



SUPER STORE SALES ANALYSIS WITH TABLEAU

Project Hand-out, Faculty Development Program – Naan Mudhalvan

Super Store Sales Analysis with Tableau

Introduction:

A superstore is a type of retail store that offers a wide variety of products and services, typically in a large, warehouse-like setting. These stores are often found in suburban areas and are designed to meet the needs of a wide range of customers. Superstores typically carry a wide range of products, including groceries, household items, electronics, clothing, and more. They also often have additional services such as pharmacies, optical centers, and banks. Superstores are known for their low prices and wide selection of products, and they are often considered to be a one-stop-shop for many consumer needs. Some popular examples of superstore chains include Walmart, Vmart, DMart, and Jiomart.

The Superstore dataset is a sample dataset provided by Tableau that contains information on sales of products from a fictional retail store called "Superstore." The data includes information on customer demographics, sales by product category, and other information that can be used to analyze and understand the performance of the store. The dataset includes fields such as Order Date, Ship Date, Profit, Sales, Quantity, Discount, etc. which can be used to create different types of visualization and perform analysis in Tableau.

Our task is to analyse the sales data and identify weak areas and opportunities for Super Store to boost business growth.

Business Questions

- Which Category is Best Selling and Most Profitable?
- What are the Best Selling and Most Profitable Sub-Category?
- Which is the Top Selling Sub-Category?
- Which Customer Segment is Most Profitable?
- Which is the Preferred Ship Mode?
- Which Region is the Most Profitable?
- Which City has the Highest Number of Sales?

Who is Reading this Analysis?

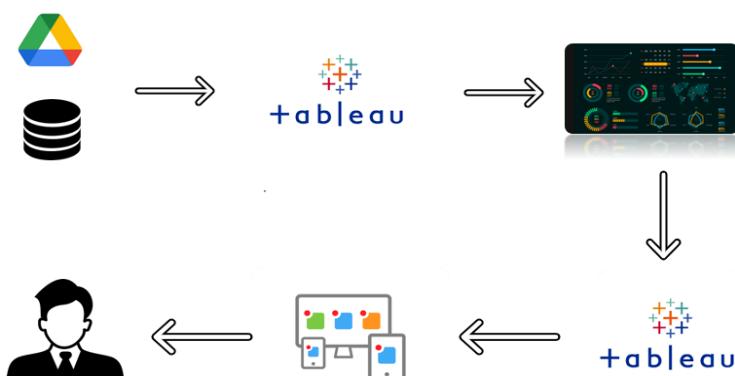
- We assume that the Super Store is a family business and is owned by 1 or 2 owners who are very invested in their business.
- Owners are likely not people who are expert in reading charts or interpreting statistical analysis, so our analysis will be in layman terms and easy to understand.

Project Flow

To accomplish this, we have to complete all the activities listed below,

- Define Problem / Problem Understanding
 - Specify the business problem
 - Business requirements
 - Literature Survey
 - Social or Business Impact.
- Data Collection & Extraction from Database
 - Collect the dataset,
 - Storing Data in DB
 - Perform SQL Operations
 - Connect DB with Tableau
- Data Preparation
 - Prepare the Data for Visualization
- Data Visualizations
 - No of Unique Visualizations
- Dashboard
 - Responsive and Design of Dashboard
- Story
 - No of Scenes of Story
- Performance Testing
 - Amount of Data Rendered to DB
 - Utilization of Data Filters
 - No of Calculation Fields
 - No of Visualizations/ Graphs
- Web Integration
 - Dashboard and Story embed with UI With Flask
- Project Demonstration & Documentation
 - Record explanation Video for project end to end solution
 - Project Documentation-Step by step project development procedure

Technical Architecture:



Milestone 1: Define Problem / Problem Understanding

Activity 1: Specify the business problem

Refer Project Description

Activity 2: Business requirements

Business requirements for a retail store like Superstore would include tracking sales and revenue data, having an effective inventory management system, understanding customer demographics, tracking sales by product category, having a sales and marketing strategy, having an efficient supply chain management system, having a robust fulfillment and delivery system, and having a dedicated customer service team. These requirements are necessary to ensure that the store is able to meet the needs of its customers, understand its performance and target the right audience. Specific requirements may vary depending on the business model, size and the target market of the store.

Activity 3: Literature Survey (Student Will Write)

A literature survey for Superstore would involve reviewing existing research on topics related to the retail industry such as sales and revenue analysis, inventory management, customer demographics, product categories, sales and marketing strategies, supply chain management, fulfillment and delivery, and customer service. The survey would involve defining a research question, identifying relevant sources, reading and analyzing the literature, organizing and summarizing the literature, identifying gaps in the literature, and providing recommendations for Superstore based on the findings. The goal of the literature survey is to understand the current state of knowledge on a topic and to identify areas where further research is needed. It also helps to build upon the existing knowledge and avoid duplication of effort.

Activity 4: Social or Business Impact.

Social Impact:Creating jobs and economic opportunities for community members.Supporting local suppliers and promoting local products

Business Model/Impact: Higher market share and brand recognition,Improved customer satisfaction and loyalty

Milestone 2: Data Collection & Extraction from Database

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes and generate insights from the data.

Activity 1: Collect the dataset

Please use the link to download the dataset: [Link](#)

Activity 1.1: Understand the data

Data contains all the meta information regarding the columns described in the CSV files. we have provided 5 CSV files:

1. sample-superstore

2. superstore-people
3. superstore-returns

Column Description for sample-superstore:

1. Order Date: The date that the order was placed.
2. Ship Date: The date that the order was shipped.
3. Ship Mode: The shipping method used for the order (e.g. Regular Air, Next Day Air, etc.).
4. Customer ID: A unique identifier for the customer who placed the order.
5. Customer Name: The name of the customer who placed the order.
6. Segment: The segment that the customer belongs to (e.g. Consumer, Corporate, etc.).
7. City: The city where the customer is located.
8. State: The state where the customer is located.
9. Postal Code: The postal code of the customer's address.
10. Product ID: A unique identifier for the product that was ordered.
11. Category: The category that the product belongs to (e.g. Furniture, Office Supplies, etc.).
12. Sub-Category: A more specific category that the product belongs to (e.g. Chairs, Paper, etc.).
13. Sales: The total sales for the order.
14. Quantity: The number of units of the product that were ordered.
15. Discount: The percentage discount that was applied to the order.
16. Profit: The profit made from the order.
17. Country: The country where customer is located
18. Region: The region of the country where the customer is located (e.g. West, East, South, etc.).
19. Order ID: A unique identifier for the order.

Column Description for superstore-people:

1. Region: The region of the country where the customer is located (e.g. West, East, South, etc.).
2. Regional Manager: Each region could have a regional manager who is responsible for managing the sales and operations of the Superstore locations within that region.

Column Description for superstore-returns:

1. Order ID: The unique identifier for the original order
2. Returned: Customer returned the order(e.g. Yes,No)

Activity 2: Storing Data in DB & Perform SQL Operations

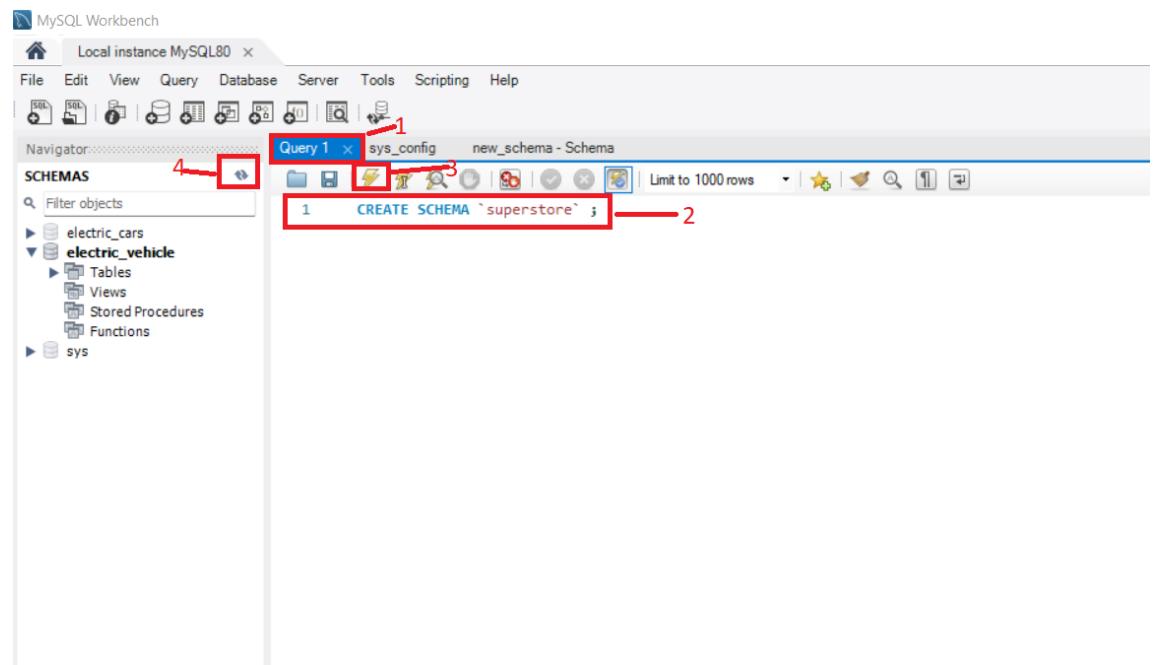
Go to My SQL Workbench

Go to new query sheet named Query 1 by default

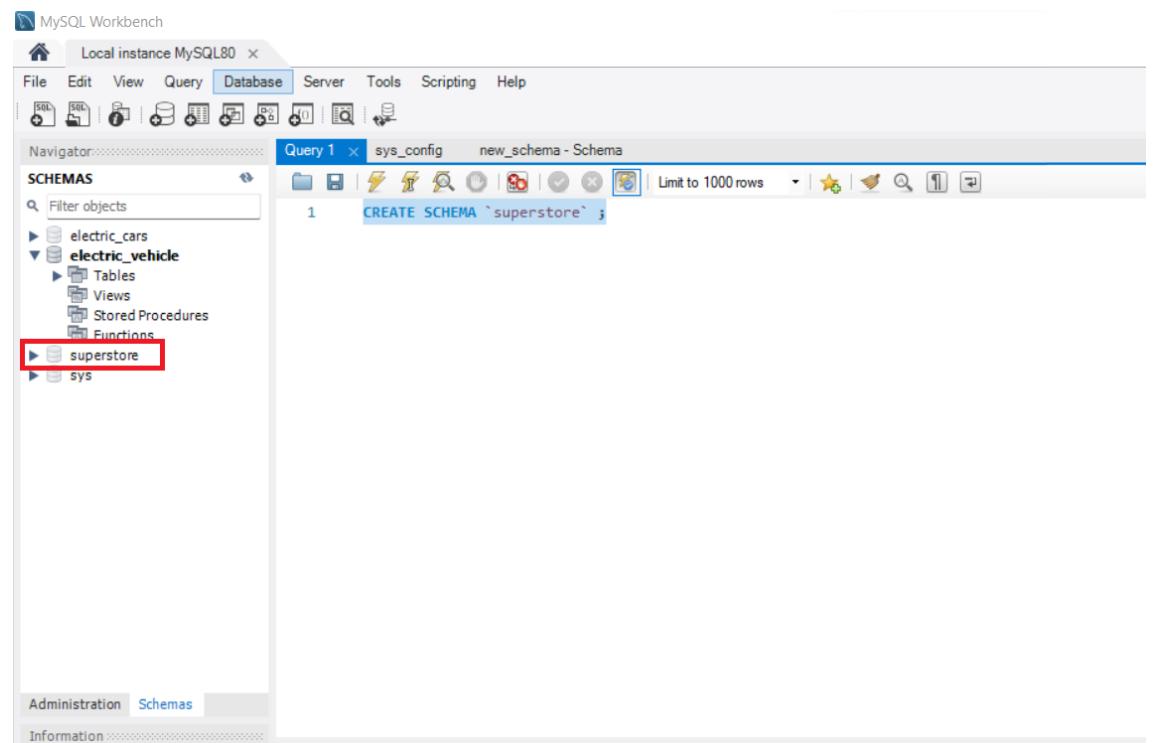
Write command to create new database schema “ CREATE SCHEMA `superstore` ; ”

Run the Query

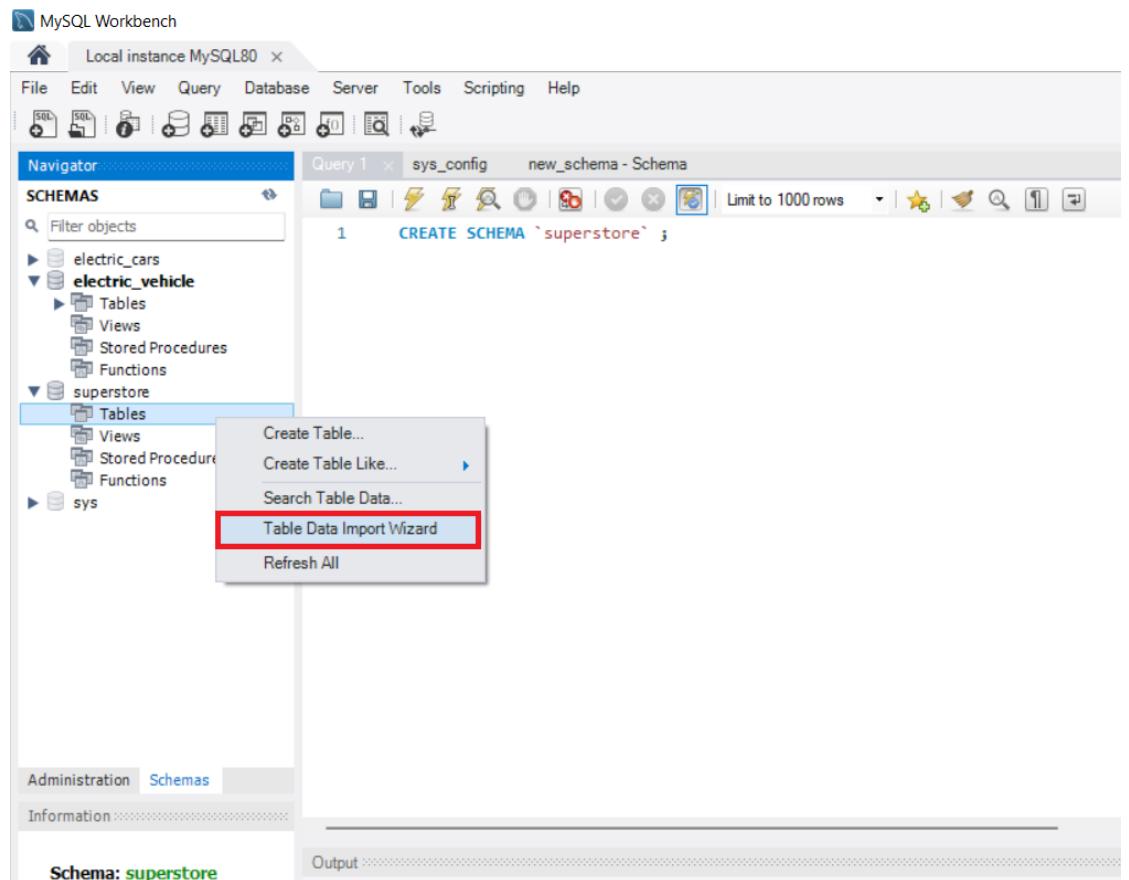
Refresh the schemas



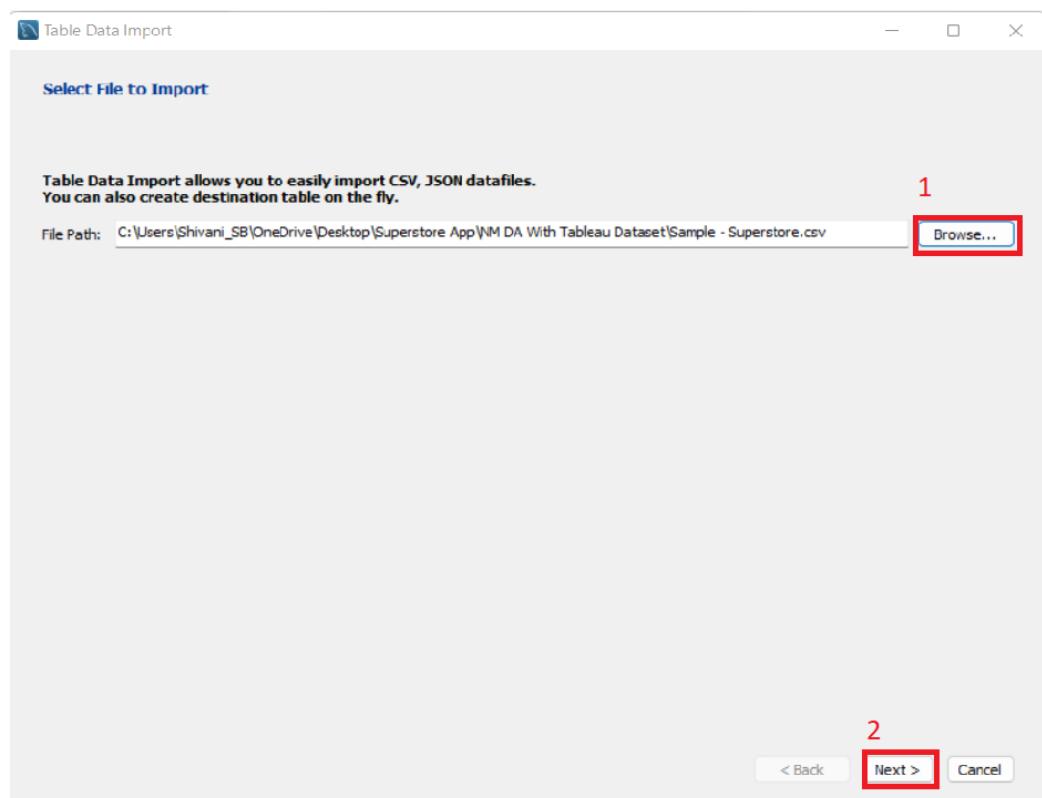
After refreshing you'll get new schema created with name "superstore"



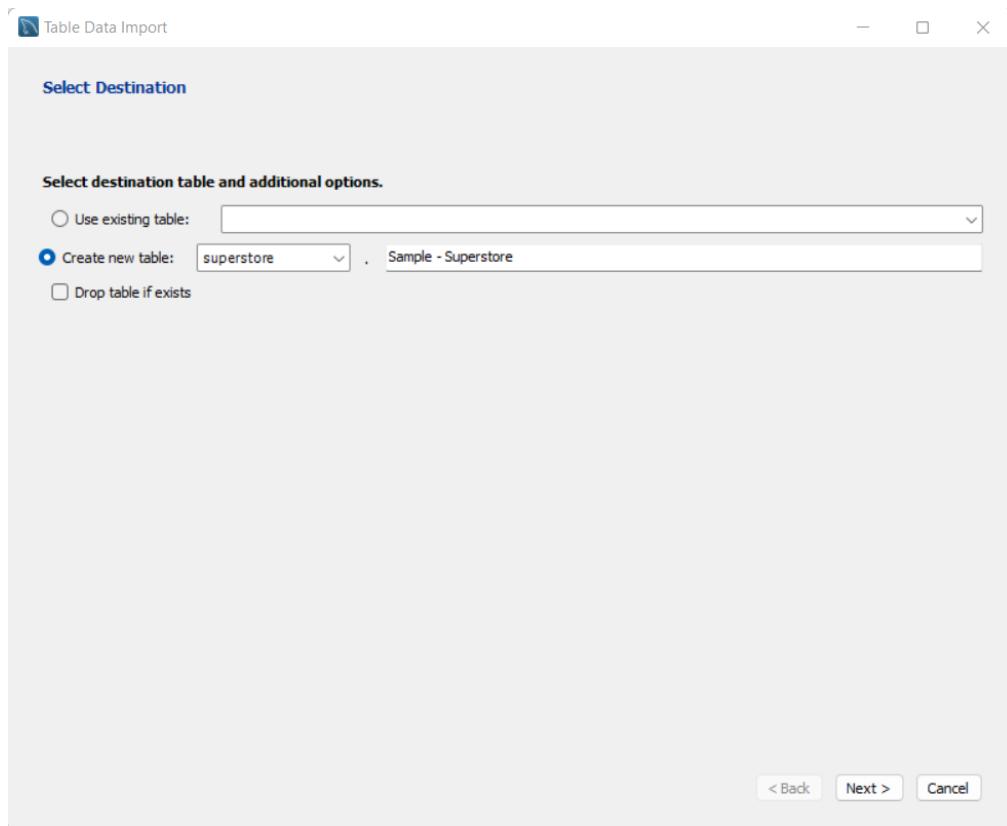
Now go to the dropdown of newly created schema, go to tables right click on it and select Table Data Import Wizard



Now browse csv file you want to upload as a table in your database an



We want to create a new table, so we will choose to create a new table and click on next.



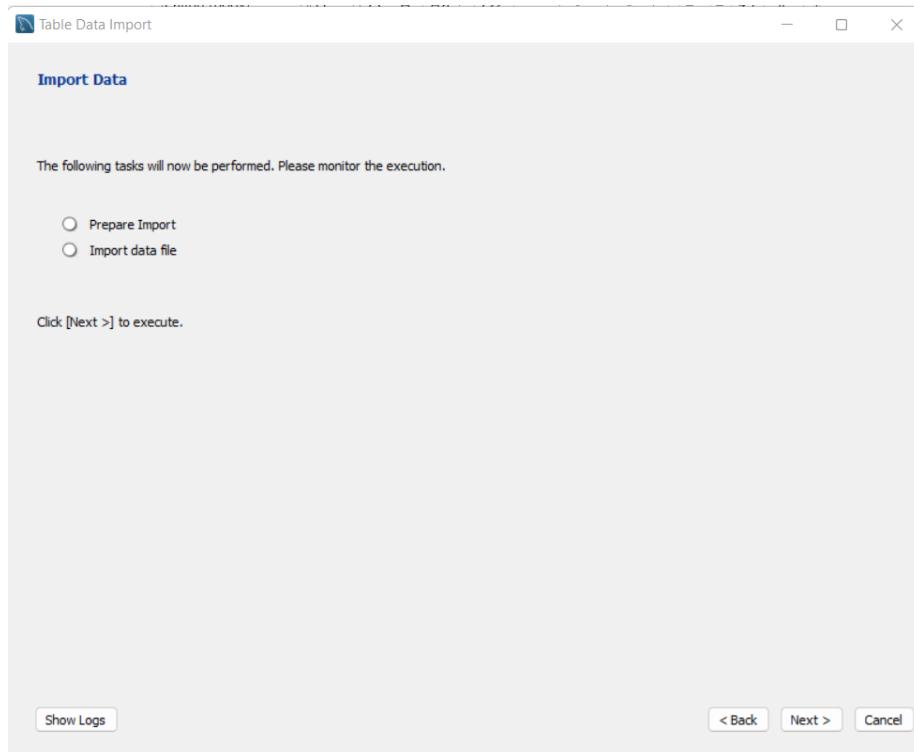
Now you can configure import setting if you want to like encoding and datatype of any particular column or if you want to unselect any column that you don't want to be there in the data table, here I've not made any changes just click on Next

This screenshot shows the 'Configure Import Settings' step of the import process. At the top, it says 'Detected file format: csv' with a file type icon. Below that, the 'Encoding:' dropdown is set to 'utf-8'. Under the heading 'Columns:', there is a table with two columns: 'Source Column' and 'Field Type'. All seven columns ('i>Row ID', 'Order ID', 'Order Date', 'Ship Date', 'Ship Mode', and 'Customer ID') have their checkboxes checked and are mapped to 'text' field types. Below this table is a preview of the data in a grid format:

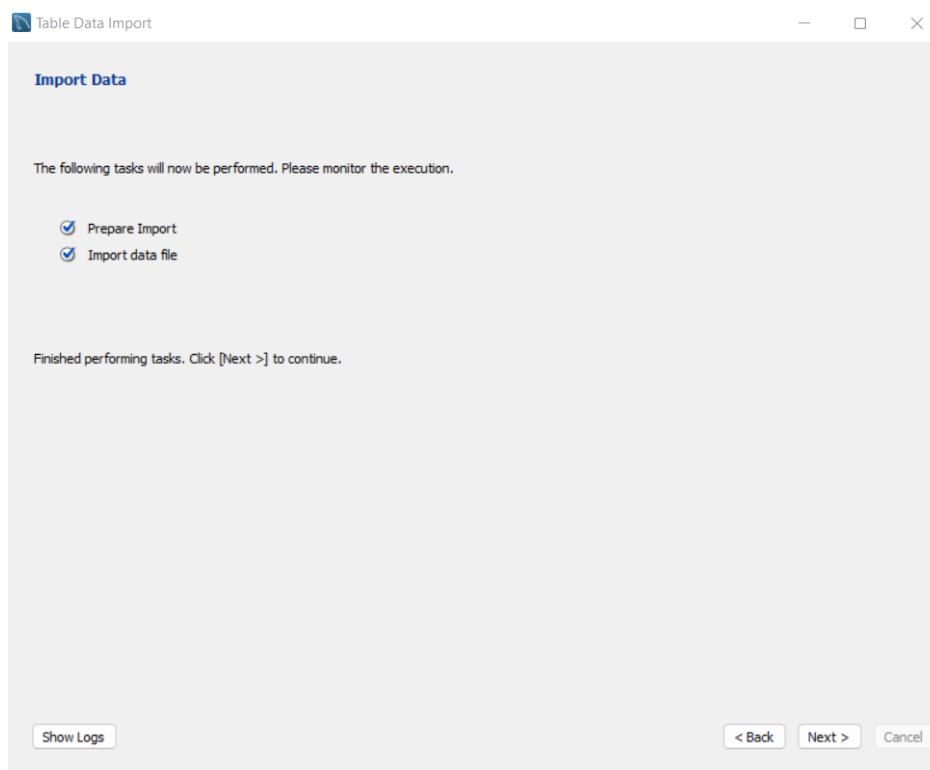
i>Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer ...	Segment	Country/R...	City
1	US-2019-10...	03-01-2019	07-01-2019	Standard Cl...	DP-13000	Darren Pow...	Consumer	United States	Houston
2	US-2019-11...	04-01-2019	08-01-2019	Standard Cl...	PO-19195	Phillina Ober	Home Office	United States	Naperville
3	US-2019-11...	04-01-2019	08-01-2019	Standard Cl...	PO-19195	Phillina Ober	Home Office	United States	Naperville
4	US-2019-11...	04-01-2019	08-01-2019	Standard Cl...	PO-19195	Phillina Ober	Home Office	United States	Naperville

At the bottom right of the window are three buttons: '< Back', 'Next >', and 'Cancel'.

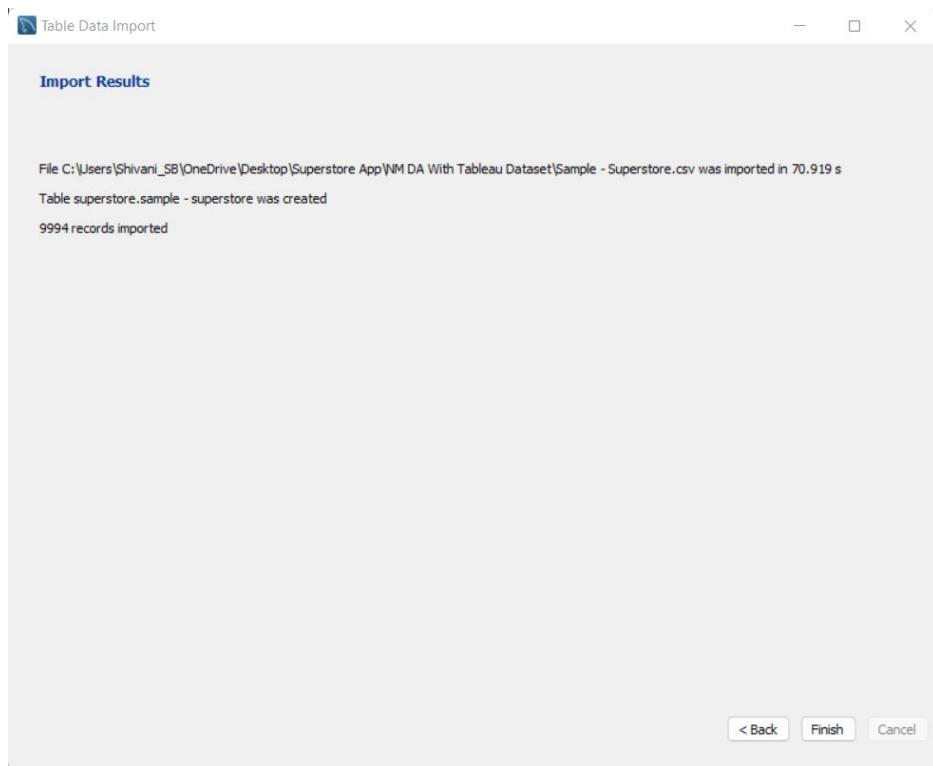
Now you can see the operations that will be performed while your data is being imported just click on Next to start importing, It might take a few minutes.



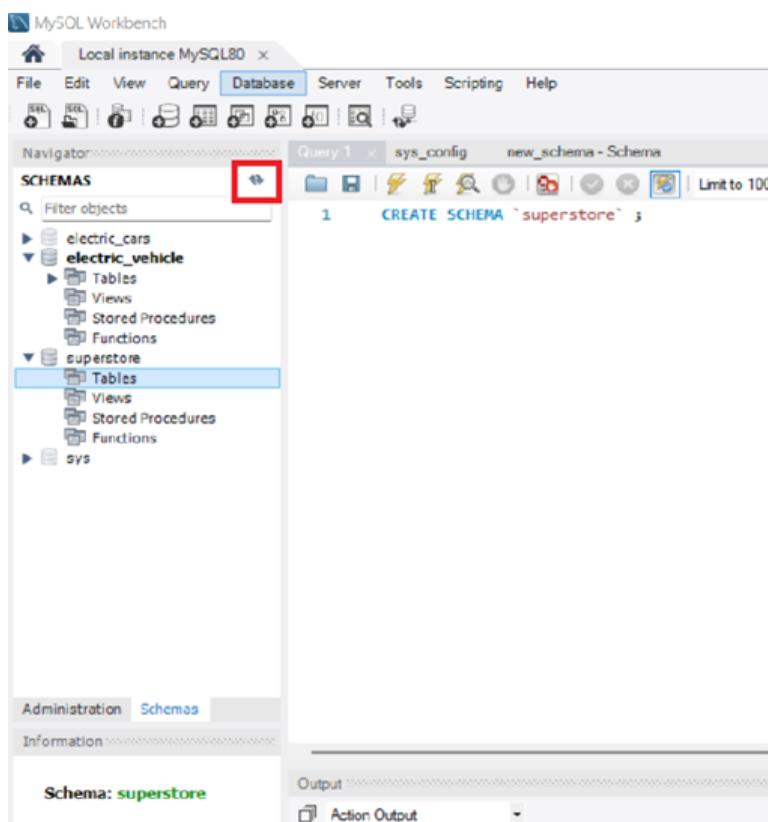
After import is done you can see both operations will be checked now just click Next to see import results



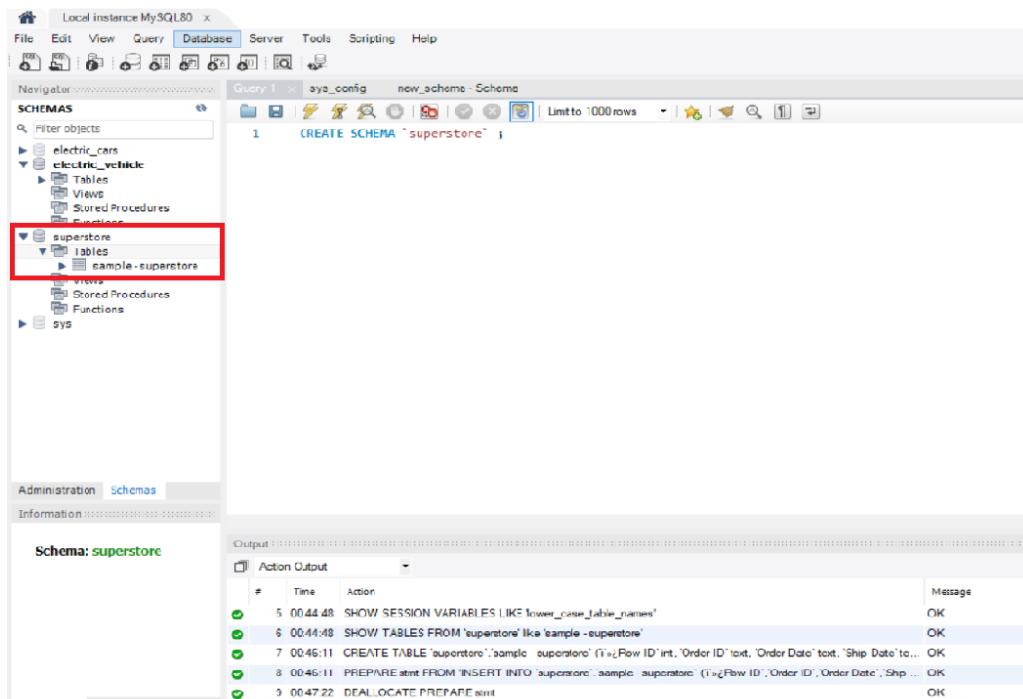
Now you can see the import results, click on Finish



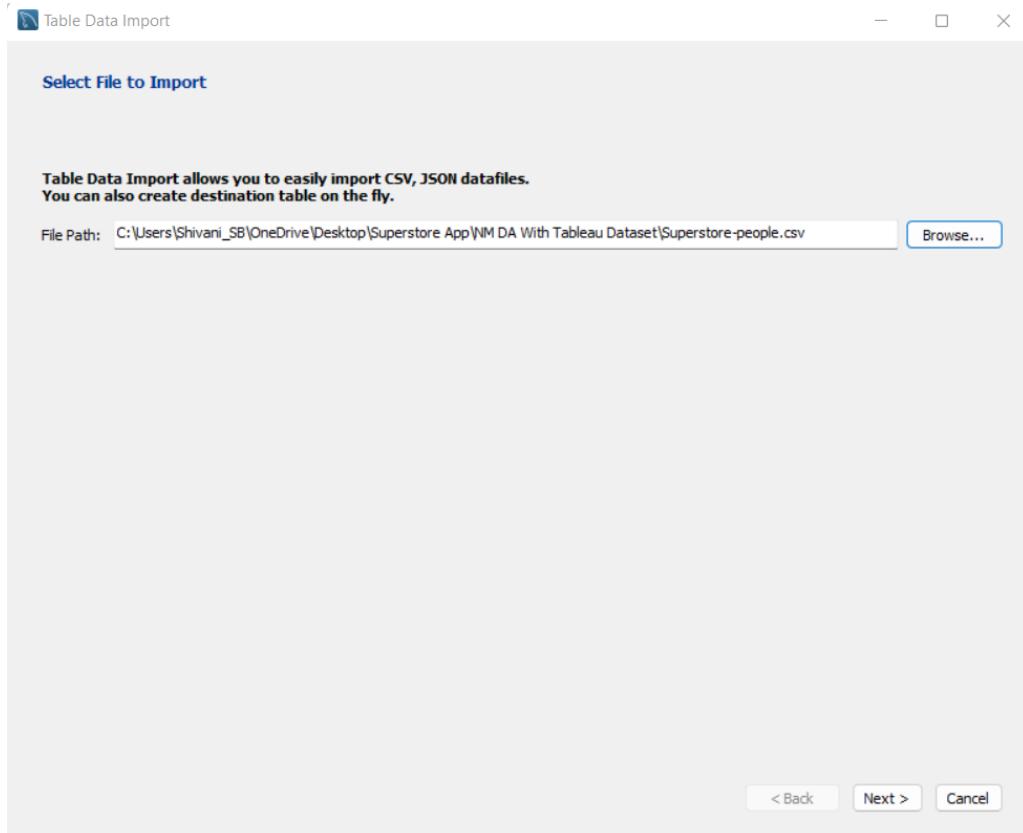
Now you just need to refresh Schemas

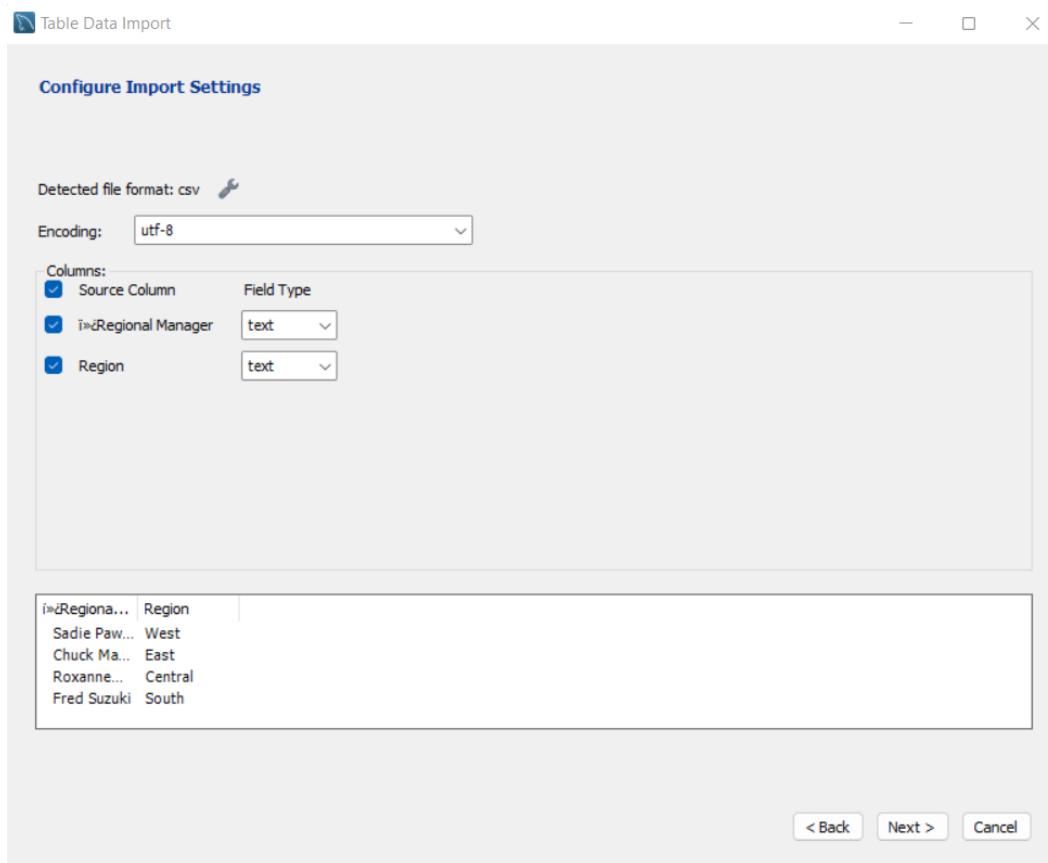
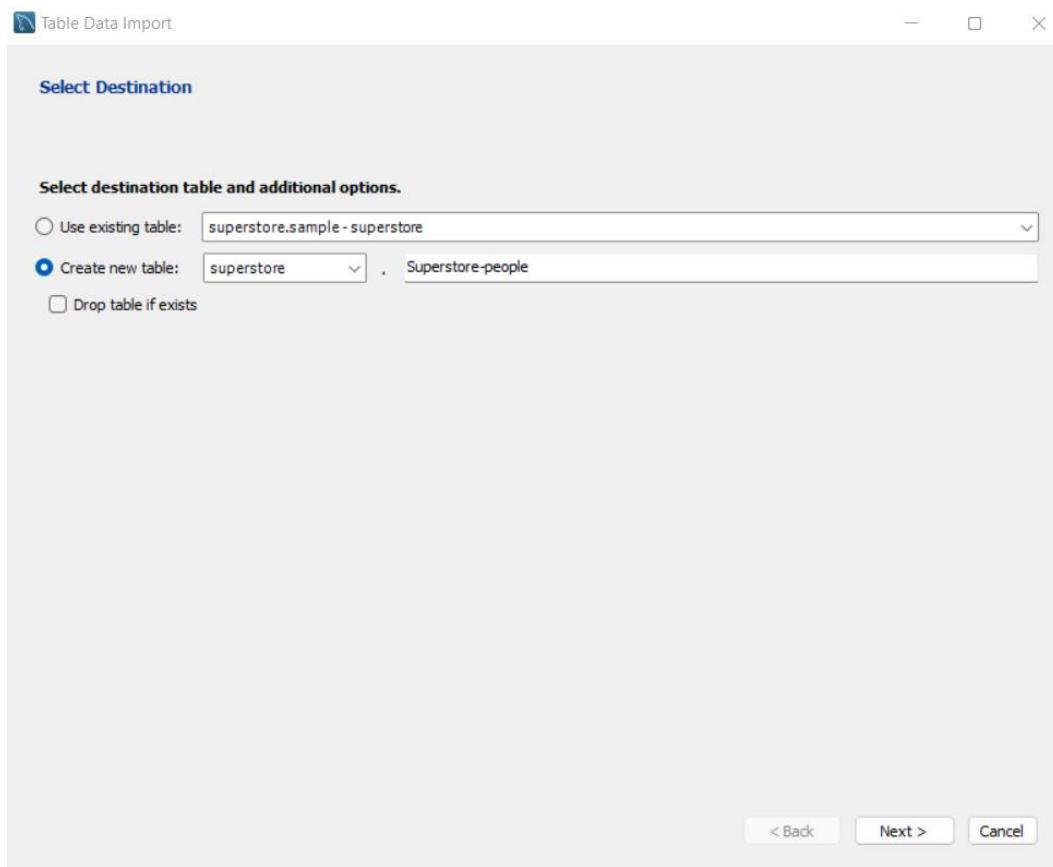


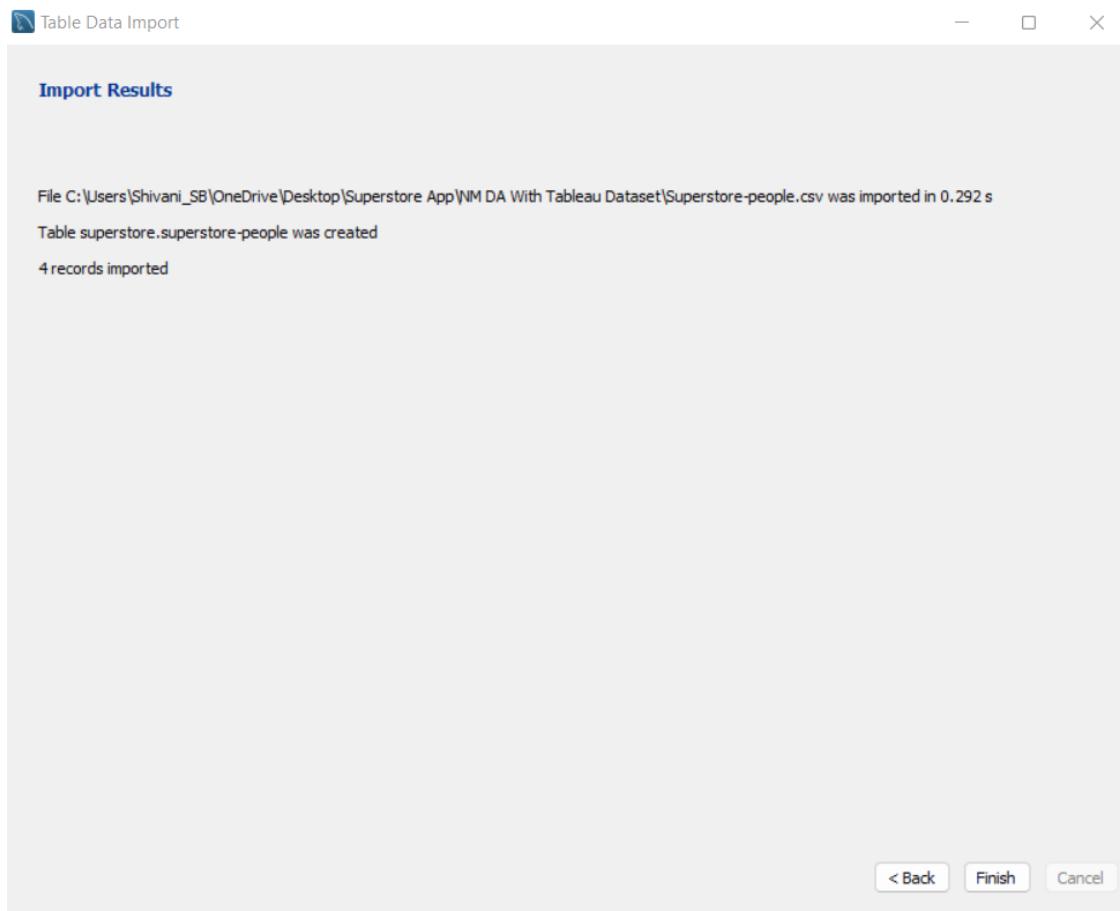
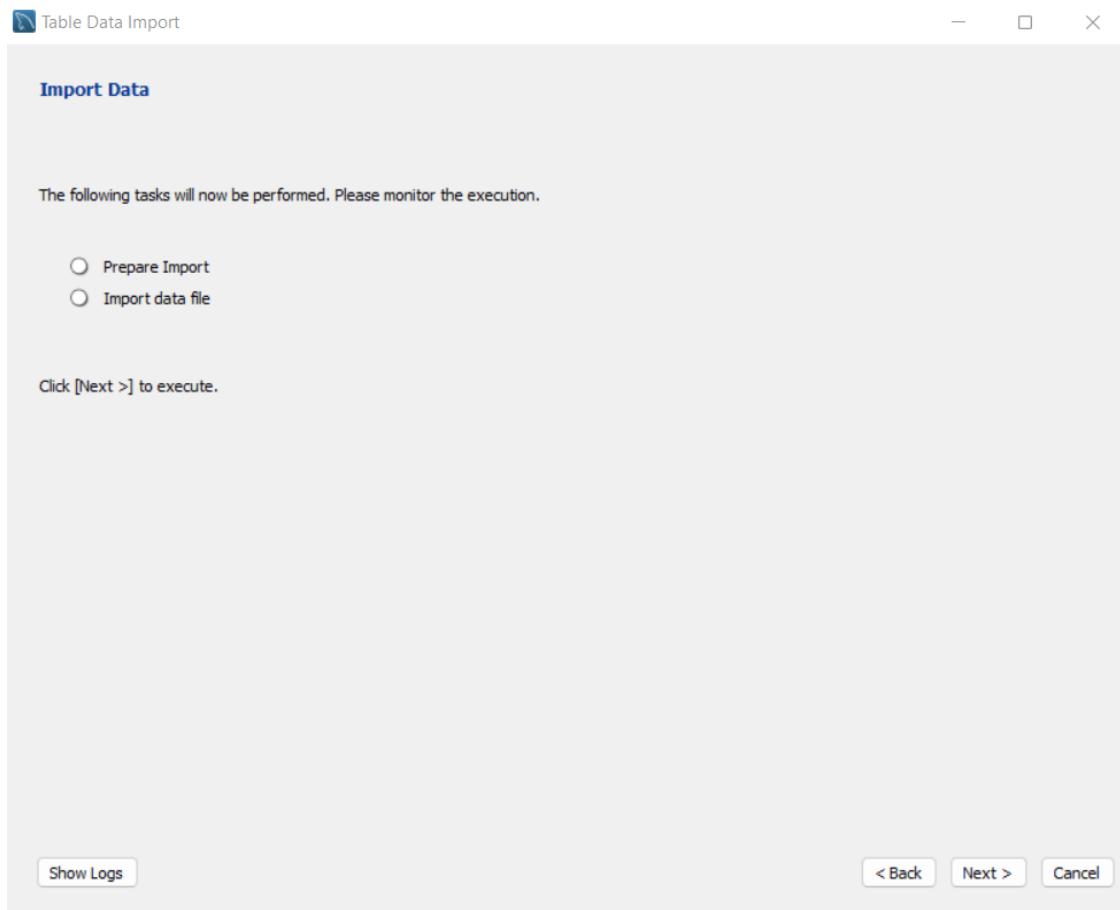
Now under tables you can see your newly created table

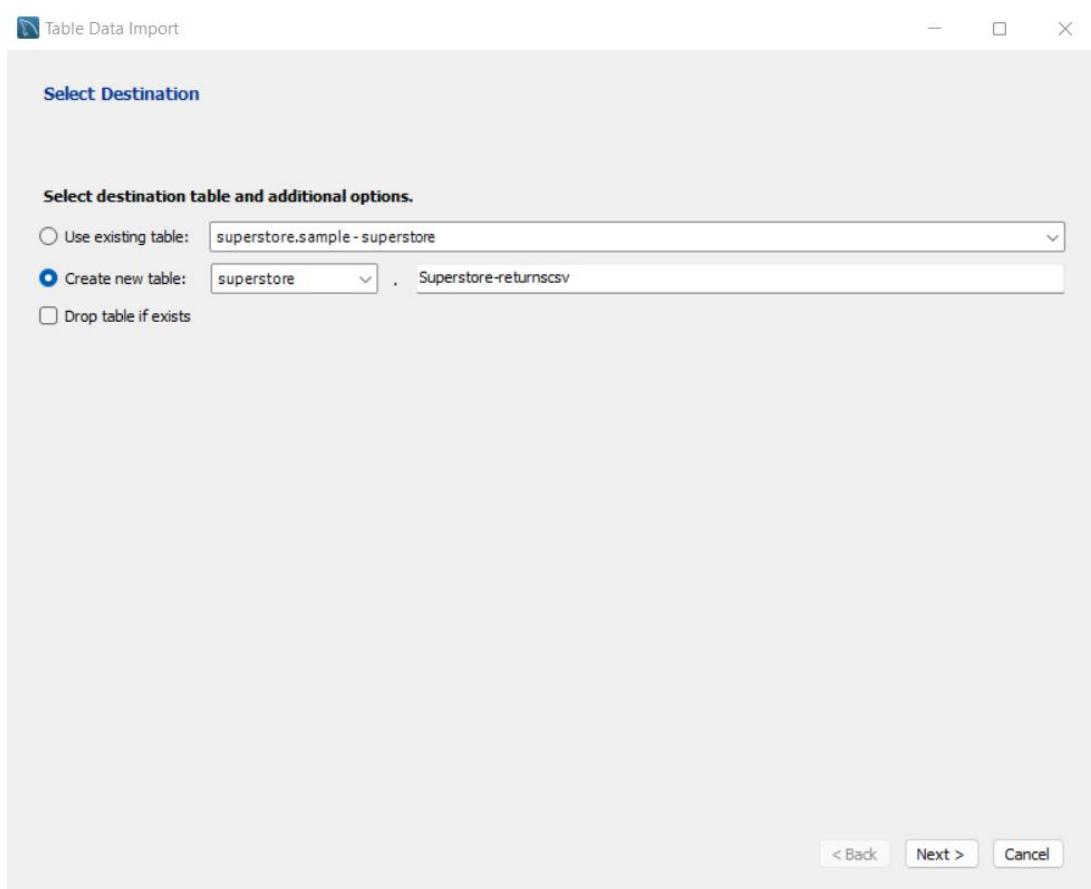
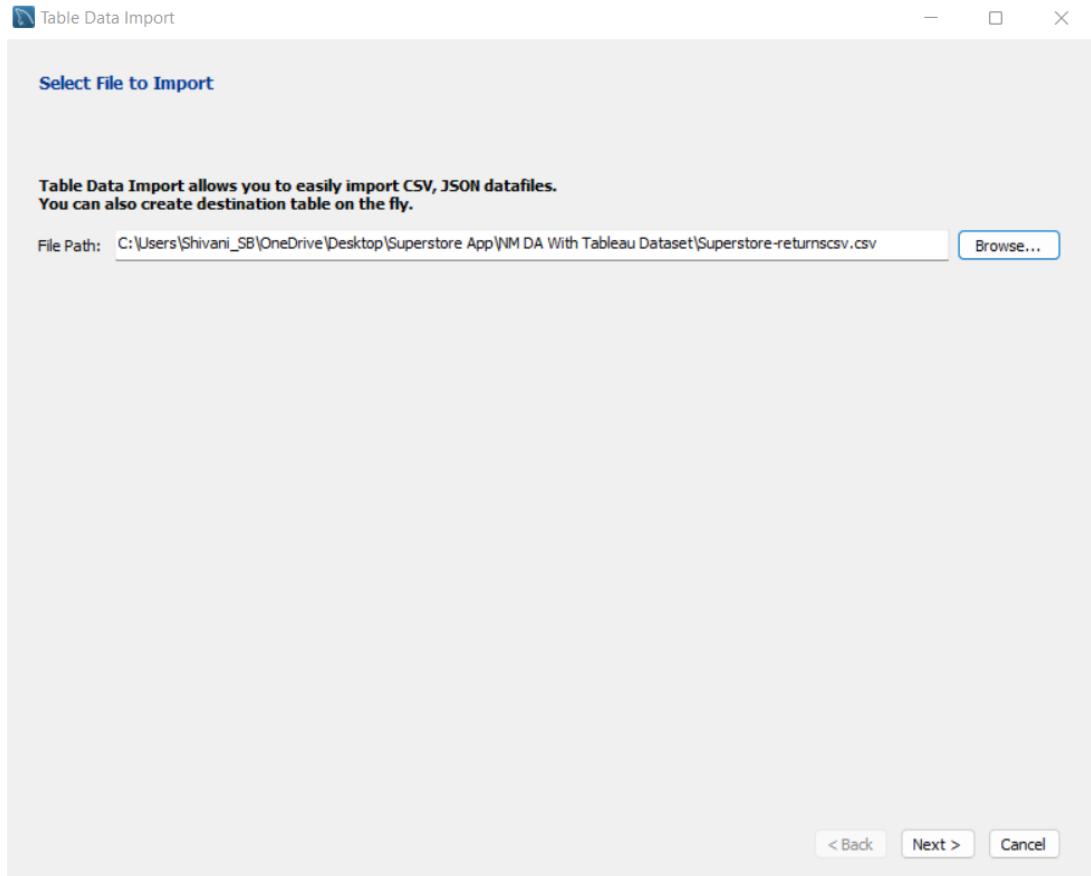


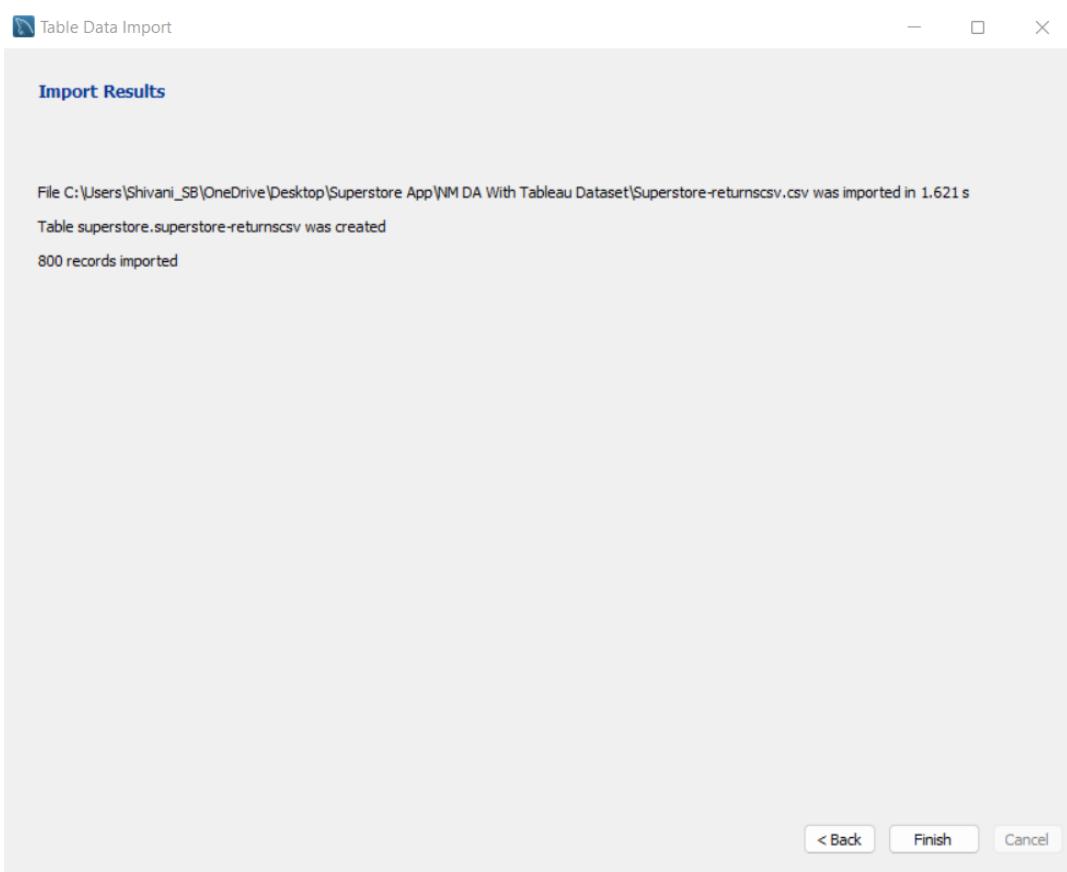
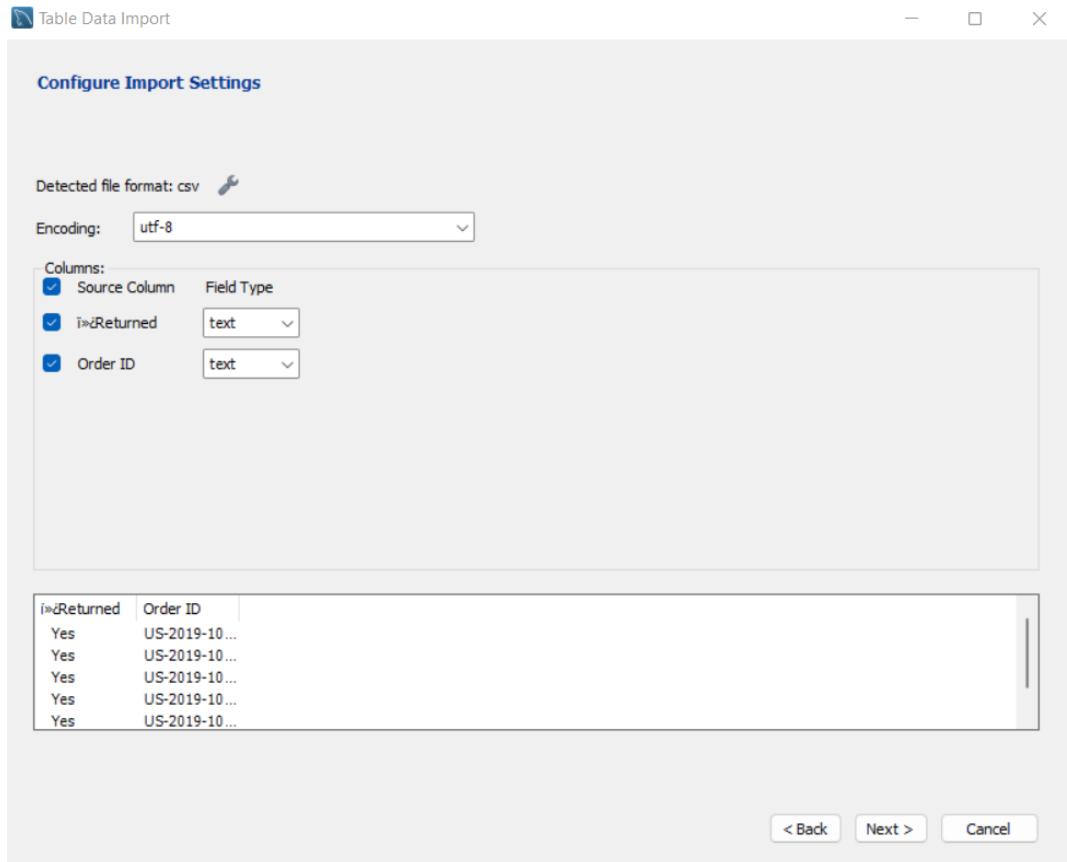
Now similarly import other two tables also











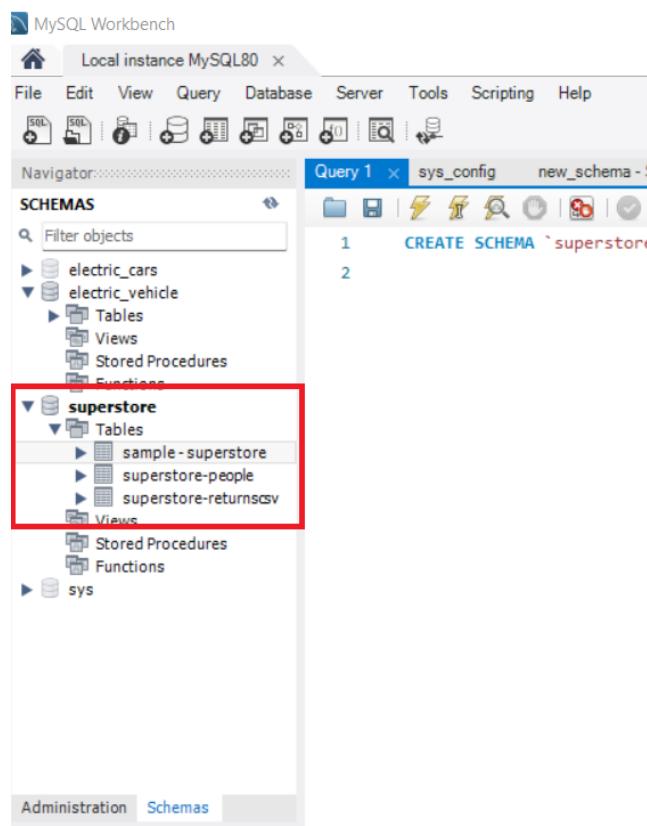
Now again refresh Schemas and see whether tables got created

The screenshot shows the MySQL Workbench interface. In the Navigator pane, the 'superstore' schema is selected, indicated by a red box around its entry. The 'Tables' node under 'superstore' is also highlighted with a red box. The Query Editor displays the SQL command: `CREATE SCHEMA `superstore` ;`. Below the Query Editor, the Output pane shows the execution log with several entries, all marked as 'OK'. The status bar at the bottom indicates 'Information'.

If you have only one Schema in your Database then it'll be your default schema and it'll be highted as bold , if you have multiple schemas then to use the particular schema you have to make it a default schema. For that you need to go to the schema you want to use right click and set as default schema.

The screenshot shows the MySQL Workbench interface with the 'superstore' schema selected in the Navigator pane. A context menu is open over the 'superstore' schema entry. The 'Set as Default Schema' option is highlighted with a red box. Other options in the menu include 'Load Spatial Data', 'Filter to This Schema', 'Schema Inspector', 'Table Data Import Wizard', 'Copy to Clipboard', 'Send to SQL Editor', 'Create Schema...', 'Alter Schema...', 'Drop Schema...', 'Search Table Data...', and 'Refresh All'. The status bar at the bottom indicates 'Information'.

Now you can see superstore schema is highlighted



Now let us perform some Queries to know more about our data

SQL Queries

1. Show all the data in superstore table

The SELECT statement is the most commonly used command in Structured Query Language. It is used to access the records from one or more database tables and views. It also retrieves the selected data that follow the conditions we want.

You need to write the query, “ select * from `sample - superstore`”; “ select it and run using run only selected query option. The results will appear on the canvas as the result grid.

The screenshot shows the MySQL Workbench interface with the following details:

- File Bar:** File, Edit, View, Query, Database, Server, Tools, Scripting, Help.
- Schemas:** Local instance MySQL 8.0, sys, sys_config, new_schema - Schema.
- Query Editor:** Contains the following SQL code:


```

1 CREATE SCHEMA `superstore` ;
2 • select * from `sample - superstore`;
3
      
```
- Results Grid:** Shows the results of the second query. The grid has columns: InvRow ID, Order ID, Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country/Region, City, State/Province, Postal Code, Region, Product ID, Category. The data includes rows for various customers like Darren Powers, Phillip Ober, and Maria Elzond, with details such as address, city, state, and product categories like Office Supplies and Furniture.
- Right Panel:** Contains tabs for Results, Editor, Form Editor, Field Types, Query, and Data.

2. Show all the data in people table

You need to write the query, “ select * from `sample - people`”; “ select it and run using run only selected query option. The results will appear on the canvas as the result grid.

The screenshot shows the MySQL Workbench interface with the following details:

- File Bar:** File, Edit, View, Query, Database, Server, Tools, Scripting, Help.
- Schemas:** Local instance MySQL 8.0, sys, sys_config, new_schema - Schema.
- Query Editor:** Contains the following SQL code:


```

1 CREATE SCHEMA `superstore` ;
2 • select * from `sample - superstore`;
3 • select * from `superstore-people`;
      
```
- Results Grid:** Shows the results of the third query. The grid has columns: InvRegional Manager, Region. The data includes rows for Sadie Pawthorne (West), Chuck Magee (East), Roxanne Rodriguez (Central), and Fred Suzuki (South).

3. Show all the data in returns table

You need to write the query, “ select * from `superstore-returnsCSV`”; “ select it and run using run only selected query option. The results will appear on the canvas as the result grid.

Query 1 x sys_config 2 new_schema - Schema

CREATE SCHEMA `superstore` ;
 select * from `sample - superstore`;
 select * from `superstore-people`;
 select * from `superstore-returnsCSV`;

Results

IDReturned	Order ID
Yes	US-2019-100762
Yes	US-2019-100867
Yes	US-2019-102652
Yes	US-2019-103373
Yes	US-2019-103744
Yes	US-2019-103744
Yes	US-2019-103940
Yes	US-2019-103940
Yes	US-2019-103940

4. Show different region from superstore table

The SQL DISTINCT command is used with SELECT key word to retrieve only distinct or unique data. You need to write query “SELECT DISTINCT Region FROM `sample - superstore`; ” , execute the query and you can see the results.

Query 1 x sys_config 2

CREATE SCHEMA `superstore` ;
 select * from `sample - superstore`;
 select * from `superstore-people`;
 select * from `superstore-returnsCSV`;
 SELECT DISTINCT Region FROM `sample - superstore`;

Result Grid

Region
Central
East
South
West

5. Show data from superstore table from category technology

A WHERE clause in SQL is a data manipulation language statement. You need to write query “SELECT * FROM `sample - superstore` WHERE Category='Technology'; ” , execute the query and you can see the results.

The screenshot shows a SQL Server Management Studio window titled 'Query 1 < sys_config'. The query is:

```

1 • CREATE SCHEMA `superstore`;
2 • select * from `sample - superstore`;
3 • select * from `superstore-people`;
4 • select * from `superstore-returnsCSV`;
5 • SELECT DISTINCT Region FROM `sample - superstore`;
6 • SELECT * FROM `sample - superstore` WHERE Category='Technology';

```

The results grid displays data for Technology products across various regions. Key columns include Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country/Region, City, State/Province, Postal Code, Region, Product ID, and Category.

Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country/Region	City	State/Province	Postal Code	Region	Product ID	Category	
13	US-2019-167199	06-01-2019	10-01-2019	ME-17230	Maria Espar	Home Office	United States	Henderson	Kentucky	42420	South	TEC-PH-1000439	Technology	
14	US-2019-167199	06-01-2019	10-01-2019	Standard Class	ME-17230	Maria Espar	Home Office	United States	Henderson	Kentucky	42420	South	TEC-PH-1000497	Technology
18	US-2019-153405	09-01-2019	13-01-2019	Standard Class	MS-17830	Melanie Seite	Consumer	United States	Laredo	Texas	78041	Central	TEC-AC-1000106	Technology
32	US-2019-162775	06-01-2019	10-01-2019	Second Class	CS-12230	Chris Seemon	Corporate	United States	Bossier City	Louisiana	71111	South	TEC-AC-10003174	Technology
34	US-2019-103366	15-01-2019	17-01-2019	First Class	EH-13990	Erica Hadley	Consumer	United States	Roswell	Georgia	30076	South	TEC-AC-10003428	Technology
38	US-2019-155402	06-01-2019	10-01-2019	Standard Class	DL-13916	Defne Latchford	Consumer	United States	Philadelphia	Pennsylvania	19134	East	TEC-PH-10004645	Technology
49	US-2019-157627	06-01-2019	10-01-2019	Standard Class	SD-20485	Shirley Daniels	Home Office	United States	Denton	Texas	75249	South	TEC-PH-10003561	Technology
76	US-2019-155502	26-01-2019	31-01-2019	Standard Class	SD-20485	Shirley Daniels	Home Office	United States	Alexandria	Virginia	22304	South	TEC-PH-10004833	Technology
77	US-2019-155402	26-01-2019	31-01-2019	Standard Class	SD-20485	Shirley Daniels	Home Office	United States	Alexandria	Virginia	22304	South	TEC-PH-10004833	Technology
85	US-2019-140795	01-02-2019	03-02-2019	First Class	BD-11500	Bradley Drucker	Consumer	United States	Green Bay	Wisconsin	54302	Central	TEC-AC-10001432	Technology
88	US-2019-123400	07-02-2019	09-02-2019	Standard Class	ER-13910	Erik Retkev	Consumer	United States	Gent Perle...	Florida	33710	South	TEC-PH-10002990	Technology
96	US-2019-131905	06-02-2019	09-02-2019	First Class	ND-18460	Nell Dutch	Corporate	United States	Chesapeake	Virginia	23320	South	TEC-PH-10001615	Technology
97	US-2019-131905	06-02-2019	09-02-2019	First Class	ND-18460	Nell Dutch	Corporate	United States	Chesapeake	Virginia	23320	South	TEC-PH-10003645	Technology
98	US-2019-107788	07-02-2019	12-02-2019	Standard Class	CS-17790	Cecilia Kiriakov	Corporate	United States	Iuden	New Jersey	07046	West	TEC-AC-10003710	Technology

6. Show details of orders from east region who purchased phones

Here we have used where Clause with two conditions and we have used AND operator. You need to write query “`SELECT * FROM `sample - superstore` WHERE Region= 'East' and `Sub-Category` = 'Phones' ;`”, execute the query and you can see the results.

The screenshot shows a SQL Server Management Studio window titled 'Query 1 < sys_config2'. The query is:

```

1 • CREATE SCHEMA `superstore`;
2 • select * from `sample - superstore`;
3 • select * from `superstore-people`;
4 • select * from `superstore-returnsCSV`;
5 • SELECT DISTINCT Region FROM `sample - superstore`;
6 • SELECT * FROM `sample - superstore` WHERE Category='Technology';
7 • SELECT * FROM `sample - superstore` WHERE Region= 'East' and `Sub-Category` = 'Phones';

```

The results grid displays data for phones purchased in the East region. Key columns include Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country/Region, City, State/Province, Postal Code, Region, Product ID, Category, and Sub-Category.

Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country/Region	City	State/Province	Postal Code	Region	Product ID	Category	Sub-Category
O19-115791	16-01-2019	18-01-2019	Second Class	DL-13315	Defne Latchford	Consumer	United States	Philadelphia	Pennsylvania	19134	East	TEC-PH-10004614	Technology	Phones
O19-143707	01-03-2019	05-03-2019	Standard Class	HR-14770	Halle Redmond	Home Office	United States	New York City	New York	10035	East	TEC-PH-10003655	Technology	Phones
O19-127964	03-03-2019	08-03-2019	Standard Class	AP-10720	Anne Pryor	Home Office	United States	New York City	New York	10035	East	TEC-PH-10004700	Technology	Phones
O19-122336	13-04-2019	17-04-2019	Second Class	JO-15895	Jonathan Doherty	Corporate	United States	Philadelphia	Pennsylvania	19140	East	TEC-PH-10000702	Technology	Phones
O19-118304	29-04-2019	04-05-2019	Standard Class	KH-16360	Katherine Hughes	Consumer	United States	Dover	Delaware	19901	East	TEC-PH-10000376	Technology	Phones
O19-010373	18-05-2019	24-05-2019	Standard Class	BS-11755	Bruce Stewart	Consumer	United States	Cleveland	Ohio	44105	East	TEC-PH-10002885	Technology	Phones
O19-105767	23-05-2019	27-05-2019	Standard Class	AR-10510	Andrew Roberts	Consumer	United States	Philadelphia	Pennsylvania	19134	East	TEC-PH-10003092	Technology	Phones
O19-010342	30-05-2019	01-06-2019	First Class	LW-16205	Laurie Workman	Corporate	United States	New York City	New York	10024	East	TEC-PH-10003505	Technology	Phones
O19-141796	21-06-2019	21-06-2019	Same Day	JG-15160	James Galang	Consumer	United States	Long Beach	New York	11561	East	TEC-PH-10001578	Technology	Phones
O19-130624	21-06-2019	24-06-2019	First Class	TB-21280	Toby Bräuhardt	Consumer	United States	New York City	New York	10024	East	TEC-PH-10003963	Technology	Phones
O19-101146	25-06-2019	25-06-2019	Standard Class	DB-10045	David Brooks	Consumer	United States	Lakewood	New Jersey	8701	East	TEC-PH-10004539	Technology	Phones
O19-154963	22-06-2019	06-06-2019	Standard Class	AA-10045	Andrea Andreoli	Consumer	United States	Chestnut	Pennsylvania	19013	East	TEC-PH-10004059	Technology	Phones
O19-107944	02-07-2019	06-07-2019	Standard Class	EH-13945	Eric Hoffmann	Consumer	United States	Plainfield	New Jersey	7060	East	TEC-PH-10002923	Technology	Phones
O19-161883	11-07-2019	15-07-2019	Standard Class	IA-16780	Yvonne Ametivens	Corporate	United States	Harkmark	New Jersey	76011	East	TEC-PH-10003885	Technology	Phones

7. Show product details of top 5 most profitable products

The ORDER BY clause in SQL will help us to sort the records based on the specific column of a table. This means that all the values stored in the column on which we are applying the ORDER BY clause will be sorted, and the corresponding column values will be displayed in the sequence in which we have obtained the values in the earlier step. Using the ORDER BY clause, we can sort the records in ascending or descending order as per our requirement. The records will be sorted in ascending order whenever the ASC keyword is used with ORDER by clause. DESC keyword will sort the records in descending order. Here we are going to order by Profit.

We can retrieve limited rows from the database, here we want top 5 only so we will set the limit to 5. You need to write a query “`SELECT `Product ID`, Category, Sales, Profit FROM `sample - superstore` order by Profit DESC LIMIT 5;`”, execute the query and you can see the results.

The screenshot shows a MySQL Workbench interface with a query editor and a result grid. The query editor contains the following SQL code:

```

Query 1 | sys_config 2
CREATE SCHEMA `superstore`;
select * from `sample - superstore`;
select * from `superstore-people`;
select * from `superstore-returnscsv`;
SELECT DISTINCT Region FROM `sample - superstore`;
SELECT * FROM `sample - superstore` WHERE Category='Technology';
SELECT * FROM `sample - superstore` WHERE Region= 'East' and `Sub-Category` = 'Phones' ;
SELECT `Product ID`, Category, Sales, Profit FROM `sample - superstore` order by Profit DESC LIMIT 5;

```

The result grid displays the following data:

Product ID	Category	Sales	Profit
TEC-CO-10004722	Technology	17499.95	8399.976
TEC-CO-10004722	Technology	13999.96	6719.9808
TEC-CO-10004722	Technology	10499.97	5039.9856
OFF-BI-10000545	Office Supplies	9892.74	4946.37
OFF-BI-10001120	Office Supplies	9449.95	4630.4755

The last row of the result grid is highlighted with a red box, and the word "Result" is written in red next to it.

8. Show no of order placed from different region where maximum discount was offered

The HAVING clause places the condition in the groups defined by the GROUP BY clause in the SELECT statement. Here we will group by region as we want count of different regions and we have condition that we want max discount so max function will be used with having clause.

You need to write query “ SELECT COUNT(`Order ID`), Region FROM `sample - superstore` GROUP BY Region HAVING max(Discount); ” , execute the query and you can see the results.

The screenshot shows a MySQL Workbench interface with a query editor and a result grid. The query editor contains the following SQL code:

```

Query 1 | sys_config 2
CREATE SCHEMA `superstore`;
select * from `sample - superstore`;
select * from `superstore-people`;
select * from `superstore-returnscsv`;
SELECT DISTINCT Region FROM `sample - superstore`;
SELECT * FROM `sample - superstore` WHERE Category='Technology';
SELECT * FROM `sample - superstore` WHERE Region= 'East' and `Sub-Category` = 'Phones' ;
SELECT `Product ID`, Category, Sales, Profit FROM `sample - superstore` order by Profit DESC LIMIT 5;
SELECT COUNT(`Order ID`), Region FROM `sample - superstore` GROUP BY Region HAVING max(Discount);

```

The result grid displays the following data:

COUNT(`Order ID`)	Region
2323	Central
2848	East
1620	South
3203	West

The last row of the result grid is highlighted with a red box, and the word "Result" is written in red next to it.

9. Insert new regional manager from northwest region into people table

SQL INSERT statement is a SQL query. It is used to insert a single or a multiple records in a table.

You need to write the query, “ insert into `superstore-people` (`Regional Manager` , Region) values ('Rahul','Nortwest'); ”

`select * from `superstore-people`;`

select both and run using run only selected query option. The results will appear on the canvas as the result grid.

```

Query 1 sys_config
1 • CREATE SCHEMA `superstore` ;
2 • select * from `sample - superstore`;
3 • select * from `superstore-people`;
4 • select * from `superstore-returnsCSV`;
5 • SELECT DISTINCT Region FROM `sample - superstore`;
6 • SELECT * FROM `sample - superstore` WHERE Category='Technology';
7 • SELECT * FROM `sample - superstore` WHERE Region= 'East' and `Sub-Category` = 'Phones' ;
8 • SELECT `Product ID`, Category, Sales, Profit FROM `sample - superstore` order by Profit DESC LIMIT 5;
9 • SELECT COUNT(`Order ID`), Region FROM `sample - superstore` GROUP BY Region HAVING max(Discount);
10 • insert into `superstore-people` ('i>Regional Manager', Region) values ('Rahul','Northwest'); 1
11 • select * from `superstore-people`;
--
```

Result Grid | Filter Rows: Export: Wrap Cell Content: Result

i>Regional Manager	Region
Sadie Pawthorne	West
Chuck Magee	East
Roxanne Rodriguez	Central
Fred Suzuki	South
Rahul	Northwest

10. Rename column name

The ALTER TABLE statement in Structured Query Language allows you to add, modify, and delete columns of an existing table. Here, We can modify and rename the columns using alter table command, you need to write query as “ALTER TABLE `superstore-people` RENAME COLUMN `i>Regional Manager` TO Regional_Manager;

`select * from `superstore-people`;"`

select both and run using run only selected query option. The results will appear on the canvas as the result grid.

```

Query 1 sys_config
6 • SELECT * FROM `sample - superstore` WHERE Category='Technology';
7 • SELECT * FROM `sample - superstore` WHERE Region= 'East' and `Sub-Category` = 'Phones' ;
8 • SELECT `Product ID`, Category, Sales, Profit FROM `sample - superstore` order by Profit DESC LIMIT 5;
9 • SELECT COUNT(`Order ID`), Region FROM `sample - superstore` GROUP BY Region HAVING max(Discount);
10 • insert into `superstore-people` ('i>Regional Manager', Region) values ('Rahul','Northwest');
11 • select * from `superstore-people`;
12 • ALTER TABLE `superstore-people` RENAME COLUMN `i>Regional Manager` TO Regional_Manager; 1
13 • select * from `superstore-people`;
--
```

Result Grid | Filter Rows: Export: Wrap Cell Content: Result

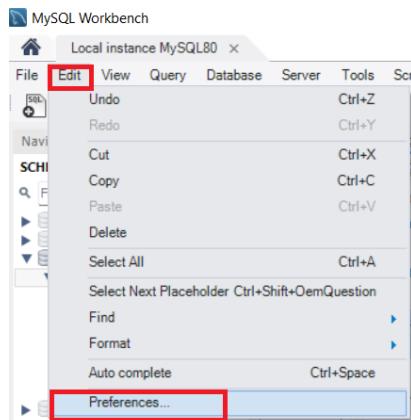
Regional_Manager	Region
Sadie Pawthorne	West
Chuck Magee	East
Roxanne Rodriguez	Central
Fred Suzuki	South
Rahul	Northwest

11. Update name of regional manager from northwest region into people table

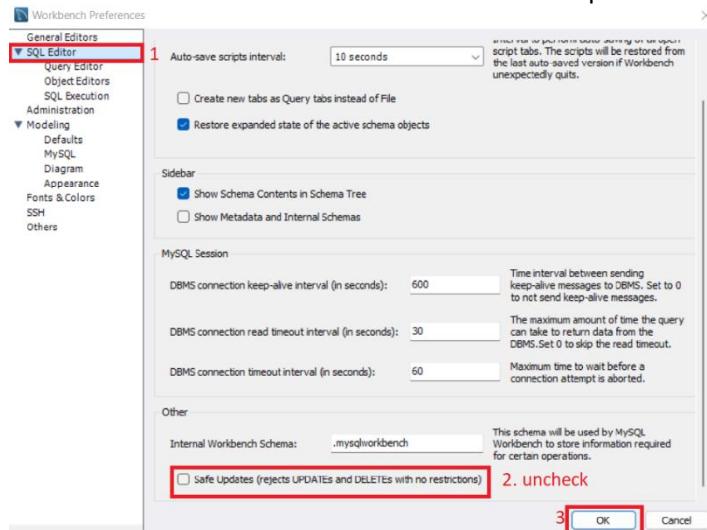
SQL UPDATE statement is used to change the data of the records held by tables. Which rows is to be update, it is decided by a condition. To specify condition, we use WHERE clause. Here we are changing name of regional manager that belongs to northwest region.

We are working in safe mode by default so it won't allow us updating table we need to disable safe mode first.

Go to edit -> preferences



Go to SQL editor scroll down and uncheck safe updates and click on okay.



Now you just need to run the query " UPDATE `superstore-people` SET Regional_Manager = 'Naveen' WHERE Region = 'Northwest'; select * from `superstore-people`;"

select both and run using run only selected query option. The results will appear on the canvas as the result grid.

Regional_Manager	Region
Sadie Pawthorne	West
Chuck Magee	East
Roxanne Rodriguez	Central
Fred Suzuki	South
Naveen	Northwest

12. Delete record for Northwest region from table people

The SQL DELETE statement is used to delete rows from a table. Generally DELETE statement removes one or more records from a table. We can delete the row that is having region as Northwest you need to write query as "DELETE FROM `superstore-people` WHERE Region = 'Northwest';

```
select * from `superstore-people`;"
```

select both and run using run only selected query option. The results will appear on the canvas as the result grid.

Regional_Manager	Region
Sadie Pawthorne	West
Chuck Magee	East
Roxanne Rodriguez	Central
Fred Suzuki	South

SQL Query Script

```

CREATE SCHEMA `superstore` ;

select * from `sample - superstore`;

select * from `superstore-people`;

select * from `superstore-returnsCSV`;

SELECT DISTINCT Region FROM `sample - superstore`;

SELECT * FROM `sample - superstore` WHERE Category='Technology';

SELECT * FROM `sample - superstore` WHERE Region= 'East' and `Sub-Category` = 'Phones' ;

SELECT `Product ID`, Category, Sales, Profit FROM `sample - superstore` order by Profit DESC LIMIT 5;

SELECT COUNT(`Order ID`), Region FROM `sample - superstore` GROUP BY Region HAVING max(Discount);

insert into `superstore-people` (`Regional Manager`, Region) values ('Rahul','Northwest');
select * from `superstore-people`;

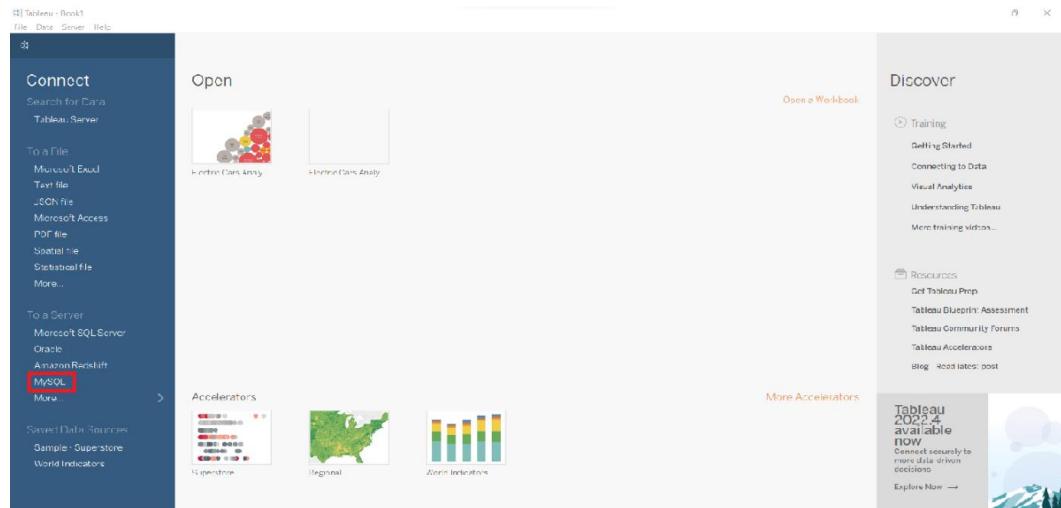
ALTER TABLE `superstore-people` RENAME COLUMN `Regional Manager` TO Regional_Manager;
select * from `superstore-people`;
UPDATE `superstore-people`SET Regional_Manager = 'Naveen' WHERE Region = 'Northwest';
select * from `superstore-people`;
```

```
DELETE FROM `superstore-people` WHERE Region = 'Northwest';
select * from `superstore-people`;
```

Activity 3: Connect DB with Tableau

Database Connection

Launch tableau desktop, on the left menu you can see different options for connection, now select MySQL under To a Server, if MySQL is not there please go to more and now you can select MySQL from a variety of options provided.



Now to create database connection you need to provide some details, these details you can get in your workbench under administration.

MySQL

General Initial SQL

Server
localhost

Port
3306

Database
superstore

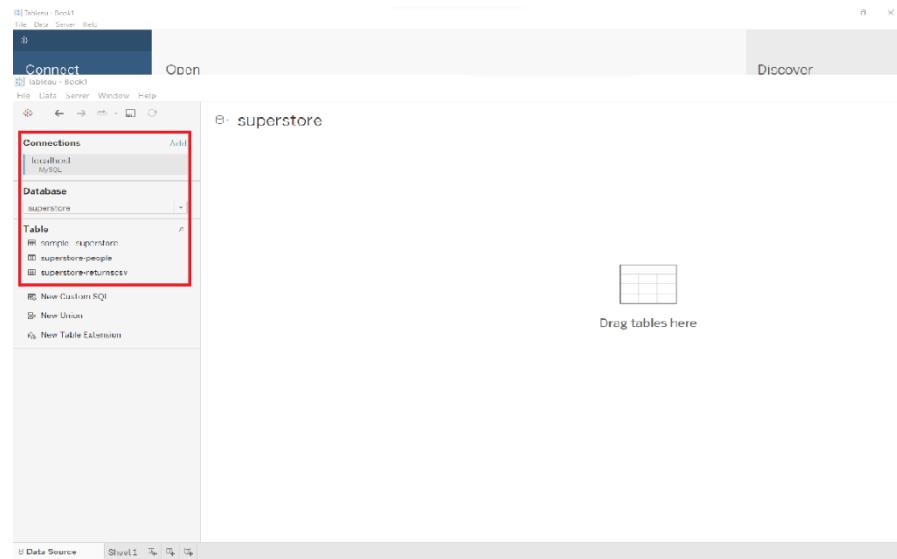
Username
root

Password

Require SSL

Sign In

After giving details we need to just click on Sign In, and your connection will be established



Drag and drop sample-superstore table to canvas

The screenshot shows the Tableau Data Source interface. On the left, under 'Database' and 'superstore', the 'sample - superstore' table is selected and highlighted with a red box. A blue arrow points from this selection to the 'Fields' pane on the right. The 'Fields' pane lists fields: 'Name' (sample - superstore), 'Type' (All), 'Field Name' (e.g., 'InvRow ID', 'Order ID', 'Order Date'), 'Physical Table' (sample - superstore), and 'Remote...' (All). Below the Fields pane is a preview table with columns: InvRow ID, Order ID, Order Date, Order ID, Order Date, Ship Date, Ship Mode, Customer ID, and Customer Name.

Now let's drag next table sample- people on to the canvas and join with superstore data and form relationship, both have region in common so we will form relation based on region

The screenshot shows the Tableau Data Source interface. Under 'Database' and 'superstore', the 'sample - superstore' and 'superstore-people' tables are selected and highlighted with red boxes. A blue arrow points from the 'superstore-people' selection to the 'Relationships' pane. In the 'Relationships' pane, 'sample - superstore' is joined with 'superstore-people' using the 'Region' field as the common key. The 'Operator' is set to 'Equal'. The preview table shows the joined data with columns: All, sample - superstore, All, superstore-people, and Regional Manager.

Now let's drag next table sample- returnscsv on to the canvas and join with superstore data and form relationship, both have Order Id in common so we will form relation based on Order Id

The screenshot shows the Tableau Data Source interface. Under 'Database' and 'superstore', the 'sample - superstore' and 'superstore-returnscsv' tables are selected and highlighted with red boxes. A blue arrow points from the 'superstore-returnscsv' selection to the 'Relationships' pane. In the 'Relationships' pane, 'sample - superstore' is joined with 'superstore-returnscsv' using the 'Order ID' field as the common key. The 'Operator' is set to 'Equal'. The preview table shows the joined data with columns: All, sample - superstore, All, superstore-returnscsv, and 'Is Returned'.

Milestone 3: Data Preparation

Activity 1: Prepare the Data for Visualization

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency.

Data preprocessing can be performed in many ways using many different steps depending on your data here, we are going to do some part of data preparation on our data.

Rename Columns: Here you can see that there are some marks in Row Id column name, Let's rename it.

The screenshot shows the Tableau Data Editor interface. At the top, the title bar says 'Tableau - Book1'. Below it, the 'Connections' section shows a connection to 'localhost MySQL' with the database set to 'superstore'. The 'Tables' section lists 'sample - superstore', 'superstore-people', and 'superstore-returnscsv'. In the main area, a tree diagram shows 'sample - superstore' branching into 'superstore-people' and 'superstore-returnscsv'. Below this, a data grid for 'sample - superstore' is displayed with 20 fields and 9994 rows. A red box highlights the 'sample - superstore' table in the top navigation bar. Another red box highlights the 'Row ID' column in the data grid. A third red box highlights the context menu option 'Rename...' for the 'Row ID' column.

You need to select sample superstore data first then go to the column you want to rename, right click and select rename and give a new name to the column, your column will be renamed.

The screenshot shows the Tableau Data Editor interface after renaming the 'Row ID' column. The 'sample - superstore' table is selected in the top navigation bar. The data grid shows the same 20 fields and 9994 rows, but the 'Row ID' column has been renamed. A red box highlights the 'sample - superstore' table in the top navigation bar. Another red box highlights the 'Row ID' column in the data grid. A third red box highlights the context menu option 'Renamed' for the 'Row ID' column.

Similarly if we go to returnscsv data there also we need to rename the first column, performing the same steps. select sample-returnscsv table from dropdown and rename first column

The screenshot shows the Tableau Data Source interface. On the left, the connections pane shows a MySQL connection named 'localhost'. The database is set to 'superstore'. In the center, a diagram shows a relationship between 'sample - superstore' and two other tables: 'superstore-people' and 'superstore-returnscsv'. Below this, the 'superstore-returnscsv' table is selected, displaying 2 fields and 800 rows. A context menu is open over the 'IsReturned' column, with the 'Rename' option highlighted.

After renaming

The screenshot shows the Tableau Data Source interface after renaming. The 'superstore-returnscsv' table now has a single field named 'Returned'. The data values are identical to the original 'IsReturned' column: Yes, Yes, Yes, Yes, Yes, Yes.

Type	Field Name	Phys...	Rem...
Abc	Returned	super...	IsRet...
Abc	Order ID (superstore-returnsc...	super...	Order ...

Again select sample-superstore table from dropdown and let's check type of each column, here order date is abc type that is string type but date should have date type, we need to change the type of order date. Right click on abc and select date.

B- sample - superstore+ (superstore)

```
graph LR; localhost[localhost MySQL] --> sample[sample - superstore]; sample --> people[superstore-people]; sample --> returns[superstore-returnscsv]
```

Connections

localhost MySQL

Database

superstore

Table

- sample - superstore
- superstore-people
- superstore-returnscsv

New Custom SQL

New Union

New Table Extension

Connection

Live Extract

Filters 0 | Add

1

sample - superstore 20 fields 9994 rows

2

#	sample - superstore					
Row ID	Order ID	Number (decimal)	Date & Time	Ship Mode	Customer ID	Customer Name
84	US-2019-115161	31-0	✓ Date	First Class	LC-17050	Liz Carlisle
82	US-2019-134103	30-0	✓ String	Standard Class	MV-18190	Mike Vittorini
83	US-2019-134103	30-0	✓ Default	Standard Class	MV-18190	Mike Vittorini
81	US-2019-100328	28-0	Geographic Role	Standard Class	JC-15340	Jasper Cacioppo
78	US-2019-117163	27-01-2019	02-02-2019	Standard Class	EJ-13720	Ed Jacobs
79	US-2019-117163	27-01-2019	02-02-2019	Standard Class	EJ-13720	Ed Jacobs

Now you can see the type is changed to date. Similarly we need to change the type of ship date also as it is also a string we need to convert to date.

File Data Server Window Help

sample - superstore+ (superstore)

Connection Live Extract

Filters 0 | Add

Connections

- localhost MySQL

Database

- superstore

Table

- sample - superstore
- superstore-people
- superstore-returnscsv

New Custom SQL

New Union

New Table Extension

```

graph LR
    A["sample - superstore"] --> B["superstore-people"]
    A --> C["superstore-returnscsv"]
  
```

sample - superstore

Row ID	Order ID	Order Date	Ship Date & Time	Customer ID
101	US-2019-168368	11-02-2019	15:00	GA-14725
102	US-2019-168368	11-02-2019	15:00	GA-14725
103	US-2019-127614	11-02-2019	15:00	✓ Default
104	US-2019-127614	11-02-2019	15:00	Geographic Role
105	US-2019-168368	11-02-2019	15:02-2019	Second Class
100	US-2019-125799	08-02-2019	09-02-2019	First Class
				NM-18445

After changing typed you can see both as date type denoted with calendar symbol.

File Data Server Window Help

Connections Add

localhost MySQL

Database superstore

Table sample - superstore
superstore-people
superstore-returnscsv

New Custom SQL
New Union
New Table Extension

sample - superstore (superstore)

superstore-people

superstore-returnscsv

Connections Connection
Live Extract

Filters 0 Add

sample - superstore		20 fields 9994 rows			
Name	sample - superstore				
Row ID	101	US-2019-168368	11-02-2019	15-02-2019	Second Class
	102	US-2019-168368	11-02-2019	15-02-2019	Second Class
	103	US-2019-127614	11-02-2019	15-02-2019	Standard Class
	104	US-2019-127614	11-02-2019	15-02-2019	Standard Class
	105	US-2019-168368	11-02-2019	15-02-2019	Second Class
	106	US-2019-125759	08-02-2019	09-02-2019	First Class
Type	Field Name	Phys...	Rem...		
Row ID	sample...	1x1Ro...			
Order ID	sample...	Order ...			
Customer ID	sample...	Order ...			

You can check with other types also like location should have geographic role in our data all location related columns are correctly in geographic role denoted by globe symbol.

The screenshot shows the Tableau Data Source view. On the left, the 'Connections' pane lists 'localhost MySQL' and the 'Database' pane shows 'superstore'. Under 'Table', there are three entries: 'sample - superstore', 'superstore-people', and 'superstore-returnscsv'. Below these are options for 'New Custom SQL', 'New Union', and 'New Table Extension'. The main area displays a preview of the 'sample - superstore' table, which contains 20 fields and 9994 rows. The table structure includes columns for Name, Type, Field Name, Physical Name, and Remarks. The preview shows data for 'Country/Region', 'City', 'State/Province', and 'Postal Code'. At the top right of the preview, there are sorting icons (ascend, descend, and refresh) and a row limit of 100. The bottom right corner of the preview has a small icon with a red border.

From the setting symbol on the right top of table view, we can perform sorting in ascending or descending order and see hidden fields also.

This screenshot is identical to the one above, showing the Tableau Data Source view with the 'sample - superstore' table preview. However, a 'Sort Fields' dialog box is now open at the bottom right of the preview area. The dialog contains several options: 'Show aliases', 'Show hidden fields', 'Sort Fields:', and a list of sorting rules: 'Data source order', 'A to Z ascending', 'Z to A descending', 'A to Z ascending per table', and 'Z to A descending per table'. The 'Z to A descending per table' option is highlighted with a red border. The rest of the interface remains the same, including the connection details and table structure.

Now at the bottom of your screen there are some symbols first one is to create New Worksheet, second one is to create new Dashboard, third one is to create new Story, and by default you have one sheet now you can continue on sheet 1 or if you want to create you can create by clicking 1st icon.

Tableau - Book1

File Data Server Window Help

Connections Add

localhost MySQL

Database superstore

Table

- sample - superstore
- superstore-people
- superstore-returnsCSV

New Custom SQL

New Union

New Table Extension

sample - superstore 20 fields 9994 rows

Name sample - superstore

Fields

Type	Field Name	Phys...	Rem...
#	Row ID	sample...	1% Ro...
Abc	Order ID	sample...	Order ...
File	Customer Data	sample...	Order ...

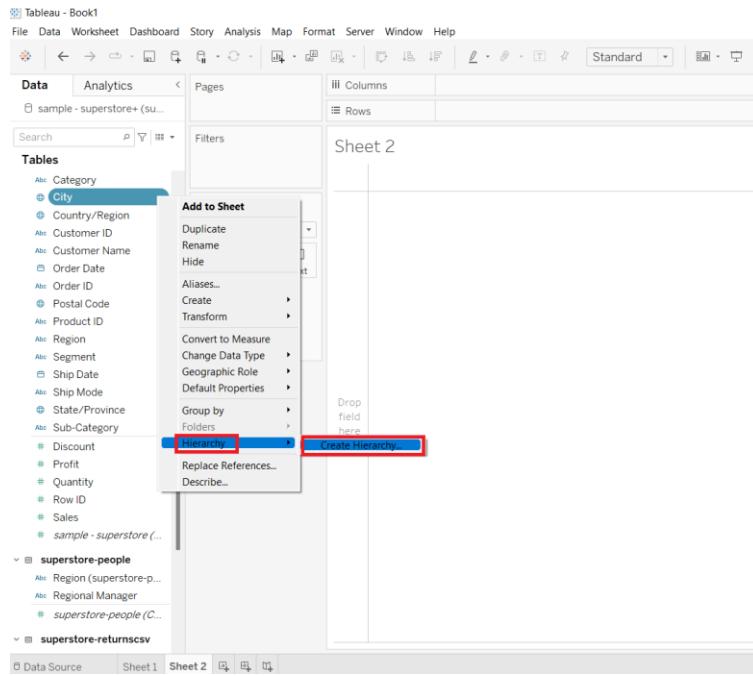
1 2 3

Data Source Sheet1

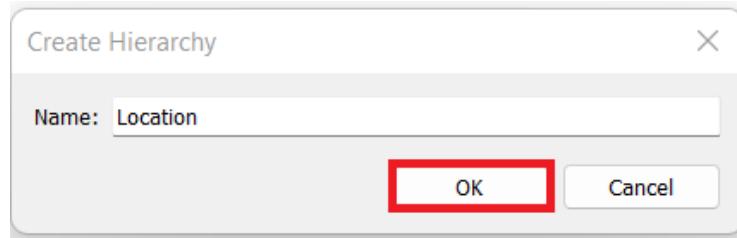
Now we'll perform Hierarchy where we can form a hierarchy of columns, like there are 5 columns related to location City, Country/Region, Postal Code, Region, State/Province we can form a hierarchy of these columns. For that we need to go to the City column on the left side where we can see the data source and click on the dropdown.

The screenshot shows the Tableau interface with the title "Tableau - Book1". The top menu bar includes File, Data, Worksheet, Dashboard, Story, Analysis, Map, Format, Server, Window, and Help. Below the menu is a toolbar with various icons. The left sidebar is titled "Tables" and lists several data sources: "sample - superstore+" (selected), "superstore-people", and "superstore-returnscsv". A red box highlights the "City" field in the "sample - superstore+" list. The main workspace is divided into four quadrants: "Pages" (top-left), "Filters" (top-right), "Marks" (bottom-left), and "Sheet 2" (bottom-right). The "Marks" quadrant contains options for Automatic, Color, Size, Text, Detail, and Tooltip. A "Drop field here" placeholder is visible in the bottom-right quadrant. The status bar at the bottom shows "Data Source" and "Sheet 2".

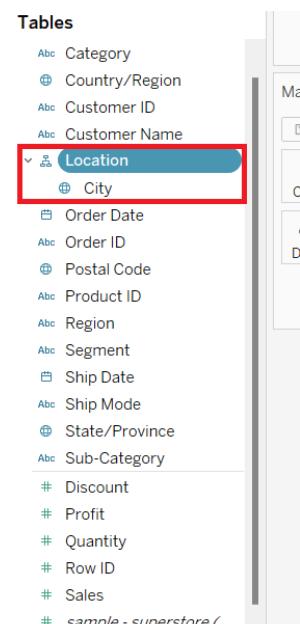
From Dropdown select Hierarchy then click on hierarchy



then a dialog box will appear give name to the hierarchy (as location or any other) and click on okay



Hierarchy will be created and you can see the city under the location hierarchy.



Now drag and drop Country/Region, Postal Code, Region, State/Province also under this hierarchy. You can see your created hierarchy

Tables

Abc	Category
Abc	Customer ID
Abc	Customer Name
▼	Location
⊕	City
⊕	Country/Region
⊕	Postal Code
⊕	State/Province
Abc	Region
⊕	Order Date
Abc	Order ID
Abc	Product ID
Abc	Segment
⊕	Ship Date
Abc	Ship Mode
Abc	Sub-Category
#	Discount
#	Profit
#	Quantity
#	Row ID
#	Sales
#	sample - superstore (...)

Note:- Data Preparation is not limited to this only it depends on your data, we have performed some basic operations on the data.

Milestone 4: Data Visualization

Data visualization is the process of creating graphical representations of data in order to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

Activity 1: No of Unique Visualizations

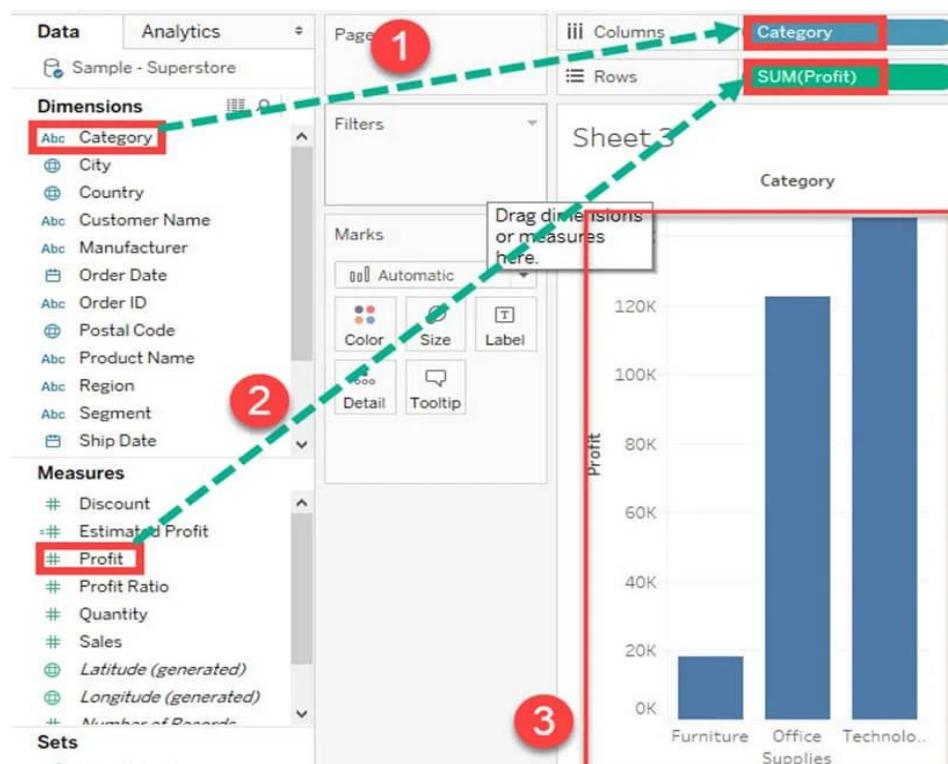
The number of unique visualizations that can be created with a given dataset. Some common types of visualizations that can be used to analyze the performance and efficiency of Superstore Sales include bar charts, line charts, heat maps, scatter plots, pie charts, Maps etc. These visualizations can be used to compare performance, track changes over time, show distribution, and relationships between variables, breakdown of revenue and customer demographics, workload, resource allocation and location of hotels.

Activity 1.1: Bar Chart:

A bar chart can compare the data across different categories. The height of the bars represents the measured value of each category. It can be represented as vertical and horizontal type bar charts. The procedure to create a bar chart is given as follows.

Step) Go to a new worksheet.

- Drag 'Category' into Column.
- Drag 'Profit' into Rows.
- It creates a bar chart by default.

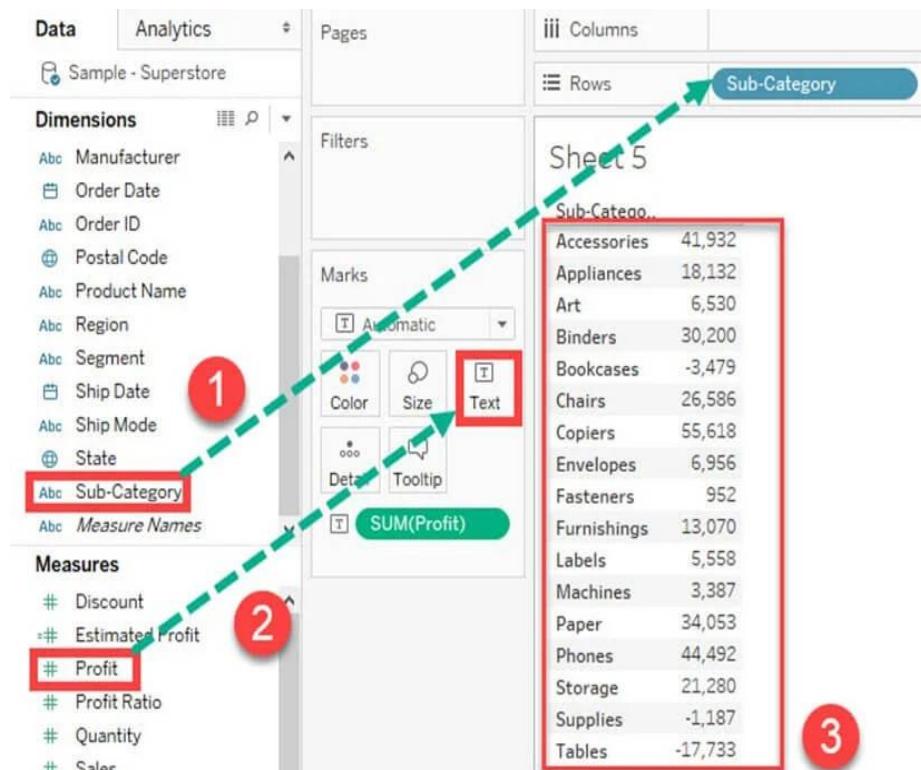


Activity 1.2: Text Tables:

Text tables are used to show the exact value of measures across the different dimensions. A text table is also called a Pivot Table. It groups the dimensions and measures by default. The procedure to design a text table is given as follows.

Step) Go to a new Worksheet.

- Drag 'Sub-Category' into Rows.
- Drag 'Profit' into the textbox present in the marks card.
- It creates a text table by default.

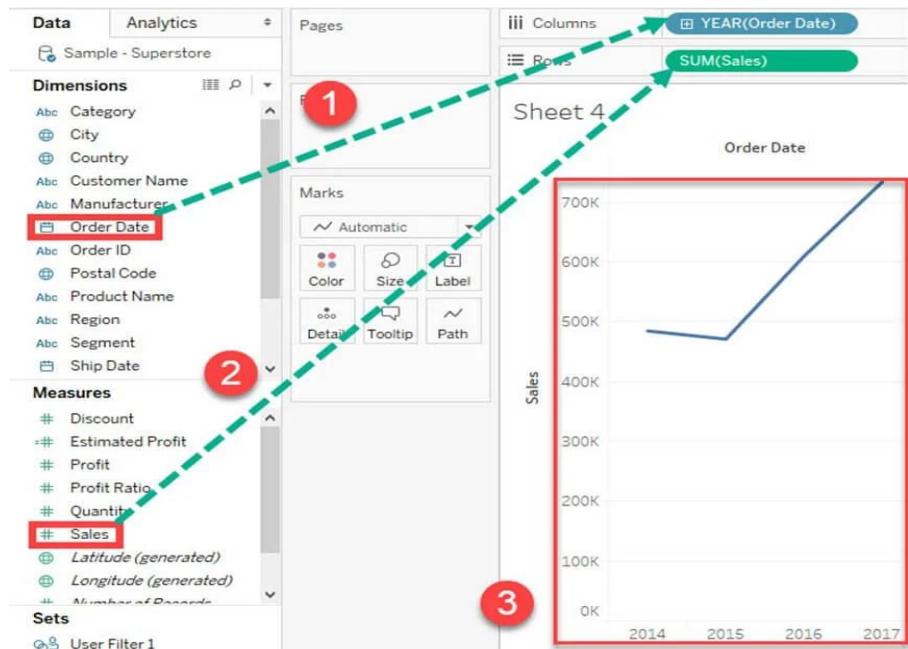


Activity 1.3: Line chart:

A Line Chart should be used to compare the data over the different periods. A line chart is created by the series of dots. These dots represent the measured value in each period. The procedure to create a line graph is shown below.

Step) Go to a new Worksheet.

- Drag 'Order Date' into Columns.
- Drag 'Sales' into Rows.
- It creates a line chart by default.

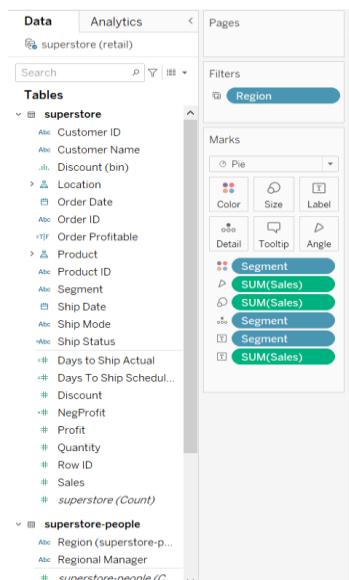


Activity 1.4: Pie Chart:

A pie chart can show the segment-wise data. It can show the contribution of measure over different members in a dimension. The angle of pie determines the measured value. Different colors can be assigned to pie to represent the members in a dimension.

Step 1) Go to a new Worksheet

- Select Segment and sales from the data pane.

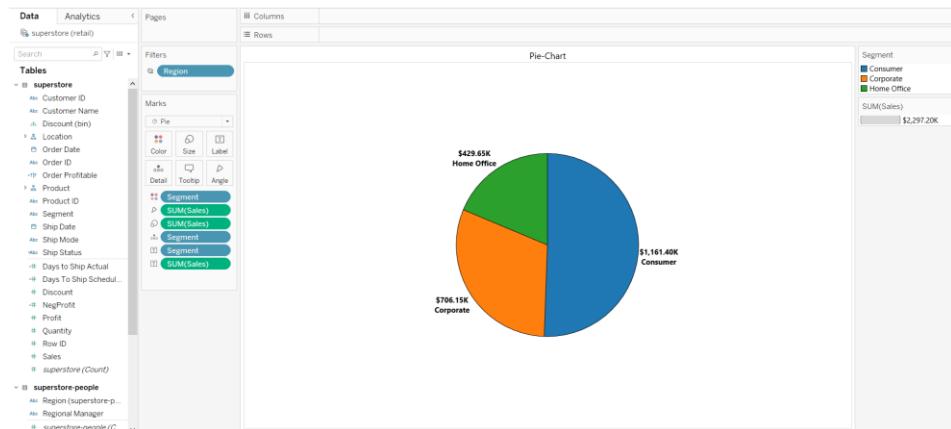


Step 2) Click on the 'Show Me' button present in the top right corner of the worksheet.

- Select 'Pie Chart' from the list.



It creates a Pie Chart as shown below.

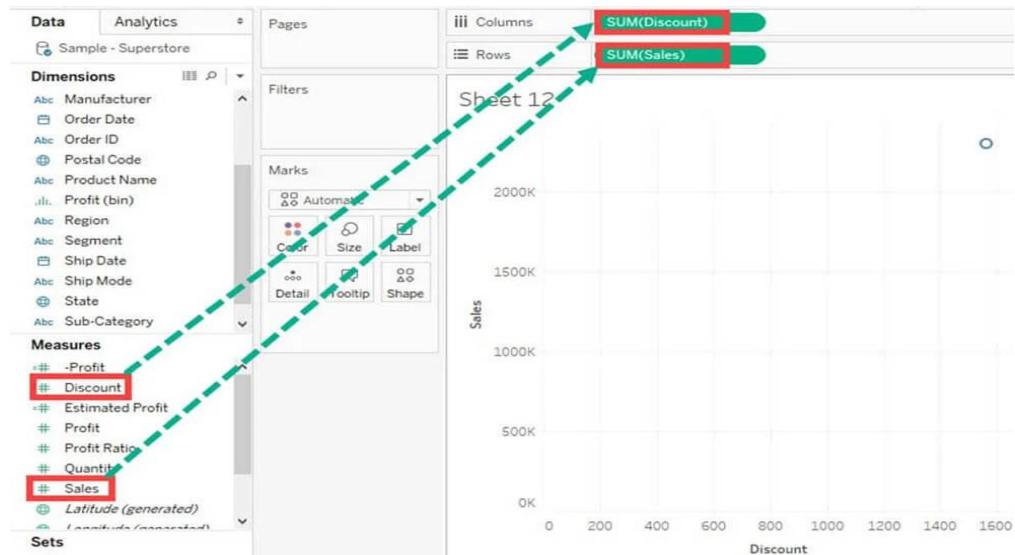


Activity 1.5: Scatter Plot:

The relationship between two measures can be visualized using a scatter plot. A scatter plot is designed by adding measures in both x and y-axis. This can show the trend or relationship between the measures selected. A scatter plot can be designed by following the below procedure.

Step 1) Go to a new Worksheet.

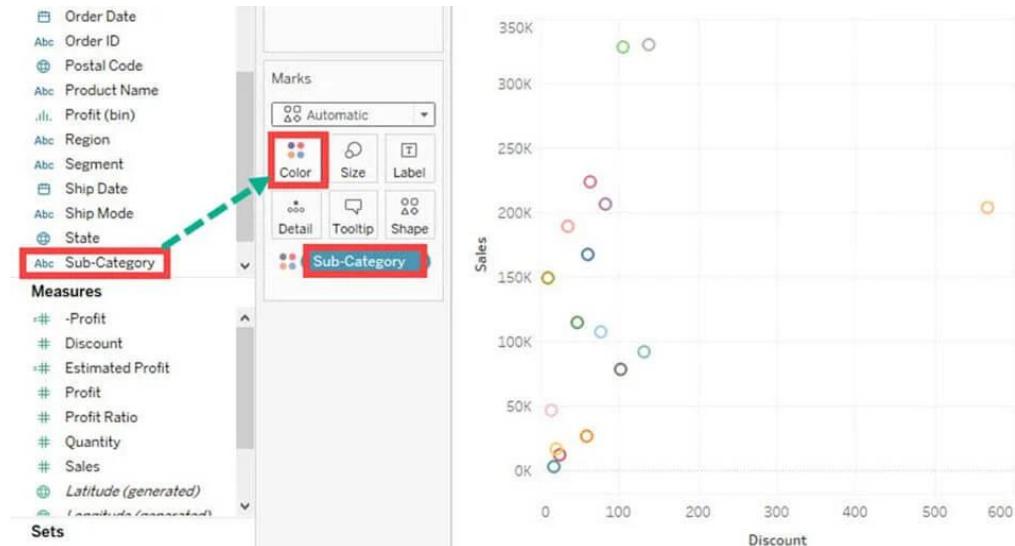
- Drag 'Discount' into Columns.
- Drag 'Sales' into Rows.



This creates a scatter plot by default.

Step 2)

- Drag 'Sub-Category' into Color icon present in the Marks card.
- It creates a scatter plot showing the relationship between discount and sales for each sub-category.



Activity 1.6: Area Chart:

Area charts can represent any quantitative (measure) data over different periods of time. It is basically a line graph where the area between line and axis is generally filled with color. The procedure to create an area chart is given below.

Step 1) Go to a new Worksheet.

- Hold the control key on the keyboard and select 'Order Date' and 'Quantity'.

Dimensions

- Abc Category
- ⊕ City
- ⊕ Country
- Abc Customer Name
- .Abc Discount (bin)
- Abc Manufacturer
- ⊕ **Order Date**
- Abc Order ID
- ⊕ Postal Code
- Abc Product Name
- .Abc Profit (bin)
- Abc Region

Measures

- =# -Profit
- # Discount
- =# Estimated Profit
- # Profit
- # Profit Ratio
- # Quantity**
- # Sales
- ⊕ Latitude (generated)
- ⊕ Longitude (generated)

Step 2)

Click on the 'Show Me' option present at the top right corner of the worksheet.

Select the Area chart icon as shown in the figure.



Step 3)

Drag 'Region' from the dimension pane and add it in the color icon of the Marks card.

This creates an area chart as shown in the image.

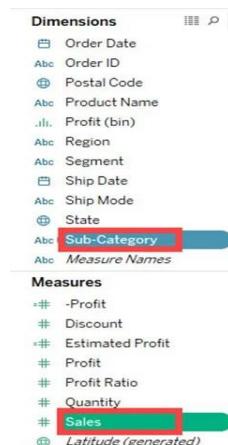


Activity 1.7: Bubble Chart:

A bubble chart visualizes the measures and dimensions in the form of bubbles. The sizes of the bubbles determine the size of measure value for effective visualization. The color of bubbles can be set to differentiate the members present in a dimension. The procedure to create a bubble chart is given below.

Step 1) Go to a new Worksheet.

- Hold on the Control key on the keyboard.
- Click on 'Sub-Category' and 'Sales'.



Step 2)

- Click on the 'Show Me' option present in the top right corner of the worksheet.
- Select the 'Packed bubbles' icon as shown in the image.



It creates a bubble chart as shown below

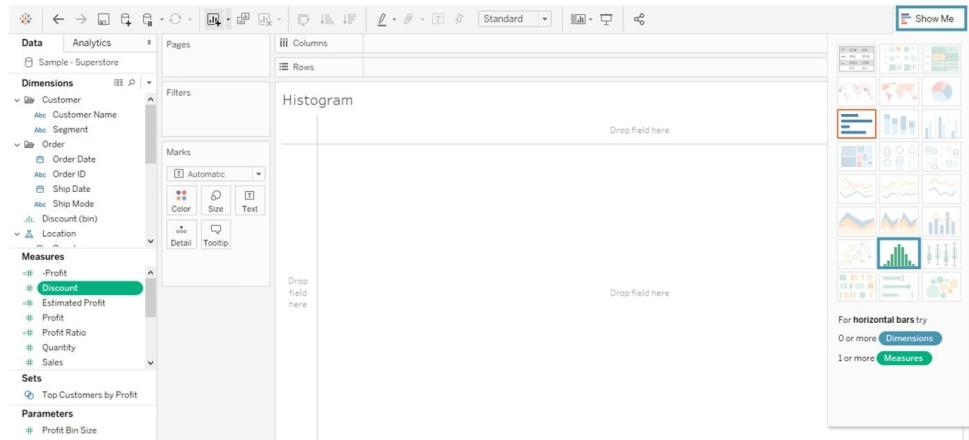


Activity 1.8: Histogram:

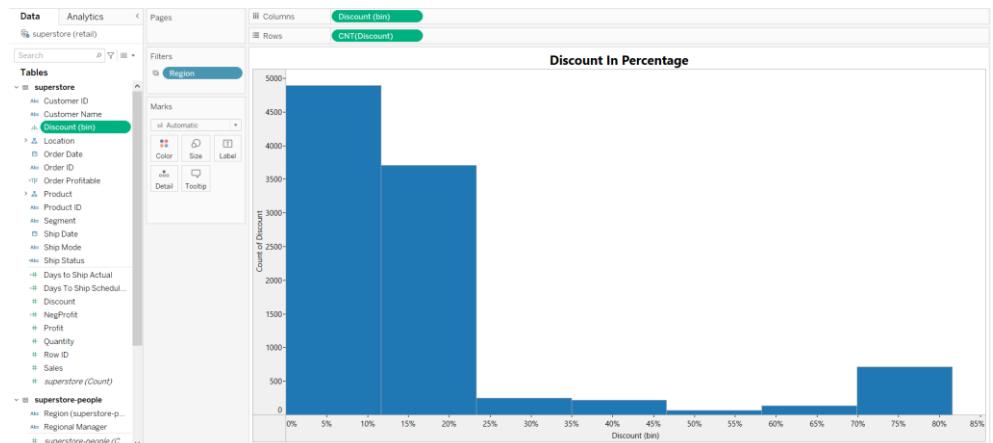
Histograms show how your data is distributed across distinct groups. By grouping your data into specific categories (also known as "bins"), then plotting the number of records in a category as a vertical bar, you can quickly see which bins the majority of your data falls in. The histogram is your best option for visualizing how data fall into categories. For example, the number of customers by company size, student performance on an exam, and frequency of a product defect.

The procedure to create a histogram is shown below.

- Go to a new Worksheet.
- Select Discount from the measures.
- Click on the Show Me button present in the top right corner of the worksheet.



Select the Histogram icon as shown in the image.



It creates a histogram chart.

Activity 1.9: Dual Axis Chart:

Dual axis charts can be used to visualize two different measures in two different chart types. A date column and two measures are necessary to build a dual axis chart. The different scales used in the graph helps the user to understand both measures. The procedure to create a dual axis chart is shown below.

Step 1) Go to a new Worksheet.

- Hold the control key and select 'Order Date', 'Sales' and 'Quantity'.

The screenshot shows a list of dimensions and measures. Dimensions include Category, City, Country, Customer Name, Discount (bin), Manufacturer, Order Date (highlighted with a red box), Order ID, Postal Code, Product Name, Profit (bin), and Region. Measures include -Profit, Discount, Estimated Profit, Profit, Profit Ratio, Quantity (highlighted with a red box), Sales (highlighted with a green box), and Latitude (generated). A horizontal bar at the bottom indicates the selection of 'Quantity' and 'Sales'.

Dimensions

- Category
- City
- Country
- Customer Name
- Discount (bin)
- Manufacturer
- Order Date**
- Order ID
- Postal Code
- Product Name
- Profit (bin)
- Region

Measures

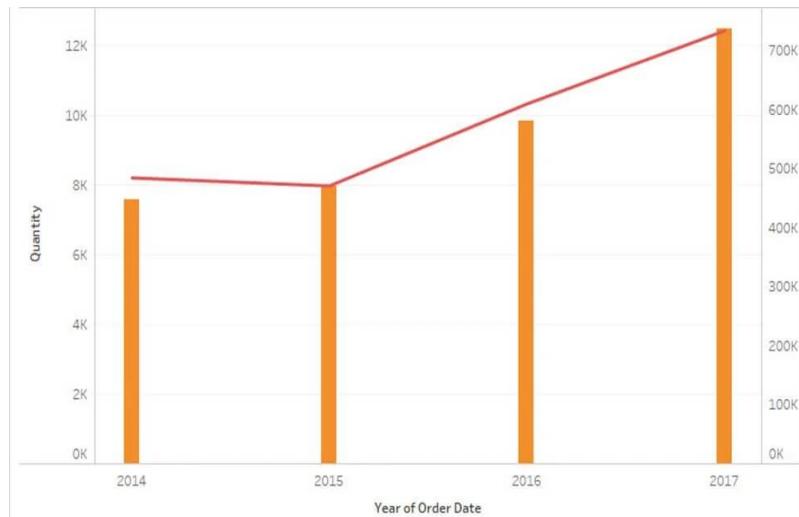
- Profit
- Discount
- Estimated Profit
- Profit
- Profit Ratio
- Quantity**
- Sales**
- Latitude (generated)

Step 2)

- Click on the 'Show Me' option present in the top right corner of the Worksheet.
- Select 'dual combination' icon as shown in the image.



It creates a dual axis chart as shown in the figure.

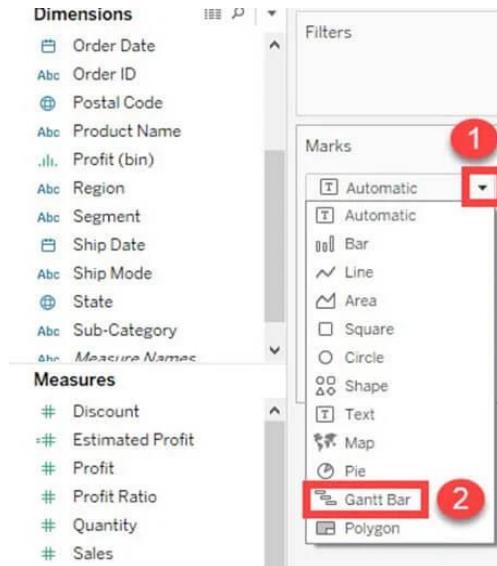


Activity 1.10: Gantt chart:

A Gantt chart can show the comparison of data between categories. It can be used to identify the time taken for each process. In the below example, time taken for shipment by each type of ship mode is given. The procedure to create a Gantt chart is given as follows.

Step 1) Go to a new Worksheet.

- Click on the drop down button in the Marks Card.
- Select 'Gantt Bar' from the list.



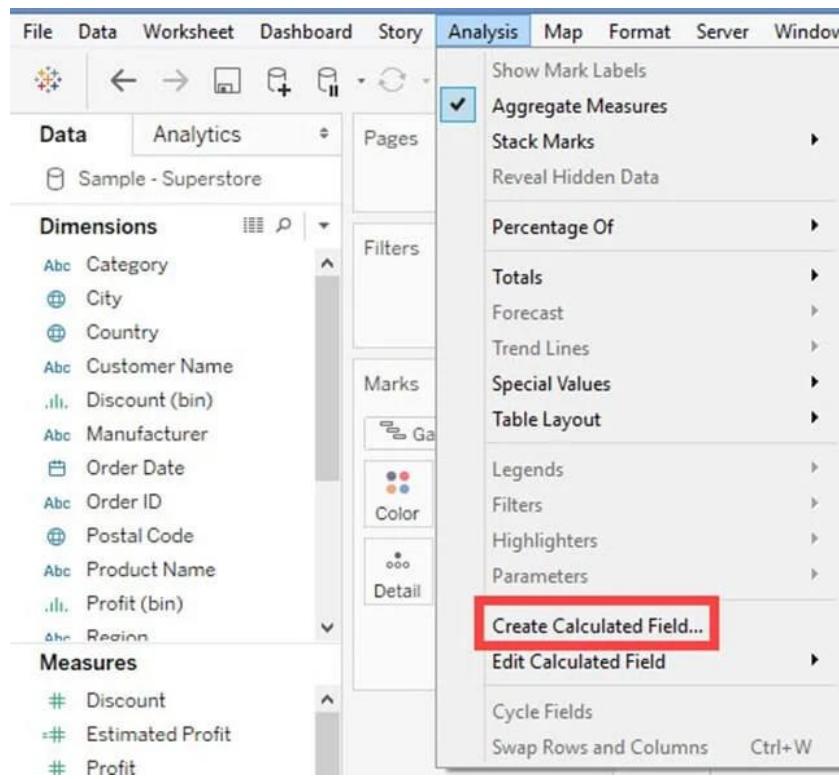
Step 2)

- Drag 'Order Date' into Columns.
- Right Click on the Order date and select the 'Day' as shown in the image.

The screenshot shows the Tableau interface with the 'Order Date' dimension selected and being dragged into the columns shelf. The context menu for 'Order Date' is open, with the 'Day' option highlighted and selected. The menu also includes options like 'Filter...', 'Format...', 'Show Header', 'Include in Tooltip', 'Show Missing Values', and 'Edit in Shelf'.

Step 3)

- Click on 'Analysis' Present in the Menu bar.
- Select 'Create Calculated Field' present in the list.



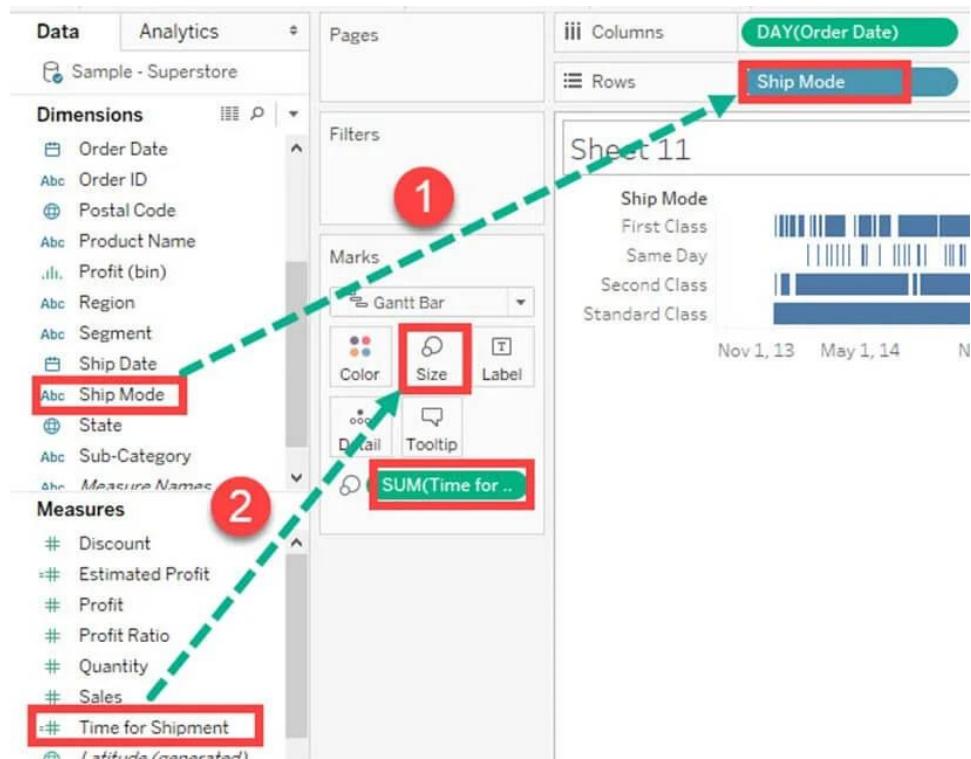
Step 4)

- Enter the name of calculated field as 'Time for Shipment'
- Type the formula as shown in the image to create a difference between order date and ship date.
- Click on OK.

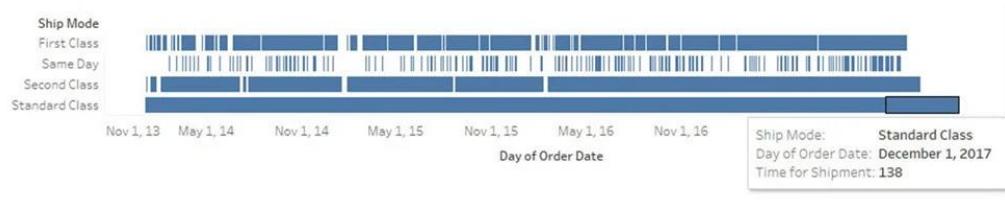


Step 5)

- Drag 'Ship Mode' into Rows.
- Drag 'Time for Shipment' into the size icon present in the marks card.



Step 6) This creates a Gantt chart. It shows the time taken for each shipment across different ship mode.



Activity 1.11: Heat Map:

A heat map can visualize the data in the form of size as well as color on different measures. Two different measures can be visualized simultaneously using heat maps. One measure can be assigned to size whereas another measure can be assigned to the color of the heat map. The procedure to create heat map is given as follows:

Step 1) Go to a new Worksheet.

- Hold the control key on the keyboard and select 'Sub-Category' and 'Sales' from the data pane.

Dimensions

- Abc Manufacturer
- Order Date
- Order ID
- Postal Code
- Product Name
- Region
- Segment
- Ship Date
- Ship Mode
- State
- Sub-Category**
- Measure Names

Measures

- # Discount
- =# Estimated Profit
- # Profit
- # Profit Ratio
- # Quantity
- Sales**
- # Latitude (generated)
- # Longitude (generated)
- # Number of Products

Sets

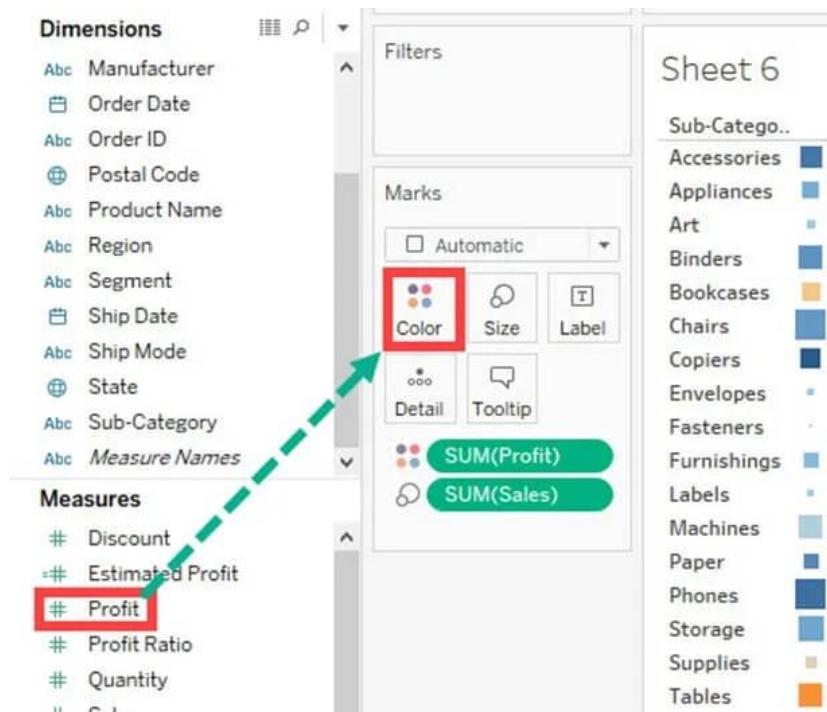
- User Filter 1

Step 2)

- Click on the 'Show me' button present in the top right corner of the worksheet.
- Select the 'Heat Map' icon as shown in the image.

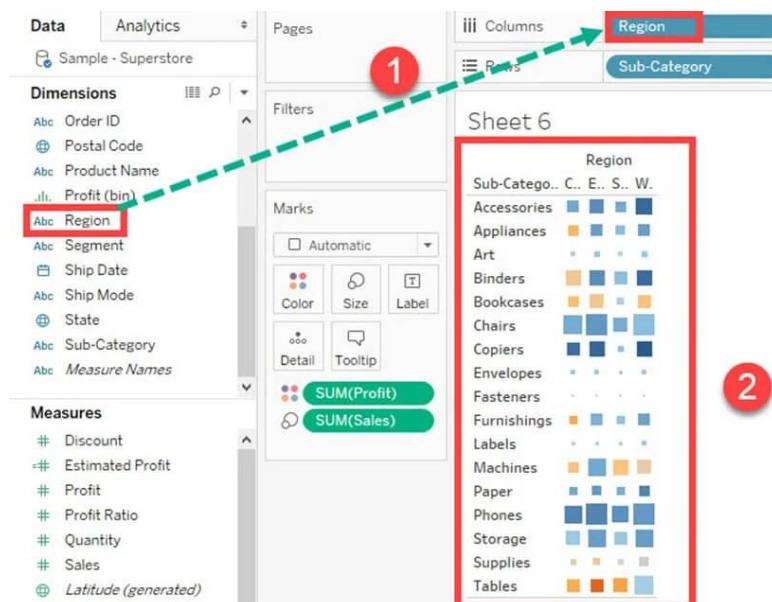


Step 3) Drag 'Profit' into the Color box.



Step 4)

- Drag 'Region' into Columns.
- This will create a Heat map. It can be used to visualize Sales and Profit across different dimensions.

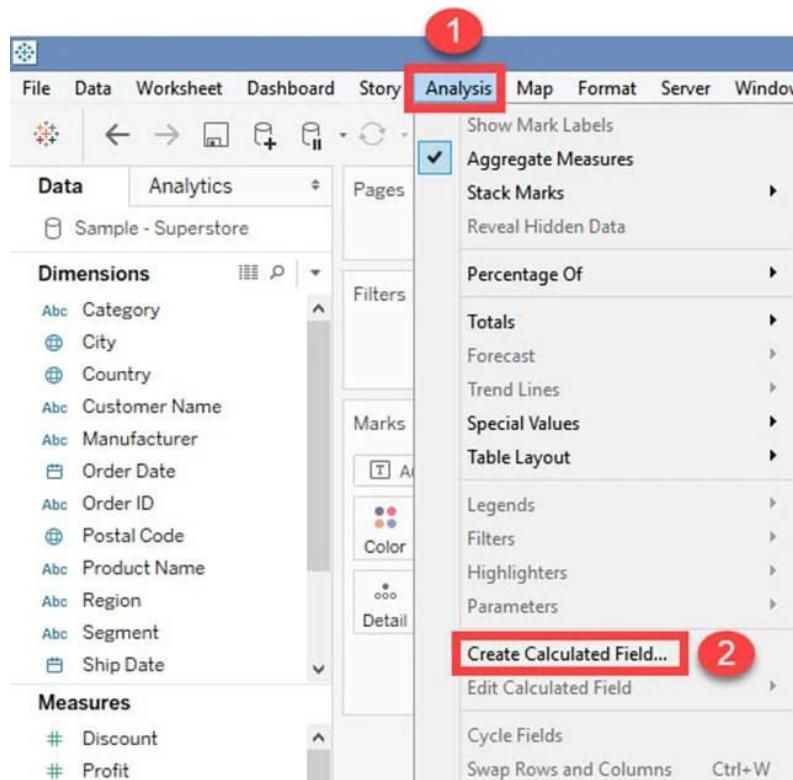


Activity 1.12: Bullet Chart:

A bullet chart can be used as a gauge or indicator to show the performance of measures. Two measures can be compared to each other using the bullet graph. For example, if we are having estimated profit and actual profit we can compare both of them using a bullet chart. The procedure to create a bullet chart is given as follows.

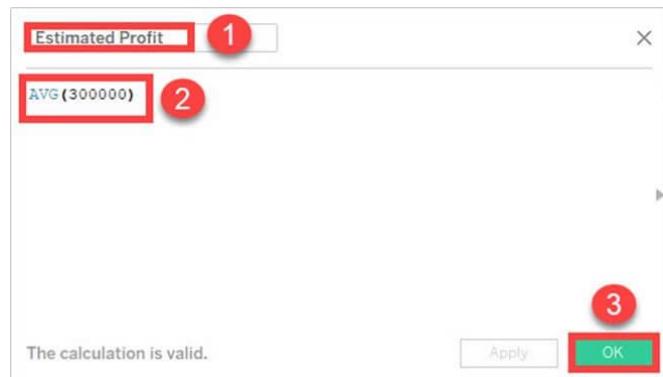
Step 1) Go to a new Worksheet.

- Click on 'Analysis' present in the Menu bar.
- Select 'Create Calculated Field' from the list.



Step 2) It opens a Calculated Field Window.

- Enter a name for the calculated field. In this example, it is named as 'Estimated Profit'.
- Type the estimated value of the measure. In this example, Profit is taken as the measure. So the calculated field is created for estimated profit.
- Click on OK.



Step 3) Go to measures in the Data pane. Hold the control key on the keyboard and select 'Estimated Profit' and 'Profit'.

Step 4)

- Click on the 'Show me' button present in the top right corner of the worksheet.
- Select the bullet graph icon as shown in the image.



It creates a bullet graph as shown below.

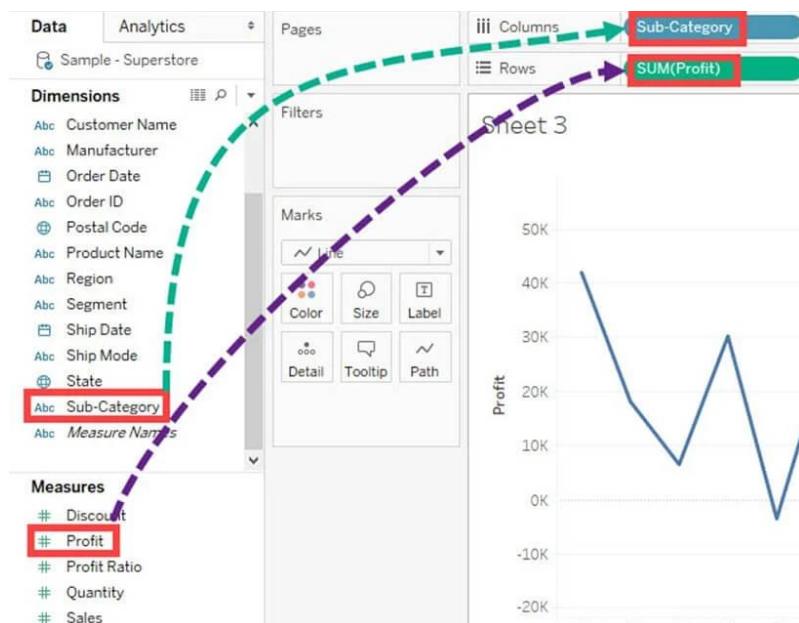


Activity 1.13: Pareto Chart

A Pareto chart consists of both bar and line graphs. The same measure is used to create the graphs but the measure values are manipulated differently. The purpose of using Pareto Chart in Tableau is to identify the contribution of members present in a field. For example, Profit contributed by different subcategory of products in a retail store can be analyzed using Pareto Chart. It can show the top members and their contribution. The procedure to create a Pareto Chart is given as follows.

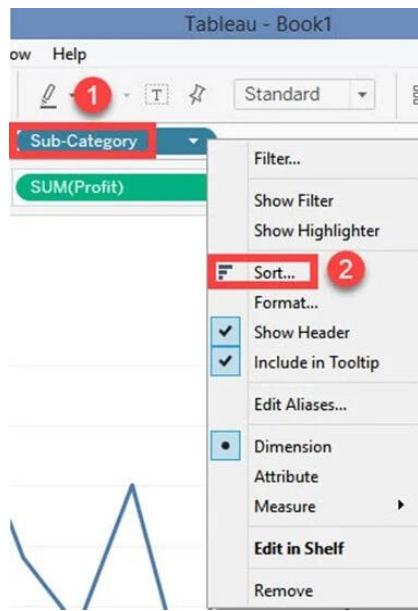
Step 1) Go to a new Worksheet.

- Drag 'Sub-Category' into Columns.
- Drag 'Profit' into Rows.



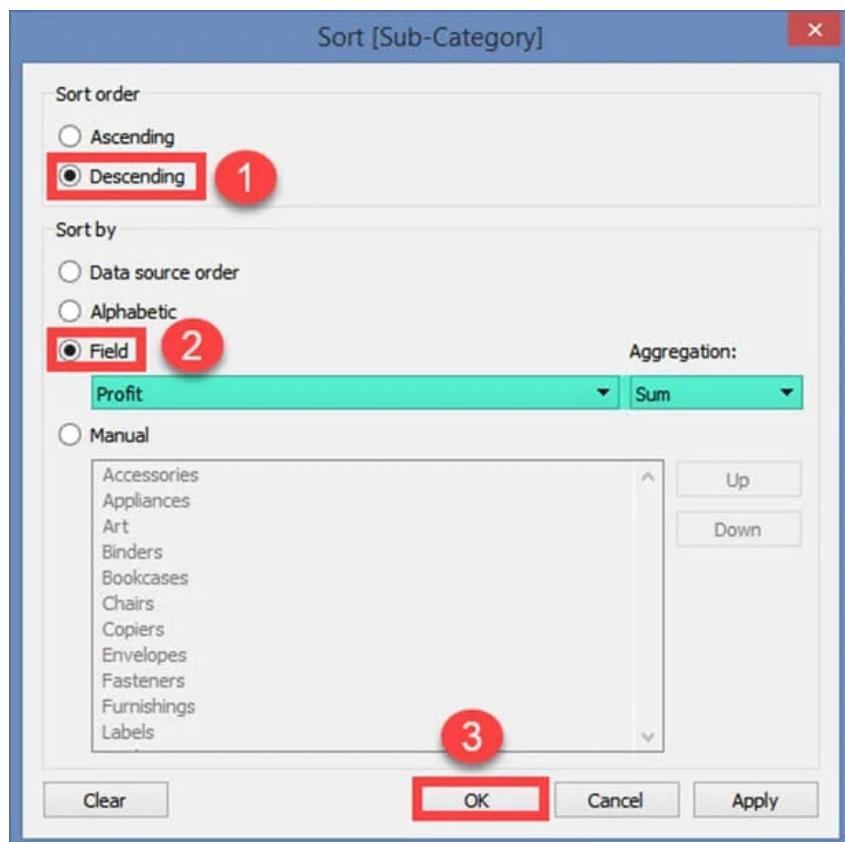
Step 2) Right click on 'Sub-Category'.

- Select 'Sort' option from the list.



Step 3) It opens a Sort Window.

- Click on 'Descending' in Sort order.
- Select 'Field' in 'Sort by' section. Select the field as Profit and choose 'Sum' as aggregation.
- Click on OK.

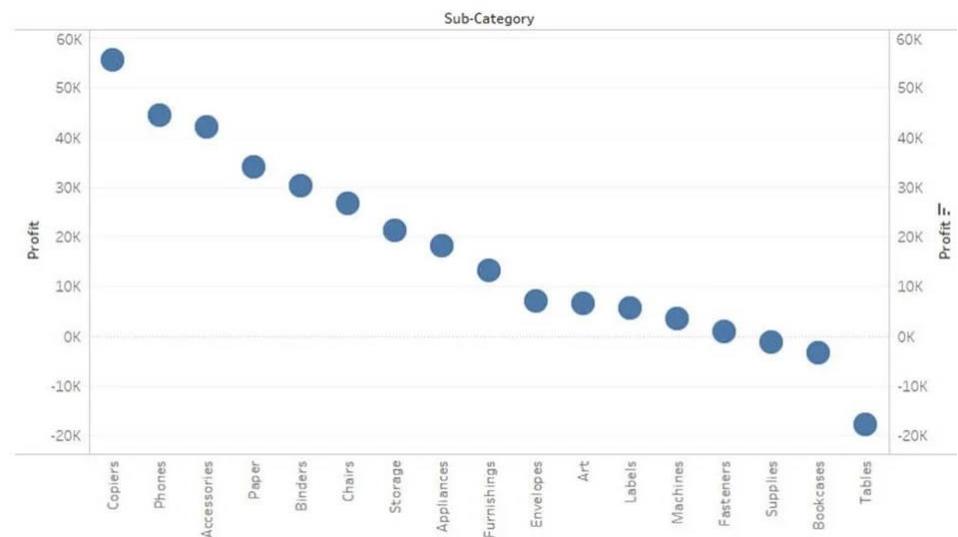


Step 4) Drag 'Profit' again into Rows.

- Right-click on the newly added 'Profit' and Select 'Dual axis' option.

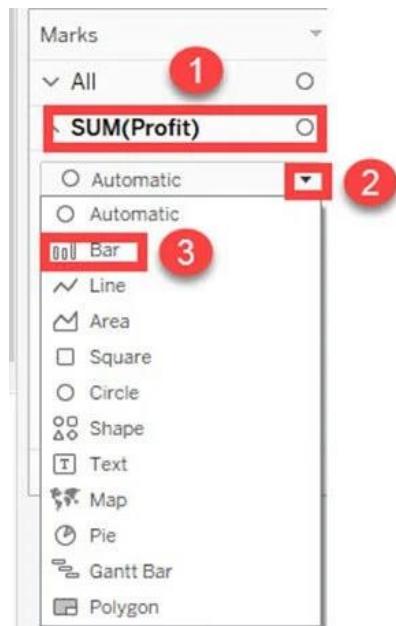


It merges the x-axis of both measures and converts the visualization as given below.



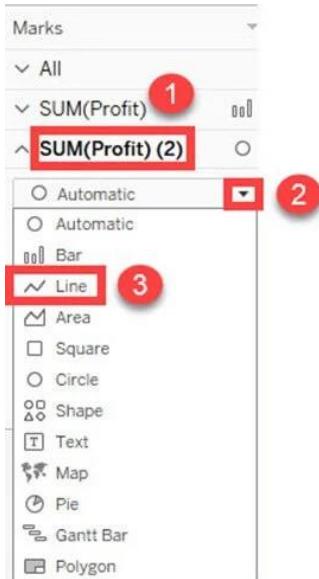
Step 5) Go to the Marks Card.

- Select 'SUM(Profit)' for the marks card list.
- Click on the drop-down button as shown in the image.
- Select 'Bar' as the chart type.



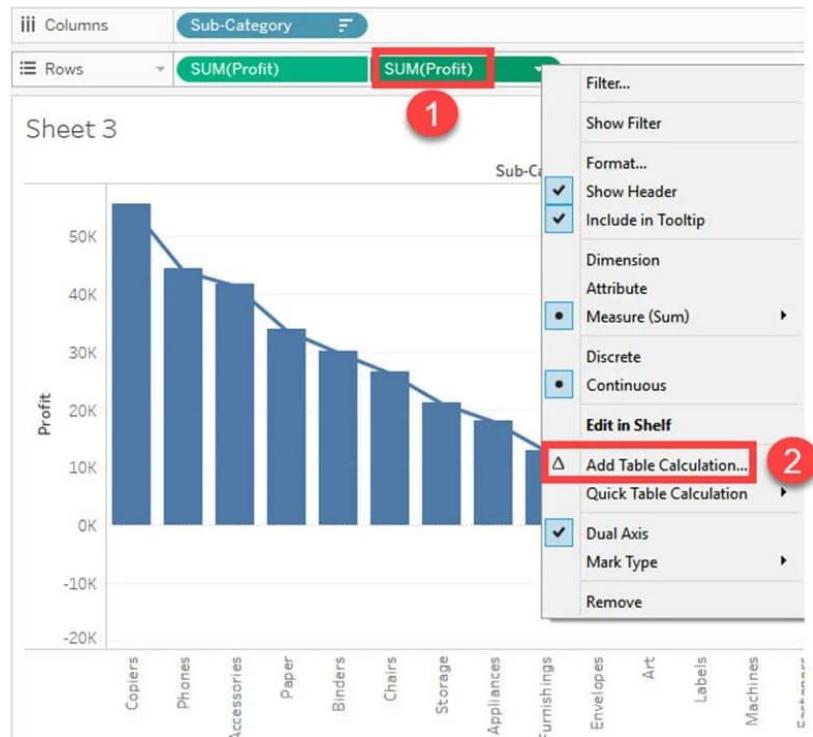
Step 6) Select ‘SUM(Profit)(2)’ from the marks card list.

- Select the drop-down button as given in the image.
- Click on ‘Line’ from the list.



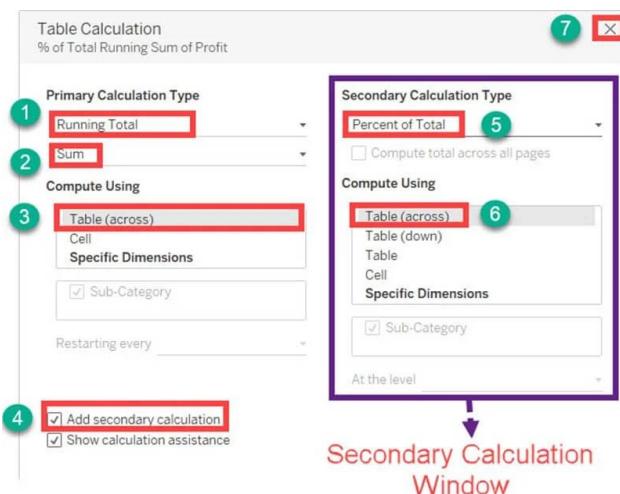
Step 7) Select ‘SUM(Profit)’ on the right side of rows as shown in the image.

- Right click on it and select ‘Add Table Calculation’ from the list.



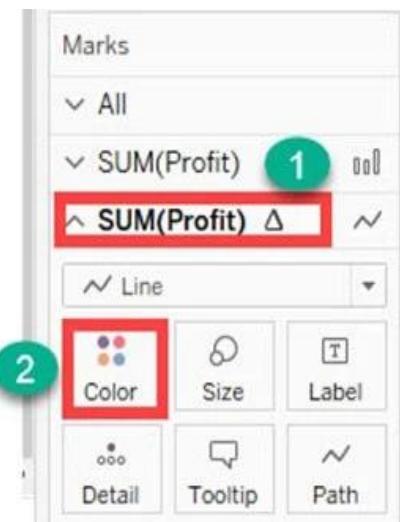
Step 8) It opens the Primary Calculation Type window.

- Select 'Running Total' from the drop-down.
- Select 'Sum' as aggregation from the drop-down.
- Click on Compute Using 'Table (across)'.
- Check in the 'Add Secondary Calculation' box.
- It expands the window for 'Secondary Calculation Type'. Select 'Percent of Total' from the dropdown list.
- Select on Compute Using 'Table (across)'
- Now close the Window by clicking on the close icon as shown in the image.



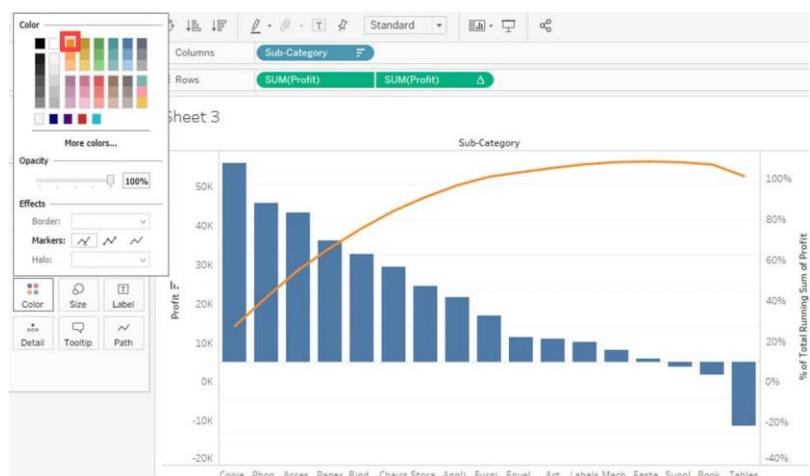
Step 9) Go to the last marks card namely ‘SUM (Profit)’.

- Click on the color icon present in the marks card.



Step 10) Select any color of your choice.

- This changes the color of the line present in the graph.

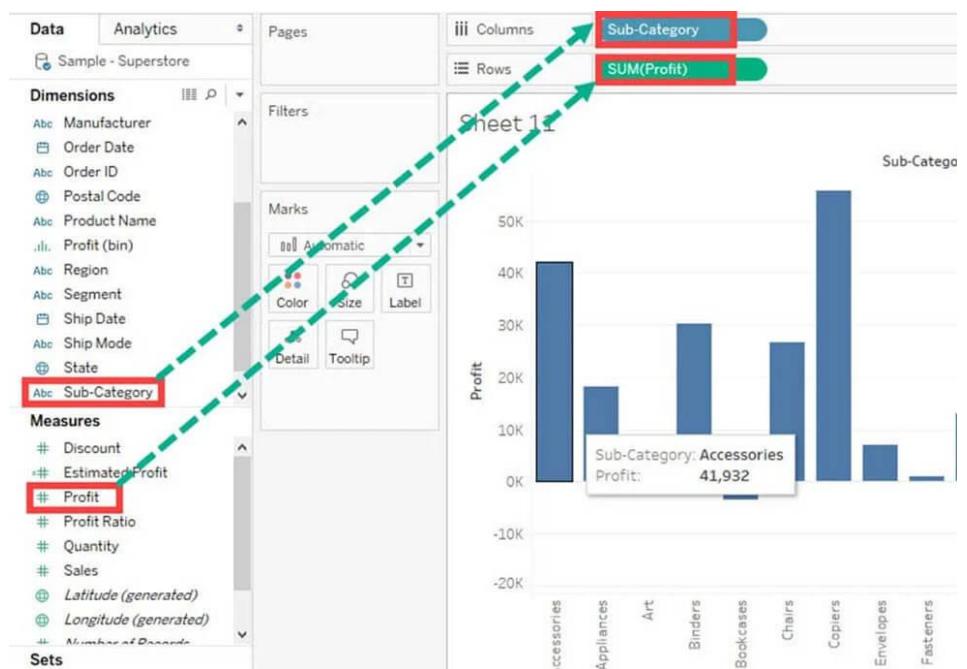


Activity 1.14: Waterfall Chart:

Waterfall Chart can visualize the cumulative effect of a measure over dimension. It can show the contribution of growth or decline by each member in a dimension. For example, you can see the contribution of profit by each sub-category using a waterfall chart. The waterfall chart can be designed in Tableau by following the given procedures.

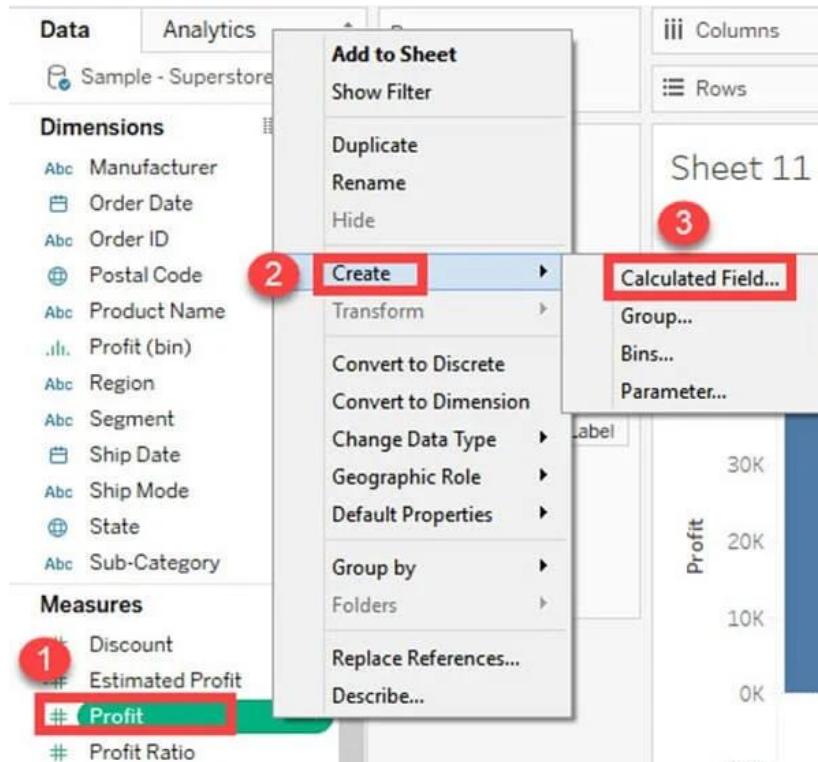
Step 1) Go to a new Worksheet.

- Drag Sub-Category into Columns
- Drag Profit into Rows.



Step 2)

- Right click on 'Profit' present in the Measures Pane.
- Choose 'Create' from the list.
- Select the 'Calculated Field' Option.

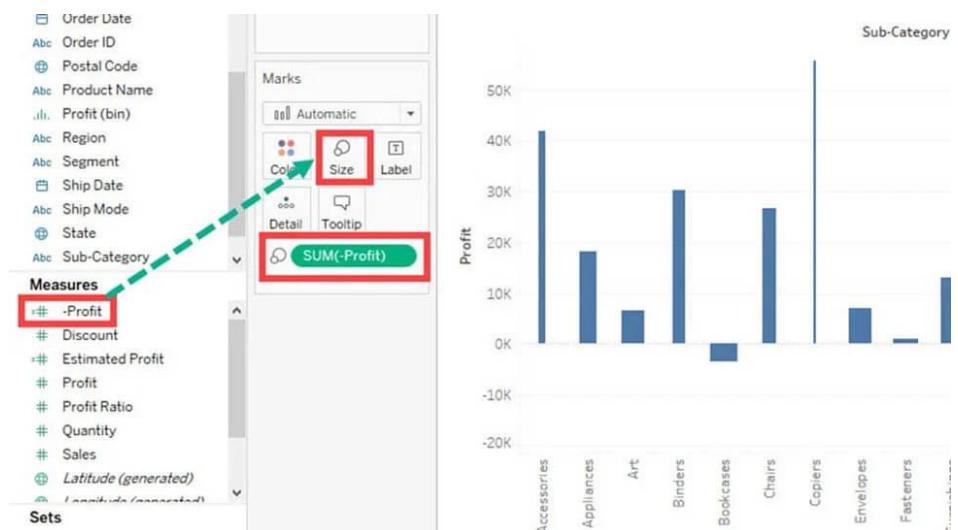


Step 3) It opens the ‘Calculated Field’ Window.

- Enter the name of the Calculated field as ‘-Profit.’
- Write the formula as shown in the image.
- Click on OK.

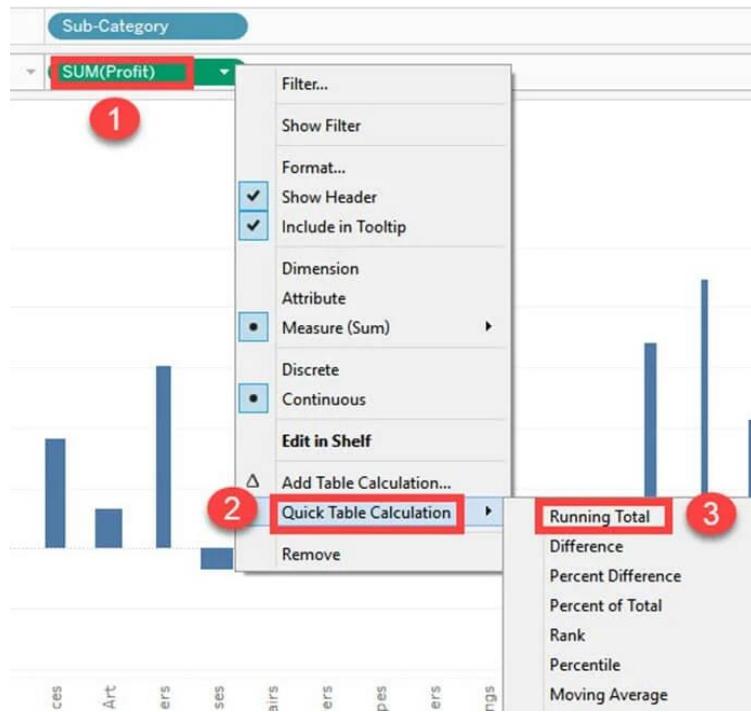


Step 4) Drag the newly created calculated field ‘-Profit’ into the size option present in the marks card.



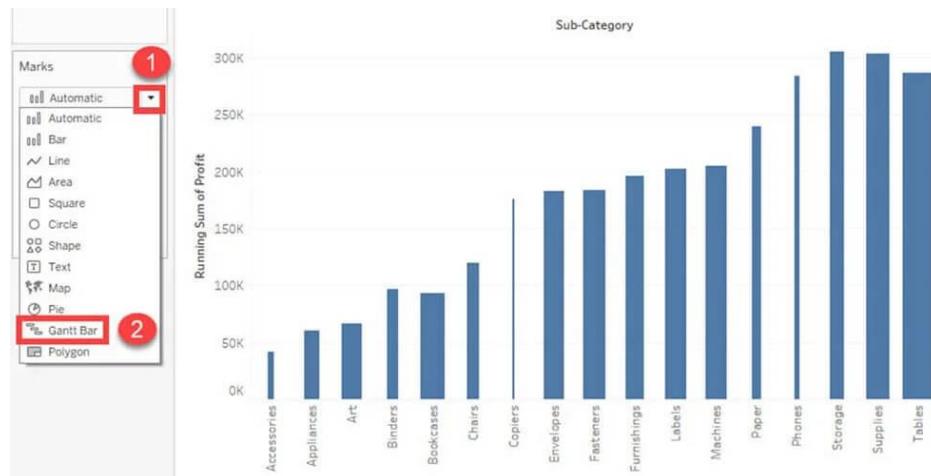
Step 5)

- Right, Click on ‘SUM (Profit)’ present in the Rows.
- Select ‘Quick Table Calculation’ from the list.
- Click on ‘Running Total’ option.

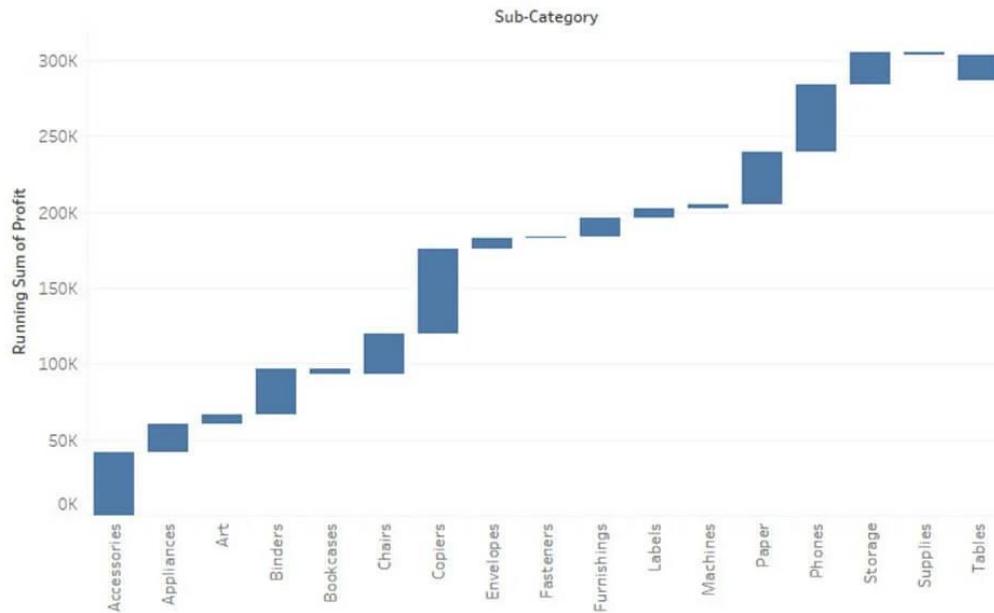


Step 6)

- Click on the drop-down option present on the marks card.
- Select 'Gantt Chart' from the list.



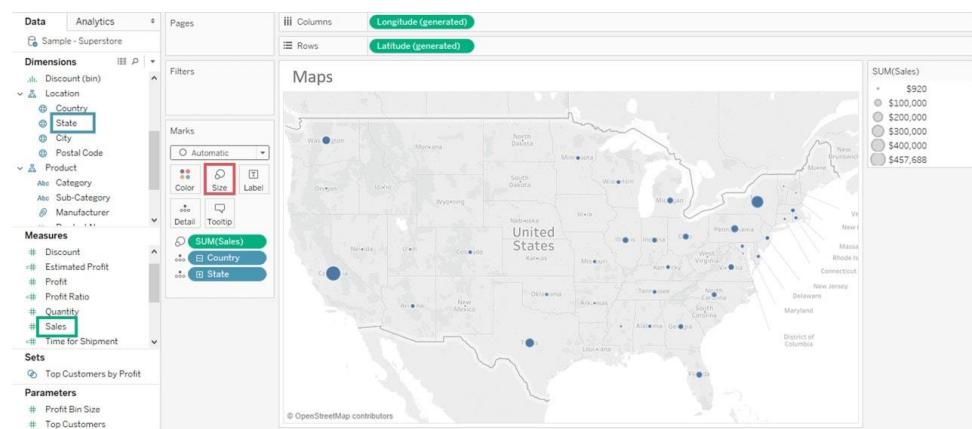
This creates a waterfall chart as shown below.



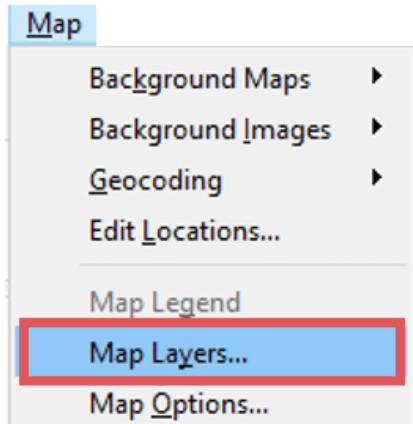
Activity 1.15: Maps

These are a no-brainer for visualizing any kind of location information, whether it's postal codes, state abbreviations, country names, or your own custom geocoding. Maps highlight geographic trends in a format everyone knows and understands. If you have geographic information associated with your data, maps are a simple and compelling way to show how location correlates with trends in your data. For example, insurance claims by state, product export destinations by country, car accidents by zip code, and custom sales territories.

- Navigate to a worksheet.
- In the Data pane, under Dimensions, double-click State.
- From Measures, drag Sales to Size on the Marks card. The data points on the map update to show the number of sales proportionally.



Select Maps, followed by Map Layers.



In the Map Layers pane, do the following:

- Click the Style drop-down and select Normal.
- Under Map Layers, clear Country/Region Names.



Activity 1.16: Box and Whisker Plots

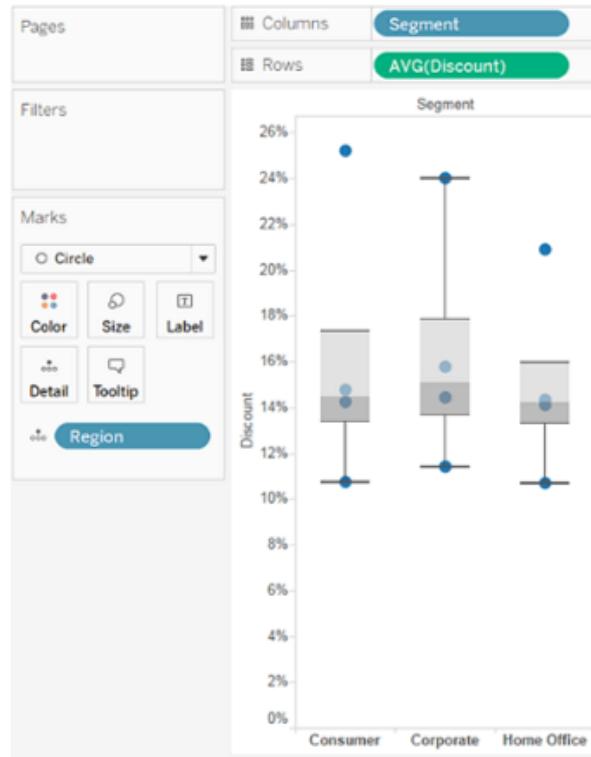
Box-and-whisker plots, or boxplots, are a common way to show distributions of data. The name refers to the two parts of the diagram: the box, which contains the median of the data along with the 1st and 3rd quartiles (25% greater and less than the median), and the whiskers, which typically represents data within 1.5 times the interquartile range (the difference between the 1st and 3rd quartiles). The whiskers can also be used to also show the maximum and minimum points within the data. Use box-and-whisker diagrams to understand your data at a glance. See how data is skewed towards one end and identify outliers in your data. For example, comparing scores between sites, analyzing data before and after a process change, or examining data from duplicate machines manufacturing the same products.

Open a new Worksheet.

- Drag the Segment dimension to Columns.
- Drag the Discount measure to Rows. Tableau creates a vertical axis and displays a bar chart—the default chart type when there is a dimension on the Columns shelf and a measure on the Rows shelf.
- Drag the Region dimension to Columns, and drop it to the right of Segment. Now you have a two-level hierarchy of dimensions from left to right in the view, with regions (listed along the bottom) nested within segments (listed across the top).
- Click Show Me in the toolbar, then select the box-and-whisker plot chart type.
- Drag Region from the Marks card back to Columns, to the right of Segment.



Tableau displays a box plot. The horizontal lines are flattened box plots, which is what happens when boxplots are based on a single mark. Box plots are intended to show a distribution of data, and that can be difficult when data is aggregated, as in the current view.



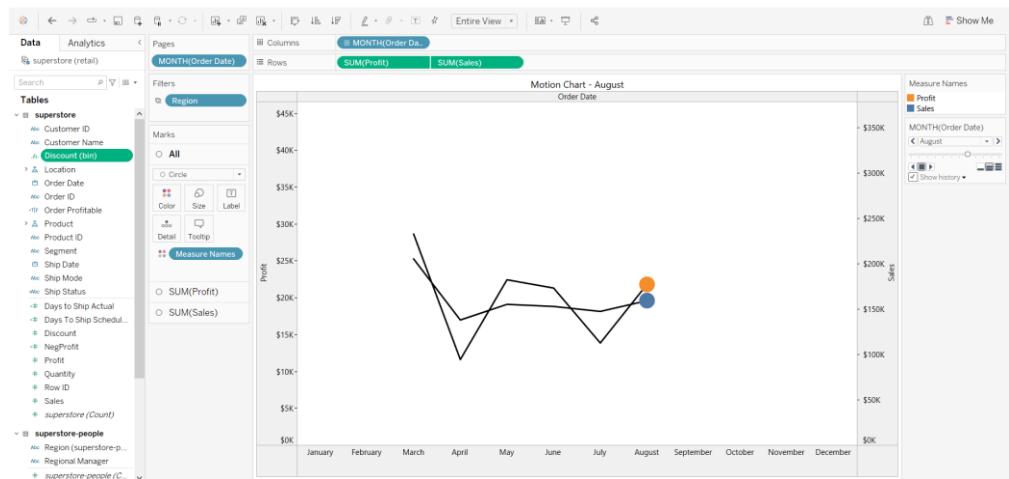
To disaggregate data, select Analysis > Aggregate Measures. This command turns aggregation on or off, and because data is aggregated by default in Tableau, the first time you select this command, it disaggregates the data. Now, instead of a single mark for each column in the view, you see a range of marks, one for each row in your data source.



Activity 1.17: Motion Charts

Motion chart is used to show the data using X-axes and Y-axes that display the change over time by showing the movement of the data points as well as variations in the color of the lines. The motion chart has the advantage to view the trail of how the data has changed over time. Motion chart needs only one Time Dimension and one Measure in tableau.

- Create a Trend Chart with X-axis as the Order Date (in the format of Month) and Sales and Profit are the Measures
- All you need to do to make the Motion Chart is drag Order Date over to the Pages shelf and change the format again to match with the X-axis.
- Change the Mark Type from Automatic to Circle.
- Go to Show History and select Trails to view the trend change. Your Motion Chart is ready for launch.



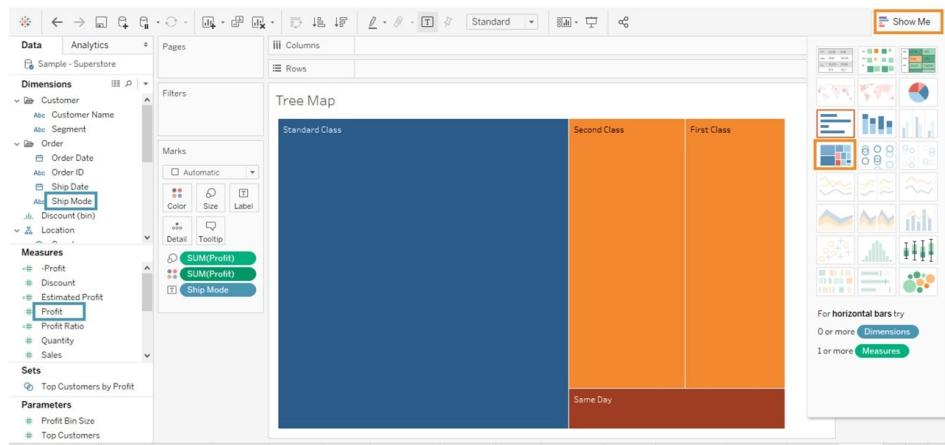
Press on the arrow buttons to see the motions, change the Show History customizations, the speed etc.

Activity 1.18: Tree Maps

Treemaps relate different segments of your data to the whole. By nesting rectangles within others, treemaps show how individual data points fit in a hierarchy. As the name of the chart suggests, each rectangle is subdivided into smaller rectangles, or sub-branches, based on its proportion to the whole. They make efficient use of space to show the percent total for each category. When a dataset can be broken down in many different ways, a treemap might be the best way to show what categories most of the data falls in. For example, storage usage across computer machines, managing the number and priority of technical support cases and comparing fiscal budgets between years.

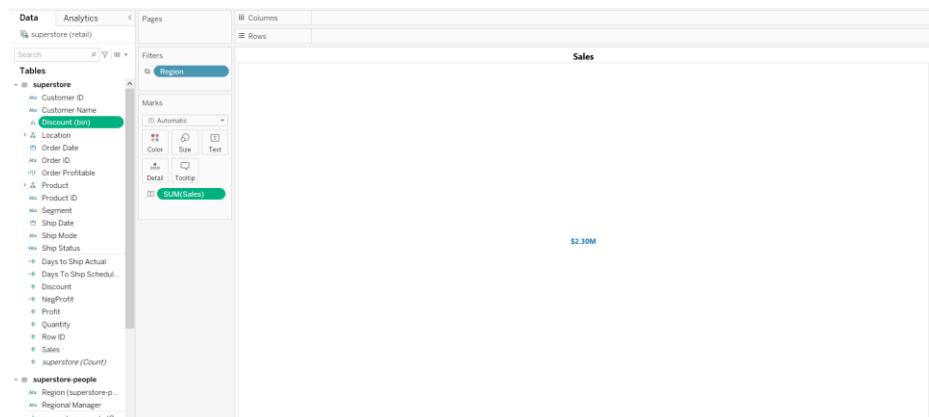
To achieve this objective, the following are the steps.

- Drag and drop the measure Profit two times to the Marks Card. Once to the Size shelf and again to the Color shelf.
- Drag and drop the dimension Ship Mode to the Label shelf. Choose the chart type TreeMap from Show Me and the following chart appears.



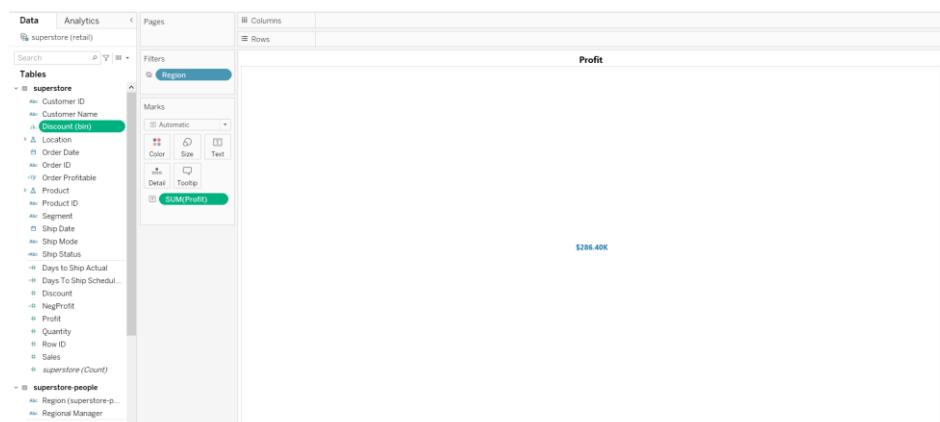
Activity 1.19: Sales

Calculate total no of Sales. The Superstore data set is a data set that simulates the sales data of a company that sells office products and furniture. The "Sales" column contains the total revenue generated from each sale. Drag sales in Text label and format the data for sales such as currency,display units etc.



Activity 1.20: Profit

Calculate total no of Profit . The "Profit" column contains the profit generated from each sale. Drag profit in Text label and format the data for profit such as currency,display units etc.



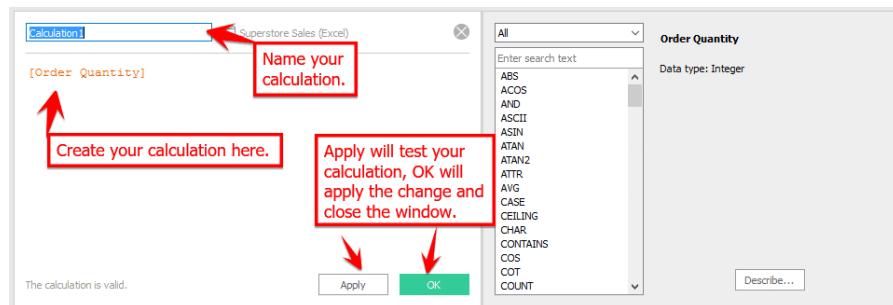
Activity 1.21: Profit Ratio

To sum all of the profit figures as well as sum all of the sales figures and then divide by the totals, the calculation on Tableau calculated field looks like: $\text{Sum}([\text{Profit}])/\text{Sum}([\text{Sales}])$. Using calculations in Tableau is where you truly start to take your analysis to the next level.

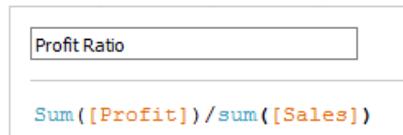
Create A Calculated Field In Tableau

Here are the simple steps you can follow on how to add calculated field in Tableau and when you want to create a calculation:

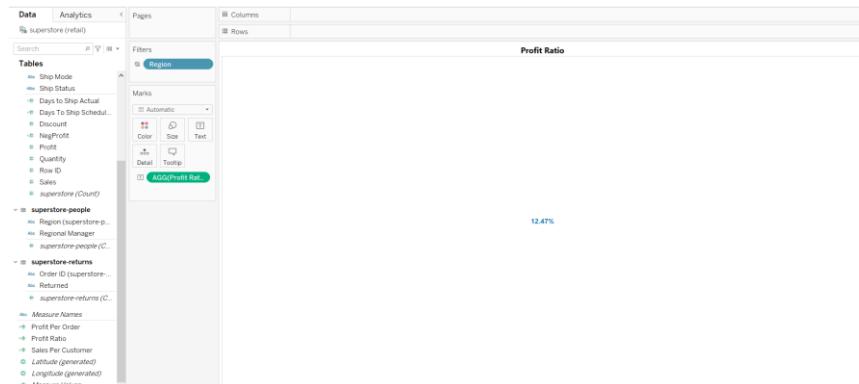
- Right click on create > Tableau Calculated Field, or click the drop down arrow next to Dimensions.
- Name your new Tableau calculation by writing over the box that says Calculation #.
- In the Tableau calculated field window below, write your calculation.
- Click Apply to test the effect of your, and click Ok to apply the changes and to close out of the calculation window.



To sum all of the profit figures as well as sum all of the sales figures and then divide by the totals, the calculation on Tableau calculated field looks like: $\text{Sum}([\text{Profit}])/\text{Sum}([\text{Sales}])$. Tableau now knows to sum the figures first and then calculate the ratio, rather than sum all the individual ratios.



Drag profit ratio in Text label and format the data for profit ratio in percentage.

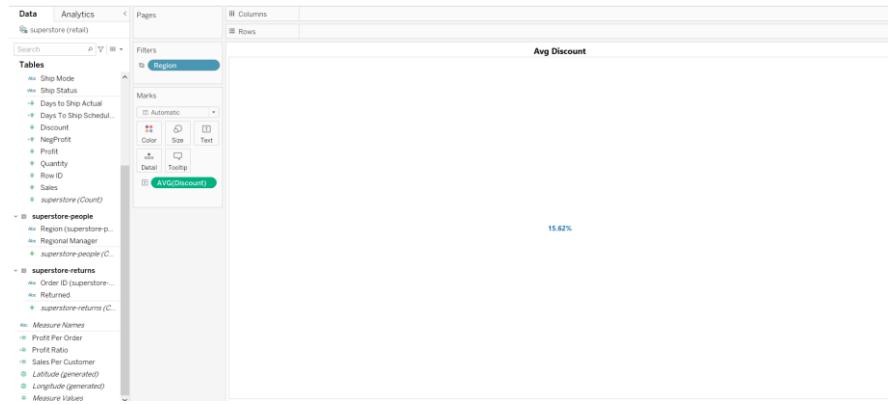


Activity 1.22: Average Discount

The average discount in the Superstore data set refers to the average percentage reduction in the price of a product, given as a decimal value.

We can use the built-in tableau function 'AVG' to calculate the average.

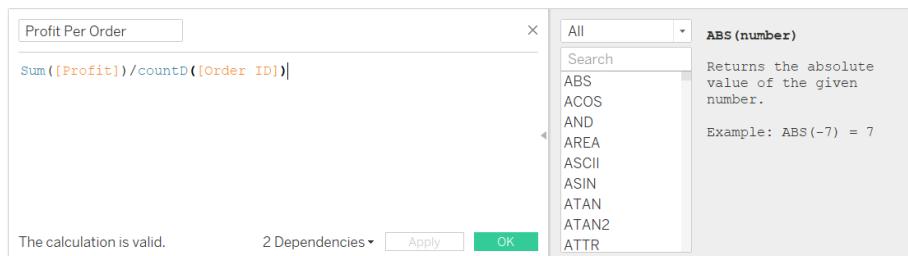
Drag Average Discount in Text label and format the data for Average Discount in percentage.



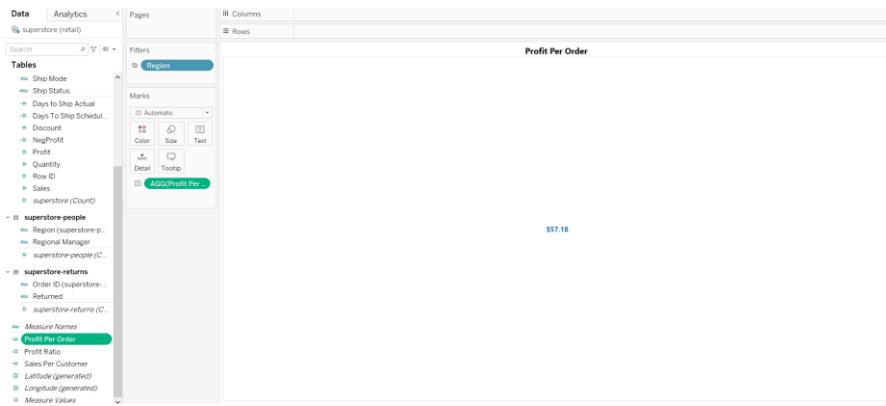
Activity 1.23: Profit per order

Profit per order, also known as profit margin, is a metric that represents the amount of profit earned for each order placed. It can be calculated by taking the profit generated from each sale and dividing it by the number of orders. This metric can be used to understand the profitability of each order and how it varies across different segments of the business such as regions, products, and customers.

Create a calculated field for profit per order by right-clicking anywhere in the data pane and selecting "Create Calculated Field" and then using the formula: $\text{SUM}([\text{Profit}])/\text{COUNTD}([\text{Order ID}])$

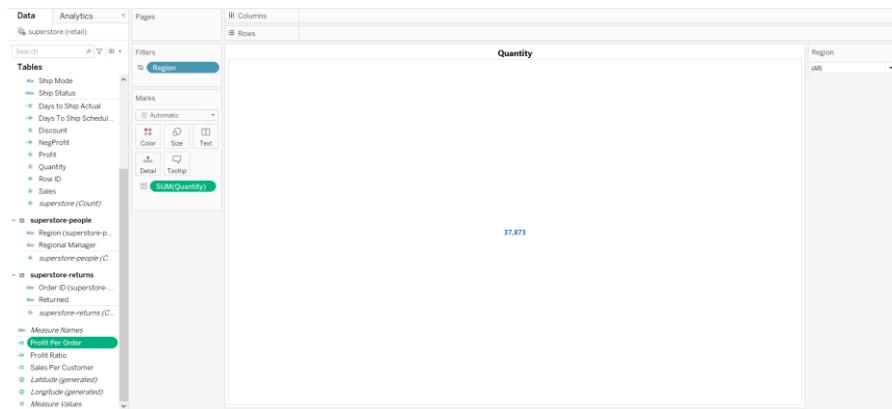


Drag Profit per order in Text label and format the data for Profit per order such as currency,display units etc.



Activity 1.24: Quantity

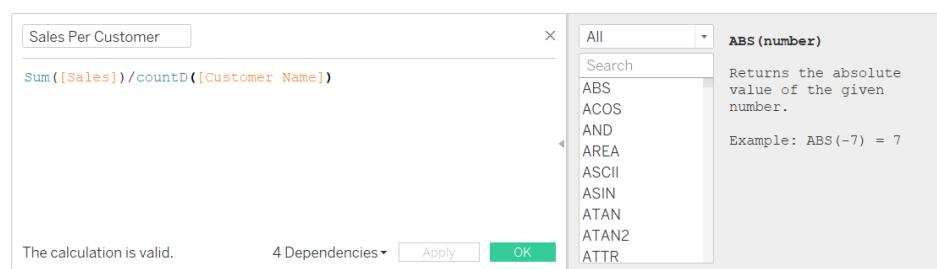
The "Quantity" column contains the number of products sold in each transaction. Drag Quantity in Text label



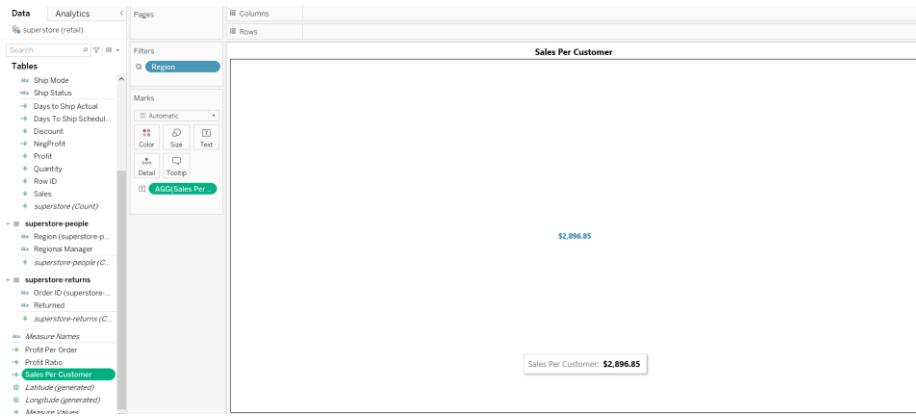
Activity 1.25: Sales per customer

Sales per customer is a metric that represents the average amount of sales generated by each customer. It can be used to understand the value of a customer to the business and how it varies across different segments of the business such as regions, products, and time.

Create a calculated field for sales per customer by right-clicking anywhere in the data pane and selecting "Create Calculated Field" and then using the formula: $\text{SUM}([\text{Sales}])/\text{COUNTD}([\text{Customer Name}])$.



Drag Sales per customer in Text label and format the data for Sales per customer such as currency, display units etc.



Activity 1.26: Sales By Segment

Sales by segment is a metric that represents the total sales generated by different segments of the business.

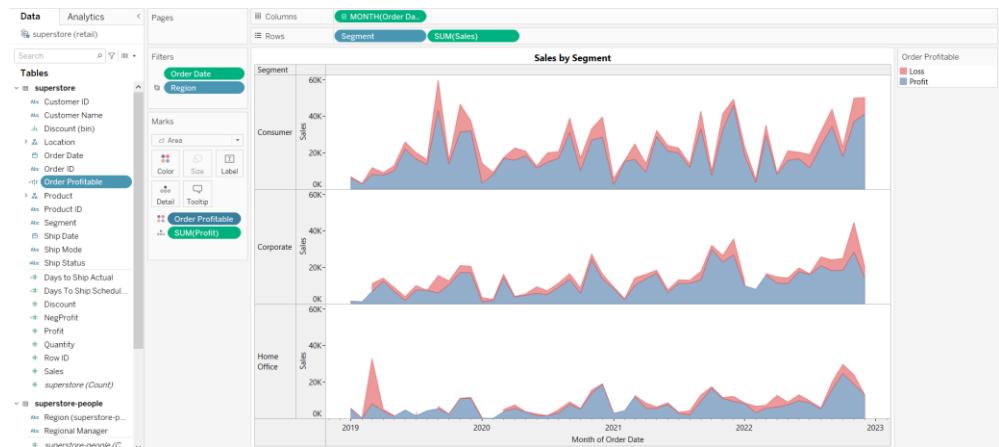
Create a calculated field for Order Profitable by right-clicking anywhere in the data pane and selecting "Create Calculated Field" and then using the formula:`{fixed [Order ID]:sum([Profit])}>0`



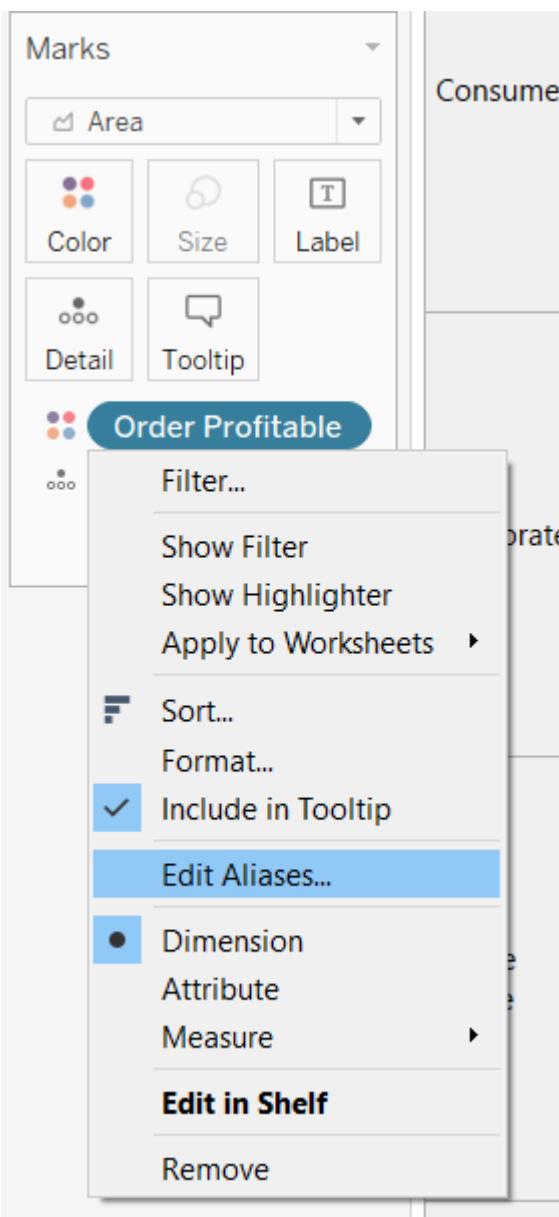
An area chart is a great way to visualize sales by segment. An area chart is a type of chart that represents data as a series of points connected by a line, with the area below the line filled in with a color or pattern. This type of chart is useful for showing changes in data over time or comparing data across different segments.

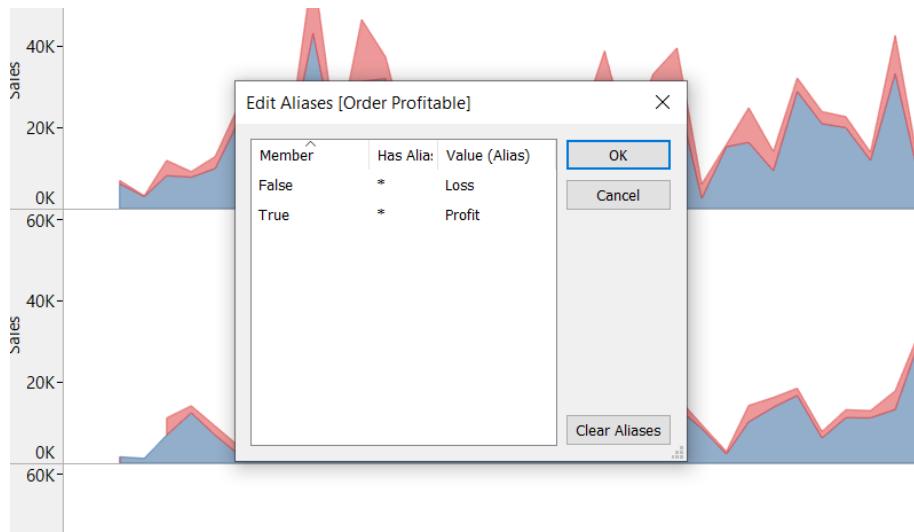
To create an area chart in Tableau to show sales by segment, you can follow these steps:

- Connect to your data source and drag the "Segment" dimension to the Columns or Rows shelf.
- Drag the "Sales" measure to the opposite shelf.
- Drag the "Order Date" dimension to the Columns shelf change date to month.
- Change the chart type to "Area" by clicking on the Show Me button or in the top right corner of the screen.
- Drag Profit in detail and order profitable in color



Select order profitable to edit original alias in profit & loss





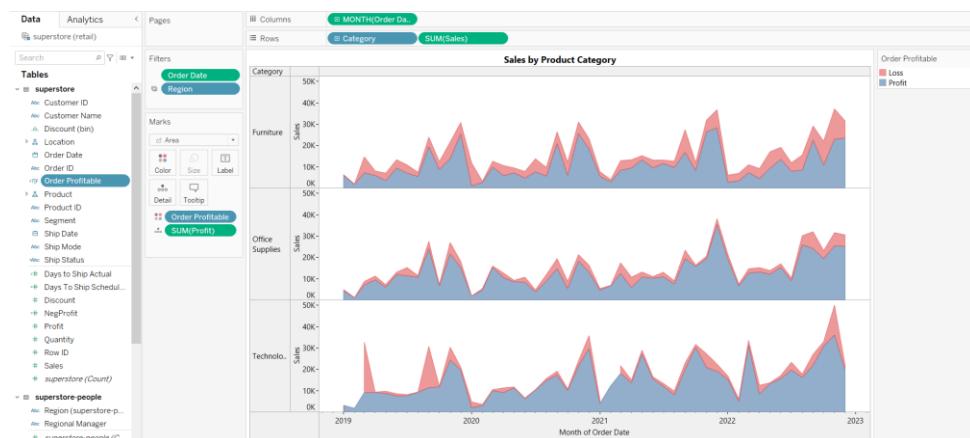
Activity 1.27: Sales By Product Category

Sales by Product Category is a metric that represents the total sales generated by different categories of products.

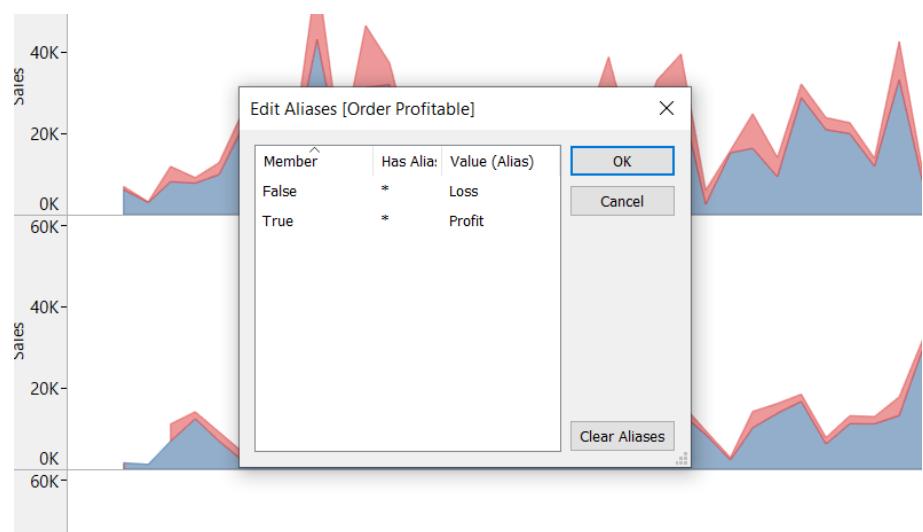
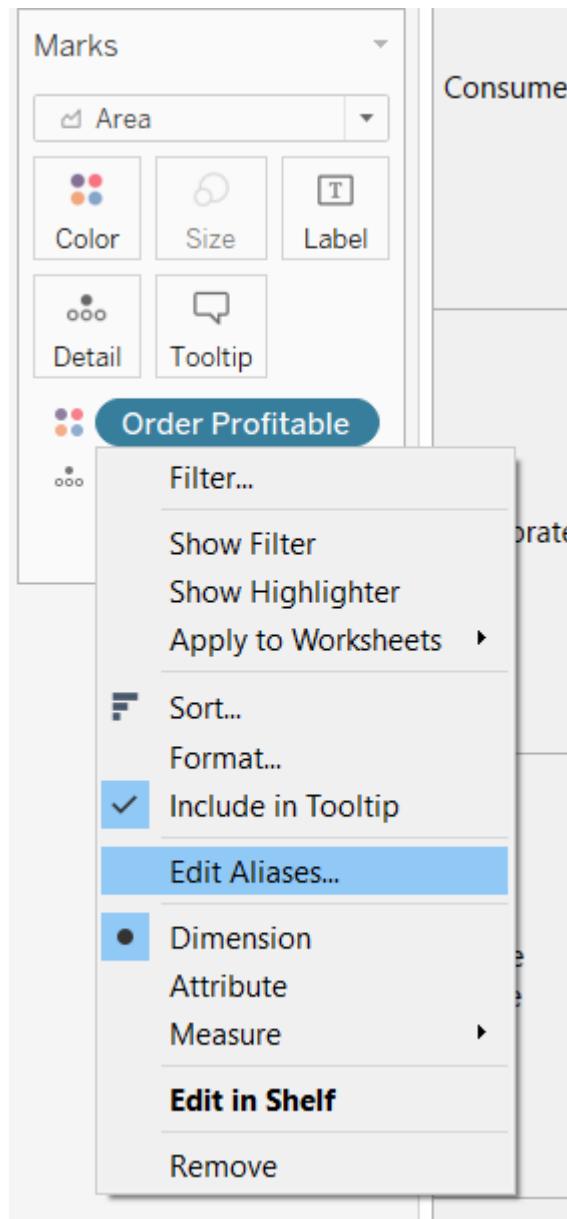
An area chart is a great way to visualize sales by Product Category. An area chart is a type of chart that represents data as a series of points connected by a line, with the area below the line filled in with a color or pattern. This type of chart is useful for showing changes in data over time or comparing data across different Product Categories.

To create an area chart in Tableau to show sales by Product Category, you can follow these steps:

- Connect to your data source and drag the "category" dimension to the Columns or Rows shelf.
- Drag the "Sales" measure to the opposite shelf.
- Drag the "Order Date" dimension to the Columns shelf change date to month.
- Change the chart type to "Area" by clicking on the Show Me button or in the top right corner of the screen.
- Drag Profit in detail and order profitable in color



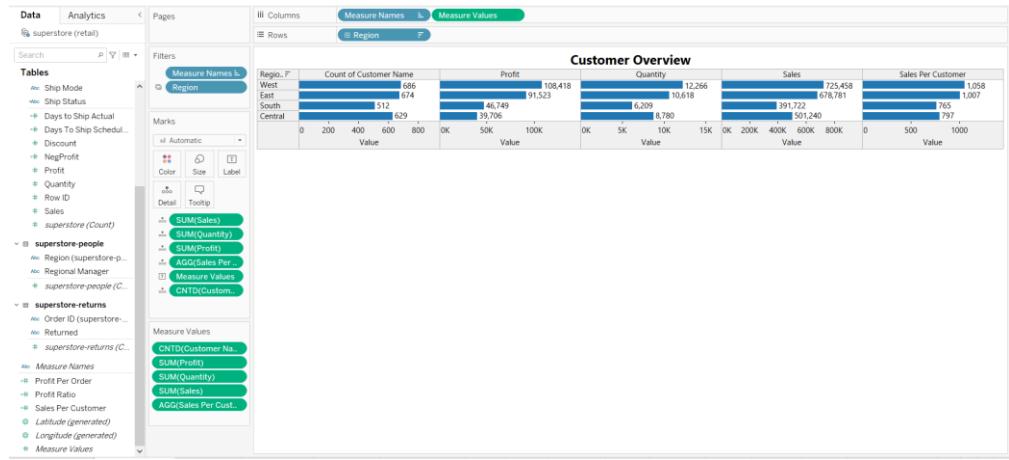
Select order profitable to edit original alias in profit & loss



Activity 1.28: Customer Overview

Customer overview provides valuable insights into how different customers are contributing to the business in terms of sales, profit, quantity, sales per customer & count of customers by customer name.

Check the below screenshot and drag all data points accordingly.

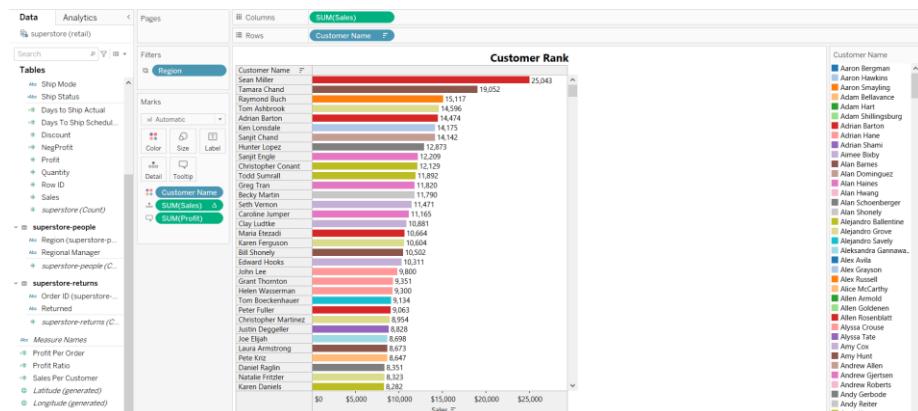


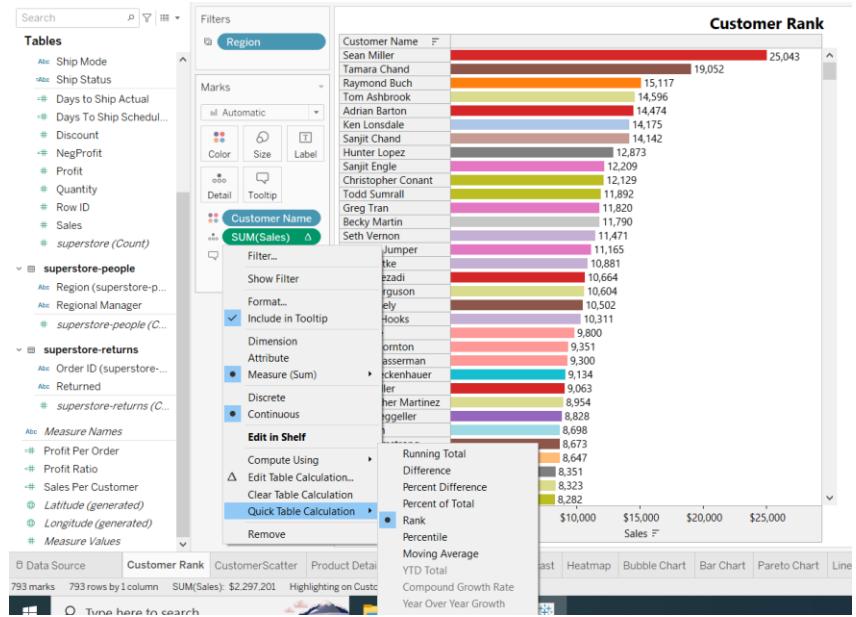
Activity 1.29: Customer Rank

Create a customer rank to see which customers are the most valuable to the business in terms of sales, profit, or other metrics. Here are some steps to create a customer rank in Tableau:

- Drag the "Customer Name" dimension to the Rows or Columns shelf.
- Drag the measure you want to use for ranking (e.g. Sales, Profit, etc) to the opposite shelf.
- Right-click on the measure and select "Rank" in the context menu.
- You can change the sorting order of the rank by clicking on the sort icon on the right side of the measure.
- Change the chart type to "Bar" chart.

Check the below screenshot and drag all data points accordingly.





Activity 1.30: Sales and profit by customer

Sales & profit by customer means analyzing the performance of individual customers in terms of their sales and profit. This analysis can help in identifying the most valuable customers, as well as those who may need more attention. With this analysis, it is possible to understand the relationship between sales and profit for each customer, which can help in identifying patterns in customer behavior and making data-driven decisions.

Scatter plot of sales and profit by customer can show which customers have high sales but low profit. Identifying these patterns can help in making decisions about pricing, discounts, and customer engagement strategies.

Check the below screenshot and drag all data points accordingly.



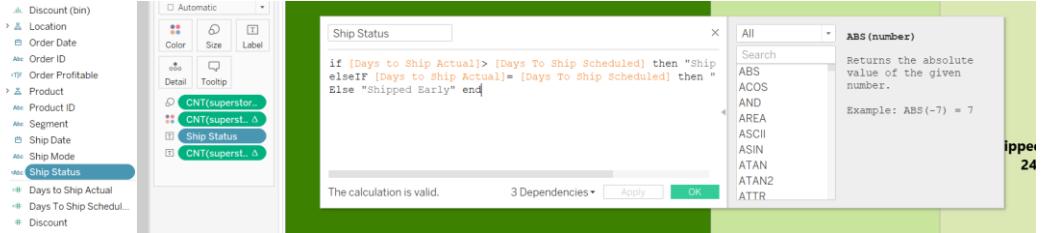
Activity 1.31: Shipping Status

Shipping status refers to the current state of a package or shipment as it moves through the shipping process. This can include information such as the package's origin and destination, the carrier that is handling the package, and the current location of the package. The shipping status can also include

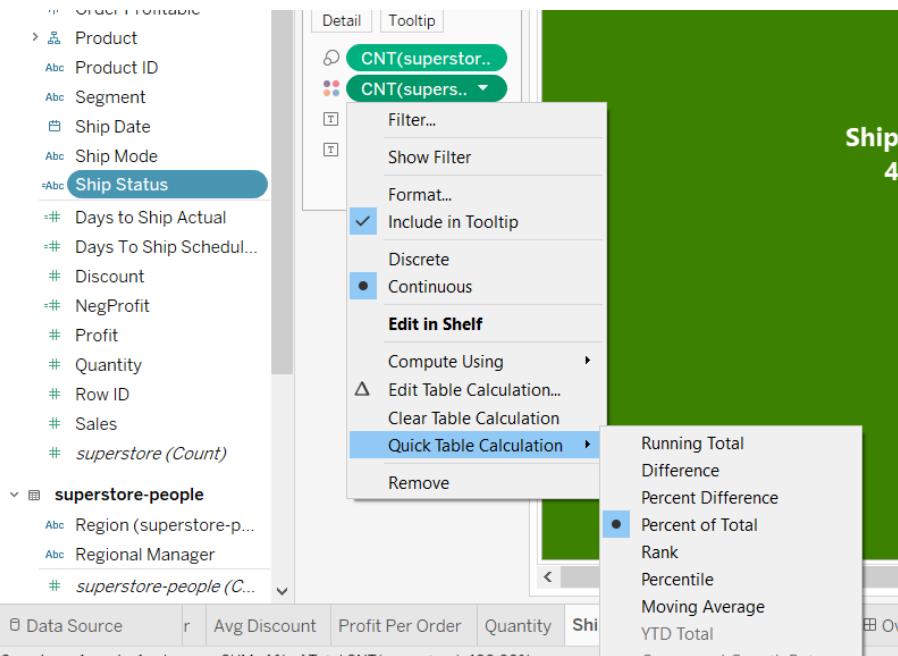
details such as the estimated delivery date and any updates about the package's journey.

Create a calculated field for ship status by right-clicking anywhere in the data pane and selecting "Create Calculated Field" and then using the formula:

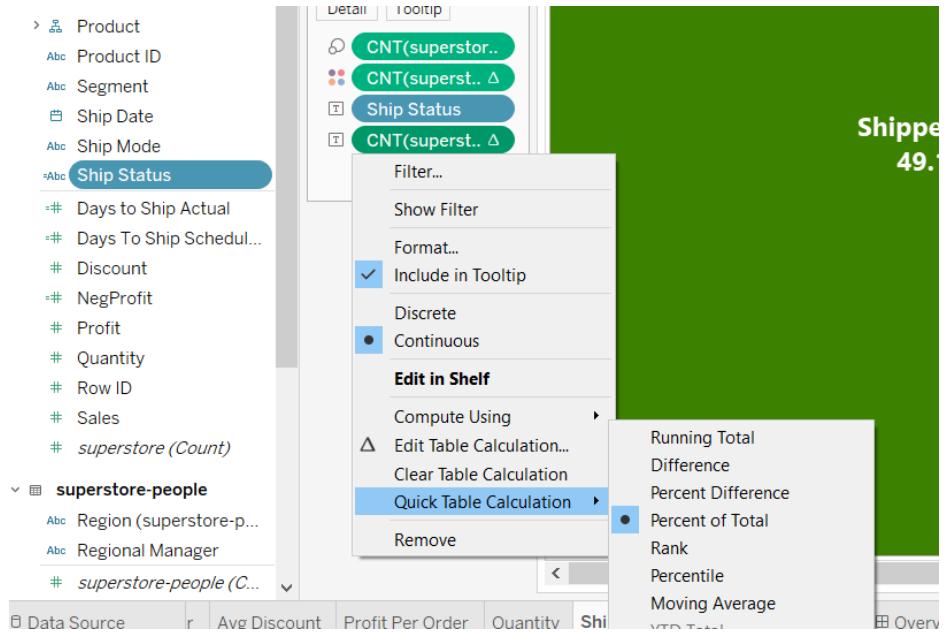
```
if [Days to Ship Actual]> [Days To Ship Scheduled] then "Shipped Late"
elseIf [Days to Ship Actual]= [Days To Ship Scheduled] then "Shipped On
Time"Else "Shipped Early" end
```



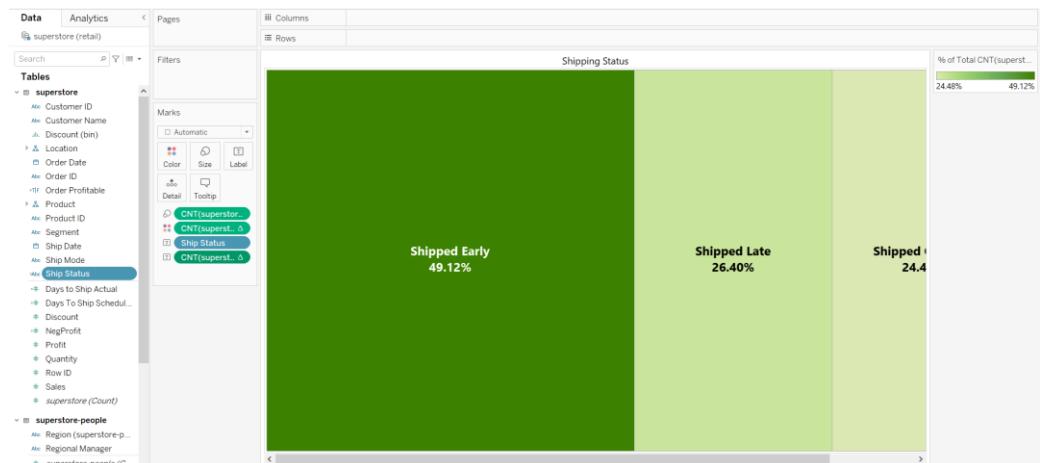
- Drag superstore count to size & colors
- Drag Ship status & count of superstore to text label
- Create a quick table calculation for the count of superstore (color) change to percent of total.



- Create a quick table calculation for the count of superstore (text label) change to percent of total.



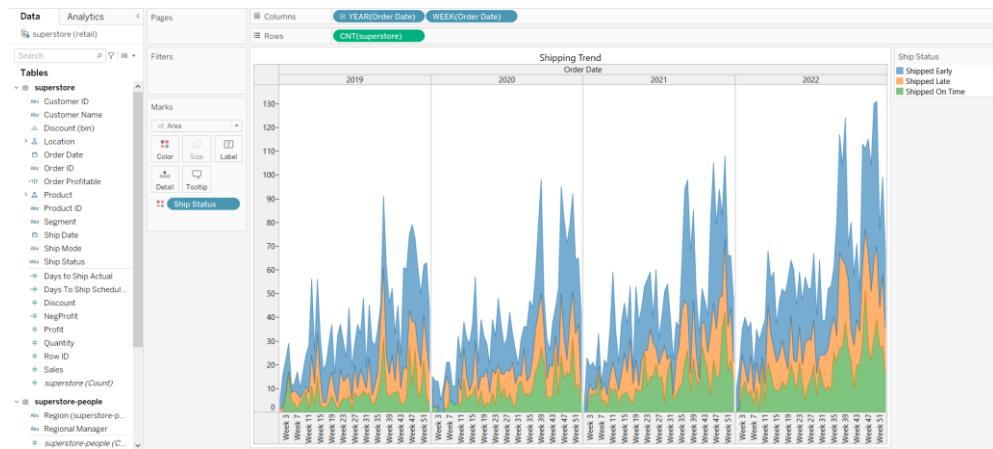
Check the below screenshot and drag all data points accordingly to create treemaps



Activity 1.32: Shipping Trend

Shipping trend refers to the pattern or direction of change in shipping activity over a period of time. You can use various types of charts such as line chart, area chart, or column chart to analyze shipping trends over time.

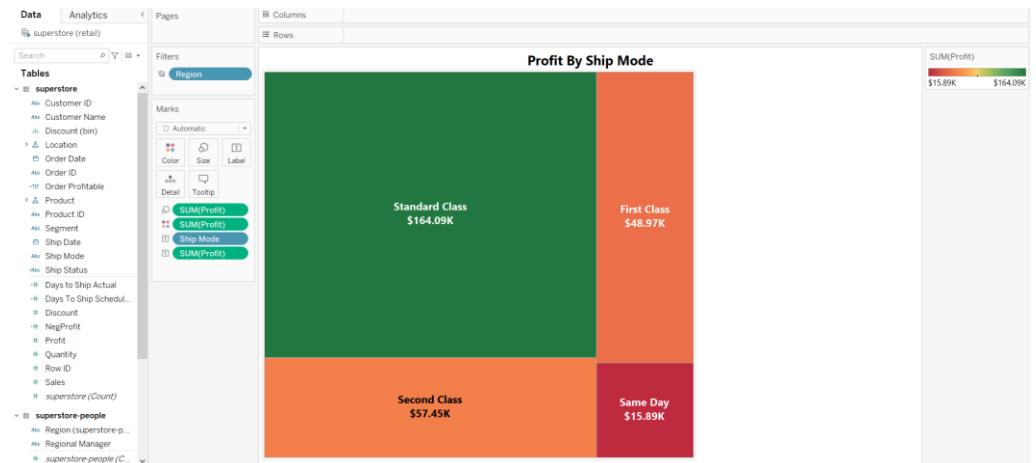
Check the below screenshot and drag all data points accordingly to create an area chart.



Activity 1.33: Profit by ship mode

Profit by ship mode is an analysis that looks at the profit generated from different shipping methods or carriers. It allows you to see which shipping options are most profitable for your business, and which may be costing you money. This information can be used to make data-driven decisions about shipping strategies, such as which shipping options to offer customers, which carriers to use, and how to negotiate better rates with carriers.

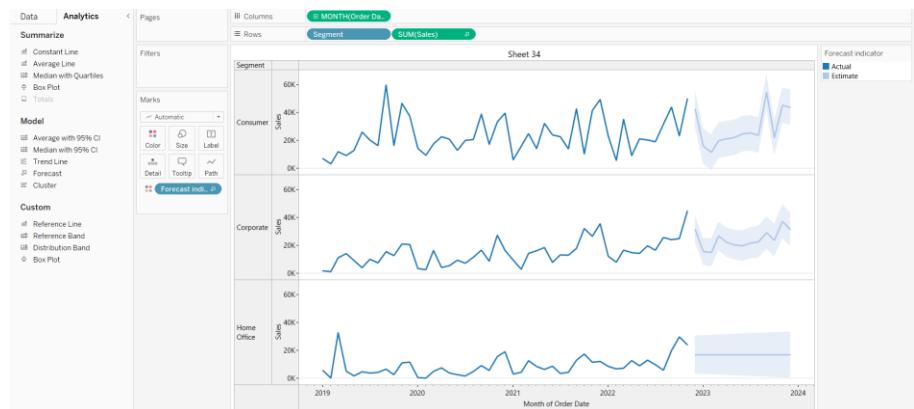
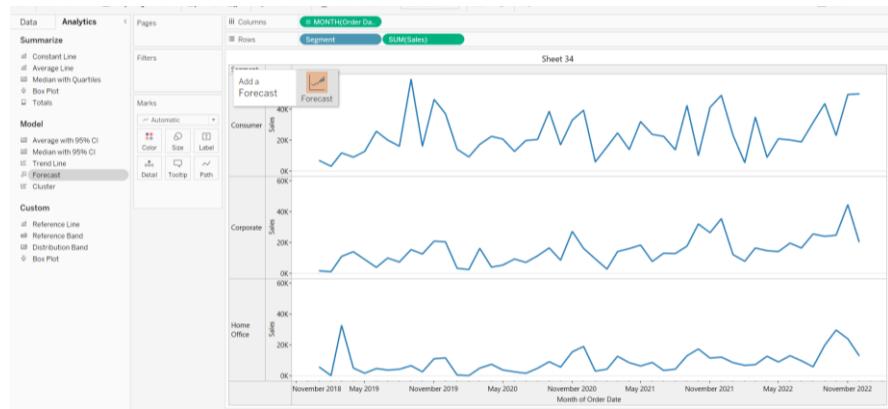
Check the below screenshot and drag all data points accordingly to create a Tree Map.



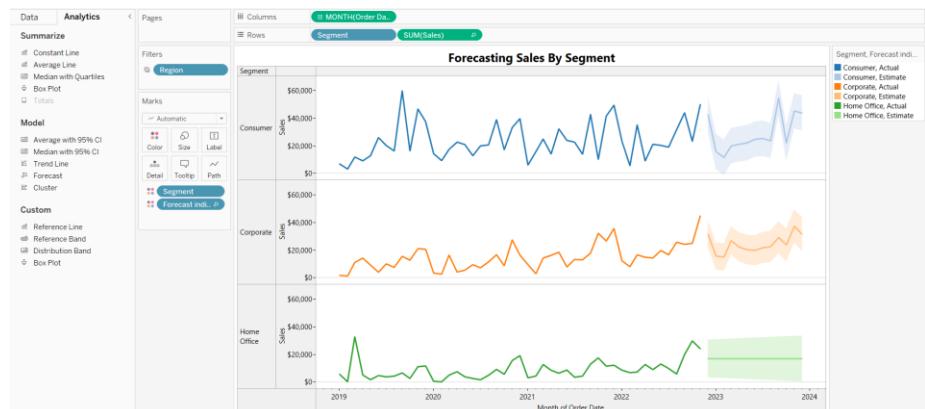
Activity 1.34: Forecasting By Sales Segment

Forecasting sales by segment involves using historical sales data to predict future sales for specific segments of your customer base.

- Drag the "Segment" dimension to the Rows shelf.
- Drag the "Sales" measure to the Rows shelf.
- Drag the "Order date" to the Rows shelf and change it to month.
- Create a line chart to visualize the historical sales data by segment
- Use the built-in forecasting function in Tableau to predict future sales by segment, go to the analytics pane, drag forecast into line chart and select forecast.



Drag Segment to color to get the actual & estimate forecasting by segment.



Activity 1.35: Product Detail Sheet

Profit by ship mode is an analysis that looks at the profit generated from different shipping methods or carriers. It allows you to see which shipping options are most profitable for your business, and which may be costing you money. This information can be used to make data-driven decisions about shipping strategies, such as which shipping options to offer customers, which carriers to use, and how to negotiate better rates with carriers.

Create a calculated field for Days to ship actual by right-clicking anywhere in the data pane and selecting "Create Calculated Field" and then using the formula: DATEDIFF('day',[Order Date],[Ship Date])

The screenshot shows the Tableau Data Editor interface. On the left, the data pane lists fields: Product, Product ID, Segment, Ship Date, Ship Mode, Ship Status, Days to Ship Actual, Days to Ship Scheduled, Discount, NegProfit, Profit, Quantity, Row ID, Sales, superstore (Count), superstore-people, and Region (superstore-people). In the center, a dialog box titled 'Days To Ship Actual' contains the formula 'DATEDIFF('day',[Order Date],[Ship Date])'. Below it, a message says 'The calculation is valid.' and '7 Dependencies'. On the right, a tooltip for the 'ABS(number)' function is displayed, with the text: 'Returns the absolute value of the given number. Example: ABS(-7) = 7'. The tooltip also includes a table of values for ABS.

Create a calculated field for Days To Ship Scheduled by right-clicking anywhere in the data pane and selecting "Create Calculated Field" and then using the formula:

CASE [Ship Mode]

WHEN "Same Day" THEN 0

WHEN "First Class" THEN 1

WHEN "Second Class" THEN 3

WHEN "Standard Class" THEN 6

END

The screenshot shows the Tableau Data Editor interface. On the left, the data pane lists fields: Product, Product ID, Segment, Ship Date, Ship Mode, Ship Status, Days to Ship Actual, Days to Ship Scheduled, Discount, NegProfit, Profit, Quantity, Row ID, Sales, superstore (Count), superstore-people, and Region (superstore-people). In the center, a dialog box titled 'Days To Ship Scheduled' contains the formula 'CASE [Ship Mode] WHEN "Same Day" THEN 0 WHEN "First Class" THEN 1 WHEN "Second Class" THEN 3 WHEN "Standard Class" THEN 6 END'. Below it, a message says 'The calculation is valid.' and '6 Dependencies'. On the right, a tooltip for the 'ABS(number)' function is displayed, with the text: 'Returns the absolute value of the given number. Example: ABS(-7) = 7'. The tooltip also includes a table of values for ABS.

Check the below screenshot and drag all data points accordingly to create a Text Table.

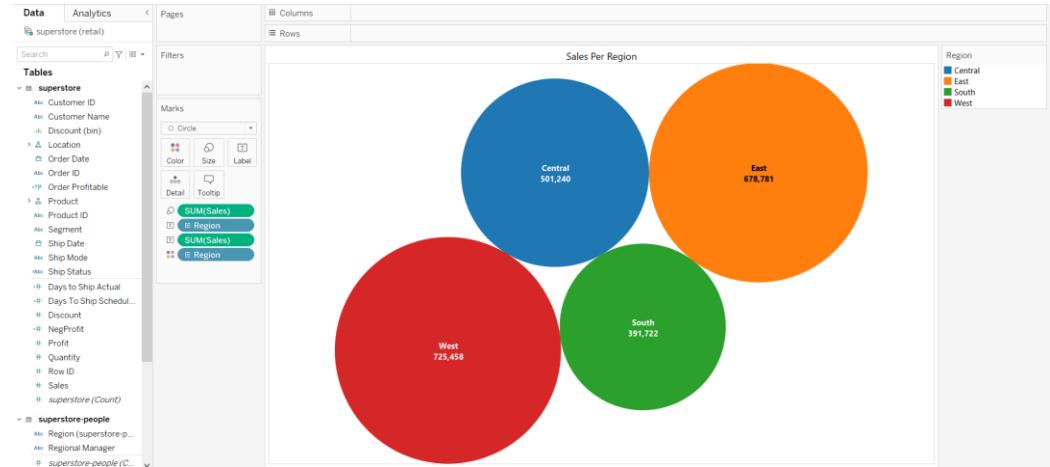
The screenshot shows the Tableau Data Editor interface with a 'Text' table type. The columns are: Order ID, Customer Name, Year of Order, Ship Mode, Sales, Quantity, Discount, Profit, Profit Ratio, Days to Ship Actual, and Days to Ship Scheduled. The data is as follows:

Order ID	Customer Name	Year of Order	Ship Mode	Sales	Quantity	Discount	Profit	Profit Ratio	Days to Ship Actual	Days to Ship Scheduled
US-2019-100006	Dennis Kane	2019	Standard Class	378	3	2%	110	29.0%	5	5
US-2019-100090	Ed Braxton	2019	Standard Class	699	9	40%	-19	-2.7%	12	8
US-2019-100279	Scott Williamson	2019	Standard Class	22	2	0%	11	48.0%	6	4
US-2019-100328	Michael Mutsch	2019	Standard Class	91	6	20%	32	35.0%	8	4
US-2019-100328	Jasper Caxoppel	2019	Standard Class	4	1	20%	1	33.3%	5	6
US-2019-100363	Jim Mitchell	2019	Standard Class	21	5	40%	8	36.1%	12	14
US-2019-100391	Barry Wenzl	2019	Standard Class	15	2	0%	7	46.0%	6	4
US-2019-100578	Kurt Miller	2019	Standard Class	697	11	90%	62	8.9%	24	16
US-2019-100670	Jeffrey Elstrot	2019	Second Class	129	8	0%	18	13.7%	8	4
US-2019-100762	Nat Gelles	2019	Standard Class	509	11	0%	219	42.1%	24	20
US-2019-100853	Jennifer Braxton	2019	Standard Class	73	6	100%	-125	-171.6%	12	10
US-2019-100860	Cindy Stewart	2019	Second Class	19	5	0%	9	48.0%	3	4
US-2019-100867	Eugene Middlebrand	2019	Standard Class	322	6	20%	20	6.3%	6	5
US-2019-100877	Howard Ratner	2019	Standard Class	402	3	20%	23	7.5%	5	4
US-2019-100895	Stewart Wisnicky	2019	Standard Class	605	7	0%	177	29.2%	18	12
US-2019-100916	Frank Hawley	2019	Standard Class	789	10	0%	123	15.6%	18	15
US-2019-100972	Dennis Bolton	2019	Second Class	166	3	0%	80	48.0%	3	5
US-2019-101000	Matt Collier	2019	First Class	2	1	80%	-6	-265.0%	1	2
US-2019-101175	Howard Shandor	2019	Standard Class	101	6	20%	-1	-1.0%	8	5
US-2019-101266	Michael Moore	2019	Second Class	13	2	0%	6	48.0%	3	3
US-2019-101364	Tamara Willingham	2019	Standard Class	297	13	20%	100	33.8%	6	4
US-2019-101392	Ann Steele	2019	Standard Class	269	7	0%	70	26.0%	6	6
US-2019-101476	Shirley Daniels	2019	Standard Class	8	3	20%	1	13.8%	6	4
US-2019-101482	Benjamin Patterson	2019	Standard Class	60	4	0%	28	46.0%	5	5
US-2019-101476	Shirley Daniels	2019	First Class	70	1	0%	30	43.0%	1	1
US-2019-101560	Chris Selenick	2019	Second Class	542	19	0%	111	20.4%	12	12
US-2019-101602	Mick Crebaga	2019	First Class	804	8	50%	-31	-3.9%	2	6
US-2019-101602	Mick Crebaga	2019	Standard Class	2	1	70%	-1	-70.0%	6	4
US-2019-101833	Frank Gastineau	2019	Second Class	34	3	0%	17	50.0%	3	5
US-2019-101931	Todd Sumrall	2019	First Class	1,253	17	35%	19	1.5%	5	15
US-2019-102008	Russell Applegate	2019	Standard Class	49	1	0%	24	50.0%	6	4
US-2019-102071	Parker Gardineer	2019	Standard Class	236	7	0%	44	18.8%	12	12
US-2019-102245	Howard Ratner	2019	Standard Class	29	4	0%	14	47.0%	5	5
US-2019-102274	Dave Hallsten	2019	Standard Class	866	13	0%	323	37.3%	24	20
US-2019-102295	Erica Hackney	2019	Second Class	121	1	20%	-18	-15.0%	3	2

Activity 1.36: Sales Per Region

Sales per region is a metric that measures the amount of sales generated in a specific geographic region.

Check the below screenshot and drag all data points accordingly to create a bubble chart.



Milestone 5: Dashboard

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data, and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

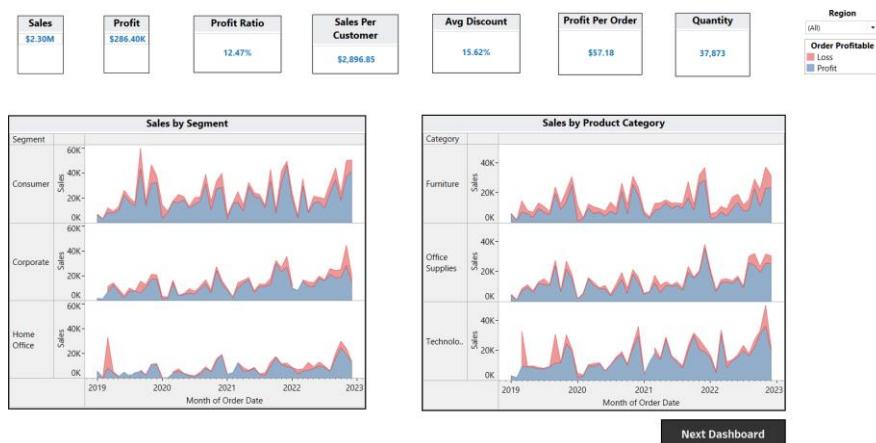
Activity :1- Responsive and Design of Dashboard

The responsiveness and design of a dashboard for Superstore Sales Analysis is crucial to ensure that the information is easily understandable and actionable. Key considerations for designing a responsive and effective dashboard include user-centered design, clear and concise information, interactivity, data-driven approach, accessibility, customization, and security. The goal is to create a dashboard that is user-friendly, interactive, and data-driven, providing actionable insights to improve the performance and efficiency of Superstore.

Create a dashboard: Once you have created graphs in multiple worksheets, you can create a dashboard by clicking the "New Dashboard" button. You can then arrange the worksheets on the dashboard and add objects like text, images, and web pages.

Format and style your dashboard: Tableau allows you to format and style your dashboard in a variety of ways, including customizing colors, fonts, and backgrounds.

Overview Dashboard

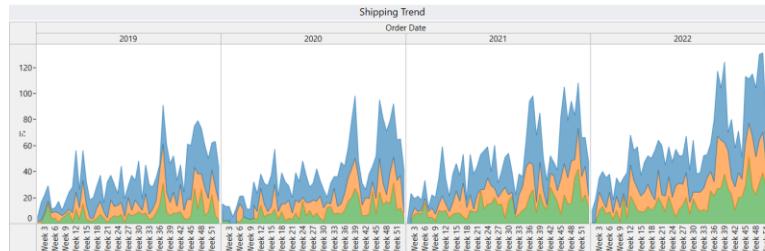


Order Order Quantity Shipping Status Shipping Trend Sales Per Region Overview-D1 Customer Analysis-D2 Shipment Trend-D3 Forecasting & Product Details... Ship Mode & Quantity of order... Superstore Story Clustering

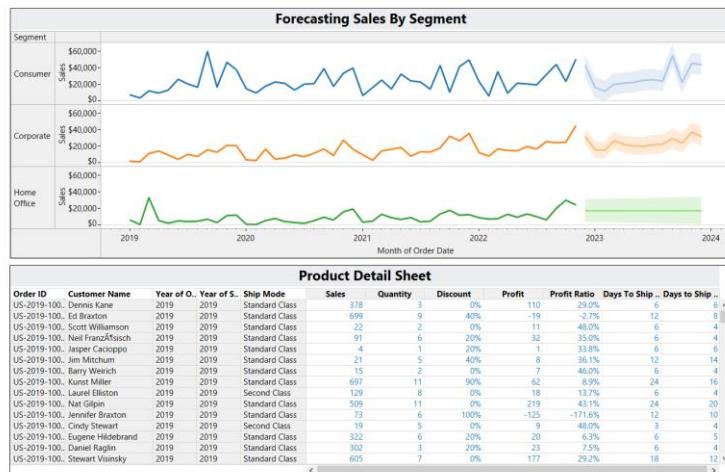
Customer Analysis Dashboard


[Previous Dashboard](#)
[Next Dashboard](#)
[Order](#) [Quantity](#) [Shipping Status](#) [Shipping Trend](#) [Sales Per Region](#) [Overview-D1](#) [Customer Analysis-D2](#) [Shipment Trend-D3](#) [Forecasting & Product Details...](#) [Ship Mode & Quantity of order...](#) [Superstore Story](#) [Clustering](#) [...](#) [←](#) [→](#) [☰](#)

Shipment Trend Dashboard

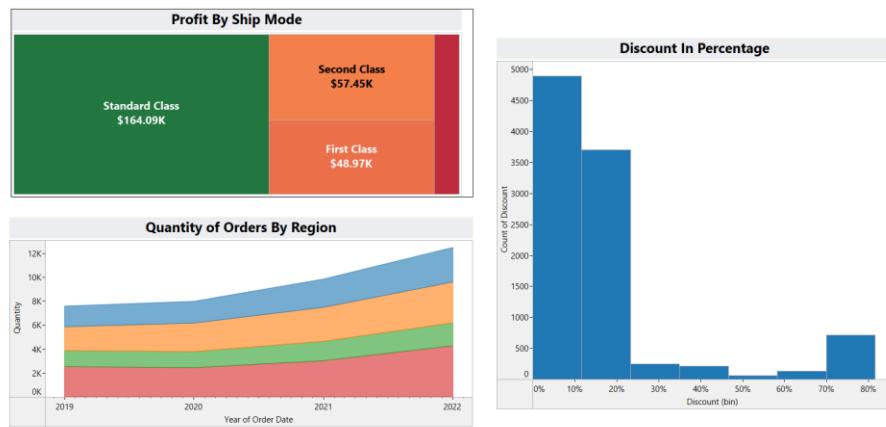

[Previous Dashboard](#)
[Next Dashboard](#)
[Order](#) [Quantity](#) [Shipping Status](#) [Shipping Trend](#) [Sales Per Region](#) [Overview-D1](#) [Customer Analysis-D2](#) [Shipment Trend-D3](#) [Forecasting & Product Details...](#) [Ship Mode & Quantity of order...](#) [Superstore Story](#) [Clustering](#) [...](#) [←](#) [→](#) [☰](#)

Forecasting & Product Details



Order Quantity Shipping Status Shipping Trend Sales Per Region Overview-D1 Customer Analysis-D2 Shipment Trend-D3 Forecasting & Product Details Ship Mode & Quantity of order... Superstore Story Clustering

Ship Mode & Quantity of order By Region



Order Quantity Shipping Status Shipping Trend Sales Per Region Overview-D1 Customer Analysis-D2 Shipment Trend-D3 Forecasting & Product Details Ship Mode & Quantity of order... Superstore Story Clustering

Dashboard Link:-

https://public.tableau.com/app/profile/shivam.shivhare3038/viz/tableausuperstore_16747386787460/Overview-D1?publish=yes

Milestone 6: Story

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

Organize your story: Once you have created your worksheets and dashboards, you can organize them into a story by creating a new storyboard. You can arrange the worksheets and dashboards on the storyboard in a logical sequence that follows the story you want to tell.

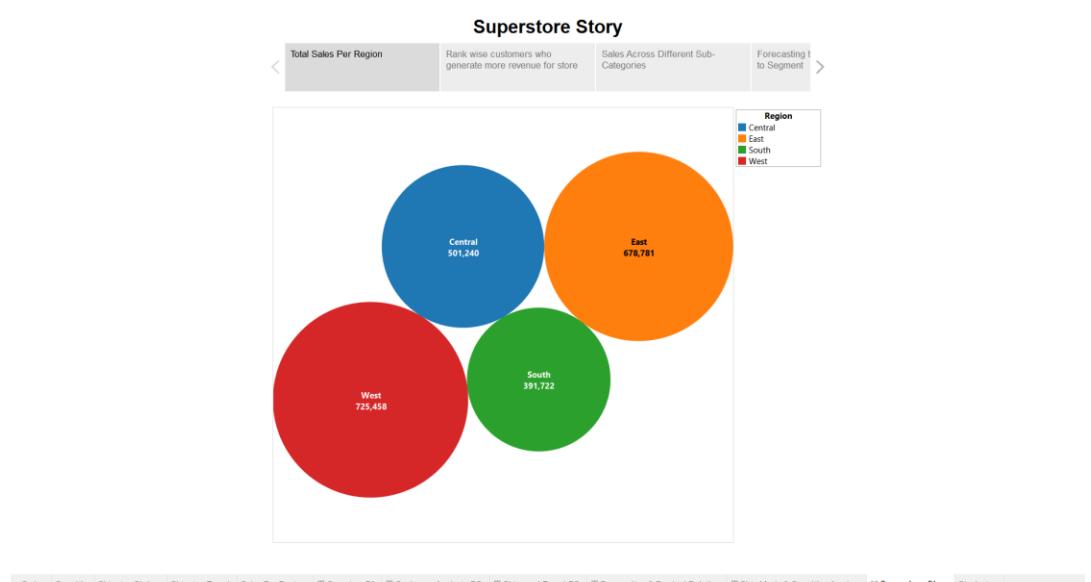
Add text and annotations: To help guide your audience through the story, you can add text and annotations to the storyboard. This can include headings, captions, and callouts that highlight key insights or findings.

Format and style your story: To make your story more visually appealing, you can format and style the worksheets and dashboards on the storyboard. This can include customizing colors, fonts, and backgrounds to match your organization's branding.

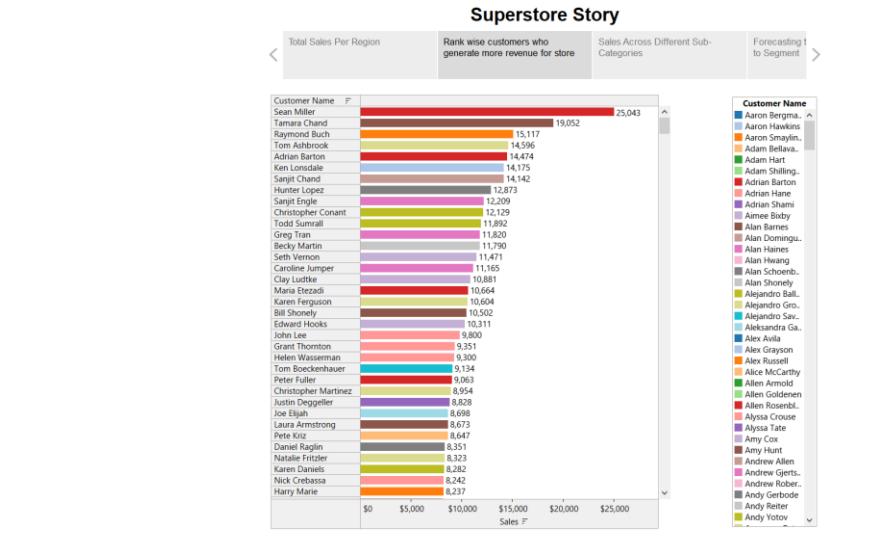
Activity:1- No of Scenes of Story

The number of scenes in a storyboard for a data visualization analysis of the performance and efficiency of Superstore Analysis will depend on the complexity of the analysis and the specific insights that are trying to be conveyed. A storyboard is a visual representation of the data analysis process and it breaks down the analysis into a series of steps or scenes.

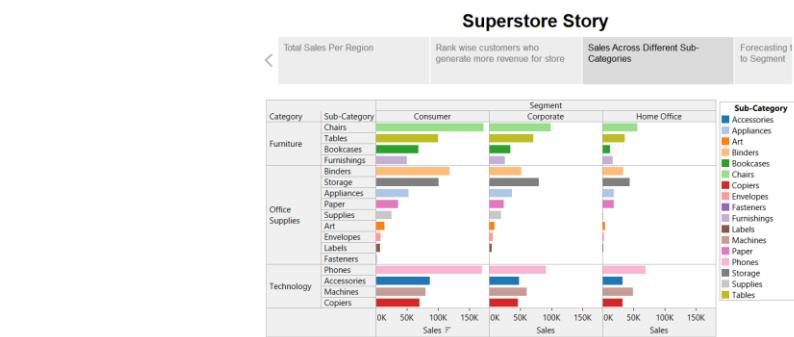
Total Sales For Region



Rank wise customers who generate more revenue for store

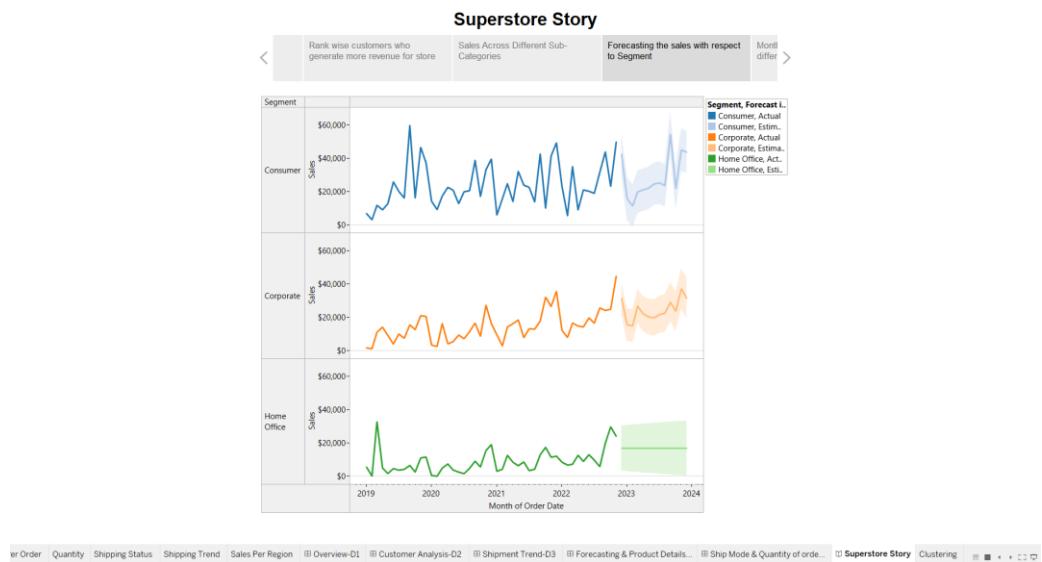


Sales Across Different Sub-Categories



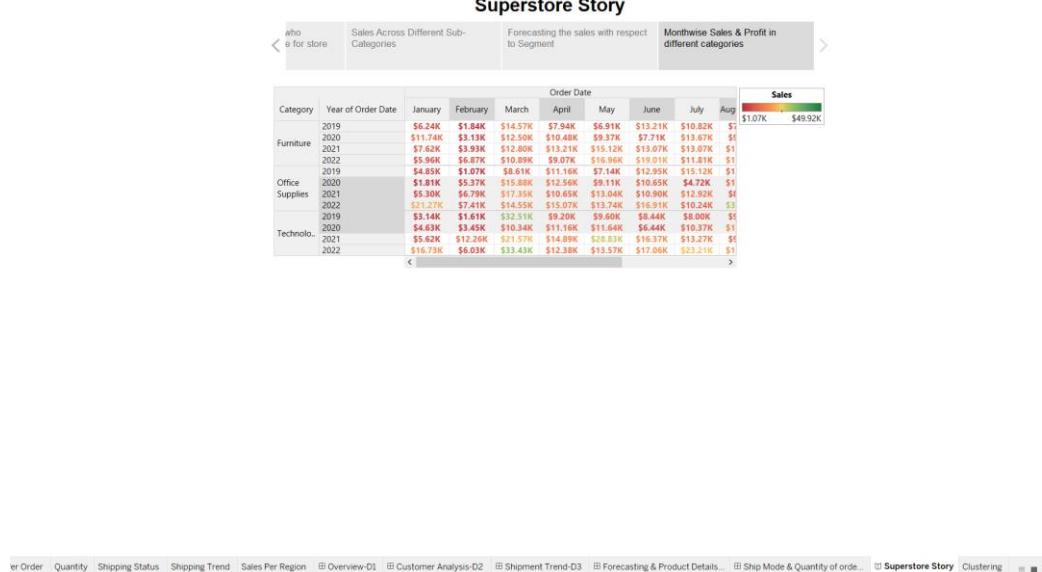
Order Quantity Shipping Status Shipping Trend Sales Per Region Overview-D1 Customer Analysis-D2 Shipment Trend-D3 Forecasting & Product Details... Ship Mode & Quantity of order... Superstore Story Clustering

Forecasting the sales with respect to Segment



Order Quantity Shipping Status Shipping Trend Sales Per Region Overview-D1 Customer Analysis-D2 Shipment Trend-D3 Forecasting & Product Details... Ship Mode & Quantity of order... Superstore Story Clustering

Month Wise Sales & Profit in different categories



Order Quantity Shipping Status Shipping Trend Sales Per Region Overview-D1 Customer Analysis-D2 Shipment Trend-D3 Forecasting & Product Details... Ship Mode & Quantity of order... Superstore Story Clustering

Story Link:

https://public.tableau.com/shared/Z9XDZSMM6?:display_count=n&:origin=viz_share_link

Milestone 7: Performance Testing

Activity 1: Amount of Data Rendered to DB

- The amount of data that is rendered to a database depends on the size of the dataset and the capacity of the database to store and retrieve data.
- Open the MySQL Workbench, go to the database then click to expand the tables, select the table and click on (i) button to get the information related to the table such as column count, table rows etc.

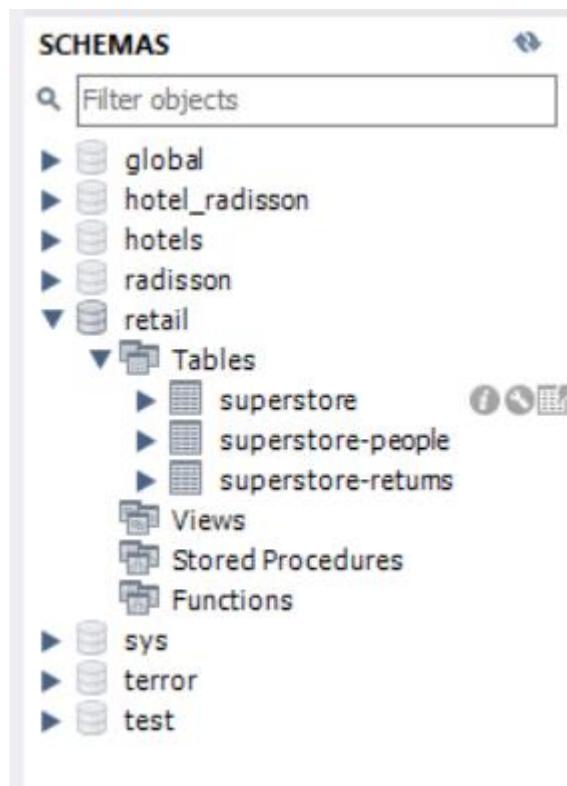


Table Details

- Engine: InnoDB
- Row format: Dynamic
- Column count: 20
- Table rows: 8829
- AVG row length: 298
- Data length: 2.5 MB
- Index length: 0.0 bytes
- Max data length: 0.0 bytes
- Data free: 4.0 MB
- Table size (estimate): 2.5 MB
- File format:
- Data path: C:\ProgramData\MySQL\MySQL Server 8.0\Data\retail\superstore.ibd
- Update time: 2023-01-26 08:29:13
- Create time: 2023-01-26 08:17:17

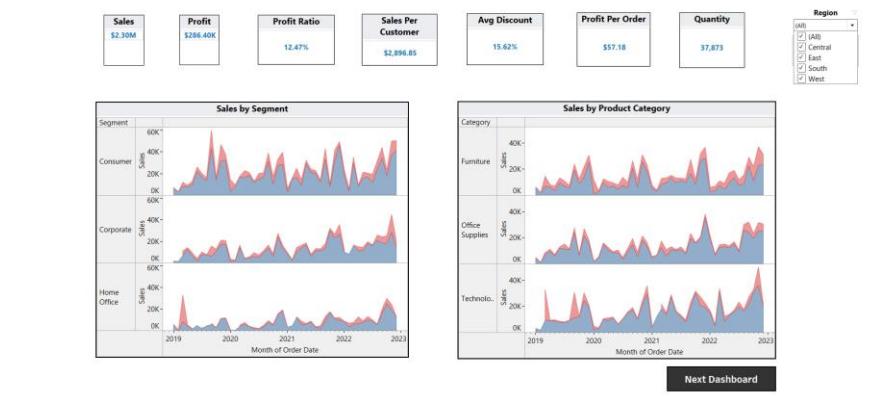
Table Details

- Engine: InnoDB
- Row format: Dynamic
- Column count: 2
- Table rows: 4
- AVG row length: 4096
- Data length: 16.0 KB
- Index length: 0.0 bytes
- Max data length: 0.0 bytes
- Data free: 0.0 bytes
- Table size (estimate): 16.0 KB
- File format:
- Data path:
- Update time: 2023-01-25 22:30:30
- Create time: 2023-01-25 22:30:29

Table Details

- Engine: InnoDB
- Row format: Dynamic
- Column count: 2
- Table rows: 800
- AVG row length: 81
- Data length: 64.0 KB
- Index length: 0.0 bytes
- Max data length: 0.0 bytes
- Data free: 0.0 bytes
- Table size (estimate): 64.0 KB
- File format:
- Data path:
- Update time: 2023-01-25 22:29:51
- Create time: 2023-01-25 22:29:21

Activity 2: Utilization of Data Filters



Order Quantity Shipping Status Shipping Trend Sales Per Region Overview D1 Customer Analysis-D2 Shipment Trend-D3 Forecasting & Product Details Ship Mode & Quantity of order Superstore Story Clustering

Activity 3: No of Calculation Fields

Workbook	Number of Calculation Fields
superstore (retail)	21
superstore (Count)	25

Activity 4: No of Visualizations/ Graphs

1. Bar Chart:
2. Text Tables:
3. Line chart:
4. Pie Chart:
5. Scatter Plot:
6. Area Chart:
7. Bubble Chart:
8. Histogram:
9. Dual Axis Chart:
10. Gantt chart:
11. Heat Map:
12. Bullet Chart:
13. Pareto Chart
14. Waterfall Chart:
15. Maps
16. Box and Whisker Plots
17. Motion Charts
18. Tree Maps
19. Sales
20. Profit
21. Profit Ratio
22. Average Discount
23. Profit per order
24. Quantity
25. Sales per customer
26. Sales By Segment
27. Sales By Product Category
28. Customer Overview
29. Customer Rank
30. Sales and profit by customer
31. Shipping Status
32. Shipping Trend
33. Profit by ship mode
34. Forecasting By Sales Segment
35. Product Detail Sheet
36. Sales Per Region

Milestone 8: Web integration

Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others.

Publishing dashboard and reports to tableau public

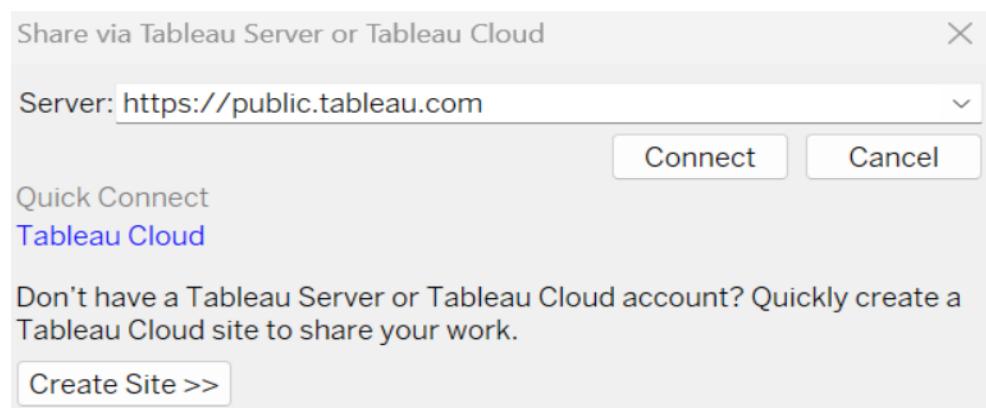
Sharing via Tableau Server/Public

To make your Tableau Dashboard publicly accessible, the best way is to share it via Tableau Server/Public. The steps are as follows:

Step 1: Open Tableau Desktop, and click on the Server button.

Step 2: After selecting the Server, click on Tableau Public and select Save to Tableau Public.

Go to Dashboard/story, click on share button on the top ribbon

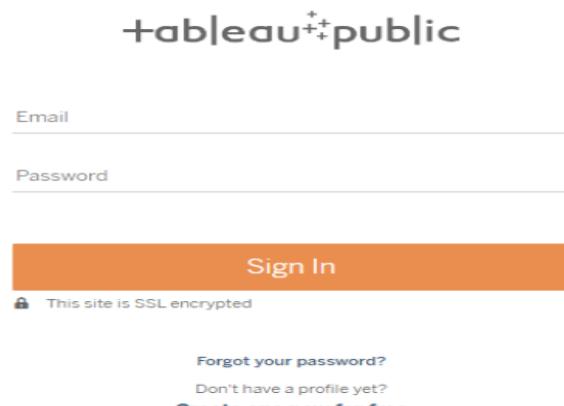


Give the server address of your tableau public account and click on connect.

Step 3: A dialog box pops up asking to enter your credentials for Tableau Public. If you don't have it, then you can create a profile.

Step 4: Upon entering the credentials, a new dialog box will open. Select Create Data Extract and click on Extract.

Once you click on connect it will ask you for tableau public user name and password



Once you login into your tableau public using the credentials, the particular visualization will be published into tableau public

Note: While publishing the visualization to the public, the respective sheet will get published when you click on share option.

Step 5: Once done, repeat step 2, to view the final embedded Tableau Dashboard on the browser.

Step 6: Click on Edit if you want to add a title or any description for the Tableau Dashboard.

Step 7: Save it and share it with the Tableau Viewer.

Activity 1: Dashboard and Story embed with UI With Flask

Explanation video link: (Reference Video to Embed Dashboard & Story)

https://drive.google.com/file/d/1UCDtjXj8Fl4bgHeNZs2EZFYuJzVF_gwN/view?usp=sharing

Download Flask Application Files:- [Link](#)

The screenshot shows a website for 'SB-Store'. At the top, there's a navigation bar with links for Home, About, Dashboard, Story, Contact, and social media icons. The main content area has a heading 'Get analytics solution for your store that help grow businesses' with a subtext 'We are one stop solution to all your business requirements'. Below this is a 'GET STARTED' button. To the right, there's an illustration of two people at a desk with a large screen displaying charts and graphs.



Learn more about us

We provide analytics solution for your Superstore.

- Get Sales Analytics that can help increase sales.
- Get Profit and order Analytics.
- Get future forecasting in order to make better decision for your business.

Analytics overview

Sales, profit and order analytics

Analytics that can help you understand trends and patterns for sales, profit and order of your business.

Stunning Dashboards

Get insightful dashboards.

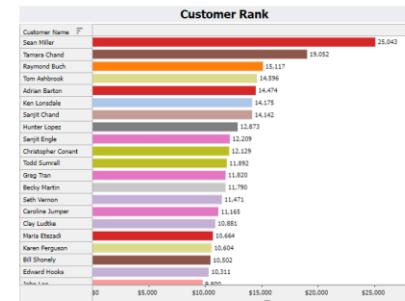
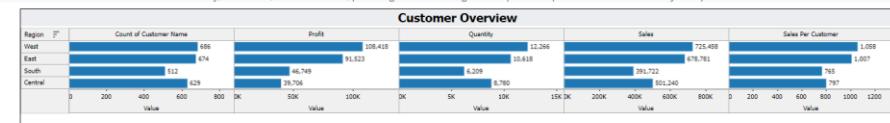
Narrative Stories

Amazing stories natted to get better understanding of the business.



Dashboard

Dashboard for Superstore Sales Analysis crucial to ensure that the information is easily understandable and actionable. Key considerations for designing a responsive and effective dashboard include user-centered design, clear and concise information, interactivity, data-driven approach, accessibility, customization, and security. The goal is to create a dashboard that is user-friendly, interactive, and data-driven, providing actionable insights to improve the performance and efficiency of Superstore.



Previous Dashboard

Next Dashboard

Product Gallery

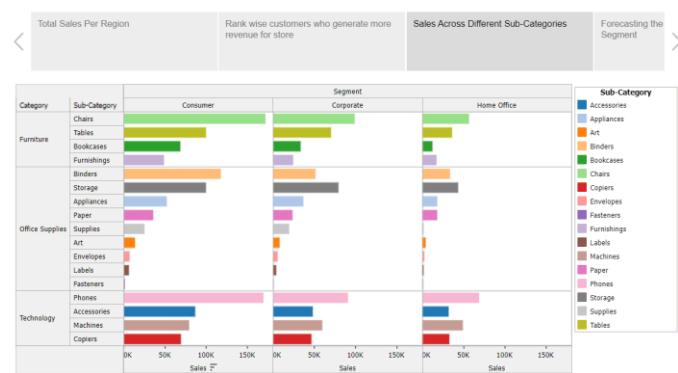
There's a wide variety of product in different categories. We have all the products that your office may need from furniture to stationery supplies to technical products like printer, keyboard,etc.



Story

The number of scenes in a storyboard for a data visualization analysis of the performance and efficiency of Superstore Analysis will depend on the complexity of the analysis and the specific insights that are trying to be conveyed. A storyboard is a visual representation of the data analysis process and it breaks down the analysis into a series of steps or scenes.

Superstore Story



SB-Store

Home About Dashboard Story Contact

Products

Our supermarket sells products from different category. Our products are shipped worldwide to variety of customer groups.

Furniture
Wide variety of quality furniture products

Office Supplies
All the things that you need for your office

Technology
Wide variety of products in technology

Contact Us

Location: Hyderabad, Telangana
Email: info@thesmartbridge.com
Call: +1 8888 888888 88

Please feel free to reach out to us in case of any enquiry.

SB-Store

Home About Dashboard Story Contact

Location: Hyderabad, Telangana
Email: info@thesmartbridge.com
Call: +1 8888 888888 88

Please feel free to reach out to us in case of any enquiry.

SB-Store

Hyderabad, Telangana, India
Phone: +1 8888 888888 88
Email: info@thesmartbridge.com

Useful Links

- > Home
- > About us
- > Services
- > Terms of service
- > Privacy policy

Our Services

- > Web Design
- > Web Development
- > Product Management
- > Marketing
- > Graphic Design

© Copyright Smartbridge. All Rights Reserved

Milestone 9: Project Demonstration & Documentation

Below mentioned deliverables to be submitted along with other deliverables

Activity 1:- Record explanation Video for project end to end solution

Activity 2:- Project Documentation-Step by step project development procedure

Create document as per the template provided