Tableau

What is a field?

→ A field, also known as column, is a single piece of information from a record in a dataset, i.e. column names in a dataset is called field in tableau.

Now fields can be quantitative or qualitative.

1, 2, 3, etc.   'Bus', 'Auto', 'Person', etc.

In tableau, quantitative fields are referred to as Measures, and Qualitative fields are referred to as Dimensions.

Qualitative field :- (i) Describes or categorizes data. (Describes categories of data)

(ii) Tells you what, when, or who.

(iii) Slices the quantitative data.

Quantitative field :- (i) Numerical data. (Measure categories of data)

(ii) Provides the measurement for qualitative category.

(iii) Can be used in calculations.

Dimensions and measures are building blocks of Tableau charts.

On visualising a dataset of wildlife strikes which is a measure the Measures of the dataset, we have found that (i) as we added more dimensions to the view, the single bar representing all of the wildlife strikes was sliced into smaller categories of data.

(ii) Using the filter allowed us to focus on a specific subset of data.

Moving dimensions and measures in and out of the view changes the resulting chart. It is useful to view different aspects of data.

"What does a row of data contain?" is a simple but important question, and it can have complex answer.

Why does knowing data granularity matter?

→ Data granularity refers to the level of detail for a piece of data, wherever you are looking. As data become less granular, we might describe it as an aggregation, or as an aggregated data. Aggregation refers to how data is combined.

The level of granularity or aggregation in a row or chart affects the questions we can ask of the data, and the discoveries we can make.

If your questions reach the the limits of your data:

(i) Raise data quality issues with your Data Steward.

(ii) Describe the data questions you hope to explore and answer.

(iii) Provide detailed feedback so that others can develop a plan to collect more and better data.

Measures placed in a view are aggregated by SUM, which means that the data for that field in all of the rows is combined.
& Measures can also be aggregated as average, median, count or count distinct.

Dimensions break down the aggregated total into smaller totals by category.

What is an aggregate function? What is granularity?

→ Data is generated and analyzed at many different levels of granularity. Granularity is the level of detail of the data. For example, when looking at graduation data, granularity would describe whether a row in the data set represents a single person or the graduating class of a university.

What if we wanted to identify a value for the graduating class as a whole – that is, decrease the granularity?

Aggregation and granularity are complementary concepts. Aggregation is a mathematical operation that takes multiple values and returns a single value: operations like sum, average, count, or minimum. This changes the data to a lower granularity.

The field or fields that determine what makes up a row are the grouping fields (in Tableau Prep).

COMBINING AGGREGATED DATA

Oftentimes, data at different levels of granularity needs to be combined. For example, to give a comprehensive way for geospatial analysis, city-level data often needs to be combined with country-level data.

Joins are a common way to combine data. But if we perform a join across different levels of granularity, we will inevitably introduce issues of data replication an or loss.

If the data on joining two tables, the data gets repeated then, we sum a particular field with repeated data then we would g.get the sum of multiplied multiplied; that may require us to average+ perform a diff. aggregation (such as average sum or min), or use an LOD expression to fix the number of ants per year, thus preventing accidental over-counting :

{FIXED [Field1] : MIN([Field2]) }.

Granularity can be increased by adding rows in a table.

How data is represented in Tableau?

→ Each field in the Tableau Data pane has an icon to the left showing the field's data type. Fields are assigned both a data type and role automatically when you're connect to data in Tableau. We can manually change the type in the data grid or Data pane.

⅄ **Country, State, City** } Hierarchy → The Country, State, and City fields have been organized into a hierarchy. This useful for creating drill-downs in charts.

Abc **Customer Name** } Blue text field → Customer name is a categorical, or, qualitative, field. A blue icon indicates that the field is discrete, which means it's data that contains separate parts.
Most dimensions are discrete fields.

=T|F **Market Selector** ! } Invalid field: The red exclamation point on this calculated field indicates a problem. The It can mean a field is missing, or, in this case, that the calculated field is broken.
(in red)

📅 **Ship date** } Date: The ship date icon is a date field. Date can be discrete or continuous. The date in this image is blue, which tells us it is discrete.
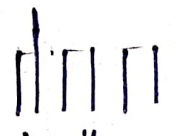
=# **Profit Ratios** } ~~Calculated value field:~~ Profit Ratios is a calculated field. The green icon indicates it is a continuous field.
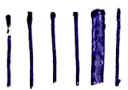
⊕ Latitude (generated)
⊕ Longitude (generated)
=# Number of records
Measure values
} Tableau generated fields: Sometimes they are added to the view automatically, depending on the type of the chart you are building.
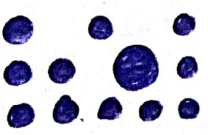
📅🕐 → Discrete data and time field.

Tableau already have built-in capabilities to spot light and dark colors, large and small shapes, groups and orientations of objects. These are referred to as pre-attentive attributes.

See which attributes work best with given data types:
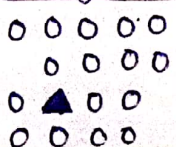
1.	2.
 length	 Width
Quantitative? Excellent Qualitative? Poor	Quantitative? Limited Quantitative? Poor

3.	4.
 Orientation	 Size
Quantitative? limited Qualitative? Poor	Quantitative? Limited Qualitative? Poor.

5.	6.
Enclosure	Shape
Quantitative? Poor Qualitative? limited	Quantitative? Poor Qualitative? Excellent.

7. ● ● ●

Quantitative? Excellent

Qualitative? Poor

Position

8. ● ●
    ●   ●●●
    ●   ●●●
            ●●●

Quantitative? Poor

Qualitative? Excellent

Grouping

9. O O O O
    O O O O
    O O O O
    O O O O
    O O **O** O
    O O O O

Quantitative? Poor

Qualitative? Excellent

Color Hue

10. ◍ ◍ ◍
     ◍ ◍ ●

Quantitative? Limited

Qualitative? Poor

Color intensity
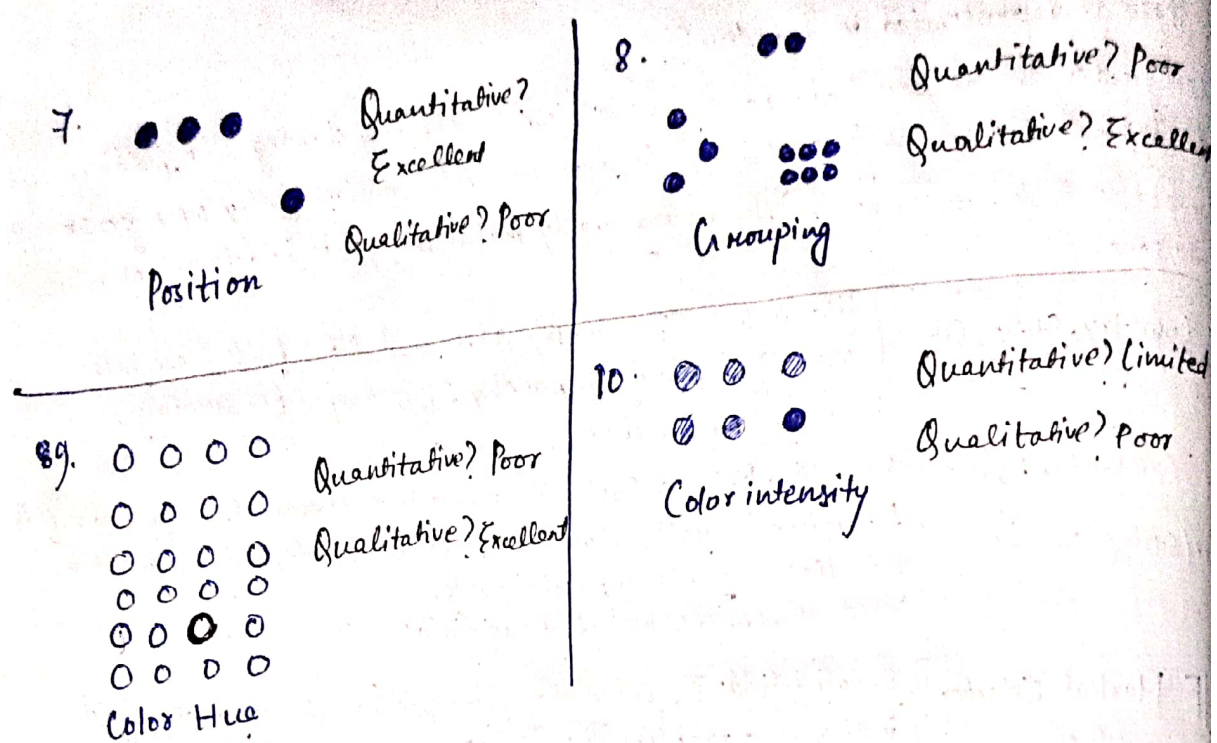
06/05/2020

Visualization in Tableau is possible through dragging and dropping Measures and Dimensions onto these different shelves.

Rows and Columns: Represent the x and y-axis of your graphs/charts.

Filter: Filter help us view a strained version of our data. For example, instead of seeing the combined Sales of all the categories, we can look at a specific one, such as just Furniture.

Pages: Pages work on the same principle as Filters, with the difference that we can actually see the changes as we shift between the Pages values. (It is located just ~~before~~ above p filter.)

Pages in the tableau are used when any of the fields need a broader or detailed visualization than a simple aggregated visualization.

Marks: The Marks property is used to control the mark types of our data. We may choose to represent our data using different shapes, sizes or text. The Marks card provides us with control over how the data is displayed in the view. Using this, we may change the level of detail as well as the appearance of the marks without affecting the headers built by fields on Columns and Rows.

Show Me:

Show Me is the brain of Tableau!
When we drag and drop fields onto the visualization area, Tableau makes default graphs for us, but we can change these by referring to the Show Me option. It contains; pie chart, bar graphs, a histograph, etc.

The DB-Ch Rule : When you drag and drop a field either in rows or columns, Discrete data is colored blue while continuous data is coloured green.

Headers are created when we use discrete field.

~~Axis is a created when we use continuous field.~~

We can show and hide headers, axes anytime.

Hiding headers can be useful when we are working with multiple measures.

Axes are created when we place a measure or continuous field on the Rows or Columns shelves. By default, the values of the measure field are displayed along a continuous field.