

* TABLEAU DESKTOP *

* Introduction To Tableau :-

- It is a BI (Business Intelligence) Tool

* Tools offered by Tableau :-

- 1) Tableau Desktop
 - 2) Tableau Online
 - 3) Tableau Server
 - 4) Tableau Reader
 - 5) Tableau Public
- } Used to create different types of report (Dashboards)
- } Publishing report (After report someone manager has to see report), so we use the 3 tools to publish.

- Tableau has great compatibility with many databases & application like

- Spreadsheet (Excel)
- Databases
- Bigdata (Hadoop, etc)
- Data warehouse
- Cloud Application
- Cloud Databases
- 50+ different databases

- Tableau Report can be published on different devices & has compatibility with

- 1) Desktop
- 2) Tablets
- 3) Mobiles

* Tableau divides the data into two Types / Parts

1) Dimensions :-

- Text
- Dates
- Geographical Data.

2) Measures :-

- Numbers (quantities)
- Figures
- Float values.
- Tableau Support 24 different types of charts used to prepare a report.

} calculated fields.

• Gartner → company which analyzes the performance of BI Tools all over world.

- Tableau is one of best

current version of Tableau.

2021-L-0

* Connecting to Different Database :-

- Go to connect to Data → To a file → different type of file & servers available.
- From here we connect to different files & servers to get the data & make the report.

* Tableau Add Excel File

↳ Excel File → Having more than one sheet in single file

- we connect to excel file → Go to connect data → excel file → select excel file.

- Here excel file has more than one spreadsheet in a file, they can be blended together by dragging & dropping & having relationship b/w them.

- The Data blending i.e. connecting spreadsheet b/w each other with common column.

* LIVE vs EXTRACT :-

- Live → Live i.e. when we are connected to live current data
- Extract → Takes snap of our previous data.
 - Data which is previously available i.e. of previous company or of previous year.
 - Also called as static data in some cases.
- When we connect with a excel worksheet with live connection & working on it.
- After some time when we update our excel sheet with some records add/deleted, to update these directly in tableau we only have to click refresh button.

Refresh is available at

- i) At Data source → table below → At top there is a refresh button
 - ii) In sheet we are working → in Data → below there is a DB database → right click & refresh, our sheet of report will be updated automatically.
- For Extract the data modification done in source or a DB cannot be refreshed or it has to be refreshed by in sheet → click on DB logo → in data → right click → extract → edit refresh
 - Data will be updated.

- Also when published using extract connection the update will be seen automatically after 15 min until it is forcefully refreshed

- If live we see Database as



← single (live)

- If Extract we see Database & file also have the icon as



← Double

(Extract)

* DATA TYPES IN Tableau :-

- When a DB or worksheet is connected to tableau & having different fields / attributes divided in Dimensions & Measures.

- When we open a data or view data , above each field we see a small icons indicating Data type.

1) → Number Data type (1,2,...)

2) → Character datatype / string. (Alpha numeric also).

3) Calender icon → Date datatype ()

4) Geographic datatype → Globe symbol ()
(state, country, zip code, city)

5) Boolean values → True / False. (T/F)

6) Date & time Values →

- On clicking on to these icons of any ~~field~~ field → we can change datatype of any field from here → But if we change the format & total structure & values of that particular column will get change.

- To keep that same , the datatype is recommended not to change in any condition except as per app'n needed.

* VIEW DATA :-

- There are different method to view data in tableau.
- 1) Drag + drop the sheet in the Data Source.
- 2) In data source → click on right part of sheet name icon present.
 - ↳ If we want to see custom i.e. top 10, 200, etc data we can type in Box & see.
 - ↳ And if we enter the no. of record which exceed more than present records it will show maxm available all records of a sheet (or source from where we got it).

- If we want to hide some of columns :-

- Open the data & view data by drag + drop in data source tab below → over a column heading there is a drop down → there we have hide option.
 - ↳ Near datatype.
- To unhide → on top click → Show hidden file → column are visible in faint color → again dropdown → select unhide.

COLUMN FORMATTING :-

i) rename :- Just double click on column Name to rename.
OR By Right click.

ii) Reset - Right click → click on Reset Name → original name will appear

iii) Copy Column | Entire Data from Tableau :-

- Control + A in data → ^{Paste} Copy in excel

- Click on column heading → entire column selected →
copy → Paste in excel.

iv) Split | custom Split :-

- To split the column data based on the separator they have in b/w them.

i) Split :- Click on column drop down → Go to split
→ This take Space as automatic splitter (delimiter)
if separates two in different column & keep the original column as well.

ii) Custom Split :- To split on basis of a Delimiter / spaces b/w two texts in a column (eg: , - / . ' etc)

Right click on drop down → custom split → enter separator → split off First | Last | All → ok
↓ ↓ ↓
split left text split right text split all in separate columns.

SORT :-

- To Sort Columns | Data | Table :-

- In DATA SOURCE TAB → when we drag data → one option wj there Sort fields → To sort table options
 - i) Data Source Order → Sort as per the source of data wj there.
 - ii) A to Z Ascending → columns get sort form A to Z
 - iii) A to Z Descending → sort column from Z to A
 - iv) A to Z Ascending as per Table → As per table means when two or more table are join together then data of all table are displayed. Based on Name of table sorting wj done.
 - v) A to Z Descending as per table → Table wj sorted as per descending order.

SORTING FIELD DATA :-

- Sorting field has a small icon in name of a field itself → These can sort as per Asc to Desc

icon



* DRILL DOWN & HIERARCHIES :-

- Suppose we are creating a Hierarchy as City, State, country, zipcode.

Country → State → City → Zipcode

- Here a Hierarchy is created from County to Zipcode

- This is also called Drill Down i.e. Going in depth or at the ~~lower~~^{higher} Granularity

Drill Down $\propto \frac{1}{\text{Granularity}}$

Granularity $\propto \frac{1}{\text{Aggregation}}$

Year → Months → Weeks → Days

Category → Subcategory → Product Name

Date Field
Hierarchy are
auto created.

Type & ex.
of
Hierarchy
Drilling Down

* How to Make Hierarchy :-

→ To make this

- 1) Drag a subcategory in dimension field on the Category field → It will ask you to name the Hierarchy created → we can add then product add & drag & drop again to category (or Hierarchy already created)

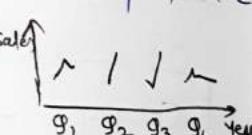
- The placing of drag & drop decides the position & rank of fields so drag & place accordingly

- Now when we drag & drop that Hierarchy field in a Column / Row to create a chart → it gives a \oplus sign so that we can expand → i.e. it gets drilled down & drilled up depending upon our application.

* DATE Hierarchy :-

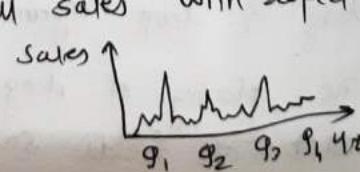
- When we drag & drop in Row / column
- There is auto \oplus sign generated i.e. automatic hierarchy is generated from Year to Day i.e. Drilled down.
- By expanding we can go from Year to Day wise sales

① Discrete Type :- BLUE COLOR → Year & sales column

- When we drag & drop in fields → we can see the line chart in discrete manner by default → i.e. there will different sales report for that particular individual year, quarter, month or week. 

② Continuous Type :- To get continuous type

- click on drop down of Year on Year tab → select the second year i.e. second year in list → This is the continuous type

- This gives us the ~~chart~~ in continuous manner i.e. the chart we get here is on a single display → & gives us the understanding on overall sales with respect to all the other years & days. 

To Remove Hierarchy :-

- Select Hierarchy → Go to Drop down → Remove Hierarchy. (✓)

SORTING IN WORKSHEET :-

- There are different types.

- 1) Quick Sort 1 :- Available on the chart titles itself
→ click on chart title → there is icon → available to sort & default sort → keep of clicking on icon.
- 2) Quick Sort 2 :- Swap the fields i.e. Rows → columns & vice versa.
Present on top tool bar →

3) Sort by Tool Bar :-

- First click on measures i.e. Sales in Row/column →
Then on Tool bar → Then on top tool bar click on
sort by subcategory icon

- This will sort according to category

4) Sorting from Pill :-

- Click on category → drop down → sort → can sort
base on Alphabetical order → Asc / Desc. → on category
- We can also do advanced sorting → dropdown → go to sort
→ sort by → field → Here we can sort the chart
by another field (Profit / Quantity, etc)
- Just we have drag & drop all details of suppose profit in
Labels to see value at top of bar.

NOTE :- If we want to know if the particular field is sorted or not, into the front of the particular field in row/column there will be a logo

(category → If this, then sorting is applied.)

5) Sorting by MARKS card / tab :-

- It is card available in sheet to apply color / size / label / detail / Tool tip.
- If we drag & drop any field in marks card.
- Select that field → Go to top tool bar where we can sort the particular field → & change the chart.
- This is a alternative to dropdown sort of a field & then choosing a different field to sort by.

Example :- If we draw the chart with Subcategory v/s Sum of Sales

- But sort above chart base on Quantity Sold.
- Then drag Quantity sold & put in Mark Card → Labels → click on dropped field → & then click on sort tab top toll bar

GROUPING :-

- Used to Group one or more data in one package.
 - Suppose you have created a chart in a sheet → sheet has many bars → if we want to combine in a group any two bars → select the bars with **ctrl** on → right click → group → they will grouped by values also & will be displayed as one bar.
↳ Visually grouping
 - Group** - The grouped total data will be saved as a dimension in your Data tab in a sheet at left
 - This is used for comparing the group of data with the rest of data to find some conclusions.
 - Group** - If we want to expand the other → remaining data other than grouped → they can be combined or can be expanded for respective individual value also → Go to dropdown of field → self select → exclude / include other.
 - The saved grouped dimension can be used for further or can be deleted if we want.
- Alternate :- Group can be done by drop down list from the field name in data tab also. → window grouping.
- The Grouped Field saved in Data tab can also be used in Hierarchy also → by dragging it into a hierarchy
↓
Create
↓
Group-

* MEASURE NAMES & VALUES :-

- When we drag & drop connect the tableau with our data & we set, there are some autogenerated fields created in tableau.
 - These fields are mentioned by Italic font.
- i) Number of Records (#) → Measure
- It gives count of records present in the data.
 - This field can be used for the making report & chart.

To Get Grand Total → Go to Analysis → Totals → Get column Grand Totals.

2) Measure Names ^{values} (#) → These are auto generated
Aggregate functions (sum, Avg, Count, etc.)

NOTE :- If we drag & drop any field in ~~the~~ the
Marks card → Text → This gives us table format
Report in a sheet.

- And one another field used to get data should be in either Row/ column
- • Row or Column both should not be occupied
~~or if so it gives graph (chart)~~

3) Measure Names (Abc) (Dimension) :-

This is the name of measure which has to be counted as a value.

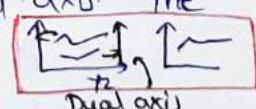
e.g. Sales, Discount, Profit, etc → These are names
↓ ↓ ↓ of measure
sum(sales) sum(discount) sum(), etc → These are
Avg(sales) Avg() measure values.

Note :- Just by double clicking on Measure Name also we get the result with values in sheet.

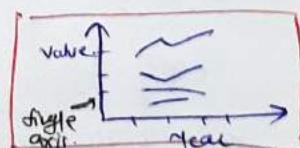
Note :- If we double click on Measure value it gives us graph chart directly.

* Measure Name Usage :-

- When we draw a report / chart for more than one quantity ~~to~~ with respect to a field, we can bring any two quantity on a single chart with dual axis. The other chart will be individual.



- To bring more than one & more than two chart in one single chart then we take help of measure name & measure value.



- Drag & drop measure ~~value~~ ^{name} say in filter (if needed)
→ To filter out which we need to display → drag & drop measure value in column / row → we get chart (disturbed)
→ To make it proper → drag & drop measure value in colour
→ in mark cards → To label (apply legends) again drag & drop measure value in label in mark card → we get multiple chart in single with single axis.

* Discrete v/s Continuous :-

- When we add data source we get Dimension & Measure
- The Dim/Measure whose datatype (#, Atx, etc) are in Blue colour → Discrete fields (Distinguished)
(Green)/Grey colour → continuous fields (Aggregated)
- Furniture, Table, etc. \uparrow sum, max, Avg, etc.

Blue → Discrete
Green → Continuous

- When we have a chart

- i) When we drag & drop Discrete field (eg. Category, Subcat) in color mark card → The graph will be coloured according to the distinguished sub fields.
- ii) similarly if we drag & drop continuous field (mostly measure) it gives a pattern colour (single colour) applied in low to dark shade (ie. min to max value).

↑
Some of applied for Map Also Re. Map chart

- For Map

- i) Dimension → Discrete → If drag & drop dim. in colour marks → we get dim (distinguished) colour
- ii) Measures → continuous → Drop in colour mark → we get Heat map / Filled chart

For Filter

- i) If Drag → Drop → Dimension (Discrete) Field → filter
 - It allows us to filter with what we want
- ii) Drag → Drop → Measures (continuous) field → filter
 - It allows us option → To get (sum, count, Avg etc)
 - Also advanced filter options.

* PARAMETERS :- + SETS

- Example Suppose you a chart of subcat v/s sales having 50 subCategories.

- We want top n sub cat → Drag + Drop sub cat in filter → Go to dropdown → By field → Select Top n we want.

- To get different n → we have repeat this always
- To avoid this we use Parameters.

Parameters with filter :-

- To get user defined top N → Again drag drop in subcat in filter → Go to By field → dropdown Top 10 → Create → Select parameter → Here we can apply → Select current value → Applying range.

Parameters with sets :-

- Parameters with set is used to get top N rows highlighted with others also in chart.
- Sets are created to particular field with parameters.

Select the Parameter field for which we have created parameter & now want to create set.

- Go to field (In Filter / Dimension) → drop down → create set → Top → By field → Select parameter name we already created → ok.
- Inf — Drag the set → put in color in mark card.
- Make sure there is no field in filter. → or may be there if required.

COMBINE :-

- Used to combine two or more dimensions together.
- When we want to combine two fields together (e.g. Categ. & Sub Categ.) to determine maxm value for say suppose sale for a Categ. in combination with subcategory we need to use combine.
- STEPS
 - Select the fields we want combine by using ctrl → Go to drop down → create → combine field & then sort to get lowest & highest.
 - We can combine as many fields to get values & map from well sorted values (Asc / Desc).

SETS :-

- These are custom fields (dim or measures), subset of fields made by based on some conditions.
- This is mostly used to find top values, or any group of data to be highlighted from field.

Two ways to create:

- 1) Visual selection → If a graph → bar/map/scatter, etc. are created → from those we want to select some & give different colors (or any) → then select them by control or by selecting by mouse → we get tool tip → create set & it will created → drag that set name & drop in color mark → it will get different colour.
- 2) From field → Go to field → drop down → select create → set → different opn (General / condn / Top) → Select say top → By field → put top we want → or which

TITLE & CAPTION :-

- Title of Graph is taken as sheet name by default.
- To Rename → double click. → edit.

We can give sheet name, ~~or~~ Data source name, etc. → from Insert in title diag. box.

Caption :- By default at bottom → Go to worksheet → show caption
 ↑ - we can insert if we want to display any info here.

To format → Go to title → click → drop down → ^{or} right click → format.

EXPORTING :-

- Export file as an image, pdf, or to excel, ppt etc we can do.
- If we want to export data from which chart is prepared
 - Go to worksheet → export
 - Get Data to Excel
 - Get Image also
 - Save as → pdf

GRANULARITY

- Breaking down into small pieces.
- To go in more details of any value (sales, quantity etc) to get more granularity → drilling down.
- Granularity is achieved by dragging & dropping different fields in to Mark Card:

WORKSHEET INTERFACE

Overall GUI of Tableau.

* MANAGING METADATA :-

- How to create hierach
 - How to rename field
 - Hide / unhide field (dim/measures)
 - How to insert calculation field
- We can create a folder in which we can add fields
- Right click in data tab → create group →
right click again → create folder.
- How to create calculated field :-
- Right click in Data tab → calculated field →
Here we add some form to measures | dim fields &
create new custom calculation field.
 - Here in calculation field → there are different methods
(sum, avg, datiff, round, trim, etc.) where we can set
different function & create a calculated field from
other fields

Example :-

sum ([profit]) + sum ([sales])

round ([profit], 1) ⇒ rounds upto 1
decimal value.

CROSS DATABASE JOIN

- If we want to combine the files from different sources / connection / DB & use them to prepare reports, it is possible
- Example :- Excel + DB SQL Oracle + notepad file, etc.
(combination of files)

For Notepad (csv file) → Select Text as connection
in Tableau

DATA BLENDING

- It is a method of combining data from multiple data sources into a single view.
- Data Blending Require at least one common field between the data sources -
- If field names which are common ~~are~~ have different names, the relationship can be done manually & if have same name, auto generated relationship occurs.

Data menu → Relationships → pick Data source
→ we can change it by dragging down (which to keep primary)
→ select automatic / custom
↑ ↑
Same fields Different field names

- Now when we come to prepare our view
- Select one data source → when view is created first with that ~~do~~ data source → It becomes primary & blue tick is seen on it.
- Now when we drag & drop the some fields from other data sources → the other will be treated secondary & Orange tick is seen.

* FROM 360 Digit Mgt *

- Data blending is method for combining data from multiple sources.
- It occurs on sheet by sheet basis. Data blending brings in additional information from secondary source & displays in single view with data from prim. source.
- Let's say there are two data sources A & B

Data Source A
Daily Transaction

Data Source B
Monthly Target

- As we see → Daily Transaction sales of a superstore of B gives the monthly target file. The source A is at much ~~high~~ layer level of detail (higher granularity) than B. When we perform joining on these it causes wrong results & also decrease the performance of tableau.

- Now at this case Blending is soln, it is a advance version on Left Join. What it does is it will perform a separate query (view) to database A (datasource) & separate query of datasource B, just the view will be in a single window or combined or together based on Link or relation b/w a particular common field b/w both the datasources.
- Firstly we call / connect the two data source with a Tableau.
- Now when we use any one of data source initially to make a view, becomes a primary data source with blue tick on a barrel seen on left at its name.
- Then when we switch of other data source Tableau automatically understand that we are performing blending & tries to find the link b/w two sources based on common field present b/w them.
- There are two ways to make link or relation b/w sources
 - i) Auto link generated if names of field are common, if field name are not same then try to rename & make same name to other field, it creates auto link.
 - ii) Manually Go to data menu → Blend Relationship

→ Here we have to choose prim. data source (or prim. is chosen automatic if field of source are already in view) then add the relation manually to fields.

NOTE

→ The fields with which we have blended data source or made links must be present in a view (row/column shelf) compulsorily. otherwise blending provides incorrect result.

Blending / link fields have a symbol

⊗ → This indicates fields are linked & are used in a view

⊖ → Linked but not yet used in a view.

2) If we want to establish relationship on Date Field, if Date in both sources are at different granularities / Level of detail then they must be established very carefully.

i.e. ✓ Year ↔ Year

✓ Month ↔ month

✓ Quarter ↔ Quarter

⊗ Year ↔ month → This is wrong blending

Data Blending View Result :-

- Data blending gives us the view based on Left Join i.e.

All the data from primary data source & matching data from the secondary data source.

- Those are not matched in secondary shows up as Null.

- If our req. is to view all data from sec. data source & matched data from prim. source → In this case after prim. source → i.e. first drag & drop the fields of Secondary which were earlier will now treated as primary & primary earlier will be treated as secondary if result will (view) be altered as accordingly.

- If we want to emulate the Inner Join Scenario by blending is possible, just click on null values → It will ask what we want to do → click on Alter Data → It will give us a view based on matching data from both sources only i.e. case of Inner JOIN.
- We can also emulate the Left Join but cannot view or emulate outer join scenario.

Important Points

- 1) Unlike joins, data blending keeps data sources separate & simply displays their info. together. This is ideal when data is at different levels of granularity.
- 2) The order in which fields are used determines which data source is primary vs secondary data source.
- 3) Prim. data source in view is denoted by blue check mark, any second. data source & fields from secondary data source have orange check mark.
- 4) Linking fields can be automatically determined based on shared field names or can be manually created.
- 5) Data blending behaves similarly to left Join, which may result in missing data from second. data source.
- 6) Data Sources can be blend on a per sheet basis i.e. on every sheet the prim. & sec. source can be changed as per requirement.
- 7) We can have multiple secondary data sources but only one prim. data source.

* WAYS TO COMBINE DATA :-

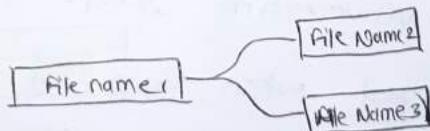
1) JOIN :- We have two tables → Remerge by join to get big table combining all fields of both table (source).

- One thing while joining is very imp. that we should know which table to use as left & which to use as right.
- Because based on type of join the result varies.
- Here Join is happened at datasource level itself.

2) UNION :- This is used when the two or more data source have same fields & same data types.

- Unlike joins it doesn't merge the field & forms the big table with all fields.
- It will add all data with similar column headings & avoid the duplicates.
- Union works only when data come from same connections i.e. (Excel → both sources, CSV → both sources)
- Can't union if one source is Excel & other as ms access or Text file.

3) Relationship (Noodles) :-



- Newly introduced in Tableau.
- Just drag & drop Tableau automatically find common field & best way to join or make link b/w the tables.
- Once link is created it doesn't merge like join immediately but holds that info into datasource & leave them separated.
- They are joined on the visualization level or dashboard level.

Types of Relationship :-

- i) Many to many relationship
- ii) Many to one — n —
- iii) One to many — n —
- iv) One to one — n —

- Tableau builders recommend to use relationship as your first choice / approach to combine your data because it makes your data prep & analysis easier & more intuitive (असरकारी). Use joins when we absolutely need it.

only Joins were there in Tableau version before
pre - 2020.2 version

4) Blending :-

- Used mostly when the two data sources or more than two sources are from ~~for~~ different connections.
(Ex: Excel, Text & Access)
- The data source are joined at again ~~visualisation~~ visualization level.
- Used when data is at different level of detail & at different granularity.
- The secondary data source have the aggregated values for the measure at viz level of detail.

+ DATA BLENDING CALCULATIONS :-

- Suppose by blending & using two data sources we have created a view based on link field as 'state' & profit of both sources.
- Now if we want to find out the total ~~Profit~~ i.e. sum of profit of both & create view.
for this we have to create one calculation field & we calculate as
- To create calculated field use any one of data source right click → create calculated field.

NOTE :- If one of view

- If our view while summing profit, if one of the data source view or both have null values then it give wrong result by just sum.

$$\boxed{\text{Null} + \text{Value} = \text{Null}}$$

$$\boxed{\text{Null} + \text{Null} = \text{Null}}$$

- To convert null value to zero '0' use.

$$\text{ZN}(\text{sum}([\text{Profit}])) + \text{ZN}(\text{sum}([\text{datasource name}].[\text{Profit}]))$$

$\text{ZN} \rightarrow \text{convert null to zero}$

- If any of data source don't have null value in it then don't use ZN.

FILTERS-SHELF :-

- When we create a view & then we want to see the view based on some two or three parameters only then we have to use filter shelf.
- For Ex:- We want to see the sales of different product region wise. There are different region like East, West, North, South.
- If we want to see for North, South only we have to drag & drop region into filter shelf to do that.
- There are different option when we drag any field of dimension or measure. let see -

① General :-

- i) Selection from list
 - List out what we want to see , by check boxes-
- ii) Custom value list
 - Create custom list by typing

② Wildcards :-

- It filter with a pattern
- like start with 'a' , 'A' , etc.

Condition: - ⚫ First drag of drop dim field in filter.

- This allows the filtering of data by dimension field based on field like quantity, sum, avg of measure field.

- We can select By field → select Aggregation type → also put computation (<, >, =) or some range of values (min & max) or By applying some formula also.

To get range of values (min & max), to see for any aggregation like (sum, avg, count, etc) we can select field → click on Load → it gives min to max range for our reference & then we can enter our condition value.

④ Top :-

- To get top N items based on any measure field like sum(sales), sum(quantity), sum(profit)
avg(sales), count(quantity), etc.

- Depending upon our appn.

DATE FILTERS

- When we drag & drop date into filter we have option to filter date.
- Select Range of date → There is previous | Next | current year.
- We can put custom year also.
- There is also Starting date.

Interactive filters

- Interactive filters are those which are mostly used for dashboard level.
- We can create filter from field dropdown box for any field we req. & it changes the view of a chart.
- We can format the filter based on req by drop down on a filter for getting different layouts.

Advance Interactive filters

- Say we want to create filter for Categ & Sub-Categ & relate them with each other -
 - i.e. they must be dependent to each other.
 - To get this in filter at right of sub-cat which we created → go to drop down → select only relevant values.
 - We get the filtered ~~field~~ & data for particular category selected only.

- similarly for Date Filter
- for date also we can create filters in same way.
- methods sorted column \rightarrow users \rightarrow filtering and filtering

WHERE Tableau Filters :-

① Summary level Filter :-

- say suppose we have created view of Profit vs sub-cat & there some sum(profit) -ve & +ve both.
- To filter the sum(profit) > 0 or < 0 , $= 0$ then we use this filter.
- In this we will drag & drop measure field i.e. profit in the filters shelf then
- select sum \rightarrow At least (मिनीममी) \rightarrow At most (माइक्सम)
- we select at least & put zero (0).

② Record Level Filter :-

- In this we will select All values instead of sum(profit)
- i.e. API values \rightarrow get values of particular data & not aggregated value.
- i.e. This allows that if we select At least value \leq this will do sum of only +ve profit & display in view
- To sum, Avg \rightarrow depend on what view we made of i.e. Subcat vs sum/avg (sales)
- In turn on that basis of view \rightarrow All values in filter \rightarrow it adds the +ve profit.

- This filter is very useful to get sum of only +ve profit or only -ve profit depending on what pills are present & view & make better decisions.

DATA SOURCE Filters

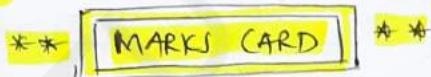
- This means whatever data we are connecting to data source → filtering those fields we use this. & the data inside those fields can also be filtered.
- Data source filter → How to apply → right click on data source (name of file) → edit filter → add here the field which we want to filter → In that field select data what we need to filter.
- Now as this is ~~at~~ filtering at data source level this will filter all the worksheet in tableau where it is connected.
- Not recommended / or used to show some part of data source only.
- The filtered part is also not shown in filters shelf.

Filter From Legends :-

- This is another method to filter by selecting the legends of different colour we want
- Just click on a legend we want → tool tip displays → press keep only → we can see filter is applied → The filter applied is seen in filter shelf
- To remove → drag filter shelf pill & throw out.

FORMAT FILTERS

- When created filter → Go to drop down → select format → Here we can format filters.
- But mostly the formatting done is seen when we create a dashboard



1. SIZE

- Used to change the size / width of graph which is in the view.
- If we drag & drop the measures field in size card that will get max size of part bar / line for max value & min for min value.

2. COLOR

- Used to make graph more meaningful & more colourful
- Different fields of dim / measure can drag & drop in color to sort / filter by colour.

3. Label

- It is also called Text
- Text is also present at top tool bar [T]
- This label / Text is used to label / display value of that particular bar / line point.
- If we drag any dimension like Region (country) we get labelled to with that field.
- Say suppose we have mentioned label of sum (sales), sum (profit), sum (quantity) as a label in view.
- But it displays only value of above aggregate fields.
How can we recognize which is what.
- To given names click on Label → Label appearance → Text → click on → From here we apply name to distinguish

Ex Total Sale : <(sum sale)>

↑
Add this

↑
This already present there,
don't disturb it

- We can also format label text here → Bold → colour → font, etc.

MIN/max

- We can also label for min & max value → create
Follow same proced. above → marks to label →
Select min/max

SELECTED

- There is one more option there as selected → selected means when we click to particular part of view in pan that will be applied as label

Highlighted

- When we select by drag multiple parts in view pan that will be displayed with label.

4. TOOLTIP :-

- Tooltip is when we select or set cursor on part of a view that gives tooltip with different info like

EX :-

category : furniture
sales : \$ 200k

- To add extra tooltip with same level of view we can without changing, we can do, just drag & drop the measures or dim field in Tool tip & it will give additional info to the graph / chart that is not in pan / view.

- When we click on tool tip → we get editor → to format the tool tip → color, font, etc.

→ To add extra item like datasource name, date, time

→ click on insert in the drag box displayed.

5. SHAPES :-

- Shape is used to give a different shape to a chart type
- This is mostly for a Scatter Chart
- When scatter chart is created shape card will be appeared in marks card.
- Then for different fields we can drag & drop them in the shape card we can see diff shape to distinguish them in scatter chart like (\square , $*$, o , Δ , \diamond , etc).

Note - We can Add more custom shapes by downloading the file, extracting it \rightarrow copying all \rightarrow folder \rightarrow paste in \rightarrow Tableau Repository folder in document \rightarrow shapes \rightarrow Paste. \rightarrow The shapes will be in Tableau.

6. DETAIL

- Detail will be same as color only difference that we will get different section in a view when we drag & drop any field wrt that field data we get detailed view
- cause we will not get as different colors.

FORMATTING

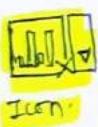
→ Go to format menu bar → format
or click on pill dropdown → format.

1) FONTS

- To change the font in the view.  ← Icon.
different option when we click on sheet → entire sheet ~~of~~ font can changed.
↳ Worksheet, pane, Header, Tooltip, Title can be changed.
- * - We can also format base on fields in view
- * eg. Click on field dropdown → select field → format we want to change.

NOTE

In drop down near clear sheet there is a option called clear worksheet formatting by which we can undo all formatting applied to worksheet



- [OR] - If we right click on any part in view & select ~~format~~
format, we can format the sheet of that selected part as well.

2) ALIGNMENT :-

- To change alignment of text / Label / Tooltip, etc.

Icon → 

Pane → changes alignment of numbers present in view other than header.

Header → To change alignment of Headers.

↳ Row headers, column Headers → different options are there, we have to select that

3) SHADING :-

- Shading is used to give background colour to the sheet / row / column

i) sheet → entire sheet will get background colour including rows & columns.

→ Pane → pane is nothing but wherever the numeric data is present | shading to it will be done.

→ Row banding → This mean alternate color to rows.

→ Headers → Row & column Headers or titles.

ii) column → To give colour to column data or column header is possible.

iii) Rows → similarly for rows.

4) BORDERS :-

Icon →

To provide borders to different part of view

Sheet | Row | column

5)

Formatting

- For any formatting there are three options.

SHEET	ROWS	COLUMNS
<u>Default :-</u> • <u>Worksheet :-</u> This formats the entire worksheet.	<u>Default :-</u> • All formatting done for only <u>Rows</u> & not total worksheet.	Similar as Row it will take for column.
• <u>Pane :-</u> formats only the data / numeric data in the view other than column, Row & label headings.	<u>Example</u> <u>Shading :-</u>	
• <u>Header :-</u> formats header / Title & Label of Rows & column - If chart then formats the vertical & Horiz axis labels.	<u>Default :-</u> <u>Heading :-</u> Apply shading to only Header or title or vertical axis of view.	
<u>TOTAL :-</u> For Subtotals	<u>Pane :-</u> Apply to values of data other than heading - row wise only.	
• <u>Pane :-</u> To format only values of Grand Total Row & Column wise combinedly.	<u>Grand Total :-</u> Unlike for sheet only Total of Row wise is formatted.	
• <u>Header :-</u> formats on headers of Grand total in Row as well as column.		
<u>Grand Total :-</u> Grand Total.		
• <u>Pane :-</u> Format Grand total values only (All)		
• <u>Header :-</u> Formats Row & column Header of Grand Total		
<u>Row Divide :-</u> For Borders		
• <u>Pane :-</u> Apply a border divider to rows having values.		
• <u>Header :-</u> Format borders only to headers in rows.		
- Same for Column Divider		

NOTE :- Going for drop down in fields in formatting, the separated formatting for Pkls in view can be done.

CALCULATIONS :-

- Actually this is used to create a calculation field to create new calculation which will be saved as a separate field in data source & can be dragged in view.
- It can be created as → right click in data source → create → calculation field.
- There are various field we can make for calculations like difference, addn, float, ceiling, floor, round, etc.

The created field be shown with symbol in datasource as

=# → If it is a measure field created.

=Abc → Dimension field created.

Number functions :-

- To format the number there in any field.
- To do this Go to dropdown in apill → Go to format → pane → dropdown to Number → There are different option to format.

STRING FUNCTIONS

- function like

- Upper
- Lower
- Trim
- Left
- Right
- Patterns.



This is also done by making calculation field in Data source.

* Split

- we to split the text into different strings

- select field we want split in data source → ~~transform~~ → click on drop down → transform → custom split → use the separator (-, space, ., comma, etc) → separate Field will be created for this field.

- instead of custom split if we select split it splits all & determines automatically the separator.

TABLE CALCULATIONS :-

Two types of calculation are there

1) DataBase Related

2) Table Related.

↳ Based on the output we are getting in the view, & when we do calculation on those, they are called Table Related calculation.

Example:- Say Suppose we need percentage of Sales by category.

- Create a view as Category v/s sum of Sales.

To get % of Total Sales → Go to drop down to sum (Sales) pill → Quick Table Calculation → Select Percent of Total → make Label on to display

- Quick Table Calculation are Inbuilt Table Calculations in Tableau.

Table Related Calculation made on Pill will shown by a inserted triangle  Δ

Ex

sum(sales) Δ

← Pill

↑ This indicated that Pill has a Table related calculation.

Right click anywhere ~~out of pane~~ & Select Summary → we get summary of all in the view → like count, sum, avg, min, max, etc.

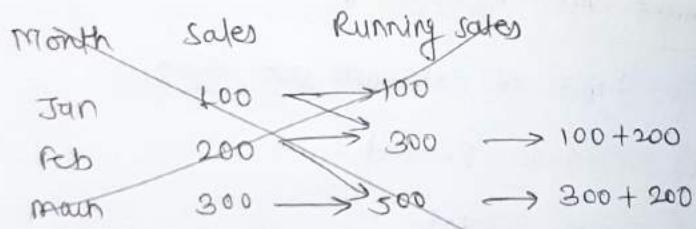
Jump

Running Total :-

Tricks

This gives us the cumulative total of all the values of that field.

Ex



Month	Sales	Running sales
Jan	100	100
Feb	200	300 → 100 + 200
Mar	300	500 → 100 + 200 + 300

- This is also available in dropdown of a pill → Table calculation.

COMPUTE USING

- It is advanced option related to Table Calculation.
- This option is also available when we drop down the measure field in the view.
- But this is imp that it will be only visible when we perform first some table calculation e.g. percent by total then compute using will be visible for that field.

[OR]

- Compute by using this can be done by ~~down down~~ drop to fill → Add table calculation → then perform own calculation

Alternative method

Different calculation are

for Percent by Total → Compute Using]

- i) Table down → It create a difference in sales for previous value. (Creates percentage based on Grand Total)
- ii) Pane (~~down~~) → Also give same as Table down for single view.
- iii) Pane (down) → This create the percentage based on subtotal for sub category.

More understanding of Compute

- i) Table Across → measures of apply % based on row wise values i.e. Row wise total & % based on that.

Example :-

	12.8%	20.6%	13.6%	Grand 100%
				Add of all.

2) Table Down :-

This gives the value of σ_0 based on total of from top to bottom column wise.

Example :-

12.6 - 1.

13.6 - 1.

18.7 - 1.

:

:

Grand Total = 100 +.



Ignores the subtotal values in between & gives total based on Grand total down column.

3) Table :-

This compute based on total table i.e. total will calculated from Across & Down Total.

Example:-

13.6 - 1. - - - - 18.1.

14.6 - 1. 13.79 + 16.1.

10.1 - 1. 19.1. 46.1.

----- 100.00 -----

Total



Total will of all values in table.

4) Pane Down :-

This takes / computes σ_0 based on subtotal in the view or data in view. if not on grand total like table down.

5) Pane → Same like Pane down.

[OR] Also we can mention based on σ_0 fields present in the view like % of Category wise sales or % of Subcategory wise Sales.

CONDITIONAL FORMATTING :-

Index function → $\boxed{\text{Index ()}}$

- This is used to give numbering to the data in view.
- This is created by using calculated fields + type Index ()

Top N :- To find Top 10 suppose → create field as

$\boxed{\text{Index ()} \geq 10}$ → Drag & drop in color mask

Last function → $\boxed{\text{Last ()}}$

- To find the number data from lower to maxⁿ value in data.

RANKS

- It is used to give ranks to any number related measure.
- For these also we have to create a calculated field. by right click on data source
- In the calculated field dialog box → there are diff. function available → In drop down → select Table Calculation → Here there are different Rank functions.

Ex ~~Rank~~ $\text{Rank}(\text{sum [Sales]})$

- When we drag & drop that into the view → it gives rank from max to min of sales.
- If there are repetitive no. then same rank is provided to it + the no. next is skipped → e.g.

Sales	Rank
100	1
150	2
150	2
200	4

Rank-Dense ($\text{sum}([\text{sales}])$) → This gives all no. without skipping any rank.

eg.

Sales	Rank
50	4
60	3
70	2
70	2
80	1

Rank-Modified ($\text{sum}([\text{sales}])$)

- When there is same no. it give rank with highest no. & skip lowest

Skipped → 2 (lowest)

Sales	Rank
80	1
70	3
70	3
60	4

Rank-Unique ($\text{sum}([\text{sales}])$)

- If there are same values it gives the rank based on Avg. value (a or A)
a → lower rank
A → higher rank

Prod.	Sales	Rank
a	10	5
b	20	4
C	30	2
c	30	3
d	40	1

Logical Functions :-

If Else Condition

Example :- If sum ([Profit]) > 15000 Then
 'High Profit'
 Else

'Low Profit' End ← we must write end
 at last

↑
 This can be done in calculation field & then
 drop that field in marks card for result
 color wise | size wise | etc.

Example :- If sum ([Profit]) ≥ 1000 Then 'Good'
 Else If sum ([Profit]) = 1000 then 'OK'
 Else 'Poor'
 End

CROSSTAB

- Crosstab in tableau is called a Text table, which shows data in textual form. The chart is made of one or more dimension field & one or more measure field
- It is created by placing one dimension field (e.g. Category) in row shelf & another dimension (e.g. Region) in column shelf & then placing measure field (e.g. sum(sales)) in marks card at Label / Text
- We can analyze data by % of Total or running total by using compute.

INTRODUCTION TO CHARTS

8-

- We have ~~24~~ different type of chart available in the Tableau
- These are present in **Show Me** at right top corner in a Tableau worksheet.
- When we open Show me → and take cursor to any chart → below it shows that min requirement for that chart to get view → i.e. how many dimension field & measures are needed.
- If we select a dim. field & second ~~field~~ any Field by pressing control & open show me ~~tab~~ tab tableau gives us some suggestion that what chart can be prepared as per your selection.

1. SCATTER CHARTS

Requirement :- 2 to 4 measures → Row & column shelf
 0 to more dimensions ← In marks card.

Example :-

- Drag sum(sales) in Row shelf
- sum(profit) in Column shelf
- Category in colour mark card.

} creates scatter chart

Dropline :- The create drop line for selected point scatter chart
 click on an shape in chart → right click → dropline → show dropline

Drop line allows to see x & y axis values.



we can also edit drop line.

WORD MAPS :-

Word Map

- This is a map based only on texts & their size of text.
- Requirement → 1 Dimension → Drag & Drop in text card.
1 or 2 measures → Drag & Drop in color/size.
- By default when we do this this gives us a tree map, but got marks card → Go to automatic map dropdown → Select Text 

LINE CHARTS :-

Two Types of line chart

- 1) Discrete chart ← individually separate & distinct (Blue Pills)
- 2) Continuous chart ← forming an unbroken whole, without interruption. (Green Pills)

Requirement :-

- 1 Date Field → mostly line chart are used for date fields.
- OR
- 0 or more dimensions
- + 1 or more measures.

- Drag date field in column & measure field in rows shelf → Date field has auto generated hierarchy

Year → Quarter → Month → Day

- + Drop down any dimension in marks card (color/detail)

REFERENCE LINES :-

- Use to compare the values for a particular value like Average Sales.
- Sales done below Avg of Sales done above avg.

- How to get this line.

- Right click on the axis of a measure eg. (sum(sales))
- add reference line → select scope → entire table
- pan / cell → select what we want → ok

OR

Go to left in Data → New Data there wj

Analytics tab → custom → Reference line.

* Reference Band :-

- This create a Band from f to value ie. band created 'from value' to 'To value'

* Reference Distribution :-

- We get computation of value bases on percentiles, percentages, etc. i.e. it create a band based on f. of say sum (sales) or avg (sales).

BUBBLE CHARTS :-

Requirement :- 1 or more dimensions

1 or ~~more~~ 2 measures.

- size of bubble represent measure field & colour

represent dimensions

- For Bubble chart eg. sum(sales) → put in size and sub-categ → put in label.

sub-categ → put in colors.

- All fields are put in mark card only for bubble chart.

BAR CHART & STACKED BAR CHART

8-

Requirement for Bar chart :-

- 0 or more dimension
- 1 or more measures.

Stacked Bar :-

1 or more Dimension

1 or more measures.

Side by Side Bar :-

- 1 or more Dim
- 1 or more measures

- Require compulsary 3 fields for side by side Bar.

TREE MAPS :-

Requirement :-

- 1 or more Dim
- 1 or 2 measures

- measure defines the sizes of boxes.
- In this all the fields are added to make could only select as in automatic chart type.
- OR select free map from the show me tab.

BUMP CHART

:-

Customchart

- It is a simple line chart where the Rank changes according to time (Date).
- It let us know the trend of an measure whether it changes its value in a particular year from low to high or vice versa.

How to get ?

- Drag & drop Year (Date) in column shelf & sum(sales) / sum(profit) etc. in the rows shelf , put catg / region / segment in colour card.
- Go sum(sales) dropdown → quick table calcn → Rank → Go again → sum(sales) pill dropdown → Compute using → select catg / region which in card we draged earlier there will be option in compute using.

FUNNEL CHART

Custom chart → root in show me tab

- Used to represent sales/profit/quant at different stages/sector/segments, etc. etc.
- In funnel Topmost is called as Head & Lowermost is called as Neck part.
- It is just a bar chart type which is expanded & size is added to card with a measure, & select view type as a entire view.

Advanced funnel chart :-

- Funnel is also a custom chart to make as not available in show me chart.
- To get from this funnel chart follow steps
 - i) Drag & drop sales or profit or quant in the rows shelf of one dimension like segment or catg. or sub-catg in column shelf.
 - ii) we will get a bar chart → sort it in descending order → Go to marks card → select the chart type from automatic to area → half of funnel is created.
 - iii) Second half will be created by using -ve sales or measure field that we have in edit view → right click in data → create calculation field for example - [sales].
 - iv) Drag that created field in column shelf before the first sales pill & we will get funnel full created.
 - v) To label them in marks card two pill card will be created, select any one & put dimension pill we have in row shelf to label & for other pill card put sales in label & we will get funnel created.
- Funnel will be always from maxm value to minm value sorted properly.

WATERFALL CHART :-

→ Also custom chart.

- This is used if we want to determine, say how our company profit ^{or sales, etc.} goes increasing ^{or decreasing} from one part to another in form of waterfall chart.

- Waterfall chart is created from Gant chart

- How to create waterfall chart?

- i) Drag & drop say, profit in rows shelf & subcat. in column shelf
- ii) We get bar chart → sort it by ↑, ↓ → form marks card select Gant chart from automatic → then drag & drop again profit to date
- iii) We get a waterfall chart → to give color or label again drag & drop profit in colour or label card.

PIE CHART :-

Requirement → 1 or more dimension

1 or 2 measures

MAP CHART :-

Requirement → 1 Geographic dim

0 or more dimension

0 or 1 measures.

- * - Some time we could not get the match of the geographic data in map → for this → right click on map → Edit locations → select here county we want or keep as select from fields.

Map Options :-

- Right click on map → view Toolbar → This allows maximising, search, fit window for the map.

Map Layers :-

- These are used to custom different option for the map.
- To change style color, add boundaries, satellite view, different data layers, etc.
- To get this → Right click on map → Maplayer
OR Go to map toolbar at top → Maplayer

Custom Territories

- This nothing but creating a group / set of states & analysing it.
- Territories may be state or city.
- For this we will make groups by selecting state by dragging or by $ctrl +$ make group (\oplus) → In this way we make no. of group.
- These all groups created will be save in a field newly created in data.
- Now clear sheet → bring that field & drop to make map → If not created → Go to automatic chart in maps card & ~~select~~ select map → then we can see different sales / profit done groupwise.
- For field created newly → go to drop down → edit → we can give name to groups.

Highlighting

Icon →

- This is used to highlight a data in the view.
- The data highlighted we want to can be customized from the icon above below the tool bar.
- Suppose, we have a view in the tableau like catg & segment in column shelf & sales, region in row shelf.
- To highlight go to above icon → first enable highlighting
→ Then select the dim. we want to highlight / or set of combination of dimension we need to highlight → click on any part then the selected part will be highlighted for all parts.

ACTION FILTERS

Icon

- The action filter is used to filter out the data when click any part of one sheet, the data will be filtered out in view of second sheet (Target sheet)
- How to perform this?

Steps

- 1) Make sure you have two views created in two sheets & first determine what is your source sheet & target sheet.
Target is generally the → Source sheet ⇒ Catg v/s sales } i.e. Target is Target sheet ⇒ Subcat v/s Sales } at higher granularity.
- 2) Go to worksheet tab → Action filter → Add Action
→ Filter → Select source sheet → Run action on
→ Three type are there (Hover, Select, Menu) → select Target sheet → Go to
Clearing the section will → show all values → ok

- Now when we click on category view of any part it will filter the data in second sheet.
- This is very imp to create the dashboards.

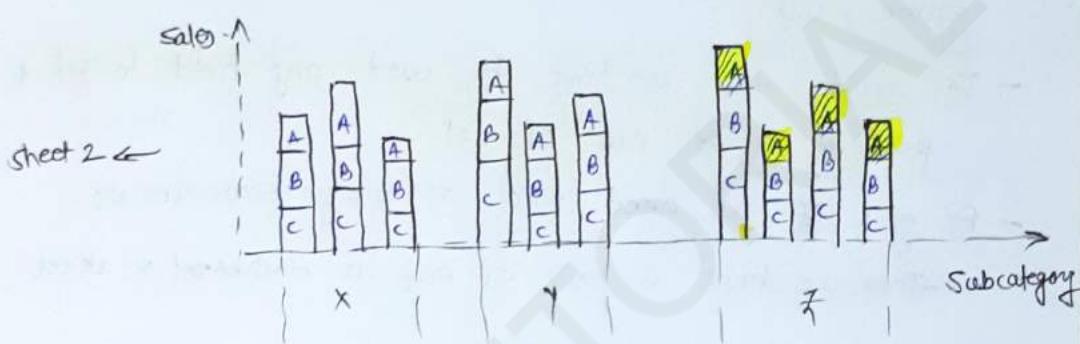
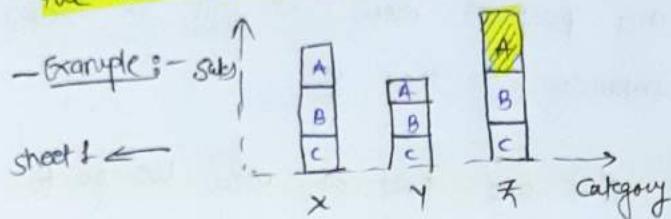
There are three type for Run Action on

- 1) Select → This is used when we click the part of view filtered will be applied.
- 2) Hover → This is used when we just take cursor on any part of view like hyperlink
But for this in clearing selection will →
select "keep filtered values"
- 3) Menu → This used when we click on part of view the filter will be present inside tooltip, when we click to tool tip, we get the target sheet filtered.

* To edit filter → go to worksheet → Action → double click on filter name we already created.

* Actions Highlight :-

- There is another option in the action after filter is the Highlight → This is used when we highlight the part of view then some highlights will be filtered out in the another view.



- This filter works as per above ex. when we highlight the A region of a catg. the same region A in subcat for Z is selected out.

- How to do?

Again Go to worksheet → Action → Highlight → select source sheet → select Target sheet → select Run Action on (sheet 2) → ok.

- Highlight filter & all action filter we have to apply again when we go to Dashboard design.

- Dashboard tab has its own Action filter which is diff. from worksheet action. Dashboard action filter is used when we go to dashboard making.

Action URL :-

- This is used to give a hyperlink to any part of a view so that it will take us to another file or on a website.
- When we click on any part of view we will be directed to a webpage containing to that url.
- Mostly when we click any part of view we go to same url
- To avoid this we have to add any field to get put in that part of url.
- For eg : If we need wiki of every state in us
 - Then we have a view of map in dashboard or sheet.
 - To get Action url → Go to sheet → Go to worksheet → Actions → Go to Url → Select sheet name → put url → click ok

eg. URL → <https://en.wikipedia.org/wiki/California>

Now here instead to get wiki of any state when we click then we should

Now, to get the wiki of that state which we click in view then we modify url as

URL → <https://en.wikipedia.org/wiki/<state>>

- To get state we have to click on this

Again to get the action in dashboard we have to use Action Rule in the Dashboard Tab, so Action applied on worksheet will not apply in dashboard.

DASHBOARD

- Dashboard is used for combining views of different sheets into a single place.
- There are many options in dashboard to make it interactive & talkative.
- There are mainly two options
 - i) Tiled → When we drag & drop sheet in this it cannot be tiled manually / automatically.
 - ii) Floating → These can be tiled & customized & placed manually.
- There are different "object" to drag & drop in dashboard available like:
 - i) Horizontal → Convert dashboard view into two vertically viewers which are auto-adjusted, when any graph gets compressed or of lower size after applying filter.
 - ii) Vertical → Splits this makes view into two horiz. views.
 - iii) Blank → Insert a blank box where we can type anything.
 - iv) Navigation
 - v) Text → To insert text
 - vi) Image → for inserting image.
 - vii) Webpage → To get webpage connected in dashboard.
 - viii) Extension

STORY POINT :-

- Sequential way of visualizing Analyzed data we used Story Point to do this.
- In story we can different sheet views as well as a dashboard as an story.
- Just we have to drag & drop the sheet or dashboard into the story.
- The caption is generated for every story we create where we can add text info we want.

DEVICE DESIGNER :-

- It allows us to design layout in such a way that it will be visible in mobile display, desktop, ipad, etc.
- Go to Device preview → Here we can select the "Device type" as Default / Desktop / Tablet / Phone → then we can choose "Model" → then Add design Layout → layout will be chosen & added to preview
- In Layout Desktop → select "custom" →

TABLEAU ONLINE :-

- To publish the data online or to send the report prepared to manager all over world we need tableau online.

website → online.tableau.com

TABLEAU BY KIRILL

Dimensions → These are independent variables

Measures → Depend variable.

- While dragging & dropping a field any of shelf be careful you drop according to level of detail / granularity or we can say hierarchy
- Always drop for eg. first category, then subcategory or first region then category or first region then sales representative
- If we drop first subcateg. & then catg. it will give result by drill down we say it is not correct or it is mixed up.
- If dropped by mistake you can swap the more aggregated field first.
- * To drag without removing pill already from any shelf to any other shelf or card do that by pressing Ctrl. That will keep it in both places. Instead of going to the dim & measures to do this we can follow this method.

DATA EXTRACT

Live Connection :- Allows you real time data

Extract Connection :- Extract is kind of batch which needs to be refreshed from time to time to get the updated data.

- Extract are saved subset of data that you can use to improve performance or to take advantage of Tableau functionality not available or supported in your original data
- When we create extract of data, you can reduce total amount of data by using filters or configuring other limits.

How to Data Extract ?

- Connect our file with tableau → Go to Sheet → In data → right click on data source name connected → Extract Data → Extract → choose location to save → as we save → our symbol changes in data → before datasource name → ie. like 

- Also in folder new file has been created

Top Notes

- Now as we created extract of our data → The new file is created where we saved → Now what extract does is that when we did any changes in the source data we got from we will not be affected in our view bcz of that.
- Still if we want to change in our view with extract → right click on data source → extract → refresh → our view will be updated.
- Also why we use extract is that for very large data having many records it is very hard to use live connection as if data keep changing live connection allow directly to connect with source data & our tableau view changes if this cause overload.
- Extract allows us to create separate extract file which can also keep data filtered from the main data source. Allow tableau to work faster & query faster.
- Still if we want to go to live connection → Go to right click on data source name → & uncheck the "Use Extract" → we can go live directly.

Benefits of Extract

- i) Speeding up workbook through optimization: Extract is embedded (fixed) within workbook. Thus, extract are often much faster than live connection, especially when querying complex visualizations, eg. large data sets or many filters. Therefore, extracts are great for flow of visual analytics.

ii) Offline:- As extract is embedded, your data source can be utilized when you are offline or your network connection is poor.

iii) Privacy:- Hiding certain fields within your data source & then creating extract, you can hide fields that aren't used, ensuring data can be viewed only people we allow.

+ TIME SERIES:

- For Time series analysis over a period w.r.t some measure when we drag a date field in column shelf & measure in row shelf we get a timeline line chart.

- Now, here first we get a date field ~~is~~ will be in

Blue Color → Indicates it is a discrete field i.e. it acts like a dimension.

(Dimension)
(Discrete)
(More aggregated)
(Less granular)

- It is a more aggregated field ~~than~~ & at less granular.

- When we select month from pill i.e. from first section we select month then it shows aggregated value of month i.e. for all year ~~is~~ it shows 12 month aggregating all year value in single month.

- This is aggregation of a value which is caused due to dim. field for you

Green Color

→ To convert this dim i.e. blue field into measure, click to dropdown of pill & select second section Year or month

- This gives a perfect timeline with more granularity & Level of detail which gives us the timeline for every year & with separated months as we roll our cursor
- This indicates that timeline is more granular & less aggregated.

** Time **

* Aggregation, Granularity & Level of Detail (LOD) :-

- When we make any view tableau automatically does the aggregations like sum(sales), sum(profit) for say different dim like catg. wise, region wise or for range of years/months/quarters, etc.
- How does it work actually?

Tableau will always aggregate measures at level of Granularity of your worksheet

Example :- Say we view we have by Month v/s sales, hence it aggregate all values for particular month of a year.

Concept behind Tableau is that measures get aggregated if dimensions specify the level of granularity.

$$\text{Aggregation} \times \frac{1}{\text{Granularity}}$$

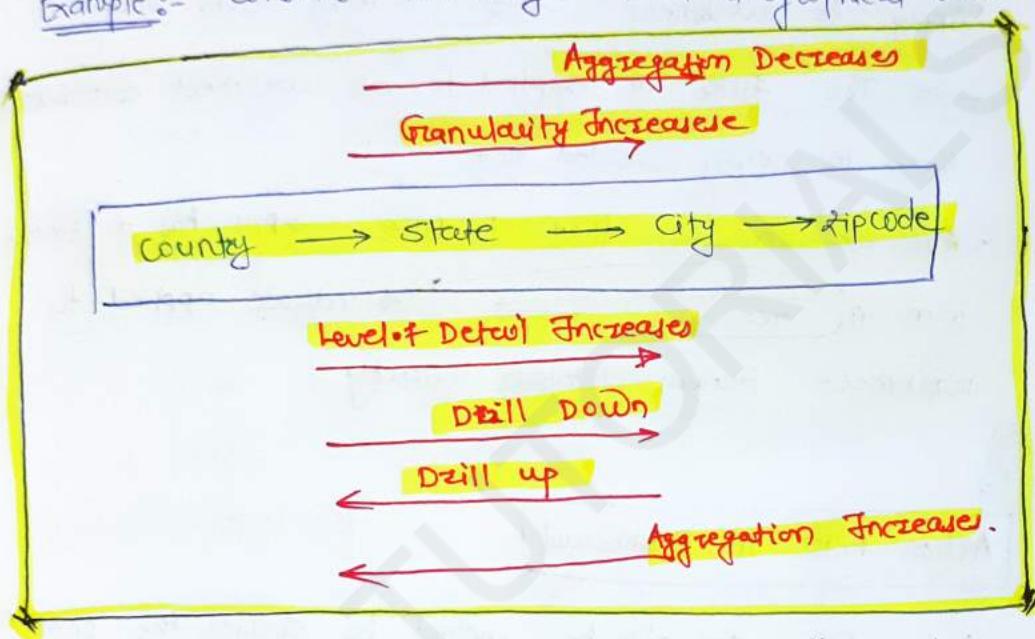
- When we uncheck the "Aggregate Measures" from the **Analytics** tab we get the view for each value of **states**. From marks tab → change from line / chart to circle o or square □
- The level of Granularity in the view can be increased by dragging any of dimension in the marks card at "color" or "Detail" card. By doing this our view becomes more granular again.
- The "color" card also increases level of granularity but gives us with more color differentiating each part of view, but "Detail" card allows mainly the level of granularity increase in the view without changing any of color in the view.
- While using Extract refresh to get data in tableau view changed as per source changed, first data after data change in source say in excel | csv it must be saved after data change.

Relationship, Joint & Blending :-

- We will study some Hierarchy first →

→ Hierarchy of the dim. decides that the the Granularity & level of detail we go in view

Example:- Consider Hierarchy of a Geographical dim.



- For Hierarchy when we created → then when we drop county in view → we get + sign where we can increase the granularity or Drill down or increase Level of Detail & decrease Aggregation
- When we use the profit or any other dim. field in the color card to get lower or higher. We can custom our values to show to minm or dark shade color or maxm shade colour by double clicking on the color legends → Advanced → apply start & end value & centre value.

Filter :-

- When we are working on a multiple sheet & we want to add a filter say year wise & if we want to apply for all sheets.
- In first worksheet while working → drag the year date pill in filter card → drop down → "Apply to worksheet" → "All using Data source" → This filter will be applied to all worksheet automatically even to newly created one.
- Also we can see that in filter → pill has a small icon which indicate applied to worksheet through database directly.

Jump

Action Filter in Dashboard

- Action filter can also be applied by clicking the sheet we dropped in dashboard → click on that view → Go to dropdown → "Use as filter" → This act as a action filter → Type "filter".

Action Filter → Highlighting

- In dashboard to filter & highlight from view to change other view we use ~~to~~ these two types mainly
- Let us see diff. b/w them.

- Say that if we filter \rightarrow then what filter or "Action filter" does is that it filter data for selected part or data & temporarily deletes the remaining data \rightarrow As we do this our view connected to that get changed to filtered data only. This is as simple as that.
- Unlike in the "Action Highlighting" \rightarrow Here we do not filter data \rightarrow we just highlight the other ^(target) new data for selected ~~or~~ data in first (source) view.
But applying Action Highlighter does not highlight view in target sometimes. This is because the data in target view is shared by one or more data in source view, where tableau gets confused & doesn't highlight anything.

To overcome this we have to make the target view or worksheet more granular based on source view fields. When tableau is ready or can distinguish the highlighted part then only we get highlighted view in Target sheet.

Joins | Blending | Relationship

8-

Before Tableau 10
revision

JOINS:

Use JOINS when we want to combine data at row level & datasource must be same (i.e. all from either excel or csv, no combn allowed)

- We know there are different types of Join we use in SQL as well as Tableau.
- Join works on 2 or more tables for 1:one or more matching fields in \leftrightarrow tables.
- If there is only single matching field the join will be performed on

$$\leftarrow \text{Table 1}, \text{common column name} = \langle \text{Table 2} \rangle \cdot \text{common column name}$$

- But if there are more than one field in both table which matches then,
 - i) If we perform joining only on one column then it gives error or gives us result but totally wrong.
It can be done correctly if one of table field have PK i.e. unique fields.
 - ii) But if tables don't use PK & many duplicates in both table for 2 matching field then we have to always perform join on both fields i.e.

on $\langle \text{Table 1} \rangle \cdot \langle \text{column T1}_1 \rangle = \langle \text{Table 2} \rangle \cdot \langle \text{column T2}_1 \rangle$,

and $\langle \text{Table 1} \rangle \cdot \langle \text{column T1}_2 \rangle_2 = \langle \text{Table 2} \rangle \cdot \langle \text{column T2}_2 \rangle$

Blending :-

- This is mainly used when the data we have are from different source & when we have data from same source also but the two data or files are at different level of granularity or detail where we cannot perform join operation.
- The two data sources we want to blend are on same excel file with different excel worksheet then first call the excel file drag first the data source & again add the data ^{source} with same worksheet & this time drag another sheet which we want to blend & we can see the two datasources in the tableau sheet at the left corner in data.
- We can also not perform join even if level of granularity same but datasets are from from different source eg. (one from excel other from csv / one from excel other from SQL) , then we have to compulsarilly use only blending.
- Data Blending uses Left Join by Default depending upon Prim. Data source (Left Table) & Secondary data source (Right Table).
- Aggregation of data in blending happens on level of granularity of the view. Tableau sends separate query to both the data sources.
- Aggregation happens before blend
- Blending happens at worksheet Separately.

Blending Example :-

- Let say we have sales data & Target data which are at different level of details & we have to perform Blending on that.
- Now we have the blending & view done for sum data source on Month vs sales for different categories i.e. we have Month (continuous field, Green color) in column & Category then sum(sales) in Row shelf.
- Now we have to blend & make link b/w the dim fields on view i.e. month & category. Perform Join ~~on~~ or make blending over these dim fields.
- Then, drop target sales in row shelf after sales & we can compare target & sales achieved sales, then ~~on~~ combine them by dual axis, just make sure that axis are synchronized i.e. they have same axis fig. i.e. same axis division.

~~W.B.~~

To synchronize → right click on vertical axis
→ click on "Synchronize Axis".

- To bring chart front or back i.e. sent back & bring front that depend which is last pill in shelf, swap the pill back or front to change view to send field back or to bring front.

* Calculation Field In Blending :-

More :- While performing blending on a any field it is compulsory that it should be a view but not compulsory it should in shelfs only it can be placed into the marks card at "Detail" or at "Color" card.

- To perform calculation field while blending it can be done -
- Right click on any field → calculation field → Create → Type here fields we need → Here we can take field from two sources.

Example

$$\text{sum}([\text{Sales}]) - \text{sum}([\text{SalesTarget}])$$

↑
This measure is from
prim data source

↑ [we can also drag +
drop field in calculation]
This measure is from
secondary data source -

- This can be dragged i.e. calculation field can be dragged & dropped into view for viz. but as we are performing blending, make sure all those pills are there in viz for which we made join relation for blending, mainly in the marks card. If not there it gives us wrong aggregation result.

Map Using Long & Latitude Values :-

- Map in tableau can be viewed by various methods like following Geographical dimensions

- i) Country name
- ii) State name
- iii) City name
- iv) Zip code
- v) Longitude & Latitude values

* - Now first 4 are very easy to just drag drop, just make sure data type is Geographical 

* - For Latitude & longitudinal also the data type must be compulsory Geographical only

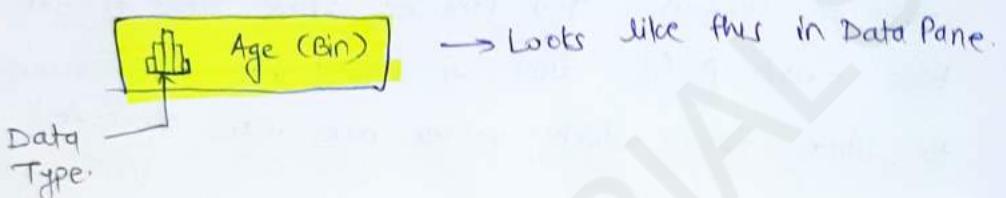
* - To draw map with Lat & Long value make sure that we drop the Longitude value in column & Latitude value in Rows shelf only.

BINS IN TABLEAU :-

- Sometimes it is useful to convert continuous measure (or a numeric dimension) into bins.
- Any discrete field already as a dimension field in Tableau can be considered as a set of bins.
For ex, say we create view with Profit on Rows & state on columns. You consider state field as set of bins - each profit value is sorted into bin corresponding to state from which values are recorded.
- If we want to see value of profit assigned to bins without reference of dim field, we can do that by creating numeric bin, which each bin consisting of a range of value.
- Example:- Say suppose we have Age as dimension measure in data ~~is~~ having values from 1 to 70 now we want to plot a view based on group of 5 years combined. Then we have to create a bin for this range of 5. That is we create (0-5) as one bin, (5-10) as second & so on.
- Keep that in mind that after creating a bin that new field created will be treated as dimension field & not measure as we create its conversion from continuous to discrete.
- Also keep in mind that Bin fields cannot be used in calculations. But we can create calculated field that will replicate a bin of specific size.

Create Histogram from Binned Dimension?

- If we create binned dimension, we can use it as starting point for creating a histogram.
- When we create a Bin it is a dim field with range of value & considered as discrete field & the pill looks like this.



- Tableau can also suggest the bin size for us depending upon the no. of records we have in our data source.
- Or we can manually give a size depending on our approach.

Alternate way to Create Bin

- We can create bin by another method also, just by clicking on a measure field for which we want create bin → Go to "Show me" → Click on Histogram chart.
- This creates an automatic bin for us and displays a chart as a Histogram.
- The column contains auto generated field called "Bin" which can be seen in dim. in data pane, & in row shelf it takes the count of total records.
- Later if we want to set interval → we can edit the bin from newly created field in pane.
- Also the chart / view we get is bin created with a continuous field, to create discrete click on bin field dropdown in column shelf, then select discrete.

PARAMETERS

- Parameter is a workbook variable such as number (float or Integer), Date or string that can replace a constant set value in to a range value or calculation or filter.
- Parameters once created in any worksheet are global for the workbook & can be used in any worksheet.
- Bins can also use a parameter value → go to bin field → Edit pin → size of bin → dropdown → Select the parameter which we have created already.

DATA PREPARATION

- The data that we import into Tableau for analysing must be in a proper format as well organized manner.
- There should not be other heading than column heading, no grand total should be there, no pivoting of table should be there.
- Data should be properly in column with column heading & records in rows.

DATA INTERPRETER

→ Data Interpreter is only available for excel, pdf, csv, google sheet.

- This is function available in Tableau when we upload or connect any file to tableau.
- It is used to clean data from excel, csv, pdf & Google Sheet with Data Interpreter.
- The file we upload may have things like Titles, stacked / merged headers, notes, empty rows & columns, white space.
- When we analyse this in tableau, it is very difficult for tableau to interpret your data. That's where Data Interpreter help us.
- It is automatic cleaner pops up when we upload / connect the file, just check that box & things happen.
- When we click on ~~Read~~ "Review the result", we get file that tableau has made changes.

DATA PIVOT

- Most of time the record we have in excel, csv, pdf file is stored or have in cross-tab format & is not good to get interpret by Tableau.
- To convert those record into a columnar format we use data Pivot tab.
- After we connect data to the tableau & can view the records in the grid or data pane → select the two or more column we want to pivot → click drop down arrow next to the column name → select "Pivot".
New columns called "Pivot Field Names" & "Pivot Field Values" are created & added to data source. These columns replace the original columns selected for pivot.

SPLIT

- We to split the column into more than one or more column

- Example :-

First Name	split	First Name	Last Name
Abhijit Kadam	→	Abhijit	Kadam

- Split is available at column header drop down & require a separator as (space, ., ., /, - , etc).

METADATA

- Info all about the column fields we have in our data in grid.

Highlighter

→ Newly introduced in Tableau

10.0V39

ATAC

- When we quickly want to find the marks or set of marks on view without using other marks like color card, label card in view we use highlighter.
- This is a new feature in Tableau which allows us quickly find & highlight data while maintaining all data in the view (without disturbing any colors / labels / sizes).
- The work of highlighter can also be done by the legends with colors, but if there so many data then we cannot use legends as we cannot type search in legends.
- Unlike in highlighter we can search and filter what we need to see. It acts like a action highlighter filter.
- How to use highlighter
 - It is mainly used by dragging & dropping a discrete field i.e. dimension into a "Detail" in mark card
 - Go to dropdown of that pill → choose highlighter
 - As we select to drop in label it will not affect the view of map → just may make it more granular.
 - Then filter will act as highlighter we can search & that part will be highlighted in view.

CLUSTERS

- It is a powerful feature in tableau that allows you to easily group similar dimension members in a view.
- This helps us create statistically based segments which provide insight into how different groups are similar as well as how they perform from each other.
- The cluster → is in "Analytics" pane → drop + drag that in view + it automatically creates cluster → we can edit it → it will be present in marks card → drop down → edit → we can make custom groups of cluster we need → or also add drag + drop more measures or dim. field in cluster.
- Mostly used on a scatter plot views.

Trend Lines

- These are the line of best fit for set of data to represent behaviour in certain pattern.
- It is available in "Analytics" pane can be used by drag + drop + wed with cluster for each cluster how we can use the trend line.
- Trend lines are mainly drawn for a scatter plot cluster can be save → drag + drop created cluster in marks card in data pane → it will be saved as a group → rename → This cluster act like an dimension & can be used for different visualization in a view.

CROSS DATABASE JOIN → Tableau 10 New

- In tableau before the version release of Tableau 10 we cannot perform Join on the Cross database i.e. one data source from Excel / CSV + other from SQL Server or any other.
- Now with Tableau 10 we can join the data sources from different databases.

Ex :- Excel with CSV

Excel with SQL Server

SQL Server with MySQL, etc--

- This overcomes that drawback, because previously to do this we used Blending & Blending had some limitations like it won't be able to perform extract, can't publish datasource as when published link b/w prim & sec. data source w/ lost bcz blending w/ done on worksheet level, etc-
- By Cross database join the color code w/ given to each source joined so that fields can be identified easily that they are from which source name.

TABLEAU ADVANCED

GROUPS & SETS

A) GROUP

Dimension to Measure & vice versa

- NOTE** :- In tableau if by some error if tableau shows some of field in green datatype i.e. it is a measure field & displays it in dimension, then we have fix it.
- If it is a measure drag & place it to measure from dim.
 - If it is dimension & it is still shown by green datatype & in dimension only it is placed then drag that place in measures & again drag & place in dim. (Green comes to blue i.e. discrete field) & get convert to dimension.
 - This mostly happen when period / Date is in only year i.e. no. (2015, 2016, etc), Tableau confuses & sees that like measures.

GROUP

From → ⚡

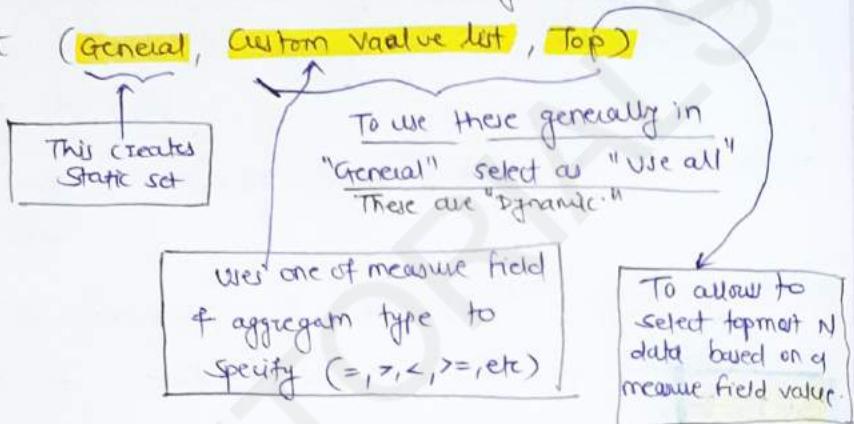
- We can group our data or create a group to combine related members in field.
- Groups are useful for both correcting data type errors like (California, CA, Calif. these are same, so group them) Also Group answer the "what if" question eg. what if the East & South regions are ~~are~~ grouped.

SETS :-

- We can use sets to compare & ask questions about a subset of data. Sets are custom field that define subset of data based on some condn.
- We can make sets more dynamic & interactive by using them in Set Actions.
- There are two types of set
 - 1) Static Set → These can be done from view by selecting one or more items → hover over → create set
 - 2) Dynamic Set → These can be done by clicking dropdown on a dimension field in data pane → create → set → set can be then used with parameter to make it dynamic.

(Creating Dynamic Set) :-

- Static set are fixed & created by selecting item from visuals & cannot be changed.
- Dynamic set can't be created from visuals, they should be created from dim. field in data pane & then they involved measure field to set them.
- Once we select create \rightarrow set \rightarrow we get 3 options to make set (**General**, **Custom Value List**, **Top**)



Combining Set :-

- More than one set created can be used in two combination to analyse the data.
- But make sure that they must be created from the same dimension & for same measure. Can use different measures field.
- To create combined set \rightarrow click dropdown of set created already \rightarrow create \rightarrow combined set

Created sets in calculation field

- We can we created set in calculation field also.
- Say suppose we created set A for customer having
Salary > 50000 & < 20000

Then the can be used in calculation field A

for eg.

```
If ([set2 < 20k]) then "Small sales"  
Elseif ([set2 > 50k]) then "Large sales"  
Else "medium"  
End.
```

ADVANCED TABLE CALCULATION

Example :- Coal Mining

- Table calculation are performed after the aggregation of view is done.
- Table calculation are done at visualization level & not have any relation with data source level.
- Whereas, calculation field is performed at the data source level.
- Various Table Calculation :-

i) Running Total :-

- It is done to get the cumulative total as the chart advances.
- After clicking then by "compute using" we can apply with what we want either by "cell" or by "field"

e) Difference :-

- Difference gives us the subtraction of second value & first value & successively increase with extend of view.
- By using "compute using" we have options like same as above
- Also by using "relative to" we have
 - i) Previous → Difference from previous value & cumulative
 - ii) Next → Difference from next value & cumulatively
 - iii) First → All value difference with only first value
 - iv) Last → All with only last value.

- Difference also has "compute Using"

i) **Table (down)** → difference is done from top to down where values will be subtracted from top data to lower data successively.

ii) **Table (Across)** → difference is done from left to right successively.

iii) **Table (across then down)** → from left to right then top to bottom.

iv) **Table (down then across)** → vice versa as above.

v) **Pane (down)** → for that particular level of detail or pane in top to bottom.

vi) **Cell** → cell wise difference.

* How to Use & Save Table Calculation :-

- Table calculation that are done at visualization level can be saved as a calculation field.
- Just take the pill that we have performed table calculation & drag & drop into the data pane → it will be saved as calculation field & can be used as a field in other calculation.
- The saved field will be always used as a ~~T~~ platform as a Table calc only as it will have Δ in its pill.
- The save calc → open edit → Here we have "default Table Calculation" → here we can edit "compute using" → it is always better to save by field which is on axis → that will give us if field is on Y axis that is taken as Table across → if on X axis taken as Table down → with that select field so that in case if we change orientation of view the table calc is depend on field & will not change the Table calc.

* Reference line for Table calculation field :-

- Now when we need to add ref line to table calc field there are two methods very imp.

i) Right click on ~~set~~ axis created by Table Calⁿ :-

- Right click on the axis that is created or related to the axis , because to add ref line to the table & set value on that axis we have to do this
- If we double click from Analytics pan set will not show w the option to get ref line correctly.

ii) Drag Ref line view & select Table Calⁿ :-

- Drag Ref line from Analytics pan in view & select field we want to add ref line to then Select .

MOVING AVERAGE

- Moving Average is the ~~field~~ → table calculation to determine the average of data points by creating a series of averages of different subset of full data set.
- It can be created in calculated field → Table calculation
- Window-Avg()
- Window-Avg() takes three arguments:

Window-Avg ([Field name] measure field / calculated field), -7, 0)

This is a field of which we need to determine moving average

i.e. average will be of last 7 + current value
i.e. total (8)

This is used to choose offset of data we have
-ve to take first & +ve of next or (Last)

Upto where we need the average
If (0) then i.e. current point & that value also taken to average.

Example :-

Sr-No	Total Sales	moving average	formula used
1	10	$\rightarrow 10/1$	window-Avg([Total Sales], -2, 0)
2	15	$\rightarrow 10+15/2$	
3	20	$\rightarrow 10+15+20/3$	last 2 values + current value
4	25	$\rightarrow 15+20+25/3$	
5	30	$\rightarrow 20+25+30/3$	For these as they are first values so general avg is taken only.
6	35	$\rightarrow 25+30+35/3$	
7	40	$\rightarrow 30+35+40/3$	
8	45	$\rightarrow 35+40+45/3$	
9	50	$\rightarrow 40+45+50/3$	
10	60	$\rightarrow 45+50+60/3$	

Vimp Note

- Moving Average are always or mostly taken for the "Y-axis field" or we can say where the graph is expanding or running OR "Table Across"

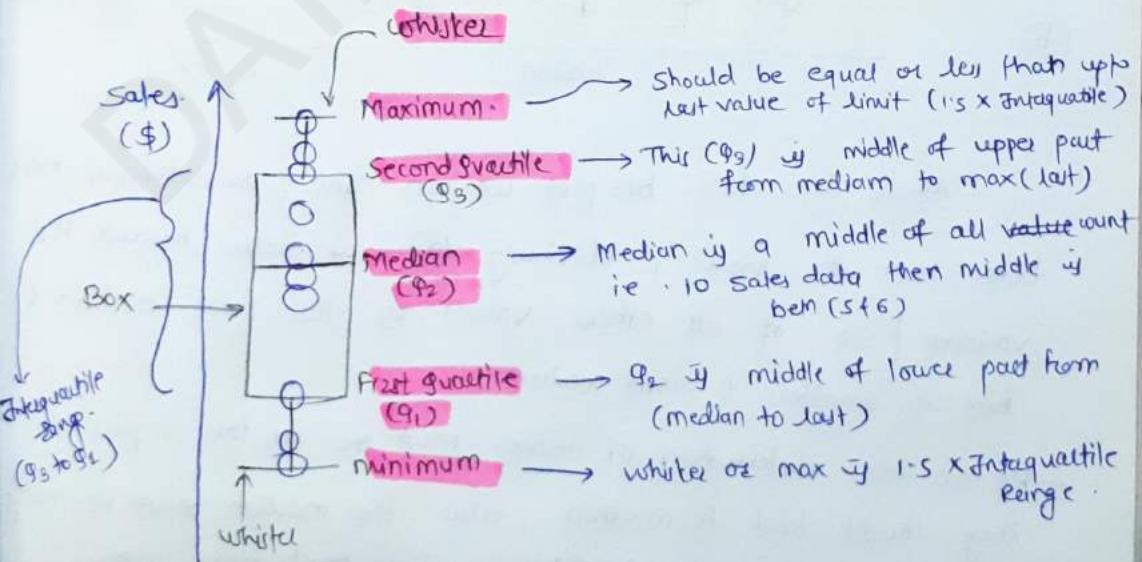
ADVANCED DATA PREPARATION

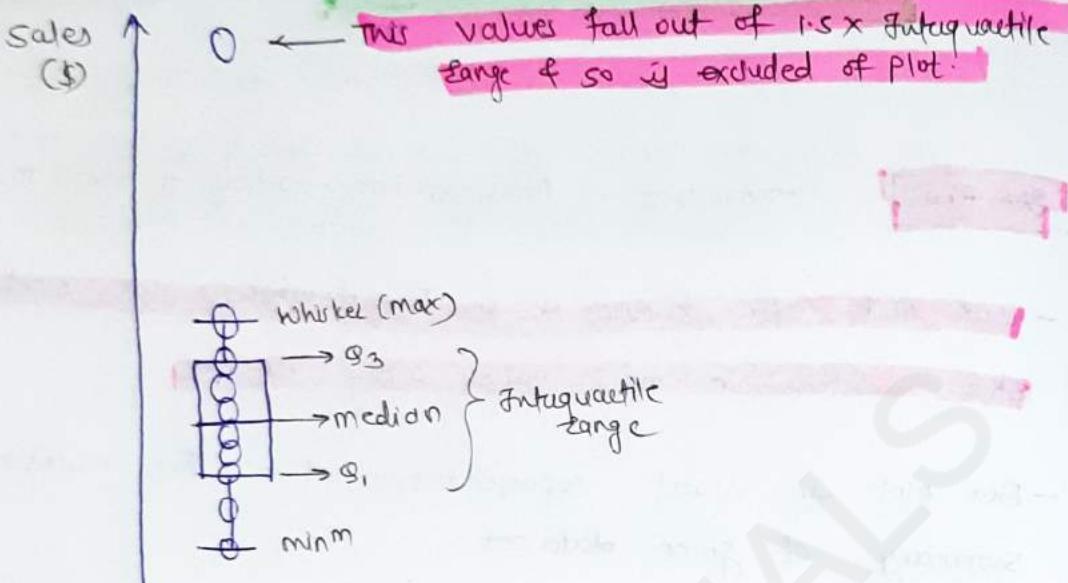
BOX PLOT

→ Available in Analytics Pan → drag to view to see

- Box plot also known as a Box & whisker plot, used to show distribution of values along an axis.
- Box plot is visual representation of a five number summary of given data set.
- Five no. summary includes:
 - i) Minimum
 - ii) First quartile (Q_1)
 - iii) Median (Second quartile) (Q_2)
 - iv) Third quartile (Q_3)
 - v) Maximum.

- Consider an example, we have a view of ^{State} ~~Settler~~ v/s sales analysis without aggregation.

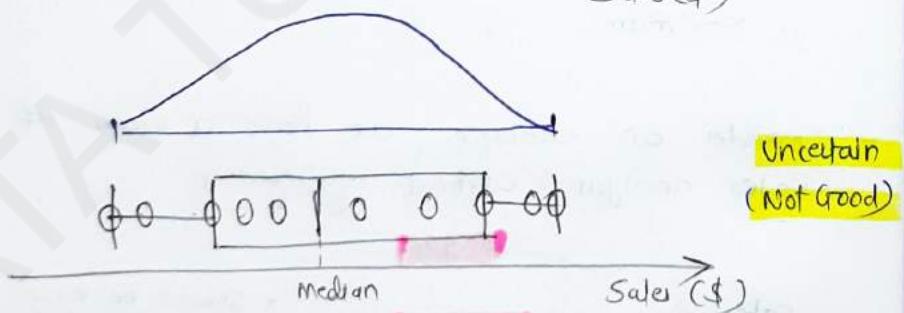




(a)

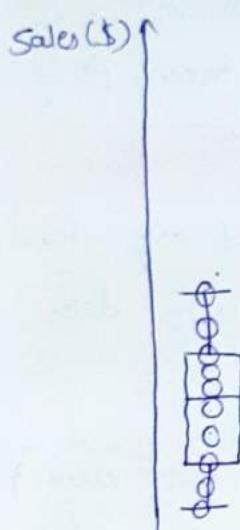


(b)



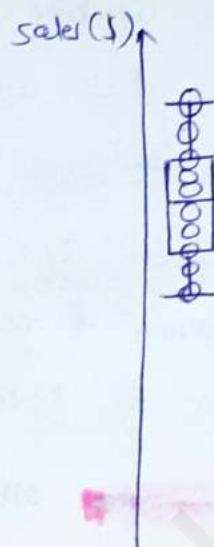
- Above are two box plots we can see, so which is more better from the view, obviously (a) is better because the variance / SD of all other values is less from median & box is smaller & more values.
- We can see distribution of values must be as low as possible, they should tend to median, also the median must be mostly at the higher side of sales, as we need more sales.

one more example



(a)

[Not Better]



(b)

[Better]

- for above example we can see the median & quartile range create box as compact as it can , but out of these which is more likely to be Good
- so (b) is Good because the median value must be at the higher side of sales , as we need our company to make maxm sales .
- (a) looks like there ate very low sales of company .

Larger Data Sets

- When we are working with large data sets, the tableau will start working slow or taken plenty of time to load.
- What we do here is there if we want only few ~~no. of~~ ^{out} no. of column & rows of all or plenty data then we use a Data Source filter.
- **Data Source filter** allows us to filter out rows & columns which are not reqd for analysis & makes tableau faster.
- Also there are some columns which do not have any values in their rows. How to find them?
- Mostly for measure field which have datatype number (#) will appear with column name & (#) if it have values in it, if it doesn't have value at all it shows datatype for measure field as (Abc) which indicates that column is total empty as tableau can't understand what datatype to use for null value column it uses (Abc)
- We can find such in metadata tab & ~~hide~~ all those columns.

NOTE :- We get rid of all the blank column because while performing pivot it gives totally wrong results.

* Advanced Level Blending VAMP

- This is very important that blending must be done correctly for time line series.
- Always blend your data at the level of year i.e. it will take the correct aggregation for both sources for that particular year only.
- Next then perform relationship based on view or level of detail i.e. ~~per~~ as per quarter, month, week, etc.
- Always check your source data whether the monthly are given uniformly or they are at quarterly level as per that we have to perform blending relationship.
- In Blending always remember that tableau sends query to both databases, aggregates the measure field at the level of detail of view then displays it on the view.

Inp Note :- Always check the data connection blend relationship which we used earlier for say (A) datasource as primary, in next worksheet if we used (B) as prim. then blend relationship should be changed accordingly as we require.

* SALES PER CAPITA :-

- Per Capita means per person.
- It is commonly used in statistics, economics & business to report on an per person sales / profit, etc.
- Per Capita is often used to compare economic indicators of countries with different population sizes.

eg

$$\boxed{\text{sum(sales)} / \text{sum(population)}}$$

* **NOTE**

- For calculation field in blending data, the aggregation is very important, blending always done for aggregation of data at level of detail of view.

* ANIMATIONS IN TABLEAU :-

- Animation to a view in tableau can be done by giving any field into the "PAGE MATRIX".
- Here we can play video, move page accordingly & also increase & decrease speed of video.

Annotation :-

- Annotation in tableau are used to give or do give reader attention to a point in the view.
- It can be done by right click on view → annotation
- Two option mainly are
 - i) Point :- To mark & give some name to a particular point or part in view.
 - ii) Area :- This can specify the particular point or information to all view.
- Annotation can be made dynamic by selecting → annotations → e.g. "Area" → right side "Filter" → select the field changing → By filter on field changes Annotation also changes.

PAGE MARK with history



- In page mark there is a history option which is a very powerful option which allows us to see the how the particular part of view starts from start to end & allow us to show the path it travelled.
- It has various option in its drop down:



- LOD expressions are used to run complex queries involving many dimensions of data source level instead of bringing ~~them~~ in all data to Tableau Interface.

Types of LOD:

+)**FIXED LOD:**

- This LOD expression computes values using the specified dimension without reference to any other dimension in the view.
- Simply this LOD expression aggregates value only at the dimension specified in the LOD expression calculation. Unlike Include & Exclude LOD, fixed LOD does not take into consideration dimension in the view & so it is very powerful.
- When we add LOD expression in a view, Tableau must reconcile two levels of details — one in view, & one in your expression.
- The behavior of LOD expression in a view varies upon whether expression LOD is "coarse" or "fine", or some level of detail of view.

Example of "Coarser" View :-

- An expression has coarser level of detail than view when it references a subset of dimension in the view.
- Consider the view without LOD expression has the dimensions [category] & [segment].
- Now we create LOD expression calculate field as { Fixed [segment] : sum([sales]) }
- Now when add above LOD expression in view, the expression have coarser view level than view already had. Expression is based on single dimension [segment], whereas view is based on two dim. [segment] & [category].
- This result the view contains certain replicated values that appear multiple times.
- Replicated values will be useful for comparing specific values against average values within a category.
- When values are being replicated, changing aggregation for relevant field in LOD expression (for ex from Sum to Avg or vice versa) will not change result of aggregation.

Example of "Fine" View :-

- The LOD expression has finer level of detail than view when it references to superset of dimension in view.
- When we use finer expression than in view already have tableau will aggregate result at level of view already has

For example

{Fixed [segment], [category] : sum ([sales])}

- Now when view has Level of detail of [segment] only if we use above LOD expression in the view.
- We get only the aggregation average value automatically by the tableau.
- Fixed LOD can have either Coarser, finer or at same level of detail the view already has.

INCLUDE LOD :-

- Include LOD expression compute values using specified dimensions in addition to whatever dimension are in the view.
- In other words Include expression calculates the specified dimension in your syntax in addition to dimension already in your view.
- Include expression is for calculating higher level of detail (~~more~~ finer) without changing your view.
- Include LOD also allows us to add dimensionality to my view without adding dimension to columns, rows or details button.

Example

```
{Include [customer name] : sum([sales]) }
```

EXCLUDE LOD :-

- This LOD expression will subtract dimension from the view level of detail.
- It excludes the values of dimension Δ from the view
 \uparrow
measure.
- Exclude LOD are useful for "percent of total" or "difference from overall average" scenarios.

LOD by KIRILL

- LOD expression brings out the information at different level of granularity we need.

$$\text{Aggregation} \propto \frac{1}{\text{Granularity}}$$

- Dimensions field decides the level of granularity of measures field does aggregation over that level of granularity.
- Types of LOD

b) INCLUDE LOD

- It allows us to include additional dimension in your calculations.
- It allows us to go at higher level of granularity from level of aggregation which is already in view.

Example :-

- view consist of level of detail at [Segment] & [Category]
Sales
- Include allows us to go drill down to say [subcategory] level without changing original level of visualization.
- visualization will remain at [Segment] & [category] but [subcategory] will be added by calculation to include more level of granularity.

e) EXCLUDE LOD

- When we have our level of detail of visualization at say for eg. [country], [state], [city].
 - + We want to go to ^{lower} ~~higher~~ level of granularity & higher level of aggregation we use Exclude LOD
- This also will not disturb original level of visualization.
Exclude LOD will added by calculation.
- Allows us exclude the dimension from the level of detail of visualization without changing original view detail.

③) FIXED LOD

- This type of LOD won't rely on original level of granularity or aggregation of your visual visualization.
- Unlike 'Include' & 'Exclude' it doesn't depend upon the original level of visualization.
- Fixed LOD, it doesn't matter from which actual visualization you are calling this formula.
- Include & Exclude LOD are totally relative/ dependent on level of visualization originally you are at whereas Fixed LOD doesn't matter as it is absolute & totally independent at what level your original visualization is.

Example of LOD :-

① INCLUDE LOD :-

- This allows to add up the dimension to the original level of detail view.
- Let say our view is at an level of [Region], [Category] vis Avg [Profit]
- Include LOD allows us to increase granularity by adding [subcategory] to original view without changing level of detail of view.

Name → {Include [subcategory] : (sum [Profit])}

→ Inp that we used here sum (profit), what this will do is that for particular category it will have subcategory, it will sum (profit).

→ Now replace new field to Avg ([Profit]) we have, then this gives us sum (include) → change this to Avg (include)

→ The average we get here is $\Rightarrow \frac{\text{total sum (subcategory profit)}}{\text{total no. of subcategory in particular category}}$.

→ The only avg (profit) gives us $\Rightarrow \frac{\text{Total sum (Profit)}}{\frac{\text{Total no. of orders}}{\text{(Total no. of records)}}$

Inp - simply Include LOD allows us to increase granularity & decrease aggregation without changing level of visualization of original view. because it lowest granular level at view.

- Include LOD we either take Avg (sales) or sum (sales) as we add dim to view value will be same for avg from

② Exclude LOD :-

- This LOD allows us to come down to level of granularity & increase aggregation without changing level of visualization.

- Example :-

Let say our view is at the level below.

[country], [state], [city] v/s sum(profit),

- Now we want to get sum(profit) state wise without changing the current level of view then we use as exclude LOD

Exclude LOD \leftarrow Name.

{ Exclude [city] : sum([sales]) }

- Here it is very imp that we are excluding [city] to get [state] sales, very imp.

- Now what we get in our visualization when we drag above calculation LOD in our \$ label card is that for every city in map sum [profit] for state will be there, which is repeated for all the city in particular state.

- The pill which we dragged will show below,

ATTR (Exclude LOD)

- Why we get to see here "ATTR" & not "sum" because already we are at level of detail in view at [city] level which is already granular, when we drag a LOD [state] level, it is at one step down granular & will

have higher value than original level & also value will be repeated & so will always show "ATTR".

③ FIXED LOD

- This type of LOD does not depend on what level your view is.
- You can explicitly drag the type of LOD in to view & gain the Insights.
- This is mostly used LOD because unlike Include & Exclude it is independent on level of visualization we have.

+ Advanced Mapping Technique :-

SHAPE WITH POLYGON

- To draw a polygon like Rectangle, Triangle, pentagon, hexagon, etc we need three requirements.
 - i) $x + y$ axis coordinates that can generate a point so that we can connect them
 - ii) Direction → path of order they need to follow.
 - iii) shape 2D

- How to draw?

- Get the file connected that has a coordinates of $x + y$ to get connected.
- Put $x + y$ in column f row shelf & from analysis \rightarrow menu off the aggregation.
- Then we the path put the shape if we have one multiple shape present in file in colour mark.
- Then select chart from "automatic" to "Polygon" in marks card.
- The drop path measure in a "path card" in the marks card.

* HOW TO CHANGE BACKGROUND IMAGE WITH PARAMETER

- To change the background images in a worksheet we can do that by using a parameter.
- Consider you have two images then you need two parameters to control it.
- Firstly To add images → Go to → Map menu → Background Images → select Datasource we want to add images to worksheet in → Add Image → Provide Name → → Browse Image → ~~Select~~ ^{Select} X axis Name in X field → And insert here the X pixel size → (size available in properties of image we have → To get → right click on image) → then do same for Y axis → click OK.
- Make sure that when we have to add image to Background , your visualization must contain X & Y axis both measures only , because it can't take the Dimension field there.
- This type / function is mostly used to add Ground plan, building plans to get done.
- To this we need X & Y axis with range of values of the pixels of image.

FORECASTING

* Time Series :-

Moving Average (MA)

- MA is used in finance & statistics.
- Smooths out fluctuations in data by calculating an average within a given window of time periods.
- The bigger the window smoother the line in visualization.

Eg. Let say we take out moving average of sales for 20 days & 30 days.

- 20 days less smoother
- 30 days more smoother than 20 days.

$$\text{Moving Average} = \frac{\text{Sum (Field / measure)}}{\text{no. of periods}}$$

Syntax

g :- Window_Avg (sum([sales]), -4, 0)

Average of four previous rows with current row

If 0 then current row included
If -1 → current value excluded

Types of Moving Average

1) Simple moving Average (SMA)

- Each data point is equally weighted.

2) Exponential moving Average (EMA)

- Recent data is given more weight.

[SMA] :-

- When we calculate the MA it can also be done in table calculation through pill dropdown & also by using formula mentioned before.

- Let's see by pill dropdown.

→ select pill dropdown in measure field → quick table calculation → moving average → It default calculates for ^{before} 2 days including current value (ie. Total = 3) → To edit → again click pill dropdown → edit table calculation → Select & choose values we want.

- we have in edit → Null if there are not enough value → If we uncheck this we get our time line with null values at start if the window average doesn't get the value we require for moving average.

By using calculation field we can control it by parameters.

→

window-Avg (sum([sales]), - [Parameter 1],
[Parameter 2])

Parameter 1 ←

- This argument takes how many last or next rows to include → -ve (Last) & +ve (Next)
- As our seq is last so we used - [Parameter]

Parameter can have values → from L to 180
or any value.

& with this we can enter if also create MA dynamic

Parameter 2

- This takes whether to add current value or not
0 → Add
-1 → Exclude.
- Create parameter with values 0 & -1 & use there.

NOTE → Here mostly what happen if we choose 0 it adds value i.e. if parameter L we choose is say 30 but this value is also included if it is actually 31 it shows on visualization.

- This can be modified & edited easily by changing our formula.

- Let name of parameters be

1) Parameter 1 → select MA

2) Parameter 2 → Add current value

window_Avg (sum [sales]),

(CASE [Add current value])

when 0 then - ([Select MA] - L)

Else - [Select MA]

END),

[Add current value]

1st Argument

But for this parameter

If value is 0
then it add extra
value so we sub
L from Parameter 2

If value is 1
it excludes so select
as it is.

2nd Argument

* How to change Date By using Parameter Control

- We can change End Date & Start Date of particular timeseries very easily.
- We can do it by creating first two parameters name them as for eg. Start Date & End date.
 - Select data type as Date → monthly create from date field dropdown → create → parameter → +
it takes automatic range values min & max
- Do same for parameter name "End date"
- Create calculate field as → Drag field in filter → +
 - [Date] >= [start date]
 - AND [Date] <= [End Date]
 - Select "True"

* DISCRETE DATE

• Data Benefits from Discrete Date :-

- 1) Mostly used in financial Data → change in stock price
- Discrete fields in the view can be used with the continuous fields to get some very better insight.
- Use continuous field for date & select from drop down pill as an discrete to get newly result.

Year on Year Difference

:- Take difference b/w two years.

- This is important when we want to analyze the sales/profit on ~~year~~ monthly & day basis.
- And put month & day discrete on the column & put year discrete on color marks.
- Sales on the Rows.
- Now to perform the double ~~calculation~~ Table calculations like eg.
 - 1) find running total of sales &
 - 2) simultaneously calculate difference from each year sales.
- For this in sum(sales) drop down go to "Add Table calc" Select Running Total → The below secondary calculation assistance check box we can see → click of simultaneously

we can add second Table calculation as say "Difference from" select from what to what we want difference. We can select difference from dimension also, like "Month" → difference each month of all year, "Year" → All values of one year to another year.

- This is very imp to add two table calculation one single go for one single pill.

Monitoring change in values / When connected to Live Data

- very important to study as we get to study how to control our visualization when we are connected to live data.
- like it changes everyday if our source is getting refreshed automatically.
- Download ~~the~~^{historical} data from Any finance website like Yahoo, Tata, etc. from previous yr to current date.

→ Date Part function

- Returns a specified date part as an integer -

Syntax :- Datepart (date-part , date , [start_of_week])
 ↑ Optional.

where,

date-part → Any part of particular date like Year, Month, day of year, day, week, hour, minute, second.

date → must be specified as # dd-mm-yyyy #

Example 3-

will always give integer value

Datepart ('Year', #02-10-2015#) \Rightarrow Gives 2015

Datepart ('Month', #02-10-2015#) \rightarrow Gives 10

Datepart ('Quarter', #02-10-2015#) \Rightarrow Gives 4

Datepart ('dayofyear', #02-10-2015#) \Rightarrow 275

n ('day', n) \rightarrow 02

n ('week', n) \Rightarrow 40.

- To compare or get synchronize months of each year we have to use datepart only.

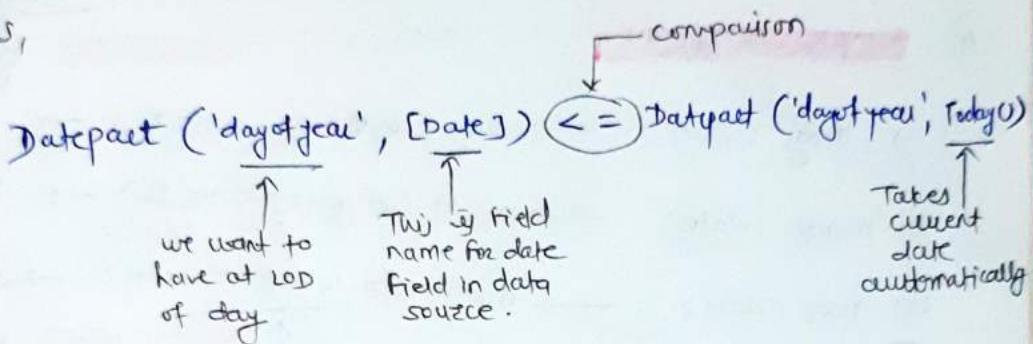
Example

our viz has \rightarrow two Year.

2020 & 2021
 \downarrow \downarrow
 All months Till June only.

- Now the condition is that we ~~do~~ need to see value of 2020 year upto June only & also for that particular date i.e. let say 02/06/2021
- And also we want to update it as our data gets refreshed i.e. now we have 04/06/2021 then same day of 2020 yr should also get updated.

- for this we have to use a calculation field
as,



- put this calc. in filter shelf & say only "True" values.

- We get sync. date for both year & will be updated automatically as we open on next few days after.

* Dynamic Date Values Function

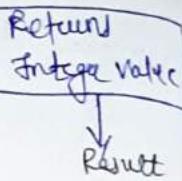
1) Today()

- Returns current date i.e. system date
`dd/mm/yyyy`

2) Now()

- Returns current date & time
• `ss : mm : hh dd/mm/yyyy`

* Date Functions :-



A) Simple Functions :-

- i) Day (date) → Day (#16-09-2015#) → 16
- ii) Month (date) → month (#16-09-2015#) → 9
- iii) Year (date) → month (#16-09-2015#) → 2015

These are hard coded values,
we can also add here fields
of date directly or a
parameter value.

- Similarly Datepart function can be used
→ Date field / Hard date / parameter

* Datepart (date-part, date, [start-of-week])

- can be 'Year', 'Quarter' → Returns (1 to 4)
- 'dayofyear' → Returns (1 to 365)
- 'week' → Returns (1 to 52)
- 'weekday' → Returns (1 to 7)
- 'Day' → Returns (1 to 31)
- 'Month' → Returns (1 to 12)

* DATENAME (date-part, date, [start-of-week])

- Just returns string instead of integer.

- To convert string to date or time

PICKED 1 AND 2

- **Makedate** (year, month, date)

eg. Make Date (2015, 04, 23)

Result → # 23/04/2015 #

- **Make Time** (hour, minute, second)

g. makeTime (02, 10, 30)

Result → # 02:10:30 #

- **Makedatetime** (date, time) → To concatenate date & time field in one column.
↑ ↑
datefield time field

- * **DATEADD** (date-part, interval, date)

- Adds interval to specified date-part of date.

g. DATEADD ('month', 5, #02-01-2012#)

Result ⇒ # 02-06-2012 #

- Add 5 month to date.

- * **DATEDIFF** (date-part, date1, date2, [startofweek])

- Returns difference b/w two dates.

eg. Datediff ('day', #02-01-2020#, #02-02-2020#)

Result → 31

FORECAST THEORY :-

- Tableau has integrated forecasting functionality.
- It is available directly in Analytics pane.
- Tableau uses exponentially smoothening system for forecasting.
- Forecasting cannot work on views using:
 - i) Table Calculations
 - ii) Disaggregated functions
 - iii) Percent calc.
 - iv) Grand Total or subtotal.
 - v) Date values with aggregation set to exact date.
- Forecasting exponential smoothening put more weight on the recent data.
- To overcome the some limitation of forecasting in tableau other models like 'ARIMA models' can be performed by integrating R into Tableau
- Past + Present events → Identify Patterns → Apply pattern to the future.
- Predictions are more accurate, the closer they are to the present.

* TREND & SEASONALITY

Trend :-

- A direction or pattern upwards & downwards in which data goes
- At least 5 observations are required to create basic model of forecasting with trend.

Seasonality

A pattern repeats itself after every interval of timespan.

- It may year patterns (Temp of earth)
- OR daily patterns (Parking facility)

Timespan \Rightarrow Level of Aggregation for data values.

(Year, Quarter, Month, Day)

- Requires atleast 2 seasonal cycles for basic trend.

* MODEL CLASS

- For Forecasting we can choose the model as an exponential smoothing & ARIMA (Auto-regressive Integrated moving average)
- R integration gives you freedom to create any model your data & your goals require.

Data Requirement for Forecasting :-

- Required data :- L measure , L Date dimension
- OR L measure , L integer dimension.
- A date dimension works best because it specifies clear order (intuitive).
- Integer dimension can be Day or Week no. an observation was made to put in.

* PARETO CHART

- It is a dual axis chart used to highlight dimension member that are having biggest impact to the measure in question-
 - It contains both Bar & a Line graph, where individual values are represented in descending order by bars, & ascending cumulative total is represented by line.
 - Pareto chart made an observation that 80% of effects come from 20% of the causes.
- Ex:- → 80% of profit derive from 20% of the available product
→ 80% of land was owned by 20% of the population
- Before preparing the chart first decide that what question we want to answer.

* How to Build Pareto Chart

- A) Built for state :-
 - i) Drag & drop state field in the column shelf & the sales field in row shelf.
 - ii) For Sales field → create table calculation → edit table calc → Running total → specific dim "state" → Add sec calc → % of total → specific dim "state"

iii) Then create calc. field As ~~or~~ name ~~no.~~ of state

Index() | size()

This is also a table calc. → what this do that
Index() gives us row no. of state from first to last
size() gives total no. of record of state available.
→ am saying state because we have to add state
later in detail card, → The dim present in detail
card will be responsible for Index & size calculation.

iv) Drag & drop of state field in the column shelf
→ edit table calc. → compute using \rightarrow "state".

v) Then drag state already present in column shelf to
the details card.

Tip
vi) If graph shows weird path → set the state
field on sum (sales) \rightarrow descending.

vii) Now we get % of state count on x axis (cumulative)
or Running Total % of sales running % on y axis.

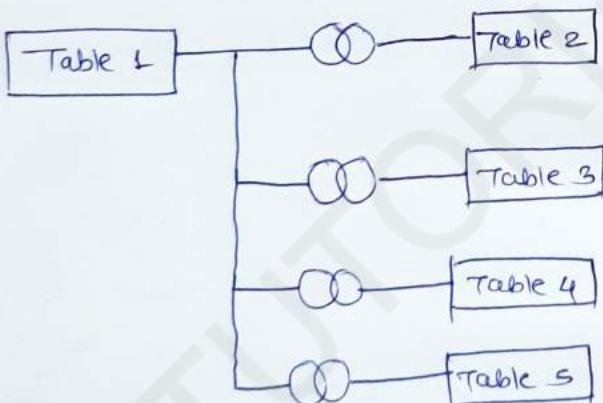
viii) Add sum sales to row shelf \rightarrow make chart
dual axis & for **sum (sales)** \rightarrow select chart type
as Bar chart.

ix) Then add set line for 80% & 20% on x &
y axis.

* JOINS & SCHEMAS

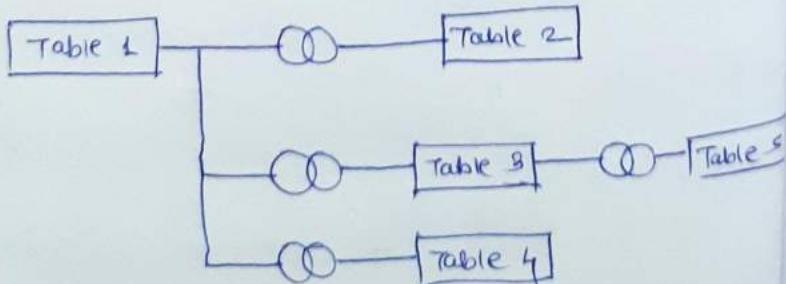
- We know that joins can be done in physical layer of the data tab.
- Joins can be done to generate the schemas.

Example



The above is an Example of "STAR SCHEMA" & there is a fact table "Table t" at centre & all other Table are joined to Table t.

Example



Above is an example of a "SNOWFLAKE SCHEMA", since there is a fact table surrounded by dim table & intun they are surrounded by other table.

→ MAP TYPES :- CUSTOM MAPS

- 1) Hexagon Maps
- 2) Tiled Maps -
- 3) Area Tiled Maps

→ MakeDate Function :-

Syntax :- makedate (year, month, date)

- Returns date value constructed from specified year, month & date.

makedatetime (date, time) → combines date & time.

- The above are mostly used when our year, month are in string datatype in different columns i.e. year & month column, to convert them to date we use the makedate() function.

- If year & month is in string, convert to number & then use the above function.

HORIZON CHART

- This is chart which is an alternative used for a multiple Area Chart.
- These are mainly used to show the measure by breaking it into the group of say (0 to 20), (20 to 40) to 100 in percentage over a particular dimension like country, state or category.

RADIAL BAR CHART

* YTD, MTD, STD

YTD :- YTD is period starting from beginning of the current year & continuing upto present day.

- Present day can be any day of that particular year
e.g. 10 Aug. 2021 → If Year we want is 2021.

STD :- Similar to YTD is also the sales for the Quarter for which our present day is

Q1	Q2	Q3	Q4
Jan Feb Mar Apr	May Jun July Aug		
Jan Feb Mar	Apr May June	July Aug Sep	Oct Nov Dec

Eg. our present day is 10 Aug, then we get sales for quarter Q3 for that particular year.

MTD :- MTD directly gives sales for the given month from the start of month to present date of same month.

* Two ways

1) Date field → Year, Datepart, Month.

2) Using DateTrunc Function.

YTD 2

↓ Take field name

If ~~[Date]~~ <= Today() ← condn must satisfy i.e.
Today must be ↑ than Date
field values.

And

$\lceil \text{Year} ([\text{Date}]) = \text{Year} (\text{Today}()) \rceil$ then sum ([sales])
End.

OR

→ OR

$\text{Datepart} ('year', [\text{Date}]) =$
 $\text{Datepart} ('year', \text{Today})$

IF [Date] <= Today()

And

[Date] >= DateTrunc ('year', Today()) then ---

End

OR

IF DateDiff ('year', [orddate], today()) = 0

then sum([sales]) End.

only sees the
Year of
order date.

→ what this does that is gives
TID of sales irrespective of the
todays date i.e. no condition of
todays dates => than orddate.

Prior YTD
If DateDiff ('year', [orddate], today()) = 1

then sum([sales]) End.

→ This gives TTD of full prior year.

YTD

If ~~[orddate]~~ <= today() And
DateDiff ('year', [orddate], today()) = 0 → This cond
allow to get
till date TTD
then [sales]
End.

GTD

If [Date] <= Today()

And

Datpart ['quarter', [Date]] = Datpart ['quarter', Today()]

And

Year([Date]) = Year(Today())

Then [sales]

End.

[OR]

If [Date] <= Today()

And

Datediff ['quarter', [Date], Today()] = 0

then [sales] End.

NOTE * → with Datediff() function by
 $=1 \rightarrow$ we get prior YTD
 $=2 \rightarrow$ prior to prior YTD
 & so on.

* With Datpart() we have add Year(Date) = Year(Today())
 because Year we have specify of which year we need to
 find GTD because if not specified it calculates GTD for
 all year of ~~Today()~~ quarters of all years.

Hence mostly try to use Datediff.

MTD

JF

[Date] <= Today()

And

Datediff ['month', [Date], Today()] = 0

then

[sales]

For Prior

use = 1

For Previous Year Same month

use = 13

* For Last 6 months MTD

* To find Last N months MTD

Datediff ('month', Datetunc('month', [orderdate]),

{max(Datetunc('month', [orderdate]))}) < [N]
How many months

This Datetunc gives the months of all year in the data available.

→ This gives the current month or the (LTD) max month available.

HISTORY YTD & STD

- If we require to determine the YTD, STD --- for the HISTORY data or dynamic data then it can be done by using parameter control.
- Set Parameter to 'Date' datatype
if select a proper date initially i.e. default date to the parameter -
- Then while adding into the calculation we can get the calculation as -


```

if [Orderdate] <= today() → substitute here parameter
DateDiff('year', [Orderdate], parameterdate/parameter) = 0
      
```

 then


```

[sales]
      
```

 End.
- YTD & MTD, STD we can also get for the different measures like 'Profit', 'Sales' by using parameter control for these if substituting that parameter calculation in [sales]

E let Parameter Name \Rightarrow "Dynamic Select Measure"
 \Rightarrow if have ~~String~~ value list as
 'Sales' & 'Profit'

Calculation Field \Rightarrow Dynamic Measure

\Rightarrow Case [select Measure]
 when 'Sales' then [sales]
 when 'Profit' then [Profit]
 End.

Put this [Dynamic measure] in [sales] place in above calculation

* LYTD :-

- When we use YTD ie. Year to Date sales , this is for the current year data ie. for eg.

Today() = 16-May-2021

- Then YTD calculates e.g. [sales] till 16-May-21 from 01-Jan-21

- What if we want the same [sales] over date from 01-Jan-20 to 16-May-20 i.e. of last year to compute the YoY ie. (Year over Year) growth percentage.

- This last year to date (LYTD) can be obtained in tableau to get YoY in the analysis dynamically how growth is happening.

How to calculate LYTD :-

If

Year([orderdate]) = Year(today()) - 1

And

[orderdate] <= Dateadd('Year', -1, today())

then

[sales]

End

OR other formulae

If

year ([orderdate]) <= Dateadd ('Year', -1, today())

And

Datediff ('Year', [orderdate], today()) = 1

then

[sales]

end.

+ **LSTD**

If

[orderdate] <= Dateadd ('Year', -1, today())

And

Datediff ('quarter', [orderdate], today()) = 4

then

[sales]

end.

[OR]

If

[orderdate] <= Dateadd ('quarter', -1, today())

and

Year ([orderdate]) = Year (today ()) - 1

and

month ([orderdate]) <= month (today ())

and

Quarter ([orderdate]) = Quarter (Today ())

then

[sales] end.