

Data exploration

Collecting data

Differentiate between data formats and structures

- ▶ **Video:** Discover data formats
5 min
- ▢ **Reading:** Data formats in practice
10 min
- 📖 **Practice Quiz:** Self-Reflection: Unstructured data
2 questions
- ▶ **Video:** Understanding structured data
1 min
- ▢ **Reading:** The structure of data
10 min
- ↔ **Ungraded Plugin:** Differentiating Data Types
30 min
- ▢ **Reading:** Data modeling levels and techniques
20 min
- 📖 **Practice Quiz:** Test your knowledge on data formats and structures
4 questions

Explore data types, fields, and values

Weekly challenge 1

Data modeling levels and techniques

This reading introduces you to data modeling and different types of data models. Data models help keep data consistent and enable people to map out how data is organized. A basic understanding makes it easier for analysts and other stakeholders to make sense of their data and use it in the right ways.

Important note: As a junior data analyst, you won't be asked to design a data model. But you might come across existing data models your organization already has in place.

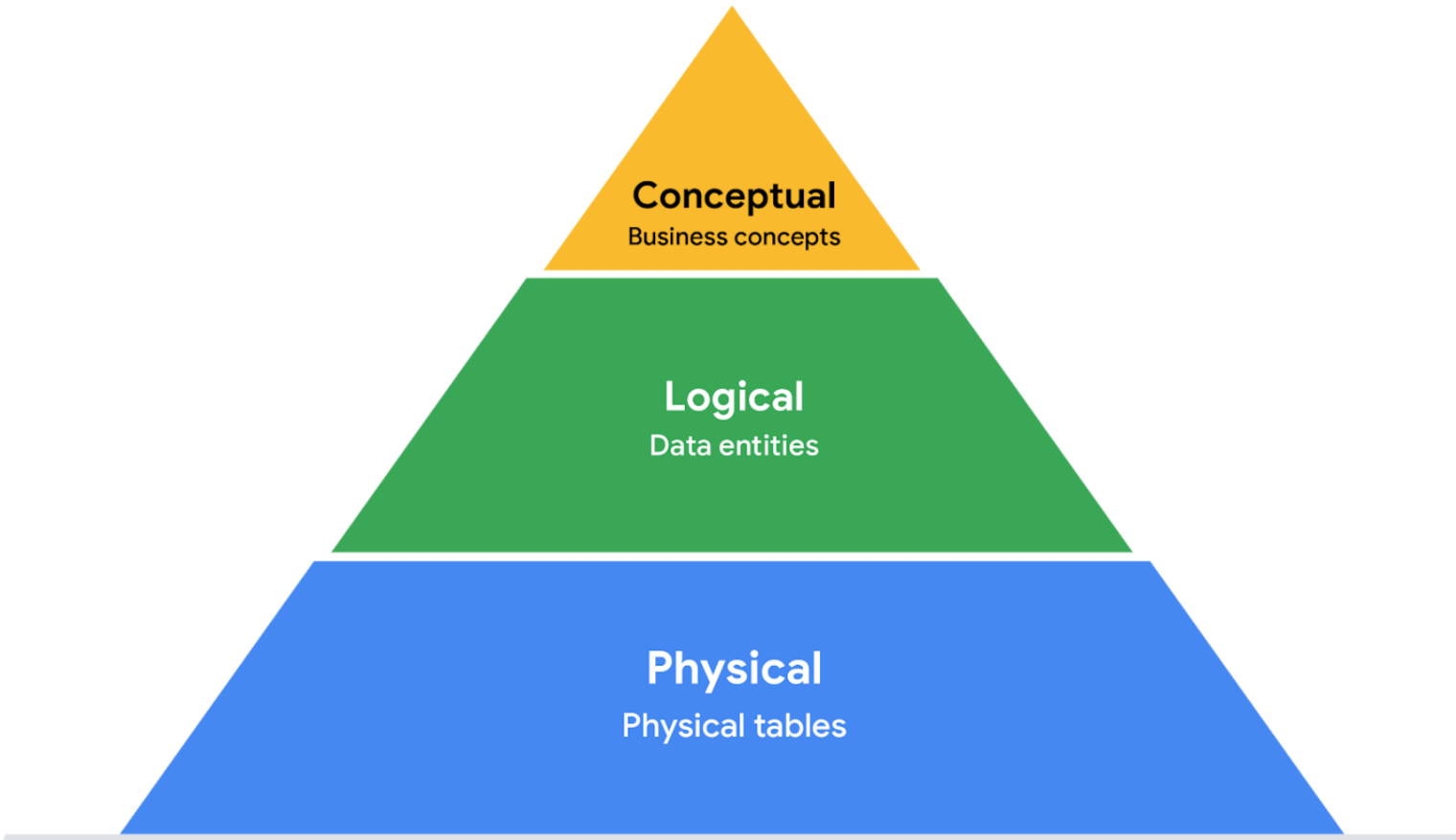
What is data modeling?

Data modeling is the process of creating diagrams that visually represent how data is organized and structured. These visual representations are called **data models**. You can think of data modeling as a blueprint of a house. At any point, there might be electricians, carpenters, and plumbers using that blueprint. Each one of these builders has a different relationship to the blueprint, but they all need it to understand the overall structure of the house. Data models are similar; different users might have different data needs, but the data model gives them an understanding of the structure as a whole.

Levels of data modeling

Each level of data modeling has a different level of detail.

The three most common types of data modeling



1. **Conceptual data modeling** gives a high-level view of the data structure, such as how data interacts across an organization. For example, a conceptual data model may be used to define the business requirements for a new database. A conceptual data model doesn't contain technical details.
2. **Logical data modeling** focuses on the technical details of a database such as relationships, attributes, and entities. For example, a logical data model defines how individual records are uniquely identified in a database. But it doesn't spell out actual names of database tables. That's the job of a physical data model.
3. **Physical data modeling** depicts how a database operates. A physical data model defines all entities and attributes used; for example, it includes table names, column names, and data types for the database.

More information can be found in this [comparison of data models](#). ↗

Data-modeling techniques

There are a lot of approaches when it comes to developing data models, but two common methods are the **Entity Relationship Diagram (ERD)** and the **Unified Modeling Language (UML)** diagram. ERDs are a visual way to understand the relationship between entities in the data model. UML diagrams are very detailed diagrams that describe the structure of a system by showing the system's entities, attributes, operations, and their relationships. As a junior data analyst, you will need to understand that there are different data modeling techniques, but in practice, you will probably be using your organization's existing technique.

You can read more about ERD, UML, and data dictionaries in this [data modeling techniques article](#) ↗.

Data analysis and data modeling

Data modeling can help you explore the high-level details of your data and how it is related across the organization's information systems. Data modeling sometimes requires data analysis to understand how the data is put together; that way, you know how to map the data. And finally, data models make it easier for everyone in your organization to understand and collaborate with you on your data. This is important for you and everyone on your team!