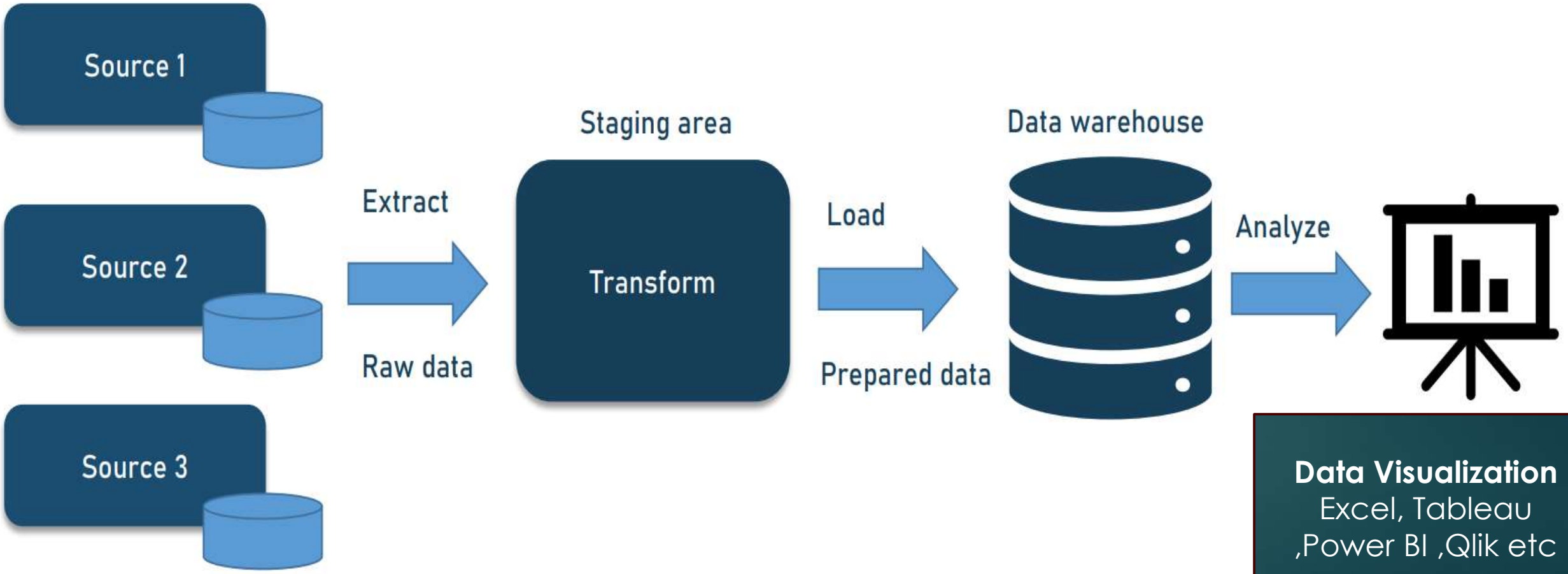
Tableau For Data Analytics

ZERO TO HERO BY NAMASTE SQL

Data Analytics -> The Big Picture



What is data visualization

Data visualization is the process of representing data and information graphically through charts, graphs, maps, and other visual tools.

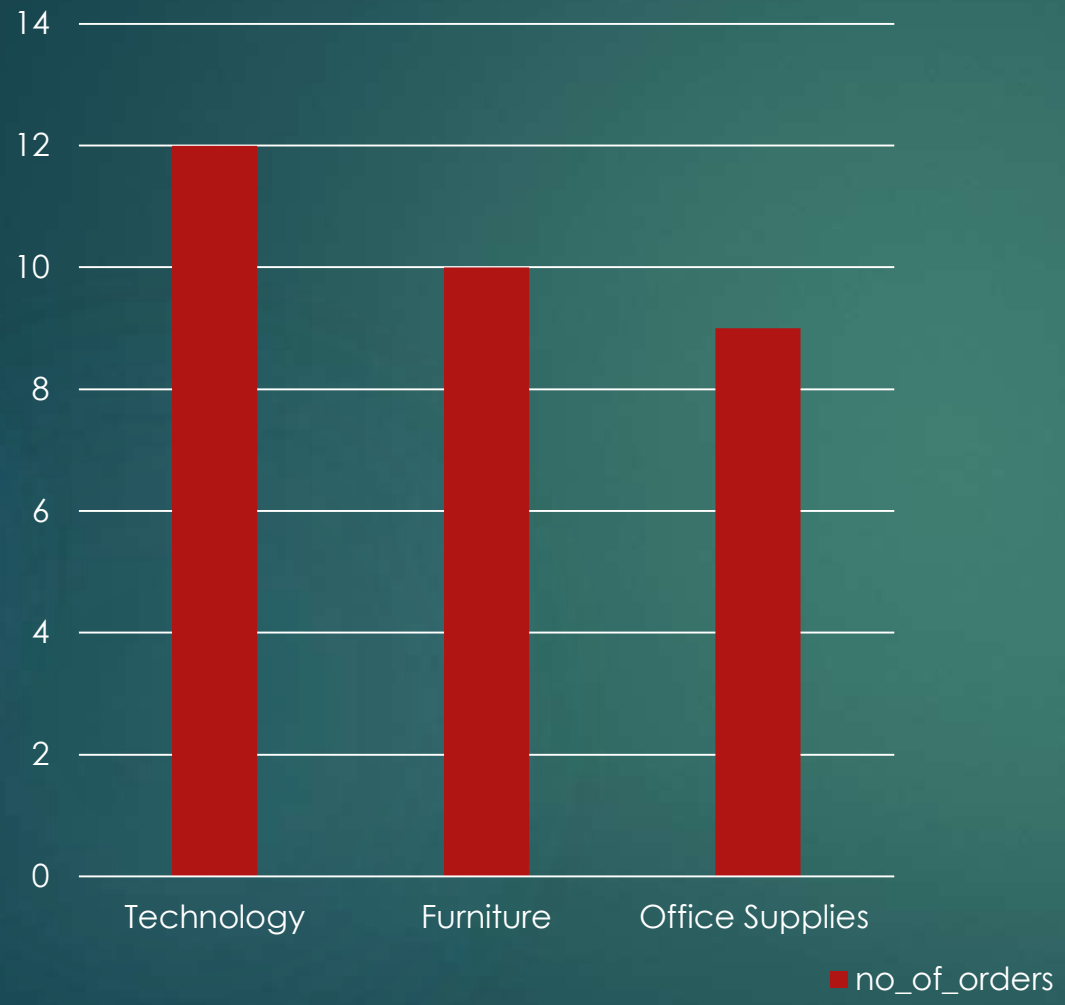
Why it is required

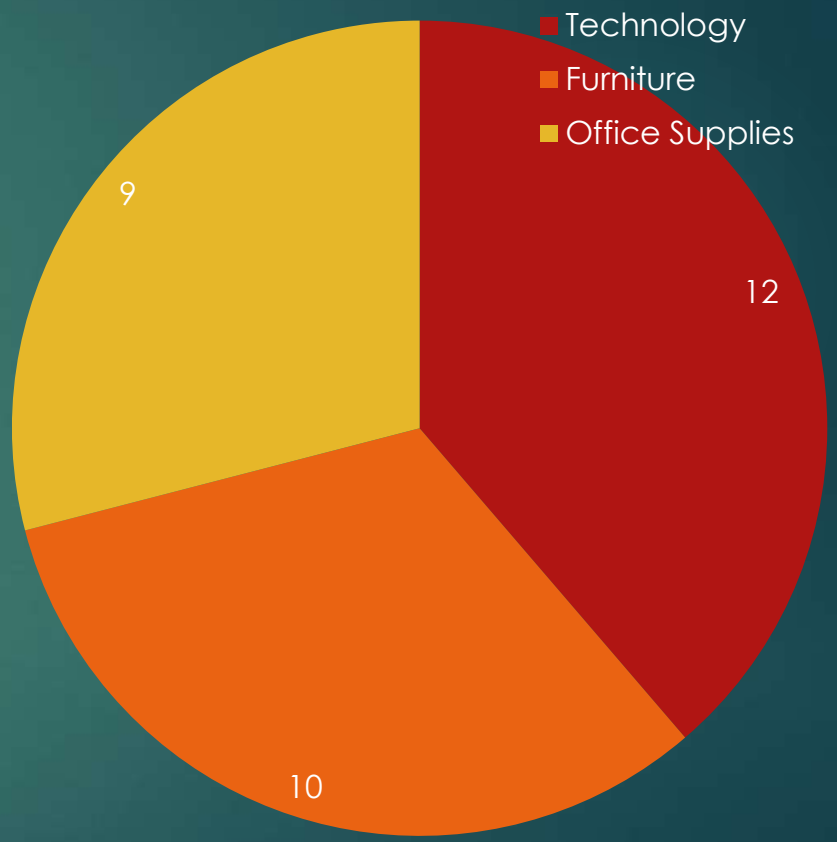
- Simplifies Complex Data
- Finding solution to problems
- Identify trends , patterns and relationships, comparative analysis
- Fast decision making
- Highlights Important Information
- Improves Retention and Recall

Data in raw form

order_id	order_date	category	ship_mode	customer_i	customer_	segment	country	city	state	postal_cod	region
CA-2020-152156	08-11-2020	Furniture	Second Cl	CG-12520	Claire Gute	Consumer	United Stat	Hendersor	Kentucky	42420	South
CA-2020-152156	08-11-2020	Furniture	Second Cl	CG-12520	Claire Gute	Consumer	United Stat	Hendersor	Kentucky	42420	South
CA-2020-138688	12-06-2020	Office Supplies	Second Cl	DV-13045	Darrin Van	Corporate	United Stat	Los Angele	California	90036	West
US-2019-108966	11-10-2019	Furniture	Standard C	SO-20335	Sean O'Dc	Consumer	United Stat	Fort Laude	Florida	33311	South
US-2019-108966	11-10-2019	Office Supplies	Standard C	SO-20335	Sean O'Dc	Consumer	United Stat	Fort Laude	Florida	33311	South
CA-2018-115812	09-06-2018	Furniture	Standard C	BH-11710	Brosina Hc	Consumer	United Stat	Los Angele	California	90032	West
CA-2018-115812	09-06-2018	Office Supplies	Standard C	BH-11710	Brosina Hc	Consumer	United Stat	Los Angele	California	90032	West
CA-2018-115812	09-06-2018	Technology	Standard C	BH-11710	Brosina Hc	Consumer	United Stat	Los Angele	California	90032	West
CA-2018-115812	09-06-2018	Office Supplies	Standard C	BH-11710	Brosina Hc	Consumer	United Stat	Los Angele	California	90032	West
CA-2018-115812	09-06-2018	Office Supplies	Standard C	BH-11710	Brosina Hc	Consumer	United Stat	Los Angele	California	90032	West
CA-2018-115812	09-06-2018	Furniture	Standard C	BH-11710	Brosina Hc	Consumer	United Stat	Los Angele	California	90032	West
CA-2018-115812	09-06-2018	Technology	Standard C	BH-11710	Brosina Hc	Consumer	United Stat	Los Angele	California	90032	West
CA-2021-114412	15-04-2021	Office Supplies	Standard C	AA-10480	Andrew All	Consumer	United Stat	Concord	North Caro	28027	South
CA-2020-161389	05-12-2020	Office Supplies	Standard C	IM-15070	Irene Mad	Consumer	United Stat	Seattle	Washingto	98103	West
US-2019-118983	22-11-2019	Office Supplies	Standard C	HP-14815	Harold Pav	Home Offic	United Stat	Fort Worth	Texas	76106	Central
US-2019-118983	22-11-2019	Office Supplies	Standard C	HP-14815	Harold Pav	Home Offic	United Stat	Fort Worth	Texas	76106	Central
CA-2018-105893	11-11-2018	Office Supplies	Standard C	PK-19075	Pete Kriz	Consumer	United Stat	Madison	Wisconsin	53711	Central
CA-2018-167164	13-05-2018	Office Supplies	Second Cl	AG-10270	Alejandro	Consumer	United Stat	West Jord	Utah	84084	West
CA-2018-143336	27-08-2018	Office Supplies	Second Cl	ZD-21925	Zuschuss	Consumer	United Stat	San Franci	California	94109	West
CA-2018-143336	27-08-2018	Technology	Second Cl	ZD-21925	Zuschuss	Consumer	United Stat	San Franci	California	94109	West
CA-2018-143336	27-08-2018	Office Supplies	Second Cl	ZD-21925	Zuschuss	Consumer	United Stat	San Franci	California	94109	West

Data in the form of charts





Orders data dashboard

Gartner Report

Gartner®

Magic Quadrant for Analytics and Business Intelligence Platforms 2023

What is Tableau?

- Tableau is a powerful data visualization tool used for converting raw data into an understandable format through interactive and shareable dashboards.
- It provides a wide range of visualization options to create charts, graphs, maps, and other visual representations of data.
- Tableau is known for its user-friendly interface, allowing users to quickly and easily connect to various data sources, perform data analysis, and share insights.
- It is widely used in business intelligence to help organizations make data-driven decisions.

Tableau Ecosystem

Tableau offers a variety of products tailored to different needs and use cases in data visualization and business intelligence. Here are the main product offerings from Tableau

1. Tableau Desktop Public
2. Tableau Desktop Professional
3. Tableau Server
4. Tableau Online
5. Tableau Reader
6. Tableau Prep

Download and install Tableau Desktop

1. Tableau Desktop Public: https://www.tableau.com/products/public/download

2. Tableau Desktop Professional : https://www.tableau.com/products/desktop/download

3. Tableau Public: https://public.tableau.com/app/discover

Tableau Data Types

Data Type	Icon
String Values (Text)	Abc
Integer Values (Decimal & Whole Numbers)	#
Date Values (DD/MM/YYYY or MM/DD/YYYY)	
Date & Time Values (DD/MM/YYYY HH:MM:SS)	
Boolean Values (True or False)	T|F
Geographic Values (Region, Postal code etc.)	

Dimensions

Dimensions are qualitative fields that typically contain categorical data. They describe the "what," "who," "where," and "when" of your data. Dimensions are used to segment and categorize data, and they often form the rows and columns of your data table. They help in slicing and dicing your data to drill down into specific details.

- **Examples of dimensions include:**

1. Names (e.g., customer names, product names)
2. Dates (e.g., order dates, birthdates)
3. Geographic locations (e.g., countries, cities)
4. Categories (e.g., product categories, departments)

- **In Tableau:**

1. Dimensions are displayed in blue in the data pane.
2. They are used to create headers, labels, and axes in charts.
3. They are usually discrete (distinct, separate values).

Measures

Measures are quantitative fields that typically contain numerical data. They represent the metrics or quantities that you want to analyze. Measures are used to perform calculations, aggregate data, and provide the values you want to analyze and visualize.

- **Examples of measures include:**

1. Sales revenue
2. Profit
3. Quantity sold
4. Scores or ratings

- **In Tableau:**

1. Measures are displayed in green in the data pane.
2. They are used to create axes and determine the size, length, or position of marks in charts.
3. They are usually continuous (can take any value within a range) and can be aggregated (e.g., sum, average, count).

Discrete vs continues fields in Tableau

- **Discrete fields** are categorical and contain distinct, separate values. They represent things that can be counted in distinct units and are often used to segment and categorize data. Discrete fields are typically used for creating headers/labels. Discrete fields are represented in blue.

- **Continuous fields** are quantitative and represent measurable quantities that can take any value within a range. They are typically used for axes that show a range of values, such as time series, measures, and numerical ranges. Continuous fields are represented in green.

Live vs Extract Connection

Live connection means that Tableau directly queries the data source in real time whenever you create or interact with a visualization. This ensures that you are always working with the most current data.

1. Real-Time Data: Always provides the most up-to-date data.
2. No Storage Required: Data is not stored locally, reducing storage requirements on your machine.
3. Dynamic Updates: Any changes in the data source are immediately reflected in Tableau.

Extract involves creating a snapshot of the data at a specific point in time and storing it in Tableau's proprietary format (.hyper or .tde file). The data is imported into Tableau, and all operations are performed on this static copy.

1. Improved Performance: Extracts are optimized for fast querying, often resulting in better performance, especially with large datasets.
2. Offline Access: Extracts can be used offline without needing constant access to the original data source.
3. Reduced Load on Data Source: Since the data is extracted, it reduces the load on the original data source.

File Types in Tableau

1. **Tableau Workbook (.twb):** This file contains the definitions of the data sources, worksheets, dashboards, and stories but does not contain the actual data. It's a lightweight file that saves references to the data.

1. **Tableau Packaged Workbook (.twbx)**: This is a packaged version of a .twb file that includes the workbook along with any data sources and images used. It's useful for sharing workbooks with others who may not have access to the original data sources.

1. **Tableau Data Extract (.hyper and .tde)**: These files contain a snapshot of your data optimized for fast query performance in Tableau. .hyper is the newer format, replacing .tde.

Tableau Order of Operations

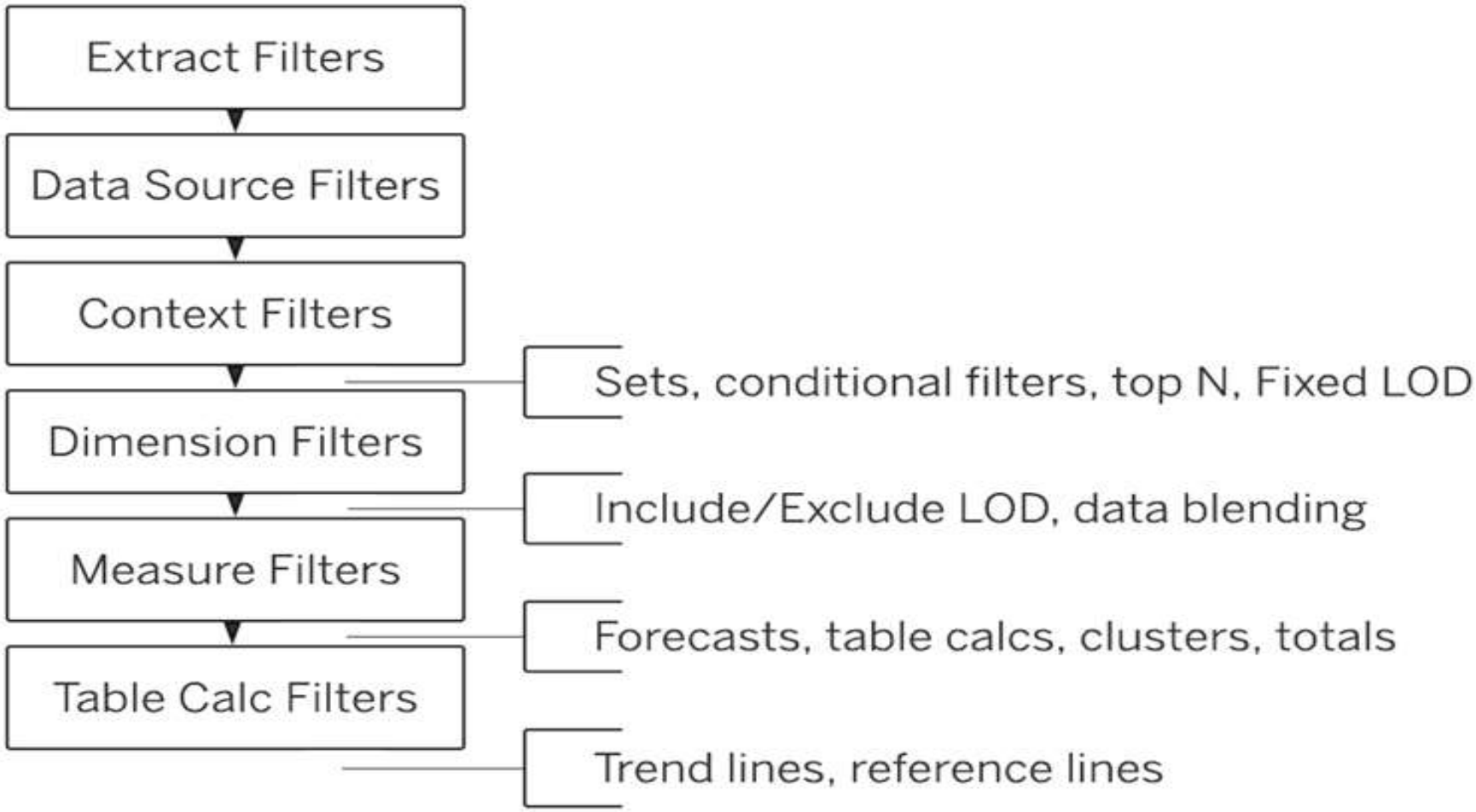


Tableau fundamentals

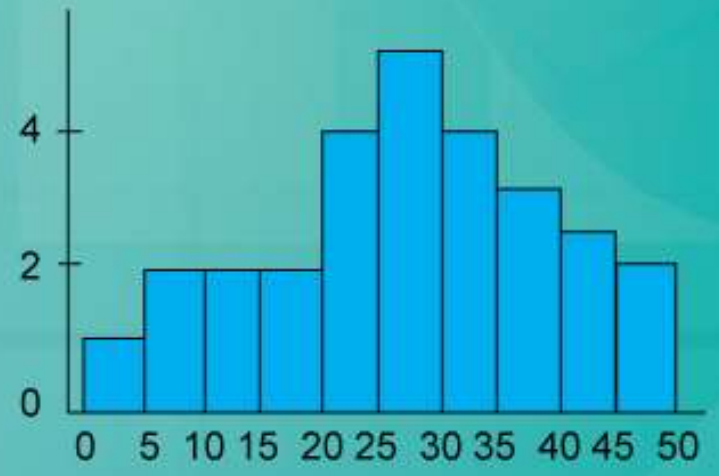
- Filters dimension/measures
- Bar chart / line chart -> formatting
- Marks types
- Data Granularity In Tableau
- Working with dates
- Measure name / measure values
- Calculated columns-> dimensions/measures/aggregated measures
- Dashboard overview and save to tableau public

Chart Types

- Tables/cross tab
- Bar / stacked bar/ side by side bar
- Line / area charts
- Pie chart
- Dual axis chart
- Combined axis chart
- Highlight table
- Histogram
- Scatter plot

Definition

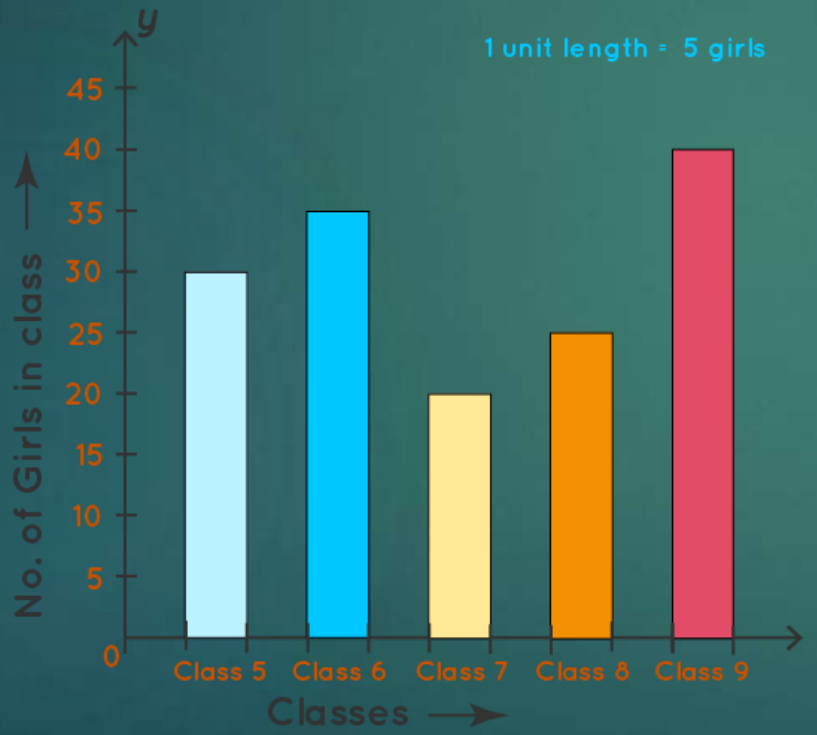
Histogram A graphical distribution of data arranged into discrete groups. Although similar in appearance to a bar graph, a histogram deals with continuous data.



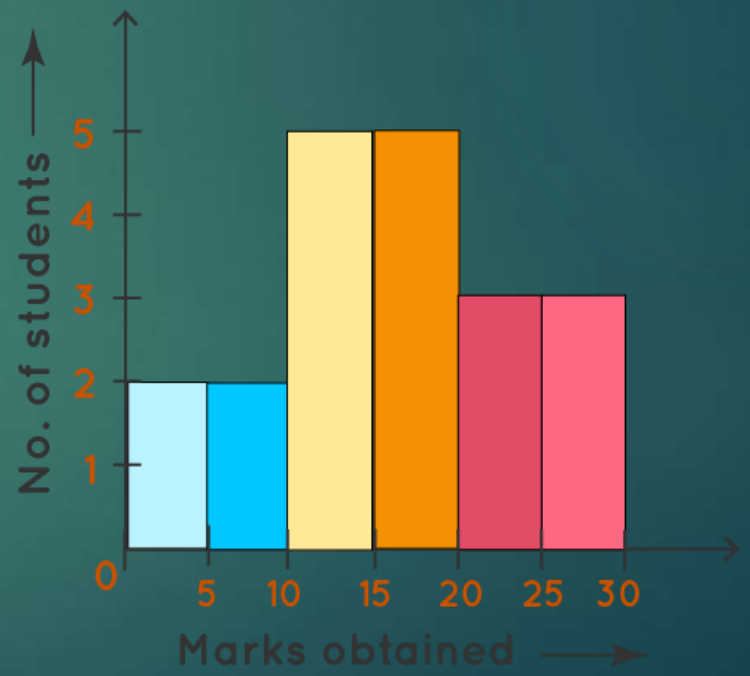
Difference Between Bar Chart and Histogram

cuemath
THE MATH EXPERT

Bar Graph

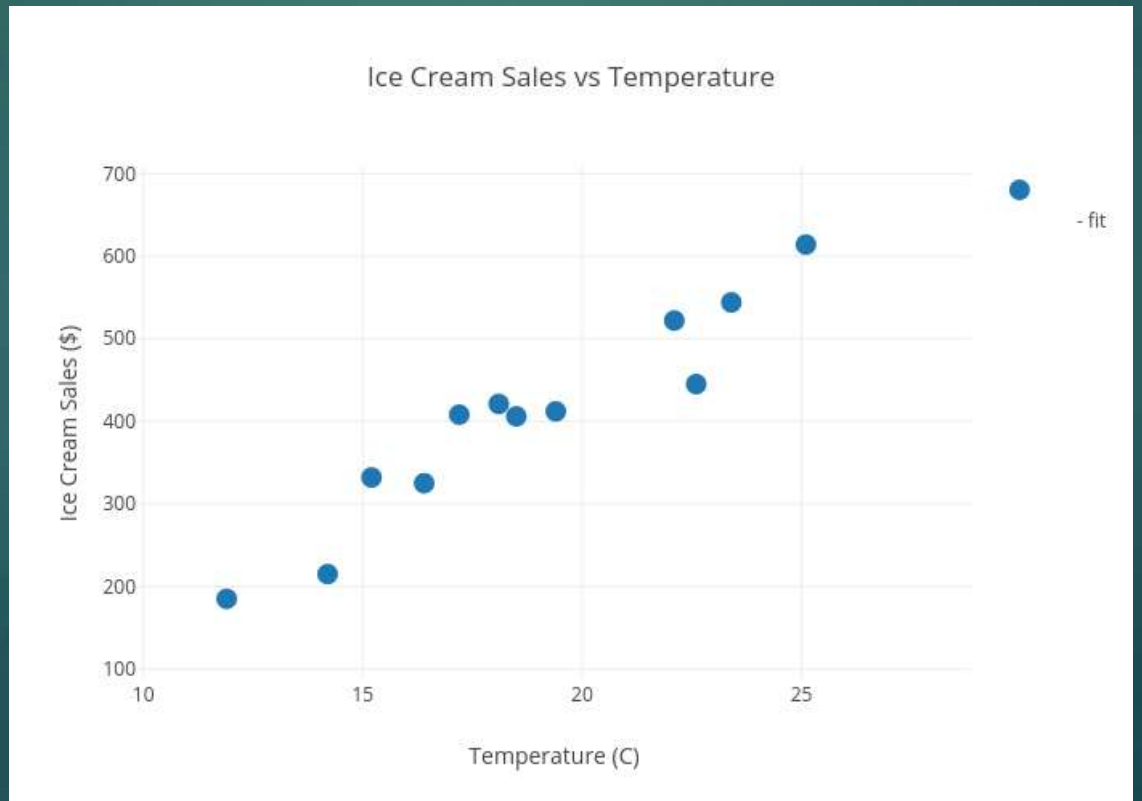


Histogram



Scatter Plots

- A scatter plot (aka scatter chart, scatter graph) uses dots to represent values for two different numeric variables. The position of each dot on the horizontal and vertical axis indicates values for an individual data point. Scatter plots are used to observe relationships between variables.



Calculated fields

1- Row level calculations
2- Aggregated calculations
3- LOD
4- Quick table calculations

Row level:

1. String -> split , contains , concat
2. dates -> datediff
3. Logical -> if else
4. boolean

Parameters

1- Top N /Index
2- Stacked bar chart(sc / region)
3- calculations -> choose dimension

Groups

A tableau group is a set of multiple members combined in a single dimension to create a higher category of dimension.

1. From chart labels / marks
2. From dimension directly
3. Using find option -> product name

Groups are static

Sets

Sets in Tableau are used to create subsets of data based on certain conditions defined by the user. For example, a set can be created for having a subset data of top 10 customers with the highest sales.

1. Fixed sets
2. Dynamic sets
3. Combined sets

Combined sets

1. Top 10 and bottom 10 by profit -> union
2. Top 10 by sales and bottom 10 by profit -> intersect

Hierarchy and drill down

Maps

- filled maps
- Symbol maps
- Custom territory
- Dual axis
- Create role (region)
- Postal code -> give geo role
- Custom geo coding

Unrecognize, ambiguous , manual latitude ,longitude

Table calculations

Table calculations are transformations you can apply to the aggregated values in a visualization. They are a special type of calculated field that computes on the local data in Tableau based on what is currently in the view.

Table calculations are impacted by

- Layout
- Direction (addressing)
- Scope (partitioning)
- Filters

Table calculations rely on two types of fields: addressing and partitioning fields. The key to understanding table calcs is to know how these fields work.

- **Partitioning fields define the scope:** They break up the view into multiple partitions or sub-views. The table calculation is then applied to the marks within each partition.
- **Addressing fields define the direction:** They define the "direction" that the calculation moves (for example, in calculating a running sum, or computing the difference between values).

- Quick table calculations
- Additional table calculations -> index, rank , running sum etc

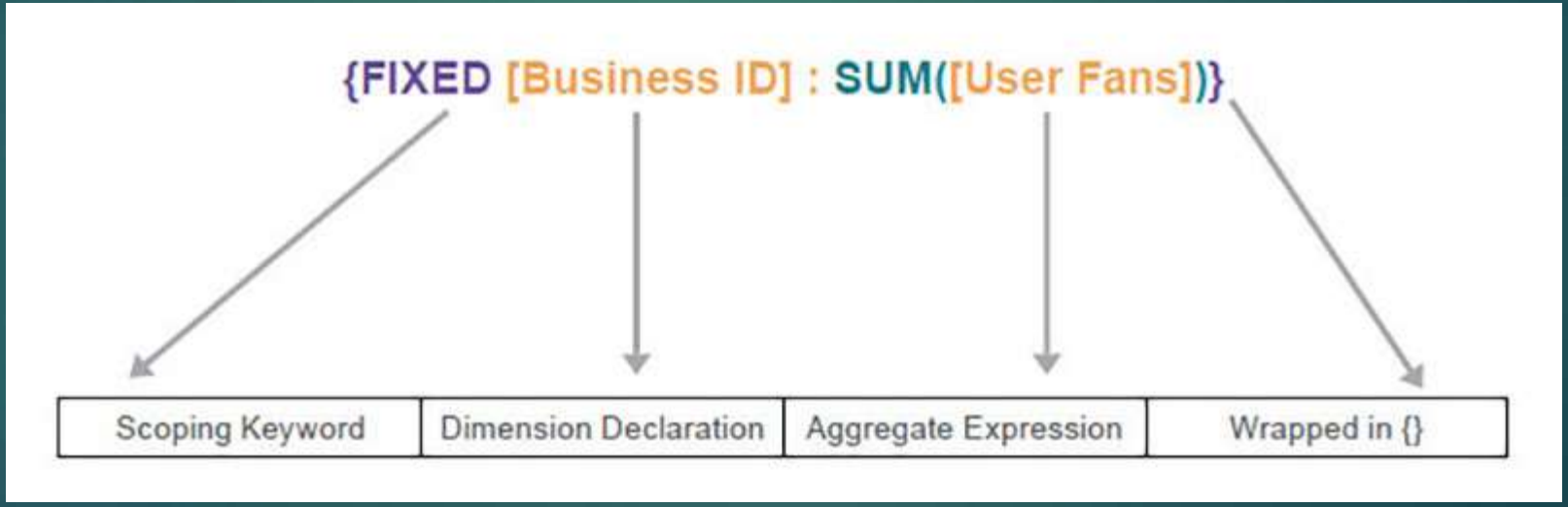
LOD Expressions

Level of Detail expressions (also known as LOD expressions) allow you to compute values at the data source level and the visualization level. However, LOD expressions let you control the granularity you want to compute. They can be performed at a more granular level (INCLUDE), a less granular level (EXCLUDE), or an entirely independent level (FIXED).

Types of LOD

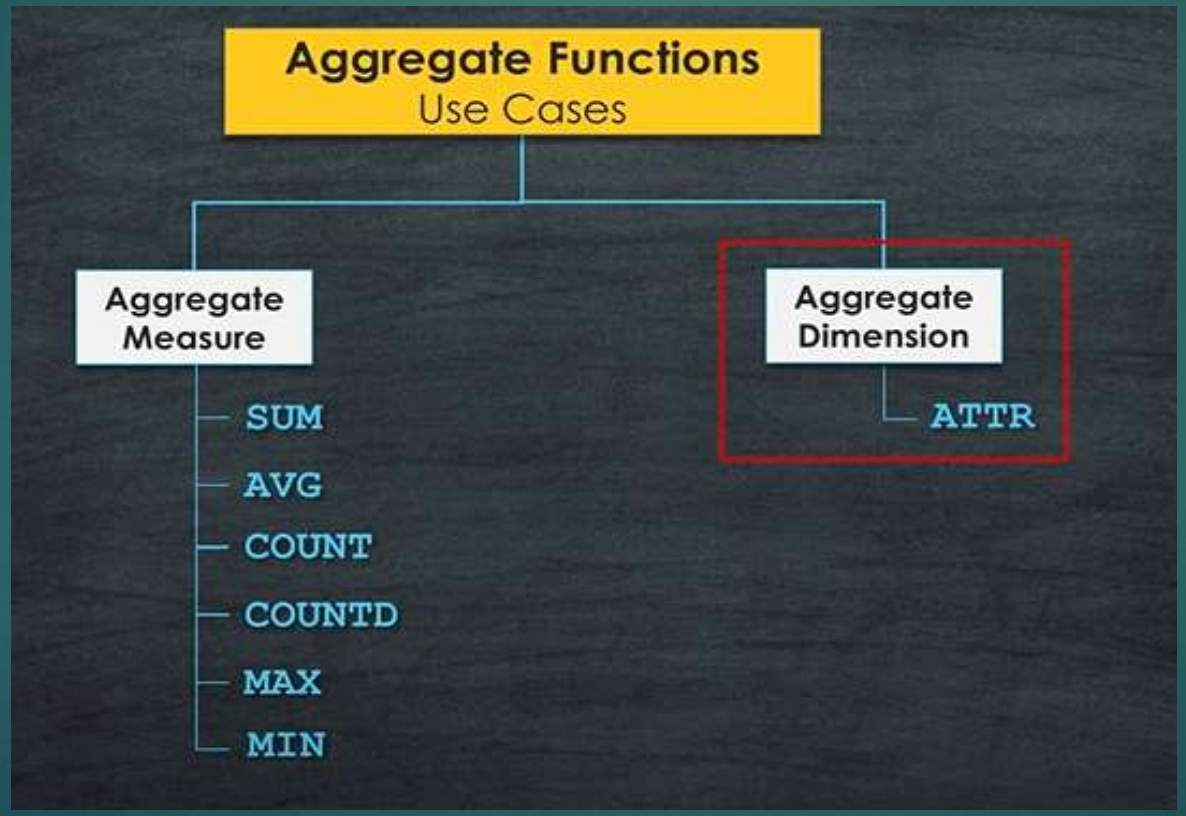
- Fixed : Independent of the view
- Include: Add to view
- Exclude : minus from view

Syntax :

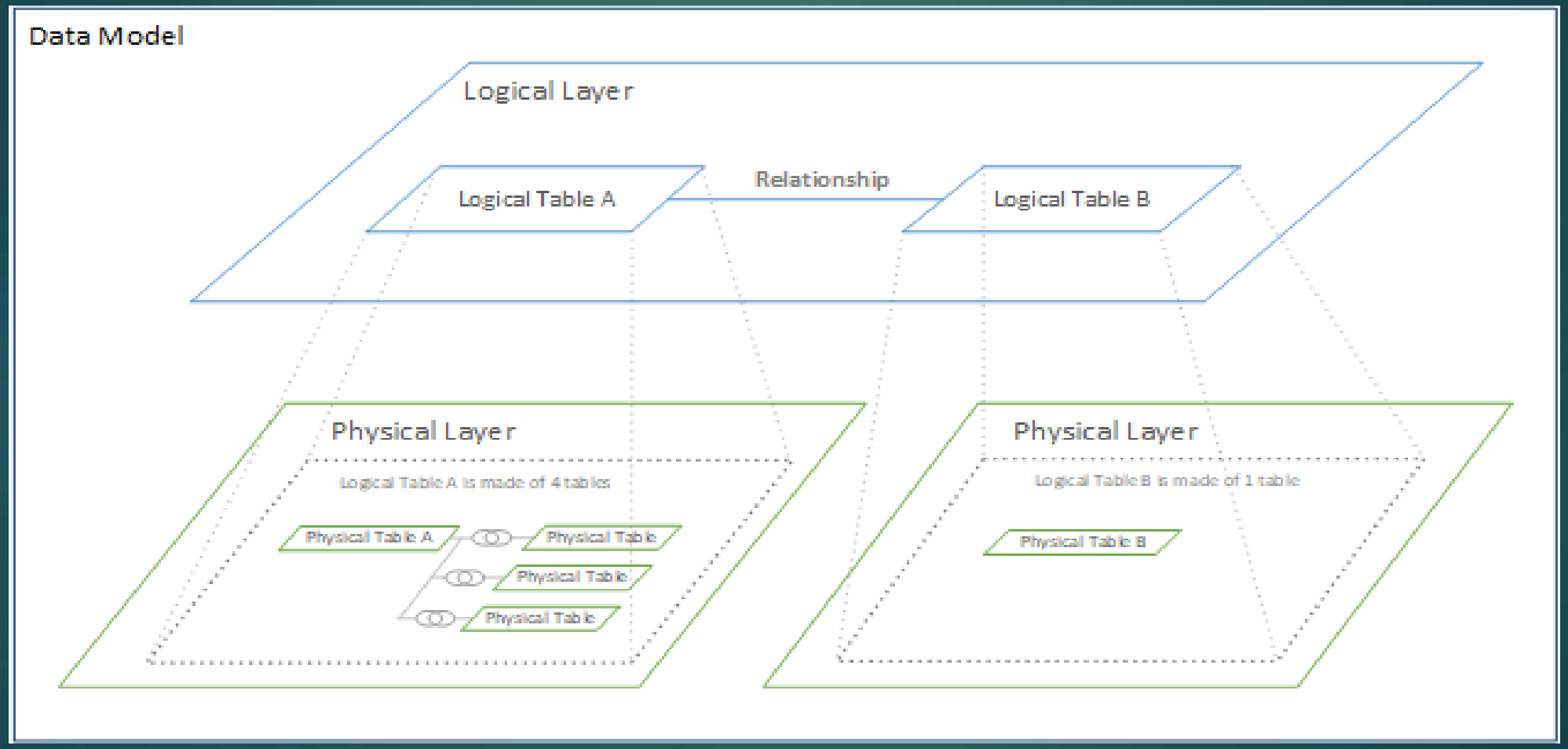


ATTR (Attribute)

- ATTR() is a special Tableau function that effectively returns a record-level result as an aggregation. If there are multiple values of the record-level field in the current context then ATTR() will return * instead of a single value.



Data Model in Tableau



relationship -> joins-> blending

	Blend	Join	Relationship
Visual cue			
# data sources	At least 2	1	1
# connections	At least 1 per data source	Can be multiple	Can be multiple
Structure	Retain original table	New combined form	Retain original table
How tables are combined	Separate queries; linking field	Physical join	Logical; context specific
How numbers are aggregated	Based on primary	Based on join level granularity	Smart; context specific

Union, pivot unpivot

Add a Slide Title - 5