

Basic Data visualization (with Tableau)

Unit 8

Working Draft

Nepal Data Literacy Program, 2019

Organized by



Supported by





EVERY DAY WE CREATE
**2,500,000,
000,000,
000,000**

(2.5 QUINTILLION) BYTES OF DATA

*This would fill 10 million blu-ray discs,
the height of which stacked, would measure
the height of 4 Eiffel Towers on top of one another.*



BIG DATA:

Data stored grows
4X FASTER THAN WORLD ECONOMY



Substantial shift in
**ECONOMIC POWER AND SOURCE
OF ECONOMIC VALUE**



Increasing quantity of data allows for
MORE QUALITATIVE APPROACH



Visualization Source:
<https://webbizkb.com>



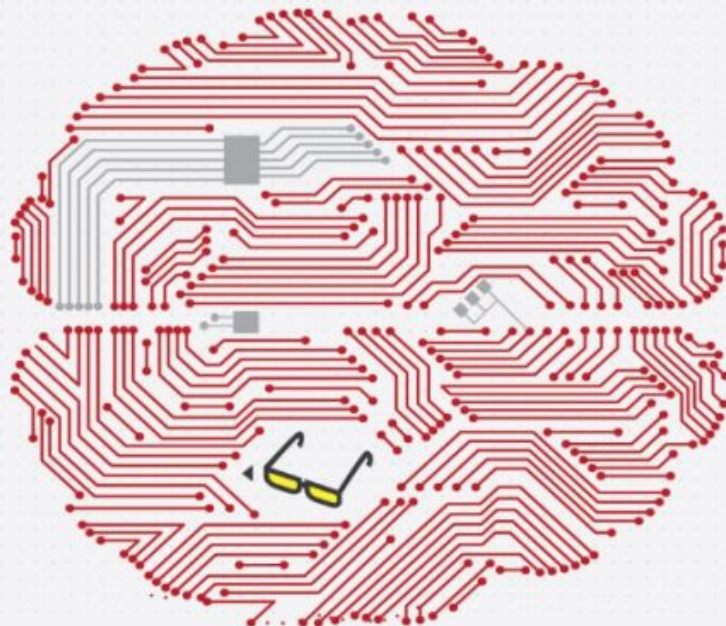
90% ▶

OF INFORMATION TRANSMITTED
TO THE BRAIN IS **VISUAL.**

VISUALS ARE PROCESSED

60,000X

FASTER IN THE BRAIN THAN TEXT.



DATA: VISUAL



DATA: TEXT

SOURCE: FORRESTER CSO INSIGHTS 2012
linjaCOM, from The Noun Project

Comparatively

Are these texts effectively communicating the message ?

“Our brains have the ability to process visuals a lot faster than text. It’s been reported that 70 percent of all our sensory receptors are in our eyes, and that we can usually get the sense of a visual scene in less than 1/10 of a second. That’s a lot faster than how long it typically takes us to read and comprehend text-only information.”

“According to the University of Minnesota, visual data is processed by human brains 60,000 times faster than written data.”



What makes good
visualization ?

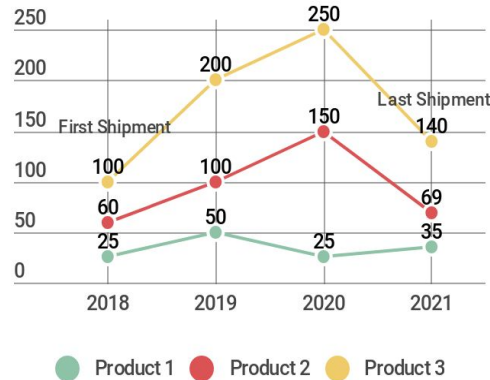
Don't feel pressure to include all the information you have about your chart in the form of text.

You don't want to overwhelm your viewer with paragraphs of copy, bulky legends, axes labels, and additional words.

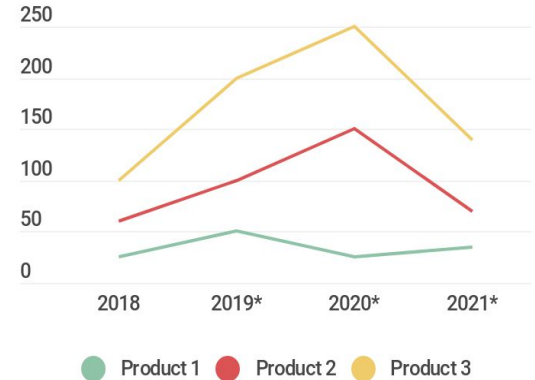


Bad, Vague Title

Text explaining the data stated below.



Good, Memorable Title



Staying Focused

- **What's the point?**

Make sure to include only the data you want the reader to remember

- **Simplify the numbers!**

When possible, reduce numbers to simplest form.

- **Make the angle clear**

Titles and labels should be specific, easy to understand and true to the data, including citing the source of the data

- **Set the scene!**

Styles and colors should aid understanding, not distract readers

Basic Design Concepts

- **Simplicity:** Choose a maximum of three colors and fonts and stick with them consistently.
- **Brevity:** Keep text short and to the point.
- **Creativity:** Incorporate playful design that relates to the topic.
- **Two dimensions:** Avoid 3D graphics: they distort data and look.
- **Clarity:** Label clearly, specify units, use a legend when necessary.

Basic Design Concepts:

Remove
to improve
(the **data-ink** ratio)

Tools to generate visualisation (freely available or has a public version)

Google charts



Infogram

The Infogram logo, which features the text "infogr.am" in white lowercase letters inside a red speech bubble shape.

infogr.am

Tableau



Note: This is not the exclusive list of tools to generate data visualisation

Advantages of Tableau

1. Powerful visualization tool which makes analyzing data fast and easy, beautiful and useful.
2. Easy to learn. Access to a huge Tableau community and free resources
3. Data can be connected from 40 different sources including Excel spreadsheets
4. Tableau dashboards can be shared within the organization via desktop or mobile browsers
5. No coding required for all the amazing functionalities including built in table calculations for complex analyses
6. Amazing self service analytics without the need for IT intervention

Limitations of Tableau Public

1. Tableau Desktop only connects to Microsoft Access, Excel or Text files
2. Works can only be saved to Tableau's public server
3. Storage capacity is limited to 50 megabytes per named user
4. Datasource size is limited to 100,000 records
5. Tableau public workbooks can be viewed and downloaded by anyone which are not ideal when working with private data

Tableau public version is free

Now, if you haven't installed Tableau public.
Install it now !!!!

<https://online.tableau.com>

The learners will be using **Tableau public**, it has a free downloadable format of the Tableau website for the exercises.



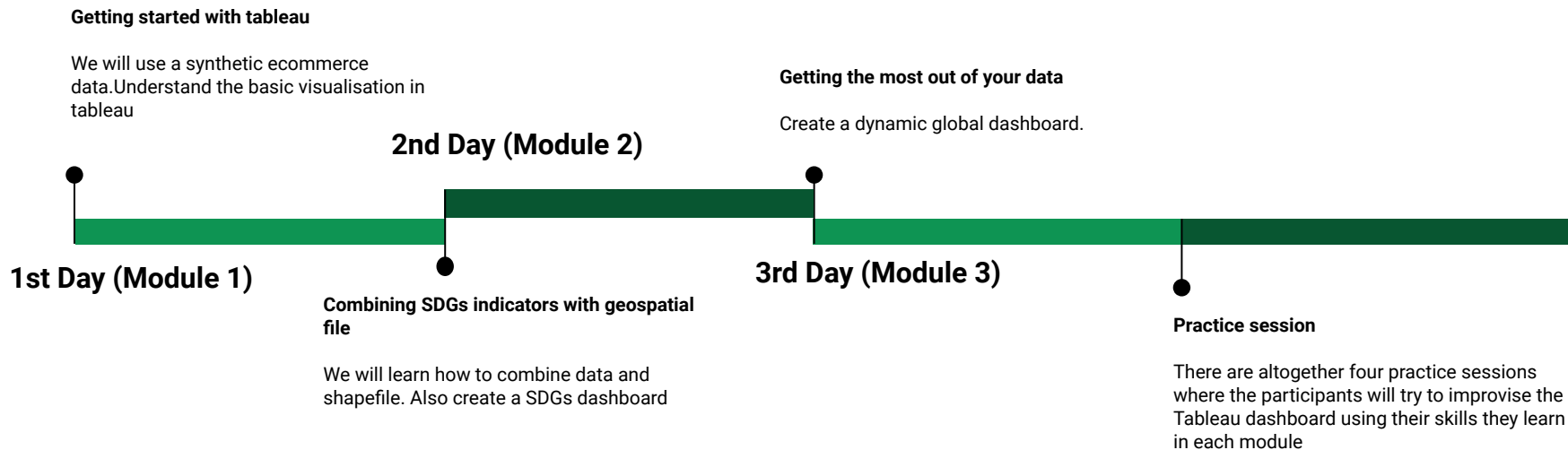
Objectives of the Unit

At the end of the unit, participants will be able to

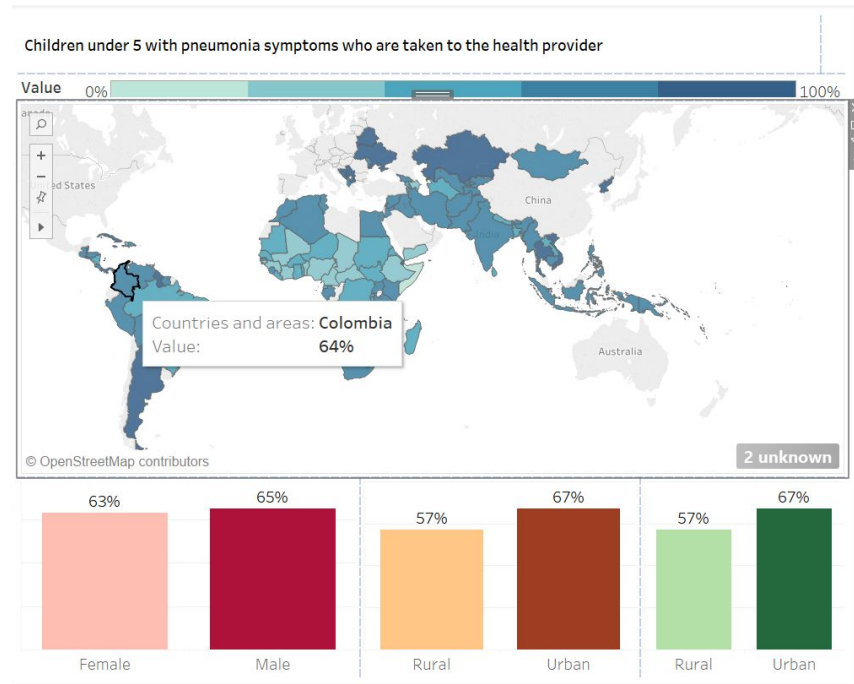
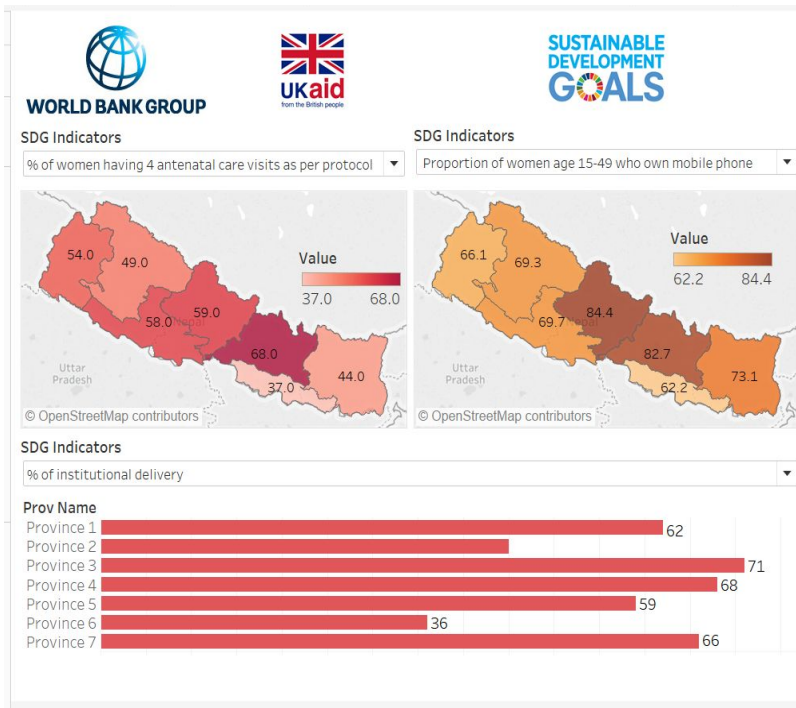
- Work with the data intuitively, and present it in a visually appealing manner.



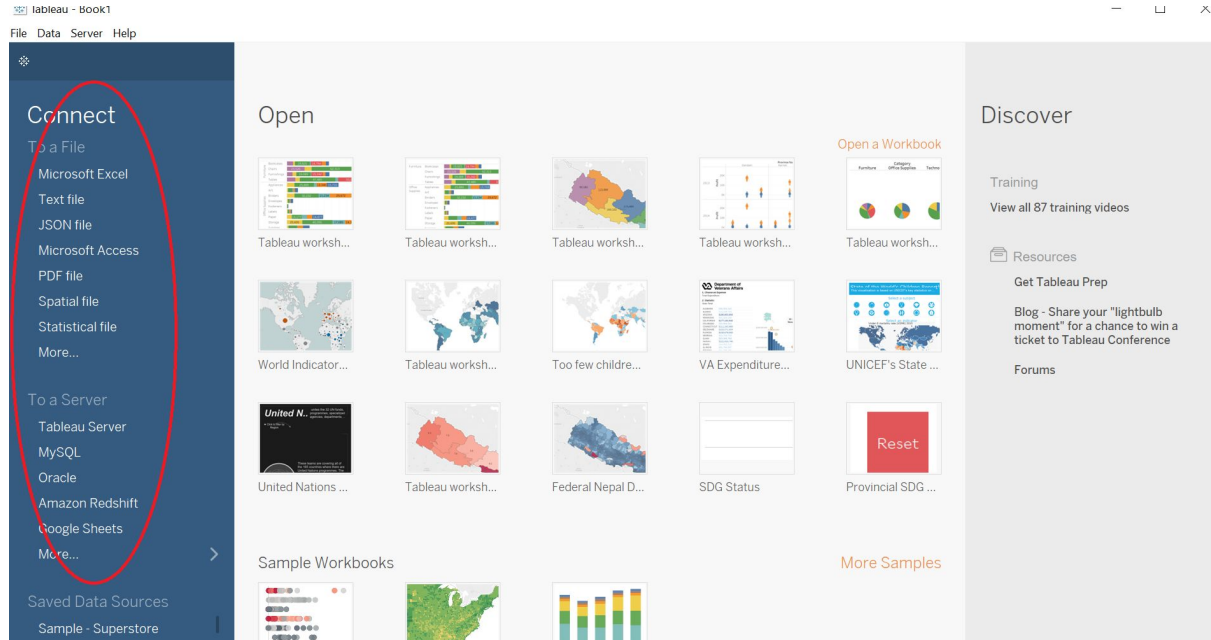
Snapshot of the session



Data is beautiful



Start screen



In the Connect pane on the left, we can see the wide variety of data sources Tableau connects to natively.

For this module, we'll connect to the **Province_ecommerce_data.csv**.

Data source page has four main areas. Blank white space (canvas) allows you to drag and drop one or more tables to the canvas area.

The Grid allows you to review rows of data, also you can modify your data sources from the Grid.

Tableau - Book1

File Data Server Window Help

Connections [Add](#)

Province_ecommerce_data
Microsoft Excel

Sheets

☐ Use Data Interpreter
Data Interpreter might be able to clean your Microsoft Excel workbook.

Sheet1

New Union

Sheet1 (Province_ecommerce_data)

Connection
☒ Live ☐ Extract

Filters
0 [Add](#)

Sort fields Data source order

☐ Show aliases ☐ Show hidden fields 1,000 rows

Sub-Category	Category	Number of...	Manufactur...	Customer ID	Order Date	Ship Mode	Ship Date	Product Na...	Order ID	Prov
File holders	Furniture	1	Dragon Furnit...	9853071700	1/2/2013	Standard Class	1/7/2013	Dragon Furnit...	7250	Pr
Desks	Furniture	1	Apex Imperial	9856997154	1/3/2013	Standard Class	1/8/2013	Apex Imperial...	6416	Pr
Table holders	Furniture	1	Royal designs s	9853289507	1/7/2013	Standard Class	1/11/2013	Royal designs...	2835	Pr
File holders	Furniture	1	Annapurna	9856771858	1/8/2013	Standard Class	1/12/2013	Annapurna M...	7522	Sl
Table holders	Furniture	1	AMD	9853318062	1/9/2013	Second Class	1/13/2013	AMD Two-Ton...	3897	Pr
Table holders	Furniture	1	AMD	9854441957	1/10/2013	Second Class	1/13/2013	AMD Value U-...	1990	Pr
Table holders	Furniture	1	Other	9854105353	1/10/2013	Standard Class	1/17/2013	24-Hour Roun...	4166	Pr

Go to Worksheet

Data Source Sheet1

Click on the **metadata grid**: It helps you to understand the fields (columns)

Tableau - Tableau workshop 190817_Practice session

File Data Server Window Help

Province_ecommerce_data

Connection
☒ Live ☐ Extract

Filters
0 Add

Province_ecommerce_data.csv

Sort fields Data source order ▼

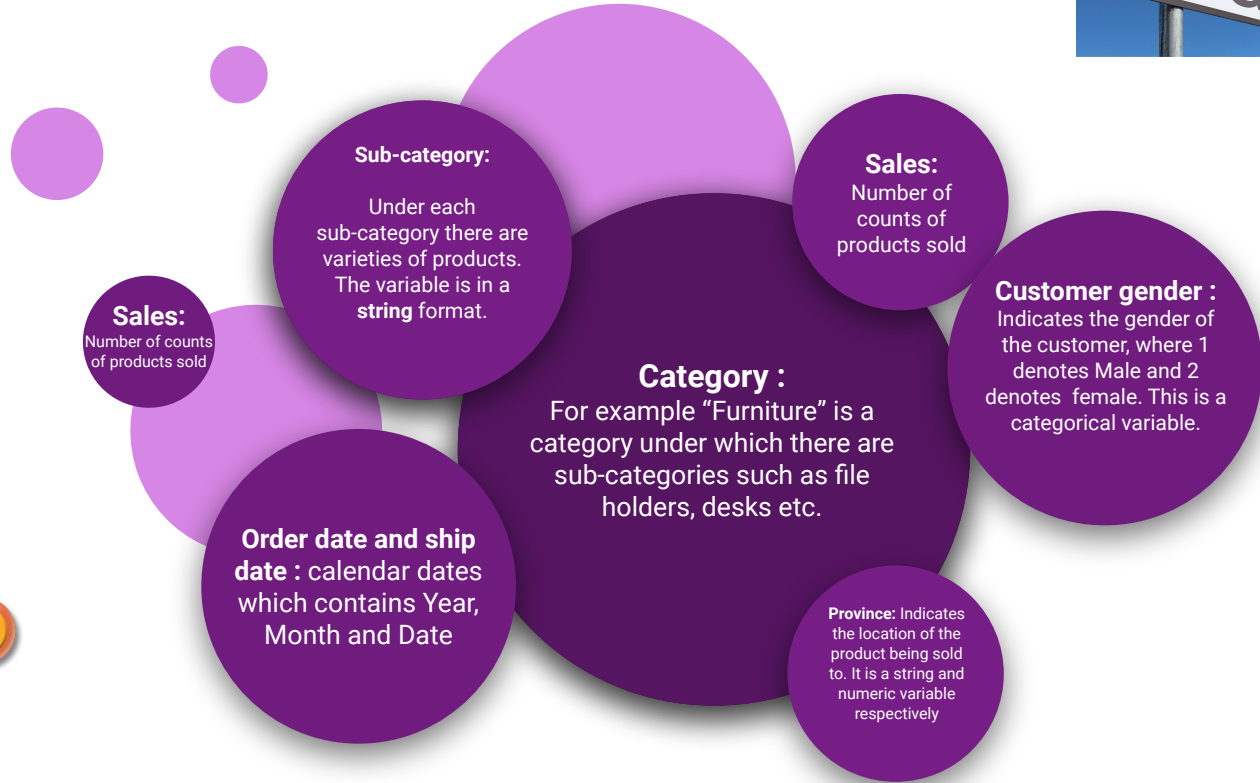
☐ Show hidden fields

Field Name	Table	Remote Field Name
Sub-Category	Province_ecommerce_data.csv	Sub-Category
Category	Province_ecommerce_data.csv	Category
Number of records	Province_ecommerce_data.csv	Number of records
Manufacturer	Province_ecommerce_data.csv	Manufacturer
Customer ID	Province_ecommerce_data.csv	Customer ID
Order Date	Province_ecommerce_data.csv	Order Date
Ship Mode	Province_ecommerce_data.csv	Ship Mode
Ship Date	Province_ecommerce_data.csv	Ship Date
Product Name	Province_ecommerce_data.csv	Product Name

Data Source Sheet 1

Know your data- Module 1

(Province ecommerce)



Note: Province ecommerce is a synthetic data produced for the training purpose.

Tableau working window

The screenshot shows the Tableau interface with the 'Dimensions' and 'Measures' panes on the left. The 'Dimensions' pane lists fields like Category, Customer ID, Manufacturer, Order Date, Order ID, Payment method, Product Name, Province Name, Ship Date, Ship Date 1, Ship Mode, Sub-Category, and Measure Names. The 'Measures' pane lists fields like Area, Cost per unit, Customer gender, Number of records, Price per unit, Profit per unit, Sales, Total sales revenue, Latitude (generated), Longitude (generated), Number of Records, and Measure Values. The main workspace shows a blank sheet with a red overlay explaining the difference between dimensions and measures.

01 Dimensions

- Dimensions are categorical fields
- For example : payment types, province name
- Discrete values. Don't total sum

02 Measures

- Metrics/numerical fields that we analyze.
- For example: Sales, Profit, Cost
- Continuous values

No. Not all columns containing numerical data are considered measures. A order date, for example, makes more sense as a dimension. Say your model has a column that contains order date in a series. Knowing that a certain product was ordered in 20000 years in total is not particularly interesting. However, knowing which products/categories were ordered in which date is useful

The screenshot displays the Tableau Desktop interface. On the left, the **Data** pane shows a data source named 'Sheet1 (Province_ecom...)' with a list of fields. The **Dimensions** list includes 'Category', 'Customer ID', 'Manufacturer', 'Order Date', 'Order ID', 'Payment method', 'Product Name', 'Province Name', 'Ship Date', 'Ship Date 1', 'Ship Mode', 'Sub-Category', and 'Measure Names'. The **Measures** list includes 'Area', 'Cost per unit', 'Customer gender', 'Number of records', 'Price per unit', 'Profit per unit', 'Sales', 'Total sales revenue', 'Latitude (generated)', 'Longitude (generated)', 'Number of Records', and 'Measure Values'. The **Marks** card is set to 'Automatic' and includes 'Color', 'Size', 'Text', 'Detail', and 'Tooltip' options. The main workspace is labeled 'Sheet 1' and contains two 'Drop field here' prompts. On the right, a floating 'Show Me' panel is visible, featuring a grid of chart thumbnails and buttons for 'For stacked bars try', '1 or more Dimensions', and '1 or more Measures'.

Tableau - Book1

File Data Worksheet Dashboard Story Analysis Map Format Server Window Help

Data Analytics

Sheet1 (Province_ecom...)

Dimensions

- Category
- Customer ID
- Manufacturer
- Order Date
- Order ID
- Payment method
- Product Name
- Province Name
- Ship Date
- Ship Date 1
- Ship Mode
- Sub-Category
- Measure Names

Measures

- Area
- Cost per unit
- Customer gender
- Number of records
- Price per unit
- Profit per unit
- Sales
- Total sales revenue
- Latitude (generated)
- Longitude (generated)
- Number of Records
- Measure Values

Marks

Automatic

Color Size Text

Detail Tooltip

Columns

Rows

Sheet 1

Drop field here

Drop field here

Drop field here

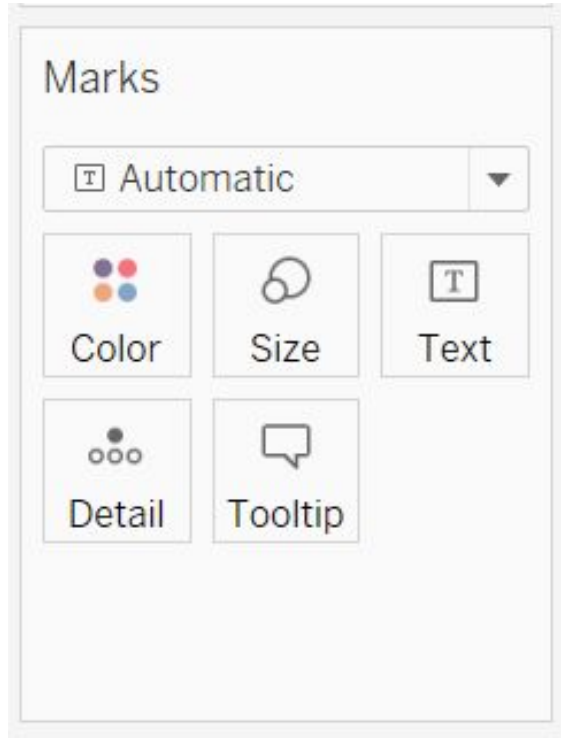
Drop field here

Show Me

For stacked bars try

1 or more Dimensions

1 or more Measures



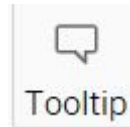
Add color on dimensions or measures
(as per their value)



Change size of dimensions or measures
(as per their value)



Add details to the chart
(as per their value)



The user hovers the pointer over an item,
without clicking it, and a tooltip may
appear—a small "hover box" with
information about the item being hovered
over.

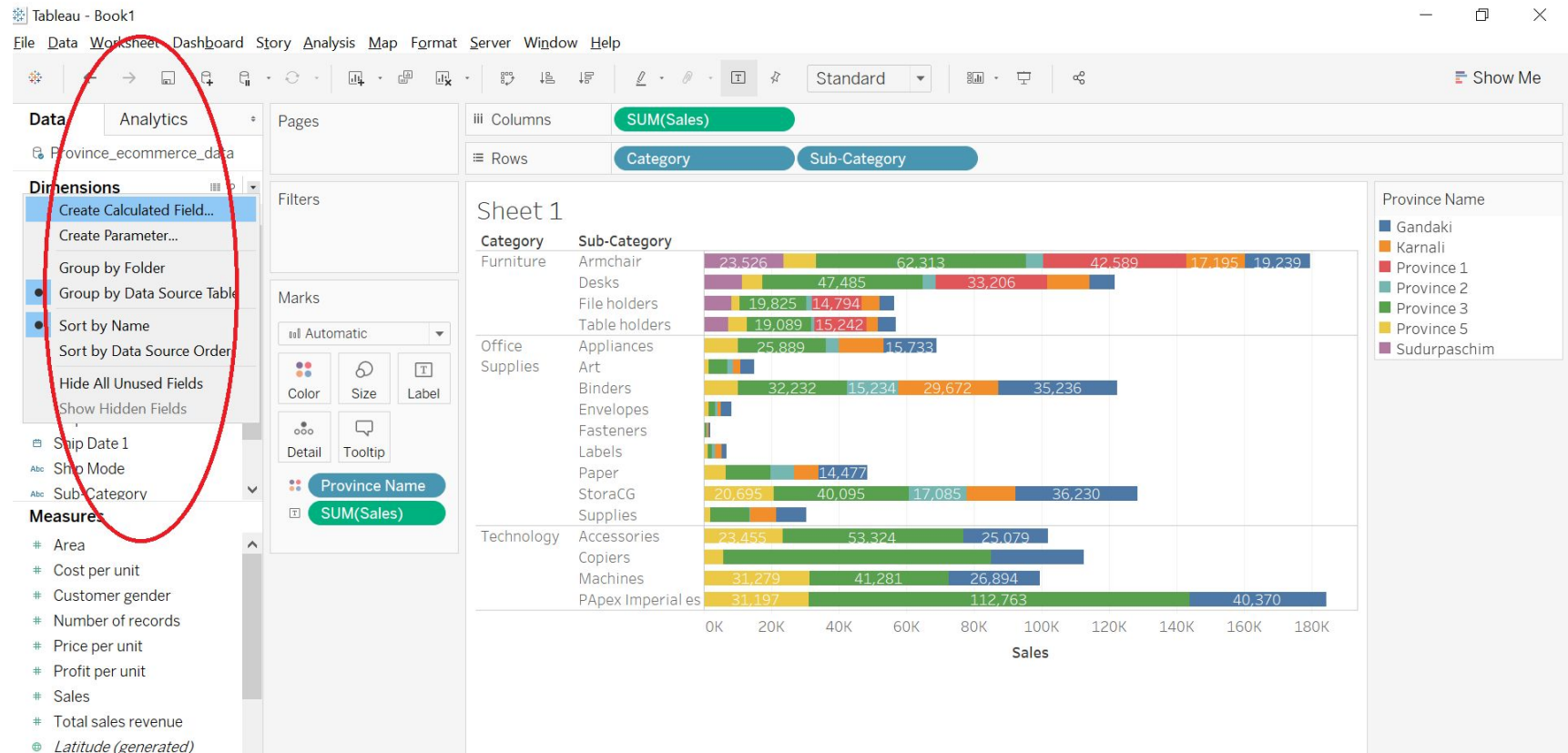
Example of data visualization in action



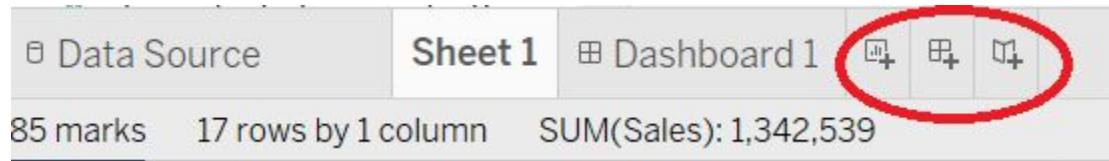
Source: Tableau

Sometimes your data source does not contain a field (or column) that you need for your analysis.

For example, your data source might contain fields with values for total profit. For this reason you will have to create a calculated field



At the lower section of the tableau screen you will notice options to create new sheets Or create a dashboard Or a story :



You will also notice a set of statistics displayed at the bottom

Now, we begin the hands-on exercise

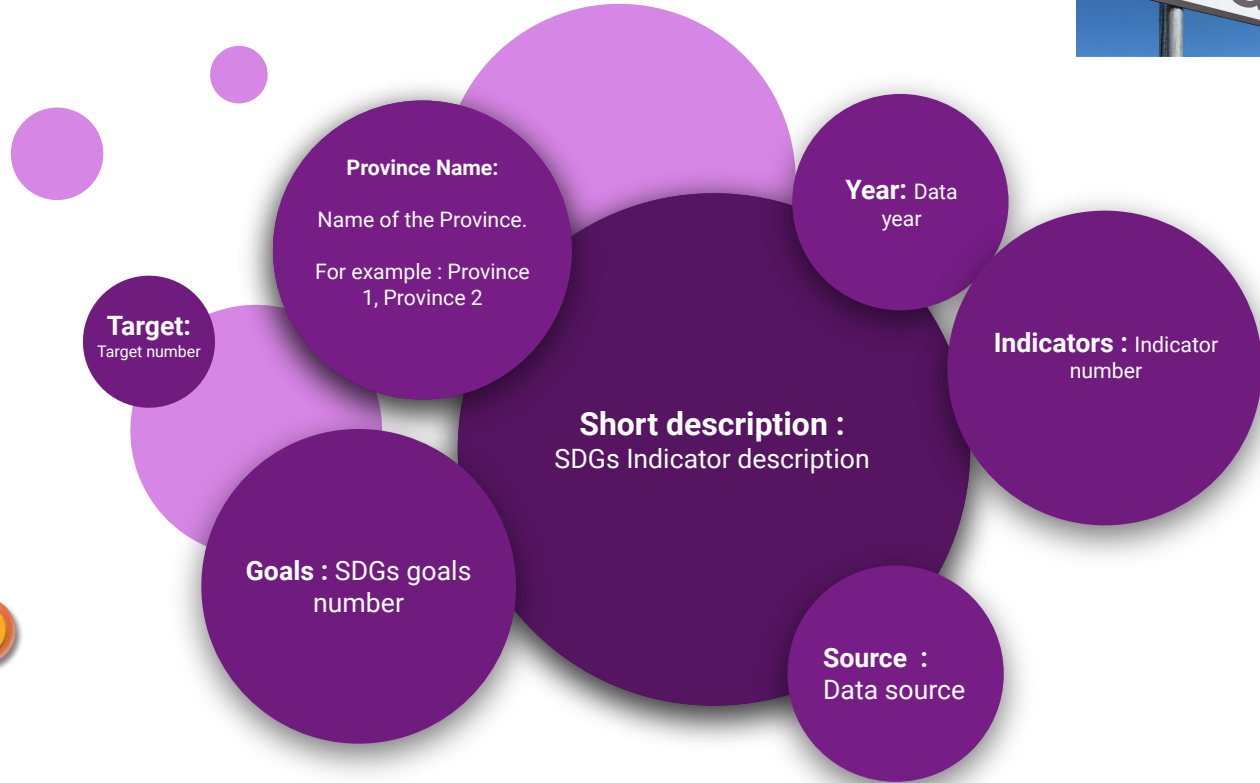
(Please open your student guide note)

Analytics



Know your data- Module 2

(Province ecommerce)



Note: Province ecommerce is a synthetic data produced for the training purpose.

Create a simple descriptive visualisation

