

Lesson 05: Dataset Alchemy Handling Multiple Datasets for Rich Insights

Overview

In this exercise, you will learn to use joins and unions. The lesson focuses on practical application with real-world scenarios. You will explore how joins and unions can effectively summarize data. The hands-on approach aims to develop proficiency in using them for efficient decision-making.

Instructions

- Use **sales_data.csv**, **sales_data_store_b.csv**, **customer_feedback.csv** and **product_information.csv** for your reference
- Work through all questions individually or within your group
- Divide your time evenly among the questions to ensure comprehensive understanding
- Utilize tableau official documentation and online resources to enhance your problem-solving abilities

Tasks

You are a retail chain analyst who uses unions in Tableau to consolidate sales data from multiple stores and joins to combine customer feedback with product information for comprehensive sales performance and satisfaction analysis.

Task 1

Consolidate Sales Records: Utilize Tableau's union feature to combine the input CSV files containing monthly sales data from each store, creating a unified dataset for total sales analysis.

- Choose **Text File** and open **sales_data.csv** and **sales_data_store_b.csv**
- Click on New Union option
- Drag the **sales_data** and **sales_data_store_b** into the union table that appeared after choosing New Union and click on **Apply**
- Go to the sheet to start analyzing the data. Drag **Date** to columns, **TotalSalesAmount** (aggregated as SUM) to rows, and select the bar chart to see total sales over time
- Click on the '+' sign to split the date into quarter, months, and date

Task 2

Merge Product Feedback with Catalog: Apply a join in Tableau between the customer feedback dataset and the product catalog dataset on the product ID, facilitating an analysis of customer satisfaction across different product categories.

- Drag **product_information.csv** to the workspace
- Double click on **customer_feedback.csv**. This will automatically create an inner join between two tables.
- Navigate to a new worksheet and drag **Category** to columns and **FeedbackScore** (aggregated as AVG) to rows to analyze average feedback scores by product category

Discussion Questions (Optional)

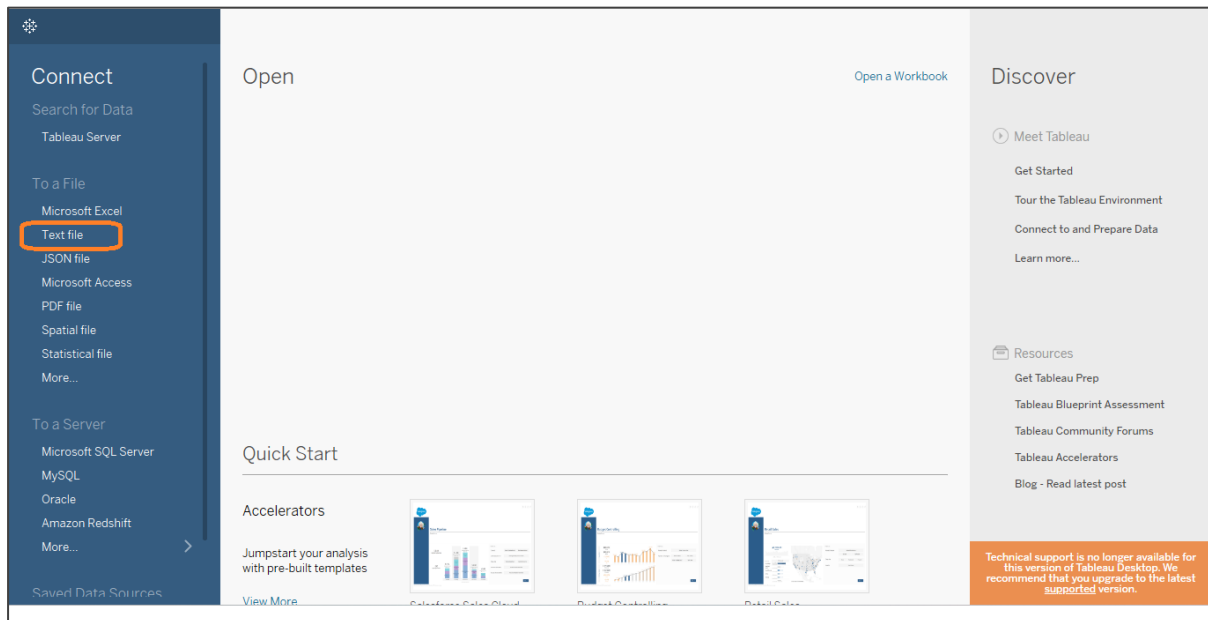
If time permits, discuss the following questions:

- How does the integration of sales data from multiple sources affect decision-making in retail management?
- What are the potential challenges and solutions in merging customer feedback data with product information for a holistic analysis?

Answer Key

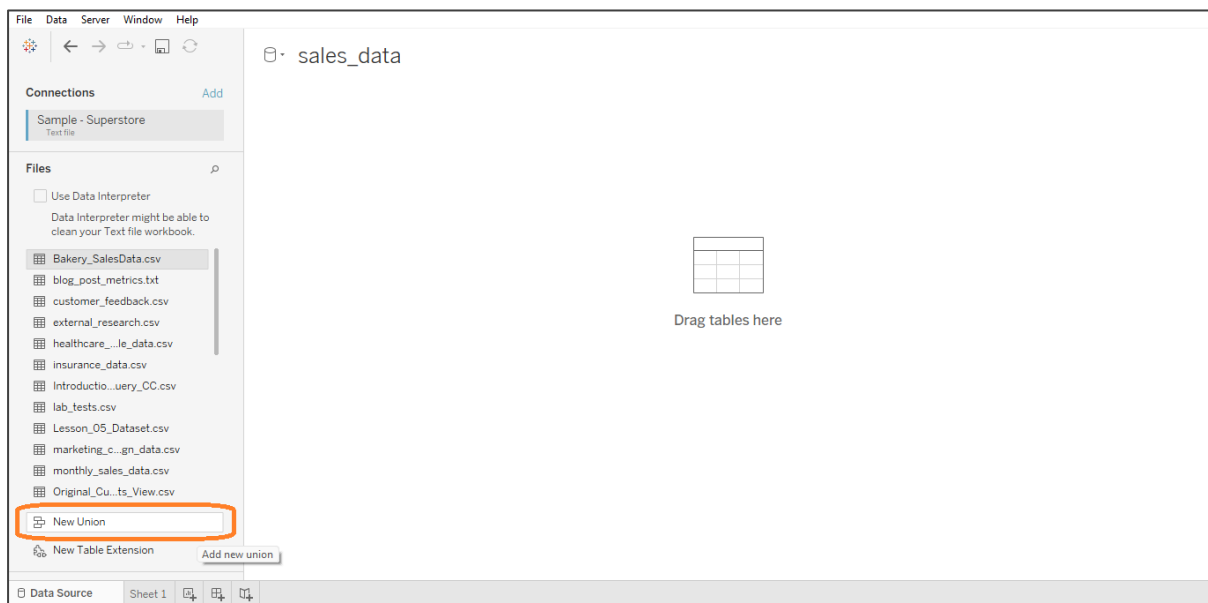
Task 1

- Choose **Text File** and open **sales_data.csv** and **sales_data_store_b.csv**

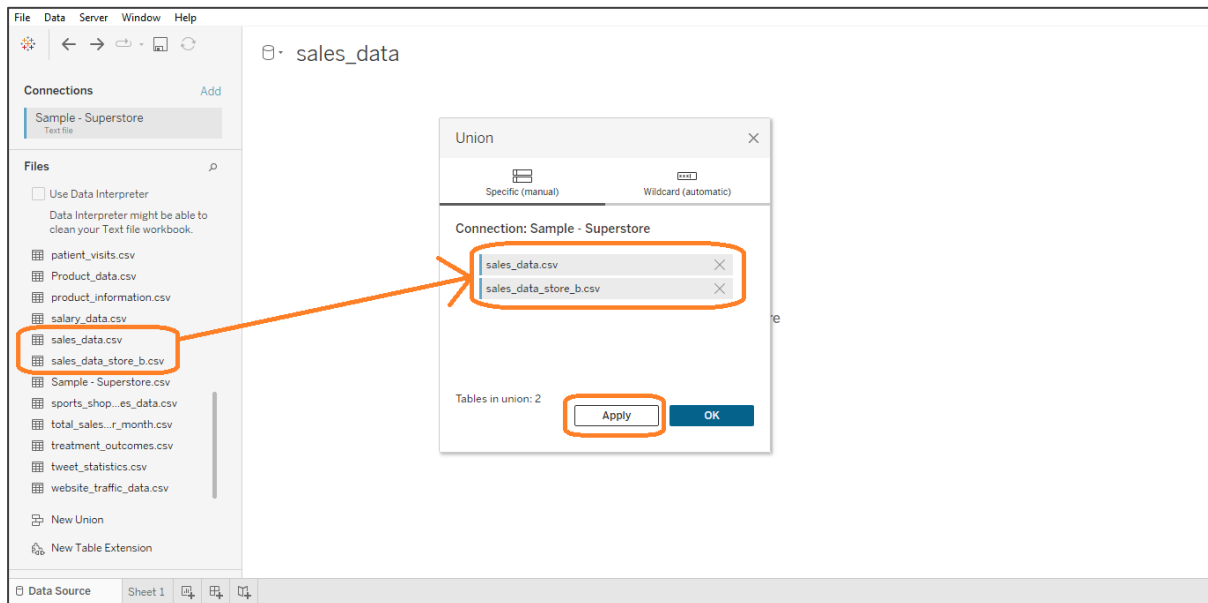


Note: In the data pane, you will see both datasets are listed.

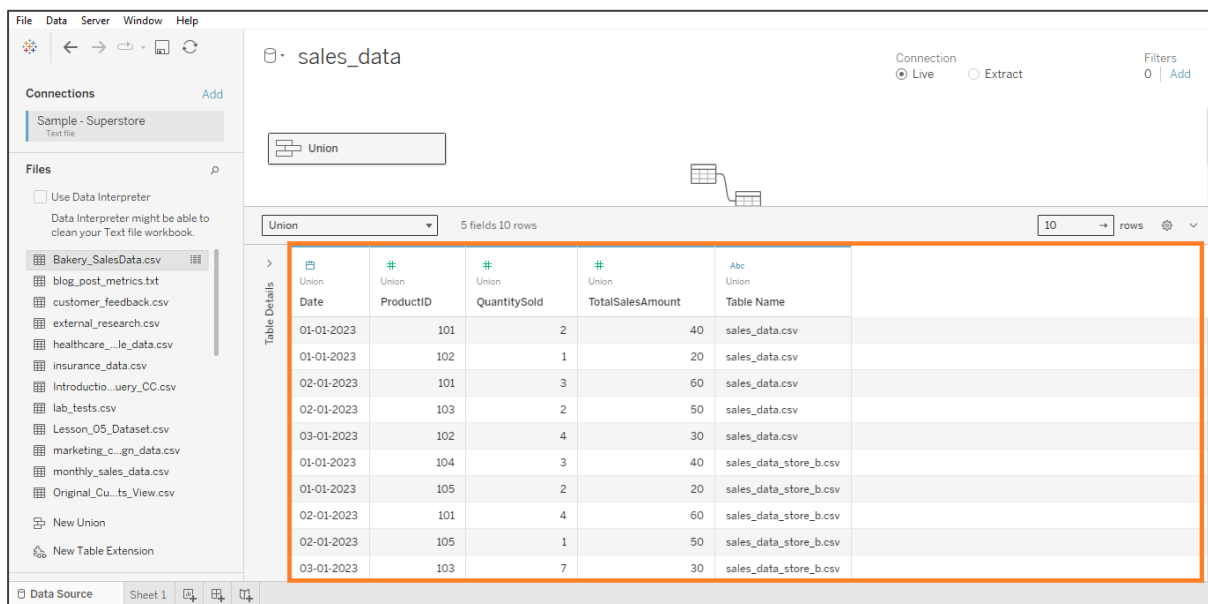
- Click on New Union option



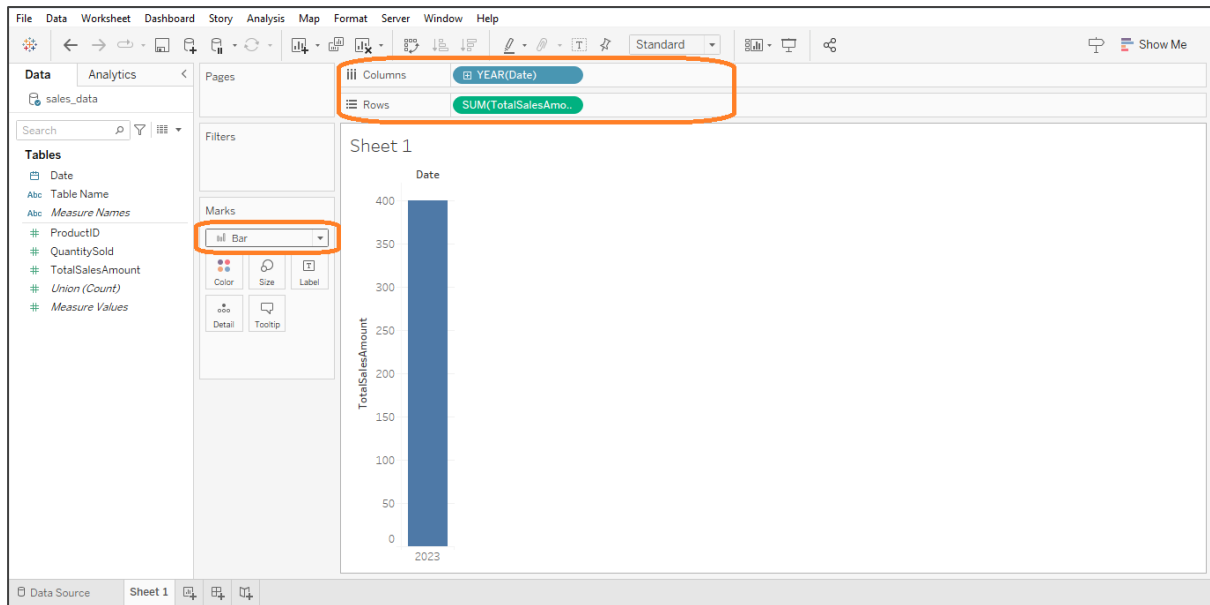
- Drag the **sales_data** and **sales_data_store_b** into the union table that appeared after choosing New Union and click on **Apply**



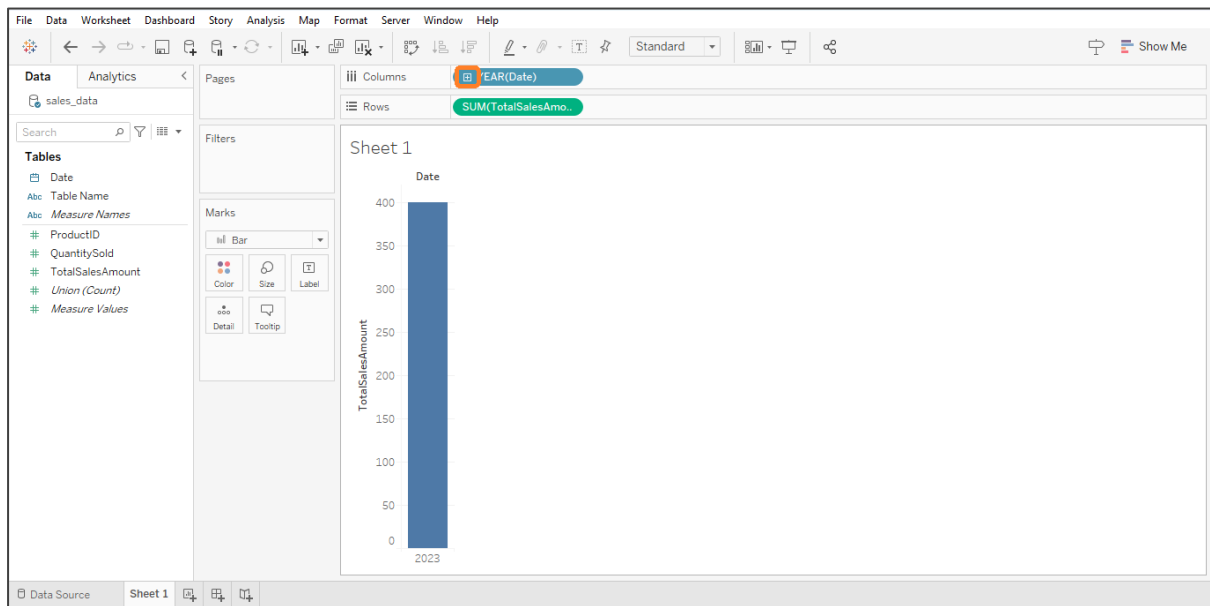
The output appears as shown below:



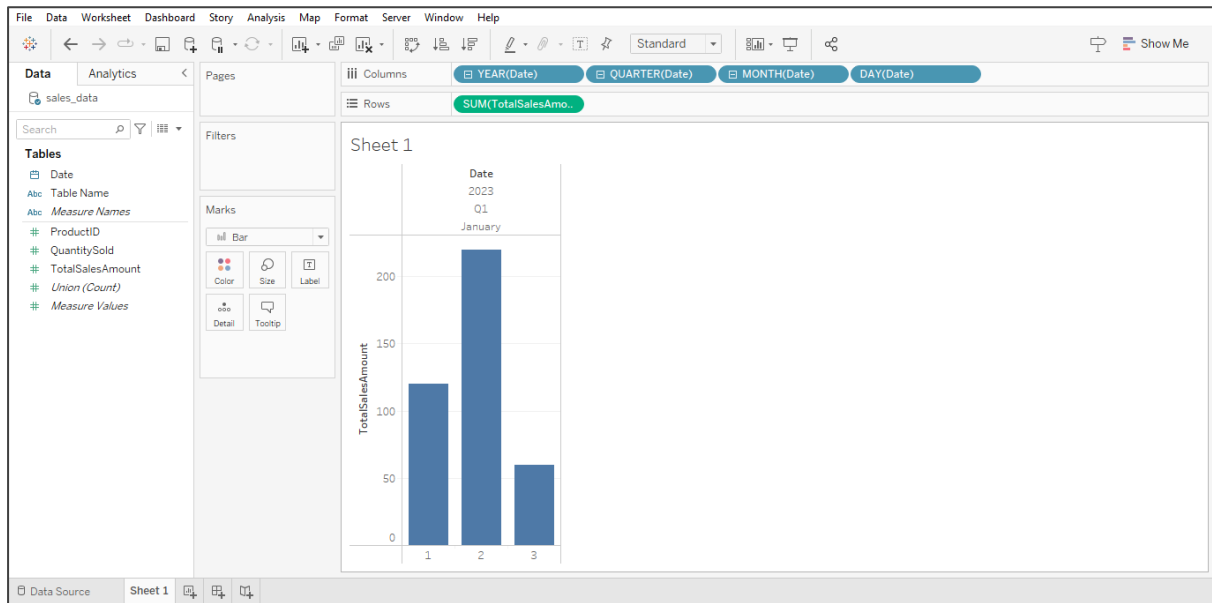
- Go to the sheet to start analyzing the data. Drag **Date** to columns, **TotalSalesAmount** (aggregated as SUM) to rows, and select a bar chart to see the total sales over time.



- Click on the '+' sign to split the date into quarter, months, and date



The output appears as shown below:



Task 2

Note: Go to Data Source tab and add **customer_feedback.csv** and **product_information.csv** to workspace using the **connect to data** option

- Drag **product_information.csv** to the workspace

The screenshot shows the Tableau Desktop interface. The 'Connections' pane on the left lists 'product_information.csv' as a text file. The 'Files' pane on the left lists various CSV files. The 'product_information.csv' file is highlighted in the 'Files' pane. The 'product_information.csv' file is also listed in the 'Connections' pane. The 'product_information.csv' file is highlighted in the 'Files' pane. The 'product_information.csv' file is highlighted in the 'Files' pane.

product_information

Connection: ☒ Live ☐ Extract

Filters: 0 | Add

product_information.csv

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

product_information.csv 4 fields 6 rows

#	product_information.csv	product_information.csv	product_information.csv	product_information.csv
Product ID	Product Name	Category	Price	
101	Television	Electronics	20.00000	
102	Bed	Home Goods	20.00000	
103	Refrigerator	Electronics	25.00000	
104	Shirt	Clothing	20.00000	
105	Utensils	Home Goods	22.50000	

- Double click on **customer_feedback.csv**. This will automatically create an inner join between two tables.

The screenshot shows the Power BI interface with the 'product_information+' workspace. In the 'Files' pane on the left, 'customer_feedback.csv' is highlighted with an orange box. The main area displays a relationship diagram with 'product_information.csv' and 'customer_feedback.csv' connected by an inner join line. Below the diagram, a preview of the joined data is shown with 5 rows. The columns are 'Customer ID', 'ProductID (customer fe...', and 'Feedback Score'.

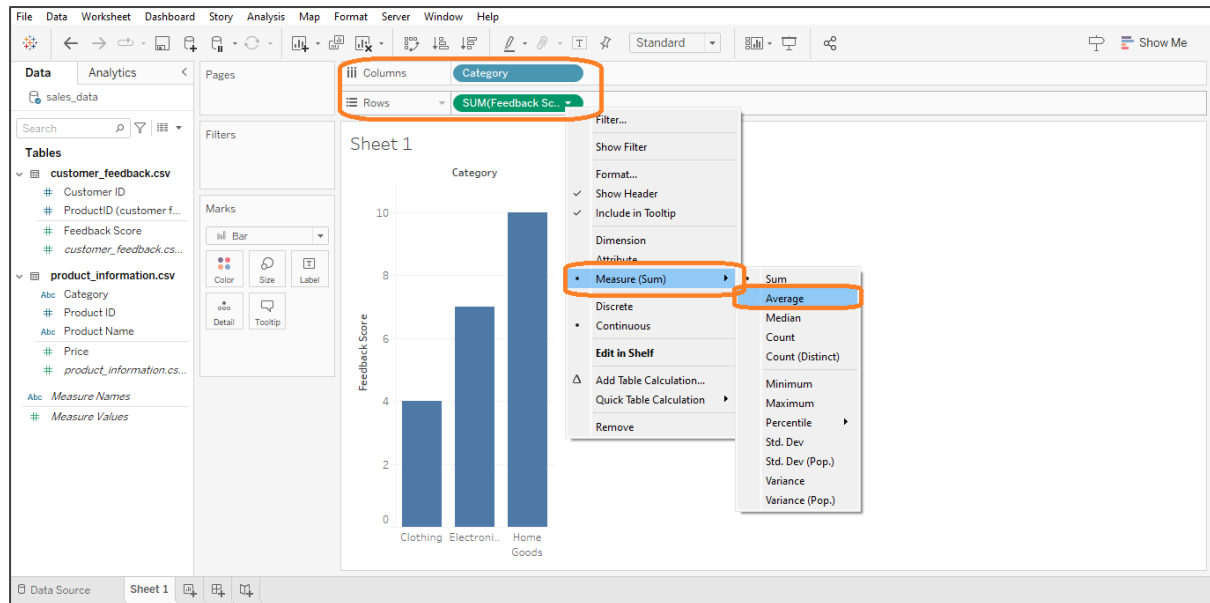
Customer ID	ProductID (customer fe...	Feedback Score
1	101	4
2	102	5
3	103	3
4	104	4
5	105	5

The output table will appear as shown below:

The screenshot shows the same Power BI interface, but now the output table is visible and highlighted with an orange box. The table has 5 rows and 3 columns: 'Customer ID', 'ProductID (customer fe...', and 'Feedback Score'. The data is as follows:

Customer ID	ProductID (customer fe...	Feedback Score
1	101	4
2	102	5
3	103	3
4	104	4
5	105	5

- Navigate to a new worksheet and drag **Category** to columns and **FeedbackScore** (aggregated as AVG) to rows to analyze average feedback scores by product category



The output appears as shown below:

