Lesson 05 Demo 03
Data Blending

Objective: To demonstrate, how to manually set up a relationship, create visualizations, and switch between primary and secondary datasets for analysis in Tableau

Tools required: Tableau Desktop

Prerequisites: None

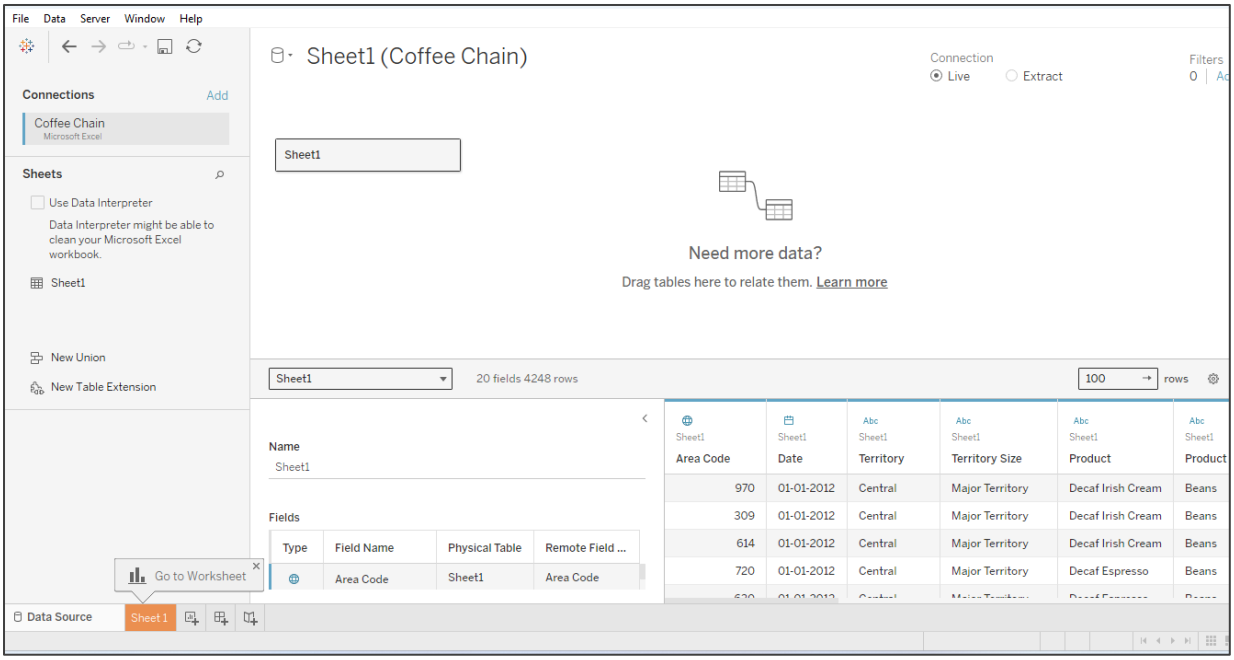
Note: Download the **Coffee Chain** and **Office City** dataset from the Resource Material of the LMS

Steps to be followed:

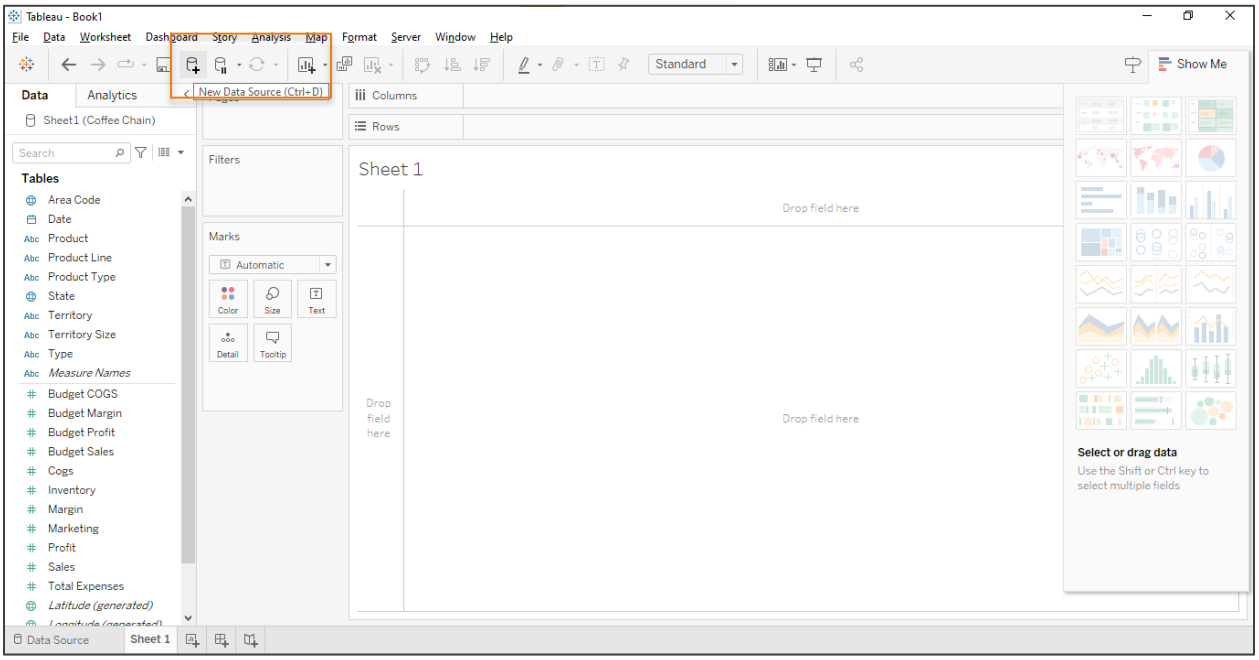
1. Import the dataset
2. Create another relationship between the datasets manually
3. Create a visualization
4. Switch the primary and secondary datasets

Step 1: Import the dataset

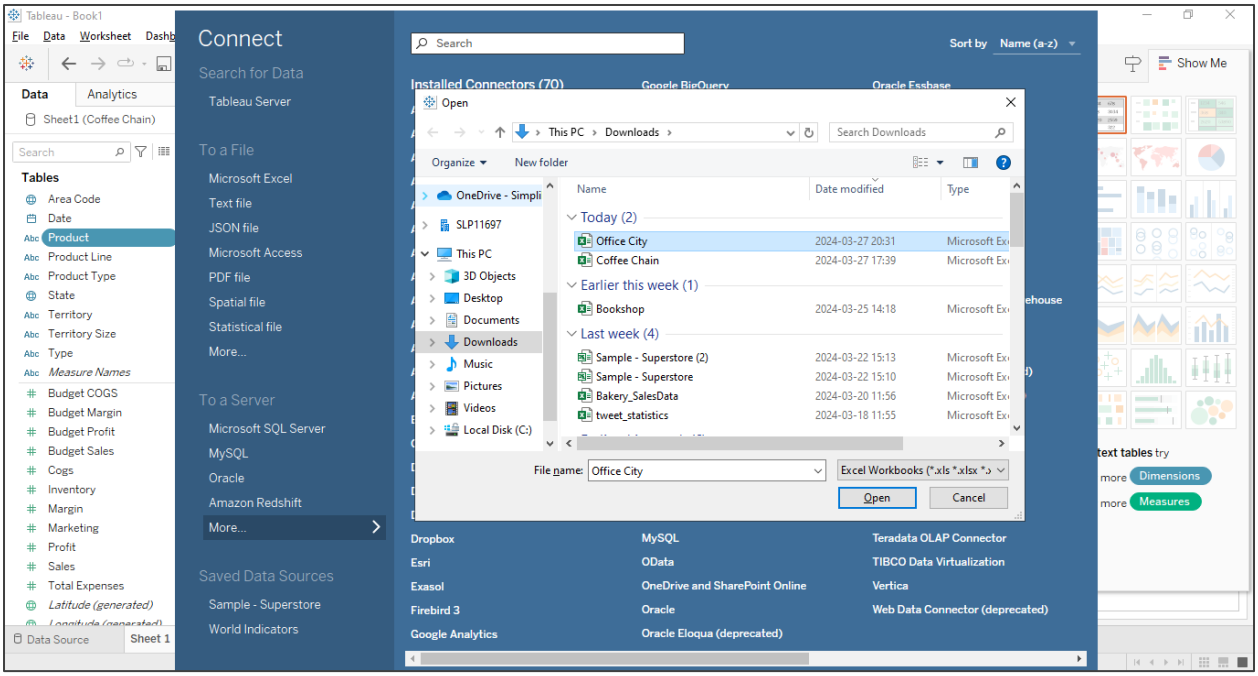
1.1 Open a new Tableau file and connect the **Coffee Chain** dataset. The Data Source view will appear as shown below:



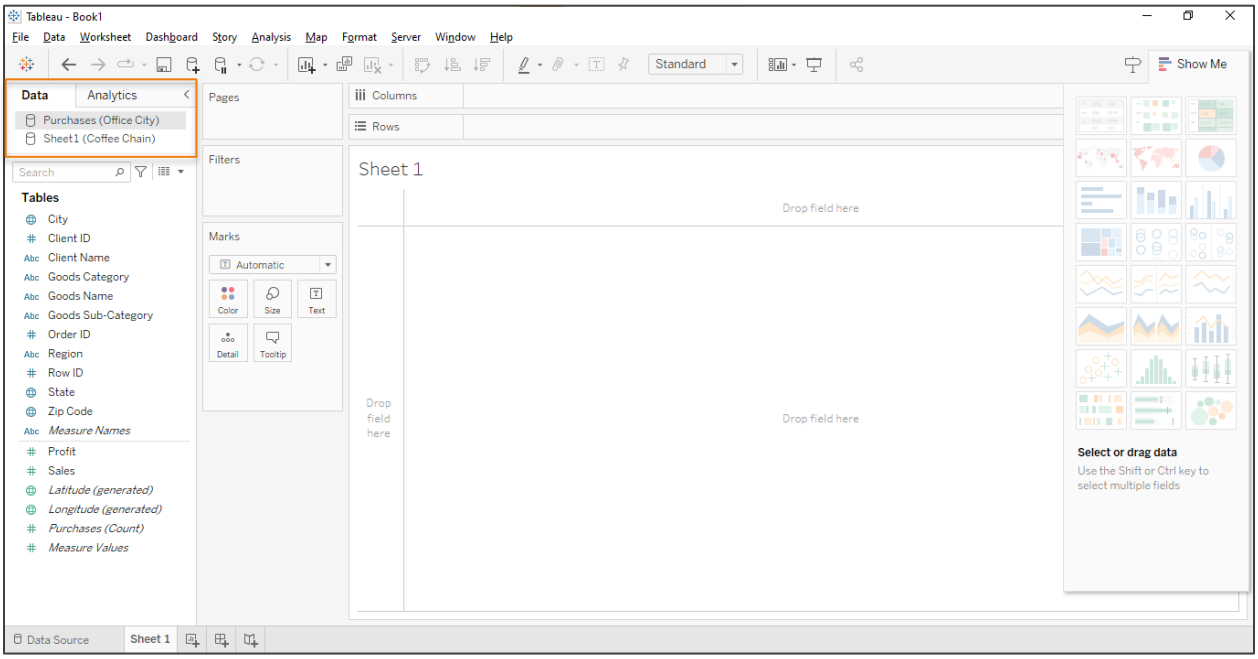
1.2 Select **Sheet1** and click on **New Data Source**, as shown below:



1.3 Select the **Microsoft Excel** and click on **Open**. Choose the file (**Office City**) from your local drive. Tableau will load and establish a connection with the file.

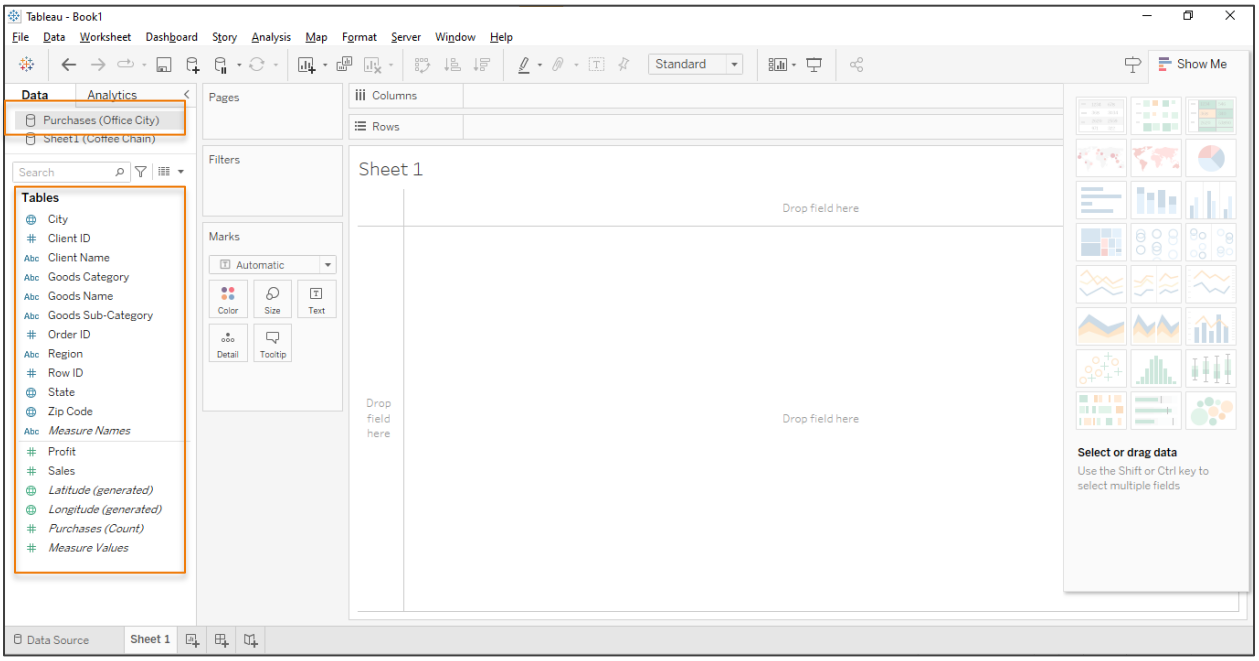


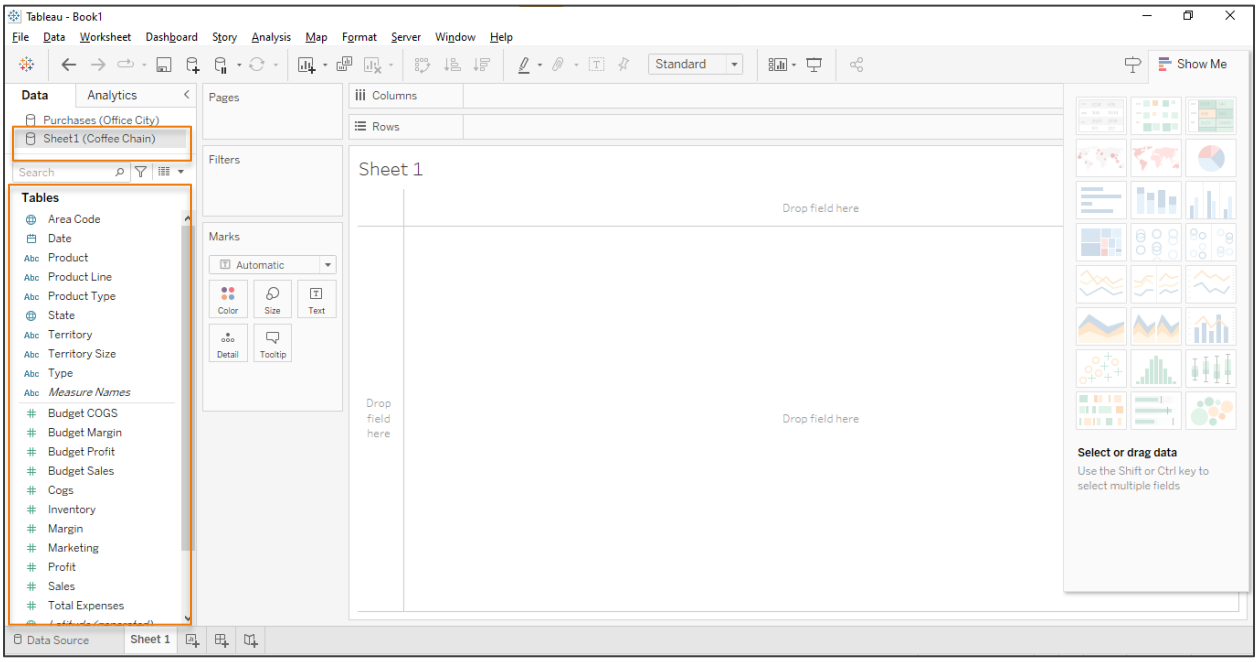
In Sheet 1, the connection of two datasets with Tableau will be shown as below:



Note: You can switch between the two datasets easily by clicking on them. Notice how the dimension and measure fields change. Thus, Tableau has not connected (or linked) the two datasets; rather, it has kept them separate.

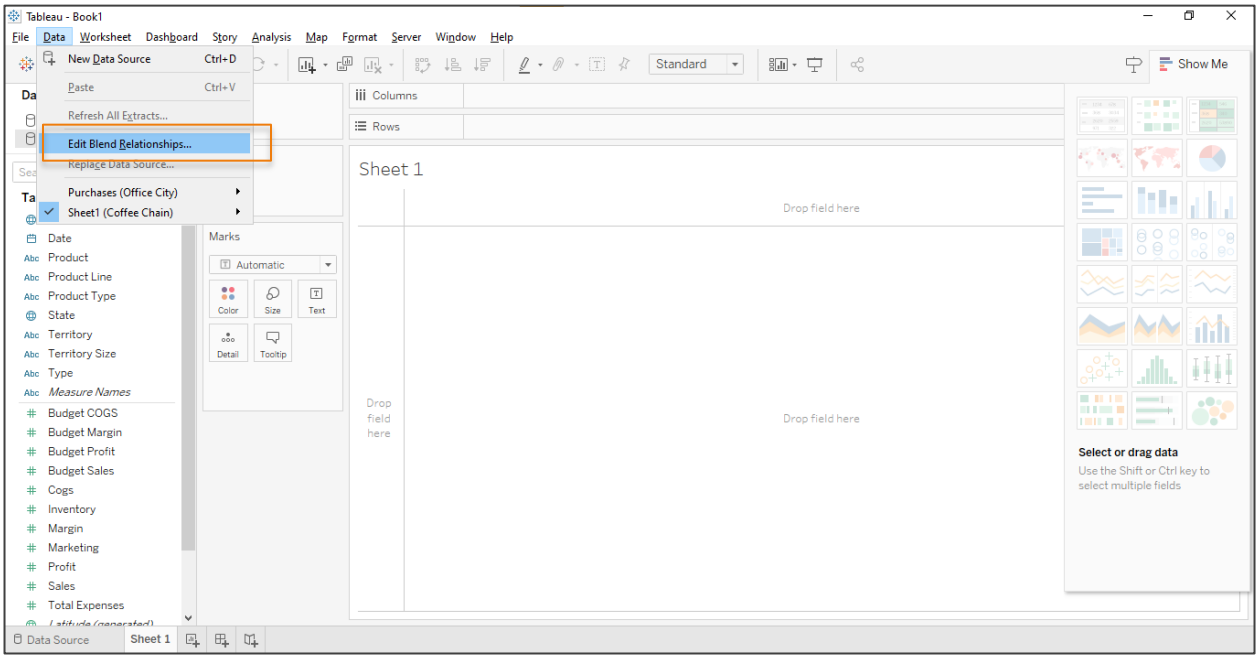
1.4 Click on **Purchases (Office City)** to view the first set of fields below. Click on **sheet1 (Coffee Chain)** to view the second set of fields as shown below:



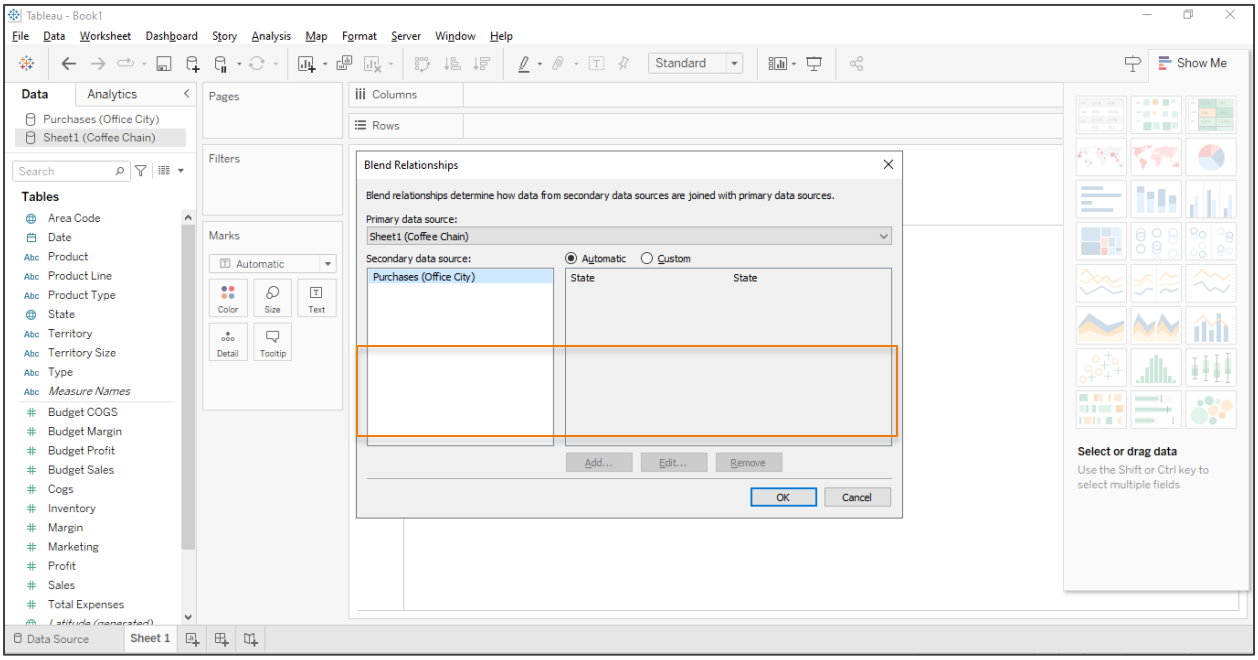


Note: It is important to note that this way, we can connect a Tableau file with multiple datasets even though they do not have any relationship or a common primary key between them. However, if you try to connect multiple files using data model section, you must mandatorily create a relationship between the data tables.

1.5 Go to **Data** and select **Edit Blend Relationship** in the menu to view the automatically created linkage between the tables in the backend



Note: In the Blend Relationship window, State is already used by Tableau to link or blend the two datasets.

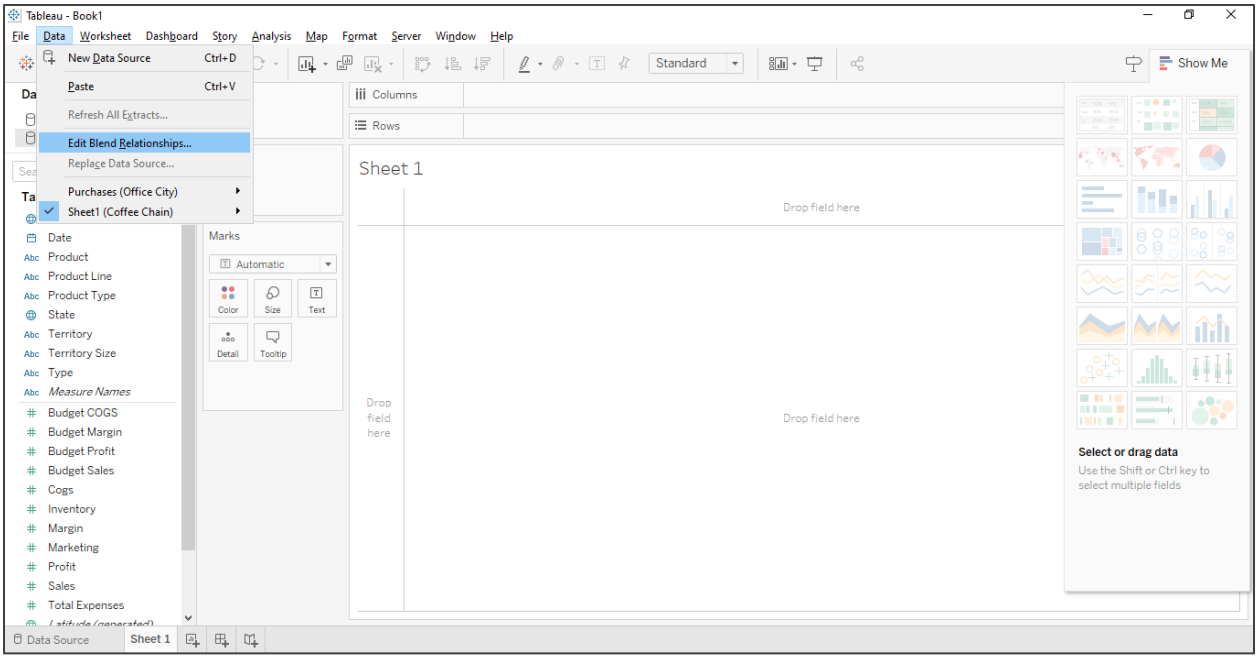


Step 2: Create another relationship between the datasets manually

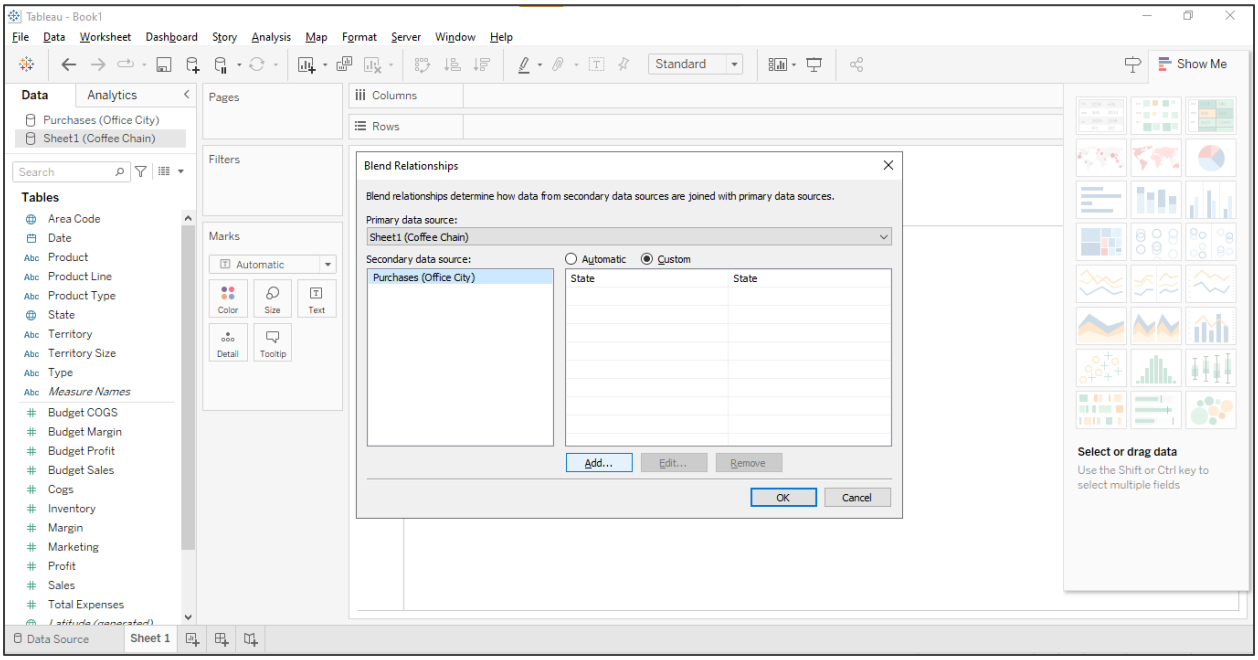
Note: Observe that the **Coffee Chain** dataset has a field called **Territory**, and the **Office City** dataset has a field named **Region**. Check the aliases for **Territory** and **Region**, and it turns out that they are the same as below. It is just that the field names are different.

1. Territory Alias:
    1. Central
    2. East
    3. West
    4. South
2. Region Alias:
    1. Central
    2. East
    3. West
    4. South

2.1 Click on **Data** and select **Edit Blend Relationship** to create a manual relationship between the **Territory** and **Region** fields

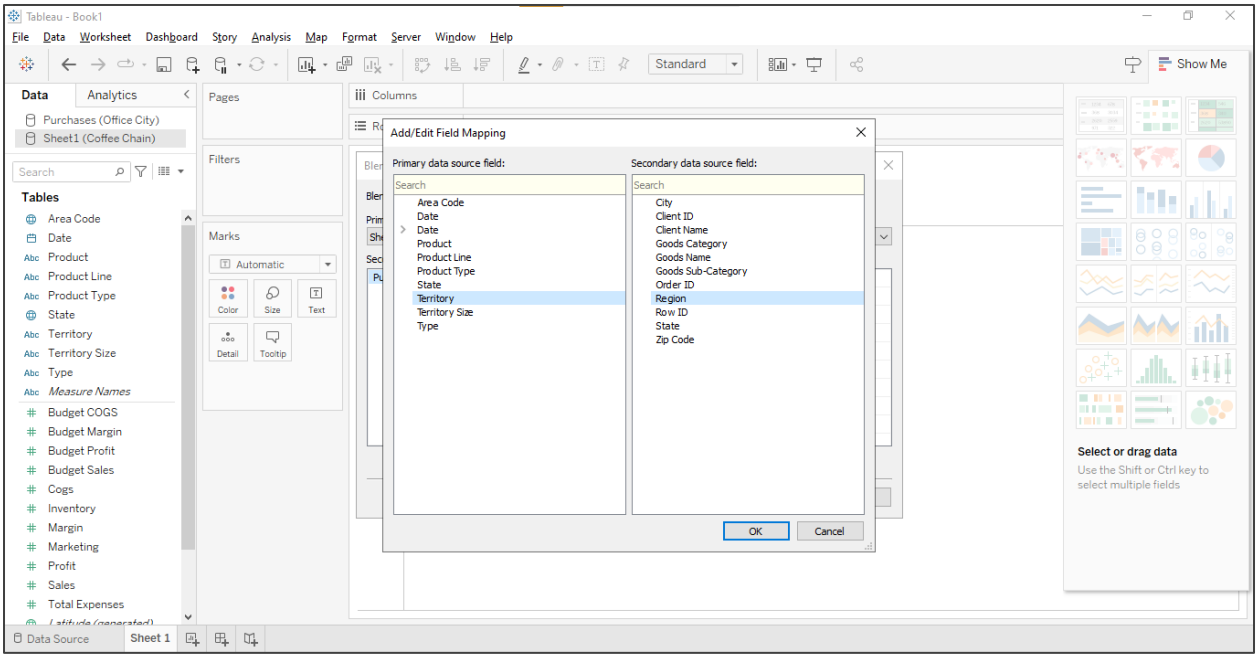


2.2 Select the **Custom** radio button and click on **Add**

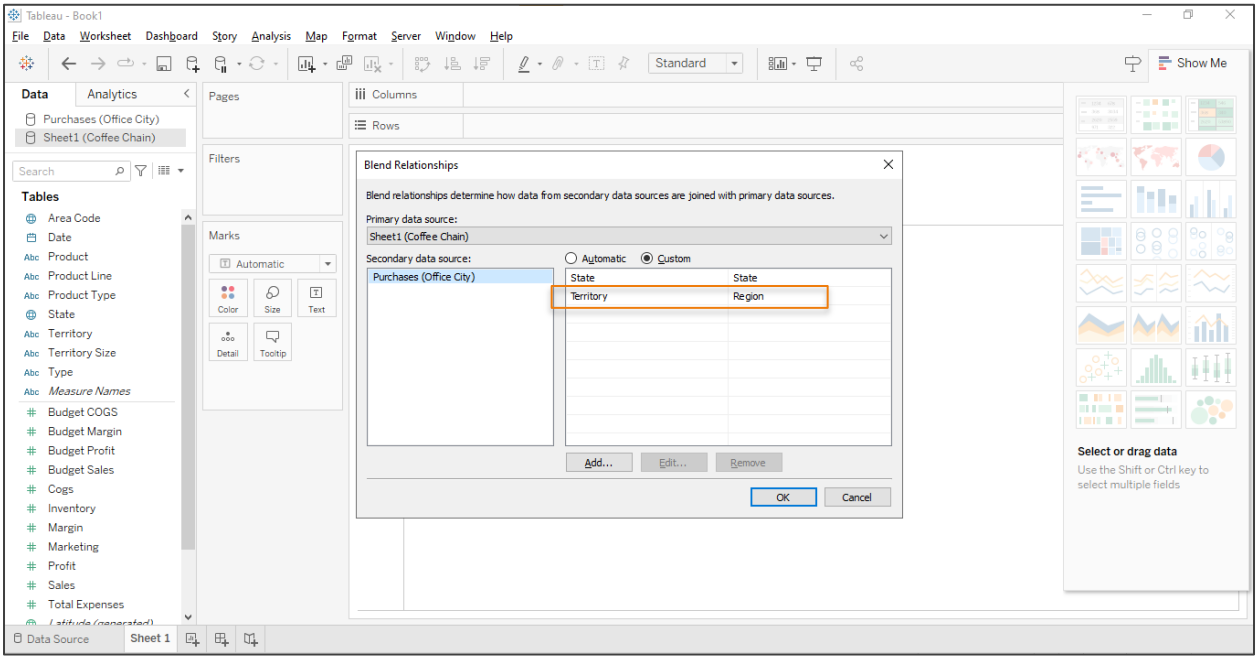


Note: You can change the Primary and Secondary data source from the dropdown, if you need to.

2.3 **Add/Edit field Mapping** window opens as below, click on **Territory** and **Region** as shown below, and then click on **OK**



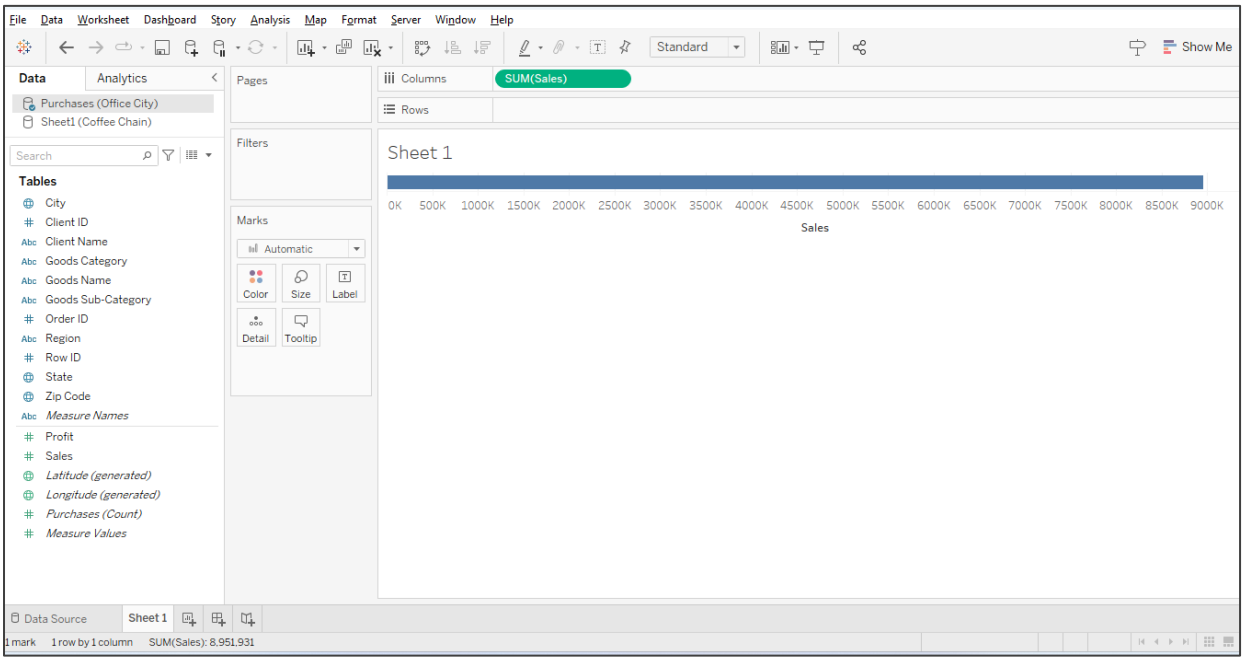
2.4 You will see that another relationship has been added to the window. Click on **OK** to close this window.



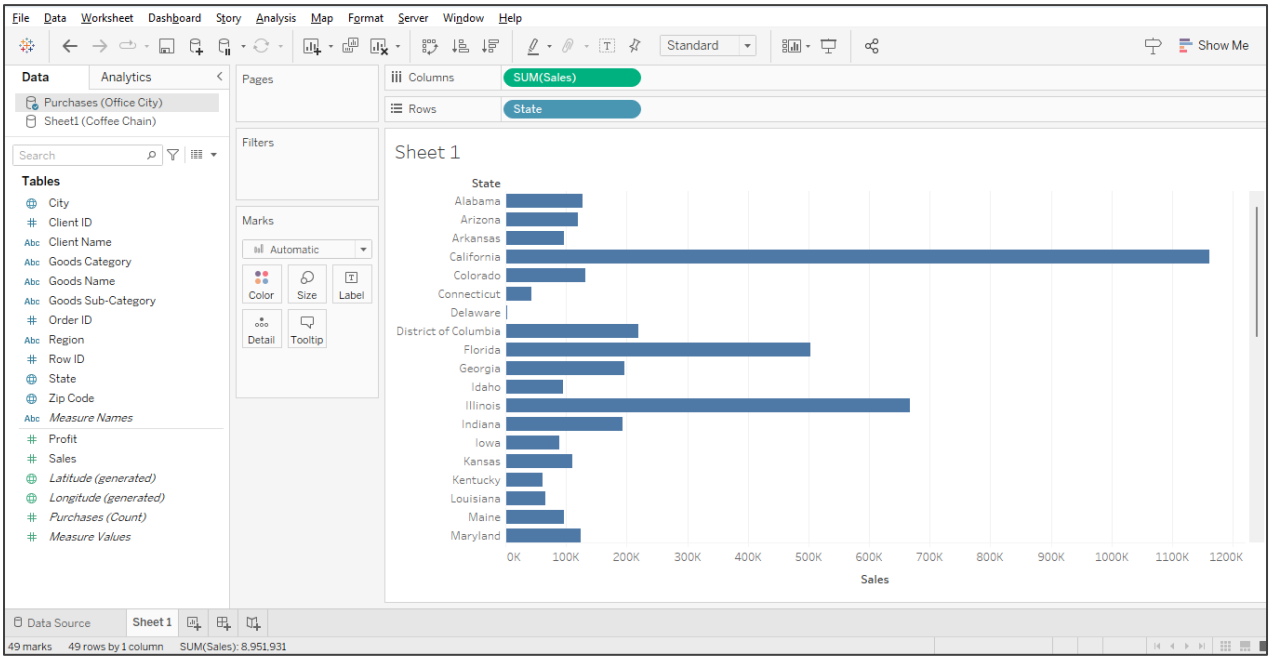
Note: There is another easy way to create the relationship between **Territory** and **Region**. This is by changing the name of one of the fields and making the names of the fields the same. This will instigate Tableau to automatically create the relationship between these fields as they now have the same name.

Step 3: Create a visualization

3.1 Select the **Purchase (Office City)** Dataset by clicking on it and drag **Sales** to Columns

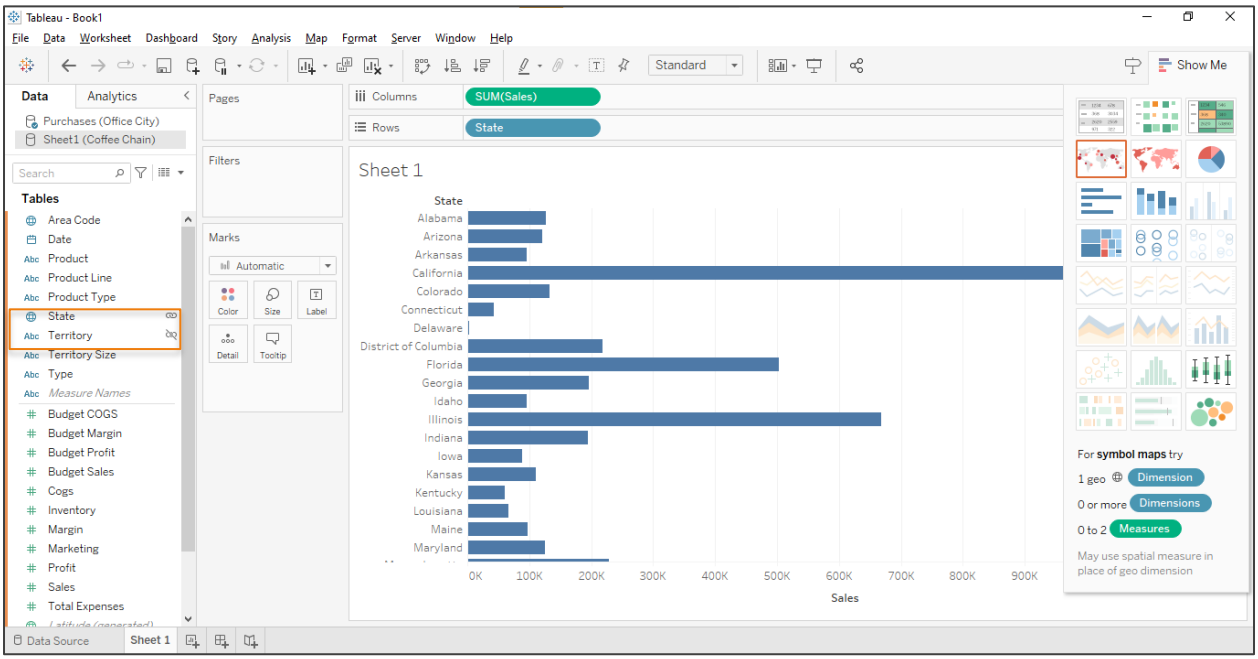


3.2 Drag **State** field of **Office City** to Rows



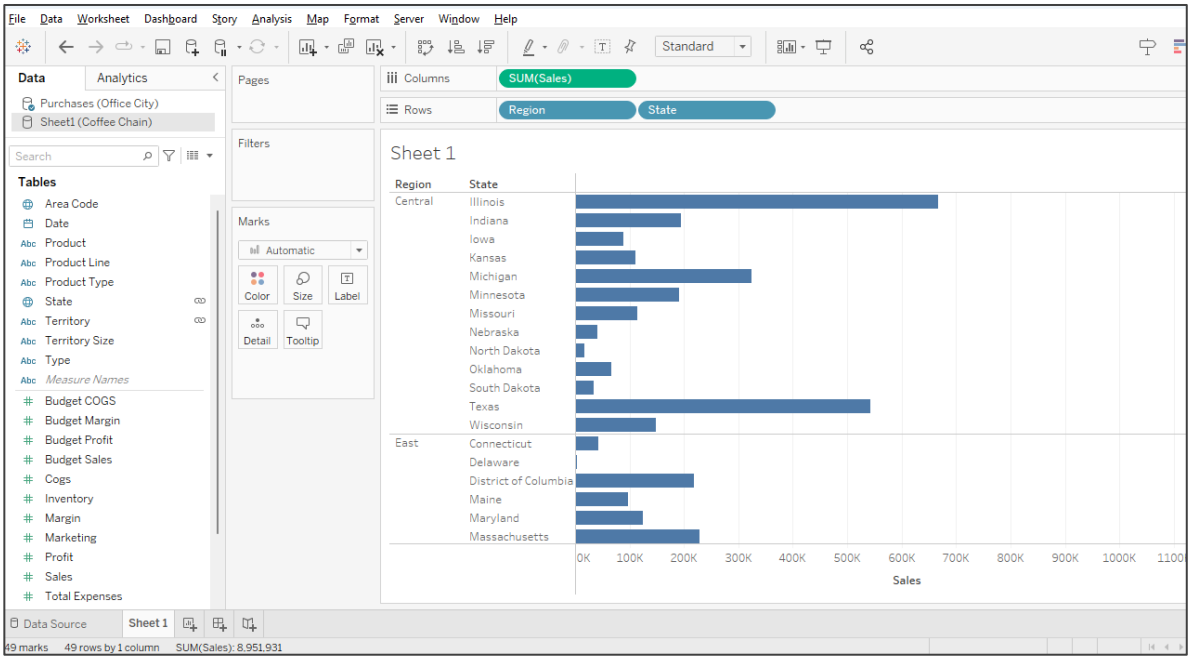
Observation: As you move the **Sales** field from Office City to rows, you will observe a blue tick mark against the Office City dataset in the Data pane. This is an indication that Tableau will now categorize Office City as the primary dataset, that is, the dataset it has used to start creating visualizations.

3.3 Switch to the **Coffee Chain** dataset (which is the secondary dataset in this view) by clicking on it

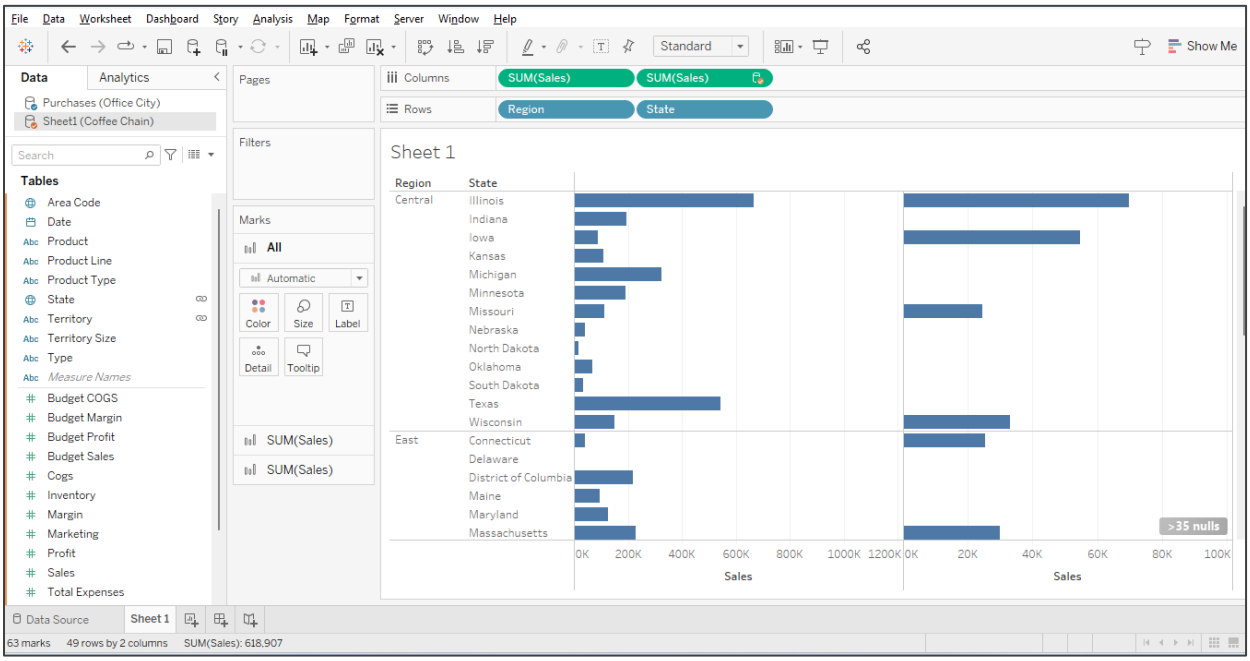


Observation: You will notice two links against the State and Territory dimensions as below. The State link is active, and the Territory link is inactive. You may click on the **Territory** link and make it active.

3.4 We can also make the **Territory** link active using the **Region** field from **Office City** in the row.



3.5 Now, bring the **Sales** from **the Coffee Chain** dataset into the **columns**. This will create the chart below:

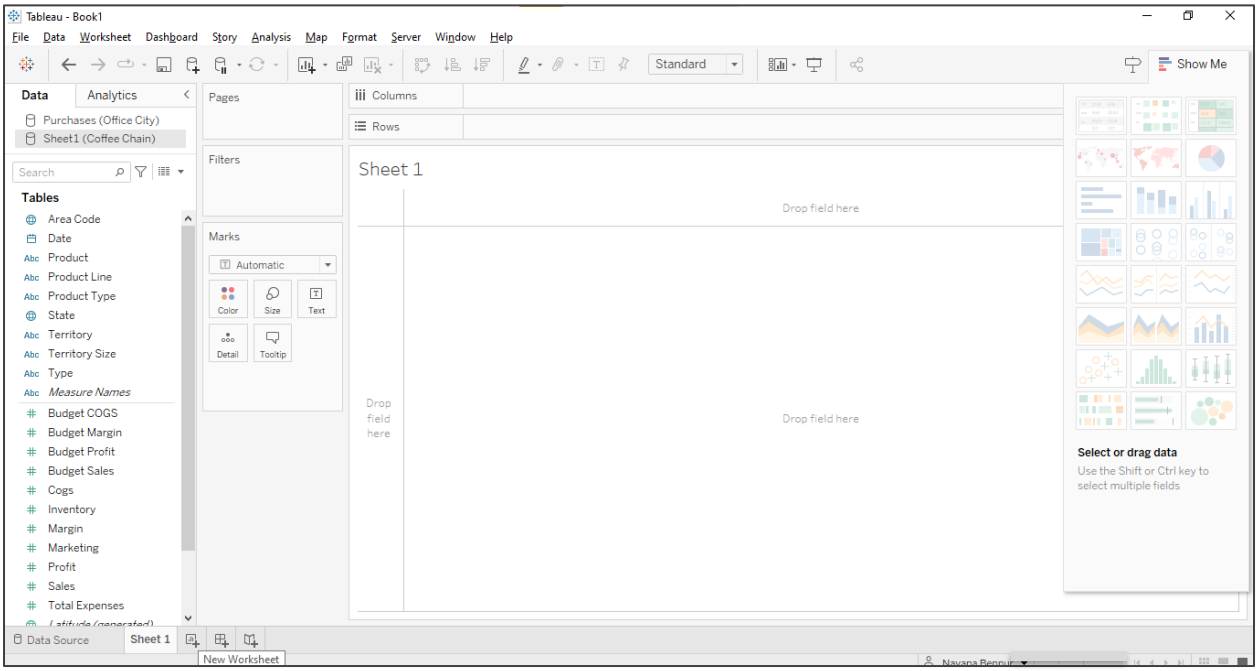


Observation: In the second bar chart of the Coffee Chain dataset, there are 35 states with no sales transactions, shown as nulls. The dataset is marked as secondary, and Tableau uses states from the primary dataset first and then

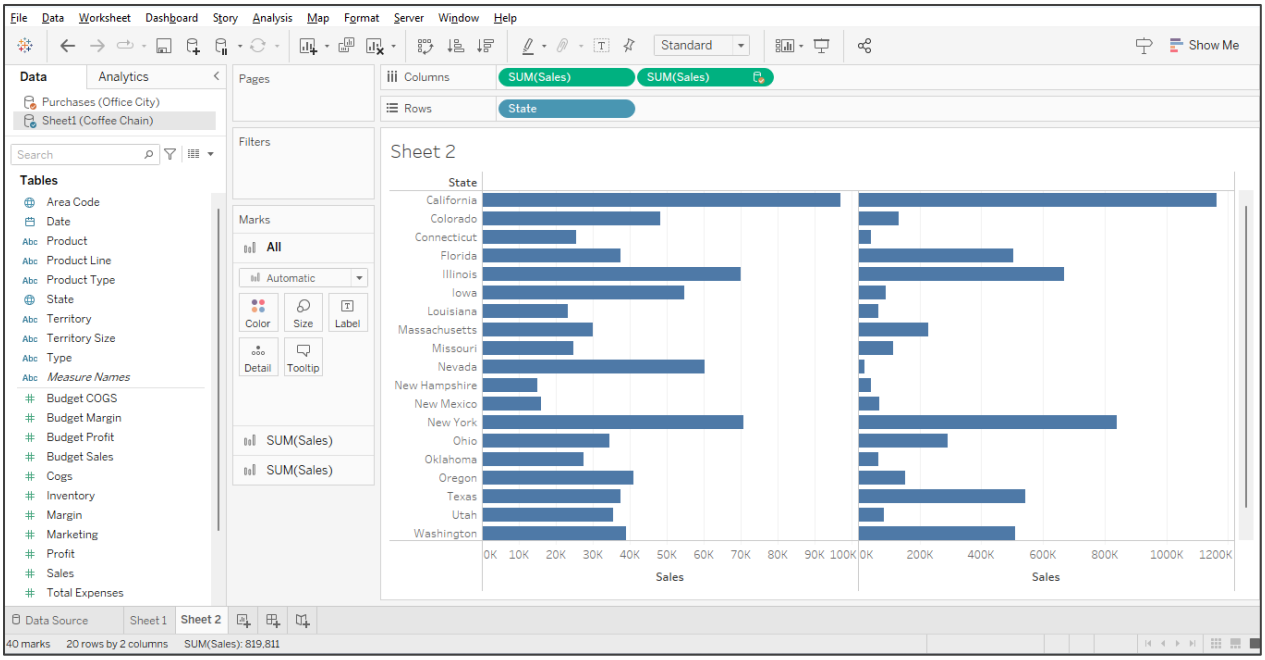
maps sales from the secondary dataset only for existing states, leaving out states with no sales.

Step 4: Switching the Primary and Secondary Datasets

4.1 Open a new Sheet



4.2 Use **Sheet1(Coffee Chain)** as the primary dataset and **Purchases (Office City)** as the secondary dataset to create the same visualization to viewing the **Sales**, Then, observe the difference between them.



Observation: This time, only the States that exist in the Coffee Chain are plotted in the bar charts. The Sales of the states that exist in the Office City are not plotted as only the States from the Coffee Chain are used to create visualization (being the Primary Dataset).

With these steps, you can enable efficient data management and analysis in Tableau, ensuring correct relationship setup and tailored visualizations.