

26th Nov 2021

in Product Management Data

Anatomy of a Data Product



A data product is an application or tool that uses data to help businesses/users improve their decisions. Data products provide the user easy, non-technical access to Data science methodologies – meaning it can abstract the complications of data science to provide a concrete output to the users to make decisions that are critical to their work.

➤ Types of Data Products

Data products themselves can be classified under two categories:



📌 Decision Supporting

Products of this type are mostly used to derive insights from the underlying data. They can either:

1. Provide raw data access

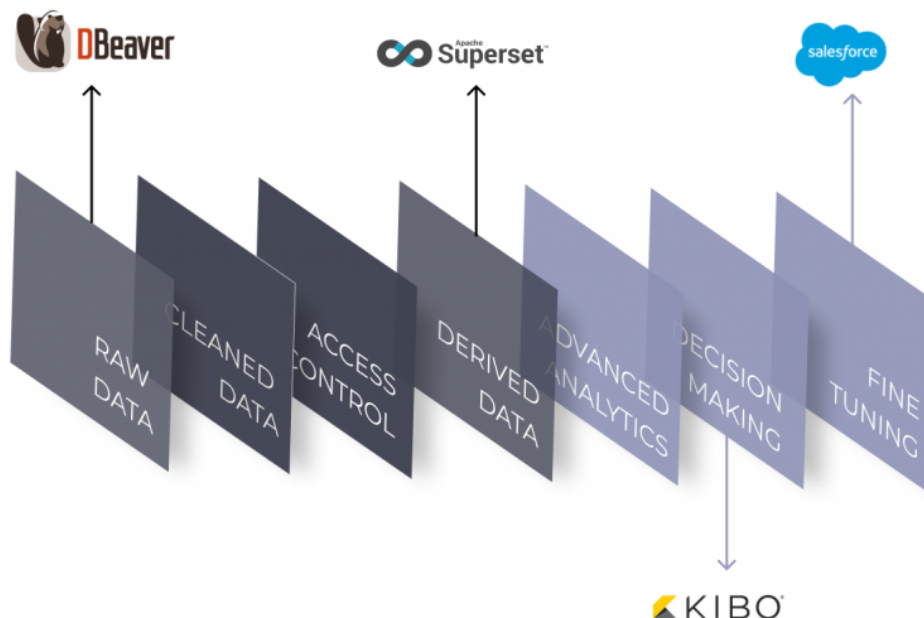
Where the user processes the data on their end to generate insights. Think of database access tools like DBeaver, MySQL Workbench, etc.

2. Provide derived data access

Where a large portion of the data is cleaned up, and the user has access to only the cleaned-up data points. Advanced Analytics run on that data can also feature in the data points available. Tools like Tableau, PowerBI, StreamLabs, Plotly Dash, etc. fall under this category. They mostly are visualization tools with an ability to process data at their end OR allow the user to process some data through complex queries.

📌 Automated Decision-Making

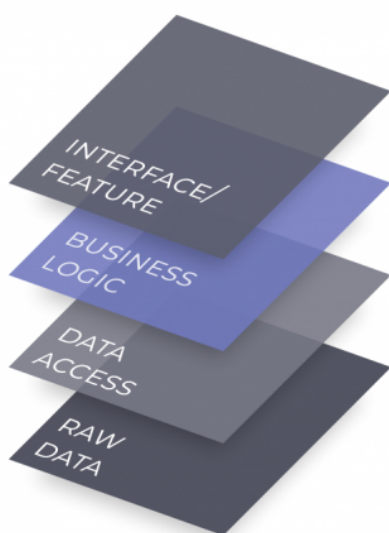
Products under this type take decisions on behalf of the user. Almost always, there is a Data science element involved. Here, the users' role is only in monitoring and tweaking the outputs from the machine. Recommender systems, workforce planning systems are typical examples of these types of products.



There are many tools that are hybrids: Google Analytics with Google Ads can provide an easy way to analyze the user journey and, at the same, provide segments to target through Google Ads.

➤ Anatomy of a data product

Now that we know the types, let's understand the anatomy of a data product:



Anatomy of a Product

INTERFACE/FEATURE LAYER

Interactive layer - contains features, interfaces etc.
User facing features and interfaces are coded at this layer.

BUSINESS LOGIC LAYER

Contains the business logics - like who should see what?
And other data processing elements required for features

DATA ACCESS LAYER

Manages the access levels for each data point. Also
transforms the data in a usable format.

RAW DATA LAYER

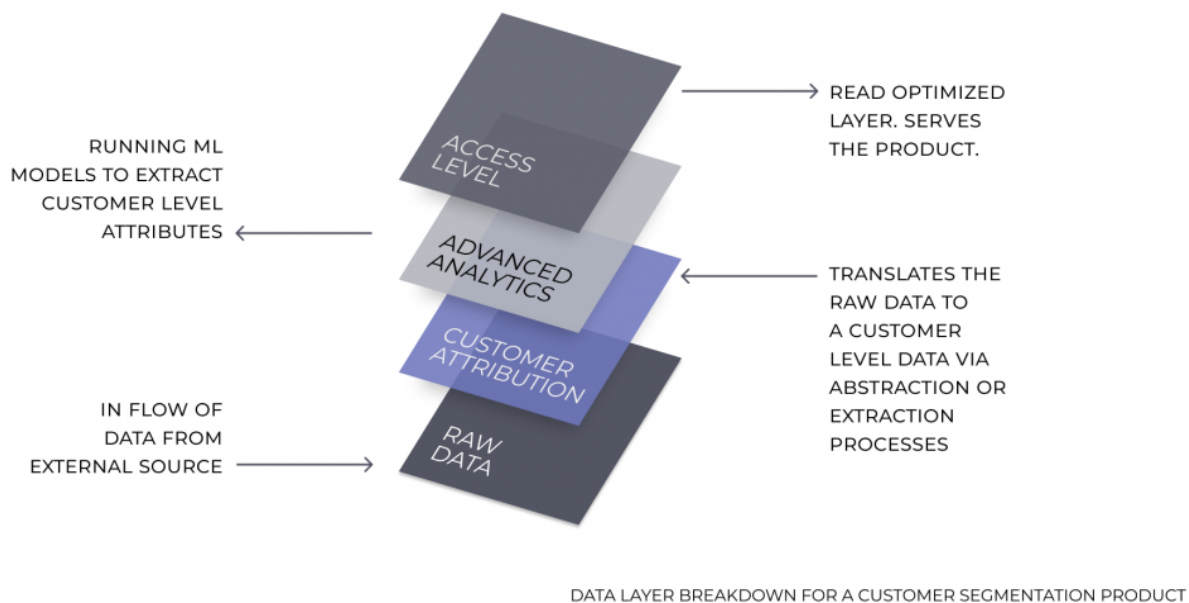
All data is stored here.

I'll explain each layer in simple terms, but you can skip this if you already know what each layer caters to...

Starting from the bottom:

Raw Data Layer

This is the underlying data that is used to build the product. All necessary data for the functioning of the product is available here. It is important to ensure that this layer is broken down based on the type of product that you are building. In the Customer segmentation tool, the underlying data itself can be broken down as per the level of processing and the availability of the data:



The significant thing to note here is that the access level shown above is only a read-optimized storage space – for fast retrieval of data.

Data access layer

This layer handles the question – Who has access to what? While it might seem like data governance, it has a bigger role – it supplies the data to the next layer to take the right decisions based on the data.

Note that the Data supporting tools directly short-circuit to the Interface/Feature layer from here. There is no business logic for those products.

Business-logic layer

This is where the configurations of the knobs and switches lie. Think of the recommendation tools – here is where the interface option of, say, “Disable

already viewed products in the recommendation' switch controls the output from the incoming data from the previous layer.

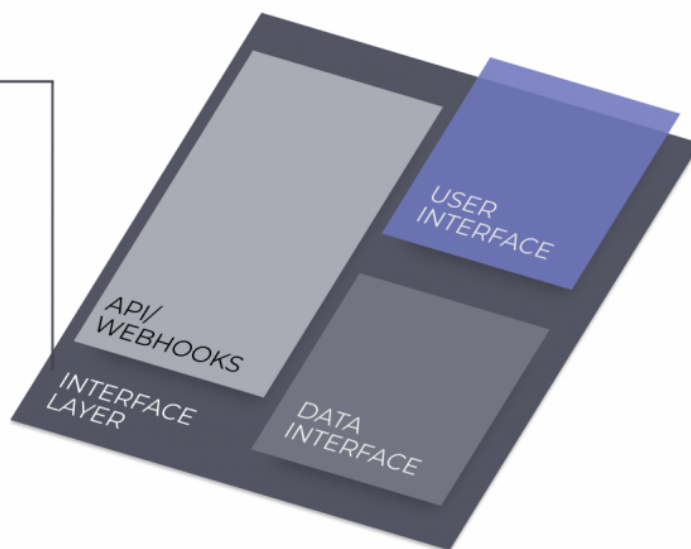
👉 Interface/Feature layer

It might sound a bit of a stretch to combine both the interface and feature here – as the feature layer is technically vertical (a feature requires data from the bottom two layers, combined with the business & interface layer), but I keep it here to demonstrate a point about how these features are envisioned – top to bottom.

This layer serves as a front-end for all the data & data-driven decisions made... Note that data itself can flow between all these layers. E.g. a recommendation product will have an API in the feature layer that will go from the business layer, into the access & the raw data layer... from there the output will get filtered from the business logic layers (the knobs and switches) and again, it'll go down to retrieve the product images & other presentation contents.

INTERFACE LAYER ←

Acts like an intermediary between the product and the consumers. even if that consumer is another product or service



Git Gud, from today. Subscribe:

Subscribe

substack

↘ Challenges in Data Product Management

While every product made on the planet made today can be broken down as per the above anatomy, the difference, when it comes to product management, has to be noted and understood. Traditional product management focuses on the top two layers: Interfaces & Business logic. We define features in those two layers.

However, with data products, it is important to understand the underlying two layers and how they affect the two layers. E.g., if you're working on building a customer segmentation tool and want to make it 'AI-enabled', you've to know the underlying data that is used to make a segment. Additionally, when it comes to 'pretty features' like NLP for dynamic audience creation, knowing which set of attributes are used to create a segment and thus how they'll be considered together, based on the input has to be known.

We have a tendency to cover for the average and handle the extremes at a later point. However, as data products, by definition, are decision critical, we've to be mindful of the extremes and handle those cases as well.

On the flip side, ensuring your core functionalities and features are reliable is also critical. Taking the above example, if you were to enable an NLP-powered customer segmentation tool, you'll always hit the bottleneck of query response time (which happens at the data layer). In fact, as a primary driver to your product, the NLP feature will be the bane of the tool.

My advice — start small. Cover your fundamentals. The Interface layer will be tempting to work on, but long-term reliability and customer experience will only be attained through strengthening the foundational layers.

—

Thank you for reading. If you enjoyed this article, don't forget to clap & follow!

📌 You may also like

→ [Product Management is hard.](#)

→ [5 Common Pitfalls to Avoid as Product Manager](#)

→ [3 Ways to Dominate Your Organization's Data](#)

→ [New to Product Management? Here are 6 foundations you should build!](#)

→ [Building a Careem Super-app — Part 1](#)

→ [The Startup Product Managers Manifesto](#)

You can get articles like this and more, right in your inbox. Subscribe to my newsletter "The Trifecta"!

Subscribe

substack

Share:

