

About

When working with data, you often find that to answer a question, you need two separate tables containing different, but related, data.

Join links two tables to create a single worksheet with columns from both. The change applies to your current worksheet and does not affect the original tables.

Example

If your Sales Detail table contains sales information and your Customer Master table contains customer details, Join creates a worksheet with all information about sales and customers.

Using just Sales Detail, you can easily find all sales over a certain amount. With Customer Master, you can find all customers within a given category. Both are simple filter operations. But what if you want to find all customers within a given category with sales over \$100?

Join the tables.

1. Open **Sales Detail**, then click **Join**.

2. Choose **Customer Master** as the foreign table to join with. It's best practice to start with the largest table as your base table, and join in the smaller foreign table.



3.

In **Join on columns**, choose two *matching* columns from the dropdowns, one from each table. For this example, choose **customer_key**. The colors next to the column names represent the column's data type, like integer or text. Data types need to match in order to join.

Trans ID	Date	Time	SKU	UPC
1000012	12/31/12	13:57:51	146841	4369505631
1000021	12/31/12	05:07:39	300423	999999901596
1000031	12/31/12	17:56:09	299952	999999901125
1000031	12/31/12	06:40:29	300038	999999901211
1000038	12/31/12	13:53:15	300442	999999901615
1000044	12/31/12	21:49:57	231800	7266860100
1000044	12/31/12	21:43:27	300306	999999901479
1000044	12/31/12	19:29:48	101982	1200081003
1000049	12/31/12	01:15:09	302015	999999903188
1000050	12/31/12	01:33:03	110746	2172
1000051	12/31/12	00:03:58	299118	999999900291
1000056	12/31/12	08:26:06	225783	7143934030

4.

You can add a suffix to the column names and labels you'll be importing. This will help as you dive deeper into the analysis.

5.

Click **Join configuration**, then **Keep all rows**. There are a few different ways to configure the join. The default matches all values exactly, but you can choose to join on a closest match.

Keep all rows finds matching rows in the foreign table and adds columns to matching rows in the base table. All rows are kept from the current table, whether or not there is a match. Rows in the imported table without a match are not added.

Keep only rows that have a match retains only matching rows from the base table. Columns are added to matching rows in the base table. This is the most commonly used type of join.

With both configurations, if there are multiple matching rows in the base table, you can choose to bring in either only the first, or all, matching rows.

Keep only rows that do not have a match retains only the rows in the base table that do not have a match in the foreign table. Columns are not added to rows in the base table.

6.

You also can configure how to include or exclude rows. To hide columns irrelevant to your analysis, click **Choose imported columns to leave out**.

Click **Hide All** or **Show All**, or drag and drop specific columns.

The screenshot shows the 'Timeline' configuration panel. At the top right is a 'Run all' button. The main section is titled 'Join on columns' and contains two identical entries: a blue box with '123' and a text box with 'customer_key (Customer Key)'. Below this is a link 'Add another column'. Further down are three sections: 'Add suffix to imported column names', 'Join configuration', and 'Choose imported columns to leave out'. The 'Choose imported columns to leave out' section is expanded, showing a search bar 'Search column by name', a table with 8 visible and 4 hidden columns, and buttons for 'Hide All' and 'Show All'. The visible columns are 'Last Name', 'First Name', 'Customer Key', and 'Gender'.

7.

Click **Run**, and that's it! You've got a worksheet with all the columns you need to perform the analyses you need.

Takeaways

- Join two tables to create a larger worksheet that includes data from both.
- Depending on your analysis, configure to include or exclude rows.
- Hide columns irrelevant to your analysis.
- Manipulate the resulting worksheet the same way you would manipulate the base table.

