Tableau Fundamentals

By: D.B.Biradar

@2021

URL: https://www.scodeen.in/

Email: hr@Scodeen.in

Address:

SCODEEN GLOBAL, GIRINAGAR HOSINGSOCIETY, FLATNO25, BALAJINAGAR, DHANKAWADI, KATRAJ, PUNE-43 Maharashtar, INDIA
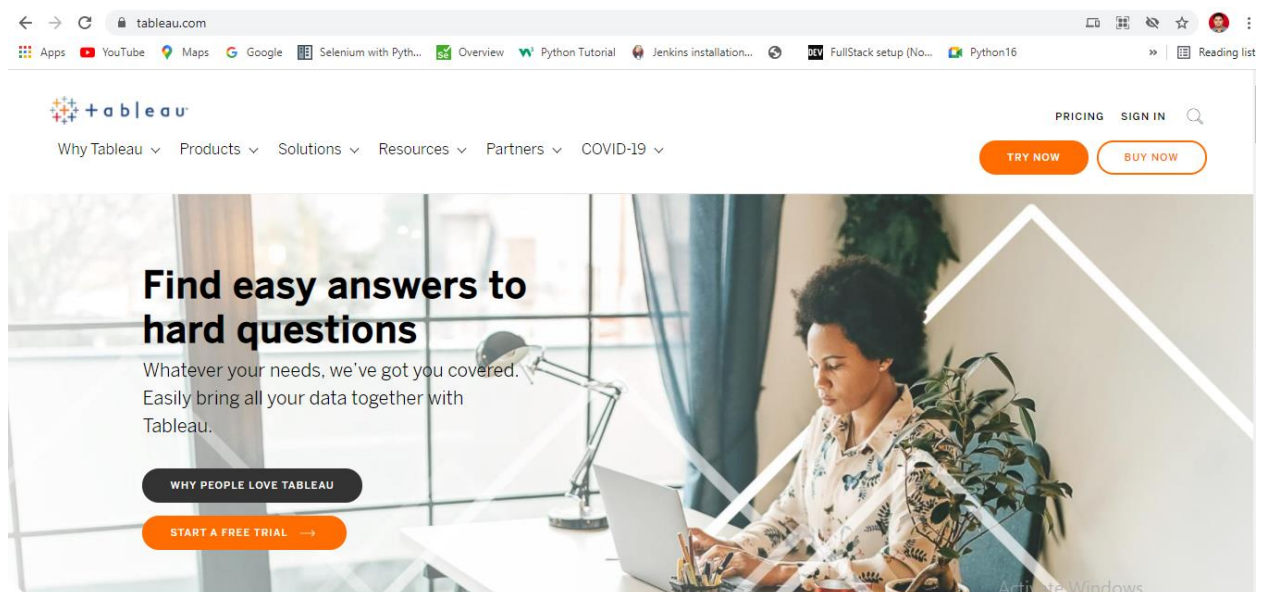
Tableau:

- ❖ Tableau is a powerful data visualization tool used in the business intelligence industry.
- ❖ It simplifies the raw data into a very easily understandable format.
- ❖ It visualizes and creates interactive, sharable dashboards.
- ❖ Tableau is not required any technical or programming skills.
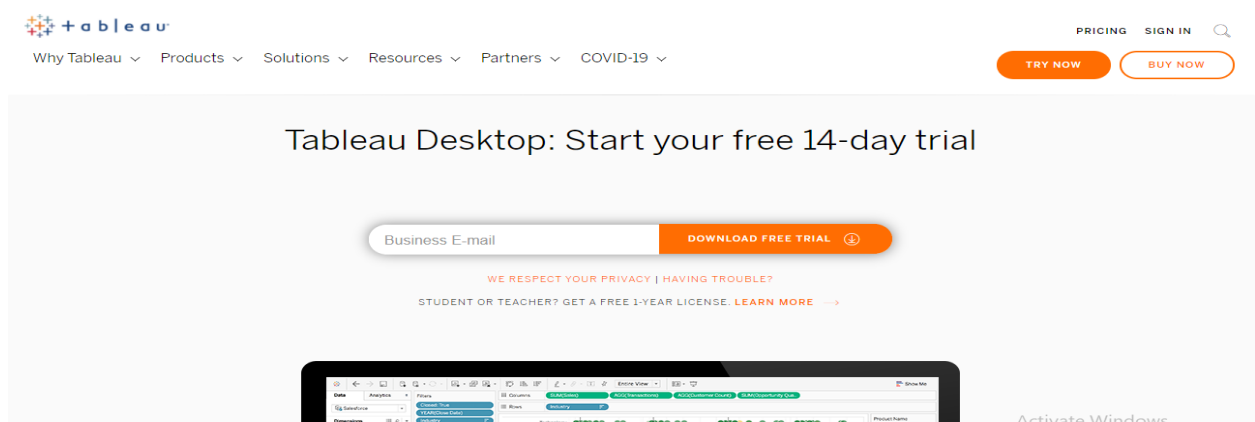
Tableau Installation:

Navigate to the below URL and download 14 days trial version of Tableau Desktop

URL: https://www.tableau.com/

1. Click on Free trail

2. After click on free trail it will navigate to next window

3. Enter your email id and click on download then it will download below file

   File: TableauDesktop-64bit-2021-2-0

4. Click and run it will install tableau.

There are three main features of Tableau, such as:

1. **Data Blending:** Data blending is an essential feature in Tableau. It is used when we combine related data from multiple data sources, which you want to analyze together in a single view, and represent in the form of a graph.

2. **Real-time Analysis:** Real-Time Analysis makes users able to quickly understand and analyze dynamic data when the Velocity is high, and real-time analysis of data is complicated. Tableau can help extract valuable information from fast-moving data with interactive analytics.

3. **The Collaboration of Data:** Data analysis is not isolating task. That's why Tableau is built for collaboration. Team members can share data, make follow up queries, and forward easy-to-digest visualizations to others who could gain value from the data. Making sure everyone understands the data and can make informed decisions is critical to success.

Tableau tools:

1. **Tableau Desktop:** Tableau Desktop establishes connectivity between the Data Warehouse and other various types of files. The dashboards and the workbooks created here can be either shared locally or publicly. It allows us to code and customizes reports. Right from establishing the stories, charts to blending them all to form a dashboard, all the necessary work is created in Tableau Desktop.

2. **Tableau Public:** This Tableau version is specially built for cost-effective users. The word '**Public**' means that the created workbooks cannot be saved locally. They should be kept on the Tableau's public cloud, which can be accessed and viewed by anyone.

3. **Tableau Online:** Its functionality is similar to the Tableau server, but data is stored on the server that hosted on the cloud, which is maintained by the Tableau group. There is no storage limit on the data which is published in the Tableau Online. Tableau Online creates a direct link over 40 data sources who are hosted in the cloud such as the **Hive, MySQL, Spark SQL, Amazon Aurora**, and many more.

4. **Tableau Server:** The software is correctly used to share the workbooks, visualizations, which is created in the Tableau Desktop application over the organization. To share dashboards in the Tableau Server, you should first publish your workbook in the Tableau Desktop. Once the workbook has been uploaded to the server, it will be accessible only to the authorized users.

5. **Tableau Reader:** Tableau Reader is a free tool which allows us to view the visualizations and workbooks, which is created using Tableau Desktop or Tableau Public. The data can be filtered, but modifications and editing are restricted. There is no security in Tableau Reader as anyone can view workbook using Tableau Reader.

There are several types of data terminologies in Tableau, such as:

1. **Bookmark:** A .tbm document in the bookmarks folder in the Tableau repository that contains a single worksheet. It helps in improving data analysis. Unlike, web browser bookmarks, .tbm files are a compatible way to display various studies quickly.

2. **Workbook:** A workbook is a file with .twb extension that holds one or more worksheets as well as dashboards and stories.

3. **Dashboard:** The dashboard is a combination of several views that are arranged on a single page. In Tableau, dashboards are used to observe and compare a variety of data together, and also it allows interacting with other worksheets.

4. **Data Source Page:** Data Source is a page where you can set up your data source. Does this data source page generally consist of four main areas? Join area, left pane, a preview area, and metadata area.

5. **Worksheet:** The worksheet is a collection of sheets. It's a place where you build views of your data by dragging various fields onto the shelves.

6. **Dimensions:** Dimension is commonly known as a field of categorical data. Dimensions hold discrete data such as members and hierarchies that cannot be aggregated. It also contains characteristic values such as dates, names, and geographical data. The dimensions used to reveal details of your information.

7. **Measures:** measures are the measurable quantities of the data, which can be analyzed by a dimension table. Measures are stored in a table which contains foreign keys referring uniquely to the associated dimension tables. The table supports data storage at the atomic level and thus, allows the number of records to be inserted at one time. **For example**, a Sales table can have a product key, customer key, promotion key, items sold, referring to a specific event.

8. **Filters shelf:** Filter shelf is located on the left side of the workbook. Filters shelf is used to exclude the data from a view by filtering it using both dimensions and measures.

9. **Pages shelf:** Page shelf is on the left side of the view. With the help of the page shelf, you can split a view into a sequence of pages based on the values and members in a continuous or discrete field. Adding a field with the pages shelf is similar to adding a field in rows shelf. For each new row, a new page is created.

10. **Marks card:** Marks card is on the left side of the worksheet. The user can drag fields to the control mark properties such as color, type, shape, size, label, detail, and tooltip.

Difference between .twb and .twbx extension:

.twb	.twbx
A .twb is an xml document that contains the information about your dashboards, sheets, and stories.	It doesn't contain any data so to share your workbook, and you have send both workbook file and the data source file of Tableau.
A .twbx is a package of files "compressed" together.	It includes data source file and any other file used to produce the workbook, including images.

Difference between traditional BI tools and Tableau:

Traditional BI tools	Tableau
Traditional BI architecture has hardware limitations.	Tableau doesn't have dependencies.
It based on the complex set of technologies.	Tableau based associative search that makes it fast and dynamic.
It does not support in-memory, multi-thread, and multi-core computing.	It supports in-memory when used with advanced technologies.
It has a predefined view of data.	It uses predictive analysis for various business operations.

Data types in Tableau:

Tableau expresses fields and assigns data types automatically. If the data source appoints the data type, Tableau will use that data type. If the data source doesn't individually assign a data type, Tableau will assign one.

Tableau consists of the following data types:
- ❖ Date values
- ❖ Text values
- ❖ Numerical values
- ❖ Date and time values
- ❖ Boolean values (True or False conditions)
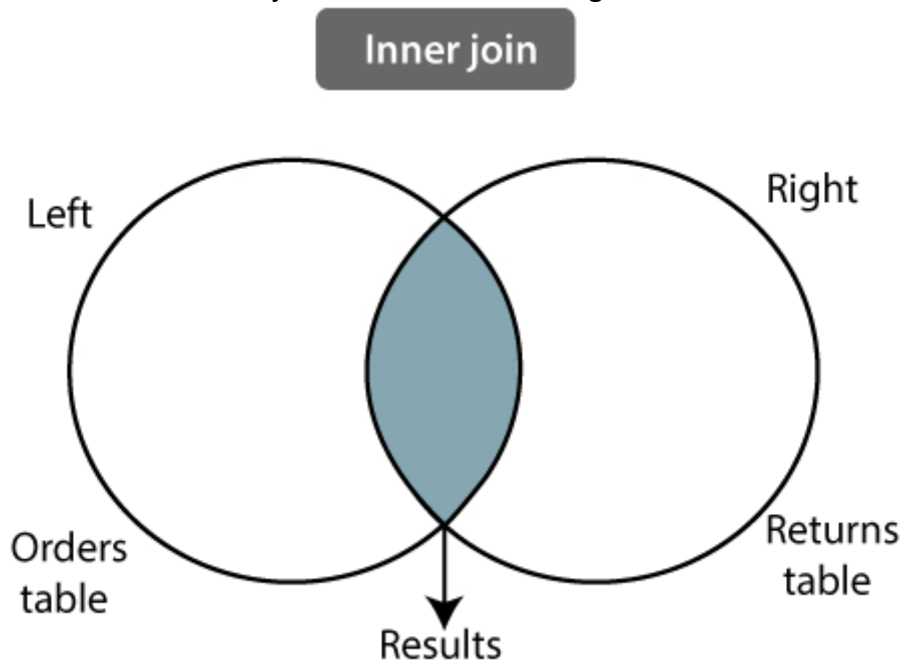- ❖ Geographic values (longitude and latitude used for maps)

Joins in Tableau:

The joins in Tableau are the same as SQL joins. There are five main joins, such as:

1) **Cross Joins:** Cross join produces rows which combine each row from the first table with each row from the second table.
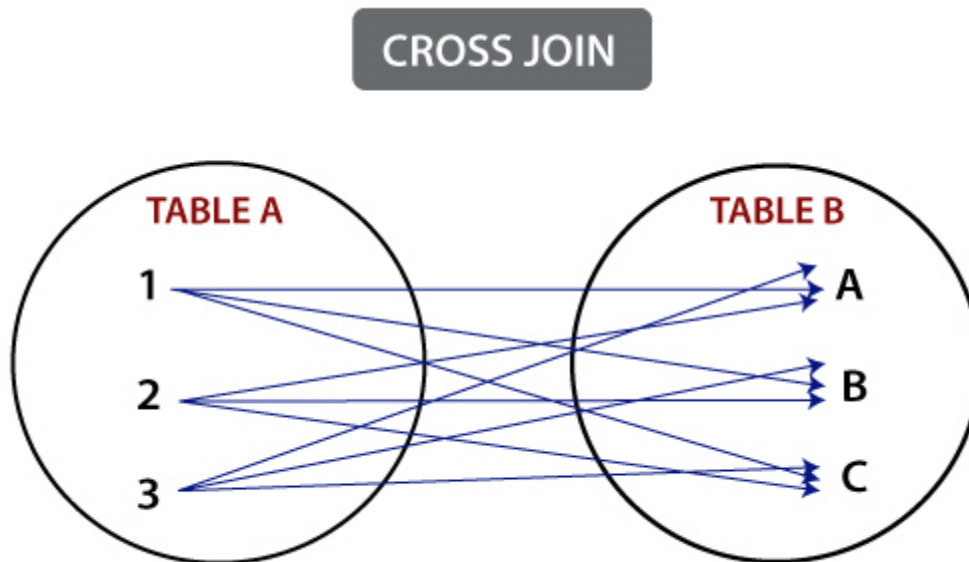

2) **Inner Joins:** An inner join returns the matching rows from the tables that are being merged.

3) Left Outer Join: The left outer join returns matching rows from the tables being joined, and also non-matching rows from the left table in the result and places **NULL** values in the attributes that come from the right table.
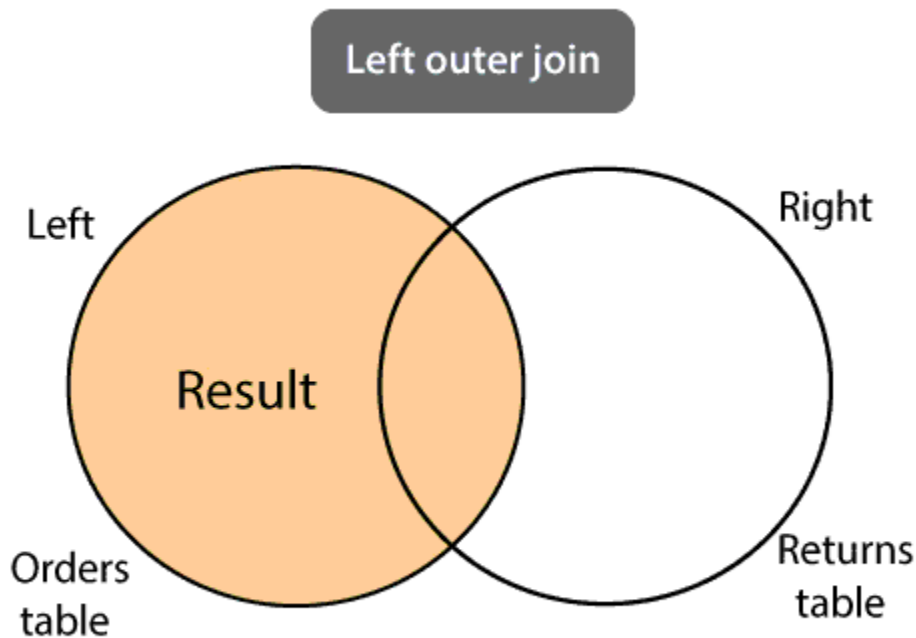
4) Right Outer Join: The right outer join operation returns matching rows from the tables being merged, and also non-matching rows from the right table as the result and places **NULL** values in the attributes that come from the left table.
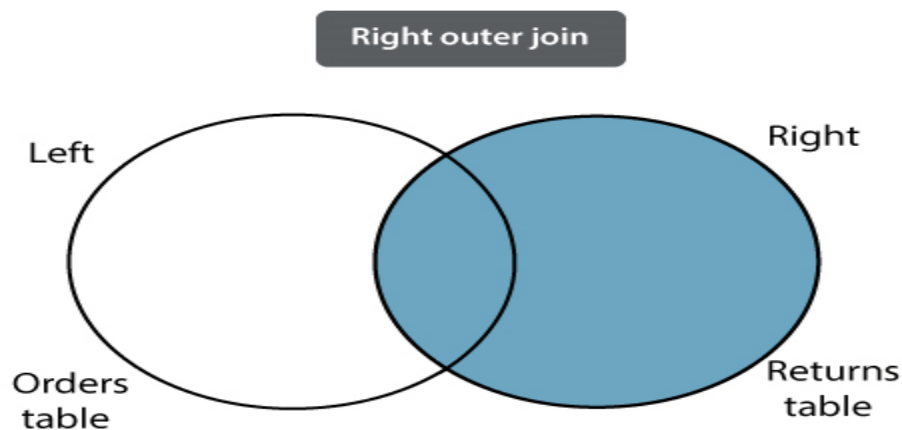
4) **Full Outer Join:** The full outer join is used to combine tables. As a result, it contains all values from both tables.

Full outer

Left   Right

Orders   Returns
table   table
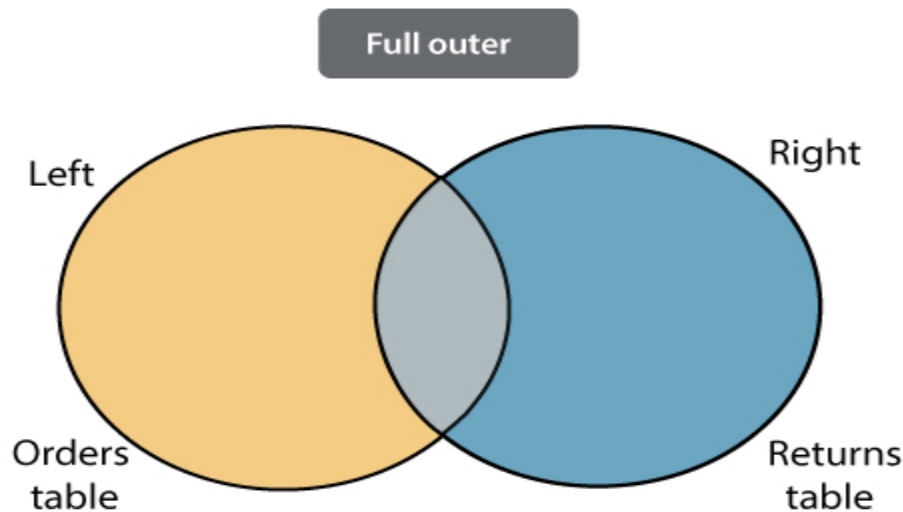
**In Tableau, you can join an at most 32 tables.**

Different types of connections:

We can connect the datasets in two ways, such as:

1. **Live:** Live connection sends queries to your database and retrieves data. These queries will return whatever data is currently in the database.
2. **Extract:** Extracts connection saved subsets of data that use to improve performance or to take the advantages of Tableau functionality which are not supported or available in your original data.

Different types of Shelves on Tableau Dashboard:

Shelves are the named areas placed on the left and top of the view in the worksheet. You can build a view by drag and drop fields onto the shelves. Some shelves are available when you select a specific mark type.
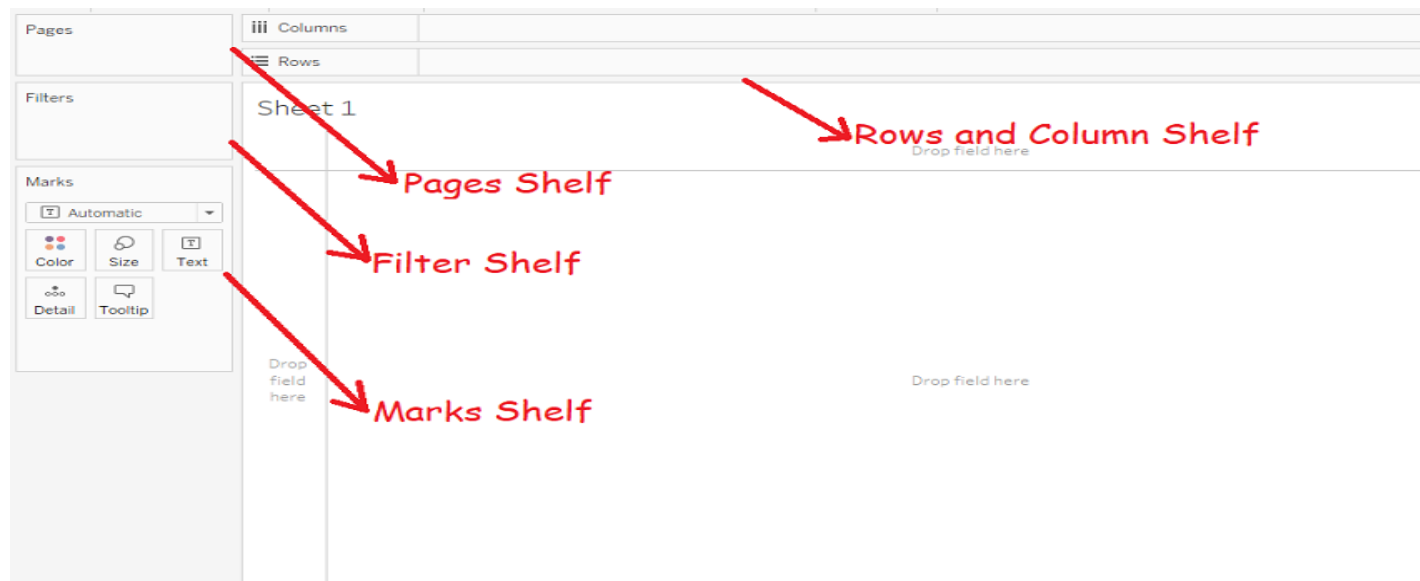
Dual axis:

The dual-axis is 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.

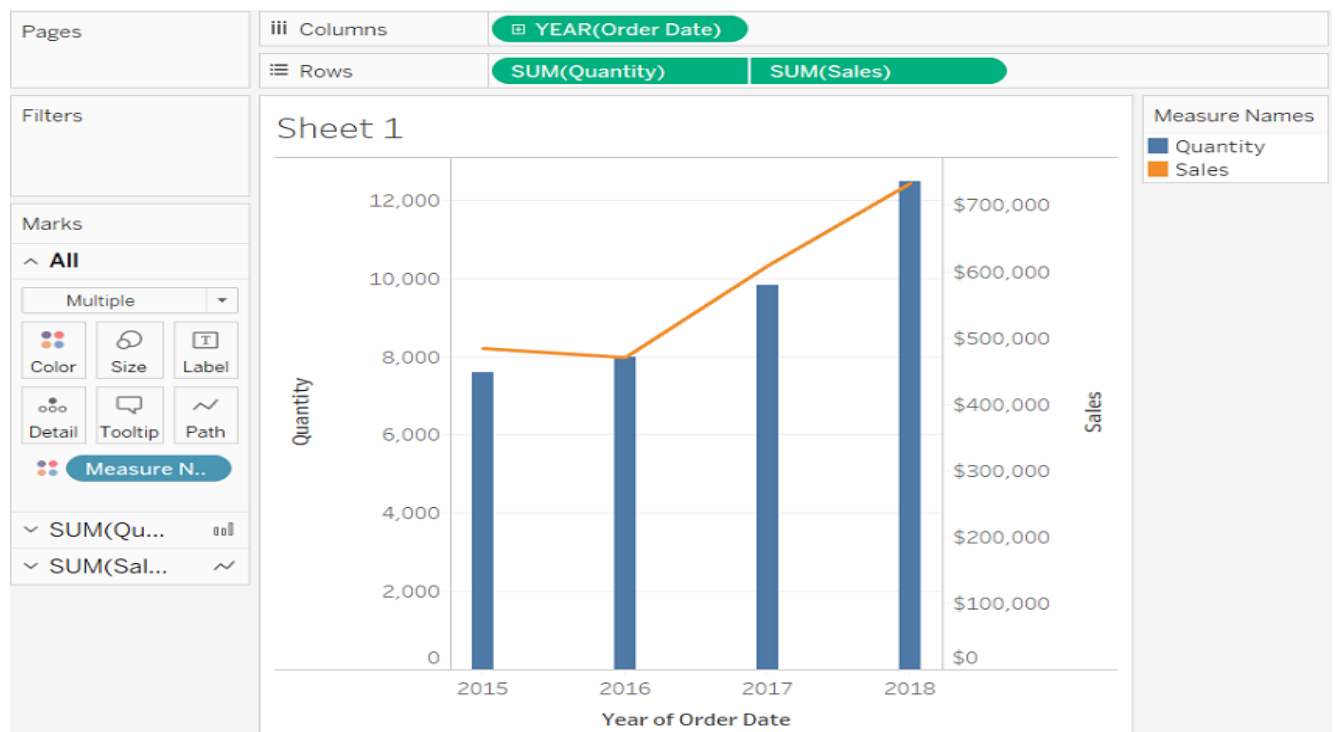The different scales are used in the graph that helps the user to understand both measures.
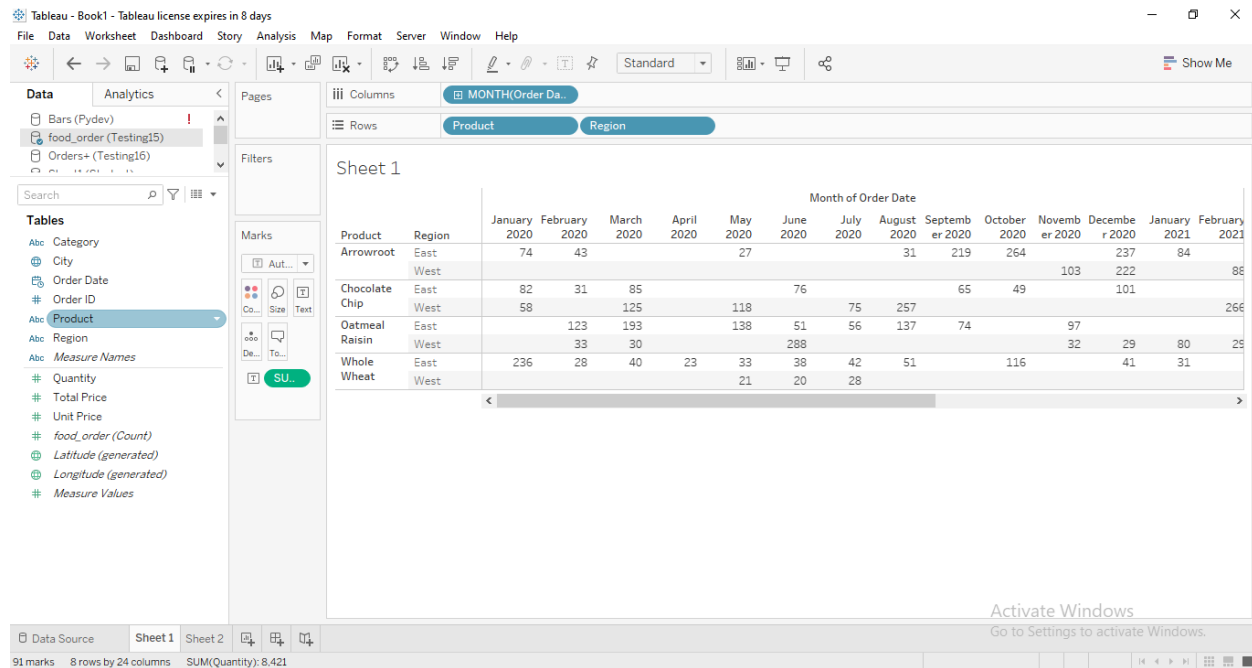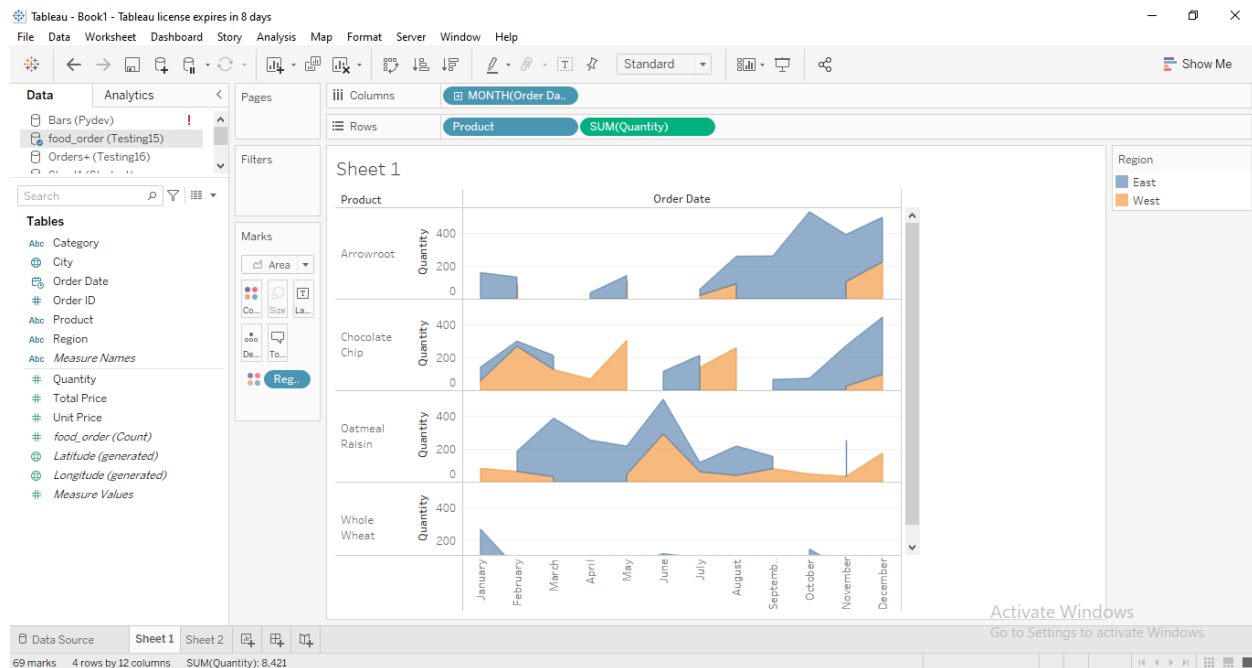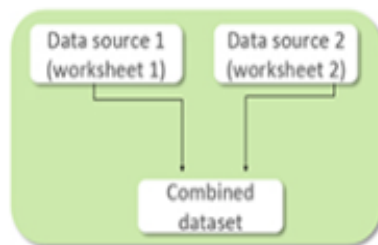
Tableau Text Report:

Tableau Chart Report:

Data Joining	**Data blending**
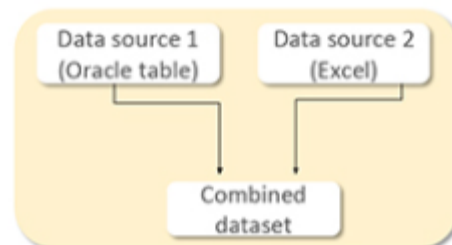
Difference between Data Joining and Data blending:

Data joining	Data blending
Data joining is used when you are combing the data from the same source.	Data blending is required two completely defined data sources in a report.

Tableau Story

A story is a sheet which contains a sequence of dashboards or worksheets that work together to deliver information.You can create stories to show how the facts are connected, provide context, and demonstrate how decisions relate to outcomes or make a compelling case. Each sheet in a story is known as a story point.

Different Types of Tableau files:

❖ **Bookmarks:** It contains only single worksheet and its easy way to share your work.
❖ **Workbooks:** Workbook can hold one or more dashboards and worksheets.
❖ **Packaged workbooks:** It contains the workbook along with any supporting local file data and background images.
❖ **Data extraction files:** Data extraction files are a local copy of a data source or a subset.
❖ **Data connection files:** Data connection file is a small XML file that contain various connection information.

Tableau Testing Document:

Basically to Test any Tableau report or BI report we need the below documents.

1. SRS/BRS Document about the report
2. Data Usage Sheet (Specifies the Attribute/Column details about the report).
3. Report Wireframe (Blue print of report)
4. Data Sample sheet (Data populated in report format)

Tableau Report Wireframe/Blue Print:

RPT001_TAX Payment Summary Report													

Date Generated:
Financial Year Start Date:
Financial Year End Date:
Tax Refrence Number:

Tax Account Number	Title	First Name	Last Name	Address Line 1	Address Line 2	Zone	State	Country	Tax File Number	Total Gross Tax	Total Net	Total Tax Withhel	Payer's PAN Number	Payer's Name

Tableau Test Cases:

Test Case Name	Test Step	Description(Design Step)	Expected Result(Design Step)
TC001_Generate_RPT001_Tax Payment Summary Report_From Tableau Server.	1	Login to tableu Server https://scodeen.sharepoint.com/sites/WX98770/SitePages/Home.in	Login Sholud be succesful
	2	Navigate to Report icon and click.	Report Icon should be able to click present on Scodeen
	3	Verify all the reporting view thumbnails are present.	The reporting Thumbnails should be present.
	4	Navigate to RPT001 Tax Payment Summary report and click and open.	RPT001 Tax Payment Summary report should be present.
	5	Verify the user is able to download the report in the below format for the selected report Image Data Crosstab(.xlsx,.CSV) PDF Powerpoint	User should be able to download the report in the below format for the selected report Image Data Crosstab(.xlsx,.CSV) PDF Powerpoint
TC002_Verify_RPT001_Tax Payment Summary Report_Logic.	1	All the Tax payment related information is given for calender year	The report should contains all the tax payment related information for a given calender year.
TC003_Verify_RPT001_Tax Payment Summary Report_Layout.	1	Verify Tax payment report is populated as per report wireframe.	Tax payment summary report should populate as per report wireframe.
TC004_Verify_RPT001_Tax Payment Summary Report_Attributes	1	Verify the below attributes from Tax Payment Summary Report from Tableau against DWH/BAL. Tax Account Number Title First Name Last Name Address Line 1 Address Line 2 Zone	All the attributes should be present as per report logic and report wireframe as against DWH/BAL v/s Tableau.
TC005_Verify_RPT001_Tax Payment Summary Report_Completeness.	1	Generate the report RPT001 Tax Payment Summary Report.	Report should generate succesfully.
	2	Verify that the report has picked all the attributes from source which satisfy the report logic	Report should pick all the attributes satisfying report logic.
	3	Verify the count of records in report matches with the no of records present in tableau report.	The count of records in report should match with the no. of records present in tableau report.

Different Tableau Validation:

Basically we need to perform the same kind of validation as we have performed in ETL but it varies in some aspects which is highlighted below.

1. Report Generation
2. Report logic.
3. Report Layout (Wireframe Testing)
4. Attribute validation (data validation)
5. Count validation.
6. Duplicate validation.
7. Report Completeness

In Tableau we need to do testing whatever the end point or end layer of ETL. I mean to say that here the data complete data has been stored in either it any be DWH (data warehousing) or BAL (Business application layer) In real time practice.

But in general we can extract the data from any source i.e. excel, .csv or any RDBMS   etc, in tableau and we can generate the report and create charts as per requirement.

***************************Thank You ******************************

Interview Questions:

1. What is tableau? Different tools of tableau?
2. Which tableau tool you have used in your project?
3. What Are the Data Types Supported in Tableau?
4. How will You Understand Dimensions and Measures?
5. What is meant by 'discrete' and 'continuous' in Tableau?
6. What are the different validations you will perform with respect to tableau report?
7. What are the different Document will you use in Tableau testing?
8. How will you validate the tableau data against DWH?
9. What Are the Filters? Name the Different Filters in Tableau.
10. There Are Three Customer Segments in the Superstore Dataset. What Percent of the Total Profits Are Associated with the Corporate Segment?
    Follow these steps:

    Drag segment field to the rows shelf. Here, segment consists of Consumer, Corporate, and Home Office
    Double-click on the profit field under Measures.
    Right-click on SUM (Profit) under marks card, select Quick Table Calculation and click on Percent of the total.
11. What Are the Different Joins in Tableau?
12. What is the Difference between Joining and Blending?
13. What are the Difference Between a Live Connection and an Extract?
14. What is a Calculated Field, and How Will You Create One?
15. How Can You Display the Top Five and Bottom Five Sales in the Same View?
16. We can display it using the In/Out functionality of sets.
    Follow these steps:

    Drag the Customer Name field to Rows shelf and Profit field to Columns shelf to get the visualization.
    Create a set by right-clicking on the Customer Name field. Choose to create an option and click on Set.
    Provide the name 'Top Customers' to the set. Configure the set by clicking on Top tab, selecting By field, and filling the values as Top, 5, Profit, and Sum.
    Similarly, create a second set called 'Bottom Customers' and fill the By Field values as Bottom, 5, Profit, and Sum.
    Select these two sets and right-click on it. Use the option Create Combined Set. Name it 'Top and Bottom Customers' and include all members of both sets. Pull the Top and Bottom Customers onto Filters.
17. What are the Difference Between Tree maps and Heat Maps?
    A **heat map** is not only defined by color, but you can also use its size. Here we define the size by sale by dragging the Sales tab to Size under marks card, comparing profit and sales through the color and size.

A **Tree map** is used to represent hierarchical data. The space in the view is divided into rectangles that are sized and ordered by a measure.

18. What is the Difference Between .twbx And .twb?
19. Explain the Difference Between Tableau Worksheet, Dashboard, Story, and Workbook?
20. What Will the Following Function Return?
Left (3, "Tableau")

Choose the correct answer:

Tab
Eau
Error
None of the above
It will return an error because the correct syntax is: left(string, num_chars). So, it should be: Left ("Tableau," 3)

21. What is the Rank Function in Tableau?
22. The ranking is assigning something a position usually within a category and based on a measure. Tableau can rank in several ways like:
rank
rank_dense
rank_modified
rank_unique

23. How Can You Schedule a Workbook in Tableau after Publishing It?
24. When you're signed in to Tableau Server, go to Content > data sources or Content > Workbooks, depending on the type of content you want to refresh.
Select the checkbox for the data source or workbook you want to refresh, and then select Actions > Extract Refresh.
In the Refresh Extracts dialog, select Schedule a Refresh, and complete the following steps:
Select the schedule you want.
If available, specify whether you want a full or incremental refresh.

25. How Do You Handle Null and Other Special Values?
If the field contains null values or if there are zeros or negative values on a logarithmic axis, Tableau cannot plot them. Tableau displays an indicator in the lower right corner of the view, and you can click the indicator and choose from the following options:

Filter Data
Excludes the null values from the visualization using a filter. In that case, the null values are also excluded from any calculations used in the view.
Show Data at Default Position
Shows the data at a default location on the axis.

26. How Can You Embed a Webpage in a Dashboard?
27. Follow these simple steps to embed a webpage in a dashboard:
    Go to dashboard
    Double click the 'Webpage' option available under 'Objects.'
    Enter the URL (here https://en.wikipedia.org/wiki/) of the webpage in the dialog box that appears
    You can see the webpage appears on the dashboard.
28. How Do You Make the Webpage Dynamic?
29. Begin by bringing Map by Sales into view. It shows the state's name and its sales.
    Go to the dashboard.
    Double click the 'Webpage' option available under 'Objects.'
    Do not provide a URL in the dialog box that appears and click on Ok.
    Click on the Dashboard in the menu and select 'Action.'
    Click on 'Add Action' and select 'Go to URL.'
    Enter 'https://en.wikipedia.org/wiki/' under the URL option. Click on the arrow adjacent to it and select 'State.'
    Click on 'Select option' and hit 'Ok.'
    Now, when you click on any state like California, it brings up the California Wikipedia page. This is how to make it dynamic.
30. How Can You Optimize the Performance of a Dashboard?
    There are multiple ways to optimize the performance of the dashboard like:

    Maximize the number of fields and records. You can exclude unused fields from your visualization or use extract filters.
    Limit the number of filters used, by avoiding quick filters and using action and parameter filters instead. These filters reduce query loads.
    use Min/Max instead of Average because average functions require more processing time than Min/Max
    Use boolean or numerical calculations more than string calculations. Computers can process integers and boolean much faster than strings.
    Boolean > int > float > date-time > string
31. Which Visualization Will Be Used in the given Scenarios?
    To show aggregated sales totals across a range of product categories and subcategories
    To show the duration of events or activities
    To show quarter wise profit growth
    We would use the following visualizations for the given scenarios:

    Tree map
    Gantt chart
    Waterfall chart
32. What is the difference between Traditional BI Tools and Tableau?
33. What are the different datatypes in Tableau?

34. What are Measures and Dimensions?
35. How many maximum tables can you join in Tableau?
36. What are the different connections you can make with your dataset?
37. What are shelves?
38. What are sets?
39. What are groups?
40. What is a hierarchical field?
41. What is Tableau Data Server?
42. What is Tableau Data Engine?

It's an analytical database designed to achieve instant query response, predictive performance, integrate seamlessly into existing data infrastructure and is not limited to load entire data sets into memory

Tableau Data Engine is not really in-memory technology
43. What is disaggregation and aggregation of data?
44. Mention what are different Tableau files?