# DATA VISUALIZATION – WEEK 2

## TABLEAU DASHBOARD

A dashboard is a consolidated display of many worksheets and related information in a single place. It is used to compare and monitor a variety of data simultaneously. The different data views are displayed all at once.

While creating a dashboard, we can add views from any worksheet in the workbook along with many supporting objects such as text areas, web pages, and images. Each view you add to the dashboard is connected to its corresponding worksheet. So when you modify the worksheet, the dashboard is updated and when you modify the view in the dashboard, the worksheet is updated.

The purpose of a dashboard is to tell a story. Knowing the purpose of the story that you are telling is imperative to making it memorable. Some of these attributes are as follows:

- They're simple and focused.
- They create an emotional response.
- Listeners can relate to the actors in them.
- They have good, real data.

### Best practices for dashboard design

Dashboard design should highlight data and decoration should be limited, to add to a reader's understanding. Much has been written about dashboard design for data visualization.

The following are some key points that you should consider when designing dashboards:

- **Keep it simple**: Your dashboard needs to answer only one question. If you can construct it so elegantly that it answers several questions, that's great, but according to Stephen Few, the maximum number of data points that any of us can remember at once is three.
- **Keep in mind the audience**: Consider what devices will be used to read or consume the data story and content.
- **The method of consumption**: Will users interact with virtualization on their computer, mobile phone, or tablet? Size your dashboard accordingly.
- **Performance**: Like any other web application, the amount of time it takes for a consumer to get what they need is a major factor of performance. Design your visualizations and data sources so that they perform efficiently.
- **Start big and end small**: Always place the most aggregated (macro) data points or summary metrics on the upper left, and guide users for granular, actionable data points.
- **Use colors, graphics**, and fonts that are appropriate for your subject matter and consumers. Choose a simple, single color or complementary palette colors.

The following objects can be added to your dashboard:

- Horizontal layout containers: You can drag worksheets that you want next to each other into these containers.
- Vertical layout containers: You can drag other objects and worksheets that you want to stack from top to bottom into these containers.
- Text objects: These can be used to add titles and calls to action. Though you can't add field tokens to text objects, you can add parameter tokens, which will be discussed in depth in the next chapter.
- Images: You can use these to browse logos or branding elements that add richness to your dashboard. You can also use images as links to web pages, since each image can have a URL attribute.
- Web page objects: These can be used to add content from the Internet.
- Blank objects: These objects can be used to control space. Blank containers are transparent. So, the background colors that you are using will show through them.

*Exercises on Dashboard based on sales superstore and Co2 emission datasets.*

# MAPS VISUALIZATION

As we saw last week, with Location-Based Data, Maps Can Help Tell the Story based on the location. Maps are more than just a new layout for your data; they're topical. People make associations with different geographies, so by putting data on a map, you're already telling a story.

## Why maps matter

Mapping lets you see the implications of your data in ways not detectable on a standard spreadsheet, linear graph or pie chart. Since we are all familiar with maps, using maps in visualization immediately orients your audience to the data. Maps provide context that leads to better ways to prioritize, plan and execute your objectives.

Does one product category outshine another in a specific state or region? Which category had more sales in one region in the super sales data we had been working with? Which countries emit most Co2 per capita?

Answering questions like these becomes fundamentally easier and more accurate when you see your data on a map. Incorporating these maps into dashboards is where value skyrockets. One of the most powerful visual techniques is to relate different views of information. So when you combine your maps with time trends and other charts on a dashboard, the result is a much richer understanding of your information.

Incorporating maps into your day-to-day analysis, reports and dashboards requires the following capabilities:

- Create maps fast. Insist on making maps as quickly as you create a bar chart.
- Add relevant data for depth. Overlay related data and demographic details for robust analysis.
- Drill into maps for answers. Real-time exploration reveals answers to detailed questions.
- Customize with ease. Add specialized maps when that meets your needs.

*Exercises based on sales superstore and Co2 emissions datasets.*

# TABLEAU CALCULATIONS

The data sources that you query to tell stories do not always have all the data points that you need. Tableau Public's data engine allows you to create new mathematical calculations of varying complexity, from basic multiplication functions to sophisticated aggregations with specific levels of detail that add color, context, and key insights to your data stories.

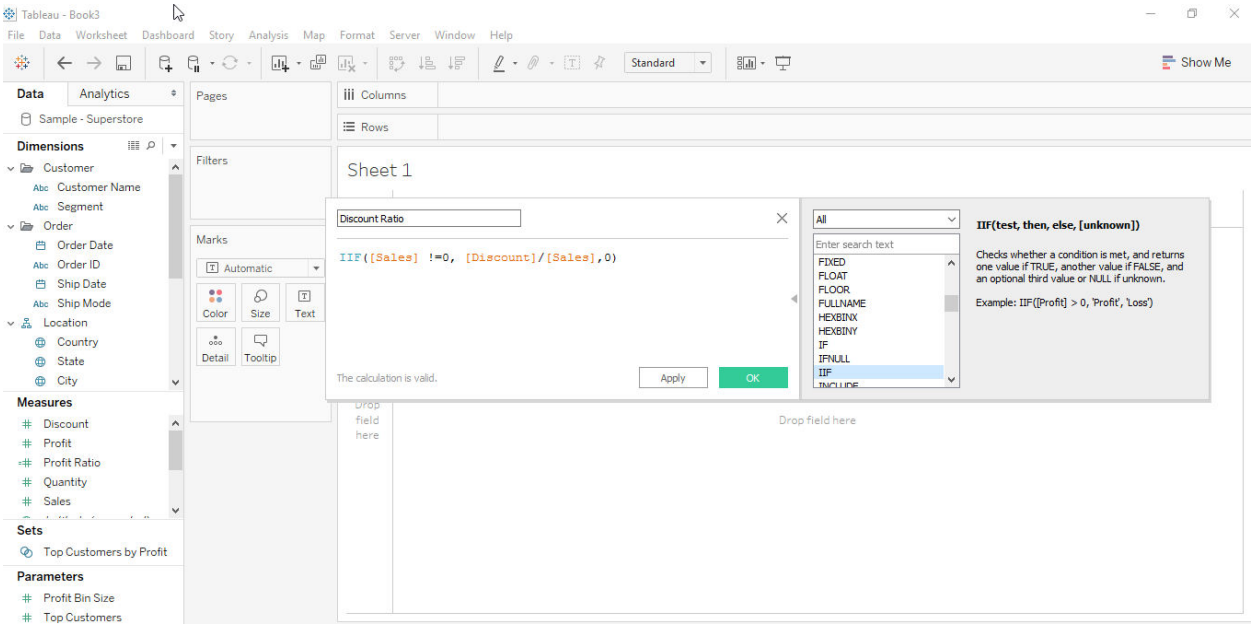
## OBJECTIVES

- To create and edit calculated fields.
- Types of calculations.
- Number functions
- String functions
- Date functions
- Type conversions
- Aggregate functions
- Logic functions

**Calculated Fields:** A calculated field that you create exists only in your data source in Tableau software; it does not exist in the file that you originally queried.

If your underlying data doesn't include all the fields you need to answer your questions, you can create new fields in Tableau and then save them as part of the data source. You created calculated fields in Tableau by defining a formula that is based on existing fields and other calculated fields, using standard functions and operators.

You create a calculated field as below.




**Types of calculations:** There are several types of calculations that you can create. Some of them are common to other data management tools, but some, such as user calculations, are unique to Tableau Public. When creating calculated fields, the functions that you can apply on a field depend on the type of that field.

The different types of functions that you can perform in Tableau Public are as follows:

- **Number**: This includes geometric, trigonometric, and rounding functions, among others.
- **Strings**: This includes options to find characters, measure string length, split and parse fields, and match character strings.
- **Date**: This includes functions for duration, the addition and subtraction of dates, the truncation of dates, and the identification of date parts.
- **Type conversion**: This allows you to convert fields into different types without modifying the metadata of the source field.
- **Logical**: This includes the powerful IF, CASE, and ISNULL and allows you to tell Tableau Public how to group fields or relate them to parameters.
- **Aggregate calculations**: These are the most commonly used functions; they include sum, count, avg., min, max, and median.
- **User functions**: These are more commonly used for Tableau Server rather than Tableau Public; they allow you to create calculated fields that operate on usernames or groups.
- **Table calculation**: This will be detailed in the next chapter; it allows you to compute aggregations based on data points that are visible in the current window.

**Number Function:** Number functions include the following, in order of decreasing precedence:

- **ABS**: This takes the absolute value of a number.
- **CEILING**: This rounds up a decimal to its nearest integer of equal or greater value.
- **FLOOR**: This rounds a decimal down to its nearest integer of equal or lesser value.
- **MAX** and **MIN**: These take the maximum and minimum values of an expression across all records or maximum and minimum of two expressions for each record.
- **ROUND**: This allows you to specify how many decimals to round up a float number to.
- **ZN**: Returns expression if it is not null, otherwise returns zero.
- **DIV**: This produces the whole number product of a division statement.
- **PI**: This produces the numeric value of pi.

**Date Time Function:** The most commonly performed analysis is the measurement of change over time, and the prevalence of date functions is high. The date functions include the following functions:

- **DATEADD**: Adds an increment to the specified date and returns the new date. Th increment is defined by the interval and the date_part.
- **DATEDIFF**: This calculates the elapsed number of specified dateparts between two dates.
- **DATENAME**: This produces the name of the specified datepart of a date. An example that demonstrates the implementation of this function is, DATENAME ('month', #7/21/2015#) = "July".
- **DATEPARSE**: This turns a string into a date in the format that you specify.
- **DATEPART**: This is like DATENAME. It returns the datepart of a specified date but returns a numeric value. An example that demonstrates the implementation of this function is, DATEPART ('month', #7/21/2015#) = 7.
- **DATETRUNC**: This rounds up a date to the first date of the datepart that you specify. An example that demonstrates how to implement this function is, DATETRUNC ('month, #7/21/2015#) = '7/1/2015'.
- **DAY**: This returns the day of the month of a given date. An example that demonstrates how to implement this function is, DAY (#7/21/2015#) = 21.
- **ISDATE**: This tests whether a string that you enter is a date and produces a Boolean value (true or false).
- **MONTH**: This produces the numerical month of a date.
- **TODAY** and **NOW**: These produce the date and datetime of the current moment by using the time settings on your computer.
- **YEAR**: This produces the year of the current date.

## String Functions:

- **CONTAINS**: This has a Boolean output. It tests whether a field contains the specified string.
- **FIND**: This finds the place where a string of characters is located within a field.
- **LEFT, MID, and RIGHT**: These often work in conjunction with FIND and LEN when extracting fixed or variable strings of characters from a field.
- **LEN**: This produces the length of a field.
- **MIN** and **MAX**: These are commonly used on numeric fields. They produce the numerically minimum or maximum values in sequences respectively.
- The **REGEX** expressions: These features were introduced in Tableau Public 9.0. They extract, match, and replace variable strings within fields and are similar to the LIKE function in ANSI SQL.
- **REPLACE**: This replaces a sequence with a specified value.
- **TRIM, LTRIM, and RTRIM**: These trim leading or lagging spaces from a string.
- **UPPER and LOWER**: These are commonly used to normalize the contents of a field.

## Aggregate Functions:

- **SUM**: Returns the sum of all the values in the expression. SUM can be used with numeric fields only. Null values are ignored.
- **Average**: Returns the average of all the values in the expression. AVG can be used with numeric fields only. Null values are ignored.
- **Median**: Returns the median of a single expression. MEDIAN can be used with numeric fields only. Null values are ignored.
- **COUNT** and **COUNTD**: These count the number of dimension members and the number of distinct dimension members respectively; they are typically performed on dimensions rather than measures.
- **Minimum** and **Maximum**: Returns the minimum and maximum of a single expression across all the records or the maximum of two expressions for each record.
- **Percentile**: Aggregate calculation that returns the percentile value from the given expression corresponding to the specified number. Valid values for the number are 0 through 1.
- **Standard Deviation**: This is the square root of the variance for a data set, and it is the unit of measurement of distance from the mean within a data set; it is commonly represented by the Greek letter, sigma
- **Variance**: This is the average distance from the mean value, and it's used to measure the distribution of a data set.

**Logical Functions:** There are several logic functions in Tableau. Some of these functions are sub functions or parts of others. We will focus on the following major functions:

- **IF, followed by THEN, ELSE-IF, or ELSE**: This tests whether a condition is met and show the result if it is met as well as other conditions that need to be tested and results that need to be produced in case none of the conditions are met.
- The **IIF** statement: Checks whether a condition is met, and returns one value if TRUE, another value if FALSE, and an optional third value or NULL if unknown.
- The **IFNULL** function: Returns expression 1 if not null, otherwise returns expression 2.
- The **CASE** statement: Finds the first value that matches the expression and returns the corresponding result.
- **AND, OR, and NOT**: Perform a logical conjunction for AND, logical disjunction for OR and logical negation for NOT.
- **END**: This is critically important, as it terminates the loops of the IF and CASE statements. Tableau Public will tell you in case you need to add it and have not done so.

**Operators**

Type any necessary operators into your formula. All standard operators such as addition (+), subtraction (–), multiplication (*), and division (/) are supported.

Operators are colored black in the formula.

**Parameters**

Parameters are placeholder variables that can be inserted into calculations to replace constant values. When you use a parameter in a calculation, you can then expose a parameter control in a view or dashboard to allow users to dynamically change the value.

*Exercises based on sales superstore and Co2 emissions datasets.*