**Tableau-TDS**

**Creating live connection to data source:**

**Explain the differences between using live connections versus extracts**:-

Live and extracts are two ways you can make the data connection to the tableau. Live allows you real-time data while extracts are kind of batch which needs to be refreshed from time to time to get the updated data.

Live:- Direct connection between tableau and data source.

Downside-

1. Time taking usage
2. Downgrade in performance
3. Big data cannot be prepared

Upside- Real-time data analysis

Extract:-Makes a local copy of data source or snapshot of data source.

Adv-

1. Tableau is no longer connected to data source
2. Columnar data storage
3. High performance
4. High storage value

**Note:-**In Tableau use **Filters -> Records** to Improve Performance (Particularly **Context filters**).

Filters:-

Data source filters:- Filter applied both in live and extract.

Extract filters:- Only applied on extract.

**Tableau Order of Operation:-**

Top to bottom priority

EXTRACT FILTER

DATA SOURCE FILTER

CONTEXT FILTER

DIMENSION FILTER

MEASURE FILTER

TABLE CALCULATION FILTER

Note :- In tableau **Blue Pill-Discrete and Green Pill- Continuous.**

**Aggregate, Join, or Union Data**

**Aggregate and Group Values:-** When the data is in Granular Level ,Sometimes you’ll need to adjust the **granularity** of some data, either to reduce the amount of data produced from the flow, or to align data with other data you might want to join or union together. Aggregate option will help in arranging the some data into groups , this depends on the data type( String, number and dataset).

Fields are distributed between the **Grouped Fields** and **Aggregated Fields** columns based on their data type. Click the group or aggregation type (for example, AVG or SUM) headings to change the group or aggregation type. In the data grids below the aggregation and group profile, you can see a sample of the members of the group or aggregation. Any cleaning operations that are made to the fields are tracked in the **Changes** pane.

**Join:-** The data that you want to analyse is often made up of a collection of tables that are related by specific fields. Joining is a method for combining the related data on those common fields. The result of combining data using a join is a table that’s typically extended horizontally by adding fields of data.

Joining is an operation you can do anywhere in the flow. Joining early in a flow can help you understand your data sets and expose areas that need attention right away.

**Tableau supports the following join types:**

|  |  |
| --- | --- |
| **Join Type** | **Description** |
| Left https://help.tableau.com/current/prep/en-us/Img/prep_leftjoin.png | For each row, includes all values from the left table and corresponding matches from the right table. When a value in the left table doesn't have a corresponding match in the right table, you see a null value in the join results. |
| lnner https://help.tableau.com/current/prep/en-us/Img/prep_innerjoin.png | For each row, includes values that have matches in both tables. |
| Right https://help.tableau.com/current/prep/en-us/Img/prep_rightjoin.png | For each row, includes all values from the right table and corresponding matches from the left table. When a value in the right table doesn't have a corresponding match in the left table, you see a null value in the join results. |
| Left Only https://help.tableau.com/current/prep/en-us/Img/prep_leftonlyjoin.png | For each row, includes only values from the left table that don't match any values from the right table. Field values from the right table show as null values in the join results. |
| Right Only https://help.tableau.com/current/prep/en-us/Img/prep_rightonlyjoin.png | For each row, includes only values from the right table that don't match any values from the left table. Field values from the left table show as null values in the join results. |
| Not Inner https://help.tableau.com/current/prep/en-us/Img/prep_noninnerjoin.png | For each row, includes all of the values from the right and the left table that don't match. |
| Full https://help.tableau.com/current/prep/en-us/Img/prep_fulljoin.png | For each row, includes all values from both tables. When a value from either table doesn't have a match with the other table, you see a null value in the join results. |

**Union:-**

Union is a method for combining data by appending rows of one table onto another table. For example, you might want to add new transactions in one table to a list of past transactions in another table. Make sure the tables you union have the same number of fields, the same field names, and the fields are the same data type.

**Similar to a join, you can use the union operation anywhere in the flow.**

**Note:-** To maximize the performance a single union can have a maximum of 10 inputs. If you need to union more than 10 files or tables, try unioning files in the Input Step.

## **DIFFERENCE BETWEEN TABLEAU RELATIONSHIPS AND JOINS**

**RELATIONSHIPS**

* In Tableau, relationships are called the Logical layer and Joins are called the physical layer.
* Relationships are not Joins. Relationships keep the tables separate and do not merge them into a single table.
* In the earlier versions of Tableau, in the absence of a relationships, this task would have required **data blending** with multiple sources and would therefore be slow. Using relationships, we can do this in one data source. Also, data blending is limited to a worksheet, but relationships are available for the entire workbook. Relationships will increase the performance of your dashboard.

**JOINS**

* A Join between two tables merges the tables and creates a single table.
* Joins between two tables are created based on the common field names. To create a join, you need to double click on the table in the logical/relationship layer and add another table next to it.

**Compared to Relationships, Joins have some disadvantages**

1. Joins are static and once made, will affect the data in the entire workbook.
2. Depending on the join type you use, you may lose unmatched data
3. When using left or right join, the unmatched data will not be shown.This problem can be resolved by using outer join, but here again, if your tables are huge, this may result in slow performance. Due to this reason, Outer joins are not recommended.

**Note:-**All these problems will be resolved by using Relationships. You can still use joins but use them when absolutely necessary.

**18-05-2022**

## **Types of Charts and Graphs**

### **Bar Chart**

Bar charts are one of the most common data visualizations. You can use them to quickly compare data across categories, highlight differences, show trends and outliers, and reveal historical highs and lows at a glance. Bar charts are especially effective when you have data that can be split into multiple categories.

For Bar charts-

0 or more Dimensions

1 or more measures

### **Line Chart**

The line chart, or line graph, connects several distinct data points, presenting them as one continuous evolution. Use line charts to view trends in data, usually over time (like stock price changes over five years or website page views for the month). The result is a simple, straightforward way to visualize changes in one value relative to another.

For line chart-

1 date

0 or more Dimensions

1 or more measures

### **Pie Chart**

Pie charts are powerful for adding detail to other visualizations. Alone, a [pie chart](https://www.tableau.com/data-insights/reference-library/visual-analytics/charts/pie-charts) doesn’t give the viewer a way to quickly and accurately compare information. Since the viewer has to create context on their own, key points from your data are missed. Instead of making a pie chart the focus of your dashboard, try using them to drill down on other visualizations.

For pie chart-

1 or more Dimensions

1 or 2 measures

### **Maps**

Maps are a no-brainer for visualizing any kind of location information, whether it’s postal codes, state abbreviations, country names, or your own custom geocoding. If you have geographic information associated with your data, maps are a simple and compelling way to show how location correlates with trends in your data.

For maps

1 geo dimension

0 or more Dimensions

1 or more measures

### **Density Maps**

Density maps reveal patterns or relative concentrations that might otherwise be hidden due to an overlapping mark on a map—helping you identify locations with greater or fewer numbers of data points. Density maps are most effective when working with a data set containing many data points in a small geographic area.

For density maps-

1 Geo

1 or more Dimensions

1 or 2 measures

### **Scatter Plot**

Scatter plots are an effective way to investigate the relationship between different variables, showing if one variable is a good predictor of another, or if they tend to change independently. A scatter plot presents lots of distinct data points on a single chart. The chart can then be enhanced with analytics like cluster analysis or trend lines.

For scatter plot -

0 or more Dimensions

2 or 4 measures

### **Gantt Chart**

[Gantt charts](https://www.tableau.com/learn/articles/how-to/gantt-chart) display a project schedule or show changes in activity over time. A Gantt chart shows steps that need to be completed before others can begin, along with resource allocation.

For gannt chart-

1 date

1 or more Dimensions

0 or 2 measures

### **Bubble Chart**

Although bubbles aren’t technically their own type of visualization, using them as a technique adds detail to scatter plots or maps to show the relationship between three or more measures. Varying the size and color of circles creates visually compelling charts that present large volumes of data at once.

For bubble chart-

1 or more Dimensions

1 or 2 measures

### **Treemap**

Treemaps relate different segments of your data to the whole. As the name of the chart suggests, each rectangle in a treemap is subdivided into smaller rectangles, or sub-branches, based on its proportion to the whole. They make efficient use of space to show percent total for each category.

For more types of charts, visual examples, tips, and information, download our whitepaper. In this paper, you’ll learn about different chart (and graph) types—from bar charts to density maps to box-and-whisker plots. You'll also learn when to use one chart over another, along with tips on how to leverage these chart types for maximum impact.

For tree map-

1 or more Dimensions

1 or 2 measures

**Default properties of Data field:-**

The default properties men includes default settings for aggregation, comments, number formatting ,colour, shape and total( based on the type of field)

**Data types of Data field:-**

We saw that there are seven basic data types in tableau namely, String, Numeric( Whole and Decimal), Boolean, Date, Date & time, Geographic and mixed or cluster.