



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

What Is Data Modeling?

BY [WILL HILLIER](#), UPDATED ON SEPTEMBER 5, 2023

16 mins read

Data can be messy.

This is often a headache for [data analysts](#). But understanding a dataset's structure and the relationship between different data points can also help you manipulate it to meet your business intelligence needs.

To figure all this out, data analysts typically use a process known as data modeling. Data modeling allows you to dive deep into data, helping design, implement, and manage complex database systems. Data models also keep data analysts, software designers, engineers, and other stakeholders on the same page, ensuring everyone's needs are being met.

These benefits sound great, but what exactly does data modeling involve? Why is it important? And what different types of data models exist?

Learn other data analyst skills in CareerFoundry's [free 5-day data short course](#).

In this introductory guide, we'll answer all your questions, including:

1. [What is data modeling?](#)
2. [Why is data modeling important?](#)
3. [Types of data models](#)



This month, **save up to €1,400** on our career change

1. What is data modeling?

Data modeling is the process of mapping how data moves from one form or component to another, either within a single database or a data management system.

Data modeling is a fundamental [data wrangling](#) and design task that should happen before any database, computer system, app, algorithm, or other data structure is created. By defining the relationship between different data elements and visually representing these, data modeling helps analysts build systems that are fit for purpose.

If this feels a bit abstract, it might help to think of data modeling as a little like designing a new building. Before an engineer constructs a block of apartments, they must first understand the different elements required and how these interact. Where will the windows and doors go? Where should the pipes come in and out of the building? Crucially, how do all these elements relate to one another? Only with these details mapped out can the engineer hope to create a solid structure that does what it is supposed to do.

Similarly, data modeling helps data analysts define everything they need to know about their data, from data formats and data flows to data handling functions. Only once they have all this information can they hope to create a solid structure that meets their needs.

Data modeling doesn't just serve an upfront purpose, either. Once a database is up and running, the model acts as an all-important reference guide. This allows future data engineers and analysts to understand the principles underlying the database's original design and construction, how it works, and how data is shared between different systems. This is important because no system or database stays the same.

🌟 This month, **save up to €1,400** on our career change

2. Why is data modeling important?

As we explained in section one, the main reason why data modeling is so important is that it informs data structures that are fit for purpose.

However, data modeling offers numerous additional benefits. Some of these include:

- **It provides insights:** While data modeling is the foundation of effective data structures, it also provides useful insights before you reach this point. For example, you'll quickly learn to spot where data are missing or incorrect. By generally improving the understanding of data, data modeling can help high-level decision-making, even before a database or structure is up and running.
- **It tackles core data wrangling tasks:** Data modeling forces analysts to standardize data, establish hierarchies, and generally make the data more consistent and usable. All these tasks are for key [data cleaning](#). So by modeling your data, you are effectively killing two birds with one stone—creating a structural blueprint and tidying your data.
- **It improves communication:** Data modeling involves having a clear understanding of how different stakeholders will use data, what kinds of reports they'll need, and so on. Data modeling inherently encourages clearer communication between different groups, ensuring everyone understands their role and how the data will impact their and others' work.
- **It saves resources:** Designing a database upfront (before you invest time and money in building it) reduces unnecessary duplication of tasks. It also ensures the database won't miss important functionality and it minimizes data storage requirements by identifying and eradicating duplicate data.
- **It supports compliance:** Every organization has statutory data protection

🌟 This month, **save up to €1,400** on our career change

3. Types of data models

When exploring different data models, you'll find there are many individual data models designed for specific data modeling tasks. These range from network models to relational models, and more. However, if you're new to the concept of data modeling, a more helpful distinction at this stage is the different categories of data models.

Broadly speaking, these categories focus on what's known as a model's level of abstraction, or how close to the real world the model is. So, for instance, at a high level of abstraction, a data model will describe the overall structure of a database or information, without focusing on the detail. Meanwhile, at a low level of abstraction, a data model will provide a granular schematic of how a system should be implemented, on a task-by-task basis.

When categorizing data models in this way, there are three main options to choose from, each building on the one that comes before it.

Let's take a look.



A visualization of conceptual, logical, and physical data models, highlighting the different levels of abstraction between them. Source: [Stack oOverflow](#)

Conceptual data modeling

At the highest level is the conceptual data model. A simplified, loosely defined representation of a data system, a conceptual model's purpose is to define a structure's main entities (or table types) and the relationships between them.

The conceptual model is the first step in any data modeling project. It helps designers



This month, **save up to €1,400** on our career change



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

important to getting right, as Analytics rounds often upon details you will develop your own detailed logical and physical models.

Logical data modeling

Building on the conceptual data model, a logical data model is a more thorough representation of a system. It is the first model that describes the data's attributes (or the characteristics that define all items) and keys (sets of attributes that help uniquely identify rows and their relationship to the tables within the model).

A logical data model is useful when you're trying to understand the detailed requirements of a system or network of systems before committing to a full system implementation.

Physical data modeling

Lastly, the physical data model builds on the logical data model. The physical model is a detailed system representation, defining the specific data elements such as table and column names, accounts, indexes, and different data types.

Most commonly created by database administrators and software developers, the physical model outlines the parameters required for your database or database management system (DBMS), including software and hardware. As your end goal, the physical data model is tied to specific database implementations and database management systems.

4. The data modeling process

Okay, now we understand the different types of data models, what does the process



This month, **save up to €1,400** on our career change



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

You may come across different variations of the data modeling process. But they all

tick off the same general steps:

1. Determine the purpose of the model

First, determine the purpose of your model. What problem is the model trying to solve? And what specific requirements need to be met?

For example, if you're planning to use the data for [predictive analytics](#), the model should be designed to reflect this by focusing on the elements most relevant to this task. Determining a focus and clear set of goals will help you to identify the pertinent entities in the model and the relationships between them.

2. Identify the main entities in the model (and their attributes)

The next step is to identify the main entities in your model. Entities are the 'things' in your data that you are interested in.

For example, if you're tracking customer orders, the main entities will be customers and orders. Meanwhile, you'll also need to identify each entity's attributes or values. In this case, customer attributes could be first and last names, and telephone numbers. Attributes for orders, meanwhile, might be the price of the order, what the item is, or the item's SKU.

3. Define the relationships between entities

Once you've identified the main entities in the model, you'll need to define the



This month, **save up to €1,400** on our career change

4. Identify integrity rules and constraints

Integrity rules and constraints keep data accurate and consistent while ensuring that they satisfy the functions of your database.

For instance, the data must be organized logically, and easy to retrieve, update, delete and search. If you've ever [played around with Microsoft Excel](#), you'll be more familiar with the idea of rules and constraints than you realize.

For example, if a column has the 'NOT NULL' constraint, it means the column cannot store NULL values. In practice, this might mean that an order has to have a customer name and a product number to be valid.

5. Identify data that needs to be included in the model

Next up, you'll need to identify any data that needs to be included in the model.

This can be easily achieved by creating a diagram or sample to help you spot where there are gaps in your existing data. The additional data required might be data you already have access to, or it could be external data that needs funneling into your model.

6. Create, validate and update the model

The final step is to create and test the model, using appropriate sample data to ensure it meets the requirements outlined in step one.

Testing the model against real-world data will ensure it accurately presents the correct information and confirm that the model is performing as intended. You may need to

🌟 This month, **save up to €1,400** on our career change

Relational models

A relational model is one in which data is organized into tables, using unique columns and rows. Relational models also use what are known as keys; unique identifiers for each record in the database.

Customer Information			
CustomerID	FirstName	LastName	Address
C0001	John	Smith	123 Example Str.
C0002	Susan	Hopkins	45 Sample Blvd.

A simple relational model showing columns, rows, and keys. Source: phoenixnap.com

While relatively straightforward to create, and useful for easily manipulating and retrieving data, relational models are also quite rigid. This rigidity means they are not always useful for highly complex data analytics tasks, but they remain popular for their ability to bring order to chaotic datasets.

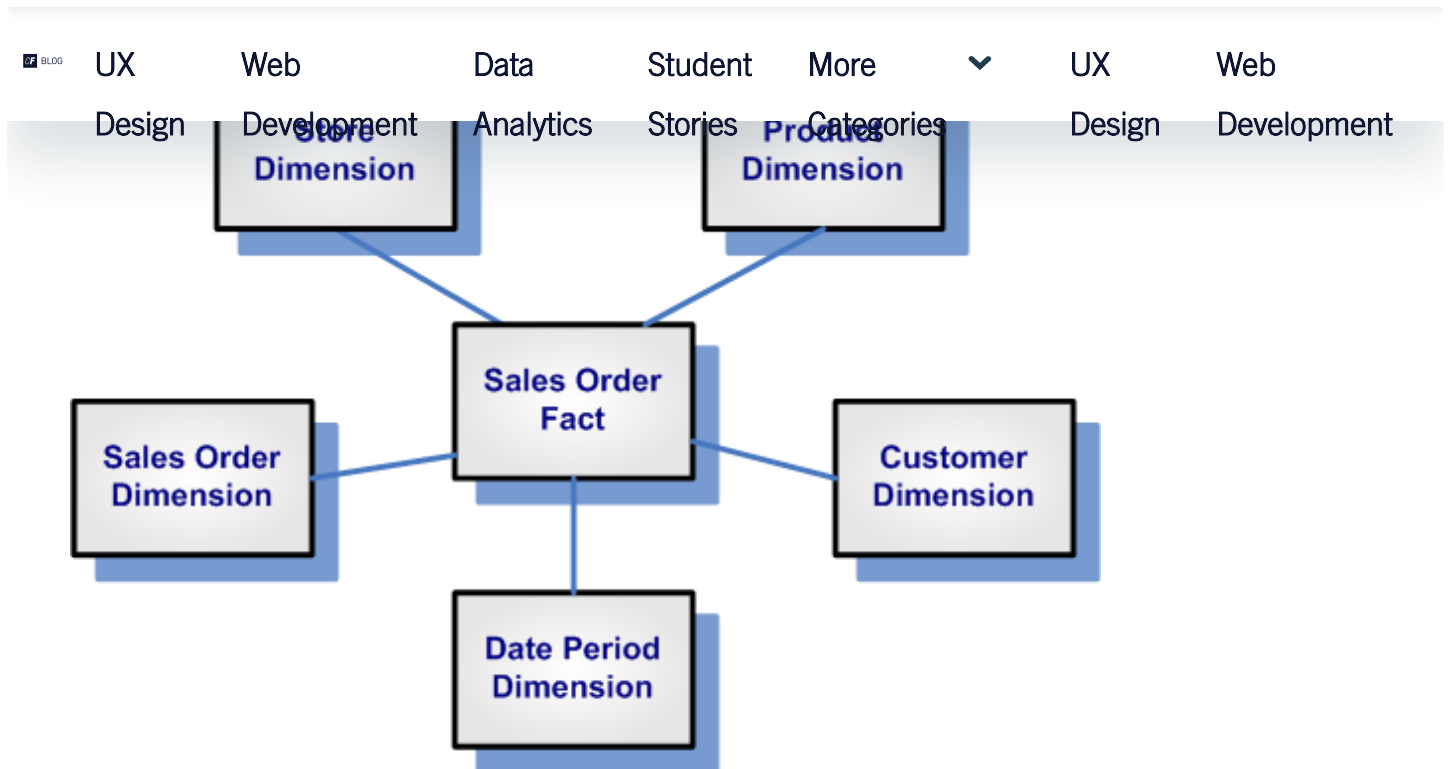
In the real world, relational modeling is often used to organize information like financial transactions, products, calendar events, or any other ordered list of items.

Related reading: [What is a relational database?](#)

Dimensional models

A dimensional data model is one in which data is organized into 'dimensions' and 'measures'.

🌟 This month, **save up to €1,400** on our career change



A very simple dimensional model showing the various dimensions of a single sales order. Source: bi-insider.com

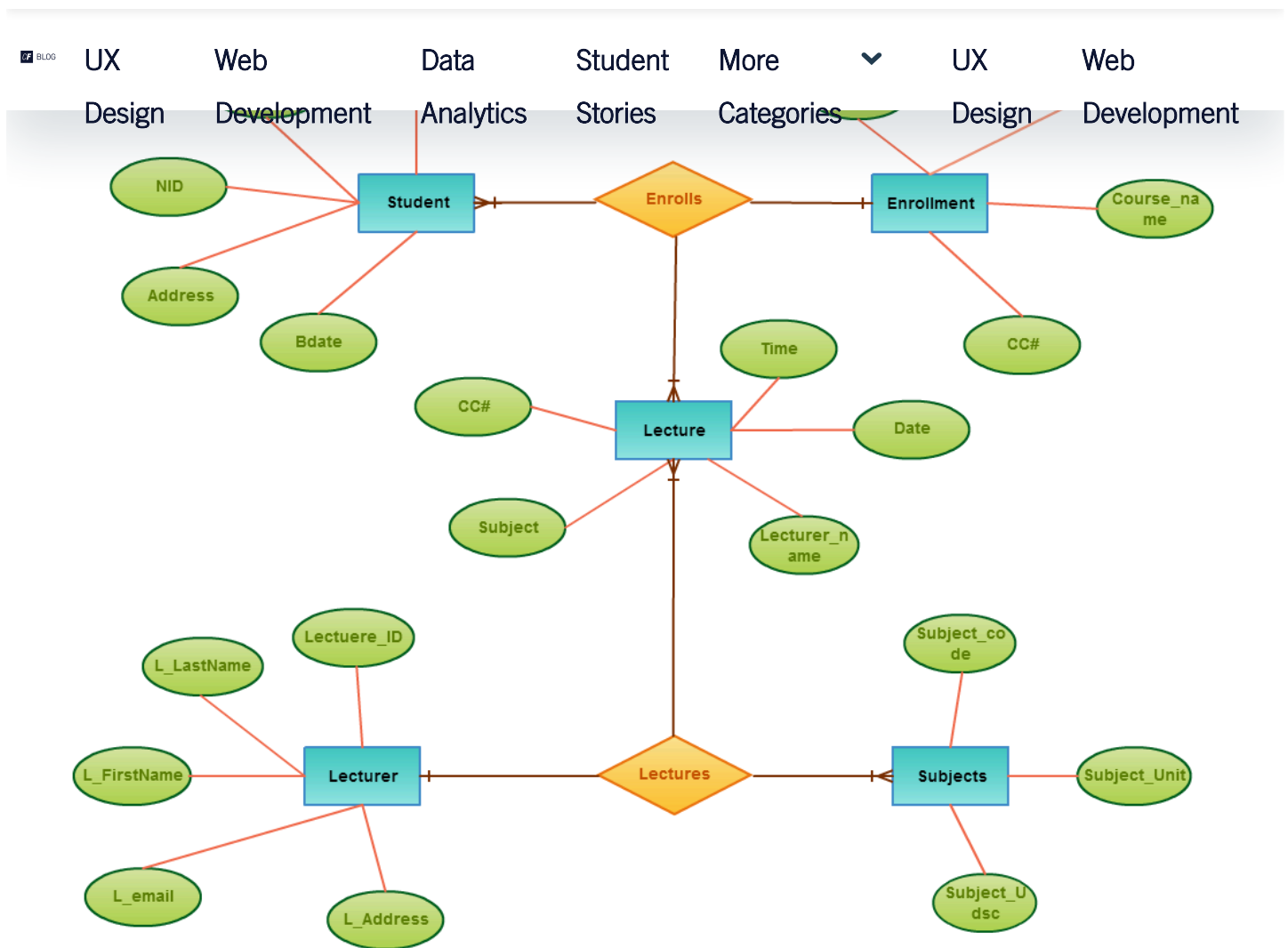
This kind of model is often used in business since it lacks the rigidity of relational models and can be used to map more complex relationships. While this offers greater flexibility it does make the model more challenging to develop. While dimensional models can present more valuable insights, the trade-off is that it is usually harder to extract these.

In the real world, dimensional models are used for many analytics tasks, such as analyzing customer behavior, sales data, or financial trends over time.

Entity-relationship models

An entity-relationship model (or ER model) is a variation of the relational model. It is often used to describe the structure of the relationship between entities within a specific domain area.

🌟 This month, **save up to €1,400** on our career change



A simple ER model for a student enrolment system. Source: [creately.com](https://www.creately.com)

The items of interest we want to track might include people, products, or events. Meanwhile, relationships describe the connections between these entities, such as that between a person and a product, or a product and an event.

In the case of the example shown here, the relationship mapped is between students, enrolments, lectures, and subjects.

In the real world, entity-relationship models are often used to represent the structure of data in software applications, data warehouses, or other information systems.

6. Data modeling tools (and how to choose

★ This month, **save up to €1,400** on our career change



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

However, most data modeling tools follow the same general principles. Namely, a good tool simplifies database design, considers your business rules, and minimizes the risk of unnecessary mistakes.

Some common data modeling tools include:

- [erwin Data Modeler](#)
- [ER/Studio](#)
- [SQL Database Modeler](#)
- [DbSchema](#)
- [IBM Infosphere Data Architect](#)
- [Archi](#)

How to choose a data modeling tool

The best data modeling tool for a given purpose depends on many factors, including your specific needs, the size and complexity of the data set, your organization's strategic objectives, and the available budget.

The best place to start is researching and finding out which tools are available. Once you've created a shortlist, here are some questions to ask yourself:

- Does the tool balance intuitive design (for the average user) with more advanced functionality (for more technical team members)?
- Is high performance important? Is the tool fast enough? Will it work under pressure in the real world?
- Data models require regular amends as your data and the situation changes. Does the tool allow you to update the model easily, or it is a cumbersome task?



This month, **save up to €1,400** on our career change



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

data professionals to see which tools they prefer. Finally, draw up a shortlist of tools and evaluate each one based on its features, price, and user reviews.

7. Benefits and challenges of data modeling

By mapping the relationships between data elements and the rules that govern them, data modeling can help you design an efficient and effective model for your database or DBMS.

There are many benefits to data modeling, including the ability to:

- Organize data in a way that is easy to understand and use
- Reduce data redundancy and improve consistency
- Improve data retrieval and storage
- Share data between different systems
- Improve the quality of data by providing a clear and consistent view

Meanwhile, there are some challenges associated with data modeling, too. These include the need for:

- Careful planning and design to ensure the data model meets the needs of both the system and the business
- Skilled personnel who can understand and manipulate the data model
- Adequate resources (both time and money) to support the data modeling process

Overall, though, solving these challenges is a small price to pay. Trying to cut corners is a false economy and will result in much higher costs in the long run!

8 Summary



This month, **save up to €1,400** on our career change



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

start thinking about the database design and how you will implement it. A good model can lead to a well-performing, scalable database that will serve your organization for many years.

As we've shown, data modeling is a valuable skill with broad applications. To learn more about a potential career in data analytics, why not try this [free, 5-day Data Analytics Short Course](#)? Sign up for five daily lessons direct to your inbox. Alternatively, check out the following introductory data analytics articles:

- [A complete guide to time series analysis and forecasting](#)
- [Data transformation: A beginner's guide](#)
- [What is data mining?](#)

What You Should Do Now

1. Get a hands-on introduction to data analytics and carry out your first analysis with our [free, self-paced Data Analytics Short Course](#).
2. Take part in one of our FREE [live online data analytics events](#) with industry experts, and [read about Azadeh's journey](#) from school teacher to data analyst.
3. Become a [qualified data analyst](#) in just 4-8 months—complete with a [job guarantee](#).
4. 🌟 This month, save up to 20% on our career change programs with our tuition reduction offer. [Book your application call](#) and invest in your future today!

This article is part of:

🌟 This month, **save up to €1,400** on our career change



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

Will is a freelance copywriter and project manager with over 15 years' experience helping firms communicate all things tech- and education-related. His words have been published in print and online, including in the Daily Telegraph, TES, and across other education sector media. He's also known for his ability to complete a Rubik's Cube in under five seconds, but it has to be seen to be believed.



Related Data Analytics Articles



Data Analytics

7 Best Data Warehouse Tools to Explore in 2025

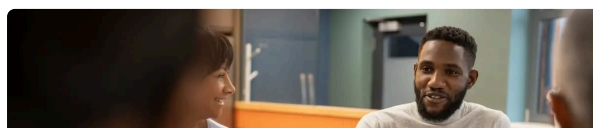
December 18, 2024 - 9 minutes read



Data Analytics

8 of the Best Data Analyst Courses in Germany & How You Could Take Them for Free

May 15, 2024 - 16 minutes read



🌟 This month, **save up to €1,400** on our career change



UX

Web

Data

Student

More



UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

CAREERFOUNDRY

What is CareerFoundry?

CareerFoundry is an online school for people looking to switch to a rewarding career in tech. Select a program, get paired with an expert mentor and tutor, and become a job-ready designer, developer, or analyst from scratch, or your money back.

[Learn more about our programs →](#)




PROGRAMS TO CHANGE YOUR CAREER

[Product Design](#)[UX Design](#)[Full Stack Development](#)

COMPANY

[About Us](#)[Job Guarantee](#)[For Businesses](#)

This month, **save up to €1,400** on our career change

 BLOG

UX

Web

Data

Student

More

▼

UX

Web

Design

Development

Analytics

Stories

Categories

Design

Development

[Intro to Frontend Development](#)

[Intro to Data Analytics](#)

[Intro to Digital Marketing](#)

[Intro to Product Management](#)

ADVANCED COURSES

[Voice User Interface Design](#)

[UI for UX Designers](#)

[Frontend Development for Designers](#)

[Animation for UI Designers](#)

[Python for Web Developers](#)

[Cloud Computing for Web Developers](#)

[Machine Learning with Python](#)

[Data Visualizations with Python](#)

We'd love to hear from you!

[Connect with a program advisor](#)

[TERMS AND CONDITIONS](#) | [PRIVACY POLICY](#) | [COOKIE POLICY](#) | [IMPRINT](#) | [SECURITY](#)

 This month, **save up to €1,400** on our career change