

7.1.3

Determine Entity Relationships

So far, we've looked over our data and familiarized ourselves with it, and we have a better grasp on the connections between the data files through primary and foreign keys. It's quite a lot to take in, but we've gotten pretty far—let's keep up the momentum!

Another crucial part of getting a database ready is preparing a solid foundation. We're doing part of that by familiarizing ourselves with the data, but how will it be represented in SQL? SQL organizes data into tables, and each CSV's data will live in its own table: six CSV files will mean six SQL tables. Let's look more into these tables and their properties.

Table Structure

When working in Excel and Visual Basic for Applications (VBA), we're working directly with worksheets with data. In SQL, the same worksheets we have been exploring are organized into tables instead. They are similar to DataFrames in that they have headers and indexes, with data in columns and rows. Take a look at the following images.

Table 1

Headers
data
data

Table 2

Headers
data
data
data

Table 3

Headers
data
data
data

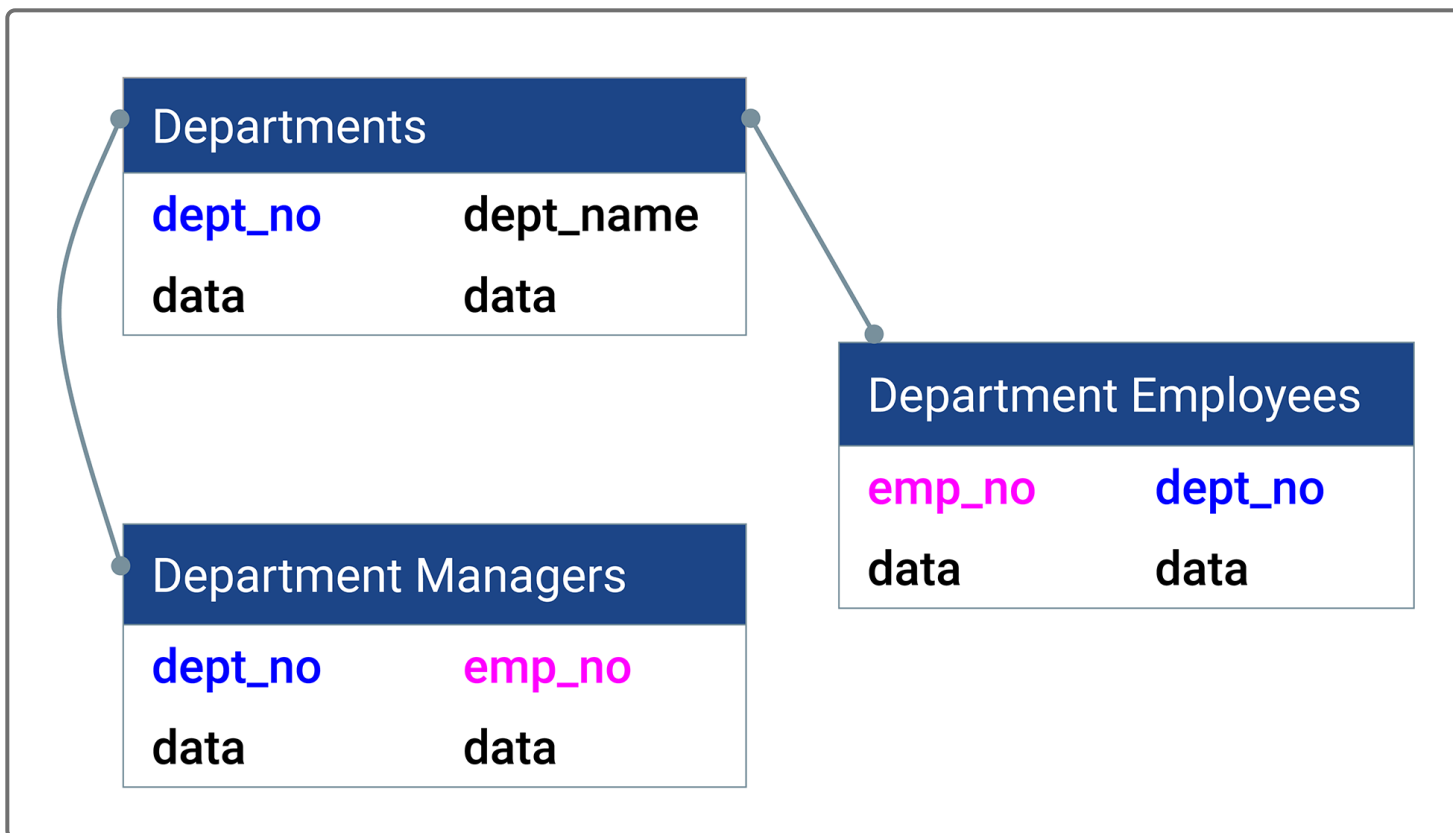
Next we'll cover how table structure comes into play when creating an entity relationship diagram.

Entity Relationship Diagrams (ERDs)

An **entity relationship diagram (ERD)** is a type of flowchart that highlights different tables and their relationships to each other. The ERD does not include any actual data, but it does capture the following pertinent information from each CSV file:

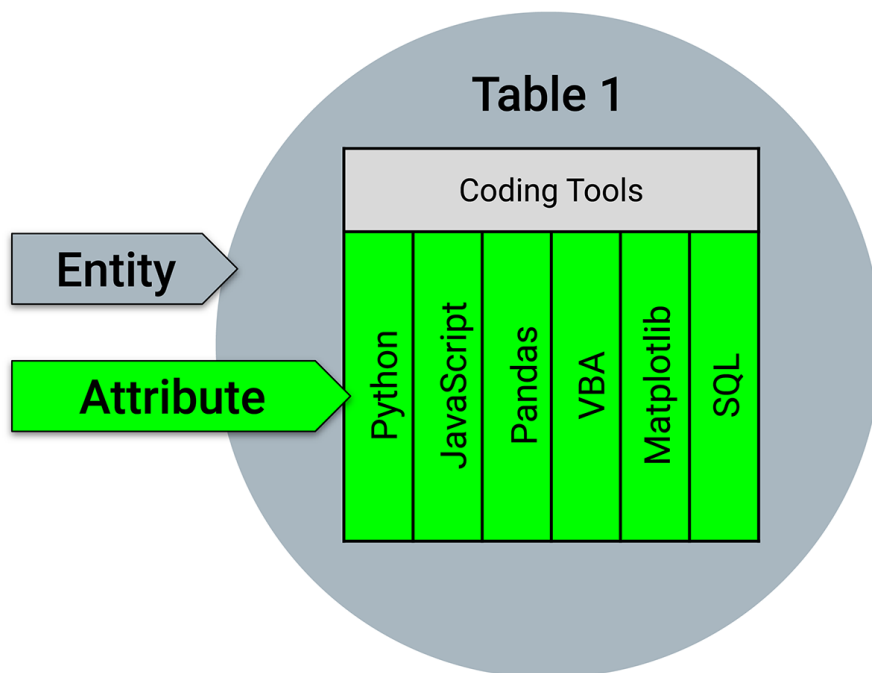
- Primary keys
- Foreign keys
- Data types for each column

The ERD also shows the flow of information from one table to another, as captured in the image below:



In addition to creating new databases, ERDs are used to document existing databases. The visual representation of the tables gives a deeper understanding of the data and the database as a whole.

When creating a diagram, we need to fully understand all of the data being inserted. Database components include tables, known as **entities**, with data, known as **attributes**.



Data types include Booleans, integers, and varying characters (i.e., within a string).

There are three types of ERDs: **conceptual**, **logical**, and **physical**. Each one builds upon the other—you need the conceptual ERD to build a logical ERD to build a physical ERD. We'll learn how to create ERDs later in this module.

© 2020 - 2022 Trilogy Education Services, a 2U, Inc. brand. All Rights Reserved.