**TABLEAU INTERVIEW QUESTIONS ASKED ME**

**What is the maximum limits for Tableau**

File size: unlimited, there is no practical file size limitation. Customers have created extracts of multiple terabytes but the best practice is to keep everything as small as possible coming into Tableau.

Columns coming into Tableau: unlimited

Rows: 10 million

Tables joined: 32

Columns in a worksheet: default is 6 for each column and row shelve but can be increased to 16

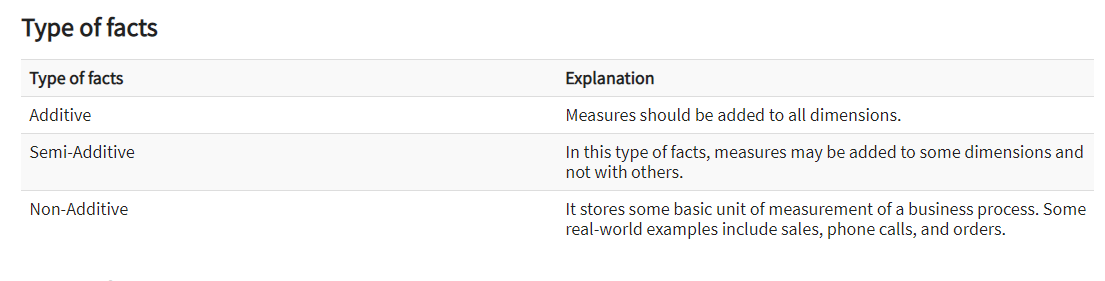
* Analysis-Table Layout- Advanced

**How do you Optimize a Tableau dashboard for performance?**

**How do you set up security in Tableau?**

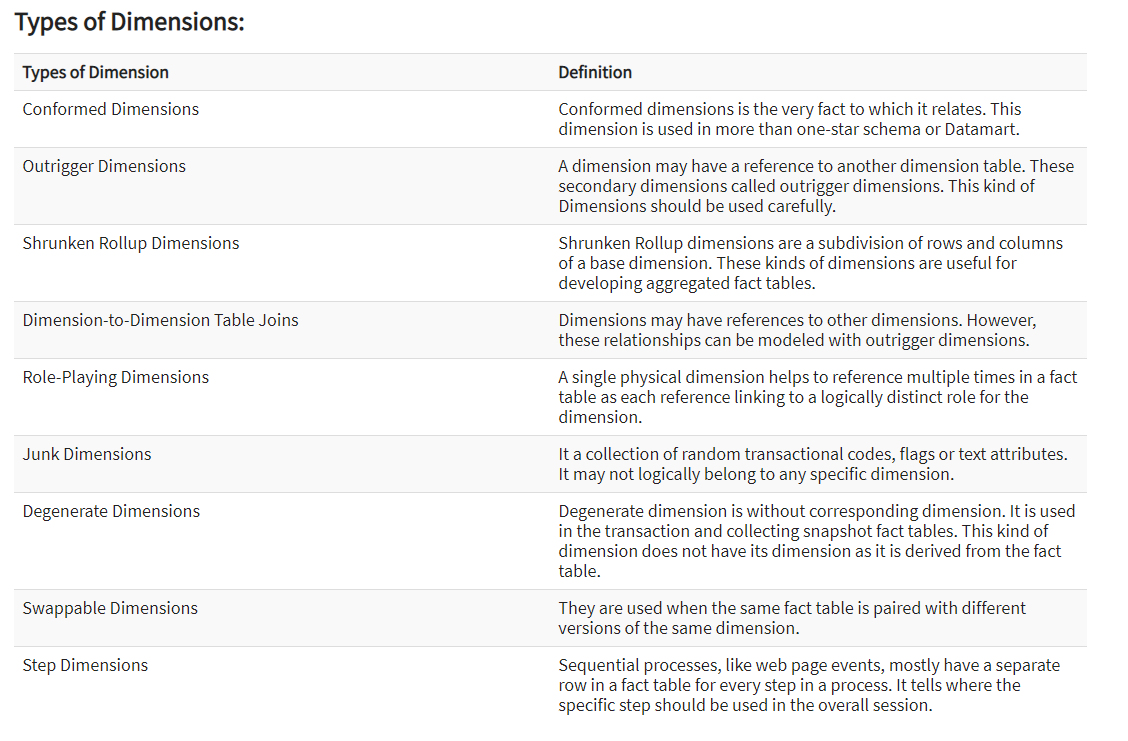
**What are Reference Lines?**

**What type of Fact Tables are there? Explain**



**What type of dimensions are there? Explain**

* Slowly Changing Dimension
  + i.e. name or address can change slowly over a period of time.
    - Type 1 New value replaces the old value. No history record is kept
    - Type 2 A new row is added and the old history record is dated
    - Type 3 Add two new columns one for old and one for new
* Junk
  + Is used to avoid having a large number of foreign keys in the Fact Table
  + A single table with a combination of different and unrelated attributes
  + Created to manage Foreign keys created by rapidly changing dimensions
* Conformed
  + Helps to create consistent reporting across the enterprise
  + Dimension that is shared among multiple Fact tables/Data Marts and has the same meaning, i.e.
    - Time Dimension
    - Product Dimension
    - Staff Dimension
* Degenerate Dimension
  + Also called Fact Dimension
  + Constructed from columns in the Fact table instead of the dimension table
  + Useful dimensional data is sometimes stored in a fact table to reduce duplication
  + i.e. A Sales Fact table that contains a Purchase Order field and a carrier tracking number field in the Fact table
  + Degenerate dimensions are frequently used to support drill through actions because the drill through action in SSAS requires that you select the attribute from a dimension
* Reference Dimensions
* Role Play Dimension
  + A dimension table the has multiple valid dimensions with a Fact table
  + A Fact table may include foreign keys for both ship date and delivery date. The same dimension attributes apply to each foreign key so the same dimension tables can be joined to the foreign keys



**What are parameters?**

* Dynamic values that can replace constant values in calculations, filters and reference lines
* Parameters are static
  + The range of values are static
  + Needs to be used in calculation(calculated field), filter or reference line
* Parameter values
  + Can be defined by the desktop user
  + Entered as an input from the report consumer
  + Populated with the values of a field from the data source
* Parameter allowable values
  + All Value Parameter
  + Range parameter
  + List parameter
    - You must specify the list of values
* After you create your parameter you can use it becomes available for a value in your filter values drop down

**PARAMETER STEPS**

* Create parameter with list of values
* Create Calculated field
* Use parameter in calculated field
  + CASE [**Parameter name “select Dimension”]**

WHEN “Customer Name” THEN [Customer Name]

WHEN “State” THEN [State]

WHEN “Subcategory” THEN [Subcategory]

END

* + CASE [**Parameter name “select Measure”]**

WHEN “Sales” THEN SUM([Sales])

WHEN “Profit” THEN SUM([Profit])

END

**Do parameters allow for a drop-down list?**

* Yes but you must specify the list manually or use All
* The list is single valued. You are not able to choose multiple values
* The List can be used in a case statement in a Calculated field

**What is the difference between filters and parameters?**

**What is the difference between a regular calculation and a Table calculation?**

* A regular calculation i.e. Sales-Profit is from the data source
  + Computation is handled by the data source
* A table calculation is represented by the delta symbol ^ on the pill
  + A secondary calculation done on top of the result set
  + Computation is done by Tableau
  + See Pre defined Quick table calculations. Which are available depends on the data on your view
    - Running Total
    - Difference
    - % of Total
    - Rank
    - Percentile
    - Moving Average
    - Compound Growth Rate
    - YTD Total
    - Year Over Year Growth
  + Or you can manually create a Table calculation
    - Create calculated field
    - Choose Table Calculation on drop down
  + A Table calculation ca be dragged into the Measures shelf for future use

**What is the difference between Live and extract connections?**

**Live:**

* Real time updates as the data is updated
* Performance is only as fast as your database

**Extract:**

* Snapshot of the data
* Allows for more functionality of Tableau
* Faster performance than live
* Best practice
* Extract is a .hyper file (used to be .tde)

**Filter Shelf possible filters (independent of each other)**

* Dimension Filter
* Measure Filter
* Date Filter
* Conditional Filters
* Top/Bottom Filters

**What is a context filter**

* When you have multiple filters in the filter shelf they are independent of each other.
* To make the filters dependent on each other choose “Add to context”
* The higher dimension in the Hierarchy gets filtered first.
  + i.e. Category – Top 10
  + i.e. Furniture – Top 10
* **Context filter** is located on the filter shelf. It allows you flexibility for filtering the data even further such as
  + TOP N
  + **Improves performance** − If you set a lot of filters or have a large data source, the queries can be slow. You can set one or more context filters to improve the performance.
  + **Creates a dependent numerical or top N filter** − You can set a context filter to include only the data of interest, and then set a numerical or a top N filter.
  + Right click and choose “Add to context”
  + Tableau creates a table for this context filter

**. What is the disadvantage of Context Filters?**

* The Context Filter is not frequently changed by the user—if the Filter is changed, the database must be recomputed and the temporary table has to be rewritten, slowing performance.
* When we set a dimension to context, Tableau creates a temporary table that will require a reload each time the view is initiated. For Excel, Access, and text data sources, the temporary table created is in an Access table format. For SQL Server, MySQL, and Oracle data sources, we must have permission to create a temporary table on our server. For a multidimensional data source, or cubes, temporary tables are not created, and Context Filters defined which Filters are independent and which are dependent.

**Extract Filter**

* In the data source sheet the top right is the Extract filter
* Allows you to narrow the data source in Tableau
* Can be used anywhere without any connection to database

**Normal/Basic Filters**

* Dimension Filters
* Measure Filters
* Date Filters

**Global Filter**

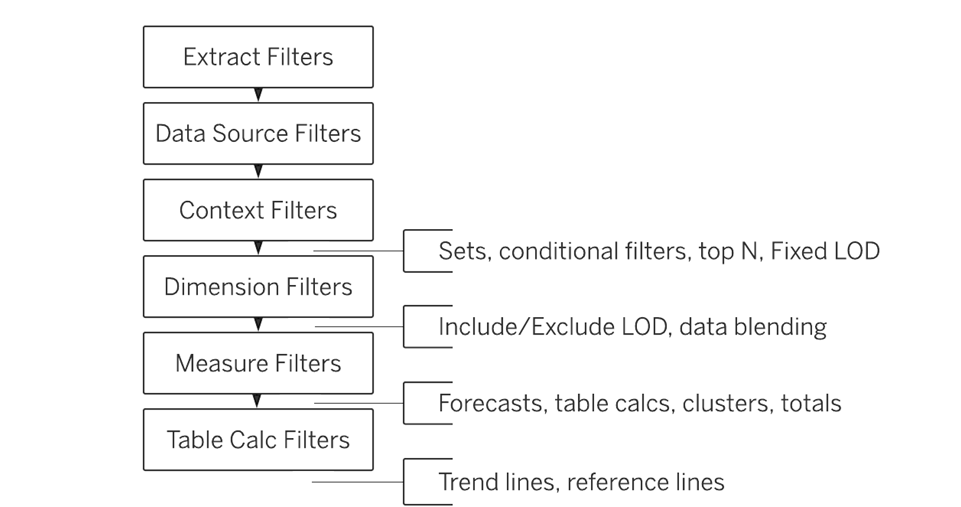
* Right click on filter name and choose all using this data source

**User Filter** (for user level security)

* Sign into the server
* Right click on server and choose Create User Filter

**Filters Order of Operations**

* Extract Filters
  + Used only if the user extracts the data source
  + After choosing the extract radio button then choose edit to bring up the extract data filter window
* Data Source Filters (Global filter)
  + On the data source sheet on the top right choose add filter
* Context Filters
  + Filter shelf dependent filters
* Set Filters
* Dimension Filters
* Measure Filters
* Table Calc Filters



**Dimensions**

* By default, dimensions are discrete(blue)
* definable

**Measures**

* By default, measures are continuous(green)
* infinite

**File types**

.twb – Metadata only. Analysis cannot be performed

.twbx – Metadata and data source. Analysis can be performed

### ****5. What is the latest version of Tableau Desktop?****

Tableau Desktop 2019.4 (as of November 6, 2019)

### ****What is the difference between heat map and treemap?****

* A heat map is a great way to compare categories using color and size. In this, we can compare two different measures.
* A treemap is a very powerful visualization, particularly used for illustrating hierarchical (tree-structured) data and for visualizing a part of or a whole relationship.

**What is a Marks Card**

* The area where you drag the fields to control the mark properties such as type, color, size, shape, ect

**What are groups**

* Used to combine dimensions into higher level categories
* You can group dimensions of several states into a region
* Marked with a paper clip in the data pane
* Is one dimensional and used to create a higher level category by using lower level category members
  + Certain Sub-categories can be grouped by category
  + Group stays the same even if numbers change
  + Cannot be used in calculated fields

**SET**

* Can have conditions and can be grouped across multiple dimensions/measures
  + Top sales and profit can be clubbed together for different categories by creating a set
* If numbers change the set members change
* Can be used in calculated fields
* Set symbol is shown as a double ring
* Constant set
  + Value does not change
* Computed set
  + Values are dynamic and can change
* Combined sets
  + Two related sets based on the same dimension

**What is a BIN**

* To create a histogram, you must first create a bin
  + When you choose a histogram in show me it automatically creates bins
* Organize values of a measure into Bins
* i.e. You have a measure that holds the ages of customers from 18 to 90
  + if you want to analyze the data by age groups you would use bins

**How do you test your work in Tableau?**

* Build a table chart and use the filter shelf to verify the results of your formulas

**SQL QUESTIONS**

**What are DDL and DML statements**

**What are the limitations on views? (i.e. update, insert, delete)**

**What is the difference between a Unique Key and Primary Key**

Unique Key:

* Allows for NULLS

Primary Key:

* Does not allow for NULLS
* Only one primary key per table

**What is a foreign Key?**

**What is the difference between RowNumber Over and Rank Over**

**Can RowNumber Over do the same as Rank Over**

**I have a table named Employee with EmployeeID, StartDate, EmployeeName, EmployeeGender, EmployeeState**

* The same employeeID is duplicated because the employee had several start dates
* How can I retrieve only the latest start date along with other columns

**Give an example of a full outer join**