**TABLEAU**

# NOTE

* **AGGREGATION - Top, Bottom, Non-Null**
* **ANALYSIS – Highlighter**
* **ANALYTICS PANE - Custom | Model - Lines – Drag to View -SELECT CORRECT OPTION**
* **AXIS – Double Click – Format**
* **AXIS - Edit X Y Axes**
* **AXIS – Right Click – Hide Field Labels**
* **AXIS - X | Y Labels | Marks cards | Filters | Special - DO NOT MATCH ALWAYS AUTOMATICALLY**
* **CALCULATED FIELD - Drop down carat - Top of Data Pane**
* **CALCULATED FIELD - Right Click measure – Create Calculated Field**
* **CALCULATED FILED – DATEPART - SUNDAY = 1 SATURDAY = 7**
* **DASHBOARD - Right Click – Use as Filter**
* **DASHBOARD – Right Click Visual – Filters – Choose Filters**
* **DASHBOARD – Utilize Top Left for important KPIs**
* **DATA - Convert between dimensions and measures / discrete and continuous**
* **DATA - discrete dimensions and continuous measures are most common combinations of data roles**
* **DATA - Measure Names | Measure Values ARE AUTOMATICALLY GENERATED ALL FIELDS ABOVE**
* **DATA - Tableau assigns data roles to fields automatically. Good practice to review and adapt**
* **FILETR - Measure Names | Measure Values TO ADD MORE FILEDS (INCLUDING CALCULATED)**
* **FILTER – Customize – Show Apply Button**
* **FILTER - Drag to Filter Card or Filter Directly (gets added to Filter card)**
* **FILTER - Fields in Marks Cards can be used as conditions for Filters – BUT CHECK**
* **FILTER – More than one filter – Right Click – Show Relevant Values**
* **FILTER - Order - Extract | Data source | Context | Dimension | Measure**
* **FILTER – Right Click Measure or Dimension – Show Filter**
* **FILTER – Single Value Dropdown**
* **MARKS - Adding to Marks Cards adds to Row or Column**
* **MARKS - Cards – Edit type – Separate for each measure**
* **MARKS – Change chart type in dropdown**
* **MARKS – Edit Label | Change Format | Alignment**
* **MARKS - Show Mark Labels button**
* **MARKS - Drag order for Treemaps**
* **QUICK TABLE CALCULATIONS - Triangle**
* **TABLE The field that is listed first in the rows shelf will appear first**
* **VISUAL - Can Drag directly to View**
* **DATA – Create Bins**
* **VISUAL – Exclude Null**
* **VISUAL – Lasso Select – Drag to Bottom**
* **VISUAL – Right Click value – Edit Alias**
* **VISUAL - Show Me button – Ctrl + 1**
* **VISUAL – Tableau will not show if too crowded**
* **WORKSHEET – Right Click – Duplicate as Crosstab**
* **FILTER – Exclude Null**
* **TOOLTIPS - Edit format**
* **TOOLTIPS - Add Visualization**
* **VISUALIZATION - Drag pills to the chart area for quick colour**
* **DATA - Use CTRL key to multi-select pills to drag on and off the canvas**
* **DATA - CTRL + F - Search field name**
* **DATA - Type field names on inline formula**
* **DATA - Right-click (OPTION) and drag a field to Rows, Columns, or the Marks card**
* **DATA - CTRL (CMD) -drag to quickly duplicate fields**
* **DATA - Hide fields not being used/show hidden fields - Data Pane - Dropdown carat**
* **DATA - Create fast hierarchy - Drag one pill on top of another pill - drag other pills – rename Hierarchy**
* **DATA – Drill down Hierarchy – Click + Plus icon**
* **VISUALIZATION - Drag x or y axis to chart area for colour**
* **KPIs - Right Click Measure values and add to sheet**
* **DATA – Sort**
* **FILTER – Apply to Selected Worksheets**
* **DASHBOARD – Keep track of filters and which visualizations they impact**
* **DATA – Type of Data shows before value**
* **DATA – Link icon to link 2 or more datasets**
* **MARKS – Change Label colour to match – Match Mark Color**
* **DATA – Group by Folder**
* **VISUAL – Multiple pills in Columns for side-by-side (Compute using Pane across)**
* **REFERENCE LINE – Analytics Tab – Label Value**
* **MARKS – Colour – Right Click Dimension – Show Highlighter**
* **DATA – Dimensions with more values in Columns (X) – Dimensions with fewer values in Rows (Y)**
* **DATA – Column Chart – Measure (green) in Rows – Dimensions (blue) in Columns**
* **DATA – Try moving pills behind and in front of each other and from and between rows and columns**
* **TOOLTIP – Preview**
* **TOOLTIP – Add Filter – Add to Context**
* **MARKS – Colour – Advanced – Set Start and End**
* **MAPS – Mixed – Drag 2nd geometry parameter to canvas – Add Marks card Layer**
* **Maps – Mixed – Drag marks card to bring to front and back**
* **MAPS – Style**
* **DATA - Data Source Locale**
* **DATA - Workbook Locale**
* **ROWS - Type it in**
* **Save ad-hoc or Quick Table Calculation into Data Pane – Right Click or drag itt from Canvas into Data Pane – give it a name**
* **DATA – Add header of categorical variable to Table – Drag measure values to column (double click) – can sort counts**
* **DATA – Summary Card – Right click to add more summary values**
* **DATA - Create Matrix – Drag same values to Rows and Columns – Disaggregate**
* **DATA - Table - Drag pill to Columns - Double click pills**
* **DATA - Can sort by different categories**
* **DATA - Sorting from a field label gives a non-nested sort by default**
* **DATA - Sorting from an axis gives a nested sort by default**
* **DATA - Sort menu on Pill**
* **DATA - Combine Sets**
* **DATA - Filter Sets**
* **DATA - Order of Filters in Tableau**
* **DATA - 2 Green Bars on y-axis - Side by Side - Adds Measure values to Columns and Measure names to Rows and Filters**
* **CTRL + Drag (Mac: CMD + Drag) - copy pills in Marks**
* **Sample Superstore Excel-based data source - C:\Users\<user name>\Documents\My Tableau Repository\Datasources\2020.3\en\_US-US\Sample - Superstore.xls**
* **DATA – Combo / Dual Axes Chart – Can only sync second axis**

# Create and format visualizations

* Drag and drop dimensions and measures on the canvas, shelves and cards.
* Canvas - where your visualizations will appear.
* Columns - correspond to the x axis of your view.
* Rows - correspond to the y axis.
* Pages shelf lets you break a visualization into several pages, e.g. one page for each neighbourhood.
* Filters shelf lets you filter your data, and you will learn more about this in a next chapter.
* Marks field contains marks cards and marks types.
* Marks cards encompass color, size, and shape: these let you add context and detail to your view.
* Marks types - You can change the type of marks displayed in the view to fit your analysis better.
* Informative titles
* Colours and large fonts
* Legends
* Adjust axes and titles
* Create tooltips
* Can format at both Workbook and Sheet level
* **Dual Axes – Drag to top and right | Right click and choose Dual Axes**
* **Right Click on Y Axis – Choose Synchronise Axes**
* **Hide axes**
* **Centre title**
* **Edit Axes names**
* **Add colours to dimension**

# Workbook vs Sheet

|  |  |
| --- | --- |
| **WORKBOOK** | **SHEET** |
| .twbx | Similar to Excel tab |
| Organise, save share and publish | Displayed along workbook bottom |
| Multiple sheets | 1. Worksheet |
| Similar to whole Excel file | 1. Dashboard |
|  | 1. Story |

# Dashboard vs Story

|  |  |
| --- | --- |
| **DASHBOARD** | **STORY** |
| **Worksheet can be placed in a Dashboard** | **Dashboard can be placed in a Story** |
| Collection of several views | Dashboards can be bookmarked to create stories |
| Easy to compare data | Sequence of visualizations to tell a narrative |
| Uncovers key insights | Each individual visualization is called a Story Point |
| Automatically connected to worksheets | 1. Dashboard |
| Drill down and do advanced | 1. Story |
| Views can be connected – 1 view is interactive filter |  |

* Drag different Worksheets to Dashboards overlay
* Can move | float legend and filters
* Use visualisations | dashboards as interactive filters
* **Add Filter - Click visualisation | dashboard – Analysis toolbar – Filters**
* Drag different Dashboards to Story

# Discrete

* Blue fields are discrete, or categorical fields, they have individually separate and distinct values.
* Examples include room type, neighbourhood, and the ID number of the listing.
* **Discrete means "individually separate and distinct."**

# Continuous

* Green fields are continuous fields, treated as an infinite range.
* Examples are the number of reviews per month, room price, or the longitude of the location.
* **Continuous means "forming an unbroken whole, without interruption".**

# Dimensions

* Colour coded Blue
* Dimensions, positioned at the top, contain qualitative values, such as names or dates.
* E.g. in dataset: Neighbourhood, Room Type, or number of reviews per month.

# Measures

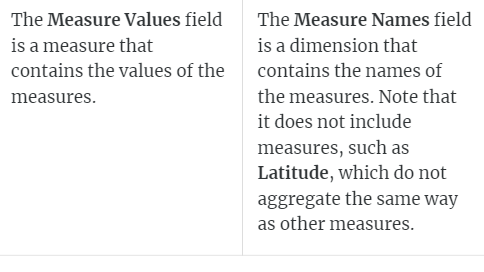
* Colour Coded Green
* Measures, positioned under the dimensions, contain numeric quantitative values that you can measure, and aggregate.
* E.g. in dataset: Price, Number of minimum nights, and Total number of reviews.

# Dimensions / Measures / Discrete / Continuous

* Measures - usually metrics, or numerical data
* Measures are aggregations – up to the granularity set by the dimensions in the view
* Measure value depends on the context of the dimensions
* Dimensions are usually categorical fields - set the granularity, or LOD in the view
* Dimensions used to build the view will determine # of marks
* Continuous numbers can take on any value in a range
* Discrete numbers have distinct, separate values
* **Dimensions are Discrete in Blue -** come out onto the view as themselves
* **Measures are Continuous in Green -** come out onto the view as aggregates
* Continuous pill creates an axis - automatically fill the entire view along that direction
* Discrete pill creates a label with panes for each value - as much / little room as required
* Continuous pill on the color shelf will create a gradient
* Discrete pill on the color shelf will create a color palette
* Map - Measure on color defaults to a filled map
* Map - Dimension on color defaults to a symbol map
* Map - gradients or palettes depends on if the pill is continuous or discrete
* Dates - options at the top are Discrete date parts - are treated like categories
* Dates – options at bottom are Continuous date truncations - treated as an ongoing progression along an axis
* Filtering - Discrete dimension - options related to the specific list of values for that pill
* Filtering - Continuous measure - asks first if want to filter at the row-level or aggregate level - then brings up options for continuous ranges



# Measure Values and Measure Names



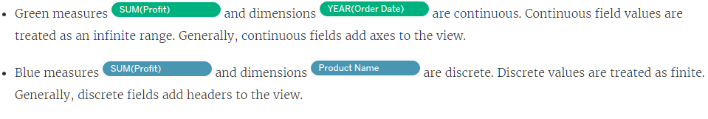
* What is the Measure Names field? - It is a dimension that contains the names of the measures
* What is the Measure Values field? - It is a measure containing the values for the measures
* When you add Measure Values to a view, Tableau creates a Measure Values card that lists the measures in the data source with their default aggregations, such as sum
* You can change the aggregation for measures in the card and remove individual measures from the view by dragging them out of the Measure Values card
* When you add Measure Names to a view, the measure names appear in the view as row or column headers, depending on whether you add it to Rows or Columns
* Fifty on Rows and sixteen on Columns
* What do you need to do so two or more measures share an axis? - Add Measure Names to Rows or Columns and display at least two measures

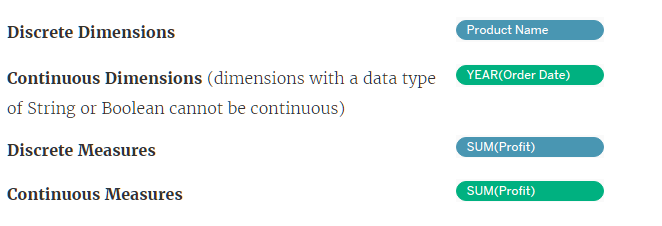
# Attribute (ATTR)

* ATTR() can indicate that there are multiple values from the secondary data source.
* Dimensions added to Tooltip on the Marks card are automatically wrapped in ATTR() because dimensions on Tooltip must be aggregated.
* ATTR() will display \* as there is more than 1 value and the view or the values need to be adjusted.
* Like other aggregations, ATTR() can be used to change a non-aggregate value to an aggregate value to resolve aggregation errors in the calculation.
* When a calculation that returns numeric data contains ATTR(), if there are multiple values in the ATTR(), the calculation will return NULL rather than an asterisk.

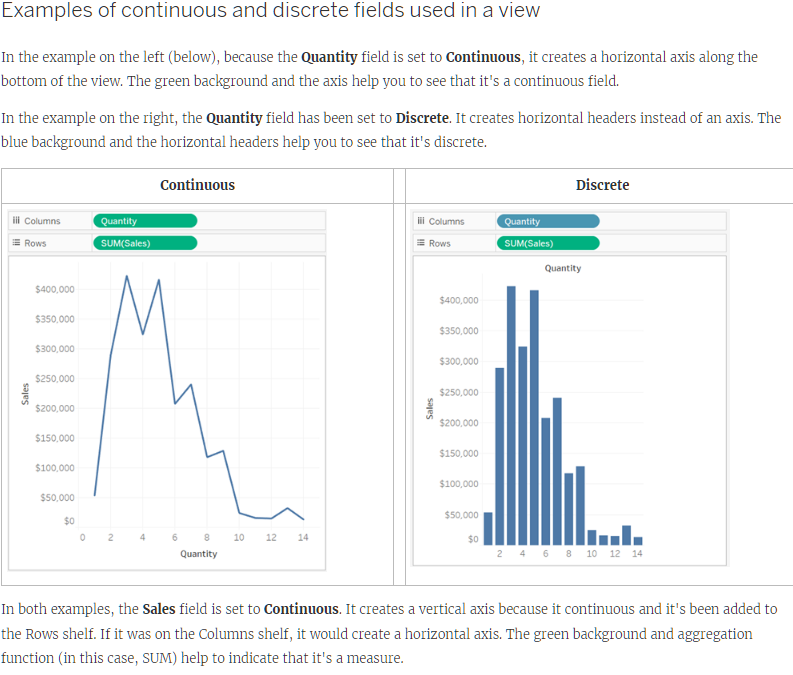
# Data roles in Tableau

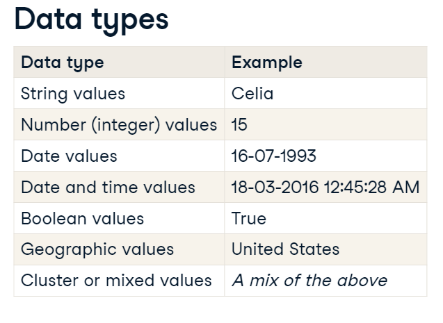
* Discrete dimensions and continuous measures are the more common combinations of data roles.
* E.g. eye color and sex, and height and weight, respectively.
* Less common E.g. discrete measures (shoe size and age) and continuous dimensions (date).



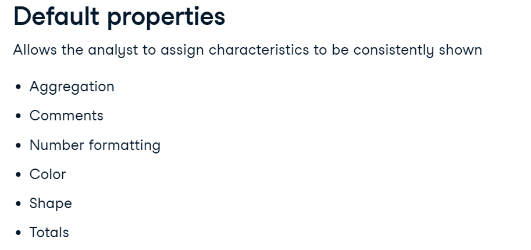




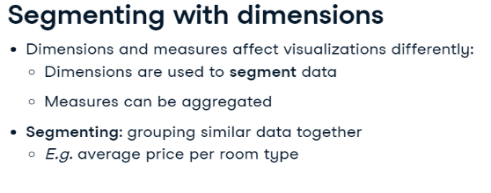




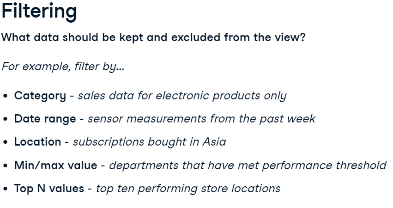
# Default properties

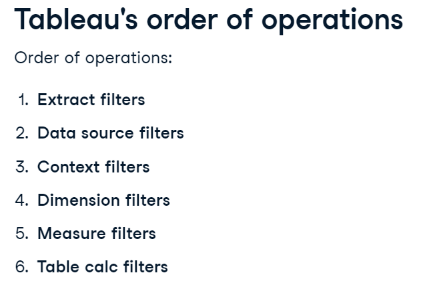


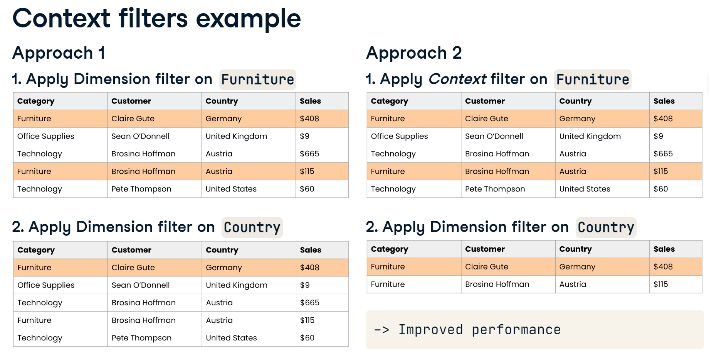
# Segmenting with dimensions



# Filters



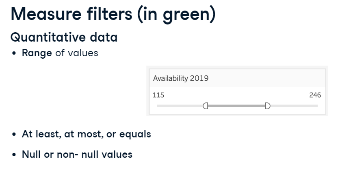




# Dimension filters (in blue)

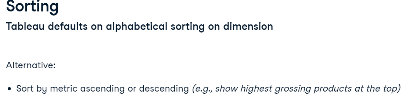


# Measure filters (in green)

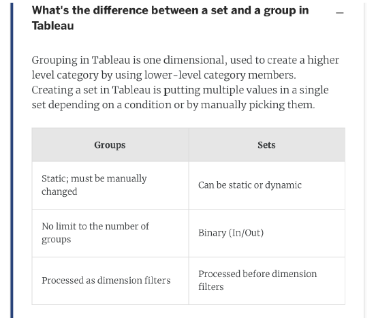


* Add Context Filter to Context – Right Click Pill in Filter shelf card - Black
* Add Segment Filter – Connect to data source
* Compare data between multiple data sources by using a common field - establish relationship - if fields have same names, Tableau will automatically establish blend relationship – if fields are not named the same, create relationships or map fields to one another
* Filters – All using data source – enable filtering across multiple data sources

# Sorting



# Sets & Groups



# Aggregation

* Default aggregation for Measures is SUM
* Can only aggregate Dimensions with MIN, MAX, COUNT and COUNT DISTINCT
* Aggregating a Dimension creates a temporary Measure
* All Dimension aggregations can be applied to Measures but not vice-versa

# Calculated Fields

* Create new Field - Measure or Dimension
* Analysis Tab – Create Calculated Field
* Use Functions
* Enter name of Field and add Formula
* Can be edited in dropdown
* Right click Measure and Choose Calculated Field

# Geographical Data

* Filled Map | Symbol map
* Geocoding – Globe icon
* Drag Country (globe icon) to View
  + Automatically creates map and geo data
  + Automatically adds Country to Marks Cards
* Edit Map layers in Map tab

# Date Data

* Calendar Icon
* Date hierarchy
* Top is Dimension – Discrete – Blue – Aggregates
* Bottom is Measure – Continuous – Green – Timeline
* DATEDIFF
* 
* DATEPART SUNDAY = 1 SATURDAY = 7
* 

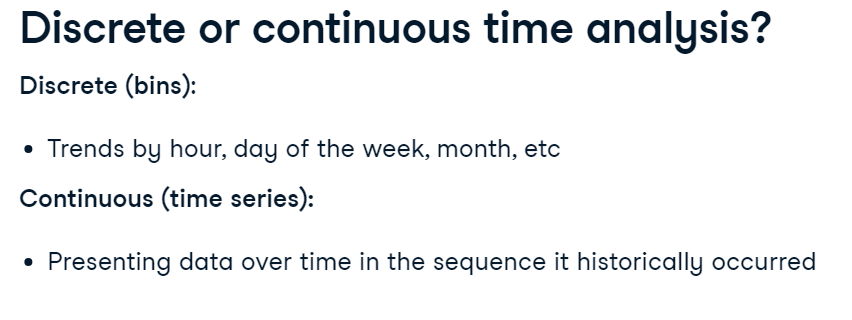
# Reference Lines, Trend Lines, Forecasting

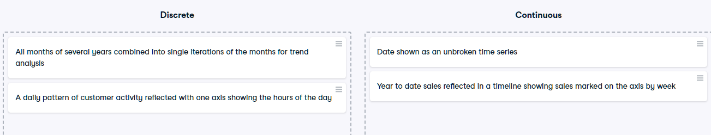
* Reference line drawn on a chart representing another measure or point of reference E.g. AVG
* Reference line – Analytics pane - Custom
* Trend line - used to predict the continuation of a certain trend
* Trend line – Analytics pane - Model
* Forecasting - predicting the future value of a measure using mathematical models
* Forecasting – needs a time dimension and a measure
* Forecast – Analytics pane – Model

# Data Preparation

* When a numeric value is brought into Tableau, it's placed by default in the Measures section
* Move numeric fields that shouldn’t be aggregated to the Dimensions section
* Check Default Properties – Number Format - Custom
* Fit Width
* Edit Alias
* Add Highlighter - Analysis tab
* Show Filter – Customize – Show Apply Button
* Create Calculated Field – Drop Down Carat at top of Data Pane

# Discrete vs Continuous Time Analysis



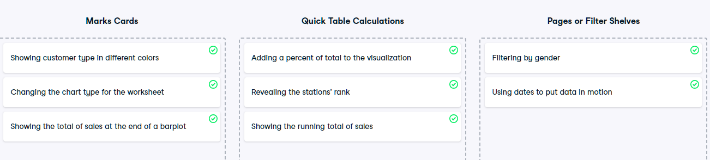


# KPI Dashboard

* Key Performance Indicators are measurable values that track a company's key business objectives.
* Turn off the field label for the columns by right clicking on it and then selecting Hide Field Labels for Column.
* Add KPI charts to pre-formatted dashboard.
* Worksheets - fit the entire view - charts will fill up whatever space is available in the dashboard.
* Turn off the title by clicking down arrow and then unchecking the title.
* When adding sheets to the dashboards, filters are automatically brought in - Remove duplicate.
* Bring in new filters - clicking the down arrow in grey menu upper right corner - Filters, select the one you want.
* Each filter drop-down - Apply to Worksheets - All Using this Data Source – when a filter is applied, every sheet using this data source will filter its content accordingly
* Select visualization - grey menu - Select Use as Filter - Click the funnel to make it solid so it's used as a filter.

# Quick table calculations

* A table calculation is a calculation that you can apply to the values in a visualization.
* Examples include running total, difference, percent of total, and many more.
* These predefined calculations are calculated based on what is currently in the visualization.
* They do not consider any measures or dimensions that are filtered out of the view.
* Table calculations are defined by their scope and direction.
* The scope defines the group on which the calculation is performed.
* The direction defines how the table calculation moves within the scope.
* Options are across, down, down then across and so on.
* Quick table calculations are table calculations that you can apply quickly to your visualization.
* They are applied to the visualization with the most typical scope and direction settings.
* When the menu is accessed, only calculations that are possible with your data are available

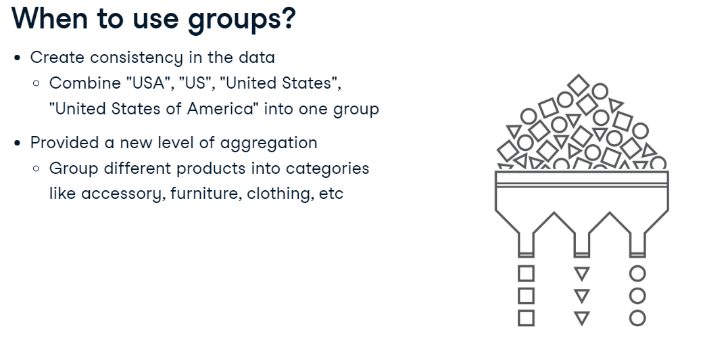


# Ranking

* QUICK TABLE CALCULATION – Rank
* TABLE CALCULTION – Edit
* Tableau applies Competition Ranking - Data with same value will get same rank

# Groups

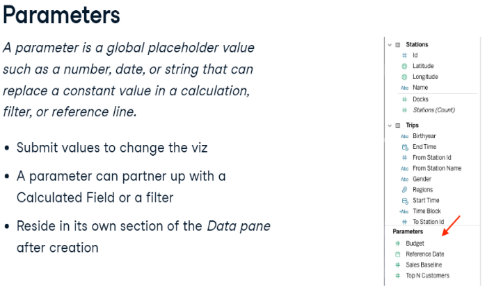


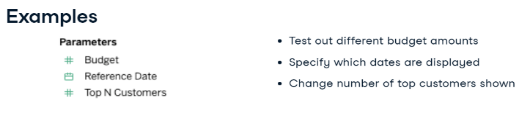


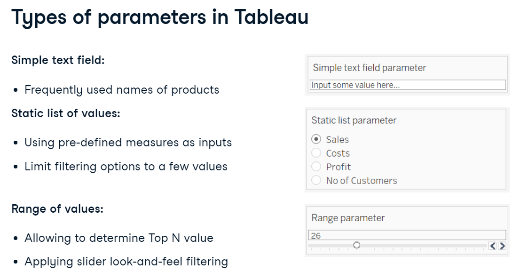
# Groups – Maps

* Lasso Selection
* Paperclip
* Edit Groups

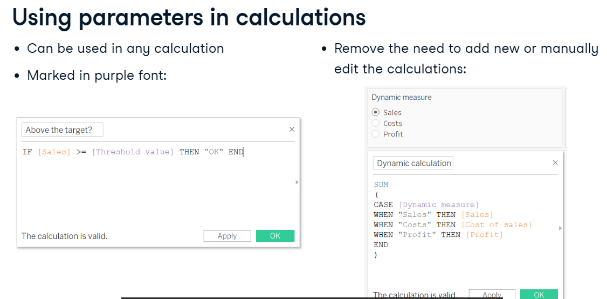
# Parameters



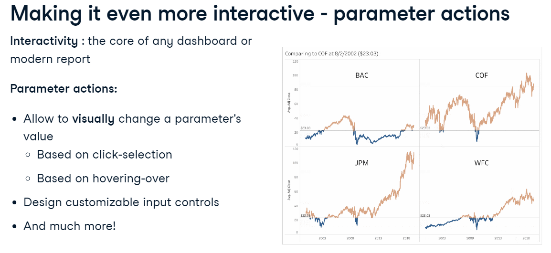




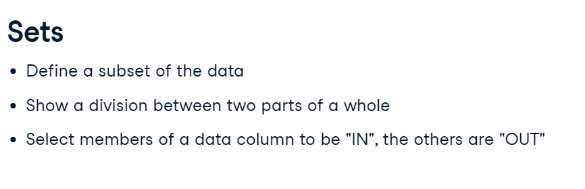
* **Create Parameter in Data Pane**
* **Set Start, End, Steps, Current Value / List / Range**
* **List – Can select Single Value List in Parameter Controls**
* **Connect to Filter -Top / Bottom / List / Condition**
* **Connect to Calculated Field – True / False check**
* **Replace pill in Visual with Parameter**
* **Show Parameter**
* **Edit Axis name with Parameter**
* **Edit Tooltip to update name**
* **Switch Parameters**

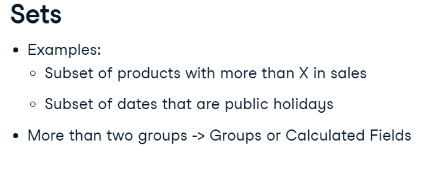






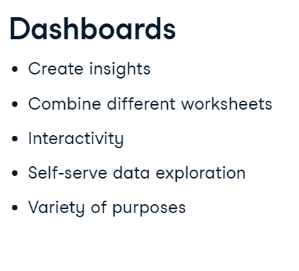
# Sets

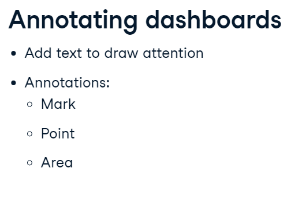


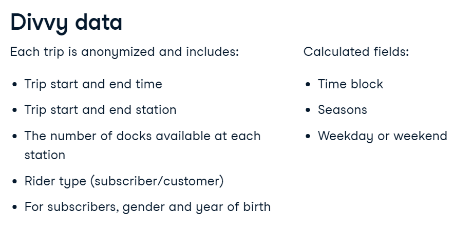


* Joined Circle
* Include / Exclude set

# Dashboards

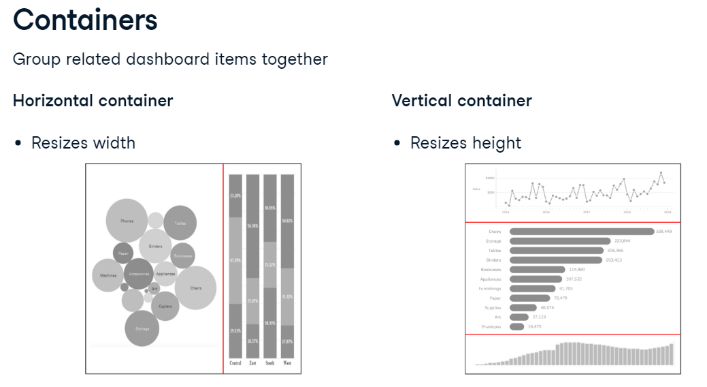






* Size - Make use of entire available space
* Worksheet - Fit - Entire View
* Edit / Hide Axis
* Use as Filter - Also filters worksheets
* VISUAL - Right Click - Annotate
* Annotation – Format
* Visualization - Show Filters (affecting the viz)
* Filter – Apply to Worksheets / All using datasource

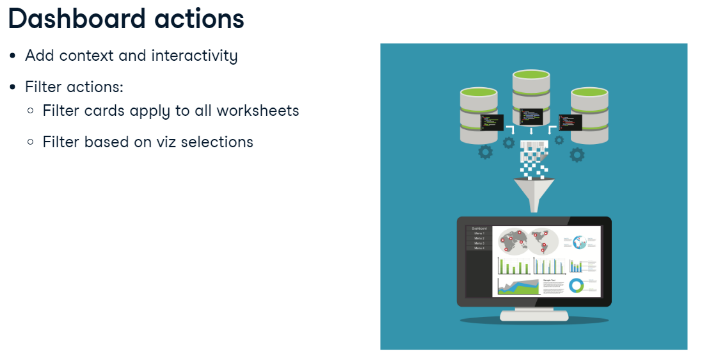
# Dashboard Containers



# Show/Hide button

* Use to display legends / filters in a floating container
* Use text / images and tooltips for accesibility

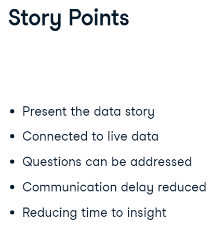
# Dashboard Actions

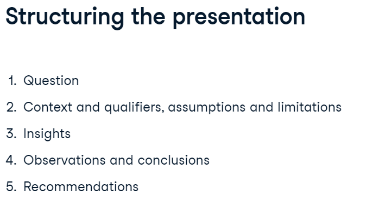


# Communicating Insights Visually



# Story Points







* Story – Story Points – Navigator
* Format / Edit Viz
* Annotate
* Legend – Line under Pen / Dashboard Actions
* Edit Axis – Fixed range
* Add Filter Action from Source to Target dashboard
* For dashboards that you plan to include in your story, use the Fit to option under Size on the Dashboard pane. It will resize your dashboard so that it’s the right size for the story you’re creating

# Mobile

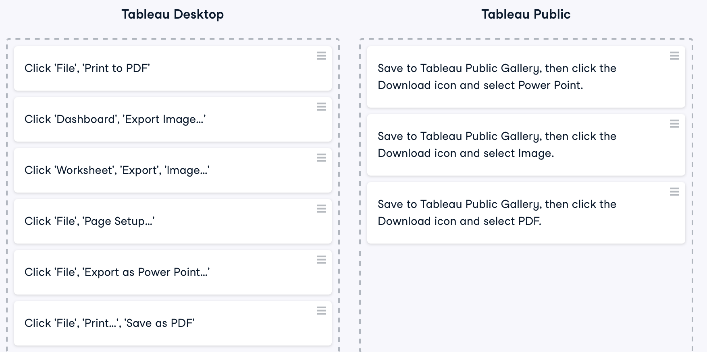
* Mobile View - Navigation – Edit

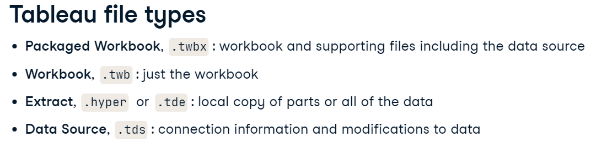
# Animation

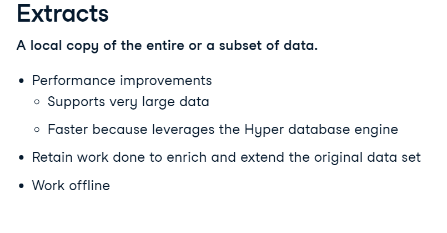
* Simultaneous
* Sequential
* Format – Animations

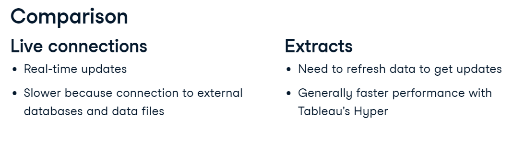
# Save / Export / Share / Extracts

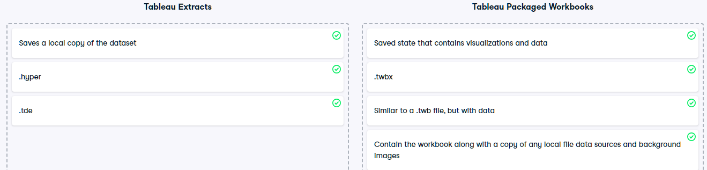
* Options

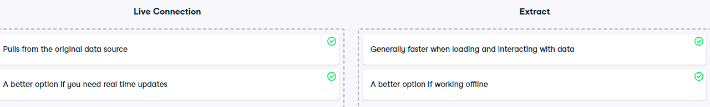




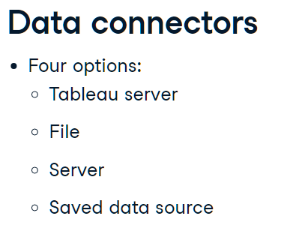


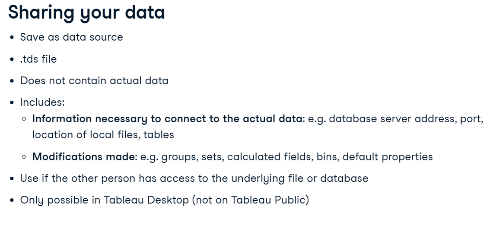






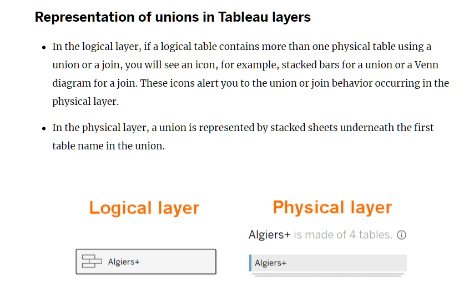
# Data Connectors

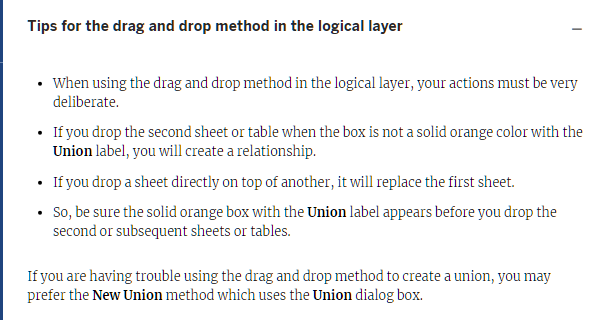


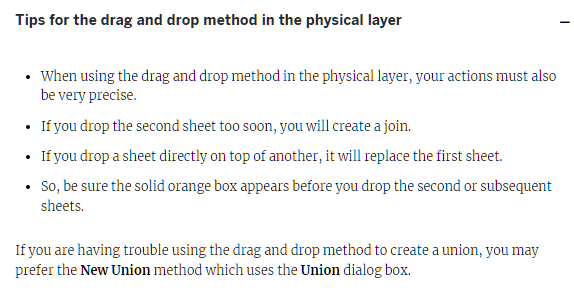


# Physical and Logical layers

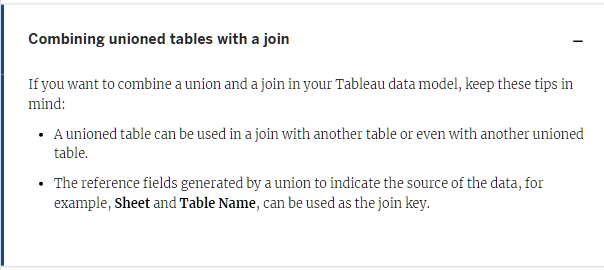
* Default view in the Data Source page canvas is the logical layer of the data source
* You combine data in the logical layer using relationships (or noodles)
* Think of this layer as the relationships canvas in the Data Source page
* Next layer is the physical layer
* You combine data between tables at the physical layer using unions (and joins)
* Each logical table contains at least one physical table in this layer
* Think of the physical layer as the joins and unions canvas in the Data Source page

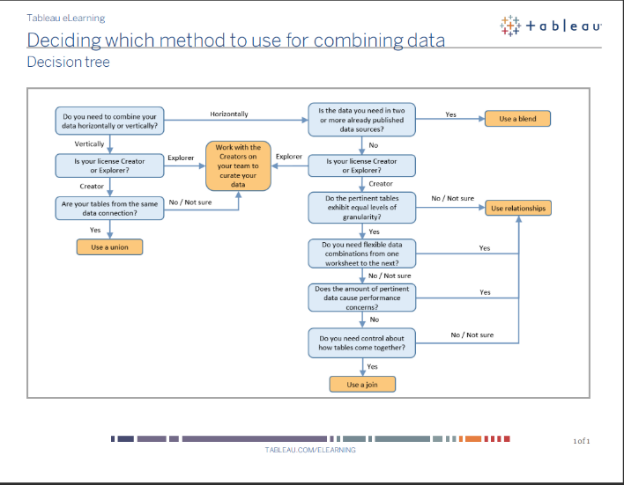






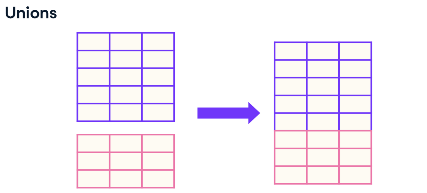
# Combining Data





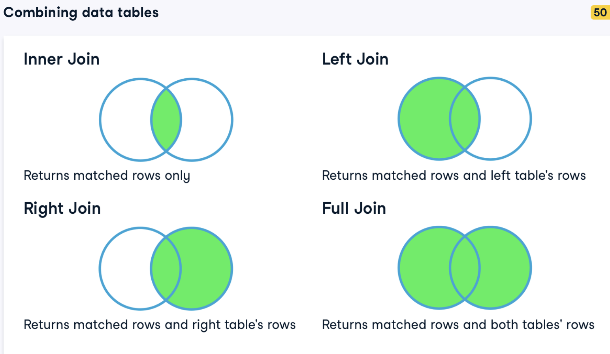
# Blends

# Unions



* Data Sources page
* Connect Data
* Right Click Table – Open – Drag to designated Union area
* Check Union
* Verify number of records
* Edit Union – Add / Remove

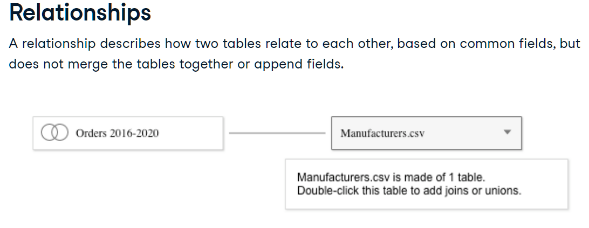
# Joins

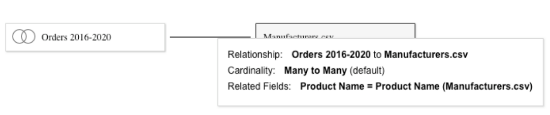


* Right Click Table – Open – Drag to Join area (outside designated Union area)
* Select Join type
* Select Join key

# Relationships

* A relationship describes how two tables relate to each other, based on common fields, but does not merge the tables together or append fields.
* When a relationship is created between tables, the tables remain separate, maintaining their individual level of detail and fields.
* One or more related fields are selected to define the relationship.
* Tableau suggests thinking of a relationship as a contract between two tables.
* Tableau will bring data from both using that contract to decide what join would be most appropriate in the given context.



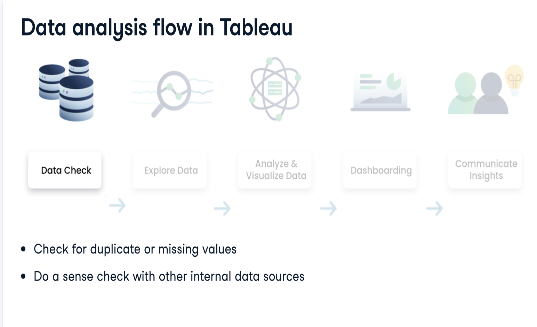


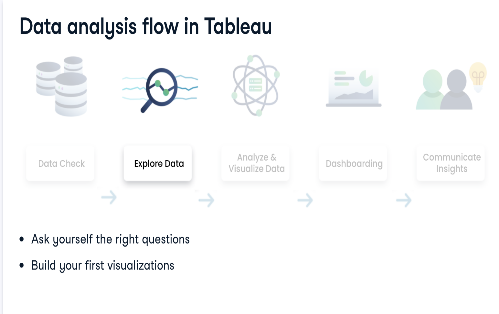
# Joins vs Relationships

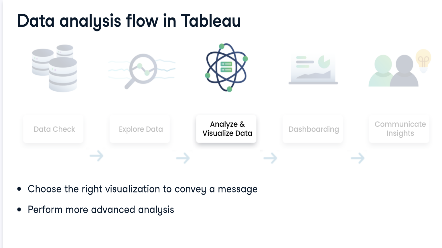


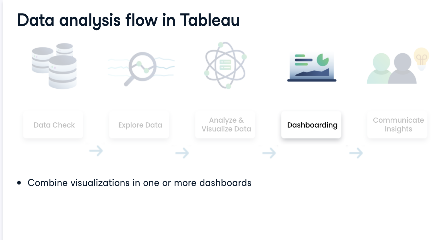
* Make connection to new Data source
* Drag to Canvas
* Edit Relationship
* Match on one or more fields

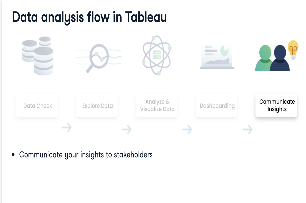
# Data Analysis Flow in Tableau



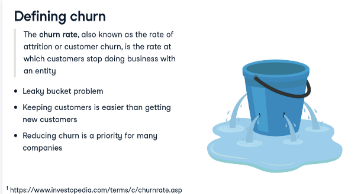


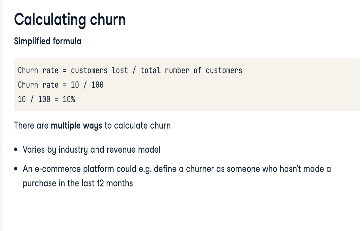






# Churn





# Waterfall (bridge) chart

* Explains the net change in value between two points, split over categories
* Typically starts at a baseline of zero; then, there are a series of bars that present category contribution to the total.
* Positive values can easily be distinguished from the negative ones by the use of a categorical or graded color palette
* Exposes the complexity hidden behind an aggregated number
* Downside is that it works well with only a limited number of categories
* Running Total
* Gantt chart
* Add value to data and add negative sign
* Add grand total

# Heat map / Highlight Table

* Used for presenting density and comparisons, often in a matrix form, relying on the use of colors to communicate the values.
* Heat maps have various forms and types. The most frequently used are colored geo-maps, for example, to illustrate the density of the population, but they also are often used in web analytics to analyze where on the screen visitors click the most.
* Another use case is a matrix comparison
* Advantage of heat maps is that it gives almost an instant high-level picture through their use of easy-to-understand color gradations
* However, without appropriate labels, it might be challenging to interpret when a high level of detail is required
* Has 2 versions – Heat Maps and Highlight Table
* Add pill to detail and colour - change marks type to square - edit colour - centre value

# Scatter plot

* Show the relationship and correlation between two numerical variables plotted simultaneously along both the X and Y axis.
* Often used in exploratory data analysis or when we need to plot the data on a quadrant
* Can be multidimensional, with the use of color, size, and shape, and easily store quite some data, with a high number of data marks
* Powerful to present a correlation between two measures
* However, in many cases, scatter plots can be tricky to communicate the data insights with, especially when data marks are plentiful and when there is no immediate correlation legible from the chart.

# Pie Chart

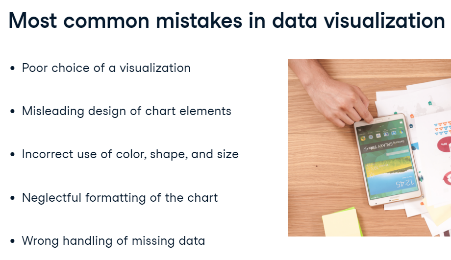
* Proportion / Angle – Rows
* Size / Colour – Column
* Maximum dimensions = 5

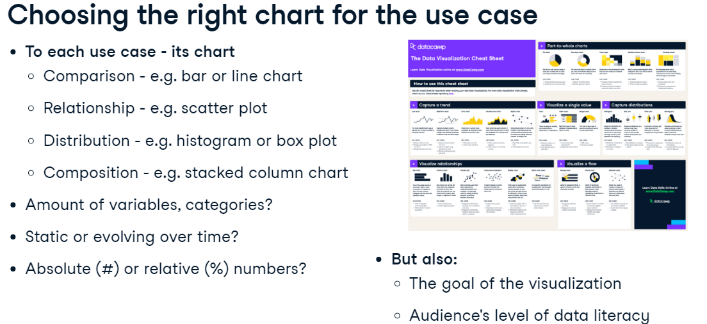
# Bar-in-bar charts

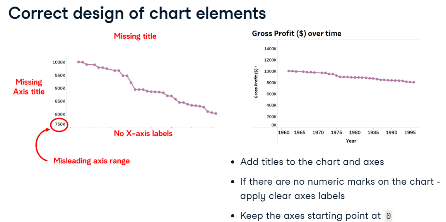
* visually show the same information as a side-by-side chart
* More compact
* More organised
* More immediately informative as a result
* Start with side-by-side chart
* Move Measure Names from Rows to Colour and Size - both measures have
* one stacked bar
* Analytics menu - Stack Marks - Off - both measures start at 0
* Edit Colour and Size

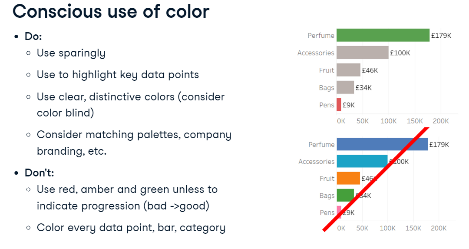
# Bullet graphs

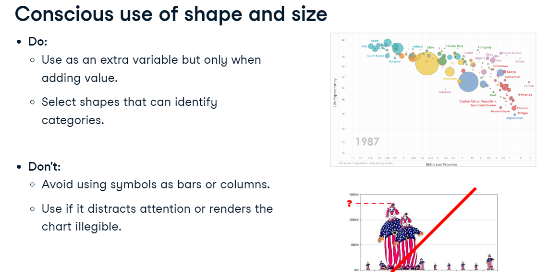
* One measure is in an entirely different format than the other
* One measure is shown as a horizontal bar
* Another measure is shown as a vertical reference line, perpendicular
* to the horizontal line
* Reference bands that measure the different stages of progress
* toward the quota
* Start side-by-side chart
* Remove Measure values from Columns
* Remove Measure names from Rows and Filters
* Drag Actual pill to Columns
* Drag Target pill to Details Marks Card
* Right Click Actual axis - Add Reference Line - Per Cell - Target -
* Average
* Right Click Actual axis - Add Reference Line - Distribution - Per Cell
  + Sum - 25, 50,75 % of Target - Average
* Decrease size of Actual Bar
* Calculated field - Achieved Target - Colour
* Calculated filed - % Achieved - LabelData Visualization

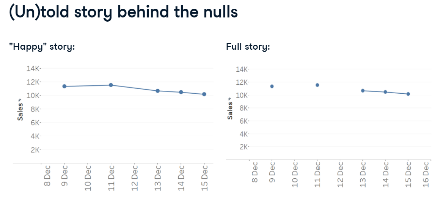


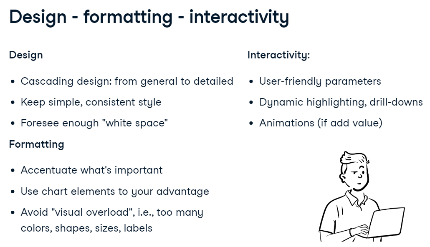


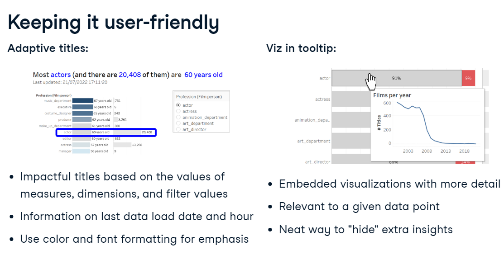


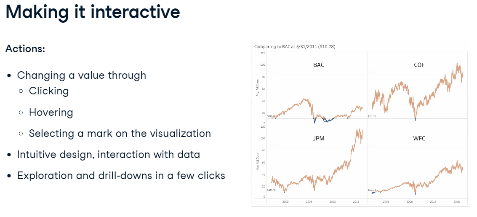


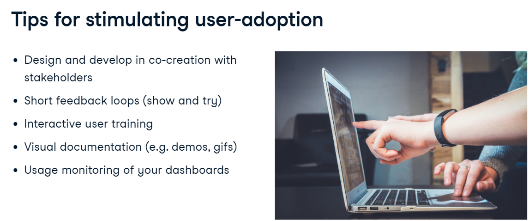












# Dashboard Actions

* Select Source and Target sheets
* Run action…
* Clearing selections will..

# Viz in Tooltips

# Connect to URL

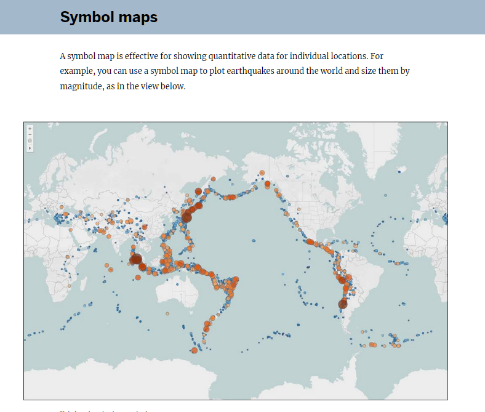
* Calculated Filed "https://en.wikipedia.org/w/index.php?search="+[Title]
* Dashboard Webpage Object
* Action – URL

# Word / Tag cloud chart

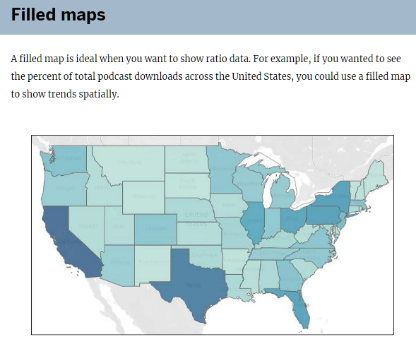
* Key words enriched with size and / or colour
* Start with Treemap
* Convert Marks to Text

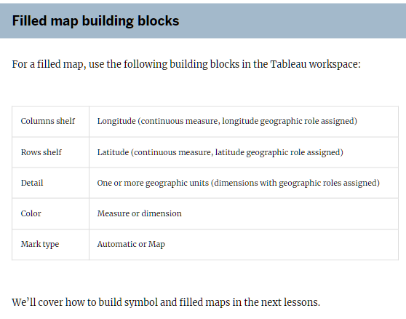
# Mapping



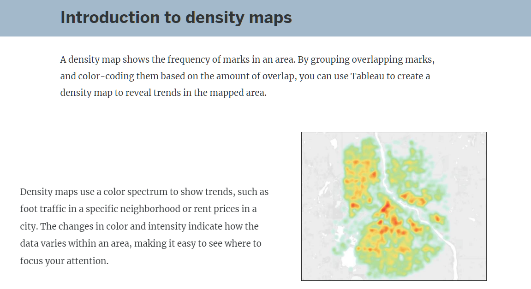


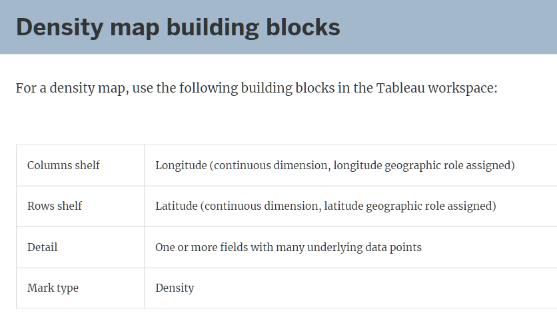




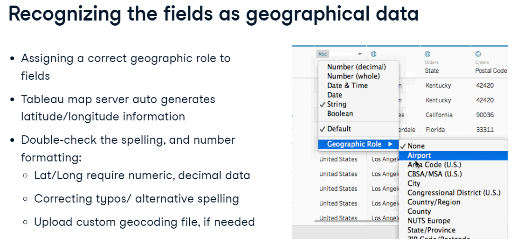


* Symbol maps can show up to two quantitative values per location – Size and Colour
* Pin icon - Fix map data from changing



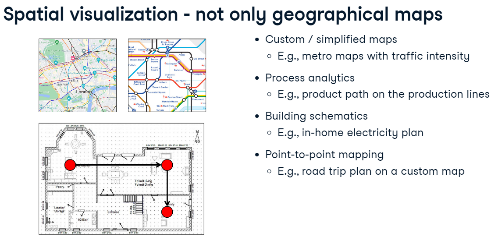


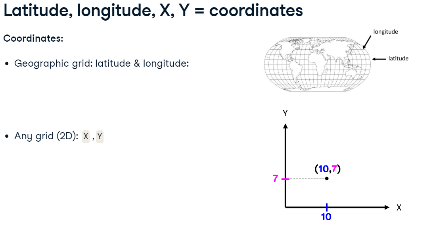
* By adjusting the combination of Color, Intensity, and Size, you can make the trends in a density map much more apparent

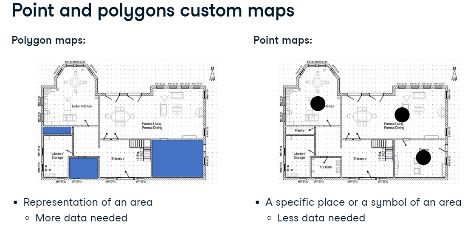


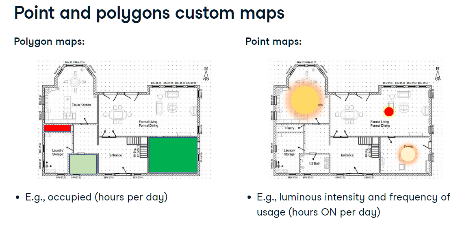
* Colour Palettes
* Opacity
* Size
* Maps – Background – Dark
* Add time filter – Add to Pages shelf
* Marks type – Density
* Density Colour
* Density Intensity
* Dual Axis
* Layering – Multiple Rows
* Multiple Marks for each map
* Overlay – Right Click Rows and select Dual Axis
* Charts on top of Maps

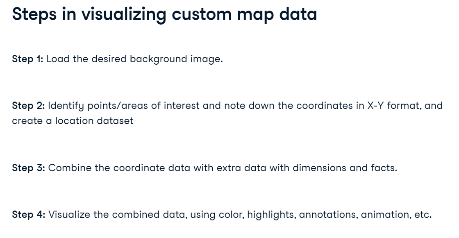
# Custom Maps





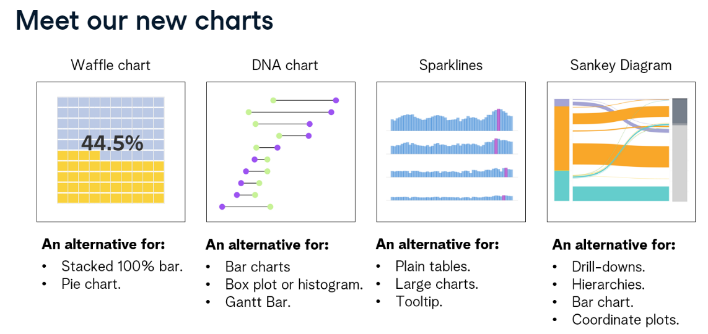


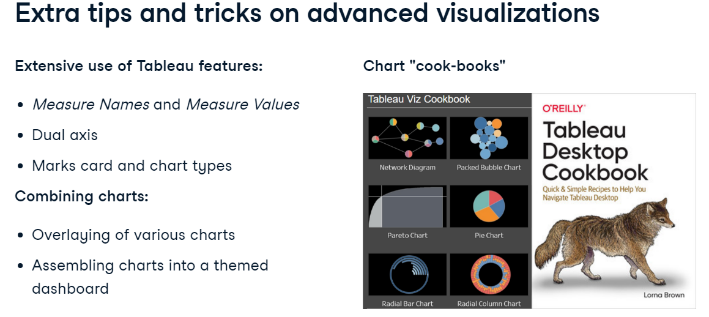


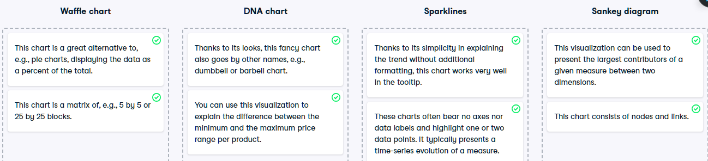


* Add custom background
* Add custom location file
* Use paths

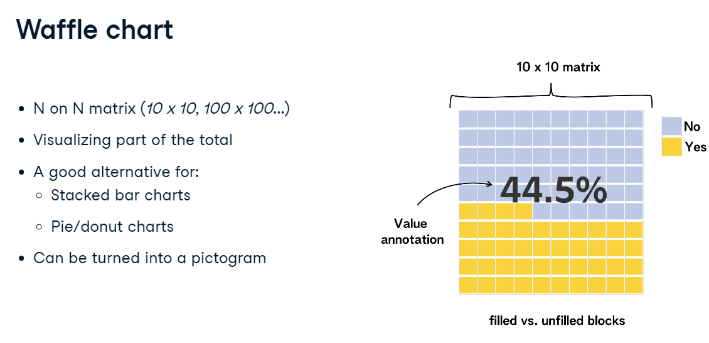
# Advanced Charts







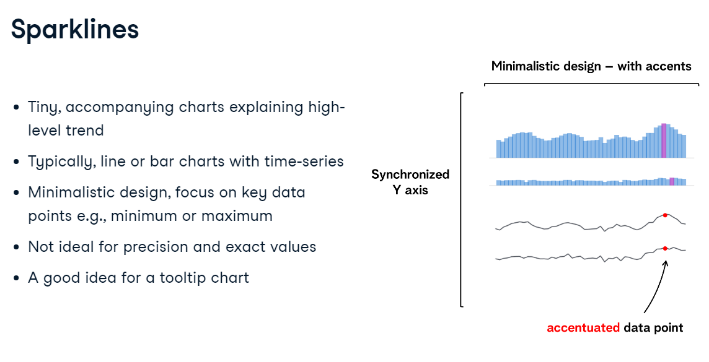
# Waffle Chart



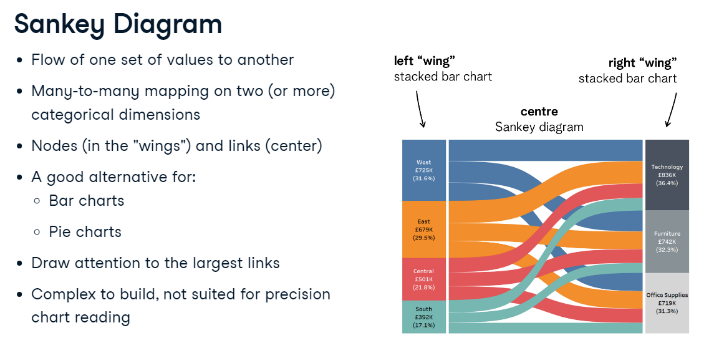
# DNA Chart



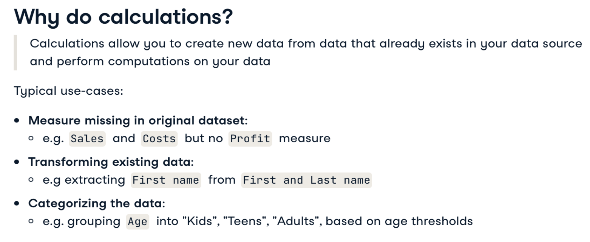
# Sparklines Chart

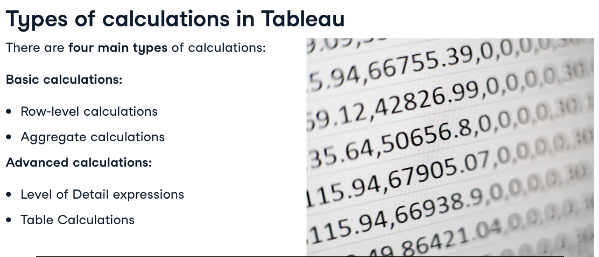


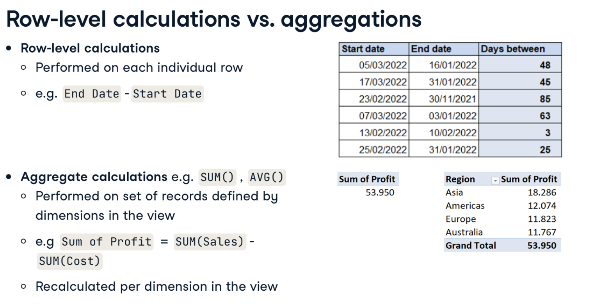
# Sankey Chart



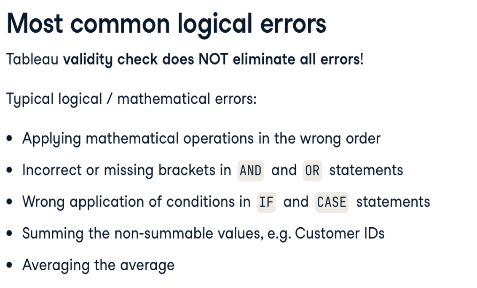
# Calculations



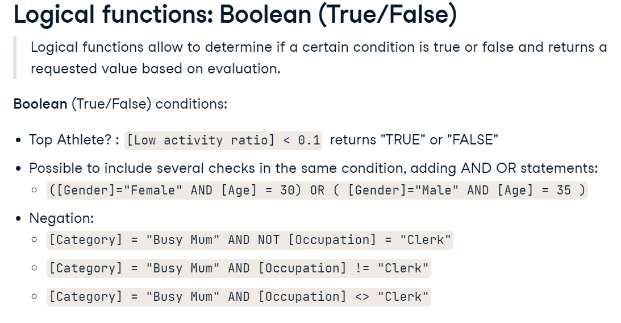


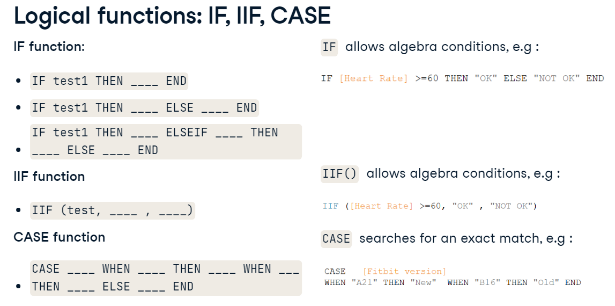


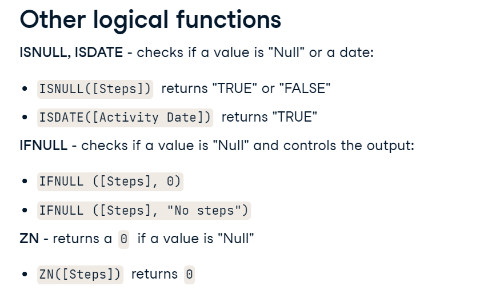




# Logical Functions



.

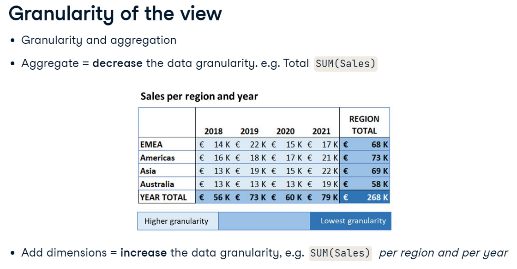


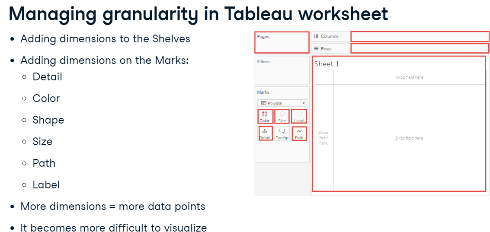


# Text Functions

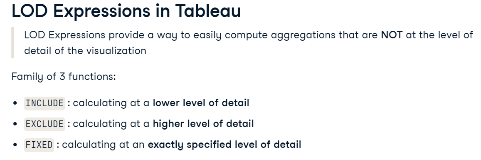
* Transform – Split
* Alias 1 and 0

# Granularity

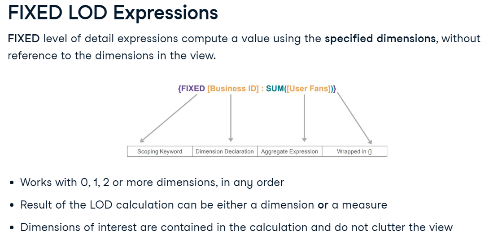


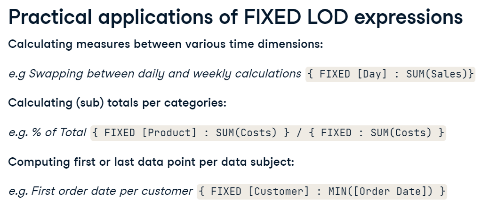


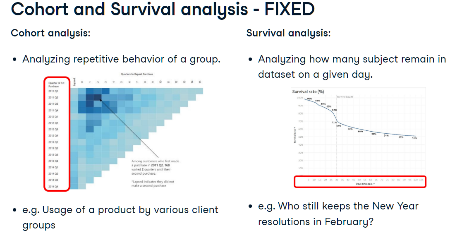
# LOD Expressions (Level of Detail)



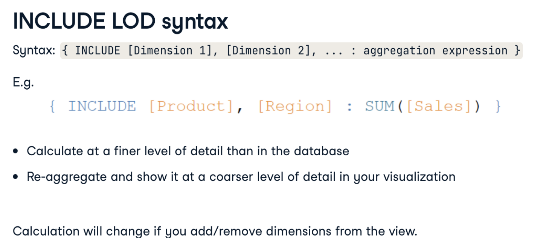
# FIXED LOD Expression

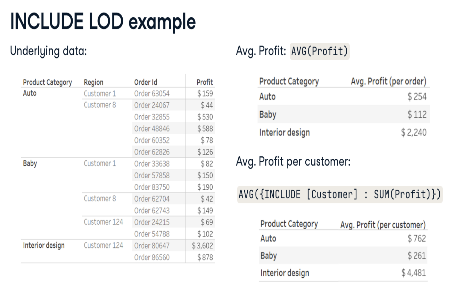




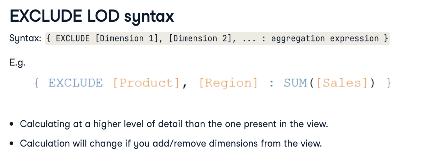


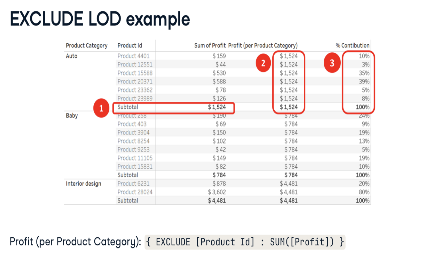
# INCLUDE LOD Expression

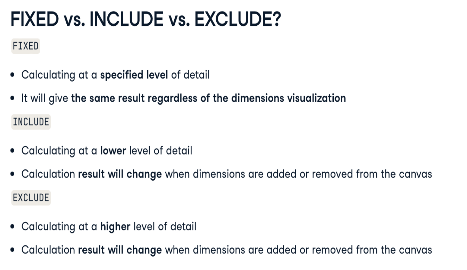


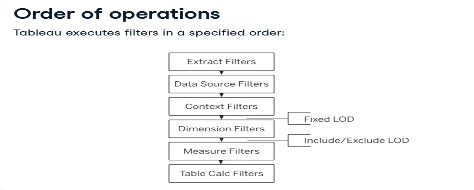


# EXCLUDE LOD Expression

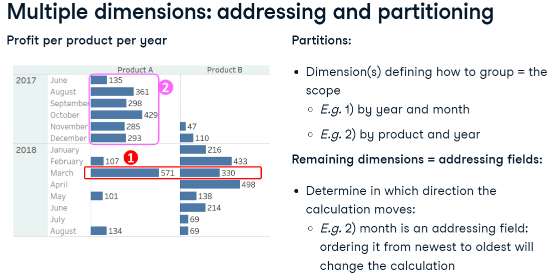


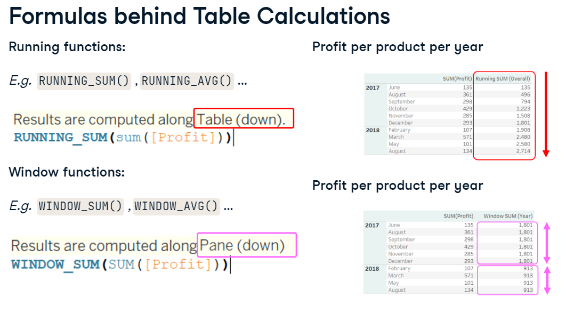


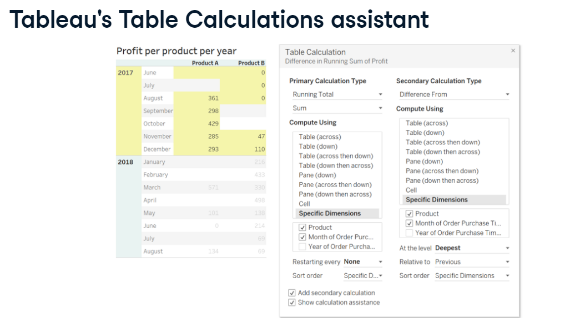


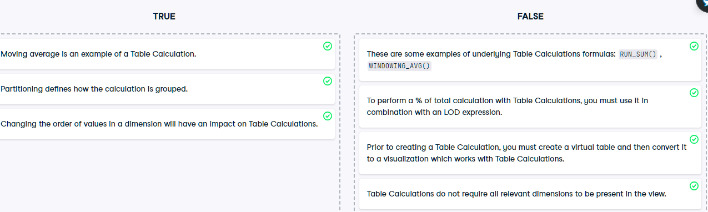


# Table Calculations



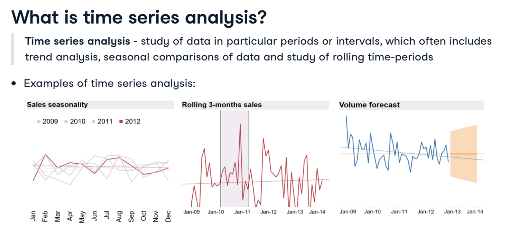


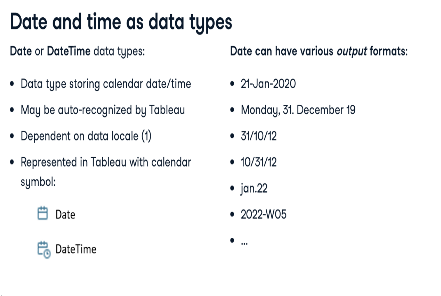


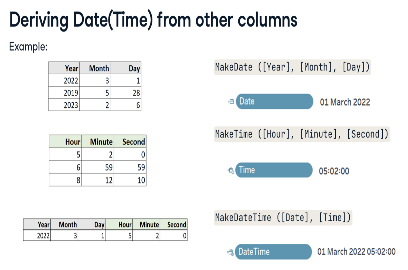


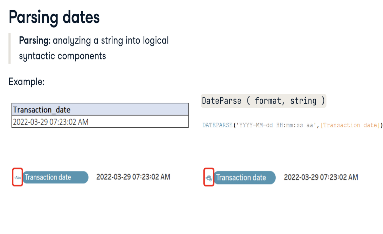
* Ranking default – Table across
* If desired Ranking dimension in Columns – Choose Table Down

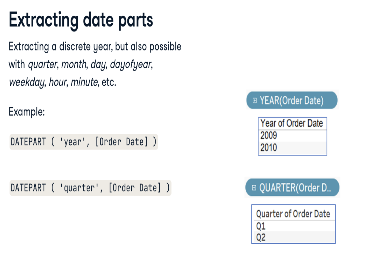
# Time Series Analysis

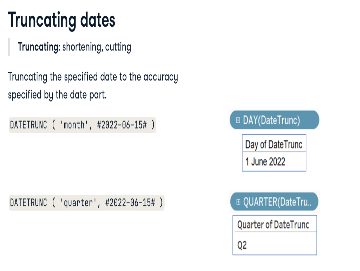


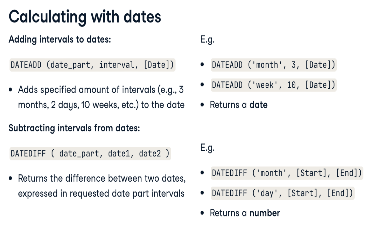




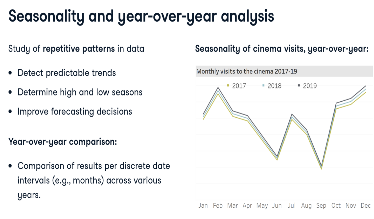


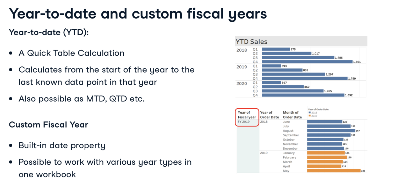


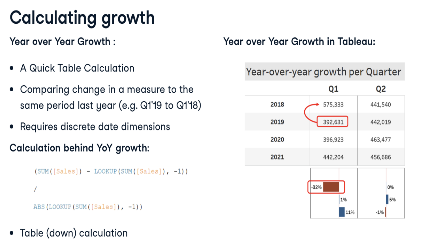


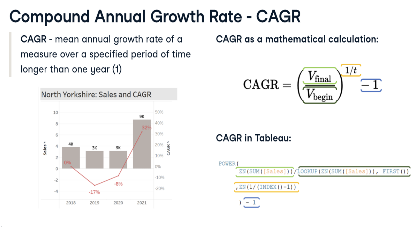


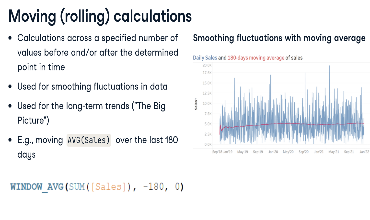


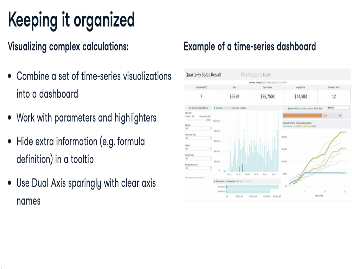












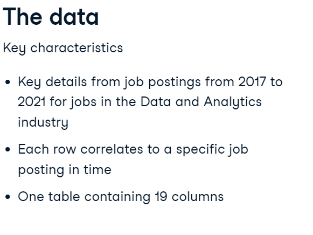
* Calculate Moving Average – Right Click measure in Rows – Quick Table Calculation – Moving Average
* Adapt Moving Average – Double click measure in Rows – Adapt start value (-5 for 5 week average)
* START OF WEEK - DATEDIFF('week',[Order Date],[Return Date], 'sunday')

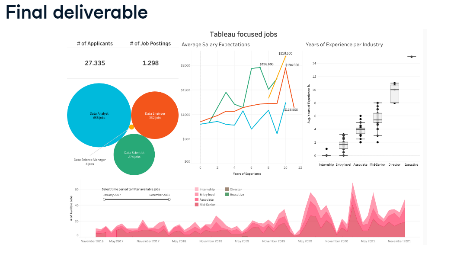
# Custom Date

* Create custom continuous-date-value and discrete-date-part on Pill
* Data source – Calendar default – ISO-8601 week based calendar
* Create Custom Date Hierarchy from discrete-date-parts

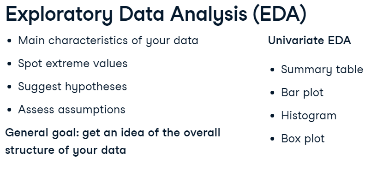
# Case Study: Job market analysis in Tableau

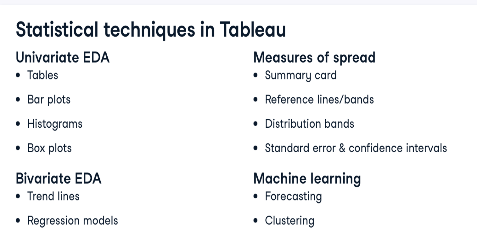






# Exploratory Data Analysis

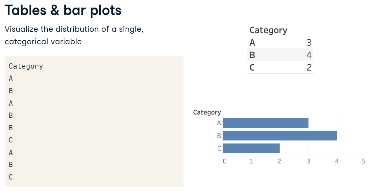




# Statistics

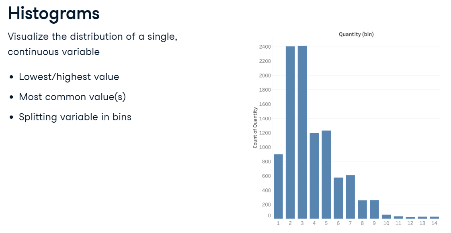


# Tables & Bar plots



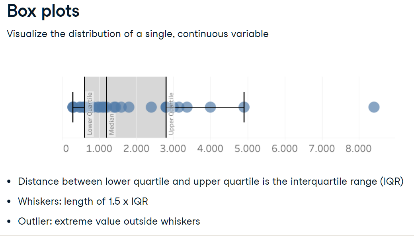
# Histogram

* Compute using Pane across for side-by-side comparison
* Bar plot for a continuous variable
* Set pill bin sizes – Columns (Continuous if large data)
* Set same pill to Count - Rows
* Select pill – Show me - Histogram
* Set Parameter for Bin size

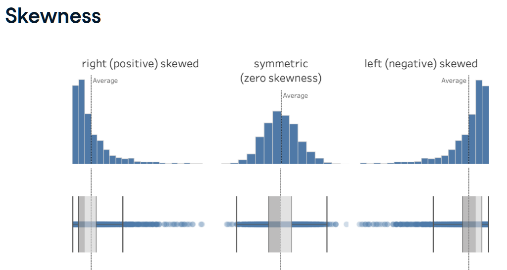


# Box plot

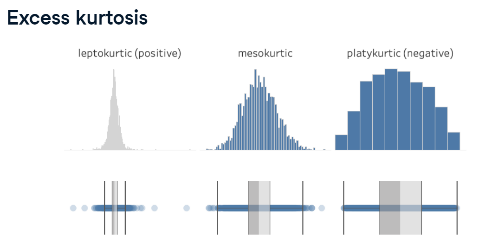
* Standardised chart for distribution of data based on a five-number summary for a Continuous variable
* Spot trends and differences - Compare distributions among multiple categories
* Data distribution, variance, symmetry or skewness, outliers
* Whiskers or Hinges (Tableau) - Minimum, Maximum, Median, First quartile, Third quartile
* Does not show average (mean)
* Disaggregate data - Analysis Menu
* Disaggregate data - Drag to Details Marks
* Sort by different fields

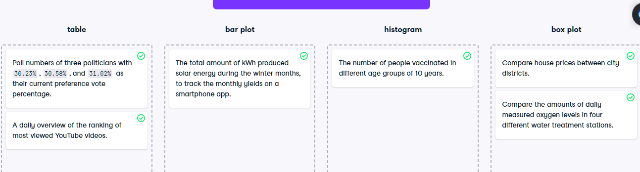


# Skewness

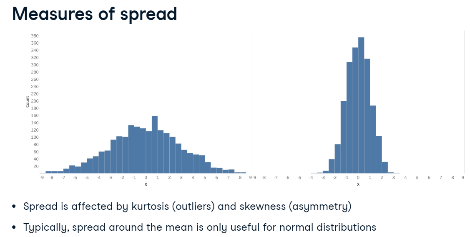


# Kurtosis (Variance)

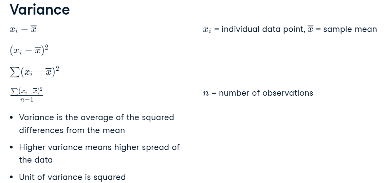




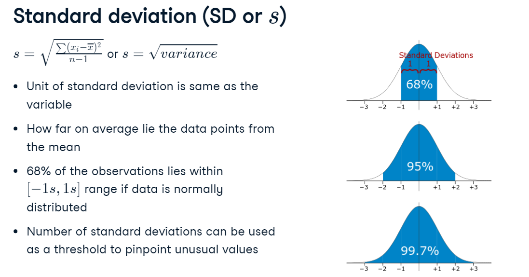
# Spread



# Variance (Kurtosis)



# Standard Deviation

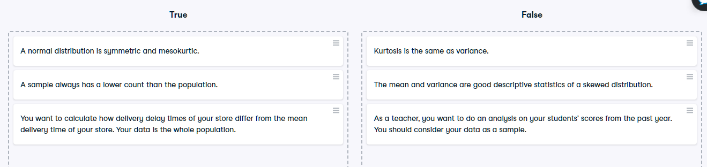


* 68% lies within 2 Sd from mean (1 Sd on each side)– Normal distribution
* Sd is affected by outliers
* Leptokurtic (many outliers) – IQR (50% of data) – Better spread measure (not influenced by outliers)

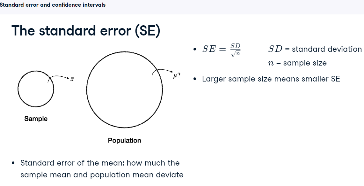
# Spread – Sample vs Population

* Tableau considers as sample by default
* Difference denominators
* Sample: n-1
* Population: N





# Standard Error (SE)



# Confidence Interval (CI)

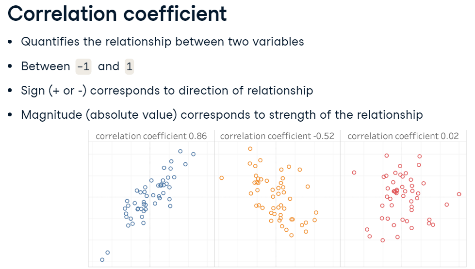


* Sample Mean – Variance – Sd – SE – CI
* Edit All Trend Lines - Show Confidence Band

# Scatter plot

* Relationship between 2 variables
* Variable of interest is placed on Y-axis by convention
* 2 ways to disaggregate your data.
* Drag unique row identifier to Detail mark – then drag pills to the Rows an Column shelves (aggregates points with more then 1 measurement)
* Analysis menu – turn off Aggregate Measures - safest way - each row represented as a single point (does not require a unique identifier - but impossible to exclude data points if leaving out extreme values)

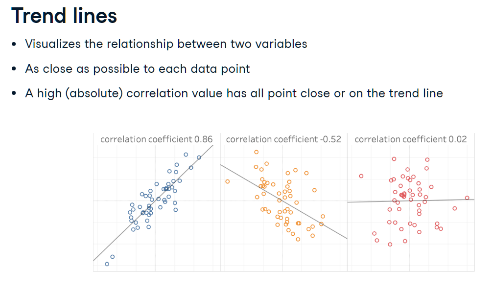
# Correlation

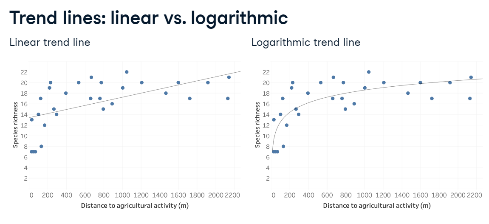


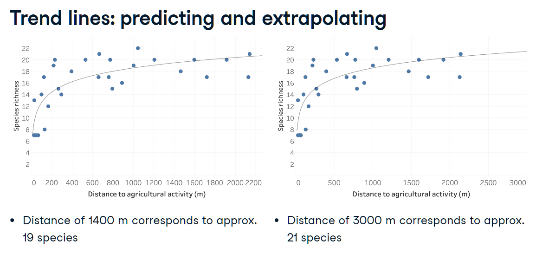


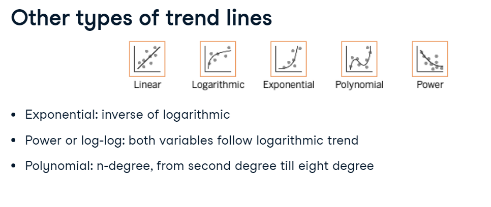
# Trend Lines

* Right click Canvas – Show Trend line
* Analytics Menu – Trend line
* Selecting parts of data points changes Trend line

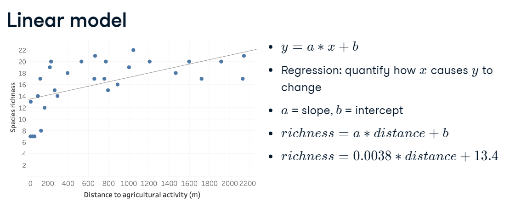


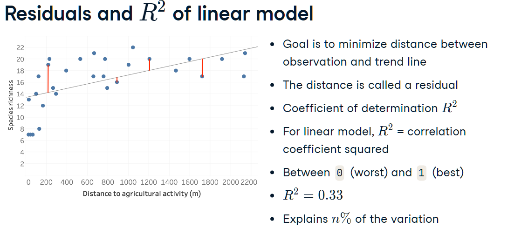




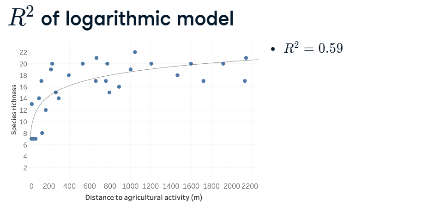


# Assessing Trend Lines



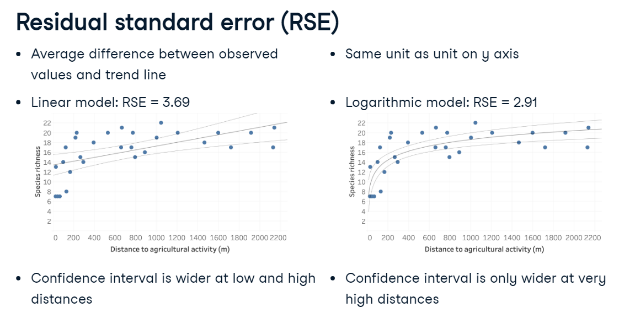


* 33% of variation in richness is explained by the distance – poor fit



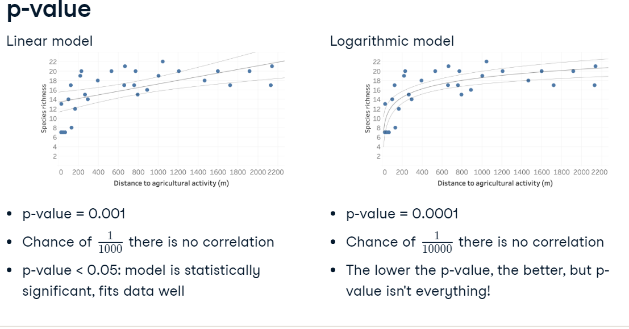
* 59% of variation in richness is explained by the distance – better fit
* Tableau shows R squared as natural log – should be exponentiated

# Residual Standard Error (RSE)

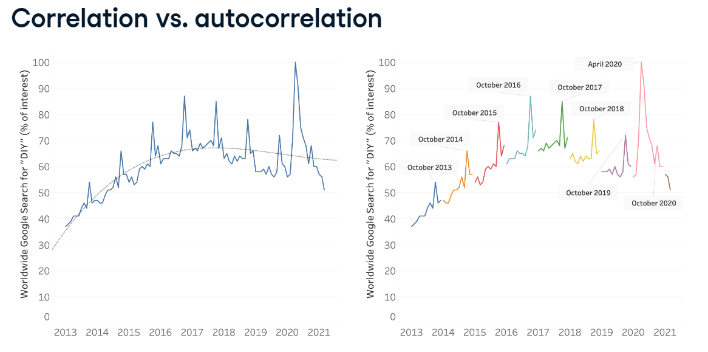


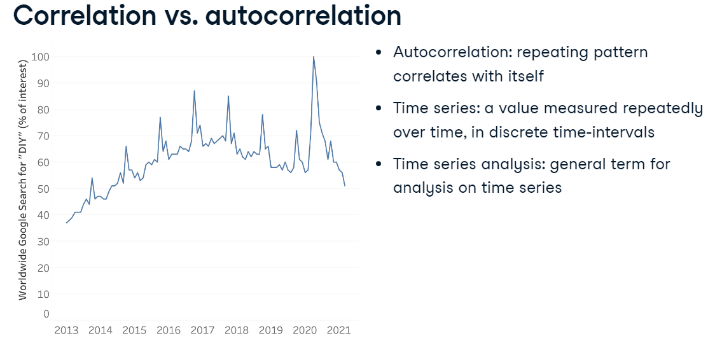
* Linear Model RSE 3.69 - differs 3 – 4 species from the observed value
* Logarithmic Model RSE 2.91 – differs 2 - 3 species from the observed value

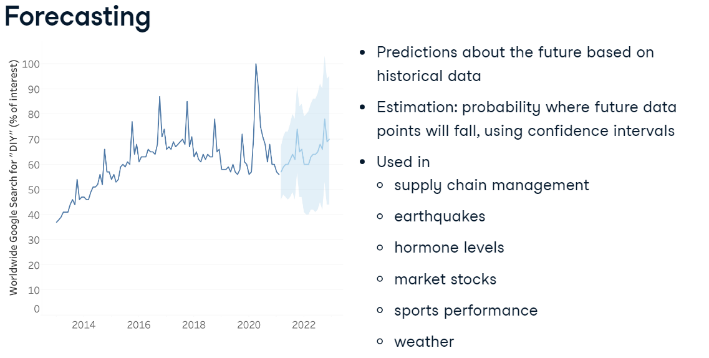
# p-value

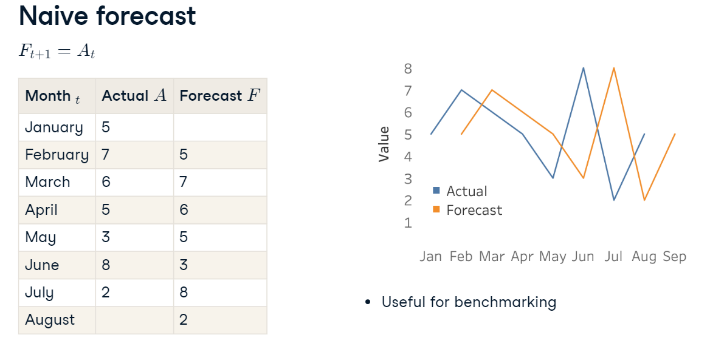


# Forecasting

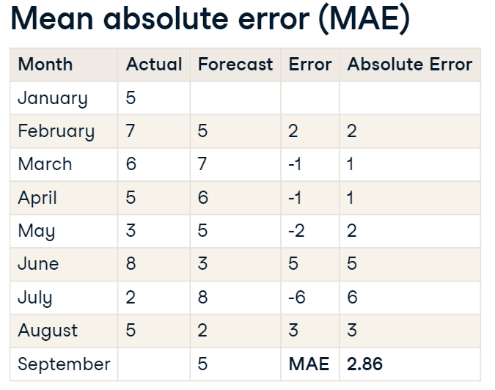


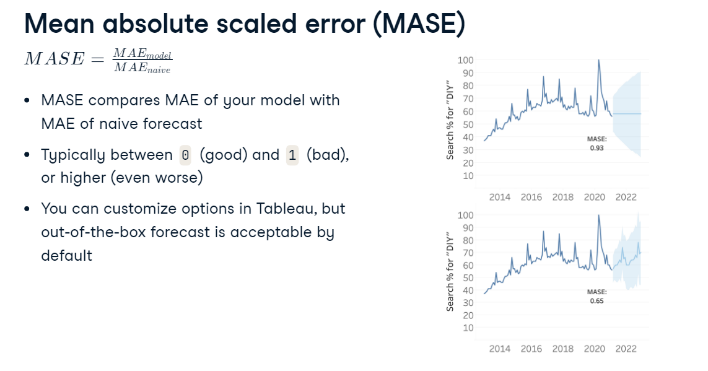












* Change Forecast Indicator to ATTR

# k-means Clustering

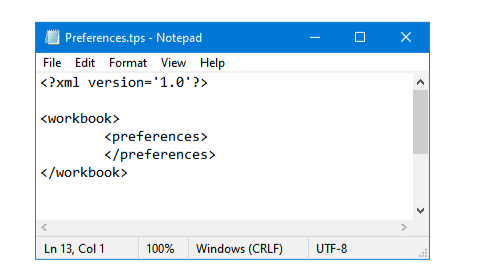
* Clustering is an example of unsupervised machine learning.
* E.g. Differentiating between tissue types or segmenting customers in different groups
* Don't know beforehand what tissue types or customer groups are present in the data
* Let the unsupervised algorithm figure out which data points are more similar to each other
* Out is a set of clusters.
* k-means clustering can be applied to one, two, or more variables – k = number of clusters you want
* E.g. k = 3 if you want to split your data into three categories good, medium, and bad
* Also let Tableau suggest number of clusters if k isn't specified
* All distances between the random centres and each data point are measured
* Each data point is assigned to the centre it is the closest to - each centre is then moved to the new centre of the points assigned to it
* The process is iterative: all distances between each data point and the new centre are measured again, and the data points are assigned accordingly
* Once the centre stops moving between iterations, the final clusters are set
* Can’t assess the quality of the clustering result by comparing actual and predicted values
* Two metrics are used to assess the clustering algorithm: between-group sum of squares and within-group sum of squares
* Between-group sum of squares measures the separation between the clusters as the sum of squared distances between each cluster’s centre, and the average value of the data set - The larger the value, the better the separation between clusters
* Within-group sum of squares quantifies the cohesion of clusters as the sum of squared distances between the centre of each cluster and the individual data points in the cluster - The smaller the value, the more cohesive the clusters
* Create Matrix – Drag same values to Rows and Columns – Disaggregate



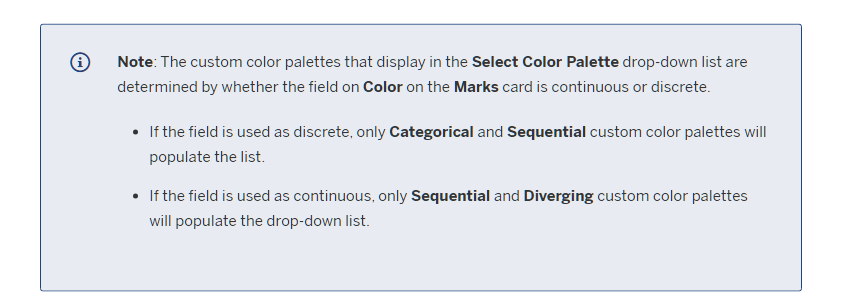


# Custom colour palettes

* Documents - My Tableau Repository - Preferences.tps (text editor)
* Between the opening and closing workbook tags, add an opening preferences tag, <preferences>, and a closing preferences tag, </preferences>



* Categorical color palette contains several distinct colours that can be assigned to fields that behave in a discrete fashion
* Sequential colour palette shows a single color, varying in intensity, used for continuous fields (measures)
* Diverging palette visualizes two ranges of numerical values using color intensity to show the magnitude of the number and the actual color to show which range the number is from - difference between positive and negative numbers

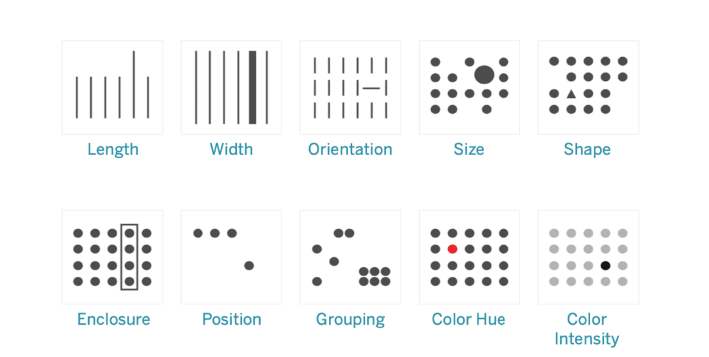


# Custom shapes

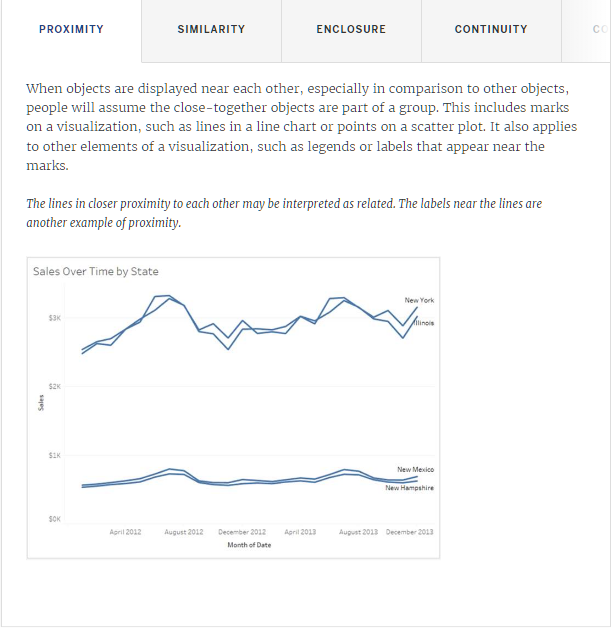
* Documents – My Tableau Repository - Shapes – Create new folder
* Place transparent shapes in new folder – 32X32 pixels recommended

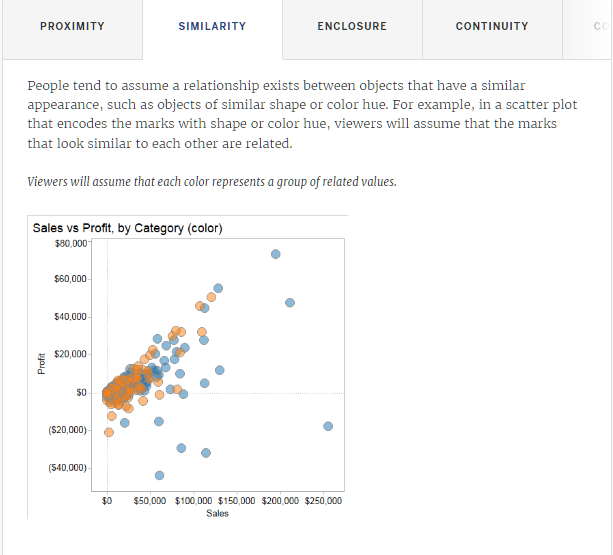
# Pre-attentive attributes

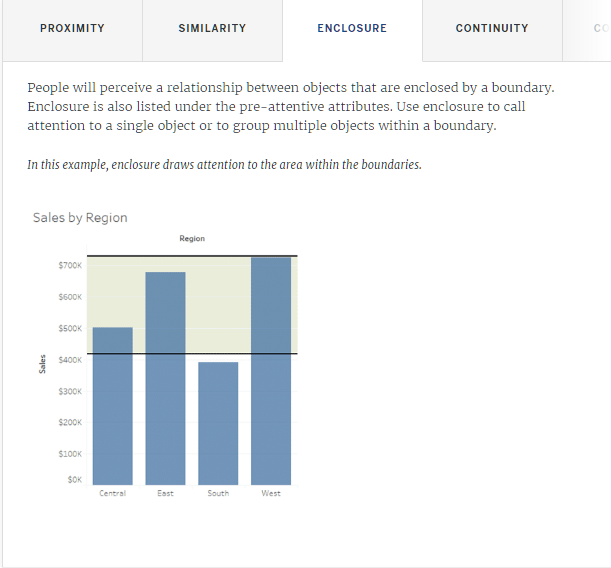
* Visual analytics leverages pre-attentive attributes to guide you down the most useful paths.
* "Pre-attentive attributes" refer to information we can process visually almost immediately, before sending the information to the attention-processing parts of our brain.

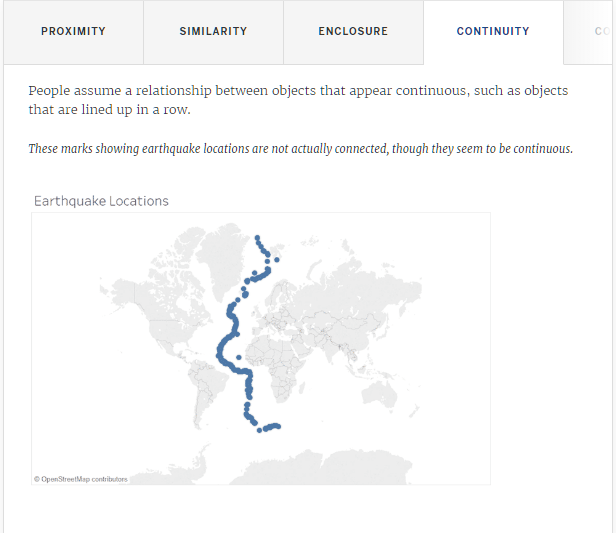


# Gestalt principles - Five principles of grouping







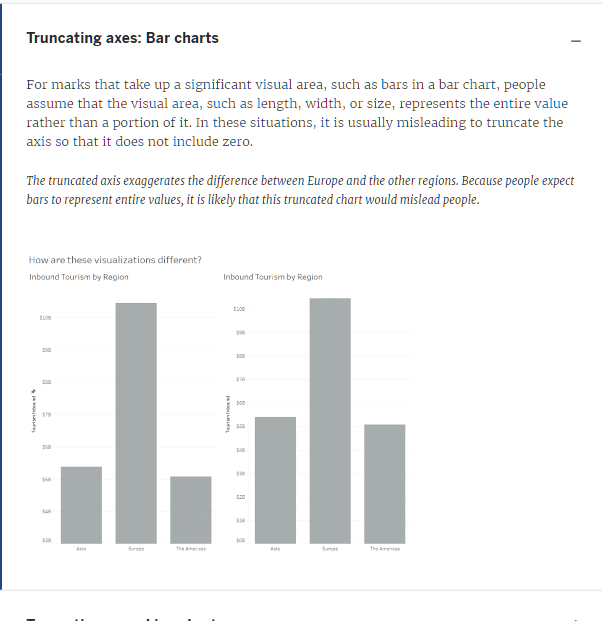




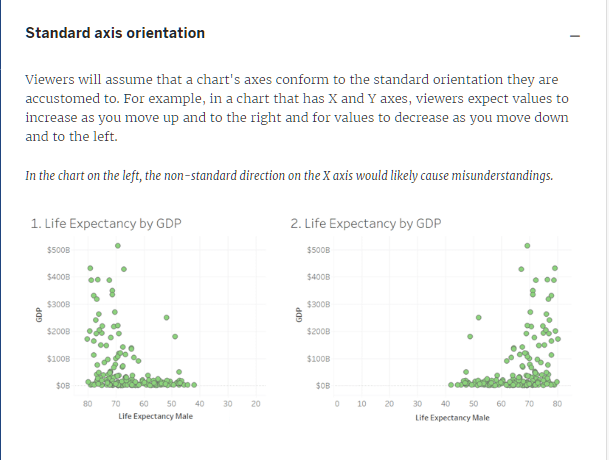
# Considering the visual area of marks

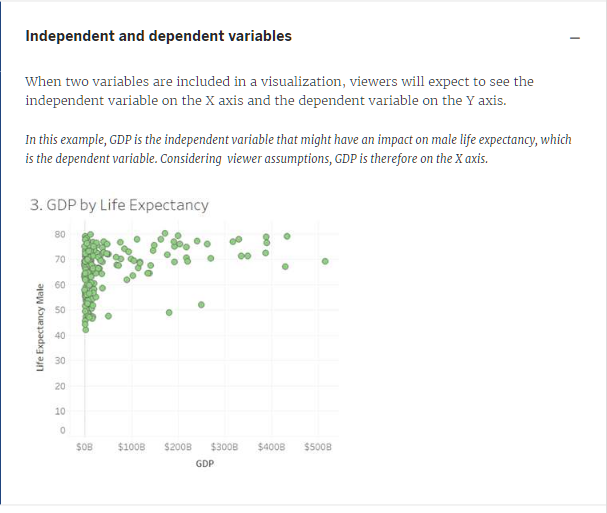
* When interpreting length, width, and size, people will assume that the visual area of a mark represents its magnitude
* They will assume that an increase in visual area reflects an increase in a quantitative value

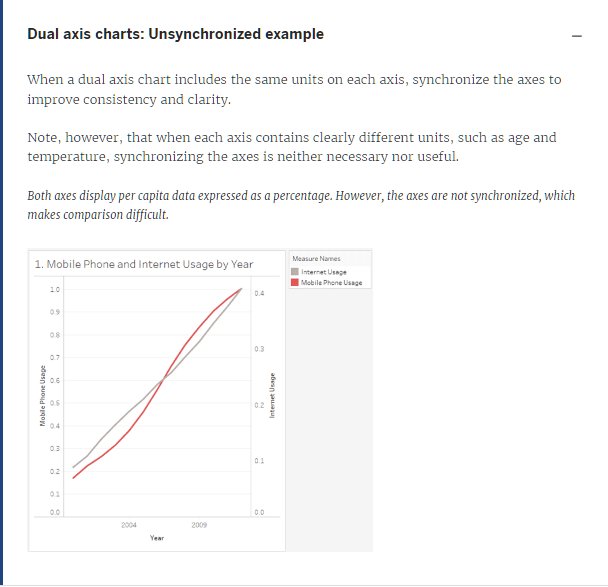
# Displaying and configuring axes

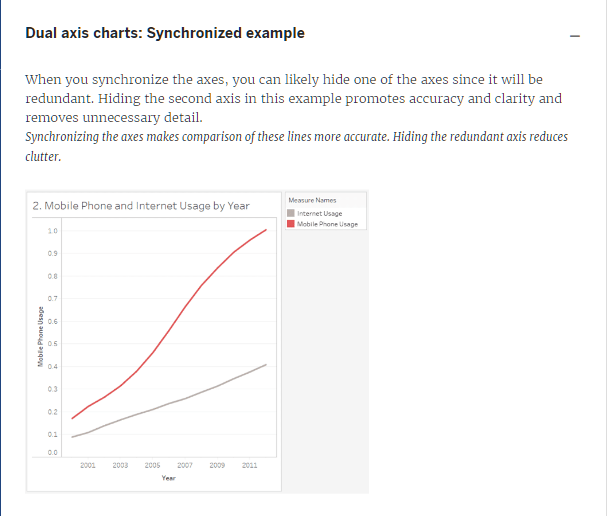
* Most viewers have standard expectations about how axes behave in a chart
* Their experience leads to assumptions about how axes use visual area to convey information

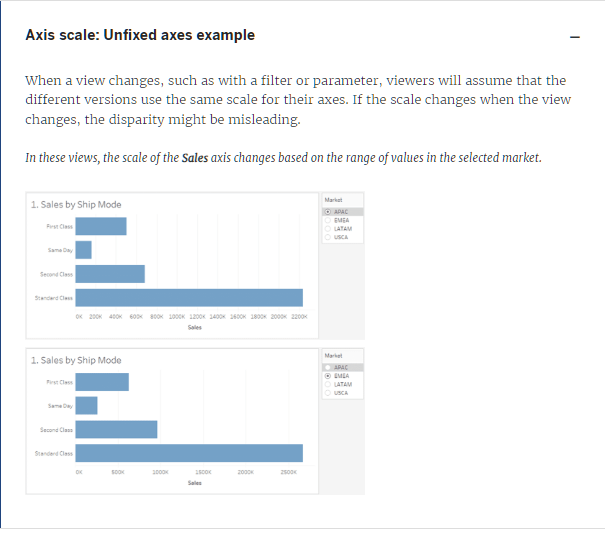


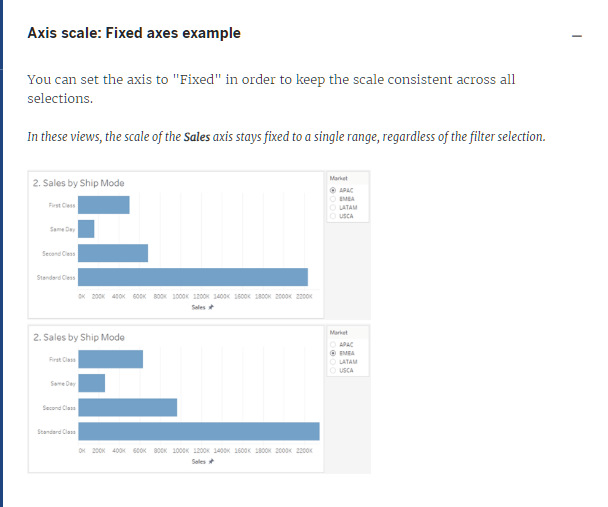












# Considering expectations about color

* Cultural and natural assumptions
* Natural expectations
* Unintentional color relationships

# Tableau Prep Builder

* Use Data Interpreter for Excel files
* Sub-tables
* Filed names
* Merged cells
* Empty columns or rows