Tableau Fundamentals course notes

Contents

1 Tableau Fundamentals

1.1 Connect to and Customize Data

1.1.1 Connecting to data

Live connections vs. extracts:

- Connecting live, which is the default connection, leaves the data in the database or source file. This is best when you want to leverage a high performance database's capabilities, or to get up-to-the-second changes in data visualized in Tableau

- You can alternatively choose to extract the data into Tableau's high performance in-memory data engine. This option is useful when you connect to a slow database or when you want to take query load off critical systems

The types of flat files you can connect to include Microsoft Excel files, text files, JSON files, PDF files, and more. Once you've connected to a data source, Tableau opens to a workbook with the data loaded. This is the Data Source page. Metadata describes your data. It includes the names of fields and data types as well as aggregation information and default display properties. Tableau assigns metadata to the data that you bring into Tableau; you can edit the metadata in a way that better supports how you want the fields to look and function when you're building a view. If you need to reference something quickly or wish to see some of the name changes that you've made, click the Manage metadata button to switch to the metadata grid view. The metadata grid displays the fields in your data source as rows so that you can quickly examine the structure of your Tableau data source and perform routine management tasks, such as renaming fields or hiding multiple fields at once.

1.1.2 Customizing a Data Source

When analyzing data, you may need to do some cleanup and organization as you work with the underlying source data. This could include customization like connection information, organizational or metadata changes, attributes, or aliases. Note that Tableau preserves the customization you make, but it does not change the underlying source data.

Tableau automatically predicts whether your fields are dimensions or measures, but sometimes you may want to re-categorize a dimension as a measure or vice versa. To make the field names more intuitive, Tableau allows you to rename any of the database fields. To clarify the meaning of the labels associated with discrete dimension members in the view, Tableau allows you to create aliases, or alternate names. You can edit a field's default properties, such as color, number format, or aggregation. This way, every time you add the field to the view, it will maintain its assigned properties.

1.2 Organize Data and Create Filters

1.2.1 Creating Groups in Your Data

A group lets you combine several members of a single dimension into a single data point or category type, creating a new dimension field that didn't originally exist in your data (Groups in Tableau are represented by a paper clip). Groups can be created using each of the following methods:

- Create a group based on a field in the Data pane
- Create a group based on labels in the view
- Create a group based on marks in the view

Once a group as been created within the Create Group dialog, you can check Include 'Other' to group all remaining members into a group. The group will be assigned the name Other and can be renamed as needed.

1.2.2   Creating Hierarchies in Your Data

Organizing fields in your data into hierarchies can help you analyze how they relate to one another. In Tableau, a hierarchy is an arrangement of data fields in a hierarchical format with an "above" and "below" structure. When a hierarchy exists in your data, you can see how fields relate to one another in tiers, view the structure of it, and access the data in the shelf and within the visualization. One of the most useful ways to navigate hierarchies is to drill down using the [+] or drill up using the [-] to hide or display the related fields.

1.2.3   Understanding Filtering in Tableau

Filtering does not remove or modify data, it just changes the data that appears in your view. When you connect to data, you can choose to extract it, saving a snapshot of how it looks in your workbook and ultimately reducing the number of times Tableau queries the data source. As part of this process, you can filter out data that isn't relevant to your analysis.

Data source filters reduce the amount of data being fed into Tableau and restrict what data the viewer sees. With certain access rights, the viewer can view all of the underlying data in a data source. If sensitive data hasn't been excluded from the data source, use data source filters to filter it out. For systems that rely heavily on partitions or indexing, data source filters may yield tremendous control over the performance of queries issued by Tableau.

All filters in Tableau are computed independently. If you want one filter to be applied before other filters, make it a context filter so it will be processed first. When you use multiple filters on your data, the filters get applied in a very specific order, called the order of operations (Figure 1).
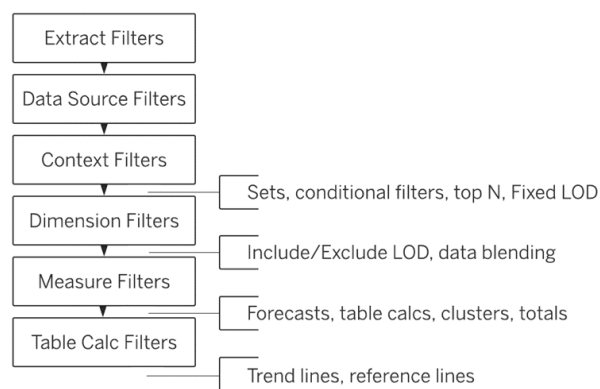
Figure 1: Tableau filtering order

1.3   Filtering Your Data

A dimension filter restricts categorical data in your view. The members present in a dimension can be included or excluded using this filter. The easiest and the most basic way of filtering dimension data on a view is to include or exclude data points, also referred to as marks, from your view. When you use the Keep Only and Exclude options, dimensions are automatically added to the Filters shelf. Another way of creating a dimension filter is to drag the dimension field from the Data pane to the Filters shelf. When you do this, the Filter dialog box appears which allows you to specify the type of filter you want to create. If you no

longer want a filter in your view, you have the option to remove it or hide it from the view. You can also remove a filter by dragging it out of the Filters shelf.

- You can create a quick filter by right-clicking the field in the Data Pane and selecting Show Filter. The field will automatically appear in the Filters shelf and the view. You cannot do this in Tableau Online;
- By default, a filter only applies to the worksheet that it is created in. But you can change that by right-clicking the filter in the Filters shelf, and selecting Apply to Worksheets;
- Use all allows you to specify that all members are used as an input to the filter. This means that if the list of members change, the list will update to reflect the changes. Select from list limits the filter to the members you select initially.

1.3.1 Using Sets to Highlight Data

List of Figures

List of Tables