

Data Visualization Using Tableau



Dataset Alchemy: Handling Multiple Datasets for Rich Insights



Quick Recap



- Tableau's Analytics pane boosts visualizations with live stats, trends, and predictions.
- Reference Lines and Bands mark key data points or intervals for deeper insights.
- Distribution and Box Plot tools display the range and essential statistics of the data.
- Parameters change values on the fly, enabling a more interactive and thorough data exploration.

Engage and Think



You are buying a new phone and are considering the top three best-selling models: the iPhone 15 series, Samsung Galaxy S24 series, and iPhone 13 series. To make an informed decision, you decide to gather insights like camera quality, battery life, performance, display quality, and price through research and add your findings for each phone to different tables.

How would you combine the data from each table to compare the specifications and choose the model that best fits your needs and budget?

Learning Objectives

By the end of this lesson, you will be able to:

- 🔗 Apply relationships to combine data from multiple tables to improve data analysis efficiency
- 🔗 Utilize joins and unions in Tableau to create integrated datasets for visualization and analysis
- 🔗 Combine diverse and complex datasets to uncover hidden insights by using data blending





Relationships in Tableau

Relationships in Tableau

It is a dynamic, flexible way to combine data from multiple tables for analysis.



- It defines the relationship between two tables based on common fields without merging the data.
- It is a way to create a data model using primary keys (common fields) between multiple tables.
- It is created in the logical layer of the data model section.

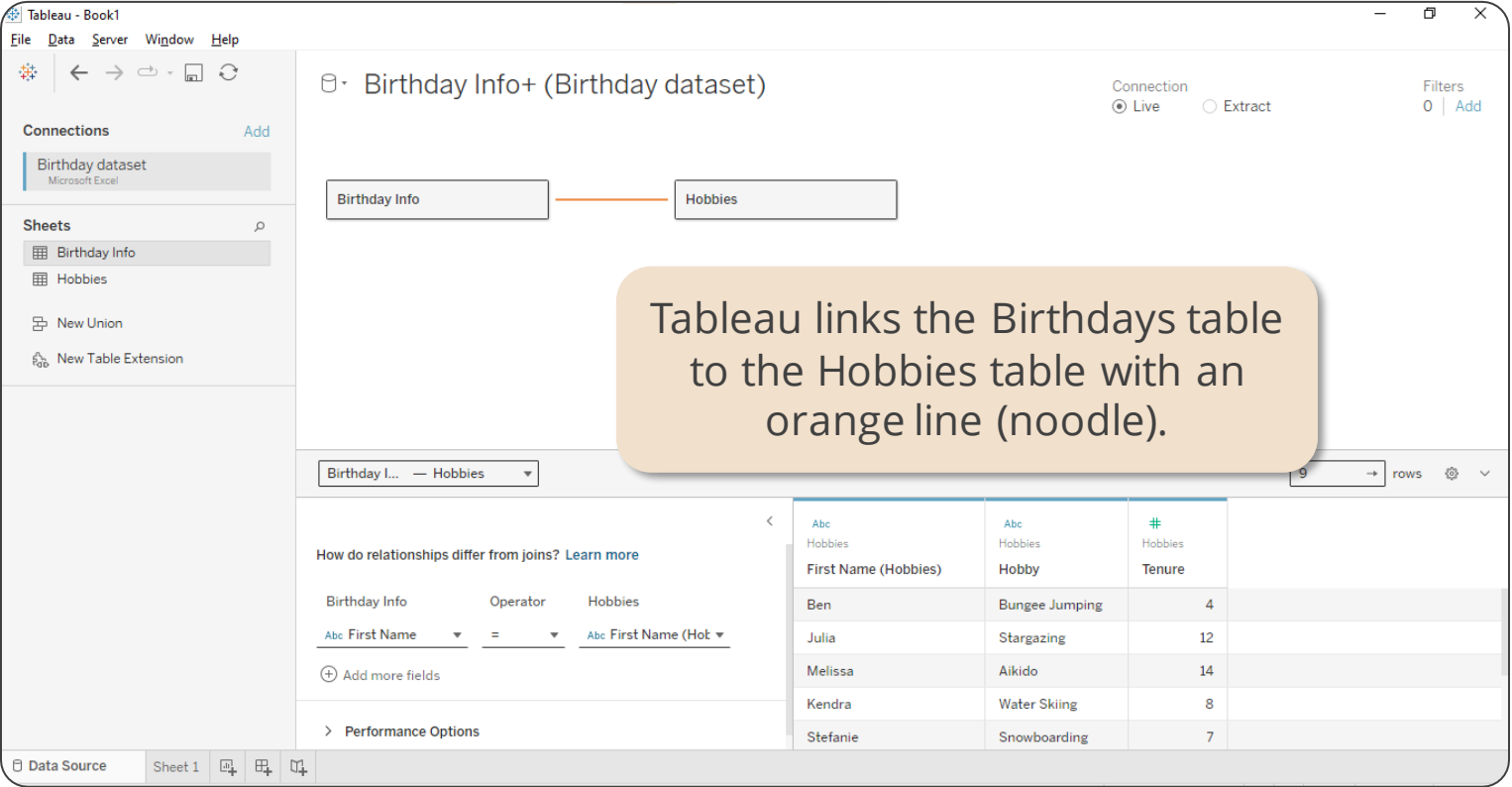
The logical layer of the data model specifies how tables are related to each other based on common fields.

Relationships: Example

Consider two tables: Birthdays and Hobbies, each containing information about colleagues. You can use this data to find birthday present ideas based on their hobbies.

Birthday info			
First Name	Birhday Mont	Date	Age
Ben	January	5	3
Julia	January	10	3
Melissa	March	7	2
Kendra	August	4	2
Stefanie	July	24	3
Mark	September	22	4
Chris	October	8	2
Sarah	October	26	3
Alonso	December	12	3

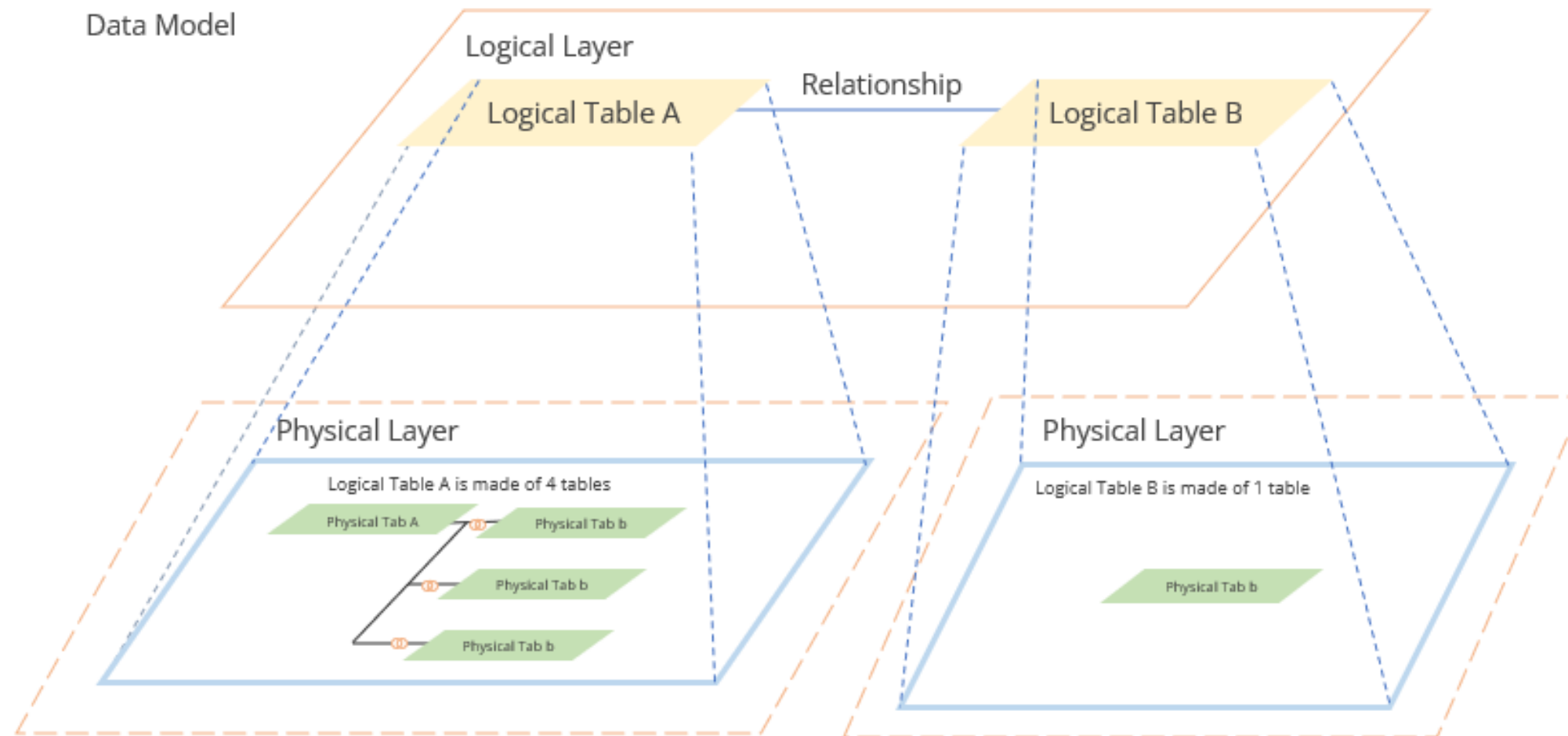
Hobbies		
First Name	Hobby	Tenure
Ben	Bungee Jumpir	4
Julia	Stargazing	12
Melissa	Aikido	14
Kendra	Water Skiing	8
Stefanie	Snowboarding	7
Mark	Painting	21
Chris	Running Marat	9
Sarah	Wood Carving	15
Alonso	Cooking	2



Clicking on the noodle reveals that Tableau has correctly inferred the relationship between the tables based on their shared field: [First Name].

Data Model Layers

The Logical Layer refers to how data is logically organized and related within Tableau, while the Physical Layer refers to the actual data sources and how data is stored physically.



Data Model Layers

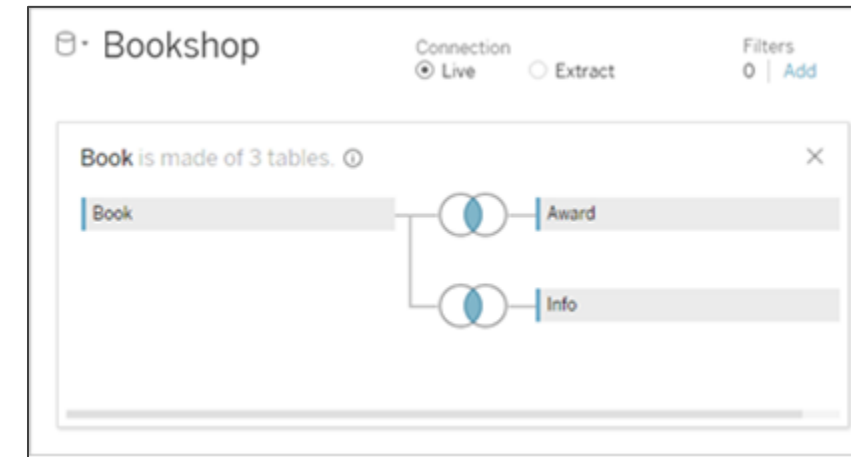
When navigating through the data source in Tableau, it is essential to understand the distinction between the logical and physical layers. For example:

Logical layer view



The primary view of a multi-table data source is the logical layer, where tables are interconnected and combined using relationships (noodles).

Physical layer view



Double-clicking a logical table reveals its physical tables, which can be joined or unioned. For instance, the Book logical table consists of three joined physical tables: Book, Award, and Info.

Logical vs. Physical Layers

Understanding the disparities between the logical and physical layers in Tableau is crucial for optimizing data modeling and analysis:

Logical layer	Physical layer
Created in the Relationships canvas (default view) on the data source page	Created in the Join or Union canvas in the data source (which opens by a double click on any icon for a table)
Creates logical relationships	Creates joins between different tables
Does not create a new entity, it only establishes a relationship between them	Creates a new entity for the table as it changes the structure of the table, like adding more columns to the table
Called the logical layer as it only establishes a logical relation between the tables using the primary keys	Called the physical layer as it changes the physical entity of the table, like the output tables created after joining, which are new tables and not the same as the input tables

Demo: Relationship in Tableau



Duration: 10 minutes

Demonstrate the process of establishing relationships in the provided bookstore dataset between the book and info tables.

DEMONSTRATION

Quick Check

What does a relationship between two tables in a database describe?

- A. It establishes a hierarchy between the tables.
- B. It describes how the tables are connected based on common fields.
- C. It merges the tables into a single table.
- D. It separates the tables into distinct entities.





Joins and Union

Joins

They combine data from two or more tables that have a relationship based on a common field. This allows you to create a single, unified dataset for analysis.

Tableau - Book1

File Data Server Window Help

Connections: Bookshop (Microsoft Excel)

Sheets: Use Data Interpreter (Data Interpreter might be able to clean your Microsoft Excel workbook.)

Author, Award, Book, Checkouts, Edition, Info, Publisher, Ratings, Sales Q1, Sales Q2, Sales Q3, Sales Q4, New Union, New Table E

Book+ (Bookshop)

Connection: Live (selected), Extract

Filters: 0 | Add

Book is made of 2 tables. ⓘ

Book

Join dialog box:

- Inner, Left (selected), Right, Full Outer
- Data Source: Book
- Info
- Book ID = BookID (Info)
- Add new join clause

Book: 8 fields 58 rows

58 rows

Book ID	Title	Auth ID	BookID (Info)	Genre	Series ID
MM424	(im)Mortality	CW626	MM424	Young Adult	null
NR695	9803 North Millwor...	CS190	NR695	Mystery	IR-ESP
AM124	A Horrible Human ...	KD840	AM124	Fiction	null
AK974	Adventures of Kaya	WH901	AK974	Young Adult	null
AD222	Alanna Saves the Day	BH149	AD222	Childrens	null

Fields:

Type	Field Name	Physical Table	Remote Field Name
Abc	Book ID	Book	BookID

Go to Worksheet

Data Source Sheet 1

One must be familiar with primary and secondary keys to understand joins.

Primary vs. Secondary Key

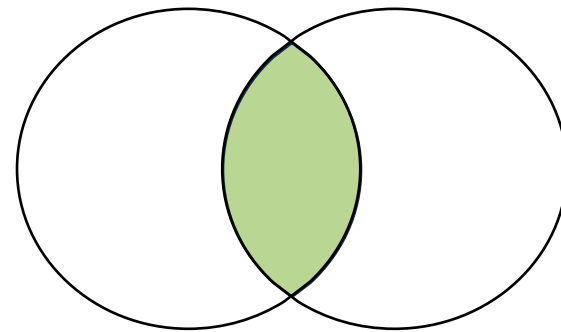
Joins can be performed using primary keys and secondary keys to combine data from multiple tables.

Primary key	Secondary key
A primary key identifies and accesses individual rows in a table.	Secondary keys, also known as foreign keys, establish relationships between tables, enforcing referential integrity.
It uniquely identifies a record. A primary key column has all unique values, which act as an identifier for a record (row) for a dimension.	It is the field in the table that is the primary key for another table.
A table can contain one and only one primary key.	A table can contain more than one foreign keys.
No two rows can carry duplicate values for a primary key.	A foreign key can contain duplicate values.

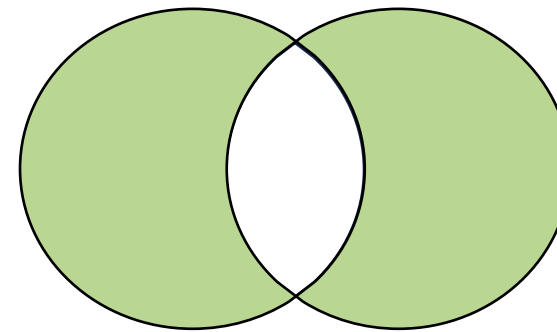
Types of Joins

There are four types of joins in Tableau.

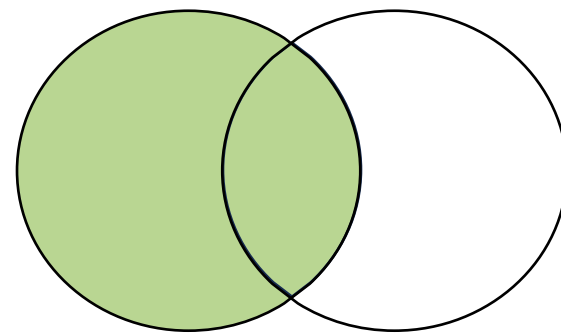
Inner join



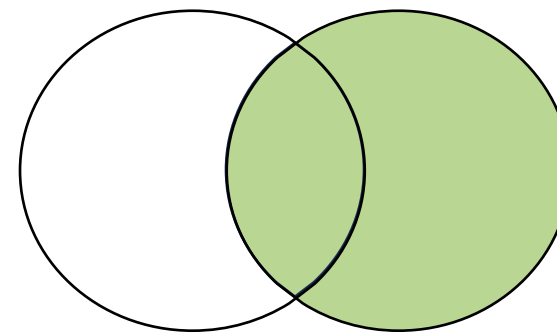
Full outer join



Left join



Right join



Types of Joins: Example

Consider two key tables in your Tableau dashboard: Product sales and product profits. Identify how Tableau joins can help you get insights from the given data and analyze your products' performance.

Table: Product sales

Product ID	Quantity	Sales
6	2	678
7	14	4,324
8	6	564
9	25	3,476
10	10	986
11	8	453
12	7	834
13	13	923
14	32	5,454
15	15	623

Table: Product profits

Product ID	No Of Transactions	Profits	Profits/ Transactions
1	2	230	115
2	4	1,500	375
3	2	321	161
4	10	945	95
5	5	345	69
6	7	154	22
7	5	200	40
8	8	221	28
9	6	1,890	315
10	10	200	20

Primary key: Product ID

Inner Join

It returns rows that have matching values in both tables.

Table: Product sales

Product ID	Quantity	Sales
6	2	678
7	14	4324
8	6	564
9	25	3476
10	10	986
11	8	453
12	7	834
13	13	923
14	32	5454
15	15	623

Table: Product profits

Product ID	No Of Transactions	Profits	Profits/ Transactions
1	2	230	115.00
2	4	1500	375.00
3	2	321	160.50
4	10	945	94.50
5	5	345	69.00
6	7	154	22.00
7	5	200	40.00
8	8	221	27.63
9	6	1890	315.00
10	10	200	20.00

Result:

Product ID	Quantity	Sales	No Of Transactions	Profits	Profits/ Transactions
6	2	678	7	154	22.00
7	14	4324	5	200	40.00
8	6	564	8	221	27.63
9	25	3476	6	1890	315.00
10	10	986	10	200	20.00

The resulting table includes the data corresponding to the primary key, which means that only common products from both tables are included.

Left Join

It returns all rows from the left table and the matched rows from the right table.

Table: Product sales

Product ID	Quantity	Sales
6	2	678
7	14	4324
8	6	564
9	25	3476
10	10	986
11	8	453
12	7	834
13	13	923
14	32	5454
15	15	623

Table: Product profits

Product ID	No Of Transactions	Profits	Profits/ Transations
1	2	230	115.00
2	4	1500	375.00
3	2	321	160.50
4	10	945	94.50
5	5	345	69.00
6	7	154	22.00
7	5	200	40.00
8	8	221	27.63
9	6	1890	315.00
10	10	200	20.00

Result:

Product ID	Quantity	Sales	No Of Transactions	Profits	Profits/ Transations
6	2	678	7	154	22.00
7	14	4324	5	200	40.00
8	6	564	8	221	27.63
9	25	3476	6	1890	315.00
10	10	986	10	200	20.00
11	8	453	N/A	N/A	N/A
12	7	834	N/A	N/A	N/A
13	13	923	N/A	N/A	N/A
14	32	5454	N/A	N/A	N/A
15	15	623	N/A	N/A	N/A

The resulting table contains all the values from the product sales and the corresponding matches from the product profits.

Right Join

It returns all rows from the right table and the matched rows from the left table.

Table: Product sales

Product ID	Quantity	Sales
6	2	678
7	14	4324
8	6	564
9	25	3476
10	10	986
11	8	453
12	7	834
13	13	923
14	32	5454
15	15	623

Table: Product profits

Product ID	No Of Transactions	Profits	Profits/ Transactions
1	2	230	115.00
2	4	1500	375.00
3	2	321	160.50
4	10	945	94.50
5	5	345	69.00
6	7	154	22.00
7	5	200	40.00
8	8	221	27.63
9	6	1890	315.00
10	10	200	20.00

Result:

Product ID	Quantity	Sales	No Of Transactions	Profits	Profits/ Transactions
1	N/A	N/A	2	230	115.00
2	N/A	N/A	4	1500	375.00
3	N/A	N/A	2	321	160.50
4	N/A	N/A	10	945	94.50
5	N/A	N/A	5	345	69.00
6	2	678	7	154	22.00
7	14	4324	5	200	40.00
8	6	564	8	221	27.63
9	25	3476	6	1890	315.00
10	10	986	10	200	20.00

The resulting table contains all values from the product profits and corresponding matches from product sales.

Full Outer Join

It includes all values from both tables; any unmatched values from either table will result in null values.

Table: Product Sales

Product ID	Quantity	Sales
6	2	678
7	14	4324
8	6	564
9	25	3476
10	10	986
11	8	453
12	7	834
13	13	923
14	32	5454
15	15	623

Table: Product Profits

Product ID	No Of Transactions	Profits	Profits/ Transactions
1	2	230	115.00
2	4	1500	375.00
3	2	321	160.50
4	10	945	94.50
5	5	345	69.00
6	7	154	22.00
7	5	200	40.00
8	8	221	27.63
9	6	1890	315.00
10	10	200	20.00

Result:

Product ID	Quantity	Sales	No Of Transactions	Profits	Profits/ Transactions
1	N/A	N/A	2	230	115.00
2	N/A	N/A	4	1500	375.00
3	N/A	N/A	2	321	160.50
4	N/A	N/A	10	945	94.50
5	N/A	N/A	5	345	69.00
6	2	678	7	154	22.00
7	14	4324	5	200	40.00
8	6	564	8	221	27.63
9	25	3476	6	1890	315.00
10	10	986	10	200	20.00
11	8	453	N/A	N/A	N/A
12	7	834	N/A	N/A	N/A
13	13	923	N/A	N/A	N/A
14	32	5454	N/A	N/A	N/A
15	15	623	N/A	N/A	N/A

The resulting table will combine all the rows from both tables, regardless of matching IDs.

Joins vs. Relationship

After learning about joins and relationships, it is important to understand the differences between them.

Relationships	Joins
They are defined between logical tables in the logical layer of the data model section.	They are defined between the physical tables in the physical layer.
They do not require defining a join type.	They require defining the join type.
They act like a container for tables that are joined or unioned.	They are merged into their logical table.
Rows are not duplicated.	A joined table can have duplicated data.

Quick Check

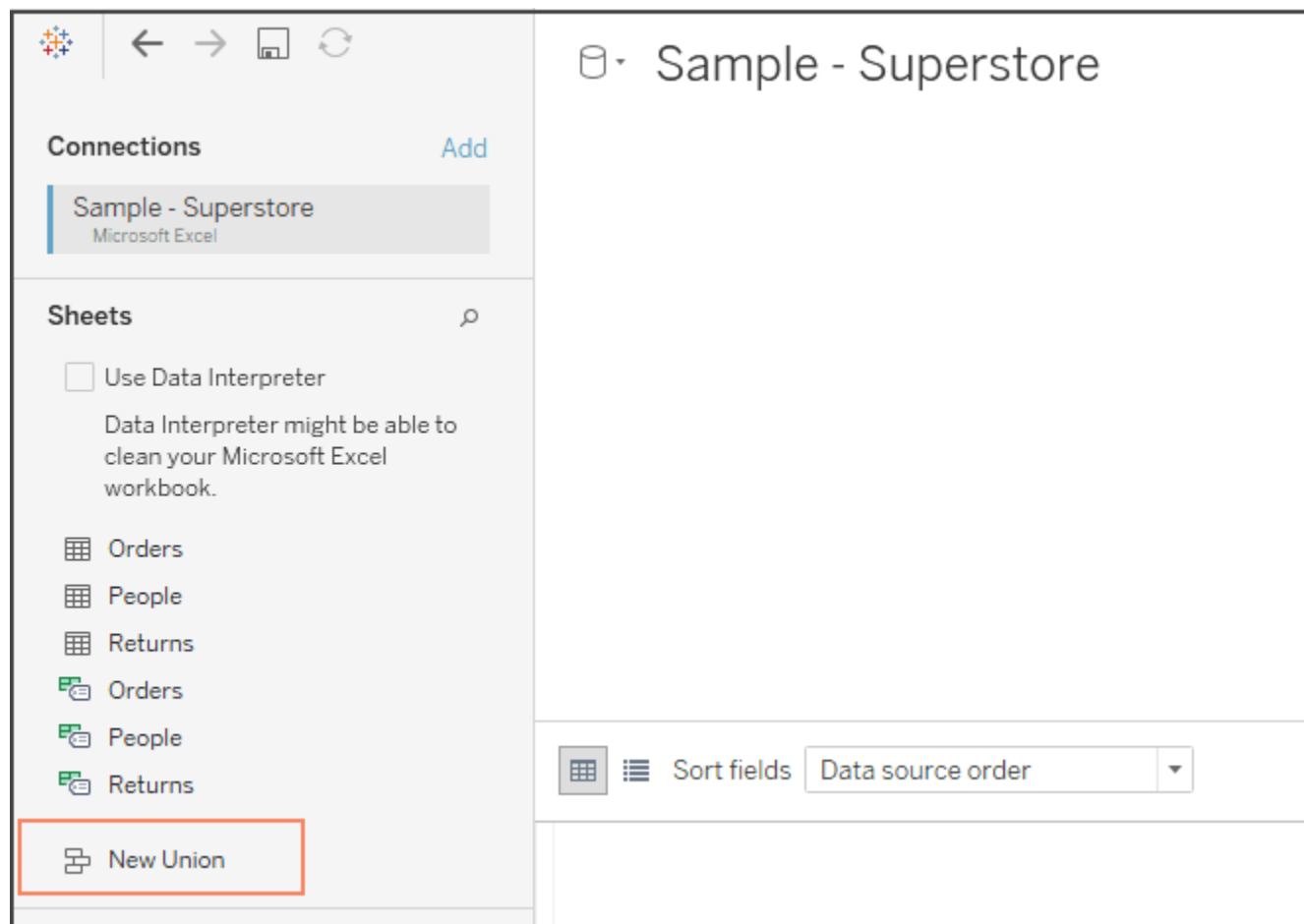


What does the result table include when performing the inner join in Tableau?

- A. Only the data from the primary table
- B. Only the data from the secondary table
- C. Data corresponding to the primary key, which is common to both tables
- D. Data that is unique to each table, excluding the primary key

Union

It is a method for combining data by appending rows of one table onto another table.



Criteria for Union

The **data structure** should be the same, which means:

- The **number of columns** should be the same.
- The **names of the fields** should be the same.
- The **data types** should be the same.

Union: Example

Consider a scenario where a retail business with three locations wants to analyze overall sales performance for a specific period across all locations.

Table -1

Date	Customer	Items	Type
04-May-23	Lane	5	Credit
10-May-23	Chris	6	Credit
28-May-23	Juan	1	Credit

Table -2

Date	Customer	Items	Type
01-Jun-23	Lisa	3	Credit
13-Jun-23	Isaac	4	Cash
22-Jun-23	Sam	2	Credit

Table -3

Date	Customer	Items	Type
11-Jul-23	Mario	2	Credit
14-Jul-23	Wei	1	Cash
26-Jul-23	Jim	7	Cash

Result Table

Date	Customer	Items	Type
04-May-23	Lane	5	Credit
10-May-23	Chris	6	Credit
28-May-23	Juan	1	Credit
01-Jun-23	Lisa	3	Credit
13-Jun-23	Isaac	4	Cash
22-Jun-23	Sam	2	Credit
11-Jul-23	Mario	2	Credit
14-Jul-23	Wei	1	Cash
26-Jul-23	Jim	7	Cash

Demo: Joins and Union



Duration: 20 minutes

Demonstrate the process of joins and unions in the provided bookstore dataset to enrich and consolidate the data for comprehensive analysis.

DEMONSTRATION

Quick Check



What are the criteria for performing a union operation in Tableau?

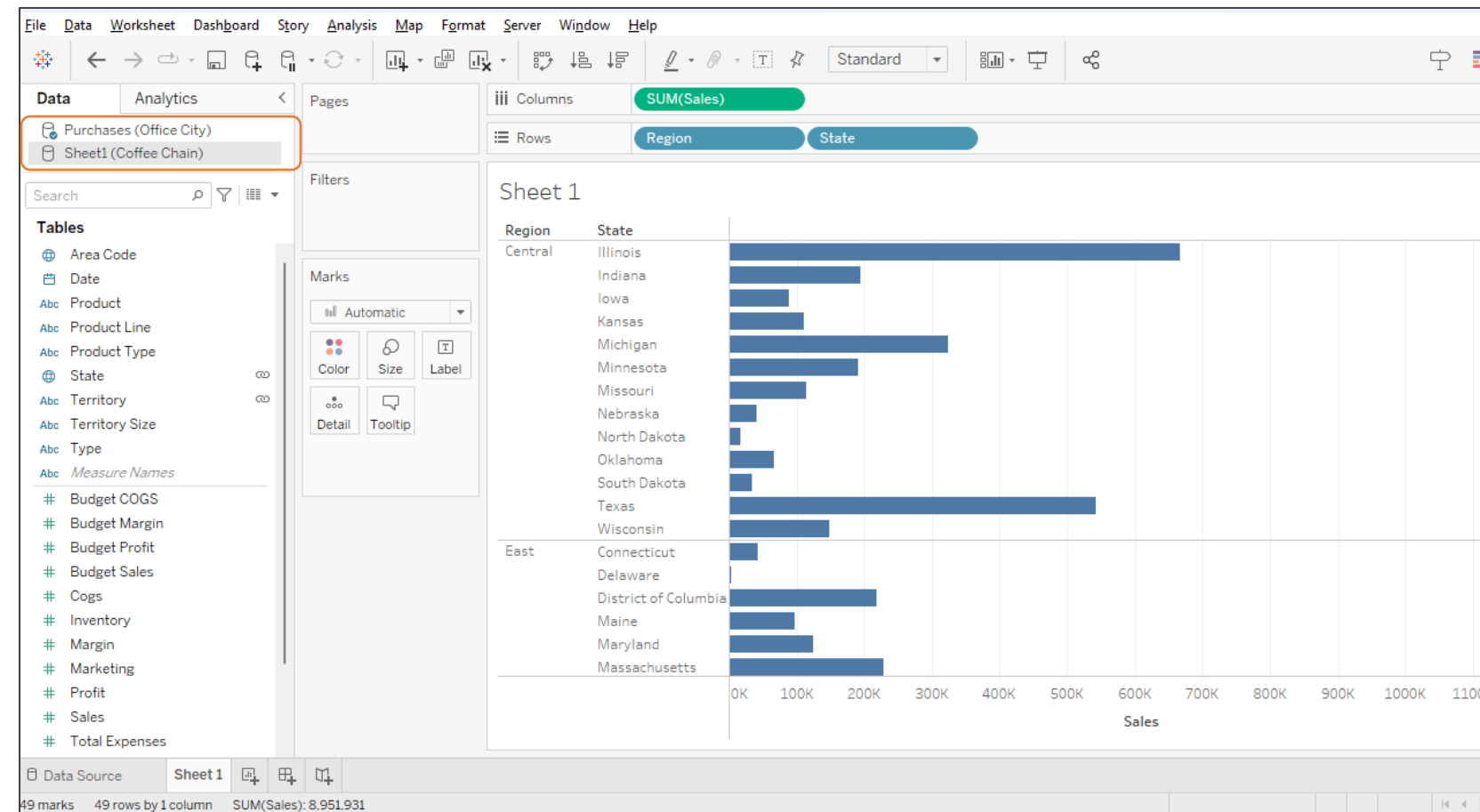
- A. The number of rows should be the same in both datasets.
- B. The number of columns should be the same in both datasets.
- C. The names of the fields should be different in both datasets.
- D. The data types should be different in both datasets.



Data Blending

Data Blending

It refers to combining data from various sources and displaying it on one single screen based on one common field.



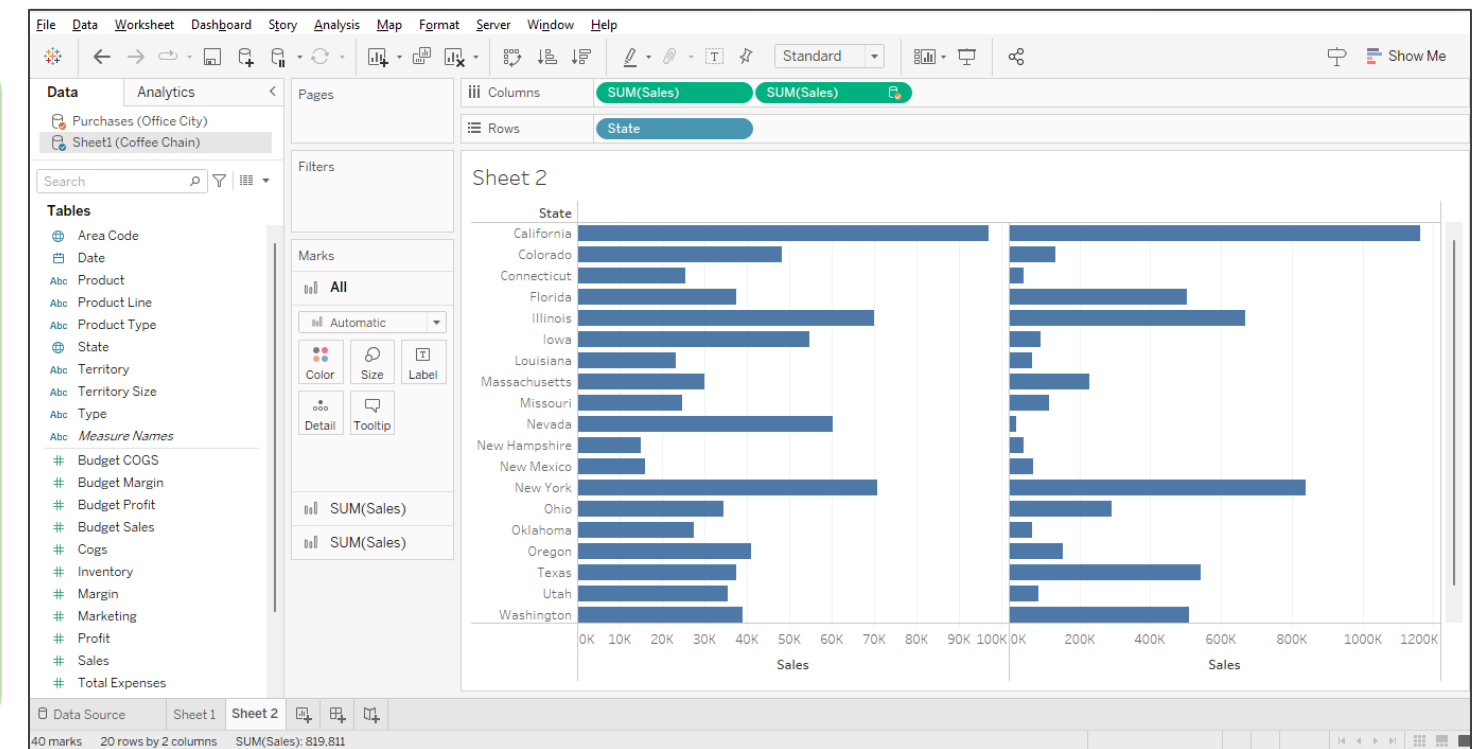
Unlike Relationships, Joins, or Append operations, it does not physically combine the data. It allows one to display data from primary and secondary data sources.

Understanding Data Blending

In data blending, the focus is on the primary data, while the secondary data provides additional context.

The results of data blending are as follows:

- **Query each data source independently** (per your requirement)
- **Aggregate the results** to the required level
- Present the results visually together in the View



Example: The Local Coffee Shop Uses Data Blending Approach



Scenario

The local coffee shop wants to analyze customer data to personalize its marketing, enhance customer engagement, and improve store operations. However, the data is stored in different formats and sources, making it difficult to extract valuable insights.



Solution

The local coffee shop has implemented the data blending technique to blend data from different sources and formats into a single view. By blending data such as customer demographics, purchase history, and store performance, the coffee shop can gain a holistic view of its customers and operations. This enables them to create targeted marketing campaigns, understand customer preferences, and optimize store layouts and operations for better customer satisfaction.

Demo: Data Blending



Duration: 25 minutes

Demonstrate the process of data blending in the provided Coffee Chain and Office City datasets and switch between primary and secondary datasets for analysis..

DEMONSTRATION

Quick Check



Which of the following statements is true about data blending in Tableau?

- A. It requires at least one common field between datasets.
- B. It physically combines the data from the datasets.
- C. To blend data, add the secondary data source from the file menu.
- D. The additional information from the secondary data source is not displayed in the view.

Guided Practice



Overview

Duration: 20 minutes

In this exercise, you will learn to use joins and unions in Tableau to consolidate sales data from multiple stores to combine customer feedback with product information for comprehensive sales performance and satisfaction analysis.

GUIDED PRACTICE

Key Takeaways

- A Relationship is a dynamic, flexible way to combine data from multiple tables for analysis.
- Tableau supports different types of joins, including inner, full outer, left, and right joins.
- Joins merge datasets based on common fields, combining their data for analysis.
- Data blending is a process in which data from several sources are combined.
- Data blending does not physically combine the data, unlike Relationships, Joins, or Append.



Practice Project

Healthcare Analytics

In this project, you will optimize patient healthcare by leveraging Tableau to integrate and analyze diverse data sources. The goal is to improve operational efficiency and research progress within a healthcare organization's analytics team. You will merge patient records, clinical trial results, and external research to generate insights that enhance patient care and inform decision-making processes.



Additional Resources

- [The Tableau data model](#)
- [Union your data!](#)
- [Join your data](#)
- [Use relationships for multi-table data analysis](#)



Q&A

