

UNDERSTANDING BASIC JOINS

Joins combine tables by adding more columns of data across similar row structures. This can cause data loss or duplication if tables are at different levels of detail and joined data sources must be fixed before analysis can begin.

When **joining** tables, the fields that are joined on must be of the **same data type**. If we change the data type after performing the join on the tables, the join will break.

Fields used in the join clause cannot be removed without breaking the join

UNDERSTANDING BASIC JOINS

Data blending simulates a **traditional left join**.

The **main difference** between a **Join** and **Blend** is when the **aggregation** is performed.

A **join** first combines the **data at a row level** and then carries out the **aggregation**.

A **blend aggregates** and **then combines the data**.

In most scenarios, traditional table joins are the best choice as they tend to be most performant as the work is done by the database rather than the local machine.

Table joins are better when tables have a 1:1 relationship (i.e., there is only one record for each value in the linking fields in each table).

If the tables have a 1:many or many: many relationship this creates duplicated data after the join

Duplicated data can be resolved after the table join is created, but the increased amount of data may cause performance issues.

DETAILS OF THE DATA SOURCE USED FOR JOINS

We will be using a fictitious spreadsheet Understanding_Joins.xlsx having 2 tables viz Projections and Actuals for explaining joins

Projections Table

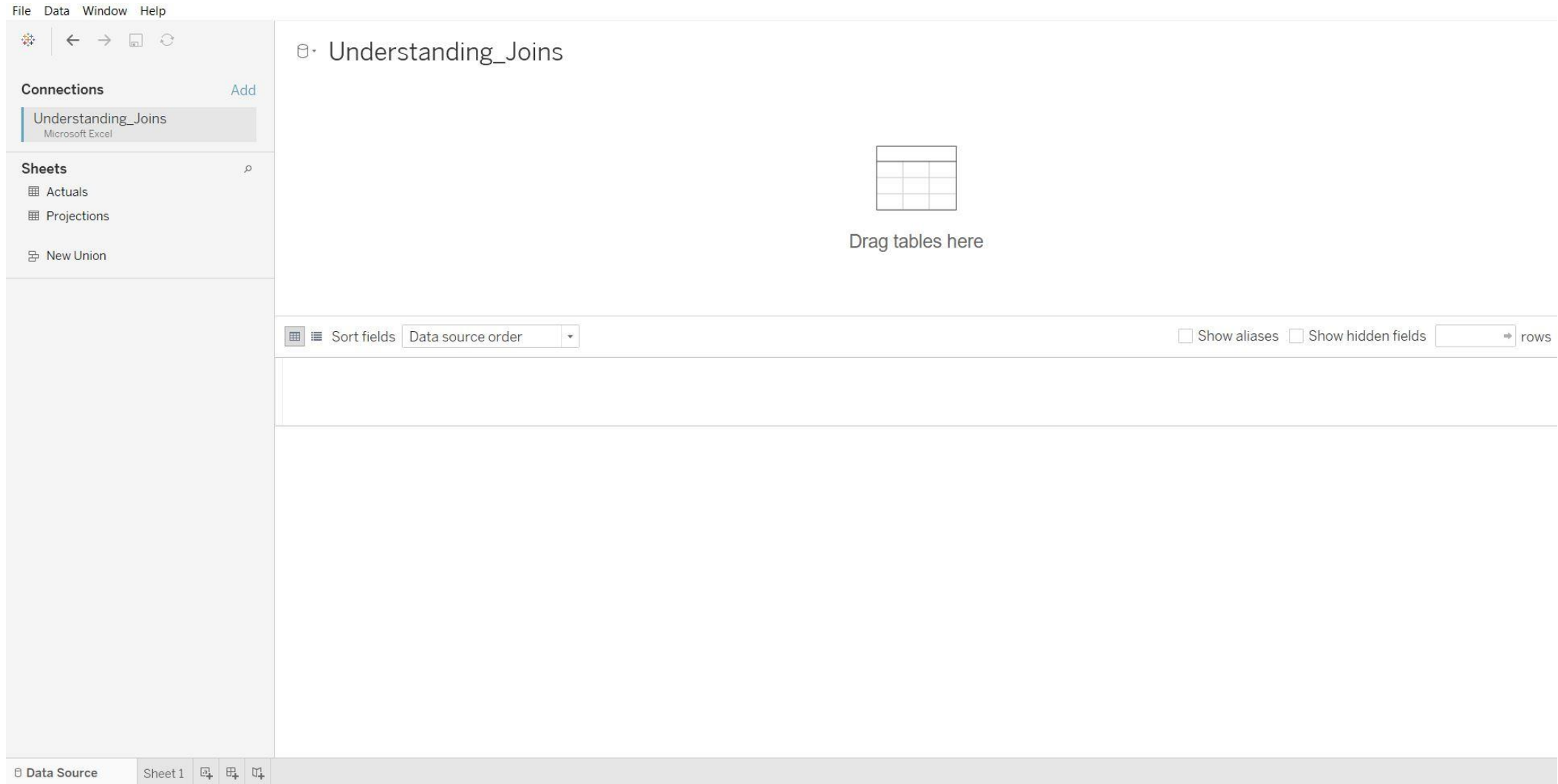
Team	Place	Projected Score
Team A	Place A	225
Team B	Place B	345
Team C	Place C	175
Team D	Place D	267
Team E	Place E	300
Team H	Place H	145
Team I	Place I	254

Actuals Table

Team Name	Place Name	Actual Score	Entered By
Team A	Place A	175	Person 1
Team B	Place B	300	Person 2
Team C	Place C	225	Person 3
Team D	Place D	190	Person 4
Team E	Place E	256	Person 5
Team F	Place F	335	Person 6
Team G	PlaceG	195	Person 7

STEPS TO CREATE A JOIN

Step 1: Connect to the relevant data source or sources e.g.: Understanding_Joins.xlsx



STEPS TO CREATE A JOIN

Step 2: Drag the first table **Projections** to the canvas

File Data Window Help

← → ↺ ↻

Connections [Add](#)

Understanding_Joins
Microsoft Excel

Sheets [p](#)

Actuals


Projections

New Union

Projections (Understanding_Joins)

Filters
0 [Add](#)

Projections



Need more data?
Drag tables here to relate them. [Learn more](#)

Sort fields Data source order

☐ Show aliases ☐ Show hidden fields 7 rows

Team	Place	Projected Score
Team A	Place A	225
Team B	Place B	345
Team C	Place C	175
Team D	Place D	267
Team E	Place E	300
Team H	Place H	145
Team I	Place I	254

[Go to Worksheet](#)

Data Source Sheet 1

STEPS TO CREATE A JOIN

Step 3: Select **Open** from the menu or **double-click** the first table to open the join canvas (**physical layer**)

File Data Window Help

Connections [Add](#)

Understanding_Joins
Microsoft Excel

Sheets

- Actuals
- Projections
- New Union

Projections (Understanding_Joins)

Filters 0 [Add](#)

Projections

- Open...
- Rename
- Remove
- Field names are in first row
- Generate field names automatically
- Convert to Union...

Need more data?
Drag tables here to relate them. [Learn more](#)

Sort fields Data source order

☐ Show aliases ☐ Show hidden fields 7 rows

Team	Place	Projected Score
Team A	Place A	225
Team B	Place B	345
Team C	Place C	175
Team D	Place D	267
Team E	Place E	300
Team H	Place H	145
Team I	Place I	254

Go to Worksheet

Data Source Sheet 1

STEPS TO CREATE A JOIN

Step 4: Double-click or drag the **Actuals** table to the join canvas

The screenshot shows a data tool interface with a sidebar on the left and a main workspace. The sidebar has two sections: 'Connections' with a link to 'Understanding_Joins' (Microsoft Excel) and an 'Add' button; and 'Sheets' with a list of 'Actuals', 'Projections', and a 'New Union' button. The main workspace is titled 'Projections+ (Understanding_Joins)' and contains a diagram showing two tables, 'Projections' and 'Actuals', connected by a join symbol (two overlapping circles with a red exclamation mark). A 'Join' dialog box is open in the center, showing four join types: 'Inner' (selected), 'Left', 'Right', and 'Full Outer'. Below the join types, there are columns for 'Data Source' and 'Actuals'. A search bar is visible above the 'Data Source' column. At the bottom of the workspace, there is a table with columns: 'Team Name', 'Place Name', 'Actual Score', 'Entered By', 'Team', 'Place', and 'Projected Score'. The table has two rows of data, one for 'Actuals' and one for 'Projections'.

Connections [Add](#)

Understanding_Joins
Microsoft Excel

Sheets

- Actuals
- Projections
- New Union

Projections+ (Understanding_Joins)

Projections is made of 2 tables. ⓘ

Projections Actuals

Join

- Inner
- Left
- Right
- Full Outer

Data Source Actuals

Search

Place

Projected Score

Team

Create Join Calculation...

Sort fields Data source

Team Name	Place Name	Actual Score	Entered By	Team	Place	Projected Score
Actuals	Actuals	Actuals	Actuals	Projections	Projections	#
Actuals	Actuals	Actuals	Actuals	Projections	Projections	Projected Score

Show a

STEPS TO CREATE A JOIN

Step 5: Select the connecting field i.e., **Place** from the drop-down of Left side table

The screenshot illustrates the process of creating a join in Power BI. On the left, the 'Connections' pane shows 'Understanding_Joins' from 'Microsoft Excel'. The 'Sheets' pane lists 'Actuals', 'Projections', and 'New Union'. The main area displays 'Projections+ (Understanding_Joins)' with a diagram showing 'Projections' and 'Actuals' tables connected by a join icon. A 'Join' dialog box is open, showing the 'Left' join type selected. The 'Data Source' table has a dropdown menu open with 'Place' selected. The 'Actuals' table has a dropdown menu open with 'Place' selected. The background shows a preview of the data tables:

Team Name	Place Name	Actual Score	Entered By	Team	Place	Projected Score
Abc	Abc	Abc	Abc	Abc	Abc	#

STEPS TO CREATE A JOIN

Step 6: Select the connecting field i.e., **Place Name** from the drop-down of Right-side table

The screenshot shows a data tool interface with a sidebar on the left and a main workspace. The sidebar has two sections: 'Connections' and 'Sheets'. The 'Connections' section shows a connection named 'Understanding_Joins' from 'Microsoft Excel'. The 'Sheets' section shows two sheets: 'Actuals' and 'Projections', and a 'New Union' button. The main workspace shows a diagram of two tables, 'Projections' and 'Actuals', connected by a join symbol. Below the diagram is a 'Join' dialog box. The dialog has four tabs: 'Inner', 'Left', 'Right', and 'Full Outer'. The 'Inner' tab is selected. Below the tabs are two columns: 'Data Source' and 'Actuals'. The 'Data Source' column has the value 'Place'. The 'Actuals' column has a dropdown menu open, showing a search bar and a list of fields: 'Actual Score', 'Entered By', 'Place Name', 'Team Name', and 'Create Join Calculation...'. The 'Place Name' field is selected. Below the dialog is a table with columns: 'Team Name', 'Place Name', 'Actual Score', 'Entered By', 'Team', 'Place', and 'Projected Score'. The table has two rows of data: one from the 'Actuals' table and one from the 'Projections' table.

Projections+ (Understanding_Joins)

Projections is made of 2 tables. ⓘ

Projections Actuals

Join

Inner Left Right Full Outer

Data Source Actuals

Place = Search

Actual Score
Entered By
Place Name
Team Name
Create Join Calculation...

Place Name

Actuals	Actuals	Actuals	Actuals	Projections	Projections	Projections
Team Name	Place Name	Actual Score	Entered By	Team	Place	Projected Score
Actuals	Actuals	Actuals	Actuals	Projections	Projections	Projections

STEPS TO CREATE A JOIN

Step 7: When finished, close the join dialog and join canvas.

← → ↺ ↻

Connections Add

Understanding_Joins
Microsoft Excel

Sheets ⌵

Actuals

Projections

New Union

Projections+ (Understanding_Joins)

Filters
0 Add

Projections is made of 2 tables. ⓘ

Projections

Actuals

Join

Inner

Left

Right

Full Outer

Data Source

Actuals

Place

=

Place Name

Add new join clause

Sort fields

Data source

☐ Show aliases ☐ Show hidden fields 5 rows

Actuals	Actuals	#	Actuals	Actuals	Projections	Projections	#
Team Name	Place Name	Actual Score	Entered By	Team	Place	Projected Score	
Team A	Place A	175	Person 1	Team A	Place A	225	
Team B	Place B	300	Person 2	Team B	Place B	345	
Team C	Place C	225	Person 3	Team C	Place C	175	
Team D	Place D	190	Person 4	Team D	Place D	267	
Team E	Place E	256	Person 5	Team E	Place E	300	

STEPS TO CREATE A JOIN

Step 8: Now we see the final join created

←

→

↺

Connections

Add

Understanding_Joins

Microsoft Excel

Sheets

Actuals

Projections

New Union

Projections+ (Understanding_Joins)

Filters
0 | Add

Projections is made of 2 tables. ⓘ

Projections

Actuals

Sort fields

Data source order

☐ Show aliases

☐ Show hidden fields

5

⇒

rows






Actuals	Actuals	#	Actuals	Projections	Projections	#
Team Name	Place Name	Actual Score	Entered By	Team	Place	Projected Score
Team A	Place A	175	Person 1	Team A	Place A	225
Team B	Place B	300	Person 2	Team B	Place B	345
Team C	Place C	225	Person 3	Team C	Place C	175
Team D	Place D	190	Person 4	Team D	Place D	267
Team E	Place E	256	Person 5	Team E	Place E	300

JOIN TYPES

In general, there are four types of joins that we can use in Tableau:
inner, left, right, and full outer.

If we aren't sure what join type, we want to use to combine data from multiple tables, it is preferred to use relationships.

Not all databases support all join types. If an option is unavailable in the join dialog, it is likely due to a constraint from the data source.

Join Type	Result
Inner 	<p>When you use an inner join to combine tables, the result is a table that contains values that have matches in both tables.</p> <p>When a value doesn't match across both tables, it is dropped entirely.</p>
Left 	<p>When you use a left join to combine tables, the result is a table that contains all values from the left table and corresponding matches from the right table.</p> <p>When a value in the left table doesn't have a corresponding match in the right table, you see a null value in the data grid.</p>
Right 	<p>When you use a right join to combine tables, the result is a table that contains all values from the right table and corresponding matches from the left table.</p> <p>When a value in the right table doesn't have a corresponding match in the left table, you see a null value in the data grid.</p>
Full outer 	<p>When you use a full outer join to combine tables, the result is a table that contains all values from both tables.</p> <p>When a value from either table doesn't have a match with the other table, you see a null value in the data grid.</p>
Union 	<p>Though union is not a type of join, union is another method for combining two or more tables by appending rows of data from one table to another. Ideally, the tables that you union have the same number of fields, and those fields have matching names and data types. For more information about union</p>

SUMMARY OF JOINS

Given below are some of the important points w.r.t Joins:

1. The **visual cue** for a **join** is a **Venn Diagram**
2. We need **at least 1 data source** to create a **join**
3. We can have **multiple connections** while creating **joins**
4. The **structure** created during a join is a **new combined form**
5. **Tables** are **combined** via a **Physical join**
6. Numbers are **aggregated** based on the **join level granularity**