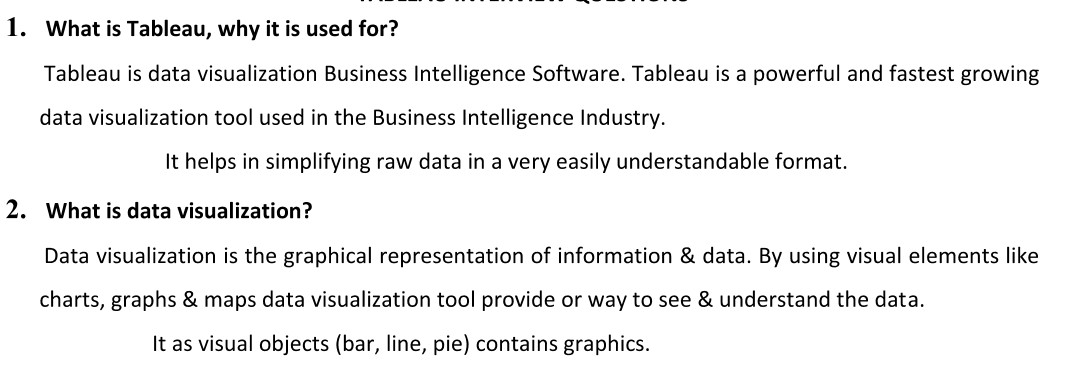
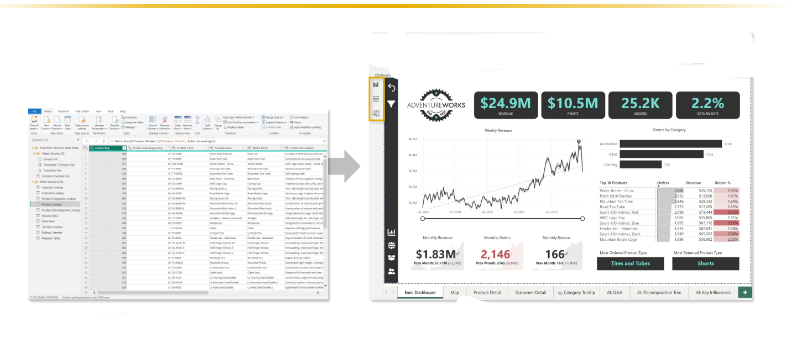
Tableau



enabling data-driven decision-making



Downloading Tableau

Tableau Desktop - Professional version

Option - 1

Tableau public download link

<https://www.tableau.com/products/public/download>

Option 2

Tableau Desktop download link for windows users

<https://www.tableau.com/products/desktop/download>

1. Tableau Desktop download after 20 and try to experience the feature which missed in public version

How to connect data source in Tableau

1. Explain this with Tableau data source connectivity page

TYPES OF CONNECTIONS

-LIVE

\*IT S A REAL TIME UPDATES

-EXTRACT

\*IT S SNAPSHAT OF COPY OF DATA

\* SCHEDULE REFRESH

15-04-2024

Types of extensions

-twb-tableau workbook -

* This file type is a plain XML file that contains all the necessary information for Tableau to connect to data sources, define visualizations, and create dashboards.
* .twb files do not include the actual data. They contain references to data sources, queries, and instructions on how to visualize the data.
* Since .twb files only contain references to data sources, they are generally smaller in size compared to .twbx files.

-twbx-tableau packeged workbook - save with packaged workbook with datasource - large memory

* This file type is a packaged version of a Tableau workbook. It contains the workbook itself (.twb) along with any external data sources and images used within the workbook.
* .twbx files bundle everything together into a single file, making it easier to share and transport workbooks, especially when they rely on external data sources.
* .twbx files are larger in size compared to .twb files because they include the actual data and any other resources needed for the workbook.
* They are essentially zip archives that contain the .twb file along with a folder containing any external files referenced in the workbook.

-tds-tableau datasource - only save data connection information

* TDS files are XML files that represent a Tableau data source. They contain metadata about the data source, including information about the connection to the data, data schema, and any transformations applied.
* Unlike TWB files, TDS files do not contain visualizations or worksheets. They are specifically used to define and manage connections to data sources.
* TDS files can be created and shared independently of workbooks. They allow users to maintain consistent connections to data sources across multiple workbooks.

-tdsx-tableau packaged datasource

* TDSX files are packaged versions of TDS files. They include the TDS file itself along with any local data files associated with the data source.
* Similar to TWBX files, TDSX files bundle everything needed for the data source into a single file, making it easier to share and transport.
* TDSX files are useful when you need to share a data source along with its associated data, ensuring that others can connect to the data source without needing access to the original data files.
* They are essentially zip archives that contain the TDS file along with any necessary local data files.

-tbm-tableau bookmark -

TBM (Tableau Bookmark):

* TBM files are essentially snapshots of Tableau workbooks or views.
* They capture the state of a particular workbook or view at a specific point in time, including the layout, filters, parameters, and data displayed.
* TBM files allow users to save and share specific views or configurations of Tableau workbooks with others.
* They are useful for sharing insights, collaborating on analyses, or documenting particular states of data visualization projects.

.Hyper exten

Fir extracting data and save in local machine

* tps-tableau preference

* You can create you own color palette

<?xml version='1.0'?>

<workbook>

<preferences>

<color-palette name = "test" type = "regular">

<color>#033445</color>

<color>#033445</color>

<color>#CD5C5C</color>

<color>#C0C0C0</color>

<color>#FA8072</color>

<color>#FF0000</color>

</color-palette>

</preferences>

</workbook>

17-04-2025

DATA TYPES

-NUMBER

\*WHOLE - 19

\*DECIMAL - 19.5

-STRING - tesafsdf455

-DATE - 25/01/2025

-DATE TIME - 25/01/2025 07:52AM

-GEOGRAPHY - city,state

-BOOLEAN - True, False



DIMENSION -

In Tableau, dimensions are essentially categorical variables or attributes that provide context or categories for your data. They are typically qualitative data that you can use to segment, group, or filter your data.

For example, if you're analyzing sales data, dimensions could include things like "Product Category," "Customer Segment," or "Region." These dimensions are used to break down the data and provide a basis for comparison or analysis.

-qualitative values

-it can't be aggregated

-eg name,id,rollno

It represent in blue color.

Measure -

In Tableau, measures are numerical values that you typically perform calculations on or aggregate in some way. Measures represent quantitative data and are used to perform mathematical operations like sum, average, count, etc. on your data.

For example, in a sales dataset, measures could include "Sales Revenue," "Profit," "Quantity Sold," or "Average Sales Price." These measures provide the numerical values that you want to analyze or visualize.

-quantitative values

-it can be aggregated

-eg sales, profit,withdrawals

It represent in green color.

discrete-individually separate and segment ,finite values.

continuous-forming an unbroken whole without any interpretation.

It represent in blue color.

It represent in green color.

Workseet

-it contains data,analytics,shelves,single view,legends

Row shelf, Column shelf,filter shelf,pages shelf

Page shelf - similar to flip book animation

Change order date to drop down and choose date value month

Change chart type to circle

marks

-color

-size

-label

-detail

-tooltip

-angle -

-path -

-shapes -

How to Plot the chart

default chart for visualization-bar chart

default chart for visualization while using date-line chart

Preview of the menu options

Create simple chart

1. Region wise sale in Bar chart and stacked with category splitup and Horizantal bar
2. Line chart with order data sale
3. Text chart - with order data year wise sale
4. Area char - with order data wise sale

Customized cart

donut chart

lollipop chart

Butterfly chart

=====================================

1.Combine -Is only works on dimension view.(Its help to concatenate the two different dimension)

2.Group - Its a one dimensional view,Help to higher level category by the lower level

Its works on dimension and measure level.

3.Drill Down (Hierarchy) -Default year,

4.Set -Set is a sub set of data.With set we can find top N or Botttom N.

5.Parameter - A parameter is a workbook variable such as a number, date,

or string that can replace a constant value in a calculation, filter, or reference line.

We can give dynamic option to client.

A parameter is a workbook variable such as a number, date,

or string that can replace by a constant value

Calculated Field

Create a new field, based on previous data with similar or few contion.

We can create a new data(Diemsnion or measure) based on the existing data with syntax in calculated field.

IF [Select]=1 THEN [Sales]

ELSEIF [Select]=2 THEN [Profit]

ELSEIF [Select]=3 THEN [Discount]

END

Dynamic Title for Visualization:

* Scenario: Create a parameter for users to select a region, and then use it to dynamically change the chart title.
* Calculation:

'Sales by Region: ' + [Select Region Parameter]

Using calculation How to map start and end data

[Start Date]<=[Order Date]

AND

[End Date]>=[Order Date]

=================================================

Filter - Help to reduce and minimze the datas.

Types - 1.Data Source Filter - it's a filter where we can remove the specific info from data source itselft. ex emp removal

2.Extract Filter - it's filter where we can remove the info after taking copy of data source in a local mechine - done at sheet, on right click on data

3.Quick Filter -

a) Dimension Filter - based on dimension

b) Measure Filter - based on measure

4.Context Filter : Grey Colour -

5.Cascading Filter - Help to a relevant value between parent and child filter.

Filter Executing Order :

1.Extract Filter - Region -

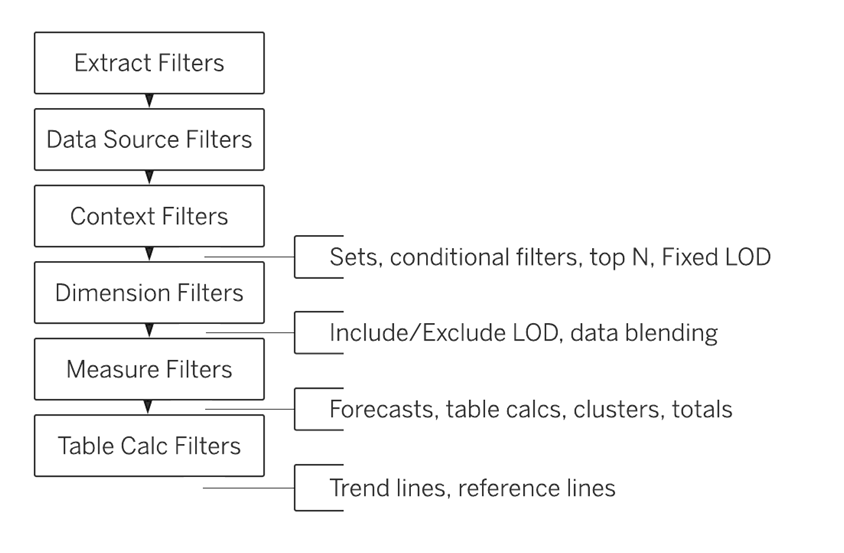
2.Data Source Filter -

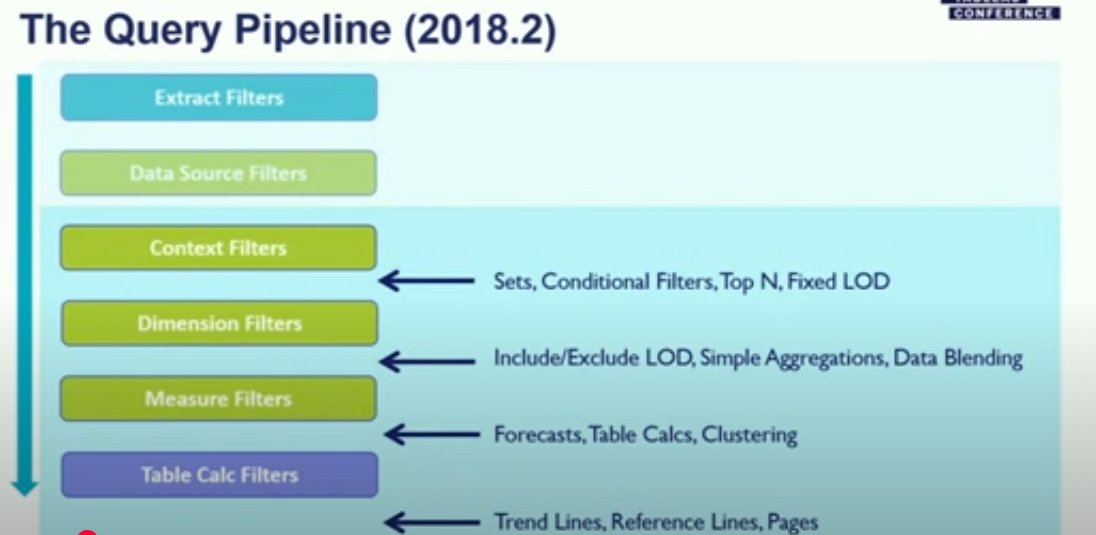
3.Context Filter -

4.Dimension Filter -

5.Measure Filter -

6.Table Calculation Filter





in filter - in default it will be in only this workbook so filtering data will apply only current worksheet we have also

\* All using related datasource -

 indicating that the filter is being applied to multiple data sources.

\* All using this datasource - only applied on this data source

\* seleted worksheet - customized selection

\* only worksheet - this workbook only

====================================================

Map and how to add city with logitude and latitude

Sample source available here:>> D:\Srini Doc\Tableau\class 4

Axis

====

DUAL AXIS

A **dual axis** in Tableau allows you to overlay two measures on a single chart while maintaining independent axes for each measure. This is useful when you want to compare two related metrics that have different scales or units.

BLENDED AXIS

Blended axes in Tableau are used to display multiple measures on the same axis, making it easier to compare values side by side in a single chart

SYNCHRONIZING AXIS

The **synchronization axis** in Tableau is a feature that aligns the scales of two or more measures when they are displayed on a dual-axis. It ensures that both axes are directly comparable by using the same scale.

ANALYTICS

CONSTANT LINE

AVERAGE LINE

TOTALS

FORECAST

CLUSTER

Quick table calculation

Running total

difference

percent of difference - Prev month sales - current month sales = value / Prev month sales \* 100

percent of total - Individual value / total value \*100

Rank -

percentile-(rank-1/toral no row-1)\*100 - Keep the Rank in ASC

moving average

==============================================

FUNCTIONS

NUMBER FUNCTIONS

ROUND - if the value is going beyond .5 then it converts to the next values

ABS - Absolute use to convert the given value to positive values

CELI - Ceiling function convert the given value to the next value if decimal value exist

FLOOR - Floor function simply remove the decimal

SIGN - using this we can identify the negetive values

SQUARE - simply square the values

SQRT - simply square root the values

============================================

AGGREATE FUNCTION

SUM

AVG

MAX

MIN

COUNT -10,10,20,20,30 = 5

COUNTD-10,10,20,20,30 = 3

================================

LOGICAL & GENERAL FUNCTION

==>>> IN

IN >> sum(quantity) in (28,33) will return true or false

if [Sub-Category] in ('Art','Chairs') THEN TRUE

END

==> AND == > OR

AND >> IF (ATTR([Market]) = "New Business" AND SUM([Sales]) > [Emerging Threshold] )THEN "Well Performing"

IF SUM([Sales])>10000 AND SUM([Profit]) > 1 THEN 'Good Progress'

END

OR >> IF [Profit] < 0 OR [Profit] = 0 THEN "Needs Improvement" END

IF SUM([Sales])>10000 OR SUM([Profit]) > 1 THEN 'Good Progress'

END

==> CASE

CASE [Region]

WHEN 'West' THEN 1

WHEN 'East' THEN 2

ELSE 3

END

==>> may not working in Measure so apply below workaround

CASE [Sales\_Aggregate] WHEN 100 THEN "Low" WHEN 500 THEN "Medium" ELSE "High" END

==> IF

IF SUM([Sales])>10000 OR SUM([Profit]) > 1 THEN 'Good Progress'

END

IF [Region]='East' THEN 1

ELSEIF [Region] ='West' THEN 2

ELSEIF [Region]='South' THEN 3

ELSE 4

END

==> IIF

IIF >> IIF(test, then, else, [unknown]

Task 1

|  |  |
| --- | --- |
|  | sales category |
| < 1000 | low sales |
| >= 1000 | Avg sales |
| >= 10000 | high sales |

Null functions

===========

IF NULL >> IFNULL(expr1, expr2) IFNULL([Profit], 0)

IS NULL >> ISNULL(expression) ISNULL([Profit])

ZN >> ZN(expression) Returns <expression> if it is not null, otherwise returns zero.

Example: ZN(sum([Profit]))

STRING FUNCTION

=================

UPPER ==> convert to upper case

LOWER == > convert to lowercase

LEFT == >

RIGHT

MID

SPLIT

CONTAINS

STARTSWITH

ENDSWITH

FIND

FINDTH

Task-1

convert upper case customer name into Proper format eg Stven smith

Task - 2

email

student@gmail.com

Student2@yahoo.com

Student3@hotmail.comEmail

Student4@hcl.com

gmail

yahoo

Hotmail

hcl

=========================================================================

Table Calculations

================

FIRST

LAST

INDEX

SIZE

RANK

RANK\_DENSE

RANK\_MODIFIED

RANK\_UNIQUE

RANK\_PERCENTILE

Rank sample data set - D:\Srini Doc\Tableau\class 7

DATE FUNCTION

=================

DATE

Returns a date given a number, string, or date <expression>.

DATE([Employee Start Date]) - if the date along with time we can convert with date

DATE("September 22, 2018")

DATE("9/22/2018")

DATE(#2018-09-22 14:52#)

TODAY

NOW

DATE DIFF 01-Jan-2024 - 03-April-2025 - 2 days (day), 1 Month (Month),1 quarter (quarter)

DATE PART - 05-Jan-2024 - 2

DATE NAME -

DATE TRUNC-19/aug/2024 - aug-01-2024 (month) - 01-jan-2024 (year), july-01-2024

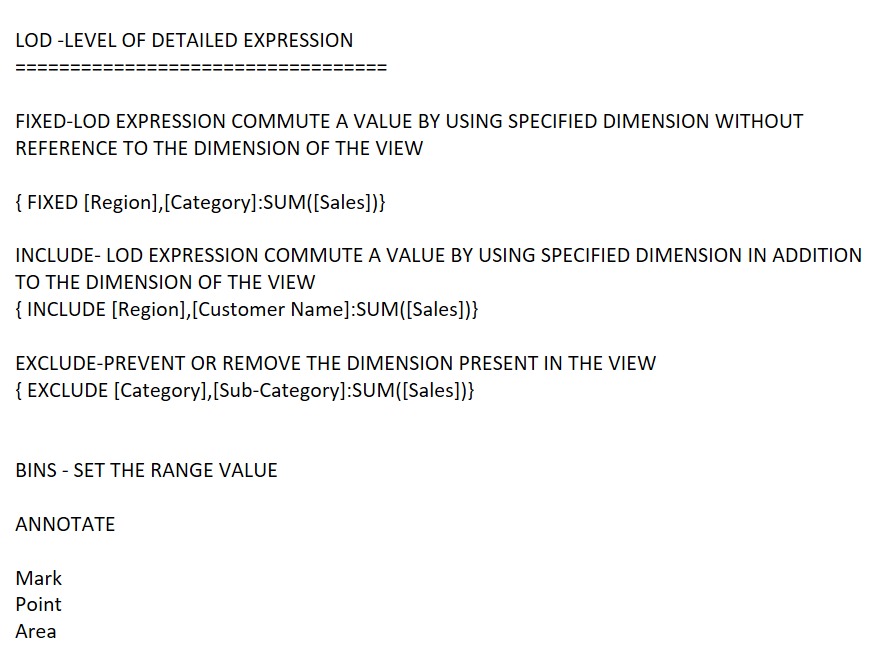
01-08-2024 01/12/2023 01-01-2024

01-01-2024 01-01-2023

01-07-2024 01-10-2023

DATE PARSE

29-04-2025



Using all of these option we can make some annotation on the point or corresponding area

Data Modelling :-

Data modeling in Tableau refers to the process of organizing and structuring data to make it suitable for analysis and visualization. This involves defining relationships between different data sources, creating calculated fields, and ensuring that the data is in the right format for Tableau to use effectively.

1.Joins

2.Relationships

3.Union

duplicate allowed- similar to union all

4.Custom SQL

5.Blending - different data source

6.Cross DB Join - it allows to combine table across two different datasource, using this we can perform some type of join whereas we can't do it in blending

=====================================

Cross DB Join

Join the tables from different data source

Data Blending

1.Combine the data from the different data source. 1.Excel and 2.Orcle and 3.My SQL

Default join Left Outer

=========================================================================

Blending - two diff database

\* Linking field

\* Blending can be performed worksheet to worksheet

\* Data source should be at least 2

\* Retains the original structure - Table columns are separate for each of the data source

\* At least 1 connection per data source

\* primary data source denotes in blue colour and sec data source denotes in orange colour

\* In blending the matching value gets aggregated based on left table linking field. if the value is non aggregable then it returns \*. Example - Queries will be sent to each of the data source separately and second table get aggregated then returns the values.

\* numbers are aggregated based on primary linking field

Data blending at a glance

* Data blending occurs on a sheet-by sheet basis.
* The order in which fields are used determines which data source is the primary versus the secondary data source.
* The primary data source is indicated with a blue check mark, any secondary data sources and fields from secondary data sources have an orange check mark.
* Linking fields can be automatically determined based on shared field names or the blend relationship can be manually created.
* Data blending behaves similarly to a left join, which may result in missing data from the secondary data source.
* Asterisks (\*) may appear. This indicates multiple dimension values in a single mark, because data blending takes aggregated results and combines them in the view.
* A secondary data source can be used to re-alias the field values in a primary data source. For more information, see [Alias Field Values Using Data Blending](https://help.tableau.com/current/pro/desktop/en-us/multipleconnections_edit_primary_aliases.htm)

Joins

\* diagram represent the joining

\* it will not work worksheet to worksheet so drawback would be we should use the same collection not like blending.

\* Join upto 32 tables

\* Recommended within 1 data source

\* can use calculated fields to match

\* Connection can be multiple

\* It produces the duplicates if we have multiple match (based on granularity).

\* Does not retain the original table instead make physical join - Table columns are not separate and it produce only one

\* Duplicated item produces aggregated values Ex - sum(sales) which can be avoid by using Lod, example - fixed customer\_id:min(salary)

How multiple join works

\* if we put product table with existing join of (customer (Left) customer order) then this result first generated as seperate table first and then new table (product detail) get joined.

Relationship

\* Noddle joining

\* Introduce after 2020.2

\* after 2020.4, can use calculated fields to match

\* Recommended within 1 datasource

\* Connection can be multiple

\* Retain original table

\* Join and aggregate - Logical- context specific this smart to know which field you are combining

