

LAB 1: CONNECTING TO DATA



Table of Contents

Learning Outcome..... 2

Task 1: Connect to an Excel File..... 2

Task 2: Change and Save a Data Source 6

Task 3: Join Tables and Build a View..... 10

Task 4: Blend Data from Two Sources 12

Task 5: Union Data and Merge Fields 18

Learning Outcome

At the end of this session, learners will be able to:

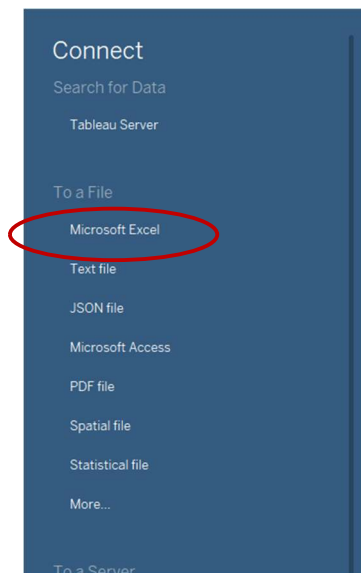
- Connect to an Excel File
- Create a data source file for sharing
- Create a table join
- Create a view that uses joined tables
- Blend data from two sources

Task 1: Connect to an Excel File

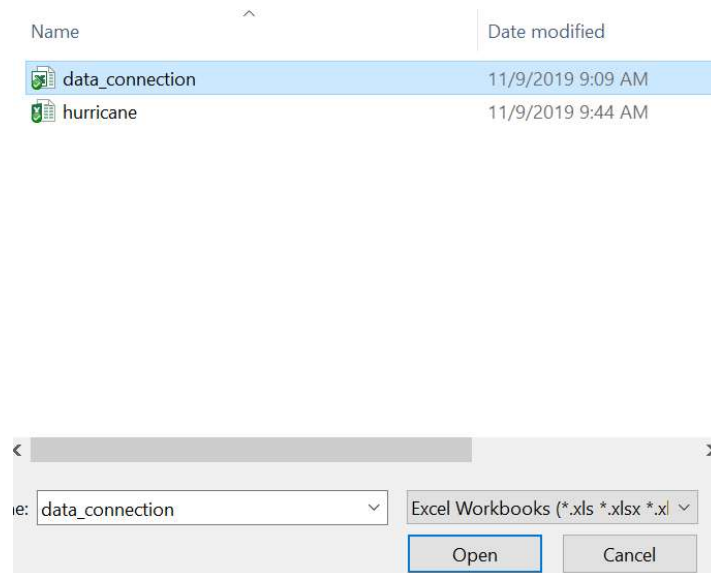
In this exercise, you will connect to an Excel file that contains customer order data to prepare for analysis in Tableau. Connect to the file, give the connection a more friendly name, and edit some field name and a data type.

1. Open Tableau Desktop, and on the Start page, click **Microsoft Excel**.

Step 1

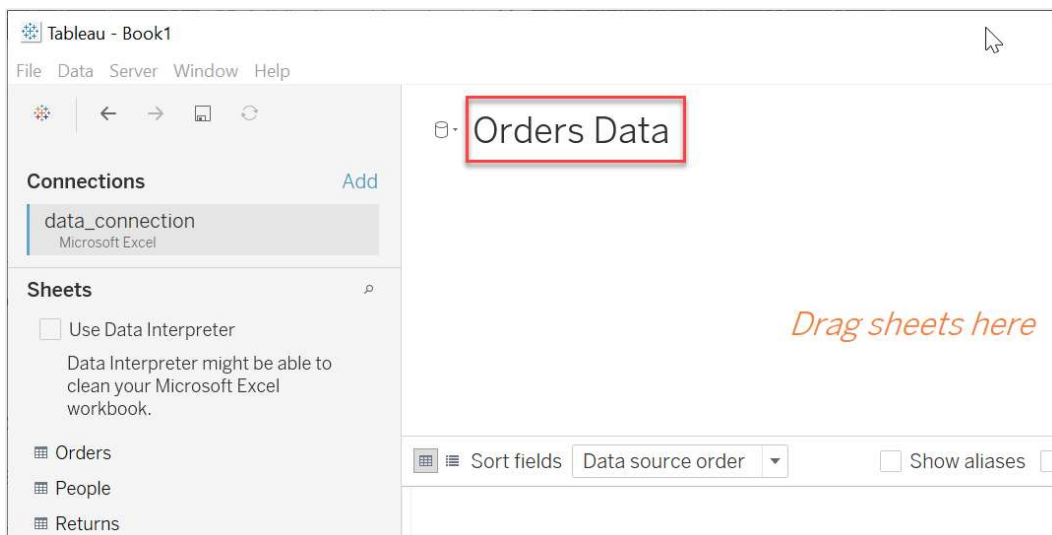


Step 2

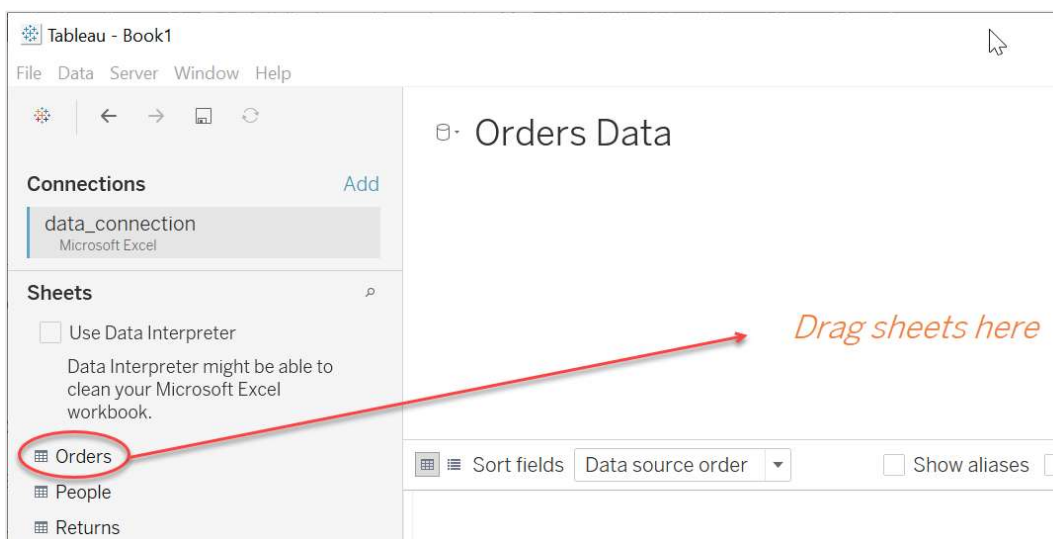


2. In the Open dialog box, navigate to **data_connection.xls**, select it, and click **Open**.

- At the top of the Data Source page, next to the data source icon, rename the connection name to **"Orders Data"**.



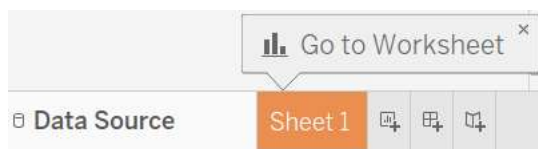
- Under **Sheets**, click and drag **Orders** to the **"Drag sheets here"** area.



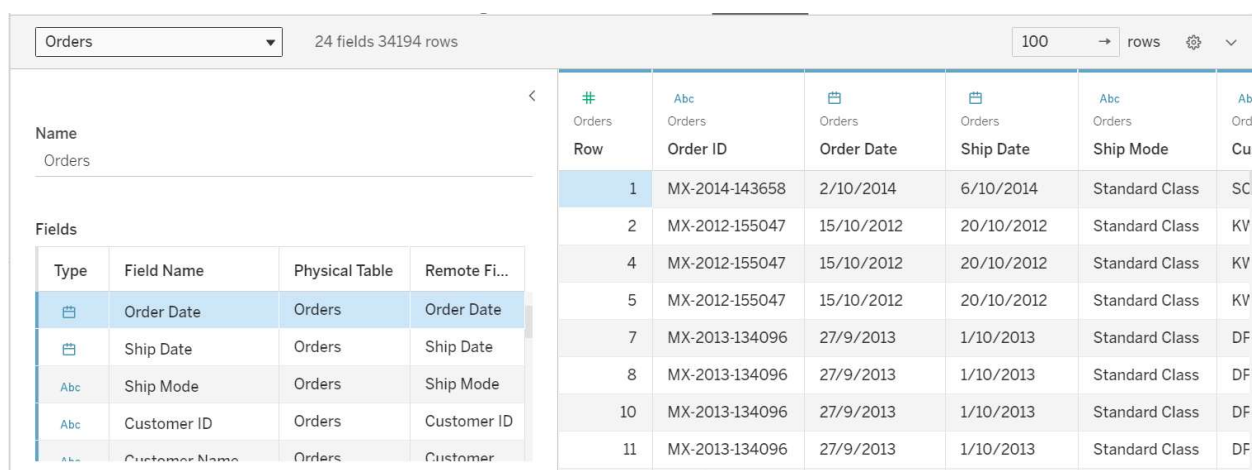
- By default, when Tableau connect to a data, it creates a **live** connection to the data.



- To make the grid area easier to see, close the small **Go to Worksheet** window near the bottom of the screen, near the **Sheet 1** tab.



- Under the Data Source tab, the Metadata panel is located on the left panel, while the Preview panel is on the right.



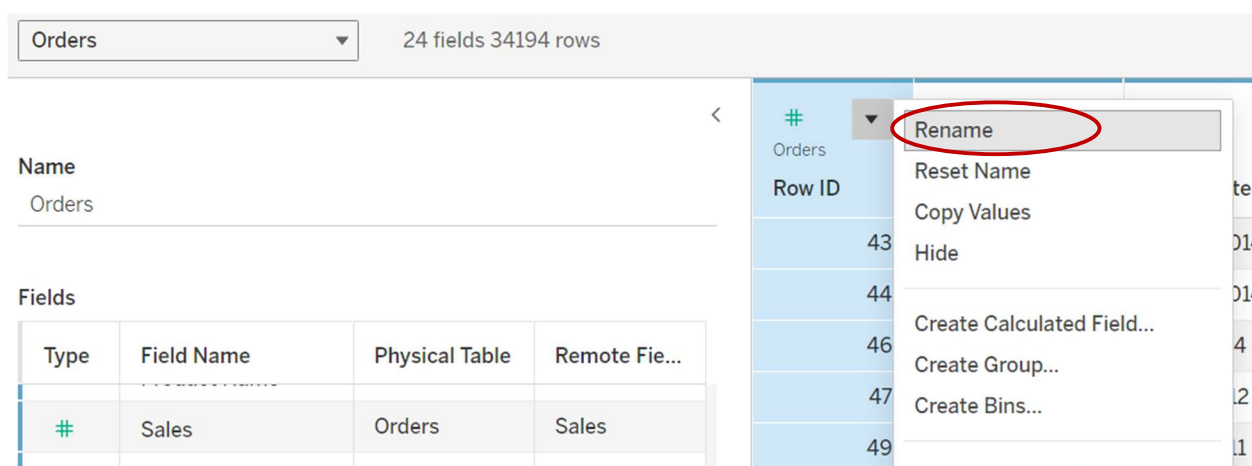
Orders 24 fields 34194 rows 100 rows

Name	Orders
Orders	

Type	Field Name	Physical Table	Remote Fi...
Order Date	Order Date	Orders	Order Date
Ship Date	Ship Date	Orders	Ship Date
Ship Mode	Ship Mode	Orders	Ship Mode
Customer ID	Customer ID	Orders	Customer ID
Customer Name	Customer Name	Orders	Customer Name

Row	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name
1	MX-2014-143658	2/10/2014	6/10/2014	Standard Class	SC	
2	MX-2012-155047	15/10/2012	20/10/2012	Standard Class	KV	
4	MX-2012-155047	15/10/2012	20/10/2012	Standard Class	KV	
5	MX-2012-155047	15/10/2012	20/10/2012	Standard Class	KV	
7	MX-2013-134096	27/9/2013	1/10/2013	Standard Class	DF	
8	MX-2013-134096	27/9/2013	1/10/2013	Standard Class	DF	
10	MX-2013-134096	27/9/2013	1/10/2013	Standard Class	DF	
11	MX-2013-134096	27/9/2013	1/10/2013	Standard Class	DF	


- Click the drop-down arrow next to **Row**, click **Rename**, change the name to **Row ID**, and press Enter.

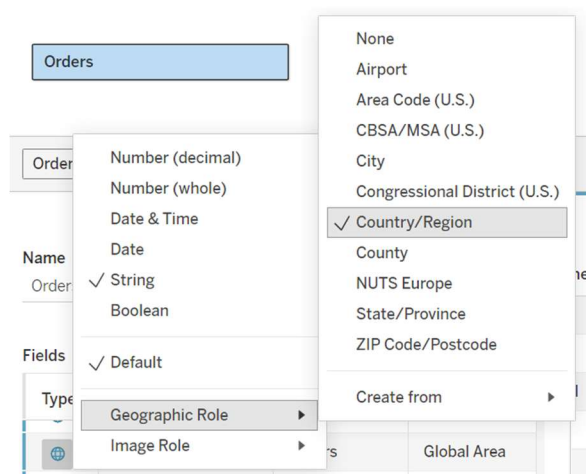


Orders 24 fields 34194 rows

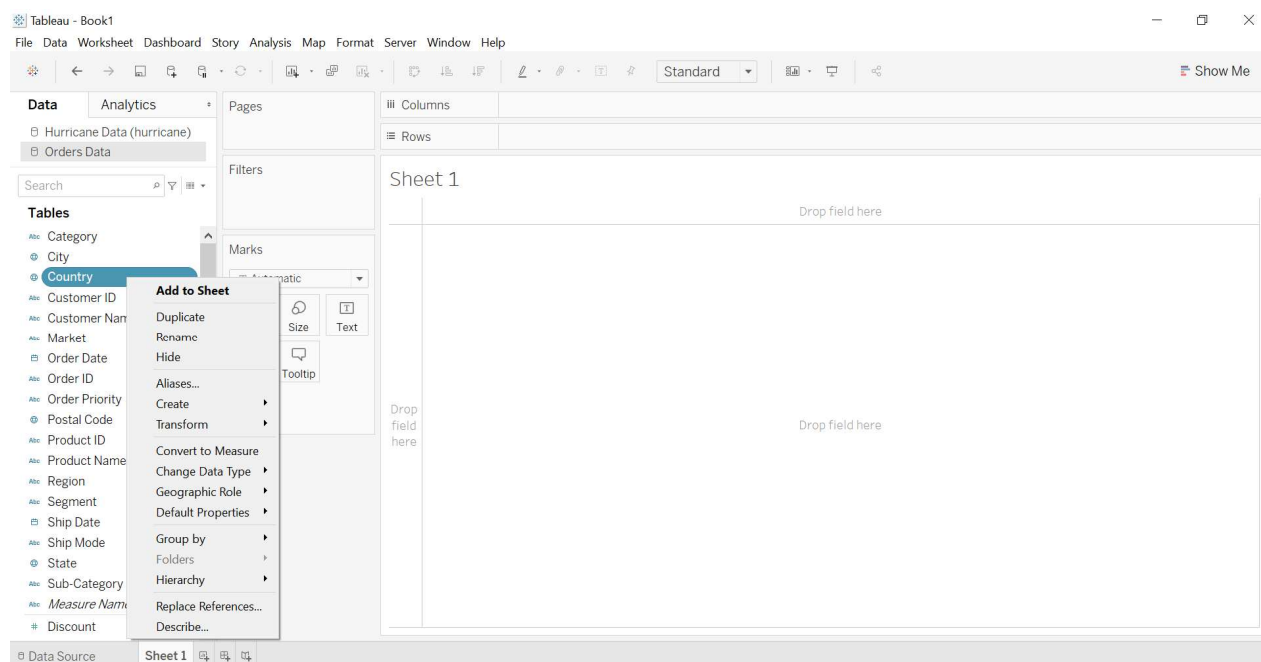
Name	Orders
Orders	

Type	Field Name	Physical Table	Remote Fie...
Row ID	Sales	Orders	Sales

- Follow the same steps to rename **Global Area** to **Country**.
- To the left of the Country field, click the **string** icon  select **Geographic Role** → **Country/Region**.



11. Click the **Sheet 1** tab to go to a blank worksheet.
12. Click on the drop-down arrow next to one of the field names, such as **Country**, and note the menu items. You can make all the same metadata changes in this view that you can on the Data Source page. There are additional options, such as changing a **Measure to a Dimension** or **Dimension to a Measure**.

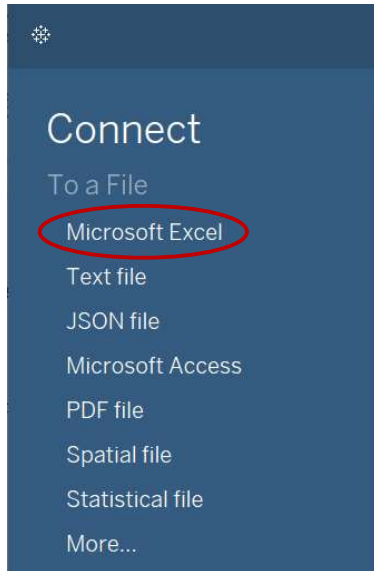


Task 2: Change and Save a Data Source

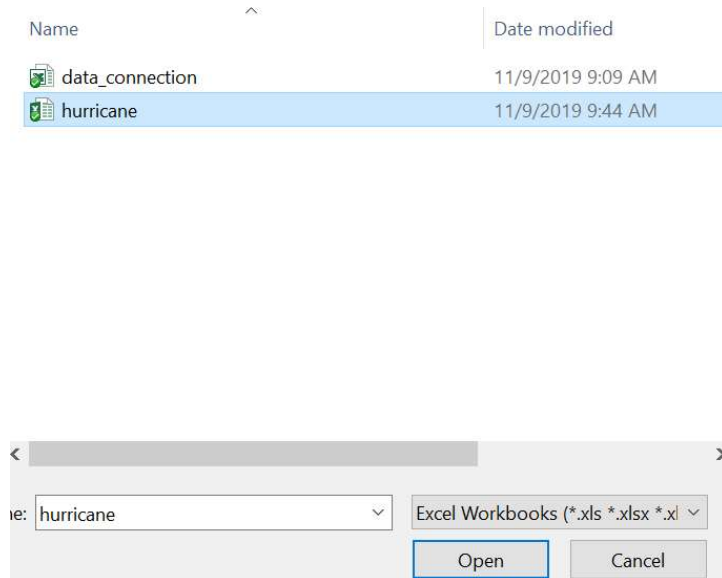
In this exercise, you have a Microsoft Excel file with data about hurricanes that you want to share with some coworkers for building workbooks in Tableau. You need to make some changes to the data source before it is ready to share so that your colleagues can use the data for analysis.

1. Open Tableau Desktop, and on the Start page, click **Microsoft Excel**.

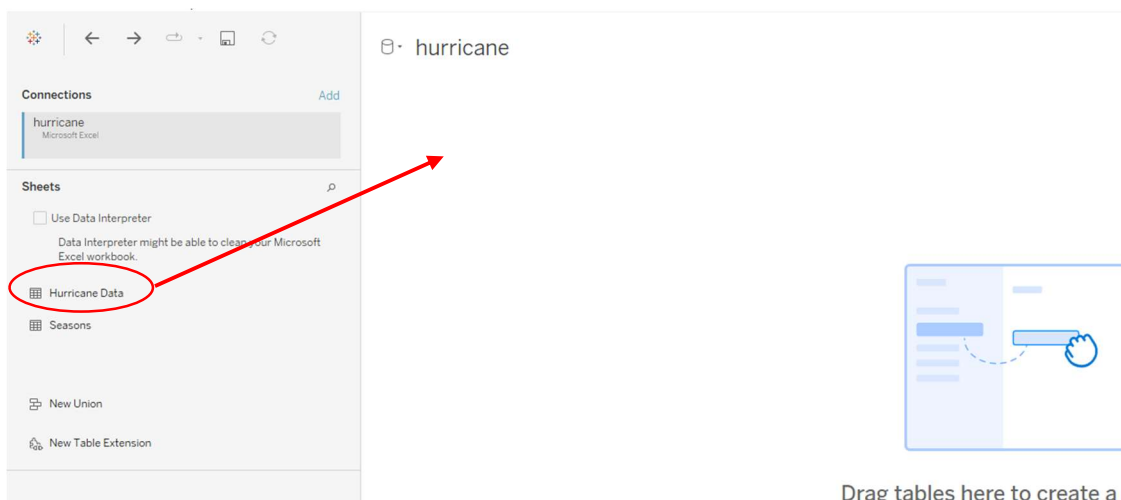
Step 1



Step 2

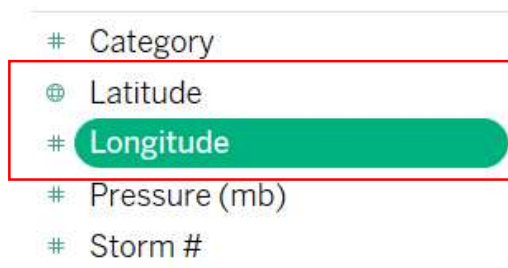


2. In the Open dialog box, navigate to **hurricane.xlsx**, select it, and click **Open**.
3. Drag the **Hurricane Data** sheet to the data area.

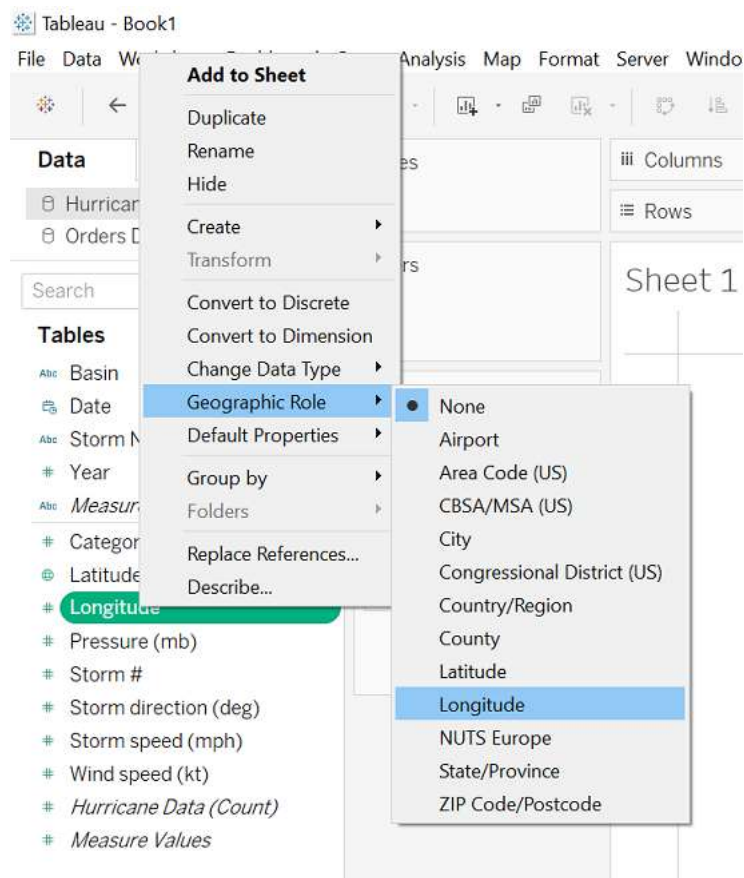


4. Click **Sheet 1** to open the Tableau Desktop workspace.

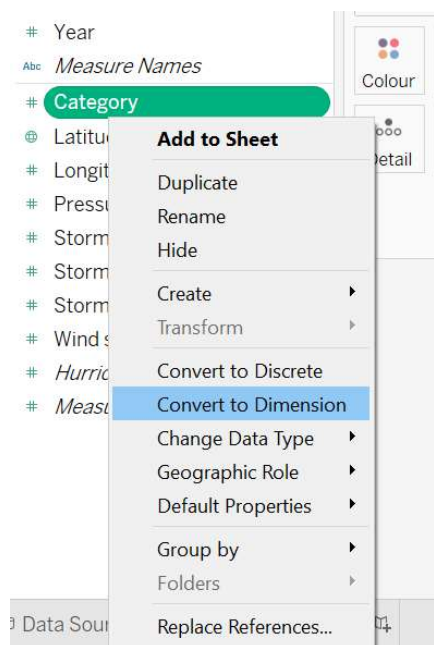
- Right click the **Lat (deg)** field and rename it to **Latitude**. Repeat the same steps to rename **Long (deg)** field to **Longitude**.



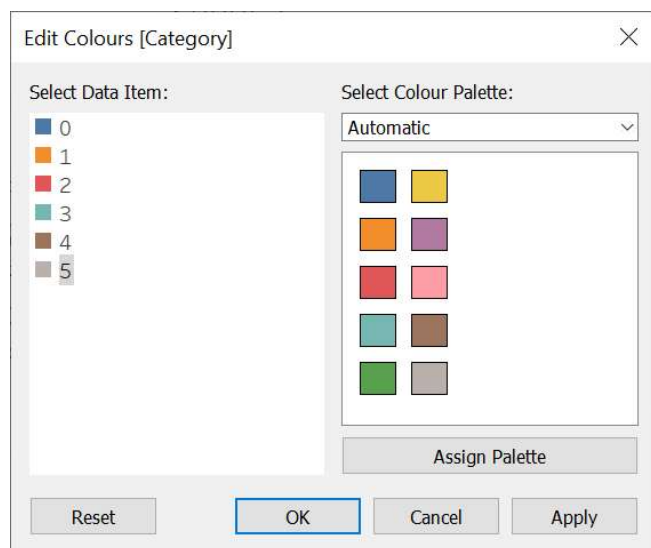
- Right click **Longitude**, select **Geographic Role** → **Longitude**.



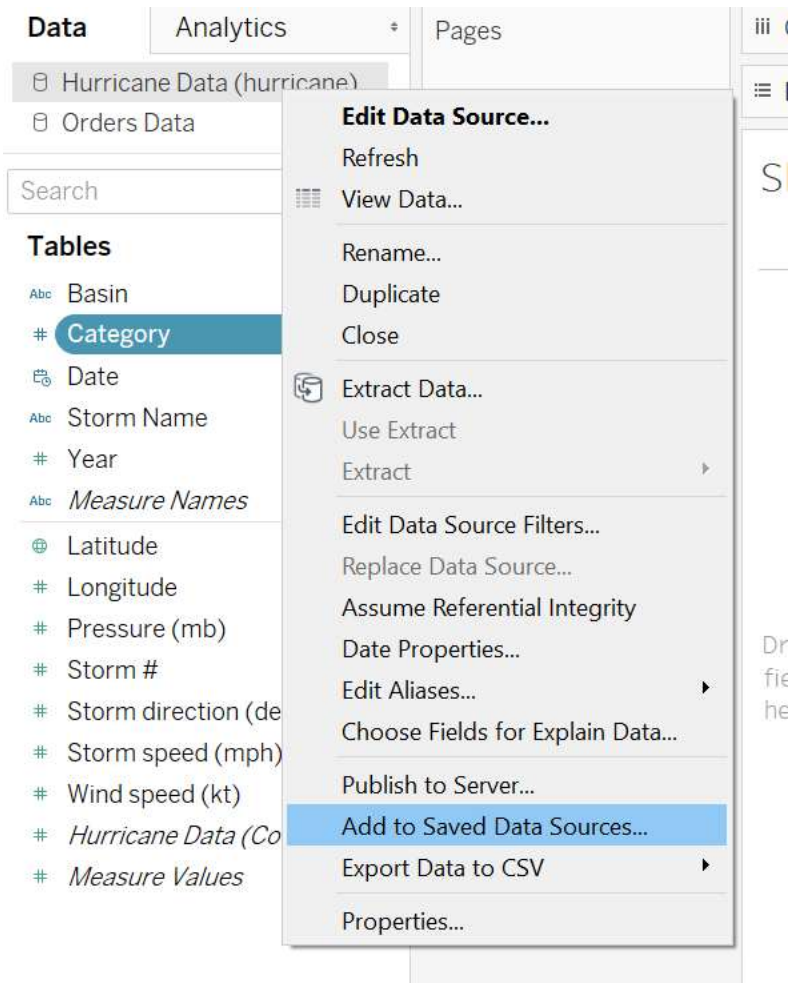
- Drag the **Category** field from **Measures** to **Dimensions**.



8. Right-click **Category** select **Default Properties** → **Colour**. Under **Select Data Item**, assign **brown** colour for **Category 4** and **gray** colour for **Category 5**. Click **OK**.



9. On the **Data** tab, right-click on **Hurricane Data (hurricane)**, and click **Add to Saved Data Sources**.

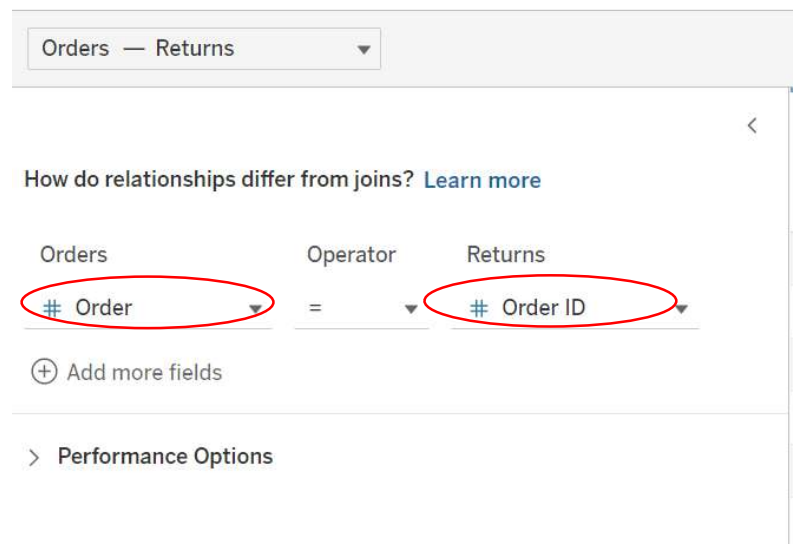


10. Save the Data Source file to your preferred location. Close Tableau Desktop (you don't need to save the workbook). The data source (.tds) file stored the metadata edits, data extract and calculated fields information. You can now share the data source file with your co-workers.

Task 3: Join Tables and Build a View

In this exercise, you will build a view that shows a company's orders that were returned. The data source has separate tables for orders and returns. You will join the two tables and build a view that shows the records that are present in both tables.

1. Open the **join_tables_starter.twbx** workbook. Click the **Data Source** tab in the lower left corner.
2. Click and drag the **Orders** table to the data area. Click and drag the **Returns** table to the same area, exactly to the right of the **Orders** table.
Tips: You can also double-click to move the table. You might find double-clicking easier.
3. In the **Edit Relationship** window, make sure the join type is **=**.
 - a. Under **Orders**, select the **Order** field from the drop-down list.
 - b. Under **Returns**, select the **Order ID** field from the drop-down list.



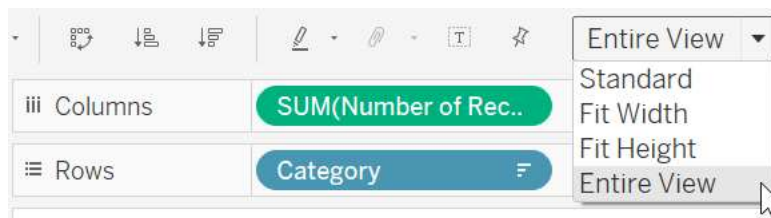
Scroll to the right in the data preview pane to see the records added from the Returns table.

4. Click **Sheet1** tab at the left bottom corner to open the chart designer.

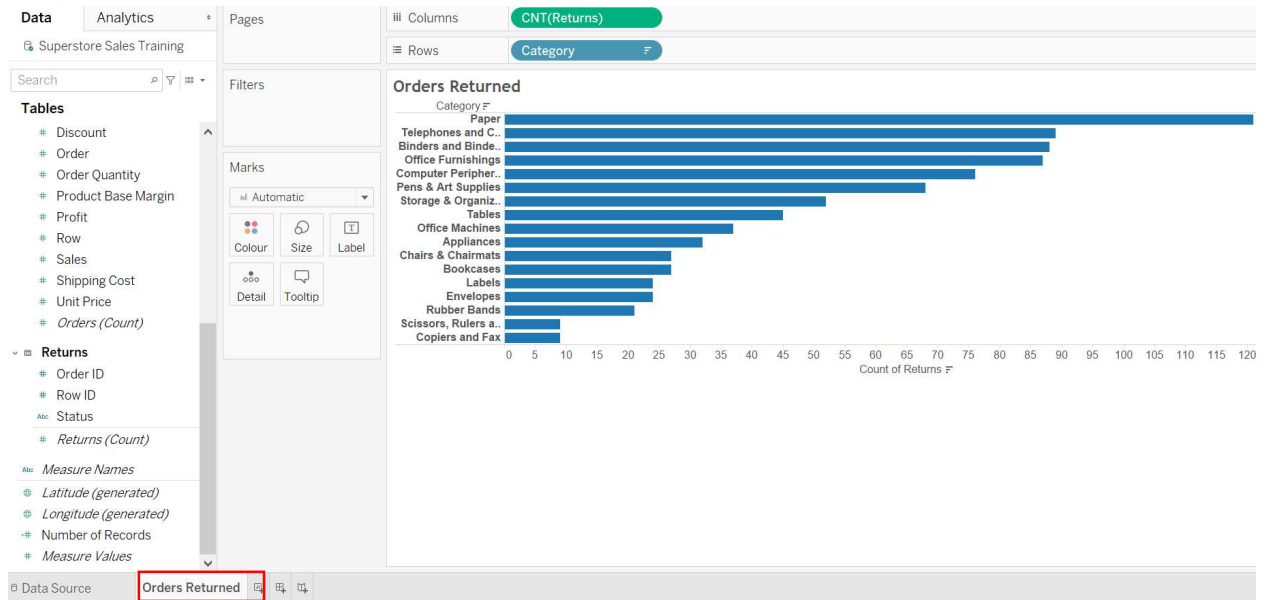
5. Drag **Returns (Count)** to the **Columns** shelf. As we have created an inner join, this includes only records with entries in both tables. These are the records for orders that were returned.
6. Drag **Category** to the **Rows** shelf.
7. On the **Number of Records** axis, click the **Sort** icon to sort in descending order.



8. In the **Fit** box, change the view to **Entire View**.



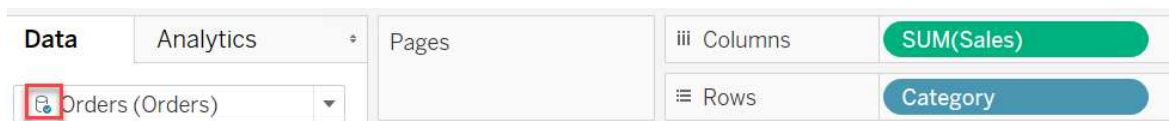
9. Rename the worksheet as **Orders Returned**.



Task 4: Blend Data from Two Sources

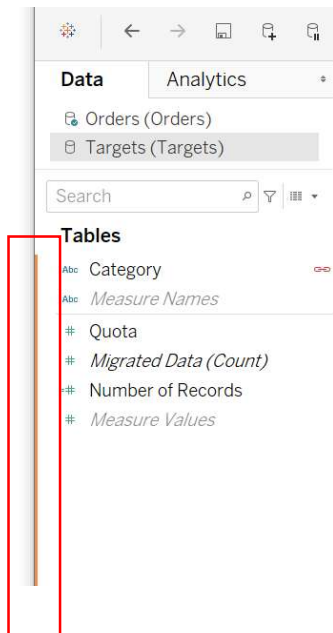
In this exercise, you will build an Orders view that uses Orders as the primary data source. Then you will build a view with Targets as the primary data source.

1. Open the **blend_data_starter.twbx** workbook.
2. Click on the **Explore Orders** tab. Notice that Explore Orders uses the **Orders** data source. The Orders data source includes **Labels, Rubber Bands** and **Storage & Organization**.
3. Click on the **Explore Targets** tab. Notice that Explore Targets uses the **Targets** data source. The Targets data source includes **Décor**, which was not present in the Orders data source. It does not include Labels, Rubber Bands, or Storage & Organization, where were present in the Orders data source.
4. Click on the **Orders** tab. Notice that neither data source has been selected as the primary data source yet.
5. In the **Data** pane, select the **Orders** data source.
6. Drag **Sales** to the **Columns** shelf.
7. Drag **Category** to the **Rows** shelf.



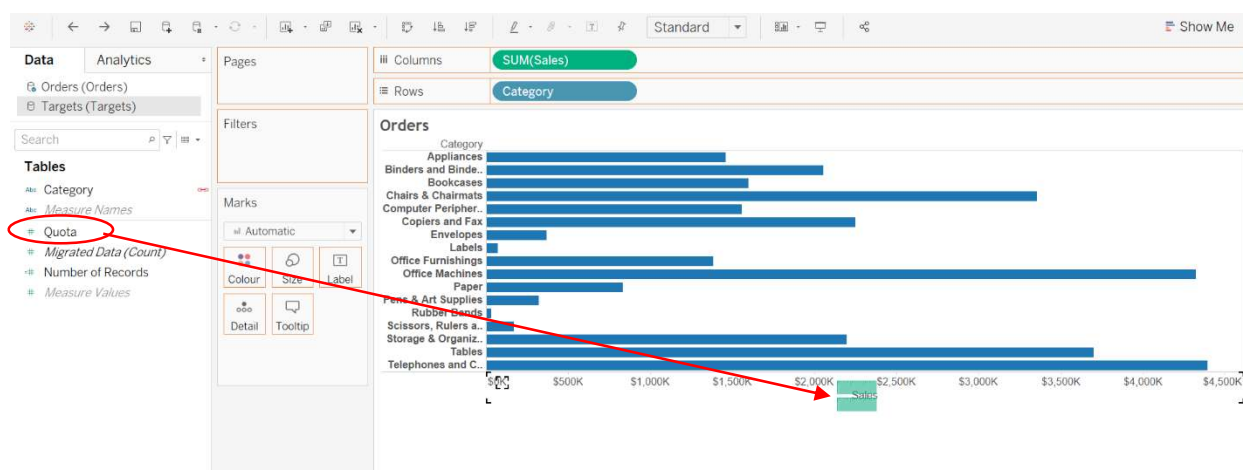
Notice that the Orders data source now has a blue check mark in the Data Pane, indicating that it is the primary data source.

8. On the **Orders** worksheet, in the **Data** pane, select the **Targets** data source.

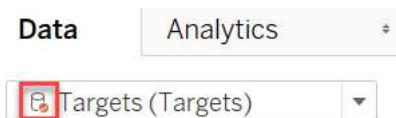


Notice that a vertical orange bar appears on the left side of the Data pane, indicating that this will be the secondary data source.

9. Drag **Quota** to the **Sales** axis, and drop it when two horizontal green bars appear.

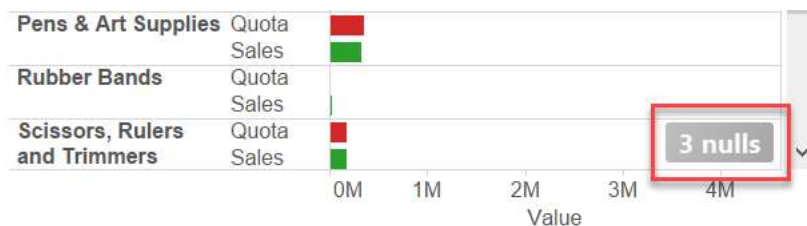


10. Notice that the **Targets** data source now has an orange check mark in the Data pane, indicating that it is the secondary data source.

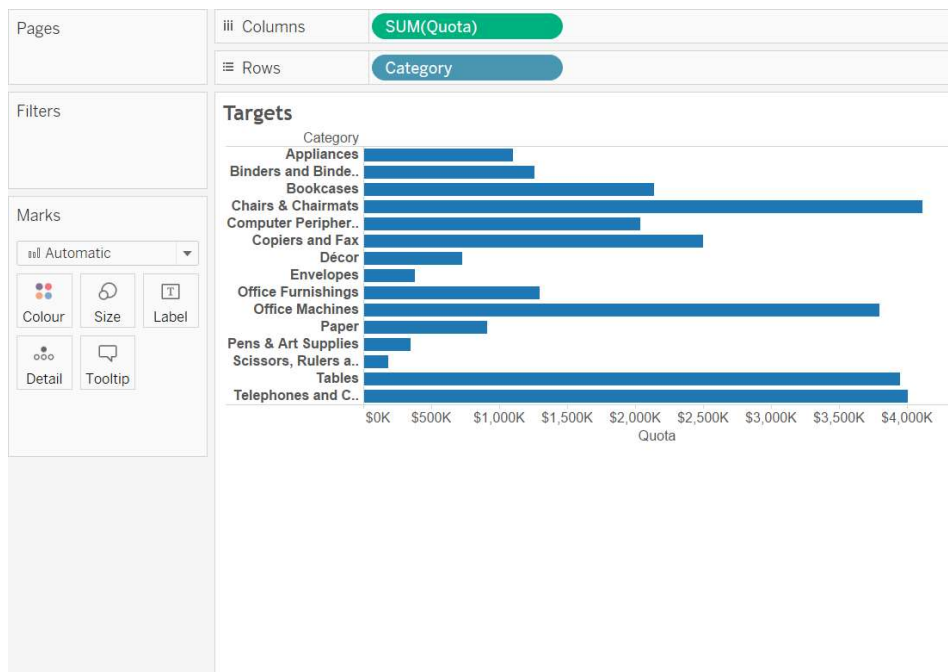


11. On the **Orders** worksheet, in the **Data** pane, select the **Orders** data source.

12. Drag **Measure Names** to **Colour** on the **Marks** card.
13. On the **Marks** card, click **Colour** → **Edit Colours**.
14. Select **Quota**, and then pick a different colour. Do the same for **Sales**. Click **OK**.
15. In the bottom right corner of the view, click the **null** indicator, and then click **Show data at default position**.

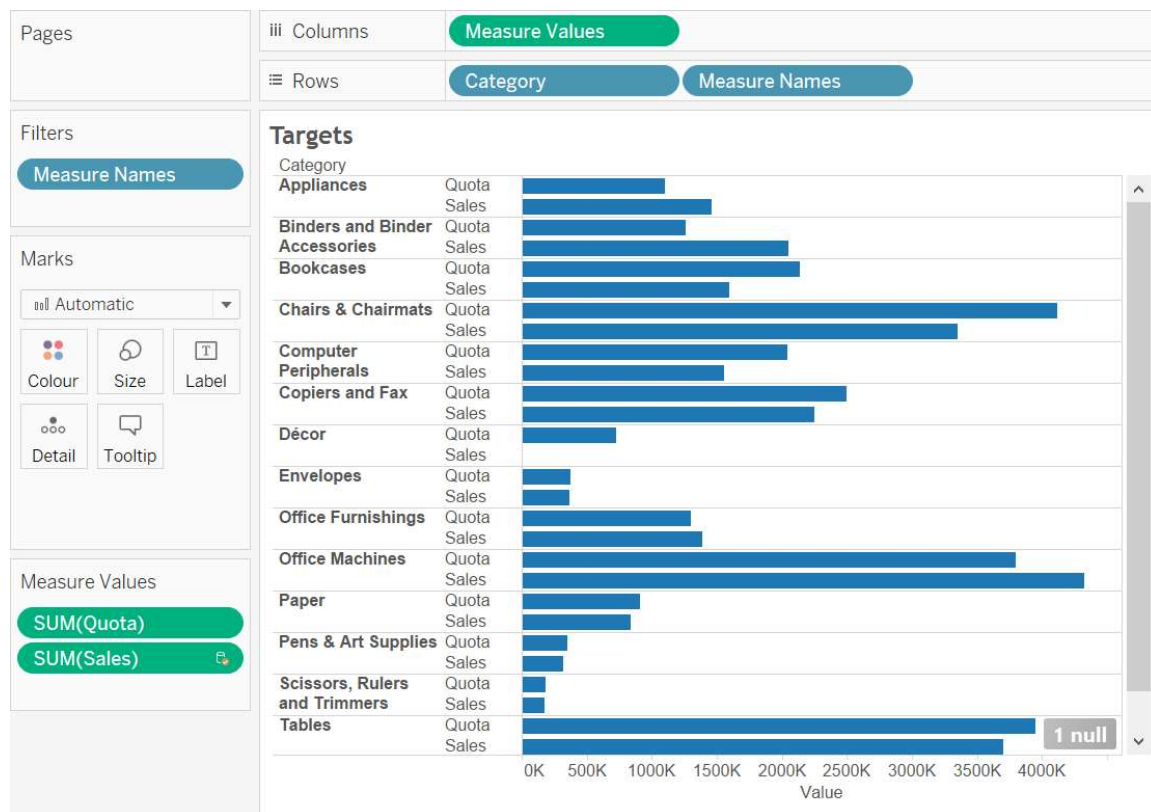


16. Click the **Targets** tab.
17. In the **Data** pane, select the **Targets** data source.
18. Drag **Quota** to the **Columns** shelf.
19. Drag **Category** to the **Rows** shelf.



20. On the **Targets** worksheet, in the **Data** pane, select the **Orders** data source.

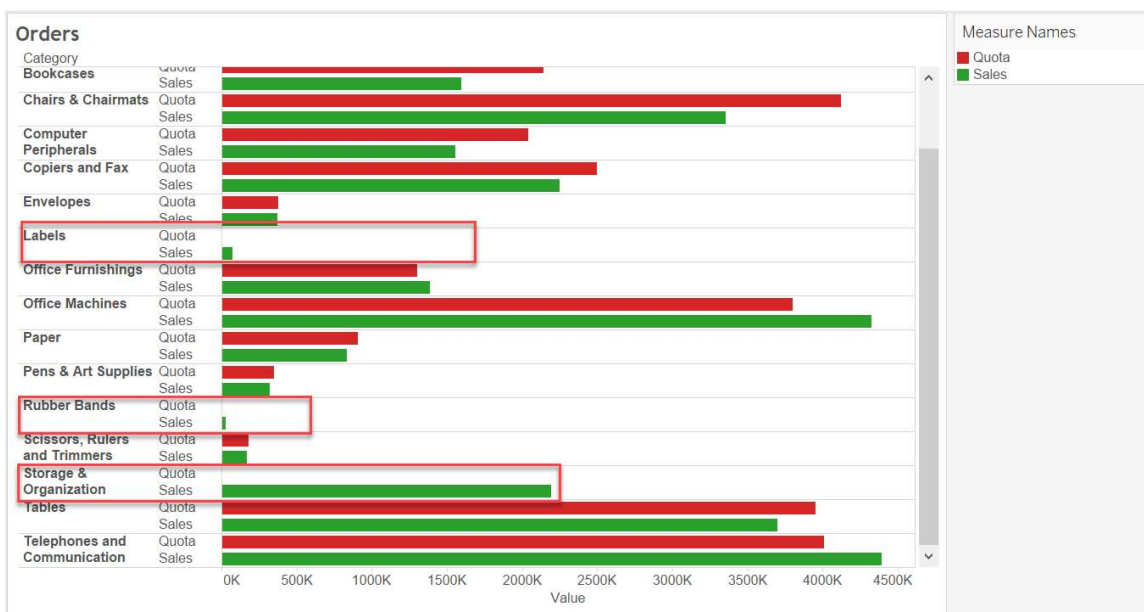
21. Drag **Sales** to the **Quota** axis, and drop it when two horizontal green bars appear.



22. On the **Targets** worksheet, in the **Data** pane, select the **Targets** data source.
23. Drag **Measure Names** to **Colour** on the **Marks** card
24. On the **Marks** card, click **Colour** → **Edit Colours**.
25. Select **Quota**, and then pick a different colour. Do the same for **Sales**. Click **OK**.
26. In the bottom right corner of the view, click the **null** indicator, and then click **Show data at default position**.

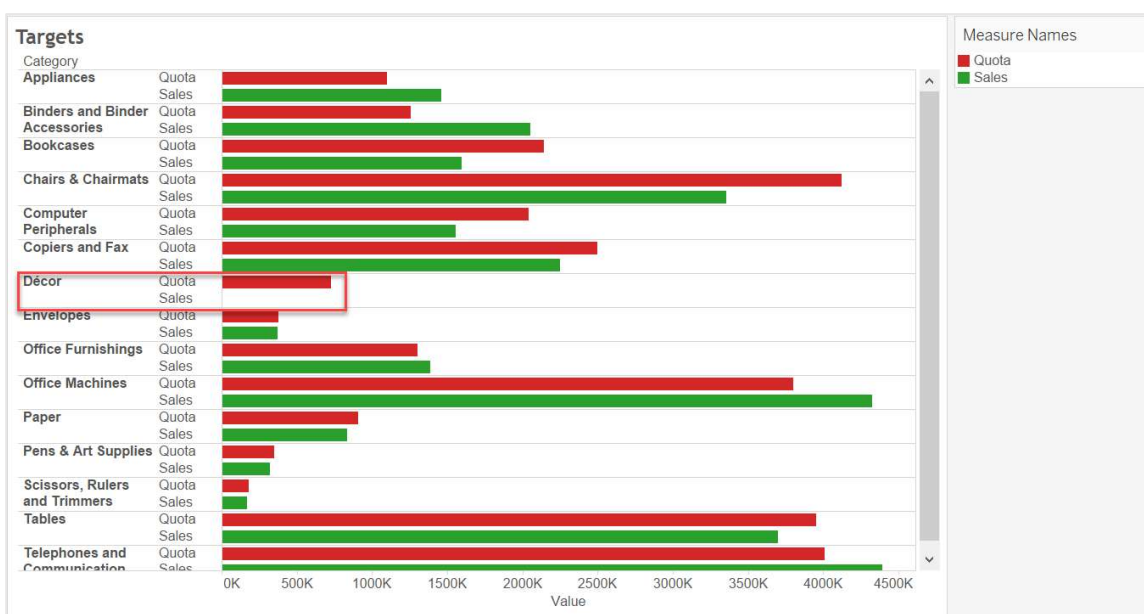
Results of the Orders and Targets view

1. Click the **Orders** tab.



Notice that the results include **Labels**, **Rubber Bands**, and **Storage & Organization**, which are categories present only in the Orders data source. They have no Quota bars because they are not present in the Targets data source.

2. Click the **Targets** tab.



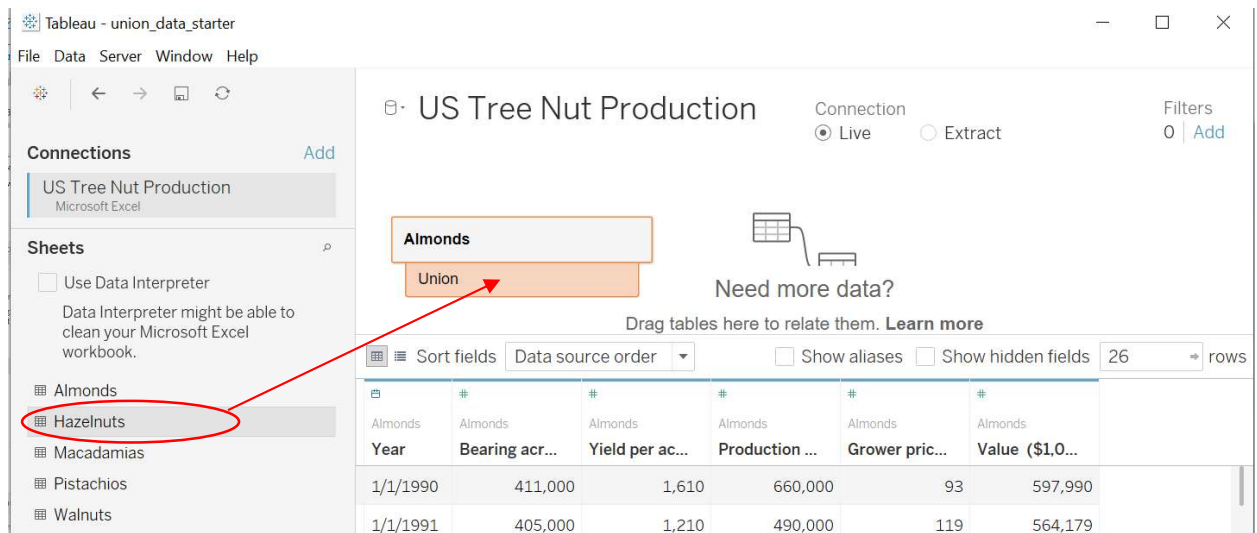
Notice that the results include **Décor**, which is present in the Targets data source but not in the Orders data source. It has no Sales bar because it is not present in the Orders data source.

Notice that the results do not include Labels, Rubber Bands, or Storage & Organization, which were present only in the Orders data source.

Task 5: Union Data and Merge Fields

In this exercise, you will compare agricultural data that was collected over a period of many years using one table. You will need to create a manual union and merge any mismatched fields.

1. Open the **union_tables_starter.twbx** workbook. Click the **Data Source** tab in the lower left corner.
2. Click and drag the **Almonds** table to the data area. Click and drag the **Hazelnuts** table to the same area, when you see the **Drag table to union** box, drop it directly below the Almonds until the **Union box** appears.



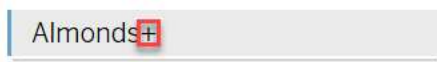
Once that is done, you should see the following:



Place your mouse cursor over it, you should see a down-arrow, **click** on the **down-arrow** and select **Open** from the pop-up menu.



On the canvas, notice the plus (+) sign at the end of the **Almonds** table.



3. On the canvas, click the drop-down arrow next to **Almonds +**, and select **Edit Union**.

4. From the left panel, under **Sheets**, drag the sheets: **Macadamia**, **Pistachios**, and **Walnuts** to the Union window, and drop each sheet inside the union box.



Click **OK**.

5. From the Metadata pane, click the **Yield per acre (pounds)** column to highlight it. Press the **Ctrl** key and click the **YPA in pounds** column.
6. With both columns highlighted, right-click one of them, and click **Merge Mismatched Fields**.
7. Widen the newly merged column to view the entire title.

Type	Field Name	Phys...
	Year	Almon...
	Bearing acreage (acres)	Almon...
	Yield per acre (pounds) & YPA in pounds	Project
	Production (1,000 pounds)	Almon...

8. Double-click the **Yield per acre (pounds) & YPA in pounds** column header to rename the column to **Yield per acre (pounds)**.

9. Repeat the same step to rename the **Sheet** column to **Tree nut variety**.
10. Click **Sheet1** tab at the left bottom corner to open the chart designer.
11. Drag **Year** to the **Columns** shelf.
12. Drag **Yield per acre (pounds)** to **Rows**.
13. Drag **Bearing acreage (acres)** to **Rows** to the left of the **Yield per acre** field.
14. Drag **Tree nut variety** to **Colour** on the **Marks** card.
15. In the **Fit** box, change the view to **Entire View**.

