

You can pick up where you left off. Just join a new session and we'll reset your deadlines.

Join a session

Back to Week 1

Lessons

Prev

Next

Specialization Introduction: Excel to MySQL: Analytic Techniques for Business

Introduction to Managing Big Data with MySQL

Lesson 1: Problems Databases Solve

Lesson 2: Database Design Tools

Database Design Tools That Will Help You Learn SQL Faster 4 min

How Entity-Relationship Diagrams Work 8 min

Database Structures Illustrated by Entity-Relationship Diagrams 6 min

Relational Schemas 9 min

Lesson 3: Building Your Own Entity-Relationship Diagrams

Lesson 4: Building Your Own Relational Schemas

Lesson 5: Test Your Understanding

In this video, we're going to go over the basic building blocks of a relational

How Entity-Relationship Diagrams Work

Have a question? Discuss this lecture in the week forums.

Interactive Transcript

English

0:04 In this video, we're going to go over the basic building blocks of a relational database and learn how to make entity relationship diagrams. Remember I said that databases are collections of tables that each have their own unified theme. Another thing that I mentioned that we will really emphasize more strongly over the next few videos is that these tables are linked by one or more columns with the same values. These columns allow you to connect rows in different tables.

0:29 As an example, we can see here that the Customer Contact column of the Orders table and the ContactID column of the Customer Contact table are populated by the same types of values. Whenever there is a value of cid1 in a row of the Customer Contact column of the orders table. We can find all the information about that Customer Contact by looking up cid1 in the ContactID column of the Customer Contact table. Figuring out exactly what information you want to collect and exactly how you can organize information to link tables, can be tricky. So database architects often draw diagrams called entity relationship diagrams or Diagrams to communicate the structure of the database they intend to implement to the rest of their team before they actually make the database.

1:12 This is what an Diagram might look like. The person who made this one color coated theirs so that it kind of looks like a bunch of flowers at least to my eye. Normally, they look like a collection of ovals, rectangles and diamonds collected by lines. Clearly, Diagrams are not fine art. What Diagrams are supposed to represent are the conceptual relationships in your data. What is related to what? And how? You will see, even if you never design a database in your life, being able to read these diagrams can be very helpful for writing tricky queries. In addition, looking at lots of Diagrams to get a sense of how different data and databases can be organized, will help you build an intuition for the types of data relationships you will work with in your own databases. Let's dive in and learn what the pieces of these diagrams mean.

1:57 The boxes in the Diagram are called Entities. They represent categories of data your database will keep track of. Each box is one category. Eventually, each entity will probably become a table when the database is made.

2:12 The ovals in the Diagram are called Attributes. These represent aspects of each category or entity that will be recorded. Attributes will likely become the columns in the table that will be built around that entity. Each attribute in the Diagram has to be connected to at least one entity. Further, according to the rules of set theory, each attribute must be unique for that entity. So you shouldn't, in theory at least, have two college start dates in your student entity, for example.

2:42 Let's step away from the Diagrams for a moment to discuss more terminology. Each time you collect data relevant to an entity, those data are collectively called an entity instance. You can think of entity instances as the rows in a data table.

2:57 The most important attribute of the entity is the attribute that has a unique value for each entity instance. Formally, they are called Unique Attributes. Informally, they are also sometimes called unique keys or unique identifiers.

3:11 Unique attributes or keys or identifiers, in table form, would be a column that has a different unique value for every single row. They're like the ID columns we saw earlier that allow us to link tables together. Every entity must have at least one attribute that serves as a unique key or identifier. With that vocabulary under a belt we can now return back to our diagrams. You indicate unique attributes in an entity by underlining their title in the oval. In our diagram, StudentID is the unique attribute for STUDENT, and CollegeName is the unique attribute for COLLEGE.

3:46 This means that every STUDENT will have their own unique StudentID, and every COLLEGE will have its own unique name.

3:53 If there are multiple attributes underlined in an entity, any one of them could potentially be used as a unique key or identifier to link the table with other tables when the database is implemented.

4:04 Every entity or category has to be related to at least one other entity in the database. You use a diamond and lines to represent and describe these relationships. The word in the diamond spells out the nature of the relationship. And the diagram shown here, STUDENT attend COLLEGE.

4:22 Now here's where things get really important for an analyst who will be querying a database tree for information. Each relationship between entities in the Diagram shows how many instances of one entity are associated with how many instances of another entity. If you thing back to when you were asking questions about the EGGER'S ROAST COFFEE spreadsheet. Many of our questions about whether data entry was correct or an error could have been answered if we knew the nature of the relationships between the columns. For example, does each company have one contact person, or multiple contact people? Turns out that there more than one by the way. Is each order supplied from exactly one distribution center, or can there be multiple distribution centers? How many addresses can be associated with a given EGGER'S ROAST COFFEE employee?

5:06 These limitations on the relationships between two entities, which, once the entities have been made into tables, also ends up meaning the relationships between the values in the columns that link the two tables together, are called cardinality constraints. And they're indicated by the little symbols next to the entity rectangles in our Diagram.

5:24 There are a couple of different formats of Diagrams. But in the format we are using in this course, the symbol closest to the rectangle represents the Maximum number of instances of the entity that can be associated with another entity. And the symbol farthest away from the rectangle, represents the Minimum number of instances of that entity that can be associated with another entity. To understand the relationships between entities in the arrow diagrams that have this format. All we start from one entity, follow the connecting line and then interpret the symbols closes to the second entity. So if we start with college in our diagram here, we can see that each COLLEGE is attended by students.

6:01 These symbols here are the ones closest to students. So they are the ones we will use to interpret relationship in this direction. If the symbol next to an entity is a straight line, that means one. If the symbol looks like a bird's foot, that means more than one.

6:15 Remember that the symbol farthest away from the second entity indicates the Minimum while the symbol closest to the second entity equals the Maximum. So the way we would read this diagram is that, each COLLEGE must be attended by at least one STUDENT but can be attended by many students.

6:33 The way we would read the relationship in the opposite direction is that each students attends a minimum of 1 but no more than 1 COLLEGE.

6:44 Sometimes the Diagrams provide more specific cardinality constraints than just one and many. When they do, those constraints are indicated by specific numbers in parenthesis. There are two important things to know about the notation for Specific Cardinality Constraints. First, the numbers in parentheses take precedence over the bar symbol in the minimum constraint. Second, and pay attention to this one because it is potentially confusing, the numbers in the parentheses are always written with the minimum number on the left and the maximum number on the right. In order to try and be consistent with written languages that always write from left to right. Importantly, the numbers are always written in this left to right direction even if the numbers in parenthesis refer to a relationship that is depicted from the right to left direction. I know, that's a little strange, but let me show you what I mean.

7:31 If we read this Diagram in this direction which is from right to left, despite the fact that there is a bar symbol that usually equals 1. And despite the fact that the numbers in parenthesis are written so that the minimum number is on the left and the maximum number is on the right. We interpret the diagram to mean that each college must be attended by a Minimum of 1,000 students and a Maximum of 10,000 students.

7:55 One symbol to know about is the M which I have also sometimes seen as an N. It means Infinite. If the 10,000 was replaced by an M or an N, that would mean that the COLLEGE could be attended by an infinite number of students.

8:10 Another symbol to know about is the circle. The circle means Optional. So if we read this Diagram in a left to right direction, we would learn that each STUDENT could attend as many as two colleges, but doesn't have to attend any.

8:24 By the way, as a small detail, for this interpretation to make sense, there would have to be some other attributes in the database not pictured here. That could be used to explain, why the STUDENT was relevant to the database even if he or she wasn't attending a COLLEGE. For example, perhaps they transferred nor dropped out. So now you understand how to read the most fundamental parts of an Diagram. In the next video, we'll talk through some Diagrams, that will give you a sense of some of the more complicated organizations you might see in databases you work with in real life.

Downloads

Lecture Video	mp4
Subtitles (English)	WebVTT
Transcript (English)	txt

Would you like to help us translate the transcript and subtitles into additional languages?