

Lesson 05 Demo 02

Joins and Unions

Objective: To demonstrate the process of joins and unions to enrich and consolidate the data for comprehensive analysis

Tools required: Tableau Desktop

Prerequisites: None

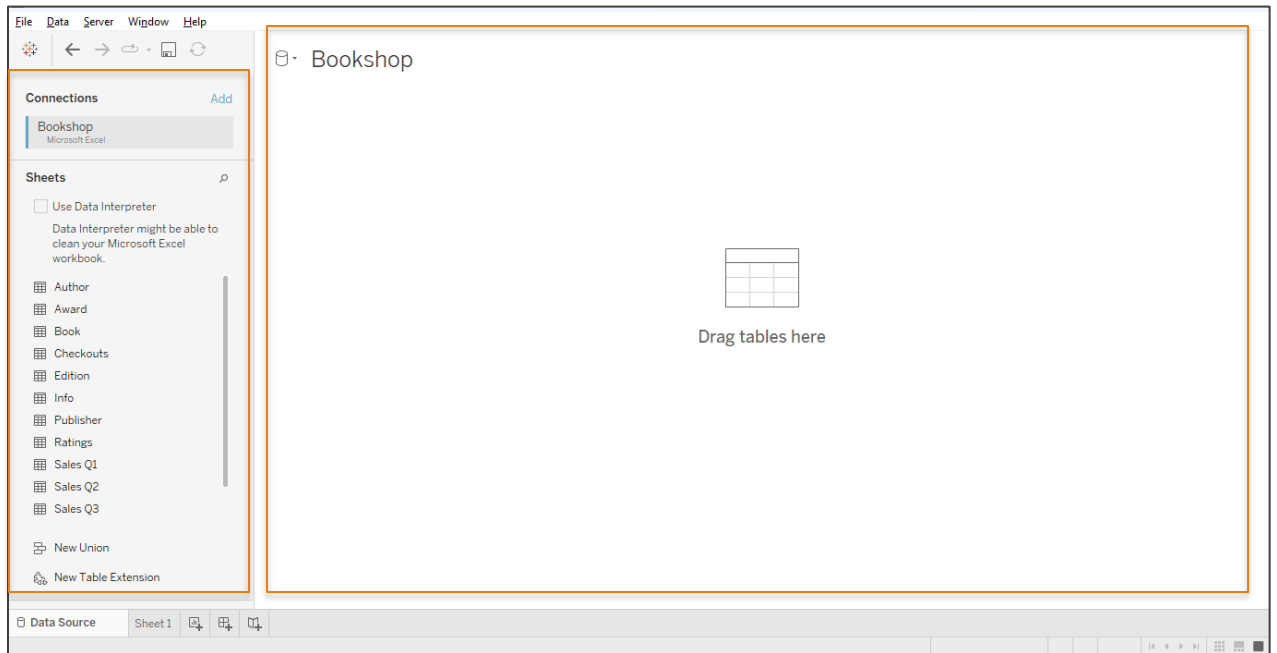
Note: Download the **Bookshop** dataset from the Resource Materials of the LMS

Steps to be followed:

1. Import the dataset
2. Create a join between the Book table and the Info table
3. Create a union

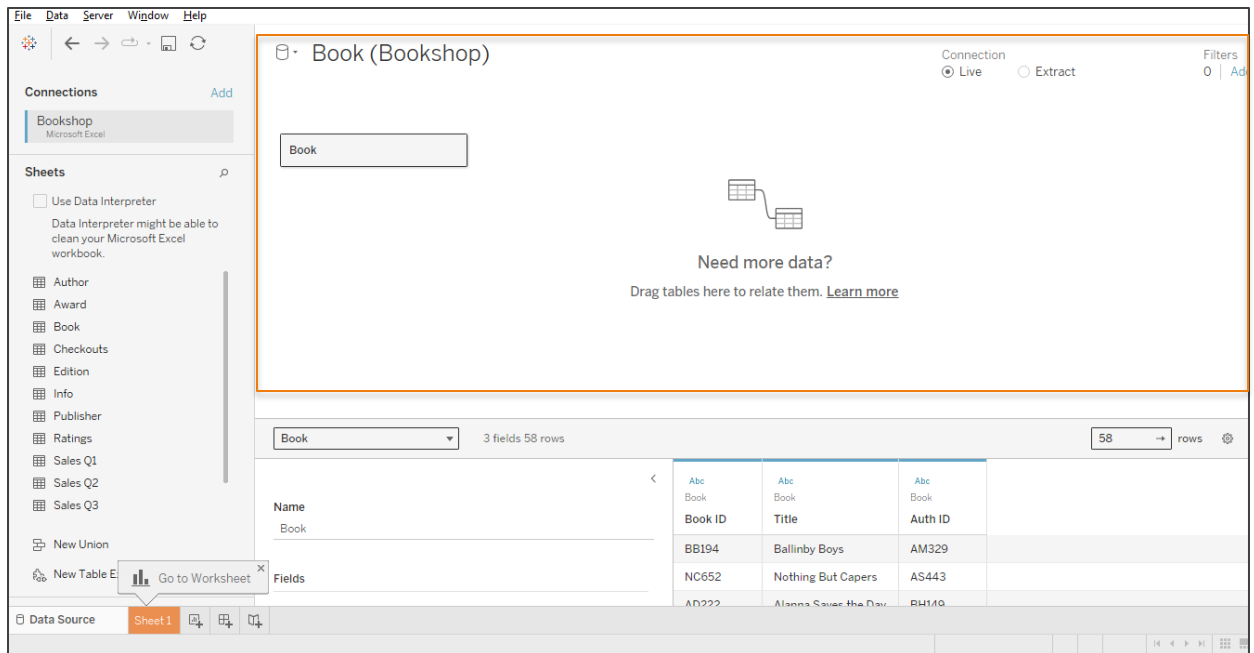
Step 1: Import the dataset

- 1.1 Open a new Tableau file and connect with the **Bookshop** dataset. The Data Source view will look as below.



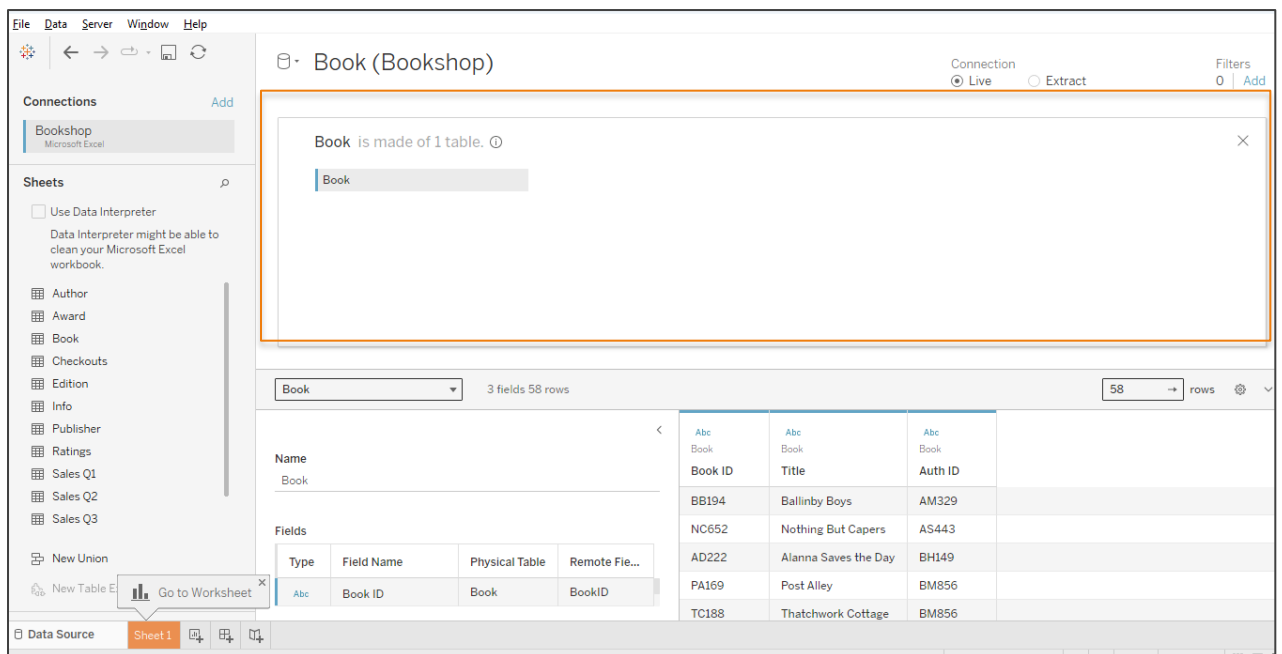
Note: All the tables are listed on the **Connection pane** (on the left). The white canvas that you see on the Data Source view is called the **Data Model section**.

- 1.2 Drag the **Book** table into the data model. The logical layer gets activated as you drag a table into the data model section.



Step 2: Create a join between the Book table and the Info table

- 2.1 Double-click on the **Book** Table icon. The physical layer (the gray box) of the Book Table will open as shown below:



Note: In this physical layer, we can create Joins between tables.

2.2 Now, drag the **Info** Table and put it inside the physical layer of the **Book** Table. The physical layer of the **Book** Table looks as shown

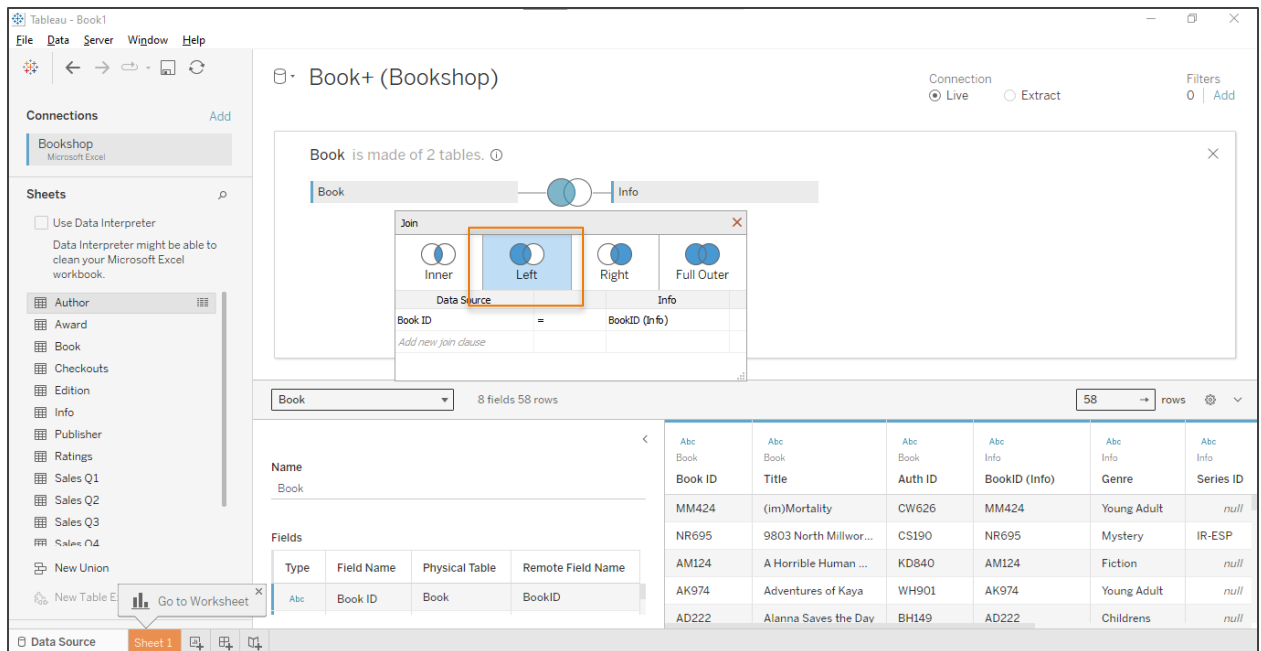
Note: The Join icon connects the two tables in the physical layer. This Join icon can be changed to represent the different types of Joins.

below.

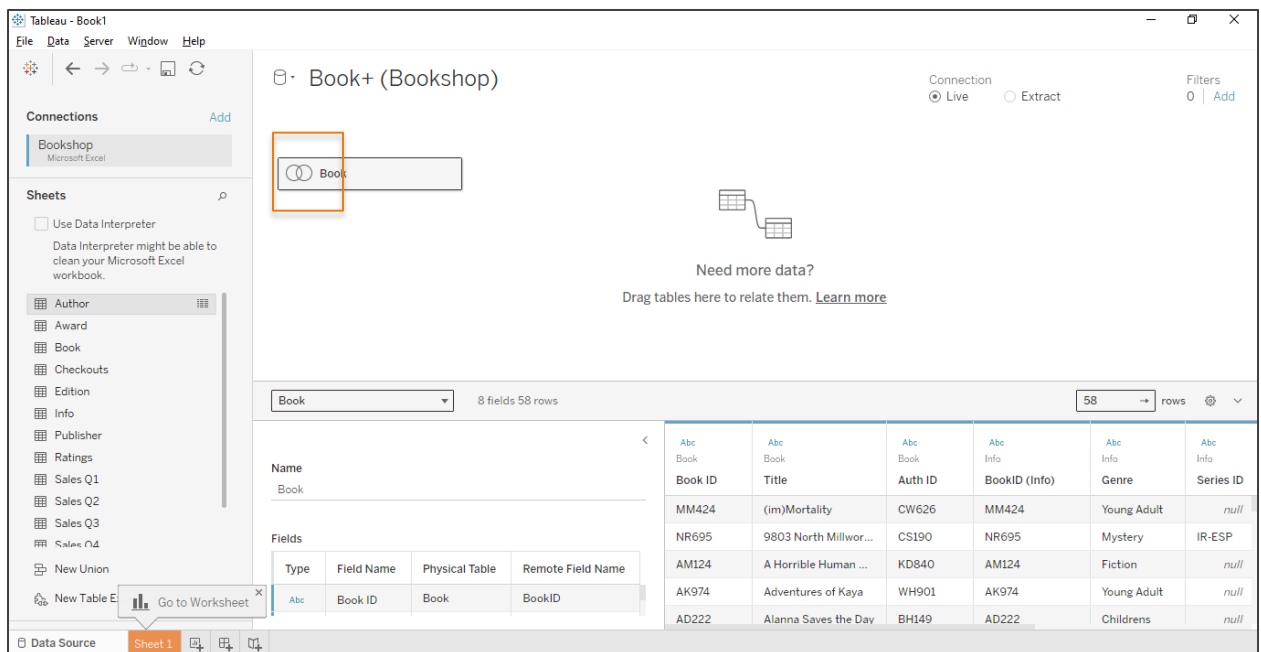
The screenshot shows the Tableau interface with a workbook named 'Book+ (Bookshop)'. The 'Connections' pane on the left shows 'Bookshop' as a Microsoft Excel connection. The 'Sheets' pane shows a list of tables: Author, Award, Book, Checkouts, Edition, Info, Publisher, Ratings, Sales Q1, Sales Q2, Sales Q3, and Sales Q4. The 'Book' table is selected in the 'Physical Table' column of the 'Fields' pane. The 'Info' table is selected in the 'Remote Field Name' column. The 'Join' icon (a circle with a horizontal line) is placed between the two tables. The 'Data Source' pane shows the 'Book' table with 8 fields and 58 rows. The 'Info' table is also visible in the 'Data Source' pane. The 'Data Source' pane shows the 'Book' table with 8 fields and 58 rows. The 'Info' table is also visible in the 'Data Source' pane.

Book ID	Title	Auth ID	BookID (Info)	Genre	Series ID
BB194	Ballinby Boys	AM329	BB194	SciFi/Fantasy	null
NC652	Nothing But Capers	AS443	NC652	Nonfiction	null
AD222	Alanna Saves the Day	BH149	AD222	Childrens	null
PA169	Post Alley	BM856	PA169	Fiction	null
TC188	Thatchwork Cottage	BM856	TC188	Fiction	null

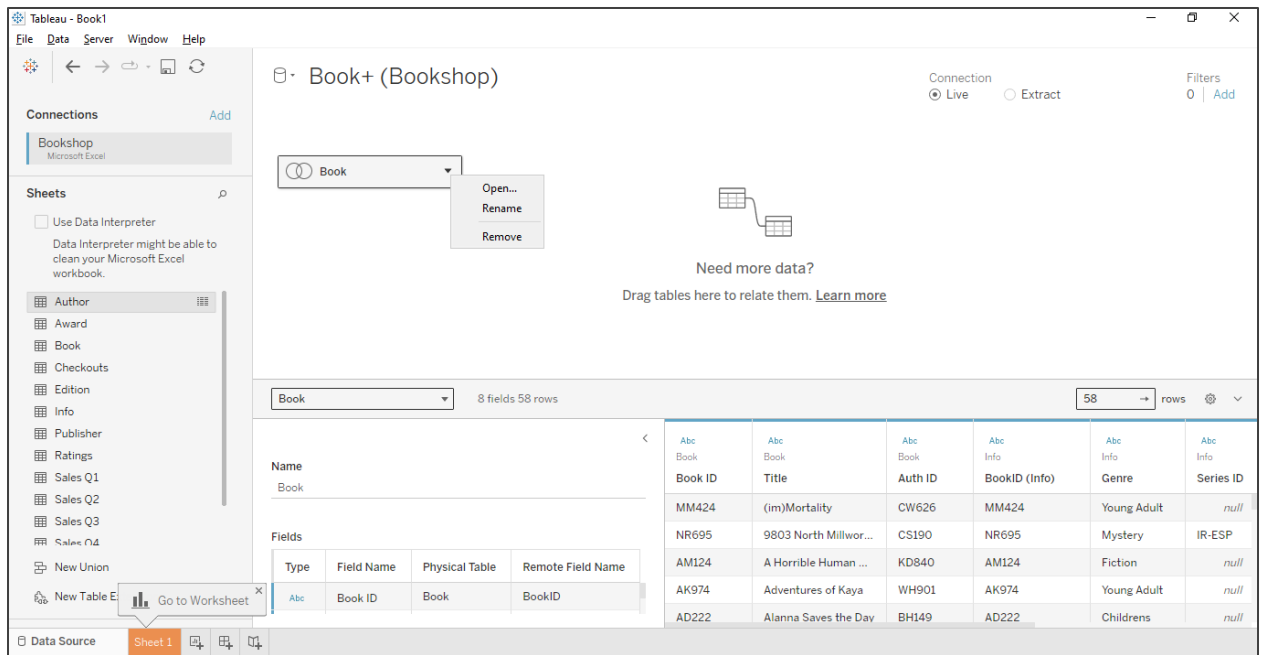
2.3 Left-click on the Join icon to view the four different types of Joins available in Tableau. Select the **Left Join** for **Book** and **Info** Table and close the Join option window and the physical layer.



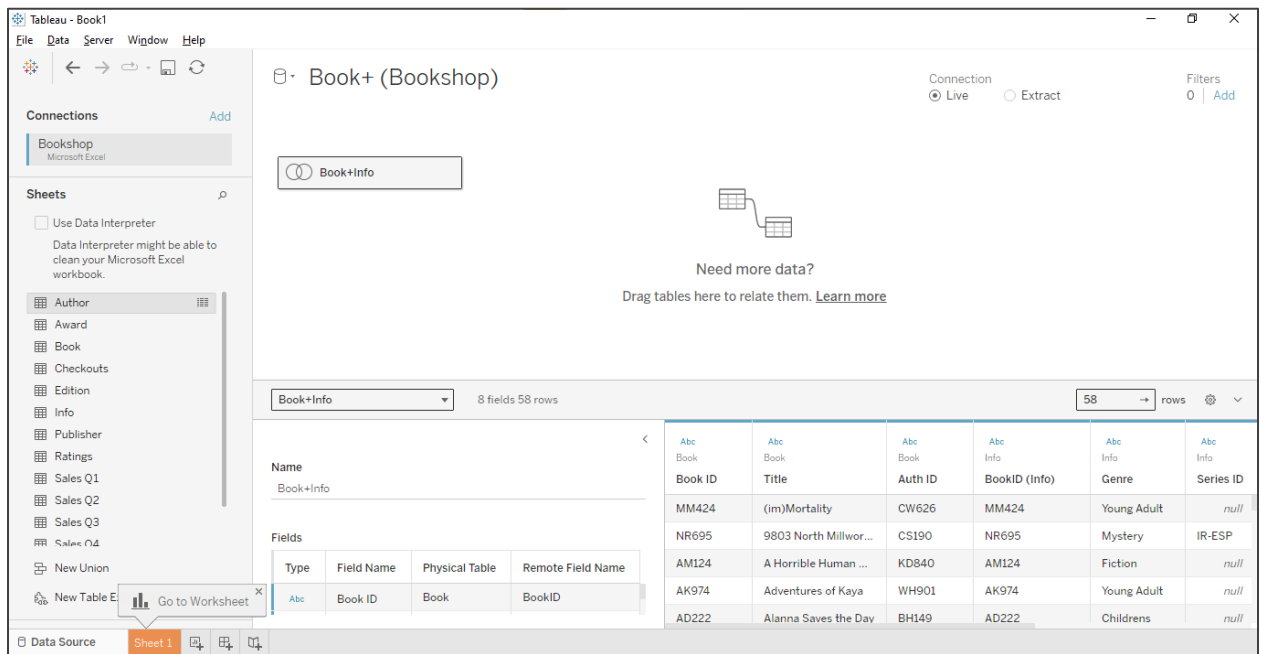
Observation: The **Book** Table will have a small icon indicating a Join is present in this table.



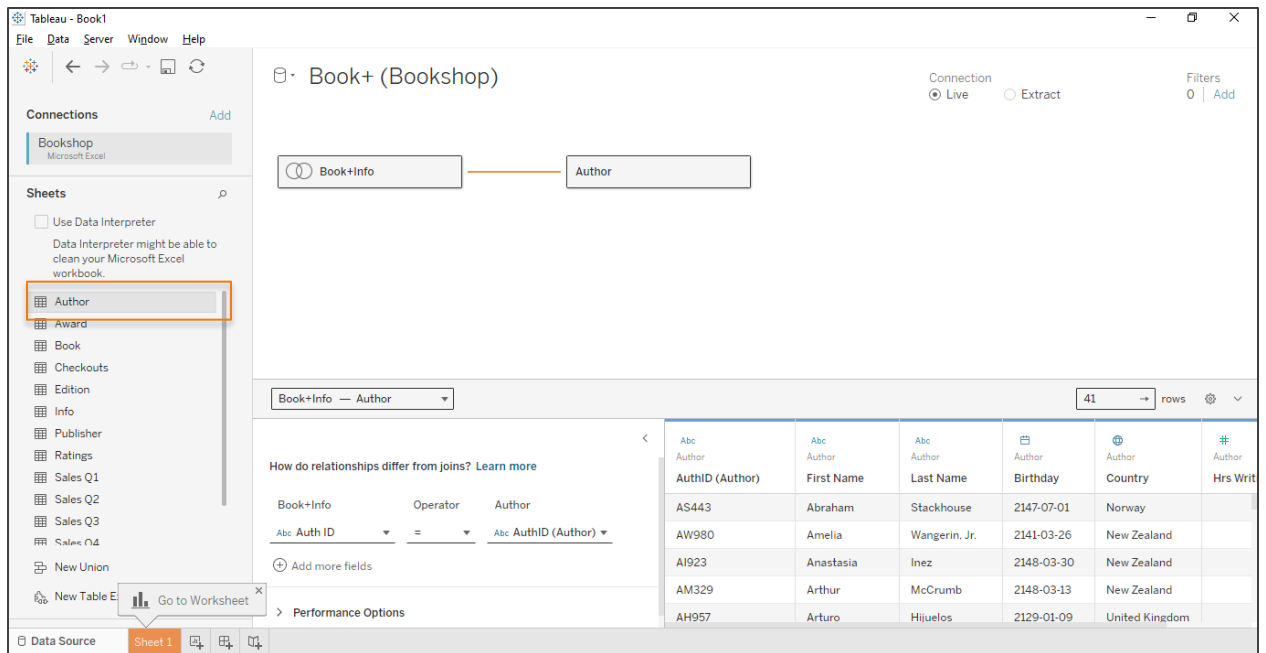
2.4 Click the **Book** Table icon and then choose **Rename** to rename this table to **Book+Info** Table



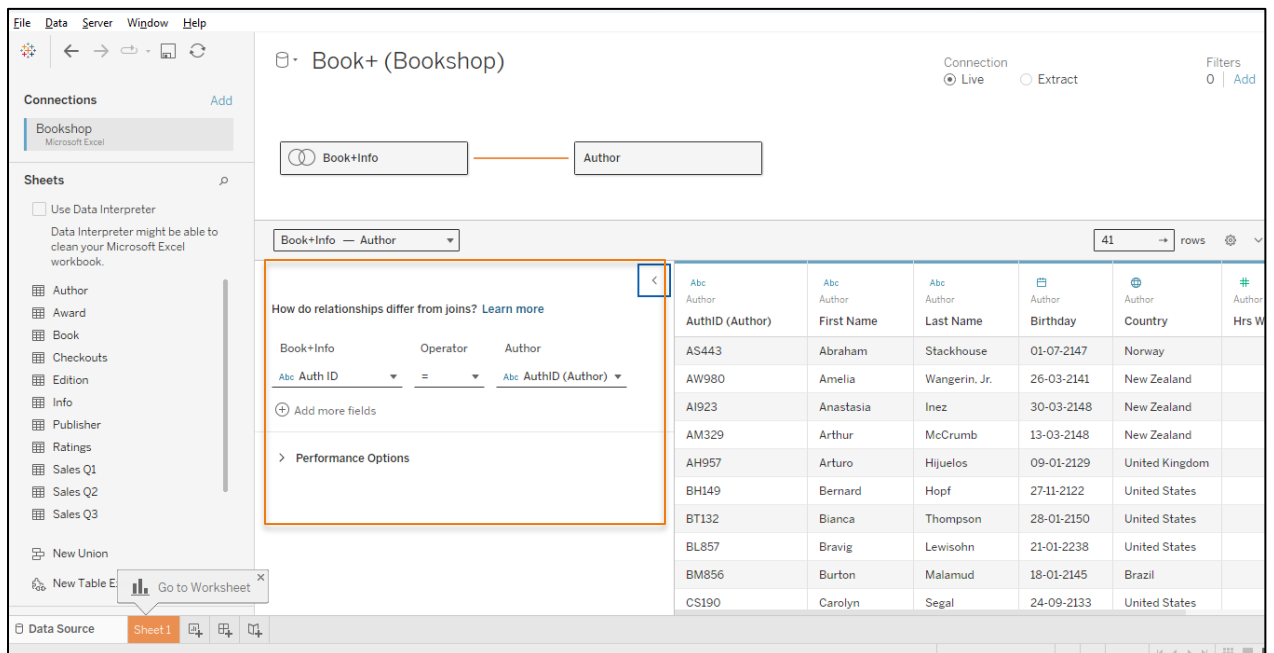
The **Book+Info Table** is shown below:



2.5 Create a relationship between the **Book+Info Table** and the **Author Table** and drag the **Author Table** into the logical layer to create the relationship between them



2.6 You can also view and edit details of the existing relationship from the Edit Relationship section (on the left side of the data view). You can also add more fields to the existing relationship.



Note: The data visible in the Data Sheet is flowing from the Table, which is selected in the Data Model Section. You can also drag the Data Sheet pane above or below to better view.

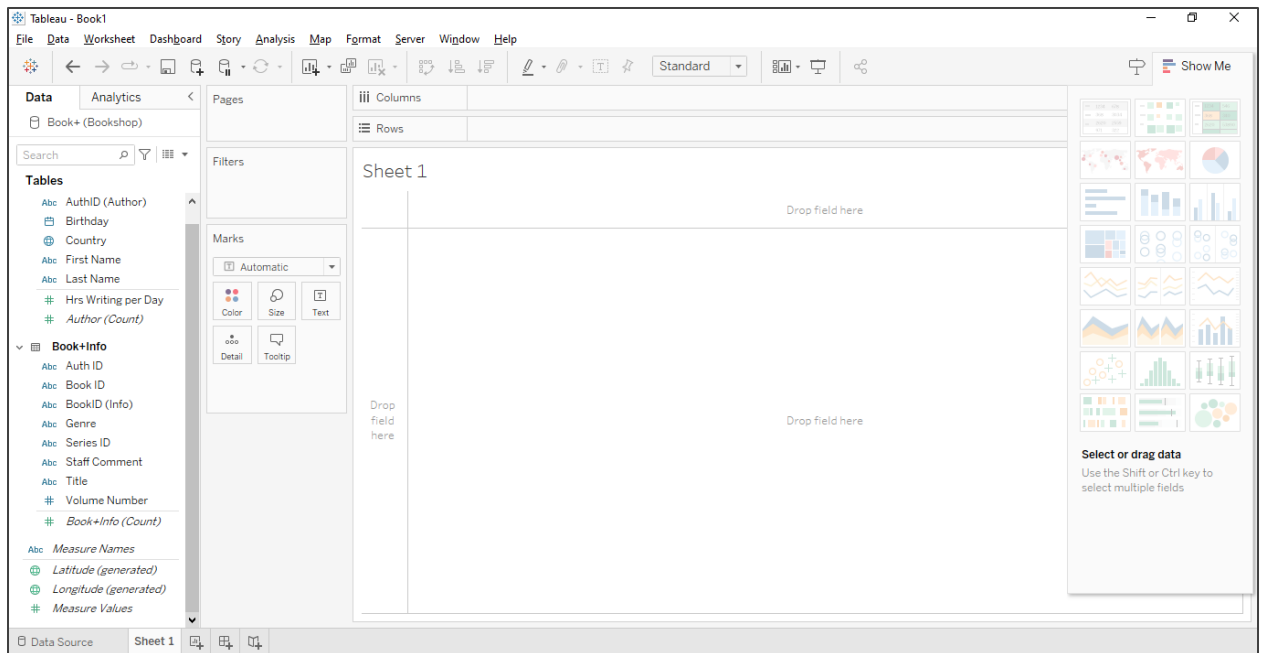
2.7 Click on the **Book+Info** Table icon. The Data Sheet will appear on-screen as shown below.

The screenshot shows the Power BI Desktop interface with the 'Book+Info' table selected in the Data Model section. The Data Sheet pane displays the table's data, including columns for Book ID, Title, Auth ID, BookID (Info), Genre, and Series ID. The interface also shows the 'Connections' pane with 'Bookshop' selected and the 'Sheets' pane with 'Sheet1' selected.

Type	Field Name	Physical Table	Remote ...
Abc	Book ID	Book	BookID
Abc	Title	Book	Title
Abc	Auth ID	Book	AuthID
Abc	BookID (Info)	Info	BookID (I...
Abc	Genre	Info	Genre
Abc	Series ID	Info	SeriesID

Book ID	Title	Auth ID	BookID (Info)	Genre	Series ID
MM424	(im)Mortality	CW626	MM424	Young Adult	
NR695	9803 North Millworks Road	CS190	NR695	Mystery	
AM124	A Horrible Human with the H...	KD840	AM124	Fiction	
AK974	Adventures of Kaya	WH901	AK974	Young Adult	
AD222	Alanna Saves the Day	BH149	AD222	Childrens	
AY135	And I Said Yes	EK605	AY135	Fiction	
BB194	Ballinby Boys	AM329	BB194	SciFi/Fantasy	
BF889	Banana Slug and the Glass H...	GG800	BF889	Childrens	
BC244	Banana Slug and the Lost Cow	HB302	BC244	Childrens	
BR858	Banana Slug and the Mossy ...	MW974	BR858	Childrens	

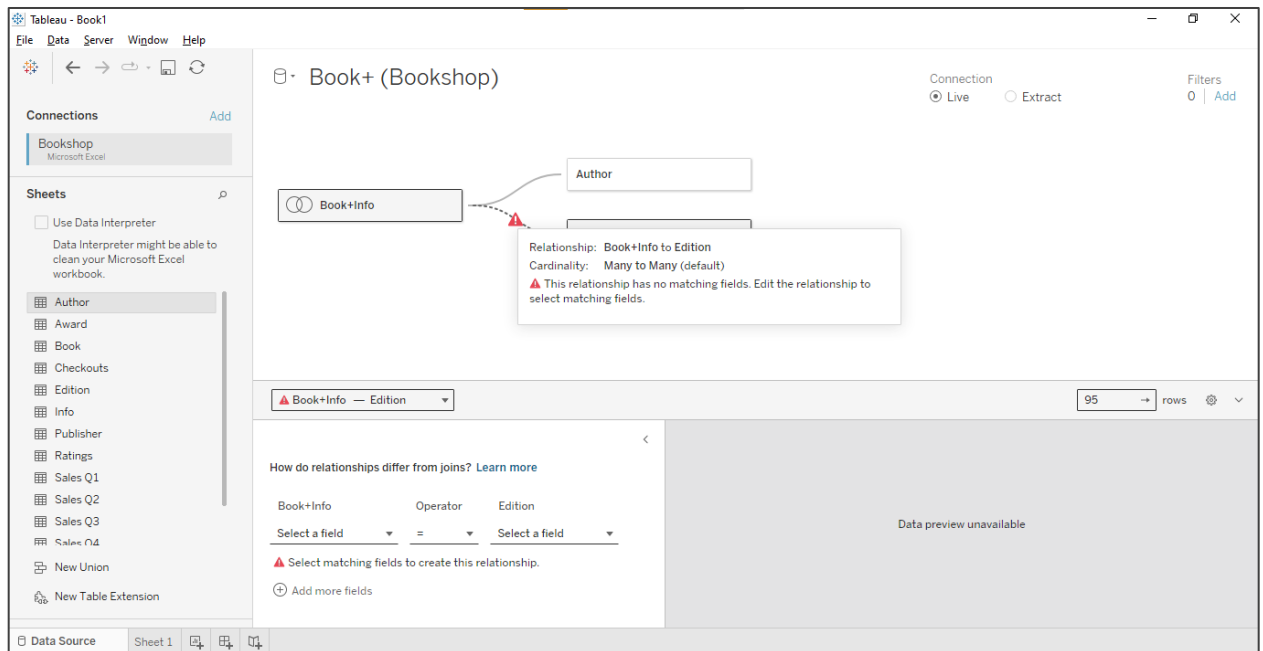
2.8 Select **Sheet1**. The Data pane displays the data from the various tables as shown below:



Note: You can see multiple tables with names listed with their fields (marked as Dimension and Measures for each respective table), and the Tableau-generated fields are listed at the bottom of the Data pane:

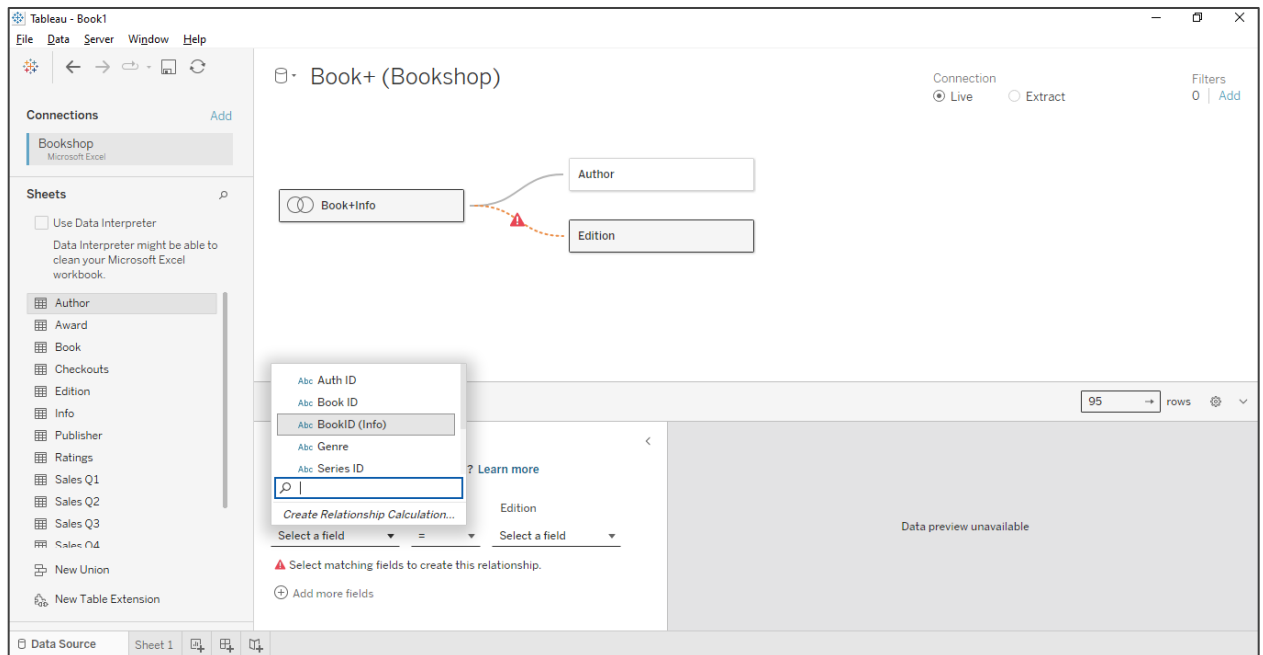
- Measure Names
- Longitude (generated)
- Latitude (generated)

2.9 Click on the Data Source and join it with the **Edition** table. A message, as shown below, will show up when you move your cursor over the broken relationship.

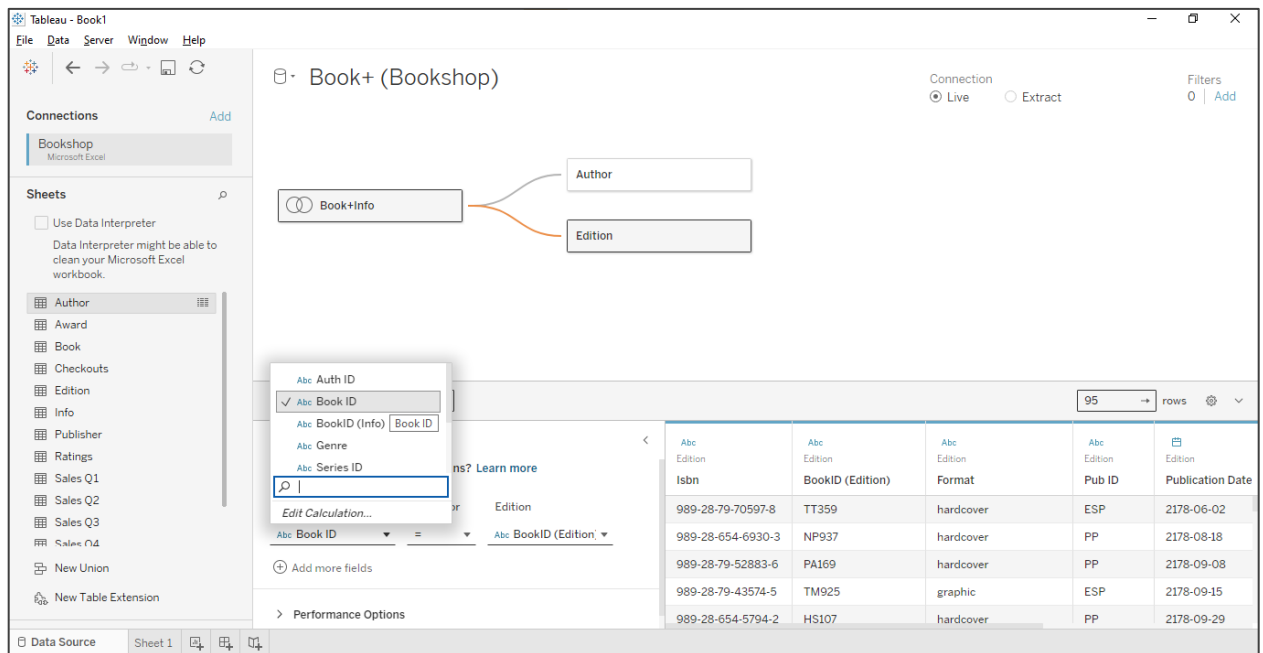


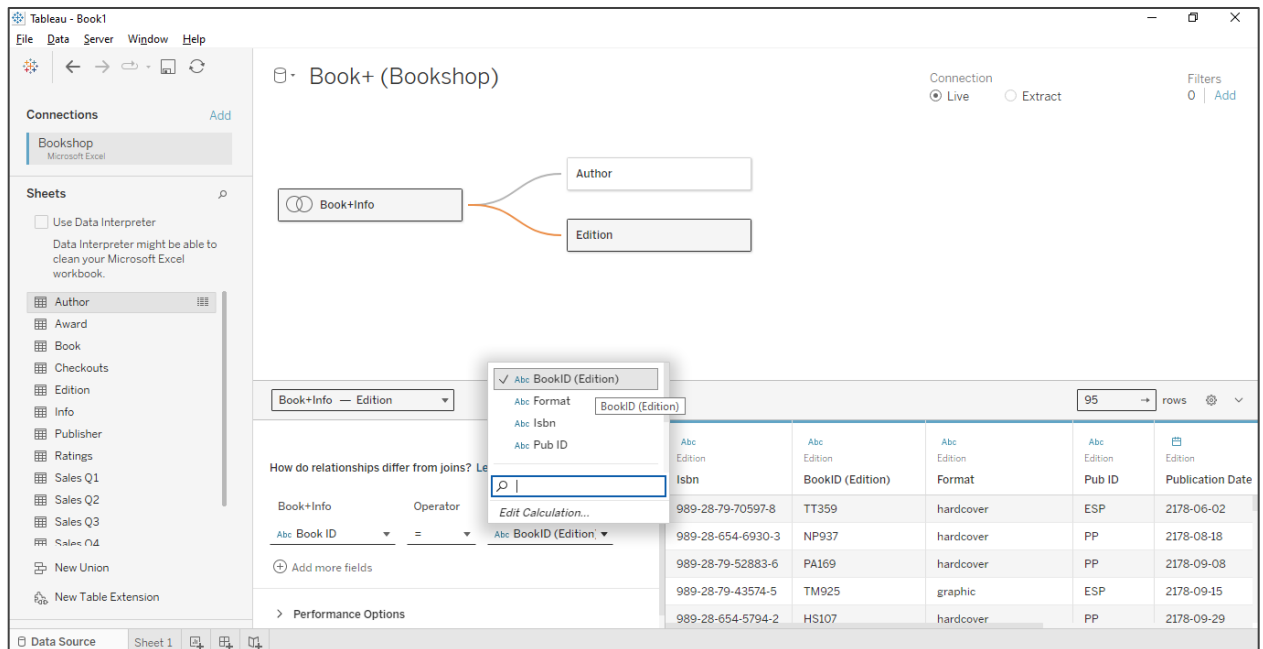
Observation: First, Tableau tries to establish a connection between the **Book and Info** table, but it fails. In the screenshot shown above, the dotted link in red indicates that the relationship could not be established. Refer to the view below:

Note: The lack of unique fields between the two tables (**Book+Info** and **Edition**) prevented the creation of this relationship. This will be confirmed if you click on the red dotted line to view the relationship details in greater detail. While the Edition table only has one book ID field, the Book+Info table has multiple book ID field names. Tableau is therefore unable to select which fields to use when building a relationship.

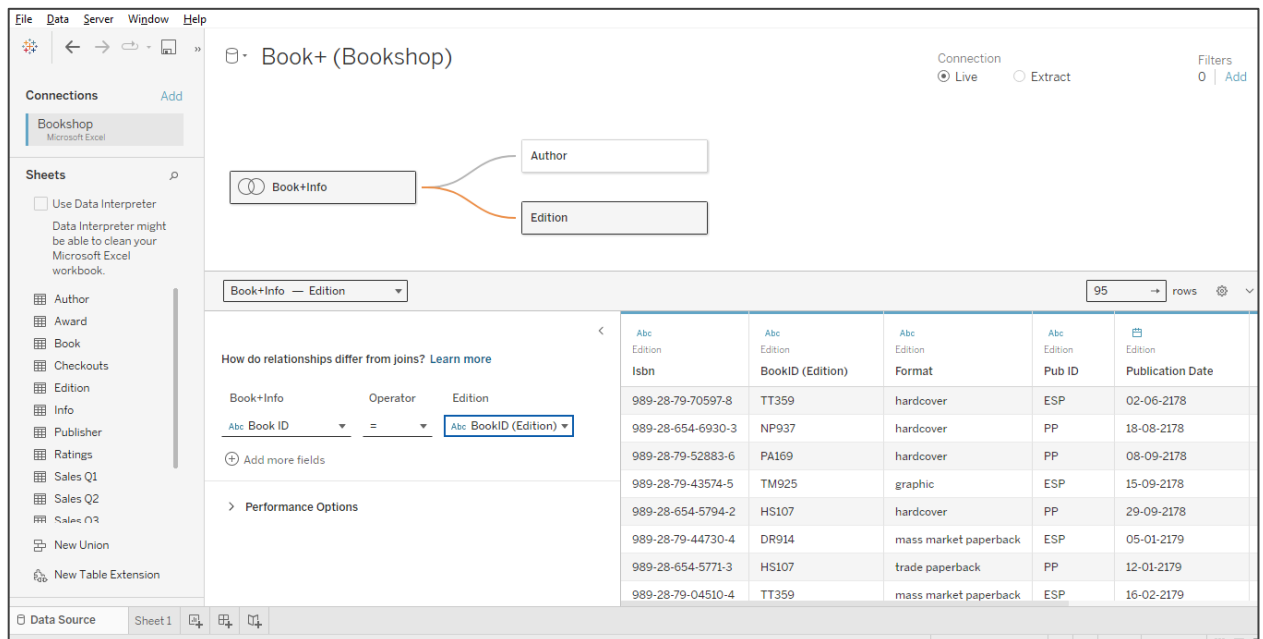


2.10 Now, create this relationship manually. Click on the dropdown under the **Book+Info** Table and select the **Book ID** coming from the **Book** table and the **Book ID** from the **Edition** table

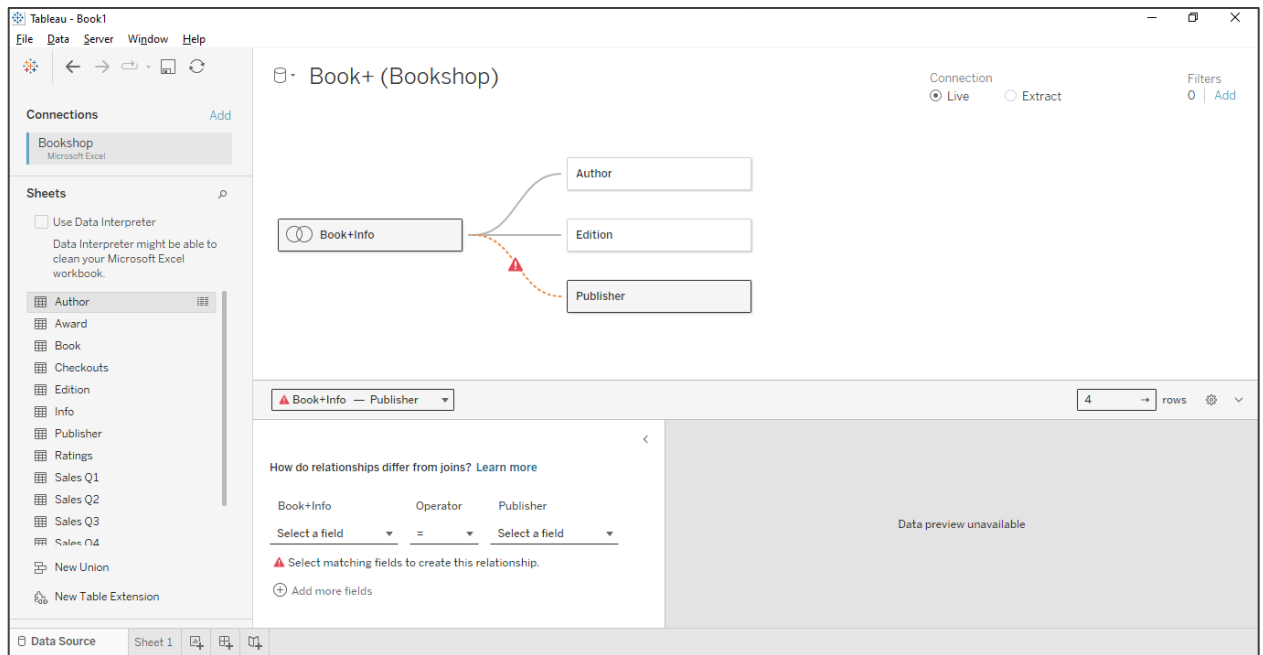




Observation: By doing this, a relationship between **Book+Info** and **Edition** is created and validated.

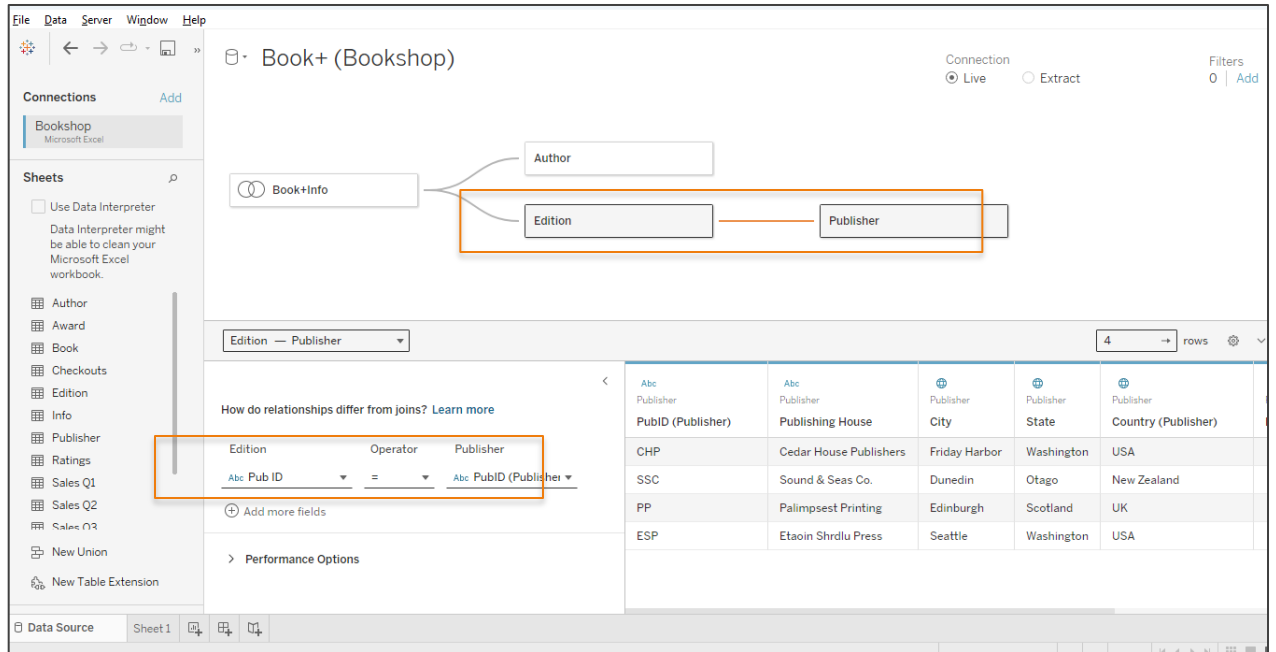


2.11 Now, let us bring the **Publisher** Table into the logical layer. It will not create a relationship between the **Book** and the **Info** Table.



Note: The reason for this is that Tableau is unable to establish a relationship between the two tables because there isn't a matching field.

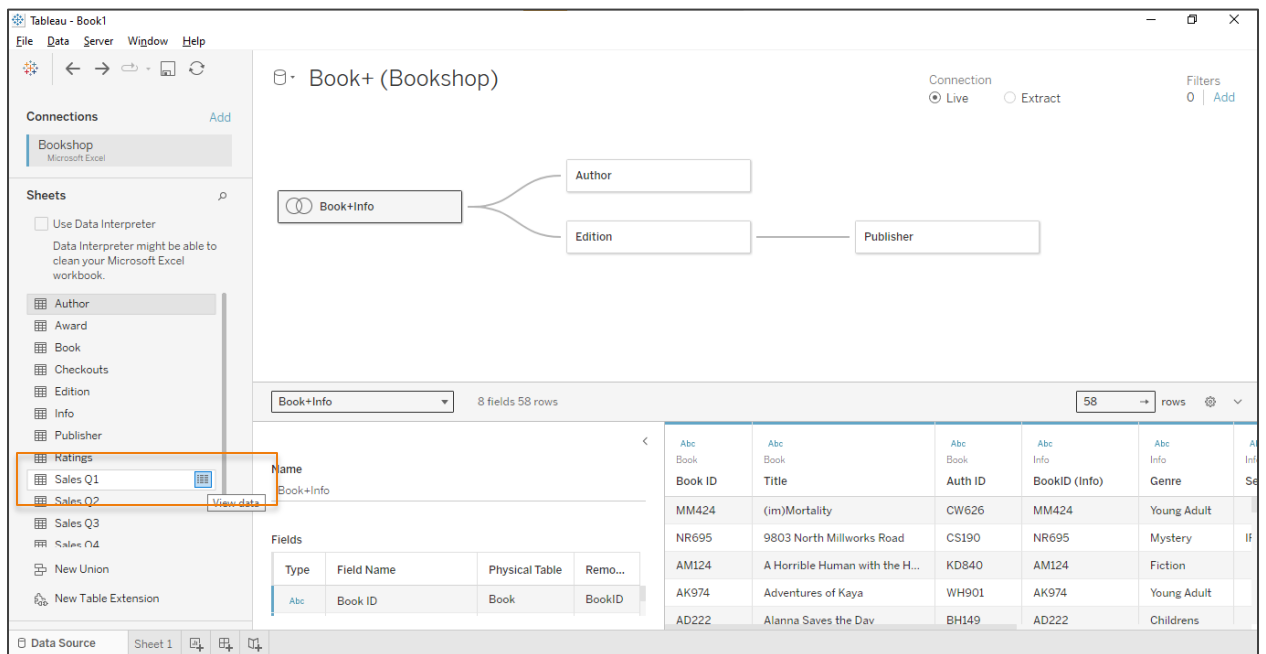
2.12 Now, drag the **Publisher** Table and bring it closer to the **Edition** Table. A relationship will be established as both tables have a common primary key, **PubID**.



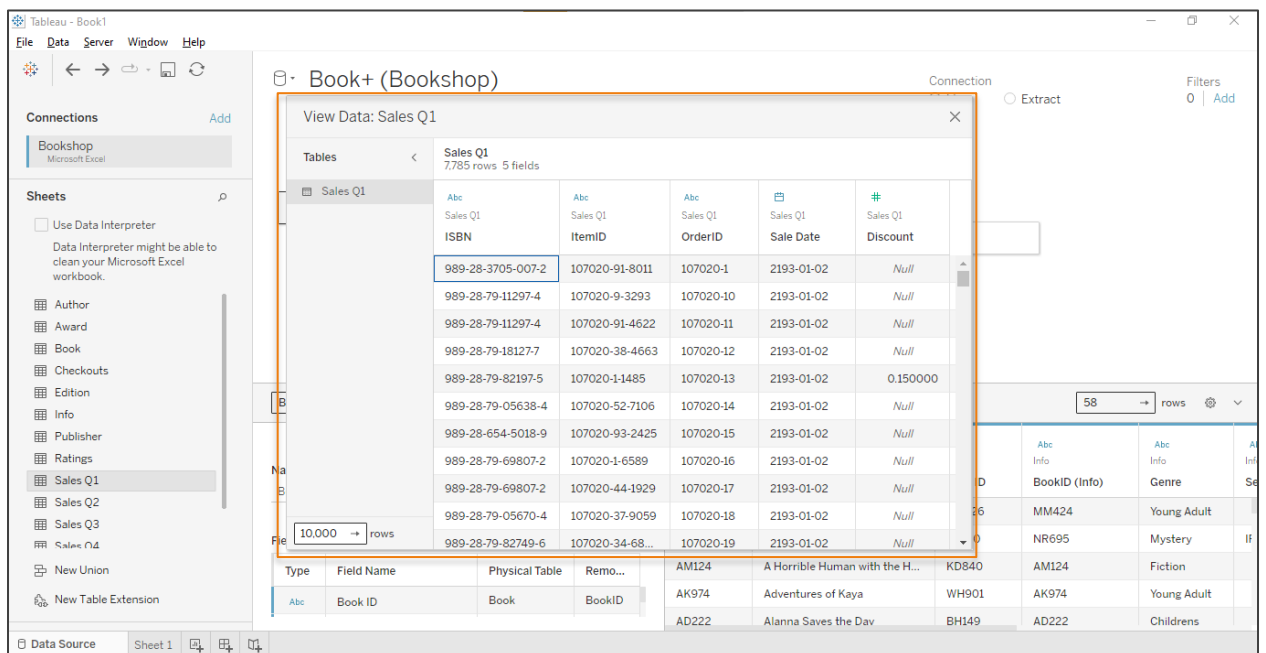
Step 3: Create a Union

Note: We have Q1, Q2, Q3, and Q4 sales tables available. These four quarters of data need to be concatenated using Union.

- 3.1 On **Sales** Tables, click on the **View Data** icon (highlighted below) and observe that it has an **ISBN** field. This is common with the edition table.

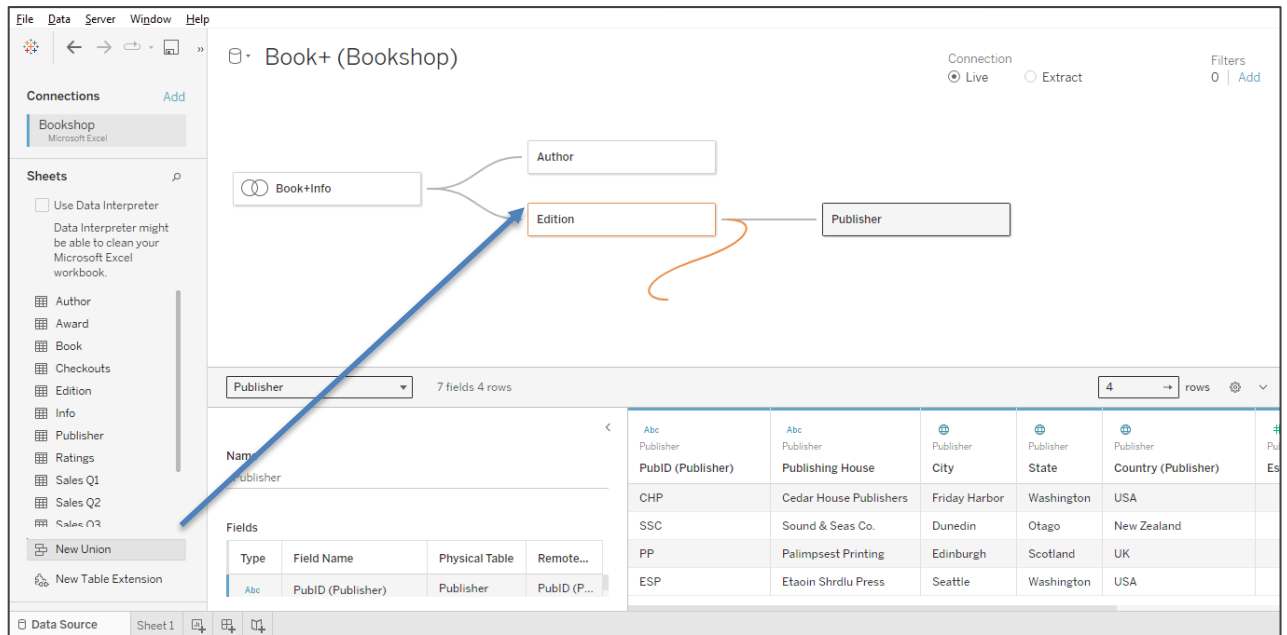


3.2 Click on the **View data** icon to open the window as shown below. You can validate that the ISBN field is present in all sales tables and that they have the same structure (same column numbers and names).

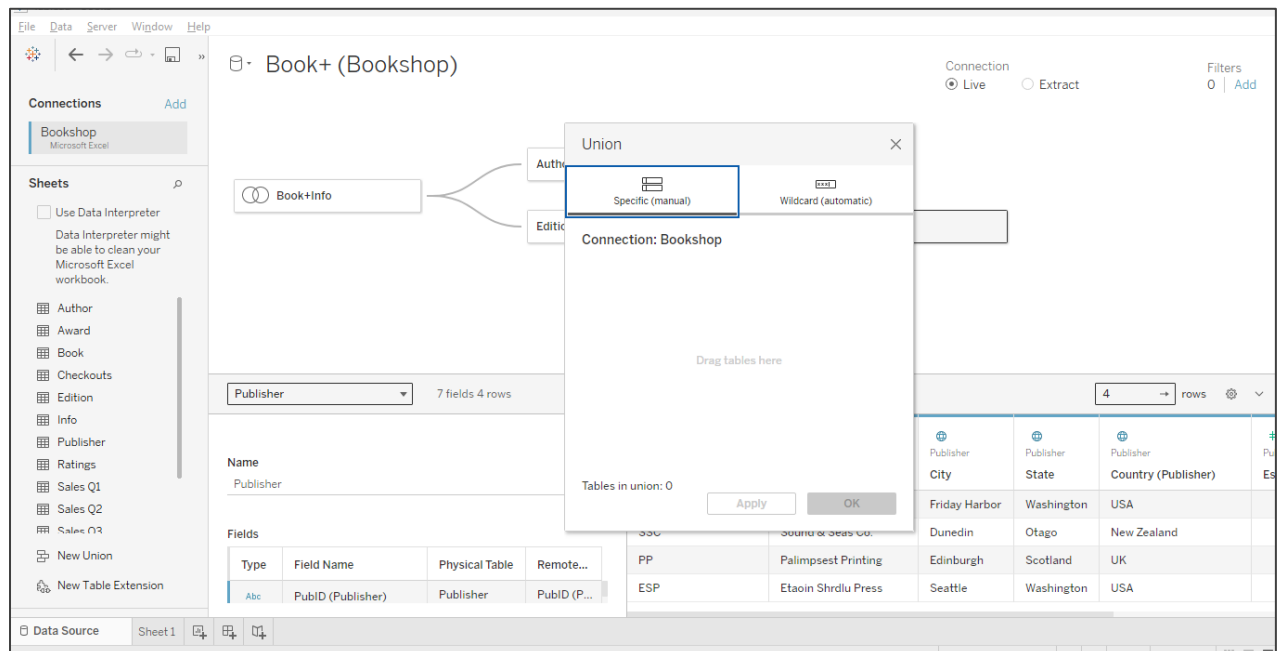


Observation: It implies that one can establish a relationship between the Edition table and the Sales table (which is unified for every quarter). Prior to that, though, all the sales tables must be united.

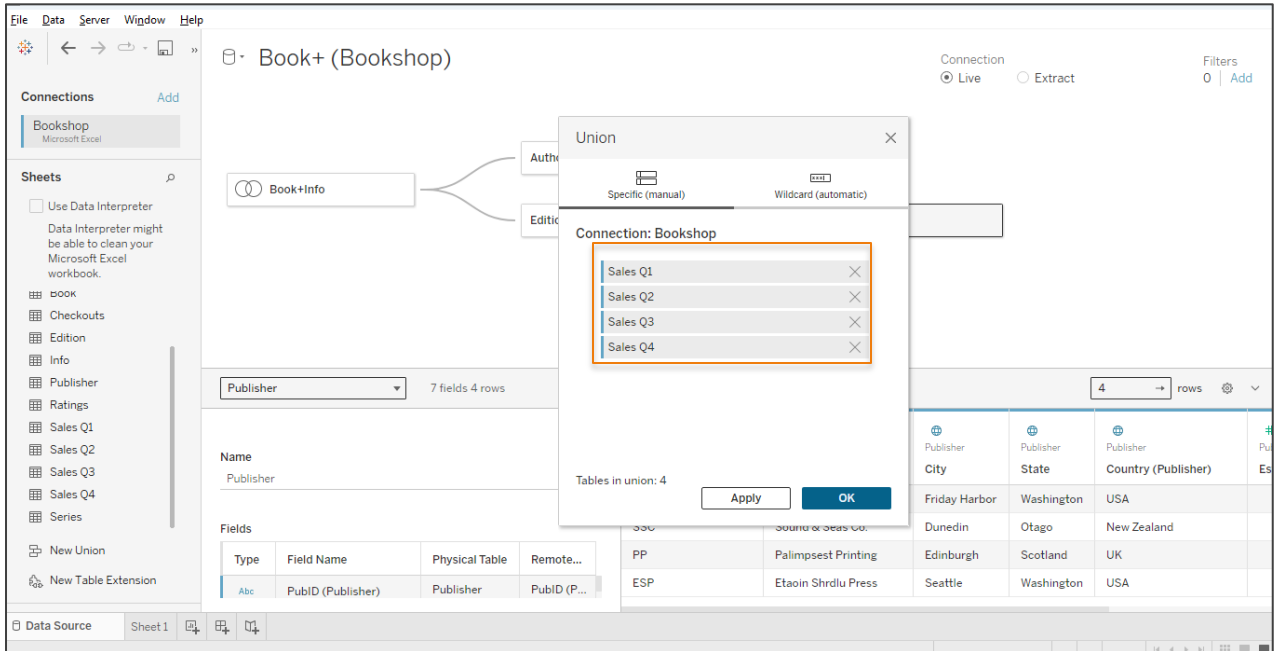
3.3 Drag a **New Union** icon (at the end of the Connections pane) closer to the **Edition** table icon and drop it



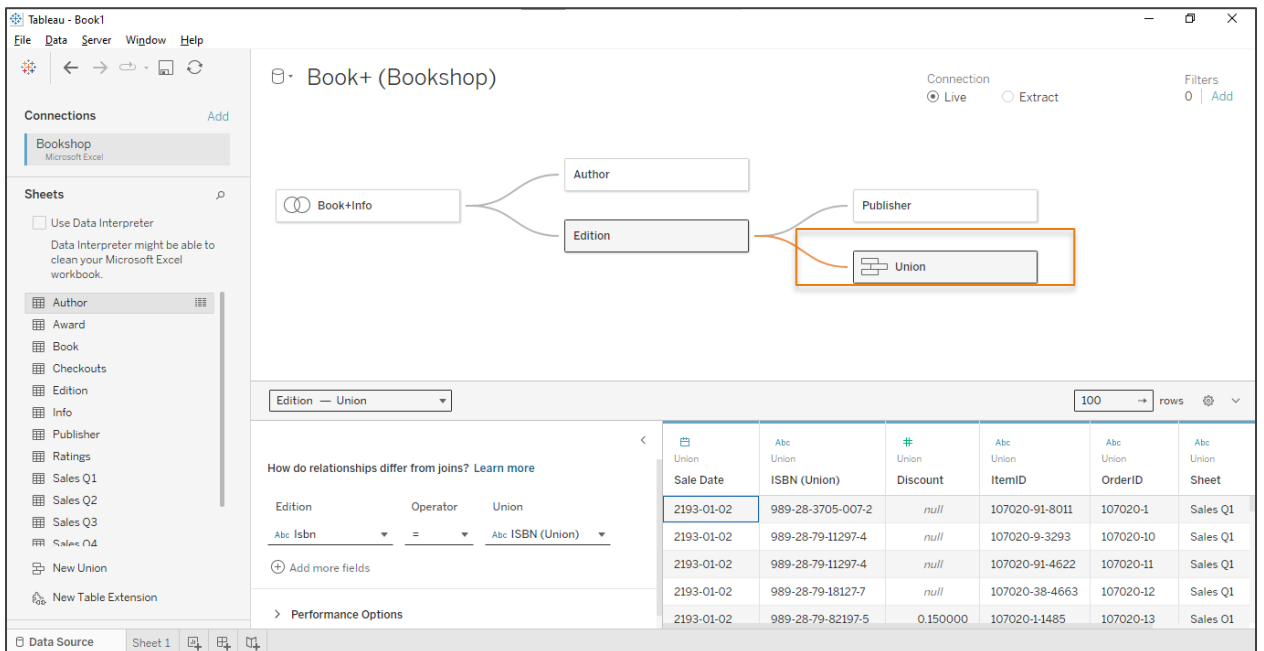
This will open the **Union** window as shown below.



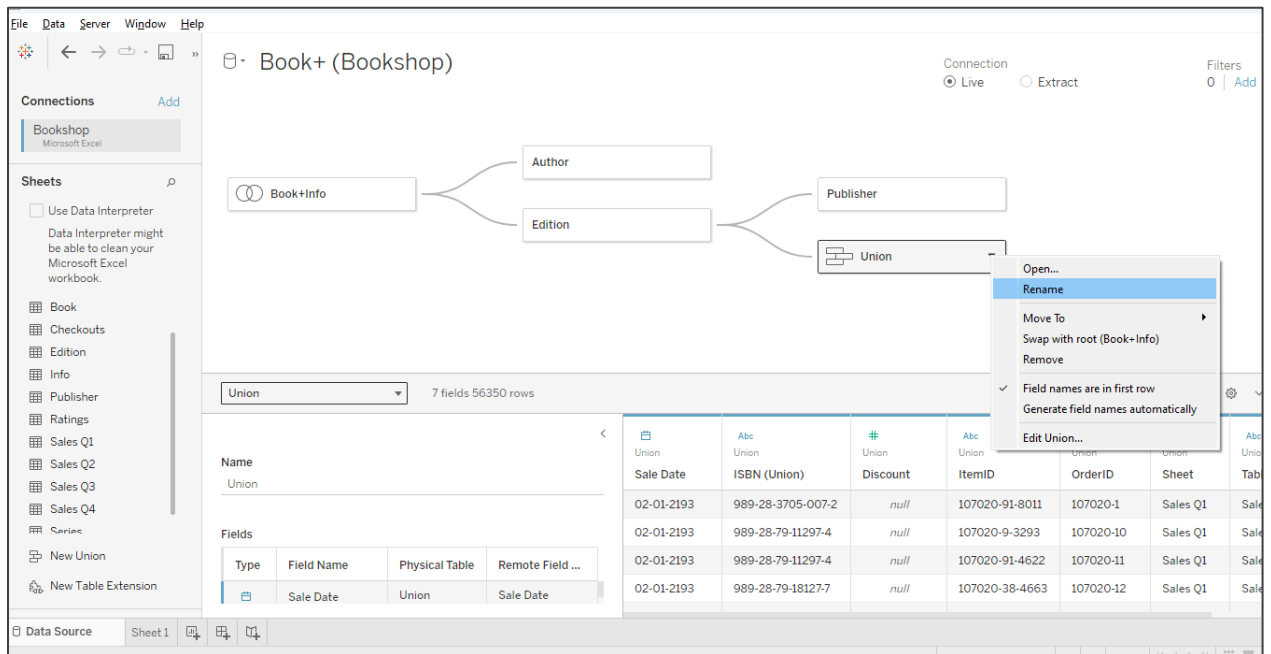
3.4 Drag the **Sales** Tables (Sales Q1, Sales Q2, Sales Q3, and Sales Q4) to create a union into this window, as shown below:



3.5 Click on **Apply** and **OK**. A Union field is created as shown below with the Union icon indicating that it contains Unioned tables:



3.6 Right-click on **Union** and rename it as **Sales**



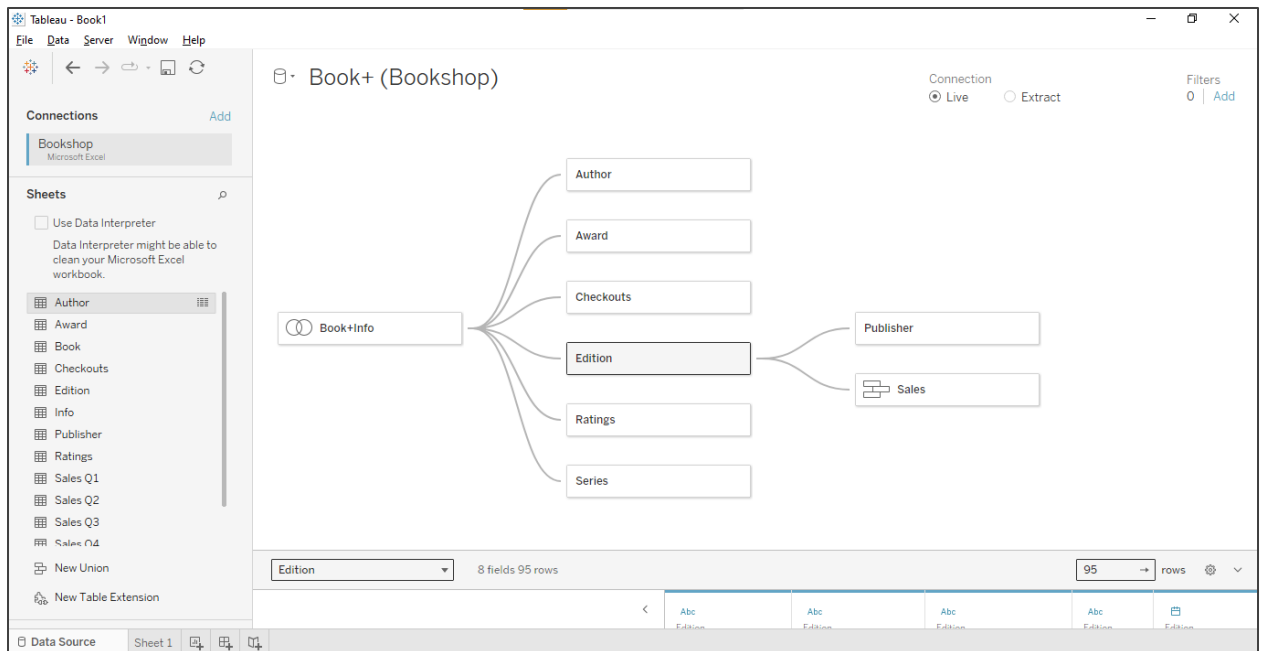
Note: Connect the below tables with the Book+Info Table

- Checkouts using Book ID
- Ratings using Book ID

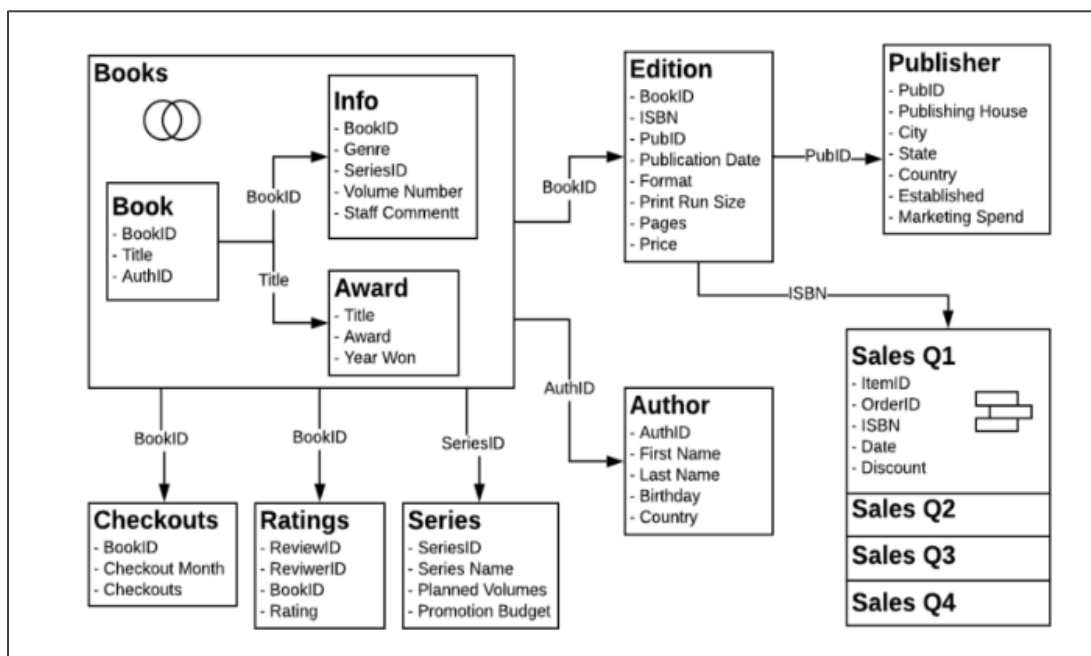
For both above tables, you need to create the relationship manually since BookID exists in duplicate in the Book+Info table.

- Series: using Series ID
- Awards: using the title

The final data model is created and looks as below.



The diagram for the data model looks as below.



With these steps, users can effectively connect and consolidate data in Tableau, enabling comprehensive analysis through joins and unions between tables.