SQL Joins and Unions

This lesson introduces the concept of joins. As you'll explore in this lesson, a join combines two or more tables based on a column that both tables share.

So far, you've used SQL commands and statements to query a single table at a time. Now, you'll learn how to combine multiple tables in a database so that you can view all related data together.

Say you have information about your customers in two tables.

One table contains data about each customer's purchases, and the other table contains data about each customer's demographics.

If you could view all of your customer information in one table, it would be easier to see which customers made which purchases. You may even see trends; for example, you may see that certain demographic groups are more likely to purchase certain products.

Join clauses

First, what is a join? A join involves combining two or more tables, as described here:

Join: A clause that combines two or more tables that are linked by a shared column

This module focuses on joining two tables together because this is the most common scenario; however, it's also possible to join many tables together.

For the tables to be combined, they must share a common column. In the image below, for example, there are two tables. Combining these two tables isn't possible because you can't tell which columns in Table 2 correspond to which customers in Table 1. You can't identify any common relationship between the two tables.

Table 1	
Customer	**Purchase Date**
Sheila M.	6/6/2021
Omar Q.	5/4/2021
Leyla B.	4/28/2021
Shannon F.	6/1/2021
Justin A.	5/30/2021

Table 2	
Product Category	**Delivery Date**
Kitchen Appliance	5/9/2021
Electronics	5/6/2021
Food & Grocery	6/10/2021
Baby Clothes	6/16/2021
Shoes	6/11/2021

On the other hand, the tables in the diagram below can be combined because now they do have a common relationship: the Customer column. The Customer column is the link between the two tables. In Table 1, you can see that Leyla B. has a purchase date of April 28, 2021. And in Table 2, you can see that Leyla B. has a delivery date of May 6, 2021.

Table 1	
Customer	Purchase Date
Sheila M.	6/6/2021
Omar Q.	5/4/2021
Leyla B.	4/28/2021
Shannon F.	6/1/2021
Justin A.	5/30/2021

Table 2	
Customer	Delivery Date
Omar Q.	5/9/2021
Leyla B.	5/6/2021
Shannon F.	6/10/2021
Sheila M.	6/16/2021
Justin A.	6/11/2021

When you combine the two tables by joining on the Customer column, you get the table below. This is much easier to view than two separate tables. Now you can see at a glance that Leyla B. placed her order on April 28, 2021, and received her delivery on May 6, 2021.

Joined Table		
Customer	Purchase Date	Delivery Date
Sheila M.	6/6/2021	6/16/2021
Omar Q.	5/4/2021	5/9/2021
Leyla B.	4/28/2021	5/6/2021
Shannon F.	6/1/2021	6/10/2021
Justin A.	5/30/2021	6/11/2021

You can think about joining tables as adding columns from one table to the other table, using the common column as the link.