

# Lab 01 Get Familiar with Tableau

COMP7507 Visualization and Visual Analytics

Sept 13, 2022

## 1. Goal

The goal of this lab session is to get familiar with the interface and basic operations of Tableau.

## 2. Download and Installation

**Tableau software:**

Please refer to Lab 0 - Tableau Installation.

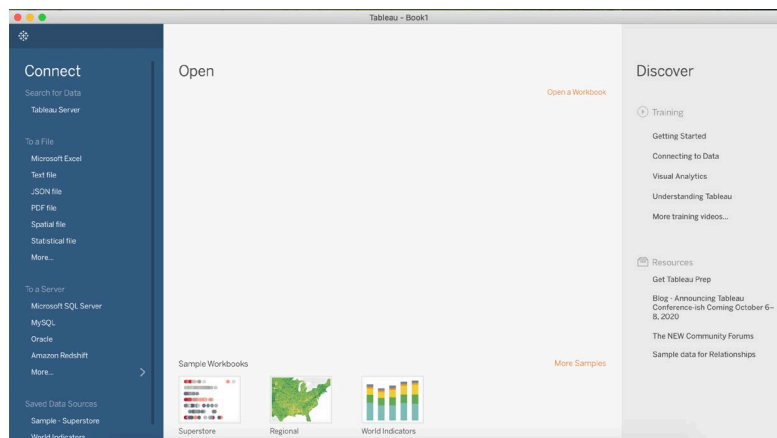
**Dataset:**

There is one data file for this lab session: [Global\\_Superstore.xls\[1\]](#). Please download it from Moodle.

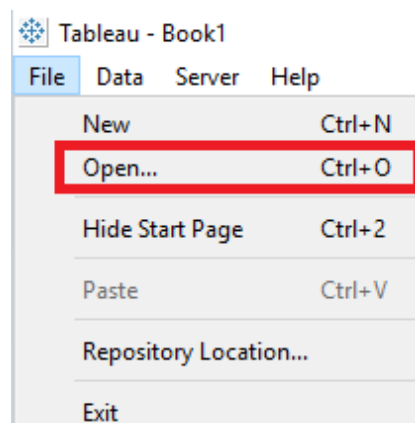
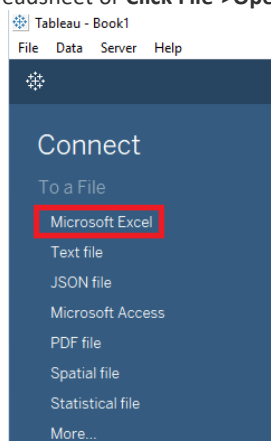
## 3. Getting started with Tableau

### 3.1. Connect to Data

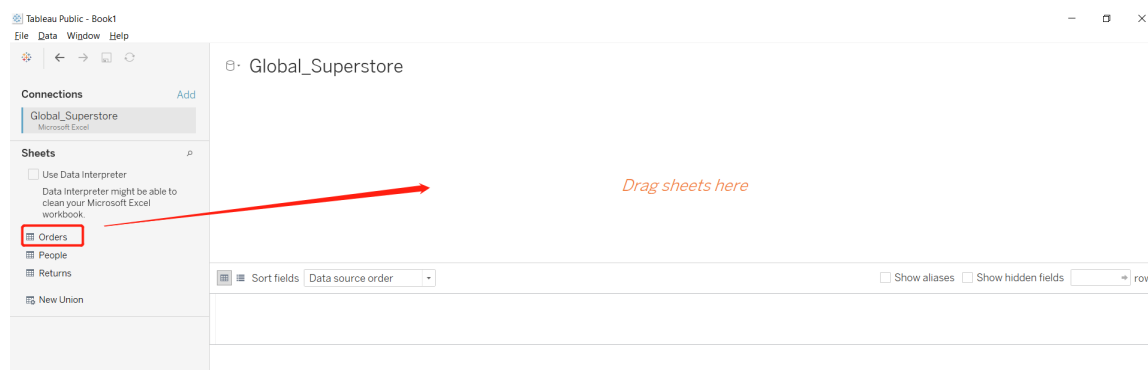
**User Interface.** The figure below shows the welcome screen, where we can connect to new data, connect to saved data sources, or open recently used workbooks. In the Connect pane, we can see a wide variety of data sources (Tableau connects natively, e.g., *excel*, *text files*, *relational databases*, *online data sources*, and many others.)



**Import Data.** In this lab session, we will connect to a global superstore data which is an Excel file. This data set contains transactions of customers purchasing specific products. Click **Microsoft Excel** on the Connect pane to load an excel spreadsheet or **Click File->Open** to load the excel file (Global\_Superstore.xls).



Once the Excel data is loaded, we can choose which sheets or tables we'd like to use. Here, we drag "Orders" out into the canvas. If we would like to add another table, such as "Returns", we could double-click or drag it out as well.



### 3.2. Data Preparation

Once the data is imported, we can see a preview of the data. Note that the types of data fields are different: Row ID is considered as a number, Order Date and Ship Date are considered as dates, etc. We can rename columns here or even change data types, such as changing Row ID to a string.

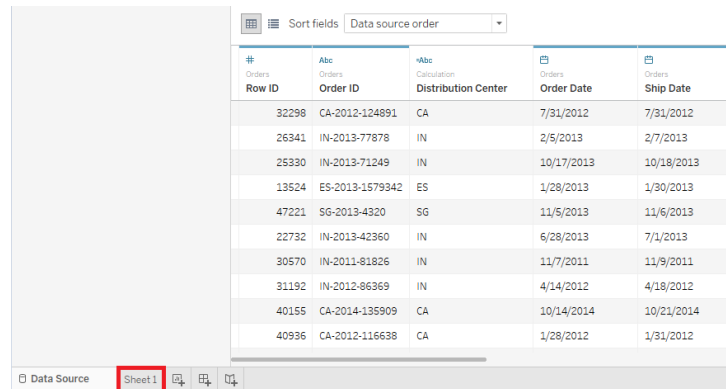
Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment
32298	CA-2012-124891	7/31/2012	7/31/2012	Same Day	RH-19495	Rick Hansen	Consumer
26341	IN-2013-77878	2/5/2013	2/7/2013	Second Class	JR-16210	Justin Ritter	Corporate
25330	IN-2013-71249	10/17/2013	10/18/2013	First Class	CR-12730	Craig Reiter	Consumer
13524	ES-2013-1579342	1/28/2013	1/30/2013	First Class	KM-16375	Katherine Murray	Home Office
47221	SG-2013-4320	11/5/2013	11/6/2013	Same Day	RH-9495	Rick Hansen	Consumer
22732	IN-2013-42360	6/28/2013	7/1/2013	Second Class	JM-15655	Jim Mitchum	Corporate
30570	IN-2011-81826	11/7/2011	11/9/2011	First Class	TS-21340	Toby Swindell	Consumer
31192	IN-2012-86369	4/14/2012	4/18/2012	Standard Class	MB-18085	Mick Brown	Consumer
40155	CA-2014-135909	10/14/2014	10/21/2014	Standard Class	JW-15220	Jane Waco	Corporate
40936	CA-2012-116638	1/28/2012	1/31/2012	Second Class	JH-15985	Joseph Holt	Consumer

The Order ID field in this dataset has three parts: the distribution center code (e.g. CA), the year (e.g. 2012), and the product ID (e.g. 124891). If we would like to split this field and keep only the distribution center code, it is easy to achieve in Tableau. First, click on the drop-down next to the field name and select "Split". Then we have a column for each of those pieces. We can use drop-down again to delete splits 2 and 3 and just keep the 1st. Let's rename the field as "Distribution Center".

Order ID	Order ID - Split 1	Order ID - Split 2	Order ID - Split 3
CA-2012-124891	CA	2012	124891
IN-2013-77878	IN	2013	77878
IN-2013-71249	IN	2013	71249
ES-2013-1579342	ES	2013	1579342
SG-2013-4320	SG	2013	4320
IN-2013-42360	IN	2013	42360
IN-2011-81826	IN	2011	81826
IN-2012-86369	IN	2012	86369

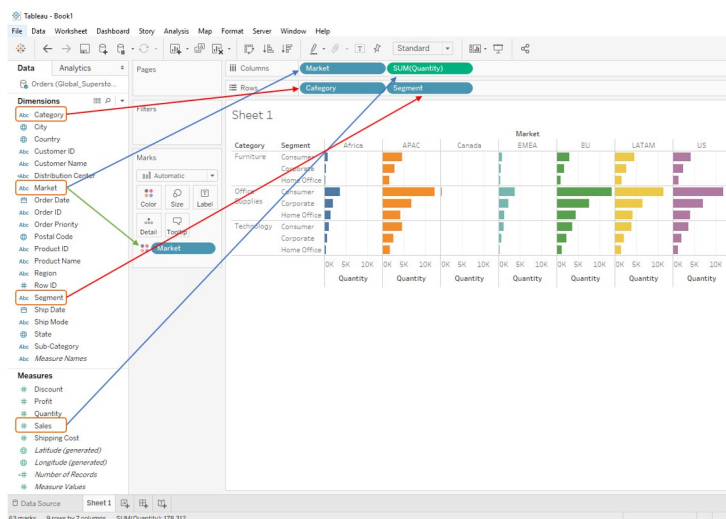
### 3.3. Measure Names and Measure Values

We click on our sheet tab at the bottom line and enter the workspace.



Row ID	Order ID	Distribution Center	Order Date	Ship Date
32298	CA-2012-124891	CA	7/31/2012	7/31/2012
26341	IN-2013-77878	IN	2/5/2013	2/7/2013
25330	IN-2013-71249	IN	10/17/2013	10/18/2013
13524	ES-2013-1579342	ES	1/28/2013	1/30/2013
47221	SG-2013-4320	SG	11/5/2013	11/6/2013
22732	IN-2013-42360	IN	6/28/2013	7/1/2013
30570	IN-2011-81826	IN	11/7/2011	11/9/2011
31192	IN-2012-86369	IN	4/14/2012	4/18/2012
40155	CA-2014-135909	CA	10/14/2014	10/21/2014
40936	CA-2012-116638	CA	1/28/2012	1/31/2012

Now, let's see how easy it is to start building something. Let's bring **Category** and **Segment** to Rows, **Quantity** to Columns, **Market** to Columns, and **Market** to Color, as well. It's easy to create a visualization (as shown below) of how the sales are looking per category, customer segment, and market, in terms of the number of items sold. We can also quickly observe that Canada is an emerging market to be developed.



On the left of the data window, we observe that data is divided into two parts: **Measure Names (with blue icons)** and **Measure Values (with green icons)**.

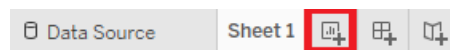
Abc	Order ID	#	Discount
Abc	Order Priority	#	Profit
🌐	Postal Code	#	Quantity
Abc	Product ID	#	Sales
Abc	Product Name	#	Shipping Cost
Abc	Region	🌐	Latitude (generated)
#	Row ID	🌐	Longitude (generated)
Abc	Segment	#	Orders (Count)
📅	Ship Date	#	Measure Values
Abc	Ship Mode		
🌐	State		
Abc	Sub-Category		
Abc	Measure Names		

**Measure Names** are categorical fields, in this case, fields such as date, customer, and Category. These are fields that we want to slice and dice our numerical data by. Measure Names are often discrete. Discrete fields create labels in the chart and are color-coded blue in the data pane and the view.

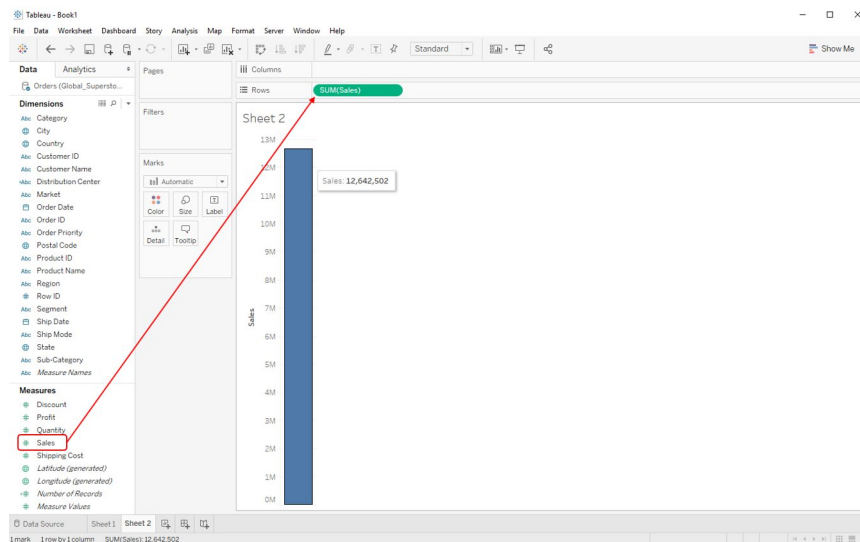
**Measures Values**, on the other hand, are the metrics. They are the numbers we want to analyze. Measures are often continuous. Continuous fields create axes in the chart and their pills are color-coded green.

### 3.4. Build Views

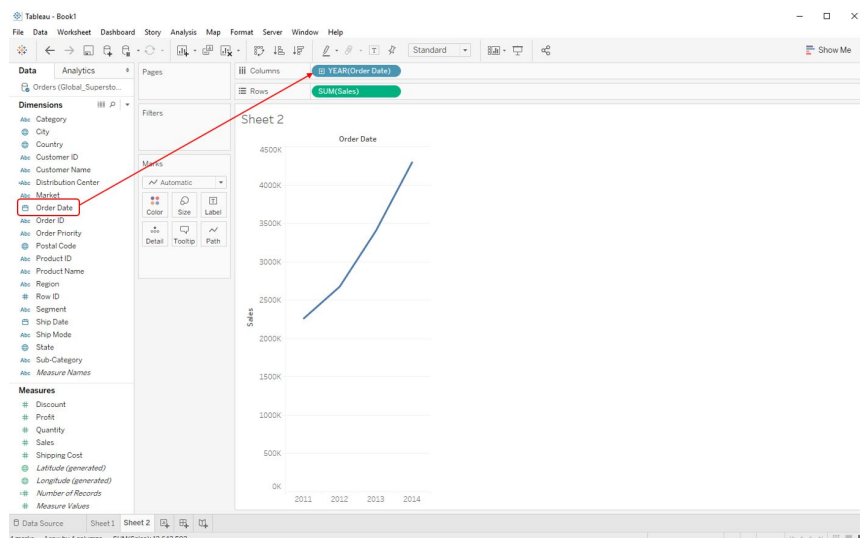
Suppose we are interested in the total sales number. First, we create a new sheet via clicking the new worksheet icon as shown below.

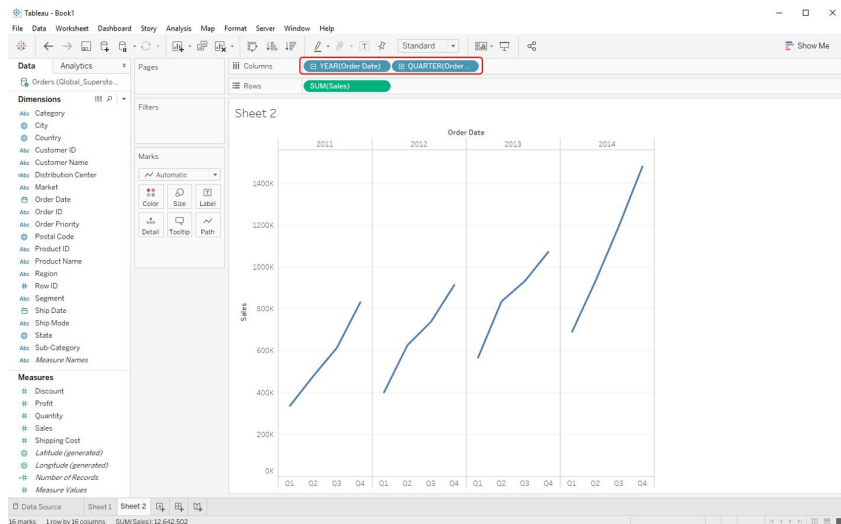


Let's drag Sales to the view. We can see that Tableau queries the database and returns a single result giving us the sum of Sales. We can see that this company has done about 12.5 million in sales.

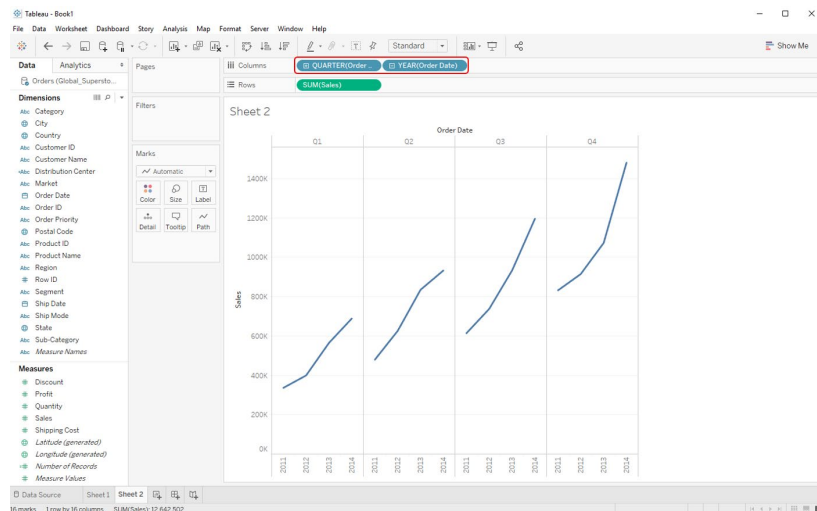


If we want to see this over time, we can drag the Order Date to the top of the view. Tableau aggregates our dates at the year level. We can choose to expand this with the plus (+) symbol. Now we see both quarters and years in view.

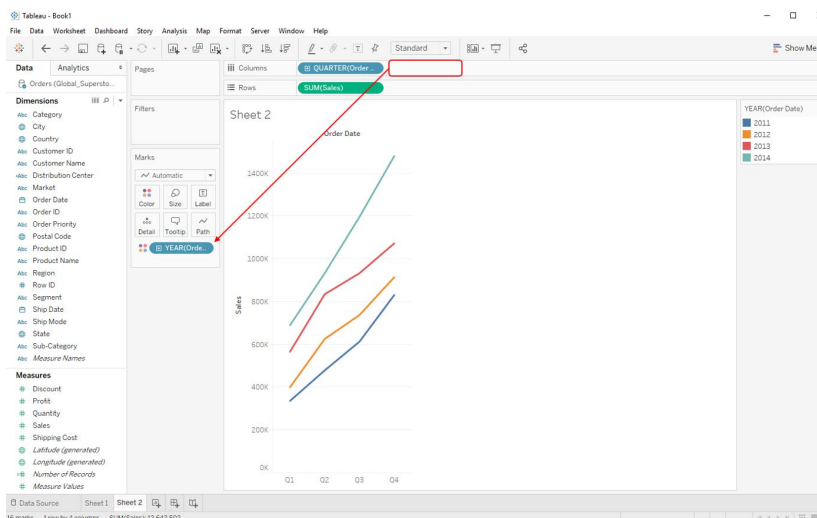




To see how all our quarters are doing over the years, we can easily drag the YEAR item in the Columns and move it behind QUARTER (as shown below). Now we can compare how our growth looks by a quarter across the years.



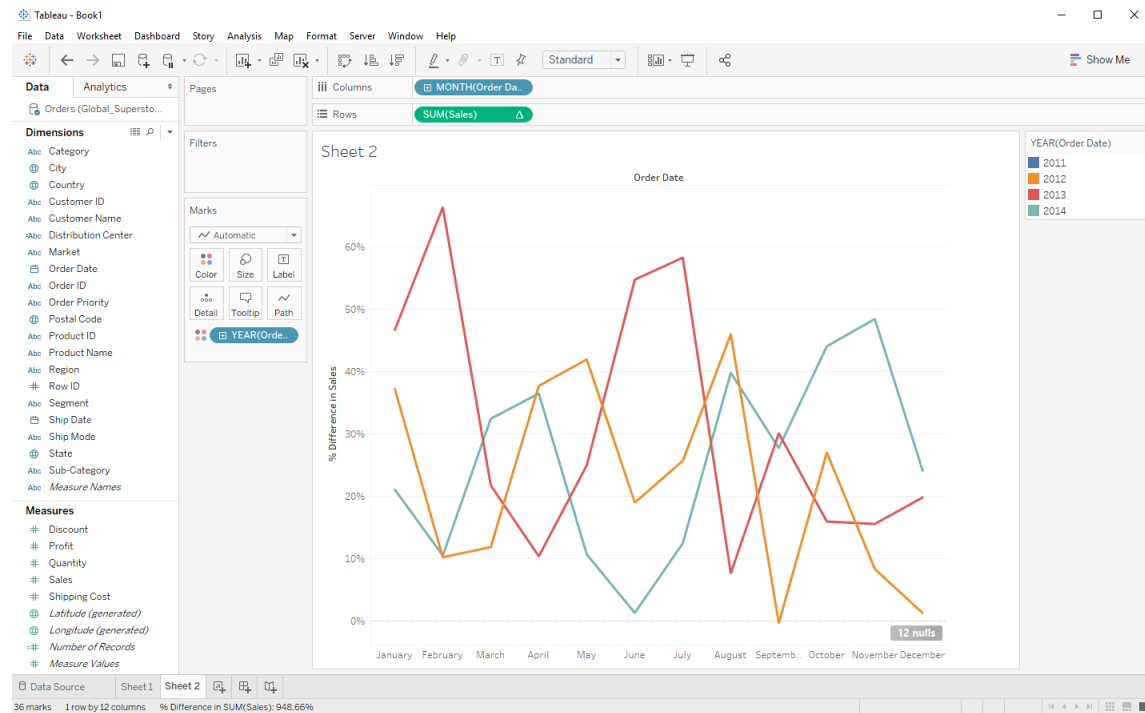
Moving Year to Color shows us all the years on top of each other.



## 5. Tasks for you

Tableau also provides quick table calculations for the imported data. Here are two simple tasks for you:

- (1) Try to visualize the data by months instead of by years.
- (2) Try to explore how to use Tableau to calculate “Year over Year Growth” and visualize it as shown below.
- (3) Upload the screenshot of the resulting graph in (2) to Moodle (**by Sept 27, 2022**).



## 7. References

- [1] The data source and this tutorial are adapted from <http://www.tableau.com/learn/training>
- [2] <https://onlinehelp.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-home.htm>
- [3] <https://public.tableau.com/s/resources>
- [4] [https://public.tableau.com/s/resources?qt-overview\\_resources=1](https://public.tableau.com/s/resources?qt-overview_resources=1)
- [5] <https://community.tableau.com/docs/DOC-9135>