Tableau Lecture 1: Intro to Data Visualization & Tableau

**Agenda**

* Part A : Introduction to Data Visualization and Tableau
  + Discover what data visualization is, and how we can use it to better understand data.
  + We acquaint ourselves with the history of data visualization and get introduced to Tableau. Then, we get an idea about the various Tableau product offerings.
  + We also learn about the pros and cons of Tableau.
  + Finally, we see how to download and install Tableau Public.
* Part B : Tableau Public GUI and Basics
  + In this section, you will get a brief idea about the Tableau UI components, different data types and various fields in Tableau.

Netflix Viz & Dashboard Demonstration

Dataset : [Netflix Movies & TV shows](https://public.tableau.com/en-us/s/resources?qt-overview_resources=1#qt-overview_resources)

Demonstration : <https://drive.google.com/file/d/1efZEbXNxJo46WL64cs6ON7gaevvsilta/view?usp=sharing>

Tableau Public Dashboard : <https://public.tableau.com/app/profile/tino3819/viz/NetflixDashboard_16521889408740/Dashboard1>

**What is Business Intelligence (BI)**

BI combines business analytics, data mining, data visualization, data tools and infrastructure, and best practices to help organizations to make more data-driven decisions.

**Why BI?**

It helps companies make better decisions by showing present and historical data within their business context.

**What is Data Visualization?**

* It is the graphical representation of information and data.
* By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.
* Humans are good at deriving knowledge from visualizations.

**Brief History**

* Until the early 21st century, Database, Excel, Access etc were used to produce numbers and data.
* The main idea behind Tableau’s creation was to make the database industry interactive and comprehensive.
* Tableau is a popular data visualization and business intelligence tool used for reporting and analyzing vast volumes of data.
* Tableau was founded by Pat Hanrahan, Christian Chabot, and Chris Stolte from Stanford University in 2003.

**Over the years..**

* Tableau has been named a Leader in the Gartner Magic Quadrant for Analytics & Business Intelligence Platforms for the 10th consecutive year.
* Tableau Software has a market capitalization of $14.61 billion and generates $982.95 million in revenue each year.



**Pros of Tableau**

* Quick and interactive visualizations
* Easy to use for non programmers
* High performance
* Mobile friendly
* Extensive customer resources (Tableau Community)
* Working with different data sources
* Easy to upgrade

**Cons of Tableau**

* Focuses primarily on visualization and cannot work with uncleaned data. In order to efficiently use Tableau, you need to do proper data cleaning in the underlying database first.
* Lacks data modeling and data dictionary capabilities for Data Analysts. This means that you've to separately maintain your metrics definitions elsewhere.
* Lack of version control and collaboration when building data logic and dashboard.

**Tableau Product Suite**

* Tableau Prep
* Tableau Desktop
* Tableau Server
* Tableau Online
* Tableau Reader
* Tableau Public

**References:**

* [Comparison of Product Suite](https://public.tableau.com/views/TableauProductSuite/ProductSuite?:embed=y&:showVizHome=no&:display_count=y&:display_static_image=y&:bootstrapWhenNotified=true)
* [Tableau Pricing](https://www.tableau.com/pricing/teams-orgs)
* [Understanding License types of Tableau](https://help.tableau.com/current/blueprint/en-us/bp_license_types.htm)
* [Tableau Desktop vs Tableau Public](https://www.edureka.co/blog/tableau-desktop-vs-tableau-public-vs-tableau-reader/#:~:text=Tableau%20Desktop%20is%20meant%20for,%2C%20bloggers%2C%20students%20and%20more)

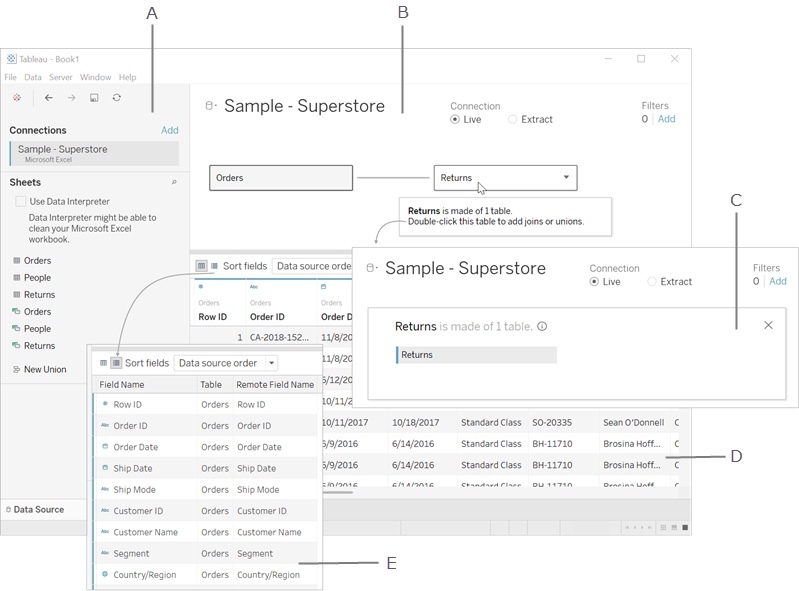
**Tableau Public Installation**

* Pre-read Document: [link](https://docs.google.com/document/d/14IYhrwVt6p5ChF7mXT9yZ8JrwCVJ95RB/edit?usp=sharing&ouid=100659516601446935794&rtpof=true&sd=true)
* Tableau Public download link : [link](https://www.tableau.com/en-gb/products/public/download)
* Tableau Public Desktop installation steps: [link](https://help.tableau.com/current/desktopdeploy/en-us/desktop_deploy_download_and_install.htm#install-tableau-desktop-public-edition)

**Tableau Public GUI**

Loading dataset in Tableau and Data Source view

Dataset: [link](https://docs.google.com/spreadsheets/d/1yEDocbFhGgZ1BWQO8FHy1jYg4tW8YP8X/edit?usp=share_link&ouid=100659516601446935794&rtpof=true&sd=true)



1. Left pane - Displays the connected data source and other details about your data.
2. Canvas: logical layer - The canvas opens with the logical layer, where you can create relationships between logical tables.
3. Canvas: physical layer - Double-click a table in the logical layer to go to the physical layer of the canvas, where you can add joins and unions between tables.
4. Data grid - Displays first 1,000 rows of the data contained in the Tableau data source.
5. Metadata grid - Displays the fields in your data source as rows.

**Note:** Canvas will be explained in future lectures.

**Reference:** <https://help.tableau.com/current/pro/desktop/en-us/environment_datasource_page.htm>

**Data Fields**

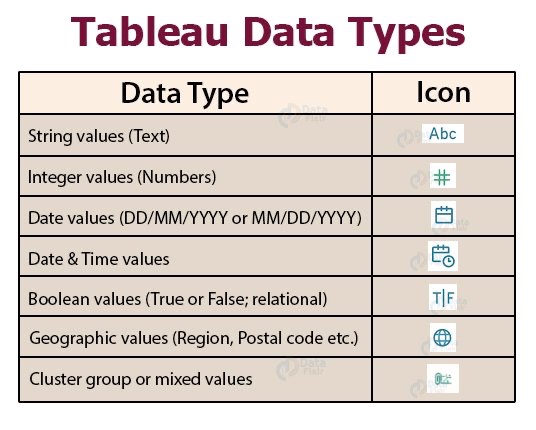
* After you connect to your data and set up the data source with Tableau, the data source connections and fields appear on the left side of the workbook in the Data pane.
* The term "fields" refers to columns.
* When you connect to a new data source :
  + Each field is automatically assigned a Data Type (such as integer, string, date).
  + Tableau assigns each field in the data source as dimension or measure in the Data pane, depending on the type of data the field contains.
    - Discrete Dimension or Continuous Measure (more common)
    - Continuous Dimension or Discrete Measure (less common)

**Reference:**

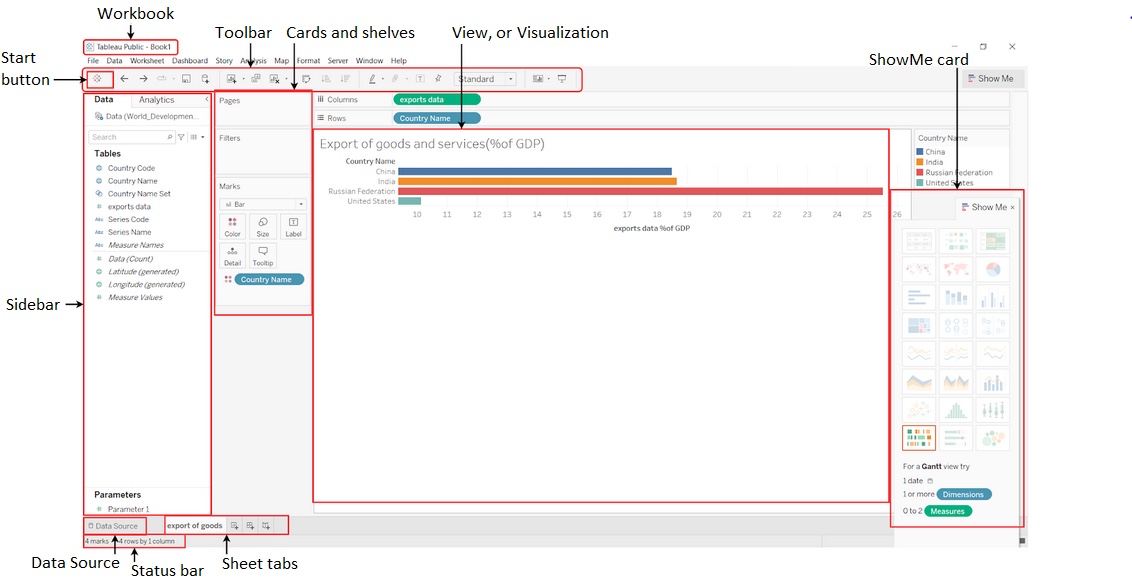
* <https://help.tableau.com/current/pro/desktop/en-us/datafields_typesandroles.htm>
* <https://help.tableau.com/current/pro/desktop/en-us/datafields_understanddatawindow.htm>

**Data Type**

* Tableau expresses fields and assigns data types automatically.
* If the data source appoints the data type, Tableau will use that data type.
* If the data source doesn't individually assign a data type, Tableau will assign one



**Tableau View page**



**Measure:**

* A measure is a field that is a dependent variable; that is, its value is a function of one or more dimensions.
* Tableau treats any field containing numeric (quantitative) information as a measure.

**Dimension:**

* Dimension is a field that can be considered an independent variable.
* By default, Tableau treats any field containing qualitative, categorical information as a dimension.

**Rule of Thumb :**

Generally,

* the measure is the number;
* the dimension is what you “slice and dice” the number by.

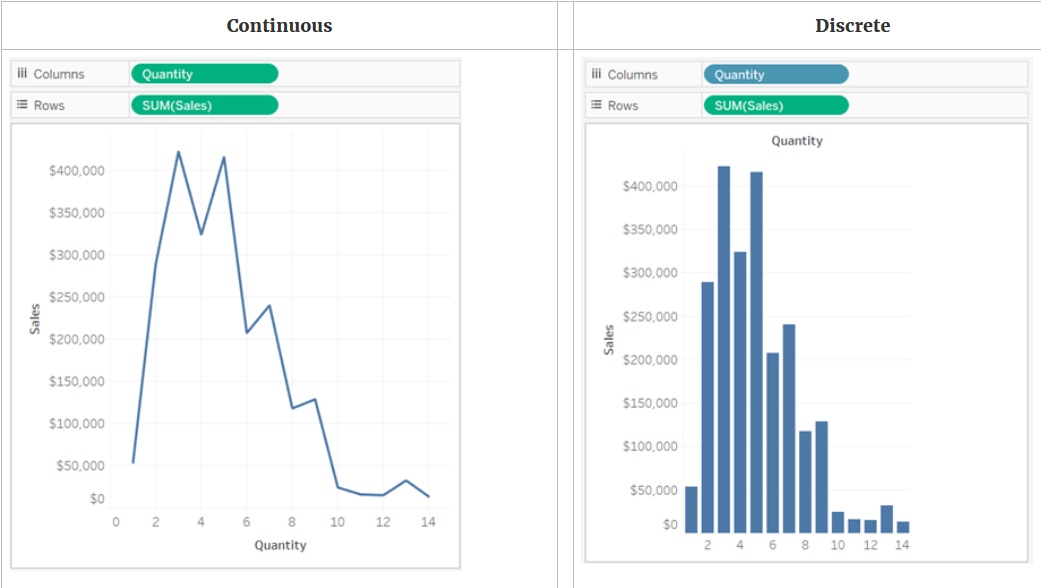
**Discrete and Continuous fields**

* Tableau represents data differently in the view depending on whether the field is discrete (blue), or continuous (green).
* Continuous and Discrete are mathematical terms.
  + Continuous means "forming an unbroken whole, without interruption"
  + Discrete means "individually separate and distinct"



**Rule of Thumb :**

* Discrete fields draw headers; Continuous fields draw axes.
* Discrete fields can be sorted; Continuous fields cannot.
* Blue color field indicates Discrete Field
* Green color field indicates Continuous Field



* In the example on the left (above), because the Quantity field is set to Continuous, it creates a horizontal axis along the bottom of the view.
  + The green background and the axis help you to see that it's a continuous field.
* In the example on the right (above), the Quantity field has been set to Discrete. It creates horizontal headers instead of an axis.
  + The blue background and the horizontal headers help you to see that it's discrete.
* In both examples, the Sales field is set to Continuous.
  + It creates a vertical axis because it is Continuous and it's been added to the Rows shelf.
  + If it was on the Columns shelf, it would create a horizontal axis.
* The green background and aggregation function (in this case, SUM) help to indicate that it's a Measure.
* The absence of an aggregation function in the Quantity field name helps to indicate that it's a Dimension.

**Converting Measure to Dimension**

* You can convert a field from a measure to a dimension in the current view.
* If you want the change to affect all future uses of the field in the workbook, you can convert a field in the Data pane from a measure to a dimension.

To convert a measure to a dimension in the Data pane, do either of the following :

* Click and drag the field from the measures area in the Data pane and drop it into a dimensions area (above the line).
* Right-click (control-click on a Mac) the measure in the Data pane and select Convert to Dimension.

If you place a field that you converted from a measure to a dimension on a shelf, it now produces headers instead of an axis.

**Business problem 1:**

Dataset : [sample superstore](https://docs.google.com/spreadsheets/d/1Kt-ZZqiJ9483tEpOShNYRvmbAgnC0eN5/edit?usp=share_link&ouid=100659516601446935794&rtpof=true&sd=true)

Determine the total sales value of each category.

* Use the Orders table
* Drag Category to Columns
* Drag Sales to Rows
* Switch to Entire View
* Click on Label and check "Show marks labels"

