Singapore, India, Australia and the United States. Currently, Whit lives in Singapore and enjoys scuba diving, snowboarding, traveling and learning about other cultures.

For inquiries, contact [whit@datastreamx.com](mailto:whit@datastreamx.com)