Basic Data visualization (with Tableau)

Unit 8

Working Draft

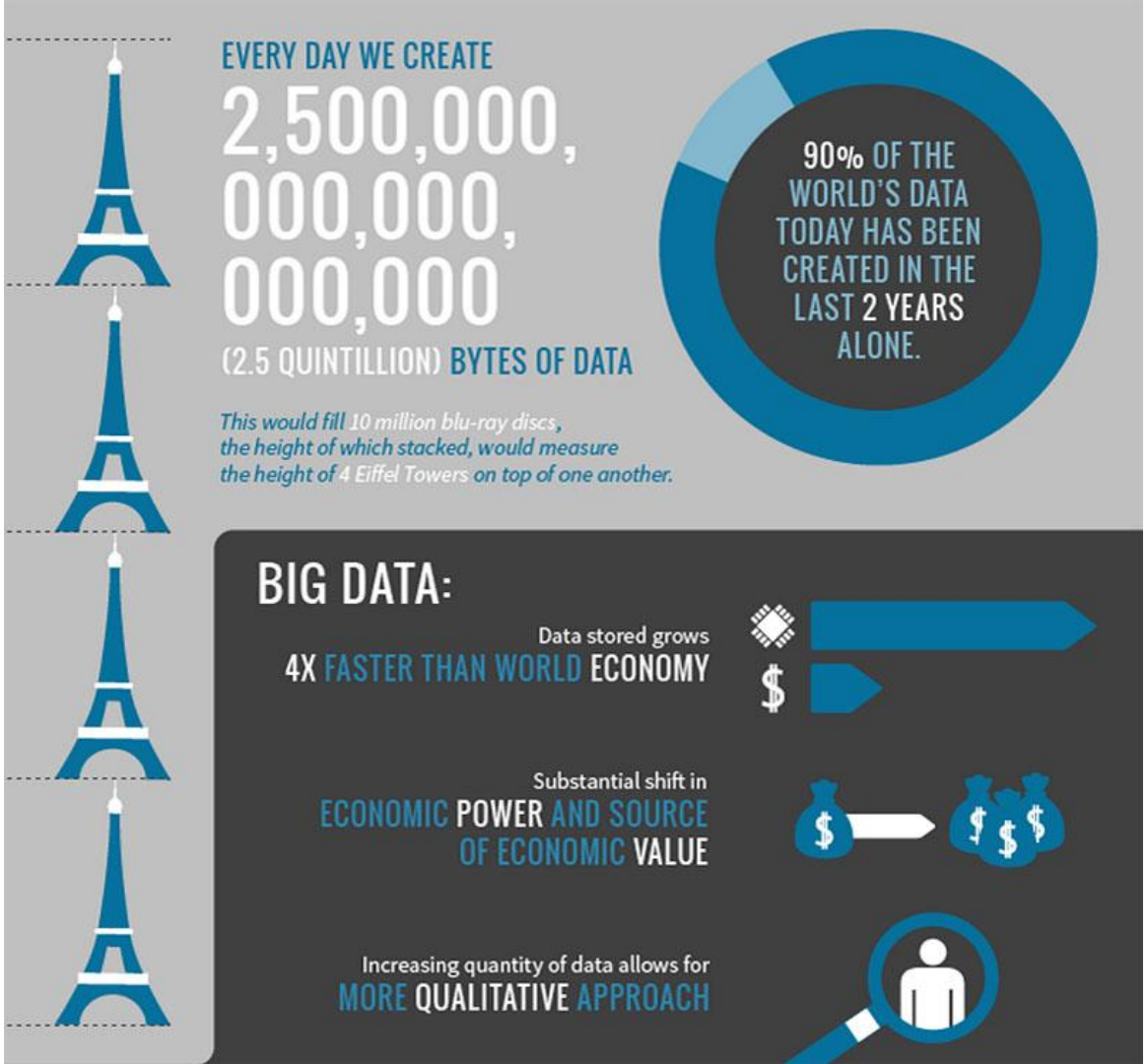
Nepal Data Literacy Program, 2019

Organized by

THE WORLD BANK
IBRD • IDA

Supported by

UKaid
from the British people



Visualization Source:
https://webbizkb.com

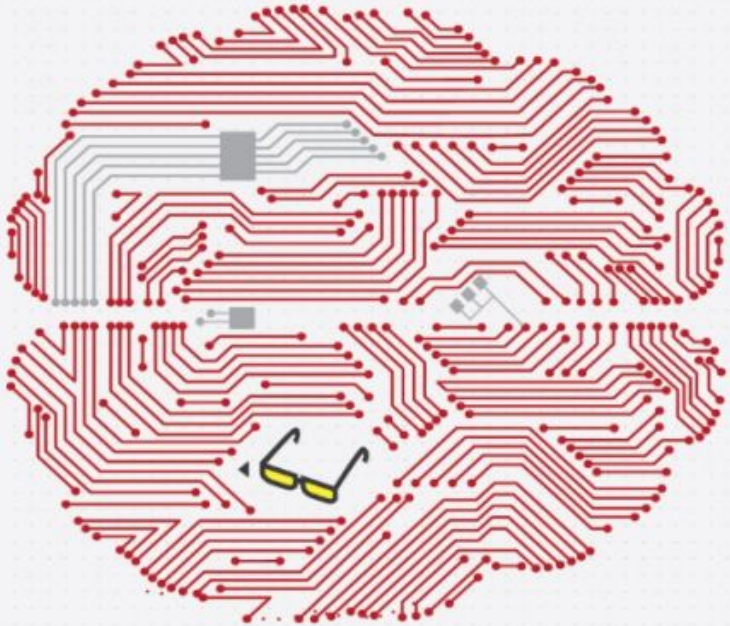
90%

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 ?

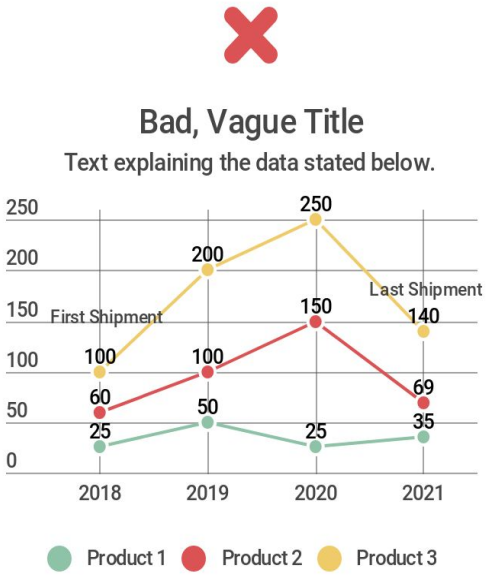
“Our brains have the ability to process visuals a lot faster than text. It’s been reported that 70 percent of all our sensory receptors are in our eyes, and that we can usually get the sense of a visual scene in less than 1/10 of a second. That’s a lot faster than how long it typically takes us to read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

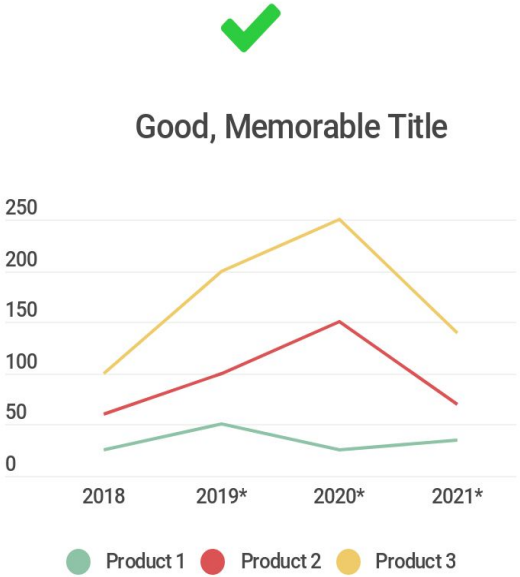
“According to the University of Minnesota, visual data is processed by human brains 60,000 times faster than written data.”

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.





www.infogram.com

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:

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

Google charts

Infogram

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

Source: https://medium.com/datadriveninvestor/10-reasons-why-tableau-rocks-797687ff0dd0

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

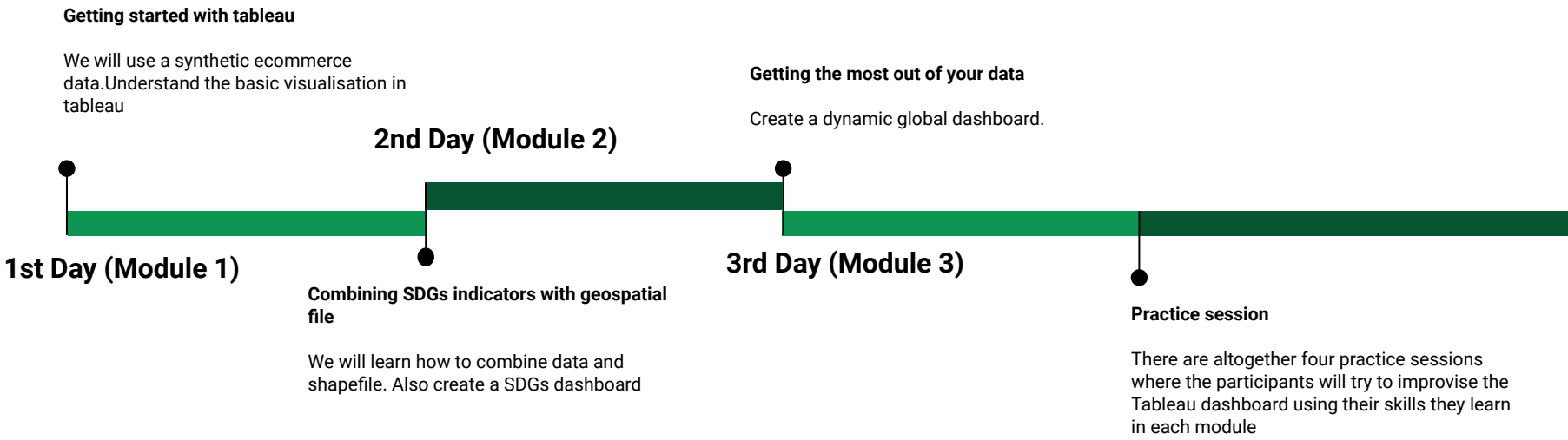
The learners will be using **Tableau public**, it has a free downloadable format of the Tableau website for the exercises.

Objectives of the Unit

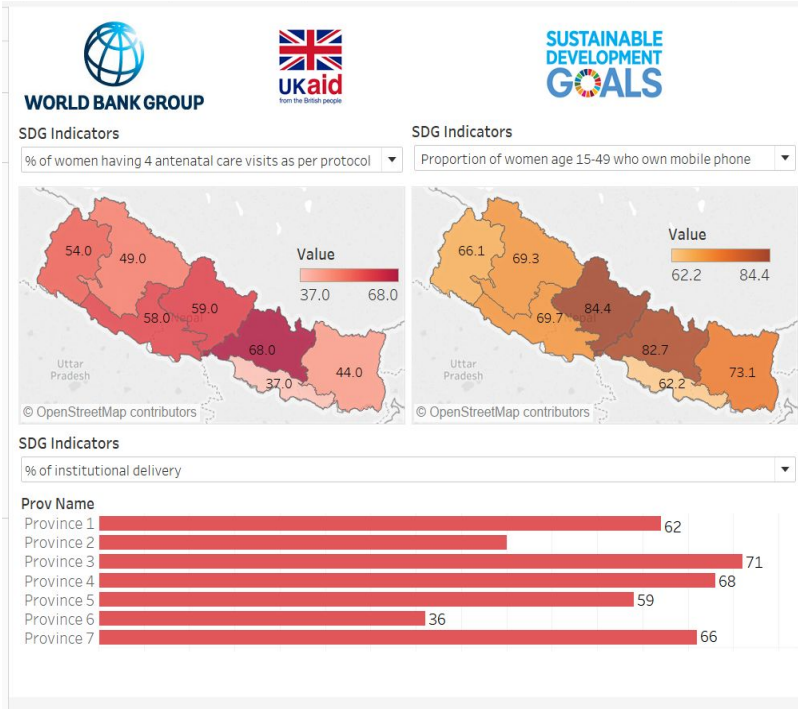
At the end of the unit, participants will be able to

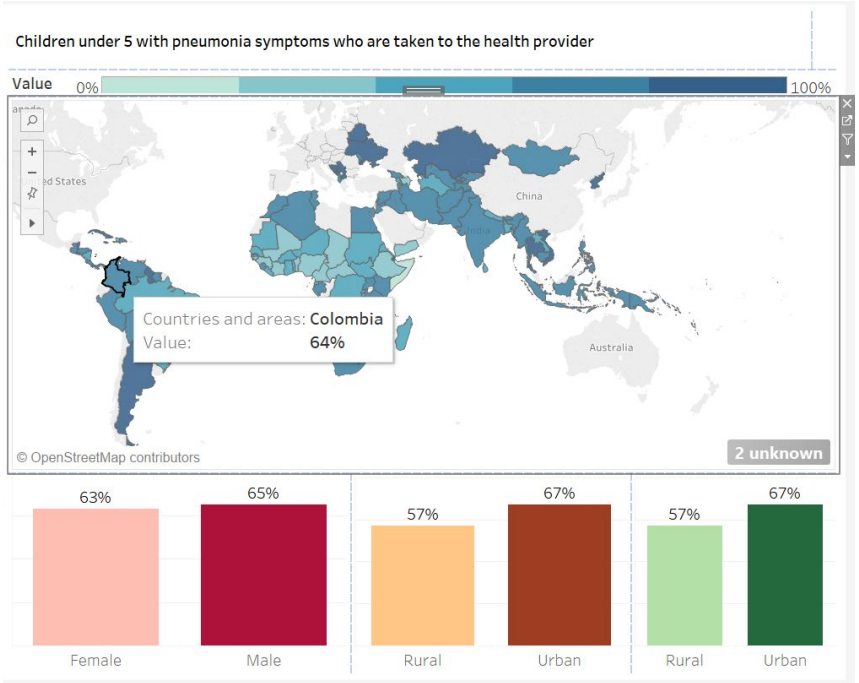
- Work with the data intuitively, and present it in a visually appealing manner.

Snapshot of the session

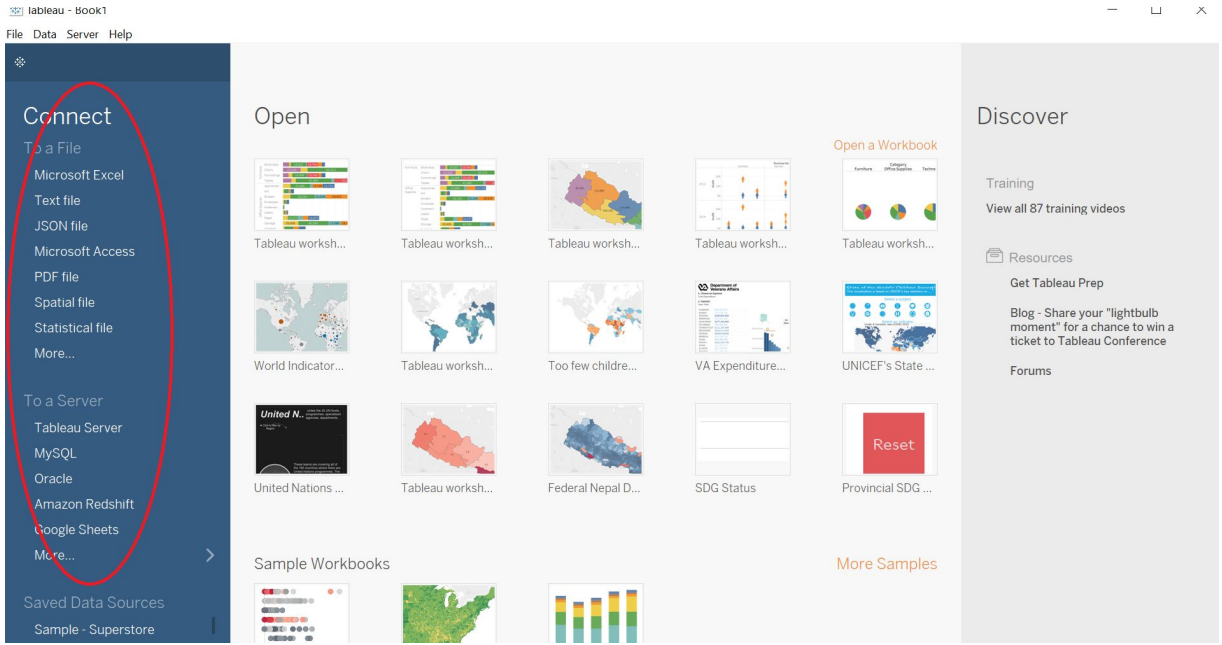


Data is beautiful





Start screen

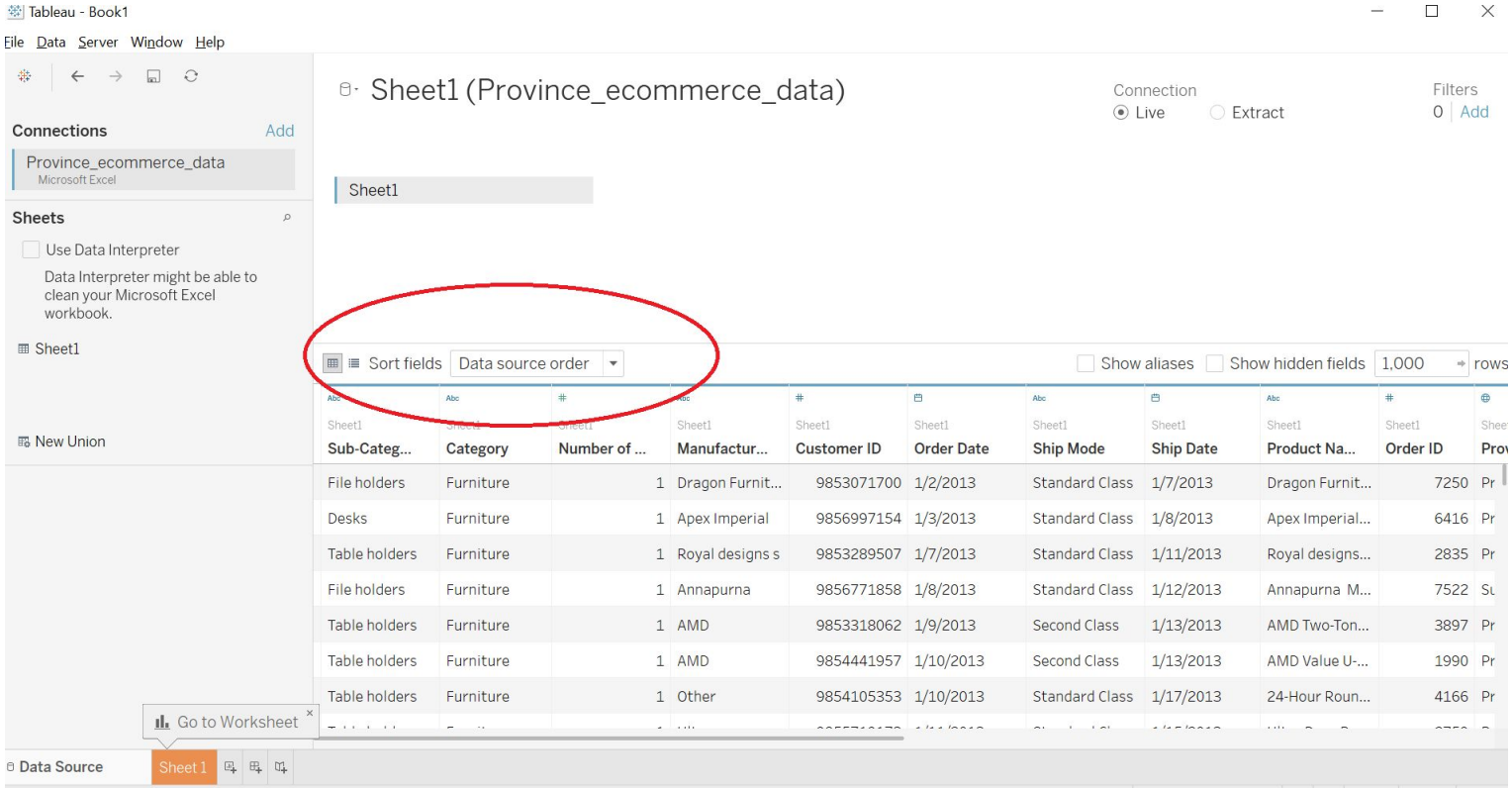


In the Connect pane on the left, we can see the wide variety of data sources Tableau connects to natively.

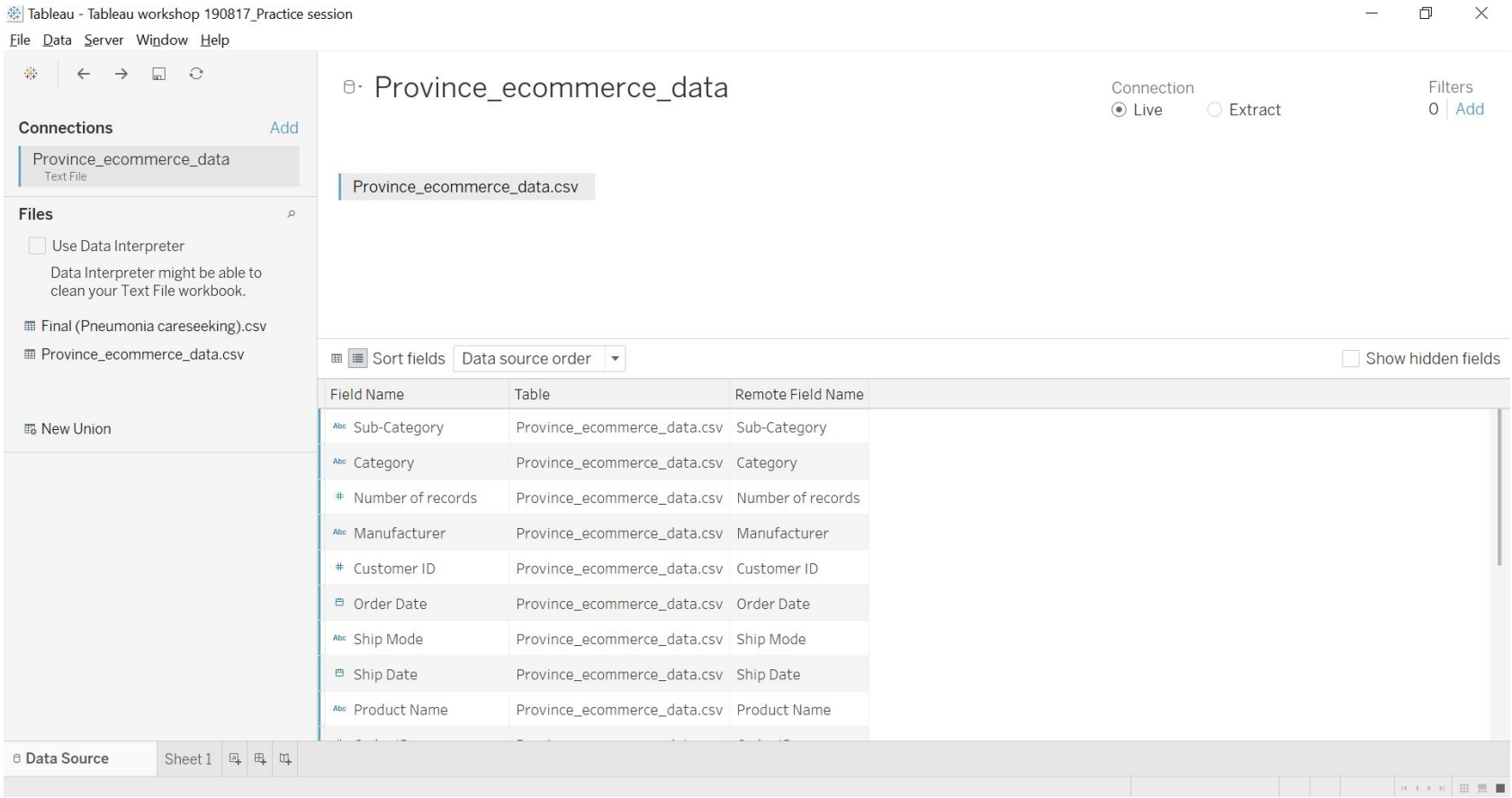
For this module, we'll connect to the **Province_ecommerce_data.csv**.

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.



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



Know your data- Module 1

(Province ecommerce)

Sub-category:

Under each sub-category there are varieties of products. The variable is in a **string** format.

Sales:
Number of counts of products sold

Sales:
Number of counts of products sold

Customer gender :
Indicates the gender of the customer, where 1 denotes Male and 2 denotes female. This is a categorical variable.

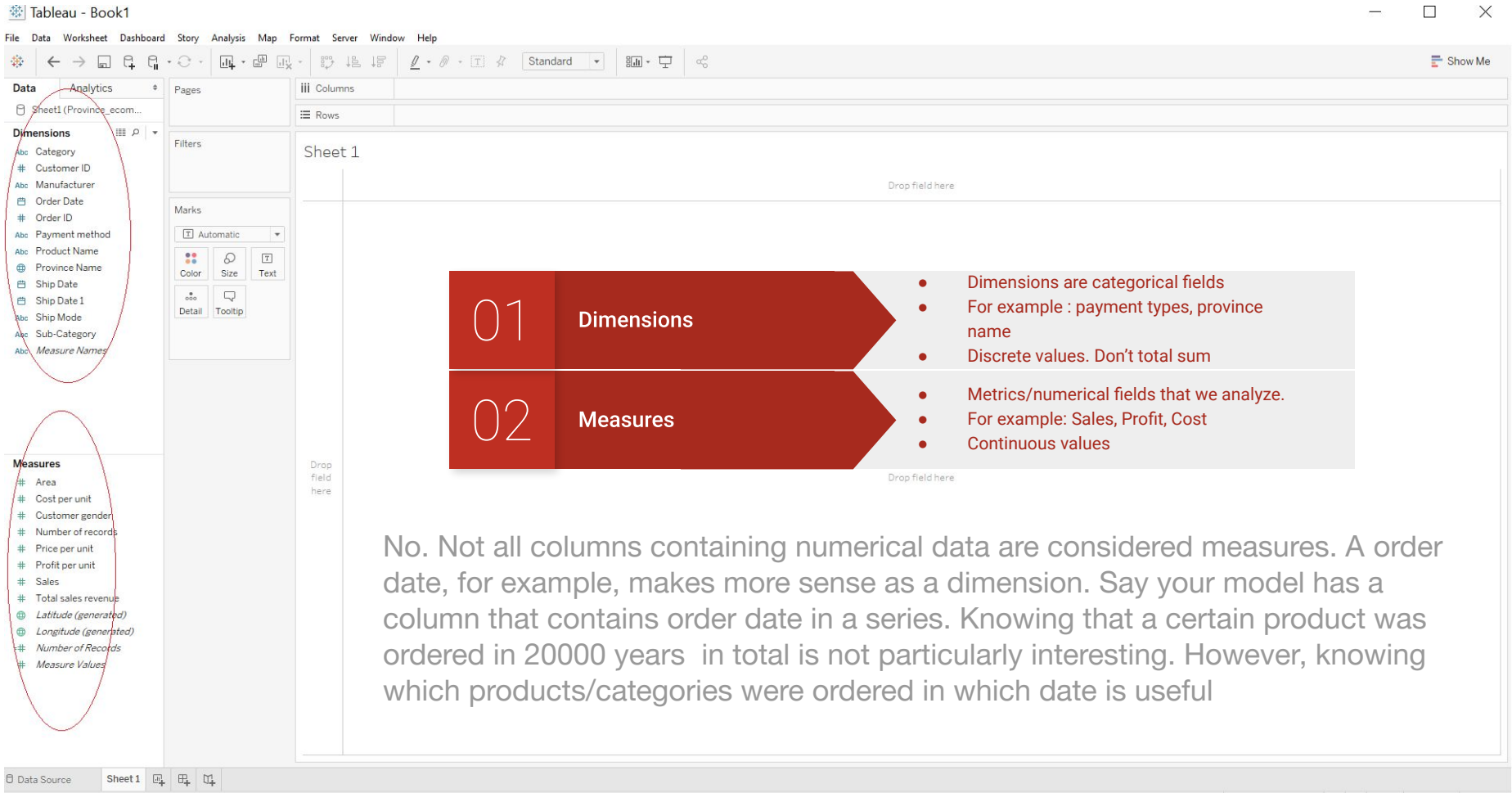
Category :
For example “Furniture” is a category under which there are sub-categories such as file holders, desks etc.

Order date and ship date : calendar dates which contains Year, Month and Date

Province: Indicates the location of the product being sold to. It is a string and numeric variable respectively

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

Tableau working window



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

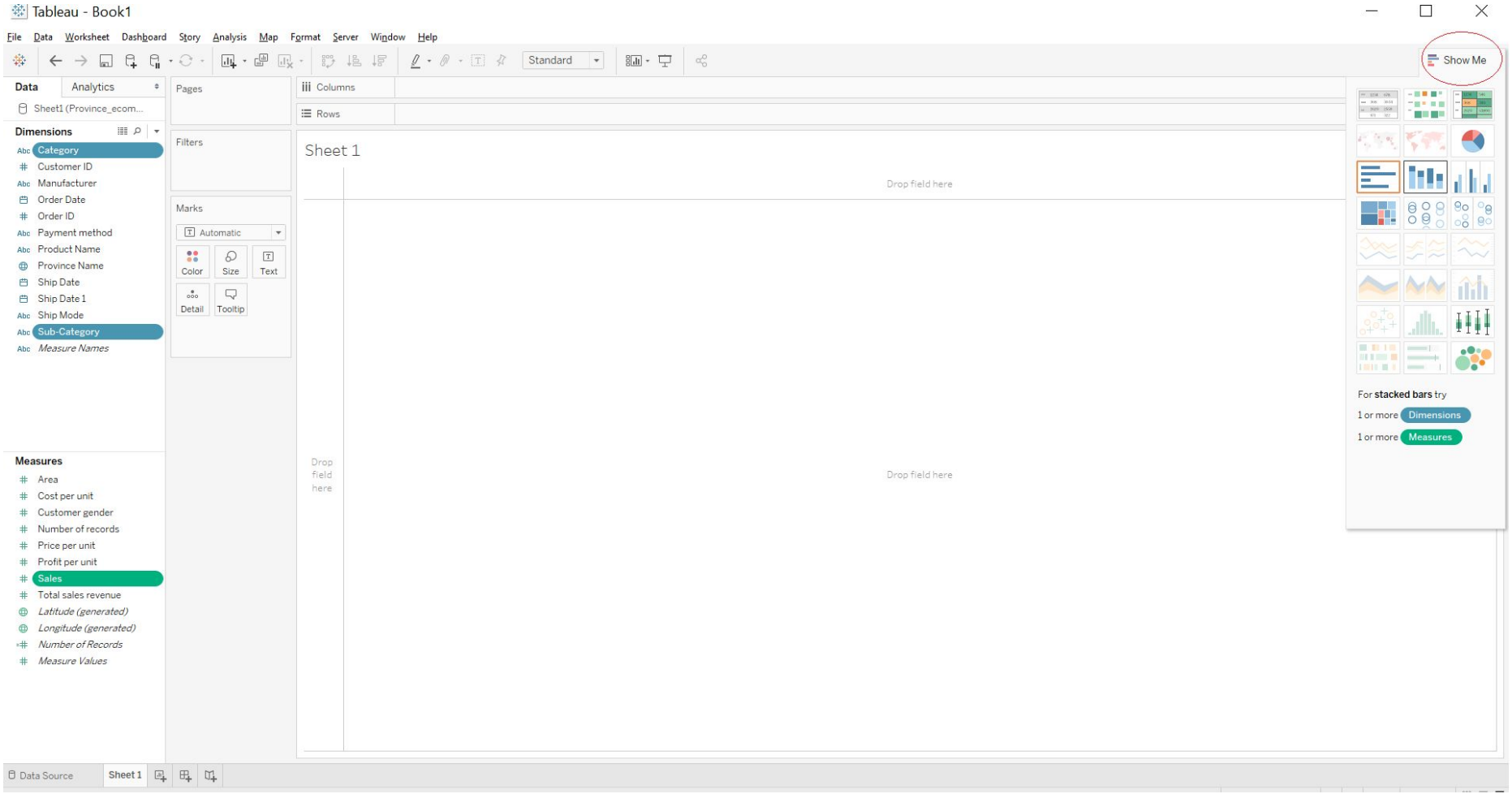
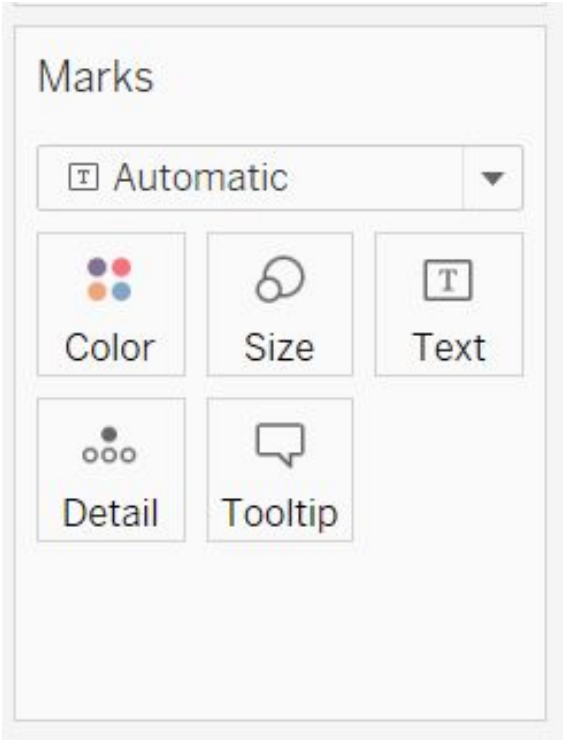


Tableau - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Standard
Show Me
Data
Analytics
Sheet1 (Province_ecom...
Dimensions
Category
Customer ID
Manufacturer
Order Date
Order ID
Payment method
Product Name
Province Name
Ship Date
Ship Date 1
Ship Mode
Sub-Category
Measure 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
Pages
Filters
Marks
Automatic
Color
Size
Text
Detail
Tooltip
Columns
Rows
Sheet 1
Drop field here
Drop field here
Drop field here
For stacked bars try
1 or more Dimensions
1 or more Measures
Data Source
Sheet 1



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)

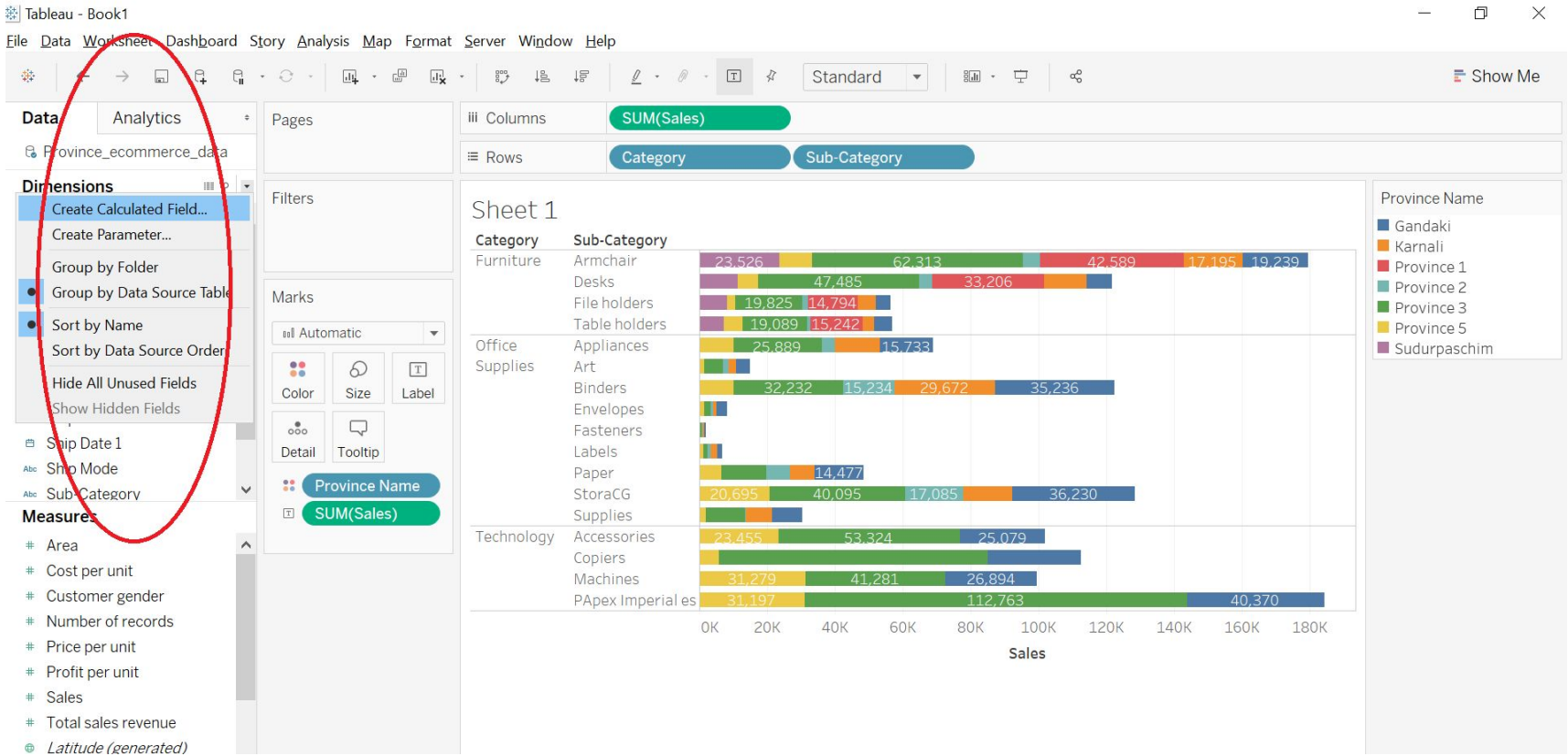
The user hovers the pointer over an item, without clicking it, and a tooltip may appear—a small "hover box" with information about the item being hovered over.

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



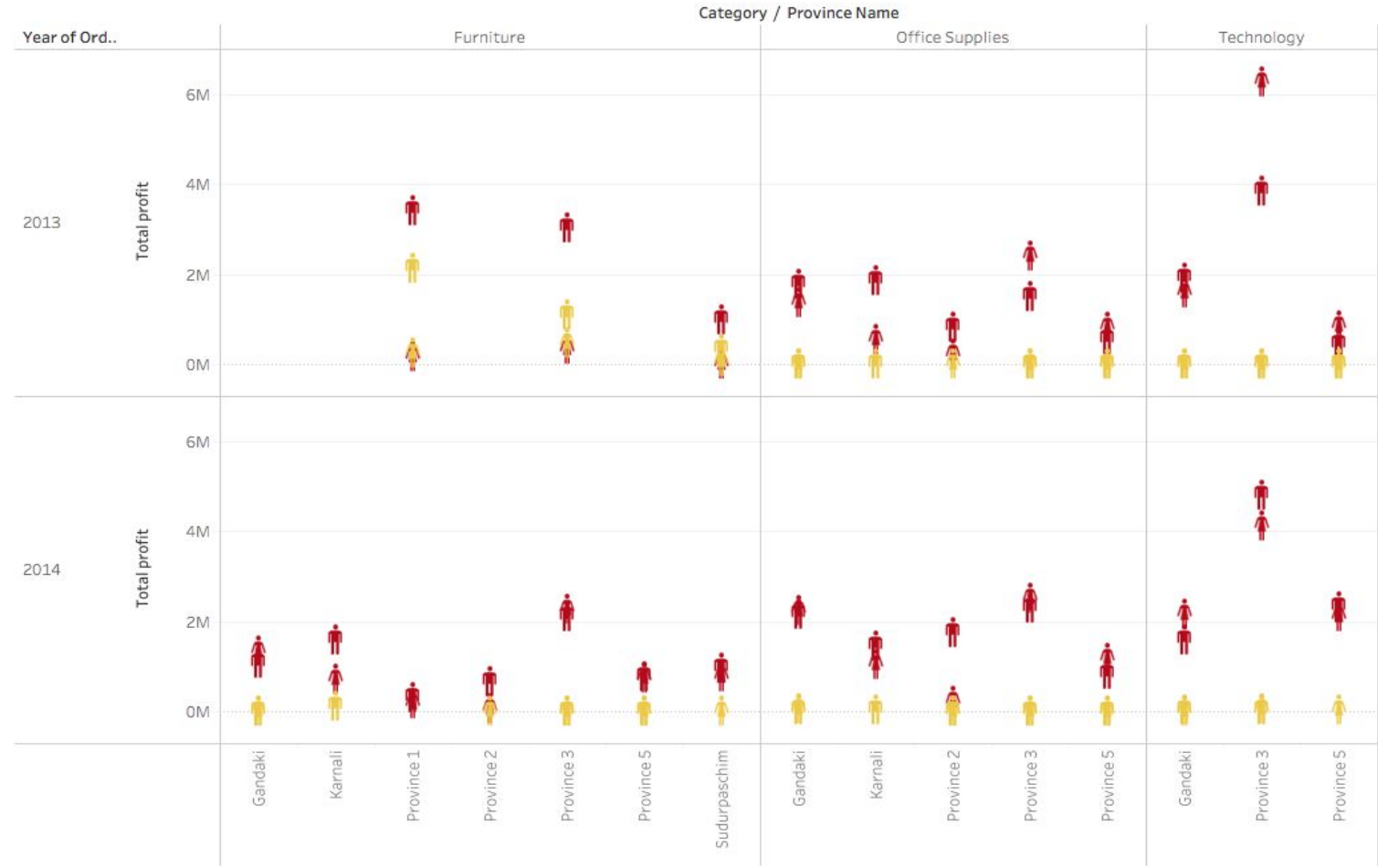
At the lower section of the tableau screen you will notice options to create new sheets Or create a dashboard Or a story :

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


Category / Province Name
Year of Ord..
Furniture
Office Supplies
Technology
2013
2014
Total profit
6M
4M
2M
0M
Gandaki
Karnali
Province 1
Province 2
Province 3
Province 5
Sudurpaschim
Customer gender
Male
Female
Payment method
Cash
epayment

Know your data- Module 2

(Province ecommerce)

Province Name:

Name of the Province.

For example : Province 1, Province 2

Year: Data year

Target: Target number

Indicators : Indicator number

Short description :
SDGs Indicator description

Goals : SDGs goals number

Source :
Data source

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

Create a simple descriptive visualisation